Universidad Central "Marta Abreu" de Las Villas.
Facultad de Matemática Física y Computación
Tesis de Maestría en Inteligencia Artificial

TESIS PRESENTADA EN OPCIÓN AL TÍTULO ACADÉMICO DE MASTER EN CIENCIA DE LA COMPUTACIÓN

Usando la Teoría de las Probabilidades para el cálculo de la Certidumbre en un Sistema Basado en Casos.

Autor: Lic. Andrés Tellería Rodríguez

Tutor: Msc. Iliana Gutiérrez Martínez.

Cotutor: Dr Rafael E. Bello Pérez.

2000-2001

Introducción

Existen en la actualidad diversas formas de representar el conocimiento, a ellas se les concede significativa importancia a la hora de diseñar un sistema experto porque éstas definen la estructura interna de todo tipo de información. Especialmente, debido al énfasis impuesto en la representación y dominio del conocimiento es que los sistemas que utilizan técnicas de Inteligencia Artificial para lograr conocimiento experto se conocen como Sistemas Basados en Conocimiento. En dependencia de sus características, los Sistemas Basados en el Conocimiento entre otros pueden ser: Sistemas de Conocimiento Basados en Reglas(SBCR), Sistemas de Conocimientos Basados en Probabilidades(SBCP), Sistemas de Conocimiento Basados en Redes Expertas, Sistemas Basados en Modelos o Sistemas de Conocimientos Basados en Casos(SBCc).

Un sistema de representación flexible y expresivo debe ser capaz, entre otras cosas de representar explícitamente y manipular conocimiento incierto e incompleto. En los Sistemas de Conocimientos Basados en Casos, por ejemplo, el razonamiento se apoya en el hecho de que los humanos actúan intuitivamente sobre la base de información parcial, por tanto, es necesario disponer de mecanismos de inferencia que permitan extraer conclusiones sugeridas pero no aseguradas(garantizadas) a partir de premisas dadas. A estos mecanismos que trabajan bajo las circunstancias anteriores se les han denominado Sistemas de R*azonamiento Bajo incertidumbre, Sitemas de Razonamiento Aproximado, Sistemas de Razonamiento Inexacto*, etc. [Bux89, Dut85, Wal96,Tur92]. Los términos: incertidumbre, aproximado, inexacto son relativos a la duda, ambigüedad, imprecisión, inconsistencia y vaguedad que existe en el conocimiento presente en estos sistemas.

En los Sistemas Basados en Casos la incertidumbre está presente en casi todos los componentes del conocimiento: en los rasgos usados para describir los casos(vocabulario), en las relaciones entre los casos y sus rasgos, en la forma de evaluar la semejanza, y en la manera que la biblioteca de casos se organiza y se accede [Dut85, Bon91]. En tales sistemas esta incertidumbre se puede originar desde diversas fuentes [Zad78]

Resumen

Introducción

Existen en la actualidad diversas formas de representar el conocimiento, a ellas se les concede significativa importancia a la hora de diseñar un sistema experto porque éstas definen la estructura interna de todo tipo de información. Especialmente, debido al énfasis impuesto en la representación y dominio del conocimiento es que los sistemas que utilizan técnicas de Inteligencia Artificial para lograr conocimiento experto se conocen como Sistemas Basados en Conocimiento. En dependencia de sus características, los Sistemas Basados en el Conocimiento entre otros pueden ser: Sistemas de Conocimiento Basados en Reglas(SBCR), Sistemas de Conocimientos Basados en Probabilidades(SBCP), Sistemas de Conocimiento Basados en Redes Expertas, Sistemas Basados en Modelos o Sistemas de Conocimientos Basados en Casos(SBCc).

Un sistema de representación flexible y expresivo debe ser capaz, entre otras cosas de representar explícitamente y manipular conocimiento incierto e incompleto. En los Sistemas de Conocimientos Basados en Casos, por ejemplo, el razonamiento se apoya en el hecho de que los humanos actúan intuitivamente sobre la base de información parcial, por tanto, es necesario disponer de mecanismos de inferencia que permitan extraer conclusiones sugeridas pero no aseguradas(garantizadas) a partir de premisas dadas. A estos mecanismos que trabajan bajo las circunstancias anteriores se les han denominado Sistemas de R*azonamiento Bajo incertidumbre, Sitemas de Razonamiento Aproximado, Sistemas de Razonamiento Inexacto*, etc. [Bux89, Dut85, Wal96,Tur92]. Los términos: incertidumbre, aproximado, inexacto son relativos a la duda, ambigüedad, imprecisión, inconsistencia y vaguedad que existe en el conocimiento presente en estos sistemas.

En los Sistemas Basados en Casos la incertidumbre está presente en casi todos los componentes del conocimiento: en los rasgos usados para describir los casos(vocabulario), en las relaciones entre los casos y sus rasgos, en la forma de evaluar la semejanza, y en la manera que la biblioteca de casos se organiza y se

accede [Dut85, Bon91]. En tales sistemas esta incertidumbre se puede originar desde diversas fuentes [Zad78]

Tradicionalmente, los métodos de representación de conocimiento han intentado modelar la incertidumbre mediante diferentes teorías como la probabilidad y/o la teoría de conjuntos difusos. La primera refleja la naturaleza estocástica de adopción de medidas mientras la segunda capta la subjetividad de comportamiento humano. Los métodos de decisión estocásticos tal como análisis estadístico no miden la imprecisión en el comportamiento humano, más bien, estos métodos modelan conocimiento incompletos o inexacto sobre el mundo externo. La teoría de los conjuntos fuzzy propicia los métodos que constituyen medios adecuados para modelar la incertidumbre que provienen de fenómenos que no son ni estocásticos ni aleatorios [Che92].

En la actualidad todas estas técnicas son muy utilizadas en el marco teórico, sin embargo, existen escasos sistemas computacionales que consideren el manejo de este tipo de información. Por otra parte, la organización y el acceso a la misma constituyen uno de los principales problemas de los Sistemas Basados en el Conocimiento existentes. Es por ello que se propone crear un Sistema computacional Basado en Casos donde además de manipular conocimiento incierto, garantice una buena organización de la información y un acceso eficiente a la misma. Para lograr este objetivo general se proponen los siguientes objetivos específicos:

- Hacer un estudio del alcance de las técnicas de actuales de manejo de la incertidumbre.
- Determinar una estructura óptima de organización de los casos que permitiera un uso eficiente de una o varias de estas técnicas.
- Crear un modelo matemático que fuese capaz de adaptar el uso de al menos una de ellas a la problemática actual del manejo de la incertidumbre en un Sistema Basado en Casos.
- Construir un Sistema Computacional para implementar el modelo planteado.
- Mostrar la aplicabilidad del sistema a través de una aplicación real.

Estructura General del Trabajo

El trabajo está desarrollado en tres capítulos: en el Capítulo I se resumen las técnicas fundamentales para el manejo de la incertidumbre que existen en la actualidad. En el Capítulo II se analiza la problemática de la incertidumbre en los Sistemas Basados en Casos y se plantea un modelo matemático que hace uso de la teoría de la probabilidad para la determinación de la incertidumbre en una base de casos. Por último en el Capítulo III se describe el Sistema computacional TreeData y se exponen los resultados obtenidos al procesar una base real.

Novedad Científica:

Se propone una vía matemática para el cálculo de la certidumbre en un Sistema de Conocimientos Basados en casos, ello permite completar la base original dejándola lista para la recuperación y adaptación.

Proporciona desde el punto de vista computacional un nuevo componente que no existía: el TreeStructView, importante en la creación, visualización y recorrido de un árbol de decisiones.

Se presenta una técnica que logra un acceso rápido a los casos almacenados en una base dada.

Valor Práctico:

Las facilidades brindadas por el software implementado para inferir soluciones usando tratamiento probabilístico bajo incertidumbre en un sistema de conocimiento tipo TreeData (programa para inferir decisiones), no están presente en otros sistemas implementados antes en nuestro país. Además puede ser usado en la carrera Ciencia de la Computación en la disciplina Inteligencia Artificial, tanto en Pregrado como en Postgrado cuando se impartan los temas referentes a Sistema Basados en el Conocimiento(SBC). En particular los Sistemas de Conocimiento Basados en Casos(SCBC).

Capítulo I

El desarrollo de técnicas para el manejo de incertidumbre en los Sistemas Expertos ha sido uno de los temas que ha constituido el centro de la atención de muchos especialistas en el campo de la Inteligencia Artificial. Esto se debe a que en aquellos dominios correspondientes a las ciencias poco formalizadas, donde se trabaja con información que no puede ser dada de forma precisa, el manejo de la incertidumbre se hace necesario si se pretende que el sistema represente lo más fielmente posible la realidad existente. Analizaremos seguidamente algunas cuestiones fundamentales acerca de la incertidumbre.

I.1 Incertidumbre.

El término incertidumbre se entiende como duda, ambigüedad, cuestionable, problemático, no seguro.

La incertidumbre puede manifestarse de diversas formas y ser provocada por diferentes causas, que se relacionan como sigue:

- Información o conocimiento impreciso.
- Información incompleta.
- Conceptos o palabras inexactos.
- Agregación de la información desde múltiples fuentes.
- Relación de causa y efecto no absoluta.

La primera de estas causas se refiere al hecho de tener que hacer inferencias a partir, por ejemplo, de datos de los cuales no estamos completamente seguros(por errores de medición, de transmisión, etc.)

La segunda causa es motivada porque simplemente no se tiene toda la información o resulta muy costoso obtenerla o considerarla. La incompletitud de la información puede ser provocada por no tener acceso a todos los datos

relativos a las variables del problema o por que sea muy voluminosa la cantidad de datos a considerar y en la práctica no se puedan usar todos.

La tercera causa se refiere a conceptos imprecisos, no claros que han sufrido algún cambio.

La cuarta causa está presente cuando la variable es estudiada de manera no uniforme ej: diferentes instrumentos, diferentes expertos, los cuales no tienen por qué dar el mismo valor para la misma.

En el último caso este tipo de causas pueden aparecer cuando los expertos son incapaces de establecer relaciones fuertes entre las premisas y las conclusiones.

I.2 Requerimientos para un modelo de razonamiento bajo incertidumbre.

Seguidamente se presenta una lista de requerimientos que debería ser satisfecho por un formalismo ideal para representar incertidumbre y hacer inferencias bajo incertidumbre.

1. Las reglas de combinación no deben estar basadas sobre hipótesis globales de independencia de las evidencias.
2. Las reglas de combinación no deben asumir la exhaustividad y exclusividad de las hipótesis.
3. Debe haber una representación explícita de la cantidad de evidencia y los rasgos para confirmar o para refutar cualquier hipótesis dada.
4. La representación debe permitir al usuario describir la incertidumbre al nivel de detalle disponible.
5. Debe haber una representación explícita de consistencia.
6. Debe haber una representación explícita de ignorancia.
7. Debe haber una clara distinción entre un conflicto en la información(violación de consistencia) e ignorancia sobre la información.
8. Debe haber una medida de incertidumbre de segundo orden.

9. La representación tiene que ser natural al usuario(para poder suministra los datos e interpretar los resultados).

10. La sintaxis y semántica de la representación de la incertidumbre debe ser cercana a las reglas de combinación.

11. Se debe poder comparar pares de certidumbre.

12. Se debe poder seguir el proceso de agregación y propagación de la incertidumbre a través del proceso de razonamiento para resolver conflictos o contradicciones y/o para explicar la confirmación de hipótesis.

I.3 Enfoques para la representación de Incertidumbre

Los enfoques existentes para representar la incertidumbre pueden ser divididos en dos grandes categorías: caracterizaciones simbólicas y caracterizaciones numéricas. En el primer caso la incertidumbre se representa cualitativamente como un determinado criterio y el razonamiento se produce a partir de la consideración de un conjunto de avales asociados a los conocimientos y a los hechos esquivando el tratamiento matemático. Como técnicas dentro de esta tendencia podemos citar esquemas de razonamiento tales como: La lógica predefinida [Rei80], la lógica no monotónica [McD80] y la circunscripción [McC80]. En el segundo caso la idea fundamental consiste en describir la incertidumbre mediante una cierta medida numérica susceptible de ser utilizada en un conjunto de operaciones mediante las cuales pueda realizarse el proceso de inferencia de nuevos hechos a partir de evidencias y hechos conocidos. La ventaja de esta vía está en la posibilidad de desarrollar un aparato matemático que permita tratar de objetivizar algo tan eminentemente subjetivo como el razonamiento en estas circunstancias.

Las técnicas desarrolladas bajo esta teoría se conocen como *Técnicas Estadísticas para el manejo de la Incertidumbre*.

A grandes rasgos se puede decir que los enfoques simbólicos se han dirigido a problemas con enfoque impreciso; mientras que los enfoques numéricos se pueden dividir en los orientados a problemas con información imprecisa y la lógica borrosa para distintos problemas.

Enfoques Simbólicos:

- Lógica no monotónica(McDermott, 1980)
- Circunscripción(McCarthy, 1980)
- Default lógica(Recter, 1980)

Enfoques Numéricos:

- Análisis bayesianos [Dud80]
- Factores de certeza [Shor75]
- Teoría de Dempster – Shafer [Har95]
- Conjuntos Rugosos. [Paw82, Orl98, Paw95a, Paw95b].
- Redes Bayesianas [Che85]
- Sistemas borrosos [Zad65, Zad83]

I.4 Técnicas actuales de manejo de incertidumbre en los Sistemas Expertos.

Analizaremos en este capítulo el estado del arte de las técnicas actuales de manejo de incertidumbre en los Sistemas Expertos.

Valoraremos matemáticamente cuatro medidas de incertidumbre que pueden ser usadas en Sistemas Expertos, las mismas serán: Probabilidades Bayesianas, los Factores de Certeza, las Funciones de Creencia y la Lógica Difusa.

Las cuatro teorías difieren en los cálculos que ellas utilizan para definir, actualizar y combinar las medidas de incertidumbre y los juicios de independencia, es por ello que antes de aplicar cualesquiera de ellas es preciso determinar cuál es la mejor forma de modelar la incertidumbre para el problema particular y cómo podemos estimar, combinar y actualizar las medidas de incertidumbre.

En el caso de los Sistemas Expertos que sólo están interesados en un dominio estrecho de aplicación, es posible formular estrategias de asignación especial, modelos y patrones de razonamiento los cuales serán apropiados para esa aplicación. Muchas de estas estimaciones especiales podrán ser hechas por expertos del dominio y codificadas en el Sistema. Además un usuario podrá realizar algunas de estas estimaciones, pero el Sistema Experto debe ser capaz de guiarlo en este proceso.

I.5 Medidas de Incertidumbre en los Sistemas Expertos.

Las medidas de incertidumbre que son utilizadas en un Sistema Experto pueden provenir de varias fuentes. Algunas podrán ser medidas objetivas basadas en frecuencias relativas o modelos estadísticos establecidos y otras podrán ser suministrada por el experto del dominio o podrán ser hechas por el usuario del

Sistema. Todas estas medidas de incertidumbre necesitan ser combinadas por el sistema experto para hacer inferencias y tomar decisiones.

Un Sistema Experto puede ser considerado como un consultante, que suministra varios modelos, estimaciones, extrae otras desde el usuario, combina todos los juicios y finalmente informa al usuario de las conclusiones que se arriban por el sistema si este está de acuerdo con las estimaciones hechas por él. El Sistema Experto debe ser capaz, además de justificar sus estimaciones de incertidumbre y sus procesos de razonamiento y modificar algunas de sus hipótesis y estimaciones cuando sea solicitado por el usuario para hacer sus modelos y tomar las conclusiones más convincentes. Esto significa que las medidas de incertidumbre en las cuales las conclusiones se basan, deben ser aceptadas por el usuario.

A continuación se presentan estas teorías analizadas desde el punto de vista de los criterios expresados anteriormente.

I.5.1 Probabilidades Bayesianas.

La más desarrollada y mejor entendida teoría de la incertidumbre es la teoría Bayesiana . El modelo de probabilidad bayesiano ha sido usado en diferentes Sistemas Expertos como [PROSPECTOR] [Rei80, McD80] y más recientemente [And89].

Antes de definir una probabilidad hace falta una estructura previa que se conoce con el nombre de espacio probabilístico. El epacio probabilístico a su vez consta de dos elementos un conjunto **Ω** y una clase A de subconjuntos de **Ω**.

Se dice que se tiene definida una probabilidad en el espacio probabilístico(Ω, A) si a cada subconjunto de A se le asocia un número real, comprendido entre cero y uno, de tal forma que se cumplen las siguientes axiomas:

P1. $P(\Omega) = 1$

P2. $P(\bigcup_{j \in J} A_j) = \sum_{j \in J} P(A_j)$ si $A_k \cap A_j = \phi$ *para todo* $k \neq j$

Otras propiedades que se deducen de las anteriores son las siguientes:

PD1. $P(\overline{A}) = 1 - P(A)$

PD2. $P(\phi) = 0$

PD3. $P(A) \leq P(B)$ *si* $A \subset B$

PD4. $P(A) \leq 1$

PD5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

PD6. $P(\bigcup_{j \in J} A_j) = \sum_{j \in J} P(A_j)$ *si J es numerable*

PD7. $P(A \cap B) \leq \min[P(A), P(B)] \leq \max[P(A), P(B)] \leq P(A \cup B)$

La ecuación PD5 es utilizada en el Capítulo III del presente Trabajo para el cálculo de la Certidumbre.

Un tipo especial de probabilidad es la probabilidad condicionada que se relaciona con las probabilidades no condicionadas mediante la regla de Bayes:

$$P(A \cap B) = P(A / B)P(B)$$

(1.1)

Si las probabilidades no condicionadas $P(A \cap B)$ y $P(B)$ han sido asignadas con $P(B) \neq 0$, entonces la probabilidad condicional $P(A/B)$ es determinada de manera única por la regla de Bayes que se utiliza para calcular probabilidades condicionadas mediante probabilidades no condicionadas.

Si una medida de probabilidad P no condicional es especificada, la previsión(o esperanza) de una variable aleatoria X, denotada por $P(X)$, puede ser calculada $P(X)=\int xdP$, donde P es la función de densidad de X. Cuando el espacio probabilístico Ω es finito:

$$P(X) = \sum_{w \in \Omega} P(w) \cdot w$$

(1.2)

donde w representa los posibles valores de X.

La teoría Bayesiana es aplicable en la práctica en algunos tipos de eventos seleccionados o variables cuyas probabilidades pueden ser asignadas de forma precisa o añadiendo algún juicio de independencia o aplicando las reglas de la teoría para calcular otras probabilidades o previsiones. Por ejemplo si $\{A_1, A_2, \ldots, A_k\}$ es un conjunto exhaustivo de hipótesis mutuamente

excluyentes y B es un evento observado, uno podría hacer estimaciones precisas de probabilidades a priori $P(A_i)$ y probabilidades condicionadas $P(B/A_i)$ para $1 \le i \le k$ y usarlas para calcular las probabilidades predictivas:

$$P(B) = \sum_{i=1}^{k} P(B / A_i) \cdot P(A_i)$$

(1.3)

Después de observado el evento B, siendo $P(B) \neq 0$ se puede actualizar la incertidumbre sobre las hipótesis para calcular sus probabilidades a posteriori usando la regla de Bayes. No vale la pena en la teoría bayesiana tratar de calcular $P(A_i/B)$ de esta forma, ya que sería más fácil asignar las cantidades $P(A_i)$ y $P(A_i/B)$ directamente.

En principio el enfoque bayesiano puede ser aplicado en algunos problemas, pero se dificulta en la práctica por la necesidad de realizar muchas estimaciones precisas de probabilidades para determinar un modelo de probabilidad completo, chequear que estas son coherentes y que ellas determinan un modelo único de probabilidad.

La principal dificultad de la teoría Bayesiana está en que necesita modelos precisos de probabilidad, no puede modelar adecuadamente la ignorancia, información parcial, expresiones inciertas en lenguaje natural o conflictos entre opiniones de expertos.

I.5.2 Factores de certeza.

Otra forma de medir la incertidumbre es mediante los factores de certeza que nacieron en torno al sistema MYCIN[Sho76]. Para definir el factor de certeza, CF, es necesario definir previamente otras medidas MC y MI.

MC(h, e): Medida de credibilidad. Mide el decrecimiento relativo de incredibilidad de una hipótesis h debido a una evidencia o información e.

Definición intuitiva:

$$MC(h,e) = \frac{[1 - P(h)] - [1 - P(h/e)]}{1 - P(h)} = \frac{P\left(h/e\right) - P(h)}{1 - P(h)}$$

(1.4)

1-*P(h)* incredibilidad de la hipótesis h antes de conocer la evidencia e.

1-*P(h/e)* incredibilidad de la hipótesis h después de conocer la evidencia e.

El denominador *1-P(h)* lo convierte en un término relativo en vez de un valor absoluto.

Sin embargo, los autores de este concepto modificaron la idea intuitiva limitándola a valores positivos, es decir, dándole a esta medida un valor cero cuando al aplicar la fórmula anterior se obtiene un valor negativo. Por ello la definición formal es como sigue:

$$MC(h,e) = \begin{cases} 1 \quad si \;\; P(h) = 1 \\ \dfrac{\max[P(h), P(h)] - P(h)}{1 - P(h)} \quad en \;\; otro \;\; caso \end{cases}$$

(1.5)

Luego $0 \leq MC(h, e) \leq 1$.

MI(h, e): Medida de incredibilidad, mide el decremento relativo de la credibilidad de una hipótesis h debido a una evidencia o información e.

Definición intuitiva

$$MI(h,e) = \frac{P(h) - P(h/e)}{P(h)}$$

(1.6)

Definición formal

$$MI(h,e) = \begin{cases} 1 \quad si \;\; P(h) = 0 \\ \dfrac{P(h) - \min[P(h/e), P(h)]}{P(h)} \quad en \;\; otro \;\; caso \end{cases}$$

(1.7)

Donde $0 \leq MI(h, e) \leq 1$.

Finalmente el factor de cereza se define mediante la expresión:

CF (h, e)= MC(h, e) – MI(h, e)

(1.8)

donde $-1 \leq CF(h, e) \leq 1$.

Verificándose las siguientes propiedades:

1. Si h es cierto supuesto e, entonces

MC(h, e) = 1 y MI(h, e) = 0

Por lo tanto: CF = 1.

2. Si h es falsa supuesto e, entonces

MC(h, e) = 0 y MI(h, e) = 1

Por lo que: CF = –1

3. CF(h, e) + CF($\overline{h}$, e) = 0

I.5.3 Funciones de creencia.

La teoría de las funciones de creencia fue iniciada por [Dem68] en una serie de trabajos en los años 60 y desarrollada por [Sha76]. Smet ha desarrollado una variante interesante de esta teoría llamado "el modelo de creencia transferible". Otros trabajos se encuentran desarrollados en [AI-].

Una función de creencia es una función de valores reales definida en todos los subconjuntos de un espacio de posibilidad Ω la cual puede ser escrita en la forma $C_r(A) = \sum_{B \subseteq A} m(B)$ la credibilidad, $C^r(A)$, de un conjunto A, que puede ser interpretada como una cota inferior de probabilidad de A, ya que no solo se suman las probabilidades de los sucesos en la que la ocurrencia de A es segura donde m es una función de probabilidad de masa en los subconjuntos de Ω es decir $m[\Omega] = 0,\ m(B) \geq 0$ para todo $B \subseteq \Omega$ y la sumatoria $\sum_{B \subseteq \Omega} m(B) = 1$. Aquí $\subseteq$ denota inclusión de conjuntos no necesariamente estricta(en la teoría de Smet m(ϕ) puede ser positiva). Las probabilidades conjugadas superiores pueden estar definidas de la forma siguiente:

$$Pl(A) = 1 - Cr(A^c) = \sum_{B \cap A \neq \phi} m(B)$$

(1.9)

La plausibilidad

$Pl(A) = \sum_{B \cap A \neq \phi} m(B)$ que puede ser interpretada como una cota superior de la probabilidad de que ocurra A, ya que se suman tanto las probabilidades de los casos en los que A es dudoso como la de los casos en que A es seguro. Nótese que las dos medidas acotan la probabilidad.

Entre las propiedades más importantes de estas medidas destacan:

1. $Pl(A) \geq C_r(A)$.
2. $Pl(\varnothing) = C_r(\varnothing) = 0$
3. $Pl(\Omega) = C_r(\Omega) = 1$
4. $C_r(A)+C_r(\overline{A}) \leq 1$
5. $Pl(A)+Pl(\overline{A}) \geq 1$
6. $Pl(A)+C_r(\overline{A}) = 1$

La propiedad 6 da lugar a otra definición de plausibilidad que es dada por

$Pl(A) = 1 - C_r(\overline{A})$

Las funciones de creencia son frecuentemente generadas por un mapeo multivaluado el cual es una función A desde un espacio fundamental

$\Psi = \{\psi_1, \psi_2, ...\psi_n\}$ de subconjuntos de Ω. Para simplificar asumamos que los conjuntos $A(\psi_1), A(\psi_2), ..., A(\psi_n)$ son bien definidos. Esto posibilita asignar una medida de probabilidad bayesiana P sobre Ψ utilizando información frecuencial o juicios subjetivos y esto genera una función de creencia sobre Ω a través de $m(A(\psi_i)) = P(\psi_i)$.

Las funciones de creencia son frecuentemente estimadas usando un mapeo multivaluado y estimando probabilidades Bayesianas. En otros problemas la asignación de probabilidad m_i se puede realizar directamente. Por ejemplo puede ser información frecuencial acerca de los resultados $B_1,\ldots,B_n$ de pruebas previas donde los resultados son registrados como subconjuntos B_i de Ω.

Dempster y Shafer proponen el uso del intervalo [Cr(A), Pl(A)] en lugar de una medida de incertidumbre de A. Este intervalo juega un importante papel práctico ya que permite medir la ignorancia y la incertidumbre de A.

I.5.4 Lógica Difusa.

La Lógica Difusa es un súper conjunto de la lógica convencional (booleana) que ha sido extendida para manipular el concepto de verdad parcial, es decir valores de verdad entre "completamente verdadero" y "completamente falso". Fue introducida por Zadeh [Zad65] en los años 60 como una vía de modelar incertidumbre en el lenguaje natural.

En la teoría tradicional de conjuntos, un elemento está o no en un conjunto. En otras palabra, el grado de pertenencia de un elemento a un conjunto es binario 0 o 1. En la teoría de los conjuntos difusos el concepto de pertenencia es generalizado. Un elemento x en un conjunto difuso A puede estar parcialmente en ese conjunto, con un cierto grado de pertenencia $\mu_A(x)$. Esta generalización de grados de pertenencia permite una modelación más real de

muchas situaciones. Por tanto, un conjunto difuso está formado por varios elementos que tienen asociado un grado de pertenencia.

Definición(después de [Zad65])

Sea U un frame de discernimiento(conjunto de todos los posibles valores x de un atributo). Un conjunto difuso A en U es caracterizado por una función de pertenencia $\mu_A : U \to [0,1]$. El valor $\mu_A(x)$ para $x \in U$ representa el grado de pertenencia de x a A. La función característica μ_A puede ser entendida como una medida del grado de compatibilidad de x con el concepto A.

Es importante entender que los conjuntos difusos que se definen son arbitrarios y muy fácilmente modificados para permitir al sistema o al experto del dominio calibrar la repuesta a lo que ellos creen apropiado.

Operaciones Básicas.

1. ***Unión e intersección.*** En la teoría de conjuntos tradicional la unión e intersección tienen significado análogo al de la Lógica Difusa. La Unión de conjuntos difusos incluye una operación de máximo como sigue:

$$A \cup B = \{\max[\mu_A(x), \mu_B(x)] | x\}$$

(1.10)

La operación Intersección incluye la operación de mínimo.

$$A \cap B = \{\min[\mu_A(x), \mu_B(x)] | x\}$$

(1.11)

Estas operaciones dan el mismo resultado que las operaciones de $\cup$ e $\cap$ cuando usamos conjuntos no difusos. En la lógica difusa las operaciones de unión e intersección son usadas para mezclar información del mismo tipo.

2. ***Relación***. Un nivel de construcción más alto en la lógica difusa que ha sido usado para cuantificar o combinar interacciones entre diferentes tipos de información es la relación.

Sean

$$A = \sum \mu_A(x_i) | x_i$$
$$B = \sum \mu_B(y_j) | y_j$$

La relación se define como:

$$R = A \times B = \sum \sum \mu_R(x_i, y_j) | (x_i, y_j)$$

(1.12)

donde para variables discretas:

$$\mu_R(x_i, y_j) = \min[\mu_A(x_i), \mu_B(y_j)]$$

(1.13)

Esta operación es análoga a la unión de distribución de probabilidades.

3. ***Composición.*** Un segundo nivel de construcción el cual es usado para actualizar información es llamado composición. Su utilidad puede ser expresada de la siguiente forma: dada una relación R cual es el subconjunto inducido mediante la aplicación del conjunto difuso x. Explícitamente,

$$y = x \cdot R$$
$$R = A \times B, \quad A \in U, \quad B \in V, \quad x \in U, \quad y \in V$$

Computacionalmente la composición incluye tomar el máximo de algunos mínimos (es decir $\max_y[\min\langle \mu_x(y), \mu_R(x,y) \rangle]$). La aplicación de la composición puede resumirse como sigue: si existe una relación conocida entre dos piezas de información y existe una opinión de otra del mismo

dominio de las dos primeras piezas, esta opinión debe ser usada para actualizar el estado actual de la información.

Combinación de Evidencias.

Un método que usa varias de las primitivas discutidas para combinar evidencias de forma significativa ha sido propuesto y desarrollado posteriormente. Este es llamado promediación pesada difusa(FWA). Este mecanismo toma el valor y la importancia relativa(peso) de la información de cada nodo para ser combinada. El FWA puede ser visto como una unión de intersecciones.

4. ***Promediación pesada difusa(FWA).***

$$Belief = \frac{\sum \left(Belief_i \cdot Weight_i \right)}{\sum Weight_i}$$

(1.14)

5. Adición difusa.

$$A + B = \left\{ \min[\mu_A(x), \mu_B(y)] \middle| (x + y) \right\}$$

(1.15)

6. ***Multiplicación difusa***.

$$A \cdot B = \left\{ min[\mu_A(x), \mu_B(y)] \middle| x \cdot y \right\}$$

(1.16)

7. División difusa.

$$A/B = \left\{ \min[\mu_A(x), \mu_B(y)] \middle| x/y \right\}$$

(1.17)

I.5.5 Conjuntos Rugosos.

Sea U un conjunto finito, no vacío llamado universo y sea I una relación binaria definida sobre U. Denotemos por $I(x)$ el conjunto de todos los $y \in U$, tal que yIx. Si I es reflexiva y simétrica, es decir xIx para todo $x \in U$ y xIy implica yIx para todo $x, y \in U$ entonces I es una relación de tolerancia. Si I es además transitiva, es decir xIy y yIz implican xIz entonces I es una relación de equivalencia. En este caso $I(x)=[x]_I$ es decir $I(x)$ es una clase de equivalencia de la relación I. Si I es una relación de tolerancia y xIy entonces x y y son llamados similares con respecto a I(I-similares). Mientras que si I es una relación de equivalencia y xIy entonces x, y se dicen indiscernibles con respecto a I (I-indiscernibles).

Se definen las siguientes operaciones:

$$I_*(X) = \{x \in U : I(x) \subseteq X\}$$

(1.18)

$$I^*(X) = \{x \in U : I(x) \cap X \neq \Phi\}$$

(1.19)

que asigna a cada subconjunto X del universo U dos conjuntos $I_*(X), I^*(X)$ llamados aproximaciones inferior y superior de X respectivamente. Al conjunto

$$BN_I(X) = I^*(X) - I_*(X)$$

(1.20)

se le denomina región frontera de X.

Si la región frontera de X es vacía, es decir $BN_I(X) = \Phi$, el conjunto X es llamado exacto con respecto a I.

Por tanto los conjuntos rugosos pueden ser usados para modelar la vaguedad.

La vaguedad puede ser caracterizada numéricamente definiendo el siguiente coeficiente:

$$\alpha_I(X) = \frac{\left|I_*(X)\right|}{\left|I^*(X)\right|}$$

(1.21)

donde $\left|X\right|$ denota la cardinalidad del conjunto X.

La incertidumbre desde la perspectiva de los conjuntos rugosos se maneja a través de una función de pertenencia la cual es definida empleando la relación *I* de la siguiente manera:

$$\mu_X(x) = \frac{\left|X \cap I(x)\right|}{\left|I(x)\right|}$$

(1.22)

Esta función nos da el grado de pertenencia de un conjunto elemental a un concepto.

Capítulo II

Análisis del lugar de la incertidumbre en los Sistemas Basados en Conocimiento.

En los Sistemas Basados en Conocimiento que consideran la existencia de incertidumbre, la que puede aparecer tanto en la formación del propio conocimiento como en los datos que aparecen al formular un problema concreto, el enfoque de trabajo usual ha sido realizar el proceso de razonamiento sin considerar la incertidumbre existente y paralelamente realizar un proceso de cálculo de la incertidumbre asociada con la conclusión. En este proceso la combinación de incertidumbres provenientes de diversas fuentes se hace necesaria al menos en dos niveles. En el primero se deriva una conclusión utilizando una pieza del conocimiento almacenado en la base de conocimiento y es necesario combinar los diferentes valores de las incertidumbres asociadas con las hipótesis en las que se basa esa pieza de conocimiento y la incertidumbre de la propia pieza de conocimiento para determinar la incertidumbre de la conclusión alcanzada. En el segundo nivel es necesario combinar diversas conclusiones (cada una con su propia incertidumbre) para alcanzar la respuesta. A este proceso se le denomina propagación de la incertidumbre desde los datos a la respuesta.
Las técnicas anteriormente expuestas en el Capítulo I han sido usadas en [Paw95a, Paw95b] y otros Sistemas Expertos. A continuación se describe el uso de una de ellas en un Sistema Basado en Casos.

II- 1 Sistemas Basados en Casos. Incertidumbre

Definición
Los Sistemas Basados en Casos son Sistemas Basados en el Conocimiento que utilizan, como método de solución de problemas, el Razonamiento Basado en Casos. El Razonamiento Basado en Casos es un paradigma de solución de problemas que difiere de otros enfoques y técnicas en que es capaz de utilizar el conocimiento específico adquirido en situaciones previas y utilizarlo en la situación presente. Un problema nuevo se resuelve buscando en su memoria un caso similar resuelto en el pasado. Además incrementa su conocimiento almacenando el nuevo caso para ser usado en situaciones futuras. Esto permite que su conocimiento se mantenga actualizado en todo momento [Gar92].
Componentes de un Sistema Basado en Casos
Los componentes principales de un Sistemas Basados en Casos son:
La Base de Conocimiento o Base de Casos:
Es la que contiene los casos con las descripciones de los problemas resueltos y también los no resueltos previamente; estos últimos son aquellos problemas a los que se les dio una "solución" que posteriormente mostró ser errónea y son útiles para prevenir errores en el futuro. Si luego de la ejecución ocurrió algún error en la

predicción contienen además la explicación de las anomalías, las estrategias de reparación y referencias al próximo resultado.

En el procedimiento de recuperación de casos :

En él se determinan los casos que mayor similitud guardan con el patrón de búsqueda. Esto puede realizarse por semejanza parcial o por analogía. En el primer método, se emplea una función de semejanza la cual determina una medida numérica del grado de similitud del objeto de la memoria con respecto al nuevo problema. En el segundo, se busca un elemento de la memoria cuya descripción pueda hacerse igual a la del nuevo problema, mediante el reemplazamiento de los valores de los rasgos diferentes a partir de la equivalencia de valores, definida en una red semántica.

El Módulo o Procedimiento de Adaptación de las Soluciones:

Es el que se encarga de encontrar los valores de los rasgos desconocidos a partir de los casos recuperados mediante un algoritmo capaz de adaptar las soluciones de los casos semejantes para completar el nuevo problema. Atendiendo a dicho módulo de adaptación se distinguen los dos tipos de sistemas basados en casos: los interpretativos y los solucionadores de problemas. Un sistema Basado en Casos interpretativo toma un problema y su solución y utiliza los casos para proponer un argumento para justificar esa solución. Un solucionador de problemas deriva soluciones al nuevo problema a partir de soluciones viejas a problemas similares.

II-1.1 Etapas de Razonamiento Basado en Casos

Asignación de índices.

El sistema recibe el caso de entrada y analiza sus características y atributos más relevantes (índices) para poder acceder a los casos de mayor similitud en base de casos. A los índices usados para medir la similitud entre los casos se les denomina rasgos predictores, mientras que a los índices de los cuales deseamos obtener su valor se les denomina rasgos objetivos.

Acceso a casos similares.

El siguiente paso es el acceso a casos similares en la base de casos a través del procedimiento de recuperación. Esto se realiza usando los rasgos predictores encontrados en el paso anterior. De esta forma se obtienen los casos más similares al actual. En muchas ocasiones es necesario reducir aun más el conjunto de casos para obtener los más relevantes.

Creación de la solución.

Una vez seleccionados los casos similares, se efectúa el procedimiento de adaptación, que consiste en la modificación y combinación de las soluciones de los casos similares para formar una nueva solución, una interpretación, o una explicación dependiendo de la tarea que lleve a cabo el sistema.

Evaluación de la solución.

La solución obtenida en la fase anterior, representa una solución potencial o tentativa. Es necesario hacer una evaluación de la solución propuesta antes de darla

al usuario, es decir, se intenta ver cuáles son las cualidades y debilidades de la solución para evaluar su utilidad.

Asignación y almacenamiento del caso nuevo.

Una vez creada la solución y evaluada, se procede a entregar la solución al mundo real. Además, se puede crear un nuevo caso. Este caso se forma no sólo de la solución creada, sino del caso o problema original, la solución encontrada, las justificaciones y explicaciones (si las hay). El caso nuevo se le asignan índices y se almacena en la base de casos.

Explicación, reparación y prueba.

Si la solución fracasa, es importante para el sistema obtener y analizar información de retroalimentación para no repetir los mismos errores. Si ocurre algo inesperado el sistema debe tratar de explicarlo. Después se trata de reparar la solución con base en esta explicación y se regresa al paso de evaluación.

Más información acerca del tema se puede encontrar en [GAR94, KOL93].

II.1.2 Ventajas y Desventajas del Razonamiento Basado en Casos sobre otras tecnologías.

El RBC como nuevo método de solución de problemas para el desarrollo de sistemas basados en conocimiento presenta de un determinado número de ventajas entre las que se pueden citar:

Ventajas.

- La adquisición del conocimiento. La unidad básica del conocimiento es el caso y no la regla. Los seres humanos por lo general articulan su conocimiento mediante ejemplos de problemas y soluciones anteriores, más que por medio de reglas específicas y abstractas.
- Permite proponer soluciones a problemas rápidamente. Esto lo logra evitando el tiempo necesario para derivar estas respuestas a partir de cero.
- Aprendizaje. A diferencia de los sistemas basados en reglas convencionales, si la misma situación se presenta repetidamente, no se tiene que construir o generar la misma solución a partir de cero cada vez que se presente el mismo problema.
- Propone soluciones en dominios no entendidos completamente por el sistema.
- Ofrece un medio de evaluación de soluciones cuando no se cuenta con un método algorítmico de evaluación.
- Se centra en las características o partes más importantes de un problema.

Desventajas.

- Confía ciegamente en los casos previos almacenados en su memoria para intentar proponer su solución.
- Puede ser que no recupere el caso más apropiado para la solución del nuevo caso.

Estas desventajas podrían eliminarse con la introducción de un mecanismo de manejo de la incertidumbre en estos sistemas.
El desarrollo de este mecanismo constituye el objetivo fundamental del presente trabajo y a el le dedicaremos los próximos epígrafes.

II.1.3 Modelo General de un Sistema Basado en Casos sin Manejo de Incertidumbre.

Descripción de un caso.

Un caso de la base es descrito a partir de los valores que se le asignan a los rasgos predictores y los rasgos objetivos. Los rasgos predictores son los que determinan los valores de los rasgos objetivos. Un caso O_t que tenga n rasgos predictores y p rasgos objetivos se describe del siguiente modo:

$$O_t\left(x_1(O_t),...,x_i(O_t),...,x_n(O_t), y_1(O_t),...,y_j(O_t),...,y_p(O_t)\right)$$

(2.1)

donde: x_i, y_i : rasgo predictor y objetivo respectivamente.

$x_i(O_t)$: valor del rasgo predictor $i = 1, n$ para el caso O_t t =1, m de la base

$y_j(O_t)$: valor del rasgo objetivo $j = 1, p$ para el caso O_t t =1, m de la base.

A los valores que toman cada rasgo se les llaman categorías y de forma general son en número igual o mayor a 2 y es precisamente a través de esas categorías que se enlazan los rasgos predictores con el o los rasgos objetivos.
Una base de casos es, como se dijo, una colección de casos que puede ser representada a través de una tabla de decisión, en la cual, las columnas son etiquetadas por variables que representan los rasgos predictores y objetivos (decisiones) y las filas representan los casos.
La Tabla 2.1 muestra un ejemplo de una base de casos. Aquí los casos se representan por $O_1, ..., O_m$, los rasgos predictores por: $x_1, x_2, ..., x_n$ y la decisión por y_1. El valor del rasgo x_i en el caso O_t se representa por $x_i(O_t)$ y el valor de la decisión en dicho caso se representa por $y_1(O_t)$.

Caso	x_1	**…**	x_n	y_1	y_2	**⋯**	y_p
O_1	$x_1(O_1)$	⋯	$x_n(O_1)$	$y_1(O_1)$	$y_2(O_1)$	⋯	$y_2(O_1)$
…	…	…	…			…	
O_m	$x_1(O_m)$	⋯	$x_n(O_m)$	$y_1(O_m)$	$y_p(O_m)$	⋯	$y_p(O_n)$

Tabla 2.1. Tabla de decisión que representa una base de casos.

En lo adelante la base de casos que representemos usando esta definición le nombraremos Base de casos original y en la medida que se calculen otros parámetros con los datos almacenados en ella se irá completando hasta formar la base de casos que usaremos en el módulo de recuperación.

Módulo de Recuperación

En el módulo de recuperación es donde se realiza el acceso a casos similares en la base a través del procedimiento de recuperación, haciendo uso de los rasgos predictores a que hicimos referencia anteriormente. De esta forma se obtienen los casos más similares al problema actual. En muchos problemas prácticos es necesario reducir el conjunto de casos para obtener los más relevantes. En el presente trabajo se propone una vía alternativa.

Los dos aspectos claves de esta fase son el algoritmo de acceso a casos y la medida de similitud entre casos. Para determinar que tan similar es un caso a otro se han desarrollado varias técnicas. La más sencilla consiste en contar el número de características similares entre los dos casos. El problema de esta técnica es que la importancia de las características varia de un contexto a otro. Otra técnica consiste en utilizar un conjunto de heurísticas que permitan determinar cuales características tienen mayor relevancia (peso) y formular una función de semejanza que involucre la similitud entre cada uno de los rasgos teniendo en cuenta el peso de los mismos. Un modelo matemático de esta última técnica es el siguiente:

Función de semejanza entre un nuevo problema a resolver y un caso O_t de la base

$$\beta(O_0, O_t) = \frac{\sum_{i=1}^{n} p_i \cdot \delta_i(O_0, O_t)}{\sum_{i=1}^{n} p_i} \qquad (2.2)$$

donde:

$O_0 = O_0(x_1(O_0),...,x_n(O_0))$: Problema a resolver

n : Número de rasgos predictores.

p_i : Peso o relevancia del rasgo *i*.

$\delta_i(O_0, O_t)$: Función de comparación entre los casos O_0 y O_t atendiendo al rasgo *i*. Esta función puede estar definida de diferentes formas, por ejemplo:

Para rasgos numéricos.

$$\partial_i(O_0, O_t) = \begin{cases} 1 & si \quad \mathrm{x}_i(O_0) = x_i(O_t) \\ 0 & e.o.c \end{cases} \qquad (2.3)$$

$$\partial_i(O_0, O_t) = \begin{cases} 1 & si \quad |x_i(O_0) - x_i(O_t)| < \varepsilon \\ 0 & e.o.c \end{cases}$$

(2.4)

ε : Valor umbral definido por el usuario.

$$\partial_i(O, O_t) = 1 - \frac{|x_i(O) - x_i(O_t)|}{\max_i - \min_i}$$

(2.5)

$\max_i, \min_i$: Valores máximo y mínimo respectivamente que alcanza el rasgo i.

Por otro lado, el algoritmo de acceso a los casos debe ser rápido y eficiente. Este depende de las técnicas de indexación usadas y su diseño se vuelve un aspecto crítico cuando la base de casos es muy grande.

Entre las técnicas de acceso a casos podemos mencionar:

Acceso exacto. Se busca el caso que encaje exactamente con el caso nuevo.

Acceso Jerárquico. Básicamente se buscan los casos en un árbol, tomando decisiones en cada nodo hasta que ya no se pueda avanzar. Si se encuentra una hoja se muestran los casos correspondientes, si se llegó hasta un nodo, se muestra todos los casos que se deriven de él. Esta última técnica es la que se usa en nuestro modelo.

Módulo de Adaptación

Una vez seleccionados los casos similares, se efectúa el procedimiento de adaptación, que consiste en la modificación y combinación de las soluciones de los casos similares para formar una nueva solución, una interpretación, o una explicación dependiendo de la tarea que lleve a cabo el sistema. Es deseable que el sistema también genere justificaciones o explicaciones que apoyen a la solución creada para el nuevo caso.

Al acceder a los casos similares, lo mejor que puede pasar es encontrar un caso exactamente igual al nuevo, pero generalmente no sucede así, y es necesario realizar el proceso de adaptación. La adaptación puede realizarse por el usuario o de manera automática por el sistema.

Si la adaptación la realiza el usuario, el sistema sólo realiza la búsqueda de los casos similares.

Si la adaptación la ejecuta el sistema, entonces debe contener algún conocimiento, tal como fórmulas o reglas. Las técnicas desarrolladas para la realización de la adaptación en forma automática han dado en general buenos resultados. En [Kol93] se mencionan los siguientes métodos de adaptación:

Reinstanciación. En este tipo de técnica se utiliza el marco o contexto de situaciones anteriores, pero con nuevo argumentos.

Ajuste de parámetros. Se ajustan los parámetros de la solución de casos anteriores de acuerdo con las diferencias entre las descripciones de los casos en cuestión.
Búsqueda local. Se realizan búsquedas dentro de jerarquías semánticas y se soluciona el problema por analogía.
En el presente trabajo se propone un nuevo enfoque donde se hace uso de la certidumbre de la solución.

II .2 Solución a la problemática de los Sistemas Basados en Casos.

Fundamentación.

Si en una base de casos se pueden definir los rasgos o variables predictores y los rasgos objetivos, los primeros pueden influir de diferente forma en los valores que toman los segundos, es decir, dentro de los rasgos que influyen sobre el objetivo (llamados también variables dependientes) existen unos que están más fuertemente correlacionados con él que otros, estos son los rasgos más predictores otros, en ocasiones, pueden correlacionarse con el objetivo solamente a través de otro u otros rasgos, serán de este modo considerados como menos predictores.

II.2.1 Incertidumbre en los Sistemas Basados en Casos

La información incompleta es omnipresente: casi toda la información que se tiene del mundo es no cierta, no completa o no precisa. La incertidumbre cala nuestra compresión del mundo real.

El propósito de los sistemas de información es modelar el mundo real. Por ello dichos sistemas tienen que ser capaces de modelar la incertidumbre.

A nivel empírico, la incertidumbre es casi una medida inseparable de casi cualquier media, resultante de una combinación de inevitables errores de medición y límites de resolución de los instrumentos de medición utilizados.

A nivel cognitivo, esta emerge de la vaguedad y ambigüedad inherente al lenguaje natural. Se ha hecho claro que hay distintos tipos de incertidumbre. La incertidumbre es un concepto multidimensional. Las teorías matemáticas clásicas para caracterizar situaciones bajo incertidumbre han sido la teoría de conjuntos y la teoría de las probabilidades [Kli95, Kli98].
El término incertidumbre se utiliza de forma general para abarcar diferentes tipos de incertidumbre: imprecisión, incompletes, vaguedad e inconsistencia.

La imprecisión es relativa a la existencia de un valor el cual no puede ser medido con la precisión adecuada; *la incompletes* a la ausencia de valores para algunas variables consideradas, *la vaguedad* a la ausencia de un límite claro para un concepto o conjunto y *la inconsistencia* describe una situación en la cual hay dos o más valores en conflicto para ser asignados a una variable [Bon85, Bos93, Par96].
Obviamente el Razonamiento Basado en Casos no escapa de la problemática de la incertidumbre. Hay posibilidades de imprecisión tanto en el conocimiento (base de casos) como en el planteamiento del problema.

El conocimiento del problema radica esencialmente en la colección de casos, luego la primera interrogante es si el caso representa o no el objetivo para el cual fue

puesto en la base, es decir ¿cuál es el grado de certeza de que la solución contenida en el caso sea válida?. Otro aspecto referido con el conocimiento del sistema es el modelo de organización de los casos. Un buen modelo de organización debe permitir almacenar los casos quizás no como un todo, y luego recuperar los casos más semejantes reconstruyendo los mismos, aquí la fuente de incertidumbre esta dada por la medida en que el modelo permite descomponer y recomponer los casos.
Los resultados planteados en este trabajo están orientados a resolver el primer aspecto, es decir determinar la certidumbre de los valores representados en la base y una vez determinada, manipularla en los módulos de recuperación y adaptación de forma tal que se solucionen las siguientes interrogantes:

1. Supongamos que estamos trabajando con tres rasgos predictores: *FORMA, TAMAÑO* y *COLOR* y que por algún método dado, se ha determinado que el rasgo *FORMA* es el rasgo más importante y por tanto el que mayor peso tiene. Sin embargo, los valores que ha tomado ese rasgo tienen muy bajo nivel de certidumbre, surge entonces la próxima interrogante ¿es correcto asignar un peso grande a un rasgo del cual estamos tan inseguros?
2. Considérese el rasgo *COLOR* con los valores *{AZUL, ROJO, VERDE, NEGRO}*, tener en cuenta la incertidumbre, significa que el rasgo *COLOR* toma un valor en el caso almacenado y otro en la descripción del problema a resolver, pero sobre ambos valores hay dudas sobre su certeza. Por ejemplo*: COLOR=AZUL (0.9)* en el caso *y COLOR=ROJO (0.6)* en la descripción del problema. La situación anterior nos obliga a pensar en dos aspectos. El primero es que ahora la función de comparación tendrá como argumentos valores sobre los que existen dudas y por lo tanto el valor de la función de comparación entre estos colores dadas las certidumbres 0.9 y 0.6 no tiene que ser igual que en el caso en que ambos valores de certidumbres sean 1. Considerar el valor de la certidumbre en los valores dados a los rasgos nos obliga a redefinir las funciones de comparación de rasgos definidas originalmente. De aquí surge la siguiente interrogante: ¿cómo construir funciones de comparación de rasgos considerando incertidumbre en los valores dados a estos?.

3. Supongamos que los casos recuperados con mayor nivel de semejanza son O_5 y O_{15} (ambos con el mismo nivel de semejanza) y que en ambos casos el valor del rasgo objetivo es diferente (situación de conflicto). ¿Por qué caso adaptar?

Hemos señalado algunos momentos del razonamiento basado en casos donde resulta necesario tener en cuenta la manipulación de la incertidumbre. En el siguiente epígrafe se presentan algunas ideas preliminares sobre la elaboración de un modelo para la manipulación de la incertidumbre en los SCBC.

II.2.2 Modelo para la Manipulación de la Incertidumbre en una Base de Casos

Descripción del Problema

Se resuelve el problema de la toma de decisiones utilizando el razonamiento Basado en Casos en condiciones de incertidumbre. Desde la perspectiva del enfoque basado en casos los problemas de decisión se describen de la forma siguiente:

- Un conjunto de rasgos que describen la situación S (rasgos predictores).
- Un conjunto de posibles decisiones a tomar D (rasgos objetivos).
- La incertidumbre presente en los rasgos predictores.
- La incertidumbre presente en los rasgos objetivos.

Luego, los casos en la base se componen de un conjunto de rasgos predictores de la forma (valor, certidumbre) y uno o más rasgos objetivos de la forma (valor, certidumbre), donde los valores son discretos y de cardinalidad finita.
Un nuevo problema se describe en términos de los valores que toman los rasgos predictores en la nueva situación donde el agente tiene que tomar la decisión y el sistema debe ser capaz de inferir qué decisión debe tomar a partir de la experiencia acumulada en la base de casos. Luego el modelo a desarrollar tiene que:

1. Dar una vía para calcular la incertidumbre presente en los casos almacenados en la base de casos (pues usualmente la base de casos se crea a partir de fuentes donde no están explícitamente registrados los valores de incertidumbre). Es decir, debemos encontrar para cada valor $x_i(O_t)$ dado al rasgo predictor i (i=1,n) en el caso O_t (t=1,m) una medida $\zeta_i(O_t)$ que denota la certidumbre de dicho valor y

para cada decisión $y_j(O_t)$ dado al rasgo objetivo j (j=1,p) en el caso O_t (t=1,m), una medida $\eta_j(O_t)$.

2. Desarrollar un procedimiento de recuperación, para la selección de los casos más semejantes al nuevo problema a partir de las descripciones planteadas en el caso y la situación donde el agente debe tomar la nueva decisión, considerando que en la descripción de ambas situaciones hay incertidumbre.
3. Desarrollar un proceso de adaptación que a partir de las decisiones tomadas en los casos recuperados y la incertidumbre asociada determine cuál debe ser la decisión a tomar y con qué certidumbre se toma.

Estructura General del Modelo Desarrollado

El cálculo de la certidumbre asociada al valor de cada rasgo se calcula a partir de dos magnitudes. La primera de ellas tiene en cuenta la información almacenada en la base de casos y representa la certidumbre de que un rasgo tome un valor dado de su dominio conociéndose la decisión tomada. La segunda magnitud representa la certidumbre del valor conocida su fuente de obtención (exacta, criterios de expertos, instrumentos de medición) y por tanto puede variar de uno a otro caso. Ambas magnitudes son calculadas como probabilidades condicionadas y se les denominan *Primera* y *Segunda Probabilidad* respectivamente (para abreviar P_1 y P_2).

La probabilidad P_1 se calcula usando el árbol de decisión que describe la relación de dependencia entre los valores de las variables predictoras y objetivos. Para la construcción del árbol se calculan los niveles de asociación de las variables predictoras respecto a las variables objetivos, usando el Test Ji-Cuadrado [MAR91, CHA94]. De esta forma en el árbol sólo estarán presentes aquellos rasgos cuyo nivel de asociación este por encima de un valor predefinido por el experto. Estos son los rasgos que se consideran dependientes. El resto de los rasgos se consideran independientes entre sí y con respecto a los rasgos objetivos y para ellos la probabilidad P_1 de un valor se calcula como su frecuencia de aparición en la base. Esta idea conduce a la creación de un Modelo de dependencia con síntomas relevantes según [Cas89].

La creación del árbol permite simplificar el proceso de selección de los casos más semejantes, pues permite no tener que considerar todos los casos en el proceso de recuperación como se ilustra en el epígrafe II.3. Por otra parte los niveles de asociación calculados permiten realizar la asignación de los pesos o su modificación, en caso de que los pesos hayan sido previamente asignados.

Una vez calculadas ambas probabilidades para calcular la certidumbre total $\zeta_i(O_t)$ se combinan éstas una co-t-norma g.

$$\zeta_i(O_t) = g(P_1, P_2)$$

(2.6)

Para el nuevo caso esta certidumbre se calcula teniendo en cuenta sólo P_2.

Una vez calculada la certidumbre de todos los valores para los rasgos predictores y objetivos y la certidumbre de los valores del nuevo caso, se procede a realizar los procesos de razonamiento (recuperación y adaptación).

El procedimiento de recuperación se basa en una función de semejanza la cual compara la descripción del nuevo problema O_0 con las descripciones de los casos O_t almacenados en la base, considerando la importancia de cada rasgo o peso w_i, los valores de las comparaciones a nivel de rasgo ∂_i y las certidumbres de los valores comparados.

$$\beta(O_0, O_t) = f_1(w_i, \delta_i(O_0, O_t), m_i)$$

(2.7)

$$m_i = f_2(\zeta_i(O_0), \zeta_i(O_t))$$

(2.8)

Una vez seleccionados los casos más semejantes es posibles que estos propongan decisiones diferentes por lo que se hace necesario determinar cuál de las decisiones debe tomarse. Esta elección se realiza teniendo en cuenta el grado de semejanza calculado en el procedimiento anterior y a la certidumbre de la decisión de los casos recuperados.

$$\mu(\beta(O_0, O_t), \eta_j(O_t))$$

(2.9)

II.3 Modelo Matemático para la Determinación de la Certidumbre

Este modelo está enfocado a la existencia de un solo rasgo objetivo y_1, aunque pudiera ser extendido al caso en que existan múltiples decisiones.

II.3.1 Fundamentos Matemáticos para la Construcción del Árbol de Decisión

Sea y_1 la decisión (rasgo objetivo), la cual toma valores en conjunto: $\{y_{11}, \ldots y_{1\theta}\}$ donde θ es la cardinalidad del conjunto de valores que toma dicho rasgo objetivo.

Sean x_i (i=1,n) el conjunto de rasgos predictores, los cuales toman los valores en los conjuntos $\{x_{i1},...x_{i\eta_i}\}$ donde η_i es la cardinalidad del conjunto de valores que toma el rasgo predictor x_i.

Para cada variable predictora x_i determinamos la tabla de contigencia que la relaciona con la variable objetivo y_1 de la siguiente forma:

$x_i \diagup y_1$	y_{11}	....	y_{1q}	....	$y_{1\theta}$	
x_{i1}	$O_{11}(x_i)$	....	$O_{1q}(x_i)$	....	$O_{1\theta}(x_i)$	$\sum_{j=1}^{\theta} O_{ij}(x_i)$
....	....	....	....	....	.....	.....
x_{ip}	$O_{p1}(x_i)$	...	$O_{pq}(x_i)$		$O_{p\theta}(x_i)$	$\sum_{j=1}^{\theta} O_{pj}(x_i)$
...	....	...	.....	....	.....	.....
$x_{i\eta_i}$	$O_{\eta_i 1}(x_i)$	....	$O_{\eta_i q}(x_i)$	....	$O_{\eta_i \theta}(x_i)$	$\sum_{j=1}^{\theta} O_{\eta_i j}(x_i)$
	$\sum_{i=1}^{\eta_i} O_{i1}(x_i)$	....	$\sum_{i=1}^{\eta_i} O_{iq}(x_i)$	...	$\sum_{i=1}^{\eta_i} O_{i\theta}(x_i)$	n

Tabla 2.2 .Tabla de contingencia que relaciona las variables x_i y y_1

Donde:

$O_{pq}(x_i)$ es el número de casos que toman simultáneamente los valores x_{ip} y y_{1q}, $p=1,\eta_i, \quad q=1,\theta$

Los valores $O_{pq}(x_i)$ que se muestran en la Tabla 2.2, constituyen las frecuencias observadas. Sin embargo, de acuerdo a las reglas de probabilidad estas frecuencias debían ser :

$$E_{pq}(x_i) = \frac{\sum_{i=1}^{\eta_i} O_{iq}(x_i) \cdot \sum_{j=1}^{\theta} O_{pj}(x_i)}{n}$$

(2.10)

que son las llamadas frecuencias esperadas o teóricas.

Estas frecuencias esperadas deben ser comparadas con las frecuencias observadas $O_{pq}(x_i)$ en la Tabla 2.2

Las diferencias $O_{pq}(x_i) - E_{pq}(x_i)$ se llaman residuales, se elevan al cuadrado para evitar la compensación de diferencias positivas y negativas y se dividen por las frecuencias esperadas $E_{pq}(x_i)$ para establecer magnitudes relativas. Resulta en el estadístico:

$$\chi^2(x_i) = \sum_{p=1}^{\eta_i} \sum_{q=1}^{\theta} \frac{\left(O_{pq}(x_i) - E_{pq}(x_i)\right)^2}{E_{pq}(x_i)} \tag{2.11}$$

Si la hipótesis fundamental de independencia es cierta, este estadístico tiene distribución aproximadamente igual a la Chi-cuadrado, con grados de libertad determinado por el producto (C-1)*(R-1), donde C y R son el número de filas y columnas de la Tabla 2. La idea de los grados de libertad es aquí clara pues este es efectivamente, el número de celdas de la Tabla 2 que "podrían llenarse libremente" si están fijados los totales marginales de filas y columnas.

La prueba Chi-cuadrado tiene realmente muchas limitaciones y los principales detractores llegan incluso a decir que el único caso en que él puede ser aplicado con fiabilidad, es el caso de las tablas 2*2. Una de las principales limitaciones es que suministra muy poca información sobre la forma y sobre la estrechez de la asociación entre las variables. Por ello junto con el Chi-cuadrado o en su lugar se utilizan otros estadísticos o pruebas que hablan de la forma o fortaleza del enlace.

Otras medidas de asociación entre variables aleatorias discretas nominales y ordinales.

En muchas investigaciones, más que discernir la dependencia de dos variables interesa la naturaleza y fortaleza de la asociación. Los indicadores que miden esto se llaman *medidas de asociación*. Ninguna medida resume adecuadamente todos los tipos de asociación posibles. Las medidas se diferencian por su interpretación, y por la forma en que ellas pueden reflejar asociaciones perfectas o parciales. Las medidas difieren además por la forma en que ellas se afectan por otros factores, por ejemplo los totales marginales. Asi, hay medidas que son muy "sensibles a los marginales", en el sentido que su valor está muy influenciado por la distribución de los totales marginales de filas y columnas. Tales medidas reflejan por tanto información sobre los marginales, además de sobre la asociación.

Una medida de asociación determinada puede tener un valor bajo para una tabla dada, pero por ello no significa que las variables objeto de estudio no estén relacionadas, sino que ellas no están relacionadas en la forma a la que es sensible dicha medida. Por ello ninguna medida individual es la mejor para todas las situaciones. Al seleccionar una, debe tenerse en cuenta el tipo de datos, la hipótesis de interés así como las propiedades de cada medida. A veces se calcula un gran número de medidas y luego se referencian aquellas que más "ayudan a respaldar" las hipótesis del investigador, como si fueran las únicas que se hubieran calculado; pero esto no es correcto a menos que haya un verdadero análisis de cuáles medidas son las que se necesita observar.

Se verán primero algunas medidas de asociación para variables discretas en dependencia del tipo de datos a tratar.

Medidas nominales

Se consideran aquí medidas que asumen apenas que las dos variables de la tabla están medidas nominalmente. Como tales, estas medidas pueden suministrar solamente alguna indicación sobre la estrechez de la asociación y no pueden indicar casi nada sobre la dirección o cualquier otra cosa de la naturaleza de la relación. Las medidas se pueden clasificar en dos tipos: aquellas que están basadas en el Chi-cuadrado y aquellas que se fundamentan en la lógica de reducción proporcional del error (PRE, Proportional Reduction Error, como se le dice clásicamente).

Medidas nominales basadas en Chi-cuadrado.

Como se explicó antes, el test Chi-cuadrado en sí, no proporciona una buena medida del grado de asociación entre las dos variables; pero como está tan expandido el uso del Chi-cuadrado en la dócima de independencia, se ha estimulado la definición de medidas de asociación basadas en el Chi-cuadrado tratando de minimizar la influencia del volumen de la muestra y de los grados de libertad, así como restringir el rango de los valores de la medida al intervalo [0,1].

Estas medidas ayudan entonces a comparar los resultados del Chi-cuadrado en tablas diferentes cuando hay variación de las dimensiones y de los volúmenes de las muestras. Sin estas correcciones, es absolutamente inadmisible comparar con el Chi-cuadrado tales tablas.

El llamado coeficiente Phi, modifica el Chi-cuadrado dividiéndolo por el volumen de la muestra y extrayendo la raíz cuadrada del resultado:

$$\Phi(x_i) = \sqrt{\frac{\chi^2(x_i)}{n}} \tag{2.12}$$

Para tablas en las cuales una dimensión es mayor que 2, Phi no yace necesariamente entre 0 y 1, ya que el valor del Chi-cuadrado puede ser más grande que el volumen de la muestra. Por tanto Phi sólo queda estandarizado en el intervalo [0,1] en tablas en las cuales R=2 o C=2 (R y C denotan siempre el número de filas y columnas).

Para obtener una medida que debe hacer siempre entre 0 y 1 Pearson[ZAM97] sugirió el uso de:

$$C(x_i) = \sqrt{\frac{\chi^2(x_i)}{\chi^2(x_i) + n}} \tag{2.13}$$

que es conocido como coeficiente de contingencia. El valor de este coeficiente está siempre entre 0 y 1; pero generalmente no puede llegar a ser 1. De hecho, el máximo

valor de C depende sobre el número de filas y columnas. Por ejemplo para una tabla 4x4, el máximo valor de C es 0.87.
Cuanto mayor sea el valor de $C(x_i)$, mayor será el grado de asociación , o lo que es lo mismo cuanto menor sea el valor de $C(x_i)$, mayor será el grado de asociación.
Ahora se define la siguinte magnitud

$$p(x_i) = 1 - C(x_i)$$

(2.14)

Si el valor $p(x_i) \leq \alpha$ (donde α es un nivel de significación prefijado de antemano, normalmente 0.05) la relación es estadísticamente significativa. El mejor predictor es aquel que tenga el valor más pequeño de $p(x_i)$ que sea menor o igual que alfa (α).

Medidas Ordinales

Lógicamente, las variables con nivel de medición ordinal, pueden ser analizadas en tablas de contingencia con las medidas de asociación descritas anteriormente para variables simplemente nominales; pero otras medidas reflejan información adicional sobre la base del ranqueo.

Al considerar la clase de relaciones que pueden existir entre dos variables ordinales, se arriba a la noción de dirección de la relación y al concepto de "correlación generalizada". Las variables están positivamente correlacionadas cuando los casos con valores bajos de una de las variables tienen también valores bajos de la otra y los casos con valores altos de una de las variables los tiene también altos en la otra. Dicho en otras palabras, la correlación es positiva cuando la "relación de dependencia tiende estadísticamente a ser creciente". Las variables correlacionadas negativamente muestran el comportamiento inverso: a más altos valores de la primera variable corresponden los más bajos valores de la segunda.

Para tablas con dos variables ordinales hay varias medidas de asociación que se basan en la comparación de los valores de ambas variables en todos los pares posibles de casos u observaciones. Un par de casos se dice "concordante" si los valores de ambas variables para un caso son más altos (o ambos más bajos) que los valores para el otro caso. El par de casos se dice "discordante" si el valor de una variable para un caso es mayor que el correspondiente valor para el otro caso y la dirección está invertida en la segunda variable. Cuando los casos tienen valores idénticos sobre una o las dos variables se dicen "ligados". De esta forma, para cualquier par de casos con medidas de las variables X e Y, el par puede ser concordante o discordante o ligado, esto último en tres formas posibles: ligado respecto a X pero no de Y, ligado respecto a Y pero no de X y ligado respecto a X e Y.

II.3.1.1 Creación del Árbol de Decisión a partir del Nivel de Significación de las Variables

Sea K el conjunto de rasgos predictores x_i con nivel de significación $p(x_i) \leq \alpha$ (el valor de α es el valor mínimo que puede tomar el nivel de significación y lo asigna el experto). Luego en este conjunto se encuentran los rasgos predictores que están más asociados al rasgo objetivo.

$$K = \{x_i : p(x_i) < \alpha\}$$

(2.15)

El árbol de decisión se construye siguiendo el siguiente algoritmo:

Algoritmo para la Construcción del Árbol de Decisión

A1: Situar en la raíz del árbol los casos agrupados por la decisión.

A2: $K' = K$

A3: Buscar en el conjunto K' el rasgo predictor que esté más asociado con los valores del rasgo objetivo y_1 y expandir la raíz del árbol por los valores asociados a dicho rasgo predictor.

A4: Eliminar el rasgo seleccionado de K'

A5: Repetir mientras $K' \neq \phi$.

i) Dentro del conjunto K' elegir el conjunto de los rasgos predictores que estén más asociados con el rasgo predictor que se encuentra en las hojas. Reducir K' a ese conjunto.

ii) Si $K' = \phi$ terminar. En caso contrario elegir dentro de K' el rasgo predictor más asociado con el rasgo predictor que se encuentra en las hojas (determinando el grado de asociación como se mostró anteriormente) y expandir cada hoja del árbol por los valores del rasgo predictor seleccionado.

iii) Eliminar el rasgo seleccionado de K'

Por la forma en que se construye el árbol se deriva que en la raíz del árbol se encuentran todos los casos agrupados por el valor de la decisión. Estas agrupaciones, según las técnicas de los conjuntos rugosos, constituyen los *conjuntos conceptos*. Por otra parte si en el nivel $l = l_i$ aparece el rasgo x_i, entonces en este nivel el número de ramas será η_i que es la cantidad de valores que toma dicho rasgo. La rama correspondiente al valor x_{ij} será la rama $b = b_j$. Cada rama conduce a un nodo

donde se encuentran agrupados por el valor del rasgo correspondiente los casos derivados del nivel anterior. Dichos nodos, también para las técnicas de los conjuntos rugosos, constituyen los llamados *conjuntos indiscernibles*.

II.3.1.2 Cálculo de la primera probabilidad.

La probabilidad $P_1(x_i)$ la calcularemos a partir de la frecuencia $O_{pq}(x_i)$ que aparece en las tablas de contingencia asociadas a las variables a medida que realizamos el recorrido en profundidad del árbol de decisión, para ello construimos la ley de probabilidades teniendo en cuenta la dependencia de los valores de los rasgos predictores con el rasgo objetivo y el resto de los rasgos predictores.

Así denotaremos por $P_1(x_i)$ las probabilidades correspondientes al rasgo x_i, dependiendo del comportamiento de ese rasgo con el rasgo objetivo y el resto de los rasgos predictores. Ella es calculada mediante:

$$P_1\left(x_i(O_t) = x_{ij}\right) = \sum_{i=1}^{k} P\left(x_i(O_i) / x_j(O_i)\right) \cdot P\left(y_j(O_i)\right)$$

(2.16)

Caso 1: Cálculo de la primera probabilidad para los valores de los rasgos predictores independientes (que no aparecen en el árbol de decisión).

En el caso que x_i constituya un rasgo independiente, su distribución de probabilidades la construimos a partir de las distribuciones de frecuencias del rasgo x_i en los casos de la tabla original ¡independiente del resto de los rasgos, mediante la distribución de frecuencias y la expresión que usaremos es:

$$P_1\left(x_i\left(O_t\right)=x_{ij}\right)=\frac{\left|x_{ij}\left(O_t\right)\right|}{\left|n\right|}$$

(2.17)

Por ello se toman los criterios:

Caso 2: Cálculo de la primera probabilidad para los valores de los rasgos predictores dependientes (que aparecen en el árbol de decisión).

En el casos que los rasgos sean dependientes procedemos de la siguiente forma:

- Determinar las dependencias entre las variables predictoras y el rasgo objetivo y el resto de los rasgos predictores, con el valor del nivel de significación de cada rasgo se crea el árbol de decisiones.
    - Una vez establecido el primer nodo se expande éste teniendo en cuenta cual es el rasgo que con él se enlaza.
    - Se repite hasta llegar a la hoja o rasgo menos predictor, que será el de menos peso.
- Calculamos los diferentes valores de las probabilidades $P(x_i)$ asociadas a cada rasgo x_i dividiendo el valor de la frecuencia que aparece en cada celda entre el total marginal de la columna que contiene a dicha celda en la tabla de contingencia.
- Los valores de la probabilidad calculada se almacenan en la tabla o en la propia base de forma tal que pueda ser utilizada más tarde en el cálculo de la Certidumbre total asociada al rasgo.

Las Tablas de Contingencia y el árbol de decisiones permiten también inferir directamente la además de la primera probabilidad otros datos que aparecen en ellas como por ejemplo los % de casos distribuidos en las filas y las columnas. Estas tablas tomadas por columna dan la probabilidad de dependencia del rasgo predictor respecto al objetivo y tomadas por fila dan la probabilidad de dependencia del rasgo objetivo respecto al predictor.

Caso 3: Cálculo de la primera probabilidad para los valores del rasgo objetivo (decisión).

Para los valores que no se reflejan en las relaciones del árbol de dependencia y por ende no aparecen en las tablas de contingencia, la primera probabilidad se calcula de forma parecida a como se realiza con los rasgos independientes, pero ahora la fórmula 2.18 toma la forma:

$$P_1\left(y_i\left(O_t\right)=y_{ij}\right)=\frac{\left|y_{ij}\left(O_t\right)\right|}{\left|n\right|}$$

(2.18)

Es decir, usamos la frecuencia con que el rasgo toma un valor determinado de categoría se divide ese valor de frecuencia entre el número total de frecuencias para todas las categorías de ese rasgo (total marginal de la fila correspondiente donde se encuentra la celda), pues en este caso el rasgo independiente y el rasgo objetivo no estarán directamente correlacionados y si existen casos que no se les ha podido determinar el valor de esta primera probabilidad para el rasgo objetivo se procede de manera similar al caso del rasgo independiente, descrito anteriormente pero con la fórmula 2.18.

II.3.2 Cálculo de la segunda probabilidad para los rasgos predictores y objetivos.

Como se sabe que el cálculo de la primera probabilidad condicionada o no de los rasgos no constituye una medida de la certidumbre del valor ya que una fuente importante de incertidumbre es la forma en que se obtienen dichos valores. Es por ello que para determinar la certidumbre total se debe calcular otra probabilidad P_2, que para el caso de los rasgos predictores está muy relacionada con la forma en que se obtienen sus valores. En el caso de los rasgos objetivos usaremos algunas ideas de la teorías probabilísticas adaptadas a nuestro caso.

Determinación de P2 para los rasgos predictores.

Como se ha hecho notar, el término incertidumbre o razonamiento inexacto se emplea en la IA para referirse a un amplio rango de situaciones donde la información relevante es deficiente por diversas causas, atendiendo a ello se consideran los casos siguientes:

Caso 1: El valor de la variable es tomado de una fuente en la cual no hay motivos para no confiar.

$$P_2\left(x_i\left(O_t\right) = x_{ij}\right) = 1$$

(2.19)

Caso 2: La variable toma el valor a partir del criterio de varios expertos.

Consideremos la existencia de n expertos: $a_1, a_2, ..., a_n$, asignando el valor del rasgo x_i del caso O_t. Dado un valor particular x_{ik} de x_i consideremos además que $a_r\left(x_{ik}\right)$ denota la respuesta del experto *r* sobre ese valor. Asumamos que $a_r\left(X_{ik}\right) = 1$ cuando el valor se afirma y $a_r\left(x_{ik}\right) = 0$ cuando el valor se niega. Entonces

$$P_2\left(X_i\left(O_t\right) = x_{ij}\right) = \frac{\sum_{r=1}^{n} c_r \cdot a_r\left(x_{ij}\right)}{\sum_{r=1}^{n} c_r}$$

(2.20)

donde c_r es el grado de competitividad del experto.

Caso 3: La variable toma el valor a partir de un instrumento de medición.

$$P_2\left(x_i\left(O_t\right) = x_{ij}\right) = P$$

(2.21)

donde *P* es la certidumbre asociada al instrumento de medición utilizado y constituye una probabilidad por si sola.

Determinación de la segunda probabilidad, P_2 para los rasgos objetivos.
Dada la base de casos descrita en el inicio del presente Capítulo, que puede incluso ser representada a través de una tabla donde se encabezan las columnas con los distintos rasgos predictores seguidos cada uno de ellos por las probabilidades calculadas y las filas con los casos de la base, se calculan las probabilidades para cada caso al rasgo objetivo, sacando un promedio de las segundas probabilidades de todos los rasgos predictores según la expresión:

$$P_2(y_1(O_t) = y_{1k}) = \frac{\sum_{i=1}^{n} P_2(x_i(O_t) = x_{ij})}{n}$$

(2.22)
y con ello queda lista para el cálculo de la Certidumbre de los rasgos.

Cálculo de la Certidumbre Total
Si en una base de casos se calcula una primera probabilidad condicionada por la relación entre los rasgos predictores al objetivo y se calcula una segunda con la información proveniente de diferentes fuentes o apoyados en la propia teoría probabilística se puede calcular la Certidumbre total, como una función de ambas de manera que la base quede mejor definida y lista para la recuperación y adaptación ante cualquier aparición de un nuevo caso a solucionar.
Se calcula la certidumbre total de cada valor para cada caso usando la expresión: $\xi_i(O_t) = g(P_1, P_2)$ como ya se había explicado en la parte de la fundamentación**,** donde $\xi_i(O_t)$ es la certidumbre del valor del rasgo *i* para el caso O_t y g puede ser cualquiera de las funciones de combinación que han sido utilizadas en la literatura para propagar incertidumbre y que se conocen como t-conormas arquimedeanas por ejemplo:

$$g(x, y) = x + y - xy$$

(2.23)

$$g(x, y) = \frac{x + y}{1 - xy}$$

(2.24)

entre otras.
Estas funciones de combinación cumplen las siguientes propiedades:

1. Si $0 < x, y < 1$ entonces $0 < g(x, y) < 1$
2. $g(x, y) = g(y, x)$
3. $g(x, g(y, z)) = g(g(x, y), z)$

4. $g(x,0) = x$

5. Si $x \leq y$ entonces $g(x,z) \leq g(y,z)$

6. $g(x,1) = 1$

Esta última propiedad resulta importante ya que garantiza la seguridad absoluta cuando cualquiera de las dos probabilidades asignadas es 1.

Recuperación de un nuevo caso

Para el nuevo caso esta certidumbre se calcula teniendo en cuenta sólo P_2.

Una vez calculada la certidumbre de todos los valores para los rasgos predictores y objetivos y la certidumbre de los valores del nuevo caso, se procede a realizar los procesos de razonamiento (recuperación y adaptación).

El módulo de recuperación consta de dos etapas:

- Etapa de acceso.
- Etapa de recuperación.

En la etapa de acceso haremos uso de la estructura jerárquica de representación de los casos. Esta estructura nos permitirá reducir el número de casos a comparar pues sólo consideraremos aquellos casos que para los rasgos mas predictores (los que están en el árbol) tengan igual valor que en el problema planteado. Esto no significa que vayamos a considerar todos los rasgos predictores que están en el árbol pues si todos son considerados pudieran no tenerse en consideración casos que difieran en pocos rasgos significativos. Por eso hay que definir hasta que nivel del árbol vamos a considerar. El siguiente algoritmo muestra este proceso:

Algoritmo para el Acceso a los casos más relevantes

Este algoritmo está basado en el uso de una heurística de preferencia conocida como preferencia de características sobresalientes [Kol89]

Algoritmo de Acceso a los Casos más Relevantes

Entrada: Problema a resolver: O_0

Nivel a considerar: level

Conjunto de rasgos más predictores: K

Salida: Conjunto de casos potenciales para la recuperación: P

A1: $P = \phi$

A2: Para *l=1, level*:

i) Seleccionar el rasgo más que se encuentra en el nivel l. Sea este x_i

ii) Si $x_i(O_0) = x_{ip}$ entonces $P = P \cap X_{ip}$

De esta forma en el conjunto P nos quedan los casos potenciales para la etapa de recuperación, en la cual se determina la semejanza entre ellos y el nuevo problema. En este proceso se tienen en consideración no solo los más predictores, sino todos los rasgos iniciales y la certidumbre de los valores que ellos toman. A la primera función de semejanza mostrada anteriormente se le realiza una transformación a través del siguiente algoritmo.

Algoritmo para la Transformación de la Función de Semejanza en el Módulo de Recuperación

Entrada: P, O_0 nc (número de casos más semejantes a recuperar)

Salida: S Conjunto de casos más semejantes

A1: Para todo $O_t \in P$ determinar

$$\beta(O_0, O_t) = \frac{\sum_{i=1}^{n} p_i' \cdot m_i(O_0, O_t) \cdot \delta_i(O_0, O_t)}{\sum_{i=1}^{n} p_i'}$$

(2.2.5)

$$m_i(O_0, O_t) = \begin{cases} \dfrac{\xi_i(O_0) + \xi_i(O_t)}{2} & si \ \ (\xi_i(O_0) \wedge \xi_i(O_t)) \geq \varepsilon \\ 0 & e.o.c \end{cases}$$

(2.26)

$$p_i' = \begin{cases} 1 - p(x_i) & \text{si el peso no ha sido asignado con anterioridad} \\ p_i - p(x_i) & e.o.c. \end{cases}$$

(2.27)

donde:

n: Número de rasgos predictores.

p_i': Peso o relevancia del rasgo *i*.

A2: Agrupar los *nc* casos más semejantes en el conjunto *S*.

Como se puede observar en la fórmula anterior si para el problema o para el caso con que estamos comparando, el valor del rasgo que se está considerando está por debajo de un umbral predeterminado, dicho rasgo no se toma en consideración ya que estaríamos comparando valores cuyo nivel de certidumbre está por debajo de lo permisible.

II.3.1 Optimización de la adaptación.

Dado un determinado caso la adaptación se efectúa determinando el caso óptimo de los casos recuperados. El criterio de optimalidad se da a través de una función que combina el valor de la semejanza para dicho caso y la certidumbre de su solución. Esto se realiza a través de la fórmula:

$$\mu(O_t) = \alpha\beta(O_0, O_t) + (1-\alpha)\eta_1(O_t)$$

(2.28)

donde α es un parámetro que se selecciona según el criterio del experto. A medida que α tiende a 1 significa que se le está dando mayor importancia a la semejanza que a la certidumbre de la solución.

Capítulo II

Análisis del lugar de la incertidumbre en los Sistemas Basados en Conocimiento.

En los Sistemas Basados en Conocimiento que consideran la existencia de incertidumbre, la que puede aparecer tanto en la formación del propio conocimiento como en los datos que aparecen al formular un problema concreto, el enfoque de trabajo usual ha sido realizar el proceso de razonamiento sin considerar la incertidumbre existente y paralelamente realizar un proceso de cálculo de la incertidumbre asociada con la conclusión. En este proceso la combinación de incertidumbres provenientes de diversas fuentes se hace necesaria al menos en dos niveles. En el primero se deriva una conclusión utilizando una pieza del conocimiento almacenado en la base de conocimiento y es necesario combinar los diferentes valores de las incertidumbres asociadas con las hipótesis en las que se basa esa pieza de conocimiento y la incertidumbre de la propia pieza de conocimiento para determinar la incertidumbre de la conclusión alcanzada. En el segundo nivel es necesario combinar diversas conclusiones (cada una con su propia incertidumbre) para alcanzar la respuesta. A este proceso se le denomina propagación de la incertidumbre desde los datos a la respuesta.
Las técnicas anteriormente expuestas en el Capítulo I han sido usadas en [Paw95a, Paw95b] y otros Sistemas Expertos. A continuación se describe el uso de una de ellas en un Sistema Basado en Casos.

II- 1 Sistemas Basados en Casos. Incertidumbre

Definición
Los Sistemas Basados en Casos son Sistemas Basados en el Conocimiento que utilizan, como método de solución de problemas, el Razonamiento Basado en Casos. El Razonamiento Basado en Casos es un paradigma de solución de problemas que difiere de otros enfoques y técnicas en que es capaz de utilizar el conocimiento específico adquirido en situaciones previas y utilizarlo en la situación presente. Un problema nuevo se resuelve buscando en su memoria un caso similar resuelto en el pasado. Además incrementa su conocimiento almacenando el nuevo caso para ser usado en situaciones futuras. Esto permite que su conocimiento se mantenga actualizado en todo momento [Gar92].
Componentes de un Sistema Basado en Casos
Los componentes principales de un Sistemas Basados en Casos son:
La Base de Conocimiento o Base de Casos:
Es la que contiene los casos con las descripciones de los problemas resueltos y también los no resueltos previamente; estos últimos son aquellos problemas a los que

se les dio una “solución” que posteriormente mostró ser errónea y son útiles para prevenir errores en el futuro. Si luego de la ejecución ocurrió algún error en la predicción contienen además la explicación de las anomalías, las estrategias de reparación y referencias al próximo resultado.

En el procedimiento de recuperación de casos :

En él se determinan los casos que mayor similitud guardan con el patrón de búsqueda. Esto puede realizarse por semejanza parcial o por analogía. En el primer método, se emplea una función de semejanza la cual determina una medida numérica del grado de similitud del objeto de la memoria con respecto al nuevo problema. En el segundo, se busca un elemento de la memoria cuya descripción pueda hacerse igual a la del nuevo problema, mediante el reemplazamiento de los valores de los rasgos diferentes a partir de la equivalencia de valores, definida en una red semántica.

El Módulo o Procedimiento de Adaptación de las Soluciones:

Es el que se encarga de encontrar los valores de los rasgos desconocidos a partir de los casos recuperados mediante un algoritmo capaz de adaptar las soluciones de los casos semejantes para completar el nuevo problema. Atendiendo a dicho módulo de adaptación se distinguen los dos tipos de sistemas basados en casos: los interpretativos y los solucionadores de problemas. Un sistema Basado en Casos interpretativo toma un problema y su solución y utiliza los casos para proponer un argumento para justificar esa solución. Un solucionador de problemas deriva soluciones al nuevo problema a partir de soluciones viejas a problemas similares.

II-1.1 Etapas de Razonamiento Basado en Casos

Asignación de índices.

El sistema recibe el caso de entrada y analiza sus características y atributos más relevantes (índices) para poder acceder a los casos de mayor similitud en base de casos. A los índices usados para medir la similitud entre los casos se les denomina rasgos predictores, mientras que a los índices de los cuales deseamos obtener su valor se les denomina rasgos objetivos.

Acceso a casos similares.

El siguiente paso es el acceso a casos similares en la base de casos a través del procedimiento de recuperación. Esto se realiza usando los rasgos predictores encontrados en el paso anterior. De esta forma se obtienen los casos más similares al actual. En muchas ocasiones es necesario reducir aun más el conjunto de casos para obtener los más relevantes.

Creación de la solución.

Una vez seleccionados los casos similares, se efectúa el procedimiento de adaptación, que consiste en la modificación y combinación de las soluciones de los casos similares para formar una nueva solución, una interpretación, o una explicación dependiendo de la tarea que lleve a cabo el sistema.

Evaluación de la solución.

La solución obtenida en la fase anterior, representa una solución potencial o tentativa. Es necesario hacer una evaluación de la solución propuesta antes de darla al usuario, es decir, se intenta ver cuáles son las cualidades y debilidades de la solución para evaluar su utilidad.

Asignación y almacenamiento del caso nuevo.

Una vez creada la solución y evaluada, se procede a entregar la solución al mundo real. Además, se puede crear un nuevo caso. Este caso se forma no sólo de la solución creada, sino del caso o problema original, la solución encontrada, las justificaciones y explicaciones (si las hay). El caso nuevo se le asignan índices y se almacena en la base de casos.

Explicación, reparación y prueba.

Si la solución fracasa, es importante para el sistema obtener y analizar información de retroalimentación para no repetir los mismos errores. Si ocurre algo inesperado el sistema debe tratar de explicarlo. Después se trata de reparar la solución con base en esta explicación y se regresa al paso de evaluación.

Más información acerca del tema se puede encontrar en [GAR94, KOL93].

II.1.2 Ventajas y Desventajas del Razonamiento Basado en Casos sobre otras tecnologías.

El RBC como nuevo método de solución de problemas para el desarrollo de sistemas basados en conocimiento presenta de un determinado número de ventajas entre las que se pueden citar:

Ventajas.

- La adquisición del conocimiento. La unidad básica del conocimiento es el caso y no la regla. Los seres humanos por lo general articulan su conocimiento mediante ejemplos de problemas y soluciones anteriores, más que por medio de reglas específicas y abstractas.
- Permite proponer soluciones a problemas rápidamente. Esto lo logra evitando el tiempo necesario para derivar estas respuestas a partir de cero.
- Aprendizaje. A diferencia de los sistemas basados en reglas convencionales, si la misma situación se presenta repetidamente, no se tiene que construir o generar la misma solución a partir de cero cada vez que se presente el mismo problema.
- Propone soluciones en dominios no entendidos completamente por el sistema.
- Ofrece un medio de evaluación de soluciones cuando no se cuenta con un método algorítmico de evaluación.
- Se centra en las características o partes más importantes de un problema.

Desventajas.

- Confía ciegamente en los casos previos almacenados en su memoria para intentar proponer su solución.
- Puede ser que no recupere el caso más apropiado para la solución del nuevo caso.

Estas desventajas podrían eliminarse con la introducción de un mecanismo de manejo de la incertidumbre en estos sistemas.
El desarrollo de este mecanismo constituye el objetivo fundamental del presente trabajo y a el le dedicaremos los próximos epígrafes.

II.1.3 Modelo General de un Sistema Basado en Casos sin Manejo de Incertidumbre.

Descripción de un caso.

Un caso de la base es descrito a partir de los valores que se le asignan a los rasgos predictores y los rasgos objetivos. Los rasgos predictores son los que determinan los valores de los rasgos objetivos. Un caso O_t que tenga *n* rasgos predictores y *p* rasgos objetivos se describe del siguiente modo:

$$O_t\left(x_1(O_t),...,x_i(O_t),...,x_n(O_t), y_1(O_t),...,y_j(O_t),...,y_p(O_t)\right)$$

(2.1)

donde: x_i, y_i : rasgo predictor y objetivo respectivamente.

$x_i(O_t)$: valor del rasgo predictor $i = 1, n$ para el caso O_t t =1, m de la base

$y_j(O_t)$: valor del rasgo objetivo $j = 1, p$ para el caso O_t t =1, m de la base.

A los valores que toman cada rasgo se les llaman categorías y de forma general son en número igual o mayor a 2 y es precisamente a través de esas categorías que se enlazan los rasgos predictores con el o los rasgos objetivos.
Una base de casos es, como se dijo, una colección de casos que puede ser representada a través de una tabla de decisión, en la cual, las columnas son etiquetadas por variables que representan los rasgos predictores y objetivos (decisiones) y las filas representan los casos.
La Tabla 2.1 muestra un ejemplo de una base de casos. Aquí los casos se representan por $O_1, ..., O_m$, los rasgos predictores por: $x_1, x_2, ..., x_n$ y la decisión por y_1. El valor del rasgo x_i en el caso O_t se representa por $x_i(O_t)$ y el valor de la decisión en dicho caso se representa por $y_1(O_t)$.

Caso	x_1	**…**	x_n	y_1	y_2	**…**	y_p
O_1	$x_1(O_1)$	…	$x_n(O_1)$	$y_1(O_1)$	$y_2(O_1)$	…	$y_2(O_1)$
…	…	…	…			…	

O_m	$x_1(O_m)$	...	$x_n(O_m)$	$y_1(O_m)$	$y_p(O_m)$	...	$y_p(O_n)$

Tabla 2.1. Tabla de decisión que representa una base de casos.

En lo adelante la base de casos que representemos usando esta definición le nombraremos Base de casos original y en la medida que se calculen otros parámetros con los datos almacenados en ella se irá completando hasta formar la base de casos que usaremos en el módulo de recuperación.

Módulo de Recuperación

En el módulo de recuperación es donde se realiza el acceso a casos similares en la base a través del procedimiento de recuperación, haciendo uso de los rasgos predictores a que hicimos referencia anteriormente. De esta forma se obtienen los casos más similares al problema actual. En muchos problemas prácticos es necesario reducir el conjunto de casos para obtener los más relevantes. En el presente trabajo se propone una vía alternativa.

Los dos aspectos claves de esta fase son el algoritmo de acceso a casos y la medida de similitud entre casos. Para determinar que tan similar es un caso a otro se han desarrollado varias técnicas. La más sencilla consiste en contar el número de características similares entre los dos casos. El problema de esta técnica es que la importancia de las características varia de un contexto a otro. Otra técnica consiste en utilizar un conjunto de heurísticas que permitan determinar cuales características tienen mayor relevancia (peso) y formular una función de semejanza que involucre la similitud entre cada uno de los rasgos teniendo en cuenta el peso de los mismos. Un modelo matemático de esta última técnica es el siguiente:

Función de semejanza entre un nuevo problema a resolver y un caso O_t de la base

$$\beta(O_0, O_t) = \frac{\sum_{i=1}^{n} p_i \cdot \delta_i(O_0, O_t)}{\sum_{i=1}^{n} p_i} \qquad (2.2)$$

donde:

$O_0 = O_0(x_1(O_0),...,x_n(O_0))$: Problema a resolver

n : Número de rasgos predictores.

p_i : Peso o relevancia del rasgo *i*.

$\delta_i(O_0, O_t)$: Función de comparación entre los casos O_0 y O_t atendiendo al rasgo *i*. Esta función puede estar definida de diferentes formas, por ejemplo:

Para rasgos numéricos.

$$\partial_i(O_0, O_t) = \begin{cases} 1 & si \quad \mathrm{x}_i(O_0) = x_i(O_t) \\ 0 & e.o.c \end{cases} \tag{2.3}$$

$$\partial_i(O_0, O_t) = \begin{cases} 1 & si \quad |x_i(O_0) - x_i(O_t)| < \varepsilon \\ 0 & e.o.c \end{cases} \tag{2.4}$$

ε : Valor umbral definido por el usuario.

$$\partial_i(O, O_t) = 1 - \frac{|x_i(O) - x_i(O_t)|}{\max_i - \min_i} \tag{2.5}$$

$\max_i, \min_i$: Valores máximo y mínimo respectivamente que alcanza el rasgo i.

Por otro lado, el algoritmo de acceso a los casos debe ser rápido y eficiente. Este depende de las técnicas de indexación usadas y su diseño se vuelve un aspecto crítico cuando la base de casos es muy grande.

Entre las técnicas de acceso a casos podemos mencionar:

Acceso exacto. Se busca el caso que encaje exactamente con el caso nuevo.

Acceso Jerárquico. Básicamente se buscan los casos en un árbol, tomando decisiones en cada nodo hasta que ya no se pueda avanzar. Si se encuentra una hoja se muestran los casos correspondientes, si se llegó hasta un nodo, se muestra todos los casos que se deriven de él. Esta última técnica es la que se usa en nuestro modelo.

Módulo de Adaptación

Una vez seleccionados los casos similares, se efectúa el procedimiento de adaptación, que consiste en la modificación y combinación de las soluciones de los casos similares para formar una nueva solución, una interpretación, o una explicación dependiendo de la tarea que lleve a cabo el sistema. Es deseable que el sistema también genere justificaciones o explicaciones que apoyen a la solución creada para el nuevo caso.

Al acceder a los casos similares, lo mejor que puede pasar es encontrar un caso exactamente igual al nuevo, pero generalmente no sucede así, y es necesario realizar el proceso de adaptación. La adaptación puede realizarse por el usuario o de manera automática por el sistema.

Si la adaptación la realiza el usuario, el sistema sólo realiza la búsqueda de los casos similares.

Si la adaptación la ejecuta el sistema, entonces debe contener algún conocimiento, tal como fórmulas o reglas. Las técnicas desarrolladas para la realización de la

adaptación en forma automática han dado en general buenos resultados. En [Kol93] se mencionan los siguientes métodos de adaptación:
Reinstanciación. En este tipo de técnica se utiliza el marco o contexto de situaciones anteriores, pero con nuevo argumentos.
Ajuste de parámetros. Se ajustan los parámetros de la solución de casos anteriores de acuerdo con las diferencias entre las descripciones de los casos en cuestión.
Búsqueda local. Se realizan búsquedas dentro de jerarquías semánticas y se soluciona el problema por analogía.
En el presente trabajo se propone un nuevo enfoque donde se hace uso de la certidumbre de la solución.

II .2 Solución a la problemática de los Sistemas Basados en Casos.

Fundamentación.

Si en una base de casos se pueden definir los rasgos o variables predictores y los rasgos objetivos, los primeros pueden influir de diferente forma en los valores que toman los segundos, es decir, dentro de los rasgos que influyen sobre el objetivo (llamados también variables dependientes) existen unos que están más fuertemente correlacionados con él que otros, estos son los rasgos más predictores otros, en ocasiones, pueden correlacionarse con el objetivo solamente a través de otro u otros rasgos, serán de este modo considerados como menos predictores.

II.2.1 Incertidumbre en los Sistemas Basados en Casos

La información incompleta es omnipresente: casi toda la información que se tiene del mundo es no cierta, no completa o no precisa. La incertidumbre cala nuestra compresión del mundo real.

El propósito de los sistemas de información es modelar el mundo real. Por ello dichos sistemas tienen que ser capaces de modelar la incertidumbre.

A nivel empírico, la incertidumbre es casi una medida inseparable de casi cualquier media, resultante de una combinación de inevitables errores de medición y límites de resolución de los instrumentos de medición utilizados.

A nivel cognitivo, esta emerge de la vaguedad y ambigüedad inherente al lenguaje natural. Se ha hecho claro que hay distintos tipos de incertidumbre. La incertidumbre es un concepto multidimensional. Las teorías matemáticas clásicas para caracterizar situaciones bajo incertidumbre han sido la teoría de conjuntos y la teoría de las probabilidades [Kli95, Kli98].
El término incertidumbre se utiliza de forma general para abarcar diferentes tipos de incertidumbre: imprecisión, incompletes, vaguedad e inconsistencia.

La imprecisión es relativa a la existencia de un valor el cual no puede ser medido con la precisión adecuada; *la incompletes* a la ausencia de valores para algunas variables consideradas, *la vaguedad* a la ausencia de un límite claro para un concepto o conjunto y *la inconsistencia* describe una situación en la cual hay dos o más valores en conflicto para ser asignados a una variable [Bon85, Bos93, Par96].

Obviamente el Razonamiento Basado en Casos no escapa de la problemática de la incertidumbre. Hay posibilidades de imprecisión tanto en el conocimiento (base de casos) como en el planteamiento del problema.

El conocimiento del problema radica esencialmente en la colección de casos, luego la primera interrogante es si el caso representa o no el objetivo para el cual fue puesto en la base, es decir ¿cuál es el grado de certeza de que la solución contenida en el caso sea válida?. Otro aspecto referido con el conocimiento del sistema es el modelo de organización de los casos. Un buen modelo de organización debe permitir almacenar los casos quizás no como un todo, y luego recuperar los casos más semejantes reconstruyendo los mismos, aquí la fuente de incertidumbre esta dada por la medida en que el modelo permite descomponer y recomponer los casos.

Los resultados planteados en este trabajo están orientados a resolver el primer aspecto, es decir determinar la certidumbre de los valores representados en la base y una vez determinada, manipularla en los módulos de recuperación y adaptación de forma tal que se solucionen las siguientes interrogantes:

4. Supongamos que estamos trabajando con tres rasgos predictores: *FORMA, TAMAÑO* y *COLOR* y que por algún método dado, se ha determinado que el rasgo *FORMA* es el rasgo más importante y por tanto el que mayor peso tiene. Sin embargo, los valores que ha tomado ese rasgo tienen muy bajo nivel de certidumbre, surge entonces la próxima interrogante ¿es correcto asignar un peso grande a un rasgo del cual estamos tan inseguros?
5. Considérese el rasgo *COLOR* con los valores *{AZUL, ROJO, VERDE, NEGRO}*, tener en cuenta la incertidumbre, significa que el rasgo *COLOR* toma un valor en el caso almacenado y otro en la descripción del problema a resolver, pero sobre ambos valores hay dudas sobre su certeza. Por ejemplo*: COLOR=AZUL (0.9)* en el caso *y COLOR=ROJO (0.6)* en la descripción del problema. La situación anterior nos obliga a pensar en dos aspectos. El primero es que ahora la función de comparación tendrá como argumentos valores sobre los que existen dudas y por lo tanto el valor de la función de comparación entre estos colores dadas las certidumbres 0.9 y 0.6 no tiene que ser igual que en el caso en que ambos valores de certidumbres sean 1. Considerar el valor de la certidumbre en los valores dados a los rasgos nos obliga a redefinir las funciones de comparación de rasgos definidas originalmente. De aquí surge la siguiente interrogante: ¿cómo construir

funciones de comparación de rasgos considerando incertidumbre en los valores dados a estos?.

6. Supongamos que los casos recuperados con mayor nivel de semejanza son O_5 y O_{15} (ambos con el mismo nivel de semejanza) y que en ambos casos el valor del rasgo objetivo es diferente (situación de conflicto). ¿Por qué caso adaptar?

Hemos señalado algunos momentos del razonamiento basado en casos donde resulta necesario tener en cuenta la manipulación de la incertidumbre. En el siguiente epígrafe se presentan algunas ideas preliminares sobre la elaboración de un modelo para la manipulación de la incertidumbre en los SCBC.

II.2.2 Modelo para la Manipulación de la Incertidumbre en una Base de Casos

Descripción del Problema

Se resuelve el problema de la toma de decisiones utilizando el razonamiento Basado en Casos en condiciones de incertidumbre. Desde la perspectiva del enfoque basado en casos los problemas de decisión se describen de la forma siguiente:

- Un conjunto de rasgos que describen la situación S (rasgos predictores).
- Un conjunto de posibles decisiones a tomar D (rasgos objetivos).
- La incertidumbre presente en los rasgos predictores.
- La incertidumbre presente en los rasgos objetivos.

Luego, los casos en la base se componen de un conjunto de rasgos predictores de la forma (valor, certidumbre) y uno o más rasgos objetivos de la forma (valor, certidumbre), donde los valores son discretos y de cardinalidad finita.
Un nuevo problema se describe en términos de los valores que toman los rasgos predictores en la nueva situación donde el agente tiene que tomar la decisión y el sistema debe ser capaz de inferir qué decisión debe tomar a partir de la experiencia acumulada en la base de casos. Luego el modelo a desarrollar tiene que:

4. Dar una vía para calcular la incertidumbre presente en los casos almacenados en la base de casos (pues usualmente la base de casos se crea a partir de fuentes donde no están explícitamente registrados los valores de incertidumbre). Es decir, debemos encontrar para cada valor $x_i(O_t)$ dado al rasgo predictor i (i=1,n) en el

caso O_t (t=1,m) una medida $\zeta_i(O_t)$ que denota la certidumbre de dicho valor y para cada decisión $y_j(O_t)$ dado al rasgo objetivo j (j=1,p) en el caso O_t (t=1,m), una medida $\eta_j(O_t)$.

5. Desarrollar un procedimiento de recuperación, para la selección de los casos más semejantes al nuevo problema a partir de las descripciones planteadas en el caso y la situación donde el agente debe tomar la nueva decisión, considerando que en la descripción de ambas situaciones hay incertidumbre.
6. Desarrollar un proceso de adaptación que a partir de las decisiones tomadas en los casos recuperados y la incertidumbre asociada determine cuál debe ser la decisión a tomar y con qué certidumbre se toma.

Estructura General del Modelo Desarrollado

El cálculo de la certidumbre asociada al valor de cada rasgo se calcula a partir de dos magnitudes. La primera de ellas tiene en cuenta la información almacenada en la base de casos y representa la certidumbre de que un rasgo tome un valor dado de su dominio conociéndose la decisión tomada. La segunda magnitud representa la certidumbre del valor conocida su fuente de obtención (exacta, criterios de expertos, instrumentos de medición) y por tanto puede variar de uno a otro caso. Ambas magnitudes son calculadas como probabilidades condicionadas y se les denominan *Primera* y *Segunda Probabilidad* respectivamente (para abreviar P_1 y P_2).

La probabilidad P_1 se calcula usando el árbol de decisión que describe la relación de dependencia entre los valores de las variables predictoras y objetivos. Para la construcción del árbol se calculan los niveles de asociación de las variables predictoras respecto a las variables objetivos, usando el Test Ji-Cuadrado [MAR91, CHA94]. De esta forma en el árbol sólo estarán presentes aquellos rasgos cuyo nivel de asociación este por encima de un valor predefinido por el experto. Estos son los rasgos que se consideran dependientes. El resto de los rasgos se consideran independientes entre sí y con respecto a los rasgos objetivos y para ellos la probabilidad P_1 de un valor se calcula como su frecuencia de aparición en la base. Esta idea conduce a la creación de un Modelo de dependencia con síntomas relevantes según [Cas89].

La creación del árbol permite simplificar el proceso de selección de los casos más semejantes, pues permite no tener que considerar todos los casos en el proceso de recuperación como se ilustra en el epígrafe II.3. Por otra parte los niveles de asociación calculados permiten realizar la asignación de los pesos o su modificación, en caso de que los pesos hayan sido previamente asignados.

Una vez calculadas ambas probabilidades para calcular la certidumbre total $\zeta_i(O_t)$ se combinan éstas una co-t-norma g.

$$\zeta_i(O_t) = g(P_1, P_2)$$

(2.6)

Para el nuevo caso esta certidumbre se calcula teniendo en cuenta sólo P_2.

Una vez calculada la certidumbre de todos los valores para los rasgos predictores y objetivos y la certidumbre de los valores del nuevo caso, se procede a realizar los procesos de razonamiento (recuperación y adaptación).

El procedimiento de recuperación se basa en una función de semejanza la cual compara la descripción del nuevo problema O_0 con las descripciones de los casos O_t almacenados en la base, considerando la importancia de cada rasgo o peso w_i, los valores de las comparaciones a nivel de rasgo ∂_i y las certidumbres de los valores comparados.

$$\beta(O_0, O_t) = f_1(w_i, \delta_i(O_0, O_t), m_i)$$

(2.7)

$$m_i = f_2(\zeta_i(O_0), \zeta_i(O_t))$$

(2.8)

Una vez seleccionados los casos más semejantes es posibles que estos propongan decisiones diferentes por lo que se hace necesario determinar cuál de las decisiones debe tomarse. Esta elección se realiza teniendo en cuenta el grado de semejanza calculado en el procedimiento anterior y a la certidumbre de la decisión de los casos recuperados.

$$\mu(\beta(O_0, O_t), \eta_j(O_t))$$

(2.9)

II.3 Modelo Matemático para la Determinación de la Certidumbre

Este modelo está enfocado a la existencia de un solo rasgo objetivo y_1, aunque pudiera ser extendido al caso en que existan múltiples decisiones.

II.3.1 Fundamentos Matemáticos para la Construcción del Árbol de Decisión

Sea y_1 la decisión (rasgo objetivo), la cual toma valores en conjunto: $\{y_{11}, \ldots y_{1\theta}\}$ donde θ es la cardinalidad del conjunto de valores que toma dicho rasgo objetivo.

Sean x_i (i=1,n) el conjunto de rasgos predictores, los cuales toman los valores en los conjuntos $\{x_{i1},...x_{i\eta_i}\}$ donde η_i es la cardinalidad del conjunto de valores que toma el rasgo predictor x_i.

Para cada variable predictora x_i determinamos la tabla de contigencia que la relaciona con la variable objetivo y_1 de la siguiente forma:

$x_i \diagup y_1$	y_{11}	....	y_{1q}	....	$y_{1\theta}$	
x_{i1}	$O_{11}(x_i)$	....	$O_{1q}(x_i)$	....	$O_{1\theta}(x_i)$	$\sum_{j=1}^{\theta} O_{ij}(x_i)$
....	....	....	....	....	.....	.....
x_{ip}	$O_{p1}(x_i)$	...	$O_{pq}(x_i)$		$O_{p\theta}(x_i)$	$\sum_{j=1}^{\theta} O_{pj}(x_i)$
...	....	...	.....	....	.....	.....
$x_{i\eta_i}$	$O_{\eta_i 1}(x_i)$	....	$O_{\eta_i q}(x_i)$	....	$O_{\eta_i \theta}(x_i)$	$\sum_{j=1}^{\theta} O_{\eta_i j}(x_i)$
	$\sum_{i=1}^{\eta_i} O_{i1}(x_i)$	....	$\sum_{i=1}^{\eta_i} O_{iq}(x_i)$	...	$\sum_{i=1}^{\eta_i} O_{i\theta}(x_i)$	n

Tabla 2.2 .Tabla de contingencia que relaciona las variables x_i y y_1

Donde:

$O_{pq}(x_i)$ es el número de casos que toman simultáneamente los valores x_{ip} y y_{1q}, $p=1,\eta_i, \quad q=1,\theta$

Los valores $O_{pq}(x_i)$ que se muestran en la Tabla 2.2, constituyen las frecuencias observadas. Sin embargo, de acuerdo a las reglas de probabilidad estas frecuencias debían ser :

$$E_{pq}(x_i) = \frac{\sum_{i=1}^{\eta_i} O_{iq}(x_i) \cdot \sum_{j=1}^{\theta} O_{pj}(x_i)}{n}$$

(2.10)

que son las llamadas frecuencias esperadas o teóricas.

Estas frecuencias esperadas deben ser comparadas con las frecuencias observadas $O_{pq}(x_i)$ en la Tabla 2.2

Las diferencias $O_{pq}(x_i) - E_{pq}(x_i)$ se llaman residuales, se elevan al cuadrado para evitar la compensación de diferencias positivas y negativas y se dividen por las frecuencias esperadas $E_{pq}(x_i)$ para establecer magnitudes relativas. Resulta en el estadístico:

$$\chi^2(x_i) = \sum_{p=1}^{\eta_i} \sum_{q=1}^{\theta} \frac{\left(O_{pq}(x_i) - E_{pq}(x_i)\right)^2}{E_{pq}(x_i)} \tag{2.11}$$

Si la hipótesis fundamental de independencia es cierta, este estadístico tiene distribución aproximadamente igual a la Chi-cuadrado, con grados de libertad determinado por el producto (C-1)*(R-1), donde C y R son el número de filas y columnas de la Tabla 2. La idea de los grados de libertad es aquí clara pues este es efectivamente, el número de celdas de la Tabla 2 que "podrían llenarse libremente" si están fijados los totales marginales de filas y columnas.

La prueba Chi-cuadrado tiene realmente muchas limitaciones y los principales detractores llegan incluso a decir que el único caso en que él puede ser aplicado con fiabilidad, es el caso de las tablas 2*2. Una de las principales limitaciones es que suministra muy poca información sobre la forma y sobre la estrechez de la asociación entre las variables. Por ello junto con el Chi-cuadrado o en su lugar se utilizan otros estadísticos o pruebas que hablan de la forma o fortaleza del enlace.

Otras medidas de asociación entre variables aleatorias discretas nominales y ordinales.

En muchas investigaciones, más que discernir la dependencia de dos variables interesa la naturaleza y fortaleza de la asociación. Los indicadores que miden esto se llaman *medidas de asociación*. Ninguna medida resume adecuadamente todos los tipos de asociación posibles. Las medidas se diferencian por su interpretación, y por la forma en que ellas pueden reflejar asociaciones perfectas o parciales. Las medidas difieren además por la forma en que ellas se afectan por otros factores, por ejemplo los totales marginales. Asi, hay medidas que son muy "sensibles a los marginales", en el sentido que su valor está muy influenciado por la distribución de los totales marginales de filas y columnas. Tales medidas reflejan por tanto información sobre los marginales, además de sobre la asociación.

Una medida de asociación determinada puede tener un valor bajo para una tabla dada, pero por ello no significa que las variables objeto de estudio no estén relacionadas, sino que ellas no están relacionadas en la forma a la que es sensible dicha medida. Por ello ninguna medida individual es la mejor para todas las situaciones. Al seleccionar una, debe tenerse en cuenta el tipo de datos, la hipótesis de interés así como las propiedades de cada medida. A veces se calcula un gran número de medidas y luego se referencian aquellas que más "ayudan a respaldar" las hipótesis del investigador, como si fueran las únicas que se hubieran calculado; pero esto no es correcto a menos que haya un verdadero análisis de cuáles medidas son las que se necesita observar.

Se verán primero algunas medidas de asociación para variables discretas en dependencia del tipo de datos a tratar.

Medidas nominales

Se consideran aquí medidas que asumen apenas que las dos variables de la tabla están medidas nominalmente. Como tales, estas medidas pueden suministrar solamente alguna indicación sobre la estrechez de la asociación y no pueden indicar casi nada sobre la dirección o cualquier otra cosa de la naturaleza de la relación. Las medidas se pueden clasificar en dos tipos: aquellas que están basadas en el Chi-cuadrado y aquellas que se fundamentan en la lógica de reducción proporcional del error (PRE, Proportional Reduction Error, como se le dice clásicamente).

Medidas nominales basadas en Chi-cuadrado.

Como se explicó antes, el test Chi-cuadrado en sí, no proporciona una buena medida del grado de asociación entre las dos variables; pero como está tan expandido el uso del Chi-cuadrado en la dócima de independencia, se ha estimulado la definición de medidas de asociación basadas en el Chi-cuadrado tratando de minimizar la influencia del volumen de la muestra y de los grados de libertad, así como restringir el rango de los valores de la medida al intervalo [0,1].

Estas medidas ayudan entonces a comparar los resultados del Chi-cuadrado en tablas diferentes cuando hay variación de las dimensiones y de los volúmenes de las muestras. Sin estas correcciones, es absolutamente inadmisible comparar con el Chi-cuadrado tales tablas.

El llamado coeficiente Phi, modifica el Chi-cuadrado dividiéndolo por el volumen de la muestra y extrayendo la raíz cuadrada del resultado:

$$\Phi(x_i) = \sqrt{\frac{\chi^2(x_i)}{n}} \tag{2.12}$$

Para tablas en las cuales una dimensión es mayor que 2, Phi no yace necesariamente entre 0 y 1, ya que el valor del Chi-cuadrado puede ser más grande que el volumen de la muestra. Por tanto Phi sólo queda estandarizado en el intervalo [0,1] en tablas en las cuales R=2 o C=2 (R y C denotan siempre el número de filas y columnas).

Para obtener una medida que debe hacer siempre entre 0 y 1 Pearson[ZAM97] sugirió el uso de:

$$C(x_i) = \sqrt{\frac{\chi^2(x_i)}{\chi^2(x_i) + n}} \tag{2.13}$$

que es conocido como coeficiente de contingencia. El valor de este coeficiente está siempre entre 0 y 1; pero generalmente no puede llegar a ser 1. De hecho, el máximo

valor de C depende sobre el número de filas y columnas. Por ejemplo para una tabla 4x4, el máximo valor de C es 0.87.
Cuanto mayor sea el valor de $C(x_i)$, mayor será el grado de asociación , o lo que es lo mismo cuanto menor sea el valor de $C(x_i)$, mayor será el grado de asociación.
Ahora se define la siguinte magnitud

$$p(x_i) = 1 - C(x_i)$$

(2.14)

Si el valor $p(x_i) \leq \alpha$ (donde α es un nivel de significación prefijado de antemano, normalmente 0.05) la relación es estadísticamente significativa. El mejor predictor es aquel que tenga el valor más pequeño de $p(x_i)$ que sea menor o igual que alfa (α).

Medidas Ordinales

Lógicamente, las variables con nivel de medición ordinal, pueden ser analizadas en tablas de contingencia con las medidas de asociación descritas anteriormente para variables simplemente nominales; pero otras medidas reflejan información adicional sobre la base del ranqueo.

Al considerar la clase de relaciones que pueden existir entre dos variables ordinales, se arriba a la noción de dirección de la relación y al concepto de "correlación generalizada". Las variables están positivamente correlacionadas cuando los casos con valores bajos de una de las variables tienen también valores bajos de la otra y los casos con valores altos de una de las variables los tiene también altos en la otra. Dicho en otras palabras, la correlación es positiva cuando la "relación de dependencia tiende estadísticamente a ser creciente". Las variables correlacionadas negativamente muestran el comportamiento inverso: a más altos valores de la primera variable corresponden los más bajos valores de la segunda.

Para tablas con dos variables ordinales hay varias medidas de asociación que se basan en la comparación de los valores de ambas variables en todos los pares posibles de casos u observaciones. Un par de casos se dice "concordante" si los valores de ambas variables para un caso son más altos (o ambos más bajos) que los valores para el otro caso. El par de casos se dice "discordante" si el valor de una variable para un caso es mayor que el correspondiente valor para el otro caso y la dirección está invertida en la segunda variable. Cuando los casos tienen valores idénticos sobre una o las dos variables se dicen "ligados". De esta forma, para cualquier par de casos con medidas de las variables X e Y, el par puede ser concordante o discordante o ligado, esto último en tres formas posibles: ligado respecto a X pero no de Y, ligado respecto a Y pero no de X y ligado respecto a X e Y.

II.3.1.1 Creación del Árbol de Decisión a partir del Nivel de Significación de las Variables

Sea K el conjunto de rasgos predictores x_i con nivel de significación $p(x_i) \leq \alpha$ (el valor de α es el valor mínimo que puede tomar el nivel de significación y lo asigna el experto). Luego en este conjunto se encuentran los rasgos predictores que están más asociados al rasgo objetivo.

$$K = \{x_i : p(x_i) < \alpha\}$$

(2.15)

El árbol de decisión se construye siguiendo el siguiente algoritmo:

Algoritmo para la Construcción del Árbol de Decisión

A1: Situar en la raíz del árbol los casos agrupados por la decisión.

A2: $K' = K$

A3: Buscar en el conjunto K' el rasgo predictor que esté más asociado con los valores del rasgo objetivo y_1 y expandir la raíz del árbol por los valores asociados a dicho rasgo predictor.

A4: Eliminar el rasgo seleccionado de K'

A5: Repetir mientras $K' \neq \phi$.

iv) Dentro del conjunto K' elegir el conjunto de los rasgos predictores que estén más asociados con el rasgo predictor que se encuentra en las hojas. Reducir K' a ese conjunto.

v) Si $K' = \phi$ terminar. En caso contrario elegir dentro de K' el rasgo predictor más asociado con el rasgo predictor que se encuentra en las hojas (determinando el grado de asociación como se mostró anteriormente) y expandir cada hoja del árbol por los valores del rasgo predictor seleccionado.

vi) Eliminar el rasgo seleccionado de K'

Por la forma en que se construye el árbol se deriva que en la raíz del árbol se encuentran todos los casos agrupados por el valor de la decisión. Estas agrupaciones, según las técnicas de los conjuntos rugosos, constituyen los *conjuntos conceptos*. Por otra parte si en el nivel $l = l_i$ aparece el rasgo x_i, entonces en este nivel el número de ramas será η_i que es la cantidad de valores que toma dicho rasgo. La rama correspondiente al valor x_{ij} será la rama $b = b_j$. Cada rama conduce a un nodo

donde se encuentran agrupados por el valor del rasgo correspondiente los casos derivados del nivel anterior. Dichos nodos, también para las técnicas de los conjuntos rugosos, constituyen los llamados *conjuntos indiscernibles*.

II.3.1.2 Cálculo de la primera probabilidad.

La probabilidad $P_1(x_i)$ la calcularemos a partir de la frecuencia $O_{pq}(x_i)$ que aparece en las tablas de contingencia asociadas a las variables a medida que realizamos el recorrido en profundidad del árbol de decisión, para ello construimos la ley de probabilidades teniendo en cuenta la dependencia de los valores de los rasgos predictores con el rasgo objetivo y el resto de los rasgos predictores.

Así denotaremos por $P_1(x_i)$ las probabilidades correspondientes al rasgo x_i, dependiendo del comportamiento de ese rasgo con el rasgo objetivo y el resto de los rasgos predictores. Ella es calculada mediante:

$$P_1\left(x_i(O_t) = x_{ij}\right) = \sum_{i=1}^{k} P\left(x_i(O_i) / x_j(O_i)\right) \cdot P\left(y_j(O_i)\right)$$

(2.16)

Caso 1: Cálculo de la primera probabilidad para los valores de los rasgos predictores independientes (que no aparecen en el árbol de decisión).

En el caso que x_i constituya un rasgo independiente, su distribución de probabilidades la construimos a partir de las distribuciones de frecuencias del rasgo x_i en los casos de la tabla original ¡independiente del resto de los rasgos, mediante la distribución de frecuencias y la expresión que usaremos es:

$$P_1\left(x_i\left(O_t\right)=x_{ij}\right)=\frac{\left|x_{ij}\left(O_t\right)\right|}{\left|n\right|}$$

(2.17)

Por ello se toman los criterios:

Caso 2: Cálculo de la primera probabilidad para los valores de los rasgos predictores dependientes (que aparecen en el árbol de decisión).

En el casos que los rasgos sean dependientes procedemos de la siguiente forma:

- Determinar las dependencias entre las variables predictoras y el rasgo objetivo y el resto de los rasgos predictores, con el valor del nivel de significación de cada rasgo se crea el árbol de decisiones.
    - Una vez establecido el primer nodo se expande éste teniendo en cuenta cual es el rasgo que con él se enlaza.
    - Se repite hasta llegar a la hoja o rasgo menos predictor, que será el de menos peso.
- Calculamos los diferentes valores de las probabilidades $P(x_i)$ asociadas a cada rasgo x_i dividiendo el valor de la frecuencia que aparece en cada celda entre el total marginal de la columna que contiene a dicha celda en la tabla de contingencia.
- Los valores de la probabilidad calculada se almacenan en la tabla o en la propia base de forma tal que pueda ser utilizada más tarde en el cálculo de la Certidumbre total asociada al rasgo.

Las Tablas de Contingencia y el árbol de decisiones permiten también inferir directamente la además de la primera probabilidad otros datos que aparecen en ellas como por ejemplo los % de casos distribuidos en las filas y las columnas. Estas tablas tomadas por columna dan la probabilidad de dependencia del rasgo predictor respecto al objetivo y tomadas por fila dan la probabilidad de dependencia del rasgo objetivo respecto al predictor.

Caso 3: Cálculo de la primera probabilidad para los valores del rasgo objetivo (decisión).

Para los valores que no se reflejan en las relaciones del árbol de dependencia y por ende no aparecen en las tablas de contingencia, la primera probabilidad se calcula de forma parecida a como se realiza con los rasgos independientes, pero ahora la fórmula 2.18 toma la forma:

$$P_1\left(y_i\left(O_t\right)=y_{ij}\right)=\frac{\left|y_{ij}\left(O_t\right)\right|}{\left|n\right|}$$

(2.18)

Es decir, usamos la frecuencia con que el rasgo toma un valor determinado de categoría se divide ese valor de frecuencia entre el número total de frecuencias para todas las categorías de ese rasgo (total marginal de la fila correspondiente donde se encuentra la celda), pues en este caso el rasgo independiente y el rasgo objetivo no estarán directamente correlacionados y si existen casos que no se les ha podido determinar el valor de esta primera probabilidad para el rasgo objetivo se procede de manera similar al caso del rasgo independiente, descrito anteriormente pero con la fórmula 2.18.

II.3.2 Cálculo de la segunda probabilidad para los rasgos predictores y objetivos.

Como se sabe que el cálculo de la primera probabilidad condicionada o no de los rasgos no constituye una medida de la certidumbre del valor ya que una fuente importante de incertidumbre es la forma en que se obtienen dichos valores. Es por ello que para determinar la certidumbre total se debe calcular otra probabilidad P_2, que para el caso de los rasgos predictores está muy relacionada con la forma en que se obtienen sus valores. En el caso de los rasgos objetivos usaremos algunas ideas de la teorías probabilísticas adaptadas a nuestro caso.

Determinación de P2 para los rasgos predictores.

Como se ha hecho notar, el término incertidumbre o razonamiento inexacto se emplea en la IA para referirse a un amplio rango de situaciones donde la información relevante es deficiente por diversas causas, atendiendo a ello se consideran los casos siguientes:

Caso 1: El valor de la variable es tomado de una fuente en la cual no hay motivos para no confiar.

$$P_2\left(x_i\left(O_t\right) = x_{ij}\right) = 1$$

(2.19)

Caso 2: La variable toma el valor a partir del criterio de varios expertos.

Consideremos la existencia de n expertos: $a_1, a_2, ..., a_n$, asignando el valor del rasgo x_i del caso O_t. Dado un valor particular x_{ik} de x_i consideremos además que $a_r\left(x_{ik}\right)$ denota la respuesta del experto *r* sobre ese valor. Asumamos que $a_r\left(X_{ik}\right) = 1$ cuando el valor se afirma y $a_r\left(x_{ik}\right) = 0$ cuando el valor se niega. Entonces

$$P_2\left(X_i\left(O_t\right) = x_{ij}\right) = \frac{\sum_{r=1}^{n} c_r \cdot a_r\left(x_{ij}\right)}{\sum_{r=1}^{n} c_r}$$

(2.20)

donde c_r es el grado de competitividad del experto.

Caso 3: La variable toma el valor a partir de un instrumento de medición.

$$P_2\left(x_i\left(O_t\right) = x_{ij}\right) = P$$

(2.21)

donde *P* es la certidumbre asociada al instrumento de medición utilizado y constituye una probabilidad por si sola.

Determinación de la segunda probabilidad, P_2 para los rasgos objetivos.
Dada la base de casos descrita en el inicio del presente Capítulo, que puede incluso ser representada a través de una tabla donde se encabezan las columnas con los distintos rasgos predictores seguidos cada uno de ellos por las probabilidades calculadas y las filas con los casos de la base, se calculan las probabilidades para cada caso al rasgo objetivo, sacando un promedio de las segundas probabilidades de todos los rasgos predictores según la expresión:

$$P_2(y_1(O_t) = y_{1k}) = \frac{\sum_{i=1}^{n} P_2(x_i(O_t) = x_{ij})}{n}$$

(2.22)
y con ello queda lista para el cálculo de la Certidumbre de los rasgos.

Cálculo de la Certidumbre Total
Si en una base de casos se calcula una primera probabilidad condicionada por la relación entre los rasgos predictores al objetivo y se calcula una segunda con la información proveniente de diferentes fuentes o apoyados en la propia teoría probabilística se puede calcular la Certidumbre total, como una función de ambas de manera que la base quede mejor definida y lista para la recuperación y adaptación ante cualquier aparición de un nuevo caso a solucionar.
Se calcula la certidumbre total de cada valor para cada caso usando la expresión:
$\xi_i(O_t) = g(P_1, P_2)$ como ya se había explicado en la parte de la fundamentación, donde $\xi_i(O_t)$ es la certidumbre del valor del rasgo *i* para el caso O_t y g puede ser cualquiera de las funciones de combinación que han sido utilizadas en la literatura para propagar incertidumbre y que se conocen como t-conormas arquimedeanas por ejemplo:

$$g(x, y) = x + y - xy$$

(2.23)

$$g(x, y) = \frac{x + y}{1 - xy}$$

(2.24)

entre otras.
Estas funciones de combinación cumplen las siguientes propiedades:

7. Si $0 < x, y < 1$ entonces $0 < g(x, y) < 1$

8. $g(x, y) = g(y, x)$

9. $g(x, g(y, z)) = g(g(x, y), z)$

10. $g(x,0) = x$

11. Si $x \leq y$ entonces $g(x,z) \leq g(y,z)$

12. $g(x,1) = 1$

Esta última propiedad resulta importante ya que garantiza la seguridad absoluta cuando cualquiera de las dos probabilidades asignadas es 1.

Recuperación de un nuevo caso

Para el nuevo caso esta certidumbre se calcula teniendo en cuenta sólo P_2.

Una vez calculada la certidumbre de todos los valores para los rasgos predictores y objetivos y la certidumbre de los valores del nuevo caso, se procede a realizar los procesos de razonamiento (recuperación y adaptación).

El módulo de recuperación consta de dos etapas:

- Etapa de acceso.
- Etapa de recuperación.

En la etapa de acceso haremos uso de la estructura jerárquica de representación de los casos. Esta estructura nos permitirá reducir el número de casos a comparar pues sólo consideraremos aquellos casos que para los rasgos mas predictores (los que están en el árbol) tengan igual valor que en el problema planteado. Esto no significa que vayamos a considerar todos los rasgos predictores que están en el árbol pues si todos son considerados pudieran no tenerse en consideración casos que difieran en pocos rasgos significativos. Por eso hay que definir hasta que nivel del árbol vamos a considerar. El siguiente algoritmo muestra este proceso:

Algoritmo para el Acceso a los casos más relevantes

Este algoritmo está basado en el uso de una heurística de preferencia conocida como preferencia de características sobresalientes [Kol89]

Algoritmo de Acceso a los Casos más Relevantes

Entrada: Problema a resolver: O_0

Nivel a considerar: level

Conjunto de rasgos más predictores: K

Salida: Conjunto de casos potenciales para la recuperación: P

A1: $P = \phi$

A2: Para *l=1, level*:

i) Seleccionar el rasgo más que se encuentra en el nivel l. Sea este x_i

ii) Si $x_i(O_0) = x_{ip}$ entonces $P = P \cap X_{ip}$

De esta forma en el conjunto P nos quedan los casos potenciales para la etapa de recuperación, en la cual se determina la semejanza entre ellos y el nuevo problema. En este proceso se tienen en consideración no solo los más predictores, sino todos los rasgos iniciales y la certidumbre de los valores que ellos toman. A la primera función de semejanza mostrada anteriormente se le realiza una transformación a través del siguiente algoritmo.

Algoritmo para la Transformación de la Función de Semejanza en el Módulo de Recuperación

Entrada: P, O_0 nc (número de casos más semejantes a recuperar)

Salida: S Conjunto de casos más semejantes

A1: Para todo $O_t \in P$ determinar

$$\beta(O_0, O_t) = \frac{\sum_{i=1}^{n} p_i' \cdot m_i(O_0, O_t) \cdot \delta_i(O_0, O_t)}{\sum_{i=1}^{n} p_i'}$$

(2.2.5)

$$m_i(O_0, O_t) = \begin{cases} \dfrac{\xi_i(O_0) + \xi_i(O_t)}{2} & si \ \ (\xi_i(O_0) \wedge \xi_i(O_t)) \geq \varepsilon \\ 0 & e.o.c \end{cases}$$

(2.26)

$$p_i' = \begin{cases} 1 - p(x_i) & \text{si el peso no ha sido asignado con anterioridad} \\ p_i - p(x_i) & e.o.c. \end{cases}$$

(2.27)

donde:

n: Número de rasgos predictores.

p_i': Peso o relevancia del rasgo *i*.

A2: Agrupar los *nc* casos más semejantes en el conjunto *S*.

Como se puede observar en la fórmula anterior si para el problema o para el caso con que estamos comparando, el valor del rasgo que se está considerando está por debajo de un umbral predeterminado, dicho rasgo no se toma en consideración ya que estaríamos comparando valores cuyo nivel de certidumbre está por debajo de lo permisible.

II.3.1 Optimización de la adaptación.

Dado un determinado caso la adaptación se efectúa determinando el caso óptimo de los casos recuperados. El criterio de optimalidad se da a través de una función que combina el valor de la semejanza para dicho caso y la certidumbre de su solución. Esto se realiza a través de la fórmula:

$$\mu(O_t) = \alpha\beta(O_0, O_t) + (1-\alpha)\eta_1(O_t)$$

(2.28)

donde α es un parámetro que se selecciona según el criterio del experto. A medida que α tiende a 1 significa que se le está dando mayor importancia a la semejanza que a la certidumbre de la solución.

Capítulo III.

Implementación de un programa. Prueba con una base de casos real.

Las ideas expuestas con antelación constituyen la base para el diseño del sistema implementado "TreeData" o programa para el cálculo de la Certidumbre en un Sistema de Conocimiento Basado en Casos (SBC), el cual además permite la recuperación de casos semejantes cuando se introduce un nuevo problema.

En el presente Capítulo se exponen las características del software elaborado, una breve descripción de su funcionamiento y se utiliza una base de casos real para probar la efectividad del programa.

III.1 Requerimientos y Características del Software.

La Programación Orientada a Objetos (POO) es el método de implementación en el que los programas se organizan como colección corporativas de objetos, cada uno de los cuales representa una instancia de una clase, y cuyas clases son, todas ellas relacionadas, miembro de una jerarquía de clases unidas mediante relaciones de herencia.[GRA91]
El enfoque orientado a objetos ha sido un paso evolutivo en el análisis, diseño e implementación de Software. Este paradigma se presta por sus características para lograr un software, poco resistente a los cambios y fácilmente mantenible; estas son características valiosas en cualquiera de los sistemas y sobre todo en los sistemas "inteligentes", los cuales deben mostrar una buena adaptación al medio en que son utilizados. La utilización de las técnicas de la Inteligencia Artificial favorece al desarrollo de estas formas de programación, pues permiten implementar software más contundentes que desafíen los retos que constantemente enfrentamos. Luego considerando esta situación y dada la importancia que tiene el poder implementar toda la teoría fundamentada se implementa un software programado en el lenguaje "Object Pascal" y utilizando como interfaz de programación el Borland Delphi versión 5. Como resultado se obtuvo una aplicación con las siguientes características:

Título

TreeData. Herramienta para el tratamiento probabilístico de bases de conocimiento bajo incertidumbre en los Sistemas Basados en Casos.

Requerimientos técnicos

- Computadora Personal Pentium a 100 MHz con 32 MB de memoria RAM

  (esto se garantiza por ser este precisamente el tipo de computadora donde el

software fue creado, bajo los sistemas operativos Windows 95/98/ME o WindowsNT/2000. Teóricamente debe funcionar con aproximadamente los requerimientos mínimos de Windows 95, como procesador 80486 con 16 MB RAM, pero se recomiendan 32 MB RAM como mínimo dado que el consumo de memoria del programa llega hasta los 8.5 MB con bases de conocimiento con más de 200 casos, o sea relativamente pequeñas. El tamaño de la base de conocimiento influye directamente en la cantidad realmente requerida.

- 1,5 MB de espacio libre en disco duro.
- Necesita que estén instalados los objetos ADO y OLEDB de Microsoft. La instalación de ADO y OLEDB está incluida en el MS Access 97 o 2000 de la suite ofimática MS Office 97 o 2000 y en varios programas de desarrollo de aplicaciones como Visual Basic, Visual C++ o Borland Delphi. Estos objetos pueden instalarse sin tener que instalar ninguno de los programas anteriores bajando la última versión de ADO y OLE que se pueden copiar de Internet en la dirección http://www.microsoft.com/downloads donde se encuentra siempre su última versión.

Descripción

El sistema consta de un solo fichero ejecutable: TreeData.exe y no necesita de otros archivos secundarios para su ejecución. Las bases de casos (BC) son almacenadas en ficheros de extensión mdb, lo que permite que puedan ser editadas en Microsoft Access o importado en Microsoft Excel. La edición y visualización puede realizarse completamente con el programa TreeData si se cumplen todos los requerimientos técnicos. En una versión diferente se incluirá la posibilidad de tener la base de conocimiento en cualquier sistema que permita exportarlas a través del ODBC[1] de la máquina. En cada uno de los ficheros de extensión mdb pueden existir más de una base de casos, cada una de las cuales puede ser matemáticamente representada, según lo fundamentado en el Capítulo II, como una matriz con las características de la Tabla 2.2, donde cada fila representa un registro de un caso y cada columna, representa los valores que toma un rasgo determinado cuando el caso es evaluado. A los valores tomados por los rasgos se les llama a menudo categoría del rasgo. Los rasgos y categorías pueden editarse fácilmente, incluyéndoles un nombre, una

[1] ODBC: Siglas en inglés “Open Data Base Conectivity”. Forma estándar para compartir bases de datos entre sistemas diferentes. Ejemplos: MS Access, MS SQL Server, PostgreSQL, etc.

abreviatura y una breve descripción[2], de manera que la base tenga toda la información necesaria solicitada por el sistema durante el procesamiento. Para lograr esto en el fichero mdb se crean algunas tablas especiales para describir los datos de los rasgos, categorías, así como los datos de los criterios de expertos para los rasgos que lo requieran, ellas se caracterizan por tener el signo "_" (subrayado) al inicio y final del nombre y serán a partir de entonces parte de la base de conocimiento, por lo que no deberán alterarse fuera del sistema. Al comenzar con una base de casos nueva los rasgos y categorías se determinan automáticamente, facilitando su edición al importarlas.
La última versión de TreeData siempre se encontrará disponible en el sitio web http://treedata.4t.com o http://telleria.4t.com[3]. Puede contactarse con el autor por correo electrónico a la dirección andres@telleria.4t.com o treedata@telleria.4t.com.

Partes fundamentales del diseño
Sin pretender hacer una exhaustiva presentación del diseño, se hace una descripción general de las Clases, Tipos e Interfaces fundamentales usados en el sistema, y sus relaciones más importantes. Para una mejor comprensión se hace una breve descripción de las clases fundamentales:
Clases
TDModCommon: Es la clase central, una instancia de ella dirige las operaciones del sistema. Maneja completamente la conexión con la base de datos; los demás objetos, a través de las interfaces que ella exporta, acceden a los datos en la base de casos y las tablas de rasgos, categorías y criterios de expertos. Contiene la lista de rasgos y categorías. Implementa las interfaces *IRelationalData* e *IOwner*.
TFormTreeData: Es la clase visual principal, maneja el ambiente principal y contiene los módulos de:

- Entrada de información inicial: TFrameIntro
- Edición de rasgos y categorías: TFrameRasgos
- Entrada de nuevos casos y modificación de existentes: TFrameDataInput
- Edición del árbol de decisiones y cálculo en general: TFrameTree

TFrameIntro: Entrada y modificación de casos, lo hace a través de una referencia a la tabla de casos que maneja TDModCommon.
TFrameRasgos: Entrada y modificación de Rasgos y Categorías. Contiene una copia de edición de la lista de rasgos cuando se salva indica a TModCommon que actualice la lista de rasgos que contiene, TDModCommon entonces lo graba en la base de datos.

---

[2] Por ejemplo: **Nombre**: "Escolaridad", **Abreviatura**: "x1", **Descripción**: "Esta es la escolaridad de…"
[3] En este sitio también se encuentra una copia de este documento.

TFrameDataInput: Entrada y modificación de datos de los casos, obtiene y actualiza la base a través de TDModCommon.
TFrameTree: Contiene una instancia de TTreeStructView y le provee espacio para pintarse. Maneja las operaciones de expansión y cálculo de probabilidades aunque no es realmente quien las ejecuta, indica que se ejecuten en respuesta a una orden del usuario.
TTreeStructView: Se encarga de la interfaz visual del árbol de decisiones. Contiene al nodo raíz del árbol que es una instancia de TtreeStruct. Para relacionarse con este último implementa la interfaz *ITreeStructOwner*. Esta componente no existía y tuvo que ser creada para la representación y manejo de un árbol. Está programado como un Control ActiveX, lo que permite su utilización desde la mayoría de los lenguajes de programación actuales sin programación extra. Permite acceder a todos los nodos del árbol mediante cualquier algoritmo de la I.A. que se programe, ello se hace accediendo al atributo TreeStruct(es el nodo raíz del árbol en la representación a mostrar). Cada nodo posee la siguiente estructura:

Nodo: TTreeStruct
Text: String Data: String BinData: Tobjectt Position: Trect Items[i]: arreglo
Count: integer

- Sus partes se explican a continuación:
Text: Texto y dato x_o relativo al nodo. Cuando se dibuja se pone el texto y si tiene "Data" el signo "=" y al lado el contenido de "Data".
BinData: Permite asociar a cada nodo cualquier tipo de dato útil para la manipulación del problema específico que se trate. Por ejemplo dado el caso, se le asigna Tvariable la cual realiza las sumas efectuadas de valores de los rasgos y luego el cálculo de las probabilidades P_1 y P_2.
Position: Posición en que se va a mostrar.
Items[i]: Arreglo con los nodos hijos de este nodo.
Count: Cantidad de nodos hijo que tiene este nodo.
El nodo brinda los métodos AddItem DeleteItem, necesarios para la construcción del árbol pues adicionar y borrar un nodo hijo respectivamente.
Con estos atributos se puede manipular completamente el árbol (crear, visualizar, recorrer, etc.)
El componente brinda a demás las siguientes posibilidades.

- Recorrido del árbol en preorden. O sea primero visita al nodo y luego a sus hijos, para lo cual llama al método TraverseTree, pasándole como parámetro un procedimiento del usuario que se ejecutará cada vez que se visite un nodo

pasándole el nodo en sí. Un método así tendrá acceso a todos los datos de cada uno de los nodos en el orden del recorrido.

- Parámetros visuales para dibujar el árbol, color con el que se dibujan los nodos, tamaño de un nodo, distancia entre nodos, distancia entre niveles, etc.

A continuación se continúa con la definición de otras clases importantes para más tarde mostrar cómo se relacionan a través de un diagrama de clases según las técnicas del Lenguaje Unificado de Programación, UML[*]

TTreeStruct: Contiene la información de los nodos del árbol. Cada nodo es un TTreeStruct que contiene una lista de nodos hijos, tiene además asociado un objeto de clase TVariable.

TVariable: Existe una contenida en cada nodo del árbol, realiza los cálculos de frecuencia. Los cálculos los realiza a partir de los datos en la base de casos. Para eso contiene una referencia a un *IRelationalData* (TFrameTree que es quien asocia al TVariable con cada nodo le pasa el *IRelationalData* que es TModCommon). A través de su *IDataProvider* las instancias de TVariable realizan las consultas a la base de casos.

TFormReport: Realiza los cálculos de probabilidades y los muestra, lo realiza por orden de TFrameTree. Contiene una referencia a *IRelationalData* (TModCommon).

TDataUnit: (Clase Abstracta) Representa una unidad de datos que contiene y maneja:

- Nombre
- Abreviatura
- Descripción
- Owner : *IOwner* --> Conoce como acceder a la base de casos.

TCategory: Hereda de TDataUnit. Representa datos de las categorías. Tienen una referencia a IOwner (Los rasgos) a los que notifican cualquier cambio.

TRasgo: Hereda de TDataUnit e implementa *IOwner*. Contiene una lista de instancias de TCategory (categorías del rasgo), con una referencia a *IDataProvider* (TDModCommon) que le permite salvar su definición en una base de datos (o leerla de ahí) haciéndola persistente. Su *IOwner* es el mismo *IDataProvider* (TDModCommon).

Ver el diagrama de clases según la metodología UML[4],Ver figura 3.1a del Anexo VII.

III.2 Manual del Usuario.

TreeData es un software diseñado para el tratamiento probabilístico de los Sistemas de Conocimientos Basados en Casos bajo incertidumbre, por lo que desde un inicio comenzó considerando la dependencia de evidencia en los rasgos para el cálculo de

[*] Ciclas en inglés de "Unit modelation Language"

[4] UML. Siglas en inglés de *Unified Modeling Language*,. metodología de diseño orientado a objetos.

una primera probabilidad y en función de las fuentes de los datos calcula una segunda probabilidad para más tarde combinar ambas según al regla PD5 del Capítulo I para calcular la Certidumbre Total de cada caso evaluado en los distintos rasgos pero además
con la actual versión del software Treedata implementada se puede:

- Crear, organizar y completar una base de conocimiento.
- Generar el árbol de decisiones a partir de los niveles de asociación entre las variables que describen los rasgos, el cual permite una organización más eficiente de la base de casos.
- Calcular la Certidumbre de los valores asociados a los rasgos predictores teniendo en cuenta la estructura de organización creada y las fuentes de obtención de los mismos.
- Recuperar casos similares a la descripción de un nuevo problema dado.

Entorno Visual del TreeData.

El programa está diseñado sobre la plataforma Windows y aprovecha las facilidades de su ambiente para brindar una mayor comodidad al usuario. Las tareas están dispuestas en un orden lógico y el acceso a las mismas se hace con relativa facilidad a través de barras de herramientas y del menú.

Al iniciar el programa muestra la pantalla de presentación que se puede apreciar en la figura 3.1 b. donde se facilitan los créditos e información secundaria que permite a interesados contactar con los creadores del Software.

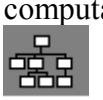
Figura 3.1 b. Pantalla de Presentación del programa.

Aquí aparece además un botón "**Continuar**" programado para que el usuario pueda continuar con la ejecución del programa pasando a una segunda forma o ventana de trabajo que se muestra en la figura 3.2 a, en la que pueden observarse las opciones:

Archivo: Da la posibilidad de iniciar directamente el trabajo abriendo una base de casos ya existente y almacenada en algún dispositivo de memoria.

Rasgos: Permite editar la definición de los rasgos y categorías. Cambiar, borrar y adicionar rasgos y categorías. (Nota: Si se hace alguna modificación esta debe ser salvada para que se haga efectiva). Al comenzar con una base de casos nueva los rasgos y categorías se determinan automáticamente. Ver figura 3.4 a.

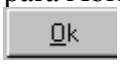
Datos: Para entrar o modificar los datos a la base de casos ya definida y que pueden ser en número únicamente limitado por la memoria de almacenaje de la computadora donde se esté corriendo el programa. Ver figura 3.4 b.


Árbol: Para construir el árbol de decisiones correspondiente a los datos almacenados y se realiza el cálculo de las probabilidades P_1 y P_2. A partir de esta opción se genera el árbol de decisiones, se le asocian a éste las tablas de contingencia y se calcula la primera y segunda probabilidad para cada rasgo a partir de la base de casos seleccionada así como se calcula la certidumbre. Ver figura 3.5.

Recuperar: Para recuperar los casos semejantes cuando se introduce un nuevo problema a resolver. Esta opción es efectiva si se ha generado el árbol de decisiones, calculado las probabilidades y Certidumbre a cada caso, así como salvado en la base de conocimiento los resultados cálculos anteriores.

Además de las opciones aparecen en la misma ventana los botones:

Explorar: Para buscar el directorio donde se encuentra almacenada la base de casos que se quiere abrir usando la misma caja de diálogo del Sistema Windows para Abrir un fichero. (Ver figura 3.2 b)



Ok: Para aceptar la base de casos localizada y desplegar una ventana como la que se muestra en la Figura 3.3.

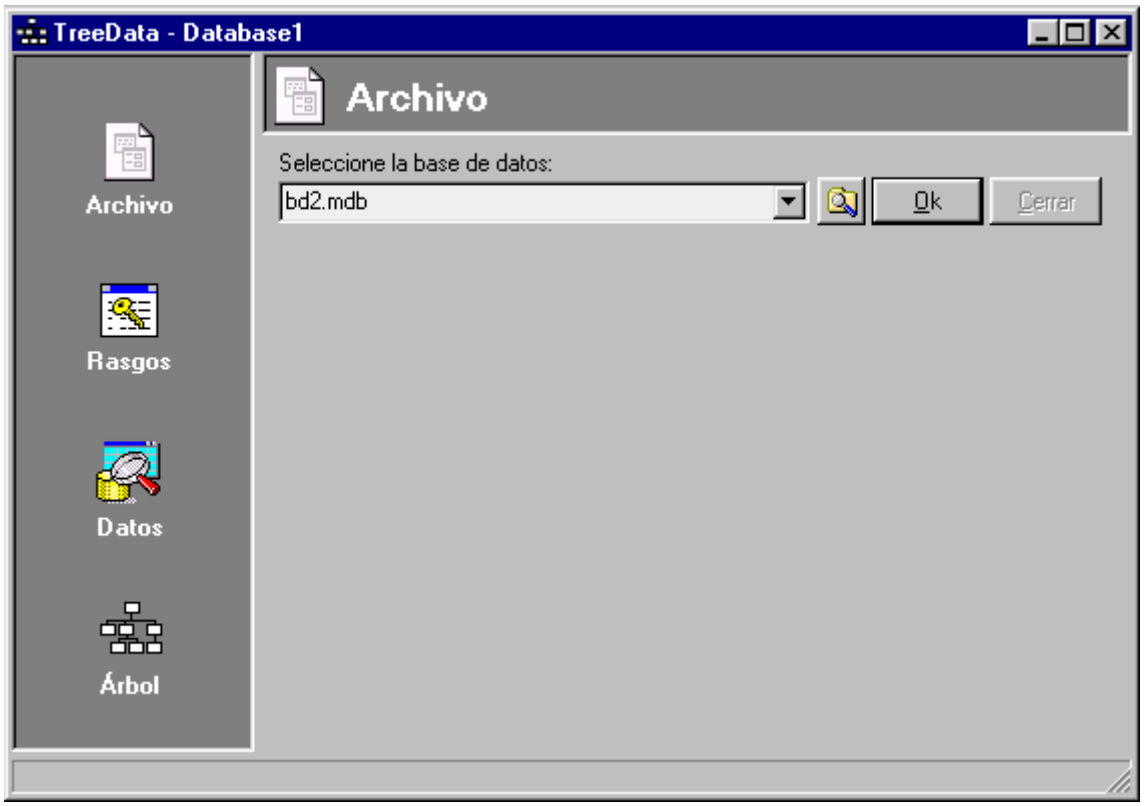


Figura 3.2 a. Pantalla de Opciones para el trabajo con el TreeData.

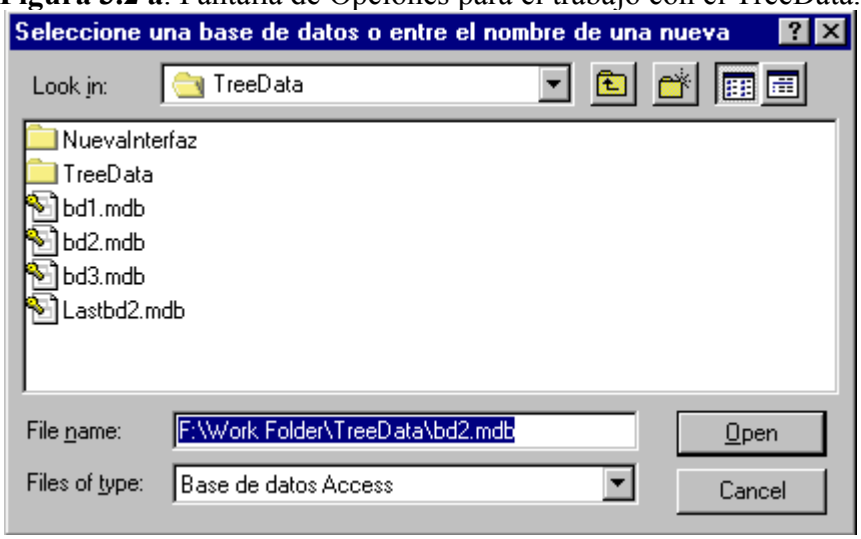


Figura 3.2.b. Cuadro de diálogo para abrir base de casos.

Procesamiento manual con una base de casos ya existente.

Al seleccionar la base de casos y aparecer esta en la barra de direcciones, se oprime el botón

e inmediatamente el programa visualiza la ventana mostrada en la figura 3.3 que muestra el nombre de las tablas contenida en la base.

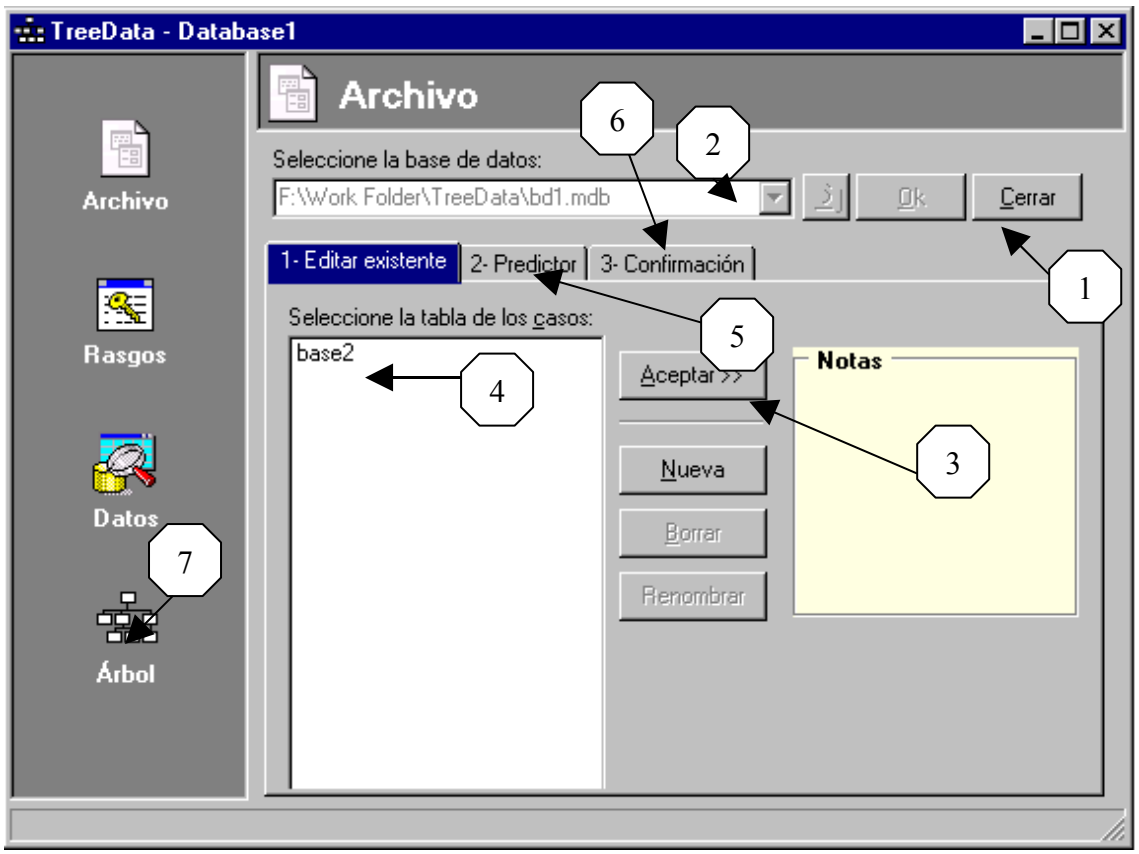


Figura 3.3. Ventana para el procesamiento de la Base de casos de forma manual.

En esta interfaz aparece una barra de dirección (2), con un botón a la derecha habilitado y dos deshabilitados, si se oprime el primero, correspondiente a Cerrar (1), la ventana vuelve a tomar la forma de la figura 3.2 y se habilitan los dos botones antes deshabilitados uno de ellos simbolizado por un icono que tiene forma de Carpeta indicando la posibilidad de acceso a directorios donde se tengan almacenadas las bases de casos. Un botón (3) con etiqueta Aceptar, que debe presionarse teniendo enfocada una tabla en el área (4) que muestra todas las tablas de casos de la base de datos.
Luego se activa al sección (5) donde se permite seleccionar el rasgo objetivo dentro de todos los rasgos de la base de casos seleccionada. Si se confirma el objetivo con (6) el programa brinda un breve resumen de las selecciones hechas y deja lista la

base para editar los rasgos, los casos y generar el árbol de decisiones, seleccionando una de las tres opciones principales respectivamente.

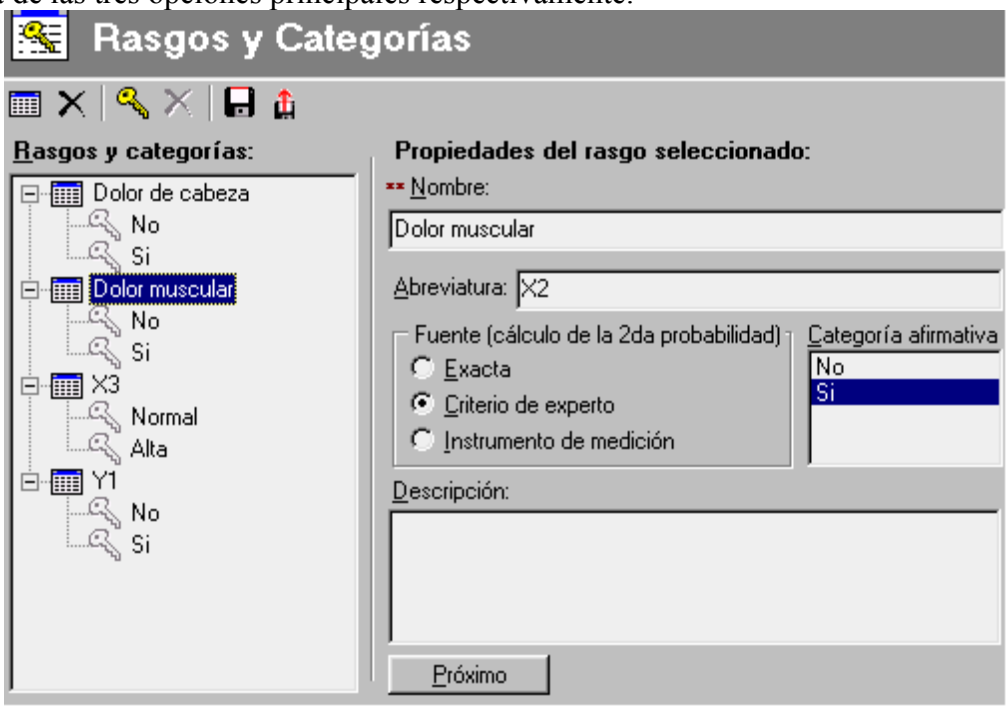
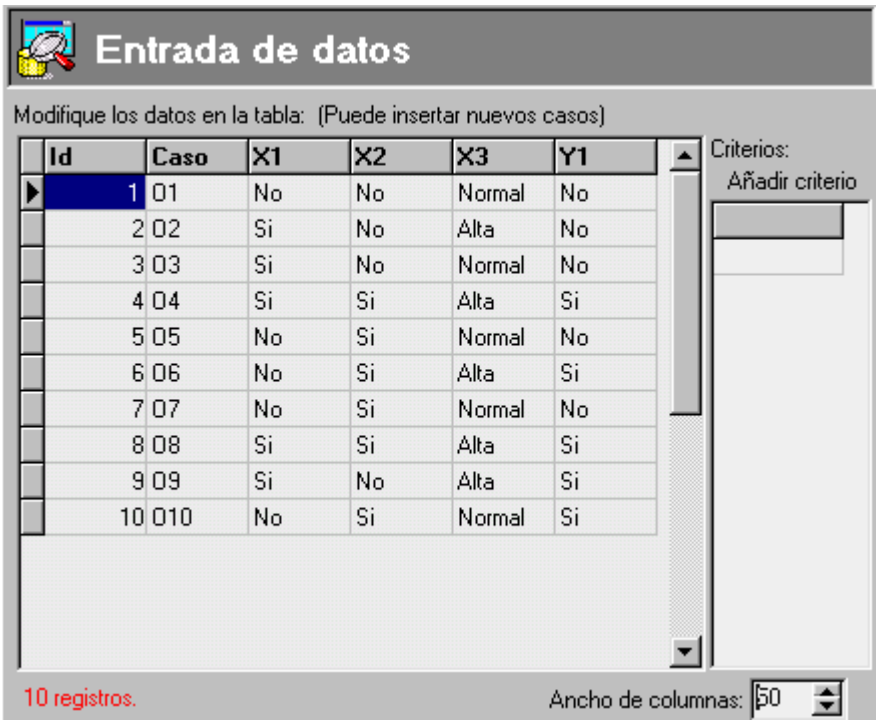


Figura 3.4 a. Opción Rasgos.

La opción Árbol se muestra en la figura 3.5, donde ahora se actualiza oprimiendo la etiqueta Actualizar. Con esta acción el rasgo objetivo se convierte en el primer nodo del árbol de decisiones en forma de rectángulo con el identificador del rasgo como etiqueta.

La figura 3.4 b muestra la tabla con todos los casos que tiene la base evaluados en todos los rasgos, en la parte inferior de la ventana (Barra de estado) se puede apreciar la cantidad de casos y un pequeño cuadro de lista en el que se puede variar las dimensiones de la tabla, si esta es de tamaño muy grande entonces se le puede reducir cambiando el número del cuadro de lista por uno menor, o mayor en caso contrario. A la derecha del área de la tabla aparece una tabla de una sola columna cuando de existir un rasgo en la base cuyo valor sea dado por el criterio de expertos, aquí se aclara el criterio dado por cada experto con 1 si es afirmado y con 0 si es negado, con solo presionar el botón “Criterio” que estará disponible en la parte superior de la tabla Añadir Criterio.



Id	Caso	X1	X2	X3	Y1
1	01	No	No	Normal	No
2	02	Si	No	Alta	No
3	03	Si	No	Normal	No
4	04	Si	Si	Alta	Si
5	05	No	Si	Normal	No
6	06	No	Si	Alta	Si
7	07	No	Si	Normal	No
8	08	Si	Si	Alta	Si
9	09	Si	No	Alta	Si
10	010	No	Si	Normal	Si

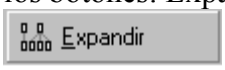
Figura 3.4 b. Edición de los datos de los casos y criterios de expertos.

La figura 3.5 muestra la ventana a la que se accede cuando se oprime el botón 7que genera el árbol de decisiones.

Una vez que se ha introducido el nodo del nivel principal del árbol se debe oprimir el botón actualizar

cuya función es colocar el rasgo seleccionado como predictor en el nodo activo.

Figura 3.5. Ventana que visualiza el árbol de decisiones y las tablas de contingencia. También en esta ventana aparece en la parte inferior un área pequeña por defecto con etiqueta estadística de los nodos hijos en la que se visualizan las diferentes tablas de contingencia asociadas al nodo del árbol activo, dicha tabla de contingencia encabeza las columnas con las categorías del rasgo objetivo y las filas con las categorías del rasgo predictor correspondiente. En la parte inferior derecha de la ventana aparecen los botones: Expandir y Borrar.



Botón para Expandir un nodo del árbol.

Botón para borrar un nodo determinado del árbol

Con el primero aparece una ventana como la de la figura 3.6, que permite seleccionar entre todos los rasgos predictores aquel que directamente esté asociado al nodo activo, de manera que se vaya formando el árbol de decisiones. Una vez construido el árbol se pueden analizar cada una de las tablas de contingencia asociadas a cada uno de los nodos con los valores de las frecuencias por caso, con esos valores se puede comprobar el valor de probabilidad P_I calculado para cada rasgo en cada nodo.

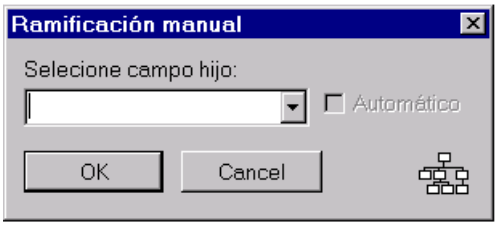


Figura 3.6. Ventana que permite seleccionar el rasgo por el cual se pretende expandir el nodo correspondiente en la creación de árbol.

Si el árbol de decisiones creado es muy extenso y no cabe en la ventana, se pueden: Ampliar la ventana activa, ampliar el área activa dentro de la ventana, disminuir el tamaño del área de la tabla de contingencia o arrastrar la parte interna del área de la ventana que contiene el árbol en cualquier dirección para poder observar aquella parte que quedó oculta en dicha ventana; pero si lo que nos interesa ver mejor es la tabla de contingencia, el área de esta puede ser también modificada en tamaño. Todo aquí queda listo ahora para realizar el cálculo de las probabilidades con solo oprimir el botón nombrado Probabilidades

, las que serán mostradas en una nueva ventana independiente, como la que aparece en la Figura 3.12 del Anexo VIII, con la posibilidad de salvarlas en la base de datos o como un fichero independiente de esta para poderlo visualizar con otras aplicaciones de Windows.

Cuando se muestra la ventana con los valores de certidumbre calculados se oprime la tecla salvar

para guardar los resultados en la base o la tecla salvar en fichero

para guardar los resultados en un fichero texto que pueda ser leído con cualquier editor de ficheros de este tipo.

III-3 Ejemplo con una Aplicación

TreeData no se consideró nunca como un software para la modelación teórica de problemas hipotéticos pues de esa forma sería muy limitado su uso, de ahí que se pretende mostrar a través de un ejemplo de la vida real como es posible que el programa trabaje en los cálculos para los que fue implementado.

III-3.1 Edición, organización y completamiento de la Base de Casos (BC).

Dados los objetivos del trabajo se hace necesario utilizar una base de conocimiento (Base de Casos BC) original con todos los casos evaluados en cada uno de sus rasgos, para ello se puede hacer uso de la opción que permite crearla en la interfaz del TreeData o utilizar una ya creada.

De revisiones realizadas [ZAM97 y DÍA99], comprobamos que desde hace algún tiempo se investiga en la Universidad Central de la Villas el problema del prediagnóstico a trastornos neuropsíquicos leves en escolares, basado en técnicas estadísticas y de inteligencia artificial. Dentro de tales técnicas figuran por ejemplo: Análisis de Clustering, Segmentación CHAID y Redes Neuronales. Pero en pocas o ninguna de ellas se atiende el problema de lo difuso, incierto o erróneo de la información recopilada para ser procesada y en casi todas se consideran los datos como fiables y posiblemente lo sean pero siempre queda un margen para la duda luego el trabajo con el conocimiento incierto o dudoso almacenado nos hizo seleccionar precisamente, para nuestro ejemplo una base de casos creada en Access con 208 casos de niños en edad escolar comprendida entre 6, 7 y 8 años. De esa base se quiere identificar y tratar aquellos casos que presentan Trastornos Neurosiquicos Leves (TNL), con la idea de aplicar una terapia adecuada capaz de curar estos trastornos, porque si bien es conocido que existen en nuestro país equipos de alta tecnología (Tomografía Axial Computarizada, Neurónica, MEDICID, Potenciales Evocados), que se cuenta además con un arsenal amplio de pruebas psicológicas, técnicas patopsicológicas y neuropsicológicas con potencialidades para contribuir al diagnóstico, a la vez que múltiples recursos de la terapia psicológica, sigue siendo una limitación la posibilidad real de aplicación masiva de estos recursos y sobre todo la insatisfecha necesidad de un diagnóstico precoz, confiable y diferenciado, sin contar con la falta de fundamentos para resolver criterios controvertidos de la terapéutica.

TRASTORNOS NEUROPSÍQUICOS LEVES (TNL)

Discapacidad identificada por encuestas y exámenes psicológicos masivos a escolares supuestamente sanos que revelan morbilidad declarada (clínica) u oculta (subclínica) o inaparente, que aparece con una prevalencia por encima del 10% hasta el 20% en poblaciones escolares menores con una diversidad de formas de presentación que implican alteraciones parciales de las funciones psíquicas, que a su vez se convierten en signos y síntomas inespecíficos propios de la morbilidad leve o subclínica e inaparente con posibilidades de que se presente con características clínico - epidemiológicas, socio - económicas y etiológicas comunes, que coinciden

en un número importante de irregularidades de conducta, (impulsividad y/o agresividad), tendencia a dificultades en la socialización, insuficiente desarrollo de la atención voluntaria, probablemente manifiesta en hiperactividad, peor desarrollo escolar, mayor positividad en etapa de lactante, tendencia a positividad en la primera infancia y preescolar, signos positivos en la etapa escolar, probablemente más frecuente en varones, y en niños con factores socioeconómicos negativos, así como en niños con mayor número de antecedentes pre - peri - post natales.
Para dar solución al problema planteado se conforma la base con 208 casos evaluados cada uno en 18 rasgos predictores como se muestra en la siguiente descripción:

Rasgo	Valores Posibles
x_1 Edad.	6, 7 y 8 Años.
x_2 Escolaridad.	1. Preescolar, 2. P. grado, 3. S.Grado.
x_3 Sexo	1. Masculino, 2. Femenino.
x_4 Color de la piel	1. Blanca, 2.Mestiza, 3. Negra
x_5 Procedencia	1. Urbana, 2. Rural
x_6 Factores Socioeconómicos	1. Positivos, 2.Negativos
x_7 Antecedentes P. Familiares. APF	1. Ninguno, 2. Algunos, 3. Bastantes
x_8 " " Preperinatales. APPP.	1. Ninguno, 2. Algunos, 3. Bastantes
x_9 " " Perinatales. APP	1. Ninguno, 2. Algunos, 3. Bastan
x_{10} Conducta agresiva o impulsiva	No, Si
x_{11} Socialización	1. Adecuada, 2. Inadecuada
x_{12} Atención voluntaria	1. Predomina, 2. No predomina
x_{13} Afectividad	1. Placentero, 2. Indiferente
x_{14} Actividad	1. Normal, 2. Hiperactivo
x_{15} Desarrollo Escolar	1. Excelente, 2. Bien, 3. Insufuciente
x_{16} Positividad en etapa de lactancia	No, Si
x_{17} Positividad en etapa Pre-Perinatal	No, Si
x_{18} Positividad en etapa escolar	1. No, 2. Si

Además se establece como rasgo objetivo aquel que permite en función de los 18 rasgos predictores decidir el prediagnóstico de cada caso, para la base actual se le nombró Predigan, (y1), éste será el único rasgo objetivo de la base, quedando:

Predigan Prediagnóstico de TNL	1. Sanos, 2. Presuntivos, 3. Dañados

Con lo cual se completa la definición de los rasgos. Aparecerán en la bases de casos 18 variables o rasgos predictores y un solo rasgo objetivo o dependiente de los primeros.
Para facilitar el trabajo con los rasgos se definieron de forma tal que todos quedaran como variables ordinales excepto x_3, x_{10}, x_{16}, x_{17} y x_{18}, definidos como variables nominales. Al representar la bases de casos se obtuvo el resultado mostrado en la Tabla 3.1 del Anexo IV. Ella se edita con el TreeData para completar los nombre de

las variables, entrar los criterios de experto y organizar cada uno de los rasgos y casos de la base, o también adicionar, cambiar, etc, los valores de los que toman los rasgos de los diferentes casos.
Cuando el fichero que contiene a la base de conocimiento anterior es localizado y cargado con el TreeData, se oprime el botón Conectar de la ventana activa para que aparezca el nombre completo de la tabla que representa a esa base y en el área de la derecha; todos sus rasgos, destacándose el rasgo objetivo (Prediagnóstico) y cada una de los rasgos predictores como se muestra en la figura 3.8 . Cuando se acepta el rasgo objetivo se puede crear el árbol de decisiones.

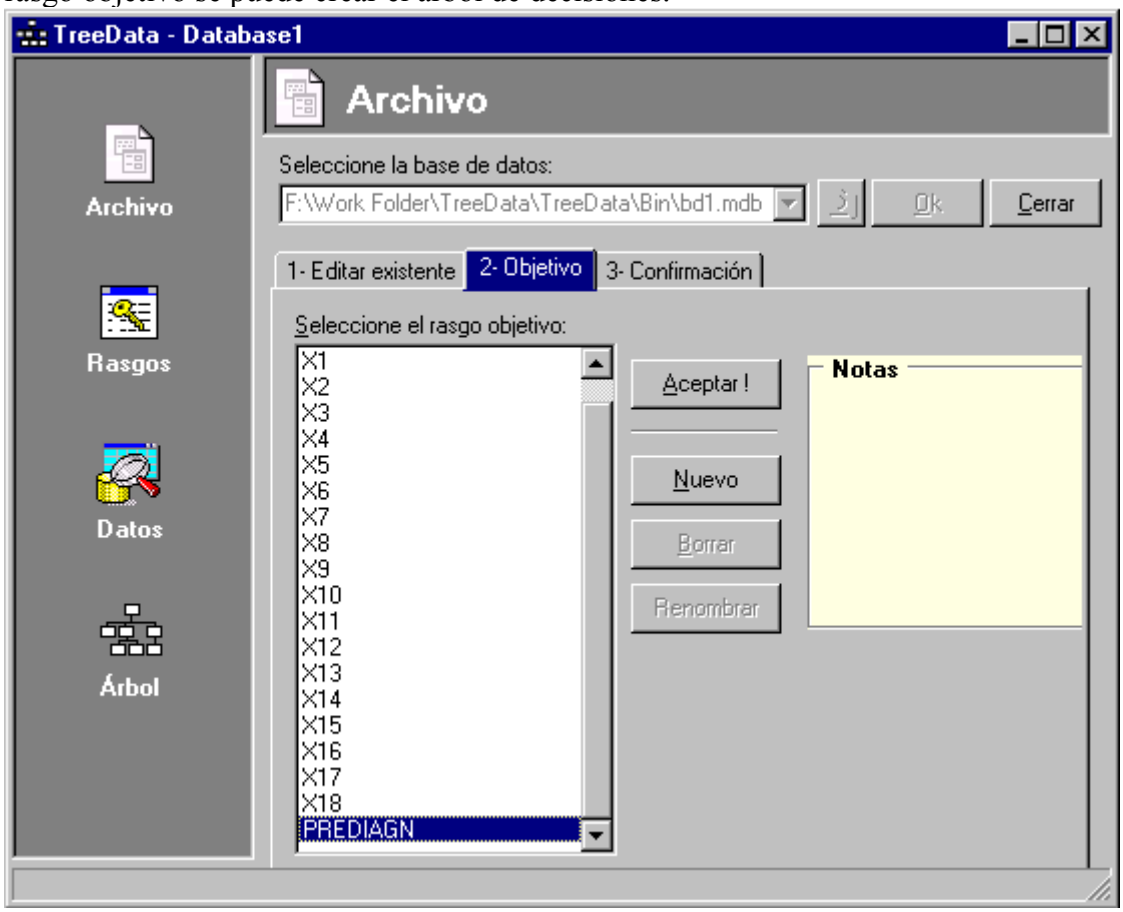


Figura 3.8. Rasgos de la base de casos bd1 con el rasgo objetivo seleccionado(Prediagn)

Por su nivel de asociación obtenido según las expresiones 2.13 y 2.14, los rasgos más predictores son:

- La escolaridad (x_2)
- La Edad (x_1)
- La Positividad en etapa de Lactancia (x_{16})
- La Positividad en la Etapa Pre-Perinatal (PPP,x_{17}).

- La Afectividad (x_{14}).

El resto no las mencionaremos aunque allí aparecen por tener un nivel de asociación bajo, lo cual las hace poco significativas.

III-3.1.1 Creación del Árbol de Decisiones. Cálculo de la primera Probabilidad.

El orden anterior en que fueron concebidos los rasgos es primordial en la creación del árbol de decisiones porque a partir de este es que se deben ir seleccionando los nodos cuando se utilice el método de creación manual, de otra forma no coincidirán los resultados de este método con los del automático.
Así, el prediagnóstico (Predigan) es el primer nodo del árbol de decisiones que se enlaza con el rasgo Escolaridad (x_2), como el más predictor y por cuyas categorías Preescolar (1) y Primer Grado (2) se enlazan: Edad(x_1), Positividad en etapa de lactancia(x_{16}) respectivamente quedando libre como primera hoja del árbol la categoría de Segundo Grado(3) del rasgo x_2. Se enlazan a demás a través de las categorías No y Si de x_{16} los rasgos Afectividad(x_{13}) y Positividad en Etapa Pre-Perinatal(x_{17}), respectivamente, los demás rasgos no se visualizan en el árbol de decisiones por no considerar su correlación con el objetivo como fuerte o media, sino de un bajo nivel de significancía. La Figura 3.7 a) y 3.7 b) del Anexo I y del Anexo II respectivamente muestran el árbol de decisiones obtenido con el TreeData.
Asociado a l árbol de decisiones se encuentran por ejemplo las tablas de contingencia mostradas en la Figura 3.9.

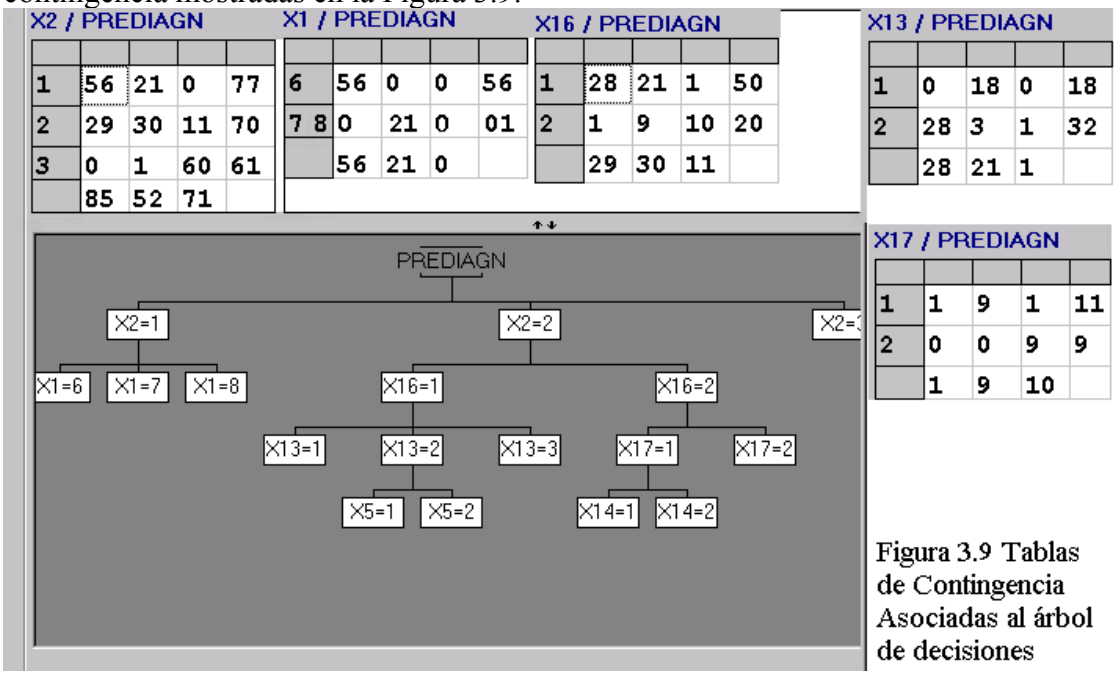


Figura 3.9 Tablas de Contingencia Asociadas al árbol de decisiones

Del árbol y las tablas asociadas a él se determinan las probabilidades de cada rasgo a través de la implementación de los criterios descritos en el Capítulo II.

A manera de ejemplo y basados en el enlace de los rasgos predictores con el objetivo y los valores de frecuencia mostrados en las tablas de contingencia, explicaremos cómo calcular la primera probabilidad para el rasgo *predictor* x_2 .

- Seleccionamos la categoría de enlace. En este caso como se trata del primer rasgo predictor las categorías de enlace serán las propias categorías del rasgo objetivo.
- Seleccionamos de la primera tabla de contingencia la primera celda. El valor de frecuencia para el rasgo x_2 en la categoría 1 del objetivo que ahí aparece (primera celda de la primera columna de la izquierda) dividido entre el total marginal de dicha columna nos da la primera probabilidad del rasgo.

$P_1\left(x_2 = 1 \middle| \Pr ediagn = 1(Sanos)\right) = \frac{56}{85} = 0{,}6588$ lo cual se interpreta como la probabilidad de que el rasgo x_2 tome valor 1(preescolar) y sea sano es 0,6588. Por otra parte, la Probabilidad de esté en primer grado y estar sano es:

$P_1\left(x_2 = 2 \middle| \Pr ediagn = 1(Sanos)\right) = \frac{29}{85} = 0{,}34$.

Para seguir ilustrando; si tomamos otro rasgo como por ejemplo x_{16}, la primera probabilidad se determina por:

$P_1\left(x_{16} = 1 \middle| x_2 = 2(\Pr imer\ grado), \Pr ediagn = 1(Sanos)\right) = \frac{28}{29} = 0{,}9655$ que se interpreta como la probabilidad de que el niño teniendo positividad en la etapa de lactancia y estando en primer grado sea sano es 0,9655 .

También de la misma forma, la probabilidad de que el niño, no teniendo positividad en la etapa de lactancia y estando en primer grado sea sano es:

$P_1\left(x_{16} = 2 \middle| x_2 = 2(\Pr imer\ grado), \Pr ediagn = 1(Sanos)\right) = \frac{1}{29} = 0{,}0345$

Para el resto de los rasgos que aparecen enlazados en el árbol de decisiones con el objetivo aún cuando no se correlacione directamente, sino a través alguna de las categorías de un rasgo determinado, TreeData calcula esas primeras probabilidades. Para los rasgos que no aparecen en el árbol de decisiones también calcula esta primera probabilidad pero esta vez utiliza la frecuencia de aparición de estas en la tabla original e ir completando la tabla cuyos resultados[5] se muestran en la figura 3,10 del Anexo V.

Para el cálculo de la segunda probabilidad P_2 el programa se basa en las expresiones fundamentadas en el Capítulo II de manera que, como el primer rasgo surge de una fuente que no hay por que dudar le asigna $P_2 = 1$, es decir $P_2\left(x_1(O_i) = 6\,años\right) = 1$, pero para otros como por ejemplo x_{10} y x_7 que se obtienen según el criterio de 3

[5] Solo aparecen los cálculos de los rasgos que caben en la ventana activa

expertos (todos con grado de experticidad 1) se obtiene utilizando la expresión 2.20 para los casos O_2 y O_{15} ejemplo $P_2(x_{10}(O_2) = Si) = 0{,}6667$ y $P_2(x_{10}(O_2) = No) = 0{,}3333$ y $P_2(x_{10}(O_{15}) = Si) = 1$ así sucesivamente. Los cálculos de esta probabilidad para todos los rasgos también fueron recogidos en la figura 3.10 del Anexo V.

III-3.1.2 Cálculo de la Certidumbre Total.

Las probabilidades calculadas no constituyen una medida de Certidumbre, por ello es que se le implemento al programa el cálculo de la Certidumbre total para cada caso usando la función $\xi_i(O_t)$ que se calcula combinando las dos probabilidades antes calculadas para cada rasgo según la expresión :

$\xi_i(O_t) = p_1(x_i(O_t)) + p_2(x_i(O_t)) - p_1(x_i(O_t)) * p_2(x_i(O_t))$

Es decir, se suman la primera y la segunda probabilidad de cada rasgo evaluada para cada caso y se le resta el producto de ambas de manera que para el rasgo x_2, por ejemplo evaluado en el caso O_{156} , esto es $x_2(O_{156})$, la primera probabilidad resulta: $P_1(x_2(O_{156}) = 1) = 0{,}2548$ y su segunda probabilidad $P_2(x_2(O_1) = 1) = 1$ por tanto aplicando la expresión que nos facilita el cálculo de la Certidumbre tenemos, $\xi_1(O_{156}) = 1$, y para cada uno de los rasgos de la tabla que tiene ambas probabilidades se procede de manera similar en todos los casos conformándose con ello la Figura 3.11 del Anexo VIII. Este anexo al igual que el anterior muestran solamente una parte de la base[5] de casos con los valores de certidumbre calculados.

Recuperación de Casos Semejantes.

Para realizar la recuperación se salvan en la base los cálculos de certidumbre efectuados, luego se oprime el botón que permite efectuar la recuperación, se visualizará entonces la Figura 3.12 del Anexo IX. Donde se entra el nuevo caso O_o, los pesos de las variables y el valor umbral de comparación. El programa utiliza entonces la función de semejanza descrita en la fundamentación matemática del Capítulo II para la identificación de los casos parecidos a O_o. Al calcular la función de semejanza para cada caso de la base comparada con el nuevo caso introducido se obtienen los resultados mostrados en la Figura 3.13 del anexo X, donde aparecerán los casos más semejantes a Oo resaltados sobre un fondo de color diferente a los que no son semejantes.

Conclusiones

Se abordó la problemática del uso de la de la teoría de las probabilidades en la toma de decisiones bajo el manejo de la incertidumbre en un Sistema Basado en Casos. Desde la perspectiva dicha teoría probabilística se logró crear un modelo teórico para el tratamiento probabilístico de tales Sistemas Basados en el Conocimiento, así como, se realizó su fundamentación matemática, el que sirvió de guía para el diseño e implementación del Software TreeData.

En las pruebas del Software TreeData fueron usadas bases de datos teóricas ya creadas con otros sistemas y bases de casos reales, con las que se obtuvieron buenos resultados, en cuanto a la eficiencia temporal en la recuperación.

La fundamentación e implementación logradas permiten, la asignación de un valor de certidumbre a los valores de los rasgos predictores y objetivos en una base de casos.

El modelo planteado, hace que se reduzcan ostensiblemente el número de comparaciones en el proceso de recuperación y que se facilite de forma significativa el proceso de adaptación.

El programa TreeData posee dentro de sus componentes fundamentales:

- TtreeStructView que no existía otra similar a ella por lo que constituye un aporte desde el punto de vista de la programación, dadas sus posibilidades en la construcción,visualización y recorrido del árbol de decisiones.
- La clase TDModCommon que maneja la conexión con la base de datos garantizando acceso a la misma y su organización. Los demás módulos, a través de las interfaces que ella exporta acceden a los datos en la base.
- La TFormTreeData que genera el árbol de decisiones para su posterior acceso aleatorio, cosa antes no conocida en otras aplicaciones (al menos en la forma en que lo hace este programa).

Las bases de casos con las que trabaja el programa pueden ser creadas en Microsoft Access, Microsoft Excel, en el Procesador estadístico SPSS o con el propio programa, lo cual demuestra su amplia interacción con otras aplicaciones.

Recomendaciones

El software elaborado puede seguir siendo perfeccionado sobre todo en la parte de la recuperación de los casos y recálculo de todas las magnitudes que intervienen en una base de casos típica. Además, durante el desarrollo y al final del trabajo se detectaron un grupo de aspectos que favorecen la continuidad del mismo y puede convertirlo en una herramienta más poderosa para el tratamiento probabilístico de los Sistemas Basados en Casos bajo Incertidumbre, por lo que se recomienda:

1. Realizar otras aplicaciones que permitan una mejor validación del sistema.
2. Comparar los resultados obtenidos con el TreeData con los obtenidos en otros sistemas que usen la tecnología del Razonamiento Basado en Casos.
3. Desarrollar implementaciones de nuevos algoritmos que utilicen otras Co-T Normas para el cálculo de la Incertidumbre que no fueron desarrolladas en el presente trabajo, así como desarrollar otros modelos para el manejo de la misma, basados en la estructura de organización desarrollada.
4. Unificar en un solo software los distintos productos utilizados en el desarrollo de este trabajo.
5. Continuar desarrollando la implementación del módulo de recuperación para este tipo de Sistemas y desarrollar el módulo de adaptación.

Referencias Bibliográficas

[Aha96] *Aha David W., Chang Li wu. " Cooperative Bayesian and Cased – Based Reasoning for Solving Multiagent Planning Tasks "*, Novy Center for Applied Research AI, Enero, 1996.

[AI-] AI-CBR . http://www.salford.ac.uk/survey/staff/IWatson/cbr01.htm.

[Alt95] Althoff, K.D. Evaluating Case-Based Reasoning Systems. In: Proc. Workshop on Case Based Reasoning: A New Force In Advanced Systems Development, pp 48-61, published by Unicom Seminars Ltd, Brunel Science Park, Cleveland Road, Uxbridge, Middlesex UB8 3P, UK, 1995.

[And89] Andersen, S. K., Olesen, K. G., Jensen, F. V., and Jensen, F. HUGIN- a shell for building Bayesian belief universes for expert systems, In proceedings of the Eleventh international Join Conference of Artificial Intelligence (IJCAI-89) volumne 2, pp 1080-1085., Detroit, Michigan Kaufmann, 1989.

[Bon85] Bonnisone P.P and P.M. Tong. Reasoning with uncertainty in Expert Systems. Int. Journal "Man Machine Studies". Vol22, pag 241-250, 1985.

[Bos93] Bosc, P. and Henry Prade. An introduction to fuzzy set and possibility theory based approaches to the treatment of uncertainty and imprecision in database management systems. Proc of second workshop uncertainty management in information system: from needs to solutions, California, 1993.

[Bre92] Breese Jhon S., " Construction of belief and Decision networks ". Computational Intelligence. Vol.8,No. 4, 1992.

[Bre95] Breese Jhon S., Heckerman D..” Decision – Theoretic Case – Based Reasoning “. IEEE, Vol. 26, No.6, Nov. 1995.

[Buc84] Buchanan B. G., Shortliffe E.H. Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic ProgrammingProjet. Adison – Wesley, Reading, MA. 1984

[Bux89] Buxton, R. Modelling uncertainty in expert systems. Informational Journal of Man-Machine Studies, 31, pp 415-476, 1989.

[Cas93] Case-Based Reasoning. http://www.ils.nwu.edu/~e_for_e/nodes//NODE-23-pg.html, 1993

[Case] Case-Based Reasoning in the Web. http://wwwagr.informatik.uni-kl.de/~lsaCBR/CBR-Homepage.html.

[Cas89] Castillo Enrique, Albarez E. “ Sistemas Expertos, Aprendizaje e Incertidumbre “,.1994.

[Cas96] Castillo Enrique, Gutiérrez J. Manuel, Hadi Ali S. “ Expert Systems and Probabilistics Network Models ”, 1996.

[Cat91] Catlett, J. On changing continuous attributes into ordered discrete attributes. Proc, of the Machine Learning- EWSL-91. European Working Session on Learning, pp 164-178, 1991.

[Cha94] “ CHAID para SPSS sobre Windows . Técnicas de segmentación basadas en razones de verosimilitud Chi-cuadrado “, Manual de usuario, SPSS Soft. Inc. 1994.

[Che92] Cheeseman, P., Hwang C-L., Fuzzy Multiple Attribute Decision Making, Methods and Applications, Springer Verlag, 1992.

[Che85] Cheeseman, P. In defense of probability. In Proceeding of the Ninth international Join Conference on Artificial Intelligence (IJCAI-85), pp 1002-1009, Los Angeles, California, Morgan Kaufmann, 1985.

[Chi92] Chmielewski, M. R., Grzymala-Busse, J.W. Global discretization of continuos attributes as preprocesing for inductive learning. Department of Computer Science, University of Kansas, TR-92-7, 1992.

[Chr92] Christiansen, A. D. Learning to predict in uncertain continuous task. Proc.

Of the 9th In. Workshop on Machine Learning, pp 72-81, 1992.

[Dem68] Dempdter A. P. A Generalization of Bayesian inference. Journal of Royal Statistical Society, 30 (Series B), pp 205-247, 1968.

[Dia99] Díaz de Villegas Á., "Tesis de Maestría",1999.

[Dud80] Duda R. O., Hart P.E. Model Design in the PROSPECTOR Consultant System for Mineral Exploration. In Michie, D., editor, Expert System in the Microelectronic Age. Edinburgh University Press, 153-167.1980

[Dut85] Dutta, A. Reasoning with imprecise knowledge in expert systems. Information Sciences, pp 37:3-24, 1985.

[Etx98] Etxeberria, R., Larranaga P., Picaza J.M. " Analysis of the behaviour of genetic algorithms when learning Bayesian network structure from data " Pattern Recognition Lett. Vol 18. 11-13, 1998.

[Ewa91] Edwards John S." Building Knowledge – Base Systems ". 1991.

[Fig94] Figueroa, J.M. Guía BIP/ISO para la expresión de la incertidumbre en las mediciones. Centro Nacional de Metrología. Mexico 1994.

[Fre87] Freznel Louis E. " Crash Course in Artificial Intelligence and Expert System ".Indianapolis,IN:Howard W. Sams, 1987.

[Fer93] Ferting K. W., Breese J. S. " Probability Intervals Over Influence Diagrams ". IEEE Vol.15, No.3, Marzo 1993.

[Fun95] Fung Robert, Del Favero B. " Applying Bayesian Networks to Information Retrieval ", ACM, Vol. 38, No. 3, Marzo 1995.

[Gar92] García,B. Razonamiento Basado en Casos.

http://www-cia.mty.itesm.mx/~bbgarcia/

[Gar90] García Luciano " Probabilidad e Inteligencia Artificial ".

[Gar94] Gardaati, S., Razonamiento Basado en Casos, Soluciones Avanzadas, Año 2, No. 13, Setiembre, pp 60-63, 1994.

[Gra91] Grady B. Object Oriented Design: with Aplications, The Benjamín CummingsPublishing Company inc. U.S.A, 1991.

[Grz91] Grzymala-Busse, J.W. On the unknown attribute values in learning from examples. Proc. Of the ISMIS-91, 6th International Symposium on Methodologies for Intelligent Systems, 368-377, Lecture Notes in Artificial Intelligence, vol. 542, Springer-Verlag, 1991.

[Grz88] Grzymala-Busse, J. W. Knowledge acquisition under uncertainty- a rough set approach. Journal of Intelligent & Robotic Systems 1, pp 3-16, 1988.

[Grz91] Grzymala-Busse, J. W. Managing Uncertainty in Expert Systems. Kluwr

Academic Publishers 1991.

[Har95] Harmanec, D. Toward a Characterization of Uncertainty Measure for the Dempster-Shafer Theory, pp 255-261. Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence (UAI'95) Montréal, PQ, Canada, August 1995.Morgan Kaufmann Publishers, 1995.

[Hwa79] Hwang, C.L., K. Yoon, Multiple Attribute Decision Making Methods and Applications: A State-of-the-Art Survey, Springer-Verlag, NY, 1979.

[Hec95] Heckerman David, Breese J., Romnelse K. " Decision Theoretic : Troubleshooting ". ACM, Vol. 38, No. 3, Marzo 1995.

[Job92] Jobson, J. D. " Applied Multivariate Data Analysis ". Vol. II: Categorical and Multivariate Methods, Springer, New York, 11-54, 1992.

[Kli95] Klir, G.J. Principles of uncertainty: What are they? Why do we need them?. Fuzzy sets and systems 74, page 15-31, 1995.

[Kli98] Klir G.J.and M.J. Wierman. Uncertainty-based information. Physica-Verlag, 1998, pag 168.

[Kol89] Kolodner, J. Judging which is the best case for a casebased reasoner. In Proceedinngs of Case Based Reaoning Workshop, 1989.

[Kol92] Kolodner, J. C., An Introduction to Case-Based Reasoning, Artificial Intelligence. Review 6, pp 3-34, 1992.

[Kol93] Kolodner, J.L., Case–Based Reasoning, Morgan Kaufmann Publishers, Inc., San Mateo, CA, 1993.

[Kon99] Kontkanen P., Millymaki P., Silander T., Tirri H., Voltonen K. Exploring the Robustness of Bayesian and Information-Theoretic Methods of Predictive Inference. The Seventh International Workshop on Artificial Intelligence and Statistics, edited by D. Heckerman and J. Whiltaker. Morgan Kaufmann Publishers, 1999

[Lam98] Lam W. " Bayesian network refinement via machine learning approach ". IEEE Trans Patt Anal Mach Int. Vol 20, No.3, 240-251, 1998.

[Lea] Leake, D.B., and Wilson, D.C., 1998 Categorizing Case-Based Maintenance: Dimensions and Directions. In, Advances in Case-Based Reasoning, Smyth, B., and Cunningham, P. (Eds.), pp.196-207. Springer-Verlag, Lecture Notes in AI #1488, Berlin. http://www.cs.indiana.edu/~leake/papers/p-98-03.pdf

[Mar98] Mardia K.V., Baczkowski A.J., Feng X., Hainsworth T.J. " Statistical methods for automatic interpretation or digitally sacanned finger prints ". Pattern Recognition Lett. Vol 18, 11-13, 1998.

[Mar91] Marques de Cantú, M. J. Probabilidades y Estadística para ciencias Químico_Biológicas. Editorial McGraw-Hill Interamericana de México, S.A. 1991.

[McC80] McCarthy, J. Circumscription: a form of non-monotonic reasoning. Artificial Intelligence, 13 (1-2), pp 27-39, 1980.

[McD80] MCDermott, D and Doyle, J. non_monotonic logic I. Artificial Intrelligence, 13(1.2), pp 41-72,19980.

[Myl93] Myllimaki P., Tirri H. " Massively parallel case-based reasoning with probabilistic similarity metrics ". Proceedings of the First European Workshop on Case-Based Reasoning. 145-154, 1993.

[Par96] Parson S. Current approaches to handling imperfect information in data and knowledge bases. IEEE trans. On Knowledge and data engineering, Vol 8, No3, June 96.

[Paw82] Pawlak, Z. Rough sets. Int. J. Computer and Information SCI. 11, pp 341-356, 1982

[Paw91] Pawlak, Z. Rough sets. Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers 1991.

[Paw95a] Pawlak, Z. Vagueness and uncertainty: a rough set perspective. Computational Intelligence, Vol 11, No 2, 1995.

[Paw95b] Pawlak, Z. Rough sets. Comm of ACM, vol 38, No 11, 1995

[Pea88] Pearl J. " Probabilistic Reasoning in Intelligent System ". San Mateo, CA: Morgan Kaufman, 1988.

[Reg93] Regazzoni C.S., Murino V., Vernazza G. " Distributed propagation of a priori constraints in Bayesian Network of Markov random fields ". IEEE Vol. 40, No. 1, Febrero 1993.

[Rei80] Reiter, R. A logic for default reasoning. Artificial Intelligence 13 (1-2), pp 81-132, 1980.

[Rod97] Rodríguez Andrés F, Vadera S. and Sucar L.E. A probabilistic model for Case Based Reasoning. University of Salford. Department of Mathematics and Computer Science. 1997.

[Sha76] Shafer, G. A Mathematical Theory of Evidence. Princeton University Press, Princeton. New Jersey, 1976.

[Sha90] Shapiro, C., John Wiley, Encyclopedia of A, vol I, 1990

[Sho73] Shortliffe, E. H., Axline, S.G., Buchanan B. G., Merigan, T.C. y Cohen, S.N. An Artificial intelligence program to advise physician regarding antimicrobial therapy. Comput. Biomed. Res.,Vol. 6,544,1973.

[Sho75] Shortliffe, E. H. y Buchanan, B. G. A model of inexact reasoning in Medicine, Mathematical Bioscience, Vol. 23, pp 351-374, 1975

[Sho76] Shortliffe, E. H., Computer-Based Medical Consultations: MYCIN. Elsevier/North-Holland, Amsterdam, London, New York,1976.

[Smy98] Smyth P. " Belief networks, hidden Markovs models, and Markov random fields: A unifying view ". Pattern Recognition Lett. Vol 18, 11-13, 1998.

[Stu96] Stuart Rusell, Naving P. "Inteligencia Artificial: Un enfoque Moderno". 1996.

[Tir96] Tirri H., Kontkanen P. and Myllymaki P. Probabilistic Instance-Based Learning. Machine Learning: Proceedings of the Thirteenth International Conference edited by L. Sailla. Morgan Kaufmann Publishers, San Francisco, CA, 1996.

[Tur92] Turban, E., Expert Systems and Applied Artificial Intelligence, Macmillan Publishing, 1992.

[Wal96] Walley, P. Measures of uncertainty in expert systems. Artificial Intelligence, 83, pp 1-58, 1996.

[Wat] Watson, Ian. Case-Based Reasoning Development Tools: A Review. http://www.cs.tulane.edu/www/cbr/papers/Watson.html.

[Wil98] Williams W.L., Wilson,R.C., HancockE.R. " Multiple graph matching with Bayesian inference ". Pattern Recognition Lett. Vol 38, 11-13, 1998

[Zad65] Zadeh L. A. Fuzzy sets. Information and Control. 8, pp 338-353, 1965.

[Zad78] Zadeh L. A. Fuzzy sets as a basis for the theory of possibility. Fuzzy Sets and Systems, 1. pp 3-28.1978.

[Zad83] Zadeh, L.A. The role of fuzzy logic in the management of uncertainty in Expert Systems, Vol 11, pp 199-227, 1983.

[Zam97] Zamora L., " Tesis de Maestria ", UCLV, 1997.

Anexo II

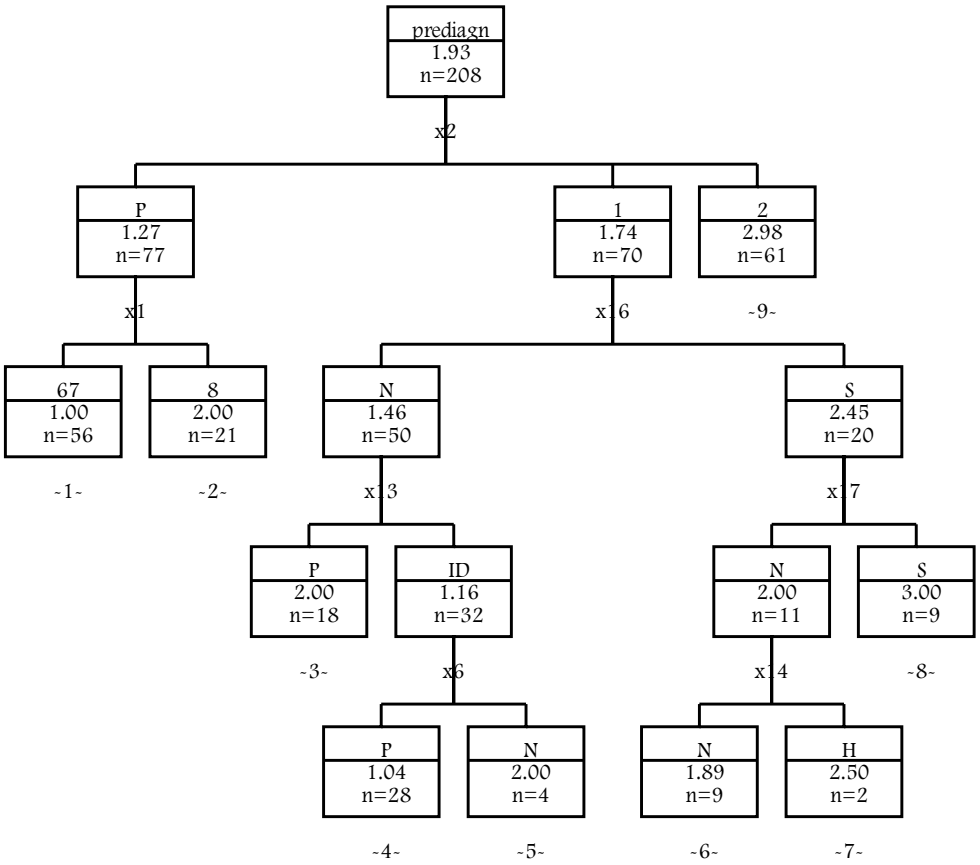


Figura 3.7 b, árbol de decisiones resumen completo con todos los rasgos enlazados.

Anexo I

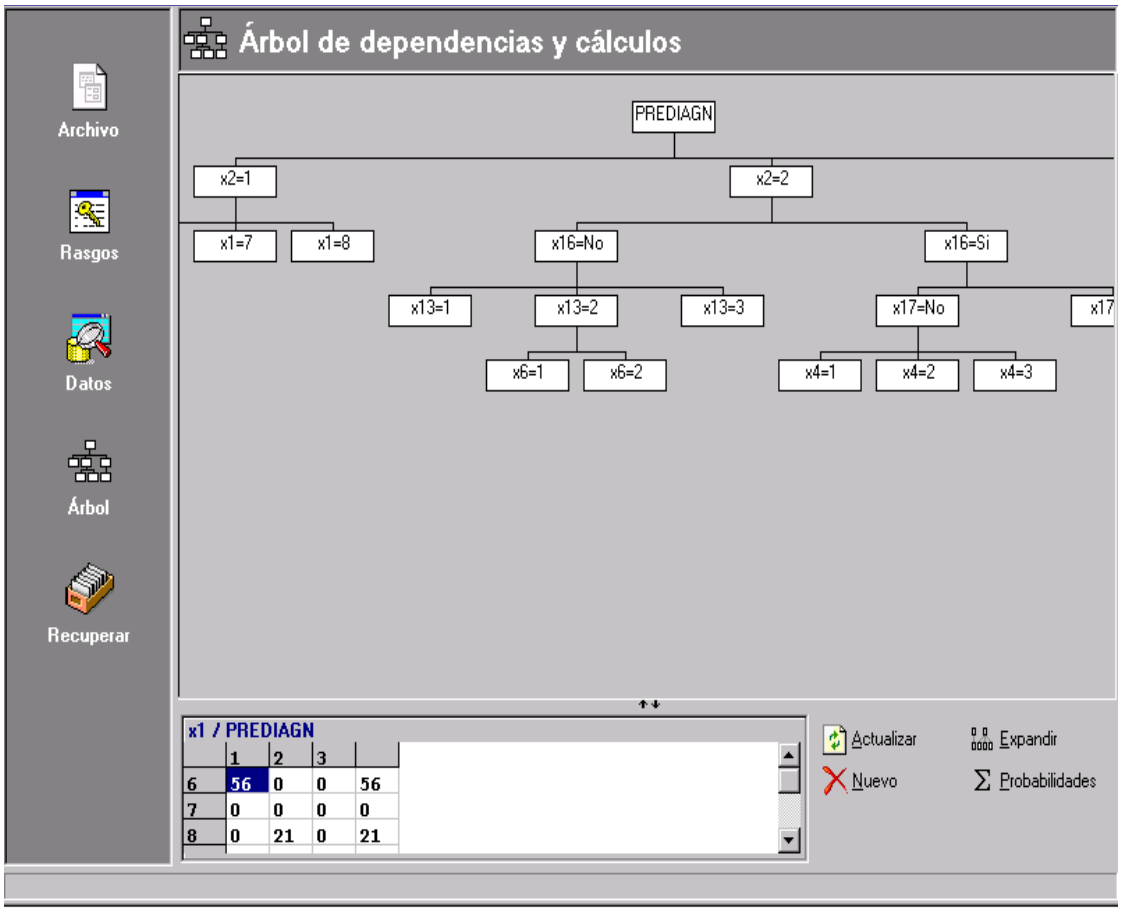
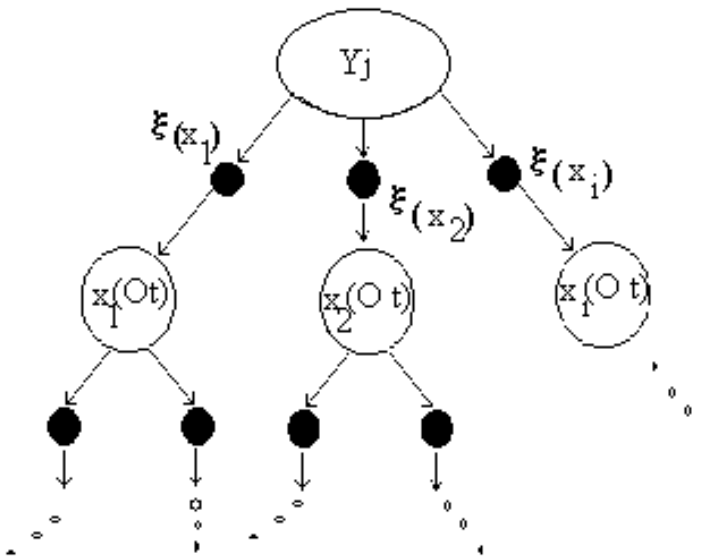


Figura 3.7 a, árbol de decisiones obtenido con el TreeData más tabla de contingencia que asocia al rasgo x_{13} contra el Predigan.

Anexo III

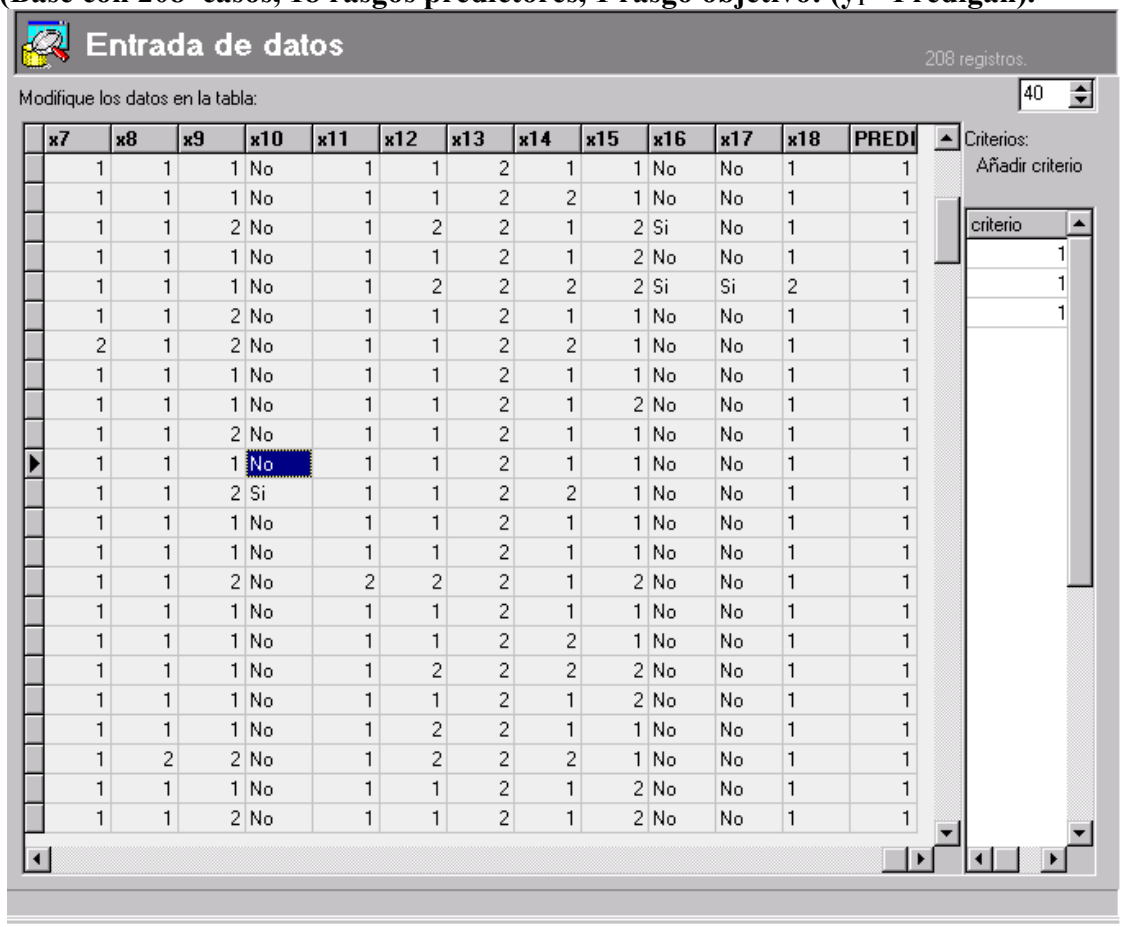


árbol de decisiones

fig. 2.1

Anexo IV

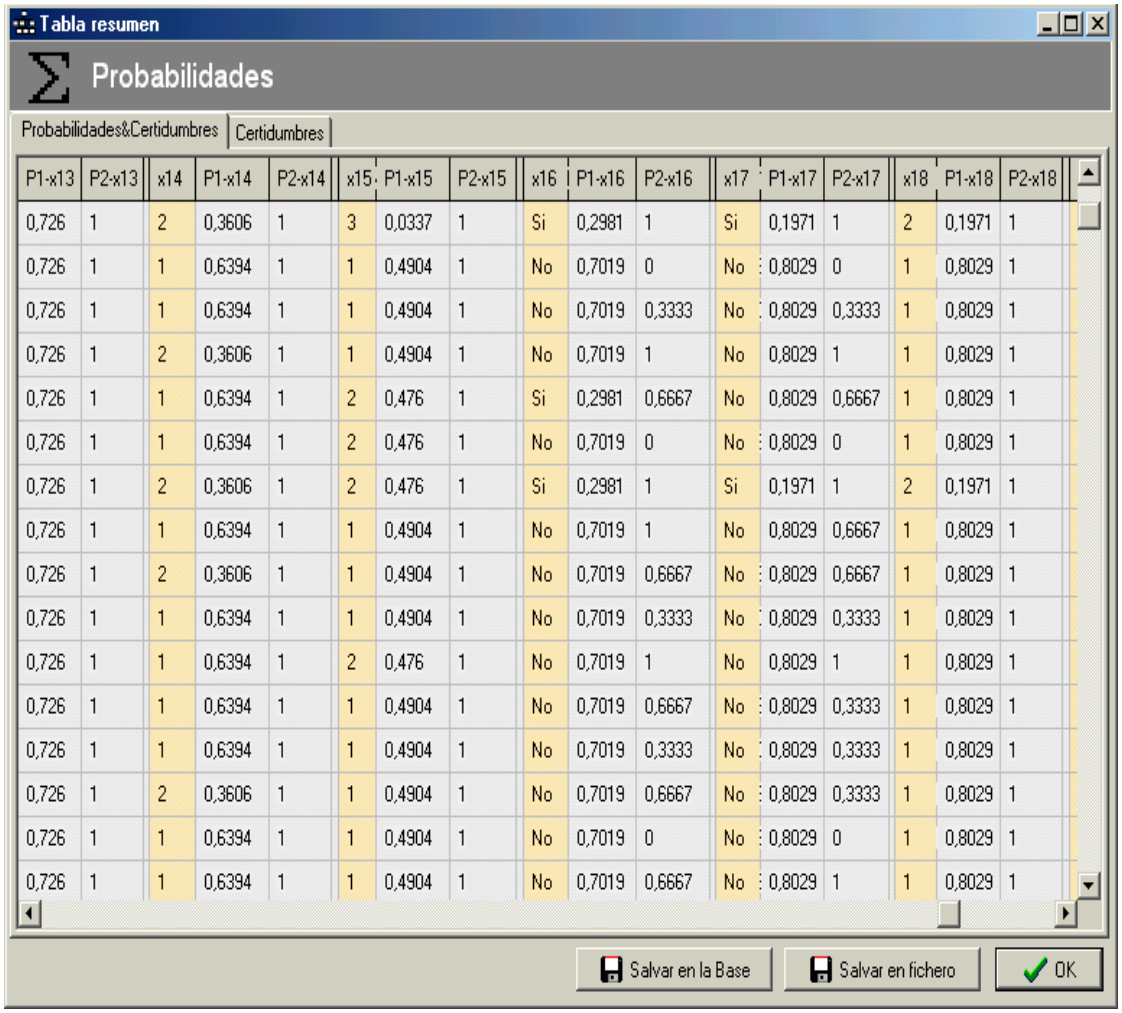
(Base con 208 casos, 18 rasgos predictores, 1 rasgo objetivo: (y_1= Predigan).



x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18	PREDI
1	1	1	No	1	1	2	1	1	No	No	1	1
1	1	1	No	1	1	2	2	1	No	No	1	1
1	1	2	No	1	2	2	1	2	Si	No	1	1
1	1	1	No	1	1	2	1	2	No	No	1	1
1	1	1	No	1	2	2	2	2	Si	Si	2	1
1	1	2	No	1	1	2	1	1	No	No	1	1
2	1	2	No	1	1	2	2	1	No	No	1	1
1	1	1	No	1	1	2	1	1	No	No	1	1
1	1	1	No	1	1	2	1	2	No	No	1	1
1	1	2	No	1	1	2	1	1	No	No	1	1
1	1	1	No	1	1	2	1	1	No	No	1	1
1	1	2	Si	1	1	2	2	1	No	No	1	1
1	1	1	No	1	1	2	1	1	No	No	1	1
1	1	1	No	1	1	2	1	1	No	No	1	1
1	1	2	No	2	2	2	1	2	No	No	1	1
1	1	1	No	1	1	2	1	1	No	No	1	1
1	1	1	No	1	1	2	2	1	No	No	1	1
1	1	1	No	1	2	2	2	2	No	No	1	1
1	1	1	No	1	1	2	1	2	No	No	1	1
1	1	1	No	1	2	2	1	1	No	No	1	1
1	2	2	No	1	2	2	2	1	No	No	1	1
1	1	1	No	1	1	2	1	2	No	No	1	1
1	1	2	No	1	1	2	1	2	No	No	1	1

Figura 3.10 Base de casos editada con el TreeData para completar.

Anexo V



P1-x13	P2-x13	x14	P1-x14	P2-x14	x15	P1-x15	P2-x15	x16	P1-x16	P2-x16	x17	P1-x17	P2-x17	x18	P1-x18	P2-x18
0,726	1	2	0,3606	1	3	0,0337	1	Si	0,2981	1	Si	0,1971	1	2	0,1971	1
0,726	1	1	0,6394	1	1	0,4904	1	No	0,7019	0	No	0,8029	0	1	0,8029	1
0,726	1	1	0,6394	1	1	0,4904	1	No	0,7019	0,3333	No	0,8029	0,3333	1	0,8029	1
0,726	1	2	0,3606	1	1	0,4904	1	No	0,7019	1	No	0,8029	1	1	0,8029	1
0,726	1	1	0,6394	1	2	0,476	1	Si	0,2981	0,6667	No	0,8029	0,6667	1	0,8029	1
0,726	1	1	0,6394	1	2	0,476	1	No	0,7019	0	No	0,8029	0	1	0,8029	1
0,726	1	2	0,3606	1	2	0,476	1	Si	0,2981	1	Si	0,1971	1	2	0,1971	1
0,726	1	1	0,6394	1	1	0,4904	1	No	0,7019	1	No	0,8029	0,6667	1	0,8029	1
0,726	1	2	0,3606	1	1	0,4904	1	No	0,7019	0,6667	No	0,8029	0,6667	1	0,8029	1
0,726	1	1	0,6394	1	1	0,4904	1	No	0,7019	0,3333	No	0,8029	0,3333	1	0,8029	1
0,726	1	1	0,6394	1	2	0,476	1	No	0,7019	1	No	0,8029	1	1	0,8029	1
0,726	1	1	0,6394	1	1	0,4904	1	No	0,7019	0,6667	No	0,8029	0,3333	1	0,8029	1
0,726	1	1	0,6394	1	1	0,4904	1	No	0,7019	0,3333	No	0,8029	0,3333	1	0,8029	1
0,726	1	2	0,3606	1	1	0,4904	1	No	0,7019	0,6667	No	0,8029	0,3333	1	0,8029	1
0,726	1	1	0,6394	1	1	0,4904	1	No	0,7019	0	No	0,8029	0	1	0,8029	1
0,726	1	1	0,6394	1	1	0,4904	1	No	0,7019	0,6667	No	0,8029	1	1	0,8029	1

Figura 3.10 Primera y Segunda probabilidades de los rasgos.

Anexo VII

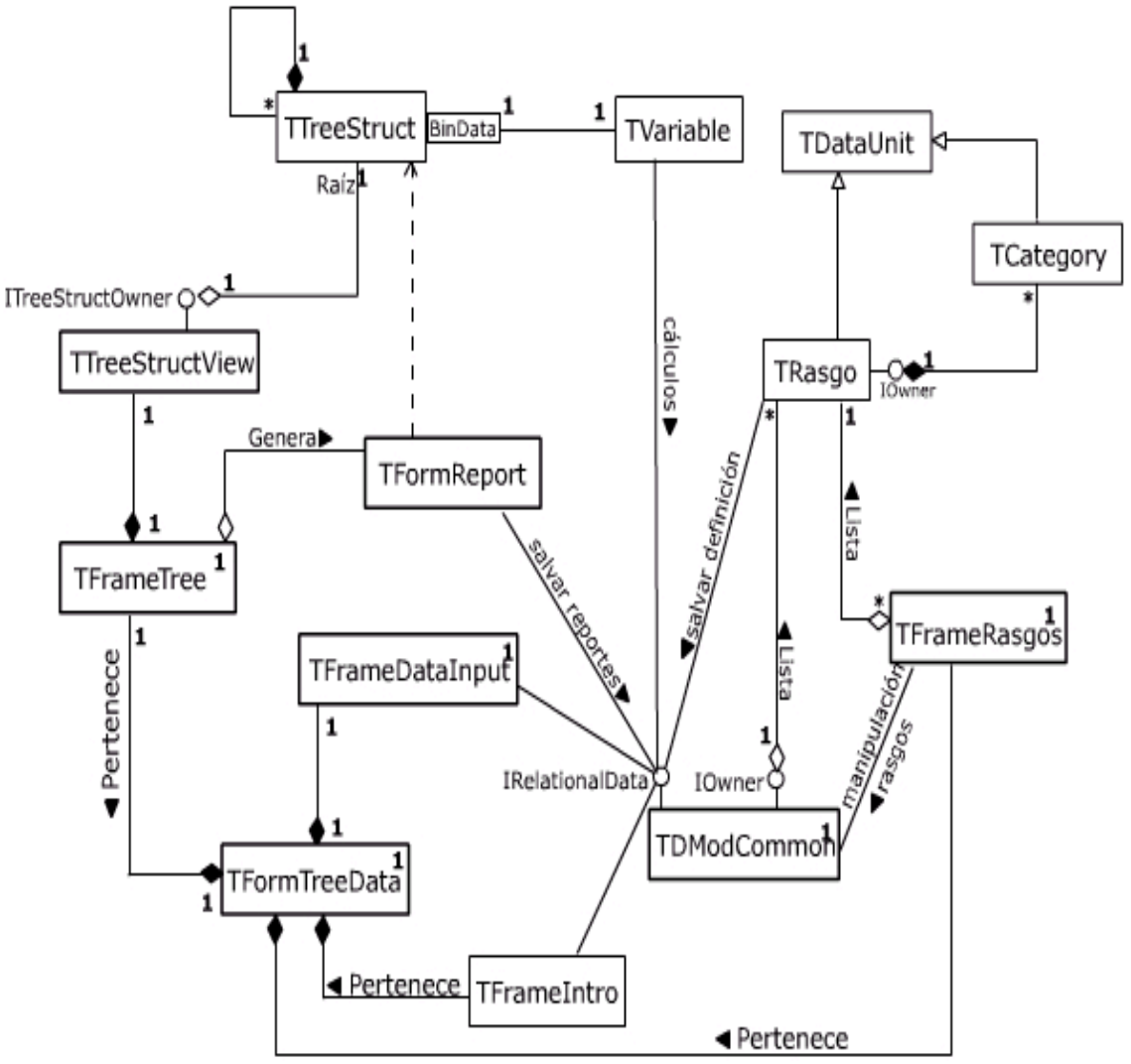
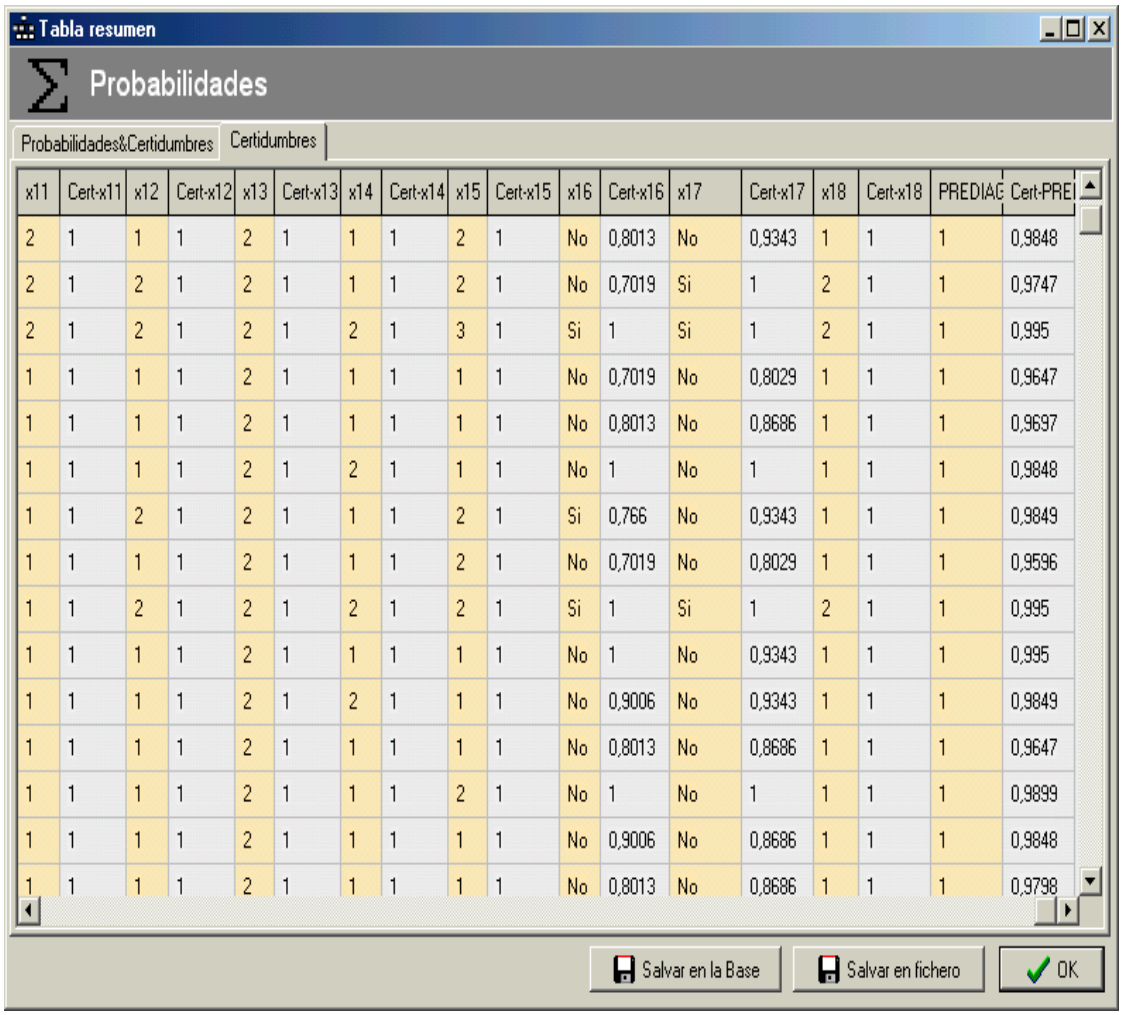


Figura 3.1a. Diagrama de clases según el Lenguaje Unificado de Programación, UML.

Anexo VIII



x11	Cert-x11	x12	Cert-x12	x13	Cert-x13	x14	Cert-x14	x15	Cert-x15	x16	Cert-x16	x17	Cert-x17	x18	Cert-x18	PREDIAG	Cert-PRE
2	1	1	1	2	1	1	1	2	1	No	0,8013	No	0,9343	1	1	1	0,9848
2	1	2	1	2	1	1	1	2	1	No	0,7019	Si	1	2	1	1	0,9747
2	1	2	1	2	1	2	1	3	1	Si	1	Si	1	2	1	1	0,995
1	1	1	1	2	1	1	1	1	1	No	0,7019	No	0,8029	1	1	1	0,9647
1	1	1	1	2	1	1	1	1	1	No	0,8013	No	0,8686	1	1	1	0,9697
1	1	1	1	2	1	2	1	1	1	No	1	No	1	1	1	1	0,9848
1	1	2	1	2	1	1	1	2	1	Si	0,766	No	0,9343	1	1	1	0,9849
1	1	1	1	2	1	1	1	2	1	No	0,7019	No	0,8029	1	1	1	0,9596
1	1	2	1	2	1	2	1	2	1	Si	1	Si	1	2	1	1	0,995
1	1	1	1	2	1	1	1	1	1	No	1	No	0,9343	1	1	1	0,995
1	1	1	1	2	1	2	1	1	1	No	0,9006	No	0,9343	1	1	1	0,9849
1	1	1	1	2	1	1	1	1	1	No	0,8013	No	0,8686	1	1	1	0,9647
1	1	1	1	2	1	1	1	2	1	No	1	No	1	1	1	1	0,9899
1	1	1	1	2	1	1	1	1	1	No	0,9006	No	0,8686	1	1	1	0,9848
1	1	1	1	2	1	1	1	1	1	No	0,8013	No	0,8686	1	1	1	0,9798

Figura 3.11 Certidumbre de cada caso.

Anexo IX

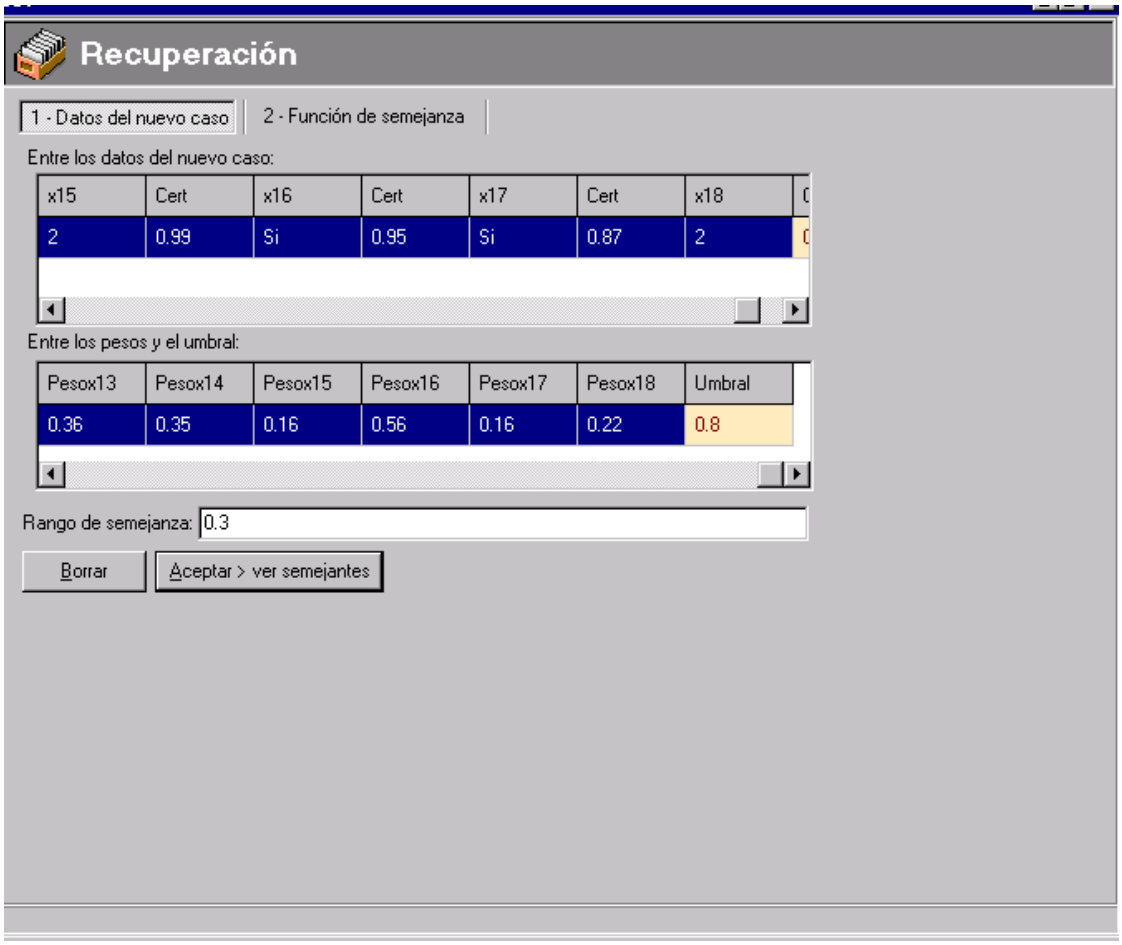


Figura 3.12 Nuevo problema para la recuperación.

Anexo X

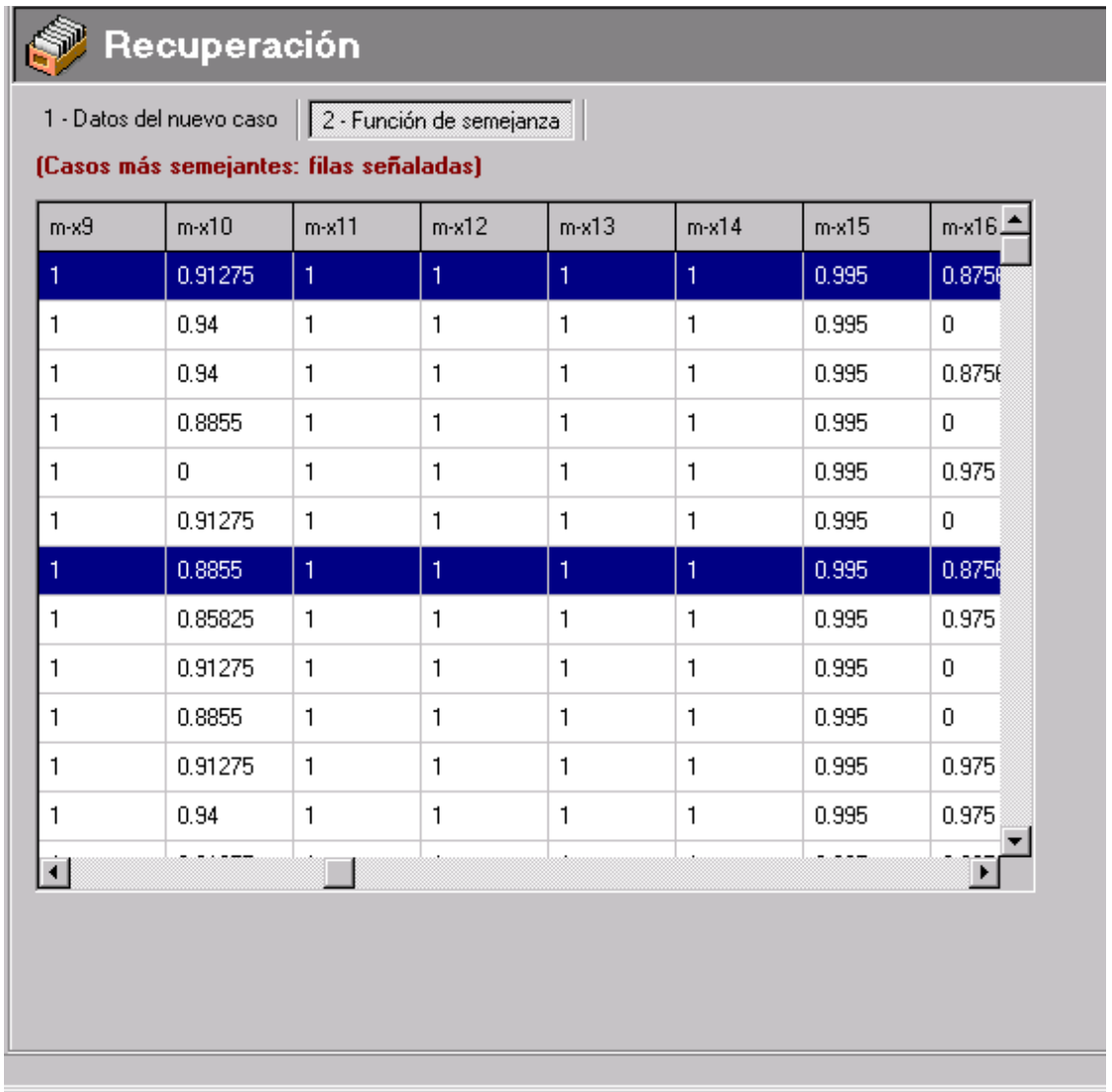


Figura 3.13. Casos recuperados en la Base similares al nuevo problema.