Benemérita Universidad Autónoma de Puebla

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS
LICENCIATURA EN MATEMÁTICAS

UNA APLICACIÓN DE LA LÓGICA DIFUSA EN PSICOLOGÍA DE CONCEPTOS

TESIS

PARA OBTENER EL TÍTULO DE
LICENCIADO EN MATEMÁTICAS

PRESENTA
SOSA PONCE ALVARO GENARO

DIRECTOR DE TESIS
MARTÍNEZ RUÍZ IVAN

PUEBLA, PUE. 02/07/2014

Para mis padres y hermana
Angela, Alvaro y Carmen.

Agradecimientos

Le agradezco en primer lugar a mis padres por darme todo ese apoyo, porque sin ellos no hubiera podido llegar a estas alturas.

También le agradezco a mi director de tesis por asesorarme en encontrar material para complementar en gran parte la sección de la aplicación.

Como olvidar agradecerle a mi hermana que me ayudo con la traducción de algunos libros.

Finalmente les agradezco a los que indirectamente me motivaron en realizar este tema.

Introducción

El propósito de esta tesis es el de dar una introducción a la lógica difusa, y mostrar la utilidad de dicha teoría en áreas de aplicación, en particular, en la educación.
El objetivo de esta tesis es el de dar a conocer la teoría de conjuntos y lógica difusa, ya que es una teoría relativamente nueva y aun tiene muchos frutos que dar.
La tesis está configurada en tres partes. En la primera presento teoría básica de Álgebras Booleanas, MV-Álgebras, Conjuntos Difuso y Lógica Proposicional Clásica para poder ingresar a las ideas de la lógica Proposicional Difusa sin mas requisitos de los expuestos en dichos capítulos.
La segunda parte está dedicada completamente a la Lógica Difusa, empezando con la Lógica Proposicional Difusa y cómo esta se extiende a una teoría más general, para poder ser más flexible al momento de ser aplicada.Por ultimo la tercera parte está dedicada a la aplicación, en la cual se expone una introducción a la Psicología de Conceptos y ejemplos de cómo se utilizan la lógica y Conjuntos Difusos.
La lectura recomendada es seguir el orden expuesto, pero para estudiantes más avanzados, les sugiero omitir los capítulos uno y tres, ya si así lo requiere regresar a estos capítulos.
Para concluir doy una lista de posibles aplicaciones en distintas áreas y quienes trabajan en ello, así como una perspectiva de un trabajo a futuro.

Índice general

Una aplicación de la lógica difusa en psicología de conceptos

Sosa Ponce Alvaro Genaro

fecha

Capítulo 1

Álgebras Booleanas y MV-Álgebras

1.1. Álgebras Booleanas

En esta sección presentamos dos maneras de definir las álgebras Booleanas, algunos ejemplos y la definición de morfismos entre álgebras Boolenas.

Definición 1.1. *Un conjunto B ordenado por la relación "$\leq$" es llamado una álgebra Booleana si cumplen:*
1. Para cada $x, y \in B$, existen un elemento $x \vee y \in B$ y un elemento $x \wedge y \in B$ con las siguientes propiedades:
(a') $x \vee y \geq x$ y $x \vee y \geq y$;
(a") $x \wedge y \leq x$ y $x \wedge y \leq y$;
(b') $z \geq x$ y $z \geq y$ implican $z \geq x \vee y$;
(b") $z \leq x$ y $z \leq x$ implican $z \leq x \wedge y$;

2. Las operaciones "$\vee$" y "$\wedge$"son mutuamente distributivas, es decir,
(a) $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$;
(b) $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$;

3. B tiene elementos unitarios, es decir, existen elemetos $0, 1 \in B$ tales que para cada $x \in B$ tenemos:

$$0 \leq x \leq 1;$$

4. B es complementado, es decir, para cada $x \in B$, existe $\bar{x} \in B$ talque:
(a) $x \vee \bar{x} = 1$;
(b) $x \wedge \bar{x} = 0$.

Note que en la definición anterior, un conjunto ordenado con las propiedades 1 es llamado una Reticula. Por lo que, una latiz distributiva complementada con elementos unitarios es un álgebra Booleana.

Definición 1.2. *Un conjunto B con dos operaciones binarias:*

$$(x, y) \mapsto x \vee y \in B$$

$$(x, y) \mapsto x \wedge y \in B$$

y una unaria

$$x \mapsto \bar{x} \in B$$

es llamada una álgebra Booleana si las siguientes propiedades se cumplen para cada $x, y, z \in B$:

1. ***Conmutatividad***
(a') $x \vee y = y \vee x$; (a")$x \wedge y = y \wedge x$;

2. ***Asociatividad***
(a') $x \vee (y \vee z) = (x \vee y) \vee z$; (a")$x \wedge (y \wedge z) = (x \wedge y) \wedge z$;

3. ***Distributividad***
(a') $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$; (a")$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$;

4. ***Elementos unitarios***
Existen dos elementos $0, 1 \in B$ tales que:
(a') $x \vee 0 = x$; (a")$x \wedge 1 = x$;
(b') $x \wedge 0 = 0$; (b")$x \vee 1 = 1$;

5. ***Complementariedad***
(a') $x \vee \bar{x} = 1$; (a") $x \wedge \bar{x} = 0$;

6. ***Idempotencia***
(a') $x \vee x = x$; (a") $x \wedge x = x$;

7. ***Absorción***
(a') $x \vee (x \wedge y) = x$*; (a")* $x \wedge (x \vee y) = x$.

Teorema 1.3. *Las definiciones 1.1 y 1.2 son equivalentes.*

Demostración. Primero demostremos que la definición 1.1 implica la definición 1.2.
Las propiedades de Distributividad y Complementariedad coinciden en ambas definiciones, por lo tanto, solo es necesario demostrar el resto de propiedades.

Veamos la conmutatividad:
$x \vee y \geq y$ y $x \vee y \geq x$ entonces $x \vee y \geq y \vee x$
$y \vee x \geq y$ y $y \vee x \geq x$ entonces $y \vee x \geq x \vee y$
así, $x \vee y = y \vee x$; del mismo modo se demuestra la otra igualdad.

Probemos la asociatividad:
$x \vee (y \vee z) \geq x$ y $x \vee (y \vee z) \geq y \vee z \geq y$ entonces $x \vee (y \vee z) \geq x \vee y$
además $x \vee (y \vee z) \geq z$ entonces $x \vee (y \vee z) \geq (x \vee y) \vee z$
$(x \vee y) \vee z \geq z$ y $(x \vee y) \vee z \geq x \vee y \geq y$ entonces $(x \vee y) \vee z \geq y \vee z$
además $(x \vee y) \vee z \geq x$ entonces $(x \vee y) \vee z \geq x \vee (y \vee z)$
así, $(x \vee y) \vee z = x \vee (y \vee z)$; análogamente para la otra igualdad.

Mostremos que los elementos unitarios cumplen con 4 de la definición 1.2:
$x \geq x$ y $x \geq 0$ entonces $x \geq x \vee 0$, además $x \vee 0 \geq x$, por lo tanto, $x \vee 0 = x$;
por 1 y 3 tenemos que:
$x \wedge 0 \leq 0$ y $x \wedge 0 \geq 0$ entonces $x \wedge 0 = 0$;
del mismo modo se puede demostrar para el elemento 1.

Veamos la idempotencia:
Como $x \geq x$ entonces $x \geq x \vee x$, además $x \vee x \geq x$, entonces $x \vee x = x$

Probemos la Absorción:
$x \geq x$ y $x \geq x \wedge y$ entonces $x \geq x \vee (x \wedge y)$, además $x \vee (x \wedge y) \geq x$, entonces $x \vee (x \wedge y) = x$; de la misma manera se demuestra la segunda igualdad.

Ahora demostremos que la definición 1.2 implica la definición 1.1:

Primero definamos la relación $\leq$ de la forma siguiente:
$x \leq y$ si y sólo si $x = x \wedge y$

Ahora demostremos que esta relación es de orden.

Reflexividad
Por 6 (a") tenemos $x = x \wedge x$, así, por definición tenemos que $x \leq x$;

Antisimetría
Si $x \leq y$ y $y \leq x$ entonces por definición tenemos:
$x = x \wedge y$ y $y = x \wedge y$, por lo tanto, $x = y$;

Transitividad
Si $x \leq y$ y $y \leq z$, por definición tenemos:
$x = x \wedge y$ y $y = y \wedge z$ entonces $x = x \wedge y = x \wedge (y \wedge z) = (x \wedge y) \wedge z = x \wedge z$, así, $x \leq z$.

Solo resta probar las propiedades 1 y 3 de la definición 1.1

Por 7 y la definición tenemos:
$x \wedge (x \vee y) = x$, es decir, $x \vee y \geq x$, análogamente para la otra desigualdad;

para (a"), tenemos que:
$x \wedge y = (x \wedge x) \wedge y = x \wedge (x \wedge y)$, por definición, $x \wedge y \leq x$; del mismo modo para la segunda desigualdad;

Si $z \geq x$ y $z \geq y$, por definición tenemos que $x = x \wedge z$ y $y = y \wedge z$ entonces:

$$x \vee y = (x \wedge z) \vee (y \wedge z) = (x \vee y) \wedge z$$

Por lo tanto, $z \geq x \vee y$;

Si $z \leq x$ y $z \leq y$, por definición tenemos que $z = x \wedge z$ y $z = y \wedge z$ entonces:

$$z = x \wedge z = x \wedge (y \wedge z) = (x \wedge y) \wedge z = z \wedge (x \wedge y)$$

Por lo tanto, $z \leq x \wedge y$;

Para terminar utilicemos 4 de la definición 1.2, $x \wedge 0 = 0$ y $x \wedge 1 = x$, por definición tenemos que $0 \leq x$ y $x \leq 1$;

Por lo tanto las dos definiciones anteriores son equivalentes.

□

El objetivo de dar estas dos definiciones es el de mostrar que se pueden abordar de distintas formas a las álgebras Boolenas, ya que la primera definición es más conjuntista y la segunda es más algebraica.

Teorema 1.4. *Sea B un algebra Booleana.Entonces para cada $x \in B$ el complemento $\bar{x}$ es único. Los elementos unitarios también son únicos.*

Demostración. Sea $x \in B$ y supongamos que existen dos complementos $\bar{x}, \bar{y} \in B$. Como ambos son complementos satisfacen 5 de la definición 1.2, así:

$$\bar{x} = \bar{x} \wedge (x \vee \bar{y}) = (\bar{x} \wedge x) \vee (\bar{x} \wedge \bar{y}) = \bar{x} \wedge \bar{y}$$

$$\bar{y} = \bar{y} \wedge 1 = \bar{y} \wedge (x \vee \bar{x}) = (\bar{y} \wedge x) \vee (\bar{y} \wedge \bar{x}) = \bar{y} \wedge \bar{x}$$

entonces $\bar{x} = \bar{y}$. Para mostrar la unicidad de los elementos unitarios es más sencillo.
Por 3 de la definición 1.1 si existiera otro elemento unitario a tal que para cada $x \in B$ se tiene $a \leq x$, en particular, $0 \in B$ tenemos $a \leq 0$ y $0 \leq a$, por lo que $a = 0$ de la misma manera se procede para demostrar la unicidad de 1.

□

Teorema 1.5. *Si B es un álgebra Booleana, entonces:*
(a) La aplicación $x \mapsto \bar{x} : B \to B$ es una involución, es decir, $x = \bar{\bar{x}}$ para cada $x \in B$;
(b) $\overline{x \vee y} = \bar{x} \wedge \bar{y}$ y $\overline{x \wedge y} = \bar{x} \vee \bar{y}$ para cada $x, y \in B$.

Demostración. (a) Por el teorema anterior, la aplicación esta bien definida y

$$x \vee \bar{x} = 1 = \bar{\bar{x}} \vee \bar{x}$$

por la unicidad tenemos que $x = \bar{\bar{x}}$;
(b) Solo se demostrará una igualdad ya que la otra se realiza de forma análoga

$$(x \vee y) \vee (\bar{x} \wedge \bar{y}) = x \vee [y \vee (\bar{x} \wedge \bar{y})] = x \vee [(y \vee \bar{x}) \wedge (y \vee \bar{y})] = x \vee (y \vee \bar{x}) = (x \vee \bar{x}) \vee y = 1$$

Por lo tanto, $\overline{x \vee y} = \bar{x} \wedge \bar{y}$

□

Note que las propiedades (b) del teorema anterior son llamadas leyes De Morgan.

Definición 1.6. *Consideremos dos álgebras Booleanas B y B'. Una función $\varphi : B \to B'$ con las propiedades:*
(a) $\varphi(x \vee y) = \varphi(x) \vee \varphi(y)$,para cada $x, y \in B$
(b) $\varphi(\bar{x}) = \overline{\varphi(x)}$,para cada $x \in B$
es llamada un Morfismo de álgebras Booleanas. Si el morfismo φ es una biyección, entonces φ es llamada un isomorfismo de álgebras Booleanas.

Teorema 1.7. *Sea E un conjunto arbitrario; el conjunto $\mathbf{P}(E)$ de todos los subconjuntos de E, junto con las operaciones $\bigcup$, $\bigcap$ y C, es un álgebra Booleana.*

Demostración. Por las propiedades de estas operaciones solo es necesario aclarar qué conjuntos son los elementos unitarios, en este caso son el conunto vacio y el conjunto E, toman el papel de $0, 1$, respectivamente.
Con esto las propiedades se verifican inmediatamente.

$\square$

Definición 1.8. *Las álgebras Booleanas $\mathbf{P}(E)$ son llamadas las álgebras Booleanas estándar.*

En lógica matemática bivalente usualmente se utiliza el álgebra Booleana de solo dos elementos $\Gamma = \{0, 1\}$.
En el conjunto Γ introducimos una relación de orden "$\leq$" de la siguiente manera:

$$0 \leq 0,\ 0 \leq 1,\ 1 \leq 1$$

Para cada par $x, y \in \Gamma$ definimos $x + y = \max\{x, y\}$, $x \cdot y = \min\{x, y\}$ y

$$\bar{x} = \begin{cases} 0 & \text{si } x = 1 \\ 1 & \text{si } x = 0. \end{cases}$$

Teorema 1.9. *El conjunto Γ junto con las operaciones $x + y$, $x \cdot y$ y $\bar{x}$ es un álgebra Booleana.*

Demostración. Se demostrará utilizando la definición 1.1.
Por las propiedades del máximo y mínimo, para cada $x, y \in \Gamma$ tenemos que:
$\max\{x, y\} \geq x$ y $\max\{x, y\} \geq y$;
$\min\{x, y\} \leq x$ y $\min\{x, y\} \leq y$;

Si $z \geq x$ y $z \geq y$ entonces $z \geq \max\{x, y\}$;
Si $z \leq x$ y $z \leq y$ entonces $z \leq \min\{x, y\}$;
Por lo tanto se cumplen las propiedades de 1 en la definición 1.1.
Ahora veamos una de las propiedades de distributividad
Sean $x, y, z \in \Gamma$, tenemos que:

$$x(y+z) = \text{mín}\{x, \text{máx}\{y, z\}\}$$
$$xy + xz = \text{máx}\{\text{mín}\{x, y\}, \text{mín}\{x, z\}\}$$

Demostremos por casos que estas dos expresiones son iguales.
(i) Si $x \leq y$ tenemos que:

$$x(y+z) = \text{mín}\{x, \text{máx}\{y, z\}\}$$

$$xy + xz = \text{máx}\{x, \text{mín}\{x, z\}\}$$

(ia) Si $x \leq z$ entonces:

$$x(y+z) = \text{mín}\{x, \text{máx}\{y, z\}\} = x$$

$$xy + xz = \text{máx}\{x, x\} = x$$

(ib) Si $x \geq z$ entonces:

$$x(y+z) = \text{mín}\{x, y\} = x$$

$$xy + xz = \text{máx}\{x, z\} = x$$

(ii) Si $x \geq y$ tenemos que:

$$x(y+z) = \text{mín}\{x, \text{máx}\{y, z\}\}$$

$$xy + xz = \text{máx}\{y, \text{mín}\{x, z\}\}$$

(iia) Si $x \leq z$ entonces:

$$x(y+z) = \text{mín}\{x, z\} = x$$

$$xy + xz = \text{máx}\{y, x\} = x$$

(iib) Si $x \geq z$ entonces:

$$x(y+z) = \text{mín}\{x, \text{máx}\{y, z\}\} = \text{máx}\{y, z\}$$

$$xy + xz = \text{máx}\{y, z\}$$

Por lo tanto, $x(y+z) = xy+xz$, del mismo modo se verifica la otra igualdad.
La propiedad 3 se cumple por definición de "$\leq$".
Solo resta probar la complementariedad.
Verifiquemos esto por casos:
(i) Si $x = 0$ entonces:
$\bar{x} = 1$, así, $x + \bar{x} = 0 + 1 = \text{máx}\{0, 1\} = 1$ y $x \cdot \bar{x} = 0 \cdot 1 = \text{mín}\{0, 1\} = 0$;
(ii) Si $x = 1$ se procede del modo anterior ya que las operaciones son conmutativas.
Por todo lo anterior se concluye que Γ con estas operaciones es una álgebra booleana.

□

Con esto concluimos esta introducción de álgebras Booleanas.

1.2. MV-Álgebras

En esta sección introduciremos la definición de MV-Álgebra, algunas propiedades basicas y un ejemplo.

Definición 1.10. *Una MV-Álgebra es un sistema $(A, \oplus, \odot, ^{-}, 0, 1)$, donde A es un conjunto no vacio, 0 y 1 elementos de A, $\oplus$ y $\odot$ son operaciones binarias sobre A, y $^{-}$ es una operación sobre A, que cumplen con los siguientes once axiomas:*

Ax.1. (a) $x \oplus y = y \oplus x$; *(b)* $x \odot y = y \odot x$;
Ax.2. (a) $x \oplus (y \oplus z) = (x \oplus y) \oplus z$; *(b)* $x \odot (y \odot z) = (x \odot y) \odot z$;
Ax.3. (a) $x \oplus \bar{x} = 1$; *(b)* $x \odot \bar{x} = 0$;
Ax.4. (a) $x \oplus 1 = 1$; *(b)* $x \odot 0 = 0$;
Ax.5. (a) $x \oplus 0 = x$; *(b)* $x \odot 1 = x$;
Ax.6. (a) $\overline{x \oplus y} = \bar{x} \odot \bar{y}$; *(b)* $\overline{x \odot y} = \bar{x} \oplus \bar{y}$;
Ax.7. (a) $\bar{\bar{x}} = x$;
Ax.8. (a) $\bar{0} = 1$.

Para los siguientes tres axiomas es necesario definir las operaciones, $\vee$ y $\wedge$, sobre A:

$$x \vee y = (x \odot \bar{y}) \oplus y$$

$$x \wedge y = (x \oplus \bar{y}) \odot y.$$

De esta forma, los axiomas son

Ax.9. (a) $x \vee y = y \vee x$*; (b)* $x \wedge y = y \wedge x$*;*
Ax.10. (a) $x \vee (y \vee z) = (x \vee y) \vee z$ *; (b)* $x \wedge (y \wedge z) = (x \wedge y) \wedge z$*;*
ax.11. (a) $x \oplus (y \wedge z) = (x \oplus y) \wedge (x \oplus z)$ *; (b)* $x \odot (y \vee z) = (x \odot y) \vee (x \odot z)$*.*

Teorema 1.11. *Sea* $(A, \oplus, \odot, ^{-}, 0, 1)$ *una MV-Álgebra arbitraria. Para cada* $x, y \in A$ *tenemos:*

(1) $x \vee 0 = x = x \wedge 1$*,* $x \wedge 0 = 0$ *y* $x \vee 1 = 1$*;*
(2) $x \vee x = x = x \wedge x$*;*
(3) $\overline{x \vee y} = \bar{x} \wedge \bar{y}$ *y* $\overline{x \wedge y} = \bar{x} \vee \bar{y}$*;*
(4) $x \wedge (x \vee y) = x = x \vee (x \wedge y)$*;*
(5) Si $x \oplus y = 0$ *entonces* $x = y = 0$*;*
(6) Si $x \odot y = 1$ *entonces* $x = y = 1$*;*
(7) Si $x \vee y = 0$ *entonces* $x = y = 0$*;*
(8) Si $x \wedge y = 1$ *entonces* $x = y = 1$*.*

Demostración. (1) $x \vee 0 = (x \odot \bar{0}) \oplus 0 = (x \odot 1) \oplus 0 = x \oplus 0 = x = x \odot 1 = (x \oplus 0) \odot 1 = (x \oplus \bar{1}) \odot 1 = x \wedge 1$;
$x \wedge 0 = (x \oplus \bar{0}) \odot 0 = 0$ y $x \vee 1 = (x \odot \bar{1}) \oplus 1 = 1$;

(2) $x \vee x = (x \odot \bar{x}) \oplus x = 0 \oplus x = x = 1 \odot x = (x \oplus \bar{x}) \odot x = x \wedge x$;

(3) $\overline{x \vee y} = \overline{(x \odot \bar{y}) \oplus y} = \overline{x \odot \bar{y}} \odot \bar{y} = (\bar{x} \oplus \bar{\bar{y}}) \odot \bar{y} = \bar{x} \wedge \bar{y}$;
La otra igualdad se realiza análogamente;

(4) $x \wedge (x \vee y) = x \wedge (y \vee x) = (y \vee x) \wedge x = ((y \odot \bar{x}) \oplus x) \wedge x = ((y \odot \bar{x}) \oplus x \oplus \bar{x}) \odot x = (y \odot \bar{x} \oplus 1) \odot x = 1 \odot x = x$
Del mismo modo se procede para mostrar la otra igualdad;

(5) $0 = x \wedge 0 = x \wedge (x \oplus y) = (x \oplus 0) \wedge (x \oplus y) = x \oplus (0 \wedge y) = x \oplus 0 = x$ De la misma manera se muestra que $y = 0$;

(6) Se utiliza el mismo procedimiento seguido en (5);

(7) Si $x \vee y = 0$ entonces $(x \odot \bar{y}) \oplus y = 0$, por (5) $y = 0$ y de hecho

$x = 0$;

(8) Se sigue de la misma forma que en (7). $\square$

En una MV-Álgebra podemos introducir una relación de orden "$\leq$".

Definición 1.12. *Si $x, y \in A$, decimos que $x \leq y$ si $x \vee y = y$.*

Teorema 1.13. *Sea A una MV-Álgebra. Para cada $x, y, z \in A$ se tienen las siguientes relaciones:*
(1) $0 \leq x \leq 1$*;*
(2) $x \leq x$*;*
(3) Si $x \leq y$ y $y \leq z$ entonces $x \leq z$;
(4) Si $x \leq y$ y $y \leq x$ entonces $x = y$;
(5) $x \leq y$ si y sólo si $x \wedge y = x$;
(6) $x \leq y$ si y sólo si $\bar{y} \leq \bar{x}$.

Demostración. (1) Por (1) del teorema anterior tenemos que $0 \vee x = x$ y $x \vee 1 = 1$, por la definición tenemos el resultado.

(2) Por (2) del teorema anterior tenemos $x \vee x = x$, por definición se tiene el resultado.

(3) Si $x \leq y$ y $y \leq z$, por definición tenemos $x \vee y = y$ y $y \vee z = z$, entonces:
$x \vee z = x \vee (y \vee z) = (x \vee y) \vee z = y \vee z = z$, por definición tenemos que $x \leq z$;

(4) Si $x \leq y$ y $y \leq x$, por defineión tenemos, $x = x \vee y = y$, así, $x = y$;

(5) Si $x \leq y$ entonces $x \vee y = y$, así, $x \wedge y = x \wedge (x \vee y) = x$;
Si $x \wedge y = x$ entonces $x \vee y = (x \wedge y) \vee y = y$, así $x \leq y$;

(6) $x \leq y$ si y sólo si $x \vee y = y$ si y sólo si $\overline{x \vee y} = \bar{y}$ si y sólo si $\bar{x} \wedge \bar{y} = \bar{y}$ si y sólo si $\bar{y} \leq \bar{x}$

$\square$

Teorema 1.14. *Las siguientes condiciones son equivalentes:*
(1) $x \leq y$*;*
(2) $y \oplus \bar{x} = 1$*;*
(3) $x \odot \bar{y} = 0$.

Demostración. (1) implica (2)
Por (6) del teorema anterior tenemos que $\bar{y} \leq \bar{x}$, es decir, $\bar{x} \vee \bar{y}$ entonces

$$\bar{x} \oplus y = (\bar{x} \vee \bar{y}) \oplus y = ((\bar{x} \odot y) \oplus \bar{y}) \oplus y = (\bar{x} \odot y) \oplus (\bar{y} \oplus y) = (\bar{x} \odot y) \oplus 1 = 1$$

(2) implica (3)
$y \oplus \bar{x} = 1$ entonces $\overline{y \oplus \bar{x}} = \bar{1}$, así, $\bar{y} \odot x = 0$

(3) implica (1)
$x \vee y = (x \odot \bar{y}) \oplus y = 0 \oplus y = y$, así, $x \leq y$
Por lo tanto son equivalentes.

□

Ejemplo 1.1. Sea $A = [0, 1]$ el intervalo unitario. Definamos las operaciones siguientes:
Para cada $x, y \in A$,
$x \oplus y = \text{mín}(1, x + y)$,
$x \odot y = \text{máx}(0, x + y - 1)$,
$\bar{x} = 1 - x$,
entonces el sistema $(A, \oplus, \odot, ^{-}, 0, 1)$ es una MV-Álgebra. Se puede verificar fácilmente que
$x \vee y = \text{máx}(x, y)$;
$x \wedge y = \text{mín}(x, y)$.

Definición 1.15. *(a) Sea $(A, \oplus, \odot, ^{-}, 0, 1)$ una MV-Álgebra, decimos que B es una MV-subálgebra de A si $B \subseteq A$, $0, 1 \in B$ y B es cerrado bajo las operaciones $\oplus$, $\odot$, $^{-}$;*
(b) Un sistema $(B, \oplus, \odot, ^{-}, 0, 1)$ es una imagen homomorfa de A si existe una función f de A en B talque $f(0) = 0$, $f(1) = 1$ y f preserva las operaciones $\oplus$, $\odot$, $^{-}$; si la función f es biyectiva, entonces f es llamado un isomorfismo de A en B.

Teorema 1.16. *Sean $(A, \oplus, \odot, ^{-}, 0, 1)$ una MV-Álgebra, B el conunto de elementos de A que son idempotentes con respecto a la operación $\oplus$. Entonces B es cerrado bajo las operaciones $\oplus$, $\odot$, $^{-}$.*

Demostración. Sean $x, y \in A$ dos elementos idempotentes con respecto a la operación aditiva en A. Entonces:
$(x \oplus y) \oplus (x \oplus y) = (x \oplus x) \oplus (y \oplus y) = x \oplus y$;

$(x \odot y) \odot (x \odot y) = (x \odot x) \odot (y \odot y) = x \odot y$;
$(x \odot y) \oplus (x \odot y) = x \odot y$;
$\bar{x} \oplus \bar{x} = \bar{x}$.
Así obtenemos que $x \oplus y$, $x \odot y$, $\bar{x}$ son elementos idempotentes de A. Por lo tanto, B es cerrado bajo las operaciones $\oplus$, $\odot$, $^{-}$. □

Observación. Toda Álgebra Booleana es una MV-Álgebra, ya que si consideramos las operaciones $\oplus = \vee$ y $\odot = \wedge$ de un Álgebra Booleana, estas operaciones cumplen con todos los axiomas de una MV-Álgebra.

Capítulo 2

Conjuntos Difusos

El objetivo de este capítulo es el de dar una introducción a la Teoría de Conjunto difusos. El término lo introdujo Lofti A. Zadeh en 1965, donde la pertenencia en un conjunto difuso no es cuestión de afirmar o negar, si no es cuestión de grado.
Por lo anterior se puede considerar que los conjuntos difusos son más flexibles que los conjuntos clasicos, por lo que, en algunas aplicaciones son mejor opción.

2.1. Conceptos Básicos

Definición 2.1. *Sea U un conjunto, que llamaremos conjunto universal. A la función $A : U \to [0, 1]$ se le llamará un $[0, 1]$-conjunto difuso, o simplemente conjunto difuso.*

Ejemplo 2.1. Las siguientes funciones estan definidas en el intervalo $[0, 80]$:

$$A_1(x) = \begin{cases} 1 & \text{si } x \leq 20 \\ \frac{(35-x)}{15} & \text{si } 20 < x < 35 \\ 0 & \text{si } x \geq 35 \end{cases}$$

$$A_2(x) = \begin{cases} 0 & \text{si } x \leq 20 \text{ o } x \geq 60 \\ \frac{(x-20)}{15} & \text{si } 20 < x < 35 \\ \frac{(60-x)}{15} & \text{si } 45 < x < 60 \\ 1 & \text{si } 35 \leq x \leq 45 \end{cases}$$

$$A_3(x) = \begin{cases} 0 & \text{si } x \leq 45 \\ \frac{(x-45)}{15} & \text{si } 45 < x < 60 \\ 1 & \text{si } x \geq 60 \end{cases}$$

donde estos conjuntos difusos representan los conceptos de Joven, Mediana edad y Viejo, respectivamente.

Definición 2.2. *Sea A un conjunto difuso definido en U.*
(1) El soporte de A es $\{x \in U |\ A(x) > 0\}$;
(2) El núcleo de A es $\{x \in U |\ A(x) = 1\}$, si el núcleo no es vacio a A se le llama conjunto difuso normal; en otro caso se le llama subnormal.
(3) Sea $\alpha \in [0, 1]$ un α-corte de A, se define como:

$$A_\alpha = \{x \in U |\ A(x) \geq \alpha\}$$

(4) El conjunto potencia difuso es $[0, 1]^U = \{f |\ f : U \to [0, 1] \quad \textit{función}\}$
(5) La cardinalidad de un conjunto difuso se puede definir de la forma:

$$|A| = \begin{cases} \sum_{x \in U} A(x) & \textit{si } U \textit{ es numerable} \\ \int_{x \in U} A(x) & \textit{si } U \textit{ es no numerable} \end{cases}$$

Una clase especial de conjuntos difusos son los intervalos difusos. Tomemos como conjunto universal a $\mathbb{R}$, E intervalo difuso se define como:

$$E(x) = \begin{cases} f_E(x) & \text{si } x \in [a, b) \\ 1 & \text{si } x \in [b, c] \\ g_E(x) & \text{si } x \in (c, d] \\ 0 & \text{en otro caso} \end{cases}$$

donde $a, b, c, d \in \mathbb{R}$, tales que $a \leq b \leq c \leq d$. f_E usualmente es una función continua, estrictamente creciente de 0 a 1 y g_E denota una función continua estrictamente decreciente de 1 a 0.
En un intervalo difuso si $b = c$ se llama un número difuso.
La representación por α-cortes de E es expresada por la igualdad

$$E_\alpha = [f_E^{-1}(\alpha), g_E^{-1}(\alpha)]$$

Si f_E y g_E son lineales, entonces $f_E(x) = \frac{x-a}{b-a}$ y $g_E(x) = \frac{d-x}{d-c}$, este tipo de intervalos difusos se les llaman trapezoidales, que son frecuentemente utilizados en aplicaciones.
Para los conjuntos difusos trapazoidales pueden ser expresados por α-cortes de la forma:

$$E_\alpha = [a + (b - a)\alpha, d - (d - c)\alpha]$$

2.2. Operaciones sobre conjuntos difusos

En esta sección se definirán los modificadores ,complementos, intersecciones y uniones, así como también algunas de sus porpiedades. Primero veamos las operaciones difusas estándar.

Definición 2.3. *Sean A, B conjuntos difusos.*
(1) La unión difusa estándar es dada por $(A \cup B)(x) = \text{máx}\{A(x), B(x)\}$;
(2) La intersección difusa estándar es dada por $(A \cap B)(x) = \text{mín}\{A(x), B(x)\}$;
(3) El complemento difuso estándar es dado por $(CB)(x) = 1 - B(x)$.

2.2.1. Modificadores Difusos

Definición 2.4. *Un modificador difuso es una función $m : [0,1] \to [0,1]$, así, para un conjunto difuso A se tiene:*

$$[m(A)](x) = m(A(x))$$

Ejemplo 2.2. Una conveniente clase de funciones, m_λ, son definadas por:

$$m_\lambda(a) = a^\lambda$$

donde $\lambda \in (0, \infty)$.Dada $a \in [0,1]$, se distinguen dos clases de modificadores.
1. Para cada $\lambda \in (0,1)$, $m_\lambda(a) > a$, m_λ incrementa a a.
Cuanto menor sea el valor de λ, más fuerte es la modificación.
2. Para cada $\lambda \in (1, \infty)$, $m_\lambda(a) < a$, m_λ decrece a a. Cuanto mayor sea el valor de λ, más fuerte es el modificador.

Definición 2.5. *Sea m un modificador difuso.*
Si $m(a) < a$, el modificador se denomina modificador fuerte.
Si $m(a) > a$, el modificador se denomina modificador débil.

Observación. Para fines prácticos se le pide a los modificadores las siguientes dos propiedades:
Ax.m1. $m(0) = 0$ y $m(1) = 1$
Ax.m2. m es una función continua.
Note que el ejemplo anterior cumple con estas dos propiedades adicionales.

2.2.2. Complementos Difusos

Definición 2.6. *Un complemento difuso es una función* $c : [0,1] \to [0,1]$ *que invierte el orden, con* $c(0) = 1$ *y* $c(1) = 0$. *Así:*

$$[c(A)](x) = c(A(x))$$

donde A *es un conjunto difuso.*

Observación. En casos prácticos es deseable considerar propiedades adicionales de los complementos difusos, de los cuales los más importantes son:
Ax.c1. c es una función continua;
Ax.c2. c es involutivo, es decir, $c(c(a)) = a$ para cada $a \in [0,1]$.

Teorema 2.7. *Sea una función* $c : [0,1] \to [0,1]$ *que invierte el orden y es involutivo. Entonces* c *es complemento difuso continuo. MÃ¡s aÃºn,* c *es biyectiva.*

Demostración. (i) Primero veamos que es complemento difuso.
Como el rango de c esta en $[0,1]$ tenemos que, $c(0) \leq 1$ y $c(1) \geq 0$, como c invierte el orden tenemos que $c(c(0)) \geq c(1)$ y por ser involutivo tenemos que $0 = c(c(0)) \geq c(1)$, así, $c(1) = 0$. Ahora por la involución tenemos que $c(0) = c(c(1)) = 1$. Por lo tanto c es un complemento difuso.
(ii) Ahora veamos que c es biyectivo.
Note que para cada $a \in [0,1]$ existe $b = c(a) \in [0,1]$ tal que $c(b) = c(c(a)) = a$. Supongamos ahora que $c(a_1) = c(a_2)$, por la involución tenemos que $a_1 = c(c(a_1)) = c(c(a_2)) = a_2$. Por lo tanto c es biyectiva.
(iii)Veamos que c es continua.
Supongamos que c es discontinua en a_0, entonces:

$$b_0 = \lim_{a \to a_0 -} c(a) > c(a_0)$$

así, existe $b_1 \in [0,1]$ talque $b_0 > b_1 > c(a_0)$, esto nos quiere decir que no existe $a_1 \in [0,1]$ talque $c(a_1) = b_1$, pero esto contradice que c es biyectiva. Por lo tanto c es continua.

□

Ejemplo 2.3. Un complemento difuso que no es continuo ni involutivo,

$$c_1(a) = \begin{cases} 1 & \text{si } a \leq t \\ 0 & \text{si } a > t \end{cases}$$

donde $a \in [0,1]$ y $t[0,1)$.
Un complemento difuso continuo que no es involutivo,

$$c(a) = \frac{1 + \cos \pi a}{2}$$

ya que $c(0{,}33) = 0{,}75$, pero $c(0{,}75) = 0{,}15$.
Una clase de complementos difusos involutivos son la clase Sugeno definida por:

$$c_\lambda(a) = \frac{1-a}{1+\lambda a}$$

donde $\lambda \in (-1, \infty)$. Para $\lambda = 0$, la función se convierte en el complemento difuso estándar.
Otra clase de complementos difusos involutivos son la clase Yager definida por:

$$c_w(a) = (1 - a^w)^{\frac{1}{w}}$$

donde $w \in (0, \infty)$. Para $w = 1$, esta función se convierte en el complemento difuso estándar, $c(a) = 1 - a$.

2.2.3. Intersecciones Difusas: t-normas

Definición 2.8. *Una Intersección difusa(o t-norma) es una función* $i : [0,1] \times [0,1] \to [0,1]$ *que cumple con las siguientes propiedades:*
Ax.i1. $i(a,1) = a$
Ax.i2. Si $b \leq d$ *entonces* $i(a,b) \leq i(a,d)$
Ax.i3. $i(a,b) = i(b,a)$
Ax.i4. $i(a, i(b,d)) = i(i(a,b), d)$
donde $a, b, d \in [0,1]$. *Así, para* A, B *conjuntos difusos:*

$$[i(A,B)](x) = i(A(x), B(x))$$

A partir de este momento nos referiremos a la intersección difusa por t-norma.

Observación. Como en casos anteriores, también se desearían unas propiedades adicionales para las t-normas, entre estas propiedades las más importantes son las siguientes:
$Ax.i5$. i es una función continua;
$Ax.i6$. $i(a,a) < a$;

$Ax.i7$. Si $a_1 < a_2$ y $b_1 < b_2$ implica $i(a_1, b_1) < i(a_2, b_2)$.

Una t-norma continua que satisface $Ax.i6.$, es llamada una t-norma Arquimediana; si además satisface $Ax.i7.$ es llamada una t-norma Arquimediana estricta.

Teorema 2.9. *La intersección difusa estándar es la única t-norma idempotente.*

Demostración. Claramente, $\text{mín}(a, a) = a$ para cada $a \in [0, 1]$. Supongamos que existe una t-norma idempotente, es decir, $i(a, a) = a$ para cada $a \in [0, 1]$. Entonces para cada $a, b \in [0, 1]$,Si $a \leq b$, entonces:

$$a = i(a, a) \leq i(a, b) \leq i(a, 1) = a$$

Así, $i(a, b) = a = \text{mín}(a, b)$. Similarmente, si $a \geq b$, entonces:

$$b = i(b, b) \leq i(a, b) \leq i(1, b) = b$$

en consecuencia, $i(a, b) = b = \text{mín}(a, b)$. Por lo tanto, $i(a, b) = \text{mín}(a, b)$ para cada $a, b \in [0, 1]$

□

Ejemplo 2.4. Sean $a, b \in [0, 1]$. Los siguientes son ejemplos de:
1. Intersección estándar $i(a, b) = \text{mín}(a, b)$;
2. Producto algebraico $i(a, b) = ab$;
3. Diferencia acotada $i(a, b) = \text{mín}(0, a + b - 1)$;
4. Intersección drástica

$$i(a, b) = \begin{cases} a & \text{si } b = 1 \\ b & \text{si } a = 1 \\ 0 & \text{en otro caso} \end{cases}$$

Teorema 2.10. *Para cada $a, b \in [0, 1]$,*

$$i_{\text{mín}}(a, b) \leq i(a, b) \leq \text{mín}(a, b)$$

donde $i_{\text{mín}}$ denota la intersección drástica.

Demostración. Desigualdad superior. Por $Ax.i1$ y $Ax.i2$,

$$i(a,b) \leq i(a,1) = a$$

por $Ax.i3$,

$$i(a,b) = i(b,a) \leq i(b,1) = b$$

Así, $i(a,b) \leq a$ y $i(a,b) \leq b$; que es, $i(a,b) \leq \text{mín}(a,b)$.
Desigualdad inferior. Por $Ax.i1.$, $i(a,b) = a$ cuando $b = 1$, y $i(a,b) = b$ cuando $a = 1$. Como $i(a,b) \leq \text{mín}(a,b)$, claramente,

$$i(a,0) = i(0,b) = 0$$

Por $Ax.i2$,

$$i(a,b) \geq i(a,0) = i(0,b) = 0$$

Por lo tanto, $i_{\text{mín}}(a,b) \leq i(a,b)$

□

2.2.4. Uniones Difusas: t-conormas

Definición 2.11. *Una unión difusa (o t-conorma) es una función* $u : [0,1] \times [0,1] \to [0,1]$ *que cumple con las siguientes propiedades:*
Ax.u1. $u(a,0) = a$;
Ax.u2. Si $b \leq d$ *entonces* $u(a,b) \leq u(a,d)$;
Ax.u3. $u(a,b) = u(b,a)$;
Ax.u4. $u(a,u(b,d)) = u(u(a,b),d)$;
donde $a,b,d \in [0,1]$. *Así, para* A, B *conjuntos difusos:*

$$u[A,B](x) = u(A(x),B(x))$$

A partir de este momento nos referiremos a la unión difusa por t-conorma.

Observación. Comparando las propiedades de las t-normas con las t-conormas vemos que son las mismas, salvo la primera propiedad.
Los requerimientos adicionales más importantes para las t-conormas son:
$Ax.u5$. u es una función continua;
$Ax.u6$. $u(a,a) > a$;
$Ax.u7$. Si $a_1 < a_2$ y $b_1 < b_2$ entonces $u(a_1,b_1) < u(a_2,b_2)$.
Una t-conorma continua que cumple $Ax.u6$. se llama Arquimediana. Si además cumple $Ax.u7.$, es llamada Arquimediana estricta.

Teorema 2.12. *La unión difusa estándar es la única t-conorma idempotente*

Demostración. Análogo a la demostración del teorema 2.9.

□

Ejemplo 2.5. Sean $a, b \in [0, 1]$
1. Unión estándar $u(a, b) = \max(a, b)$;
2. Suma algebraica $u(a, b) = a + b - ab$;
3. Suma acotada $u(a, b) = \min(1, a + b)$;
4. Unión drástica

$$u(a,b) = \begin{cases} a & \text{si } b = 0 \\ b & \text{si } a = 0 \\ 0 & \text{en otro caso} \end{cases}$$

Teorema 2.13. *Para cada $a, b \in [0, 1]$,*

$$\max(a, b) \leq u(a, b) \leq u_{\max}(a, b)$$

donde $u_{\max}$ denota la unión drástica.

Demostración. Análogo a la demostración del teorema 2.10.

□

Para un estudio más profundo se le recomienda al lector revisar el libro [4] de la bibliografía.

2.3. Relaciones Difusas

Definición 2.14. *Sea $(X_n)_{n\in I}$ con I un conjunto de índices y X_n un conjunto para cada $n \in I$. Una relación difusa es una función $R : \prod_{n\in I} X_n \to [0, 1]$. donde $\prod$ denota el producto cartesiano.*

Ejemplo 2.6. Sea R una relación difusa entre los conjuntos $X = \{\text{Nueva York, Paris}\}$ y $Y = \{\text{Beijín, Nueva York, London}\}$, que representa el concepto "muy lejos". Esta relación puede ser escrita de la siguiente forma:

$$R(X, Y) = 1/\text{NY,Beijín} + 0/\text{NY,NY} + 0.6/\text{NY,London}+$$

$$0.9/\text{Paris,Beijín} + 0.7/\text{Paris,NY} + 0.3/\text{Paris,London}$$

Para simplificar las cosas tomaremos Relaciones difusas binarias,es decir, relaciones cuyo conjunto universal es el producto cartesiano de dos conjuntos.

Definición 2.15. *Sean P y Q relaciones difusas definidas sobre $X \times Y$ y $Y \times Z$, respectivamente.*
1. La composición estándar, denotado por $P \circ Q$, produce una relación sobre $X \times Z$ definida por:

$$R(x,z) = [P \circ Q](x,z) = \max_{y \in Y} \min(P(x,y), Q(y,z))$$

para cada $(x,z) \in X \times Z$. Otras definiciones de composiciones entre relaciones difusas son en la que mín *y* máx *se sustituyen por otras t-normas y t-conormas respectivamente.*
2. La unión relacional estándar, denotada por $P * Q$, es una relación sobre $X \times Y \times Z$ definida por:

$$R(x,y,z) = [P * Q](x,y,z) = \min(P(x,y), Q(y,z))$$

para cada $(x,y,z) \in X \times Y \times Z$. De nuevo se puede sustituir la operación mín *por cualquier otra t-conorma.*
3. La inversa de una relación binaria R sobre $X \times Y$, denotada por R^{-1} es una relación binaria sobre $Y \times X$ talque $R^{-1}(y,x) = R(x,y)$ para cada $(y,x) \in Y \times X$. Claramente, $(R^{-1})^{-1} = R$.

Las relaciones difusas binarias de elementos de un conjunto X que esten relacionados con otros elementos del mismo conjunto tienen especial significado y utilidad. Es decir, se trata de funciones de la forma:

$$R : X \times X \to [0,1]$$

Estas relaciones nos permiten representar equivalencias, semejanza, compativilidad, o preferencia. Según las propiedades de R, el grado $R(x,y)$ se puede interpretar como el grado en que x es equivalente a y, x es similar a y, x es compatible a y, o x es preferido que y.
Tres de los más impotantes tipos de relaciones binarias clásicas sobre un conjunto son la equivalencia, compatibilidad y relaciones de orden, que son caracterizadas en cuatro propiedades: reflexivas, simétricas, antisimétricas, y transitivas.

Definición 2.16. *Sea R una relación difusa sobre $X \times X$. Decimos que esta relación es:*
1. Reflexiva. Si $R(x,x) = 1$ para cada $x \in X$;

2. Simétrica. Si $R(x, y) = R(y, x)$ *para cada* $x, y \in X$*;*
3. Antisimétrica. Si $R(x, y) = 1$ *y* $R(y, x) = 1$ *implica* $x = y$ *para cada* $x, y \in X$*;*
4. Transitiva (con respecto a la t-norma i). Si $i(R(x, y), R(y, z)) \leq R(x, z)$ *para cada* $x, y, z \in X$*;*

Definición 2.17. *Sea* R *una relación difusa sobre* $X \times X$*. Decimos que esta relación es:*
1. de Equivalencia si es reflexiva, simétrica y transitiva;
2. de Compatibilidad si es reflexiva y simétrica;
3. un orden parcial si es reflexiva, antisimétrica y transitiva.

Como se mencionó anteriormente, es posible profundizar en esta Teoría de conjuntos difusos, pero no es el objetivo primordial de este trabajo. Sin embargo, si el lector está interesado se recomienda consultar [4]. Con esto concluimos una breve introducción a la Teoría de conjunto difusos.

Capítulo 3

Lógica Proposicional Clásica

El objetivo de este capítulo es el de dar una breve introducción a la lógica proposicional. Así, como es el caso de los capítulos anteriores, se puede realizar un estudio más profundo, pero para nuestro objetivo, lo expuesto es lo único que necesitamos.

3.1. Lenguaje formal

Definición 3.1. *Al conjunto $X_0 = \{\neg, \vee, (,), X_1, X_2, \cdots\}$ se le llamará el alfabeto de la lógica proposicional. Los elementos de X_0 son llamados símbolos. Los símbolos $X_1, X_2, \cdots$ son llamados variables proposicionales. Los símbolos $\neg, lor$ son llamados conectivos y los símbolos $(,)$ son llamados paréntesis.*

Definición 3.2. *Una cadena finita*

$$A \equiv \sigma_1\sigma_2 \cdots \sigma_m \qquad (m \geq 1)$$

de símbolos $\sigma_1\sigma_2 \cdots \sigma_m \in X_0$ es llamada una palabra formal. El número $\lambda(A) = m$ es llamada la longitud de la palabra A.

Ejemplo 3.1. Las cadenas:
$B_1 \equiv X_2(X_4\neg$
$B_2 \equiv \neg\neg\neg\neg$
$B_3 \equiv (X_3) \vee (X_1)$
$B_4 \equiv \neg(X_1) \vee (X_2)$
son palabras formales. Las longitudes de estas palabras son $\lambda(B_1) = 4$, $\lambda(B_2) = 4$, $\lambda(B_3) = 7$ y $\lambda(B_4) = 8$.

Definición 3.3. *(a) Las palabras formales*

$$A_1 \equiv \sigma_1\sigma_2\cdots\sigma_{m_1}, \quad A_2 \equiv \tau_1\tau_2\cdots\tau_{m_2}$$

son equivalentes si $m_1 = m_2$ *(*$\lambda(A_1) = \lambda(A_2) = m$*) y* $\sigma_i \equiv \tau_i$ *para cada* i *(*$1 \leq m$*). En tal caso, lo denotaremos por:*

$$A_1 \equiv A_2;$$

Si A_1, A_2 *no son equivalentes, escribiremos* $A_1 \not\equiv A_2$*.*
(b) Si $A \equiv \sigma_1\sigma_2\cdots\sigma_p$ *y* $B \equiv \tau_1\tau_2\cdots\tau_q$ *son palabras formales, entonces uno puede definir las palabras formales* $\neg(A)$*,* $(A) \vee (B)$ *por:*

$$\neg(A) \equiv \neg(\sigma_1\sigma_2\cdots\sigma_p);$$

$$(A) \vee (B) \equiv (\sigma_1\sigma_2\cdots\sigma_p) \vee (\tau_1\tau_2\cdots\tau_q).$$

(c) Para palabras formales arbitrarias $A \equiv \sigma_1\sigma_2\cdots\sigma_p$ *y* $B \equiv \tau_1\tau_2\cdots\tau_q$ *podemos definir las palabras formales* $(A) \wedge (B)$*,* $(A) \Rightarrow (B)$ *y* $(A) \Leftrightarrow (B)$ *de la siguiente forma:*

$$(A) \wedge (B) \equiv \neg((\neg(A)) \vee (\neg(B)));$$

$$(A) \Rightarrow (B) \equiv (\neg(A)) \vee (B);$$

$$(A) \Leftrightarrow (B) \equiv ((A) \Rightarrow (B)) \wedge ((B) \Rightarrow (A)).$$

Para mayor comodidad se omitirá el uso de paréntesis si no es posible la confusión.Por lo que, se usarán la notación $\neg A$, $A \vee B$, $A \Rightarrow B$, etc. en lugar de $\neg(A)$, $(A) \vee (B)$, $(A) \Rightarrow (B)$, etc.

Definición 3.4. *Una sucesión finita de palabras formales*

$$(f_c): \quad A_1, A_2, \cdots, A_r \ (r \geq 1)$$

es llamada una construcción formal o una construcción de fórmulas si para cada indice i *(*$1 \leq i \leq r$*) la palabra formal* A_i *en* (f_c) *satisface una de las tres condiciones siguientes:*
(a) A_i *coincide con una variable proposicional (*$A_i \equiv X$*);*
(b) existe algún indice j *(*$j < i$*) talque:*

$$A_i \equiv \neg(A_j);$$

(c) existen algunos indices h, k $(h, k < i)$ talque:

$$A_i \equiv (A_h) \vee (A_k).$$

Cada palabra formal que puede ser insertada en una construcción formal es llamada una fórmula. En particular, la palabra formal A_r en (f_c) es una fórmula y (f_c) es llamada una construcción formal de la fórmula A_r.
El conjunto de todas las fórmulas es denotado por $\mathcal{L}(X_0)$ o simplemente $\mathcal{L}_0$ y es llamado el lenguaje de la lógica proposicional.

3.2. Estructura de Verdad: Versión Semántica

En esta sección definiremos la interpretación Booleana de $\mathcal{L}_0$ y basados en las interpretaciones Booleanas definiremos las tautologías y las consecuencias semánticas.

3.2.1. Interpretaciones Booleanas

Sea $\Gamma = \{0, 1\}$ el álgebra Booleana que tiene solamente dos elementos, con las operaciones:

$$\bar{x} = 1 - x;$$

$$x \vee y = \text{máx}\{x, y\};$$

$$x \wedge y = \text{mín}\{x, y\}.$$

Suponiendo que Γ es un conjunto ordenado por el orden natural de números.

Definición 3.5. *Una función $\varphi : \mathcal{L} \to \Gamma$ es llamada una Interpretación Booleana de $\mathcal{L}_0$ si satisface las siguientes dos condiciones:*
(i) $\varphi(\neg(F)) = \overline{\varphi(F)}$;
(ii) $\varphi((F) \vee (G)) = \varphi(F) \wedge \varphi(G)$,
donde F, G son fórmulas arbitrarias en $\mathcal{L}_0$.

Observación. Note que en cada álgebra Booleana, en particular en Γ, uno puede probar las igualdades:
$\bar{\bar{x}} = x$;
$\overline{x \wedge y} = \bar{x} \vee \bar{y}$;
$\overline{x \vee y} = \bar{x} \wedge \bar{y}$;

se sigue que:
$\varphi((F) \wedge (G)) = \varphi(\neg(\neg F \vee \neg G)) = \overline{\varphi(\neg F \vee \neg G)} = \overline{\varphi(\neg F) \vee \varphi(\neg G)} =$

$$= \overline{\overline{\varphi(F)} \vee \overline{\varphi(G)}} = \overline{\overline{\varphi(F)}} \wedge \overline{\overline{\varphi(G)}} = \varphi(F) \wedge \varphi(G);$$

$$\varphi((F) \Rightarrow (G)) = \varphi(\neg F \vee G) = \varphi(\neg F) \vee \varphi(G) = \overline{\varphi(F)} \vee \varphi(G);$$

$$\varphi((F) \Leftrightarrow (G)) = \varphi((F \Rightarrow G) \wedge (G \Rightarrow F)) = \varphi(F \Rightarrow G) \wedge \varphi(G \Rightarrow F).$$

Por lo tanto, para las operaciones $\wedge$, $\Rightarrow$, $\Leftrightarrow$ en $\mathcal{L}$ tenemos:

$$\varphi(F \wedge G) = \varphi(F) \wedge \varphi(G);$$

$$\varphi(F \Rightarrow G) = \overline{\varphi(F)} \vee \varphi(G);$$

$$\varphi(F \Leftrightarrow G) = (\overline{\varphi(F)} \vee \varphi(G)) \wedge (\overline{\varphi(G)} \vee \varphi(F)).$$

Definición 3.6. *Si φ es una interpretación Booleana de $\mathcal{L}_0$, entonces la fórmula $F \in \mathcal{L}_0$ es llamada φ-verdadera si $\varphi(F) = 1$, y es llamada φ-falsa si $\varphi(F) = 0$.*
Por esta razón Γ es llamado el conjunto de valores de verdad para $\mathcal{L}_0$.

Definición 3.7.
(a) Una fórmula F que es φ-verdadera para cada interpretación φ es llamada universalmente verdadera o una tautología.
(b) Una fórmula F que es φ-falsa para cada interpretación φ es llamada universalmente falsa o una antilogía.

Para concluir esta sección demos otra definición.

Definición 3.8.
(a) Un conjunto $\mathcal{H}$ de fórmulas en $\mathcal{L}_0$ se dice que es semánticamente consistente o simplemente, consistente si existe una interpretación Booleana φ tal que:

$$\varphi(H) = 1 \quad \textit{para cada } H \in \mathcal{H};$$

La interpretación φ es llamada un modelo de(para) $\mathcal{H}$ y decimos que $\mathcal{H}$ tiene un modelo.
(b) Si para cada interpretación Booleana φ existe al menos una fórmula $H \in \mathcal{H}$ tal que $\varphi(H) = 0$, entonces el conjunto $\mathcal{H}$ es llamado semánticamente inconsistente, o simplemente, inconsistente; en otras palabras, $\mathcal{H}$ es inconsistente si $\mathcal{H}$ no tiene un modelo.

3.2.2. Consecuencias Semánticas

Sea $\mathcal{H}$ un subconjunto arbitrario de fórmulas en $\mathcal{L}_0$.

Definición 3.9. *Una fórmula $F \in \mathcal{L}_0$ será llamada una consecuencia semántica o una consecuencia lógica de $\mathcal{H}$ y se denotará por:*

$$\mathcal{H} \models F$$

Si F es φ-verdadera para cada interpretación Booleana φ talque para cada fórmula de $\mathcal{H}$ es φ-verdadera. En otras palabras, $\mathcal{H} \models F$ siempre que:

$$\varphi(H) = 1 \textit{ para cada } H \in \mathcal{H} \textit{ implica } \varphi(F) = 1.$$

En particular, F es una consecuencia semántica de $\emptyset$ y escribimos

$$\emptyset \models F.$$

Si F es φ-verdadera para cada interpretación Booleana φ. En otras palabras, $\emptyset \models F$ cumple que:

$$\varphi(F) = 1,$$

para cada φ interpretación Booleana.
Las fórmulas $H \in \mathcal{H}$ son llamadas hipótesis.

Observación.
(1) Si $\mathcal{H}$ es un conjunto finito no vacio,

$$\mathcal{H} = \{H_1, H_2, \cdots, H_p\},$$

entonces en lugar de $\mathcal{H} \models F$ escribiremos:

$$H_1, H_2, \cdots, H_p \models F$$

En particular, si $\mathcal{H}$ es reducida a una sola fórmula H, es decir, $\mathcal{H} = \{H\}$, entonces en lugar de $\mathcal{H} \models F$ escribiremos:

$$H \models F$$

La expresión $\mathcal{H} \cup \mathcal{K} \models F$ se denotará por:

$$\mathcal{H}, \mathcal{K} \models F.$$

Es claro que para cada dos subconjuntos $\mathcal{H}_1, \mathcal{H}_2$ de hipótesis con $\mathcal{H}_1 \subseteq \mathcal{H}_2$ uno deduce que

$$\text{Si } \mathcal{H}_1 \models F, \text{ entonces } \mathcal{H}_2 \models F;$$

En particular, si $\models F$, entonces para cada $\mathcal{H}$ se tiene que $\mathcal{H} \models F$.
(2) En lugar de $\emptyset \models F$ escribiremos $\models F$.
(3) El conjunto de todas las consecuencias semánticas de $\mathcal{H}$, será denotado por:

$$\mathcal{C}^{sem}(\mathcal{H}) = \{F \in \mathcal{L}_0 \mid \mathcal{H} \models F\}$$

Si $\mathscr{T}$ es el conjunto de todas las tautologías, entonces es claro que

$$\mathcal{C}^{sem}(\emptyset) = \mathscr{T}.$$

Teorema 3.10. *Sea $\mathcal{H}$ un conjunto no vacio de hipótesis $\mathcal{H} = \{H_1, H_2, \cdots, H_p\}$ y definamos $H \equiv H_1 \wedge H_2 \wedge \cdots \wedge H_p$. Una fórmula $F \in \mathcal{L}_0$ es consecuencia semántica de $\mathcal{H}$ si y sólo si es consecuencia semántica de H; en otras palabras,*

$$\mathcal{H} \models F \textit{ si y sólo si } H \models F.$$

Demostración. Supongamos que $\mathcal{H} \models F$. esto quiere decir que $\varphi(F) = 1$ para cada φ interpretación Booleana tal que $\varphi(H_1) = \varphi(H_2) = \cdots = \varphi(H_p) = 1$. Ahora consideremos una interpretación Booleana φ talque $\varphi(H) = 1$. Se sigue que $1 = \varphi(H) = \varphi(H_1) \wedge \varphi(H_2) \wedge \cdots \wedge \varphi(H_p)$, tomando en cuanta que "$\wedge$" es la operación mín tenemos que $\varphi(H_1) = \varphi(H_2) = \cdots = \varphi(H_p) = 1$, entonces $\varphi(F) = 1$ (porque $\mathcal{H} \models F$). Por lo tanto, $H \models F$.
Ahora supongamos que $H \models F$ y tomemos una interpretación Booleana φ talque $\varphi(H_1) = \varphi(H_2) = \cdots = \varphi(H_p) = 1$. Tenemos que $\varphi(H) = \varphi(H_1) \wedge \varphi(H_2) \wedge \cdots \wedge \varphi(H_p) = 1$, como $H \models F$ entonces $\varphi(F) = 1$. Por lo tanto $\mathcal{H} \models F$.

□

Definición 3.11. *(1) Al conjunto de todas las interpretaciones Booleanas de $\mathcal{L}_0$ se le denotará por $\mathcal{V}$.*
(2) Sea $\varphi \in \mathcal{V}$. Definimos U_φ el conjunto de todas las fórmulas F en $\mathcal{L}_0$ talque $\varphi(H) \leq \varphi(F)$, es decir,

$$U_\varphi = \{F \in \mathcal{L}_0 \mid \varphi(H) \leq \varphi(F)\}.$$

De la definición de $\mathcal{C}^{sem}(H)$ tenemos que

$$\mathcal{C}^{sem} = \bigcap_{\varphi \in \mathcal{V}} U_\varphi.$$

Definición 3.12. *Un subconjunto U de $\mathcal{L}_0$ se dice que es cerrado bajo la MP-regla o MP-cerrado si*

$$A \in U \text{ y } A \Rightarrow B \in U \text{ implica } B \in U.$$

El conunto de todos los subconjuntos de $\mathcal{L}_0$ que tenga las propiedades:
(i) U es MP-cerrado;
(ii) U contiene todas las tautologías ($\mathscr{T} \subseteq U$);
(iii) $H \in U$.
se denotará por $\mathcal{U}_{\mathcal{H}}$.

Teorema 3.13.
(a) Si $\mathcal{S}$ es un subconjunto de $\mathcal{U}_{\mathcal{H}}$, entonces $\bigcap_{U \in \mathcal{S}} U \in \mathcal{U}_{\mathcal{H}}$;
(b) $U_\varphi \in \mathcal{U}_{\mathcal{H}}$ para cada $\varphi \in \mathcal{V}$;

Demostración. La afirmación (a) es inmediata.
(b)Sea $\varphi \in \mathcal{V}$ arbitraria. Primero veamos que U_φ es MP-cerrado. Sean A y $A \Rightarrow B$ fórmulas en U_φ. Entonces

$$\varphi(H) \leq \varphi(A) \text{ y } \varphi(H) \leq \varphi(A \Rightarrow B).$$

Debemos probar que $\varphi(H) \leq \varphi(B)$. De las desigualdades anteriores tenemos:

$$\varphi(H) \leq \varphi(A) \wedge \varphi(A \Rightarrow B) = \varphi(A) \wedge (\overline{\varphi(A)} \vee \varphi(B)) =$$

$$= (\varphi(A) \wedge \overline{\varphi(A)}) \vee (\varphi(A) \wedge \varphi(B)) = \varphi(A) \wedge \varphi(B) \leq \varphi(B),$$

Tenemos que $B \in U_\varphi$. Así, U_φ es MP-cerrado.
Si $F \in \mathscr{T}$, entonces $\varphi(F) = 1 \geq \varphi(H)$, por tanto $F \in U_\varphi$, es decir, $\mathscr{T} \subseteq U_\varphi$.
Además $H \in U_\varphi$, por lo tanto $U_\varphi \in \mathcal{U}_{\mathcal{H}}$.

□

También existe la versión sintáctica para la lógica proposicional, así como la versión más extensa que es la lógica de predicados.

Capítulo 4

Lógica Difusa

4.1. Lógica Proposicional Difusa

En el capítulo anterior se estudió el conjunto $\Gamma = \{0,1\}$ junto con las operaciones

$$B - ops : \begin{cases} \bar{x} = 1 - x; \\ x \vee y = \text{máx}\{x, y\}; \\ x \wedge y = \text{mín}\{x, y\}; \end{cases}$$

y se verificó que se ontiene con ello el álgebra Booleana con dos elementos.
Si en lugar de Γ, consideramos el conjunto $[0,1]$, entonces este conjunto no es un álgebra Booleana con las $B - ops$, ya que no se cumplen las igualdades $x \vee \bar{x} = 1$ y $x \wedge \bar{x} = 0$; esta nueva estructura algebraica es llamada por algunos autores una álgebra De Morgan.
Sin embargo, es posible dotar al conjunto $[0,1]$ de una estructura de MV-álgebra al definir las siguientes operaciones:

$$MV - ops : \begin{cases} \bar{x} = 1 - x; \\ x \oplus y = \text{mín}\{1, x + y\}; \\ x \odot y = \text{máx}\{0, x + y - 1\}; \end{cases}$$

Asimismo podemos establecer una operación "$\rightarrow$" (llamada residuación) definida por

$$x \rightarrow y = \bar{x} \oplus y$$

Esto es,

$$x \rightarrow y = \text{mín}\{1, 1 - x + y\}$$

Definición 4.1. *Sean A, B conjuntos difusos,decimos que A está contenido en B, que denotaremos por $A \sqsubseteq B$, si $A(x) \leq B(x)$ para cada $x \in [0,1]$.*

El lenguaje de la lógica proposicional difusa es en esencia el mismo lenguaje que el de la lógica proposicional clásica. Por tanto, en la construcción del lenguaje formal $\mathcal{L}_0$ de la lógica proposicional difusa empezamos con el alfabeto

$$\{\neg, \Rightarrow, (,), X_1, X_2, \cdots\}.$$

La construcción formal y las fórmulas en $\mathcal{L}_0$ se definen del mismo modo que en el capítulo anterior cambiando $\Rightarrow$ en lugar de $\vee$. El conjunto de todas las fórmulas se llamará el lenguaje $\mathcal{L}_0$ de la lógica proposicional difusa. En $\mathcal{L}_0$ introducimos las siguientes abreviaciones:

$$F \wedge G \equiv \neg((F \Rightarrow G) \Rightarrow \neg F);$$

$$F \vee G \equiv (F \Rightarrow G) \Rightarrow G;$$

$$F \& G \equiv \neg(F \Rightarrow \neg G);$$

$$F \bigtriangledown G \equiv \neg(\neg F \& \neg G);$$

$$F \Leftrightarrow G \equiv (F \Rightarrow G) \wedge (G \Rightarrow F);$$

Definición 4.2. *Una MV-interpretación de $\mathcal{L}_0$ una función $\varphi : \mathcal{L}_0 \to [0,1]$ que cumpla las propiedades:*
(i) $\varphi(\neg F) = \overline{\varphi(F)}$;
(ii) $\varphi(F \Rightarrow G) = \varphi(F) \to \varphi(G)$,
donde $F, G \in \mathcal{L}_0$ arbitrarias y "$\to$" es la operación residuación en $[0,1]$. El conjunto de todas las MV-interpratationes de $\mathcal{L}_0$ se denotará por $\mathcal{W}$.

Observación. Se pueden verificar las siguientes igualdades:

$$\varphi(F \wedge G) = \varphi(F) \wedge \varphi(G);$$

$$\varphi(F \vee G) = \varphi(F) \vee \varphi(G);$$

$$\varphi(F \& G) = \varphi(F) \odot \varphi(G);$$

$$\varphi(F \bigtriangledown G) = \varphi(F) \oplus \varphi(G).$$

Para verificar las igualdades denotaremos por $x = \varphi(F)$, $y = \varphi(G)$, así,
(1)

$$\varphi(F \wedge G) = \varphi(\neg((F \Rightarrow G) \Rightarrow \neg F)) = \overline{\varphi((F \Rightarrow G) \Rightarrow \neg F)} = \overline{\varphi(F \Rightarrow G) \to \varphi(\neg F)} =$$

$= \overline{\overline{(x \to y) \to \overline{x}}} = \overline{\overline{x \to y} \oplus \overline{x}} = (x \to y) \odot x = (\bar{x} \oplus y) \odot x = y \wedge x = x \wedge y = \varphi(F) \wedge \varphi(G)$

(2)

$$\varphi(F \vee G) = \varphi((F \Rightarrow G) \Rightarrow G) = (x \to y) \to y = \overline{x \to y} \oplus y =$$

$$\overline{\overline{x} \oplus y} \oplus y = (x \odot \bar{y}) \oplus y = x \vee y = \varphi(F) \vee \varphi(G)$$

(3)

$$\varphi(F \& G) = \varphi(\neg(F \Rightarrow \neg G)) = \overline{\varphi(F \Rightarrow \neg G)} =$$

$$\overline{x \to \overline{y}} = \overline{\overline{x} \oplus \overline{y}} = x \odot y = \varphi(F) \odot \varphi(G)$$

(4)

$$\varphi(F \bigtriangledown G) = \varphi(\neg(\neg F \& \neg G)) =$$

$$\overline{\varphi(\neg F \& \neg G)} = \overline{\overline{x} \odot \overline{y}} = x \oplus y.$$

Por lo tanto, $\mathcal{V} \subseteq \mathcal{W}$, donde $\mathcal{V}$ es el conjunto de todas las interpretaciones Booleanas.

Definición 4.3. *Sean $\varphi \in \mathcal{W}$ y $F \in \mathcal{L}_0$. Entonce, el elemento $\varphi(F) \in [0,1]$ es llamado el valor de verdad de F bajo la interpreción φ.*

Definición 4.4. *Sea $\mathcal{H}$ un conjunto difuso (de hipótesis). El conjunto $\mathcal{C}^{sem}(\mathcal{H})$ de todas las consecuencias semánticas difusas de $\mathcal{H}$ es definido por las igualdades*

$$\mathcal{C}^{sem}(\mathcal{H}) = \bigcap_{\substack{\varphi \in \mathcal{W} \\ \varphi \sqsupseteq \mathcal{H}}} \varphi = \bigwedge_{\substack{\varphi \in \mathcal{W} \\ \varphi \geq \mathcal{H}}} \varphi$$

Por lo tanto,

$$\mathcal{C}^{sem}(\mathcal{H})(F) = \bigwedge_{\substack{\varphi \in \mathcal{W} \\ \varphi \geq \mathcal{H}}} \varphi(F).$$

Para una fórmula dada F escribimos

$$\mathcal{H} \models_a F$$

si $\mathcal{C}^{sem}(\mathcal{H})(F) = a$; en este caso, decimos que F es una consecuencia semántica difusa desde $\mathcal{H}$ en grado a. En particular, si $\mathcal{H} \simeq \emptyset$ entonces

$$\mathcal{C}^{sem}(\emptyset) = \bigcap_{\varphi \in \mathcal{W}} \varphi = \bigwedge_{\varphi \in \mathcal{W}} \varphi$$

y

$$\mathcal{C}^{sem}(\emptyset)(F) = \bigwedge_{\varphi\in\mathcal{W}} \varphi(F).$$

En este caso la notación $\models_a F$ quiere decir que $\mathcal{C}^{sem}(\emptyset)(F) = a$ y F es llamada una a-tautología. Si $a = 1$, entonces en lugar de $\models_1 F$ escribimos simplemente $\models F$ y F se llamará tautología difusa.

Teorema 4.5. *Si una fórmula F en $\mathcal{L}_0$ es tautología difusa, entonces F es una tautología en el sentido clásico.*

Demostración. Sea F una tautología difusa. Tomando en cuenta que $\mathcal{V} \subseteq \mathcal{W}$ tenemos que

$$1 = \mathcal{C}^{sem}(\emptyset)(F) = \bigwedge_{\varphi\in\mathcal{W}} \leq \bigwedge_{\varphi\mathcal{V}}.$$

Se sigue que $\varphi(F) = 1$ para cada $\varphi \in \mathcal{V}$. Es decir, que F es una tautología clásica.

□

Ejemplo 4.1. Es posible demostrar que las fórmulas
$C1 \equiv F \Rightarrow (G \Rightarrow F)$;
$C2 \equiv (F \Rightarrow G) \Rightarrow ((G \Rightarrow H) \Rightarrow (F \Rightarrow H))$;
$C3 \equiv (\neg G \Rightarrow \neg F) \Rightarrow (F \Rightarrow G)$;
$C4 \equiv ((F \Rightarrow G) \Rightarrow G) \Rightarrow ((G \Rightarrow F) \Rightarrow F)$;
son tautologías difusas; por lo tanto, son tautologías clásicas.
Convendría probar que $C1$ es tautología.
Tomemos $x = \varphi(F)$, $y = \varphi(G)$, entonces:
$\varphi(C1) = \varphi(F \Rightarrow (G \Rightarrow F)) = x \to (y \to x) = \bar{x} \oplus (y \to x) = \bar{x} \oplus (\bar{y} \oplus x) = \bar{x} \oplus (x \oplus \bar{y}) = (\bar{x} \oplus x) \oplus \bar{y} = 1 \oplus \bar{y} = 1$
Por lo tanto $C1$ es una tautología difusa.
Por otro lado, es posible demostrar

$$T \equiv (X \Rightarrow \neg X) \Rightarrow \neg X$$

donde X es una variable proposicional, no es una tautología difusa, pero si es una tautología clásica.

Sea $\varphi \in \mathcal{W}$ y denotemos $x = \varphi(X)$, entonces:
$\varphi(T) = (x \to \bar{x}) \to \bar{x} = \text{mín}\{1, 1 - (x \to \bar{x}) + \bar{x}\} = \text{mín}\{1, 1 - \text{mín}\{1, 1 -$

$x + (1 - x)\} + (1 - x)\} = \text{mín}\{1, 2 - x - \text{mín}\{1, 2 - 2x\}\}$.
Ahora consideremos una MV-interpretación φ talque $x = \varphi(X) = \frac{1}{2}$,

$$\varphi(T) = \text{mín}\{1, 2 - \frac{1}{2} - \text{mín}\{1, 1\}\} = \text{mín}\{1, 1 - \frac{1}{2}\} = \frac{1}{2}$$

Por lo tanto, T no es una tautología difusa, pero si es una tautología clásica, si x toma los valores 0 o 1 se verifica que $\varphi(T) = 1$ en cualquiera de los dos casos.

Observación. El último resultado muestra que $\mathscr{T}$ el conjunto de todas las tautologías clásicas es estrictamente más grande que el conjunto $\mathscr{T}_f$ de tautologías difusas, es decir, $\mathscr{T}_f \subseteq \mathscr{T}$ y $\mathscr{T}_f \neq \mathscr{T}$.

Teorema 4.6. *Si $\mathcal{H}, \mathcal{K}$ conjuntos difusos (de hipótesis) con $\mathcal{H} \sqsubseteq \mathcal{K}$, entonces*

$$\mathcal{C}^{sem}(\mathcal{H}) \sqsubseteq \mathcal{C}^{sem}(\mathcal{K})$$

o

$$\mathcal{C}^{sem}(\mathcal{H})(F) \leq \mathcal{C}^{sem}(\mathcal{K})(F)$$

para cada $F \in \mathcal{L}_0$. Es decir, si $\mathcal{H} \models_a F$, entonces para cada $\mathcal{K}$ tal que $\mathcal{H} \sqsubseteq \mathcal{K}$ tenemos $\mathcal{K} \models_b F$ para algún $b \geq a$.

Demostración. Como $\{\varphi \in \mathcal{W} \mid \varphi \geq \mathcal{H}\} \supseteq \{\varphi \in \mathcal{W} \mid \varphi \geq \mathcal{K}\}$, entonces por definición tenemos que

$$\mathcal{C}^{sem}(\mathcal{H})(F) = \bigwedge_{\substack{\varphi \in \mathcal{W} \\ \varphi \geq \mathcal{H}}} \varphi(F) \leq \bigwedge_{\substack{\varphi \in \mathcal{W} \\ \varphi \geq \mathcal{K}}} \varphi(F) = \mathcal{C}^{sem}(\mathcal{K})(F)$$

□

Teorema 4.7. *Una fórmula F es una tautología difusa si y sólo si $\varphi(F) = 1$ para toda $\varphi \in \mathcal{W}$.*

Demostración. Por definición $\models F$ si y sólo si $\mathcal{C}^{sem}(\emptyset)(F) = 1$ y teniendo en cuanta que $\mathcal{C}^{sem}(\emptyset)(F) = \bigwedge_{\varphi \in \mathcal{W}} \varphi(F)$, se sigue que $\mathcal{C}^{sem}(\emptyset)(F) = 1$ si y sólo si $\varphi(F) = 1$ para cada $\varphi \in \mathcal{W}$

□

Teorema 4.8. *Si $\models G$, entonces $\models F \Rightarrow G$ para cada fórmula F.*

Demostración. Sea $\varphi \in \mathcal{W}$. Por el teorema anterior $\varphi(G) = 1$, así
$\varphi(F \Rightarrow G) = \varphi(F) \to 1 = 1$ (porque, para cada $x \in [0, 1]$, $x \to 1 = 1$)

□

Note que la consecuencia semántica difusa preserva Modus Ponens.

Teorema 4.9. *Si* $\models F$ *y* $\models F \Rightarrow G$*, entonces* $\models G$

Demostración. Tenemos $\varphi(F) = 1$ y $\varphi(F \Rightarrow G) = 1$ para cada $\varphi \in \mathcal{W}$, así

$$1 = \varphi(F \Rightarrow G) = \varphi(F) \to \varphi(G) = 1 \to \varphi(G)$$

se sigue que $\varphi(G) = 1$, porque $1 \to z = 1$ si y sólo si $z = 1$.

□

Para concluir esta sección conviene hacer la siguiente observación. Las fórmulas C1-C4 que se presentan en el ejemplo 4.7 permiten deducir sintácticamente cualquier teorema en Lógica Difusa. Esto es, a partir de estos resultados y sus consecuencias lógicas es posible determinar todas las fórmulas válidas.

Análogo al caso de la Lógica Clásica, la Lógica Difusa satisface varios de los resultados principales necesarios para desarrollar adecuadamente una Teoría Formal, por ejemplo el teorema de la deducción , el teorema de completitud, la consistencia, la compacidad, etc.

4.2. Lógica Difusa en el sentido amplio

Dada una fórmula F, denotaremos por $||F||$ a su grado de verdad.
El grado de verdad de una composición de fórmulas está dada por el grado de verdad de las fórmulas que la componen.

Ejemplo 4.2. Sean $F, G \in \mathcal{L}_0$, entonces

$$||F \& G|| = i(||F||, ||G||)$$

donde i es una función adecuada, llamada una función de verdad de la conjunción. Del mismo modo

$$||F \Rightarrow G|| = r(||F||, ||G||)$$

$$||\neg F|| = n(||F||)$$

respectivas funciones de la implicación y negación.
A la función i en general se le pide que sea continua (o al menos continua por la izquierda) y t-norma.
La condición de continuidad se pide para que si el grado de verdad de las fórmulas F_1 y G_1 se aproximan a los de F y G, respectivamente, entonces el grado de verdad de $F_1 \& G_1$ se aproximé al grado de verdad de $F \& G$.

Observación. Una relación entre la conjunción y la implicación es:
Si a, b, c son valores de verdad , tenemos

$$i(a,b) \leq c \text{ si y sólo si } a \leq r(b,c)$$

Entonces dada una t-norma continua i, existe una única función r tal que i y r cumplen la desigualdad anterior, la cual se denomina el residuo de i y está dada por

$$r(a,b) = \text{máx}\{c \in [0,1] \mid i(a,c) \leq b\}$$

La negación viene dada por

$$n(a) = r(a,0)$$

Ejemplo 4.3. Las operaciones de Lukasiewicz se definen como

$$i(a,b) = \text{máx}(0, a+b-1)$$

$$r(a,b) = \text{mín}(1, 1-a+b)$$

$$n(a) = 1-a$$

Las operaciones de Goguen vienen dadas por

$$i(a,b) = ab$$

$$r(a,b) = \begin{cases} 1 & \text{si } a \leq b \\ \frac{b}{a} & \text{en otro caso} \end{cases}$$

$$n(a) = \begin{cases} 1 & \text{si } a = 0 \\ 0 & \text{si } a > 0 \end{cases}$$

y las operaciones de Gödel son dadas por

$$i(a,b) = \text{mín}(a,b)$$

$$r(a,b) = \begin{cases} 1 & \text{si } a \leq b \\ b & \text{en otro caso} \end{cases}$$

$$n(a) = \begin{cases} 1 & \text{si } a = 0 \\ 0 & \text{si } a > 0 \end{cases}$$

Cuando los conjuntos difusos representan conceptos ling*ü*ísticos, como son "muy peque*ñ*o", "peque*ñ*o", medio,etc., interpretados en un contexto en particular, la contrucción resultante es llamada usualmente variable ling*ü*ística.

Definición 4.10. *Una variable lingüistica es un sistema* (v, T, X, g, m) *donde* v *es el nombre de la variable base,* T *es el conjunto de términos lingüisticos de* v*,* X *es el universo de discurso de la variable base* v*,* g *es una regla sintáctica para generar términos lingüisticos y* m *es una regla semántica que asigna a cada término lingüistico* t *su significado* $m(t)$*, que es un conjunto difuso en* X*.*

4.2.1. Proposiciones Difusas

En esta sección nos enfocaremos en proposiciones difusas simples, que se pueden clasificar en cuatro:
1. Proposiciones no condicionales y no calificadas.
2. Proposiciones no condicionales y calificadas.
3. Proposiciones condicionales y no calificadas.
4. Proposiciones condicionales y calificadas.
Para cada tipo introducimos la forma canónica y discutiremos su interpretación.

Proposiciones no condicionales y no calificadas

La forma canónica de proposiciones difusas de este tipo, p, es expresada por la oración

$$p: \ \mathcal{V} \text{ es } F$$

donde $\mathcal{V}$ es una variable que toma valores v en un conjunto universal V y F es un conjunto difuso en V que representa un predicado, como delgado, caro, bajo, normal, etc.
Dado un valor de $\mathcal{V}$, por ejemplo v, este valor pertenece a F con grado $F(v)$.

El grado de pertenencia es entonces interpretado como el grado de verdad, $T(p)$, de la proposición p. Es decir,

$$T(p) = F(p)$$

para cada valor v de la variable $\mathcal{V}$ en la proposición p.
En algunas proposiciones los valores de la variable $\mathcal{V}$ son asignados a cada elemento de un conjunto I. Es decir, la variable $\mathcal{V}$ es una función $\mathcal{V} : I \to V$, donde $\mathcal{V}(i) \in V$ es el valor para cada elemento $i \in I$. La forma canónica ahora es modificada a la forma:

$$p: \ \mathcal{V}(i) \text{ es } F,$$

donde $i \in I$.

Ejemplo 4.4. Sea I un conjunto de personas, cada una de las cuales está caracterizada por su edad, y se F un conjunto difuso que exprese el predicado "joven". Así, nuestra proposición

$$p: \ \mathcal{V}(i) \text{ es } F,$$

significa que la persona i con edad $\mathcal{V}(i)$ es joven, donde el grado de verdad de esta proposición, $T(p)$, está determinado por la ecuación

$$T(p) = F(\mathcal{V}(i))$$

Proposiciones no condicionales y calificadas

Proposiciones p de este tipo son caracterizadas por la forma canónica

$$p: \ \mathcal{V} \text{ es } F \text{ es } S,$$

o la forma canónica

$$p: \ Pro\{\mathcal{V} \text{ es } F\} \text{ es } P,$$

donde $\mathcal{V}$ y F son de la misma manera, $Pro\{\mathcal{V} \text{ es } F\}$ es la probabilidad del evento difuso "$\mathcal{V}$ es F ", S es un conjunto difuso que representa un calificador de verdad, como falso, muy verdadero, más o menos falso, etc., y P es un conjunto difuso que representa un calificador de probabilidad, como probable, muy probable, extremadamente improbable, etc.
En términos generales, las proposiciones del primer tipo son proposiciones de verdad calificada y las del segundo tipo son proposiciones de probabilidad

calificada.
En general, el grado de verdad, $T(p)$, de una proposición de verdad calificada p es dada para cada $v \in V$ por la igualdad

$$T(p) = S(F(v))$$

Ejemplo 4.5. La proposición "Tina es joven es muy cierto" es del primer tipo, donde el predicado "joven" y el calificador de "verdad muy cierto" son representados por los conjuntos difusos F y S, respectivamente.
Supongamos que la edad de Tina es de 26 años, entonces el grado de pertenencia al conjunto difuso F es de 0.87. Por lo tanto, el grado de pertenencia al conjunto difuso S es 0.76.
Así, el grado de verdad de la proposición "Tina es joven es muy cierto" es de 0.76.

Ahora discutamos sobre la proposiciones de probabilidad calificada. Para una función f de distribución de probabilidad sobre $\mathcal{V}$, tenemos

$$Pro\{\mathcal{V} \text{ es } F\} = \sum_{v \in V} f(v)F(v);$$

entonces el grado de verdad, $T(p)$, de la proposición p es dada por

$$T(p) = P(\sum_{v \in V} f(v)F(v))$$

Proposiciones condicionales y no calificadas

Las proposiciones p de este tipo son expresadas por la forma canónica

$$p: \text{ Si } \mathfrak{X} \text{ es } A, \text{ entonces } \mathfrak{Y} \text{ es } B,$$

donde $\mathfrak{X}, \mathfrak{Y}$ son variables cuyos valores estan en los conjuntos X, Y, respectivamente, y A, B son conjuntos difusos sobre X, Y, respectivamente. Esta proposición puede ser vista de la forma

$$\langle \mathfrak{X}, \mathfrak{Y} \rangle \in R$$

donde R es un conjunto difuso sobre $X \times Y$ que está determinado para cada $x \in X$ y cada $y \in Y$ por la fórmula

$$R(x, y) = \mathfrak{J}[A(x), B(y)]$$

donde $\mathfrak{J}$ es una operación binaria sobre $[0,1]$ representando una conveniente implicación difusa.

Ejemplo 4.6. Tomemos una implicación difusa particular, la implicación de Lukasiewicz

$$\mathfrak{J}(a,b) = \text{mín}(1, 1-a+b)$$

Sean $A = .1/x_1 + .8/x_2 + 1/x_3$ y $B = .5/y_1 + 1/y_2$. Entonces

$$R = 1/x_1, y_1 + 1/x_1, y_2 + .7/x_2, y_1 + 1/x_2, y_2 + .5/x_3, y_1 + 1/x_3, y_2$$

Así, $T(p) = 1$ cuando $\mathfrak{X} = x_1$ y $\mathfrak{Y} = y_1$; $T(p) = .7$ cuando $\mathfrak{X} = x_2$ y $\mathfrak{Y} = y_1$, etc.

Proposiciones condicionales y calificadas

Estas proposiciones están determinadas por la forma canónica

$$p: \text{ Si } \mathfrak{X} \text{ es } A, \text{ entonces } \mathfrak{Y} \text{ es } B \text{ es } S.$$

Una forma alternativa para este tipo de proposiciones es la determinada por la fórmula canónica

$$p: \ Pro\{\mathfrak{X} \text{ es } A \mid \mathfrak{Y} \text{ es } B\} \text{ es } P,$$

donde $Pro\{\mathfrak{X} \text{ es } A \mid \mathfrak{Y} \text{ es } B\}$ es una probabilidad condicional.

4.2.2. Inferencia a partir de Proposiciones Difusas

Inferencia a partir de Proposiciones condicionales

En esta parte describiremos generalizaciones de tres reglas de inferencia clásicas, modus ponens, modus tollens y silogismo hipotético. Estas generalizaciones están basadas en lo que llamaremos reglas de inferencia composicional.

Consideremos variables $\mathfrak{X}, \mathfrak{Y}$ que toman valores en los conjuntos X, Y, respectivamente, y supongamos que para cada $x \in X$ y $y \in Y$ las variables estan relacionadas por una función $y = f(x)$. Entonces, dado $\mathfrak{X} = x$, podemos inferir que $\mathfrak{Y} = f(x)$. Similarmente, si sabemos que los valores de $\mathfrak{X}$ están en un conjunto A dado, podemos inferir que el valor de $\mathfrak{Y}$ está en el conjunto $B = \{y \in Y | \, y = f(x), x \in A\}$. Ahora supongamos que las variables estan

relacionadas por una relación arbitraria sobre $X \times Y$, no necesariamente una función. Entonces, dados $\mathfrak{X} = u$ y una relación R, nosotros podemos inferir que $\mathfrak{Y} \in B$, donde $B = \{y \in Y | \langle x, y\rangle \in R\}$.
Observe que esta inferencia puede ser expresada en términos de funciones caracteristicas χ_A, χ_B, χ_R de conjuntos A, B, R, respectivamente, junto con la igualdad

$$\chi_B(y) = \sup_{x \in X} \text{mín}[\chi_A(x), R(x, y)]$$

para cada $y \in Y$.
Ahora procedamos al siguiente paso y supongamos que R es un conjunto difuso sobre $X \times Y$, y A', B' son conjuntos difusos sobre X y Y, respectivamente. Entonces si R y A' son dadas, podemos obtener B' por la igualdad

$$B'(y) = \sup_{x \in X} \text{mín}[A'(x), R(x, y)]$$

para cada $y \in Y$. Esta igualdad puede ser escrita en forma de composición de conjuntos difusos por

$$B' = A' \circ R$$

es llamada la regla de inferencia composicional.

Definición 4.11. *Una proposición difusa condicional p de la forma*

$$p: \ \textit{Si } \mathfrak{X} \textit{ es } A, \ \textit{entonces } \mathfrak{Y} \textit{ es } B,$$

es determinada para cada $x \in X$ y $y \in Y$ por la igualdad

$$R(x, y) = \mathfrak{J}[A(x), B(y)]$$

donde $\mathfrak{J}$ denota una implicación difusa. Dada una proposición difusa q de la forma

$$q: \ \mathfrak{X} \textit{ es } A'$$

nosotros concluimos que $\mathfrak{Y}$ es B' por la regla de inferencia composicional. Este procedimiento es llamado un modus ponens generalizado.

Observación. Considerando a la proposición p como una regla y la proposición q un hecho, el modus ponens generalizado es expresado por el siguiente esquema

$$\begin{array}{ll} \text{Regla :} & \text{Si } \mathfrak{X} \text{ es } A, \text{ entonces } \mathfrak{Y} \text{ es } B \\ \text{Hecho :} & \mathfrak{X} \text{ es } A' \\ \hline \text{Conclusión :} & \mathfrak{Y} \text{ es } B' \end{array}$$

Ejemplo 4.7. Sean los conjuntos $X = \{x_1, x_2, x_3\}$ y $Y = \{y_1, y_2\}$ de valores de las variables $\mathfrak{X}$ y $\mathfrak{Y}$, respectivamente. Sopongamos que la proposición "Si $\mathfrak{X}$ es A, entonces $\mathfrak{Y}$ es B" es dada, donde $A = .5/x_1 + 1/x_2 + .6/x_3$ y $B = 1/y_1 + .4/y_2$. Entonces dado el hecho expresado por la proposición "$\mathfrak{X}$ es A'", donde $A = .6/x_1 + .9/x_2 + .7/x_3$, encontremos usando el modus ponens generalizado inferir una conclusión en la forma " $\mathfrak{Y}$ es B'".
Usando, por ejempo, la implicación de Lukasiewicz, tenemos

$$R = 1/x_1, y_1 + .9/x_1, y_2 + 1/x_2, y_1 + .4/x_2, y_2 + 1/x_3, y_1 + .8/x_3, y_2$$

Entonces, por la regla de inferencia composicional, tenemos

$$B'(y_1) = \sup_{x \in X} \text{mín}[A'(x), R(x, y_1)]$$

$$= \text{máx}[\text{mín}(.6, 1), \text{mín}(.9, 1), \text{mín}(.7, 1)] = .9$$

$$B'(y_2) = \sup_{x \in X} \text{mín}[A'(x), R(x, y_2)]$$

$$= \text{máx}[\text{mín}(.6, .9), \text{mín}(.9, .4), \text{mín}(.7, .8)] = .7$$

Entonces concluimos que $\mathfrak{Y}$ es B', donde $B' = .9/y_1 + .7/y_2$

Definición 4.12. *La regla de inferencia Modus Tollens Generalizado es aquella determinada por el siguiente esquema*

$$\begin{array}{ll} \textit{Regla :} & \textit{Si } \mathfrak{X} \textit{ es } A, \textit{ entonces } \mathfrak{Y} \textit{ es } B \\ \textit{Hecho :} & \mathfrak{Y} \textit{ es } B' \\ \hline \textit{Conclusión :} & \mathfrak{X} \textit{ es } A' \end{array}$$

en este caso, la regla de inferencia composicional tiene la forma

$$A'(x) = \sup_{y \in Y} \text{mín}[B'(y), R(x, y)]$$

Ejemplo 4.8. Sean $X, Y, A, \mathfrak{J}$ y B como en el ejemplo anterior. Entonces R es igual al del ejemplo. Ahora supongamos que un hecho expresado por la proposición "$\mathfrak{Y}$ es B'" es dado, donde $B' = .9/y_1 + .7/y_2$. Entonces

$$A'(x_1) = \sup_{y \in Y} \text{mín}[B'(y), R(x_1, y)]$$

$$= \text{máx}[\text{mín}(.9, 1), \text{mín}(.7, .9)] = .9$$
$$A'(x_2) = \sup_{y \in Y} \text{mín}[B'(y), R(x_2, y)]$$
$$= \text{máx}[\text{mín}(.9, 1), \text{mín}(.7, .4)] = .9$$
$$A'(x_3) = \sup_{y \in Y} \text{mín}[B'(y), R(x_3, y)]$$
$$= \text{máx}[\text{mín}(.9, 1), \text{mín}(.7, .8)] = .9$$

Por lo tanto, concluimos que $\mathfrak{X}$ es A', donde $A' = .9/x_1 + .9/x_2 + .9/x_3$.

Definición 4.13. *El silogismo hipotético generalizado es expresado por el siguiente esquema*

$$\begin{array}{ll} \textit{Regla}: & \textit{Si } \mathfrak{X} \textit{ es } A, \textit{ entonces } \mathfrak{Y} \textit{ es } B \\ \textit{Hecho}: & \textit{Si } \mathfrak{Y} \textit{ es } B, \textit{ entonces } \mathfrak{Z} \textit{ es } C \\ \hline \textit{Conclusión}: & \textit{Si } \mathfrak{X} \textit{ es } A, \textit{ entonces } \mathfrak{Z} \textit{ es } C \end{array}$$

En este caso, $\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z}$ *son variables que toman valores en los conjuntos* X, Y, Z*, respectivamente,* A, B, C *son conjuntos difusos sobre* X, Y, Z*, respectivamente.*
Para cada proposición difusa condicional, existe una relación difusa, que son determinadas para cada $x \in X$*,* $y \in Y$*, y* $z \in Z$ *por las igualdades*

$$R_1(x, y) = \mathfrak{J}[A(x), B(y)]$$
$$R_2(y, z) = \mathfrak{J}[B(y), C(z)]$$
$$R_3(x, z) = \mathfrak{J}[A(x), C(z)]$$

la regla de inferencia composicional tiene la forma

$$R_3(x, z) = \sup_{y \in Y} \text{mín}[R_1(x, y), R_2(y, z)]$$

Inferencia a partir de Proposiciones condicionales y calificadas

Dada una proposición difusa condicional y calificada p de la forma

$$p: \text{ Si } \mathfrak{X} \text{ es } A, \text{ entonces } \mathfrak{Y} \text{ es } B \text{ es } S,$$

donde S es un conjunto difuso que representa un calificador de verdad, y un hecho de la forma "$\mathfrak{X}$ es A'", queremos hacer una inferencia de la forma

"$\mathfrak{Y}$ es B'".
Un método para este propósito, llamado método de resticciones del valor de verdad, consiste en los cuatro pasos siguientes:
Paso 1. Calcule el valor de verdad relativo de A' con respecto a A, denotado por $RT(A'/A)$, que es un conjunto difuso sobre el intervalo unitario definido por

$$RT(A'/A)(a) = \sup_{x:A(x)=a} A'(x)$$

para cada $a \in [0,1]$. El valor de verdad relativo $RT(A'/A)$ expresa el valor de verdad de la proposición difusa p dado el hecho "$\mathfrak{X}$ es A'".
Paso 2. Seleccione una implicación difusa conveniente $\mathfrak{J}$ por la cual la proposición difusa p es interpretada.
Paso 3. Calcule el valor de verdad relativo $RT(B'/B)$ por la fórmula

$$RT(B'/B)(b) = \sup_{a\in[0,1]} \text{mín}[RT(A'/A)(a), S(\mathfrak{J}(a,b))]$$

para cada $b \in [0,1]$. Claramente, el rol de S es el de modificar el valor de verdad de $\mathfrak{J}(a,b)$. Note que cuando S representa el calificador cierto (es decir, $S(a) = a$) para cada $a \in [0,1]$, entonces $S(\mathfrak{J}(a,b)) = \mathfrak{J}(a,b)$, y obtenemos el caso discutido anteriormente.
El valor de verdad relativo $RT(B'/B)(b)$ expresa el grado que la conclusión de la proposición difusa p es cierta.
Paso 4 Calcule el conjunto B' envuelto en la inferencia "$\mathfrak{Y}$ es B'" por la igualdad

$$B'(y) = RT(B'/B)(B(y)),$$

para cada $y \in Y$.

Ejemplo 4.9. Sea p una proposición difusa condicional y calificada de la forma

$$p: \text{ Si } \mathfrak{X} \text{ es } A, \text{ entonces } \mathfrak{Y} \text{ es } B \text{ es muy cierto,}$$

donde $A = 1/x_1 + .5/x_2 + .7/x_3, B = .6/y_1 + 1/y_2$, y S representa muy cierto, $S(a) = a^2$ para cada $a \in [0,1]$, dado un hecho "$\mathfrak{X}$ es A'", donde $A' = .9/x_1 + .6/x_2 + .7/x_3$, concluimos que "$\mathfrak{Y}$ es B'", donde B' es calculado por los siguientes pasos.
Paso 1. Calculemos $RT(A'/A)$

$$RT(A'/A)(1) = A'(x_1) = .9$$

$$RT(A'/A)(.5) = A'(x_1) = .6$$

$$RT(A'/A)(.7) = A'(x_1) = .7$$

$$RT(A'/A)(a) = 0 \text{ para cada } a \in [0,1] - \{.5, .7, 1\}.$$

Paso 2. Seleccionemos la implicación difusa $\mathfrak{J}$ de Lukasiewicz.
Paso 3. Calculemos $RT(B'/B)$
$RT(B'/B)(b) = \text{máx}\{\text{mín}[.9, S(\mathfrak{J}(.9,b))], \text{mín}[.6, S(\mathfrak{J}(.6,b))], \text{mín}[.7, S(\mathfrak{J}(.7,b))]\}$

$$= \begin{cases} (.4+b)^2 & \text{para } b \in [0, .375) \\ .6 & \text{para } b \in [.375, .475) \\ (.3+b)^2 & \text{para } b \in [.475, .537) \\ .7 & \text{para } b \in [.537, .737) \\ (.1+b)^2 & \text{para } b \in [.737, .849) \\ .9 & \text{para } b \in [.849, 1] \end{cases}$$

Paso 4. Calculemos B'

$$B'(y_1) = RT(B'/B)(B(y_1)) = RT(B'/B)(.6) = .7$$

$$B'(y_2) = RT(B'/B)(B(y_2)) = RT(B'/B)(1) = .9$$

Por lo tanto, hicimos la inferencia "$\mathfrak{Y}$ es B'", donde $B' = .7/y_1 + .9/y_2$

Capítulo 5

Psicología de Conceptos

El objetivo de este capítulo es revisar la corriente de investigación sobre la psicología de conceptos. Se describen las cuatro principales teorías de conceptos, destacando cómo cada una de ellas explican algunas importantes características de nuestras capacidades para categorizar objetos y dibujar inducciones.

En la mayoría de campos de la psicologıá y en disciplinas afines, un concepto de x (por ejemplo, el concepto de perro) se toma generalmente para ser un cuerpo de conocimientos acerca de x (por ejemplo, los perros) que se utiliza de forma predeterminada en los procesos cognitivos que financian las competencias cognitivas más altas cuando hacemos un juicio acerca de x (por ejemplo, un juicio acerca de los perros). Por lo tanto, un concepto de x es un subconjunto del conocimiento acerca de x que almacenamos en la memoria a largo plazo, o, para decirlo de otra manera, sólo una parte de nuestro conocimiento sobre nuestro concepto constituye x de x. Este cuerpo de conocimiento se utiliza por defecto, ya que se utiliza de una manera insensible al contexto: se recupera de la memoria en todos los contextos. Se viene a la mente, por así decirlo, siempre estamos pensando en su referente.

Un ejemplo puede ser útil para aclarar estas ideas. El concepto de perro es un subconjunto de nuestro conocimiento acerca de los perros. Se recupera de la memoria de largo plazo de una manera insensible al contexto, y se utiliza en los procesos de nuestras mayores competencias cognitivas. La usamos para decidir si clasificar algo como un perro, hacer inducciones acerca de los perros, a comprender frases que contienen la palabra "perro", y así sucesivamente.

Una teoría de los conceptos en los intentos de la psicología principalmente es

para identificar las propiedades que son comunes a todos los conceptos.
Los psicólogos han estado particularmente interesados en las cinco propiedades siguientes de conceptos. En primer lugar, se han tratado de determinar la naturaleza de la información que es constitutivo de los conceptos. Por ejemplo, como se verá en mayor detalle en las secciones 5.2 y 5.4, algunos psicólogos (teóricos de prototipos) sostienen que los conceptos consisten en una información estadística sobre la propiedades que son típicas y / o de diagnóstico de una clase o de una sustancia, mientras que otros (los teóricos de las hipótesis) insisten en que los conceptos consisten en información causal y / o genéricos. En segundo lugar, los psicólogos quieren determinar la naturaleza de los procesos que utilizan conceptos. Por ejemplo, algunos psicólogos han argumentado que estos procesos se basan en algún cálculo de similitud, mientras que otros no están de acuerdo. En tercer lugar, científicos cognitivos desarrollan hipótesis acerca de la naturaleza de los vehículos de conceptos. Para ilustrar, los neo-empiristas como el psicólogo Lawrence Barsalou y el filósofo Jesse Prinz sostienen que los vehículos de los conceptos son similares a los vehículos de representaciones perceptuales. En cuarto lugar, por alrededor de una década, los científicos cognitivos han tratado de identificar las áreas del cerebro que están involucradas en los conceptos que poseen. Finalmente, los científicos cognitivos han desarrollado hipótesis sobre los procesos de la adquisición de conceptos.
Como hemos visto, los conceptos se utilizan en los procesos que suscriben nuestras competencias cognitivas superiores, tales como la inducción, la categorización, la producción y la comprensión del lenguaje, y la analogía de decisiones. Mediante el desarrollo de una teoría de conceptos, explicando qué tipos de conocimiento constituyen los conceptos, qué tipos de procesos utilizan conceptos, y así sucesivamente.

5.1. Teoría Clásica de Conceptos

5.1.1. Definiciones

Hasta la década de 1970, la mayoría de los psicólogos tenían una visión simple sobre el conocimiento que constituye un concepto: Los conceptos fueron pensados para ser definiciones. De acuerdo con las versiones más comunes de la teoría clásica de los conceptos, un concepto de x representa algunas propiedades como siendo necesarias por separado y conjuntamente

suficientes para ser una x. El concepto abuela es quizás el mejor ejemplo de esta aproximación a los conceptos:
Si la gente tiene un concepto clásico de la abuela, sostienen que para ser una abuela es suficiente y necesario ser la madre de uno de los padres.
La teoría clásica de los conceptos no ha sido utilizado para explicar cómo sacar inducciones o cómo hacer analogías.
Algunos psicólogos han desarrollado versiones más complejas de la teoría clásica de los conceptos. En lugar de representar cada propiedad según sea necesaria para ser una x, un concepto de x puede consistir en una representación de cualquier combinación Booleana de propiedades siempre que esta combinación indica un condición necesaria y suficiente para ser una x. En lo siguiente, (a) ilustra las versiones simples de la teoría clásica de los conceptos, mientras que (b) ilustra las versiones más complejas:
(a) Una persona es un soltero si y sólo si es hombre, no casado y adulto.
(b) En el béisbol, una bola bateada es una bola de fair si y sólo si se instala en territorio fair entre home y primera base o entre home y tercera base, o sobre territorio fair que está delimitado por el campo fuera de la primera y tercera base, o toque primera, segunda o tercera base, o primero cae en territorio fair o más allá de la primera base o tercera base, o, sobre territorio fair, toca a un umpire o un jugador.
Las propiedades de ser hombre, no estar casado y ser adulto cada uno se toma como necesario, y juntos se considera que son suficiente para ser soltero. Por el contrario, la propiedad de ser una bola bateada que se posa sobre territorio fair entre el home y primera base o entre home y tercera base no es necesario ser una bola de fair. Es, sin embargo, necesario y suficiente que uno de los disyuntos de (b) se satisfice para que una bola sea una bola de fair.

5.1.2. Investigación sobre los conceptos clásicos

Numerosas investigaciones han examinado cómo las personas aprenden conceptos clásicos en tareas experimentales. En estos experimentos (normalmente llamados “ experimentos categoría aprendizaje ”), los participantes normalmente se les presenta con estímulos artificiales que constituyen una categoría sometida a un concepto clásico, y tienen que identificar la regla que determina la pertenencia a esta categoría. El investigador varía las condiciones de presentación (Por ejemplo, la presencia o ausencia de retroalimentación; presentación secuencial vs simultánea) y la naturaleza de la norma

(conjuntiva frente disyuntiva , etc.), mientras que los sujetos de medición velocidad y precisión de aprendizaje, que operacionalizan la dificultad de aprendizaje de una categoría definida por un tipo particular de definición.
Early comparó el aprendizaje de varios tipos de definiciones o reglas. Un sólido resultado es que la gente adquiere más fácilmente conceptos conjuntivos (rojo y cuadrado) que conceptos disyuntivos (rojo o cuadrado). Finalmente, algunos investigadores trataron de determinar una medida de la complejidad conceptual que podría predecir las dificultades de aprendizaje de las personas más o menos en definiciones o reglas complejas. De particular importancia para este último proyecto fue una secuencia de seis conceptos que son cada vez más difíciles para aprender. La investigación sobre las medidas de complejidad conceptual se ha centrado en gran parte en la explicación de por qué el aprendizaje de estos conceptos son cada vez más difíciles.
Gran parte del trabajo reciente sobre las definiciones o reglas se ha centrado en encontrar una medida de la complejidad conceptual. La medición de Feldman (longitud de descripción mínima) ha llamado mucho la atención. Él propone que "la dificultad subjetiva de un concepto es directamente proporcional a su longitud mínima de descripción Booleana, que es, a su incompresibilidad lógico ". La longitud de descripción mínima y otro principio, llamada "paridad"(es decir, cuando dos conceptos tienen la misma longitud de descripción mínima, el concepto con un número menor de casos positivos es más fácil aprender), explican la creciente dificultad de secuencia de Shepard de los conceptos; también explica por medio de la variación en las dificultades de aprendizaje 76 conceptos desarrollados por Feldman. Desde un punto de vista psicológico, Feldman interpreta este resultado como sigue: El jefe encontró que la capacidad de los sujetos para aprender los conceptos depende en gran medida de la complejidad intrínseca de los conceptos, los conceptos más complejos son más difíciles para aprender. Este efecto penetrante sugiere, en contra de las teorías de ejemplares, que el aprendizaje de conceptos críticamente implica la extracción de una simplificada o abstraída generalización de ejemplos.
Trabajos recientes, sin embargo, han arrojado serias dudas sobre esta propuesta, y las hipótesis más complejas acerca de la complejidad conceptual se han propuesto. El problema con estas hipótesis, sin embargo, es que su significado psicológico es muy claro.

5.1.3. El rechazo de la teoría clásica de los conceptos

La mayoría de los psicólogos han abandonado la teoría clásica de los conceptos ya en la década de 1970. Tres argumentos principales han sido expuestas para justificar este rechazo. En primer lugar, algunos psicólogos han argumentado que la teoría clásica de los conceptos no pueden explicar la vaguedad de categorización, es decir, por el hecho de que a veces es indeterminado si un objeto es o no es un miembro de una clase. Por ejemplo, podría ser indeterminado si algunas personas que tienen el pelo un poco, pero no mucho más a la izquierda en la cabeza, son calvos. Sin embargo, aunque generalizada, este argumento no es convincente: Una conjunción de predicados puede dar lugar a juicios de categorización vagos si los predicados mismos son vagos. Por ejemplo, como "azul" es un predicado vago, a veces será indeterminado si algo es un cuadrado azul, aunque el concepto de un cuadrado azul es un concepto clásico.
En segundo lugar, supongamos que un concepto es definido por medio de otro. Por ejemplo, la gente podría representar la acción de asesinar como la acción de matar intencionadamente que también cumple con algunas otras condiciones. A primera vista, predice que está procesando la idea de matar llevaría más tiempo de procesar el concepto de asesinar. Sin embargo, Fodor ha demostrado que este no es el caso: Estos dos conceptos son procesados en la misma velocidad.
En tercer lugar, los psicólogos descubrieron en la década de 1970 varias propiedades de nuestras decisiones de categorización que no se explican por ninguna versión de la teoría clásica de los conceptos, en particular la llamada tipicidad y efectos ejemplo (véase más adelante).

5.2. Teorías de Conceptos Prototipo

Teorías de conceptos prototipo rechazan la idea de que los conceptos representan algunas propiedades (o combinación booleana de propiedades) como necesarias y suficientes. Por lo general se propone que los conceptos son prototipos, y que un concepto de x representa cualquiera de las propiedades que son típicas de los miembros de la categoría.
Hay varias teorías de prototipos. Las teorías más simples asimilan los prototipos a la lista de propiedades típicas. Las teorías más complejas están relacionadas con las teorías de la estructura en que se distinguen atributos

de valores. Los atributos (por ejemplo, los colores, las formas) son tipos de propiedades: Determinan que los miembros de una categoría poseen una propiedad de una especie en particular. Por ejemplo, las manzanas se representan como teniendo un color.
Los valores (por ejemplo, rojo, verde, marrón) son las propiedades que posee la categoría miembros. El peso de un atributo representa la importancia de este atributo para decidir si un objeto es un miembro de la categoría, mientras que el peso de un valor representa la frecuencia subjetivamente evaluados de este valor particular entre los miembros.
Las dos teorías de prototipos brevemente descritas representan prototipos por medio de esquemas, mientras que otras teorías de prototipos representan prototipos como puntos en espacios multidimensionales.

5.2.1. Categorización y Categoría Aprendizaje

En contraste con la teoría clásica de los conceptos, las teorías de conceptos prototipo se asocian con modelos relativamente precisos de los procesos que subyacen diversas competencias cognitivas, incluyendo la categorización, la inducción, y combinación de concepto. A modo de ejemplo, revisemos el modelo de categorización de Hampton, antes de revisar algunos de los fenómenos que las teorías de prototipos tienen que explicar.
El modelo de Hampton consiste en un modelo de prototipo conceptual, una similitud medida, y una regla de decisión. Este modelo prototipo de conceptos es similar a uno hecho por Smith descrito anteriormente. Después de Hampton, la medida de similitud, $S(x, C)$, de una instancia x a una categoría C es definida en términos de valoración de $w(x, i)$, cada uno de los cuales es el peso del valor (por ejemplo, rojo) poseído por x para el atributo i del prototipo (por ejemplo, color). La medida de similaridad particular es definida por alguna forma específica de la agregación del peso $w(x, i)$ para todos los atributos relevantes. Por ejemplo,

$$S(x, C) = \sum_i w(x, i).$$

Esto significa que los prototipos suelen asumir que los juicios de categorización es influido por las propiedades tomadas independientemente uno de otro. Su configuración no importa. O, para poner el punto de otra manera, en estos modelos, las señales de categorización son independientes.

La regla de decisión de Hampton para la clasificación es una regla determinista simple,

$$S(x, C) > t \Rightarrow x \in C,$$

donde t es un criterio (o umbral) en la escala de similitud. Reglas de decisión no deterministas también se pueden utilizar, y esta regla puede ser modificada para explicar cómo la gente decide si se debe clasificar un objeto en una de dos categorías.
Por lo tanto, el modelo de Hampton del proceso de categorización implica una coincidencia de proceso entre las representaciones, es decir, el prototipo y la representación del objeto a ser categorizado, así como una medida lineal de la similitud entre el prototipo y otras representaciones. El modelo de Hampton también asume que el mismo proceso de evaluación de similitud subyace en ambos juicios de tipicidad (lo típico es un objeto de su categoría) y los juicios de categorización. Clasificaciones de tipicidad se supone que estan monotónicamente relacionada con la similitud.
La tipicidad (la medida en que un objeto posee las propiedades que son típicas de una categoría) que ha sido repetidamente demostrado que tiene una influencia extensa sobre el desempeño de las personas en una serie de tareas cognitivas. La tipicidad puede ser medida objetivamente por categorías artificiales, que puede ser medida pidiendo a la gente la lista de las propiedades de las instancias de las categorías relevantes, o se puede estimar preguntando a la gente para juzgar lo bueno que un ejemplo de un objeto en particular es ("juicios de tipicidad").
Rips, Shoben y Smith encontraron que los miembros típicos de las categorías son clasificados con mayor rapidez y precisión que miembros de la categoría atípicos: Los participantes responden más rápidamente a " un petirrojo es un pájaro " que de " un avestruz es un pájaro". Resultados similares se obtienen cuando los estímulos se presentan visualmente, por ejemplo, cuando los participantes se les muestran una imagen o un dibujo del objeto a ser categorizado, tal como un dibujo de un petirrojo. Similares hallazgos se encuentran también las categorías artificial.
La tipicidad con respecto a una categoría predice la probabilidad de ser considerado un miembro de esta categoría. Un resultado similar se ha encontrado en la lingüística. Labov ha mostrado que, en inglés Americano, artefactos se llaman "taza" o "plato" en la medida que son similares a una forma prototípica.
La tipicidad también afecta el aprendizaje de conceptos. Usando estímulos

artificiales, Posner y Keele han demostrado que, tras la adquisición de un concepto, el miembro más típico de la categoría es a veces más probable para ser considerado bien como un miembro de la categoría de los miembros de la categoría vistos durante formación, aunque este miembro más típico no se ha visto durante formación. En experimentos con categorías artificiales, los participantes aprenden la categoría de membresía de elementos típicos más rápido que la pertenencia a una categoría de elementos atípicos. Los participantes también aprenden más fácilmente a clasificar los elementos de una categoría si se han formado con los elementos típicos que si se entrenó con elementos atípicos.

Los hallazgos revisados hasta ahora son consistentes con las teorías de conceptos prototipo. Como la representación de un blanco que se supone que se ajustará con un prototipo durante la categorización, las teorías de categorización basadas en prototipos. El aprendizaje de conceptos consiste en la formación de un prototipo, teorías prototipo también esperan que la tipicidad afecte el aprendizaje de conceptos.

La idea de que los efectos de la tipicidad apoyen los conceptos prototipos a sido cuestionada desde varias direcciones. En primer lugar, Armstrong y Gleitman han argumentado que los efectos de tipicidad no muestran nada acerca de la estructura conceptual, ya que también se encuentran con que los conceptos satisfacer la teoría clásica de los conceptos.

En segundo lugar, Barsalou ha demostrado que los juicios de tipicidad no son simplemente influidos por la tipicidad (los petirojos son aves típicas), sino también por la frecuencia con un miembro de la categoría se encuentra como un miembro de la categoría (por ejemplo, con qué frecuencia se encuentran petirrojos y visto como las aves) y por la similitud de un miembro de la categoría es un miembro ideal de una categoría (la similitud de un petirrojo es un pájaro ideal). Estos hallazgos plantean un problema para los teóricos de prototipos porque estos teóricos apoyar teorías prototipo apelando en parte al hecho de que la tipicidad, como se mide por los juicios de tipicidad, predice el desempeño en tareas experimentales.

En tercer lugar, las teorías de ejemplares (ver sección 5.3) puede dar cuenta de muchos efectos de tipicidad. Por tanto, es claro si los efectos de tipicidad encontrados en las décadas de 1960 y 1970 apoyan las teorías prototipos sobre teorías ejemplares. Las investigaciones más recientes sugieren que si un prototipo o un ejemplo que se aprende en la categoría, los experimentos de aprendizaje depende de la estructura de la categoría (número de miembros de la categoría presentados durante el entrenamiento, similitudes entre los miembros de la categoría, desemejanzas entre diversas categorías) y en la etapa de la categoría de aprendizaje.

5.2.2. Inducción

Además de las tareas relacionadas con la categoriación y la categoría aprendizaje, los efectos de la tipicidad se encuentran también en tareas de inducción categórica. En tal tarea, la gente tiene que inferir si los miembros de una categoría (la meta categoría) poseen una propiedad sobre la base de que me digan que los miembros de otra categoría o de otras categorías (la categoría de fuentes o categorías) tienen esta propiedad. Por ejemplo, los participantes caben preguntarse si los gorriones tienen huesos sesamoideos dado que los petirrojos tienen huesos sesamoideos, o, equivalente, lo buena que es la siguiente inferencia:

$$\text{(a)}\ \frac{\text{Los petirrojos tienen huesos sesamoideos.}}{\text{Por lo tanto, los gorriones tienen huesos sesamoideos.}}$$

Varios hallazgos muestran que las inducciones de tipicidad influye a la gente.
Considere primero “el efecto de semejanza”. Una conclusión que se deduce de una única premisa es juzgada para ser más fuerte en la medida en que la categoría de fuente se juzga que es más similar a la categoría de destino. De este modo (a) es una inferencia mejor que la siguiente:

$$\text{(b)}\ \frac{\text{Los petirrojos tienen huesos sesamoideos.}}{\text{Por lo tanto, los pingüinos tienen huesos sesamoideos.}}$$

Considere también el “efecto tipicidad”. Una conclusión de que puede inferirse de una única premisa es juzgado para ser más fuerte en la medida en que la categoría de fuente es típico de la categoría de destino (si la categoría objetivo incluye la categoría de fuente) o de la categoría que incluye tanto la categoría objetivo y la categoría de fuente (si la categoría objetivo no incluye la categoría fuente). Consideremos, por ejemplo, las deducciones siguientes:

$$\text{(c)}\ \frac{\text{Los petirrojos tienen huesos sesamoideos.}}{\text{Por lo tanto, las aves tienen huesos sesamoideos}}$$

$$\text{(d)}\ \frac{\text{Los pingüinos tienen huesos sesamoideos.}}{\text{Por lo tanto, las aves tienen huesos sesamoideos}}$$

La inferencia (c) se juzga que es más fuerte que la inferencia (d) porque los petirrojos son una especie más típica de las aves que los pingüinos.
Dos conocidos modelos de los procesos implicados en la inducción que explican los efectos de similitud y tipicidad (así como otros efectos) suponiendo que se recuperan de la memoria los prototipos de las categorías de fuentes y de la categoría de destino que son los de Osherson y Sloman. Reviso sólo el modelo de similitud y cobertura de Osherson. En este modelo, la fuerza de la inducción es una función de la similitud promedio entre las categorías de fuentes y la categoría de objetivo y de la cobertura de las categorías de fuentes, define como la similitud promedio entre las categorías de fuentes, o bien las subclases típicas de la meta-categoría cuando la categoría objetivo incluye las categorías fuente, o la subclases típica de la categoría de más bajo nivel que incluye tanto las categorías fuente como las categorías objetivo. La similitud se determina haciendo coincidir los prototipos correspondientes.
El efecto de similitud es consecuencia de la componente de similitud en el modelo.
El efecto tipicidad es una consecuencia de la componente de la cobertura en el modelo, ya que la tipicidad de una categoría x, como petirrojos, con respecto a “ y ” una categoría más inclusiva, tales como aves, se correlaciona con la similitud entre el prototipo de x y los prototipos de las subclases típicos “ y ”.

5.3. Teorías de Conceptos Ejemplares

Teóricos ejemplares rechazan la idea de que, cuando las personas adquieren un concepto abstracto de alguna información estadística acerca de la clase que representa (por ejemplo, información sobre propiedades típicas o de diagnóstico). Más bien, proponen que las personas almacenan representaciones de los miembros de una categoría en particular (una representación de este tipo es denominado “ejemplar”), y que utilizan estas representaciones para juicios de categorización, para dibujar inducciones, y así sucesivamente. Así, para estas teorías, un concepto de los perros consiste en un conjunto de representaciones de perros en particular (Por ejemplo, una representación de Fido, una representación de Rover, etc), que son utilizado en los procesos cognitivos que subyacen a nuestras mayores competencias cognitivas. Medin y Schaffer han captado bien la esencia de las teorías ejemplares:

La idea general del modelo de contexto [el nombre de su modelo] es que los juicios de clasificación se basan en la recuperación de la información almacenada en un ejemplar. . . . Este mecanismo es, en cierto sentido, un dispositivo para el razonamiento por analogía en la medida de clasificación de nuevos estímulos que se basa en la información almacenada sobre ejemplares viejos. . . . Aunque vamos a proponer que se derivan de las clasificaciones de información del ejemplar, no asumimos que el almacenamiento y la recuperabilidad de este ejemplar es información verídica. Si los sujetos están utilizando estrategias e hipótesis durante el aprendizaje, la información del ejemplar puede ser incompleta y la relevancia de la información de dimensiones alternativas pueden variar considerablemente.

Porque la adquisición del concepto no requiere abstracción (o, en cualquier caso, requiere menos abstracción) de acuerdo con las teorías de conceptos ejemplares, el aprendizaje resulta ser más sencillo en estos puntos de vista. Por otro lado, porque la cognición implica el recuperar de la memoria a largo plazo numerosas representaciones singulares (ejemplares) y su utilización en los procesos cognitivos (por ejemplo, en el proceso de categorización subyacente), mientras que las teorías de prototipo proponer cognoscente que implica la recuperación y el uso de una sola representación, el procesamiento cognitivo es más computacionalmente intensivas de acuerdo con las teorías de ejemplares. Otra diferencia entre las teorías de prototipo y ejemplares es que las teorías prototipo asumen que los juicios de categorización a los efectos de que algo es un x, por ejemplo, un perro o una mesa y los juicios de reconocimiento para identificar a un individuo como individuo, por ejemplo, el juicio expresado por "Este es Juan", los dos implican distintos tipos de representación (respectivamente, prototipos y representaciones de datos), mientras que las teorías de ejemplares proponer que ambos tipos de juicios implican un único tipo de representación (es decir, ejemplos).

La mayoría de las teorías de ejemplares se han desarrollado en un marco espacial. Los ejemplares son representados como puntos en un espacio multidimensional, cuyas dimensiones representan las propiedades continuas de los individuos representados por los ejemplos (por ejemplo, color, altura). Una coordenada espacial a lo largo de una de estas dimensiones se corresponde a una cantidad o intensidad de la propiedad continua (por ejemplo, una determinada altura).

5.3.1. Categorización y Categoría Aprendizaje

Los modelos de conceptos de ejemplares se han desarrollado sobre todo para tener en cuenta los hallazgos en experimentos de categoría aprendizaje, y se representan así para obtener resultados más experimentales. Ahora presentamos el modelo ejemplar más conocido, el modelo de Nosofsky de contexto generalizado, antes de revisar los hallazgos más importantes que apoyan a las teorías de conceptos ejemplares.

El modelo de contexto generalizado es una extensión del modelo de contexto de Medin y de Schaffer. Se compone de un modelo ejemplar espacial de conceptos, una medida de similitud, y una regla de decisión. De acuerdo con este modelo, cada modelo representa su referente como un punto en un espacio multidimensional.

En el modelo de contexto generalizado, cada objetivo se compara con todos los ejemplares que constituyen un concepto. Por ejemplo, un perro, Fido, puede ser comparado para todos los ejemplares de perros que constituyen el concepto de alguien de perro, así como a todos los ejemplares de lobos que constituyen concepto de alguien de lobo. La similitud entre Fido y un ejemplar, por ejemplo, un ejemplo de un perro, es una función de la distancia psicológica entre Fido y el ejemplo de un perro. Esta distancia psicológica depende de la magnitud a la que Fido y el partido ejemplar en cada una de las dimensiones relevantes para categorizar a Fido: Cuanto más diferente Fido y el ejemplo están en una dimensión dada, digamos k, lo más separados que están en esta dimensión. Formalmente, para una dimensión dada, la distancia entre el objetivo y el Fido ejemplar es $|x_{t_k} - x_{E_k}|$, donde x_{t_k} es el valor del objetivo, Fido, en la dimensión k y x_{E_k} es el valor de la ejemplar en esta dimensión. Cada dimensión psicológica es ponderada: El peso de la dimensión k, w_k, mide la atención que se presta a k. Los valores más grandes de este peso capta la idea de que la falta de coincidencia a lo largo de dimensión k aumenta la disimilitud entre el modelo de un perro y Fido, lo que disminuye la probabilidad de que Fido se clasifique como un perro. Este parámetro es asumido por Nosofsky a ser dependientes del contexto. Los peso de las dimensiones sumen uno: Esto captura la idea de que la disminución de la atención a una dimensión implica el aumento de la atención a otras dimensiones.

La distancia psicológica entre Fido y el ejemplo de un perro depende de si las dimensiones relevantes son analizables. Dimensiones analizables (o separables) pueden ser atendidas independientemente una de otra. El tamaño

y el peso son analizables en las dimensiones de objetos: Podemos asistir al tamaño de un objeto, independientemente de su peso. Por el contrario, las dimensiones no analizables (o Integrales) no pueden ser atendidas independientemente una de otra. Para ejemplo, el tono, el brillo y la saturación son dimensiones no analizables.
Cuando las dimensiones no son analizables, la distancia psicológica se calcula con una métrica euclidiana:

$$d_{tE} = \sqrt{\sum_{k=1}^{n} w_k (x_{t_k} - x_{E_t})^2}.$$

Cuando las dimensiones son analizables, la distancia psicológica se calcula con una métrica city-block:

$$d_{tE} = \sum_{k=1}^{n} w_k |x_{t_k} - x_{E_k}|.$$

Más en general, la distancia entre el objetivo y el ejemplo para n dimensiones se calcula como

$$d_{tE} = c \left(\sum_{k=1}^{n} w_k |x_{t_k} - x_{E_k}|^r \right)^{\frac{1}{r}}.$$

donde r depende de si las dimensiones son analizables y c es un parámetro de sensibilidad que mide la cantidad de la distancia psicológica general entre un objetivo y un ejemplar que afecta a su similitud. La similitud entre t y E es una función exponencial de la distancia psicológica entre el objetivo y el ejemplar:

$$S_{tE} = e^{-d_{tE}}$$

Así, cuanto mayor es la distancia psicológica entre el objetivo, Fido, y el ejemplar de perro, menor será su similitud. La similitud global del objetivo, Fido, con el concepto de perro, que es, para el conjunto de ejemplos de perros, es la suma de sus similitudes con cada ejemplar de un perro.Formalmente,

$$S_{tC} = \sum_{E \in C} S_{tE}.$$

La regla de decisión es no determinista. Si dos conceptos, por ejemplo, perros y lobos, se han recuperado de la memoria a largo plazo, la probabilidad de clasificar a Fido como un perro es una función de la similitud global de Fido con el concepto de perro dividido por la suma de las similitudes generales a los conceptos de perro y del lobo. Formalmente,

$$P(t \in A) = \frac{S_{tA}}{S_{tA} + S_{tB}},$$

donde A y B son los dos conceptos relevantes. Algunas versiones del modelo generalizado de contexto S_{tA} multiplica por un parámetro (β) correspondiente a una respuesta de sesgo hacia la categoría A. Otras reglas de decisión pueden ser formalizadas por un parámetro adicional (α) que eleva S_{tA} y S_{tB} a un exponente:

$$P(t \in A) = \frac{\beta S_{tA}^{\alpha}}{\beta S_{tA}^{\alpha} + (1-\beta) S_{tB}^{\alpha}}.$$

Cuando $\alpha = 1$, la gente categoriza por la probabilidad correspondiente, cuando α es mayor que 1, la gente categoriza de una manera más determinista.

El Modelo de contexto generalizado de Nosofsky de categorización ilustra las ideas centrales de los modelos basados en modelos de los procesos cognitivos: El proceso de categorización implica comparar las representaciones de objetivos con ejemplos y calcular, de una manera no lineal, su similitud. Esa similitud normalmente se calcula de una manera no lineal en los modelos ejemplares de categorización que significa que los modelos ejemplares suponen que la configuración de propiedades, en lugar de las propiedades tomadas independientemente uno otro, es lo que importa para la categorización. Desde un punto de vista cualitativo, los modelos ejemplares de categorización están bien diseñados para dar cuenta de los efectos ejemplares. En primer lugar, vamos a considerar la ventaja de la edad de los elementos. Los participantes se les pide que aprendan la pertenencia a la categoría de los estímulos artificiales que componen dos categorías. Luego se les pide categorizar nuevos estímulos, así como los estímulos vistos durante el entrenamiento. Lo que se encontró es que los estímulos vistos durante el entrenamiento son por lo general más fáciles de clasificar a los nuevos elementos que son igualmente típicos. Para dar un ejemplo, es más fácil de clasificar mi viejo Fido que es un animal doméstico como un perro que un perro desconocido que es un perro igualmente típico. Este efecto no es predicho por las teorías de conceptos prototipo, ya que asumen que la gente abstrae de un prototipo a partir de los

estímulos que se les presentan en la fase de aprendizaje y categorizar estímulos, viejos y nuevos, mediante su comparación con el prototipo. La similitud de un elemento viejo que pertenece a una categoría dada A al conjunto de ejemplares de los miembros de A es mayor que la semejanza de un nuevo elemento a este mismo conjunto, porque el conjunto de ejemplares incluye una representación del tema antiguo, pero no hay representación del nuevo elemento.
El segundo efecto de ejemplares es la siguiente. Un miembro menos típico de una categoría se puede clasificar más rápidamente y con mayor precisión que un miembro más típico de la categoría, y su pertenencia a una categoría se puede aprender más rápidamente que la pertenencia de un miembro menos típico a la categoría, si este miembro de la categoría es similar a miembros de la categoría previamente encontradas. De nuevo para dar un ejemplo, puede ser más fácil para mí para categorizar a un perro de tres patas, como un perro que un perro cuadrúpedo si mi perro propio perdió una pierna. Medin y Schaffer establecer esto como sigue. Ellos destacan dos elementos entre los elementos de instrucción, A_1 y A_2. Estos dos elementos pertenecen a la misma categoría, A. El punto crítico es que A_1 es más similar que A_2 a un prototipo que los participantes plausiblemente abstrajeron si las teorías prototipos eran correctas. Así, las teorías del prototipo predicen que los participantes aprenderán más rápidamente la pertenencia a una categoría de A_1 que la pertenencia de A_2. Debido a que A_2 es muy similar a otros dos miembros de A y ningún miembro de la categoría alternativa, B, mientras que A_1 es muy similar a un solo miembro de A, pero a dos miembros de B, El modelo de contexto de Medin y Schaffer hace la predicción opuesta. Medin y Schaffer encontraron evidencia que apoya su predicción.

5.3.2. Modelos de Ejemplares y Conceptos de Lenguaje Natural

Las teorías de ejemplares adolecen de dos problemas importantes. En primer lugar, casi toda la investigación empírica sobre esta familia de teorías se ha centrado en actuaciones de los participantes en la categoría de experimentos de aprendizaje. Estos experimentos típicamente implican estímulos artificiales (figuras geométricas, patrones de puntos, etc) que tienen poco que ver con los objetos reales representados por conceptos fuera del laboratorio, en particular por los conceptos lexicalizados en los lenguajes naturales (por ejemplo, un perro o el agua); y las condiciones de aprendizaje en estos experimentos tienen poco que ver con la circunstancias en que los conceptos se aprenden fuera del laboratorio. En segundo lugar, las teorías de conceptos ejemplares no se han aplicado a muchas competencias de mayor reto cognitivo, una preocupación ya expresada por Murphy.

La investigación realizada por Brooks y sus colegas y por Storms y colegas mitiga el primer problema. Brooks y sus colegas han puesto de relieve el papel de los ejemplares en el conocimiento experto en medicina. Ellos tienen demostrado que el diagnóstico médico por expertos médicos es influido por la mayoría de la casos recientes que han considerado, lo que sugiere que los juicios de categorización de las categorías de enfermedades son influidos por las representaciones de los casos particulares (es decir, por ejemplares de casos concretos de enfermedades).

Storms y sus colegas han centrado su atención en los conceptos lexicalizados en las lenguas naturales. Ruts, Storms y Hampton muestran que las propiedades representadas por conceptos de categorías de orden superior funcionan como señales en lugar de independientes. Esto sugiere que los conceptos de orden superior de las categorías de objetos (vehículos, mobiliario, etc) no son prototipos. De este modo, podría ser ejemplares o quizás conjuntos de prototipos de nivel básico de categorías (por ejemplo, el concepto de muebles así de un conjunto que consta de un prototipo de sillas, un prototipo de tablas, etc.) De Boeck, Storms, y Ruts (2001) también han demostrado que para algunos conceptos expresados por los términos del lenguaje natural ("frutas", "aves", "vehículos", "Deportes") la similitud de los conceptos de las categorías que están subordinados a las categorías expresadas por estos conceptos (por ejemplo, las categorías de manzanas, plátanos, duraznos, etc, para la categoría de frutas) tienden a predecir una serie de medidas, incluyendo juicios de tipicidad, mejor que la similitud con prototipos. Es decir, cuando se evalúa si una fruta en particular es un buen ejemplo de la categoría de frutas, su similitud con los conceptos de manzanas, plátanos, melocotones, y así sucesivamente, es más importante que su similitud con un prototipo de frutas.

5.4. Teorías de Conceptos Hipotéticos

5.4.1. Conocimiento Causal y Genérico

Las Teorías de conceptos hipotéticos fueron desarrollados originalmente en la década de 1980 por los psicólogos que trabajaron en la categorización y por los psicólogos del desarrollo. Sus principios son, en parte, definidos negativamente: los teóricos de las hipótesis rechazan la idea de que los conceptos almacenan información estadística sobre las categorías, sustancias, eventos, y similares, o información acerca de miembros de la categoría en particular, las muestras, eventos, sino que también rechazan la idea de que los procesos cognitivos son impulsados por la similitud. Además, los teóricos de las hipótesis afirman que los conceptos están en algunos aspectos como en las teorías científicas (el grado en que esta analogía varía de teóricos de las hipótesis una a la otra). Como teoría científica, los conceptos constan de conocimiento que puede ser usado para explicar los acontecimientos, fenómenos o estados de cosas. Los teóricos de las hipótesis sostienen que el conocimiento causal, conocimiento nomológico, el conocimiento funcional y el conocimiento genérico son todos utilizados en la explicación, y, por lo tanto, que los conceptos constan de estos tipos de conocimiento. Por último, los conceptos se supone que se utiliza en los procesos que son de algunas maneras similares a las estrategias de razonamiento utilizadas en la ciencia. Los teóricos de las hipótesis suele aludir a inferencias para la mejor explicación o de inferencias causales para ilustrar los tipos de procesos que están definidos sobre conceptos.
Para ilustrar, una hipótesis de los perros podría ser de forma causal, por ejemplo, los perros agitan sus colas porque son felices y funcional por ejemplo, que los perros ladran para defenderse, a sí es el conocimiento acerca de los perros. Tal conocimiento podría utilizarse para explicar el comportamiento de los perros o predecir causalmente cómo un perro en particular se comportaría en circunstancias específicas.
Recientemente, los investigadores interesados en razonamiento causal y en razonamiento bayesiano han caracterizado los compromisos de las teorías de hipótesis con más detalle. Los psicólogos que trabajan en el razonamiento causal proponen que las hipótesis son (o son muy parecidos a) redes de Bayes causales. Una red de Bayes causal es una representación particular de las relaciones causales entre las variables. Sin entrar en detalles aquí, una red de Bayes causal para un conjunto de variables que específicamente tengan algunas limitaciones, mantenga en la distribución de probabilidad de estas

variables. En particular, si la variable V es causalmente relacionada con la variable W, entonces, una vez que los valores para todas las otras variables W que han sido fijos, hay una modificación del valor de V (una "intervención") que modifica la distribución de probabilidad de W. Las redes de Bayes se asocian con algoritmos que determinar los efectos de las intervenciones y con algoritmos de aprendizaje que se puede inferir relaciones causales de las correlaciones. Por lo tanto, la propuesta es que nuestros conceptos representan las relaciones causales como redes de Bayes causales y que son involucrados en los procesos cognitivos que son similares a los algoritmos causales sobre las redes de Bayes.
Psicólogos influidos por el bayesianismo han criticado este enfoque. Tenenbaum y sus colegas tienen argumentós que, aunque las redes de Bayes causales puede ser una forma adecuada de caracterizar el contenido de conceptos tales como perro, tabla, y agua, son insuficientes para caracterizar el papel de las teorías de la cognición por completo.
En particular, el enfoque de redes de Bayes causales a las teorías de hipótesis no explican por qué en algunas teorías de dominio particulares son definidas sobre variables similares y por qué las relaciones causales son similares. Por ejemplo, los conceptos de las enfermedades tienden a distinguir los síntomas de los factores patógenos y asumir que este último causa el primero. Tenenbaum y sus colegas proponen que las hipótesis forman una jerarquía y que cada hipótesis pertenece a un nivel dado de la jerarquía que está limitado por una hipótesis en el nivel inmediatamente superior. Tenenbaum y sus colegas también argumentan que los procesos psicológico que utilizan las hipótesis son poco probables que sean similares a los algoritmos causales asociados con las redes de Bayes. En su lugar, proponen procesos cognitivos que están mejor representados como las inferencias bayesianas.
Además de especificar con mayor detalle las hipótesis apelando a los recientes trabajos en inferencias causales o en inferencias bayesianas, los teóricos de las hipótesis también insisten en la importancia de los conocimientos genéricos en nuestras representaciones conceptuales de las categorías, sustancias, eventos y similares.

5.4.2. Categorización

La teoría de hipótesis ha sido muy reforzada por el trabajo en el desarrollo del esencialismo psicológico. El esencialismo psicológico tiene la hipótesis de que la gente cree que la pertenencia a alguna clase está determinada por la posesión de una esencia (incluso cuando la gente son incapaces de especificar la naturaleza de esta esencia). Además, se plantea la hipótesis que la gente cree que esta esencia hace que miembros de la categoría tengan las propiedades que son características de la categoría a la que pertenecen. Por ejemplo, Keil en su obra clásica sugiere que incluso los niños pequeños asumen que la pertenencia a la categoría en clases naturales no está determinada superficialmente, por las propiedades observables, sino por no observables, las propiedades inmutables.

Así, los niños juzgan que un zorrillo cuya apariencia es quirúrgicamente modificada para parecer un mapache sigue siendo un zorrillo.

Los teóricos de las hipótesis también han destacado "los efectos causales" en la categorización, a saber, los fenómenos observados en tareas experimentales que se explican mejor si se supone que los participantes traigan un poco de conocimiento causal para influir en estas tareas. Ahn propone "la hipótesis causal de estado", según el cual "las personas consideran como causa características más importantes y características esenciales de efecto en su representación conceptual ". Ahn y sus colegas sugieren que las decisiones de la gente para clasificar un objeto en un punto de vista lingüístico describen la categoría artificial influidas por la centralidad causal de las propiedades que son características de la categoría. Una propiedad es más central causalmente que otra en la medida en que provoca un mayor número de propiedades que son características de la categoría. Así pues, si dos objetos poseen el mismo número de propiedades características, pero si las propiedades que posee uno de ellos son causalmente más centrales que las propiedades poseídas por el otro, el primero es más probable que se clasifique como un miembro de la categoría que el segundo. El enfoque de Rehder está en tensión con la hipótesis del status causal de Ahn. Para Rehder, nuestras creencias acerca de las relaciones causales entre las propiedades que caracterizan a una categoría determina qué propiedades esperamos estan asociados entre los miembros de la categoría. Cuando un objeto no posee las propiedades que esperamos a asociarse, juzgamos que es menos probable que pertenezca a la categoría correspondiente. Desde esta perspectiva, las propiedades causalmente centrales no son necesariamente más importante en la categorización de propiedades causalmente menos centrales (para ver esto, considere estructuras comunes de efectos causales).

5.4.3. Inducción

La investigación sobre la inducción también sugiere que la información causal es a veces utilizada para evaluar la fuerza de una conclusión inductiva o para decidir cómo hacer una inducción. Proffitt, Coley, y Medin investigaron los juicios de los expertos de los árboles (paisajistas, taxónomos, y personal de mantenimiento de parques) sobre la fortaleza de conclusiones inductivas sobre los árboles. En el primer experimento, a los expertos de los árboles se les dijo que la enfermedad afecta a A una especie de árbol, x, mientras que la enfermedad afecta a B otra especie, y. Luego preguntó: "¿ Creé usted que la enfermedad podría afectar a más de los otros tipos de árboles que se encuentran por aquí? "Le pidieron a los expertos que justificaran sus juicios. Proffitt y sus colegas encontraron que la tipicidad a menudo no afectó a los expertos acerca de si los otros árboles se verían afectados por la enfermedad. En lugar de basarse en la tipicidad de las dos especies de árboles, x e y (Como se predice, por ejemplo, por Osherson y colegas en su modelo de similitud de cobertura), el patrón de respuestas y las justificaciones suministradas indican que los expertos a menudo basan sus juicios sobre mecanismos hipotéticos causales que podrían explicar la propagación de la enfermedad. Particularmente, se consideró que una enfermedad estaría presente en muchos árboles si las especies en estudio eran ecológicamente relacionadas con muchos árboles. Proffitt, Coley, y Medin informaron que un participante explicó su respuesta de la siguiente manera: "Por ejemplo, un experto mencionó que en los robles es probable que se propague la enfermedad a través de sus raíces y de que su sistema de raíces hizo una fuerte base para la inducción".

El uso del conocimiento causal no se limita a los expertos como puede demostrarse, por ejemplo, por el efecto causal de asimetría. El efecto causal de asimetría es el siguiente: Cuando hay una explicación intuitiva causal de por qué la categoría objetivo tiene una propiedad si la categoría de fuente que tenía, cambiando la premisa y la conclusión debilita la fuerza de la inducción. Por lo tanto, la inducción

$$\frac{\text{Las Gacelas tienen retina.}}{\text{Los Leones tienen retina.}}$$

es más fuerte que la inducción

$$\frac{\text{Los Leones tienen retina.}}{\text{Las Gacelas tienen retina.}}$$

5.5. Relación entre las teorías de conceptos

¿Cuál es la relación entre las teorías de conceptos revisados en las secciones anteriores? Durante mucho tiempo, la opinión tradicional ha sido que estas teorías estaban compitiendo. Es decir, se supuso que una y sólo una de estas teorías puede ser correcta: Los conceptos eran o bien definiciones, o prototipos, o conjuntos de ejemplares, o hipótesis. Por lo tanto, los psicólogos comprometidos con diferente teorías de conceptos (por ejemplo, una teoría prototipo especial y particular a teoría ejemplar) se centró en el descubrimiento de las propiedades de más altas competencias cognitivas (por ejemplo, los efectos ejemplares reportados por Medin y Schaffer) que se explica fácilmente por la teoría que recibe la aprobación (por ejemplo, la teoría ejemplar), pero que no eran fácilmente explicables por las demás teorías (teorías de prototipo, naturalmente, no explican los efectos ejemplares). Sobre esta base, se utiliza para concluir que su propia teoría, pero no la otra teoría, era probablemente correcta.

Recientemente, sin embargo, los psicólogos y filósofos de la psicología propusieron examinar estas teorías aparentemente que compiten de manera diferente. En trabajos recientes, Machery ha sostenido que cada categoría (por ejemplo, los perros), cada sustancia (por ejemplo, agua), cada tipo de evento (por ejemplo, ir al dentista), y así sucesivamente, se representa por conceptos, para varios ejemplos, por un prototipo, por un conjunto de ejemplares, y por una hipótesis (conocida como "la hipótesis de heterogeneidad"). Por lo tanto, tenemos un prototipo de perros, un conjunto de ejemplares de perros en particular, y una hipótesis de los perros. Si esto es correcto, entonces algunas de las teorías actuales de conceptos no están compitiendo una con la otra, sino que están simplemente describen los diferentes tipos de conceptos, a saber, prototipos, ejemplares e hipótesis. Debido a que estos tres tipos de conceptos tienen poco en común, también es un error asumir que hay muchas de las propiedades generales de los conceptos, y que una teoría de conceptos debe intenta describir estos. Además, la hipótesis de heterogeneidad sostiene que los primeros prototipos, ejemplares, e hipótesis se utilizan típicamente en distintos procesos. Por ejemplo, tenemos varios procesos de categorización: un proceso de categorización basado en un prototipo, un proceso de categorización basado en ejemplares y un proceso de categorización basado en hipótesis.

¿Cuál es la evidencia para la afirmación que la memoria a largo plazo almacena prototipos, ejemplares y teorías? En primer lugar, cuando se examina treinta años de investigación sobre la categorización y la inducción, uno encuentra que en ambos ámbitos de la investigación algunos fenómenos son bien explicables si los conceptos producidos por algunas de las tareas experimentales son prototipos, algunos fenómenos se explican adecuadamente si los conceptos provocados por otras tareas experimentales son ejemplares, y sin embargo otros fenómenos bien explicables si los conceptos provocados por otras tareas experimentales son hipótesis. Si se plantea la hipótesis de que las condiciones experimentales la dependencia de un tipo de conceptos (por ejemplo, los prototipos) en lugar de otros tipos (Por ejemplo, ejemplares e hipótesis), esto proporciona evidencia de la heterogeneidad de hipótesis.

En segundo lugar, la evidencia más directa sugiere que diferentes tipos de conceptos a veces compiten por el control del comportamiento. Presentando visualmente categorías artificiales, Allen y Brooks han demostrado que la gente puede, adquirir y utilizar simultáneamente dos diferentes conceptos, un conjunto de ejemplares y algo así como una definición. En un experimento de palabras utilizadas como estímulos, Malt y su análisis de protocolo ha demostrado que muchos participantes utilizan ambos prototipos y ejemplares. En cuanto a conceptos lexicalizados, Storms, De Boeck y Ruts encontraron que los modelos de prototipos explican un porcentaje adicional de la varianza en adición a la proporción de la varianza explicada por los modelos ejemplares. El estudio de los cuerpos de conocimiento utilizados por los médicos expertos para clasificar las enfermedades y de razonar sobre ellos, Brooks y sus colegas han argumentado que los médicos expertos simultáneamente tienen cuerpos diferenciados de conocimiento acerca de las enfermedades: una hipótesis, un prototipo, y ejemplares. Este último cuerpo de hallazgos es particularmente importante porque examina los conceptos de la gente en el mundo real, no los conceptos que adquieren en condiciones altamente artificiales de los experimentos de aprendizaje de conceptos.

La hipótesis de la heterogeneidad no es la única manera de combinar las teorías de los conceptos revisados en las secciones anteriores. Machery menciona dos otras alternativas. Algunos psicólogos proponen que si una determinada categoría está representada por un prototipo o por un conjunto de ejemplares depende varias propiedades de esta categoría, incluyendo la forma diferente de otros miembros de la categoría, qué diferencia con los miembros de otras categorías, y en varias características del proceso de concepto de adquisición. En contraste con la hipótesis de heterogeneidad, esta hipótesis propone que una determinada categoría está representada por un solo concepto, pero coincide con la hipótesis de heterogeneidad al considerar que hay

diferentes tipos de conceptos.
Otros han resaltado las similitudes entre los tipos de conceptos postulada por las teorías discutidas en este capítulo. Particularmente, Danks ha demostrado que los modelos de categorización presentada por prototipo, teorías ejemplares, y la teoría de hipótesis, puede ser representado como modelos gráficos. En su opinión, pues, una teoría más general de los conceptos se llama así para, que pueda unir a los prototipos, ejemplares y teorías causales.

Capítulo 6

Aplicaciones

En este capítulo desarrollaremos un método para la evaluación de los estudiantes utilizando un sistema difuso, y para concluir se mencionan otras áreas de aplicación de la lógica difusa con su respectiva bibliografía.

6.1. Un sistema difuso para la evaluación de logros de aprendizaje en los estudiantes

La evaluación de logros de aprendizaje en los estudiantes es el proceso de la determinación de los niveles de rendimiento de cada estudiante en relación a los objetivos educativos. Un sistema de evaluación de alta calidad, ofrece motivos para la mejora individual y asegura que todos los estudiantes reciben calificaciones justas para que no se limiten a los estudiantes oportunidades presentes y futuras. Por lo tanto, el sistema debe periódicamente revisar y mejorar para asegurarse de que es preciso, justo y beneficioso para todos los estudiantes. Por lo tanto, el sistema de evaluación tiene necesidades de la transparencia, la objetividad, razonamiento lógico y fácil aplicación informática que podría ser proporcionada por el sistema de lógica difusa.
A partir de estudios anteriores, se puede encontrar que números difusos, conjuntos difusos, reglas difusas, y los sistemas de lógica difusa han sido utilizados para diversos sistemas de clasificación educativos. Weon y Kim [7] desarrollaron una estrategia de evaluación basada en funciones de pertenencia difusa. Señalaron que el sistema para que la evaluación de logros de los estudiantes debe considerar los tres factores importantes de las preguntas que se dan a los estudiantes : la dificultad, la importancia, y la complejidad.

Bai y Chen [2] señalan que el factor de dificultad es un parámetro muy subjetivo y puede causar una alegación relativa a la equidad en la evaluación. Bai y Chen [1] propusieron un método para construir automáticamente la función de pertenencia del grado de reglas difusas para evaluar el logro de aprendizaje del estudiante.
Consideremos la misma situación y ejemplo como lo hicieron Bai y Chen [2]. Supongamos que hay n estudiantes para responder m preguntas. La tasa de precisión de las respuestas de los alumnos (puntuación del estudiante en cada pregunta dividida por la puntuación máxima asignada a esta pregunta) es la base para la evaluación. Tenemos una matriz de precisión de dimensión $m \times n$,

$$A = [a_{ij}], \quad m \times n,$$

donde $a_{ij} \in [0, 1]$ denota la tasa de precisión del estudiante j a la pregunta i. La tasa de tiempo de estudiantes (tiempo consumido por el estudiante para resolver la pregunta dividido por el tiempo máximo requerido para resolver esta pregunta) es otra base para ser considerada en la evaluación. Tenemos una matriz de tiempo de dimensión $m \times n$,

$$T = [t_{ij}], \quad m \times n,$$

donde $t_{ij} \in [0, 1]$ denota la tasa de tiempo del estudiante j en la pregunta i. Demos un vector de grado

$$G = [g_i], \quad m \times 1,$$

donde $g_i \in [1, 100]$ indica la puntuación máxima designada a la pregunta i que satisface

$$\sum_{i=1}^{m} g_i = 100$$

En base a la matriz de tasa de precisión A y el vector de grado G, se obtiene la puntuación total del vector original de dimensión $n \times 1$

$$S = A^T G = [S_j], \quad n \times 1,$$

donde $s_j \in [0, 1]$ es la puntuación total del estudiante j. El rango "clásico" de los estudiantes se obtienen por la clasificación de los valores de los elementos de S en orden descendente.

Ejemplo 6.1. *Supongamos que debemos examinar a 10 estudiantes con 5 preguntas, donde la matriz de precisión, de tiempo y el vector de grado son dados de la siguiente manera*

$$A = \begin{bmatrix} 0{,}59 & 0{,}35 & 1 & 0{,}66 & 0{,}11 & 0{,}08 & 0{,}84 & 0{,}23 & 0{,}04 & 0{,}24 \\ 0{,}01 & 0{,}27 & 0{,}14 & 0{,}04 & 0{,}88 & 0{,}16 & 0{,}04 & 0{,}22 & 0{,}81 & 0{,}53 \\ 0{,}77 & 0{,}69 & 0{,}97 & 0{,}71 & 0{,}17 & 0{,}86 & 0{,}87 & 0{,}42 & 0{,}91 & 0{,}74 \\ 0{,}73 & 0{,}73 & 0{,}18 & 0{,}16 & 0{,}5 & 0{,}02 & 0{,}32 & 0{,}92 & 0{,}9 & 0{,}25 \\ 0{,}94 & 0{,}49 & 0{,}08 & 0{,}81 & 0{,}65 & 0{,}93 & 0{,}39 & 0{,}51 & 0{,}97 & 0{,}61 \end{bmatrix}$$

$$T = \begin{bmatrix} 0{,}7 & 0{,}4 & 0{,}1 & 1 & 0{,}7 & 0{,}2 & 0{,}7 & 0{,}6 & 0{,}4 & 0{,}9 \\ 1 & 0 & 0{,}9 & 0{,}3 & 1 & 0{,}3 & 0{,}2 & 0{,}8 & 0 & 0{,}3 \\ 0 & 0{,}1 & 0 & 0{,}1 & 0{,}9 & 1 & 0{,}2 & 0{,}3 & 0{,}1 & 0{,}4 \\ 0{,}2 & 0{,}1 & 0 & 1 & 1 & 0{,}3 & 0{,}4 & 0{,}8 & 0{,}7 & 0{,}5 \\ 0 & 0{,}1 & 1 & 1 & 0{,}6 & 1 & 0{,}8 & 0{,}2 & 0{,}8 & 0{,}2 \end{bmatrix}$$

$$G^T = \begin{bmatrix} 10 & 15 & 20 & 25 & 30 \end{bmatrix}$$

donde V^T denota la matriz transpuesta de V.

La importancia de las preguntas es un factor importante. Tenemos l niveles de importancia para describir el grado de importancia de cada pregunta en el dominio difuso (cada nivel es un conjunto difuso). El dominio experto (utilizando alguna de las teorías expuestas en el capítulo anterior) determina la importancia de la matriz de dimensión $m \times l$

$$P = [p_{ik}], \quad m \times l,$$

donde $p_{ik} \in [0,1]$ indica el valor de pertenencia de la pregunta i al nivel k de importancia. Para simplificar las cosas, utilizaremos cinco niveles de importancia ($l = 5$), para $k = 1$ el término lingüístico "baja", $k = 2$ para "más o menos baja", $k = 3$ para "media", $k = 4$ para "más o meno alta", $k = 5$ para "alta". Representados por los conjuntos difusos

$$f_1(x) = \begin{cases} 1 & \text{si } x \in [0{,}0.1] \\ \frac{0{,}3-x}{0{,}2} & \text{si } x \in [0.1{,}0.3] \\ 0 & \text{si } x \in [0.3{,}1] \end{cases}$$

$$f_2(x) = \begin{cases} \frac{x-0{,}1}{0{,}2} & \text{si } x \in [0.1{,}0.3] \\ 1 & \text{si } x = 0{,}3 \\ \frac{0{,}3-x}{0{,}2} & \text{si } x \in [0.3{,}0.5] \\ 0 & \text{si } x \in [0{,}0.1] \cup [0.5{,}1] \end{cases}$$

$$f_3(x) = \begin{cases} \frac{x-0{,}3}{0{,}2} & \text{si } x \in [0.3{,}0.5] \\ 1 & \text{si } x = 0{,}5 \\ \frac{0{,}5-x}{0{,}2} & \text{si } x \in [0.5{,}0.7] \\ 0 & \text{si } x \in [0{,}0.3] \cup [0.7{,}1] \end{cases}$$

$$f_4(x) = \begin{cases} \frac{x-0{,}5}{0{,}2} & \text{si } x \in [0.5{,}0.7] \\ 1 & \text{si } x = 0{,}7 \\ \frac{0{,}7-x}{0{,}2} & \text{si } x \in [0.7{,}0.9] \\ 0 & \text{si } x \in [0{,}0.5] \cup [0.9{,}1] \end{cases}$$

$$f_5(x) = \begin{cases} 1 & \text{si } x \in [0.9{,}1] \\ \frac{x-0{,}7}{0{,}2} & \text{si } x \in [0.7{,}0.9] \\ 0 & \text{si } x \in [0{,}0.7] \end{cases}$$

Notemos que los mismos cinco conjuntos difusos se aplican a la precisión, el tiempo, la dificultad, la complejidad y el ajuste de las preguntas.
La complejidad de las preguntas que indica la capacidad de los estudiantes para dar respuestas correctas es también un factor importante que considerar. El dominio experto determina la matriz de complejidad difusa de dimensión $m \times l$

$$C = [c_{ik}], \quad m \times l,$$

donde $c_{ik \in [0,1]}$ denota el valor de pertenencia de la pregunta i al nivel de complejidad k.

Ejemplo 6.2. *Para el ejemplo anterior tenemos los siguiente, dado por el dominio experto:*

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0{,}33 & 0{,}67 & 0 & 0 \\ 0 & 0 & 0 & 0{,}15 & 0{,}85 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0{,}07 & 0{,}93 & 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0{,}85 & 0{,}15 & 0 & 0 \\ 0 & 0 & 0{,}33 & 0{,}67 & 0 \\ 0 & 0 & 0 & 0{,}69 & 0{,}31 \\ 0{,}56 & 0{,}44 & 0 & 0 & 0 \\ 0 & 0 & 0{,}7 & 0{,}3 & 0 \end{bmatrix}$$

La puntuación total se obtiene entonces por la formula $S = A^T G$

$$S^T = \left[\begin{array}{cccccccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 & s_9 & s_{10} \end{array} \right]$$

$$= \left[\begin{array}{cccccccccc} 67{,}60 & 54{,}05 & 38{,}40 & 49{,}70 & 49{,}70 & 48{,}80 & 46{,}10 & 52{,}30 & 85{,}95 & 49{,}70 \end{array} \right]$$

entonces el rango clásico de los estudiantes es

$$S_9 > S_1 > S_2 > S_8 > S_4 = S_5 = S_{10} > S_6 > S_7 > S_3$$

Bai y Chen usan tres pasos para evaluar las respuestas de los estudiantes. En el primer paso, los valores de precisión y de tiempo, se obtienen como la tasa media del vector precisión de dimensión $m \times 1$,

$$\overline{A} = [a_{i\bullet}], \quad m \times 1,$$

donde $a_{i\bullet}$ denota la tasa media de precisión de la pregunta i que es obtenida por

$$a_{i\bullet} = \sum_{j=1}^{n} \frac{a_{ij}}{n}$$

y la tasa de vector tiempo medio de la misma dimensión

$$\overline{T} = [t_{i\bullet}], \quad m \times 1,$$

donde $t_{i\bullet}$ denota la tasa promedio de tiempo de la pregunta i que es obtenida por

$$t_{i\bullet} = \sum_{j=1}^{n} \frac{t_{ij}}{n}$$

A continuación, obtenemos la matriz de índice de precisión difusa de dimensión $m \times l$

$$FA = [fa_{ik}], \quad m \times l,$$

donde $fa_{ik} \in [0,1]$ indica el valor de pertenencia de la tasa media de precisión de la pregunta i al nivel k, y la matriz difusa del tiempo de dimensión $m \times l$

$$FT = [ft_{ik}], \quad m \times l,$$

donde $f_{ik} \in [0,1]$ indica el valor de pertenencia del tiempo medio de la pregunta i al nivel k.

Ejemplo 6.3. *En el ejemplo anterior, obtenemos*

$$\overline{A}^T = \begin{bmatrix} 0{,}45 & 0{,}31 & 0{,}711 & 0{,}47 & 0{,}637 \end{bmatrix}$$

$$\overline{T}^T = \begin{bmatrix} 0{,}57 & 0{,}48 & 0{,}31 & 0{,}50 & 0{,}57 \end{bmatrix}$$

Basados en los conjuntos difusos tenemos las matrices difusas de la precisión y tiempo

$$FA = \begin{bmatrix} 0 & 0{,}25 & 0{,}75 & 0 & 0 \\ 0 & 0{,}95 & 0{,}05 & 0 & 0 \\ 0 & 0 & 0 & 0{,}945 & 0{,}055 \\ 0 & 0{,}15 & 0{,}85 & 0 & 0 \\ 0 & 0 & 0{,}315 & 0{,}685 & 0 \end{bmatrix}$$

$$FT = \begin{bmatrix} 0 & 0 & 0{,}65 & 0{,}35 & 0 \\ 0 & 0{,}1 & 0{,}9 & 0 & 0 \\ 0 & 0{,}95 & 0{,}5 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0{,}65 & 0{,}35 & 0 \end{bmatrix}$$

En el segundo paso, basados en la matriz de precisión difusa, FA, la matriz difusa de tiempo, FT, y las reglas difusas, $\mathfrak{R}_D$, dadas en la forma de reglas si-entonces, obtenemos la matriz de dificultad difusa de dimensión $m \times l$,

$$D = [d_{ik}], \quad m \times l,$$

donde $d_{ik} \in [0,1]$ denota el valor de pertenencia de dificultad de la pregunta i al nivel k. Cuando el nivel de precisión, l_A, y el nivel de tiempo, l_T, son dados, se determina el nivel de dificultad, l_D, por las reglas difusas dadas

$$l_D = \mathfrak{R}_D(l_A, l_T)$$

La cual denota la importancia relativa de la tasa de precisión y tiempo, respectivamente, de W_A y W_T ($W_A + W_T = 1$), que se determinan por un dominio experto. El valor de d_{ik} se obtiene por

$$d_{ik} = \max_{\{(l_A,l_T)|\ \mathfrak{R}_D(l_A,l_T)=k\}} \{W_A \cdot fa_{i,l_A} + W_T \cdot ft_{i,l_T}\}$$

A continuación, sobre la base de la matriz de dificultad difusa, D, y la matriz de complejidad, C, con sus pesos relativos W_D y W_C ($W_D + W_C = 1$), respectivamente. Teniendo en cuenta las reglas difusas, $\Re_E$, obtenemos la matriz esfuerzo de dimensión $m \times l$, de la misma manera que se obtuvo la matriz de dificultad

$$E = [e_{ik}], \quad m \times l,$$

donde $e_{ik} \in [0, 1]$ denota el valor de pertenencia del esfuerzo para responder la pregunta i al nivel k, que es una medida del esfurzo requerido por los estudiantes para responder a la pregunta i.
A continuación, sobre la base de la matriz de esfuerzo, E, y la matriz de importancia, P, con sus pesos relativos, W_E y W_P ($W_E + W_P = 1$), respectivamente. Teniendo en cuenta las reglas difusas, $\Re_W$, se obtiene la matriz de ajuste de dimensión $m \times l$

$$W = [w_{ik}], \quad m \times l$$

donde $w_{ik} \in [0, 1]$ denota el valor de pertenenciadel ajuste de la pregunta i al nivel k. A continuación, se utiliza la siguiente fórmula para obtener el vector ajuste,

$$\overline{W} = [w_{i\bullet}], \quad m \times 1,$$

donde $w_{i\bullet} \in [0, 1]$ denota el valor de ajuste final requerido por la pregunta i obtenida por

$$w_{i\bullet} = \frac{0{,}1 \times w_{i1} + 0{,}3 \times w_{i2} + 0{,}5 \times w_{i3} + 0{,}7 \times w_{i4} + 0{,}9 \times w_{i5}}{0{,}1 + 0{,}3 + 0{,}5 + 0{,}7 + 0{,}9}$$

donde 0.1, 0.3, 0.5, 0.7, 0.9 son los centros de los conjuntos difusos.

En el tercer y último paso, nos adaptamos a los rango originales de los estudiantes. Note que Bai y Chen aplicaron su método sólo para los alumnos con la misma puntuación total. Supongamos que hay q estudiantes con un puntaje total igual. Vamos a reagrupar estos estudiantes y les numeramos desde 1 a q, así como la original tasa de precisión a_{ij} correspondiente. La suma de las diferencias (SOD) para los alumnos con la misma puntuación total se calcula por

$$SOD_j = \sum_{k=1}^{q} \sum_{i=1}^{m} (a_i j - aik) \cdot \widehat{G},$$

donde $\widehat{G} = [\widehat{g}_i]$, $\widehat{g}_i = 0{,}5 + w_{i\bullet}$.
Los nuevos rangos de los estudiantes se obtendrán por la clasificación de valores de SOD en orden descendente.

Ejemplo 6.4. *Para el mismo ejemplo, tenemos tres estudiantes S_4, S_5 y S_{10} tienen la misma puntuación total ($q = 3$). Después reorganizamos la matriz original A, así*

$$\begin{bmatrix} 0{,}66 & 0{,}11 & 0{,}24 \\ 0{,}04 & 0{,}88 & 0{,}53 \\ 0{,}71 & 0{,}17 & 0{,}74 \\ 0{,}61 & 0{,}5 & 0{,}25 \\ 0{,}81 & 0{,}65 & 0{,}61 \end{bmatrix}$$

Obtenemos el valor de SOD para estos estudiantes

$$SOD = \begin{bmatrix} 3{,}27 & -5{,}46 & 2{,}19 \end{bmatrix}$$

con el nuevo rango tenemos

$$S_4 > S_{10} > S_5.$$

Entonces el nuevo rango de los estudiantes es

$$S_9 > S_1 > S_2 > S_8 > S_4 > S_{10} > S_5 > S_6 > S_7 > S_3$$

6.2. Otras aplicaciones

Desde finales de 1980, la literatura sobre las aplicaciones de la lógica difusa ha ido creciendo tan rápidamente que no sería realista una visión global de todas las aplicaciones establecidas de lógica difusa. Identificamos únicamente las más visibles aplicaciones y proporcionamos al lector información selectiva para seguir estudio. Las circunstancias que llevaron al rápido aumento de las solicitudes de lógica difusa a finales de 1980 y principios de 1990 han sido bien caracterizados por McNeil y Freiberger.
La mayoría de las aplicaciones de la lógica difusa que se desarrollaron antes de 1995 son descrito por Klir y Yuan. Alrededor de la mitad de su libro está dedicado a aplicaciones e incluye referencias clave para cada uno de los casos de aplicación. Por otra parte, el libro tiene un índice bibliográfico, que ayuda al lector a identificar rápidamente referencias en las diversas áreas de aplicación. Algunas de las aplicaciones son genéricas en el sentido de que son aplicables a múltiples disciplinas, y algunas se han desarrollado para dominios específicos de disciplinas individuales. Entre las aplicaciones genéricas descritas en el libro son razonamiento aproximado, clasificación,

control, análisis de datos, toma de decisiones, diseño, diagnóstico, tratamiento de imágenes, optimización, reconocimiento de patrones, análisis de regresión, análisis de confiabilidad, análisis de riesgos, programación y modelado de sistemas. Entre las aplicaciones del dominio específicos son aquellos en los negocios, ciencias de la computación (informática, bases de datos, sistemas expertos, recuperación de información, etc), la ingeniería (química, civil, eléctrica, ambiental, mecánica, industrial nuclear), estudios de terremotos (sismología), la ecología, la economía, la medicina, la meteorología, la física, la psicología y las ciencias de la conducta, la robótica y la ciencia de sistemas. Zimmerman ofrece un libro más reciente dedicado a las aplicaciones de lógica difusa.

La mayoría de las aplicaciones de la lógica difusa descrita en los libros anteriormente mencionados han ampliado considerablemente, y las aplicaciones en nuevas áreas han surgido. La siguiente es una lista de algunas de estas nuevas áreas, cada una con una referencia representativa:

Química (Rouvray, G.M., ed. 1997. Fuzzy Logic in Chemistry. San Diego, CA: Academic Press.)

Finanzas (Peray, K. 1999. Investing Mutual Funds Using Fuzzy Logic. Boca Raton, FL: St. Lucia Press.)

Geografía (Petry, F.E., V.B.Robinson, y M.A.Cobb, eds. 2005. Fuzzy Modeling with Spatial Information for Geographic Problems. Berlin: Springer.)

Geología (Demicco, R.V., y G.J. Klir, eds. 2004. Fuzzy Logic in Geology. San Diego, CA: Academic Press.)

Gerencia (Carlsson, C., M.Fedrizzi, y R.Fuller. 2004. Fuzzy Logic in Management. Boston: Kluwer.)

La ciencia política (Clark, T.D., J.M.Larson, J.N.Mordeson, J.D.Potter, y M.J.Wierman. 2008. Applying Fuzzy Mathematics to Formal Models in Comparative Politics. Berlin: Springer.)

Sociología (Ragin, C.C. 2000. Fuzzy-Set Social Science. Chicago: University of Chicago Press.)

Transporte (Teodorović, D., y K.Vukadinović. 1998. Traffic Control and Transport Planning: A Fuzzy Sets and Neural Networks Approach. Boston: Kluwer.)

Conclusión

Las aplicaciones de la lógica difusa son muchas, por lo que, se puede realizar un estudio más profundo en cualquiera de las áreas mencionadas.
El proyecto a futuro es el de incluir otras teorías, por ejemplo, redes neuronales, algoritmos genéticos, sistemas dinámicos; para poder tener más herramientas en las aplicaciones del área del proceso cognitivo, también se espera tener el apoyo de profesores ajenos al área de las matemáticas, por ejemplo, en psicología o en sociología.
Aún queda un largo camino que recorrer en todas las áreas mencionadas anteriormente, al igual que en áreas emergentes, por lo tanto, espero que en un futuro existan mayores adelantos en la epistemología y poder realizar un plan de estudios más capaz que el actual.

Bibliografía

[1] Bai, S.-M./ Chen, S.-M.: *Automatically constructing grade membership functions of fuzzy rules for students evaluation*, Expert Systems with Applications, 35(3), 1408-1414,(2008).

[2] Bai, S.-M./ Chen, S.-M.: *Evaluating students learning achievement using fuzzy membership functions and fuzzy rules*, Expert Systems with Applications, 34, 399-410,(2008).

[3] Eugene Roventa/ Mircea Reghiş: *Classical and fuzzy concepts in mathematical logic and applications*, CRC Press LLC, (1998).

[4] George J. Klir: *Fuzzy sets and fuzzy logic: theory and applications*, Prentice Hall PTR, (1995).

[5] Ibrahim Saleh, Seong-in Kim: *A fuzzy system for evaluating students learning achievement*, Expert Systems with Applications 36 (2009).

[6] Radim Belohlavek/ George J. Klir: *Concepts and fuzzy logic*, Massachusetts Institute of Technology, (2011).

[7] Weon, S./ Kim, J. : *Learning achievement evaluation strategy using fuzzy membership function*, In Proceedings of the 31st ASEE/IEEE frontiers in education conference, Reno, NV (Vol. 1, pp. 19-24)(2001).

Índice alfabético