TESIS DOCTORAL

MODELO DE RESOLUCION DE INCERTIDUMBRE EN SISTEMAS EXPERTOS A PARTIR DE LA TEORIA DE LA POSIBILIDAD

por

Julio García del Real Ruizdelgado
Ingeniero de Caminos por la Universidad Politecnica de Madrid

Presentada en la

FACULTAD DE INFORMATICA

de la

UNIVERSIDAD POLITECNICA DE MADRID

para la obtención del

Grado de Doctor en Informática

MADRID , JULIO DE 1988

UNIVERSIDAD POLITECNICA DE MADRID - FACULTAD DE INFORMATICA - BIBLIOTECA

FACULTAD DE INFORMATICA · BIBLIOTECA

DEVUELVA este libro antes de la última fecha anotada.

1000193919
T-54

R. 54

TESIS DOCTORAL

MODELO DE RESOLUCION DE INCERTIDUMBRE EN SISTEMAS EXPERTOS A PARTIR DE LA TEORIA DE LA POSIBILIDAD

Presentada en la

FACULTAD DE INFORMATICA

de la

UNIVERSIDAD POLITECNICA DE MADRID

para la obtención del

GRADO DE DOCTOR EN INFORMATICA

AUTOR: D. Julio García del Real Ruizdelgado

Ingeniero de Caminos por la Universidad Politécnica de Madrid.

DIRECTORES: D. Luis Maté Hernández

D. Rafael Gonzalo Molina

MADRID, JULIO 1988

ABSTRACT.

This thesis has two different parts. In the first one, a thoroughly description of the principal systems of uncertainty is studied, taking account basically the following five models:

- The MYCIN model which allows an specific treatment of uncertainty within the Experts Systems; this is not a probabilistic convencional model,

- then the PROSPECTOR model, of specific application in the probabilistic model,

- the DEMPSTER-SHAFER model, when there is overlaping,

- the Zadeh's Possibility Theory, which allows the simultaneous treatment of of uncertainty and imprecision by means of fuzzy sets

- and finally, the possibility and neccesity measures of Dubois and Prade, as an extension of Zadeh's Theory.

In the second part, after having selected Zadeh's Possibility Theory as the best model for uncertainty problems with no real solution, a new model is proposed, based on Dubois-Prade model, with simultaneous treatment of uncertainty and fuzziness.

The main interest of this work lies on the fact that it takes into account old and desused theories as the problems are not well defined. Therefore a different alternative is offered to solve this question.

INDICE

CAPITULO 1.- INTRODUCCION.

CAPITULO 2.- PRINCIPALES MODELOS EXISTENTES.

CAPITULO 1 :

INTRODUCCION

Uno de los campos más específicos de la investigación dentro de la Inteligencia Artificial ha sido y sigue siendo el que se conoce como "Razonamiento Aproximado". Principalmente en los sistemas de ayuda a la decisión y de diagnóstico, ampliamente desarrollados, se plantean problemas cuando la información de que se dispone es parcial, incompleta o aproximada.

Las deficiencias de un cuerpo de conocimiento incompleto o mal estructurado pueden ser de muy diversos tipos. La información puede ser parcial, no totalmente digna de crédito e informaciones provenientes de distintas fuentes pueden resultar contradictorias. A su vez, la incertidumbre de la información disponible no sólo es relativa a la ocurrencia o no de determinados sucesos, sino que también afecta a las relaciones causales que se pueden dar entre diversos sucesos.

La existencia de esta incertidumbre en el conocimiento que se desea introducir como base de conocimientos en un sistema experto, plantea inicialmente dos problemas básicos. El primero es relativo a la propia representación del conocimiento disponible. Cuando éste es incierto, es necesario hacer explícita la incertidumbre del mismo. Este hecho condicionará en gran parte la elección del tipo de representación a utilizar: sistema de producción, redes semánticas, marcos, etc. La diferencia con cualquier problema para el que el conocimiento es exacto y preciso, reside en la dificultad añadida de que, cuando el conocimiento existente es impreciso o incierto, habrá que elegir unos parámetros, numéricos o no, para representar la incertidumbre. Por consiguiente, la representación que se elija deberá tener en cuenta tanto el tipo de problema a tratar como la forma en que se exprese la incertidumbre e imprecisión del conocimiento, sin que puedan considerarse separadamente ambos aspectos.

La importancia del problema de elegir una adecuada representación y estructuración del conocimiento se pone claramente de manifiesto en la siguiente frase de Shortliffe, uno de los autores del sistema MYCIN, cuyo modelo de tratamiento de la incertidumbre, mediante los factores de certeza, tuvo un gran éxito inicial y una amplia difusión, (GORD-85):

> "... Como se ha sugerido previamente, de hecho, los detalles de un modelo de razonamiento aproximado en un sistema de Inteligencia Artificial pueden carecer relativamente de importancia ya que la cuidadosa estructuración semántica del conocimiento del dominio parece disminuir la sensibilidad de las inferencias realizadas a variaciones en los valores de los números usados. En (BUCH-84) se muestra que el rendimiento de MYCIN es extremadamente insensible a amplias variaciones en los factores de certeza asignados a las reglas. Algunos investigadores han llegado a sugerir que el razonamiento aproximado puede ser manejado sin el uso de modelos numéricos, (COHE-84)".

Otro sistema experto que tuvo gran resonancia en el momento de su aparición, PROSPECTOR, y que ha llegado a constituirse en paradigma de modelo de resolución de incertidumbre, debe en gran parte su éxito al laborioso trabajo de estructuración del conocimiento en aquellos dominios específicos de la geología a los que fue aplicado. Incluso se llegó a construir una de las primeras herramientas para adquisición del conocimiento de los expertos, K.A.S., "Knowledge Adquisition System".

Una segunda cuestión básica que se plantea al tratar con un cuerpo de conocimiento incierto, es que las inferencias que se realicen a partir de este conocimiento afectado de incertidumbre proporcionarán también conclusiones inciertas. Se hace necesario, en consecuencia, establecer unos mecanismos para propagar la incertidumbre en la cadena de razonamiento.

Son cada vez más numerosos los modelos que se proponen para abordar el problema del "Razonamiento Aproximado". La existencia de tantas propuestas diferentes ha contribuído a crear un clima de cierta confusión en este campo concreto. Si en una primera época se pensaba que la forma correcta de acercarse a los problemas de incertidumbre era a través de la Teoría de la Probabilidad, más adelante surgen nuevos planteamientos. Basándose en la supuesta aparición de inconsistencias en el terreno específico de los sistemas basados en reglas por la aplicación estricta de la Teoría de la Probabilidad, aparecen nuevos modelos que se intentan fundamentar teóricamente. En éstos, siempre aparecen referencias a la Teoría de la Probabilidad. La polémica y contraste de opiniones se hace entonces manifiesta, sin que haya desaparecido. Se dan a conocer en foros de la importancia del International Joint Conference on Artificial Intelligence, en 1985, ponencias con el título "In defense of probability", (CHEE-85), o también en el mismo año, en la Second Annual Conference on Artificial Intelligence Applications, Pearl presenta un trabajo titulado significativamente "How to do with probabilities what people say you can't", (PEAR-85). A pesar de ello se siguen proponiendo nuevos modelos, nuevos enfoques. Se hace necesaria una clarificación de este panorama, principalmente en cuanto a su aplicación a problemas reales.

La primera parte de este trabajo que se presenta, se dedica a hacer una exposición y estudio detallado de las principales teorías que han sido propuestas, con un doble objetivo. Por un lado, se establecen las bases para el "nuevo" modelo que se propone, al tiempo que se ponen de manifiesto los principales problemas concretos existentes, tanto de interpretación de las medidas para valorar la incertidumbre, como de tratamiento de algunos aspectos especialmente conflictivos. Por otra parte, y es el objetivo central de esta tesis, con dicha exposición se pretende contribuir a disminuir la confusión existente, remarcando las relaciones que existen entre los distintos modelos.

Son cinco los modelos expuestos, limitándose al tratamiento de la incertidumbre en sistemas expertos cuyos conocimiento viene representado por un conjunto de reglas:

1) Se presentan en primer lugar los modelos basados en la Teoría de la Probabilidad, centrándose en el sistema PROSPECTOR, que representa el conocimiento mediante una red de inferencia, y se ajusta bastante a la Teoría de la Probabilidad. La valoración de la incertidumbre de los hechos se efectúa a través de probabilidades a priori y a posteriori, mientras que la incertidumbre de las reglas en cuantificada por dos valores, grados de suficiencia y necesidad.

2) El modelo de los factores de certeza, desarrollado inicialmente en el sistema MYCIN, es un modelo ad-hoc, aunque sus autores han intentado darle una fundamentación teórica, ésta muy criticada. Sin embargo, tiene una amplia difusión y es de muy sencilla aplicación.

3) La Teoría de la Evidencia de Dempster y Shafer supone un paso adelante, pues presenta como grandes ventajas el modelizar de forma clara y eficaz la ignorancia inherente a un cuerpo de conocimiento incompleto y permite trabajar con un conocimiento estructurado como una jerarquía de hipótesis. La incertidumbre se cuantifica mediante una doble medida de Credibilidad y Plausibilidad.

4) La Teoría de la Posibilidad y la Lógica Difusa, debidas a Zadeh, presenta el avance de considerar proposiciones que no son precisas, pues contienen predicados difusos, así como modificadores y cuantificadores expresados en lenguaje natural. Como consecuencia de esta teoría, se han desarrollado sistemas expertos en los que la valoración de la incertidumbre no es numérica, sino lingüística.

5) Finalmente se exponen los trabajos de Dubois y Prade, que desarrollan los conceptos de Zadeh, y que mediante las medidas de posibilidad y de necesidad, permiten hacer frente simultáneamente a sucesos o proposiciones difusas y a un conjunto de conocimiento incierto.

Con estos cinco modelos se presenta una muy pequeña parte de la amplia gama de teorías existentes. No se entra en la consideración de otras teorías como puedan ser la Lógica Probabilística de Nilsson, la "Theory of Endorsements" de Cohen, el esquema de inferencia del sistema INTERNIST o los razonamientos basados en lógicas no monótonas, todas ellas de indudable valor y posiblemente aplicables en muchos campos. Es prioritario, dada la amplia difusión y desarrollo que están teniendo los sistemas expertos, ofrecer una presentación sencilla de las posibilidades que existen para resolver los problemas de incertidumbre que puedan aparecer en un determinado dominio, de forma que se facilite la elección de uno u otro modelo.

Con esa simple enumeración de los cinco modelos que se comentan en el capítulo segundo, se pone ya de manifiesto la confusión existente. Son numerosos los conceptos semánticos que se citan para la representación de la incertidumbre: Probabilidad, Credibilidad, Plausibilidad, Posibilidad, Necesidad y algunos otros no mencionados. Si estos conceptos son difíciles de precisar, aún más costoso resultará diferenciar claramente unos de otros.

La segunda parte de este trabajo se dedica a la obtención de un "nuevo" modelo para trabajar con conocimiento al mismo tiempo incierto e impreciso. Inicialmente este objetivo parece en flagrante contradicción con todo lo argumentado hasta ahora. Sin embargo, no se trata de añadir un modelo más que serviría para aumentar la confusión, sino que al contrario, la propuesta que se hace podría colaborar a simplificar el antes mencionado problema de elección de modelo de tratamiento de incertidumbre.

La Teoría de la Posibilidad, en el desarrollo que de ella hacen Dubois y Prade, permite el tratamiento conjunto de conocimiento incierto e impreciso. Presenta el inconveniente de que la valoración de la incertidumbre se realiza con dos números: posibilidad y necesidad de un suceso. Parece un serio problema, puesto que se trataría de presentar modelos de fácil comprensión e implementación. Sin embargo, los conceptos de posibilidad y necesidad permiten hacer uso de todas las teorías de Zadeh aplicables a proposiciones con predicados difusos, y que están ya muy desarrolladas.

En el modelo que se propone, y adelantando aspectos que se explicarán detalladamente en el apartado dedicado a las conclusiones, se unifican ambos valores de posibilidad y necesidad en un sólo número para cuantificar la incertidumbre, que es de fácil comprensión, y además, internamente, permanecen los valores de posibilidad y necesidad, lo que hace posible la aplicación de todos los resultados de Dubois y Prade y de Zadeh.

Este nuevo método, al mismo tiempo, servirá para el mencionado objetivo de simplificar el panorama actual, puesto que quedan explicitadas de forma clara las relaciones existentes entre el modelo de Dubois y Prade y el modelo de los factores de certeza de MYCIN.

CAPITULO 2 :

PRINCIPALES MODELOS EXISTENTES

2.1.- MODELOS PROBABILISTICOS.

2.1.1.- Problemática. Primeros intentos.

El problema que se plantea al intentar modelizar el proceso de razonamiento encaminado a una toma de decisión, aspecto fundamental en cualquier problema de diagnóstico, puede ser esquematizado de la siguiente manera.

Supóngase que se trata de elegir o decidir entre una de las hipótesis de un conjunto $h_1, h_2, \ldots, h_n$ que se suponen excluyentes y exhaustivas. Para esta toma de decisión es necesario obtener una serie de evidencias o datos relativos al problema en cuestión que permitan establecer una jerarquía entre ese conjunto de hipótesis. Si se llama $\underline{e}$ al conjunto global de todas las evidencias conseguidas, se trata de obtener los valores $P(h_i|e)$ a partir de los cuales se adoptará la decisión o solución más adecuada al problema que se está considerando. Como clara ilustración, si se considera un problema de diagnóstico médico, las hipótesis h_i corresponderían a las posibles enfermedades que puede sufrir un determinado paciente y la evidencia $\underline{e}$ sería el conjunto de síntomas, datos clínicos o análisis que se considera necesario obtener con objeto de emitir un diagnóstico y, en consecuencia, determinar el tratamiento adecuado.

Por aplicación de la regla de Bayes se obtiene:

$$P(h_i|e) = \frac{P(e|h_i) \cdot P(h_i)}{P(e)} = \frac{P(e|h_i) \cdot P(h_i)}{\sum_j P(e|h_j) \cdot P(h_j)}$$

Como ejemplo de los primeros intentos de aplicación de este planteamiento probabilístico se puede citar el de Warner y otros, (WARN-61), que no pasa de ser un programa computacional aplicado al diagnóstico de enfermedades de corazón congénitas. Si el conjunto de síntomas total $\underline{e}$ se considera formado por los síntomas individuales $e_1, e_2, \ldots, e_k$, condicionalmente independientes, la ecuación anterior se convierte en:

$$P(h_i | e_1, \ldots, e_k) = \frac{P(e_1|h_i) \ldots P(e_k|h_i) \, . \, P(h_i)}{\sum_j P(e_1|h_j) \ldots P(e_k|h_j) \, . \, P(h_j)}$$

Si incluso la ausencia de un determinado síntoma puede ser significativa de cara al diagnóstico, también se tiene en cuenta:

$$P(h_i | e_1, \ldots, \tilde{e}_r, \ldots, e_k) = \frac{P(e_1|h_i) \ldots (1 - P(e_r|h_i)) \ldots P(e_k|h_i) \, . \, P(h_i)}{\sum_j P(e_1|h_j) \ldots (1 - P(e_r|h_i)) \ldots P(e_k|h_j) \, . \, P(h_j)}$$

El modelo de Warner consiste esencialmente en la aplicación de esta fórmula a un conjunto de 50 posibles síntomas y 33 enfermedades. Para ello es necesario conocer los valores de las probabilidades condicionales $P(e_i|h_j)$, que fueron obtenidos a partir de los datos clínicos y de laboratorio relativos a 1035 pacientes.

Sin embargo, en el problema de diagnóstico en particular, y, en general en cualquier problema de decisión, las evidencias e_i no se obtienen indiscriminadamente, sino de forma secuencial. Es decir, una vez obtenidas unas evidencias $\underline{e}$, los datos adicionales a conseguir vendrán determinados por el estado hasta ese momento de las probabilidades de cada una de las hipótesis h_i. El siguiente paso es, en consecuencia, considerar unas evidencias obtenidas incrementalmente. Gorry y Barnett, (GORR-68), plantean un modelo de diagnóstico secuencial, utilizando la siguiente expresión para el teorema de Bayes. Si se han conseguido un conjunto de evidencias $\underline{e}$ y las correspondientes probabilidades condicionales $P(h_i|e)$ y, como consecuencia de estos valores, se decide tomar en consideración una nueva evidencia e_1, y si se llama e' al total de evidencias, $e' = e \wedge e_1$, se tiene:

$$P(h_i | e') = \frac{P(e_1|h_i \wedge e) \, . \, P(h_i|e)}{\sum_j P(e_1|h_j \wedge e) \, . \, P(h_j|e)}$$

En cada uno de los pasos $P(h_i|e')$ sustituirá a $P(h_i|e)$ a medida que se va adquiriendo nueva evidencia y se utilizará la fórmula anterior para actualizar las probabilidades. Si se hace este proceso en tan sólo dos etapas, $e = e_1$ y $e' = e_1 \wedge e_2$ se tendría:

$$P(h_i|e_1) = \frac{P(e_1|h_i) \cdot P(h_i)}{\sum_j P(e_1|h_j) \cdot P(h_j)}$$

$$P(h_i|e') = \frac{P(e_2|h_i \wedge e_1) \cdot P(h_i|e_1)}{\sum_j P(e_2|h_j \wedge e_1) \cdot P(h_j|e_1)}$$

Se aprecian ya claramente los problemas que se derivan de esta formulación en cuanto a la necesidad de datos estadísticos para poder continuar este proceso. No sólo son necesarios los valores de las probabilidades condicionales $P(e_1|h_i)$ sino también valores del tipo $P(e_2|h_i \wedge e_1)$, es decir, las interrelaciones de las diferentes evidencias o síntomas en relación con cada una de las hipótesis, con los correspondientes problemas de independencia (e_1 y e_2 pueden ser condicionalmente independientes respecto a una determinada hipótesis y estar fuertemente relacionados en presencia de otra diferente). En cuanto a la obtención de estos valores numéricos a partir de datos estadísticos, se originan también nuevos problemas pues, por ejemplo, si se ha tenido en cuenta una muestra de 2000 individuos, pudiera suceder que, para la obtención del valor $P(e_2|h_i \wedge e_1)$, el suceso $h_i \wedge e_1$ ocurriera en 25 o 30 casos con lo que el nivel de significación de $P(e_2|h_i \wedge e_1)$ sería muy bajo.

En esta línea y ya plenamente dentro de los sistemas expertos se produce un paso importante al sustituir la obtención de todas estas probabilidades condicionales que aparecen en la regla de Bayes de muestras estadísticas, por juicios subjetivos consecuencia de la consulta al experto o expertos que intervienen en el diseño de la base de conocimientos. Se pasa así a considerar las probabilidades condicionales como probabilidades subjetivas, cuyos valores son aportados por el experto, (PEAR-82), (CHEE-85).

De todas formas, el problema, aunque simplificado, continúa existiendo por la gran cantidad de valores que sería necesario obtener del experto, que se vería abrumado e incapaz de suministrarlos suficientemente precisos.

En consecuencia, se ha tratado de enfocar el problema de diagnóstico o de la toma de decisión por otros caminos. Szolovits (SZOL-78), partiendo del esquema de diagnóstico secuencial de Gorry antes citado, hace una revisión crítica de los problemas que plantea la aplicación del teorema de Bayes. Coincide con Shortliffe (SHOR-75), en la insalvable dificultad que supone la obtención de los datos necesarios para la estricta aplicación de un modelo probabilístico. En dicho trabajo de Szolovits se hace un estudio del número de parámetros necesario, así como de las posibles simplificaciones que podrían introducirse de cara a reducir el número de datos a obtener. Finalmente pasa revista a diferentes sistemas desarrollados en el campo del diagnóstico médico, PIP, INTERNIST y CASNET, incluyendo también referencias al sistema MYCIN. Todos estos sistemas emplean métodos ad-hoc en el razonamiento probabilístico para obviar las dificultades que se derivan de un desarrollo puramente bayesiano.

Dos de los modelos más difundidos son PROSPECTOR, que intenta adecuarse en la medida de lo posible a la teoría de la probabilidad, y MYCIN que tiene lejanas reminiscencias probabilísticas. Se considera el primero de ellos en este apartado, mientras que el modelo de los factores de certeza inicialmente desarrollado en MYCIN se expondrá en el siguiente apartado 2.2., haciendo hincapié en ambos sistemas en aquellos conceptos que subyacen en su desarrollo.

2.1.2.- El sistema PROSPECTOR.

Utilizando probabilidades subjetivas y métodos probabilísticos a partir del teorema de Bayes, se desarrolla a finales de los setenta y primeros de los ochenta el sistema PROSPECTOR por Duda, Hart y otros en el Stanford Research Institute, (DUDA-76), y cuyo objetivo es la ayuda a la clasificación de prospecciones geológicas así como la determinación del lugar adecuado de perforación, entre otras varias utilidades.

No se puede considerar estrictamente como un sistema basado en reglas, puesto que su base de conocimientos no se estructura como un conjunto de reglas de producción sino en lo que sus autores llaman "inference network" o red de inferencia. En la representación del conocimiento que consideran se hace uso en parte de redes semánticas muy elementales y el grafo dirigido que se utiliza para el razonamiento puede interpretarse como un sistema de producción aunque con mecanismos específicos como pueden ser la equiparación de sentencias, el establecimiento de nodos contexto o la existencia de nodos AND y OR. Detalles de esta representación pueden encontrarse en (DUDA-78) y (DUDA-84).

De forma muy simplificada y con el objetivo de mostrar los fundamentos del razonamiento probabilístico que utiliza el sistema PROSPECTOR, se puede considerar una red de inferencia como un grafo dirigido cuyos nodos son los hechos considerados, siendo los nodos de nivel más alto las hipótesis a establecer y los nodos del nivel más bajo evidencias a proporcionar por el usuario. Los arcos que unen los nodos pueden interpretarse como relaciones causales entre los hechos representados por los correspondientes nodos. Los nodos intermedios, en consecuencia, juegan un doble papel de hipótesis o evidencia según se consideren los arcos que inciden o los que salen del nodo. Esta red de inferencia no es un árbol en ningún caso de los desarrollados como aplicación del sistema y la única condición que debe verificar es la no existencia de ciclos con el fin de evitar cadenas de razonamiento indefinidas.

Se exponen a continuación con cierto grado de detalle los fundamentos del razonamiento probabilístico utilizado en PROSPECTOR. A través de esta explicación se irá viendo de qué forma se representa la incertidumbre, cómo se propaga a través de la red de inferencia y qué parámetros son utilizados, así como su significado.

2.1.2.1.- Propagación de incertidumbre.

Considérese una regla del tipo "si e entonces h" que correspondería a un arco en la red de inferencia.

El modelo PROSPECTOR conoce antes de cualquier inicialización las probabilidades a priori de todos los nodos de la red. Se trata de actualizar la probabilidad a priori de la hipótesis h, P(h), en función del conocimiento que se tenga de la evidencia e.

Se tienen en cuenta tres situaciones diferentes:

a) se conoce el hecho e como verdadero con total certeza,

b) se sabe con certeza que e es falso, y

c) la verdad o falsedad de e no se conoce con certeza.

Aunque la situación real corresponderá casi siempre al tercer caso, las dos primeras son necesarias para la actualización de P(h) en el último.

Caso a): e es verdadero, con total certeza.

El modelo PROSPECTOR utiliza probabilidades a priori y a posteriori de forma indirecta mediante el procedimiento que se describe a continuación.

Dos expresiones sencillas del teorema de Bayes son:

$$P(h|e) = \frac{P(e|h) . P(h)}{P(e)}$$

$$P(\bar{h}|e) = \frac{P(e|\bar{h}) . P(\bar{h})}{P(e)}$$

y dividiendo ambas igualdades se obtiene

$$\frac{P(h|e)}{P(\bar{h}|e)} = \frac{P(e|h)}{P(e|\bar{h})} \cdot \frac{P(h)}{P(\bar{h})}$$

que si se utiliza la notación:

$$O(h) = \frac{P(h)}{P(\bar{h})} \qquad \text{disparidad a priori}$$

$$O(h|e) = \frac{P(h|e)}{P(\bar{h}|e)} \qquad \text{disparidad a posteriori}$$

$$LS = \frac{P(e|h)}{P(e|\bar{h})} \quad \text{grado de suficiencia}$$

proporciona la relación

$$\boxed{O(h|e) = LS.O(h)} \qquad (1)$$

la cual, conocido el coeficiente LS, cuyo significado se explicará algo más adelante, permite obtener la disparidad a posteriori a partir de la disparidad a priori.

PROSPECTOR utiliza valores de disparidad (relación entre las probabilidades a favor y en contra de la hipótesis <u>h</u>) en vez de utilizar directamente probabilidades. Obsérvese que cada uno de los valores puede obtenerse inmediatamente a partir del otro:

$$O(h) = \frac{P(h)}{P(\bar{h})} = \frac{P(h)}{1-P(h)} \;; \qquad P(h) = \frac{O(h)}{1+O(h)}$$

<u>Caso b)</u>: <u>e</u> es falso, con toda certeza.

De forma análoga, por aplicación dos veces de la regla de Bayes se obtiene:

$$P(h|\bar{e}) = \frac{P(\bar{e}|h).P(h)}{P(\bar{e})} \qquad P(\bar{h}|\bar{e}) = \frac{P(\bar{e}|\bar{h}).P(\bar{h})}{P(\bar{e})}$$

y dividiendo nuevamente

$$\frac{P(h|\bar{e})}{P(\bar{h}|\bar{e})} = \frac{P(\bar{e}|h)}{P(\bar{e}|\bar{h})} \cdot \frac{P(h)}{P(\bar{h})}$$

que llamado LN, grado de necesidad, al cociente $\frac{P(\bar{e}|h)}{P(\bar{e}|\bar{h})}$ proporciona

$$\boxed{O(h|\bar{e}) = LN.O(h)} \qquad (2)$$

que, de forma sencilla, actualiza la probabilidad de h, supuesto e falso y utilizando el coeficiente LN.

Interpretación de los coeficientes LS y LN.

El coeficiente LS se llama grado de suficiencia puesto que evalúa hasta qué punto la certeza de la evidencia e es suficiente para que se dé la hipótesis h. En efecto, valores de LS muy superiores a la unidad permiten pasar del valor previo de la probabilidad P(h) a valores mucho más altos $P(h|e)$, ecuación (1). Si e es verdadero y LS es alto, entonces el conocimiento de e aumenta la certeza en h.

Análogamente, el coeficiente LN, grado de necesidad, mide cuán necesaria es la presencia de la evidencia e para que se dé la hipótesis h. Para valores de LN muy inferiores a la unidad, se pasará de la probabilidad inicial P(h) a valores mucho más bajos, $P(h|\bar{e})$, ecuación (2). Es decir, la ausencia de la evidencia e, junto con un valor bajo del grado de necesidad LN, disminuye la certeza en la hipótesis h. Cuanto menor sea el valor de LN, mayor será la necesidad de la presencia de e para la certeza de h.

En consecuencia, ambos valores LS y LN miden la fuerza de la relación causal existente entre la evidencia e y la hipótesis h. Los dos valores, con el significado anterior, y junto con el valor de la probabilidad a priori son proporcionados por el experto con el que se diseña la red de inferencia.

Sin embargo, para cada regla "si e entonces h" los grados de suficiencia y necesidad no son independientes. De su propia definición

$$LS = \frac{P(e|h)}{P(e|\bar{h})} \qquad LN = \frac{P(\bar{e}|h)}{P(\bar{e}|\bar{h})} = \frac{1-P(e|h)}{1-P(e|\bar{h})}$$

se obtiene

$$P(e|h) = LS \cdot \frac{1-LN}{LS-LN}$$

$$P(e|\bar{h}) = \frac{1-LN}{LS-LN} \qquad y$$

$$LN = \frac{1-LS \;.P(e|\bar{h})}{1-P(e|\bar{h})}$$

estas relaciones permiten hacer las siguientes consideraciones:

.) El par LS, LN proporciona la misma información que $P(e|h)$ y $P(e|\bar{h})$. El que se pida al experto uno u otro par de valores es una cuestión puramente psicológica, en función del significado de cada uno de los parámetros y, efectivamente, parecen más claros los grados de suficiencia y necesidad.

.) $LS > 1 \Longrightarrow LN < 1$
$LS < 1 \Longrightarrow LN > 1$
$LS = 1 \Longleftrightarrow LN = 1$

.) Aparece un problema que se comentará ampliamente al exponer el modelo MYCIN. Si el experto afirma que la presencia de e aumenta la probabilidad de h, $LS > 1$, entonces la ausencia de e deberá disminuirla, $LN < 1$, en contra de la opinión generalizada de los expertos, en cuanto que si la presencia de e favorece la creencia en la hipótesis h, su ausencia no tiene por qué ser un dato en contra de la misma. Sin embargo, PROSPECTOR dispone de mecanismos correctores para eliminar éste y otros comportamientos no deseados.

Caso c): la presencia o ausencia de e no es conocida con certeza.

Esta será la situación más frecuente, bien porque siendo e un nodo terminal (hoja en la red de inferencia) el usuario no tiene total certeza sobre su "status", o bien porque e sea un nodo intermedio de la red que en el proceso de inferencia vendrá afectado por una probabilidad. En ambos casos se utilizará la notación $P(e|e')$ para expresar la incertidumbre sobre e. En la primera de las situaciones, e' significará el conjunto de evidencias no expresadas que el usuario dispone para determinar la certeza de e (obsérvese que $P(e|e')$ vuelve a ser una probabilidad subjetiva), mientras que una interpretación sencilla del segundo caso sería la siguiente disposición en la red, donde e' es un nodo terminal conocido con certeza.

$$e' \longrightarrow e \longrightarrow h$$

Se trata ahora de obtener $P(h|e')$ conocidas $P(e|e')$, $P(h|e)$ y las probabilidades a priori $P(h)$ y $P(e)$. De la teoría elemental de probabilidades se deduce:

$$
\begin{aligned}
P(h|e') &\underset{(a)}{=} P(h \wedge e|e') + P(h \wedge \bar{e}|e') = \\
&\underset{(b)}{=} P(h|e \wedge e').P(e|e') + P(h|\bar{e} \wedge e') \; . \; P(\bar{e}|e') = \\
&\underset{(c)}{=} P(h|e).P(e|e') + P(h|\bar{e}).P(e|e')
\end{aligned}
$$

(a): $h = (h \wedge e) \vee (h \wedge \bar{e})$ siendo $h \wedge e$ y $h \wedge \bar{e}$ excluyentes

(b): $P(A \wedge B) = P(A|B).P(B)$

(c): Si e es conocido como verdadero o falso con total certeza, la nueva evidencia e' relativa a e no proporciona nueva información sobre h.

En definitiva, para actualizar la probabilidad de h:

$$\boxed{P(h|e') = P(h|e).P(e|e') + P(h|\bar{e}).(1-P(e|e'))} \qquad (3)$$

Los valores $P(h|e)$ y $P(h|e)$ se calculan mediante las relaciones (1) y (2), a partir de las disparidades $O(h|e)$ y $O(h|\bar{e})$ y utilizando los parámetros LS y LN de la regla "si e entonces h" y la probabilidad a priori $P(h)$; $P(e|e')$ será introducido directamente por el usuario o se habrá obtenido de la regla "si e' entonces e" en la situación sencilla considerada anteriormente o como aplicación conjunta del proceso de inferencia para obtener la probabilidad del nodo e en casos más complejos.

La ecuación (3) es una sencilla fórmula de interpolación lineal entre los valores extremos $P(e|e') = 1$, $P(h|e') = P(h|e)$ y $P(e|e') = 0$, $P(h|e') = P(h|\bar{e})$.

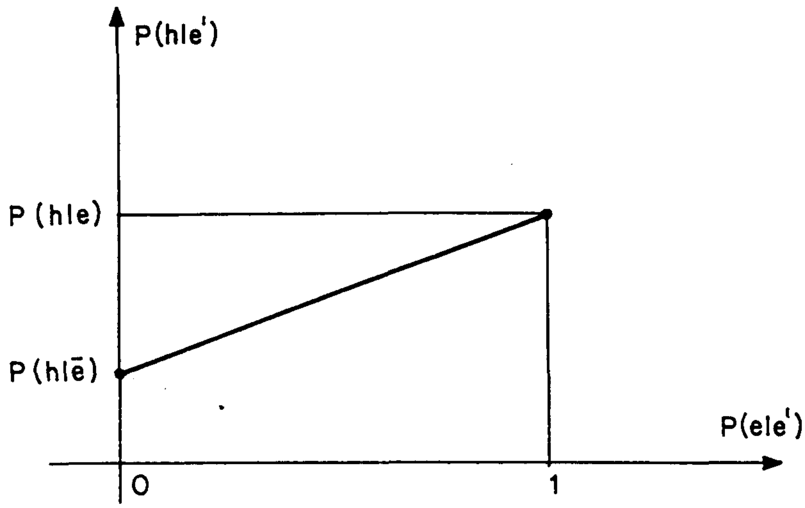


Figura 1.- Interpolación lineal para la actualización $P(h|e')$.

Sin embargo, a partir de esta relación lineal entre $P(e|e')$ y $P(h|e')$ aparecen algunos problemas de inconsistencia. En efecto, si $P(e|e') = P(e)$ se obtiene

$$P(h|e') = P(h|e).P(e) + P(h|\bar{e}).P(\bar{e}) = P(h \wedge e) + P(h \wedge \bar{e}) = P(h)$$

es decir, si e' no aporta información adicional sobre e, no se modifica la probabilidad de h. Pero los valores de P(e) y P(h) son proporcionados por el experto, la recta de interpolación también, puesto que se obtiene a partir de los valores LS, LN y P(h), y la situación habitual será que el punto de coordenadas (P(e), P(h)) no esté sobre la recta:

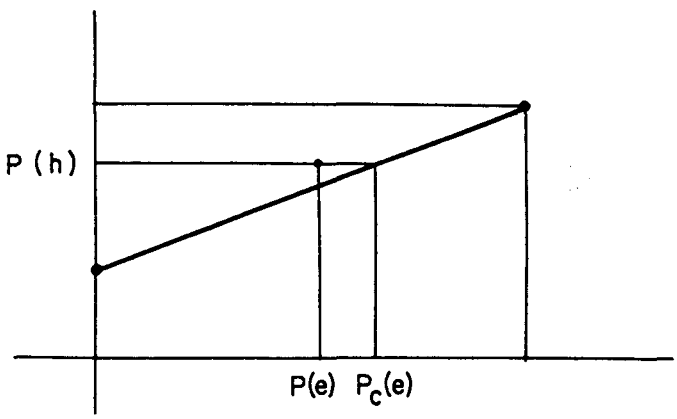


Figura 2.- Probabilidades a priori inconsistentes.

Como consecuencia de esta inconsistencia entre las probabilidades a priori y los valores LS y LN de la regla se producen comportamientos claramente contradictorios, como el siguiente. Sea $P_c(e)$ el valor de la probabilidad a priori de e consistente con la probabilidad a priori de h y supóngase que la situación es la que se refleja en la figura 2. Si, por ejemplo, $P(e) < P(e|e') < P_c(e)$ entonces $P(h|e') < P(h)$; de una evidencia e' que aumenta la credibilidad de e se deduce una disminución en la probabilidad de h, en contra de lo que debería ocurrir.

Los autores del modelo PROSPECTOR proponen muy variados mecanismos para solucionar esta inconsistencia.

El más inmediato, dado que la inconsistencia viene originada por un exceso de parámetros suministrados por el experto, sería modificar alguno de esos datos, sugiriendo que dicha modificación recaiga en aquella especificación subjetiva menos fiable. Sin embargo, al modificar alguno de los parámetros para hacerlos consistentes al considerar una regla "si e entonces h" se vería alterada la consistencia de los parámetros de aquellas reglas que inciden en los nodos e y h. No es posible por tanto tratar este problema de inconsistencia de forma local, sino que sería necesario un estudio global que afectara a toda la red de inferencia.

Otra posible solución consistiría en modificar la función de interpolación (3) de manera que fuera consistente con los valores de las probabilidades a priori. Esta modificación puede hacerse de diferentes formas.

La más simple es forzar la interpolación entre aquellos puntos conocidos, figura 3.a, logrando la consistencia. Desaparecen los problemas de consistencia, pero no hay ningún otro motivo que justifique esta modificación.

Otra posible solución en la misma línea sería la que se muestra en la figura 3.b. No se cambia la probabilidad de h si el valor de $P(e|e')$ está en el intervalo $[P(e), P_c(e)]$, interpretándose como la inexistencia de razones para dicho cambio en base a la nueva información proporcionada por la evidencia e'.

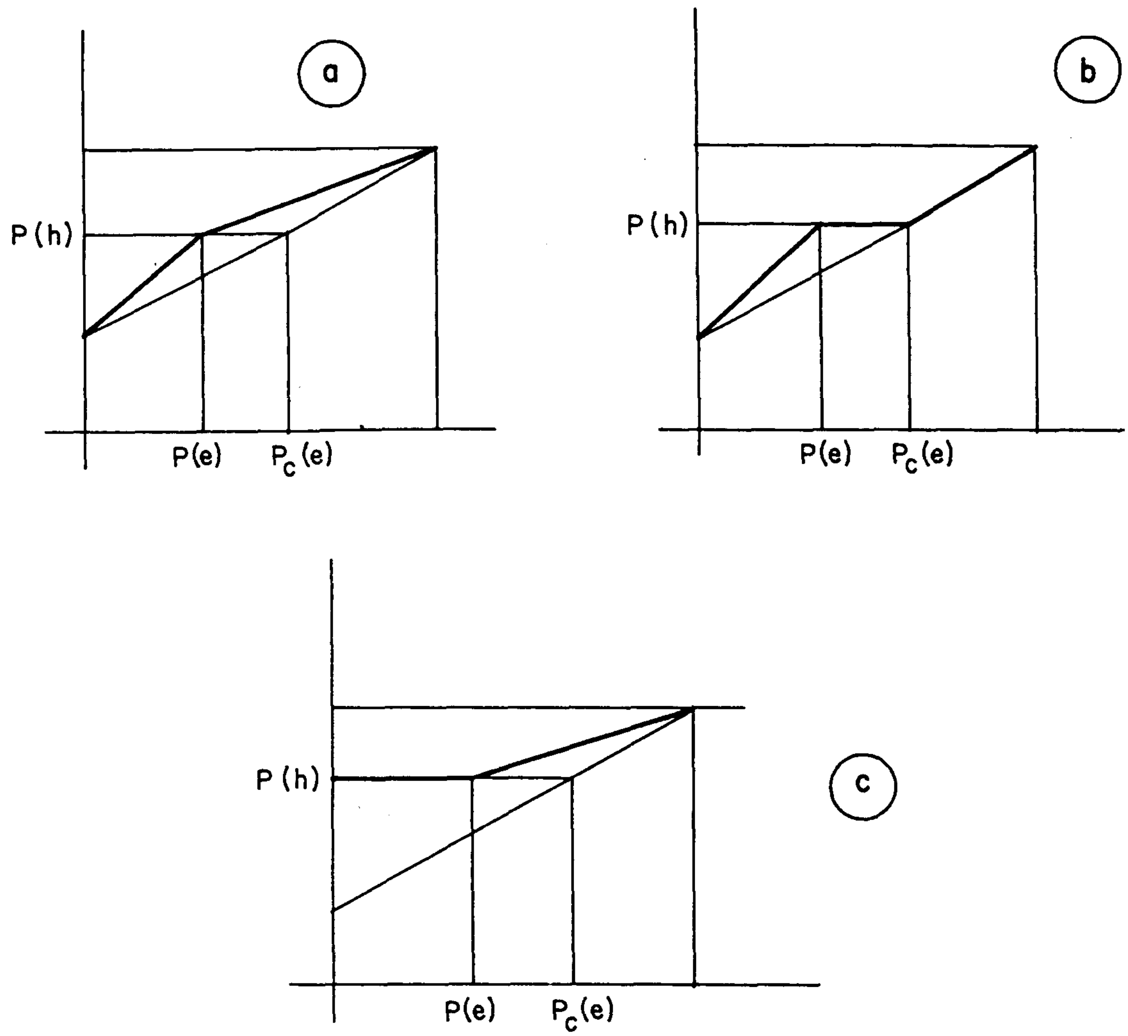


Figura 3.- Interpolación con probabilidades consistentes.

Finalmente una tercera propuesta se refleja en la figura 3.c. Si $P(e|e') < P(e)$ no se modifica la probabilidad de h. Esta última solución es coherente con el deseo de que la ausencia de la evidencia e no se constituya como evidencia en contra de la hipótesis y, en consecuencia, no se disminuye la probabilidad inicial de h.

2.1.2.2.- Combinación de evidencias.

Se considera a continuación la situación en que existen distintas reglas con la misma conclusión, "si e_i entonces h" o, en otros términos, distintos arcos en la red de inferencia entrando en el nodo que representa la hipótesis h. Cada una de estas reglas viene afectada por sus correspondientes parámetros, grado de suficiencia LS_i y grado de necesidad LN_i. Al igual que se hizo anteriormente, se tienen en cuenta diferentes casos.

Caso a): todas las evidencias e_i son verdad con certeza.

En PROSPECTOR se hace la fuerte suposición de que todas las evidencias e_i son condicionalmente independientes tanto respecto a la hipótesis $\underline{h}$ como a su negación, lo cual en términos de probabilidades significa:

$$P(e_1,\ldots,e_n|h) = \prod_{i=1}^{n} P(e_i|h)$$

$$P(e_1,\ldots,e_n|\bar{h}) = \prod_{i=1}^{n} P(e_i|\bar{h})$$

En estas condiciones, la aplicación del teorema de Bayes permite deducir:

$$O(h|e_1,\ldots,e_n) = \frac{P(h|e_1,\ldots,e_n)}{P(\bar{h}|e_1,\ldots,e_n)} = \frac{P(e_1,\ldots,e_n|h).P(h)\ /\ P(e_1,\ldots,e_n)}{P(e_1,\ldots,e_n|\bar{h}).P(\bar{h})\ /\ P(e_1,\ldots,e_n)} =$$

$$= \frac{\prod P(e_i|h)}{\prod P(e_i|\bar{h})} \,.\, O(h) = \left(\prod_{i=1}^{n} LS_i\right) .\ O(h)$$

En la notación elegida para este desarrollo se ha optado por no utilizar $e_1 \wedge \ldots \wedge e_n$, dado que PROSPECTOR considera también nodos AND y OR, con un tratamiento específico que se expondrá más adelante.

Caso b): todas las evidencias e_i son conocidas como falsas con certeza.

De forma análoga, y puesto que la independencia condicional de las evidencias e_i respecto de h y $\bar{h}$ implica la independencia de las negaciones de dichas evidencias, se obtiene:

$$O(h|\bar{e}_1,\ldots,\bar{e}_n) = \left(\prod_{i=1}^{n} LN_i\right).\ O(h)$$

Caso c): existe incertidumbre sobre la verdad o falsedad de las evidencias e_i.

Se incluyen en esta situación casos tan dispares como aquél en que algunas evidencias se conocen como verdaderas y otras como falsas, no hay información sobre alguna evidencia, así como el método general para propagar la incertidumbre a través de la red.

Se suponen conocidos los valores $P(e_i|e'_i)$, bien por indicación del usuario o por etapas anteriores del proceso. Como ya se vió anteriormente, se sabe cómo actualizar la probabilidad de h debida únicamente a las evidencias e'_i y por tanto, para cada regla incidente en h se puede calcular $O(h|e'_i)$.

Se define para cada regla una razón de probabilidad efectiva:

$$L'_i = \frac{O(h|e'_i)}{O(h)} \qquad (4)$$

La razón de probabilidad efectiva obsérvese que varía entre los valores LS si $P(e_i|e'_i) = 1$, es decir, e_i es verdad, y LN cuando $P(e_i|e'_i) = 0$, e_i falso, pasando por el valor 1 cuando $P(e_i|e'_i) = P(e_i)$ es decir, no hay aumento ni disminución en la probabilidad de e_i como consecuencia de la obtención de nuevas evidencias e'_i.

Suponiendo que las evidencias e'_i son condicionalmente independientes se considera para el conjunto de las reglas que inciden en h:

$$\boxed{O(h|e'_1,\ldots,e'_n) = \left(\prod_{i=1}^{n} L'_i\right) \cdot O(h)} \qquad (5)$$

La utilización de las razones de probabilidad efectiva proporciona el mecanismo para la propagación y actualización de la probabilidades de todos los nodos a lo largo de la red. La forma concreta de actuar es la siguiente. Se almacena con cada nodo su probabilidad a priori (o su disparidad) y con cada arco una razón de probabilidad efectiva L'_i, que inicialmente toma el valor 1, i.e., no hay ninguna información y por tanto disparidades a priori y a posteriori son iguales. Cuando la probabilidad de cualquier nodo, sea terminal o no, es modificada, se obtienen las nuevas razones de probabilidad efectivas de los arcos que salen de él, ecuación (4), y mediante la relación (5) va modificándose la probabilidad de los nodos conectados a aquéllos que se modificaron inicialmente. La aplicación repetida de este procedimiento permite que los efectos iniciales de la alteración de las probabilidades de distintos nodos se propaguen a través de toda la red.

Como ejemplo de aplicación, considérese la sencilla red de la figura 4 y supóngase que el usuario introduce información sobre la evidencia e_{23} mediante el valor $P(e_{23}|e')$ y ninguna otra información sobre el resto de los nodos de la red.

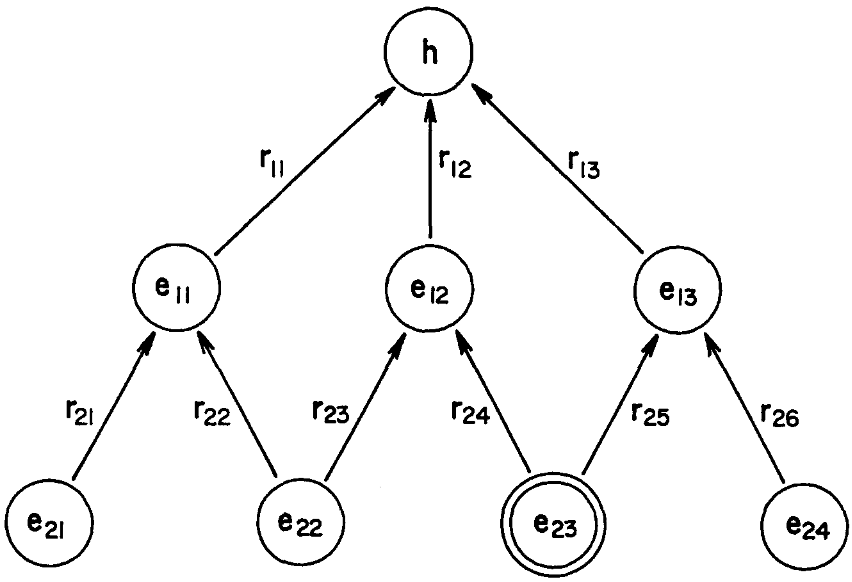


Figura 4.- Ejemplo de propagación de incertidumbre.

1. Se modifican las razones de probabilidad efectivas de los arcos r_{24} y r_{25}, obteniéndose unos nuevos valores L'_{24} y L'_{25}.

2. Resultan afectadas las probabilidades de los nodos e_{12} y e_{13}:

$$O(e_{12}|e') = L'_{23}.L'_{24}.O(e_{12}) = 1.\ L'_{24}\ .O(e_{12})$$

$$O(e_{13}|e') = L'_{25}.L'_{26}.O(e_{13}) = L'_{25}\ .1\ .O(e_{13})$$

3. Como consecuencia de la modificación de las probabilidades de estos dos nodos se modifican las razones de probabilidad de las reglas r_{12} y r_{13} y, procediendo sucesivamente de la misma manera, se van actualizando las probabilidades de todos los nodos.

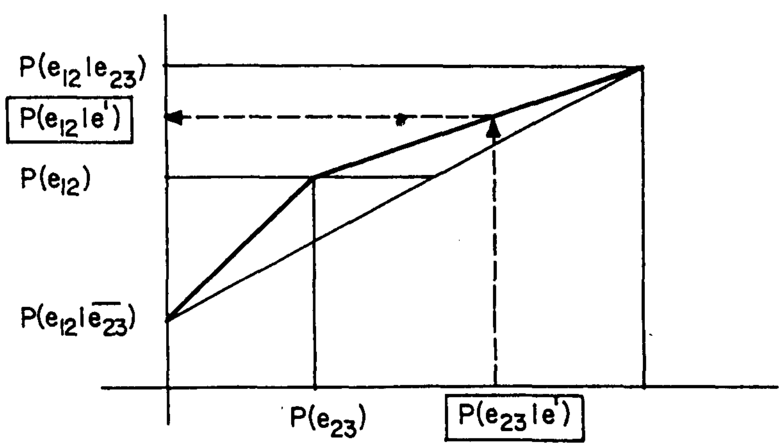
Para obtener los nuevos valores de las razones de probabilidad efectivas, definidas mediante la relación (4), se utilizan los procedimientos vistos al considerar la propagación de incertidumbre. Por ejemplo, para obtener el nuevo valor de L'_{24}:

1. Se obtienen $O(e_{12}|e_{23}) = LS_{24}\ .\ O(e_{12})$

$$O(e_{12}|\overline{e}_{23}) = LN_{24}\ .\ O(e_{12})$$

y mediante la relación $P = \dfrac{O}{1+O}$, los nuevos valores de $P(e_{12}|e_{23})$ y $P(e_{12}|\overline{e}_{23})$.

2. Con estos valores y entrando en alguna de las funciones de interpolación con el valor modificado de la probabilidad de la evidencia e_{23} se obtiene $P(e_{12}|e')$ y usando $O = \dfrac{P}{1-P}$ se obtiene $O(e_{12}|e')$.



3. La nueva razón de probabilidad efectiva es $L'_{24} = \frac{O(e_{12}|e')}{O(e_{12})}$.

En consecuencia, este conjunto de procedimientos permite de forma sencilla la propagación a través de la red de la modificación en la probabilidad de cualquier nodo de la misma.

Obsérvese que la misma regla puede ser utilizada sucesivas veces, por ejemplo si el usuario modifica su juicio y cambia una aserción hecha anteriormente. La nueva aserción anulará correctamente cualquier efecto de anteriores juicios.

2.1.2.3.- Combinación lógica de evidencias.

En la red de inferencia de PROSPECTOR existen también nodos que no corresponden a hechos, proposiciones o sentencias sino que son nodos AND y OR. Su objeto es proveer al diseñador de medios para que la red se adapte a comportamientos que se desea se produzcan. Otros mecanismos tales como la categorización de nodos como "contextos", que no será discutida en esta exposición, persiguen objetivos similares.

Un nodo tipo AND es equivalente a la consideración de una regla tipo "si $e_1 \wedge \cdots \wedge e_k$ entonces $\underline{h}$". Se debe prever por tanto un método para el cálculo de probabilidades de este tipo de combinación lógica de sentencias o proposiciones, es decir, para computar los valores de $P(e_1 \wedge \cdots \wedge e_k|e')$ y $P(e_1 \vee \cdots \vee e_r|e')$.

Con la suposición de independencia condicional se tendría

$$P(e_1 \wedge \cdots \wedge e_k|e') = \prod_{i=1}^{n} P(e_i|e') \quad y$$

$$P(e_1 \vee \cdots \vee e_r|e') = 1 - \prod_{i=1}^{r} (1-P(e_i|e'))$$

pero estas fórmulas no son utilizadas en PROSPECTOR pues, en opinión de los autores, conduce a resultados demasiado pesimistas para conjunción de evidencias y demasiado optimistas para la disyunción.

Proponen en cambio, la utilización de las fórmulas derivadas de la lógica difusa de Zadeh:

$$P(e_1 \wedge \dots \wedge e_k | e') = \min_i \left\{ P(e_i | e') \right\}$$

$$P(e_1 \vee \dots \vee e_r | e') = \max_i \left\{ P(e_i | e') \right\}$$

El comportamiento utilizando estas fórmulas no es totalmente satisfactorio, pues para la disyunción de evidencias, ninguna disminución en la probabilidad de aquellos componentes cuya probabilidad a priori no es máxima reducirá la probabilidad de la disyunción. Análogamente, la probabilidad de la conjunción nunca puede llegar a ser mayor que la probabilidad del componente que no se modifica y que es menos probable a priori.

Como alternativa, si se considera cada una de las evidencias e_i incidiendo en la hipótesis <u>h</u>, es decir considerar <u>k</u> reglas distintas, la probabilidad de la hipótesis puede llegar a ser bastante alta, aún cuando no se conozca nada sobre una o más de las evidencias e_i (i.e., no se modifique su probabilidad a priori).

La utilización que se haga de nodos OR Y AND o de arcos directamente sobre la hipótesis <u>h</u>, o incluso el uso conjunto de ambos conceptos, dependerá del comportamiento local que se desee. En la figura 5, que muestra una porción de una red de inferencia implementada en PROSPECTOR, puede apreciarse el uso que se hace de nodos OR y AND conjuntamente con el procedimiento para combinar evidencias que inciden en el mismo nodo. Concretamente, con esta mezcla de ambos tipos de combinación de evidencias se consiguen en el nodo OR comportamientos que no sería posible reflejar con uno sólo de esos mecanismos.

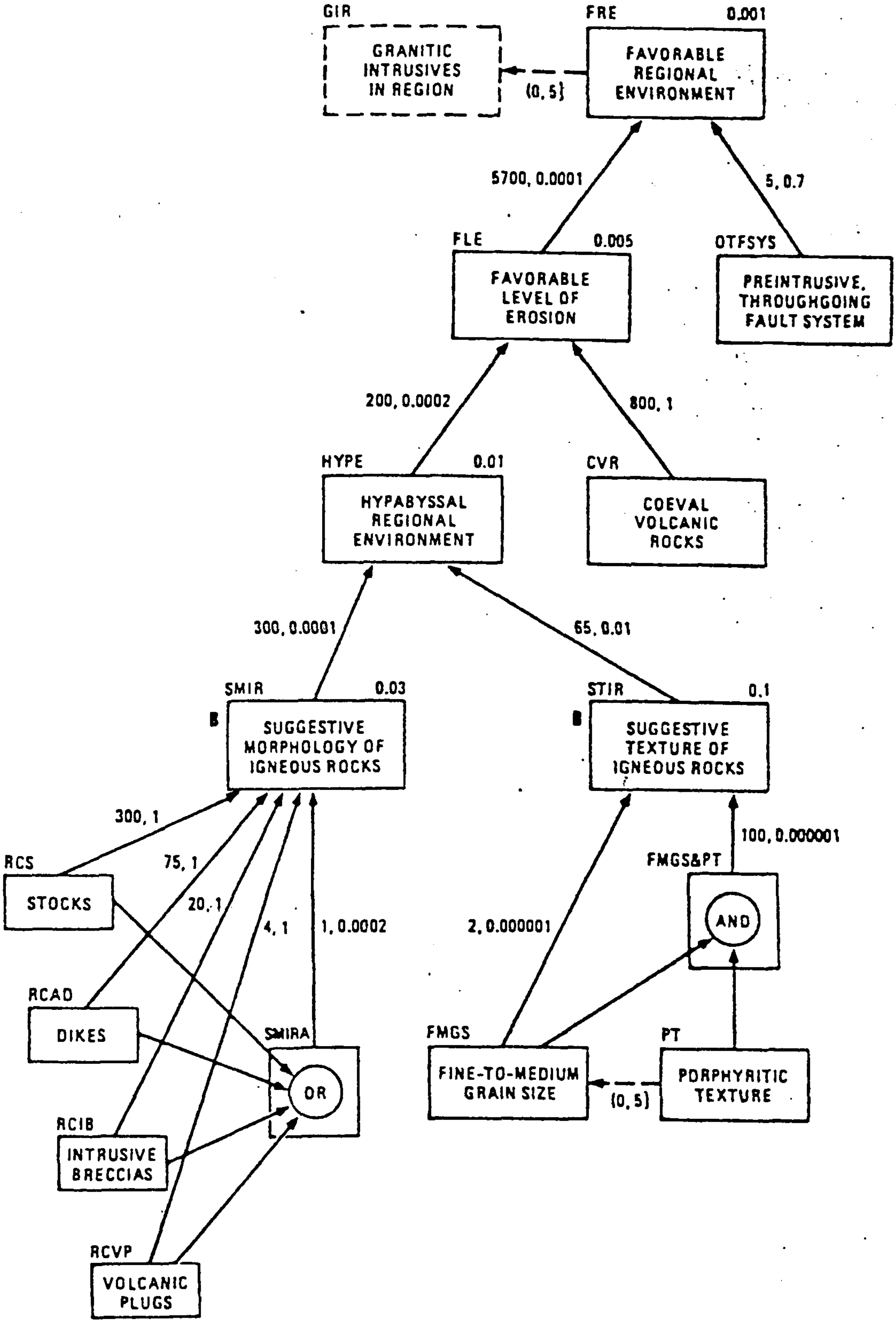


Figura 5.- Red de inferencia parcial en el sistema PROSPECTOR.

2.1.3.- Desarrollos posteriores.

Las principales críticas que se hacen a PROSPECTOR tienen relación con la suposición de independencia condicional al combinar distintas evidencias incidiendo en la misma hipótesis h. Especial importancia tiene esta suposición en el caso de que se considere un conjunto de hipótesis exhaustivo y mutuamente excluyentes, habiéndose abierto una amplia discusión teórica al respecto, (PEDN-81), (GLYM-85), (JOHN-86). Aportaciones mucho más prácticas y concretas son las de Konolige, uno de los colaboradores en el desarrollo de PROSPECTOR, (KONO-79), llegando incluso a proponer un desarrollo modificado, basado en el principio de máxima entropía, pero mucho más complejo y que no ha llegado a implementarse en ningún sistema.

Es cierto que el procedimiento que se propone en PROSPECTOR para la combinación de múltiples evidencias no deja de ser un modelo ad-hoc y sin excesivo rigor teórico, pero habría que señalar que, al igual que en otros sistemas, los inconvenientes que se derivan de los supuestos de independencia condicional se pueden evitar en gran parte al estructurar la base de conocimiento, en este caso al diseñar la red de inferencia, poniendo especial cuidado en que todas las evidencias que incidan en el mismo nodo sean realmente independientes y procurando agrupar en un mismo nodo aquéllas que no lo sean.

Con objeto de resolver este problema, Pear (PEAR-86c) ha desarrollado otro modelo basado en "bayesian networks", en los que las dependencias entre hechos se ponen de manifiesto mediante su representación explícita en una red. Otros sistemas que se han propuesto con fuerte contenido probabilístico, y en la misma línea que el anterior, son los debidos a Howard, (HOWA-81), basado en "diagramas de influencia", y los trabajos de Chow y Liu con una representación del conocimiento basada en "árboles de dependencia".

Todos estos desarrollos parten de una representación específica del conocimiento, lo que hace que sean difícilmente adaptables a un Sistema Experto basado en reglas. En cualquier caso, son mucho menos conocidos que el sistema PROSPECTOR.

2.2.- MODELO DE FACTORES DE CERTEZA. (MODELO MYCIN).

MYCIN es un sistema experto basado en reglas, desarrollado hacia el año 1975 en la Universidad de Stanford y su objetivo es proporcionar un diagnóstico en enfermedades de origen infeccioso, así como la terapia adecuada para cada paciente.

La base de conocimientos del sistema MYCIN está formado por un conjunto de reglas de la forma "si e entonces h" donde, el antecedente, a su vez, puede estar formado por la conjunción de varias premisas. En su desarrollo inicial constaba de unas 200 reglas que hacia 1978 fueron ampliadas a casi 500, la mayoría de las cuales tenían relación con enfermedades de tipo meningitis.

Como ya se comentó al introducir los diferentes modelos de razonamiento probabilístico, fueron los autores del sistema MYCIN los que con mayor énfasis argumentaron en contra de la utilización de la probabilidad clásica, y en particular de la regla de Bayes, como modelo válido para el problema del diagnóstico en el campo de la Medicina.

En el trabajo en que Shortliffe y Buchanan (S-B en adelante) exponen su modelo de razonamiento basado en los factores de certeza, (SHOR-75), comienzan por explicitar las dificultades que aparecen al intentar construir un razonamiento partiendo del teorema de Bayes y que ya se indicaron en el apartado 2.1.1. Antes de definir de manera formal los distintos parámetros que servirán de base a su modelo de razonamiento, se plantean el clarificar el sentido que pueden tener los números que el experto proporciona junto con una regla "si e entonces h, x", donde x refleja de forma subjetiva el juicio del experto sobre la intensidad o fuerza con que la observación de la evidencia e "apoya" la hipótesis h. Con este objeto pasan revista a las distintas teorías de probabilidad propuestas y prestan especial atención a la "teoría de la confirmación" desarrollada por Carnap entre 1950 y 1962, como el trasfondo filosófico adecuado en que basar las medidas que utilizarán.

S-B justifican la adecuación de esta teoría de la confirmación al razonamiento basado en los juicios cuantitativos emitidos por los expertos por contraposición a la interpretación probabilística de los números que cuantifican una regla. Por ejemplo si un experto proporciona una regla "si e entonces h, 0'7" y se interpreta el valor 0'7 como una probabilidad condicional $P(h|e) = 0'7$, se tendría $P(\bar{h}|e) = 1 - 0'7 = 0'3$. Es decir una misma evidencia e sirve al mismo tiempo como apoyo de una hipótesis h y de su negación, lo cual resulta paradójico y difícilmente aceptable por el propio experto, que alegaría que e es una evidencia en favor de h pero en absoluto puede ser tomada simultáneamente como evidencia en contra. El valor x de un regla "si e entonces h, x" interpretado como grado de confirmación parece ser más apropiado que la interpretación del mismo como probabilidad condicional. Al mismo tiempo, parece surgir la necesidad de dos medidas diferentes, una para las evidencias que confirman una hipótesis y otra para aquéllas que disminuyen su credibilidad.

Una vez que se expongan con detalle los parámetros que utilizan S-B en su modelo, esta aparente paradoja será analizada, así como diferentes críticas que se han hecho al modelo.

2.2.1.- Representación del conocimiento incierto.

Para cuantificar la incertidumbre de las reglas y hechos que constituyen la base de conocimientos, Shortliffe y Buchanan proponen la utilización de lo que llaman factores de certeza, en vez de recurrir a las probabilidades condicionales. Estos números son obtenidos del experto, y para clarificar su significado, proponen una definición formal, así como un conjunto de consideraciones para mostrar la adecuación de la misma.

Considérese una regla del tipo "si e entonces h". Debido a la necesidad antes apuntada de no confundir el apoyo que una evidencia e presta a una hipótesis h con el apoyo que esa misma evidencia presta a la negación de la misma hipótesis, se utilizan dos medidas por separado, que se llaman grado de credibilidad y grado de incredibilidad ("belief" y "disbelief" en el trabajo original), que se notarán por MB(h,e) y MD(h,e) respectivamente, y cuyo significado es el siguiente:

MB(h,e) = medida del incremento de la credibilidad en la hipótesis h, dada la evidencia e.

MD(h,e) = medida del incremento de la incredibilidad en h, conocida como cierta la evidencia e.

Es decir, se manejan por separado las evidencias que aumentan la credibilidad en la hipótesis h de aquellas evidencias que la disminuyen. En un principio el motivo de esta separación era asegurar la conmutatividad al combinar distintas evidencias que tienen relación con la misma hipótesis pero, como se explicará más adelante, posteriores modificaciones que se introdujeron en la correspondiente función de combinación, hacen innecesario este tratamiento diferenciado.

Las características de estas dos medidas se obtienen a partir de la siguiente propuesta que se hace para una definición formal de ambas: Sea P(h) la credibilidad previa en la hipótesis h antes de tomar en consideración la evidencia e y P(h|e) la credibilidad en h una vez comprobada la verdad de e. 1-P(h) podrá entonces considerarse como la incredibilidad previa en h en cuanto a la verdad de la hipótesis. Según los valores relativos de P(h) y P(h|e) se consideran los siguientes casos:

1.- $P(h|e) > P(h)$:

En esta situación aumenta la credibilidad en h. Por consiguiente MB(h,e) será mayor que cero mientras que, debido a que una única evidencia no puede simultáneamente aumentar y disminuir la credibilidad en una misma hipótesis, MD(h,e) = 0.

Puesto que el incremento de la credibilidad en h, $P(h|e) - P(h)$, coincide con la disminución de la incredibilidad en h:

$$(1-P(h)) - (1-P(h|e)) = P(h|e) - P(h)$$

se propone como medida del incremento de la credibilidad de h la disminución relativa de la incredibilidad de h:

$$MB(h,e) = \frac{P(h|e) - P(h)}{1 - P(h)} \qquad (1.a)$$

2.- $P(h|e) < P(h)$:

En este caso se produce una disminución de la credibilidad de h, o lo que es lo mismo, un aumento de la incredibilidad de h. Es decir $MD(h,e)$ 0 y, por las mismas razones anteriores, $MB(h,e) = 0$. S-B definen el incremento de la incredibilidad de h como la disminución relativa de la credibilidad en h:

$$MD(h,e) = \frac{P(h) - P(h|e)}{P(h)} \qquad (1.b)$$

3.- $P(h|e) = P(h)$:

La evidencia e ni confirma ni desfavorece la hipótesis h y en consecuencia $MB(h,e) = 0$ y $MD(h,e) = 0$.

A partir de estas dos medidas MB y MD, se define el factor de certeza $CF(h,e)$ como:

$$CF(h,e) = MB(h,e) - MD(h,e)$$

es decir, si e representa un conjunto de evidencias relacionadas con la hipótesis h, $CF(h,e)$ será el resultado del cómputo total, agrupando aquéllas que favorecen la hipótesis en $MB(h,e)$ y las que van en su contra en $MD(h,e)$.

Para el caso de una única evidencia, y de acuerdo con lo expuesto hasta aquí:

$$CF = \begin{cases} \dfrac{P(h|e) - P(h)}{1 - P(h)} & \text{si } P(h|e) > P(h) \\[2ex] \dfrac{P(h|e) - P(h)}{P(h)} & \text{si } P(h|e) \leq P(h) \end{cases}$$

En los casos extremos S-B proponen:

.) si $P(h) = 0$ entonces $MD(h,e) = 1$, $MB(h,e) = 0$ y por consiguiente $CF(h,e) = -1$.

.) si $P(h) = 1$ entonces $MB(h,e) = 1$, $MD(h,e) = 0$ y $CF(h,e) = 1$.

En este intento de dar una definición formal de las medidas que se utilizarán para cuantificar la incertidumbre empiezan ya a aparecer algunas contradicciones dentro de la propia definición. Una de ellas es clara: según estos valores propuestos, para las situaciones en que h es conocido como verdadero o falso antes de considerar cualquier evidencia, $P(h) = 1$ ó $P(h) = 0$. En estos casos, cualquiera que sea la evidencia e considerada, no hay incremento en la credibilidad o incredibilidad de h. Existe pues una contradicción entre los valores $MB(h,e) = 1$ o $MD(h,e) = 1$ propuesto y el significado que se pretende dar a través de las definiciones (1.a) y (1.b) de las medidas MB y MD como incremento de credibilidad e incredibilidad. A medida que se vaya avanzando en la exposición del modelo tal y como fue propuesto se verán las razones profundas de ésta y otras contradicciones.

Por otra parte parecen algo artificiosas las expresiones propuestas para MB y MD. MB, medida del incremento de la credibilidad de h se hace igual a la disminución relativa de la incredibilidad h, expresada en términos de probabilidades. Sería bastante más lógico definirla como incremento proporcional de la credibilidad en h, si ése va a ser su significado. Similares razonamientos deberían tenerse en cuenta para la medida del incremento de la incredibilidad MD. Siguiendo estas consideraciones se tendría:

$$MB(h,e) = \frac{P(h|e) - P(h)}{P(h)} \qquad MD(h,e) = \frac{P(h) - P(h|e)}{1-P(h)} \tag{2}$$

Las razones para definir de la forma ya indicada, (1.a) y (1.b), no se explican en el trabajo de S-B y su justificación se debería a un doble motivo. Por un lado, como consecuencia de adoptar las definiciones (1), se obtiene una forma clara para el experto de cuantificar la incertidumbre de las reglas, y además estas definiciones permiten resolver distintos problemas que se plantearon los autores al considerar los inconvenientes de los modelos probabilísticos. Ninguno de estos dos aspectos parece resolverse utilizando las definiciones (2).

La clarificación de esta doble motivación implícita, así como el adecuado funcionamiento de las tres medidas MB, MD y CF son puestas de manifiesto por los autores de MYCIN mediante las siguientes consideraciones:

1.- Intervalos de variación:

$$0 \leq MB \leq 1, \quad 0 \leq MD \leq 1, \quad -1 \leq CF \leq 1$$

2.- Comportamiento en casos extremos:

a) Si $P(h|e) = 1$, es decir, la evidencia $\underline{e}$ confirma la hipótesis $\underline{h}$ con total certeza:

$$MB = 1, \; MD = 0 \text{ y } CF = 1$$

b) Si $P(h|e) = 0$, es decir, $\underline{h}$ es falso como consecuencia de la observación de $\underline{e}$:

$$MB = 0, \; MD = 1 \text{ y } CF = -1$$

c) $P(h|e) = P(h)$, $\underline{e}$ no aporta nada al conocimiento de $\underline{h}$:

$$MB = 0, \; MD = 0 \text{ y } CF = 0$$

3.- Los factores de certeza son diferentes a las probabilidades condicionales. Utilizando probabilidades se tiene que $P(\overline{h}|e) = 1-P(h|e)$, o lo que es lo mismo, $P(h|e) + P(\overline{h}|e) = 1$. Los factores de certeza no satisfacen esta igualdad. Más concretamente:

$$CF(\overline{h},e) = - CF(h,e)$$

es decir, los factores de certeza de las hipótesis h y $\bar{h}$ no son complementarios a la unidad sino opuestos. Este comportamiento parece bastante claro, pues si el apoyo que una evidencia presta a una hipótesis es bajo, no debería ser alto el apoyo a la negación de dicha hipótesis en el caso de que exista un grado importante de ignorancia, sino que el apoyo a ambas debe ser igualmente bajo.

La comprobación de este resultado se hace para el caso $P(h|e) > P(h)$, siendo exactamente igual cuando $P(h|e) < P(h)$: si $P(h|e) > P(h)$ se tiene $P(\bar{h}|e) = 1-P(h|e) < 1-P(h) = P(\bar{h})$ y, en consecuencia,

$$MB(\bar{h},e) = 0, \quad MD(\bar{h},e) > 0$$

$$MB(h,e) > 0, \quad MD(h,e) = 0 \quad \text{y}$$

$$CF(\bar{h},e) = MB(\bar{h},e) - MD(\bar{h},e) = 0 - \frac{P(\bar{h}) - P(\bar{h}|e)}{P(\bar{h})} =$$

$$= \frac{P(\bar{h}|e) - P(\bar{h})}{P(\bar{h})} = \frac{1 - P(h|e) - (1-P(h))}{1-P(h)} =$$

$$= - \frac{P(h|e) - P(h)}{1-P(h)} = - MB(h,e) + 0 = - CF(h,e)$$

Queda así resuelto el problema ya varias veces comentado de que si e presta apoyo a una hipótesis h con una "fuerza" x, la consideración de este número x como probabilidad condicional $P(h|e)$ obliga a que $P(\bar{h}|e) = 1-x$, es decir, la misma evidencia e se constituye como evidencia respecto a la negación de la hipótesis.

Es consecuencia de este comportamiento como S-B dan ya un significado a los números que debe dar el experto cuando es preguntado sobre la fuerza con que una evidencia apoya o niega una hipótesis. El experto proporciona un número positivo, $CF > 0$, si la evidencia apoya la hipótesis; un número negativo si apoya la negación de la hipótesis, $CF < 0$, y un valor 0 si la evidencia e es independiente de la hipótesis h. Más concretamente, la pregunta que se formula al experto para obtener el factor de certeza es: "Es una escala de 1 a 10, ¿cuál es la certidumbre que usted asigna a esta conclusión?" (BUCH-84), pag. 238. Dicho número, una vez pasado a decimales, será el CF asignado a la regla. Por tanto los valores 1 y -1 corresponderán a absoluta certeza en la verdad o falsedad de h.

Obsérvese que, como consecuencia de la consideración 3 hecha anteriormente, una gran mayoría de las reglas existentes en la base de conocimientos del sistema MYCIN van afectadas de factores de certeza positivos, puesto que si una evidencia e niega una hipótesis h, lo que aparece en la conclusión de la regla es la negación h. De ahí que la pregunta al experto se haga en la escala 1 a 10, con valores sólo positivos.

Es posible ya descubrir de dónde provienen las contradicciones antes apuntadas. Como señalan Horvitz y Heckerman, (HORV-86a), (HORV-86b), hay una confusión entre la definición que se propone para las medidas MB y MD, que claramente son medidas de actualización de credibilidad ("belief updating") y en consecuencia también los factores de certeza CF, y el uso que se hace de estos últimos como medida de credibilidad absoluta ("absolute belief"). Se aprecia de forma clara que los factores de certeza, dada la manera de obtenerlos del experto, son valores de credibilidad absoluta, en contradicción con la definición propuesta.

Otras consideraciones que S-B hacen acerca del comportamiento de los factores de certeza son las siguientes:

4.- Resolución satisfactoria de la paradoja de los cuervos. Esta paradoja fue presentada por Hempel, (HEMP-65), en sus estudios sobre confirmación, para mostrar las diferencias entre este concepto y el de probabilidad. Brevemente, viene a decir que si dos proposiciones h_1 y h_2 son lógicamente equivalentes, entonces, para cualquier evidencia e deberá cumplirse que $P(h_1|e) = P(h_2|e)$. Considera la dos proposiciones h_1 = "todos los cuervos son negros" y h_2 = "todos los objetos no negros no son cuervos". Por la lógica elemental del cálculo de predicados, estas dos proposiciones son lógicamente equivalentes. Sin embargo, parece difícil aceptar que la observación de la evidencia e = "vaso verde" preste el mismo apoyo a las proposiciones h_1 y h_2. S-B señalan que su definición de factor de certeza resuelve de forma satisfactoria esta paradoja, resolución que puede encontrarse con detalle en (SHOR-76).

5.- Si se considera un conjunto de n hipótesis h_i mutuamente excluyentes, el conjunto de factores de certeza $CF(h_i,e)$ no puede ser arbitrario, sino que su suma debe respetar unos ciertos límites.

En cuanto al límite inferior, si k es el número de hipótesis de entre las n h_i que son parcial o totalmente negadas por una evidencia e, se tendrá obviamente $\sum CF(h_i,e) \geqslant -k$. Obsérvese que si como consecuencia de la evidencia e, es confirmada con total certeza una de las hipótesis, automáticamente todas las demás serán falsas con certeza y, por consiguiente, para estas hipótesis negadas por e, $\sum CF(h_i,e) = -(n-1)$.

Respecto a un límite superior para la suma de los factores de certeza de las hipótesis h_i a las que e presta apoyo, debe verificarse:

$$\sum CF(h_i,e) \leqslant 1$$

La demostración de este resultado, basada en la definición formal de los factores de certeza puede verse en (SHOR-76). En coherencia con este límite superior, los valores dados por el experto a los factores de certeza de distintas reglas cuyas conclusiones sean hechos mutuamente excluyentes deben ser tales que respeten este límite, lo cual no tiene porqué ocurrir y en cuyo caso deberán ser modificados, bien por el propio experto o mediante algún proceso de normalización.

Para concluir, se debe señalar que, aunque hasta ahora tan sólo se han considerado factores de certeza asociados a reglas del tipo "si e entonces h", MYCIN considera también factores de certeza para hechos, con el fin de poder evaluar las premisas de cada regla en el proceso de propagación de incertidumbre. En este caso el valor correspondiente será proporcionado por el usuario al ser preguntado por él en el transcurso de una sesión de consulta en el sistema.

Expuesto el tipo de parámetros que se consideran en el modelo MYCIN de factores de certeza, se pasa a exponer los mecanismos de propagación de incertidumbre que S-B proponen en su modelo de razonamiento. Una vez hecho esto se volverá a discutir el sentido de estos factores de certeza, así como las muy diversas críticas que se le han formulado.

2.2.2.- Combinación de evidencias.

Supóngase un conjunto de reglas con la misma conclusión, "si e_i entonces h", cada una de ellas con su correspondiente factor de certeza $CF(h,e_i)$. Se trata de nuevo de obtener $CF(h,e)$ donde e significa el conjunto de todas las evidencias e_i.

S-B propusieron primeramente la siguiente función de combinación:

$$MB(h,e_1 \wedge e_2) = \begin{cases} 0 \text{ si } MD(h,e_1 \wedge e_2) = 1 \\ MB(h,e_1) + MB(h,e_2) - MB(h,e_1) \cdot MB(h,e_2) \end{cases}$$

$$MD(h,e_1 \wedge e_2) = \begin{cases} 0 \text{ si } MB(h, e_1 \wedge e_2) = 1 \\ MD(h,e_1) + MD(h,e_2) - MD(h,e_1) \cdot MD(h,e_2) \end{cases}$$

Expresadas en términos de factores de certeza se obtiene:

a) e_1 y e_2 confirman la hipótesis h:

$$CF(h,e_1) = MB(h,e_1) \qquad MD(h,e_1) = 0$$

$$CF(h,e_2) = MB(h,e_2) \qquad MD(h,e_2) = 0$$

$$MD(h,e_1 \wedge e_2) = 0$$

$$MB(h,e_1 \wedge e_2) = CF(h,e_1) + CF(h,e_2) - CF(h,e_1) \cdot CF(h,e_2)$$

$$\boxed{CF(h,e_1 \wedge e_2) = CF(h,e_1) + CF(h,e_2) - CF(h,e_1) \cdot CF(h,e_2)}$$

b) e_1 y e_2 apoyan la negación de $\underline{h}$. Nótese que los dos factores de certeza $CF(h,e_1)$ y $CF(h,e_2)$ son ahora negativos:

$$CF(h,e_1) = -MD(h,e_1) \qquad MB(h,e_1) = 0$$

$$CF(h,e_2) = -MD(h,e_2) \qquad MB(h,e_2) = 0$$

$$MB(h,e_1 \wedge e_2) = 0$$

$$MD(h,e_1 \wedge e_2) = -CF(h,e_1) - CF(h,e_2) - (-CF(h,e_1)) \cdot (-CF(h,e_2))$$

$$\boxed{CF(h,e_1 \wedge e_2) = CF(h,e_1) + CF(h,e_2) + CF(h,e_1) \cdot CF(h,e_2)}$$

c) Una evidencia e_1 presta apoyo a la hipótesis $\underline{h}$, mientras que la otra e_2 es evidencia en contra de $\underline{h}$:

$$MB(h,e_1) = CF(h,e_1) \qquad MD(h,e_1) = 0$$

$$MD(h,e_2) = -CF(h,e_2) \qquad MB(h,e_2) = 0$$

$$MB(h,e_1 \wedge e_2) = CF(h,e_1) + 0 - CF(h,e_1).0 = CF(h,e_1)$$

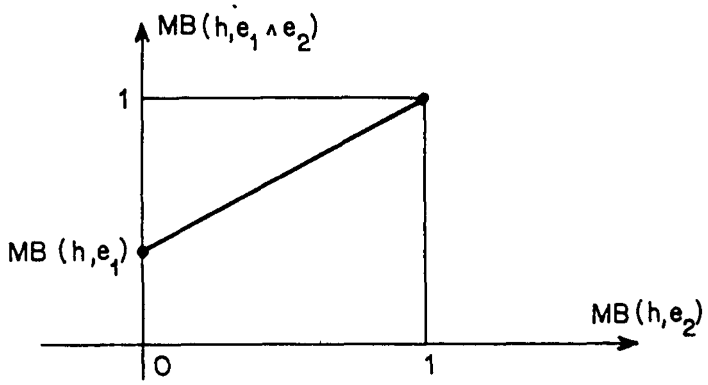
$$MD(h,e_1 \wedge e_2) = 0 + (-CF(h,e_2)) - 0.(-CF(h,e_2)) = -CF(h,e_2)$$

$$\boxed{CF(h,e_1 \wedge e_2) = CF(h,e_1) + CF(h,e_2)}$$

Sin embargo, la combinación de evidencias debe efectuarse obteniéndose por separado cada uno de los valores MB(h,e) y MD(h,e) pues se puede comprobar que con la función de combinación expresada con factores de certeza no se verifica la asociatividad.

Esta función de combinación presenta un comportamiento coherente con los conceptos de MB, MD y CF, como se puede apreciar por las siguientes consideraciones:

1.- Si se denota con e_+ el conjunto de evidencias que favorecen la hipótesis <u>h</u>, y con e_- aquéllas que disminuyen la credibilidad en <u>h</u>, se tiene que $MB(h,e_+)$ va aumentando hacia la unidad a medida que nuevas unidades de evidencia van combinándose y que alcanza el valor 1 si y sólo si una unidad de evidencia confirma la hipótesis <u>h</u> con total certeza. Este hecho es fácilmente observable si se representa gráficamente $MB(h,e_1 \wedge e_2)$ en función de $MB(h,e_2)$ considerando conocido $MB(h,e_1)$:



Igual comportamiento se tiene para $MD(h,e_-)$ por idénticas razones.

2.- La combinación de evidencias, conservando por separado las medidas MB y MD es conmutativa (independiente del orden en que se van considerando las distintas unidades de evidencia) y asociativa.

3.- Si se expresa la función de combinación de la siguiente manera:

$$MB(h,e_2) = \frac{MB(h,e_1 \wedge e_2) - MB(h,e_1)}{1 - MB(h,e_1)}$$

se obtiene una forma similar a la definición originada la medida de credibilidad MB, $MB(h,e) = \dfrac{P(h|e) - P(h)}{1 - P(h)}$, reemplazando probabilidades por MB'_s.

Expresada de otra forma:

$$MB(h,e_1 \wedge e_2) = MB(h,e_1) + MB(h,e_2) \cdot (1 - MB(h,e_1))$$

se puede interpretar, puesto que MB representa disminución relativa en la incredibilidad, como que la nueva medida MB resultado de la consideración de una nueva evidencia e_2 se verá aumentada proporcionalmente a la incredibilidad existente antes de considerar e_2.

Esta función de combinación propuesta inicialmente ha sido sin embargo modificada posteriormente debido al fuerte impacto que una única evidencia en contra de una hipótesis tenía frente a varias evidencias favoreciendo esa misma hipótesis. Considérense, por ejemplo, ocho o nueve reglas en favor de la hipótesis $\underline{h}$ con factores de certeza entre 0'4 y 0'8 y una evidencia en contra de $\underline{h}$ con un factor de certeza de 0'6. Debido a la rapidez con que converge asintóticamente hacia la unidad la función de combinación cuando ambas evidencias apoyan la hipótesis, al combinar los factores de certeza de las reglas a favor de $\underline{h}$ se obtendría un factor de certeza del orden de 0'99 y al combinar con la única evidencia que disminuye su credibilidad, el factor de certeza resultante se rebaja sustancialmente obteniéndose un valor de 0'39. Claramente se pondera excesivamente sobre todo el conjunto el efecto negativo de la única evidencia en contra de $\underline{h}$. Para mitigar este impacto se modificó la función de combinación para el caso de dos evidencias de efectos contrarios de la siguiente manera, que es la que se utiliza normalmente:

Llamando $CF(h,e_1) = x$ y $CF(h,e_2) = y$

$$CF(h, e_1 \wedge e_2) = \begin{cases} x+y - x.y & \text{si } x \text{ e } y > 0 \\ \dfrac{x.y}{1 - \min(|x|,|y|)} & \text{si } x \text{ e } y \text{ son de distinto signo} \\ x+y + x.y & \text{si } x \text{ e } y < 0 \end{cases}$$

Se ha expresado ya directamente esta nueva función de combinación en términos de factores de certeza porque de esta manera sí es asociativa y se pueden combinar las evidencias en cualquier orden. De forma accidental se obtiene una ventaja adicional puesto que no es preciso almacenar MB y MD a lo largo del proceso.

Otro aspecto importante a destacar respecto a esta forma de combinar evidencias es que implícitamente se hace la suposición de independencia condicional de las evidencias e_1 y e_2. Sin embargo, si por ejemplo e_1 implica lógicamente e_2, entonces se debería tener $CF(h,e_1 \wedge e_2) = CF(h,e_1)$ cualquiera que sea el valor de $CF(h,e_2)$. Al utilizar la función de combinación sin tener en cuenta esta dependencia de e_1 y e_2 se obtiene un valor en el que de alguna forma se ha tenido en cuenta dos veces la evidencia e_2. Los autores de MYCIN no dan solución para problemas de este tipo, salvo recomendaciones de estructurar cuidadosamente la base de conocimientos, agrupando en una sóla regla aquellas unidades de evidencias condicionalmente dependientes, en vez utilizar varias reglas, o prescribir la utilización del modelo de factores de certeza sólo en aquellos campos en que se pueda considerar la independencia condicional en la mayoría de los casos.

2.2.3.- Propagación de incertidumbre.

En todo lo desarrollado hasta este momento, al exponer y razonar sobre MB, MD y CF(h,e) se ha considerado la evidencia e como un hecho cierto y en función de la observación de su verdad se obtenía la credibilidad o incertidumbre de la hipótesis h. Sin embargo, la situación habitual será que el propio hecho e esté también afectado de incertidumbre en su conocimiento. Esto puede ser debido a dos motivos diferentes: la estructura de control del sistema ha preguntado al usuario sobre su conocimiento acerca de e y éste no tiene una certeza total sobre su verdad o falsedad y, su respuesta viene expresada por medio de un factor de certeza, o bien porque la evidencia e es a su vez conclusión de otra regla, es decir, se produce un encadenamiento de las dos reglas "si e' entonces e, CF(e,e') = x" y "si e entonces h, CF(h,e) = y":

$$e' \longrightarrow e \longrightarrow h$$

Sin ofrecer ninguna justificación, S-B proponen para la obtención del factor de certeza conjunto, CF(h,e'), las siguientes relaciones:

$$MB(h,e') = MB(h,e) \cdot \max(0, CF(e,e'))$$

$$MD(h,e') = MB(h,e) \cdot \max(0, CF(e,e'))$$

manteniendo el criterio de utilizar separadamente las evidencias a favor y en contra de la hipótesis h, aunque en esta situación es innecesaria esta separación. En términos de CF se pueden agrupar ambas relaciones en la fórmula:

$$CF(h,e') = CF(h,e) \;.\; \max\,(0,\; CF(e,e')) \qquad (3)$$

Esta relación es consistente con lo que sucedería en el caso de implicación total en cada una de las dos reglas y, según apunta Adams, (ADAM-76), puede venir sugerida por la noción intuitiva de que si $\underline{e'}$ implica $\underline{e}$ con una probabilidad $P(e|e')$ y $\underline{e}$ implica $\underline{h}$ con probabilidad $P(h|e)$ entonces $\underline{e'}$ implica $\underline{h}$ con una probabilidad $P(h|e') = P(h|e) \;.\; P(e|e')$, aunque esto sólo es cierto en unas condiciones muy particulares.

Para aquellos casos en que $CF(e,e')$ es negativo, es decir $\underline{e'}$ apoya la negación de $\underline{e}$, la fórmula (3) proporciona $CF(h,e') = 0$, es decir, no se puede afirmar nada acerca de la hipótesis $\underline{h}$. Heckerman (HECK-86), propone en esta situación otro esquema basándose en la relación $CF(\bar{h},e) = -CF(h,e)$ que, para los hechos $\underline{e}$ y $\underline{e'}$ se traduciría en $CF(\bar{e},e') = -CF(e,e')$, proporcionando un valor $CF(\bar{e},e')$ positivo. Considera en consecuencia el esquema

$$e' \longrightarrow \bar{e} \longrightarrow h$$

y aplicando la relación (3) se obtiene:

$$CF(h,e') = CF(h,\bar{e}) \;.\; \max\,(0,\; CF(\bar{e},e'))$$

En conjunto, la propuesta de Heckerman se resume en:

$$CF(h,e') = \begin{cases} CF(h,e) \;.\; CF(e,e') & \text{si } CF(e,e') \geqslant 0 \\ -CF(h,\bar{e}) \;.\; CF(e,e') & \text{si } CF(e,e') < 0 \end{cases}$$

Constituye una ampliación estricta del esquema de S-B y es aplicable a aquellos casos en que se conoce también el factor de certeza $CF(h,\bar{e})$, o lo que es lo mismo, en la base de conocimientos, además de la regla "si $\underline{e}$ entonces $\underline{h}$" se incluye también la regla "si no $\underline{e}$ entonces $\underline{h}$", cada una con sus correspondientes factores de certeza. Su comportamiento en situaciones extremas es correcta, y así, se tiene:

$$\text{si } CF(e,e') = 1 \quad \text{entonces} \quad CF(h,e') = CF(h,e)$$

$$\text{si } CF(e,e') = -1 \quad \text{entonces} \quad CF(h,e') = CF(h,\bar{e})$$

$$\text{si } CF(e,e') = 0 \quad \text{entonces} \quad CF(h,e') = 0$$

Con respecto a los resultados previsibles por la aplicación de esta fórmula, conviene señalar que en las validaciones que se hicieron del sistema MYCIN, una parte de los experimentos consistió en comparar los valores que se obtenían utilizando datos estadísticos y la definición de CF por un lado, y por otro mediante las distintas funciones de combinación. Como conclusión, se llegó a que los valores obtenidos de ambas formas no diferían significativamente, presentándose las mayores divergencias cuando existían largas cadenas de razonamiento.

2.2.4.- Combinación lógica de evidencias.

El modelo MYCIN de los factores de certeza proporciona también funciones para obtener, por intermedio de las medidas MB y MD, los factores de certeza de la conjunción o disyunción de hechos o proposiciones; $CF(h_1 \wedge h_2,e)$ y $CF(h_1 \vee h_2,e)$, a partir de los valores $CF(h_1,e)$ y $CF(h_2,e)$. La necesidad de estas funciones viene originada por la siguiente razón.

La estructura de las reglas de la base de conocimientos de MYCIN permite que el antecedente de las mismas sea la conjunción de varios hechos o premisas y, de hecho, la mayoría de ellas son de esa forma. Además, cada una de esas premisas puede ser a su vez disyunción de otras varias. Por tanto, será frecuente encontrarse con reglas de tipo

$$\text{si } (e_1 \wedge (e_2 \vee e_3) \wedge e_4) \quad \text{entonces } h, \quad (CF = x)$$

Para poder aplicar esta regla, según se vió en el apartado anterior, se deberá evaluar el antecedente de la misma y, caso de cumplirse en un cierto grado, se podrá deducir la conclusión con su correspondiente factor de certeza. Para la regla ilustrativa anterior, si $\underline{e}$ representa el conocimiento existente en el momento de la evaluación, se tendría:

$$CF(h,e) = x \cdot CF((e_1 \wedge (e_2 \vee e_3) \wedge e_4), e)$$

S-B proponen para el cálculo de factores de certeza de conjunciones y disyunciones las siguientes funciones:

$$MB(h_1 \wedge h_2,e) = \min (MB(h_1,e), MB(h_2,e))$$

$$MD(h_1 \wedge h_2,e) = \max (MD(h_1,e), MD(h_2,e))$$

$$MB(h_1 \vee h_2,e) = \max (MB(h_1,e), MB(h_2,e))$$

$$MD(h_1 \vee h_2,e) = \min (MD(h_1,e), MD(h_2,e))$$

Es fácil comprobar, a partir de la definición de factor de certeza, que dichas fórmulas, en términos de CF se convierten en:

$$CF(h_1 \wedge h_2,e) = \min (CF(h_1,e), CF(h_2,e))$$

$$CF(h_1 \vee h_2,e) = \max (CF(h_1,e), CF(h_2,e))$$

que son análogas a las utilizadas en PROSPECTOR y que serán las relaciones utilizadas por el sistema.

Adams, (ADAM-76), señala que detrás de estas fórmulas subyacen fuertes suposiciones y hacer notar algunos casos extremos, como por ejemplo que si h_1 y h_2 son hipótesis mutuamente excluyentes, entonces la conjunción $h_1 \wedge h_2$ es falsa con total certeza, independientemente de la credibilidad de cada una de ellas, en contra de lo propugnado por la fórmulas anteriores. No suele ser éste el caso, pues en MYCIN las hipótesis h_1 y h_2 son normalmente proposiciones relativas a parámetros clínicos o, más claramente, universos de discurso, diferentes. En cualquier caso, vuelve a aparecer la idea de que es la esmerada estructuración de la base de conocimientos el factor primordial para el éxito del sistema, mucho más que el procedimiento para propagar y transmitir unos valores numéricos de incertidumbre.

2.2.5.- Críticas al modelo MYCIN.

MYCIN es un sistema experto ya antiguo, muy divulgado y el modelo de factores de certeza propuesto como modelo para el razonamiento aproximado es muy popular y ha sido ampliamente utilizado en diversos sistemas expertos. Como consecuencia de todas estas circunstancias, ha generado una amplísima literatura destacando uno u otro aspecto del modelo, relacionándolo con la teoría de la probabilidad, etc. Entre todos estos trabajos merecen destacarse los debidos a Heckerman y Horvitz, (HORV-86a), (HORV-86b) y (HECK-86), por los planteamientos que proponen, y puesto que ponen de manifiesto muchas de las carencias y suposiciones implícitas en el modelo MYCIN. Esto es importante de cara al estudio de la posibilidad de aplicación del modelo de los factores de certeza en un campo determinado.

Heckerman hace notar en primer lugar que no hay relación entre la definición propuesta para factor de certeza y las distintas funciones de combinación utilizadas. Los factores de certeza son definidos, a través de las probabilidades a priori $P(h)$ y a posteriori $P(h|e)$, con la intención de reflejar el deseo de sus creadores de manejar una medida de cambio en la credibilidad. Sin embargo las funciones de combinación que Shortliffe y Buchanan proponen no se deducen de esta definición, sino que constituyen un procedimiento "ad-hoc" para combinar en distintas situaciones los factores de certeza.

Como consecuencia de no estar ligadas las funciones de combinación con la definición de factor de certeza, aparecen importantes inconsistencias que han sido señaladas por diversos autores.

Heckerman señala, por ejemplo, que de la definición de CF se deduce que la combinación de evidencias relativas a la misma hipótesis no es conmutativa, como se comprueba con el siguiente razonamiento sencillo. De los trabajos originales de S-B puede suponerse que los factores de certeza tan sólo dependen de h y e, es decir de la hipótesis cuya credibilidad se estudia y de la evidencia que provoca un cambio en la credibilidad de aquélla, sin considerar la influencia de cualquier otra evidencia e' conocida en el momento de considerar la evidencia e. Heckerman expresa este hecho de la forma siguiente

$$CF(h,e,e') = CF(h,e,\phi)$$

y llama a esta suposición axioma de modularidad, señalando las importantes consecuencias que tiene esta suposición. Con esta notación la definición de factor de certeza se reformularía como

$$CF(h,e,e') = \begin{cases} \dfrac{P(h|e \wedge e') - P(h|e')}{1 - P(h|e')} & \text{si } P(h|e \wedge e') \geqslant P(h|e') \\[2ex] \dfrac{P(h|e \wedge e') - P(h|e')}{P(h|e')} & \text{si } P(h|e \wedge e') \leqslant P(h|e') \end{cases}$$

Considérense a continuación dos evidencias e_1 y e_2 relacionadas con una misma hipótesis h, con unos valores $CF(h,e_1) = 0'9$ y $CF(h,e_2) = -0'9$. Supóngase además una probabilidad a priori $P(h|e') = 0'5$.

1.- Se combinan e_1 y e_2 en este orden. De la expresión anterior de $CF(h,e,e')$ se deduce $P(h|e_1 \wedge e') = 0'95$. En el momento de considerar la evidencia e_2, las evidencias previas serán $e_1 \wedge e'$ y $P(h|e_1 \wedge e')$ el valor de la probabilidad a priori a considerar. Nuevamente utilizando la definición de CF se obtiene $P(h|e_2 \wedge e_1 \wedge e') = 0'10$.

2.- Si ahora combinamos e_2 y a continuación e_1 se obtiene, $P(h|e_2 \wedge e') = 0'05$ y $P(h|e_1 \wedge e_2 \wedge e') = 0'90$.

Los valores obtenidos son muy diferentes. La definición de factor de certeza, en consecuencia, no asegura la conmutatividad en la combinación de evidencias.

Obsérvese que para todos estos cálculos ha sido necesario el conocimiento de la probabilidad a priori de h. S-B en su modelo no consideran en ningún momento las probabilidades a priori de cada hipótesis y, como han señalado diversos autores, su no consideración lleva a confusiones y apreciaciones incorrectas.

Así por ejemplo Adams, (ADAM-76), señala que hipótesis h_1 y h_2 tales que sus probabilidades a priori fueran muy diferentes, $P(h_1) \gg P(h_2)$ por ejemplo, y con probabilidades a posteriori también en la misma relación $P(h_1|e) > P(h_2|e)$, darían lugar a factores de certeza contrarios, $CF(h_1,e) < CF(h_2,e)$. Con unos sencillos números, si $P(h_1) = 0'8$, $P(h_2) = 0'2$ $P(h_1|e) = 0'9$ y $P(h_2|e) = 0'8$ se obtiene de la definición de CF, $CF(h_1,e) = 0'5$ y $CF(h_2,e) = 0'75$. Si, en cambio se toman $P(h_1) = P(h_2) = 0'5$ se obtiene $CF(h_1,e) = 0'8$, $CF(h_2,e) = 0'6$. Puesto que a la hora de seleccionar hipótesis MYCIN lo hace en función de los valores finales del CF de cada una de ellas, la no consideración de probabilidades a priori es un importante inconveniente.

Más importante es la confusión que apuntan Cheesman y Pearl, (CHEE-85), (PEAR-85), en relación con el problema de considerar simultáneamente una evidencia a favor y en contra de una misma hipótesis. S-B consideraban una regla del tipo "si e entonces h, 0'7", y argumentaban contra la interpretación del valor 0'7 como una probabilidad a posteriori, $P(h|e) = 0'7$, puesto que, de acuerdo con la teoría de probabilidades $P(\bar{h}|e) = 0'3$. La misma evidencia e, que en principio era una evidencia a favor de h, es ahora también evidencia en contra de h. La confusión reside en que 0'7 no debe interpretarse como una probabilidad absoluta, sino como un cambio en la credibilidad de h y la probabilidad a posteriori dependerá también del valor previo de la probabilidad de h antes de la observación de e. $P(\bar{h}|e) = 0'3$ no puede interpretarse como un aumento en la probabilidad de "no h" independiente de su valor inicial. El experto no tendría ninguna objección a un cambio en $P(\bar{h})$ de 0'9 a 0'3, pero sí sería contrario a una variación de $P(\bar{h})$ desde 0'1 a 0'3.

El esquema que plantea Heckerman parte, en contraposición con S-B, de considerar aquéllas propiedades que sería deseable cumplieran las medidas de credibilidad a utilizar, tanto en relación con su definición, como con las distintas formas de combinarse, en paralelo (evidencias relativas a una misma hipótesis) y secuencialmente (propagación de incertidumbre). Estas propiedades o axiomas los extrae Heckerman de las consideraciones que hace S-B para la definición de CF y de aquéllas propiedades que deberán verificar las funciones de combinación.

El resultado es sorprendente, pues Heckerman demuestra que las únicas medidas que satisfacen todas las propiedades deseadas son transformaciones $F(\lambda)$ de la razón de probabilidad $\lambda = P(e|h)/P(e|\bar{h})$, monótonas crecientes y tales que $F(1/x) = -F(x)$ y $\lim_{x \to \infty} F(x) = 1$.

Así por ejemplo, considerando la función

$$F_1(x) = \begin{cases} (x-1)/x & \text{si } x \geqslant 1 \\ x-1 & \text{si } x < 1 \end{cases}$$

obtiene un factor de certeza

$$CF_1 = \begin{cases} \dfrac{P(h|e) - P(h)}{P(h|e)\ (1 - P(h))} & \text{si } P(h|e) \geqslant P(h) \\[2ex] \dfrac{P(h|e) - P(h)}{P(h)\ (1 - P(h|e))} & \text{si } P(h|e) \leqslant P(h) \end{cases}$$

bastante similar a la definición de factor de certeza de MYCIN, pero que, a diferencia de éste, es consistente con todas las propiedades requeridas.

Además Heckerman, consigue descubrir algunas suposiciones implícitas en el modelo MYCIN. Así por ejemplo, demuestra que, como consecuencia del axioma de modularidad antes citado, para combinar varias evidencias relativas a una misma hipótesis h, deben ser condicionalmente independientes respecto de h y también de su negación, que raramente se verifica en aplicaciones prácticas. Asimismo prueba que, para que se mantenga la consistencia con las propiedades requeridas, la cadena de razonamiento debe tener estructura de árbol, sin caminos múltiples entre evidencias e hipótesis.

2.3.- TEORIA DE LA EVIDENCIA DE DEMPSTER-SHAFER.

Es a partir de principios de los ochenta cuando se empieza a prestar atención a lo que incluso ha llegado a llamarse Razonamiento Evidencial y en gran parte su difusión se debe a los autores del modelo MYCIN o de los factores de certidumbre, (GORD-84), (GORD-85) como un intento de superar las críticas que se hacían a su modelo, fundamentalmente su falta de base teórica y ser un modelo "ad hoc" cuyo éxito radicaba más bien en la cuidadosa estructuración de la base de conocimientos que en el propio sistema de tratamiento de la incertidumbre.

Lo que actualmente se conoce como Teoría de la Evidencia parte de los trabajos de Dempster, (DEMP-67a), (DEMP-68) posteriormente desarrollados por Shafer, (SHAF-76), en un texto que le da un fuerte fundamento teórico al método.

Su atractivo como posible método para el tratamiento de la incertidumbre se debe a los siguientes motivos:

1.- Simplicidad para capturar la incertidumbre del apoyo que las evidencias que van consiguiéndose en el razonamiento prestan a las hipótesis sobre las que se razona.

2.- A diferencia de los modelos probabilísticos bayesianos que asignan probabilidad o credibilidad de forma puntual a hipótesis individuales, el modelo evidencial permite considerar conjuntos de hipótesis a los que prestar apoyo, sin que la credibilidad depositada en ese conjunto deba ser repartida entre las hipótesis que lo componen, lo que se corresponde con la forma de razonamiento humano, especialmente en situaciones de diagnóstico, médico o industrial, donde un síntoma o fallo hace pensar en un conjunto de enfermedades o situaciones como posibles causantes pero sin poder discriminar entre ellas. Por tanto, parece muy adecuado para aquellos problemas en que sea posible trabajar con un conjunto de hipótesis jerarquizado o estructurado por niveles.

3.- Refleja muy bien la ignorancia o falta de conocimiento que se produce en un razonamiento incierto o incompleto.

4.- Siendo una extensión de la teoría de la probabilidad, supera algunas de las objecciones que se le hacen a ésta y además la contiene como caso particular, así como a la regla de combinación de evidencias propuesto por Shortliffe y Buchanan en su sistema MYCIN.

2.3.1.- Estructuración del conocimiento incierto.

El modelo evidencial de Dempster y Shafer (D-S en lo sucesivo) parte de un conjunto de hipótesis finito sobre el que razonar, que se llama marco de discernimiento, hipótesis que se suponen mutuamente excluyentes y exhaustivas y que se suele denotar por θ . Las afirmaciones o proposiciones que pueden considerarse en relación con el marco de discurso se corresponden con los subconjuntos de θ y se puede hablar de que se ha "discernido" dicha proposición por su correspondiente subconjunto.

A diferencia de los planteamientos probabilísticos, el efecto de una determinada evidencia sobre el conjunto de hipótesis no es reflejado en probabilidad o credibilidad de cada una de las hipótesis individuales, sino que se refleja en apoyo a conjuntos de hipótesis $A \subseteq \theta$, sin que el apoyo o soporte que dicha evidencia presta al subconjunto A puede ser repartido entre las hipótesis individuales que constituyen A.

Para cada evidencia considerada en el razonamiento se considera una función:

$$m : P(\theta) \longrightarrow [0,1]$$

de tal forma que para un subconjunto A formado por unas determinadas hipótesis de entre el total considerado θ , m(A) es la credibilidad o apoyo que esa evidencia presta exactamente al subconjunto A, sin que pueda precisarse más respecto a la probabilidad de cada una de las hipótesis de A. Desde la óptica de la teoría de la probabilidad esta cantidad m(A) se repartiría por igual entre todos los elementos de A, lo cual no refleja la situación real pues se supone igual probabilidad para todas esas hipótesis, cuando lo que sucede es que falta conocimiento para discriminar entre todas ellas.

Esta función, consecuencia de la consideración de una determinada evidencia y que se llama función básica de asignación de probabilidad (basic probability assignment) debe verificar las dos condiciones siguientes:

1.- $\sum_{A \subseteq \Theta} m(A) = 1$ La masa o cantidad de probabilidad total asignada debe ser la unidad.

2.- $m(\emptyset) = 0$, puesto que el conjunto de hipótesis se suponía exhaustivo

Esta forma de tratar la incertidumbre, considerando como "hipótesis" todos los subconjuntos de Θ refleja muy bien la ignorancia en el proceso de adquisición del conocimiento necesario para una toma de decisión o para la emisión de un diagnóstico. En efecto, si el marco de discernimiento es $\Theta = \{h_1, h_2, ..., h_n\}$ y una evidencia _e_ es tal que presta apoyo al conjunto formado por las hipótesis h_2 y h_4, $A = \{h_2, h_4\}$ con un grado 0,8 por ejemplo, el resto de probabilidad es asignado al conjunto total, $m(\Theta) = 0{,}2$, quedando $m(\Theta)$ como aquella porción de la probabilidad total que no ha sido asignada a ningún subconjunto de Θ . En un modelo puramente probabilístico esta cantidad sería asignada al complementario de A, es decir, a la negación de la hipótesis $\{h_2, h_4\}$. Queda por tanto obviada una de las críticas que Shortliffe y Buchanan hacían a los sistemas probabilísticos, que era la dificultad de aceptar que si una determinada evidencia prestaba apoyo a una hipótesis también era una evidencia a favor de la negación de esa hipótesis: $P(\bar{h}/e) = 1 - P(h/e)$.

Obsérvese, en consecuencia, que m(A) es aquella parte de la probabilidad total que es asignada exactamente a A y que no es posible subdividir más.

Si una determinada evidencia lo que hace es negar en un cierto grado una determinada hipótesis A (en el sentido de un elemento de $P(\Theta)$, conjunto de hipótesis ampliado) es equivalente a considerar esa misma evidencia como una evidencia a favor de la negación de la hipótesis A, es decir supondría una asignación de probabilidad a su complementario $m(\bar{A})$. Es decir, con las funciones de asignación de probabilidad se consideran siempre valores positivos comprendidos en el intervalo $[0,1]$.

Para cada unidad de evidencia aquellos subconjuntos A tales que $m(A) > 0$ se llaman elementos focales y no deben cumplir ninguna condición pudiendo tener intersección no vacía y sin tener que cubrir totalmente al conjunto θ .

A partir de esta función de asignación de probabilidad, se definen otras varias, de las cuales son de especial importancia dos: la credibilidad y la plausibilidad, que se notarán en lo sucesivo por Cr y Pl, respectivamente y que se definen de la siguiente manera:

$$Cr(A) = \sum_{B \subseteq A} m(B)$$

La credibilidad indica la cantidad de credibilidad total depositada en el conjunto A (hipótesis A), dado que la relación de inclusión es, en términos de afirmaciones, una relación de implicación. Por tanto es la suma de todas las probabilidades asignadas a los elementos focales contenidos en A, es decir, que hacen necesario A.

La plausibilidad $$Pl(A) = \sum_{B \cap A \neq \phi} m(B)$$

incorpora además aquellos subconjuntos que tienen alguna hipótesis individual en común con A y que, por tanto, hacen posible que se verifique la hipótesis A.

De la definición de credibilidad y plausibilidad se deduce inmediatamente

.) $0 \leq Cr(A) \leq 1$, $0 \leq Pl(A) \leq 1$

.) $Pl(A) \geqslant Cr(A)$

.) $Cr(\Theta) = 1$

Por tanto, para cada subconjunto de hipótesis A se tienen dos valores Cr(A) y Pl(A) que proporcionan unos límites inferior y superior para el apoyo que se puede prestar a dicha hipótesis:

$$Cr(A) \leqslant Prob(A) \leqslant Pl(A)$$

Es fácil deducir la siguiente igualdad

$$Pl(A) = 1 - Cr(\bar{A})$$

que proporciona nueva información sobre cada hipótesis A. En efecto, Pl(A) es la máxima cantidad que la evidencia considerada podría permitir creer en A, mientras que Cr(A) es la cantidad con que la evidencia presta apoyo específico al suceso A. Por consiguiente, la amplitud del intervalo $[Cr(A), Pl(A)]$ podría interpretarse como la incertidumbre con respecto a la hipótesis A, puesto que es aquella porción de credibilidad que no es depositada ni en la hipótesis ni en su negación, figura 2.3.1.

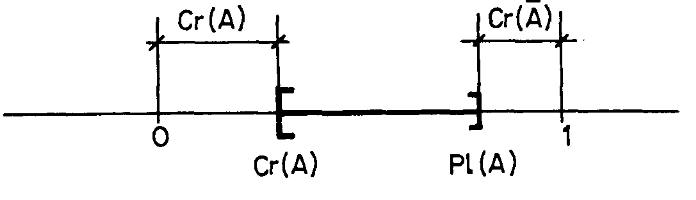


Figura 2.3.1.- Intervalo de confianza.

Este intervalo se llama intervalo de credibilidad o intervalo de confianza.

Como casos extremos se tendrían:

.) Si $Cr(A) = 0$, $Pl(A) = 1$, la ignorancia respecto a la hipótesis A es total.

.) $Cr(A) = Pl(A) = 1$, A es verdadero con total certeza.

.) $Cr(A) = Pl(A) = 0$, A es falso.

Cuando los elementos focales son conjuntos con una única hipótesis, la función de asignación de probabilidad es una distribución de probabilidad ordinaria y para cualquier hipótesis A, en el sentido amplio de hipótesis, $Cr(A) = Prob(A) = Pl(A)$. Corresponde al caso de máxima información puesto que la cantidad de probabilidad total ha sido repartida entre hipótesis individuales y, en este caso, la amplitud de cualquier intervalo de credibilidad es cero.

Son también de inmediata deducción las relaciones

$$Cr(A) + Cr(\bar{A}) \leqslant 1$$

$$Pl(A) + Pl(\bar{A}) \geqslant 1$$

dándose las igualdades sólo si ningún elemento focal tiene intersección con los sucesos A y no, A, simultáneamente. Es, en consecuencia, la existencia de hipótesis individuales que no se sabe si son de A o de no A lo que refleja la ignorancia sobre la hipótesis A.

En el caso de que los sucesos focales estén anidados formando una cadena se verifica

$$Pl(A \cup B) = Max \{Pl(A), Pl(B)\}$$

$$Cr(A \cap B) = Min \{Cr(A), Cr(B)\}$$

lo que hace que las medidas de credibilidad y plausibilidad se correspondan con las medidas de necesidad y posibilidad definidas por Dubois y Prade (DUBO-85b).

En general, las medidas de credibilidad y plausibilidad no son aditivas puesto que se cumple (SHAF-76):

$$Cr(A \cup B) \geqslant Cr(A) + Cr(B) - Cr(A \cap B)$$

$$Pl(A \cup B) \leqslant Pl(A) + Pl(B) - Pl(A \cap B)$$

En resumen, se puede concluir que este modelo de representación de la incertidumbre permite la posibilidad de estructurar la base de conocimientos de forma jerárquica y teniendo en cuenta conjuntos de hipótesis como representación de conceptos de interés semántico (categorías de enfermedades en diagnóstico médico, por ejemplo), (GORD-85) y que pueden mejorar el rendimiento del sistema.

Se han planteado distintas objecciones a este modelo, dando lugar a planteamientos alternativos o a sugerir ciertas mejoras o modificaciones. Por ejemplo Yen (YEN-86) no considera posible representar con funciones básicas de asignación de probabilidad relaciones del tipo "la probabilidad de la hipótesis A dada la evidencia <u>e</u> es 0,8" y define unas funciones más complejas de asignación múltiple de probabilidad. Debe ser un trabajo a tomar en consideración.

Por otro lado, al considerar la probabilidad, credibilidad y plausibilidad de un suceso o subconjunto A son precisamente los elementos focales "frontera" los que reflejan la ignorancia sobre A. Zadeh (ZADE-84) sugiere tener en cuenta el tamaño o "granularidad" (ZADE-79b) de los subconjuntos A y focales. Por ejemplo, si un elemento focal B no cumple la relación $B \subset A$ aún por un pequeño margen, entonces B no hace necesario A, aunque sí lo hace posible. Esta estricta interpretación de la relación de contenido da lugar a saltos bruscos en las cotas superior e inferior de Prob(A). Sugiere sustituir la habitual relación de contenido por una relación en que la transición de contenido a no contenido sea gradual, quizás mediante la utilización de conjuntos difusos.

2.3.2.- Combinación de evidencias.

Si dos fuentes de información proporcionan evidencias e_1 y e_2 favoreciendo o negando hipótesis correspondientes a un mismo marco de discernimiento, aunque no necesariamente sean las mismas hipótesis concretas, cada una de las evidencias representada por su función de asignación de probabilidad m_1 y m_2, se combinan mediante lo que se conoce como "regla de combinación de Dempster" que proporciona una nueva función de asignación de probabilidad, que se notará con $m_{12} = m_1 \oplus m_2$ y que representa el efecto conjunto de las evidencias e_1 y e_2. A partir de ella se obtendrán la credibilidad Cr_{12} y plausibilidad Pl_{12} de cada hipótesis de $P(\theta)$.

Si A_i y B_j son los elementos focales de cada una de las evidencias m_1 y m_2, a la hipótesis $C = A_i \cap B_j$ se le adjudica una probabilidad igual a $m_1(A_i) \cdot m_2(B_j)$. Puesto que la misma hipótesis C puede provenir de la intersección de otras varias parejas de elementos focales

$$m_{12}(C) = \sum_{A_i \cap B_j = C} m_1(A_i) \cdot m_2(B_j)$$

que tiene una sencilla interpretación gráfica, en la que se observa que la probabilidad asignada a C es proporcional al área sombreada y, por consiguiente, a las probabilidades asignadas por m_1 y m_2 a las hipótesis A_i y B_j:

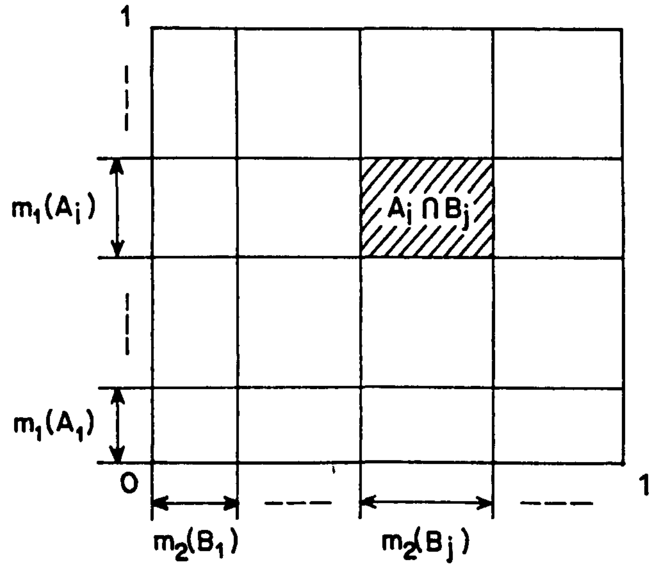


Figura 2.3.2.- Regla de combinación de Dempster.

Obsérvese que esta formulación coincide con la asignación de probabilidad a la intersección de dos sucesos independientes en la teoría de probabilidad clásica y que, se está suponiendo implícitamente la independencia de las evidencias e_1 y e_2.

De forma elemental se comprueba que

$$\sum_{C \subseteq \theta} m_{12}(C) = 1$$

pero, sin embargo, la segunda condición exigida a las funciones básicas de asignación de probabilidad, $m_{12}(\emptyset) = 0$, no se cumplirá desde el momento que existan sucesos focales A_i y B_j para cada una de las evidencias e_1 y e_2 con intersección vacía.

Dempster propone entonces normalizar la distribución m_{12}, asignando el valor 0 al conjunto vacío y modificando las asignaciones de las restantes hipótesis de manera que su suma siga siendo la unidad. Queda de forma definitiva:

$$m_{12}(C) = \frac{\sum_{A_i \cap B_j = C} m_1(A_i) \cdot m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i) \cdot m_2(B_j)} =$$

$$= \frac{\sum_{A_i \cap B_j = C} m_1(A_i) \cdot m_2(B_j)}{\sum_{A_i \cap B_j \neq \emptyset} m_1(A_i) \cdot m_2(B_j)}$$

Este proceso de normalización ha sido puesto en cuestión por diversos autores, puesto que conduce a resultados contradictorios como se verá más adelante.

Como se ha visto, esta necesidad de normalización se produce cuando las fuentes de información aportan evidencias sobre hipótesis muy diferentes, con intersección vacía. En consecuencia, el valor

$$K = \sum_{A_i \cap B_j = \phi} m_1(A_i) \cdot m_2(B_j)$$

proporciona una medida del grado de conflicto que existe entre las evidencias que se están combinando. Cuanto mayor es este valor, mayor es el conflicto entre las fuentes de información y más cuidado habrá que tener al interpretar los resultados obtenidos en el proceso de combinación. En el caso extremo de que K tome el valor 1 entonces las evidencias e_1 y e_2 son totalmente contradictorias y no existe la función m_{12}.

Se comprueba fácilmente que la regla de combinación de Dempster es conmutativa -no importa el orden en que son combinadas las evidencias- y asociativa, es decir, permite la composición de cualquier pareja de evidencias a lo largo del proceso de acumulación del conocimiento.

La regla de combinación de Dempster, expuesta hasta ahora en su forma más general, demuestra su coherencia y buen funcionamiento si se consideran situaciones concretas.

Casos particulares.

1.- La regla de Dempster es una generalización de las reglas de cálculo de la teoría de probabilidades:

Si se combinan dos asignaciones básicas de probabilidad, correspondiendo una de ellas a una distribución de probabilidad clásica (asignación de probabilidad a hipótesis individuales) y siendo la segunda una función que asigna total certeza a una hipótesis $B \subseteq \Theta$ ($m_2(B) = 1$, $m_2(A) = 0$ $\forall A \neq B$) se obtiene

$$Cr_{12}(A) = Cr_1(A \cap B) \ / \ Cr_1(B)$$

que se corresponde con la clásica regla de la teoría de probabilidades

$$Prob(A|B) = Prob(A \wedge B) \ / \ Prob(B)$$

En efecto, sean

$$\theta = \{h_1, h_2, \ldots, h_n\}$$

$$m_1 \text{ tal que } m_1(\{h_i\}) = p_i \quad \text{para } i = 1, 2, \ldots, k, \quad k < n, \sum p_i = 1$$

$$B = \{h_1, \ldots, h_r, h_{r+1}, \ldots, h_m\} \quad \text{con } h_1, \ldots, h_r \in \{h_1, \ldots, h_k\}$$
$$h_{r+1}, \ldots, h_m \notin \{h_1, \ldots, h_k\}$$

$$A = \{h_1, \ldots, h_s, h_{s+1}, \ldots, h_t\} \quad \text{con } h_1, \ldots, h_s \in \{h_1, \ldots, h_r\}$$
$$h_{s+1}, \ldots, h_t \notin \{h_1, \ldots, h_r\}$$

se obtiene

	$m_2(B) = 1$
$m_1(h_1) = p_1$	$\{h_1\} \longrightarrow p_1$
$m_1(h_2) = p_2$	$\{h_2\} \longrightarrow p_2$
$\vdots$	$\vdots$
$m_1(h_r) = p_r$	$\{h_r\} \longrightarrow p_r$
$m_1(h_{r+1}) = p_{r+1}$	$\emptyset \longrightarrow p_{r+1}$
$\vdots$	$\vdots$
$m_1(h_k) = p_k$	$\emptyset \longrightarrow p_k$

$$K = p_{r+1} + \ldots + p_k \qquad \text{(peso del conflicto)}$$

$$1 - K = 1 - (p_{r+1} + \ldots + p_k) = p_1 + \ldots + p_r = Cr_1(B)$$

$$Cr_{12}(A) = \frac{1}{1-K} \cdot \sum_{i=1}^{s} m_{12}(h_i) = \frac{p_1 + \ldots + p_s}{Cr_1(B)}$$

$$Cr_1(A \cap B) = \sum_{i=1}^{s} m_{12}(h_i) = p_1 + \ldots + p_s$$

2.- Las dos evidencias e_1 y e_2 confirman ambas la misma hipótesis, o son evidencias en contra:

	$m_2(A) = s_2$	$m_2(\theta) = 1-s_2$
$m_1(A) = s_1$	$A \longrightarrow s_1 s_2$	$A \longrightarrow s_1(1-s_2)$
$m_1(\theta) = 1-s_1$	$A \longrightarrow s_2(1-s_1)$	$\theta \longrightarrow (1-s_1).(1-s_2)$

Se obtiene $m_{12}(A) = Cr(A) = s_1 + s_2 - s_1 s_2$

$$Pl(A) = 1$$

que es la regla de combinación utilizada en el modelo MYCIN con $CF_1 = s_1$ y $CF_2 = s_2$.

3.- La evidencia e_1 presta apoyo a la hipótesis A, mientras que e_2 es evidencia en contra de A:

	$m_2(\bar{A}) = s_2$	$m_2(\theta) = 1-s_2$
$m_1(A) = s_1$	$\phi \longrightarrow s_1 s_2$	$A \longrightarrow s_1(1-s_2)$
$m_1(\theta) = 1-s_1$	$\bar{A} \longrightarrow s_2(1-s_1)$	$\theta \longrightarrow (1-s_1).(1-s_2)$

$$m_{12}(A) = Cr(A) = \frac{s_1 - s_1 s_2}{1 - s_1 s_2}$$

$$m_{12}(\bar{A}) = Cr(\bar{A}) = \frac{s_2 - s_1\, s_2}{1 - s_1\, s_2}$$

$$m_{12}(\theta) = \frac{(1-s_1)\,(1-s_2)}{1 - s_1\, s_2}$$

Con objeto de poder interpretar estos resultados considérense unos valores numéricos $s_1 = 0,4$ y $s_2 = 0,8$. Se obtiene

$$m_{12}(A) = 0,118, \quad m_{12}(\bar{A}) = 0,706, \quad m_{12}(\theta) = 0,176$$

Antes de considerar la evidencia e_2, el intervalo de confianza de la hipótesis A era $[0'4,1]$ y se ha reducido a $[0'118,0'294]$. Análogamente, antes de considerar la evidencia e_1 el intervalo de confianza de la hipótesis $\bar{A}$ era $[0'8,1]$ y se ha reducido a $[0'706,0'882]$.

El efecto conjunto de ambas evidencias, una a favor y otra en contra de una hipótesis A es reducir la credibilidad tanto en la hipótesis A como en su negación $\bar{A}$, en magnitudes respectivas de acuerdo con la fuerza relativa de cada una de las evidencias (en el ejemplo numérico, mayor disminución para la hipótesis A pues es mayor la fuerza en contra de e_2).

En general, el efecto de una evidencia apoyando una hipótesis A es una disminución de la credibilidad en las demás hipótesis (entendiéndose como hipótesis ampliadas) disjuntas con A. Si una evidencia apoya una hipótesis individual $\{h_i\}$, tras su combinación mediante la regla de Dempster, se producirá una disminución de la credibilidad de las restantes hipótesis individuales y de todas aquellas hipótesis ampliadas que no contengan a h_i.

Si se hubieran combinado de acuerdo con la fórmula utilizada por MYCIN:

$$CF_{12} = \frac{CF_1 + CF_2}{1 - \min\left\{ |CF_1|,\ |CF_2| \right\}}$$

el resultado sería

$$CF_{12} = \frac{0,4 - 0,8}{1 - 0,4} = - 0,667$$

y, en consecuencia, $Cr(A) = 0$, $Cr(\bar{A}) = 0,667$

La mayor fuerza de la evidencia en contra de e_2 niega toda credibilidad a la hipótesis A.

En el caso extremo $s_1 = s_2$, es decir evidencias a favor y en contra de igual fuerza, la regla de MYCIN daría $CF_{12} = 0$, o sea, ignorancia absoluta mientras que la regla de Dempster reduciría por igual la credibilidad en A y $\bar{A}$, pero sin anularla.

4.- La evidencia e_1 apoya la hipótesis A y la evidencia e_2 apoya una hipótesis B, más amplia que A ($A \subset B$).

	$m_2(B) = s_2$	$m_2(\theta) = 1-s_2$
$m_1(A) = s_1$	$A \longrightarrow s_1\, s_2$	$A \longrightarrow s_1(1-s_2)$
$m_1(\theta) = 1-s_1$	$B \longrightarrow s_2(1-s_1)$	$\theta \longrightarrow (1-s_1).(1-s_2)$

$Cr_1(A) = s_1$ $\qquad$ $Cr_{12}(A) = s_1$

$Cr_2(B) = s_2$ $\qquad$ $Cr_{12}(B) = s_1 + s_2 - s_1\, s_2 = s_2 + s_1(1-s_2)$

La evidencia e_2 que apoya un superconjunto de A no modifica la credibilidad de A, pero la evidencia e_1 que apoya un subconjunto de B aumenta la credibilidad de B. Ambos comportamientos resultan coherentes con los conceptos de función básica de asignación de probabilidad y credibilidad.

Sin embargo, también hay casos en que el funcionamiento de la regla de combinación de Dempster no resulta satisfactorio y conduce a resultados difícilmente aceptables. Zadeh (ZADE-86) señala el siguiente caso particular.

Considérese un conjunto de tres alternativas excluyentes, $\theta = \{a, b, c\}$ y dos evidencias que producen unas funciones básicas de asignación de probabilidad:

$$m_1(a) = 0{,}9 \qquad m_1(b) = 0{,}1 \qquad m_1(c) = 0$$

$$m_2(a) = 0 \qquad m_2(b) = 0{,}1 \qquad m_2(c) = 0{,}9$$

El resultado de la aplicación de la regla de Dempster es sorprendentemente

$$m_{12}(a) = m_{12}(c) = 0 \qquad m_{12}(b) = 1$$

es decir $Cr(b) = 1$ lo que lleva a la conclusión de que la hipótesis b es completamente cierta, cuando cada una de las dos evidencias por separado afirmaban que dicha hipótesis era altamente improbable.

Se trata de dos fuentes de información muy conflictivas, (el valor de k es 0,99, muy cercano a uno) puesto que la primera considera la hipótesis c imposible y a casi cierta y recíprocamente para la segunda evidencia. Por tanto, la hipótesis b aunque bastante improbable para ambas opiniones es la única alternativa posible.

El origen del problema está en el proceso de normalización, al eliminar la posibilidad de que ninguna de las alternativas consideradas sea correcta. Problemas de este tipo se plantean cuando las evidencias a combinar son muy contradictorias.

Dubois y Prade (DUBO-85a) proponen acertadamente dos posibles soluciones al problema planteado por Zadeh. La primera de ellas consiste en asignar a cada una de las hipótesis valores no nulos

$$m_1(a) = \varepsilon \qquad m_1(b) = k \qquad m_1(c) = 1-k-\varepsilon$$

$$m_2(a) = 1-k-\varepsilon \qquad m_2(b) = k \qquad m_2(c) = \varepsilon$$

obteniêndose

$$m_{12}(a) = m_{12}(c) = \frac{\varepsilon \cdot (1-k-\varepsilon)}{k^2 - 2\varepsilon(1-k-\varepsilon)} \qquad m_{12}(b) = \frac{k^2}{k^2 - 2\varepsilon(1-k-\varepsilon)}$$

para $k = 0{,}1$ y $\varepsilon = 0{,}01$ se obtiene

$$m_{12}(a) = m_{12}(c) = 0{,}32 \qquad m_{12}(b) = 0{,}36$$

para $k = 0{,}1$ y $\varepsilon = 0{,}001$ se obtiene

$$m_{12}(a) = m_{12}(c) = 0{,}008 \qquad m_{12}(b) = 0{,}84$$

Se obtienen resultados muy diferentes al caso considerado por Zadeh. El asignar valores aun muy pequeños a hipôtesis muy poco probables, permite seguir considerando dicha hipôtesis y ser revisada cuando se consideren nuevas evidencias.

Otra alternativa para prevenir resultados indeseables puede ser la asignaciôn de un pequeño valor al conjunto completo θ , mejor que valores no nulos a cada hipôtesis. Con este criterio

$$m_1(a) = 0 \qquad m_1(b) = k \qquad m_1(c) = 1-k-\varepsilon \qquad m_1(\theta) = \varepsilon$$

$$m_2(a) = 1-k-\varepsilon \qquad m_2(b) = k \qquad m_2(c) = 0 \qquad m_2(\theta) = \varepsilon$$

se obtiene

$$m_{12}(a) = m_{12}(c) = \frac{\varepsilon \cdot (1-k-\varepsilon)}{k^2 + 2\varepsilon(1-\varepsilon)} \qquad m_{12}(b) = \frac{k^2 + 2k\varepsilon}{k^2 + 2\varepsilon(1-\varepsilon)}$$

$$m_{12}(\theta) = \frac{\varepsilon^2}{k^2 + 2\varepsilon(1-\varepsilon)}$$

que para $k = 0{,}1$ y $\varepsilon = 0{,}01$

$m_{12}(a) = m_{12}(c) = 0,298 \qquad m_{12}(b) = 0,403 \qquad m_{12}(\Theta) = 3,3 \times 10^{-3}$

y para $k = 0,1$ y $\varepsilon = 0,001$

$m_{12}(a) = m_{12}(c) = 0,0075 \qquad m_{12}(b) = 0,85 \qquad m_{12}(\Theta) = 8,3 \times 10^{-5}$

resultados muy parecidos a los obtenidos antes.

2.3.3.- Propagación de incertidumbre.

En cualquier sistema experto cuyo conocimiento está estructurado bien como un sistema de producción constituído por un conjunto de reglas o bien mediante una red de inferencia, un aspecto primordial en el motor de inferencia que maneja dicho conocimiento es el encadenamiento de evidencias e hipótesis. El modelo de tratamiento de la incertidumbre que es inherente a hechos y reglas debe ser capaz de propagar en la cadena de razonamiento la incertidumbre asociada a cada elemento.

La teoría D-S resuelve el problema de acumular la incertidumbre cuando distintas evidencias aportan conocimiento de cara a una serie de hipótesis pertenecientes a un mismo marco de discernimiento. En términos de conocimiento estructurado en una base de reglas, la acumulación de evidencias que es capaz de tratar la regla de combinación de Dempster, supone obtener en términos de credibilidad el efecto de las evidencias e_i sobre las hipótesis H_j de un conjunto de reglas del tipo "si e_i entonces H_j" donde las hipótesis H_j se pueden considerar como subconjuntos de un determinado marco de discernimiento Θ , y en los supuestos de que las distintas evidencias son condicionalmente independientes y los hechos que reflejan son conocidos con certeza.

El problema de la propagación de incertidumbre se origina cuando los hechos e_i a su vez no son conocidos con certeza sino que ellos mismos también están sujetos a un tratamiento de incertidumbre, normalmente porque dichos hechos son conclusiones de reglas a encadenar con las consideradas antes.

En el contexto de la teoría de D-S parece claro que los hechos e_i y las hipótesis H_j corresponderán a proposiciones pertenecientes a distintos marcos de discurso y que, por tanto, un tratamiento de este problema con funciones de credibilidad y plausibilidad supondrá la consideración de muy distintos marcos de discernimiento, cada uno de ellos correspondiendo a hipótesis o decisiones intermedias.

Se han hecho algunos intentos para tratar de resolver problemas de este tipo. Se citan a continuación algunos de ellos, aunque no parece existir actualmente una metodología clara, contrastada y ampliamente aceptada.

Lowrence y Garvey, (GARV-81), son los primeros en proponer una serie de "reglas de inferencia" pero no queda nada claro si se trata de auténticas reglas de inferencia, es decir, si se utilizan distintas fuentes de información y distintos marcos de discurso para valorar la incertidumbre de las reglas "si <u>e</u> entonces <u>h</u>", por una parte, y la incertidumbre de los hechos <u>e</u>, por otra. Más bien parece que se trata de deducir la credibilidad de distintos subconjuntos de un único marco de discernimiento a partir de la credibilidad de otros subconjuntos o sucesos. En cualquier caso, no están justificadas más que por su adecuación al caso particular de que los sucesos o subconjuntos sean ciertos o falsos.

Mucho más importante parece el esfuerzo de Lu y Stephanou, (LU-84) al abordar el problema en los términos planteados anteriormente. Consideran dos marcos de discurso, uno <u>U</u> para las evidencias y otro <u>V</u> para las hipótesis, ligados por una serie de reglas del tipo "si A entonces X" donde A es un subconjunto de U y X un subconjunto de V; es decir, sucesos o proposiciones relativos a distintos marcos de discernimiento. Cada una de estas reglas viene afectada por un grado de confianza que es proporcionado por el experto que proporciona la regla al construir la base de conocimientos. Se parte de la incertidumbre de los hechos A, expresada como valores de credibilidad, Cr(A) y que se obtienen, bien a partir de unas funciones de asignación de probabilidad proporcionadas por el usuario del sistema y la correspondiente aplicación de la teoría de D-S, regla de combinación incluída, o bien como resultado de aplicar el propio proceso a la cadena de razonamiento. Estos hechos con sus credibilidades asociadas activan las correspondientes reglas y, en el nuevo marco de discernimiento V, proporcionan nuevas funciones de asignación de probabilidad. Sean

$$A \longrightarrow X, \ \vartheta_1 \qquad A, B, C, \subset U$$

$$B \longrightarrow Y, \ \vartheta_2$$

$$C \longrightarrow Z, \ \vartheta_3 \qquad X, Y, Z \subset V$$

con Cr(A) = a, Cr(B) = b, CR(C) = c. En el marco V se obtienen funciones

$$m_1(X) = a \cdot \vartheta_1 , \quad m_1(V) = 1 - a \cdot \vartheta_1 ,$$

$$m_2(Y) = b \cdot \vartheta_2 , \quad m_2(V) = 1 - b \cdot \vartheta_2 , \text{ etc.}$$

funciones que son combinadas nuevamente mediante la regla de Dempster, obteniéndose la credibilidad de los sucesos X, Y, Z, ... y de aquellos consecuencia de la aplicación de la regla de Dempster, correspondientes al nuevo marco de discurso.

Parece un enfoque bastante correcto aunque poco justificado, especialmente en lo que se refiere al paso de uno a otro marco de discernimiento. Obsérvese que, esencialmente, es el modelo MYCIN para la propagación de incertidumbre. Sin embargo, no parece muy lógico emplear en el marco U valores de credibilidad que, como consecuencia de la activación de una regla, originan asignaciones de probabilidad en el marco V. Sería más correcto utilizar también asignaciones de probabilidad para los antecedentes de las reglas, lo cual plantea problemas de aplicabilidad del método.

De todas formas, y aunque los autores proponen una serie de sencillos ejemplos para mostrar la coherencia de su método en casos muy específicos, debería ser calibrado en un sistema real, de lo cual no hay noticias.

Otro intento mucho más teórico y de resultados parciales es el debido a Dubois y Prade, (DUBO-85a). Razonando a partir de las propiedades de las funciones de credibilidad y plausibilidad definidas por Shafer obtienen unos esquemas de inferencia del tipo "modus ponens" y "modus tollens":

$$\frac{\begin{array}{l} Cr(A \longrightarrow B) \geqslant r \\ Cr(A) \geqslant s \end{array}}{Cr(B) \geqslant \max(0,\ r+s-1)} \qquad \frac{\begin{array}{l} Cr(A \longrightarrow B) \geqslant r' \\ Pl(B) \leqslant t \end{array}}{Pl(A) \leqslant \min(1,\ 1-r'+t)}$$

que combinados

$$\frac{\begin{array}{l} Cr(A \longrightarrow B) \geqslant r \\ Cr(B \longrightarrow A) \geqslant r' \\ Cr(A) \geqslant s \qquad Pl(A) \leqslant t \end{array}}{Cr(B) \geqslant \max(0,\ r+s-1) \qquad Pl(A) \leqslant \min(1,\ 1-r'+t)}$$

proporcionan unos límites inferior y superior del intervalo de confianza de la conclusión de la regla "si A entonces B".

Finalmente, se deben tener en cuenta los trabajos de Rich y Ginsberg, (GINS-84), aunque están enfocados a planteamientos de redes semánticas en un marco de razonamiento no monotóno.

2.3.4.- Implementaciones, desarrollos y extensiones.

Son bastantes los trabajos realizados para aplicar la teoría de D-S a distintos terrenos de la Ingeniería de Conocimiento. Algunos de interés son los siguientes:

Barnett, (BARN-81), hace un desarrollo estricto de la regla de combinación de Dempster con objeto de reducir el coste computacional en los supuestos

- colección de unidades de evidencia que apoyan hipótesis de un único marco de discernimiento.

- cada evidencia confirma o niega una hipótesis individual y no una disyunción de hipótesis.

Además de obtener, para este supuesto concreto, las funciones resultado de combinar las distintas evidencias, plantea un sistema de consulta dinámico (los resultados van modificándose a medida que las evidencias van obteniéndose y procesándose) de tal forma que el usuario puede, en cualquier momento, obtener información sobre

- valores de $\lambda(A)$, donde λ es una cualquiera de las funciones m, Cr, Pl y otras y A un subconjunto arbitrario de θ .
- estado general del proceso de inferencia:
  a) cuál es el peso del conflicto en la combinación de evidencias.
  b) si existe o no información totalmente contradictoria en el conjunto de evidencias que se van procesando.
  c) aspectos de decisión:
    - si existe algún suceso A tal que $m(A) = 1$ y en caso afirmativo cuál es.
    - cuál es mínimo subconjunto A tal que $Cr(A) = 1$. (Téngase en cuenta que $Cr(\theta) = 1$ siempre).

Para alguno de estos problemas deduce unas sencillas condiciones necesarias y suficientes en función de los mismos parámetros utilizados para el cálculo de la función de combinación.

En (GORD-84) se desarrolla un ejemplo basado en un conjunto de reglas extraído del sistema MYCIN. Dicho ejemplo finaliza con la obtención de los intervalos de credibilidad de cada una de las hipótesis sobre las que existen evidencias, pero nada se indica sobre cómo tomar una decisión.

Gordon y Shortliffe, (GORD-85), siguiendo la línea de Barnett, sugieren que los supuestos hechos en la implementación de este último, al considerar únicamente hipótesis individuales, se desaprovecha toda la potencialidad de la teoría de D-S de poder tratar con evidencias sobre conjuntos de hipótesis a distintos niveles de especificidad. Proponen una implementación para el caso de una estructura jerarquizada de hipótesis. Para que la computación pueda ser reducida a términos aceptables, se restringen las hipótesis de P(θ) a considerar una jerarquía estricta (dicho conjunto de hipótesis forman un árbol) y se aplica rigurosamente la regla de combinación de Dempster salvo en un caso concreto, que es justificado adecuadamente, con objeto de que todos los subconjuntos obtenidos pertenezcan a los considerados en la jerarquía de hipótesis. Como contrapartida, se obtienen tan sólo valores de credibilidad, perdiéndose la información que proporcionan los intervalos de confianza.

Una de las primeras aplicaciones de la teoría de D-S es la propuesta de Garvey y Lowrence (GARV-81), que la utilizan para decidir de qué emisor proviene una señal de radar de entre un conjunto de emisores, en base a la señal captada por diferentes sensores. Las señales procedentes de los emisores son parametrizadas y su probabilidad expresada mediante funciones básicas de asignación de probabilidad. La integración de la información proporcionada por los distintos sensores es procesada utilizando la regla de combinación de Dempster.

Es interesante también la sencilla extensión de Strat, (STRA-84). En la teoría de D-S, el marco de discernimiento está formado por un número finito de hipótesis. Strat extiende la formulación de funciones de credibilidad y plausibilidad al caso de un marco de discernimiento no discreto; es decir, cuando las hipótesis a discriminar vienen en función de una variable continua. Se obtiene una variación continua de la credibilidad y plausibilidad y se propone también una nueva formulación de la regla de combinación de Dempster para este caso. Es útil en aquellas situaciones en que se busquen decisiones basadas en evidencias inciertas o incompletas sobre magnitudes tales como tiempo, distancia, medida de sensores, etc.

Una extensión de la teoría de D-S bastante compleja es la debida a Yen, (YEN-86), que, como ya se indicó, considera insuficientes las funciones de asignación de probabilidad para representar la incertidumbre de evidencias y reglas. Parte de unas multifunciones probabilísticas y a partir de ellas reelabora toda la teoría, introduciendo probabilidades a "priori" y dos fuertes condiciones de independencia a la hora de combinar distintas fuentes de información. Es un trabajo muy teórico que necesitaría ser contrastado en situaciones reales.

Clarificador puede ser el enfoque de la teoría de D-S que realiza Zadeh, (ZADE-86), considerada en un sencillo contexto de bases de datos relacionales y que aporta luz a puntos conflictivos como son la normalización que se realiza en la regla de combinación de Dempster y la no consideración de aquellas hipótesis que al combinarlas son disjuntas.

La teoría de D-S se ha aplicado también a sistemas expertos ya comercializados como, por ejemplo, INSPECTR, (KESS-87), que está diseñado para interpretar información relativa al estado de una planta de producción de energía eléctrica de cara a identificar y diagnosticar problemas que afectan a su comportamiento. Es un sistema experto basado en reglas, cuyo motor de inferencia permite encadenamiento hacia delante y hacia atrás. Se justifica la elección del modelo D-S por las razones ya apuntadas de adecuación a requerimientos tales como tratamiento diferenciado de hechos (cuya incertidumbre puede variar de un instrumento a otro, de una planta a otra) y de reglas (cuya incertidumbre permanece relativamente constante), tratamiento adecuado de la ignorancia y facilidad para que la incertidumbre pueda ser suministrada convenientemente por el experto en el dominio. Como suele ocurrir con los sistemas comerciales, la información es siempre muy general y no se ofrecen detalles de cómo se resuelven problemas tales como la propagación de la incertidumbre al encadenar reglas o en base a qué criterios se elige un determinado diagnóstico.

Otro sistema experto que utiliza un razonamiento paralelo al de D-S es el TURBOMAC, (PARG-87), sistema de diagnóstico en maquinaria industrial con elementos de alta velocidad. Su base de conocimiento consta de 9.700 reglas y considera un conjunto de 142 posibles síntomas para concluir un diagnóstico sobre la posible presencia o ausencia de 50 causas del problema.

También considera grupos de causas. Utiliza los valores de los síntomas para determinar dos contadores del diagnóstico de cada problema o grupo de problemas: uno positivo y otro negativo según el grado de incertidumbre de que el problema esté presente o ausente, utilizando ambos contadores en el proceso de razonamiento para diagnosticar sobre problemas.

2.4.- TEORIA DE LA POSIBILIDAD. LOGICA DIFUSA.

Todos los modelos considerados hasta ahora intentan cuantificar la incertidumbre sobre la verdad o falsedad de hechos y reglas con valores numéricos, estando estos números más o menos relacionados con valores de probabilidad. Sin embargo, los hechos y reglas que forman la base de conocimientos de un sistema experto vienen dados como sentencias o proposiciones expresadas en lenguaje natural, al menos en un primer nivel del proceso de adquisición del conocimiento a partir de expertos, y previamente a su formalización en algún sistema de representación del conocimiento.

Una de las características esenciales del lenguaje natural es la utilización de predicados vagos o difusos, cuyo significado es difícil de precisar.

Tras la aparición de la teoría de los "conjuntos difusos", (ZADE-65), y su aplicación generalizada a todas las ramas de la Matemática y la Informática, a fines de la década de los setenta, el propio Zadeh utilizó estos conceptos desarrollando el llamado "Razonamiento Aproximado". Esto dió origen a lo que se conoce actualmente como "Teoría de la Posibilidad" y "Lógica Difusa", ya de amplia difusión y que ha sido objeto de atención por numerosos investigadores y especialmente ampliada por Dubois y Prade. Actualmente son numerosos los sistemas expertos que hacen uso de estos conceptos tanto para la representación del conocimiento como para la implementación de los mecanismos de inferencia.

2.4.1.- La incertidumbre y la imprecisión.

En la introducción de su trabajo "Théorie des Possibilités", Dubois y Prade, (DUBO-85b) hacen una clara distinción entre los conceptos de incierto e impreciso. Ambos tienen por origen el hecho de que el conocimiento que se posee sobre cualquier problema a tratar mediante un sistema experto es, en general, imperfecto e incompleto.

Tradicionalmente, un elemento de información se ha venido representando en múltiples ocasiones mediante una terna del tipo ⟨atributo, objeto, valor⟩. La componente valor será, en general, un predicado, entendido como un subconjunto de un universo de referencia ligado al atributo. Este, a su vez, es una determinada propiedad del sujeto. Por ejemplo, elementos de información diferentes serían:

⟨edad, Juan, 22⟩ ⟨hombre, Juan, V⟩

⟨edad, Juan, [18,24]⟩ ⟨mujer, Juan, F⟩

⟨edad, Juan, muy joven⟩ ⟨padre, Juan, Antonio⟩

La incertidumbre de un elemento de información hace referencia a la verdad de esa información, en cuanto a su conformidad con el conocimiento que se posee del mundo real. En este sentido, es normal añadir una cuarta componente a esta representación, con el fin de incluir la credibilidad del elemento de información: ⟨atributo, objeto, valor, credibilidad⟩. Por ejemplo, en el sistema MYCIN se hace de esta forma y se consideran cuaternas como

⟨Identidad, Organismo-1, Streptococo, 0,8⟩

En cambio, la imprecisión de una información está ligada al contenido de la propia información y, más concretamente, al componente valor. Así, por ejemplo, ⟨edad, Juan, 22⟩ es una información precisa, mientras que ⟨edad, Juan, [18,24]⟩ y ⟨edad, Juan, muy joven⟩ son informaciones imprecisas. Sin embargo, se observa una diferencia evidente entre estas dos últimas ternas. En ⟨edad, Juan, [18,24]⟩ el valor que toma el atributo "edad" es un subconjunto, en su sentido clásico, de un referencial que podría ser $U = [0,120] \cap N$, mientras que en la última terna el valor de ese mismo atributo viene dado por medio del predicado "muy joven", que se expresa en lenguaje natural y que no puede considerarse como un subconjunto del referencial U, aunque si sería un subconjunto difuso del mismo, en el sentido de Zadeh. En consecuencia, la información proporcionada por ⟨edad, Juan, [18,24]⟩ sería imprecisa pero no difusa, mientras que ⟨edad, Juan, muy joven⟩ sería una información difusa.

Evidentemente, una información puede ser al mismo tiempo incierta y difusa o imprecisa, como sería, por ejemplo, la proposición "la probabilidad de que x sea mucho mayor que y es 0'8":

$$\langle \text{igualdad}, (x,y), \text{mucho mayor}, P = 0'8 \rangle$$

Los conceptos de incertidumbre e imprecisión son en cierto sentido opuestos. Es comúnmente aceptado que una información tenderá a hacerse tanto más incierta cuanta más precisión se desee sobre ella. Por otra parte, las conclusiones que se deduzcan de un razonamiento con informaciones imprecisas se verán afectadas de una cierta componente de incertidumbre.

Mucho se ha argumentado a favor de la consideración de la imprecisión o vaguedad al intentar modelizar el razonamiento humano, o al enfrentarse con aspectos del conocimiento que son intrínsecamente poco adaptables a una representación precisa y cuantitativa. Zadeh, (ZADE-79a), señala que es la capacidad de razonar en términos cualitativos e imprecisos lo que diferencia la inteligencia humana de la inteligencia de una máquina, y que la forma correcta y efectiva de acercarse a campos muy complejos y mal estructurados es relajando el afán de precisión y de estructurar de forma cuantitativa y matemáticamente cualquier tema de estudio, (ZADE-75).

En el terreno específico de los sistemas basados en el conocimiento, son numerosos los inconvenientes que se derivan de un acercamiento cuantitativo mediante modelos como la teoría de la probabilidad o la lógica bivalente, como por ejemplo:

- Necesidad de suposiciones tales como la independencia condicional de hipótesis o como que éstas sean mutuamente excluyentes.

- Se fuerza al experto a ser preciso en cuanto a los números que debe proporcionar para caracterizar hechos y reglas, a veces de forma poco realista y pudiendo suceder que tales estimaciones conducen a inconsistencias que no son apreciadas por el propio experto.

Por otra parte, el conocimiento introducido en la base de conocimientos de un sistema experto, tanto en forma de hechos como de reglas, es en gran parte difuso. Zadeh, (ZADE-83a), extrae de los sistemas expertos MYCIN y PROSPECTOR reglas tipo para mostrar que, en ambos casos, las proposiciones que forman las premisas y la conclusión contienen predicados difusos y que incluso los números que se asocian a cada regla para expresar la credibilidad que merecen pueden ser interpretados como números difusos. En todo caso, si en una regla cuantificada con un grado de credibilidad, "si A entonces B, α " los hechos A y B son difusos y el valor numérico α se interpreta como un valor de probabilidad $P(B|A) = \alpha$, debería tratarse como probabilidad de un suceso difuso.

El tratamiento de la incertidumbre en sistemas expertos mediante lógica difusa ofrece las siguientes ventajas sobre los modelos numéricos:

- Permite considerar de forma adecuada la existencia de predicados difusos, en antecedentes y consecuentes de las reglas. Una gran parte de las variables que aparecen en el razonamiento humano toman valores que no son numéricos, sino lingüísticos. La utilización del concepto de variable lingüística y los procedimientos asociados, (ZADE-75), permite un tratamiento realista de esta situación.

- Permite una equiparación parcial entre el antecedente de una regla y un hecho proporcionado por el usuario. Si en la base de conocimientos existe una regla del tipo "Si X es A entonces Y es B, CF = α " y el hecho introducido es "X es A'", donde A y A' son predicados difusos, la Lógica Difusa, mediante el "modus ponens generalizado" permite deducir conclusiones acerca del valor de la variable Y.

- A diferencia de la lógica bivalente, en que los únicos valores de verdad son V o F, o de la lógica multivalente, en que el valor de verdad es un número real de un conjunto finito o infinito de valores de verdad, la Lógica Difusa permite la asignación de valores de verdad lingüísticos, como "muy cierto", "más o menos cierto", etc., considerados como subconjuntos difusos del intervalo $[0,1]$. Es posible, por tanto, la consideración y representación de proposiciones del tipo "(X es A) es τ ", donde τ es un valor de verdad lingüístico.

- El hecho anterior no es sino un caso muy particular de la posibilidad de considerar etiquetas lingüísticas que permiten modificar el significado de predicados difusos, como "muy", "más o menos", "poco", etc.

- Permite la presencia de cuantificadores difusos tales como "muchos", "alrededor del 30%", como por ejemplo en la proposición difusa "la mayoría de los suecos son rubios" e incluso ponerlo de manifiesto en proposiciones en que dicho cuantificador se encuentra de forma implícita. El hecho "los estudiantes son jóvenes" debería traducirse más correctamente por "la mayoría de los estudiantes son jóvenes" o como la regla "(si X es estudiante, entonces X es joven) es muy probable".

- Como se pone de manifiesto en este último ejemplo, un hecho puede ser cualificado a través de la Lógica Difusa, no sólo con valores de verdad sino con valores de probabilidad o posibilidad lingüísticos y difusos, de la forma "p es λ " o "p es π " con λ o π cualificadores como "poco probable" o "muy posiblemente".

2.4.2.- Representación del conocimiento difuso.

Para que sea factible realizar inferencias a partir de un conjunto de reglas y hechos difusos, es necesario, en una primera etapa, representar cada una de las proposiciones presentes, hechos o reglas, de una forma estandar. En consecuencia, se trata de representar y poner de manifiesto el significado de proposiciones de la forma:

- "X es A" donde X es una variable y A un subconjunto difuso del universo de referencia U donde toma valores la variable X.

- Proposiciones condicionales del tipo "si X es A entonces Y es B".

- Proposiciones condicionales o no en las que aparecen cuantificadores, "Q's X son A", o modificadores, "X es m A".

- Proposiciones complejas obtenidas a partir de otras proposiciones mediante conjunción, disyunción o negación.

- Proposiciones calificadas con valores de verdad lingüísticos, "p es τ ", o con valores de probabilidad o posibilidad difusos.

Representación de la proposición "X es A".

Se considera la proposición imprecisa, no difusa:

"X es un entero en el intervalo $[0,5]$"

Utilizando la notación Pos(X = u) para significar la posibilidad de que la variable X tome el valor $u \in U$, universo de referencia, se tendrá:

$$\mathrm{Pos}(X = u) = 1 \qquad \text{si } u \in [0,5] \cap N$$
$$\mathrm{Pos}(X = u) = 0 \qquad \text{si } u < 0 \text{ ó } u > 5$$

Si se considera ahora la proposición difusa

"X es un entero pequeño"

donde el conjunto de los "enteros pequeños" es el subconjunto difuso definido por la siguiente función de pertenencia, utilizando la notación de Zadeh:

$$A = \text{"enteros pequeños"} = 1/0 + 1/1 + 0'8/2 + 0'6/3 + 0'4/4 + 0'2/5$$

extendiendo la interpretación dada a la proposición no difusa anterior, se podrá considerar:

$$\mathrm{Pos}(X = 2) = \mu_A(2) = 0'8$$

o más concretamente

$$\mathrm{Pos}(X = 2 \mid X \text{ es entero pequeño}) = 0'8$$

que viene a expresar la "compatibilidad" del hecho de que X tome el valor 2 con el hecho de que "X es un entero pequeño".

Por tanto, una proposición del tipo p "X es A" se traslada en:

$$\Pi_X = A$$

donde Π_X es la <u>distribución de posibilidad</u> asociada a la variable X, o, más explícitamente:

$$\pi_X(u) = Pos(X = u \mid X \text{ es } A) = \mu_A(u)$$

donde $\pi_X: U \longrightarrow [0,1]$ se conoce como función de distribución de posibilidad asociada a Π_X.

De forma más general, X será una variable n-aria, $X = (X_1, \ldots, X_n)$ y A una relación difusa de $U = U_1 \times \ldots \times U_n$. Resulta entonces:

$$p \longrightarrow \Pi_{(X_1, \ldots, X_n)} = A$$

$$\pi_X(u_1, \ldots, u_n) = Pos(X_1 = u_1, \ldots, X_n = u_n \mid X \text{ es } A) = \mu_A(u_1, \ldots, u_n)$$

Por ejemplo, la proposición "X es Y son aproximadamente iguales" se representaría por

$$\Pi_{X,Y} = \text{aprox. igual}$$

donde "aprox. igual" es un subconjunto difuso de R x R.

En definitiva, con esta representación se trata de expresar la proposición $p \equiv$ "X es A" de forma que queden de manifiesto las restricciones que el subconjunto A impone a los valores que puede tomar la variable X.

Existen una serie de métodos muy desarrollados sobre cómo representar proposiciones más complejas que se pueden encontrar en (ZADE-72, 79a), (PRAD-82), (YAGE-80), con propuestas muy completas para el tratamiento de modificadores lingüísticos tipo m A, cuantificadores difusos Q's A, cualificación por valores de verdad, posibilidad o probabilidad, etc. En lo que sigue se expondrán tan sólo aquellos resultados de interés para la posterior explicación de los mecanismos de inferencia que se utilizan en Lógica Difusa

Es importante señalar la existencia de un procedimiento sistemático, también debido a Zadeh, (ZADE-77), para representación de proposiciones por complejas que sean éstas: se trata de un lenguaje de representación de significados denominado PRUF (Possibilistic Relational Universal Fuzzy). Sin embargo, en el caso de sistemas expertos, para la proposiciones de la base de conocimientos suele ser suficiente la simple inspección para hacer la traslación de la sentencia.

<u>Representación de proposiciones "X es A y/o Y es B".</u>

A partir de las proposiciones $p \equiv$ "X es A" y $q \equiv$ "X es B" que se trasladan en

$$p \longrightarrow \Pi_{(X_1, \ldots, X_n)} = A$$

$$q \longrightarrow \Pi_{(X_1, \ldots, X_n)} = B$$

donde A y B son subconjuntos difusos del referencial $U = U_1 \times \ldots \times U_n$, las proposiciones "p y q" y "p ó q" se trasladan, según la propuesta de Zadeh, (ZADE-79a), de la siguiente forma:

$$p \text{ y } q \longrightarrow \Pi_{(X_1, \ldots, X_n)} = A \cap B$$

$$p \text{ ó } q \longrightarrow \Pi_{(X_1, \ldots, X_n)} = A \cup B$$

La intersección y unión de conjuntos difusos se suele definir a partir de los operadores min y max, respectivamente:
$\mu_{A \cap B}(u) = \min(\mu_A(u), \mu_B(u))$, $\mu_{A \cup B}(u) = \max(\mu_A(u), \mu_B(u))$, aunque Prade demuestra, (PRAD-82) que operadores $* : [0,1] \times [0,1] \longrightarrow [0,1]$ que

verificaran las propiedades requeridas para la intersección y unión de conjuntos deberían ser necesariamente normas triangulares y conormas triangulares, respectivamente, c-duales la una de la otra, siendo c una negación fuerte. Según las normas y conormas triangulares utilizadas, se origina una amplia variedad de formas de definir las distintas operaciones. Estas diferentes maneras son utilizadas simultáneamente por algún sistema, como MILORD, que se comentará más adelante.

A partir de las proposiciones $p \equiv$ "X es A" y $q \equiv$ "Y es B", y si las variables X e Y toman valores en universos de referencia diferentes $U_1 \times \ldots \times U_n$ y $V_1 \times \ldots \times V_m$, la propuesta de Zadeh es:

$$p \text{ y } q \longrightarrow \Pi_{(X_1, \ldots, X_n, Y_1, \ldots, Y_m)} = \bar{A} \cap \bar{B}$$

$$p \text{ ó } q \longrightarrow \Pi_{(X_1, \ldots, X_n, Y_1, \ldots, Y_m)} = \bar{A} \cup \bar{B}$$

donde $\bar{A}$ y $\bar{B}$ son las extensiones cilíndricas de los subconjuntos A y B a $U_1 \times \ldots \times U_n \times V_1 \times \ldots \times V_m$.

Representación de proposiciones condicionales "Si X es A entonces Y es B" (I).

Existen múltiples propuestas para el tratamiento de este tipo de proposiciones. Algunas de ellas son éstas:

a) Siguiendo los criterios de la lógica bivalente:

"si X es A entonces Y es B" $\longleftrightarrow$ (X es no A) o (Y es B)

$$p \longrightarrow \Pi_{(X_1, \ldots, X_n, Y_1, \ldots, Y_m)} = \bar{A}^c \cup \bar{B}$$

b) Con la interpretación, (ZADE-75):

"si X es A entonces Y es B" $\longleftrightarrow$ (X es A e Y es B) o
(X es no A e Y "no importa qué")

$$p \longrightarrow \Pi_{(X,Y)} = (\bar{A} \cap \bar{B}) \cup \bar{A}^c$$

c) Una tercera interpretación propuesta por Zadeh, (ZADE-83b) es:

$$p \longrightarrow \Pi_{(X,Y)} = \overline{A}^{c} \oplus \overline{B}$$

donde $\oplus$ es la suma acotada: $\mu_{A \oplus B}(u,v) = \min(1, \mu_A(u) + \mu_B(v))$ que es consistente con la definición de implicación en la lógica de infinitos valores de verdad de Lukasiewicz: $v(p \rightarrow q) = \min(1, 1-v(p) + v(q))$.

Dubois y Prade, (DUBO-80b) aún proponen otro esquema para representar proposiciones condicionales, que se deduce con una consecuencia de la regla de inferencia difusa conocida como "modus ponens generalizado", y que se comentará en su momento.

2.4.3.- Reglas de Inferencia en Lógica Difusa.

En un principio el objetivo es, (ZADE-79a), extender a proposiciones con predicados difusos la regla esencial del cálculo de predicados: modus ponens, obteniendo un esquema de razonamiento de la forma:

$p \longrightarrow q \equiv$ "si X es A entonces Y es B"

$p' \equiv$ "X es A'"

$q' \equiv$ "Y es B'"

aunque posteriormente Zadeh, (ZADE-83a), ha desarrollado múltiples esquemas para proposiciones en las que aparecen cuantificadores.

2.4.3.1.- Modus ponens generalizado (I).

Supóngase que $p \longrightarrow q$, proposición en la que intervienen las variables X e Y está representada por la distribución de posibilidad $\Pi_{(X,Y)} = G$, siendo G un subconjunto difuso de U x V, y que la proposición p' viene representada por $\Pi_X = A'$, con A' subconjunto difuso de U. Se trata de obtener la distribución de posibilidad que representa la proposición $q' \equiv$ "Y es B'".

Zadeh, en estos supuestos, deduce una sencilla regla "compositional rule of inference", o regla de composición, basada en dos principios elementales:

.) Principio de proyección.

$$\text{de} \qquad p \longrightarrow \Pi_{(X_1, \ldots, X_n)} = F \subset U_1 \times \ldots \times U_n$$

$$\text{se deduce} \qquad r \longleftarrow \Pi_{X(s)} = \text{Proy}_{U(s)} F$$

donde X(s) es la variable $(X_{i_1}, \ldots, X_{i_s})$ y $\text{Proy}_{U(s)} F$ es la proyección de la relación difusa n-aria F, sobre $U_{i_1} \times \ldots \times U_{i_s}$, y que se define por:

$$\mu_{\text{Proy}_{U(s)} F}(u_{i_1}, \ldots, u_{i_s}) = \sup\left\{\mu_F(u_1, \ldots, u_n) \mid (u_{j_1}, \ldots, u_{j_r}) \in U_{j_1} \times \ldots \times U_{j_r}\right\}$$

siendo $(j_1, \ldots, j_r)$ la secuencia complementaria de $(i_1, \ldots, i_s)$ respecto de $(1,2,\ldots,n)$.

.) Principio de conjunción.

$$\text{de} \quad p \longrightarrow \Pi_{(X_1, \ldots, X_n, Y_1, \ldots, Y_m)} = F \subset U_1 \times \ldots U_n \times V_1 \ldots \times V_m$$

$$\text{y de} \quad q \longrightarrow \Pi_{(X_1, \ldots, X_n, Z_1, \ldots, Z_k)} = G \subset U_1 \times \ldots U_n \times T_1 \ldots \times T_k$$

$$\text{se deduce} \quad r \longleftarrow \Pi_{(X_1, \ldots, X_n, Y_1, \ldots, Y_m, Z_1, \ldots, Z_m)} = \overline{F} \cap \overline{G}$$

siendo $\overline{F}$ y $\overline{G}$ las correspondientes extensiones cilíndricas de F y G.

Como caso particular se tiene:

$$p \longrightarrow \Pi_{(X,Y)} = F \subset U \times V$$

$$q \longrightarrow \Pi_{(Y,Z)} = G \subset V \times T$$

$$r \longleftarrow \Pi_{(X,Y,Z)} = \overline{F} \cap \overline{G} \subset U \times V \times T$$

que utilizando para la intersección el operador mínimo proporciona:

$$\pi_{(X,Y,Z)}(u,v,t) = \min \left\{ \mu_F(u,v), \mu_G(v,t) \right\} \qquad (1)$$

.) <u>Regla de composición</u>.

Si se combinan ambos principios, es decir, si ahora se aplica el principio de proyección a (1), se obtiene:

$$\pi_{(X,Z)}(u,t) = \sup \left\{ \min \left\{ \mu_F(u,v), \mu_G(v,t) \right\} \mid v \in V \right\} \qquad (2)$$

que de forma esquemática se puede representar:

$$\begin{array}{rl} \text{de} & p \longrightarrow \Pi_{(X,Y)} = F \\ \text{y de} & q \longrightarrow \Pi_{(Y,Z)} = G \\ \hline \text{se deduce} & r \longleftarrow \Pi_{(Y,Z)} = F \circ G \end{array}$$

donde F o G es la composición de las relaciones difusas $F \subset U \times V$ y $G \subset V \times T$, definida como, (KAUF-73):

$$\mu_{F \circ G}(u,t) = \sup_{v} \left\{ \min \left\{ \mu_F(u,v), \mu_G(v,t) \right\} \right\}$$

Esta regla de deducción es lo que se conoce como "compositional rule of inference".

.) El "modus ponens generalizado" que se desea obtener es la aplicación de la regla de composición a una situación aún más particular:

$$p \longrightarrow \Pi_X = F \subset U$$

$$q \longrightarrow \Pi_{(X,Y)} = G \subset U \times V$$

$$\overline{}$$

$$r \longleftarrow \Pi_Y = F \circ G \subset V$$

que traducido a los siguientes términos:

de $p \longrightarrow q \equiv$ "si X es A entonces Y es B" $\longrightarrow \Pi_{(X,Y)} = G$

y de $p' \equiv$ "X es A'" $\longrightarrow \Pi_X = F = A'$

se deduce $q' \equiv$ "Y es B'" $\longleftarrow \Pi_Y = F \circ G = A' \circ G$

proporciona la deducción de la proposición difusa "Y es B'", definida por:

$$\boxed{\mu_{B'}(v) = \sup \left\{ \min \left\{ \mu_G(u,v), \mu_{A'}(u) \right\} \mid u \in U \right\}} \qquad (3)$$

Según la representación que se utilice para la proposición "si X es A entonces Y es B" se obtendrán diferentes formas para la proposición "Y es B'".

2.4.3.2.- <u>Representación de la proposición "si X es A entonces Y es B" (II).</u>

El esquema (3) del modus ponens generalizado puede ser aplicado para obtener una representación de la premisa $p \longrightarrow q$. En efecto, si en el esquema anterior se hace A = A', entonces el resultado de la inferencia será B'= B.

"si X es A entonces Y es B" $\longrightarrow \Pi_{(X,Y)} = G$

"X es A" $\longrightarrow \Pi_X = A$

"Y es B" $\longleftarrow \Pi_Y = B$

Se desea ahora conocer $\Pi_{(X,Y)}$, que también se denota por $\Pi_{Y|X}$, distribución de posibilidad condicional, que restringe los valores de Y conociendo el valor de X, ($\pi_{Y|X}(u,v)$ es la posibilidad de que Y = v, en el supuesto de que X = u). En otros términos, se trata de encontrar una solución de B = A o G, o de forma equivalente, obtener la distribución de posibilidad $\Pi_{Y|X}$ = G que verifica:

$$\mu_B(v) = \sup \left\{ \min \{ \mu_A(u), \mu_G(u,v) \} \mid u \in U \right\} \qquad (4)$$

La mayor solución de esta ecuación, en el sentido de la inclusión, (PRAD-82), (DUBO-85), en el supuesto de que A está normalizado, es:

$$\pi_{Y|X}(u,v) = \sup \left\{ x \mid \min (\mu_A(u), x) \leq \mu_B(v) \right\}$$

o de otra forma

$$\pi_{Y|X}(u,v) = \begin{cases} 1 & \text{si } \mu_A(u) \leq \mu_B(v) \\ \mu_B(v) & \text{si } \mu_A(u) > \mu_B(v) \end{cases} \qquad (5)$$

que es la implicación multivalente de Gödel: $v(p \rightarrow q) = 1$ si $v(p) \leqslant v(q)$, $v(p \rightarrow q) = v(q)$ si $v(p) > v(q)$.

2.4.3.3.- Modus ponens generalizado (II).

Si, para simplificar la notación, representamos (5) de forma abreviada por $\mu_{A \rightarrow B}(u,v)$, la ecuación (3) se convierte en:

$$\boxed{\mu_{B'}(v) = \sup \left\{ \min (\mu_{A \rightarrow B}(u,v), \mu_{A'}(u)) \mid u \in U \right\}}$$

Algunas consideraciones importantes que deben hacerse acerca de este esquema de modus ponens generalizado son las siguientes:

1.- $\mu_{AxB}(u,v) = \min (\mu_A(u), \mu_B(v))$ también es solución de (4), caso de que A esté normalizado. Sin embargo, no es la mayor solución y utilizar μ_{AxB} en vez de $\mu_{A \rightarrow B}$ conduciría a un B' más pequeño (en el sentido de la inclusión) y por tanto a un resultado arbitrariamente preciso respecto a las informaciones contenidas en las premisas.

2.- Son de fácil comprobación las siguientes propiedades, (DUBO-85), (PRADE-82):

a) A' = A y normalizado, se obtiene B' = B, de acuerdo con el modus ponens de la lógica clásica.

b) Si A' está normalizado y A' $\subset$ A, se obtiene B' = B. También es una propiedad bastante intuitiva. Por ejemplo, de

$$p \longrightarrow q \;\equiv\; \text{"Si Pedro es joven entonces Pedro es ágil"}$$

$$p' \equiv \text{"Pedro es muy joven"}$$

se deduce, por supuesto, "Pedro es ágil".

3.- Para cualquier A' se verifica:

$$\mu_{B'}(v) \geqslant \sup \left\{ \mu_{A'}(u) \mid u \in \mu_A^{-1}(0) \right\} \quad \text{para todo } v \in V$$

En particular, si existe algún $u_o \in U$ que hace $\mu_A(u_o) = 0$ y $\mu_{A'}(u_o) = 1$ se tendrá $\mu_{B'}(v) \geqslant 1$, es decir $\mu_{B'}(v) = 1$ para todo V y, en consecuencia, lo que se deduce mediante el modus ponens generalizado es B' = V, o lo que es lo mismo, Y toma valores en V: no hay ninguna deducción. La desigualdad anterior se puede interpretar como la existencia de un cierto nivel uniforme de indeterminación en la conclusión "Y es B'" cuando una parte significativa de A' no está incluída en A.

4.- Si $A' = A^C$ y existe u_o tal que $\mu_A(u_o) = 0$ entonces B' = V, es decir no se deduce ninguna conclusión, lo cual parece natural.

Todas estas propiedades, además de precisar lo que significa la inferencia con proposiciones difusas mediante el modus ponens generalizado, resultan de mucha utilidad cuando se trata de implementar este esquema de razonamiento para conjuntos difusos representados de alguna forma concreta, como distribuciones trapezoidales o mediante "curvas en S".

2.4.4.- Aplicación de la Lógica Difusa a los Sistemas Expertos.

Es conveniente indicar, en primer lugar, que se siguen desarrollando nuevas reglas de inferencia para proposiciones con predicados difusos. El modus ponens generalizado es tan sólo la regla básica, aplicable a proposiciones sencillas. Zadeh (ZADE-83b) propone múltiples reglas de inferencia con proposiciones en las que aparecen cuantificadores difusos. A título de ejemplo, se cita el denominado "intersection/product syllogism":

$$\frac{\begin{array}{l} Q_1 \ A'_s \text{ son } B'_s \\ Q_2 \ (A \text{ y } B)'_s \text{ son } C'_s \end{array}}{Q_1 \otimes Q_2 \ A'_s \text{ son } (B \text{ y } C)'_s}$$

donde Q_1 y Q_2 son cuantificadores difusos y $\otimes$ es el producto de números difusos. Son ya mecanismos de inferencia mucho más especializados, pero es conveniente saber de su existencia. Sería interesante también su implementación a casos reales concretos.

También se han desarrollado múltiples versiones de "Lógicas Difusas" en las que los valores de verdad de las distintas proposiciones son valores lingüísticos considerados como subconjuntos difusos del intervalo $[0,1]$. En este sentido, son destacables los trabajos de Baldwin, (BALD-79), (BALD-80).

En esta línea de cualificación de proposiciones mediante valores de verdad lingüísticos, debe hacerse referencia al sistema MILORD, desarrollado en Blanes por López de Mántaras y otros, (LOPE-87). Se trata de un sistema experto para el diagnóstico de neumonías, que no utiliza valores numéricos para representar la incertidumbre, sino valores lingüísticos que internamente se representan como subconjuntos difusos del intervalo unidad, convenientemente parametrizados.

Cada hecho o regla de la base de conocimientos viene calificada con un valor lingüístico τ . El esquema básico de inferencia es:

$$\frac{\begin{array}{l}\text{si (X es A) entonces (Y es B) es } \tau_1 \\ \quad \text{(X es A) es } \tau_2\end{array}}{\text{(Y es B) es } \tau_3}$$

obteniéndose τ_3 a partir de la representación de los valores lingüísticos τ_1, τ_2 y mediante unas determinadas funciones de implicación.

Sin embargo, este esquema no es el utilizado por el sistema, sino un esquema más complejo deducido de éste y de la aplicación del concepto de compatibilidad de conjuntos difusos. Este nuevo esquema es de la forma

$$\frac{\begin{array}{l}\text{si ((X es A) es } \tau_1\text{) entonces ((Y es B) es } \tau_2\text{) es } \tau_3 \\ \quad \text{(X es A) es } \tau_4\end{array}}{\text{(Y es B) es } \tau_5}$$

Un ejemplo de la compleja regla que aparece en este esquema tomada del propio sistema MILORD es:

R08004 IF 1) Community acquired pneumonia is *almost sure*
2) Bacterian disease Is *possible*
3) (no aspiration) is *very possible*
THEN [Possible]
Enterobacteria is *quite possible*

Para realizar la equiparación del hecho ((X es A) es τ_4), cuyo valor de incertidumbre viene determinado por el subconjunto difuso τ_4, con el antecedente de la regla, ((X es A) es τ_1), se utiliza la compatibilidad de conjuntos difusos de Zadeh, CP(A; A'), que es también un subconjunto difuso de $[0,1]$ y que viene definido por la expresión, ver (ZADE-79a) para más detalles:

$$\mu_{CP(A;A')}(v) = \begin{cases} \sup \left\{ \mu_{A'}(u) \mid \mu_A(u) = v \right\} & \text{si } \mu_A^{-1}(v) \neq \phi \\ 0 & \text{en otro caso} \end{cases}$$

En este concepto de compatibilidad se basa la representación de proposiciones del tipo "(X es A) es τ". Esta proposición se representa por (X es A'), donde A' es el mayor subconjunto difuso, del mismo referencial que A, tal que CP(A;A') = τ, obteniéndose:

$$\mu_{A'} = \tau \circ \mu_A$$

con el significado clásico de composición de funciones.

Si ahora "(X es A) es τ_1" viene representada por A' y la proposición "(X es A) es τ_4)"se representa por A", la equiparación de ambas proporciona CP(A';A").

El conjunto de valores lingüísticos que se tienen en cuenta en MILORD está formado por nueve términos: "imposible", "casi imposible", "débilmente posible", "posible", ..., "casi seguro", "seguro", que son con los que el experto y el usuario evalúan la incertidumbre de reglas y hechos, respectivamente.

Las funciones de implicación utilizadas, así como las operaciones de disyunción y conjunción, vienen expresadas en función de distintas normas y co-normas triangulares, y para cada uno de los nueve valores lingüísticos considerados, cada una de las operaciones es pre-computada para las diferentes normas y co-normas más frecuentes, lo que proporciona una gran flexibilidad al sistema, pues permite al experto experimentar con diferentes esquemas de razonamiento y elegir aquél que considere más adecuado al problema que se desea tratar con este modelo.

2.5.- LAS MEDIDAS DE POSIBILIDAD Y NECESIDAD DE DUBOIS Y PRADE.

Dubois y Prade, (PRAD-82), (DUBO-85b), proponen otro sistema diferente para cuantificar la incertidumbre que, al igual que ocurre en la Teoría de la Evidencia de Dempster-Shafer y en la cuantificación de la incertidumbre de reglas utilizada en PROSPECTOR, se compone de dos números o medidas diferentes, llamadas Necesidad y Posibilidad. Ambas están estrechamente relacionadas con el concepto de posibilidad de Zadeh y, por tanto, constituyen un punto de partida para el tratamiento conjunto de incertidumbre e imprecisión, es decir, de reglas y hechos que son al mismo tiempo inciertos y difusos.

En la exposición que sigue a continuación, se considerarán simultáneamente un conjunto de proposiciones $\mathcal{P}$ y una familia de conjuntos $\mathcal{P}(U)$ con el fin de interpretar y precisar el sentido de los conceptos que se vayan definiendo.

El conjunto de proposiciones $\mathcal{P}$ será tal que: i) $p \in \mathcal{P} \Longrightarrow \neg p \in \mathcal{P}$, ii) $p, q \in \mathcal{P} \Longrightarrow p \wedge q \in \mathcal{P}$. $0 = p \wedge \neg p$ será la proposición siempre falsa y $1 = \neg (p \wedge \neg p)$ la proposición siempre verdadera. La disyunción se definirá como es habitual por $p \vee q = \neg(\neg p \wedge \neg q)$ y la implicación $p \longrightarrow q$ por $\neg p \vee q$. Se dirá que dos proposiciones son incompatibles si $(p \wedge q) = 0$, i.e., si una de ellas es cierta, la otra es falsa. Se dirá también que p implica q si y sólo si $p \longrightarrow q = 1$. $(\mathcal{P}, \wedge, \vee, \neg)$ es álgebra de Boole.

Con notación conjuntista se considerará un conjunto de sucesos, subconjuntos de un referencial U, que será el suceso siempre cierto. El conjunto vacío será el suceso siempre falso. $(\mathcal{P}(U), \cap, \cup, \bar{\ })$ es también un álgebra de Boole.

No hay inconveniente en este uso simultáneo de proposiciones o sucesos, dado el isomorfismo existente entre estas dos estructuras. Además el orden parcial definido en $\mathcal{P}(U)$ por la relación de inclusión se corresponde con el orden parcial definido en $\mathcal{P}$ por el conectivo lógico de implicación.

2.5.1.- Medidas de confianza. La medida de Posibilidad.

Una medida de confianza es una función $g: \mathcal{P}(U) \longrightarrow [0,1]$ tal que:

i) $g(\emptyset) = 0$
ii) $g(U) = 1$ (1)
iii) $A \subsetneq B \Longrightarrow g(A) \leq g(B) \quad \forall A,B \subset U$
(monotonía en el sentido de la inclusión)

Si U es infinito, es habitual añadir un axioma de continuidad.

En términos de proposiciones, la definición equivalente sería: $g: \mathcal{P} \longrightarrow [0,1]$ es medida de confianza si y sólo si

i) $g(0) = 0$
ii) $g(1) = 1$
iii) si $p \rightarrow q = 1$ entonces $g(p) \leq g(q)$

Una medida de confianza no satisface en general la propiedad de aditividad, $\forall A,B \subset U$, $A \cap B = \emptyset$, $g(A \cup B) = g(A) + g(B)$, por lo que no pueden ser consideradas como una medida en el sentido matemático de este término.

Los axiomas exigidos pueden ser justificados de la siguiente forma. Considérese que la función g está asociada a una variable u que toma valores en U. g(A) significará entonces el grado de confianza en la proposición "u toma su valor en A". El axioma iii) expresa que el grado de confianza en "u toma su valor en B" es mayor que el grado de confianza en "u toma su valor en A" si $A \subsetneq B$. i) es coherente con el hecho de que u toma un valor y finalmente ii) parece natural, puesto que efectivamente hay certeza en que u toma su valor en U.

Como consecuencia inmediata del axioma iii) se deducen las siguientes desigualdades:

$$\forall A,B \subsetneq U \qquad g(A \cap B) \leq \min(g(A), g(B))$$
$$g(A \cup B) \geq \max(g(A), g(B))$$

que en términos de proposiciones:

$$\forall\ p,q \in \mathcal{P} \qquad g(p \wedge q) \leq \min\ (g(p),\ g(q))$$
$$g(p \vee q) \geq \max\ (g(p),\ g(q))$$

expresan que el grado de confianza en "$p \wedge q$ es cierto" (respectivamente "$p \vee q$ es cierto") es a lo sumo (al menos) igual al menor (mayor) de los grados de confianza en "p es cierto" y "q es cierto".

Los axiomas mínimos exigidos caracterizan a una familia demasiado amplia de medidas de confianza con la cual no es fácil trabajar. Una hipótesis natural consiste en suponer que el grado de confianza de la unión de sucesos disjuntos (de proposiciones incompatibles) es dependiente únicamente de los grados de confianza de cada uno de los sucesos (proposiciones) respectivas:

$$\forall\ A,B \subseteq U \mid A \cap B = \emptyset \qquad g(A \cup B) = g(A) * g(B) \qquad (2)$$

donde * es una operación interna sobre $[0,1]$.

La estructura de $\mathcal{P}(U)$ induce las condiciones que debe cumplir la operación *:

- La asociatividad y conmutatividad de $\cup$ obliga que * sea una operación asociativa y conmutativa.

- $\forall\ A \subset U \quad g(A) = g(A \cup \emptyset) = g(A) * g(\emptyset) = g(A) * 0$ y, por tanto,
  $\forall\ a \in [0,1] \quad a*0 = a$

- Sean $A \subseteq B$ y $C \subseteq D$. Luego $g(A) \leq g(B)$ y $g(C) \leq g(D)$. Como además $A \cup C \subseteq B \cup D$ se debe tener $g(A \cup C) \leq g(B \cup D)$. En consecuencia * debe cumplir:

$$\left.\begin{array}{l} a \leq b \\ \\ c \leq d \end{array}\right\} \Longrightarrow a * c \leq b * d$$

- Como consecuencia de esta última condición

$$\left.\begin{matrix} 0 \leq a \\ 1 \leq 1 \end{matrix}\right\} \longrightarrow \begin{matrix} 1 * 0 \leq 1 * a \leq 1 \\ 1 * 0 = 1 \end{matrix} \longrightarrow 1 * a = 1 \qquad \forall a$$

De la exigencia de estas cuatro condiciones que debe satisfacer la operación * se deduce que es una conorma triangular, (PRAD-82).

Propiedades muy útiles que se siguen de (2) son:

- $\forall A \subseteq U \qquad g(A) * g(\bar{A}) = 1 \qquad (3)$

- El axioma (2) permite calcular, caso de que U sea finito, el grado de confianza en un suceso compuesto a partir de los grados de confianza en los sucesos elementales que lo componen. Bastará definir g para cada elemento de U.

Si $U = \{u_1, \ldots, u_n\}$, $A = \{u_{i_1}, \ldots, u_{i_p}\}$ y $g_i = g(\{u_i\})$

$$g(A) = g\left(\bigcup_{j=1}^{p} \{u_{i_j}\}\right) = g_{i_1} * \ldots * g_{i_p} \qquad (4)$$

El axioma ii) se convierte ahora en

$$g_1 * \ldots * g_n = 1 \qquad (5)$$

que se verifica si $\exists\, u_i \in U \mid g_i = 1$.

Según cual sea la conorma triangular que se considere para la operación * se tendrán diferentes medidas de confianza. Si se toma para $a * b = \max(a,b)$ se obtiene:

(2) $\longrightarrow \forall A,B \subseteq U \mid A \cap B = \emptyset \qquad g(A \cup B) = \max(g(A), g(B))$ aunque este resultado, como puede probarse con cierta facilidad, es válido aunque A y B no sean disjuntos:

$$\forall A,B \subseteq U \quad g(A \cup B) = \max (g(A), g(B))$$

Las medidas de confianza que verifican esta condición se llaman medidas de posibilidad y se representan por $\Pi(A)$.

Una primera justificación del término posibilidad para designarlas radica en el hecho siguiente: puesto que max $(g(A), g(B))$ era una cota inferior del grado de confianza en el suceso $A \cup B$, al tomar $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$ se toma la más prudente de las posibles evaluaciones del grado de confianza en $A \cup B$, lo que es coherente con la semántica de los juicios de posibilidad, tanto material de factibilidad (para realizar $A \cup B$ basta realizar el más fácil de los dos, el más costoso, el más posible) como subjetivo (juicios que comprometen poco al que los emite).

Los axiomas característicos de una medida de posibilidad quedan:

$$\begin{aligned} &\Pi(\emptyset) = 0 \\ &\Pi(U) = 1 \\ &A \subseteq B \Longrightarrow \Pi(A) \leq \Pi(B) \\ &\forall A,B \subseteq U \quad \Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \end{aligned} \qquad (6)$$

Este último axioma, si U no es finito, se reemplaza por

$$\Pi(\cup A_i) = \sup \{\Pi(A_i)\}$$

El resultado (3) se convierte en

$$\forall A \subseteq U \qquad \max(\Pi(A), \Pi(\bar{A})) = 1 \qquad (7)$$

que expresa el hecho de que dos sucesos contrarios al menos uno de ellos es completamente posible. Además, cuando un suceso es considerado posible, esto no impide que el suceso contrario sea también posible. Lo que no puede ocurrir, por ejemplo, es que $\Pi(A) = \Pi(\bar{A}) = 0$ lo que se interpretaría como que una variable u que toma valores en U no toma valores ni en A ni en el subconjunto complementario.

Esta relación (7) muestra hasta qué punto son diferentes las medidas de posibilidad y necesidad. La posibilidad de un suceso, $\Pi(A)$, y la posibilidad del suceso contrario $\Pi(\bar{A})$ están ligadas entre sí mediante dicha relación, pero no es posible deducir una de la otra. En cambio, para las medidas de probabilidad se tiene $P(\bar{A}) = 1 - P(A)$.

Se tiene además:

$$\max_{i=1,n} g_i = 1 \qquad (8)$$

$$g(A) = \max \left\{ g_{i_j} \right\} \quad \text{con } A = \left\{ u_{i_1}, \ldots, u_{i_p} \right\} \qquad (9)$$

La densidad $\left\{ g_i \right\}_{i=1,n}$ se llama distribución de posibilidad asociada a la medida de posibilidad Π y se representa por $\pi(u)$. Con esta notación (9) y (8) quedan:

$$\Pi(A) = \max \left\{ \pi(u_{i_j}), \quad j = 1, \ldots, p \right\}$$

$$\max \left\{ \pi(u_i) \quad i = 1, \ldots, n \right\} = 1$$

o en el caso más general de U no finito:

$$\Pi(A) = \sup \left\{ \pi(u) \mid u \in A \right\} \qquad (10)$$

$$\sup \left\{ \pi(u) \mid u \in U \right\} = 1 \qquad (11)$$

donde $\pi(u) = \Pi(\{u\})$

En el caso de U finito, la condición (ii) equivale a que $\exists u_o \mid \pi(u_o)) = 1$. Se dirá entonces que la distribución de posibilidad está normalizada.

Es importante observar que el concepto de medida de posibilidad de Dubois y Prade coincide con el concepto de distribución de posibilidad de

Zadeh, aunque el proceso seguido es inverso. En efecto, (ZADE-78), si se tiene la proposición "X es B", donde la variable X toma valores en un universo U, si Π_X es la distribución de posibilidad asociada a la variable X y si finalmente A es un subconjunto no difuso de U, Zadeh propone

$$\Pi(A) = \text{Pos}\ (X \in A \mid X \text{ es } B) = \sup \left\{ \pi_X(u) \mid u \in A \right\} \qquad (12)$$

donde $\pi_X(u)$ tiene el significado que se comentó al introducir las distribuciones de posibilidad de Zadeh

$$\pi_X(u) = \text{Pos}\ (X = u \mid X \text{ es } B) = \mu_B(u)$$

Retomando el ejemplo que en ese momento se utilizaba, la proposición "X es un entero pequeño" induce la distribución de posibilidad

$$\pi_X(u) = \mu_{\text{"enteros pequeños"}}(u)$$

y se tendrá

$$\text{Pos}\ (X \in [0,2]\ ) = \max\ (\mu_{e.p.}(0), \mu_{e.p.}(1), \mu_{e.p.}(2))$$

2.5.2.- <u>Medidas duales de una medida de confianza</u>. <u>Medidas de Necesidad</u>.

Se ha visto que, en general, dada una medida de confianza g, g(A) y $g(\bar{A})$ no son deducibles la una de la otra, caso de las medidas de posibilidad, por ejemplo, y, por tanto, no constituyen informaciones complementarias.

Por otra parte $f: \mathcal{P}(U) \longrightarrow [0,1]$ definida por $f(A) = g(\bar{A})$ no es una medida de confianza, pues $f(\phi) = 1$, $f(U) = 0$ y es decreciente. Si se busca una función $c: [0,1] \rightarrow [0,1]$ tal que $g_c = c \circ f$ sea una medida de confianza y si además se desea recuperar g a partir de g_c como $g(A) = c\left[g_c(\bar{A})\right]$, se llega a la conclusión de que c debe ser una negación fuerte, (TRIL-79), de las cuales la más sencilla es $c(x) = 1-x$.

En consecuencia, a partir de g, medida de confianza y de c, negación fuerte, es posible construir una nueva medida de confianza g_c definida por

$$g_c(A) = c\left[g(\bar{A})\right] \qquad (13)$$

Se dice que g y g_c son c-duales. Si $c(x) = 1-x$, se dice simplemente que $g_c = g$ y g son duales.

Si además g es una medida de confianza basada en una conorma triangular $*$ se tiene

$$\forall A,B \subseteq U \mid A \cup B = U \qquad g_c(A \cap B) = c\left[c\left[g_c(A)\right] * c\left[g_c(B)\right]\right]$$

pero si $*$ es una conorma triangular y c una negación fuerte, entonces $a \perp b = c\left[c(a) * c(b)\right]$ es una norma triangular y $a * b = c\left[c(a) \perp c(b)\right]$. Se dice que $*$ y $\perp$ son c-duales. Si $c(x) = 1-x$, se dirá simplemente que son duales.

Por tanto si g es una medida de confianza basada en la conorma triangular $*$ y $\perp$ es la norma triangular c-dual de $*$, la función $g_c: \mathcal{P}(U) \longrightarrow [0,1]$ definida por $g_c(A) = c\left[g(\bar{A})\right]$ es una medida de confianza y se verifica:

$$\forall A,B \subsetneq U \mid A \cup B = U \qquad g_c(A \cap B) = g_c(A) \perp g_c(B) \qquad (14)$$

Las propiedades que verifican las medidas de confianza basada en normas triangulares son duales de las que se cumplen para las medidas de confianza basadas en conormas triangulares y fácilmente deducibles de éstas:

- $\forall A \subsetneq U \quad g_c(A) \perp g_c(\bar{A}) = 0 \qquad (15)$

- $g_c(A)$ se puede expresar en función de $g_c(U - \{u_i\}) = \hat{g}_i$:

$$U = \{u_1, \ldots, u_n\} \qquad A = \{u_{i_1}, \ldots, u_{i_p}\} \qquad \bar{A} = \{u_{i_{p+1}}, \ldots, u_{i_n}\}$$

$$g_c(A) = \hat{g}_{i_{p+1}} \perp \ldots \perp \hat{g}_{i_n} \qquad (16)$$

- La condición $g_c(\phi) = 0$ queda ahora $\hat{g}_1 \perp \ldots \perp \hat{g}_n = 0$ (17)

- $g_c(A) = 1 \Longrightarrow g_c(\bar{A}) = 0$

Si se considera la norma triangular a b = min (a,b), dual de la conorma triangular utilizada para definir las medidas de posibilidad, se obtiene:

(14) $\longrightarrow \forall A,B \subseteq U \mid A \cup B = U \quad g_c(A \cap B) = \min (g_c(A), g_c(B))$
igualdad que también se verifica para cualquier A y B.

Una medida de confianza que satisface esta condición se llama <u>medida de necesidad</u>. Se notará por N(A) y son duales de las medidas de posibilidad Π :

$$N(A \cap B) = \min (N(A), N(B)) \qquad (18)$$

$$N(A) = 1 - \Pi(\bar{A}) \qquad (19)$$

Esta última igualdad es una transcripción numérica de la dualidad de los conceptos de necesidad y posibilidad: si un suceso es necesario entonces su contrario es imposible ($N(A) = 1 \Longrightarrow \Pi(\bar{A}) = 0$).

(15) $\longrightarrow \forall A \subseteq U,\ \min (N(A), N(\bar{A})) = 0$ (20)
que expresa el hecho de que dos sucesos contrarios no pueden ser simultáneamente aún un poco necesarios: si un suceso A tiene una necesidad no nula, la necesidad del suceso contrario $\bar{A}$ ha de ser nula, pudiendo ambas tener necesidad nula.

(17) $\longrightarrow \min_{i=1,n} \{\hat{g}_i\} = 0 \quad$ con $\hat{g}_i = N(U - \{u_i\}) = 1 - \Pi(\{u_i\}) = 1 - g_i$

(16) se puede expresar como

$$N(A) = \min \left\{ \hat{g}_{i_j} \mid j = p+1, \ldots, n \right\} = \min \left\{ 1 - g_{i_j} \right\} =$$

$$= 1 - \max \left\{ g_{i_j} \right\} = 1 - \max \left\{ \pi(u) \mid u \notin A \right\} =$$

$$= \min \left\{ 1 - \pi(u) \mid u \notin A \right\}$$

que en el caso de U no finito se convierte en

$$N(A) = \inf \left\{ 1 - \pi(u) \mid u \notin A \right\} \qquad (21)$$

De las relaciones max $(\Pi(A), \Pi(\bar{A})) = 1$, min $(N(A), N(\bar{A})) = 0$ y $N(A) = 1 - \Pi(\bar{A})$ se deducen las siguientes consecuencias:

- $\Pi(A) \geqslant N(A)$

- $N(A) > 0 \Longrightarrow \Pi(A) = 1$
  $\Pi(A) < 1 \Longrightarrow N(A) = 0$

  que expresan que un suceso ha de ser completamente posible antes de ser aún un poco necesario.

Están ya claramente de manifiesto las diferencias entre las medidas de probabilidad y las medidas de posibilidad y necesidad. En efecto, un suceso que tiene probabilidad 1 es visto como cierto. No es el caso para un suceso que tiene posibilidad 1, pues el suceso contrario puede tener también posibilidad 1. Por el contrario, si la necesidad de un suceso es igual a 1, puede ser considerado como cierto, pues la necesidad del suceso contrario es cero, así como su posibilidad ($N(A) = 1 \Longrightarrow N(\bar{A}) = 0$ y $\Pi(\bar{A}) = 1 - N(A) = 0$)

Obsérvese además que si P es medida de probabilidad se verifica $P(A) + P(\bar{A}) = 1$; es decir, la probabilidad de un suceso determina completamente la probabilidad del suceso contrario, mientras que para las medidas de posibilidad se tiene $\Pi(A) + \Pi(\bar{A}) \geqslant 1$ y $N(A) + N(\bar{A}) \leqslant 1$ y, por tanto, para caracterizar la incertidumbre de un suceso, se necesitan dos números, $\Pi(A)$ y $N(A)$.

Finalmente se debe hacer notar que todos los resultados relativos a medidas de posibilidad y necesidad han sido obtenidos con el enfoque conjuntista, considerando un conjunto U donde toma valores de variable u y con sucesos $A \subseteq U$. Estos resultados tienen una traducción inmediata si se considera un conjunto de proposiciones $\mathcal{P}$. Se utilizarán de ambas formas, según el contexto requiera uno u otro tipo de expresiones.

2.5.3.- Mecanismos de inferencia.

Considérese ahora el conjunto del conocimiento representado por un conjunto de proposiciones $\mathcal{P}$, en las condiciones indicadas al comienzo de este apartado 2.5. Tanto los hechos (proposiciones) como las reglas (proposiciones del tipo $p \longrightarrow q$) serán considerados como inciertos, evaluando su incertidumbre mediante los dos valores duales de necesidad y posibilidad.

En la lógica clásica los dos procedimientos de inferencia más utilizados son el "modus ponens" y el "modus tollens".

El "modus ponens" corresponde a la primera línea de la tabla de verdad

$$\frac{\begin{array}{l} p \longrightarrow q \\ p \end{array}}{q}$$

$v(p \longrightarrow q)$	$v(p)$	$v(q)$
V	V	V
V	F	$\{V,F\}$
F	V	F
F	F	$\emptyset$

El "modus tollens", a su vez es consecuencia de la segunda línea de la tabla de verdad siguiente:

$$\frac{\begin{array}{l} p \longrightarrow q \\ \neg q \end{array}}{\neg p}$$

$v(p \longrightarrow q)$	$v(q)$	$v(p)$
V	V	$\{V,F\}$
V	F	F
F	V	$\emptyset$
F	F	V

En ambas tablas de verdad aparecen situaciones imposibles, lo que indica que los valores de verdad de $p \longrightarrow q$ y p(o de $p \longrightarrow q$ y q) no pueden elegirse independientemente. Asimismo aparecen situaciones en las que no es posible deducir nada, que se corresponden con las líneas cuyo resultado es $\{V,F\}$

Dubois y Prade proponen, a partir de las propiedades de las medidas de posibilidad y necesidad los siguientes esquemas deductivos como extensiones del "modus ponens" y "modus tollens" cuando las premisas de éstos, una regla y un hecho en ambos casos, son inciertas.

(I)
$$\frac{\begin{array}{ll} N(p \longrightarrow q) & \geq a \\ N(p) & \geq b \end{array}}{N(q) \geq \min(a,b)}$$

En efecto:
$$\left.\begin{array}{l} p \wedge q = p \wedge (\neg p \vee q) = p \wedge (p \longrightarrow q) \\ p \wedge q \longrightarrow q \end{array}\right\} \longrightarrow$$

$$\left.\begin{array}{r} N(q) \geq N(p \wedge q) = \min(N(p), N(p \longrightarrow q)) \\ N(p \longrightarrow q) \geq a \\ N(p) \geq b \end{array}\right\} \longrightarrow N(q) \geq \min(a,b)$$

(II)
$$\frac{\begin{array}{ll} N(p \longrightarrow q) & \geq a \\ \Pi(p) & \geq b \end{array}}{\Pi(q) \geq b.v(a+b > 1)}$$

donde $v(a+b>1)$ es el valor de verdad de la proposición $a+b>1$:

$$v(a+b>1) = \begin{cases} 1 \text{ si } a+b>1 \\ 0 \text{ si } a+b \leq 1 \end{cases}$$

que se deduce utilizando las mismas propiedades básicas de las medidas de necesidad aplicadas a la equivalencia lógica $\neg p \wedge \neg q = \neg q \wedge (p \longrightarrow q)$ y a la tautología $\neg p \wedge \neg q \longrightarrow \neg p$, llegándose a

$$1-\Pi(p) \geq \min(1-\Pi(q), N(p \longrightarrow q))$$

$$\begin{array}{ll} \text{(III)} & \Pi(p \longrightarrow q) \geqslant a \\ & N(p) \geqslant b \\ \hline & \Pi(q) \geqslant a . v(a+b > 1) \end{array}$$

Deducida de $p \rightarrow q = \neg p \vee q$, obteniéndose $\Pi(p \rightarrow q) = \max(1-N(P), \Pi(q))$

La combinación de $\Pi(p \longrightarrow q) \geqslant a$ y $\Pi(p) \geqslant b$ no permite deducir nada sobre $\Pi(q)$. Sin embargo, sí es posible extraer alguna conclusión. En efecto, si $\Pi(p) < 1$, entonces $N(p) = 0$ y de la fórmula de donde se dedujo el esquema (III), se obtiene $\Pi(p \longrightarrow q) = 1$. Es decir, se concluye que $\max(\Pi(p), \Pi(p \longrightarrow q)) = 1$ para que exista coherencia.

Para el "modus tollens", y mediante la utilización de procedimientos parecidos, Dubois y Prade deducen los siguientes esquemas:

$$\begin{array}{ll} \text{(IV)} & N(p \longrightarrow q) \geqslant a \\ & N(q) \leqslant b \\ \hline & N(p) \leqslant 1-(1-b) . v(a > b) \end{array} \qquad \begin{array}{ll} \text{(V)} & N(p \longrightarrow q) \geqslant a \\ & \Pi(q) \leqslant b \\ \hline & \Pi(p) \leqslant \max(1-a, b) \end{array}$$

$$\begin{array}{ll} \text{(VI)} & \Pi(p \longrightarrow q) \geqslant a \\ & N(q) \leqslant b \\ \hline & N(p) \leqslant 1-a . v(b < a) \end{array}$$

Estos esquemas de razonamiento pueden ser combinados de muy diferentes formas, tal y como reflejan en su trabajo los autores. El que es utilizado en la implementación de motor de inferencia con tratamiento de incertidumbre e imprecisión que servirá de punto de partida para el modelo que se presenta en esta tesis, es el siguiente, resultado de combinar los esquemas (I) y (V):

$$
\begin{array}{ll}
N(p \longrightarrow q) \geqslant a & \\
N(q \longrightarrow p) \geqslant a' & \\
N(p) \geqslant b & \text{con } \max(1-b, b') = 1 \\
\Pi(p) \leqslant b' & \\
\hline
N(q) \geqslant \min(a, b) & \\
\Pi(q) \leqslant \max(1-a', b') &
\end{array}
$$

Es fácil comprobar que se verifica siempre $\max(1 - N(q), \Pi(q)) = 1$.

Sobre este esquema de razonamiento es preciso hacer dos observaciones importantes. La primera de ellas hace relación a que al considerar el par $(N(p), \Pi(p))$ y, según el sentido de las desigualdades que aparecen en el esquema, el valor $N(p)$ es considerado como una cota inferior, mientras que el de $\Pi(p)$ es considerado como una cota superior. Esto es debido a que $\Pi(p)$ puede interpretarse como un valor máximo de la verdad o credibilidad relativa al hecho $\underline{p}$, mientras que $N(p) = 1- \Pi(\neg p)$ será un valor mímino, pues con la misma interpretación $\Pi(\neg p)$ sería un valor máximo acerca de la verdad de la negación de $\underline{p}$.

En segundo lugar, los valores a y a', cotas inferiores de $N(p \longrightarrow q)$ y $N(q \longrightarrow p)$, son considerados por Dubois y Prade como grados de suficiencia y de necesidad para que $\underline{q}$ sea cierto cuando $\underline{p}$ es verdadero.

Parece, de todas formas, que esquemas de inferencia de este tipo están faltos de clarificación y que sería necesario pasar de estudios teóricos a probar su funcionamiento en sistemas reales.

2.5.4.- Inferencia con proposiciones difusas.

La íntima conexión existente entre las medidas de posibilidad propuestas por Dubois y Prade, (D-P en lo sucesivo), y el concepto de distribución de posibilidad de Zadeh permite aplicar ambas a la obtención de la posibilidad y necesidad de sucesos difusos.

Siguiendo con lo expuesto al comparar la medida de posibilidad de D-P con la distribución de posibilidad de Zadeh, y con un mínimo cambio de notación, se tenía que la proposición "X es B" induce a una distribución de posibilidad $\Pi_X = B$ y para subconjuntos no difusos A se obtenía

$$\mathrm{Pos}\ (X \in A \mid X \text{ es } B) = \sup \left\{ \pi_X(u) \mid u \in A \right\} = \sup \left\{ \mu_B(u) \mid u \in A \right\}$$

Si ahora A es un subconjunto difuso de referencial U, Zadeh, (ZADE-78), propone la siguiente formulación:

$$\mathrm{Pos}\ (X \in A \mid X \text{ es } B) = \sup \left\{ \min\ (\mu_B(u), \mu_A(u)) \mid u \in U \right\}$$

dado que ahora la pertenencia de un elemento u del referencial al subconjunto A no es del conjunto $\{0,1\}$.

Desde el enfoque de D-P, si se conocen los valores de posibilidad $\pi(u)$ sobre los elementos del referencial U

$$\Pi(A) = \sup \left\{ \pi(u) \mid u \in A \right\}$$

y si ahora A es difuso:

$$\Pi(A) = \sup \left\{ \min\ (\mu_A(u), \pi(u)) \mid u \in U \right\}$$

que es exactamente la traslación de la propuesta de Zadeh.

Es fácil comprobar que los axiomas fundamentales de las medidas de posibilidad continúan siendo ciertas para conjuntos difusos. En particular

$$\Pi(A \cup B) = \max\ (\Pi(A), \Pi(B))$$
$$A \subseteq B \Longrightarrow \Pi(A) \leq \Pi(B)$$

y, por consiguiente, Π sigue siendo medida de posibilidad para conjuntos difusos.

Mediante la relación $N(A) = 1 - \Pi(\bar{A})$, que esencialmente es la definición de la medida de necesidad como dual de la medida de posibilidad, se puede ahora obtener la necesidad de un hecho difuso:

$$
\begin{aligned}
N(A) &= 1 - \Pi(\bar{A}) = 1 - \sup \left\{ \min \left(1 - \mu_A(u),\ \pi(u)\right) \mid u \in U \right\} = \\
&= 1 - \sup \left\{ \min \left(1 - \mu_A(u),\ 1 - (1 - \pi(u))\right) \mid u \in U \right\} = \\
&= 1 - \sup \left\{ 1 - \max \left(\mu_A(u),\ 1 - \pi(u)\right) \mid u \in U \right\} = \\
&= 1 - \left(1 - \inf \left\{ \max \left(\mu_A(u),\ 1 - \pi(u)\right) \mid u \in U \right\}\right) = \\
&= \inf \left\{ \max \left(\mu_A(u),\ 1 - \pi(u)\right) \mid u \in U \right\}
\end{aligned}
$$

También es posible comprobar que $N(A \cap B) = \min (N(A), N(B))$ y que $A \subseteq B \Longrightarrow N(A) \leq N(B)$, por lo que N sigue siendo una medida de necesidad.

Sin embargo, dado que para conjuntos difusos $A \cap \bar{A}$ no es necesariamente el conjunto vacío y que $A \cup \bar{A}$ no tiene por qué ser referencial U, hay algunas propiedades importantes que dejan de verificarse cuando se trata con conjuntos difusos. En particular

$$\max (\Pi(A), \Pi(\bar{A})) \leqslant 1 \quad \text{y} \quad \min (N(A), N(\bar{A})) \geqslant 0$$

En consecuencia, tampoco se verifican las importantes propiedades

$$
\begin{aligned}
N(A) > 0 &\Longrightarrow \Pi(A) = 1 \\
\Pi(A) < 1 &\Longrightarrow N(A) = 0
\end{aligned}
$$

por lo que, si se desea mantener esta relación en el par $(N(A), \Pi(A))$, deberá hacerse un proceso de normalización al calcular la necesidad y posibilidad de un suceso difuso.

Estos hechos permiten que sea posible tratar con reglas inciertas cuando el antecedente de la misma es una proposición difusa. Considérese la regla

$$p \equiv \text{"X es A"} \longrightarrow q$$

y supóngase que lo que se conoce es "X es A'", ya sea por deducción mediante el modus ponens generalizado de otras reglas que concluyen en "X es A", o bien porque el usuario del sistema ha introducido la información A' sobre el valor de X.

"X es A'" induce la distribución de posibilidad Π_X = A' y entonces se obtienen la posibilidad y necesidad del antecedente p de la regla:

$$\begin{aligned} \Pi(p) &= \text{Pos } (X \text{ es } A \mid X \text{ es } A') = \sup \left\{ \min \left(\mu_A(u), \mu_{A'}(u) \right) \mid u \in U \right\} \\ N(p) &= \inf \left\{ \max \left(\mu_A(u), 1 - \mu_{A'}(u) \right) \mid u \in U \right\} \end{aligned} \tag{22}$$

y se puede ya propagar la incertidumbre desde el antecedente a la conclusión, teniendo en cuenta la incertidumbre de la propia regla, mediante la utilización de los esquemas de inferencia que se proponían para reglas con antecedentes no difusos.

2.5.5.- Combinación de evidencias.

El procedimiento que proponen D-P para combinar la incertidumbre en el caso de un conjunto de reglas $p_i \rightarrow q$ con la misma conclusión, cuando para cada una de ellas se ha deducido el par $(N_i(q), \Pi_i(q))$ que evalúa la incertidumbre de la conclusión q es el siguiente:

Se considera la información deducida de cada regla R_i como un subconjunto difuso del referencial $\{q, \neg q\}$ de tal forma que

$$\mu_{R_i}(q) = \Pi_i(q) \quad \text{y}$$
$$\mu_{R_i}(\neg q) = \Pi_i(\neg q) = 1 - N_i(q)$$

de acuerdo con el sentido de las medidas de posibilidad y necesidad.

La existencia conjunta de distintas reglas puede entonces considerarse como distintas informaciones relativas al hecho q y, en consecuencia, podrán combinarse como intersección de subconjuntos difusos, obteniéndose, si se considera $R = \bigcap_i R_i$, el conjunto de todas las informaciones:

$$\mu_R(q) = \Pi(q) = \min_i (\Pi_i(q))$$

$$\mu_R(\neg q) = 1 - N(q) = \min_i (1 - N_i(q)) = 1 - \max_i (N_i(q))$$

llegándose a:

$$\Pi(q) = \min_i (\Pi_i(q))$$

$$N(q) = \max_i (N_i(q))$$

Sin embargo, se plantea el problema de que el par obtenido $(N(q), \Pi(q))$ no estará normalizado en el sentido $N(q) = 0$ ó $\Pi(q) = 1$, e incluso podrá suceder que tampoco se verifique $\Pi(q) \geqslant N(q)$. El proceso de normalización debe de hacerse por tanto normalizando el conjunto difuso R, con lo que se obtendrá, si $k = \max(\min_i \Pi_i(q), 1 - \max_i (N_i(q))$:

$$\Pi(q) = \frac{\min_i (\Pi_i(q))}{k}$$

$$N(q) = 1 - \frac{1 - \max_i (N_i(q))}{k}$$

Todo este conjunto de mecanismos que proponen D-P para el razonamiento con hechos y reglas inciertas, han sido implementados en un motor de inferencia que figura como anexo en su trabajo (DUBO-85b). Puesto que también se consideran hechos difusos, se hace necesaria la utilización del modus ponens generalizado como mecanismo de inferencia. Se trata, por tanto, de un motor de inferencia capaz de gestionar una base de conocimientos al mismo tiempo incierta y difusa.

CAPITULO 3 :

PLANTEAMIENTO DEL PROBLEMA. HIPOTESIS DE TRABAJO

3.1.- INTRODUCCION.

Como se expuso en el capítulo primero de este trabajo, el objetivo es la obtención de un modelo de resolución de incertidumbre que supere algunos de los problemas que aún permanecen pendientes de solución en este campo, y que ninguno de los sistemas existentes resuelven de forma satisfactoria.

En el modelo que se propone, se construye una nueva medida para evaluar la incertidumbre de los hechos que intervienen en la base de conocimiento de un sistema basado en reglas. Este modelo permite un tratamiento simultáneo de la incertidumbre de hechos y reglas, y de la imprecisión de los parámetros o atributos del conocimiento objeto de representación y tratamiento por medio de un Sistema Experto.

En consecuencia, y dada la variedad de los problemas que se plantean al construir un modelo de tratamiento y propagación de incertidumbre, y que fueron puestos de manifiesto al estudiar los distintos sistemas existentes en el capítulo segundo, será necesario abordar los siguientes aspectos:

- Parámetros empleados para representar la incertidumbre de hechos y reglas, así como su fundamentación teórica.

- Forma de representación del conocimiento impreciso.

- Incertidumbre relativa a los hechos imprecisos.

- Mecanismos de inferencia, que incluirán los tres problemas habituales:

  .) Propagación de incertidumbre al encadenar distintas reglas en el proceso de razonamiento.

  .) Combinación de incertidumbre, cuando mediante distintas reglas se deduce la misma conclusión.

.) Obtención de incertidumbre de proposiciones formadas a partir de la conjunción o disyunción de otras proposiciones.

- Procedimientos para el tratamiento conjunto de la imprecisión y la incertidumbre del conocimiento.

Para cada uno de estos problemas se propondrá una solución que deberá ser convenientemente justificada, y que será contrastada con las soluciones aportadas por modelos de resolución de incertidumbre existentes cuya fundamentación permita la comparación entre ambas soluciones.

El punto de partida del modelo que se propone es la Teoría de la Posibilidad, desarrollada por Dubois y Prade. El origen y propiedades de las medidas de posibilidad y necesidad que éstos emplean como instrumento para evaluar numéricamente la incertidumbre del conocimiento vertido en un sistema experto basado en reglas, han sido ampliamente comentados en el capítulo anterior. Como ejemplo de la aplicabilidad de esta teoría, presentan un motor de inferencia capaz de tratar con hechos y reglas que son a la vez inciertos y difusos.

Se trata de un motor con encadenamiento hacia atrás que presenta la ventaja de su sencillez en el tratamiento de la incertidumbre y la imprecisión. A continuación se expone, sin entrar en detalles no esenciales, la fundamentación de este motor, en su doble vertiente de representación del conocimiento y mecanismos básicos de inferencia. Esta presentación se debe a un doble objetivo:

1) Va a permitir analizar el comportamiento del modelo de D-P en relación con otros métodos para razonar con incertidumbre.

2) Como consecuencia de este análisis, se establecerán las hipótesis de partida, con objeto de superar aquellos aspectos que no resulten satisfactorios.

Una versión modificada de este motor, con explicaciones de las funciones esenciales utilizadas, e implementada en GCLISP, figura con el nombre de MOTOR-1 en el apéndice de esta memoria.

3.2.- Motor de Inferencia de Dubois y Prade.

Para mostrar el funcionamiento de este motor de inferencia, D-P lo aplican a una base de conocimientos muy reducida, referida a un problema de selección de personal, en función de unas determinadas aptitudes o capacidades deseables en la persona a seleccionar. La base de reglas está constituída por 14 reglas, y el árbol de deducción resultante para determinar la adecuación o no de un candidato es el que se muestra en la figura 1. En dicho árbol se incluyen también los valores numéricos que se introducen al cargar la base de conocimientos para un determinado candidato.

En este árbol se distinguen los dos tipos diferentes de conocimiento, impreciso y difuso, con lo que es capaz de trabajar el motor para realizar inferencias:

A.- Hechos: son de dos tipos:

A.1.- Difusos: El atributo correspondiente a ese hecho, en cada una de las reglas que le contienen, toma un valor que es un subconjunto difuso de un determinado referencial. Por ejemplo:

(TEST-ING CANDIDATO BUENO)

BUENO es un posible resultado de la prueba sobre los conocimientos de inglés que posee un candidato. No se da un valor numérico, sino que se proporciona un valor difuso. BUENO será un subconjunto del referencial $U = [0,20]$, que viene definido por una función de pertenencia. Para todos los valores difusos que utiliza el MOTOR-1, esta función de pertenencia viene siempre de forma trapezoidal parametrizada con cuatro valores, (a,b,α,β), figura 2.

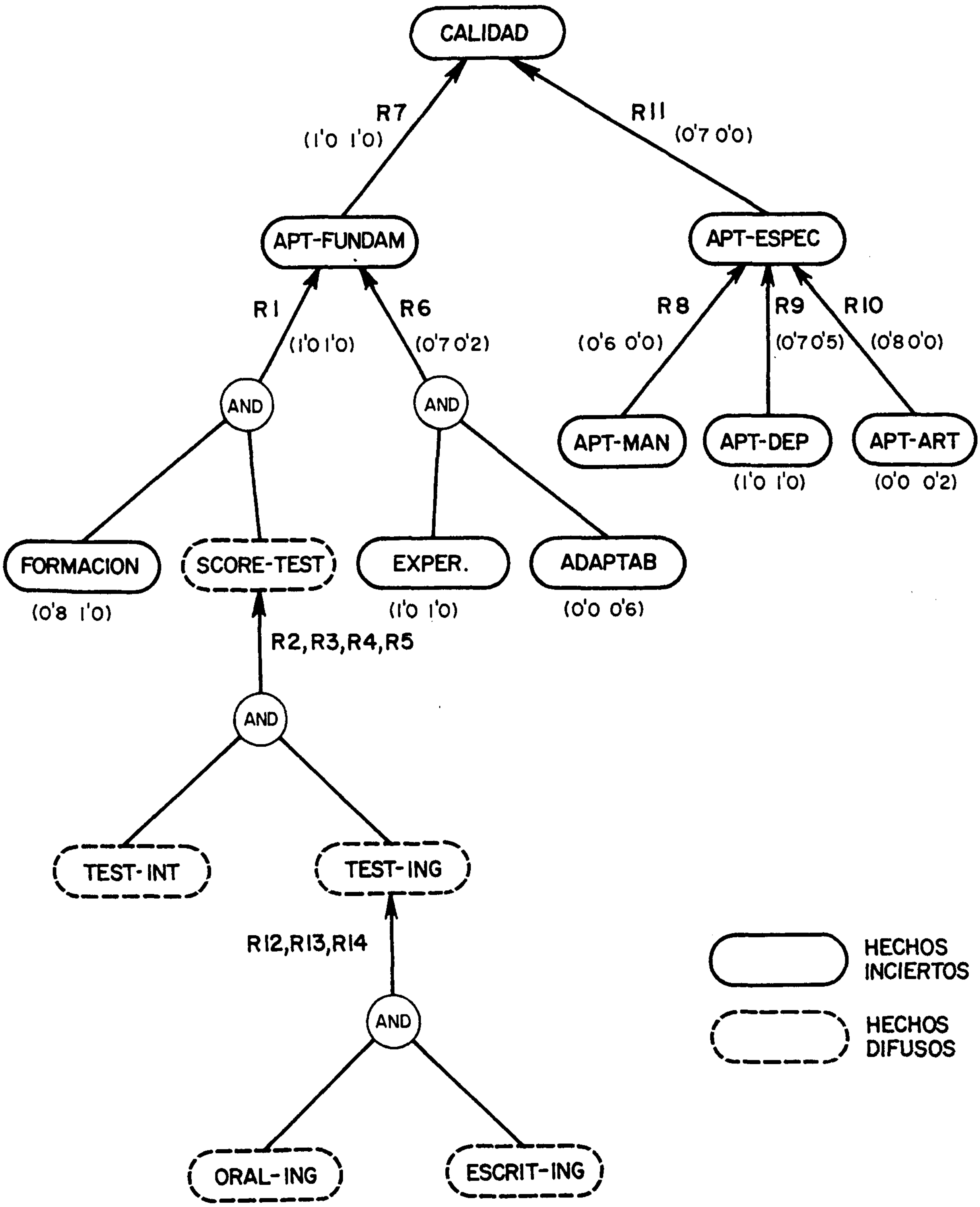


Figura 1.- Arbol de deducción en "selección de personal".

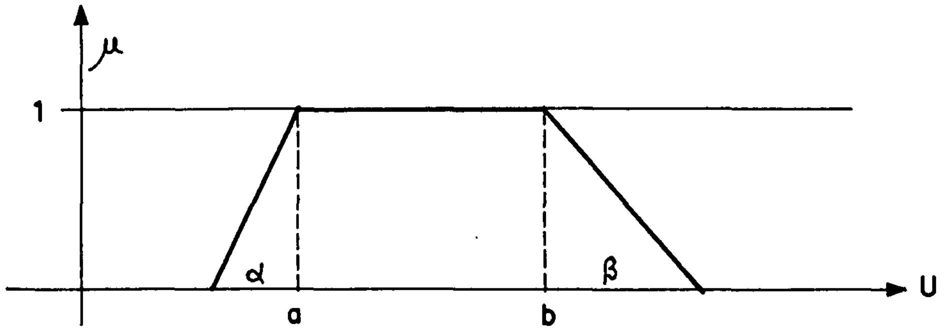


Figura 2.- Función de pertenencia parametrizada.

Así el subconjunto BUENO, corresponderá a la parametrización trapezoidal (15,16,2,1).

En otros casos, no se da un valor lingüístico sino los extremos del intervalo [a,b], como por ejemplo, 6A9, que corresponde al subconjunto (6,9,2,2).

Tanto los subconjuntos difusos que aparecen en las reglas como aquéllos valores difusos que se introducen al ejecutar el procedimiento para un determinado candidato han de ser predefinidos.

Los hechos difusos no figuran cuantificados con valores de incertidumbre, ya sean resultado de la inferencia por aplicación de reglas difusas o hechos introducidos en la base de hechos inicial.

A.2.- Inciertos: no toman un valor difuso. Su verdad o falsedad no se caracteriza por ningún valor difuso, sino que vienen afectados de unos valores de incertidumbre mediante un par (N, Π). Para aquellos hechos que no se deducen mediante aplicación de reglas, el usuario debe introducir estos dos números, aunque el motor permite también la no-respuesta, es decir, ignorancia sobre el hecho en cuestión.

B.- Reglas: también son de dos tipos:

B.1.- Inciertas: Son aquéllas cuya conclusión es un hecho incierto, pudiendo ser los hechos que forman el antecedente de la regla inciertos o difusos, e incluso combinación de ambos, caso de que exista una conjunción de premisas.

La incertidumbre de este tipo de reglas viene cuantificada con dos valores, grado de suficiencia y grado de necesidad.

B.2.- Difusas: Son reglas en las que los hechos que figuran tanto en antecedente como consecuente son difusos. Este tipo de reglas no son cuantificadas con valores de incertidumbre. En cambio, en la base de conocimientos existe todo un conjunto de reglas con los mismos antecedentes y consecuentes, diferenciándose en los valores difusos que toman los hechos o atributos que figuran en ellas. Así, en el ejemplo relativo a selección de personal, existen tres reglas de la forma:

(ORAL-ING X) y (ESCRT-INT Y) ⟶ (TEST-ING Z)

donde X, Y y Z son subconjuntos difusos. Más precisamente:

R12: (O-I 6A9) y (E-I ACEPTABLE) ⟶ (T-I PASABLE)
R13: (O-I ACEPTABLE) y (E-I ACEPTABLE) ⟶ (T-I BUENO)
R14: (O-I ACEPTABLE) y (E-I 6A9) ⟶ (T-I SATISFACTORIO)

Estas reglas son proporcionadas por el experto, que es el que fija los valores difusos de la conclusión en función de cuáles sean los valores difusos de los hechos que figuran en el antecedente. Es la multiplicidad de las reglas lo que permite no considerar valores de incertidumbre para ellas.

Los mecanismos de inferencia son totalmente diferentes según se trate de reglas difusas o inciertas.

Inferencia con reglas difusas.

Para cada una de las reglas con los mismos antecedentes y consecuentes, se deduce el valor difuso de la conclusión mediante aplicación del "modus ponens generalizado", a partir de los valores difusos reales de cada uno de los atributos que figuran en el antecedente y que, en general, serán diferentes de los valores que figuran en la regla. Estos valores reales se habrán obtenido en etapas anteriores del proceso de inferencia o bien, caso de tratarse de nodos terminales en el árbol de deducción, serán introducidos inicialmente por el usuario como Base de Hechos.

La expresión del "modus ponens generalizado" que este motor utiliza es la que se expuso en el apartado 2.4.3.3. Como consecuencia de las distintas propiedades que se verificaban y que se comentaron en aquella ocasión, los autores hacen aplicación de ellas y se obtienen cinco casos diferentes (funciones MPGP0 a MPGP4 del programa MOTOR-1).

Una primera limitación que es importante señalar es que estas funciones han sido obtenidas únicamente para el caso en que los subconjuntos difusos vengan representados por una función de pertenencia de tipo trapezoidal. Detalles de esta implementación pueden verse en (DUBO-85b) y (MART-84).

Una vez deducido, para cada una de las reglas, el valor difuso de la conclusión, el valor que resulta de la aplicación de todas ellas se obtiene como intersección, en el sentido de conjuntos difusos, de los valores difusos de cada una de ellas.

Inferencia con reglas inciertas.

- Caso 1: No hay premisas difusas en el antecedente.

  El mecanismo básico de inferencia está basado en el esquema de razonamiento que se expuso en el apartado 2.5.3., aunque con una sensible modificación, que conviene destacar:

  El esquema del que parten los autores

$$
\begin{array}{l}
N(p \longrightarrow q) \geqslant a \\
N(q \longrightarrow p) \geqslant a' \\
N(p) \geqslant b \\
\Pi(p) < b' \\
\hline
N(q) \geqslant \min\ (a,b) \\
\Pi(q) \leqslant \max\ (1 - a',b)
\end{array}
$$

estaba plenamente justificado, obteniéndose como combinación de otros esquemas de razonamiento, que a su vez, se deducían rigurosamente de las diferentes propiedades de las medidas de posibilidad y necesidad.

Sin embargo, D-P apuntan la posibilidad de sustituir los operadores mínimo y máximo por alguno de los siguientes:

$$
\left\{\begin{array}{l} \min\ (a,b) \\ \max\ (1-a',b') \end{array}\right\} \longleftrightarrow \left\{\begin{array}{l} a.b \\ 1-a'(1-b') \end{array}\right\} \longleftrightarrow \left\{\begin{array}{l} \max\ (0,\ a+b-1) \\ \min\ (1,\ 1-a'+b') \end{array}\right\}
$$

Estos dos últimos no se deducen de la Teoría de la Posibilidad, sino que, según los autores, son justificables desde el punto de vista de alguna lógica multivalente no especificada.

El motor de inferencia que construyen permite elegir entre estas alternativas. Por ocultas razones que se pondrán de manifiesto más adelante, en el ejemplo de aplicación que se está comentando, la opción que se toma es la segunda de las alternativas apuntadas, en lugar de la que se deduce razonadamente de las propiedades de las medidas de posibilidad y necesidad.

- Caso 2: Alguna o todas las premisas son difusas.

En esta situación, el consecuente de la regla es impreciso, la regla se encuentra cuantificada con valores de incertidumbre y quizás algunas de las premisas sean hechos no difusos para los que se han deducido el correspondiente par (N,Π) de incertidumbre. Para poder continuar la propagación de la incertidumbre se hace necesario obtener los valores de incertidumbre de los hechos difusos que figuran en el antecedente de la regla.

En la regla imprecisa en cuestión el hecho difuso viene acompañado con su correspondiente valor, que es un subconjunto de un cierto referencial. Su significado puede verse como una proposición difusa del tipo "X es A". Sin embargo, el resultado de la propagación de incertidumbre hasta llegar a ese hecho proporciona un valor difuso A' diferente, es decir, lo que se conoce es la proposición "X es A'".

El proceso que se sigue es sencillamente obtener la posibilidad y necesidad del hecho "X es A" cuando el conocimiento que se tiene de la realidad es "X es A'". Los valores de incertidumbre del hecho difuso se obtienen sencillamente por aplicación a los subconjuntos difusos A y A', del mismo referencial, de las relaciones

$$\Pi(X \text{ es } A \mid X \text{ es } A')$$

$$N(X \text{ es } A \mid X \text{ es } A')$$

que figuran en el apartado 2.5.4., "Inferencia con proposiciones difusas".

3.3.- Análisis del Motor de Inferencia.

Una de las consecuencias más llamativas de este motor de inferencia que merece destacarse, es la versatilidad en cuanto a la propagación de la incertidumbre, puesto que se deja una cierta libertad para elegir las funciones de propagación mediante la asignación a unas variables NORMA-1 y NORMA-2 de los operadores convenientes, que se corresponden a las distintas alternativas antes comentadas. Esta elección se realiza en la inicialización del proceso de ejecución, al mismo tiempo que se introduce la base de conocimientos. Parece el momento adecuado para tomar esta decisión, que afectará de manera importante a los resultados obtenidos. La utilización de uno u otro operador para la función de propagación estará íntimamente ligada al tipo de problema que se desee tratar con este motor de inferencia.

También destaca prontamente un hecho que parece negativo. Supóngase que se utiliza este motor de inferencia como parte de un determinado sistema experto, para el que se hayan construído una base de conocimientos que refleje el problema a tratar y la correspondiente interfaz de comunicación con el usuario. Cuanto éste debe introducir el conocimiento que posee inicialmente o cuando es preguntado por el sistema sobre algún hecho que no se puede deducir a partir de las reglas existentes en la base de conocimientos, la respuesta del usuario será de dos tipos diferentes:

- Un subconjunto difuso perfectamente determinado en el caso de hechos difusos.

- La evaluación de la incertidumbre sobre el hecho en cuestión, si éste no es difuso. Deberá proporcionar un doble valor numérico para cuantificar su credibilidad, con el significado de posibilidad y necesidad de ese hecho.

Es en este último punto donde reside uno de los principales inconvenientes de este modelo. Estos dos valores de posibilidad y necesidad serán normalmente de difícil comprensión para el usuario. Este es un profesional de aquella rama del conocimiento para la que el sistema ha sido diseñado, con el propósito de resolver problemas concretos. El usuario no conocerá tan siquiera el significado preciso de los conceptos de posibilidad y necesidad.

Uno de los principios básicos para el éxito comercial de un sistema experto es precisamente que sea de fácil utilización por el usuario y, en el aspecto concreto que ahora se plantea, que exija unas mínimas explicaciones del sistema al usuario sobre cómo debe valorar la confianza o credibilidad en el hecho sobre el que se le pide información.

Este tipo de problemas se refuerza si el aspecto que se considera es el de la interpretación de los resultados, pues las diferentes hipótesis a establecer también vendrán afectadas por una incertidumbre cuantificada con el par (posibilidad, necesidad).

Se constituye, en consecuencia, como una exigencia prioritaria la modificación del modelo de Dubois y Prade en el sentido de que la incertidumbre de los hechos se cuantifique con un único valor numérico, de significado claro. En definitiva, se trata de aumentar la sencillez para el usuario.

La existencia de un doble valor numérico en la cuantificación de las reglas inciertas no plantea los problemas que se originan con la incertidumbre de los hechos. En esta situación es el o los expertos con los que se diseña la base de conocimientos del sistema, el que debe proporcionar la valoración de la incertidumbre de cada regla. Sí será necesario dedicar una sesión de trabajo a informar de forma clara al experto del significado que tiene cada uno de los coeficientes $\underline{s}$ y $\underline{n}$, así como de los mecanismos de propagación de incertidumbre. Esto será suficiente para que el experto se sienta motivado a proporcionar de forma correcta los valores de incertidumbre de cada una de las reglas inciertas que formarán la base de conocimientos. En el modelo de D-P el significado de estos valores es:

.) $\underline{s}$ = grado de suficiencia: cuantifica la mayor o menor credibilidad en la conclusión $\underline{q}$ de una regla "si $\underline{p}$ entonces $\underline{q}$", cuando $\underline{p}$ es verdadero

.) $\underline{n}$ = grado de necesidad: mide la fuerza con que la ausencia del antecedente $\underline{p}$ apoya la verdad de la conclusión $\underline{q}$.

Conceptos análogos y con los mismos nombres aparecen en el modelo PROSPECTOR y, sin embargo, su cuantificación es muy diferente. En el modelo de D-P son números comprendidos entre 0 y 1, mientras que en PROSPECTOR estos valores podían ser mayores que la unidad, y de hecho, los valores habituales de LS, grado de suficiencia, son muy elevados.

La cuantificación que debe hacerse de estos parámetros está íntimamente ligada a la forma en que se realiza la propagación de incertidumbre, que en PROSPECTOR era una actualización de disparidades:

$$O(q|p) = LS \,.\, O(p) \qquad O(q|\neg p) = LN \,.\, O(p)$$

siendo LS y LN los grados de suficiencia y de necesidad.

En el modelo que se propone, se mantendrán estos dos valores s y n, con el sentido de D-P, así como los mecanismos de propagación de incertidumbre. El trabajar con dos valores para la incertidumbre de las reglas presenta además ventajas adicionales, pues equivale a utilizar simultáneamente dos reglas del tipo "si p entonces q, s" y "si no p entonces q, n". Está en la línea de Heckerman, (HECK-86), que se comentó al exponer el procedimiento de propagación de incertidumbre en el modelo MYCIN. En consecuencia, el considerar los dos valores s y n supone en realidad una representación más completa del conocimiento que se desea verter en la base de conocimientos del sistema.

Los que se hará preciso será una adaptación de los distintos mecanismos de inferencia utilizados por D-P puesto que el objetivo es valorar la incertidumbre de antecedentes y consecuentes de reglas con un sólo valor numérico, mientras que todas las funciones de D-P toman como argumentos el par $(N(p), \Pi(p))$ del antecedente de una regla, caso de la propagación, o pares $(N_i(q), \Pi_i(q))$, caso de la combinación de distintas reglas.

CAPITULO 4 :

PROPUESTA DE RESOLUCION DE INCERTIDUMBRE

4.1.- OBTENCION DE UN NUEVO PARAMETRO PARA LA CUANTIFICACION DE LA INCERTIDUMBRE DE HECHOS.

Como ya se expuso en el apartado anterior, uno de los objetivos es modificar el modelo de Dubois y Prade en el sentido de expresar la incertidumbre de los hechos con un único valor numérico, que de alguna forma englobe el par necesidad y posibilidad, $(N(A), \Pi(A))$ de un suceso A. Las características y propiedades de estas dos medidas se expusieron en el correspondiente apartado 2.5.

Observando algunas de estas propiedades, que serán de interés para la definición de una nueva medida a partir de las dos medidas de necesidad y posibilidad, se destacan las siguientes:

.) $N(A),\ \Pi(A) \in [0,1]$
.) $N(A) \leq \Pi(A)$
.) $N(A) = 1 - \Pi(\bar{A})$
.) $\Pi(A) = 1 - N(\bar{A})$

De acuerdo con estas dos últimas propiedades, el conocimiento sobre el hecho A caracterizado de forma numérica por el par $(N(A), \Pi(A))$ puede ser evaluado de forma equivalente por el par $(\Pi(\bar{A}), \Pi(A))$, que se podría interpretar como confianza en A y en $\bar{A}$, o lo que es lo mismo, con todos los matices ya comentados al tratar de los sistemas existentes, como falsedad y verdad relativas del hecho A.

Pero, según otra de las propiedades más sencillas de las medidas de necesidad y posibilidad, se tiene:

$$N(A) > 0 \Longrightarrow \Pi(A) = 1$$
$$\Pi(A) < 1 \Longrightarrow N(A) = 0$$

o expresado en otros términos:

$$\max\ (1 - N(A), \Pi(A)) = 1$$

que puesto de forma más conveniente para los propósitos que se pretenden, proporciona la expresión:

$$\max(\Pi(\bar{A}), \Pi(A)) = 1$$

Es importante penetrar algo más profundamente en el significado de esta expresión, puesto que va a permitir la aplicación de la lógica difusa desarrollada por Gaines, (GAIN-76), a la obtención de una nueva medida íntimamente relacionada con las medidas de posibilidad y necesidad. Si se consideran los valores $\Pi(A)$ y $\Pi(\bar{A})$ como la verdad y falsedad del hecho A, se recupera el tratamiento que hace Gaines de los valores de verdad difusos. Es decir, si F es el conjunto "falso" y V es el conjunto "verdadero", un hecho puede caracterizarse por un doble valor $(\mu_F(A), \mu_V(A))$, interpretados como el grado de pertenencia del suceso A a los conjuntos "falso" y "verdadero". De alguna forma se está considerando una lógica multivalente, con la sensible diferencia de que para cada suceso o proposición no se considera un único valor de verdad, sino un par de valores.

De acuerdo con este tratamiento de los valores de verdad, se tendrá:

.) $\mu_F(p) = 1$, $\mu_V(p) = 0$ significa que $\underline{p}$ es la proposición falsa, pues $\underline{p}$ pertenece al conjunto "falso" con grado 1, que interpretado según la definición clásica de conjunto difuso, quiere decir que hay total seguridad en la pertenencia de la proposición $\underline{p}$ al conjunto "falso" y, al mismo tiempo, hay total certeza de que $\underline{p}$ no pertenece al conjunto "verdadero".

.) $\mu_F(p) = 0$, $\mu_V(p) = 1$, de forma recíproca, significará que la proposición $\underline{p}$ es verdadera con total certeza.

.) Problemas de interpretación con este tipo de valoración de la verdad de una proposición se plantean para el caso $\mu_F(p) = 1$ y $\mu_V(p) > 0$: $\underline{p}$ es falsa con total certeza pero, al mismo tiempo hay un cierto grado de pertenencia al conjunto "verdadero".

.) El caso extremo de imposible interpretación sería cuando $\mu_F(p) = 1$ y $\mu_V(p)=1$. Evidentemente, y pasando por alto la contradicción que supone esta doble valoración, no se sabe nada acerca de la verdad o falsedad de la proposición p. En el caso anterior, aún con la misma contradicción, se puede afirmar sobre p que es más falsa que verdadera.

Gaines, al considerar esta interpretación, no se plantea estas contradicciones en casos extremos. El interés de su desarrollo radica en el tratamiento que hace de los conectivos lógicos y los resultados a los que llega.

Desde este momento puede observarse el paralelismo entre estos dos valores de Gaines y los que se consideraron anteriormente con medidas de posibilidad, $(\Pi(\bar{A}), \Pi(A))$ donde se dan exactamente las mismas contradicciones si $\Pi(A)$ se interpreta como el valor de verdad de A y $\Pi(\bar{A})$ como el grado de falsedad de A.

Partiendo de una forma modificada del principio de extensión de Zadeh, Gaines deduce el par de valores de verdad $(\mu_F(q), \mu_V(q))$ de una proposición compuesta q, a partir de los valores de verdad de las proposiciones componentes y de la interpretación clásica de los distintos conectivos en la lógica bivalente.

El principio de extensión de Zadeh, en su versión más amplia y conocida, afirma que, si se tiene una función

$$f: U_1 \times U_2 \times \dots \times U_n \longrightarrow V$$

y si $A_1, A_2, \dots, A_n$ son subconjuntos difusos de $U_1, U_2, \dots, U_n$, entonces $f(A_1 \times A_2 \times \dots \times A_n)$ es un subconjunto difuso de V. Con la suposición adicional de que las variables ligadas a cada uno de los subconjuntos A_i son no-interactivas, se obtiene:

$$\mu_{f(A_1 x \dots x A_n)}(v) = \sup \left\{ \min (\mu_{A_1}(u_1), \dots, \mu_{A_n}(u_n)) \mid f(u_1, \dots, u_n) = v \right\}$$

Gaines hace una interpretación de este principio de la siguiente forma. Si f: $U_1 \times U_2 \times \ldots U_n \longrightarrow V$ es una función y los valores que toman cada una de las variables es difuso, se tendrá:

$$\mu_V(v) = \sup \left\{ \min \left(\mu_{U_1}(u_1), \ldots, \mu_{U_n}(u_n)\right) \mid f(u_1, \ldots, u_n) = v \right\}$$

es decir, se obtiene el valor difuso de la imagen $f(u_1, \ldots, u_n)$ a partir de los valores difusos de cada una de las n variables.

En particular, Gaines considera la función implicación, $p \longrightarrow q$ de la lógica bivalente, que permite obtener el valor de verdad de la proposición $p \longrightarrow q$, $v(p \longrightarrow q)$ a partir de los valores de verdad p y q, v(p) y v(q):

$$\begin{array}{lcccl} p \longrightarrow q: & \{F,V\} & \times & \{F,V\} & \longrightarrow \{F,V\} \\ & F & , & F & \longmapsto V \\ & F & , & V & \longmapsto V \\ & V & , & F & \longmapsto F \\ & V & , & V & \longmapsto V \end{array}$$

Si ahora los valores de verdad de las proposiciones p y q no son tan sólo F y V, sino que para cada proposición p se considera el par $(\mu_F(p), \mu_V(p))$, se obtiene por el principio de extensión en su segunda forma:

$$\mu_F(p \rightarrow q) = \min\left(\mu_V(p), \mu_F(q)\right) \qquad (1)$$

$$\mu_V(p \rightarrow q) = \max\left(\min(\mu_F(p), \mu_F(q)), \min(\mu_F(p), \mu_V(q)), \min(\mu_V(p), \mu_V(q))\right)$$

De forma análoga podría obtenerse el par $(\mu_F(q), \mu_V(q))$ para el resto de los conectivos lógicos o, de forma alternativa:

•) $\neg p$ es lógicamente equivalente a la proposición $p \rightarrow f$, donde f es la proposición siempre falsa. El par $(\mu_F(\neg p), \mu_V(\neg p))$ se puede obtener a partir de las expresiones (1), considerando $\mu_F(f) = 1$, $\mu_V(f) = 0$.

$$\mu_F(\neg p) = \mu_V(p) \qquad (2)$$
$$\mu_V(\neg p) = \mu_F(p)$$

.) $p \vee q$ se puede calcular a partir de la equivalencia lógica

$$p \vee q \equiv \neg p \longrightarrow q$$

obteniéndose

$$\mu_F(p \vee q) = \min\ (\mu_F(p), \mu_F(q)) \qquad (3)$$
$$\mu_V(p \vee q) = \max(\min(\mu_V(p), \mu_F(q)),\ \min(\mu_V(p), \mu_V(q),\ \min(\mu_F(p), \mu_V(q)))$$

.) Finalmente $p \wedge q$ se obtendría de la equivalencia:

$$p \wedge q \equiv \neg(\neg p \vee \neg q)$$

Sin embargo, Gaines se centra en el caso particular de que alguno de los dos valores $\mu_F(p)$ o $\mu_V(p)$ sea la unidad. Aunque no diga nada al respecto, podría venir motivado por necesidades de comparación de los valores de verdad de dos proposiciones, como, por ejemplo:

$$p \equiv (\mu_F(p), \mu_V(p)) = (0'4 \quad 0'7) \quad \text{y}$$
$$q \equiv (\mu_F(q), \mu_V(q)) = (0'3 \quad 0'6)$$

Es sencillo transformar pares de este tipo en pares con alguno de los dos valores igual a 1, normalizando por división entre el máximo de ambos.

Nuevamente se presenta un claro paralelismo entre este caso estudiado de forma particular por Gaines y lo que sucede al considerar medidas de necesidad y posibilidad, puesto que para el par $(\Pi(\overline{A}), \Pi(A))$ se verifica que $\max\ (\Pi(\overline{A}), \Pi(A)) = 1$, es decir, alguno de los dos valores debe ser la unidad.

En el supuesto de proposiciones de este tipo, es posible un tratamiento con un único valor de verdad de la siguiente forma. Considérese la función:

$$g: [0,1] \times [0,1] \longrightarrow [0,1]$$

$$\mu_F(p), \mu_V(p) \longmapsto v(p) = \frac{1 - \mu_F(p) + \mu_V(p)}{2}$$

donde se verifica la condición max $(\mu_F(p), \mu_V(p)) = 1$.

Esta función es una biyección y por tanto permitirá trabajar con el único valor v(p) en vez de hacerlo con el par normalizado $(\mu_F(p), \mu_V(p))$.

Se demuestra a continuación la biyectividad de g, puesto que se jugará un papel importante en la medida que se definirá a partir de N(A) y Π(A), así como su relación con estas últimas. En efecto:

.) <u>g es aplicación</u>, $g(p) \in [0,1]$

$$\begin{aligned} &- \text{Si } \mu_F(p) = 1, \quad v(p) = \frac{\mu_V(p)}{2} \in [0,1/2] \\ &- \text{Si } \mu_V(p) = 1, \quad v(p) = 1 - \frac{\mu_F(p)}{2} \in [1/2, 1] \end{aligned} \qquad (4)$$

Además se observa lo siguiente:

- $v(p) = 1$ si y sólo si $(\mu_F(p), \mu_V(p)) = (0,1)$, es decir, cuando p es verdadera.

- $v(p) = 0$ si y sólo si $(\mu_F(p), \mu_V(p)) = (1,0)$, es decir, p es falsa.

- Cuando $(\mu_F(p), \mu_V(p)) = (1,1)$ se obtiene v(p) = 0'5.

Por tanto, los nuevos valores de verdad así obtenidos oscilan entre el valor 0, p falsa, hasta v(p) = 1, p verdadera, pasando por v(p) = 0'5 que significaría desconocimiento o ignorancia acerca de la verdad de p.

.) g es inyectiva:

$$g(p) = g(q) \longrightarrow \frac{1 - \mu_F(p) + \mu_V(p)}{2} = \frac{1 - \mu_F(q) + \mu_V(q)}{2}$$

- $\mu_F(p) = 1$ y $\mu_F(q) = 1$: $\mu_V(p) = \mu_V(q)$

- $\mu_F(p) = 1$ y $\mu_V(q) = 1$: $\mu_V(p) - 1 = 1 - \mu_F(q) \longrightarrow$
$\longrightarrow \mu_V(p) + \mu_F(q) = 2 \longrightarrow \mu_V(p) = \mu_F(q) = 1$

- y análogamente para los otros dos casos restantes

.) g es sobreyectiva:

Si $r \in [0,1]$, debemos encontrar un par $(\mu_F(p), \mu_V(p))$ con $\mu_F(p) = 1$ o $\mu_V(p) = 1$ de forma que $(1 - \mu_F(p) + \mu_V(p)) / 2 = r$. Si se observan las relaciones (4) se deduce:

$r \in [0,1/2]$ es imagen del par $(1,2r)$
$r \in [1/2,1]$ es imagen del par $(2(1-r),1)$

En consecuencia, es equivalente la utilización para valores de verdad del par $(\mu_F(p), \mu_V(p))$ que del valor $v(p)$, puesto que de uno de ellos se puede recuperar el otro.

Gaines, a continuación, obtiene los valores de verdad $v(q)$ para proposiciones compuestas. Por ejemplo, para los dos casos $v(\neg p)$ y $v(p \longrightarrow q)$ se obtiene:

.) $v(\neg p)$: De las relaciones (2) se deduce

$$v(\neg p) = \frac{1 - \mu_F(\neg p) + \mu_V(\neg p)}{2} = \frac{1 - \mu_V(p) + \mu_F(p)}{2} =$$

$$= \frac{1 - \mu_F(p) + \mu_V(p)}{2} = 1 - v(p)$$

.) $\underline{v(p \longrightarrow q)}$:

- Caso 1: $\mu_F(p) = 1, \mu_F(q) = 1$

Se tendrá $v(p) = \mu_V(p)/2, \quad v(q) = \mu_V(q)/2$ y

$$\mu_V(p) = 2 \cdot v(p), \mu_V(q) = 2 \cdot v(q)$$

De las relaciones (1):

$$\mu_F(p \longrightarrow q) = \min (\mu_V(p), 1) = \mu_V(p)$$

$$\mu_V(p \longrightarrow q) = \max (\min (1,1), \ldots) = 1$$

y por tanto

$$v(p \longrightarrow q) = \frac{1 - \mu_V(p) + 1}{2} = 1 - \frac{\mu_V(p)}{2} = 1 - v(p)$$

- Caso 2: $\mu_F(p) = 1, \mu_V(q) = 1$.

De la definición de la función $\underline{g}$:

$$v(p) = \mu_V(p)/2, \quad v(q) = (2 - \mu_F(q))/2 \quad \text{y}$$

$$\mu_V(p) = 2 \cdot v(p), \mu_F(q) = 2 (1 - v(q))$$

$$\mu_F(p \longrightarrow q) = \min (\mu_V(p), \mu_F(q))$$

$$\mu_V(p \longrightarrow q) = \max (\min (1, \mu_F(q)), \min (1,1), \ldots) = 1$$

y $v(p \longrightarrow q) = (1 - \min (\mu_V(p), \mu_F(q)) + 1) /2 =$

$$= 1 - \min (2 \cdot v(p), 2(1 - v(q)) /2 =$$

$$= 1 - \min (v(p), 1 - v(q)) = \max (1 - v(p), v(q))$$

- Caso 3: $\mu_V(p) = 1, \mu_F(q) = 1$

De forma análoga se obtiene $v(p \rightarrow q) = \max (1 - v(p), v(q))$

- Caso 4: $\mu_V(p) = 1, \mu_V(q) = 1$

Entonces $v(p) = 1 - \mu_F(p)/2, \quad v(q) = 1 - \mu_F(q)/2$ y

$$\mu_F(p) = 2 (1 - v(p)), \mu_F(q) = 2(1 - v(q))$$

$\mu_F(p \rightarrow q) = \min (1, \mu_F(q)) = \mu_F(q)$

$\mu_V(p \rightarrow q) = \max (\min (\mu_F(p), \mu_F(q)), \min (\mu_F(p), 1), \min (1,1)) = 1$

con lo cual $v(p \rightarrow q) = \dfrac{1 - 2 (1 - v(q)) + 1}{2} = v(q)$

pero como $v(p), v(q) \in [1/2,1]$, $1 - v(p) \in [0,1/2]$

se tiene $v(p \rightarrow q) = \max (1 - v(p), v(q))$

Por tanto, en todos los casos posibles se obtiene

$$v(p \rightarrow q) = \max (1 - v(p), v(q))$$

Con procedimientos parecidos, y siempre haciendo uso de la biyección entre $(\mu_F(p), \mu_V(p))$ y $v(p)$, Gaines llega a las relaciones:

$$\begin{aligned} v(\neg p) &= 1 - v(p) \\ v(p \wedge q) &= \min (v(p), v(q)) \qquad (5) \\ v(p \vee q) &= \max (v(p), v(q) \\ v(p \rightarrow q) &= \max (1 - v(p), v(q)) \end{aligned}$$

En consecuencia, con estos nuevos valores de verdad obtenidos mediante la función $\underline{g}$, se obtiene una lógica multivalente que fue desarrollada inicialmente por Dienes en 1949.

Dada la analogía existente entre los valores de verdad utilizados por Gaines y los pares $(\Pi(\bar{A}), \Pi(A))$, así como la interpretación dada a este último, parece que se debe intentar la misma transformación $\underline{g}$. Así se obtiene:

$$g((\Pi(\bar{A}), \Pi(A))) = \frac{1 - \Pi(\bar{A}) + \Pi(A)}{2} = \frac{N(A) + \Pi(A)}{2}$$

A partir de las medidas de necesidad y posibilidad $N(A)$ y $\Pi(A)$ se obtiene un nuevo valor que será el que se considere para evaluar la verdad del suceso A, o de una proposición $\underline{p}$

$$v(p) = \frac{N(p) + \Pi(p)}{2}$$

Todo el desarrollo hecho por Gaines es entonces aplicable a estas tres medidas:

- A partir de $v(p)$, se pueden recuperar los valores de las medidas de necesidad y posibilidad:

$0'5 \leq v(p) \leq 1$: $\Pi(p) = 1$

$\Pi(\bar{p}) = 2\ (1 - v(p))$; $N(p) = 1 - 2 + 2v(p) = 2v(p) - 1$

$0 \leq v(p) \leq 0'5$: $\Pi(\bar{p}) = 1$; $N(p) = 0$

$\Pi(p) = 2v(p)$

En resumen:

$$0'5 \leq v(p) \leq 1 \longrightarrow (N(p), \Pi(p)) = (2\ v(p) - 1,\ 1)$$

$$0 \leq v(p) \leq 0'5 \longrightarrow (N(p), \Pi(p)) = (0,\ 2\ v(p))$$

- La adecuación de esta nueva medida $v(p)$ se puede apreciar en los siguientes casos extremos:

$(N(p), \Pi(p))$	$(\Pi(\bar{p}), \Pi(p))$	v(p)	significado
(0,1)	(1,1)	0'5	ignorancia
(1,1)	(0,1)	1	certeza total
(0,0)	(1,0)	0	falsedad total

Los valores v(p) sirven, por tanto, exactamente igual que el par $(N(p), \Pi(p))$ para evaluar la incertidumbre, es decir, confianza en la verdad o falsedad de la proposición p. La situación habitual será que estos valores sean proporcionados por el usuario del sistema al introducir los datos, hechos conocidos con mayor o menor incertidumbre, al pasar un caso concreto por el sistema, o le serán preguntados en el sistema de inferencia. Parece bastante más claro el que proporcione un único valor v(p) que no un par $(N(p), \Pi(p))$ con el significado de necesidad y posibilidad que no estará claro para él, si desconoce la teoría que subyace detrás de estos conceptos. En cambio v(p) expresa la creencia de este usuario en la verdad de p con los ya comentados tres valores extremos:

$v(p) = 0 \longrightarrow$ certeza absoluta en la falsedad de p

$v(p) = 0'5 \longrightarrow$ ignorancia total sobre p

$v(p) = 1 \longrightarrow$ p verdadero con total certeza

Los valores intermedios $0 < v(p) < 0'5$ expresarán la mayor o menor credibilidad en la falsedad de p y $0'5 < v(p) < 1$, significará que se cree que p es verdadero con un cierto grado de incertidumbre.

Sin embargo, parece que se debe proporcionar al usuario una mayor claridad a la hora de evaluar la incertidumbre sobre los hechos, de forma que introduzca valores negativos si cree que el hecho es en cierta medida falso y valores positivos si cree que es verdadero. Para ello basta una simple transformación lineal, y en consecuencia, desde ahora se considerará la medida

$$V(p) = 2 \cdot v(p) - 1$$

Tras esta modificación, la relación entre esta medida V(p) y las medidas de necesidad y posibilidad N(p) y Π(p) queda:

$$\boxed{V(p) = N(p) + \Pi(p) - 1} \qquad (6)$$

$$\boxed{\begin{array}{ll} V(p) \geq 0 & N(p) = V(p); \; \Pi(p) = 1 \\ V(p) \leq 0 & N(p) = 0; \; \Pi(p) = V(p) + 1 \end{array}} \qquad (7)$$

El adecuado comportamiento de esta única medida se irá constatando a medida que se expongan los distintos mecanismos de inferencia, principalmente en la propagación de incertidumbre y combinación de distintas reglas con la misma conclusión.

4.2.- INCERTIDUMBRE DE REGLAS. MECANISMOS DE INFERENCIA.

Por las razones apuntadas al establecer los supuestos en que se basa el modelo que se propone, se mantendrá un doble valor para cuantificar la incertidumbre de una regla cuya conclusión no es difusa, con el significado de grado de suficiencia y grado de necesidad del antecedente para la conclusión.

4.2.1.- Propagación de incertidumbre.

El esquema de razonamiento del que se parte es el del modelo de Dubois y Prade en base a la doble medida de necesidad y posibilidad. De las tres alternativas que aquéllos proponían no se tomará en principio la que utilizan en el ejemplo de aplicación, sino la más justificada a partir de la Teoría de la Posibilidad, es decir, se parte del siguiente esquema:

$$\begin{array}{l} N(p \longrightarrow q) \geq s \\ N(q \longrightarrow p) \geq n \\ N(p) \geq a \\ \Pi(p) \quad b \\ \hline N(q) \geq \min\,(s,a) \qquad\qquad (8) \\ \Pi(q) \leq \max\,(1-n,b) \end{array}$$

Sin embargo, los datos de partida no son los valores de necesidad y posibilidad del antecedente p de la regla, sino su valor de verdad V(p), definido anteriormente. Es inmediato obtener el valor de verdad V(q) de la conclusión con este esquema de inferencia (8) y a partir de las relaciones (7).

$$\bullet)\ 0 \leq V(p) \leq 1 \xrightarrow[(7)]{} \begin{array}{l} N(p) = V(p) \\ \Pi(p) = 1 \end{array} \xrightarrow[(8)]{} \begin{array}{l} N(q) = \min\ (s, V(p)) \\ \Pi(q) = \max\ (1-n, 1) = 1 \end{array}$$

$$\xrightarrow[(6)]{} V(q) = \min\ (s, V(p))$$

$$\bullet)\ -1 \leq V(p) \leq 0 \longrightarrow \begin{array}{l} N(p) = 0 \\ \Pi(p) = V(p) + 1 \end{array} \longrightarrow \begin{array}{l} N(q) = \min\ (s, 0) = 0 \\ \Pi(q) = \max\ (1-n,\ V(p) + 1) \end{array}$$

$$\longrightarrow V(q) = \max\ (1-n,\ V(p) + 1) - 1$$

Dado que V(p) es negativo, se llega a unas expresiones más homogéneas para ambos casos, de la siguiente forma:

$$V(q) = \max\ (1-n,\ 1 - |V(p)|) - 1 = -\min\ (n,\ |V(p)|)$$

De forma conjunta, el esquema de propagación de incertidumbre queda de la siguiente forma:

$$\boxed{\begin{array}{l} V(p) \geq 0 \longrightarrow V(q) = \min\ (s, V(p)) \\ V(p) \leq 0 \longrightarrow V(q) = -\min\ (n,\ |V(p)|) \end{array}} \qquad (9)$$

Obsérvese que la homogeneidad antes apuntada se pone aún más de manifiesto expresando el segundo caso como

$$|V(q)| = \min\ (n,\ |V(p)|)$$

La diferencia entre ambas relaciones está en la aparición de cada uno de los parámetros s y n en las dos situaciones consideradas, lo cual es totalmente razonable:

.) $0 \leq V(p) \leq 1$. En esta situación el valor de verdad de la conclusión $V(q)$ no depende del coeficiente n, sino únicamente del grado de suficiencia s, Es un comportamiento correcto, puesto que n es una medida del grado en que se verifica la conclusión en ausencia del hecho o hechos que figuran en el antecedente, es decir cuando el antecedente es falso, y la situación que se está considerando es que p es verdadero en un cierto grado, $V(p) \geq 0$. Es lógico, en consecuencia, que sólo intervenga el grado de suficiencia s de la regla, pues este coeficiente cuantifica la verdad de la conclusión q, caso de ser verdadero el antecedente p.

.) $-1 \leq V(p) \leq 0$. Ahora $V(q)$ sólo depende del grado de necesidad n. Es exactamente el mismo comportamiento anterior y las razones para que suceda de esta forma son las mismas, puesto que ahora p es falso en un cierto grado.

La función de propagación (9) parece correcta si se tienen en cuenta las siguientes situaciones extremas:

- $V(p) = 0 \Longrightarrow V(q) = 0$. Si hay ignorancia total sobre el antecedente de la regla, no se puede deducir nada sobre la conclusión, independientemente de cúal sea la "fuerza" de la regla.

- $V(p) = 1 \Longrightarrow V(q) = s$. El antecedente p es verdadero con total certeza y el valor de la conclusión viene dado por el grado de suficiencia s, de acuerdo con el significado de este parámetro.

- $V(p) = -1 \Longrightarrow V(q) = -n$. También está de acuerdo con el significado del grado de necesidad n. Este valor será tanto mayor cuanto más necesaria sea la presencia de p para que se obtenga q. Por tanto si $n=1$, es decir absoluta necesidad de p para la verdad de q, si $V(p) = -1$, es decir p falso, se obtiene $V(q) = -1$, q falso.

- $s = 0$, $0 \leq V(p) \leq 1 \Longrightarrow V(q) = 0$. Si p no es en absoluto suficiente para q, $s = 0$, la verdad de p no aporta nada al conocimiento de q y existe ignorancia total, $V(q) = 0$.

- $s = 0'5$, $V(p) = 1 \Longrightarrow V(q) = 0'5$. De un grado de suficiencia de p para la verdad de q medio y de la verdad de p se obtiene un grado de verdad de q también medio.

- $n = 0$, $V(p) \leq 0 \Longrightarrow V(q) = 0$. El antecedente p no es en absoluto necesario para q. La falsedad de p no aporta ningún conocimiento a la verdad de q.

- $n = 0'5$, $V(p) = -1 \Longrightarrow V(q) = -0'5$. Si el grado de necesidad de p para q es medio, el grado de credibilidad en la falsedad de q también es medio.

Si ahora se representan gráficamente las relaciones (9) para distintos valores de los parámetros s y n, figuras 1 y 2, se aprecian ya algunas formas de comportamiento del modelo de propagación propuesto no muy correctas.

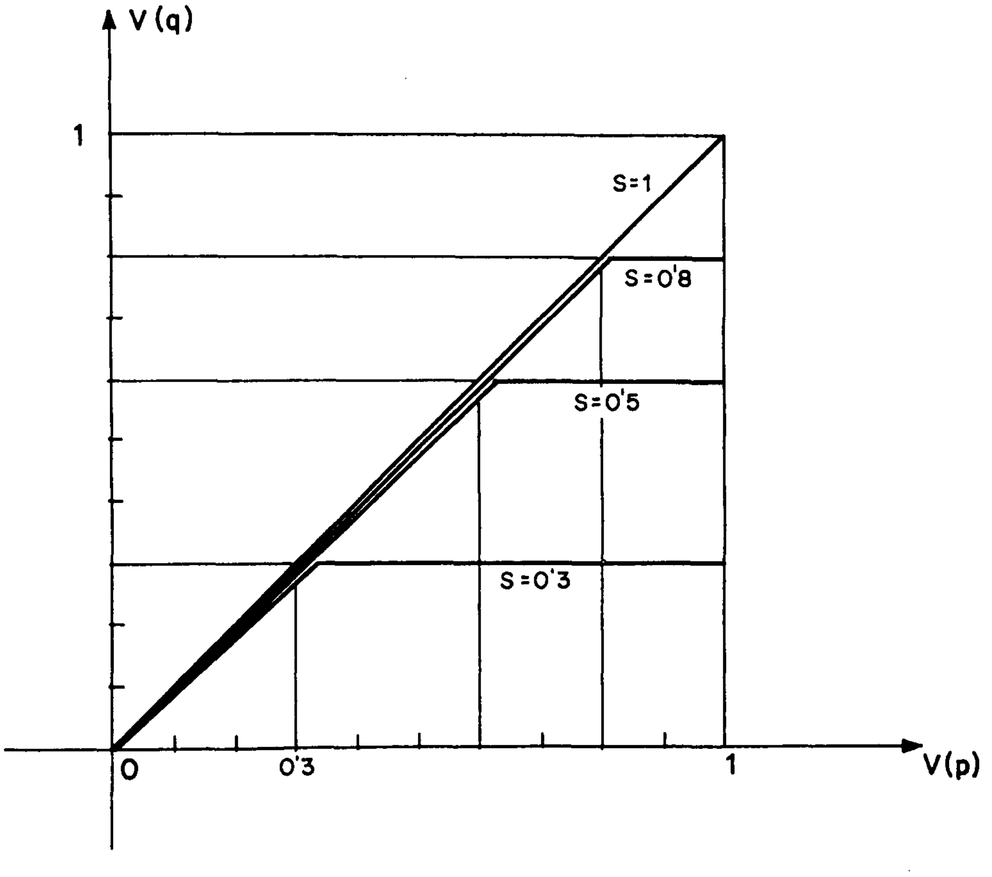


Figura 1.- Inferencia a partir de hechos ciertos.

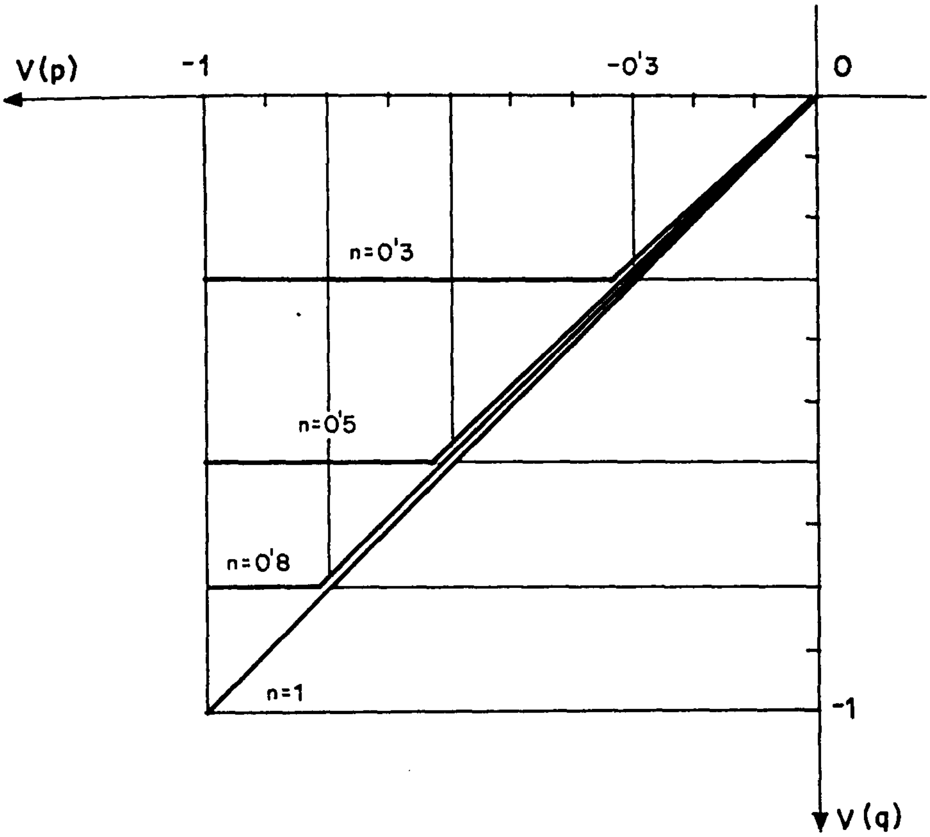


Figura 2.- Inferencia a partir de hechos falsos.

En efecto, para p cierto si $V(p) \geqslant s$ entonces $V(q) = s$, independiente del grado de verdad de p. En el caso de que $V(p) \leqslant s$ entonces $V(q) = V(p)$, independiente del valor del parámetro s. Es conveniente aclarar este hecho con algunos valores numéricos. Por ejemplo para s = 0'4, se tiene

V(p)	V(q)	Modelo MYCIN
0'2	0'2	0'08
0'4	0'4	0'16
0'5	0'4	0'20
0'7	0'4	0'28
1	0'4	0'4

Se acompañan también los valores que resultarían caso de utilizar el modelo de los factores de certeza de MYCIN considerando $V(p) = CF(p,e')$ y s $= CF(q,p)$ y con la fórmula de propagación $CF(q,e')=CF(q,p).\max(0,CF(p,e'))$ no sólo porque pone de manifiesto el incorrecto funcionamiento del esquema de propagación de incertidumbre (9), sino porque, como se expone más adelante, permitirá clarificar o sacar a la luz las razones no explicitadas por Dubois y Prade para el cambio en su esquema de razonamiento.

En primer lugar, parece bastante más correcta la forma de propagar la incertidumbre de MYCIN, puesto que se tienen en cuenta las incertidumbres del antecedente y de la regla, reforzándose, para obtener el grado de verdad de la conclusión.

Comportamientos análogos tienen lugar para valores de verdad $V(p)$ negativos, donde sólo aparece el grado de necesidad n de la regla. El paralelismo con MYCIN sigue siendo igual puesto que, en términos de factores de certeza, se estaría considerando la regla "si no p entonces q" y la expresión para propagar la incertidumbre será:

$$CF(q,e') = - CF(q,\neg p) \ . \ \max\ (0,\ CF(\neg p,e'))$$

con $CF(\neg p,e')$ jugando el mismo papel que $|V(p)|$, dado que ahora $V(p)$ es negativo.

En consecuencia, es más adecuado utilizar el producto de las incertidumbres de antecedente y regla, que no el operador min. Este es precisamente el cambio que se produce en la aplicación del modelo de Dubois y Prade. Recordando la función de propagación que estos emplean:

$$\begin{aligned} N(q) &= S\ .\ N(p) \qquad\qquad (10) \\ \Pi(q) &= 1 - n\ .\ (1 - \Pi(p)) \end{aligned}$$

y repitiendo el proceso de transformación del par $(N(p), \Pi(p))$ en la medida $V(p)$ se tiene:

$$\bullet)\ V(p) \geq 0 \xrightarrow[(7)]{} \begin{array}{l} N(p) = V(p) \\ \Pi(p) = 1 \end{array} \xrightarrow[(10)]{} \begin{array}{l} N(q) = s \,.\, V(q) \\ \Pi(q) = 1 \end{array} \xrightarrow[(6)]{} V(q) = s \,.\, V(q)$$

$$\bullet)\ V(p) \leq 0 \xrightarrow[(7)]{} \begin{array}{l} N(p) = 0 \\ \Pi(p) = V(p) + 1 \end{array} \xrightarrow[(10)]{} \begin{array}{l} N(q) = 0 \\ \Pi(q) = 1 - n \,.\, V(p) \end{array}$$

$$\xrightarrow[(6)]{} V(q) = n \,.\, V(p) = -n \,.\, |V(p)|$$

Se puede ver ya claramente que la forma de propagar incertidumbre utilizada por D-P es exactamente la misma que se emplea en el modelo MYCIN.

$$\boxed{\begin{array}{ll} V(p) \geq 0 & V(q) = s \,.\, V(p) \\ & \\ V(p) \leq 0 & V(q) = n \,.\, V(p) \end{array}} \qquad (11)$$

y que es la adoptada en el motor de inferencia MOTOR-2, que figura en el apéndice de este trabajo.

Todo este razonamiento es ahora un nuevo argumento a favor de la capacidad que tiene la biyección entre $(N(p), \Pi(p))$ y la medida $V(p)$, no sólo para conseguir el objetivo que se pretendía, sino también para analizar los comportamientos que subyacen en la Teoría de la Posibilidad de Dubois y Prade.

4.2.2.- Combinación de reglas.

Sean $\underline{n}$ reglas "si p_i entonces q; s_i, n_i" con la misma conclusión $\underline{q}$. Se trata de obtener el valor de verdad $V(q)$ de la conclusión a partir del grado de verdad $V_i(q)$ deducido mediante cada una de las reglas.

De forma similar a la que se utilizó para obtener la función de propagación de incertidumbre, simplemente se adaptará la propuesta que Dubois y

Prade hacen, para la situación que ahora se considera, en base a las dos medidas $N(p), \Pi(p)$, a la única medida $V(p)$ que se está utilizando. El procedimiento que aquéllos empleaban consistía sencillamente en:

$$N(q) = \max_i (N_i(q))$$
$$\Pi(q) = \min_i (\Pi_i(q)) \qquad (12)$$

seguido del proceso de normalización correspondiente: si $k = \max (\min_i (\Pi_i(q)), \min_i (1 - N_i(q))$ se tiene

$$\Pi(q) = \frac{\min (\Pi_i(q))}{k}$$
$$N(q) = 1 - \frac{\min (1 - N_i(q))}{k} \qquad (13)$$

Para adoptar estos valores a la única medida $V(q)$ se deberán considerar las tres situaciones siguientes:

Caso 1: $V_i(q) \geqslant 0$ para todas las reglas que concluyen en q.

Según las relaciones (7) se obtiene: $\Pi_i(q) = 1$, $N_i(q) = V_i(q)$ y para el conjunto de la reglas, según (12): $\Pi(q) = 1$, $N(q) = \max_i (V_i(q))$.

No es necesario hacer ningún proceso de normalización y se obtiene:

$$\boxed{V(q) = \max_i (V_i(q))} \qquad (14)$$

Nuevamente aparece el ya clásico problema de si, al combinarse varias fuentes de información, unas deben reforzar a las otras a medida que se van obteniendo. En el modelo MYCIN, así como mediante la regla de combinación de Dempster, sí se produce este incremento en la credibilidad a medida que van siendo combinadas las distintas evidencias. No es éste el resultado que se ha obtenido y ello se debe a la forma como Dubois y Prade deducen la forma de combinar distintas informaciones que era, según se vió en 2.5.4., considerando cada información como un conjunto difuso y hallando la intersección de todos ellos mediante el operador min.

Caso 2: $V_i(q) \leq 0$ para todas las reglas.

Se tiene ahora $N_i(q) = 0$, $\Pi_i(q) = V_i(q) + 1$ y, por aplicación de (12)

$$N(q) = 0, \quad \Pi(q) = \min_i (V_i(q) + 1)$$

quedando finalmente

$$V(q) = \min_i (V_i(q) + 1) - 1$$

que es el mismo resultado anterior, si se observa que $V_i(q) \leq 0$ y entonces

$$V(q) = \min_i (1 - |V_i(q)|) - 1 = 1 - \max_i (|V_i(q)|) - 1$$

es decir

$$\boxed{V(q) = -\max_i (|V_i(q)|)} \qquad (15)$$

Caso 3: Sean dos reglas tales que $V_1(q) \geq 0$ y $V_2(q) \leq 0$

$$N_1(q) = V_1(q), \quad \Pi_1(q) = 1 \quad y \quad N_2(q) = 0, \quad \Pi_2(q) = V_2(q) + 1$$

En esta situación sí es preciso normalizar, obteniéndose:

$$\Pi(q) = \frac{V_2(q) + 1}{k}$$

$$N(q) = 1 - \frac{1 - V_1(q)}{k}$$

con $k = \max(V_2(q) + 1,\ 1 - V_1(q))$

Se llega a

$$V(q) = \frac{V_1(q) + V_2(q)}{\max(1 - |V_2(q)|,\ 1 - V_1(q))} = \boxed{\frac{V_1(q) + V_2(q)}{1 - \min(|V_1(q)|,\ |V_2(q)|)}} \tag{16}$$

que es exactamente la fórmula utilizada en el sistema MYCIN para el caso de dos informaciones contradictorias.

Se plantea el importante problema de que estas tres fórmulas de combinación no verifican la asociatividad cuando hay evidencias a favor y en contra, como se pone de manifiesto con los siguientes valores:

$$V_1(q) = 0'6 \qquad V_2(q) = -0'5 \qquad V_3(q) = 0'4$$

$$V_{12}(q) = \frac{0'6 - 0'5}{1 - \min(0'6, 0'5)} = \frac{0'1}{0'5} = 0'2$$

$$V_{(12)3}(q) = \underline{0'4}$$

$$V_{23}(q) = \frac{-0'5 + 0'4}{1 - \min(0'5, 0'4)} = -1/6$$

$$V_{1(23)}(q) = \frac{0'6 - 1/6}{1 - \min(0'6, 1/6)} = \underline{0'52}$$

Las dos alternativas que se plantean son: 1) modificar las funciones de combinación para los casos de todas las fuentes de información a combinar, utilizando la función de combinación del modelo de factores de certeza, y 2) hacer todo el proceso con medidas de necesidad y posibilidad y una vez obtenido el resultado final, transformarlo en el valor de verdad V(q) definitivo, lo que no es complicado de calcular, aunque no permite la obtención de una fórmula sencilla para poder interpretarla. Es esta última opción la que se toma en el modelo que propone, y que es la que se ha implementado en el correspondiente motor de inferencia, MOTOR-2.

4.2.3.- Conjunción de premisas.

Cuando en una regla el antecedente es la conjunción de varias premisas, se deberá obtener el valor de verdad de dicha conjunción a partir de los valores de verdad de cada una de las premisas componentes.

Sea $p = p_1 \wedge \dots \wedge p_n$. Dubois y Prade proponen en su modelo, sin ninguna justificación explícita, las siguientes relaciones

$$\Pi(p) = \min_i (\Pi(p_i))$$

$$N(p) = \min_i (N(p_i))$$

que parece clara en el caso de la medida de necesidad, por las propiedades de ésta, pero no tanto en el caso de la posibilidad.

Si lo que se conocen son los valores de verdad de cada una de las premisas $V(p_i)$, resulta fácil la traducción de las fórmulas de D-P:

<u>Caso 1</u>: $V(p_i) \geq 0 \quad \forall i \longrightarrow \begin{array}{l} \Pi(p_i) = 1 \\ N(p_i) = V(p_i) \end{array} \longrightarrow \begin{array}{l} \Pi(p) = 1 \\ N(p) = \min (V(p_i)) \end{array}$

$$\longrightarrow V(p) = \min (V(p_i))$$

<u>Caso 2</u>: $V(p_i) \leq 0 \quad \forall i \longrightarrow \begin{array}{l} \Pi(p_i) = 1 + V(p_i) \\ N(p_i) = 0 \end{array} \longrightarrow \begin{array}{l} \Pi(p) = \min (1 + V(p_i)) \\ N(p) = 0 \end{array}$

$$V(p) = \min (1 + V(p_i)) - 1 = \min (1 - |V(p_i)|) - 1 = - \max (|V(p_i)|)$$

Y puesto que todos los $V(p_i)$ son negativos:

$$V(p) = \min (V(p_i))$$

Caso 3: $V(p_1) \geq 0, \ V(p_2) \leq 0$ $\qquad \Pi(p_1) = 1, \ N(p_1) = 1$

$$\Pi(p_2) = 1 + V(p_2), \ N(p_2) = 0$$

$$\Pi(p) = 1 + V(p_2)$$

$$V(p) = V(p_2) = \min\ (\underbrace{V(p_1)}_{\geq 0}, \underbrace{V(p_2)}_{\leq 0})$$

$$N(p) = 0$$

Por tanto, en cualquier caso $\boxed{V(p) = \min_i\ (V(p_i))}$ (17)

Nuevamente se obtienen las fórmulas de MYCIN, si se interpretan los factores de certeza $CF(p_i,e)$ como valores de verdad $V(p_i)$. Además coincide con el valor de verdad para la conjunción de proposiciones, (5), de la lógica multivalente de Dienes que fue la que se obtuvo como consecuencia de la introducción de los valores de verdad V.

CAPITULO 5 :

RESULTADOS Y CONCLUSIONES. FUTURAS LINEAS DE INVESTIGACION

En este trabajo se ha propuesto un "nuevo" modelo de resolución de incertidumbre, basado en la Teoría de la Posibilidad, desarrollando el modelo de Dubois-Prade. Al igual que éste, permite trabajar de forma conjunta con conocimiento incierto y con hechos imprecisos o difusos. Más concretamente, la parte relativa al conocimiento difuso no se modifica en ningún punto. Sí en lo referente al tratamiento de incertidumbre.

El objetivo concreto propuesto como hipótesis de trabajo fundamental, capítulo 3, se ha conseguido plenamente. Se trataba de mejorar la sencillez de utilización para el usuario, de un sistema experto con una base de conocimiento afectada de incertidumbre. Una gran mejora respecto al modelo de D-P reside en cómo el usuario debe valorar la incertidumbre del conocimiento que introduce, bien inicialmente como base de hechos, bien cuando es consultado por el sistema sobre algún suceso concreto. El usuario evalúa su confianza en un hecho con un valor numérico en el intervalo $[-1,1]$, positivo si cree que es cierto y negativo si piensa que es falso. Esencialmente, es el mismo modo de valorar la incertidumbre en el sistema MYCIN, que parece bastante claro y satisfactorio. Por supuesto que es mucho más sencillo que introducir un par de valores numéricos con el significado de posibilidad y necesidad de un suceso. En cuanto a los resultados que se obtienen del proceso de inferencia, también vienen valorados con un único número con el mismo sentido anterior, lo que simplifica también el proceso de toma de decisiones.

En lo relativo al objetivo más general que se planteaba en la introducción de este trabajo, parece que se aporta algo de claridad al complejo panorama de elección de un modelo para la resolución de la incertidumbre. Y esto se consigue en muy variados aspectos:

1) El modelo de Dubois y Prade para resolver los problemas de incertidumbre es esencialmente el mismo que el del sistema MYCIN de factores de certeza. Se han unificado las dos medidas de posibilidad y necesidad en una

única medida, con el mismo significado que los factores de certeza de MYCIN. Cuando se traducen los mecanismos de propagación y combinación de incertidumbre de D-P a esta nueva medida, los resultados que se obtienen son esencialmente las funciones de propagación y combinación del modelo MYCIN.

El motor de inferencia de D-P presentaba la ventaja de su flexibilidad, al permitir fijar los operadores en la funciones de propagación. De las tres alternativas que se citaban, no se escogía la más justificada a partir de la Teoría de la Posibilidad. Con la conversión a la nueva medida, se pone de manifiesto de forma clara, que la manera de propagar incertidumbre de D-P es exactamente la misma del modelo MYCIN. En este aspecto de propagación, el modelo D-P es tan "ad-hoc" como el sistema MYCIN, algo que ha sido objeto de crítica en múltiples ocasiones.

El modelo D-P difiere de MYCIN en cuanto a la forma de combinar incertidumbre. Para un conjunto de evidencias que apoyen todas la misma hipótesis, las formas de combinarlas son:

$$\text{MYCIN} \longrightarrow CF_1 + CF_2 - CF_1 \cdot CF_2$$
$$\text{DUBOIS-PRADE} \longrightarrow \max\,(V_1(p)),\ V_2(p))$$

Con el primero de ellos, de la combinación de varias evidencias a favor de una hipótesis, se refuerza la credibilidad que se tenía en ésta al considerar sólo una de ellas, mientras que esto no ocurre así en el modelo de D-P, tomándose la mayor de las credibilidades proporcionadas por cada una de las evidencias por separado. La elección de una u otra forma dependerá fundamentalmente del tipo de problema que se desee abordar mediante un sistema experto.

Finalmente, la forma de obtener la credibilidad de una conjunción de premisas es idéntica en ambos modelos.

De todo ello se puede concluir que D-P y MYCIN no son sistemas de tratamiento de incertidumbre distintos, sino que son esencialmente el mismo modelo.

Se sigue manteniendo la versatilidad en cuanto a la propagación de incertidumbre que poseía el modelo D-P, con la ventaja de que se ha puesto claramente de manifiesto la diferencia de cada sistema. El diseñador del sistema experto tiene mayor claridad a la hora de elegir cómo enfrentarse con los problemas de incertidumbre.

2) En sentido contrario, se amplían las posibilidades de aplicación de MYCIN. Este es un sistema muy difundido y ha sido incorporado como método de resolución de incertidumbre en herramientas específicas de Inteligencia Artificial, como el Personal Consultant Plus, por ejemplo. Pero MYCIN sólo era capaz de abordar problemas de incertidumbre.

Con la correspondencia biunívoca entre las medidas (Necesidad, Posibilidad) y la medida V(p), esencialmente un factor de certeza para hechos, es posible aplicar de forma rigurosa el modelo MYCIN a hechos difusos. Para ello basta con obtener los valores de necesidad y posibilidad de ese hecho mediante las mismas funciones que Dubois y Prade proponen, a partir de los trabajos de Zadeh, y obtener su credibilidad, en términos de factor de certeza, a través de dicha aplicación biyectiva. Si, además se desea incorporar reglas difusas, su tratamiento podría ser el que utilizan Dubois y Prade, derivado del "modus ponens generalizado".

La potencia de la biyección entre $(N(p), \Pi(p))$ y $V(p)$ ha permitido clarificar los métodos y funciones empleadas por Dubois y Prade. Pero, además, esta relación entre ambas formas de cuantificar la incertidumbre amplía las posibilidades de elección de mecanismos de propagación y combinación de incertidumbre. En efecto, las funciones de combinación de D-P, tal y como se obtuvieron expresadas en función de la medida V(p) no verificaban la asociatividad, lo que plantea problemas de importancia. Lo mismo sucedía en el sistema MYCIN antes de que fuera modificada la función de combinación, lo cual obligaba a mantener separadamente las dos medidas MB y MD, de credibilidad e incredibilidad. Si se desea mantener el esquema de combinación de DuboisPrade, no hay ningún problema. En el momento adecuado se convierte la medida V(p) en el par $(N(p), \Pi(p))$, se aplican los esquema de combinación de D-P, que sí son asociativos, y el resultado se reconvierte en una única medida. En definitiva, el modelo funciona aparentemente con un único valor de incertidumbre, tanto para la entrada de datos como para la obtención de resultados, pero si es necesario, internamente, puede trabajar con la doble medida (Necesidad, Posibilidad). Es posible, en consecuencia, la utilización de todos los resultados de la Teoría de la Posibilidad.

En el apéndice de este trabajo se adjuntan, con los nombres de MOTOR-1 y MOTOR-2, dos motores de inferencia, implementados en GCLISP. El primero de ellos es, con algunas modificaciones, el de Dubois-Prade. El segundo es la correspondiente modificación del anterior, de manera que interviene una sola medida V(p) y las funciones de propagación y combinación son obtenidas, para esta medida, en el capítulo 4. Se han probado ambos motores para una misma base de conocimientos y, efectivamente, los resultados obtenidos son los mismos.

Futuras líneas de investigación.

De acuerdo con la línea general con que se ha enfocado este trabajo y, ante la gran diversificación de métodos, se constituye como prioritario el establecer una clasificación de problemas, suficientemente estructurada, de cara a elegir el modelo de resolución de incertidumbre adecuado a cada tipo de problema. Es una tarea probablemente compleja, pero cada vez más necesaria.

En aspectos más concretos relacionados con el trabajo realizado, pueden constituir futuras líneas de investigación las siguientes:

- Evaluar y calibrar los resultados que se obtendrían con las distintas funciones de propagación de incertidumbre propuestas por Dubois y Prade. Igual comparación debería hacerse con las funciones de combinación de D-P y la del modelo MYCIN. Sin embargo, esto tan sólo sería de valor en el estudio de problemas reales, no con ejercicios académicos.

- Extender la implementación del "modus ponens generalizado", realizada en MOTOR-1 sólo con funciones de pertenencia con parametrización trapezoidal, a conjuntos difusos parametrizados como curvas en "S".

- Refinar los procedimientos para razonar con conocimiento difuso, ya que en demasiadas ocasiones se recurre a intersecciones de conjuntos sin demasiado fundamento.

- Completar y desarrollar la Teoría de la Evidencia de Dempster-Shafer en aquellos aspectos aún no resueltos, como pueda ser la propagación de incertidumbre, con el uso de diferentes marcos de discernimiento, puesto que esta teoría parece tener una gran potencialidad de cara a la resolución de problemas de incertidumbre.

Quedan aún muchos problemas por resolver en este campo, aunque se alejan mucho de la línea de trabajo seguida. Uno de los más obviados es el relativo a problemas de dependencia e independencia condicional, que o bien son resueltos con la suposición de independencia o con la propuesta de sistemas demasiado complejos para su aplicación a Sistemas Expertos.

REFERENCIAS

BIBLIOGRAFIA REFERENCIADA

(ADAM-76) Adams, J.B.
"A probability model of medical reasoning and the MYCIN model". Math. Biosci., 32, 1976.

(BALD-79) Baldwin, J.F.
"A new approach to approximate reasoning using a fuzzy logic". Fuzzy Sets and Systems, 2, 1979.

(BALD-80) Baldwin, J.F. y Pilsworth, B.W.
"Axiomatic approach to implication for approximate reasoning with fuzzy logic". Fuzzy Sets and Systems, 3, 1980.

(BARN-81) Barnett, J.A.
"Computational Methods for a Mathematical Theory of Evidence". Proceedings of the 7th International Joint Conference on Artificial Intelligence, Vancouver, 1981.

(BUCH-84) Buchanan, B.G. y Shortliffe, E.H.
"Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project". Addison-Wesley, Reading, Massachussetts, 1984.

(CHEE-85) Cheeseman, P.
"In defense of probability". Proceeding of the 9th International Joint Conference on Artificial Intelligence, 1985.

(COHE-84) Cohen, P.R.
"Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach". Pitman, London, 1984.

(DEMP-67a) Dempster, A.P.
"Upper and lower probabilities induced by a multivalued mapping". Ann. Math. Statis. 38, 1967.

(DEMP-68) Dempster, A.P.
"A generalization of Bayesian inference". J. Roy. Statis. Soc. Ser B 30, 1968.

(DUBO-85a) Dubois, D. y Prade, H.
"Combination and Programation of Uncertainty with Belief Functions. -A Reexamination-". Proceedings of the 9th International Joint Conference on Artificial Intelligence, Los Angeles, 1985.

(DUBO-85b) Dubois, D. y Prade, H.
"Théorie des Possibilités. Applications à la représentation des connaissances en informatique". Masson ed., París, 1985.

(DUDA-76) Duda, R.O., Hart, P.E. y Nilsson, N.J.
"Subjective Bayesian methods for rule-based inference systems". Proceeding of the National Computer Conference, 1976.

(DUDA-78) Duda, R.O., Hart, P.E., Nilsson, N.J. y Sutherland, G.L.
"Semantic Network Representations in rule-Based rnference systems". En "Pattern-Directed Inference Systems". Waterman, D.A. y Hayes-Roth, F. eds., Academic Press, New York, 1978.

(DUDA-84) Duda, R.O. y Reboh, R.A.
"AI and Decision Making: the PROSPECTOR Experience". En "Artificial Intelligence Applications for Business", Ablex Publishing Corporation, Norwood, New Jersey, 1984.

(GAIN-76) Gaines, B.R.
"Foundations of fuzzy reasoning". Int. J. Man-Machine Studies, 8, 1976.

(GARV-81) Garvey, T.D., Lowrence, J.D. y Fischler, M.A.
"An Inference Technique for Integrating Knowledge from Disparate Sources". Proceeding of the 7th International Joint Conference on Artificial Intelligence, 1981.

(GINS-84) Ginsberg, M.
"Non-monotonic Reasoning using Dempster's Rule". Proceedings of the 4th National Conference on Artificial Intelligence, 1984.

(GLYM-85) Glymour, C.
"Independence aseumptions and bayesian updating". Artificial Intelligence, vol. 25, nº 1, 1985.

(GORD-84) Gordon, J. y Shortliffe, E.H.
"The Dempster-Shafer Theory of Evidence". En "Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project". Addison-Wesley, Reading, Massachussetts, 1984.

(GORD-85) Gordon, J. y Shortliffe, E.H.
"A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space". Artificial Intelligence 26, 1985.

(GORR-68) Gorry, G.A. y Barnett, G.O.
"Experience with a Model of Sequential Diagnosis". Computers and Biomedical Research, 1, 1968.

(HECK-86) Heckerman, D.E.
"Probabilistic Interpretations for MYCIN's Certainty Factors". En "Uncertainty in Artificial Intelligence". L.N. Kanal y J.F. Lemmer (Eds.), Elsevier Science Publishers, North-Holland, 1986.

(HEMP-65) Hempel, C. G.
"Studies in the logic of confirmation". En "Aspects of Scientific Explanation and Other Essays in the Philosophy of Science", Free Press, New York, 1965.

(HORV-86a) Horvitz, E.J. y Heckerman, D.E.
"The Inconsistent use of Measures of Certainty in A.I. Research". En "Uncertainty in Artificial Intelligence". L.N. Kanal y J.F. Lemmer (Eds.), Elsevier Science Publishers, North-Holland, 1986.

(HORV-86b) Horvitz, E.J., Heckerman, D.E. y Langlotz, C.P.
"A Framework for Comparing Alternative Formalismes for Plausible Reasoning". Proceedings of the 5th National Conference on Artificial Intelligence, 1986.

(HOWA-81) Howard, R.A. y Matheson, J.E.
"Influence diagrams". En "Readings on the Principles and Applications of Decision Analysis", Strategic Decisions Group, California, 1981.

(JOHN-86) Johnson, R.
"Indenpendence and Bayesian Updating Methods". En "Uncertainty in Artificial Intelligence". L.N. Kanal y J.F. Lemmer (Eds.), Elsevier Science Publishers, North-Holland, 1986.

(KAUF-73 a 77) Kaufmann, A.
"Introduction à la Théorie des Sous-Ensembles Flous", Masson, París.

Vol. 1 - Eléments Théoriques de Base, 1973.

Vol. 2 - Applications à la Linguistique, á la Logique et à la Sémantique, 1975.

Vol. 3 - Applications à la Classification et à la Reconnaisance des formes, aux Automates et aux Systèmes, aux Choix des Critères, 1975.

Vol. 4 - Compléments et Nouvelles Applications, 1977.

(KESS-87) Kessler, T., Liu, G., Frogner, B. y Gloski, D.
"INSPECTR: An Interactive Power Plant Performance Diagnostic Assistant". Seminar about Expert Systems Applications in Power Plants, 1987.

(KONO-79) Konolige, K.
"Bayesian methods for updating probabilities". Apendice D en (DUDA-79).

(LOPE-87) López de Mántaras, R., Godo, Ll., Sierra, C. y Verdaguer, A. "Managing Linguistically Expressed Uncertainty in MILORD. Application to Medical Diagnosis", 1987.

(MART-84) Martín-Clouaire, R. "A fast generalized modus ponens". BUSEFAL, nº 18, Universidad Paul Sabatier, Toulouse, 1984.

(PARG-87) Parge, P., Stuart, J. y Vinson, J. "The Development and Application of TURBOMAC, an Expert Machinery Diagnostic System". Seminar About Expert Systems Applications in Power Plants, 1987.

(PEAR-82) Pearl, J. "Reverend Bayes on Inference Engines: A Distributed Hierarchical Approach". Proceedings of the 3th National Conference on Artificial Intelligence, 1982.

(PEAR-85) Pearl, J. "How to do with probabilities what people say you can't". Proceedings of the Second Annual Conference on Artificial Intelligence Applications, Miami, 1985.

(PEAR-86c) Pearl, J. "Fusion, propagation and structuring in belief networks". Artificial Intelligence, 29, 3, 1986.

(PEND-81) Pednault, E.P.D., Zucker, S.W. y Muresan, L.V. "On the Independece Assumption Underlying Subjective Bayesian Updating". Artificial Intelligence, 16, 1981.

(PRAD-82) Prade, H. "Modèles mathématiques de l'imprecis et l'incertain en vue d'applications au raisonnment naturel". Tesis Doctoral. Universidad Paul Sabatier, Toulouse, Junio 1982.

(SHAF-76) Shafer, G.
"A Mathematical Theory of Evidence". Princeton University Press, Princeton, New Jersey, 1976.

(SHOR-75) Shortliffe, E.H. y Buchanan, B.G.
"A model of inexact reasoning in medicine". Mathematical Biosciences, 23, 1975.

(SHOR-76) Shortliffe, E.H.
"Computer-Based Medical Consultations: MYCIN". Elsevier, 1976.

(STRA-84) Strat, T.M.
"Continuous Belief Functions for Evidential Reasoning". Proceedings of the 4th National Conference on Artificial Intelligence, 1984.

(SZOL-78) Szolovits, P. y Pauker, S.
"Categorical and Probabilistic Reasoning in Medical Diagnosis". Artificial Intelligence 11, 1978.

(TRIL-79) Trillas, E.
"Sobre funciones de negación en la teoría de conjuntos difusos". Stochastica, vol. III, nº 1, 1979.

(WARN-61) Warner, H.R., Toronto, A.F., Veasey, L.G. y Stephenson, R.
"A Mathematical Approach to Medical Diagnosis.- Application to Congenital Heart Disease". Journal of the American Medical Association, 177, (3), 1961.

(YAGE-80) Yager, R.R.
"Quantified propositions in a liguistic logic". Procc. Second Int. Seminar on Fuzzy Set Theory, Linz, 1980.

(YEN-86)
"A Reasoning Model Based on an Extended Dempster-Shafer Theory". Proceedings of the 5th National Conference on Artificial Intelligence, 1986.

(ZADE-65) Zadeh, L.A.
"Fuzzy Sets". Information and Control, 8, 1965.

(ZADE-72) Zadeh, L.A.
"A Fuzzy-Set Theoretic Interpretation of linguistic Hedges". Journal of Cybernetics, 1972.

(ZADE-75) Zadeh, L.A.
"The Concept of a Linguistic Variable and its Applications to Approximate Reasoning". Information Sciences, 1975.

(ZADE-77) Zadeh, L.A.
"PRUF - a meaning representation language for natural languages". En "Fuzzy Reasoning and its Applications". Mandani, E.H. y Gaines, B.R. (eds.), Academic Press, London, 1977.

(ZADE-78) Zadeh, L.A.
"Fuzzy Sets as a Basis for a Theory of Possibility". Fuzzy Sets and Systems, 1, 1978.

(ZADE-79a) Zadeh, L.A.
"A Theory of Approximate Reasoning". En "Machine Intelligence, 9" (Hayes, J.E., Michie, D. y Kulich, L.I., eds.) Wiley, New York, 1979.

(ZADE-83a) Zadeh, L.A.
"The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems". Fuzzy Sets and Systems, 11, 1983.

(ZADE-83b) Zadeh, L.A.
"A computational approach to fuzzy quantifiers in natural languages". Computers and Mathematics, 9, 1983.

(ZADE-84) Zadeh, L.A.
"Review of Books: A Mathematical Theory of Evidence". Artificial Intelligence Magazine, 1984.

(ZADE-86) Zadeh, L.A.
"A Simple View of the Dempster-Shafer Theory of Evidence and its Implications for the Rule of Combination". Artificial Intelligence Magazine, 1986.

BIBLIOGRAFIA CONSULTADA.

(ADAM-84) Adams, J.B.
"Probabilistic Reasoning and Certainty Factors". En "Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project". Addison-Wesley, Reading, Massachussetts, 1984.

(BHAT-86) Bhatnagar, R.K. y Kanal, L.N.
"Handling Uncertain Information: A Review of Numeric and Non-Numeric Methods". En "Uncertainty in Artificial Intelligence". L.N. Kanal y J.F. Lemmer (Eds.), Elsevier Science Publishers, North-Holland, 1986.

(CHAN-79) Chandrasekaran, B., Gómez, F., Mittal, S. y Smith, M.
"An approach to medical diagnosis based on conceptual schemes". Proceedings of the 6th International Joint Conference on Artificial Intelligence, Tokyo, 1979.

(CHAN-84) Chandrasekaran, B. y Gómez, F.
"Knowledge organization and distribution for medical diagnosis". En "Reading in Medical Artificial Intelligence: The First Decada". W.I. Clancey y E.H. Shortliffe (eds.), Addison-Wesley, Reading, Massachusetts, 1984.

(CHEE-83) Cheeseman, P.
"A Method of Computing Generalized Bayesian Probability Values for Expert Systems". Proceeding of the 8th International Joint Conference on Artificial Intelligence, 1983.

(CUEN-85) Cuena, J. y otros
"Inteligencia Artificial: Sistemas Expertos". Alianza Editorial, Madrid, 1985.

(DEMP-67b) Dempster, A.P.
"Upper and lower probability inferences based on a sample from a finite univariant population". Biometrika 54, 1967.

(DUBO-80) Dubois, D. y Prade, H.
"Fuzzy Sets and Systems: Theory and Applications". Academic Press, New York, 1980.

(DUDA-79) Duda, R.O., Hart, P.E., Konolige, K. y Reboh, R.A.
"A computer-based consultant for mineral exploration". Final Report, Stanford Research Institute, Project 6415, 1979.

(FRIE-81) Friedman, L.
"Extended Plausible Inference". Proceedings of the 7th International Joint Conference on Artificial Intelligence, 1981.

(GAIN-78) Gaines, B.R.
"Fuzzy and Probability Uncertainty Logics". Information and Control, 38, 1978.

(GASC-82) Gasching, J.
"Application of the PROSPECTOR system to geological exploration problems". En "Machine Intelligence" vol. 10. J.E. Hayes, D. Michie y H. Pas (eds.), New York: Halsted Press, 1982.

(GROS-86) Grosof, B.N.
"Evidential Confirmation as Transformed Probability. On the Duality of Priors and Updates". En "Uncertainty in Artificial Intelligence". L.N. Kanal y J.F. Lemmer (Eds.), Elsevier Science Publishers, North-Holland, 1986.

(HARM-85) Harmon, P. y King, D.
"Expet Systems: Artificial Intelligence in Business". John Wiley, ed., New York, 1985.

(HECK-87) Heckerman, D.E. y Horvitz, E.J.
"On the Expressiveness of Rule-based Systems for Reasoning with Uncertainty". Proceedings of the 6th National Conference on Artificial Intelligence, 1987.

(HUMM-87) Hummel, R. y Manevitz, L.M.
"Combining Bodies of Dependent Information". Proceedings of the 10th International Joint Conference on Artificial Intelligence, 1987.

(ISHI-82) Ishizuka, M., Fu, K.S. y Yao J.T.P.
"Inference Procedures with uncertainty for problem reduction method". Inform. Sciences. vol. 28, 1982.

(LOUI-86) Loui, R.P.
"Interval-Based Decisions for Reasoning Systems". En "Uncertainty in Artificial Intelligence". L.N. Kanal y J.F. Lemmer (Eds.), Elsevier Science Publishers, North-Holland, 1986.

(LU - 84) Lu, S. y Stephanou, H.
"A Set-theoretic Framework for the Processing of Uncertain Knowledge". Proceedings of the 4th National Conference on Artificial Intelligence, 1984.

(NILS-81) Nilsson, N.J.
"Problem Solving Methods in Artificial Intelligence", Mc Graw-Hill, 1981.

(PAZO-87) Pazos Sierra, J.
"Inteligencia Artificial". Ed. Paraninfo, Madrid, 1987.

(PEAR-83) Pearl, J. y Kim, J.
"A computational model for combined causal and diagnostic reasoning in inference systems". Proceeding of the 8th International Joint Conference on Artificial Intelligence, 1983.

(PEAR-86a) Pearl, J.
"A Constraint-Propagation Approach to Probabilistic Reasoning". En "Uncertainty in Artificial Intelligence". L.N. Kanal y J.F. Lemmer (Eds.), Elsevier Science Publishers, North-Holland, 1986.

(PEAR-86b) Pearl, J.
"On Evidential Reasoning in a Hierarchy of Hypotheses". Artificial Intelligence, 28, 1986.

(PEAR-87a) Pearl, J. y Verna, T.
"The Logic of Representing Dependencies by Directed Graphs". Proceedings of the 7th National Conference on Artificial Intelligence, 1987.

(PEAR-87b) Pearl, J.
"Embracing Causality in Formal Reasoning". Proceedings of the 7th National Conference on Artificial Intelligence, 1987.

(POPL-77) Pople, H.E.
"The formation of composite hypotheses in diagnostic problem solving: an excercise in synthetic reasoning". Proceeding of the 5th International Joint Conference on Artificial Intelligence, 1977.

(PRAD-83) Prade, H.
"A Synthetic View of Approximate Reasoning Techniques". Proceedings of the 8th International Joint Conference on Artificial Intelligence, 1983.

(PRAD-85) Prade, H.
"A Computational Approach to Approximate and Plausible Reasoning with Applications to Expert Systems.". IEEE Transactions of Pattern Analysis and Machine Intelligence, vol. PAMI-7, nº 3, 1985.

(SHAF-86) Shafer, G.
"Probability Judgment in Artificial Intelligence". En "Uncertainty in Artificial Intelligence". L.N. Kanal y J.F. Lemmer (Eds.), Elsevier Science Publishers, North-Holland, 1986.

(THOM-85) Thompson, T.R.
"Parallel Formulation of Evidential Reasoning Theories". Proceeding of the 9th International Joint Conference on Artificial Intelligence, 1985.

(ZADE-79b) Zadeh, L.A.
"Fuzzy Sets and information granualarity". En "Advances in Fuzzy Set Theory and Applications", Gupta, M., Ragado, R. y Yager, R. (eds.). North-Holland, 1979.

APENDICE

```
                ;*****   M O T O R - 1   *****


(DEFUN MOTOR1 ()
   (LOAD 'jinicio)
   (LOAD 'base1)
   (inicio)
   (LOAD 'jmot-ndi)
   (LOAD 'jmot-dif))



(DEFUN inicio ()
  (setq *norma-1 '* *norma-2 '*)   ;funcion de propagacion de la incertidumbre
  (atr-prop  reglas)              ;calcula ciertas propiedades de los atributos
  (atr-hecho bh))                 ;asocia a cada atributo la lista de hechos
                                  ;de la base que le contienen


(DEFUN conclusion (nreg)
  (cadr (eval nreg)) )             ;recibe como argumento el nombre de la regla
                                   ;devuelve la conclusion de la regla


(DEFUN atr-prop (reg)
  (COND ((null reg) NIL)
        (T (atr-prop-reg (conclusion (car reg)))  (atr-prop (cdr reg)))))
                   ;llama a atr-prop-reg con la parte de accion de cada regla


(DEFUN atr-prop-reg (acc)
   (setf (get 'RG-ATR (CAR ACC))
         (CONS (CAR REG) (GET 'RG-ATR (CAR ACC))) ) )
  ;asocia al atributo de la conclusion (argumento) la regla a la que pertenece


(DEFUN atr-hecho (BAS-H)
  (DO* ((bas bas-h  (cdr bas))
        (atributo (car (eval (car bas)))
                  (car (eval (car bas)))) )
      ((null bas))
      (SETF (GET 'HE-ATR atributo)
            (cons (car bas) (get 'HE-ATR ATRIBUTO))) ))
             ;a cada atributo le asocia la lista de los hechos de la bh que
             ;lo contienen


;****************************************************************************

(defun probando ()
  (reinicio 'cert)
  (reinicio 'rg-atr)
  (reinicio 'he-atr)
  (reinicio 'discr)
  (reinicio 'proviene)
  (reinicio 'npunto)
  (reinicio 'universo))


(defun reinicio (l)
  (do ((lista (symbol-plist l) (cddr lista)))
      ((null lista))
      (setf (get l (car lista)) nil)))
 ;previo a la introduccion de una nueva base de datos
```

```
;*****    funcion de ejecucion    *****

(DEFUN valor (parametros)
 (probar parametros (get 'rg-atr (car parametros))))
;Funcion maestra de ejecucion del programa


(DEFUN probar (hipotesis rg-concl)
 (LET ((flag NIL)
       (nh NIL))
  (COND ((NOT (equal (caddr hipotesis) 'OK))
                           (difusa (GET 'he-atr (car hipotesis) )))
                               ;la hipotesis es difusa: tratamiento difuso

        (T                     ;la hipotesis es no difusa

          (COND ((esta-base-nd (GET 'he-atr (car hipotesis) )))
                               ;comprueba si esta en la base
                               ;devuelve el intervalo de certeza si esta
                               ;o NIL si no esta

                ((null rg-concl) '(0.0 1.0))
                               ;si no estaba en la base y ninguna regla
                               ;concluye en ella devuelve total ignorancia

                ((encadenamiento rg-concl)
                         (setf (get 'cert (setq nh (hech-num (get 'he-atr
                                (car hipotesis))))) (comb)))
                               ;intenta realizar el encadenamiento
                               ;hacia atras

                (T  '(0.0 1.0)))))))
                               ;si no ha podido realizar el encadenamiento
                               ;devuelve total ignorancia

(DEFUN hech-num (l)
  (COND ((null l) (print 'a))
        ((igual (eval (car l)) hipotesis) (car l))
        (T (hech-num (cdr l)))))
                              ;encuentra el numero de un hecho


(DEFUN comb()
 (LET ((negposit (trinegposit (mapcar 'cdr (get 'proviene nh)) nil nil)))
  (COND ((null (car negposit)) (list (apply 'max (cadr negposit)) 1.0))
        ((null (cadr negposit)) (list 0.0 (apply 'min (car negposit))))
        (T (normalizar (apply 'max (cadr negposit))
                       (apply 'min (car negposit)))))))
                              ;combina, sin reforzar, intervalos de certeza
                              ;calificando un mismo hecho inferido por
                              ;varias reglas. Usa trinegposit


(DEFUN trinegposit (li neg posit)
  (COND ((null li) (list neg posit))
        (( > (caar li) 0.0) (trinegposit (cdr li) neg (cons (caar li) posit)))
        (T (trinegposit (cdr li) (cons (cadar li) neg) posit))))


(DEFUN esta-base-nd (hech)
  (COND ((null hech) nil)      ;la hipotesis no difusa no esta en la base

        ((igual (eval (car hech)) hipotesis) (GET 'cert (car hech) ))
                               ;si esta en la base devuelve el intervalo
                               ;de certeza

        (T (esta-base-nd (cdr hech)))))
                               ;sigue comprobando los demas hechos de la
                               ;base con el mismo atributo
```

```
(DEFUN normalizar (nec pos)
  (LET ((d (max (- 1.0 nec) pos)))
    (COND ((equal d pos) (list (redondeo2 (/ (- (+ d nec) 1.0) d)) 1.0))
          (T (list 0.0 (redondeo2 (/ pos d)))))))
     ;normaliza un intervalo del tipo (nec pos)



(DEFUN redondeo2 (n)
  (/ (round (* n 100)) 100))
        ;deja solo dos cifras decimales a un numero real


(DEFUN round (num)
  (COND ((> (- num (truncate num)) 0.5) (+ (truncate num) 1))
        (t (truncate num))))


(DEFUN encadenamiento (reg)     ;reg: reglas que concluyen a la hipotesis
                                ;solo para hipotesis no difusas
  (COND ((null reg) flag)
        ((igual hipotesis (menos2-ult (cadr (eval (car reg)))))
                                ;si la hipotesis coincide con la conclusion
                                ;de una regla ==>

             (COND ((ensayar (car reg)) (setq flag T)
                                        (encadenamiento (cdr reg)))
                                ;intenta demostrar las premisas de esa regla

                   (T (encadenamiento (cdr reg)))))
        (T (encadenamiento (cdr reg)))))
                      ;devuelve T si ha podido aplicar alguna regla
                      ;o NIL si no ha podido


(DEFUN ensayar (nrg)            ;nrg: una regla a intentar aplicar
 (LET ((x (eval nrg)) (cert-prem nil))
  (COND ((igual (setq cert-prem (int-cert-prem (car x))) '(0 1)) nil)
        (T (accion (cadr x) cert-prem)))))
                                ;si sobre las premisas de nrg tiene
                                ;algun conocimiento pasa a ejecutarla



(DEFUN int-cert-prem (r-prem)     ;r-prem: conjunto de premisas de una regla
  (LET ((i (probar (car r-prem) (GET 'rg-atr (caar r-prem) )))
        (i1 nil))
    (COND ((null (cdr r-prem)) i)
          (T (list (min (car i) (car (setq i1 (int-cert-prem (cdr r-prem)))))
                   (min (cadr i) (cadr i1)))))))
                              ;devuelve el minimo de la necesidad y la
                              ;posibilidad para cada premisa de la regla


        ;encadenamiento, ensayar y int-cert-prem intentan
         ;demostrar las premisas de reglas que concluyen
                    ;en la hipotesis


(DEFUN accion (acto c-p)
  (COND ((hecho-en-base (GET 'he-atr (car acto) ) c-p) T)
        (T (anade c-p)) T))

                              ;ejecucion del consecuente de una regla
                              ;si esta en bh --> hecho-en-base
                              ;si no esta lo añade -->añade
```

```
(DEFUN hecho-en-base (hech c-p)
 (LET ((h (menos2-ult acto)))
  (COND ((null hech) nil)
        ((igual (eval (car hech)) h)
                (setf (get 'proviene (car hech)) (append (get 'proviene
                        (car hech)) (list (cons nrg (infer c-p))))))
        (T (hecho-en-base (cdr hech) c-p)))))
                              ;devuelve nil si el hecho conclusion de la regla
                              ;no esta en la base, si no,
                              ;pone en la propiedad 'proviene de ese mismo
                              ;hecho de la base la regla nrg de la cual se
                              ;puede obtener junto al valor con que se infiere


(DEFUN menos2-ult (l)         ;devuelve la conclusion de la regla sin los
 (butlast (butlast l)))       ;valores de suficiencia y necesidad


(DEFUN anade (c-p)
 (setq x (gensym 'h))                       ;da nombre al nuevo hecho
 (set x (menos2-ult acto))
 (let   ((i nil))
  (setq i (infer c-p))                          ;calcula el valor inferido
  (setf (get 'proviene x) (list (cons nrg i)))
                                          ;inicializa la propiedad 'proviene
                                            ;de ese hecho con la regla que se
                                            ;obtiene y el valor inferido
  (setq bh (append bh (list x)))            ;lo anado a la bh
  (setf (get 'he-atr (car acto))(append (get 'he-atr (car acto)) (list x)))))
                                            ;aumenta la lista de hechos
                                            ;asociada al atributo del hecho
                                            ;que ha creado

(DEFUN infer (c-p)
  (LET* ((z (reverse acto))
         (suf (cadr z))
         (nec (car z))
         (i nil))
    (COND ((AND (< nec 0) (< suf 0))
                   (list (- 1 (cadr (setq i (calinfer (- suf) (- nec) c-p))))
                         (- 1 (car i))))
          (T (calinfer suf nec c-p)))))
                   ;devuelve el intervalo de incertidumbre del hecho inferido
                   ;ayudado por la funcion calinfer y el valor cert-prem
                   ;que se refiere al intervalo de incertidumbre de las
                   ;premisas de la regla que produce este hecho

(DEFUN calinfer (suf nec c-p)
  (list (redondeo2 (apply *norma-1 (list suf (car c-p))))
        (redondeo2 (- 1 (apply *norma-2 (list nec (- 1 (cadr c-p))))))))

(DEFUN igual (a b) (equal a b))
```

```
;*****   PARTE  DIFUSA   *****


(DEFUN difusa (hech)
 (LET ((x nil)
       (h nil)
       (val (caddr hipotesis)))

  (COND
       ((equal (cadr hipotesis) (cadr (setq h (eval (car hech)))))
                             ;el objeto difuso esta en la base en un hecho
                             ;que contiene tambien ese atributo

             (COND ((GET 'discr (caddr h) )        ;esta discretizado
                      (normalizar (neccd (eval val) (eval (caddr h)))
                                  (poscd (eval val) (eval (caddr h)))))
                   (T                             ;no esta discretizado

                       (normalizar (neccc (eval val) (eval (caddr h)))
                                   (poscc (eval val) (eval (caddr h)))))))

                       ;en ambos casos devuelve intervalo de certidumbre

       ((cdr hech)  (difusa (cdr hech)))
                            ;comprueba demas hechos que contienen a ese
                            ;atributo

       (T (setq x (probar-dif  hipotesis rg-concl))
                            ;no esta en la base: encadenamiento difuso

          (COND ((GET 'discr x)
                  (normalizar (neccd (eval val) (eval x))
                              (poscd (eval val) (eval x))))
                            ;lo ha logrado demostrar por encadenamiento
                            ;difuso, esta discretizado

                (T  '(0.0 1.0)))))))
                            ;no he podido demostrarlo por encadenamiento
                            ;difuso y tendre total ignorancia de ese hecho


(DEFUN neccd (c1 c2)
 (LET ((x nil))
  (COND ((null (cdr c2)) (max (- (- (cadar c2) 1)) (mu (caar c2) c1)))
        (T (COND ((equal 0.0 (setq x (neccd c1 (list (car c2)))))  0.0)
                 (T (min x (neccd c1 (cdr c2)))))))))
                      ;calcula la necesidad de c1 sabiendo c2 con
                      ;c1 continúa y c2 discreta


(DEFUN poscd (c1 c2)
 (LET ((x nil))
  (COND ((null (cdr c2)) (min (cadar c2) (mu (caar c2) c1)))
        (T (COND ((= 1.0 (setq x (poscd c1 (list (car c2)))))  1.0)
           (T (max x (poscd c1 (cdr c2)))))))))
                      ;igual que neccd

(DEFUN mu (pto distr)
  (COND ((AND ( <= pto (cadr distr)) ( <= (car distr) pto))  1.0)
        ((OR ( <= (+ (car (cdddr distr)) (cadr distr))  pto)
             ( <= pto (- (car distr) (caddr distr))))             0.0)
        ((<  pto  (car distr))
             (/ (float (+ pto (caddr distr) (- (car distr))))
                (float (caddr distr))))
        (T (/ (float (+ (- pto) (cadr distr) (car (cdddr distr))))
             (float (car (cdddr distr)))))))
                 ;devuelve el valor en el punto 'pto de la distribucion
                 ;definida por la cuadrupla 'distr
```

```
(DEFUN probar-dif (conclusion rg-concl)
 (LET ((y nil)
       (lval nil)
       (z nil)
       (nbp nil)
       (uni nil))
  (COND ((val-dif (GET 'he-atr (car conclusion))))
        ((null rg-concl) nil)
        (T (setq y (gensym 'd)  z (gensym '-))
           (set  y (list (car conclusion) (cadr conclusion) z))
           (setq bh (append bh (list y)))
           (setf (get 'he-atr (car conclusion)) (append (get 'he-atr (car conclusion))
                  (list y)))
           (set z (comb-dif (setq lval (encadenamiento-dif rg-concl))))
           (setf (get 'discr z) '0)
           (setf (get 'proviene y) (list rg-concl lval))
           (setf (get 'npunto z) nbp)
           (setf (get 'universo z) uni)
           z))))
                  ;prueba un hecho que es difuso mediante encadenamiento-dif

(defun val-dif (l)
       (let ((y nil))
            (cond ((null l) nil)
                  ((igual (cadr conclusion) (cadr (setq y  (eval (car l)))))
                          (caddr y))
                  (t (val-dif (cdr l))))))
                                      ;busca en la BH el valor difuso del
                                      ;hecho que se empareja con la
                                      ;conclusion.

(DEFUN comb-dif (l)
       (cond ((null (cdr l)) (eval (car l)))
             (t (inter0 (eval (car l)) (comb-dif (cdr l))))))
                                      ;devuelve el valor difuso de la
                                      ;conclusion cuando varias reglas
                                      ;concluyen en ella, como interseccion
                                      ;de cada uno de los valores difusos.


(DEFUN inter0 (e1 e2)
       (let ((md (max (caar e1) (caar e2)))
             (mf (min (caar (last e1)) (caar (last e2)))))
            (cond ((igual md (caar e1)) (inter1 md (cdr e1)))
                  (t (inter1 md (cdr e2))))))

(DEFUN inter1 (d l)
   (cond ((equal d mf) (list (list d (min (cadr (assq d e1))
                                          (cadr (assq d e2)) ))))
         (t (cons (list d (min (cadr (assq d e1)) (cadr (assq d e2))))
                  (inter1 (caar l) (cdr l))))))
```

```
(DEFUN encadenamiento-dif (reg)
      (cond ((null reg) nil)
            ((igual (butlast conclusion) (butlast (cadr (eval (car reg)))))
                (cons (aplicar-reg-dif (car reg)) (encadenamiento-dif
                                                     (cdr reg))))
            (t (encadenamiento-dif (cdr reg)))))
                                        ;realiza el encadenamiento difuso de
                                        ;las reglas que concluyen en el hecho
                                        ;devuelve una lista con los valores
                                        ;difusos que se obtienen de cada regla

(DEFUN aplicar-reg-dif (numreg)
      (let* ((x (eval numreg))
             (zf t)
             (aux nil)
             (a0 (int-cert-prem-dif (car x)))
             (a aux)
             (conclu (cadr x))
             (c (caddr conclu))
             (z (gensym '-))
             (cd (car (setq uni (get 'universo c))))
             (pasc (/ (- (float (apply '- uni)))
                      (float (setq nbp (get 'npunto c))))))
      (set z (mpg nbp))
      (setf (get 'discr z) '0) z))
                                        ;aplica el modus ponens generalizado a
                                        ;numreg


(DEFUN int-cert-prem-dif (premisas)
      (cond ((or (null premisas) (null zf)) nil)
            (t (setq aux (append aux (cddar premisas)))
               (cons (setq zf (probar-dif (car premisas) (get 'rg-atr
                          (caar premisas))))
                     (int-cert-prem-dif (cdr premisas))))))
          ;pone en aux los conjuntos difusos que figuran en las premisas
          ;devuelve los conjuntos difusos inferidos para las premisas

;********************************** MPG **********************************

(DEFUN mpg (nc)
      (let* ((delta nil)
             (incl nil)
             (one nil)
             (o-1c nil)
             (omega nil)
             (ce (eval c))
             (one-1c nil)
             (v1 (- (car ce) (caddr ce)))
             (v2 (+ (cadr ce) (car (cdddr ce)))))
            (cond ((igual (setq delta (apply 'max (mapcar 'delta1 a a0))) 1.0)
                                                                   (mpgp0 nc))
                  ((igual delta 0.0)
                       (setq incl (mapct '(lambda (x)
                                         (cond ((not (null x)) (list x))))
                                               (mapcar 'inclu2 a a0)))
                       (cond ((null incl) (mpgp1 nc))
                             ((setq omega (apply 'min incl))
                                  (setq o-1c (mu-1 omega ce))
                                  (setq one-1c (mu-1 (apply 'min one) ce))
                                  (mpgp2 nc))))
                  (t (cond ((apply 'and (mapcar 'inclone a a0))
                              (setq omega (apply 'max (mapct
                               '(lambda (x) (cond ((not (null x)) (list x))))
                                 (mapcar 'degcut3 a a0))))
                              (setq o-1c (mu-1 omega ce))
                              (mpgp3 nc))
                           ((setq one-1c (mu-1 (apply 'min one) ce))
                              (mpgp4 nc)) )))))

      ;mpg devuelve una distribucion discreta
      ;llama el mpg apropiado segun 5 casos: mpgp0,mpgp1 mpgp2, mpgp3, mpgp4
      ;con delta1 calcula  CP (A ; A')
```

```lisp
;************************ MPG particulares **********************


(DEFUN  mpgp0 (nc)
  (let ((v (+ cd (* nc pasc))))
       (cond ((eq  nc -1) nil)
             (t (append (mpgp0 (- nc 1)) (list (cons v '(1.0))))))))

(DEFUN mpgp1 (nc)
  (let ((v (+ cd (* nc pasc))))
       (cond ((eq nc -1) nil)
             ((or (<= v v1) (<= v2 v))  (mpgp1 (- nc 1)))
             (t (append (mpgp1 (- nc 1)) (list (cons v (list (mu v ce)))))))))

(DEFUN mpgp2 (nc)
  (let ((v (+ cd (* nc pasc))))
       (cond ((eq nc -1) nil)
             ((or (<= v v1) (<= v2 v))  (mpgp2 (- nc 1)))
             ((or (<= v (car o-1c)) (<= (cadr o-1c) v))
                (append (mpgp2 (- nc 1)) (list (cons v (list (mu v ce))))))
             ((and (< (car one-1c) v) (< v (cadr one-1c)))
                (append (mpgp2 (- nc 1)) (list (cons v '(1.0)))))
             (t (append (mpgp2 (- nc 1)) (list (cons v (list
                (apply 'max (mapcar '(lambda (x y)
                (apply 'max (mapcar '(lambda (z) (mug z y))
                                       (mu-1 (mu v ce) (eval x)))))
                                       a a0))))))))))

(DEFUN mpgp3 (nc)
  (let ((v (+ cd (* nc pasc))))
       (cond ((eq nc -1) nil)
             ((or (<= v v1) (<= v2 v))
                 (append (mpgp3 (- nc 1)) (list (cons v (list delta)))))
             ((and (< (car o-1c) v) (< v (cadr o-1c)))
                 (append (mpgp3 (- nc 1)) (list (cons v (list (mu v ce))))))
             (t (append (mpgp3 (- nc 1)) (list (cons v (list
                (apply 'max (mapcar '(lambda (x y)
                (apply 'max (mapcar '(lambda (z) (mug z y))
                                       (mu-1 (mu v ce) (eval x)))))
                                       a a0))))))))))


 (DEFUN mpgp4 (nc)
   (let ((v (+ cd (* nc pasc))))
        (cond ((eq nc -1) nil)
              ((or (<= v v1) (<= v2 v))
                  (append (mpgp4 (- nc 1)) (list (cons v (list delta)))))
              ((and (< (car one-1c) v) (< v (cadr one-1c)))
                  (append (mpgp4 (- nc 1)) (list (cons v '(1.0)))))
              (t (append (mpgp4 (- nc 1)) (list (cons v (list
                 (apply 'max (mapcar '(lambda (x y)
                 (apply 'max (mapcar '(lambda (z) (mug z y))
                                        (mu-1 (mu v ce) (eval x)))))
                                        a a0))))))))))
```

```
;***************************** FUNCIONES AUXILIARES *************************

(DEFUN mu-1 (y quad)
       (cond ((igual y 1.0) (list (car quad) (cadr quad)))
             ((zerop y) (list (- (car quad) (caddr quad))(+ (cadr quad)
                                                    (car (cdddr quad)))))
             ((zerop (caddr quad))
                    (cond ((zerop (car (cdddr quad))) (list (car quad)
                                                    (cadr quad)))
                          (t  (list (car quad) (+ (* (car (cdddr quad))
                                              (- 1 y)) (cadr quad))))))
             ((zerop (car (cdddr quad))) (list (+ (* (caddr quad) (- y 1))
                                        (car quad)) (cadr quad)))
             (t (list (+ (* (caddr quad) (- y 1)) (car quad))
                    (+ (* (car (cdddr quad)) (- y 1)) (cadr quad))))))
                                 ;devuelve la primera abscisa en que
                                 ;la distribucion quad toma el valor y


(DEFUN mug (z y)
       (let ((uni nil)
             (nbp nil)
             (pas nil)
             (x1 nil)
             (x2 nil)
             (y1 nil)
             (y2 nil)
             (n nil)
             (d nil))
          (cond ((null z) 0.0)
                ((get 'discr y) (setq uni (get 'universo y))
                                (setq nbp (get 'npunto y))
                                (setq pas (/ (apply '- (reverse uni))
                                             (float nbp)))
                                (setq n (truncate (/ (- z (car uni)) pas)))
                                (setq x1 (+ (car uni) (* n pas)))
                                (setq x2 (+ (car uni) (* (+ 1 n) pas)))
                                (setq y1 (cond ((setq d
                                                 (cadr (assq x1 (eval y)))) d)
                                                (0.0)))
                                (setq y2 (cond ((setq d
                                                 (cadr (assq x2 (eval y)))) d)
                                                (0.0)))
                                (/ (+ (* z (- y2 y1))
                                      (- (* x2 y1) (* x1 y2)))
                                   pas))
                ((mu z (eval y) )))))
                                 ;devuelve el grado de pertenencia de y
                                 ;en z, este o no discretizada
```

```
(DEFUN deltal (a1 a01)
       (let ((disa1 nil)
              (d1 nil)
              (d2 nil)
              (ad1 (car (get 'universo a01)))
              (pasa1 nil)
              (u1 nil)
              (a1e (eval a1))
              (a01e (eval a01))
              (u2 nil)
              (na nil) )
        (cond ((null a01) 1.0)
              ((get 'discr a01) (setq na (get 'npunto a01))
                                (setq pasa1 (/ (apply '- (reverse
                                         (get 'universo a01)))
                                               (float na)))
                                (setq disa1 (discr a1e na))
                                (cond ((and (< (assq2 1.0 a01e)
                                               (caar (last disa1)))
                                            (< (caar disa1)
                                               (assq2 1.0 (reverse a01e))))
                                       (setq d1 (cadr (assq (- (caar disa1)
                                                                pasa1)
                                                             a01e)))
                                       (setq d2 (cadr (assq (+ (caar (last disa1))
                                                                pasa1)
                                                             a01e)))
                                       (cond ((and (null d1) (null d2)) 0.0)
                                             ((null d1) d2)
                                             ((null d2) d1)
                                             ((max d1 d2))))
                                      (1.0)))
              ((setq u1 (- (car a1e) (caddr a1e))
                     u2 (+ (cadr a1e) (car (cdddr a1e))))
                (cond ((and (< (car a01e) u2) (< u1 (cadr a01e)))
                       (max (mu u1 a01e) (mu u2 a01e)))
                      (1.0))))))
                ;calcula  CP (a1, a01) para una pareja de distribuciones


(DEFUN inclone (a1 a01)
       (let ((u1 nil)
              (u2 nil)
              (a1e (eval a1))
              (a01e (eval a01)))
        (cond ((get 'discr a01)
                    (setq one (append one (list (min
                                    (mu (setq u1 (assq2 1.0 a01e)) a1e)
                              (mu (setq u2 (assq2 1.0 (reverse a01e))) a1e)))))
                    (and (<= (car a1e) u1) (<= u2 (cadr a1e))))
              (t (setq one (append one (list (min
                                    (mu (car a01e) a1e)
                                    (mu (cadr a01e) a1e) ))))
                 (and (<= (car a1e) (car a01e))
                      (<= (cadr a01e) (cadr a1e)))))))
       ;devuelve t si la parte normal (toma valor 1) de a01 esta incluida en la
       ;parte normal de a1, en caso contrario devuelve NIL.
```

```lisp
(DEFUN degcut3 (a1 a01)
  (let ((ug nil)
        (ud nil)
        (a1e (eval a1))
        (u1 nil)
        (u2 nil)
        (a01e (eval a01))
        (d nil))
      (cond ((get 'discr a01) (setq d (dint3 a1e a01e))
                              (cond ((null d) nil)
                                    ((apply 'max d))))
     (t (setq ug (abs-int (car a1e) (caddr a1e) (car a01e) (caddr a01e)))
        (setq ud (abs-int (cadr a1e) (car (cdddr a1e))
                          (cadr a01e) (car (cdddr a01e))))
        (setq u1 (- (car a1e) (caddr a1e))
              u2 (+ (cadr a1e) (car (cdddr a1e))))
        (cond ((or (null ud) (< (car ud) (cadr a1e)) (< u2 (car ud)))
                 (cond ((or (null ug) (< (car ug) u1) (< (car a1e) (car ug)))
                        nil)
                       (t (mu (car ug) a1e))))
              (t (cond ((or (null ug) (< (car ug) u1) (< (car a1e) (car ug)))
                          (mu (car ud) a1e))
                         (t (max (mu (car ug) a1e) (mu (car ud) a1e))))))))))
                                         ;devuelve el mayor valor distinto de 1
                                         ;de la interseccion entre a1 y a01



(DEFUN inclu2 (a1 a01)
  (let ((ug nil)
        (ud nil)
        (a1e (eval a1))
        (u1 nil)
        (u2 nil)
        (a01e (eval a01))
        (d nil))
      (cond ((get 'discr a01) (setq one (append one (list (min
                                    (mu (assq2 1.0 a01e)  a1e)
                                    (mu (assq2 1.0 (reverse a01e)) a1e)))))
                              (setq d (dint2 a1e a01e))
                              (cond ((null d) nil)
                                    ((apply 'min d))))
           (t (setq one (append one (list (min
                                    (mu (car a01e) a1e)
                                    (mu (cadr a01e) a1e)))))
              (setq ug (abs-int (car a1e) (caddr a1e) (car a01e) (caddr a01e)))
              (setq ud (abs-int (cadr a1e) (car (cdddr a1e))
                                (cadr a01e) (car (cdddr a01e))))
              (setq u1 (- (car a1e) (caddr a1e))
                    u2 (+ (cadr a1e) (car (cdddr a1e))))
              (cond ((or (null ud) (< (car ud) (cadr a1e))
                                   (< u2 (car ud)))
                       (cond ((or (null ug) (< (car ug) u1)
                                            (< (car a1e) (car ug))) nil)
                             (t (mu (car ug) a1e))))
                    (t (cond ((or (null ug) (< (car ug) u1)
                                            (< (car a1e) (car ug)))
                               (mu (car ud) a1e))
                             (t (min (mu (car ug) a1e)
                                     (mu (car ud) a1e)))))))))
                                        ;devuelve NIL si a01 esta incluido en
                                        ;a1, o el menor grado de interseccion
                                        ;devuelve tambien el valor de one
```

```
(DEFUN dint3 (a a0)
        (append (dinterd3 a a0) (dinterd3 a (reverse a0))))


(DEFUN dinterd3 (a a0)
  (let ((x nil)
        (x1 nil))
       (cond ((null (cdr a0)) nil)
             ((< (setq x (car (cdadr a0))) (mu (caadr a0) a))
               (cond ((< (setq x1 (cadar a0)) (mu (caar a0) a)) nil)
                     ((list x1))))
             ((< x 1.0) (dinterd3 a (cdr a0))))))


(DEFUN abs-int (a da b db)
       (cond ((igual da db) nil)
             ((list (/ (- (* a db )(* b da)) (float (- db da)))))))
                                     ;devuelve la abscisa de la interseccion
                                     ;de dos segmentos de rectas


(DEFUN dint2 (a a0)
  (append (dinterd2 a a0) (dinterd2 a (reverse a0))))


(DEFUN dinterd2 (a a0)
  (let ((x nil)
        (y nil))
       (cond ((null a0) nil)
             ((< (setq x (cadar a0))(setq y (mu (caar a0) a)))
                (cond ((< y 1.0)  (dinterd2 a (cdr a0)))))
             ((< x 1.0) (list y)))))


(DEFUN assq (x y)
   (cond ((null y) nil)
         ((equal x (caar y)) (car y))
         ((assq x (cdr y)))))
                                        ;devuelve el grado de pertenencia de x
                                        ;en y, con y discretizado




(DEFUN assq2 (x y)
   (cond ((null y) nil)
         ((equal x (cadar y)) (caar y))
         ((assq2 x (cdr y)))))
                                        ;devuelve la primera abscisa donde y,
                                        ;discretizada, toma el valor x


(DEFUN discr (a1 na)
   (let ((y nil)
         (t1 (+ ad1 (* na pasa1))))
        (cond ((eq na -1) nil)     ;uso eq
              ((zerop (setq y (mu t1 a1))) (discr a1 (- na 1)))
              ((append (discr a1 (- na 1)) (list (cons t1 (list y))))))))
                                        ;discretiza la distribucion a1


(DEFUN mapct (f l)
  (let ((val nil))
       (cond ((null l) nil)
             ((null (setq val (eval (list f '(car l)))))
                (mapct f (cdr l)))
             (t (append val (mapct f (cdr l)))))))
```

```
;******************   BASE - 1   ******************

(SETQ R1 '(   ((FORMACION CAND OK) (SCORE-TEST CAND ACEPTABLE))
               (APT-FUNDAM CAND OK 1.0 1.0)   ) )

(SETQ R2 '(   ((TEST-INT CAND 10A14) (TEST-ING CAND 6A9))
               (SCORE-TEST CAND PASABLE)))

(SETQ R3 '(   ((TEST-INT CAND 14A19) (TEST-ING CAND 6A9))
               (SCORE-TEST CAND BUENO)))

(SETQ R4 '(   ((TEST-INT CAND 10A14) (TEST-ING CAND 9A19))
               (SCORE-TEST CAND SATISFACTORIO)))

(SETQ R5 '(   ((TEST-INT CAND 14A19) (TEST-ING CAND 9A19))
               (SCORE-TEST CAND EXCELENTE)))

(SETQ R6 '(   ((EXPER. CAND OK) (ADAPTAB CAND OK))
               (APT-FUNDAM CAND OK .7 .2)))

(SETQ R7 '(   ((APT-FUNDAM CAND OK))
               (CALIDAD CAND OK 1.0 1.0)))

(SETQ R8 '(   ((APT-MAN CAND OK))
               (APT-ESPEC CAND OK .6 0.0)))

(SETQ R9 '(   ((APT-DEP CAND OK))
               (APT-ESPEC CAND OK .7 .5)))

(SETQ R10 '(   ((APT-ART CAND OK))
                (APT-ESPEC CAND OK .8 0.0)))

(SETQ R11 '(   ((APT-ESPEC CAND OK))
                (CALIDAD CAND OK .7 0.0)))

(SETQ R12 '(   ((ORAL-ING CAND 6A9) (ESCRIT-ING CAND ACEPTABLE))
                (TEST-ING CAND PASABLE)))

(SETQ R13 '(   ((ORAL-ING CAND ACEPTABLE) (ESCRIT-ING CAND ACEPTABLE))
                (TEST-ING CAND BUENO)))

(SETQ R14 '(   ((ORAL-ING CAND ACEPTABLE) (ESCRIT-ING CAND 6A9))
                (TEST-ING CAND SATISFACTORIO)))



(SETQ REGLAS '(R1 R2 R3 R4 R5 R6 R7 R8 R9 R10 R11 R12 R13 R14))


(SETQ EXCELENTE '(18 19 1 1))
(SETQ ACEPTABLE '(13 19 3 1))
(SETQ PASABLE '(11 12 3 1))
(SETQ BUENO '(15 16 2 1))
(SETQ SATISFACTORIO '(12.5 14 1.5 1))
(SETQ 7A9 '(7 9 1 1))
(SETQ BASTANTE-BUENO '(14 15 1 1))
(SETQ 10A14 '(10 14 1 1))
(SETQ 14A19 '(14 19 2 1))
(SETQ 6A9 '(6 9 2 2))
(SETQ 9A19 '(9 19 1 1))
(SETQ 11A14 '(11 14 1.5 1))
```

```
(SETF (GET 'NPUNTO 'SCORE-TEST) 30
      (GET 'NPUNTO 'TEST-ING) 30
      (GET 'NPUNTO 'PASABLE) 30
      (GET 'NPUNTO 'BUENO) 30
      (GET 'NPUNTO 'SATISFACTORIO) 30
      (GET 'NPUNTO 'EXCELENTE) 30)

(SETF (GET 'UNIVERSO 'SCORE-TEST) '(0 20)
      (GET 'UNIVERSO 'TEST-ING) '(0 20)
      (GET 'UNIVERSO 'PASABLE) '(0 20)
      (GET 'UNIVERSO 'BUENO) '(0 20)
      (GET 'UNIVERSO 'SATISFACTORIO) '(0 20)
      (GET 'UNIVERSO 'EXCELENTE) '(0 20))


(SETF (GET 'CERT 'F1) '(0.8 1.0)
      (GET 'CERT 'F5) '(1.0 1.0)
      (GET 'CERT 'F6) '(0.0 0.6)
      (GET 'CERT 'F7) '(0.0 0.2)
      (GET 'CERT 'F8) '(1.0 1.0) )


(SETQ F1 '(FORMACION CAND OK))

(SETQ F2 '(TEST-INT CAND 11A14))

(SETQ F3 '(ORAL-ING CAND  BASTANTE-BUENO))

(SETQ F4 '(ESCRIT-ING  CAND 7A9))

(SETQ F5 '(EXPER. CAND OK))

(SETQ F6 '(ADAPTAB CAND OK))

(SETQ F7 '(APT-ART CAND OK))

(SETQ F8 '(APT-DEP CAND OK))


(SETQ BH '(F1 F2 F3 F4 F5 F6 F7 F8))
```

```
;********************** EJECUCION CON BASE-1 Y MOTOR-1 ************************




(VALOR '(CALIDAD CAND OK))
(0.64 1.0)
* BH

(F1 F2 F3 F4 F5 F6 F7 F8 #:H16 #:H17 #:H18 #:D19 #:D21)
* (IMPRIME-HECHOS)
#:H16
(APT-ESPEC CAND OK)
(0.7 1.0)
#:H17
(CALIDAD CAND OK)
(0.64 1.0)
#:H18
(APT-FUNDAM CAND OK)
(0.64 1.0)
#:D19
(SCORE-TEST CAND #:-20)
((11.3333 2.22221F-01) (12.0 6.66666F-01) (12.6667 1.0) (13.3333 1.0) (14.0 1.0)
 (14.6667 3.33335F-01))
#:D21
(TEST-ING CAND #:-22)
((11.3333 2.22221F-01) (12.0 6.66666F-01) (12.6667 1.0) (13.3333 1.0) (14.0 1.0)
 (14.6667 3.33335F-01))
NIL
*
```

```
                ;*****    M O T O R - 2    *****

(DEFUN MOTOR2()
    (LOAD 'jinicio)
    (LOAD 'base2)
    (inicio)
    (LOAD 'nuevond)
    (LOAD 'jmot-dif))

;*****   FUNCION DE EJECUCION PARA LA PARTE NO DIFUSA    *****

(DEFUN valor (parametros)
 (probar parametros (get 'rg-atr (car parametros))))

;Funcion maestra de ejecucion del programa

(DEFUN probar (hipotesis rg-concl)
 (LET ((flag NIL)
       (nh NIL))
  (COND ((NOT (equal (caddr hipotesis) 'OK))
                        (conversion2 (difusa (GET 'he-atr (car hipotesis)) )))
                                ;la hipotesis es difusa: tratamiento difuso
        (T                      ;la hipotesis es no difusa

          (COND ((esta-base-nd (GET 'he-atr (car hipotesis))))
                                ;comprueba si esta en la base
                                ;devuelve el valor de verdad si esta
                                ;o NIL si no esta

                ((null rg-concl) '(-1))
                                ;si no estaba en la base y ninguna regla
                                ;concluye en ella devuelve total ignorancia

                ((encadenamiento rg-concl)
                         (setf (get 'cert (setq nh (hech-num (get 'he-atr
                                 (car hipotesis))))) (comb)))
                                ;intenta realizar el encadenamiento
                                ;hacia tras

                (T  '(-1)))))))
                                ;si no ha podido realizar el encadenamiento
                                ;devuelve total ignorancia



(DEFUN hech-num (l)
  (COND ((null l) (print 'a))
        ((igual (eval (car l)) hipotesis) (car l))
        (T (hech-num (cdr l)))))
                              ;encuentra el numero de un hecho

(DEFUN comb()
 (LET ((negposit (trinegposit
                   (mapcar 'aux (get 'proviene nh)) nil nil)))
  (COND ((null (car negposit)) (conversion2
                   (list (apply 'max (cadr negposit)) 1.0)))
        ((null (cadr negposit)) (conversion2
                           (list 0.0 (apply 'min (car negposit)))))
        (T (conversion2 (normalizar (apply 'max (cadr negposit))
                      (apply 'min (car negposit)))))))
                              ;combina, sin reforzar, valores de verdad
                              ;calificando un mismo hecho inferido por
                              ;varias reglas. Usa trinegposit


(DEFUN trinegposit (li neg posit)
  (COND ((null li) (list neg posit))
        (( > (caar li) 0.0) (trinegposit (cdr li) neg (cons (caar li) posit)))
        (T (trinegposit (cdr li) (cons (cadar li) neg) posit))))
```

```
(DEFUN esta-base-nd (hech)
  (COND ((null hech) nil)           ;la hipotesis no difusa no esta en la base

        ((igual (eval (car hech)) hipotesis) (GET 'cert (car hech) ))
                                    ;si esta en la base devuelve el valor
                                    ;de verdad

        (T (esta-base-nd (cdr hech)))))
                                    ;sigue comprobando los demas hechos de la
                                    ;base con el mismo atributo


(DEFUN normalizar (nec pos)
  (LET ((d (max (- 1.0 nec) pos)))
    (COND ((= d pos) (list (redondeo2 (/ (- (+ d nec) 1.0) d)) 1.0))
          (T (list 0.0 (redondeo2 (/ pos d)))))))
     ;normaliza un intervalo del tipo (nec pos)


(DEFUN redondeo2 (n)
  (/ (round (* n 100)) 100))
        ;deja solo dos cifras decimales a un numero real


(DEFUN round (num)
  (COND ((> (- num (truncate num)) 0.5) (+ (truncate num) 1))
        (t (truncate num))))


(DEFUN encadenamiento (reg)         ;reg: reglas que concluyen a la hipotesis
                                    ;solo para hipotesis no difusas
  (COND ((null reg) flag)
        ((igual hipotesis (menos2-ult (cadr (eval (car reg)))))
                                    ;si la hipotesis coincide con la conclusion
                                    ;de una regla ==>

             (COND ((ensayar (car reg)) (setq flag T)
                                        (encadenamiento (cdr reg)))
                                    ;intenta demostrar las premisas de esa regla

                   (T (encadenamiento (cdr reg)))))
        (T (encadenamiento (cdr reg)))))
                       ;devuelve T si ha podido aplicar alguna regla
                       ;o NIL si no ha podido


(DEFUN ensayar (nrg)                ;nrg: una regla a intentar aplicar
 (LET ((x (eval nrg)) (cert-prem nil))
  (COND ((igual (setq cert-prem (int-cert-prem (car x))) '(-1)) nil)
        (T (accion (cadr x) cert-prem)))))
                                    ;si sobre las premisas de nrg tiene
                                    ;algun conocimiento pasa a ejecutarla


(DEFUN int-cert-prem (r-prem)       ;r-prem: conjunto de premisas de una regla
  (LET ((i (probar (car r-prem) (GET 'rg-atr (caar r-prem) )))
        (i1 nil))
    (COND ((null (cdr r-prem)) i)
          (T (list  (min (car i) (car (setq i1 (int-cert-prem (cdr r-prem)))))
                 )))))
                                ;devuelve el minimo valor de verdad
                                ;de los de todas las premisas


        ;encadenamiento, ensayar e int-cert-prem intentan
         ;demostrar las premisas de reglas que concluyen
                       ;en la hipotesis
```

```
(DEFUN accion (acto vp)
  (COND ((hecho-en-base (GET 'he-atr (car acto) ) vp) T)
        (T (anade vp)) T))

                                ;ejecucion del consecuente de una regla
                                ;si esta en bh --> hecho-en-base
                                ;si no esta lo anade -->anade


(DEFUN hecho-en-base (hech vp)
 (LET ((h (menos2-ult acto)))
  (COND ((null hech) nil)
        ((igual (eval (car hech)) h)
               (setf (get 'proviene (car hech)) (append (get 'proviene
                     (car hech)) (list (cons nrg (infer vp))))))
        (T (hecho-en-base (cdr hech) vp)))))
                                ;devuelve nil si el hecho conclusion de la regla
                                ;no esta en la base, si no,
                                ;pone en la propiedad 'proviene de ese mismo
                                ;hecho de la base la regla nrg de la cual se
                                ;puede obtener junto al valor con que se infiere


(DEFUN menos2-ult (l)           ;devuelve la conclusion de la regla sin los
 (butlast (butlast l)))         ;valores de suficiencia y necesidad


(DEFUN anade (vp)
 (setq x (gensym 'h))                        ;da nombre al nuevo hecho
 (set x (menos2-ult acto))
 (let    ((i nil))
  (setq i (infer vp))                        ;calcula el valor inferido
  (setf (get 'proviene x) (list (cons nrg i)))

                                             ;inicializa la propiedad 'proviene
                                             ;de ese hecho con la regla que se
                                             ;obtiene y el valor inferido
  (setq bh (append bh (list x)))             ;lo anade a la bh
  (setf (get 'he-atr (car acto))(append (get 'he-atr (car acto)) (list x)))))
                                             ;aumenta la lista de hechos
                                             ;asociada al atributo del hecho
                                             ;que ha creado

(DEFUN infer (vp)
  (LET* ((z (reverse acto))
         (suf (cadr z))
         (nec (car z))
         (i nil))
    (COND ((AND (< nec O) (< suf O))
                (list (- 1 (cadr (setq i (calinfer (- suf) (- nec) vp))))
                      (- 1 (car i))))
          (T (calinfer suf nec vp)))))
                ;devuelve el valor de incertidumbre del hecho inferido
                ;ayudado por la funcion calinfer y el valor cert-prem
                ;que se refiere al valor de incertidumbre de las
                ;premisas de la regla que produce este hecho


(DEFUN calinfer (suf nec vp)
        (cond ((>= (car vp) O) (list (redondeo2 (apply *norma-1 (list suf
                                       (car vp))))))
              (t (list (redondeo2 (apply *norma-2 (list nec (car vp))))))))


(DEFUN igual (a b) (equal a b))

(DEFUN conversion1 (v)
        (cond ((>= (car v) O) (list (car v) 1))
              (t         (list O (+ (car v) 1)))))

(DEFUN conversion2 (l)
        (list (- (+ (car l) (cadr l)) 1)))


(defun aux (l)
        (conversion1 (cdr l)))
```

```
;*****   BASE - 2   *****

(SETQ R1 '(  ((FORMACION CAND OK) (SCORE-TEST CAND ACEPTABLE))
             (APT-FUNDAM CAND OK 1.0 1.0)   ) )

(SETQ R2 '(  ((TEST-INT CAND 10A14) (TEST-ING CAND 6A9))
             (SCORE-TEST CAND PASABLE)))

(SETQ R3 '(  ((TEST-INT CAND 14A19) (TEST-ING CAND 6A9))
             (SCORE-TEST CAND BUENO)))

(SETQ R4 '(  ((TEST-INT CAND 10A14) (TEST-ING CAND 9A19))
             (SCORE-TEST CAND SATISFACTORIO)))

(SETQ R5 '(  ((TEST-INT CAND 14A19) (TEST-ING CAND 9A19))
             (SCORE-TEST CAND EXCELENTE)))

(SETQ R6 '(  ((EXPER. CAND OK) (ADAPTAB CAND OK))
             (APT-FUNDAM CAND OK .7 .2)))

(SETQ R7 '(  ((APT-FUNDAM CAND OK))
             (CALIDAD CAND OK 1.0 1.0)))

(SETQ R8 '(  ((APT-MAN CAND OK))
             (APT-ESPEC CAND OK .6 0.0)))

(SETQ R9 '(  ((APT-DEP CAND OK))
             (APT-ESPEC CAND OK .7 .5)))

(SETQ R10 '(  ((APT-ART CAND OK))
              (APT-ESPEC CAND OK .8 0.0)))

(SETQ R11 '(  ((APT-ESPEC CAND OK))
              (CALIDAD CAND OK .7 0.0)))

(SETQ R12 '(  ((ORAL-ING CAND 6A9) (ESCRIT-ING CAND ACEPTABLE))
              (TEST-ING CAND PASABLE)))

(SETQ R13 '(  ((ORAL-ING CAND ACEPTABLE) (ESCRIT-ING CAND ACEPTABLE))
              (TEST-ING CAND BUENO)))

(SETQ R14 '(  ((ORAL-ING CAND ACEPTABLE) (ESCRIT-ING CAND 6A9))
              (TEST-ING CAND SATISFACTORIO)))



(SETQ REGLAS '(R1 R2 R3 R4 R5 R6 R7 R8 R9 R10 R11 R12 R13 R14))


(SETQ EXCELENTE '(18 19 1 1))
(SETQ ACEPTABLE '(13 19 3 1))
(SETQ PASABLE '(11 12 3 1))
(SETQ BUENO '(15 16 2 1))
(SETQ SATISFACTORIO '(12.5 14 1.5 1))
(SETQ 7A9 '(7 9 1 1))
(SETQ BASTANTE-BUENO '(14 15 1 1))
(SETQ 10A14 '(10 14 1 1))
(SETQ 14A19 '(14 19 2 1))
(SETQ 6A9 '(6 9 2 2))
(SETQ 9A19 '(9 19 1 1))
(SETQ 11A14 '(11 14 1.5 1))
```

```
(SETQ F1 '(FORMACION CAND OK))

(SETQ F2 '(TEST-INT CAND 11A14))

(SETQ F3 '(ORAL-ING CAND  BASTANTE-BUENO))

(SETQ F4 '(ESCRIT-ING  CAND 7A9))

(SETQ F5 '(EXPER. CAND OK))

(SETQ F6 '(ADAPTAB CAND OK))

(SETQ F7 '(APT-ART CAND OK))

(SETQ F8 '(APT-DEP CAND OK))


(SETQ BH '(F1 F2 F3 F4 F5 F6 F7 F8))


(SETF (GET 'NPUNTO 'SCORE-TEST) 30
      (GET 'NPUNTO 'TEST-ING) 30
      (GET 'NPUNTO 'PASABLE) 30
      (GET 'NPUNTO 'BUENO) 30
      (GET 'NPUNTO 'SATISFACTORIO) 30
      (GET 'NPUNTO 'EXCELENTE) 30)

(SETF (GET 'UNIVERSO 'SCORE-TEST) '(0 20)
      (GET 'UNIVERSO 'TEST-ING) '(0 20)
      (GET 'UNIVERSO 'PASABLE) '(0 20)
      (GET 'UNIVERSO 'BUENO) '(0 20)
      (GET 'UNIVERSO 'SATISFACTORIO) '(0 20)
      (GET 'UNIVERSO 'EXCELENTE) '(0 20))


(SETF (GET 'CERT 'F1) '(0.8)
      (GET 'CERT 'F5) '(1.0)
      (GET 'CERT 'F6) '(-0.4)
      (GET 'CERT 'F7) '(-0.8)
      (GET 'CERT 'F8) '(1.0) )
```

```
;*********************** EJECUCION CON BASE-2 Y MOTOR-2 ***********************

(VALOR '(CALIDAD CAND OK))
(0.65)
* BH

(F1 F2 F3 F4 F5 F6 F7 F8 #:H2 #:H3 #:H4 #:D5 #:D7)
* (IMPRIME-HECHOS)
#:H2
(APT-ESPEC CAND OK)
(0.7)
#:H3
(CALIDAD CAND OK)
(0.65)
#:H4
(APT-FUNDAM CAND OK)
(0.65)
#:D5
(SCORE-TEST CAND #:-6)
((11.3333 2.22221F-01) (12.0 6.66666F-01) (12.6667 1.0) (13.3333 1.0) (14.0 1.0)
 (14.6667 3.33335F-01))
#:D7
(TEST-ING CAND #:-8)
((11.3333 2.22221F-01) (12.0 6.66666F-01) (12.6667 1.0) (13.3333 1.0) (14.0 1.0)
 (14.6667 3.33335F-01))
NIL
*
```