UNIVERSIDAD CENTROCCIDENTAL
"LISANDRO ALVARADO"

Decanato de Ciencias y Tecnología
Licenciatura en Ciencias Matemáticas

"Fundamentos sobre la teoría de la posibilidad basada en conjuntos difusos"

Trabajo Especial de Grado presentado por

Br. José Ángel Torrealba

como requisito final
para obtener el título de Licenciado
en Ciencias Matemáticas
Área de Conocimiento: teoría de conjuntos difusos.
Tutor: Dra. María Luisa Capodieci

Barquisimeto, Venezuela. Abril de 2013

*A Dios primeramente. A mi madre,
Carmen Mendoza "mi primera maestra" y
mi padre Ángel Torrealba un digno ejemplo
de sacrificio, esfuerzo y trabajo.*

Agradecimientos

A **Dios**, siempre le he dado gracias por todo. En esta ocasión, por haber recibido la bendición de ser constante y perseverante durante la carrera para lograr el éxito.

A **mis padres**, que con su amor y cariño siempre me han ayudado sin condición alguna por lo cual fueron motivo para lograr el triunfo. Esto es para ustedes... Los Amo!

A **mis hermanos**, que con sus ocurrencias le dan ese toque de familia a nuestro hogar, llenando mi vida de alegría.

A mi sobrina **Yhanlexis**, que con su inocencia y sus travesuras siempre me roba una sonrisa llenandome cada día de esperanzas. Dios te bendiga!

A mi novia, **Marisela Gutierrez**, por ser un ejemplo de lucha y que con su experiencia me ayuda a crecer como persona. Cada día me siento más orgulloso de ti... Moza Te Amo!

A mis **amigos**, Luis Segovia, Erkis Rodríguez y Carlos Linárez por haberme enseñado el valor de la amistad.

A los **profesores**, Carlos Lameda, Belkis López y María Luisa Capodieci quienes de manera muy profesional con paciencia y dedicación me guiaron para lograr este trabajo.

Al personal de biblioteca y bienestar estudiantil, de una manera muy especial a la señora **Hilda Mújica** que con su amor y entrega en sus labores representó una palabra de aliento y apoyo convirtiéndola en una amiga especial para mi.

A mis **compañeros**, Naudy González, Ricardo Hernández, Andrés Pérez y Aldemar Gómez por su valiosa ayuda y colaboración en el formato LaTeX.

A todos **GRACIAS!**

Resumen

La teoría de conjuntos difusos ha permitido establecer un marco para desarrollar una teoría de posibilidad, que incluye el concepto de medida de posibilidad. A través de este trabajo se presentarán nociones básicas sobre una teoría de posibilidad basada en conjuntos difusos. Así mismo, se mostrará la utilidad del desarrollo teórico a través de su aplicación en problemas prácticos.

Palabras clave: conjuntos difusos, teoría de la posibilidad, medidas de posibilidad.

Índice

Introducción

La teoría de la posibilidad es una de las teorías sobre incertidumbre dedicadas al manejo de información incompleta. Esta teoría se enfoca principalmente en la imprecisión que es intrínseca en lenguajes naturales y que se asume "posibilística". La misma ha sido formalizada con base a la teoría de conjuntos difusos, describiendo el concepto de una distribución de posibilidad como una restricción que puede ser asignada a una variable, definiendo una medida de posibilidad, basándose en el concepto de distribución de posibilidad.

A nivel mundial, varios investigadores se han dedicado ha construir los fundamentos sobre la teoría de la posibilidad. Sin embargo, existe escasa literatura, especialmente en castellano, que presente el tema de fundamentos sobre teoría de la posibilidad basada en conjuntos difusos.

Por ello se plantea elaborar una monografía que sirva de material para la divulgación y el estudio de este tema relativamente especializado.

Los conjuntos difusos, propuestos por Lotfi Zadeh en 1965 [1], son una clase de objetos con un grado de pertenencia continuo. A partir de este concepto se han generado importantes investigaciones y desarrollos en diferentes áreas del saber y ha permitido ampliar los límites de muchas disciplinas y crear nuevas teorías.

En 1978 [2], Zadeh estableció una teoría de la posibilidad basada en conjuntos difusos, en la que define el concepto de una distribución de posibilidad como una restricción difusa que actúa como una restricción flexible sobre los valores que pueden ser asignados a una variable.

En 1980 [3], Dubois y Prade exponen información que trata sobre grados de creencia, posibilidad, probabilidad de que un elemento pertenezca a un conjunto (nítido o difuso); incluye secciones sobre medidas difusas e integrales de Sugeno, la teoría de la posibilidad propuesta por Zadeh y comparación entre posibilidad y probabilidad.

En 1988 [4], Dubois, Prade y Testemale presentan una técnica para emparejar patrones de datos tomando en cuenta la imprecisión y la incertidumbre incluidos en los valores que tienen que ser comparados en un proceso de emparejamiento, en el marco de la teoría de posibilidad planteada por Zadeh [2].

En 1988 [5], García presenta una tesis doctoral en la que propone un modelo basado en la teoría de la posibilidad de Zadeh [2] y en los desarrollos de Dubois y Prade [3] para la resolución de incertidumbre en sistemas expertos.

En 1995 [6], Klir y Yuan desarrollan información didáctica sobre la teoría de la posibilidad formulada en términos de conjuntos difusos, que incluye aspectos sobre medidas difusas, teoría de la evidencia, teoría de la posibilidad, teoría de la posibilidad basada en conjuntos difusos y teoría de posibilidad versus teoría de probabilidad.

En 2001 [7], Zimmermann presenta información sobre modelado de la incertidumbre, dependiendo de las causas es estudiada por la teoría de la posibilidad basada en distribuciones de posibilidad, medidas de posibilidad y necesidad, probabilidad de eventos difusos y posibilidad versus probabilidad.

En 2002 [8], Ross y colaboradores detallan los métodos más populares usados para trabajar con varias formas de incertidumbre, incluyendo la teoría de conjuntos difusos, teoría de la probabilidad y teoría de la posibilidad. Incluye secciones sobre borrosidad versus probabilidad, chance versus vaguedad, estructura axiomática de probabilidad y lógica difusa, tratamiento de incertidumbre e imprecisión y posibilidad versus probabilidad.

En 2006 [9], Dubois presenta un recuento sobre varias ramas de la teoría de la posibilidad y en especial sobre la denominada teoría de la posibilidad cuantitativa y sus conexiones con la teoría de la probabilidad y la estadística.

En 2006 [10], Kikuchi y Chakroborty describen la teoría de la posibilidad, discuten las razones por las que se justifica su uso en el análisis de ciertos problemas de transporte e ilustran sobre como la teoría de la posibilidad puede ser usada en expresión de ignorancia, medida de ansiedad, comparación y jerarquización de valores aproximados.

El objetivo general de este trabajo es realizar un estudio sobre la teoría de la posibilidad basada en conjuntos difusos.

Sus objetivos específicos son:

1. Revisar aspectos fundamentales sobre conjuntos difusos y la teoría de la posibilidad.
2. Estudiar la teoría de la posibilidad basada en conjuntos difusos.
3. Mostrar la utilidad de la aplicación de la teoría de la posibilidad basada en conjuntos difusos.

El trabajo consta de tres capítulos, en los que se establecen conceptos fundamentales sobre distribución de posibilidad, medidas de posibilidad y necesidad, distribuciones de posibilidad de n-atributos y aplicación de la teoría de posibilidad.

Capítulo 1

Preliminares

El trabajo pionero de Wiener y Shannon sobre la teoría estadística de la comunicación ha conducido a una aceptación universal de la creencia de que la información es intrínsecamente de naturaleza estadística y como tal, deben ser tratados por los métodos proporcionados por la teoría de probabilidades. Sin lugar a dudas, el punto de vista estadístico ha aportado una visión profunda de los procesos fundamentales implicados en la codificación, transmisión y recepción de datos, y desempeñó un papel clave en el desarrollo de la comunicación moderna, detección y sistemas de telemedición. En años recientes, sin embargo, un número de otras aplicaciones importantes han pasado a primer plano en el que las principales cuestiones se centró no en la transmisión de información, sino en su significado. En tales aplicaciones, lo que importa es la capacidad de responder a las preguntas relativas a la información que se almacena en una base de datos como en el procesamiento del lenguaje natural, la representación del conocimiento, reconocimiento de voz, robótica, diagnóstico médico, análisis de eventos raros, la toma de decisiones en condiciones de incertidumbre, análisis de imagen, recuperación de información y áreas relacionadas.

Lofti Zadeh ha propuesto que cuando nuestra principal preocupación es con el significado de información -en lugar de con su medida- el marco adecuado para el análisis de la información es posibilista en lugar de la naturaleza probabilística, lo que implica que lo que es necesario para este tipo de análisis no es una teoría de la probabilidad sino una análoga -y aún diferente- teoría que podríamos llamar la *teoría de la posibilidad*.

La teoría de conjuntos difusos proporciona una base natural para la teoría de la posibilidad, la cual juega un papel similar al de la teoría de la medida en relación con la teoría de la probabilidad. Visto de esta perspectiva, una restricción difusa puede ser interpretada como una distribución de posibilidad, cuya función de pertenencia juega el papel de una función de distribución de posibilidad y una variable difusa asociada con una distribución de posibilidad de la misma manera que una variable aleatoria se asocia con una distribución de probabilidad. En general, sin embargo, una variable puede estar asociada tanto con una distribución de posibilidad y una distribución de probabilidad, con la conexión entre las dos expresiones como el principio de consistencia posibilidad/probabilidad. Este principio -que es una expresión de una conexión débil entre posibilidad y probabilidad- se describe con mayor detalle más adelante.

La importancia de la teoría de la posibilidad deriva del hecho de que -a diferencia de lo que se ha convertido en un supuesto ampliamente aceptado- gran parte de la información en la que las decisiones humanas se basan es posibilística y no de naturaleza probabilística. En particular, la borrosidad intrínseca del lenguaje natural -que es una consecuencia lógica de la necesidad de expresar la información de forma resumida- es, en lo principal, de origen posibilística. Partiendo de esta premisa, es posible construir un lenguaje universal en el que la traducción de una proposición expresada en un lenguaje natural toma la forma de un procedimiento para calcular la distribución de posibilidad de un conjunto de relaciones difusas en una base de datos.

A continuación, se exploran algunas de las propiedades elementales del concepto de una distribución de posibilidad, motivada principalmente por la aplicación de este concepto a la representación del significado en el lenguaje natural.

1.1. El concepto de una distribución de posibilidad

¿Que es una distribución de posibilidad? Es conveniente responder esta pregunta en términos de otro concepto, a saber, la de una *restricción difusa*, con la cual el concepto de una distribución de posibilidad guarda una estrecha relación.

Definición 1.1. Sea X una variable que toma valores en un universo de discurso U, con el elemento genérico de U denotado por u y $X = u$ lo que significa que a X se le asigna el valor de u, $u \in U$.

Definición 1.2. Sea F un subconjunto difuso de U que se caracteriza por una función de pertenencia μ_F. Entonces F es una *restricción difusa* sobre X (o asociada con X) si F actúa como una restricción flexible sobre los valores que deben ser asignados a X en el sentido que la asignación de un valor u a X tiene la forma:

$$X = u : \mu_F(u) \tag{1.1}$$

donde $\mu_F(u)$ es interpretado como el grado al que la restricción representada por F es satisfecha cuando u es asignado a X. Equivalentemente, (1.1) implica que $1 - \mu_F(u)$ es el grado para el cual la restricción en cuestión debe ser estirada a fin de permitir la asignación de u a X. Un punto a ser enfatizado es que un conjunto difuso por sí mismo no es una restricción difusa, para ser una restricción difusa debe actuar como una restricción sobre los valores de una variable.

Definición 1.3. Sea $R(X)$ una restricción difusa asociada con X. Entonces, para expresar que F juega el papel de una restricción difusa en relación a X, escribimos:

$$R(X) = F$$

Una ecuación de esta forma es llamada *ecuación de asignación relacional* porque representa la asignación de un conjunto difuso (o relación difusa) a la restricción asociada con X.

Para mostrar el concepto de una restricción difusa, considere una proposición de la forma $p \triangleq X\ es\ F$, donde X es el nombre de un objeto, una variable o una proposición, y F es el nombre de un subconjunto difuso de U, como en "Ana es muy inteligente", "X es un número pequeño", "Mary es rubia es muy cierto". En forma traducida, tal proposición puede ser expresada como:

$$R(A(X)) = F \tag{1.2}$$

donde $A(X)$ es un atributo implícito de X que toma valores en U, y (1.2) significa que la proposición $p \triangleq X\ es\ F$ tiene el efecto de asignar F a la restricción difusa sobre los valores de $A(X)$.

Como un ejemplo simple de (1.2), sea p la proposición "Juan es joven", en la que joven es un subconjunto difuso de $U = [0,\ 100]$ caracterizado por la función de pertenencia:

$$\mu_{joven}(u) = 1 - S(u; 20, 30, 40) \tag{1.3}$$

donde u es la edad numérica y la S-función es definida por

$$S(u;\alpha,\beta,\gamma) = \begin{cases} 0 & para \quad u \leq \alpha \\ 2\left(\frac{u-\alpha}{\gamma-\alpha}\right)^2 & para \quad \alpha \leq u \leq \beta \\ 1 - 2\left(\frac{u-\gamma}{\gamma-\alpha}\right)^2 & para \quad \beta \leq u \leq \gamma \\ 1 & para \quad u \geq \gamma \end{cases}$$

en el que el parámetro $\beta \triangleq (\alpha + \gamma)/2$ es el punto de intersección, es decir, $S = (\beta, \alpha, \beta, \gamma) = 0.5$. En este caso, el atributo implícito *A(X)* es Edad(Juan) y la proposición "Juan es joven" tiene la forma:

$$Juan\ es\ joven\ \rightarrow R(Edad(Juan)) = joven \tag{1.4}$$

Para relacionar el concepto de una restricción difusa a la de una distribución de posibilidad, se interpreta el miembro del lado derecho de (1.4) de la siguiente manera:

"Considere una edad numérica, digamos u = 28, cuyo grado de pertenencia al conjunto difuso "joven" (como se definió en (1.3))es aproximadamente 0.7. Primero nosotros interpretamos 0.7 como el grado de compatibilidad de 28 con el concepto etiquetado

joven. A continuación, se postula que la proposición "Juan es joven" convierte el significado de 0.7 del grado de compatibilidad de 28 con joven, al grado de posibilidad de que Juan tiene 28 dada la proposición "Juan es joven" En resumen, la compatibilidad de un valor de u con joven llega a convertirse en la posibilidad del valor de u dado "Juan es joven"

Expresado en terminos más generales, el concepto de una distribución de posibilidad puede ser definido como sigue. (Por simplicidad, asumimos que $A(X) = X$)

Definición 1.4. Sea F un subconjunto del universo de discurso U el cual es caracterizado por su función de pertenencia μ_F, con el grado de pertenencia $\mu_F(u)$, interpretado como la compatibilidad de u con el concepto etiquetado F.

Sea X una variable que toma valores en U, y sea F actuando como una restricción difusa, $R(X)$, asociada con X. Entonces la proposición "X *es* F", que se traduce en

$$R(X) = F \tag{1.5}$$

asocia una *distribución de posibilidad*, Π_X, con X, que es denotada como igual a $R(X)$, es decir,

$$\Pi_X = R(X) \tag{1.6}$$

Correspondientemente, la *función distribución de posibilidad* asociada con X (o la función de distribución de posibilidad de Π_X) se denota por π_X y se define numéricamente igual a la función de pertenencia F, es decir,

$$\pi_X \triangleq \mu_F$$

Así, $\pi_X(u)$, la *posibilidad* que $X = u$, se denota como igual a $\mu_F(u)$.

En vista de (1.6), la ecuación de asignación relacional (1.5) se puede expresar equivalentemente en la forma

$$\Pi_X = F \tag{1.7}$$

dejando en evidencia que la proposición $p \triangleq X$ es F tiene el efecto de asociar X con una distribución de posibilidad Π_X que, por (1.6), es igual a F. Cuando se expresa en la forma de (1.7), una ecuación de asignación relacional define una *ecuación de asignación de posibilidad*, con el entendimiento que Π_X es inducida por p.

Ejemplo 1.1. Sea U el universo de los números enteros positivos y sea F el conjunto difuso de los enteros pequeños definidos por $(+ \triangleq \textit{union})$

$$\textit{Enteros pequeños} = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6$$

Entonces, la proposición "X es un entero pequeño" asocia con X la distribución de posibilidad

$$\Pi_X = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6 \tag{1.8}$$

en el que un término tal como 0.8/3 significa que la posibilidad que X es 3, dado que X es un entero pequeño, es 0.8.

Hay varios puntos importantes relacionados con la definición anterior que están en necesidad de comentarios.

En primer lugar, (1.6) implica que la distribución de posibilidad Π_X puede ser considerada como una interpretación del concepto de una restricción difusa y, por consiguiente, que la estructura matemática de la teoría de conjuntos difusos -y, sobre todo, el cálculo de las restricciones difusas- proporciona una base para la manipulación de las distribuciones de posibilidad por las reglas de este cálculo.

En segundo lugar, la definición implica la suposición de que nuestra percepción intuitiva de las formas en que se combinan las posibilidades está de acuerdo con las reglas de combinación de restricciones difusas. Aunque la validez de esta hipótesis no ha sido probada, parece que hay un acuerdo estrecho entre dichas operaciones básicas como la unión y la intersección de conjuntos difusos, por un lado, y las distribuciones de posibilidad asociadas con las disyunciones y conjunciones de proposiciones de la forma

"*X es F*".

En tercer lugar, la definición de $\pi_X(u)$ implica que el grado de posibilidad puede ser cualquier número en el intervalo [0,1] en lugar de sólo 0 ó 1. A este respecto, cabe señalar que la existencia de grados intermedios de posibilidad está implícita en estas proposiciones se encuentran comúnmente como "Hay una ligera posibilidad de que Marilyn es muy rica", "Es muy posible que Juan Pablo se promoverá,...", "es casi imposible encontrar una aguja en un pajar", etc.

Se podría argumentar, por supuesto, que una caracterización de un grado intermedio de la posibilidad de una etiqueta como "baja posibilidad" es comúnmente destinada a ser interpretada como "baja probabilidad". Sin lugar a dudas, este es frecuentemente el caso en el discurso cotidiano. Sin embargo, existe una diferencia fundamental entre la probabilidad y la posibilidad que, una vez comprendido mejor, dará lugar a una diferenciación más cuidadosa entre las caracterizaciones de grados de posibilidad versus grados de probabilidad especialmente en el discurso legal, el diagnóstico médico, idiomas sintéticos y, más generalmente, aquellas aplicaciones en las que un alto grado de precisión de significado es muy deseable.

Podemos presentar la diferencia entre probabilidad y posibilidad por un ejemplo sencillo.

Ejemplo 1.2. Consideremos la declaración "Antonio comió X huevos para el desayuno", con X tomando valores en $U = \{1, 2, 3, 4, ..., k\}$. Nosotros debemos asociar una distribución de posibilidad, con X interpretando $\pi_X(u)$ como el grado de facilidad con el cual Antonio puede comer u huevos. Similarmente, debemos asociar una distribución de probabilidad con X interpretando $P_X(u)$ como la probabilidad de Antonio comiendo u huevos para el desayuno. Asumiendo que empleamos algunos criterios explícitos o implícitos para evaluar el grado de facilidad con que Antonio puede comer u huevos para el desayuno, los valores de $\pi_X(u)$ y $P_X(u)$ pueden ser mostrados en la Tabla 1.1:

u	1	2	3	4	5	6	7	8
$\pi_X(u)$	1	1	1	1	0.8	0.6	0.4	0.2
$P_X(u)$	0.1	0.8	0.1	0	0	0	0	0

Tabla 1.1: Distribuciones de posibilidad y probabilidad asociadas con X.

Se observa que, mientras la posibilidad de que Antonio puede comer 3 huevos para el desayuno es 1, la probabilidad de que pueda hacerlo podría ser muy pequeño, por ejemplo, 0.1. Por lo tanto, un alto grado de posibilidad no implica un alto grado de probabilidad, ni un bajo grado de probabilidad implica un bajo grado de posibilidad. Sin embargo, si un evento es imposible, está obligado a ser improbable. Esta conexión heurística entre las posibilidades y probabilidades puede que se indique en la forma de lo que podría llamarse el principio de consistencia posibilidad/probabilidad, a saber:

Si una variable X puede tomar valores $u_1, \ldots, u_n$ con respectivas posibilidades $\Pi = (\pi_1, \ldots, \pi_n)$ y probabilidades $P = (p_1, \ldots, p_n)$, entonces el grado de consistencia de la distribución de probabilidad P con la distribución de posibilidad Π es expresada por ($+ \triangleq$ *suma aritmética*)

$$\gamma = \pi_1 p_1 + \ldots + \pi_n p_n$$

Se debe entender, por supuesto, que el principio de consistencia posibilidad/probabilidad no es una ley exacta o una relación que es intrínseca en los conceptos de posibilidad y probabilidad. Más bien es una formalización aproximada de la observación heurística que una disminución de la posibilidad de un evento tiende a disminuir su probabilidad -pero no viceversa. En este sentido, el principio es de uso en situaciones en las que lo que se conoce acerca de una variable X es su distribución de posibilidad -en lugar de su probabilidad. En tales casos -que se producen con mucha más frecuencia que aquellos en los que lo contrario es cierto- el principio de consistencia posibilidad/-probabilidad proporciona una base para el cálculo de la distribución de posibilidad de la distribución de probabilidad de X. Tales cálculos juegan un papel particularmente importante en la toma de decisiones en condiciones de incertidumbre.

En el ejemplo descrito anteriormente, la posibilidad de que X asuma un valor u se interpreta como el grado de facilidad con el que u puede ser asignado a X, por ejemplo, el grado de facilidad con que Antonio puede comer u huevos para el desayuno. Se debe entender, sin embargo, que este "grado de facilidad" puede tener o no realidad física. Por lo tanto, la proposición "Juan es joven" induce a una distribución de posibilidad cuya función de distribución de posibilidad es expresada por (1.3). En este caso, la posibilidad que la variable Edad(Juan) puede tomar el valor 28 es 0.7, con 0.7 representando el grado de facilidad con que 28 debe ser asignado a Edad(Juan) dada la flexibilidad de la restricción difusa etiquetada joven. Así, en este caso "el grado de facilidad" tiene una significancia figurativa en lugar de física.

1.2. Medidas de posibilidad

Una información adicional sobre la distinción entre probabilidad y posibilidad puede ser adquirida mediante la comparación del concepto de una medida de posibilidad con el concepto familiar de una medida de probabilidad. Más específicamente, sea A un subconjunto no difuso de U y sea Π_X una distribución de posibilidad asociada con una variable X que toma valores en U. Entonces, la medida de posibilidad de A, $\pi(A)$, se define como un número en [0,1] dada por

$$\pi(A) \triangleq \sup_{u \in A} \pi_X(u)$$

donde $\pi_X(u)$ es la función distribución de posibilidad de Π_X. Este número, a continuación debe ser interpretado como la posibilidad *que un valor de X pertenece a A*, es decir

$$\begin{aligned} Pos\{X \in A\} &\triangleq \pi(A) \\ &\triangleq \sup_{u \in A} \pi_X(u) \end{aligned} \tag{1.9}$$

Cuando A es un conjunto difuso, la pertenencia de un valor de X a A no es significativa. Una definición más general de medida de posibilidad que extiende (1.9) a conjuntos difusos es la siguiente:

Definición 1.5. Sea A un subconjunto difuso de U y sea Π_X una distribución de posibilidad asociada con una variable X que toma valores en U. La medida de posibilidad de A, $\pi(A)$, se define por:

$$\begin{aligned} Pos\{X \; es \; A\} &\triangleq \pi(A) \\ &\triangleq \sup_{u \in U} \mu_A(u) \wedge \pi_X(u) \end{aligned} \tag{1.10}$$

donde "X es A" sustituye a "$X \in A$" en (1.9), μ_A es la función de pertenencia de A, y $\wedge$ representa como siempre, el mínimo.

Según Zadeh, el término *Altura* de un conjunto difuso, se define como el supremo de la función de pertenencia, por ello (1.10) puede ser expresada por la ecuación

$$\pi(A) \triangleq Altura(A \cap \Pi_X) \tag{1.11}$$

Como una representación sencilla, consideremos la distribución de posibilidad (1.8) que es inducida por la proposición "X es un entero pequeño". En este caso, sea A el conjunto {3,4,5}. Entonces,

$\pi(A) = 0.8 \vee 0.6 \vee 0.4 = 0.8$

donde $\vee$ representa como es usual, el máximo.

Por otro lado, si A es el conjunto difuso de enteros que no son pequeños, es decir, el complemento de los enteros pequeños.

$$A \triangleq 0.2/3 + 0.4/4 + 0.6/5 + 0.8/6 + 1/7 + \ldots$$

Luego,

$$Pos\{X \; no \; es \; un \; entero \; peque\tilde{n}o\} = Altura(0.2/3 + 0.4/4 + 0.6/5)$$

$$= 0.4 \tag{1.12}$$

Cabe destacar que (1.12) es una consecuencia inmediata de la afirmación

$$X\ es\ F \Longrightarrow Pos\{X\ es\ A\} = Altura(F \cap A),$$

que se deduce de (1.7) y (1.11).

En particular, si A es un conjunto normal difuso (es decir, Altura(A) = 1), entonces, como debería esperarse

$$X\ es\ A \Longrightarrow Pos\{X\ es\ A\} = 1$$

Sean A y B subconjuntos difusos arbitrarios de U, entonces de la definición de la medida de posibilidad de un conjunto difuso (1.10), se sigue que

$$\pi(A \cup B) = \pi(A) \vee \pi(B) \tag{1.13}$$

Por comparación, la relación correspondiente para medidas de probabilidad de A, B y $A \cup B$ (si existen) es

$$P(A \cup B) \leq P(A) + P(B)$$

y, si A y B son disjuntos (es decir, $\mu_A(u)\mu_B(u) \equiv 0$),

$$P(A \cup B) = P(A) + P(B) \tag{1.14}$$

que expresa la propiedad de aditividad básica de medidas de probabilidad.

Así, en contraste con la medida de probabilidad, la medida de posibilidad no es aditiva. En cambio, tiene la propiedad expresada por (1.13), que puede considerarse como un análogo de (1.14) con + reemplazado por $\vee$.

De forma similar, la medida de posibilidad de la intersección de A y B está relacionada con la de A y B por

$$\pi(A \cap B) \leq \pi(A) \wedge \pi(B) \tag{1.15}$$

En particular, si A y B son disjuntos, (1.15) se mantiene con el signo de igualdad, es decir,

$$\pi(A \cap B) = \pi(A) \wedge \pi(B) \tag{1.16}$$

Por comparación, en el caso de medidas de probabilidad, tenemos

$$P(A \cap B) \leq P(A) \wedge P(B)$$

y

$$P(A \cap B) = P(A)P(B) \tag{1.17}$$

si A y B son independientes y no difusos. Como en el caso de (1.13), (1.16) es análogo a (1.17), en el que el producto corresponde al mínimo.

1.3. Posibilidad e información

Si p es una proposición de la forma $p \triangleq X$ es F, que se traduce en la ecuación de asignación de posibilidad

$$\Pi_{A(X)} = F$$

donde F es un subconjunto difuso de U y *A(X)* es un atributo implícito de X tomando valores en U entonces la información transmitida por p, *I(p)* se puede identificar con la distribución de posibilidad, $\Pi_{A(X)}$, de la variable difusa *A(X)*. Así, la conexión entre *I(p)*, $\Pi_{A(X)}$, *R(A(X))* y F es expresada por:

$$I(p) \triangleq \Pi_{A(X)}$$

donde

$$\Pi_{A(X)} = R(A(X)) = F$$

Por ejemplo, si la proposición $p \triangleq$ *Juan es joven*, se traduce en la ecuación de asignación de posibilidad

$$\Pi_{Edad(Juan)} = joven$$

donde joven se define por (1.3), luego

$$I(Juan\ es\ joven) = \Pi_{Edad(Juan)}$$

en la que la función distribución de posibilidad de Edad(Juan) está dada por

$$\pi_{Edad(Juan)}(u) = 1 - S(u; 20, 30, 40) \qquad u \in [0, 100]$$

De la definición de *I(p)* se sigue que si $p \triangleq X\ es\ F$ y $q \triangleq X\ es\ G$, entonces p es al menos tan informativa como q, expresada como $I(p) \geq I(q)$, si $F \subset G$. Así, tenemos un orden parcial de *I(p)* definido por

$$F \subset G \Longrightarrow I(X\ es\ F) \geq I(X\ es\ G)$$

lo cual implica que mientras más restringida es una distribución de posibilidad, más informativa es la proposición con la que está asociada. Por ejemplo, como *muy alto* $\subset$ *alto*, tenemos que

$$I(Lucia\ es\ muy\ alta) \geq I(Lucia\ es\ alta)$$

1.4. Medidas de necesidad

En teoría de la posibilidad, se define otra medida adicional, que utiliza la relación conjuntiva y, en un sentido, es dual a la medida de posibilidad:

$$N(A \cap B) = \text{mín}(N(A), N(B))$$

Entonces N es llamada la medida de necesidad. $N(A) = 1$ indica que A es necesariamente cierto (A es seguro). La relación dual de posibilidad y necesidad exige que

$$\pi(A) = 1 - N(\complement A), \qquad \forall A \subseteq \Omega$$

donde $\complement A$ es el complemento de A.

Las medidas de necesidad satisfacen la condición

$$\text{mín}(N(A), N(\complement A)) = 0$$

Las relaciones entre medidas de posibilidad y medidas de necesidad satisfacen las siguientes condiciones:

$$\pi(A) \geq N(A), \qquad \forall A \subseteq \Omega$$

$$N(A) > 0 \Rightarrow \pi(A) = 1$$

$$\pi(A) < 1 \Rightarrow N(A) = 0$$

Aquí, Ω siempre se asume como un conjunto finito.

Ejemplo 1.3. La Tabla 1.2 presenta las notas de seis estudiantes (desde A hasta E) en examenes escritos y sus funciones de pertenencia. Primero observemos que la función de pertenencia para las notas de los estudiantes 3 y 4 no es una función de posibilidad, ya que $g(\Omega) \neq 1$.

Estudiante \ Notas	A	B	C	D	E
1	0.8	1	0.7	0	0
2	1	0.8	0.6	0.1	0
3	0.6	0.7	0.9	0.1	0
4	0	0.8	0.9	0.5	0
5	0	0	0.3	1	0.2
6	0.3	1	0.3	0	0

Tabla 1.2: Notas de estudiantes y sus funciones de pertenencia.

Ahora podemos hacer diferentes preguntas:

1. ¿Qué tan confiable es la declaración del estudiante 1 que va a obtener B en su próximo exámen?
   En este caso, "A" es $\{B\}$ y "$\complement A$" es $\{A,\ C,\ D,\ E\}$

$$\begin{aligned} \pi(A) &= 1 \\ N(A) &= \text{mín}\{1 - \pi_i\} \\ &= \text{mín}\{0.2, 0.3, 1, 1\} \\ &= 0.2 \end{aligned}$$

   Por lo tanto, la posibilidad de que el estudiante 1 obtendrá B es $\pi = 1$, la necesidad es $N = 0.2$

2. Si queremos conocer la verdad de la declaración "Cualquiera de los estudiantes 1 o 2 alcanzará A o B" nuestro Ω tiene que ser definido de modo distinto. Este ahora, contiene los elementos de las primeras dos filas. Los resultados son:

$$\begin{aligned} \pi(A) &= \pi(Estudiante\ 1\ A\ o\ B\ o\ Estudiante\ 2\ A\ o\ B) \\ &= 1 \\ N(A) &= 0.3 \end{aligned}$$

3. Finalmente determinemos la verdad de la declaración "El estudiante 1 alcanzará C".
   En este caso,

$$\begin{aligned} \pi(A) &= 0.7 \\ N(A) &= 0 \end{aligned}$$

Capítulo 2

Distribuciones de posibilidad de n-atributos

Al afirmar que la traducción de una proposición de la forma $p \triangleq X\ es\ F$, es expresada por

$$X\ es\ F \longrightarrow R(A(X)) = F$$

o, equivalentemente,

$$X\ es\ F \longrightarrow \Pi_{A(X)} = F \tag{2.1}$$

estamos asumiendo tácitamente que p contiene un atributo implícito único *A(X)* cuya distribución de posibilidad está dada por el miembro del lado derecho de (2.1).

En general, p puede contener n atributos implicitos $A_1(X), \ldots, A_n(X)$, con $A_i(X)$ tomando valores en U_i, $i = 1, ..., n$. En este caso, la proposición $p \triangleq X\ es\ F$, donde F es una relación difusa en el producto cartesiano $U = U_1 \times \ldots \times U_n$, asume la forma

$$X\ es\ F \longrightarrow R(A_1(X), \ldots, A_n(X)) = F \tag{2.2}$$

o, de forma semejante,

$$X\ es\ F \longrightarrow \Pi_{(A_1(X),\ldots,A_n(X))} = F$$

donde $R(A_1(X), \ldots, A_n(X)) = F$ es una restricción difusa de n-atributos y $\Pi_{(A_1(X),\ldots,A_n(X))}$ es una distribución de posibilidad de n-atributos que es inducida por p. En forma correspondiente, la función distribución de posibilidad de n-atributos inducida por p está dada por

$$\pi_{(A_1(X),\ldots,A_n(X))}(u_1,\ldots,u_n) = \mu_F(u_1,\ldots,u_n), \qquad (u_1,\ldots,u_n) \in U$$

donde μ_F es la función de pertenencia de F. En particular, si F es un producto cartesiano de relaciones difusas de n-atributos $F_1,\ldots,F_n$ entonces el miembro del lado derecho de (2.2) se descompone en un sistema de ecuaciones de asignación relacional de n-atributos, es decir,

$$\begin{aligned} X \ es \ F \quad \longrightarrow \quad R(A_1(X)) &= F_1 \\ R(A_2(X)) &= F_2 \\ \vdots \qquad & \quad \vdots \\ R(A_n(X)) &= F_n \end{aligned}$$

Correspondientemente,

$$\Pi_{(A_1(X),\ldots,A_n(X))} = \Pi_1 \times \ldots \times \Pi_n$$

y

$$\pi_{(A_1(X),\ldots,A_n(X))}(u_1,\ldots,u_n) = \pi_{A_1(X)}(u_1) \wedge \ldots \wedge \pi_{A_n(X)}(u_n)$$

donde

$$\pi_{A_i(X)}(u_i) = \mu_{F_i}(u_i), \qquad u_i \in U_i, \quad i = 1,\ldots,n$$

y $\wedge$ denota el mínino.

Ejemplo 2.1. Consideremos la proposición $p \triangleq$ *la alfombra es amplia*, en la que amplia es una relación difusa cuyo cuadro es de la forma que se muestra en la Tabla 2.1 (con longitud y anchura expresada en centímetros).

Amplitud	*Anchura*	*Longitud*	μ
	250	300	0.6
	250	350	0.7
	.	.	.
	300	400	0.8
	.	.	.
	400	600	1

Tabla 2.1: Cuadro de Amplitud.

En este caso, la ecuación (2.2) nos lleva a la ecuación de asignación de posibilidad

$$\Pi_{(Anchura(alfombra),Longitud(alfombra))} = Amplitud \tag{2.3}$$

lo cual implica que si la compatibilidad de una alfombra cuya anchura es, digamos, 250 cm y la longitud es 350 cm con "Amplitud de la alfombra" es 0.7. Luego, la posibilidad que la anchura de la alfombra es 250 cm y la longitud sea 350 cm (dada la proposición $p \triangleq la\ alfombra\ es\ amplia$) es 0.7.

Ahora, si Amplitud se define como,

$$Amplitud = Largo \times Ancho$$

donde largo y ancho son relaciones difusas, entonces (2.3) se descompone en las ecuaciones de asociación de posibilidad

$$\Pi_{Longitud(alfombra)} = Largo$$

y

$$\Pi_{Anchura(alfombra)} = Ancho$$

El cuadro de la Tabla 2.2 nos muestra el grado de pertenencia para cada atributo: largo y ancho.

Ancho	*Anchura*	μ	*Largo*	*Longitud*	μ
	250	0.7		300	0.6
	300	0.8		350	0.7
	350	0.8		400	0.8
	.	.		.	.
	400	1		500	1

Tabla 2.2: Cuadro de Largo y Ancho en centímetros.

2.1. Distribuciones de posibilidad marginal

El concepto de una distribución de posibilidad marginal guarda una relación estrecha con el concepto de una restricción difusa marginal, que a su vez es análoga al concepto de una distribución de probabilidad marginal.

Más específicamente, tomemos $X = (X_1, \ldots, X_n)$ una variable difusa de n-atributos que toma valores en $U = U_1 \times \ldots \times U_n$, y sea Π_X una distribución de posibilidad asociada con X, donde $\pi_X(u_1, \ldots, u_n)$ denota la función distribución de posibilidad de Π_X.

Sea $q \triangleq (i_1, \ldots, i_k)$ una subsucesión de la sucesión de índices $(1, \ldots, n)$ y sea $X_{(q)}$ la variable difusa de q-atributos $X_{(q)} \triangleq (X_{i_1}, \ldots, X_{i_k})$. La *distribución de posibilidad marginal* $\Pi_{X_{(q)}}$ es una distribución de posibilidad asociada con $X_{(q)}$ que es *inducida* por Π_X como la proyección de Π_X sobre

$$U_{(q)} \triangleq U_{i_1} \times \ldots \times U_{i_k}$$

Así, por definición,

$$\Pi_{X_{(q)}} \triangleq \mathrm{Pro}_{U_{(q)}} \Pi_X$$

lo cual implica que la función distribución de probabilidad de $X_{(q)}$ se relaciona con la de X por

$$\pi_{X_{(q)}}(u_{(q)}) = \bigvee_{u_{(q')}} \pi_X(u)$$

donde $u_{(q)} \triangleq (u_{i_1}, \ldots, u_{i_k})$, $q' \triangleq (j_1, \ldots, j_m)$ es una subsucesión de $(1, \ldots, n)$ la cual es complementaria para q (es decir, si $n = 5$ y $q \triangleq (i_1, i_2) = (2, 4)$, entonces $q' = (j_1, j_2, j_3) = (1, 3, 5)$), $u_{(q')} \triangleq (u_{j_1}, \ldots, u_{j_m})$ y $\vee_{u_{(q')}}$ denota el supremo sobre $(u_{j_1}, \ldots, u_{j_m}) \in U_{j_1} \times \ldots \times U_{j_m}$.

Ejemplo 2.2. Supongamos que $U_1 = U_2 = U_3 = \{a, b\}$ y el cuadro de Π_X es dado por

Π_X	X_1	X_2	X_3	π
	a	a	a	0.8
	a	a	b	1
	b	a	a	0.6
	b	a	b	0.2
	b	b	b	0.5

Tabla 2.3: Cuadro de Π_X.

Luego,

$$\Pi_{(x_1,x_2)} = \mathrm{Pro}_{x_1 \times x_2} \Pi_X = 1/(a, a) + 0.6/(b, a) + 0.5/(b, b)$$

que en forma tabular se lee

$\Pi_{(X_1,X_2)}$	X_1	X_2	π
	a	a	1
	b	a	0.6
	b	b	0.5

Tabla 2.4: Cuadro de $\Pi_{(x_1,x_2)}$.

Entonces, de $\Pi_{(X)}$ se sigue que la posibilidad que $X_1 = b, X_2 = a$ y $X_3 = b$ es 0.2, mientras de $\Pi_{(X_1,X_2)}$ se sigue que la posibilidad de $X_1 = b, X_2 = a$ es 0.6.

Análogamente, con el concepto de independencia de variables aleatorias, las variables difusas $X_{(q)} \triangleq (X_{i_1}, \ldots, X_{i_k})$ y $X_{(q')} \triangleq (X_{j_1}, \ldots, X_{j_m})$ son no-interactivas si y sólo si la distribución de posibilidad asociada con $X = (X_1, \ldots, X_n)$ es el producto cartesiano de las distribuciones de posibilidad asociadas con $X_{(q)}$ *y* $X_{(q')}$, es decir,

$$\Pi_X = \Pi_{X_{(q)}} \times \Pi_{X_{(q')}}$$

o equivalentemente,

$$\pi_X(u_1, \ldots, u_n) = \pi_{X_{(q)}}(u_{i_1}, \ldots, u_{i_k}) \wedge \pi_{X_{(q)}}(u_{j_1}, \ldots, u_{j_m})$$

En particular, las variables $X_1, \ldots, X_n$ son no-interactivas si y sólo si

$$\Pi_X = \Pi_{X_1} \times \Pi_{X_2} \times \ldots \times \Pi_{X_n}$$

El significado intuitivo de no-interacción se puede aclarar por un ejemplo sencillo.

Ejemplo 2.3. Supongamos que $X \triangleq (X_1, X_2)$ y X_1, X_2 son no-interactivas, es decir,

$$\pi_X(u_1, u_2) = \pi_{X_1}(u_1) \wedge \pi_{X_2}(u_2)$$

Además, supongamos que para algunos valores particulares de u_1 y u_2, $\pi_{X_1}(u_1) = \alpha_1$, $\pi_{X_2}(u_2) = \alpha_2 < \alpha_1$ y por lo tanto, $\pi_X(u_1, u_2) = \alpha_2$.

Ahora, si el valor de $\pi_{X_1}(u_1)$ se incrementa a $\alpha_1 + \delta_1$, $\delta_1 > 0$, no es posible disminuir el valor de $\pi_{X_2}(u_2)$ por una cantidad positiva, por ejemplo δ_2, de tal manera que el valor de $\pi_X(u_1, u_2)$ se mantenga sin cambios.

En este sentido, un aumento de la posibilidad de u_1, no puede ser compensado por una disminución en la posibilidad de u_2 y viceversa. Por lo tanto, en esencia, no interacción puede ser visto como una forma de no compensación en el que una variación en uno o más componentes de una distribución de posibilidad no puede ser compensado por las variaciones en los componentes complementarios.

En la manipulación de las distribuciones de posibilidad, es conveniente emplear un tipo de representación simbólica que se utiliza comúnmente en el caso de los conjuntos difusos. Específicamente, asuma, por simplicidad, que $U_1, \ldots, U_n$ son conjuntos finitos, y sea $r^i \triangleq (r_1^i, \ldots, r_n^i)$ denotando una n-upla de valores tomados de $U_1, \ldots, U_n$ respectivamente. Por otra parte, sea π_i la posibilidad de r^i y sea la n-upla $(r_1^i, \ldots, r_n^i)$ que se puede escribir como la cadena $r_1^i \ldots r_n^i$.

Usando esta notación, una distribución de posibilidad Π_X se puede expresar en forma simbólica

$$\Pi_X = \sum_{i=1}^{N} \pi_i (r_1^i r_2^i \ldots r_n^i) \tag{2.4}$$

O, en caso de que se necesite un simbolo separador, como

$$\Pi_X = \sum_{i=1}^{N} \pi_i / (r_1^i r_2^i \ldots r_n^i)$$

donde N es el número de n-uplas en el cuadro de Π_X y la sumatoria se debe interpretar como la unión de los conjuntos difusos $\pi_i/(r_1^i \ldots r_n^i)$. Como ilustración, en la notación de (2.4), la distribución de posibilidad definida en la Tabla 2.3 se lee como

$$\Pi_X = 0.8aaa + 1aab + 0.6baa + 0.2bab + 0.5bbb \tag{2.5}$$

La ventaja de esta notación es que permite que las distribuciones de posibilidad sean manipuladas de manera semejante a las formas lineales en n-variables, con el entendido que, si r y s son dos-uplas y α y β son sus respectivas posibilidades, entonces

$$\alpha r + \beta r = (\alpha \vee \beta) r$$

$$\alpha r \cap \beta r = (\alpha \wedge \beta) r$$

y

$$\alpha r \times \beta s = (\alpha \wedge \beta) rs$$

donde rs denota la concatenación de r y s.

Ejemplo 2.4. Si $\Pi_X = 0.8aa + 0.5ab + 1bb$ y $\Pi_Y = 0.9ba + 0.6bb$, entonces

$$\Pi_X + \Pi_Y = 0.8aa + 0.5ab + 0.9ba + 1bb$$

$$\Pi_X \cap \Pi_Y = 0.6bb$$

y

$$\Pi_X \times \Pi_Y = 0.8aaba + 0.5abba + 0.9bbba + 0.6aabb + 0.5abbb + 0.6bbbb$$

Para obtener la proyección de una distribución de posibilidad Π_X sobre $U_{(q)} \triangleq (U_{i_1}, \ldots, U_{i_k})$, es suficiente establecer los valores de $X_{j_1}, \ldots, X_{j_m}$ en cada upla de Π_X igual a la cadena nula Λ (es decir, la identidad multiplicativa).

Ejemplo 2.5. La proyección de la distribución de posibilidad definida por la Tabla 2.3 sobre $U_1 \times U_2$ está dada por

$$\begin{aligned} \mathrm{Pro}_{U_1 \times U_2} \Pi_X &= 0.8aa + 1aa + 0.6ba + 0.2ba + 0.5bb \\ &= 1aa + 0.6ba + 0.5bb \end{aligned}$$

que está de acuerdo con la Tabla 2.4.

2.2. Distribuciones de posibilidad condicionadas

En la teoría de las posibilidades, el concepto de una distribución de posibilidad condicionada desempeña un papel que es análogo -aunque no completamente- a la de una distribución de posibilidad condicional en la teoría de probabilidades.

Más concretamente, sea una variable $X = (X_1, \ldots, X_n)$ asociada con una distribución de posibilidad Π_X, con Π_X caracterizada por una función distribución de posibilidad $\pi_X(u_1, \ldots, u_n)$ que asigna a cada n-upla en $U_1 \times \ldots \times U_n$ su posibilidad $\pi_X(u_1, \ldots, u_n)$.

Sean $q = (i_1, \ldots, i_k)$ y $s = (j_1, \ldots, j_m)$ subsucesiones de la sucesion de índices $(1, \ldots, n)$ y sea $(a_{j_1}, \ldots, a_{j_m})$ una n-upla de valores asignados a $X_{(q')} = (X_{j_1}, \ldots, X_{j_m})$.

Por definición, la *distribución de posibilidad condicionada* de

$$X_{(q)} \triangleq (X_{i_1}, \ldots, X_{i_k})$$

dada

$$X_{(q')} = (a_{j_1}, \ldots, a_{j_m})$$

es una distribución de posibilidad expresada como

$$\Pi_{X_{(q)}}[X_{j_1} = a_{j_1}; \ldots; X_{j_m} = a_{j_m}]$$

cuya función distribución de posibilidad está dada por

$$\pi_{X_{(q)}}(u_{i_1}, \ldots, u_{i_k} \mid X_{j_1} = a_{j_1}; \ldots; X_{j_m} = a_{j_m})$$

$$\triangleq \pi_X(u_1, \ldots, u_n) \mid u_{j_1} = a_{j_1}, \ldots, u_{j_m} = a_{j_m}$$

Como un ejemplo sencillo, en el caso de (2.5), tenemos

$$\Pi_{(X_2,X_3)}[X_1 = a] = 0.8aa + 1ab \tag{2.6}$$

como la expresión para la distribución de posibilidad condicionada de (X_2, X_3) dado $X_1 = a$.

Una expresión equivalente para la distribución de posibilidad condicionada que hace más clara la conexión entre

$$\Pi_{X_{(q)}}[X_{j_1} = a_{j_1}; \ldots; X_{j_m} = a_{j_m}]$$

y Π_X se puede derivar como sigue.

Supongamos que $\Pi_X[X_{j_1} = a_{j_1}; \ldots; X_{j_m} = a_{j_m}]$ es una distribución de posibilidad que consiste en aquellos términos mencionados en (2.4) en que el elemento j_1 es a_{j_1}, el elemento j_2 es $a_{j_2}, \ldots,$ y el elemento j_m es a_{j_m}. Por ejemplo, en el caso de (2.5)

$$\Pi_X[X_1 = a] = 0.8aaa + 1aab \tag{2.7}$$

Expresado en la notación anterior, la distribución de posibilidad condicionada de $X_{(q)} = (X_{i_1}, \ldots, X_{i_k})$ dada $X_{j_1} = a_{j_1}, \ldots, X_{j_m} = a_{j_m}$ se puede escribir como

$$\Pi_{X_{(q)}}[X_{j_1} = a_{j_1}; \ldots ; X_{j_m} = a_{j_m}]$$

$$= \mathrm{Pro}_{U_{(q)}} \Pi_X [X_{j_1} = a_{j_1}; \ldots ; X_{j_m} = a_{j_m}] \tag{2.8}$$

que deja de manifiesto que $\Pi_{X_{(q)}}$ (condición de que $X_s = a_s$) es una distribución de posibilidad marginal inducida por Π_X (condicionado de que $X_s = a_s$). Así, mediante el empleo de (2.7) y (2.8), obtenemos

$$\Pi_{(X_2,X_3)}[X_1 = a] = 0.8aa + 1ab$$

lo que concuerda con (2.6).

En la discusión anterior, se ha supuesto que la distribución de posibilidad de $X = (X_1, \ldots, X_n)$ está condicionada a los valores asignados a un subconjunto especificado, $X_{(s)}$, de las variables constitutivas de X. En un contexto más general, lo que podría ser especificado es una distribución de posibilidad asociado con $X_{(s)}$ en lugar de los valores de $X_{j_1}, \ldots, X_{j_m}$.

En tales casos, diremos que Π_X se particulariza al especificar que $\Pi_{X_{(s)}} = G$, donde G es una distribución de posibilidad de m-atributos dados. Cabe señalar que en el presente contexto $\Pi_{X_{(s)}}$ es una distribución de posibilidad dada más que una distribución marginal que es inducida por Π_X.

Para analizar este caso, es conveniente asumir -con el fin de simplificar la notación- que $X_{j_1} = X_1, X_{j_2} = X_2, \ldots, X_{j_m} = X_m, m < n$. Sea $\bar{G}$ la *extensión cilíndrica* de G, es decir, la distribución de posibilidad definida por

$$\bar{G} \triangleq G \times U_{m-1} \times \ldots \times U_n$$

lo cual implica que

$$\mu_{\bar{G}}(u_1, \ldots, u_n) \triangleq \mu_G(u_1, \ldots, u_m), \quad u_j \in U_j, \quad j = 1, ..., n$$

donde μ_G es la función de pertenencia de la relación difusa G.

La suposición que nos dan Π_X y G es equivalente a suponer que nos den la intersección $\Pi_X \cap \bar{G}$. De esta intersección, luego, podemos deducir la distribución de posibilidad particularizada $\Pi_{X_{(q)}}[\Pi_{X_{(s)}} = G]$ por proyección sobre $U_{(q)}$.

Así,

$$\Pi_{X_{(q)}}[\Pi_{X_{(s)}} = G] = \mathrm{Pro}_{U_{(q)}} \Pi_X \cap \bar{G} \tag{2.9}$$

Equivalentemente, el miembro de la izquierda de (2.9) puede ser considerado como la composición de Π_X y G.

Ejemplo 2.6. Consideremos la distribución de posibilidad definida por (2.5) y supongamos que

$$G = 0.4aa + 0.8ba + 1bb$$

Luego,

$$\bar{G} = 0.4aaa + 0.4aab + 0.8baa + 0.8bab + 1bba + 1bbb$$

$$\Pi_X \cap \bar{G} = 0.4aaa + 0.4aab + 0.6baa + 0.2bab + 0.5bbb$$

y

$$\Pi_{X_3}[\Pi_{(X_1,X_2)} = G] = 0.6a + 0.5b$$

Como una aplicación elemental de (2.9), veamos el siguiente ejemplo.

Ejemplo 2.7. Considere la proposición $p \triangleq$ *Juan es grande*, donde *grande* es una relación cuyo cuadro es de la forma que se muestra en la Tabla 2.5 (con la altura expresada en centímetros y el peso expresado en kilogramos).

Tamaño	*Estatura*	*Peso*	μ
	170	70	0.7
	170	80	0.8
	180	80	0.9
	–	–	–
	190	90	1

Tabla 2.5: Cuadro de Tamaño.

Ahora, supongamos que además de saber que Juan es grande, también sabemos que $q \triangleq$ *Juan es alto*, donde el cuadro de altura se da (en forma tabulada parcialmente) por la Tabla 2.6.

Altura	*Estatura*	μ
	170	0.8
	180	0.9
	190	1

Tabla 2.6: Cuadro de Altura.

La pregunta es ¿Cuál es el peso de Juan? haciendo uso de (2.9), la distribución de posibilidad del peso de Juan se puede expresar como

$$\begin{aligned} \Pi_{Peso} &= \mathrm{Pro}_{Peso}\Pi_{(Estatura,Peso)}[\Pi_{Estatura} = Altura] \\ &= 0.7/70 + 0.9/80 + 1/90 \end{aligned} \tag{2.10}$$

Una aproximación lingüística aceptable para el lado derecho de (2.10) podría ser "*algo pesado*", donde "*algo*" es un modificador que tiene un efecto específico en el conjunto difuso llamado "*pesado*". Correspondientemente, una respuesta aproximada a la pregunta sería "*Juan es un poco pesado*".

Capítulo 3

Aplicación de la teoría de la posibilidad

La teoría de la posibilidad ha sido aplicada en la solución de diversos problemas. En este capítulo se aplica la teoría de la posibilidad a un proceso de emparejamiento de patrones. Básicamente, un proceso de emparejamiento de patrones involucra un patrón que describe algunos de los requisitos y una base de conocimiento donde se recogen datos. En el caso más simple, los datos que son idénticos a la especificación expresada en el patrón se buscan y se recuperan. Cuando se usan datos nítidos, el emparejamiento de patrones es un proceso de todo o nada. Si los datos están impregnados con imprecisión e incertidumbre o si el patrón incluye especificaciones vagas, el emparejamiento entre el modelo y el dato se convierte, en una cuestión de grados. En ese caso, un requisito difuso puede ser más o menos satisfecho. Se ha desarrollado una técnica de emparejamiento de patrones difusos que se basa en medidas de posibilidad y necesidad a fin de estimar la compatibilidad entre un dato y lo que es requerido por el patrón. En este enfoque, cada componente de un dato y cada requisito elemental se asocia respectivamente con una distribución de posibilidad y un conjunto difuso, que son herramientas convenientes para la representación de la imprecisión y la incertidumbre. En la siguiente sección se presentan los antecedentes sobre el enfoque anteriormente mencionado para el emparejamiento de patrones difusos. Luego, se amplía este concepto con el fin de dar cuenta de los requerimientos elementales de importancia desigual en el patrón. Se da un ejemplo para ilustrar el procedimiento.

3.1. Emparejamiento de patrones difusos

El emparejamiento de patrones difusos toma en cuenta la representación semántica asociada a los componentes de los datos y los patrones, y proporciona una evaluación de la similitud entre un patrón y un dato, que se califica en el intervalo [0,1]. Aquí se presenta un modelo que se supone que tiene representación en forma de árbol cuyas hojas corresponden a los componentes elementales llamados átomos. Por simplicidad asumimos que estos átomos no son variables.

3.1.1. Principio general

La idea básica es la de asociar a cada átomo de un patrón la función de pertenencia de un conjunto difuso, restringiendo los valores que son más o menos compatibles con el sentido del átomo. Estos valores pertenecen a alguna escala prescrita que corresponda al intervalo del atributo que el átomo refiere. Por ejemplo, el símbolo de *altura* puede referirse a una escala de alturas, y corresponde a un subconjunto difuso de esta escala. En lugar de una escala numérica que puede tener un conjunto discreto de elementos típicos. Los datos están representados por listas cuyos componentes están asociados con distribuciones de posibilidad. Tales listas contienen valores de atributos posiblemente mal conocidos, relativos a la descripción de los objetos. Un componente en una lista hace referencia a sólo un elemento de la escala o dominio del atributo en cuestión.

La disimetría básica de la concordancia con el modelo de datos es preservada por esta convención de modelado. De hecho, un patrón difuso representa una clase mal acotada de objetos, mientras que un dato difuso representa un objeto mal conocido, cuya descripción precisa no está disponible. Sean P y D respectivamente un átomo de patrón y un componente de dato perteneciente al mismo atributo, que se van a comparar. P y D se refieren a la misma escala U para transmitir sus significados. Sea μ_P la función de pertenencia asociada al átomo P y π_D la distribución de posibilidad asociada a D. Ambos son asignaciones de U a [0,1]. Sea u un elemento de U, entonces $\mu_p(u)$ es el grado de compatibilidad entre el valor u y el significado de P. $\mu_P(u) = 1$ significa plena certeza que u es compatible con P y $\mu_P(u) = 0$ significa plena certeza de que u no es compatible con P.

En contraste, $\pi_D(u)$ es el grado de posibilidad de que u es el (único) valor del atributo que describe el objeto modelado por el dato. D es un conjunto difuso de *posibles* valores mientras P es un conjunto difuso de valores *seguramente* compatibles. Por ejemplo, $\pi_D(u) = 1$ significa que u es totalmente posible, mientras $\pi_D(u) = 0$ significa que u es totalmente imposible como valor de un atributo del objeto. En lo que sigue, μ_P y π_D siempre se supondrá que son normalizados, es decir, siempre habrá un valor que es totalmente compatible con P y un valor totalmente posible en el rango D.

3.1.2. Indices elementales de emparejamiento

Cuando la clase de valores u representado por P es un conjunto usual y D se refiere a un valor preciso $d \in U$, el emparejamiento elemental es exitoso tan pronto como $d \in P$. Cuando P es un conjunto difuso, el grado de compatibilidad se convierte en el grado de pertenencia $\mu_P(d)$. En general, cuando D es impreciso, este grado de pertenencia se convierte en demasiado impreciso y se define como un número difuso en [0,1], llamado el grado de compatibilidad de D con respecto a P, denotado $\mu_P(D)$, con función de pertenencia $\mu_{P|D}$ definido por el principio de extensión

$$\mu_{P|D}(\alpha) = sup\{\pi_D(u) \mid \mu_P(u) = \alpha\}, \quad \forall \alpha \in [0,1] \tag{3.1}$$

Esta cantidad fue introducida por Zadeh, quien interpretó $\mu_P(D)$ como un valor de verdad difuso de un predicado P dado un predicado referencial D que describe un estado real de los hechos. Esta cantidad se representa en la Figura 3.1 cuando D y P se refieren a cantidades trapezoidales difusas.

Sin embargo, la transmisión de información relevante y completa acerca de la extensión del emparejamiento entre P y D, $\mu_P(D)$ es difícil de manipular a nivel operativo. Aunque su interpretación es clara desde el punto de vista teórico, esta cantidad difusa no es fácil de entender por los usuarios. Como consecuencia, dos medidas escalares son preferidas con el fin de estimar la compatibilidad entre un patrón del átomo P y su contraparte D en la lista de componentes de referencia, es decir, un grado de posibilidad de emparejar $\Pi(P;D)$ y un grado de necesidad de emparejar $N(P;D)$, que se definen

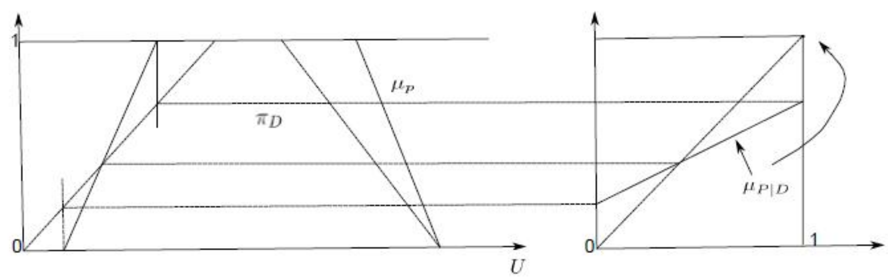

Figura 3.1: Representación gráfica del valor de verdad difuso de un predicado, presentado por Zadeh.

respectivamente por:

$$\Pi(P;D) = \sup_{u \in U} \text{mín}(\mu_P(u), \pi_D(u)) \tag{3.2}$$

$$N(P;D) = \inf_{u \in U} \text{máx}(\mu_P(u), 1 - \pi_D(u)) \tag{3.3}$$

La medida $\Pi(P;D)$ calcula en qué medida es posible que P y D se refieren al mismo valor u; en otras palabras, $\Pi(P;D)$ es un grado de solapamiento del conjunto difuso de valores compatibles con P, con el conjunto difuso de posibles valores de D. La medida *N (P;D)* calcula en qué medida es necesario (es decir, cierto) de que el valor al que se refiere D es uno de los compatibles con P, es decir, $N(P;D)$ es un grado de inclusión del conjunto de posibles valores de D en el conjunto de valores compatibles con P. La dualidad posibilidad/necesidad, es decir, que la necesidad de un evento corresponde a la imposibilidad del evento contrario, se expresa aquí por la relación

$$N(P;D) = 1 - \Pi(\bar{P};D) \tag{3.4}$$

donde $\mu_{\bar{P}} = 1 - \mu_P$ es la función de pertenencia del complemento del conjunto difuso de valores compatibles con P. Claramente, siempre tenemos que $\Pi(P;D) \geq N(P;D)$. Además hay que resaltar que *N(F,F) = 1* si y sólo si μ_F es la función de la pertenencia de un subconjunto ordinario de U que quiere decir μ_P y π_D; de lo contrario sólo tenemos $N(F;F) \geq 1/2$. En efecto, cuando dos símbolos idénticos tienen un significado difuso, no se puede estar completamente seguro de que se refieren exactamente al mismo conjunto de valores. En cualquier caso tenemos $N(s(F);F) = 1$, donde el soporte de F, *s(F)* se define como $s(F) = \{u \in U, \mu_F(u) > 0\}$.

Los casos limites donde $\Pi(P;D)$ y $N(P;D)$ toman valores 0 y 1 también son útiles para caracterizar. Para cualquier conjunto difuso F en U, sea $F^0 = \{u \in U \mid \mu_F(u) = 1\}$ el núcleo de F y *s(F)* su soporte como se definió anteriormente. Entonces se puede comprobar que

(i) $\Pi(P;D) = 0$ si y sólo si $s(P) \cap s(D) = \emptyset$

(ii) $\Pi(P;D) = 1$ si y sólo si $P^0 \cap D^0 \neq \emptyset$

(iii) $N(P;D) = 1$ si y sólo si $s(D) \subseteq P^0$

Nótese que (iii) define una inclusión más fuerte entre los conjuntos difusos que es lo usual (es decir, $\pi_D \leq \mu_P$); esto último sólo implica que $N(P;D) \geq 0.5$. La Tabla 3.1 resume las respectivas localizaciones de $\Pi(P;D)$ y $N(P;D)$ en [0,1] cuando P y/o D son precisos, imprecisos (pero nítidos) o difusos.

D / P	*Precisa* $D = \{d\}$	Imprecisa (No difusa)	Difusa
Precisa P = {p}	$\Pi = N = 1$ *si* $p = d$ $\Pi = N = 0$ *si* $p \neq d$	N = 0 $\Pi = 1$ *si* $p \in D$ $\Pi = 0$ e. o. c.	N = 0 $\Pi = \pi_D(p)$
Imprecisa (No difusa)	$\Pi = N = 1$ *si* $d \in P$ $\Pi = N = 0$ *si* $d \notin P$	$N = \Pi = 1$ *si* $D \subseteq P$ $N = \Pi = 0$ *si* $D \cap P = \emptyset$ $\Pi = 1, N = 0$	$N > 0 \Rightarrow \Pi = 1$
Difusa	$\Pi = N = \mu_P(d)$	$\Pi \geq N$	$\Pi \geq N$

Tabla 3.1

Nótese que cuando D es preciso, entonces $\Pi(P;D) = N(P;D) = \mu_P(D)$ que es entonces un número real ordinario en [0,1]. Cuando D es impreciso pero no difuso, entonces $\mu_P(D)$ es un subconjunto de [0,1] y $\Pi(P;D)$ y $N(P;D)$ son respectivamente la mínima cota superior y la máxima cota inferior del conjunto $\mu_P(D)$. De manera más general, si τ es un valor modal de $\mu_P(D)$ (es decir, $\mu_{P|D}(\tau) = 1$) entonces τ es un buen indicador escalar de la extensión del emparejamiento entre P y D, es decir, un buen representante de la cantidad difusa. El par de índices $(\Pi(P;D), N(P;D))$ es una aproximación razonable de $\mu_P(D)$ debido a los siguientes resultados:

Proposición 3.1. *Para todo P,D y τ tal que $\mu_{P|D}(\tau) = 1$, se cumple la desigualdad $N(P;D) \leq \tau \leq \Pi(P;D)$.*

Prueba.
$\Pi(P;D)$ y $N(P;D)$ se pueden obtener de $\mu_P(D)$ como sigue

$$\Pi(P;D) = \Pi(Verdad \mid \mu_P(D)), \quad N(P;D) = N(Verdad \mid \mu_P(D)) \tag{3.5}$$

donde *Verdad* es un conjunto difuso de [0, 1] tal que $\mu_{Verdad}(u) = u$, $\forall u \in [0,1]$. La primera ecuación en (3.5) se lee

$$\Pi(P;D) = \sup_{u\in[0,1]} \text{mín}(u, \mu_{P|D}(u)) \geq \text{mín}(\tau, \mu_{P|D}(\tau)) = \tau$$

Similarmente,

$$N(P;D) = \inf_{u\in[0,1]} \text{máx}(u, 1-\mu_{P|D}(u)) \leq \text{máx}(\tau, 1-\mu_{P|D}(\tau)) = \tau$$

Así, $[N(P;D), \Pi(P;D)]$ proporciona información sobre la imprecisión de $\mu_P(D)$. En conclusión, $\Pi(P;D)$ y N(P;D) tienen una clara y precisa semántica que corresponden a la naturaleza del problema de emparejamiento atómico flexible.

3.1.3. El procedimiento de emparejamiento de patrones difusos

Codificación de conocimiento y cálculo elemental de emparejamiento

Se utilizarán las funciones de pertenencia de los valores trapezoidales. Para aplicaciones prácticas, las funciones trapezoidales de pertenencia generalmente son suficientes; y de hecho, ligeras modificaciones de la forma de una función de pertenencia no afectan a la evaluación de los resultados significativamente. Una función trapezoidal se representa por una 4-upla (a,b,c,d) de los parámetros cuyo significado se explica en la Figura 3.2. c y d son no negativos.

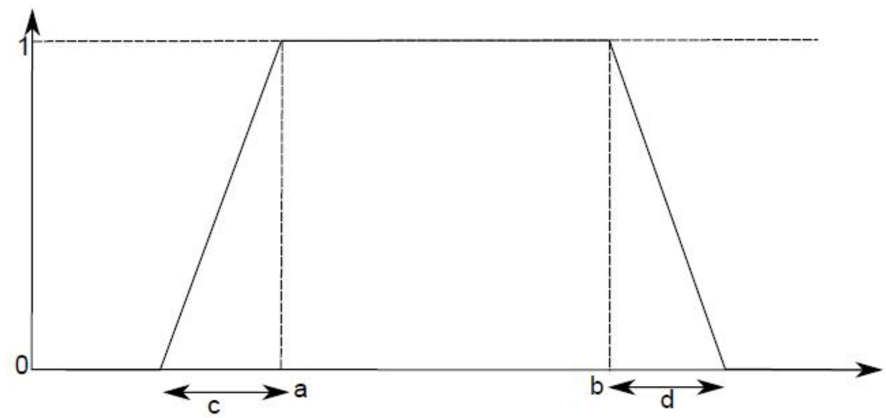

Figura 3.2: Función trapezoidal.

Se puede tener $a = b, c = 0$ o $d = 0$. Por lo tanto, se pueden representar valores precisos o imprecisos mediante 4-uplas. Sean P y D dos valores difusos cuyas funciones de pertenencia son trapezoidales y se representan por las 4-uplas (a_1, b_1, c_1, d_1) y (a_2, b_2, c_2, d_2) respectivamente. Las medidas de posibilidad y necesidad (3.2) y (3.3) son calculadas directamente a partir de las representaciones prácticas de los valores, como intersección de líneas rectas. Ver las Figuras 3.3 y 3.4.

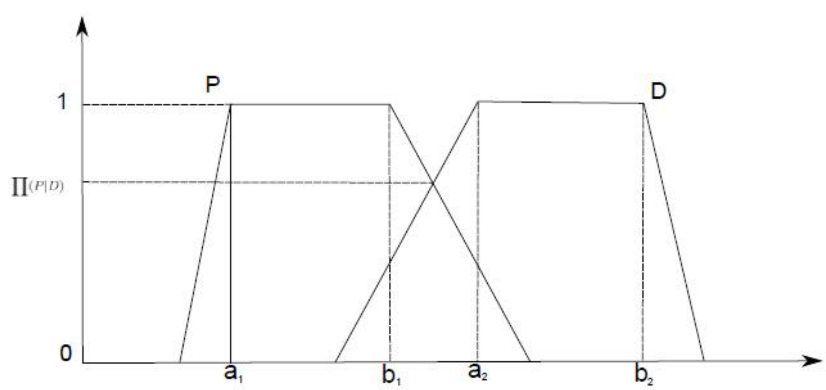

Figura 3.3: Medidas de posibilidad.

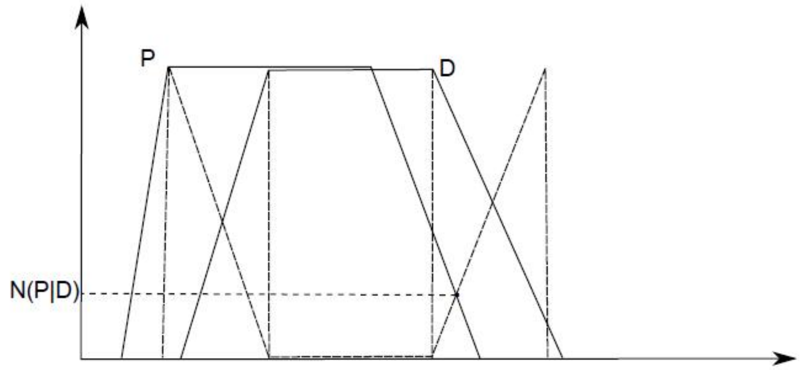

Figura 3.4: Medidas de necesidad.

Evaluación global de emparejamiento

Las medidas atómicas de posibilidad y necesidad se agregan por separado a fin de obtener dos medidas globales entre todo el patrón y todo el dato. Cuando el patrón expresa una conjunción de requisitos elementales $P_1, \ldots, P_n$, esta agregación se realiza mediante la operación mín y conserva la semántica respectiva de las medidas de posibilidad y necesidad. De hecho, tenemos

$$\Pi(P_1 \times \ldots \times P_n; D_1 \times \ldots \times D_n) = \min_{i=1,\ldots n} \Pi(P_i; D_i) \tag{3.6}$$

$$N(P_1 \times \ldots \times P_n; D_1 \times \ldots \times D_n) = \min_{i=1,\ldots n} N(P_i; D_i) \tag{3.7}$$

donde P_i y D_i se supone que están definidas en el mismo dominio U_i y donde $\times$ denota el producto cartesiano definido para dos conjuntos difusos F_i y F_j por

$$\mu_{F_i \times F_j}(u_i, u_j) = \min(\mu_{F_i}(u_i), \mu_{F_j}(u_j)) \quad \forall u_i \in U_i, \forall u_j \in U_j \tag{3.8}$$

Los atributos que aparecen en el patrón y el dato se supone que no sean interactivos. Observe que (3.6) y (3.7) suponen que todas las partes del patrón (que expresan lo que se requiere) tienen una importancia igual desde el punto de vista del usuario. Cuando el patrón expresa una disyunción de las necesidades elementales P_1 *o* P_2 *o* $\ldots$ *o* P_n la operación mín en (3.6) y (3.7) se convierte en la operación máx reemplazando $P_1 \times \ldots \times P_n$ por $P_1 + \ldots + P_n$ donde $+$ es el dual de $\times$ ($F + G = \overline{\bar{F} \times \bar{G}}$). En general, un patrón cuya estructura es un árbol de requisitos elementales pueden ser tratados mediante este enfoque.

Es de hacer notar que dos valores u_1, u_2 que pertenecen a un dominio U pueden considerarse aproximadamente igual, incluso si no son idénticos. Por ejemplo, si el patrón requiere alguien que tiene 40 años, un dato correspondiente a una persona que es de 39 puede considerarse en algunos casos como aproximadamente coincidente con la solicitud. Una igualdad aproximada puede ser convenientemente modelada por medio de una relación difusa R que es reflexiva (es decir, $\mu_R(u,u) = 1, \forall u \in U$) y simétrica ($\mu_R(u_1,u_2) = \mu_R(u_2,u_1), \forall u_1, u_2 \in U$). Cuanto más cercano estén u_1, u_2, más cerca a 1 debe estar $\mu_R(u_1,u_2)$. La cantidad $\mu_R(u_1,u_2)$ puede ser vista como un grado de emparejamiento de u_1 con u_2. R se denomina una proximidad o una relación de tolerancia. Cuando el patrón es cualquier subconjunto P de U (P puede ser difuso) pero el dato sigue siendo una constante u, la tolerancia R se puede tomar en cuenta en el grado de coincidencia (que es igual a $\Pi(P;\{u\}) = N(P;\{u\}) = \mu_P(u)$ si R es ignorado) mediante el reemplazo de P por el subconjunto ampliado $P \circ R$, definido por

$$\mu_{P \circ R}(u) = \sup_{u' \in U} \text{mín}(\mu_P(u'), \mu_R(u,u')), \quad \forall u \in U \tag{3.9}$$

La desigualdad $\mu_{P \circ R}(u) \geq \mu_P(u), \forall u \in U$ expresa que $P \circ R$ es mayor que P. Hablando en términos generales, $P \circ R$ reúne los elementos de P y los elementos fuera de P que se encuentran cercanos a un elemento de P. En el caso general, la correspondencia de D con respecto a P, tomando en cuenta que R es calificada por $\Pi(P \circ R; D)$ y por $N(P \circ R; D)$.

A continuación, se propone un método que usa pesos para modelar la importancia.

Jerarquización de los mejores emparejamientos de datos

Por lo general, el mismo patrón que expresa algún requisito, se compara con una base de datos completa, es decir, una colección de datos. Desde el punto de vista del usuario, es interesante obtener los datos que producen las mejores medidas de compatibilidad con respecto al patrón. Vamos a describir el comportamiento de nuestro sistema basado en emparejamiento de patrones difusos en ese caso. Pueden ocurrir varias situaciones (ver Tabla 3.1):

(i) No coincide con el patrón de datos, las medidas de posibilidad (y necesidad) son cero.

(ii) Uno o varios datos coinciden con el patrón con posibilidad positiva o grados de necesidad. En ese caso, nos enfrentamos al problema de la jerarquización de los datos, cada uno de ellos se caracteriza por un par de grados. Con el fin de establecer una jerarquización de varios datos con respecto a un patrón; podemos tomar en cuenta las siguientes observaciones:

1. El grado de necesidad es de mayor importancia que el grado de posibilidad, porque cuando el grado de necesidad es positivo, estamos (más o menos) seguros de que el dato coincide con el requisito.

2. Si se obtiene un grado de necesidad igual a cero y un grado posibilidad igual a 1 para cada dato, esto significa que el requisito expresado en el patrón es demasiado precisa, con respecto a los datos disponibles. Podríamos entonces desear mejorar el conocimiento sobre los datos.

3. En general, el ordenamiento de Pareto se utiliza para calificar los pares de grados, de acuerdo con la siguiente forma:

   (Π, N) es mayor que (Π', N') si y sólo si $\Pi > \Pi'$ y $N \geq N'$ o $\Pi \geq \Pi'$ y $N > N'$. Sin embargo, hay situaciones donde $\Pi(P; D) > \Pi(P; D')$ y $N(P; D) < N(P; D')$. Aquí P y D representan un patrón compuesto o datos. El ordenamiento de Pareto es sólo parcial. Así, podemos obtener varios datos que algo mejor emparejan con el patrón. Podríamos pensar en una medida de precisión para discriminar los datos. Mientras más preciso sea D, más cercano a cero será $\Pi(P; D) - N(P; D)$. Sin embargo, lo contrario es falso. Vale la pena notar que si D' y D son elementos maximales para la ordenación de Pareto, tenemos

$$N(P; D) < N(P; D') < \Pi(P; D') < \Pi(P; D)$$

lo que indica que D' es el mejor dato desde el punto de vista de la precisión (es decir, $\Pi(P;D) - N(P;D) > \Pi(P;D') - N(P;D')$) así como desde el punto de vista de la necesidad.

Proporcionar un conjunto ponderado de los datos recuperados puede no ser lo suficientemente informativo para un usuario. Puede ser también útil para explicar las razones de incertidumbre si los hay. Por ejemplo, explicando que la certeza de que coincida para un dato D tal que $\Pi(P;D) > 0$ se podría mejorar mediante la adición de más información sobre el valor de algún atributo i, es decir, restringiendo este valor a un subconjunto estricto D'_i de D_i, de forma que $N(P_i, D'_i) > 0$.

3.2. Asignación de importancia

En la Sección 3, se recordó que las medidas atómicas de posibilidad y necesidad deben agregarse por separado utilizando la operación mín en el caso de un patrón compuesto conjuntivo (véanse las fórmulas (3.6), (3.7)). Esto supone la idéntica importancia de los requisitos correspondientes a la misma importancia los requisitos correspondientes a los modelos atómicos. A continuación se presenta un método para la introducción de los pesos en agregaciones basadas en las operaciones mín o máx.

3.2.1. Mínimo y máximo ponderado

Sean $F_1, \ldots, F_n$, n conjuntos difusos en U con funciones de pertenencia $\mu_{F_1}, \ldots, \mu_{F_n}$. La unión $\cup_i F_i$ y la intersección $\cap_i F_i$ se definen a través de funciones de pertenencia $\mu_{\cup_i F_i} = \max_i \mu_{F_i}$ y $\mu_{\cap_i F_i} = \min_i \mu_{F_i}$. Otra forma de agregar las funciones de pertenencia es $\mu_{+_n F_i} = n^{-1} \sum_{i=1}^{n} \mu_{F_i}$ que define un compromiso entre las funciones de pertenencia, y se encuentra de igual manera entre la unión y la intersección.

Estas agregaciones consideran que $F_1, \ldots, F_n$ tienen igual importancia. Sin embargo, en aplicaciones prácticas, algunos F_i pueden ser menos importantes que otros en el sentido que incluso si $\mu_{F_i}(u) = 0$ por ejemplo, $\mu_{\cap_{F_i}}(u)$ no debería ser cero, debido sólo a μ_{F_i}.

En el caso de la agregación de compromiso a través de la media aritmética, la asignación de importancia se ajusta fácilmente en la agregación a través de la suma convexa.

$$\mu_{+_n F_i}(u) = \sum_{i=1}^{n} p_i \mu_{F_i}(u), \quad \forall u \in U \tag{3.10}$$

donde $p_1, p_2, \ldots, p_n$ son números no negativos tales que $\sum_{i=1}^{n} p_i = 1$ expresan la importancia relativa de los conjuntos difusos. Observe que esta operación de agregación no abarca la intersección ni la unión. Con el fin de obtener una contraparte ponderada de las operaciones mínimo y máximo, debe buscarse otro esquema.

Esto se ha hecho mediante la siguiente analogía. Observe que (3.10) se puede ver como la probabilidad de un evento difuso Φ_u definido sobre $\{1, \ldots, n\}$. Φ_u es el conjunto difuso de los F_i's que contiene u, y los p_i's definen una asignación de probabilidad sobre $\{1, \ldots, n\}$.

Por lo tanto, $\mu_{+_n F_i} = \text{Prob}(\Phi_u)$ en el sentido de Zadeh. Cambiando Prob por una medida de posibilidad o necesidad lleva a considerar los siguientes análogos de (3.10):

$$\Pi(\Phi_u) = \max_{i=1,\ldots,n} \min(\mu_{F_i}(u), w_i) \tag{3.11}$$

$$N(\Phi_u) = \min_{i=1,\ldots,n} \max(\mu_{F_i}(u), 1 - w_i) \tag{3.12}$$

(3.11) y (3.12) son versiones discretas de (3.2) y (3.3) donde los w_i definen una distribución de posibilidad en {1, ..., n} que juegan el papel de ponderación de importancia. Está claro que (3.11) produce una disyunción ponderada de los F_i's mientras que (3.12) produce un conjunto ponderado.

3.2.2. Aplicación al emparejamiento de patrones

Sean $w_1, \ldots, w_n$ los grados de importancia de los patrones $P_1, \ldots, P_n$ respectivamente. Se supone que $\forall i, w_i \in [0, 1]$; mientras mayor sea w_i, mayor será la importancia de P_i; también asumimos que $\max\limits_{i=1,\ldots,n} w_i = 1$ (normalización), es decir, los átomos más importantes se calificarán con 1.

Luego, si s_i denota un grado de emparejamiento de un componente de un dato con respecto al átomo P_i, entonces el grado correspondiente de emparejamiento s de este dato con respecto a todo el patrón (P_1 *y* $\ldots$ *y* P_n), tomando en cuenta la evaluación de importancia, se dará por

$$s = \min_{i=1,\ldots,n} \max(1 - w_i, s_i) \tag{3.13}$$

s expresa hasta que punto estamos seguros de que el conjunto difuso de requisitos importantes (definido por los w's) se incluye en el conjunto difuso de los requisitos posiblemente (respectivamente necesariamente) satisfechos por el componente de datos D_i, definido por $s_i = \Pi(P_i; D_i)$ (respectivamente $s_i = N(P_i; D_i)$) para $i = 1, \ldots, n$. Tomemos en cuenta que si todos los w_i's son iguales a 1 (igual importancia), se obtiene $s = \min\limits_{i=1,\ldots,n} s_i$; cuando $w_i = 0$, la correspondencia con los átomos P_i no se tiene en cuenta.

Así, en caso de importancia desigual de los diferentes átomos de un patrón conjuntivo, utilizamos (3.13) con $s_i = \Pi(P_i; D_i)$, $i = 1, \ldots, n$ y con $s_i = N(P_i; D_i)$, $i = 1, \ldots, n$ en lugar de la aplicación (3.6) y (3.7).

La introducción de las ponderaciones como se propone en (3.13) equivale a la modificación de los patrones P_i en P_i^* tal que

$$\mu_{P_i^*}(u) = \max(\mu_{P_i}(u), 1 - w_i)$$

y combinándolos mediante una operación de agregación simétrica que es de nuevo la operación mínimo. De hecho, las siguientes identidades son válidas y se pueden establecer mediante cálculos sencillos:

$$\min_{i=1,\ldots,n} \max(1 - w_i, \Pi(P_i; D_i)) = \min_{i=1,\ldots,n} \Pi(P_i^*; D_i)$$

$$\min_{i=1,\ldots,n} \max(1 - w_i, N(P_i; D_i)) = \min_{i=1,\ldots,n} N(P_i^*; D_i)$$

Como consecuencia de ello, usando (3.6), (3.7) y las identidades anteriores podemos afirmar que el esquema de agregación (3.13) produce los grados de posibilidad o necesidad de emparejamiento de acuerdo a si los mismos s_i's son grados de posibilidad o necesidad respectivamente. Resultados similares son válidos en el caso de una agregación disyuntiva ponderada (ver (3.11))

$$s = \max_{i=1,\ldots,n} \min(w_i, s_i)$$

Es decir, obtenemos las siguientes identidades:

$$\max_{i=1,\ldots,n} \min(w_i, \Pi(P_i; D_i)) = \max_{i=1,\ldots,n} \Pi(P_i; D_i)$$

$$\max_{i=1,\ldots,n} \min(w_i, N(P_i; D_i)) = \max_{i=1,\ldots,n} N(P_i; D_i)$$

donde $\mu_P = \min(w_i, \mu_{P_i})$

Ejemplo 3.1. En este ejemplo se trata de descripciones (más o menos aproximadas) de vehículos de segunda mano a la venta en un garaje. Los diferentes atributos implicados son la edad del carro, precio de compra (en Bolívares), consumo de gasolina y la velocidad del automóvil. Los datos disponibles se puede representar en una tabla relacional (Tabla 3.2).

	Edad	*Precio de compra*	*Consumo*	*Velocidad*
$V1$	nuevo	costoso	económico	algo rápido
$V2$	menor de 3 años	cerca de 450.000	algo económico	180-200 Km/h
$V3$	muy reciente	500.000-600.000	fuerte	rápido
$V4$	cerca de 5 años	menos de 200.000	8-9 Km/Litro	180 Km/h
$V5$	de 5 a 10	cerca de 100.000	fuerte	algo rápido
$V6$	viejo	barato	económico	no muy rápido
$V7$	nuevo	320.000-400.000	muy económico	140-160 Km/h

Tabla 3.2

Cada fila en la tabla corresponde a una lista de pares ($\langle valor\ difuso\rangle\langle atributo\rangle$); por ejemplo, la primera fila corresponde al dato

V1:((edad nuevo) (precio costoso) (consumo económico) (velocidad algo rápido))

Un dominio es asociado a cada atributo; este dominio es el conjunto de todos los valores que pueden ser tomados por el atributo. Aqui por ejemplo, el dominio asociado al atributo “edad del carro” es el intervalo [0,20]. Todos los dominios considerados en este ejemplo son escalas continuas.

Ahora, vamos a considerar el siguiente problema: una persona que quiere comprar un carro busca un carro correpondiente a diversos criterios. Sin embargo, alguno de esos criterios son más importantes que otros para el comprador considerado. Entonces él puede expresar la solicitud usando un patrón ponderado general. Por ejemplo, podemos tener la siguiente solicitud:

P: (((edad nuevo) 0.8) y ((precio moderado) 0.5) y ((consumo muy económico) 1) y ((muy rápido) 0.2))

Este patrón expresa que el rasgo más importante es un bajo consumo de gasolina (ponderacion igual a 1) y en un orden decreciente de importancia, la edad, el precio y la velocidad. Las funciones de pertenencia de los valores difusos son dados, usando la codificación de 3.1.3

1. Atributo: Edad del carro; Dominio: [0,20]

   nuevo $\begin{pmatrix} 0 & 1 & 0 & 1 \end{pmatrix}$

   muy reciente $\begin{pmatrix} 1 & 2 & 1 & 1 \end{pmatrix}$

   cerca de 5 años $\begin{pmatrix} 5 & 5 & 1 & 1 \end{pmatrix}$

   viejo $\begin{pmatrix} 10 & 15 & 2 & 5 \end{pmatrix}$

   menor de 3 años $\begin{pmatrix} 0 & 3 & 0 & 0 \end{pmatrix}$

2. Atributo: Precio de venta; Dominio: [50.000 , 1.000.000]

   costoso $\begin{pmatrix} 800.000 & 1.000.000 & 50.000 & 0 \end{pmatrix}$

cerca de 450.000	$\begin{pmatrix} 440.000 & 460.000 & 10.000 & 10.000 \end{pmatrix}$
cerca de 100.000	$\begin{pmatrix} 90.000 & 110.000 & 10.000 & 10.000 \end{pmatrix}$
barato	$\begin{pmatrix} 50.000 & 100.000 & 0 & 50.000 \end{pmatrix}$
de 500.000 a 600.000	$\begin{pmatrix} 500.000 & 600.000 & 50.000 & 500.000 \end{pmatrix}$
moderado	$\begin{pmatrix} 350.000 & 600.000 & 50.000 & 50.000 \end{pmatrix}$

3. Atributo: Consumo; Dominio [5,15]

económico	$\begin{pmatrix} 6 & 7 & 0.5 & 1 \end{pmatrix}$
algo económico	$\begin{pmatrix} 7 & 8 & 1 & 1 \end{pmatrix}$
muy económico	$\begin{pmatrix} 5 & 6 & 0 & 0.5 \end{pmatrix}$
fuerte	$\begin{pmatrix} 9 & 15 & 0.5 & 0 \end{pmatrix}$

4. Atributo: Velocidad; Dominio [100,250]

rápido	$\begin{pmatrix} 180 & 200 & 20 & 20 \end{pmatrix}$
algo rápido	$\begin{pmatrix} 150 & 180 & 20 & 20 \end{pmatrix}$
no muy rápido	$\begin{pmatrix} 120 & 140 & 10 & 10 \end{pmatrix}$
entre 140 y 160	$\begin{pmatrix} 140 & 160 & 20 & 20 \end{pmatrix}$

Los resultados obtenidos en este ejemplo con el patrón P son:

$\Pi(P;V1) = 0.5$, $\Pi(P;V2) = 0.3$, $\Pi(P;V6) = 0.2$, $\Pi(P;V7) = 0.8$ y $N(P;V7) = 0.5$

Los otros ($\Pi(P;V_i) o\ N(P;V_j)$) son iguales a cero.

En caso de igual importancia de los diferentes requisitos del patrón P, usaremos las formulas (3.6) y (3.7) y obtendremos los siguientes resultados:

$\Pi(P;V1) = 0$, $\Pi(P;V2) = 0.3$, $\Pi(P;V6) = 0$, $\Pi(P;V7) = 0.5$ y todos los demás grados son igual a cero.

Así solamente dos carros (V2 y V7), le serán propuestos al comprador en lugar de cuatro, en caso de un patrón ponderado. De hecho, en el caso ponderado, el carro V1 puede ser del interés para el comprador porque él o ella de algún modo descuidan el precio de compra. Además, el vehículo V7 es bastante preferido en el caso ponderado ($N(P;V7) = 0{,}5$). Esto es porque el comprador realmente no se molesta sobre la velocidad del vehículo. Este tipo de información, que ayuda en el proceso de discriminación, no puede ser expresado mediante el emparejamiento de patrones difusos estándar.

Referencias Bibliográficas

[1] Zadeh, L. (1965). “Fuzzy Sets”. Information and Control, pp. 338-353.

[2] Zadeh, L. (1978). “Fuzzy Sets as a Basis as a basis of a Theory of Possibility”. Fuzzy Sets and Systems, pp. 3-28.

[3] Dubois, D. & Prade, H. (1980). Fuzzy Sets and Systems, Theory and Applications. Academic Press.

[4] Dubois, D. Prade, H & Testemale, C. (1988). “Weighted Fuzzy Pattern Matching”. Fuzzy Sets and Systems, pp. 313-331.

[5] García, J. (1988). Modelo de Resolución de Incertidumbre en Sistemas Expertos a Partir de la Teoría de la Posibilidad. Tesis Doctoral, Universidad Politécnica de Madrid, 1988.POSIBILIDAD

[6] Klir, G. & Yuan, B. (1995). Fuzzy Sets Theory. Theory and Applications. Chapter 7, Possibility theory, pp. 177-211. Prentice Hall.

[7] Zimmermann, H-J. (2001). Fuzzy Set Theory and its Applications. Fourth Edition. . Chapter 8, Uncertainty Modeling, pp. 111-138. Kluwer Academic Publishers.

[8] Ross, T.; Sellers, K. & Booker, J. (2002) Considerations for Using Fuzzy Sets Theory and Probability Theory. In Ross, T.; Booker, J. & Parkinson, J. (Editors) Fuzzy logic and probability applications : bridging the gap. Chapter 5, pp. 87-104. SIAM.

[9] Dubois, D. (2006). Possibility Theory and Statistical Reasoning. Computational Statistics and Data Anal., pp. 47-69.
http://www.irit.fr/ ~ Didier.Dubois/Papers0804/D_CSDA06.pdf

[10] Kikuchi, S. & Chakroborty, P. (2006) "Place of possibility theory in transportation analysis". Transportation Research Part B 40, Elsevier, pp. 595-615
http://home.iitk.ac.in/ ~ partha/possibility

[11] Klir, G. (2006). Uncertainty and Information: Foundations of Generalized Information Theory. Chapter 2, Classical Possibility-Based Uncertainty Theory, pp.26-60. Wiley.