

Cuadernos CDTI

Centro para el Desarrollo  
Tecnológico Industrial.  
Ministerio  
de Industria y Energía.

**Sistemas Expertos:  
Consideraciones Teóricas  
y Sugerencias de uso**



Cuadernos CDTI

Centro para el Desarrollo  
Tecnológico Industrial.  
Ministerio  
de Industria y Energía.

**Sistemas  
Expertos:  
Consideraciones  
Teóricas y  
Sugerencias de uso**

Diseño Gráfico:  
Free Lance García de Paredes/Amorós.  
Fotocomposición: Amoretti  
Impresión: EPES - Ind. Gráf., S.L. -  
Alcobendas (Madrid)  
Depósito Legal: M-8244 - 1985

**CDTI Marzo 1985**

## INTRODUCCION

### SISTEMAS EXPERTOS

El objetivo de esta publicación es ofrecer a las empresas, fundamentalmente, y a las personas interesadas en las nuevas tecnologías de la información, una visión más amplia que profunda de un tema de máxima actualidad.

Este Cuaderno no intenta ser un texto sobre teoría de Sistemas Expertos, sin embargo sí se ha creído plantear con cierto detalle los fundamentos teóricos de la llamada inteligencia artificial. Así se presentan los desarrollos matemáticos detallados en "letra pequeña" para que los lectores más enterados en los procesos algebraicos puedan disfrutar de una lectura más precisa.

El lector que busque información aplicable a una empresa o trabajo va a encontrar referencias a la ingeniería, la medicina y los negocios.

En cualquier caso, no se pretende que se considere este Cuaderno como un librito de "software", sino que su lectura lleva a comprender la necesidad de explotar de otra forma las bases de conocimiento mediante técnicas informáticas.

La dificultad de pasar a las máquinas el conocimiento humano es una tarea compleja y difícil de expresar de forma algorítmica. Hay que tener en cuenta que el hombre se mueve fundamentalmente por símbolos y recibe continuamente información de forma simbólica y no algorítmica.

El futuro tecnológico informático se desarrollará cada día más, sobre la explotación de bases de conocimiento. Hay que pasar a las máquinas la experiencia acumulada en años de actividad profesional. Las grandes compañías, los seguros, la banca, la medicina, la administración, etc., tienen magníficos expertos cuya acumulación de conocimientos y experiencias no deben de acabar cuando la vida activa de estas personas termina. Se nos abre un nuevo horizonte tecnológico. Hay que aprovecharlo y estudiarlo.

RAFAEL LAMAS  
Subdirector de CDTI

<b>LOS SISTEMAS EXPERTOS: CONCEPTO, REALIZACIONES Y TECNICAS DE CONSTRUCCION</b>	6	<b>LOS SISTEMAS EXPERTOS EN INGENIERIA</b>	49
<b>0. Introducción.....</b>	6	<b>1. Introducción.....</b>	49
<b>1. Definiciones.....</b>	7	<b>2. Ampliaciones al diseño ..</b>	49
<b>2. Antecedentes históricos</b>	8	Concepto.....	49
<b>3. Las realizaciones más ca- racterísticas.....</b>	10	El enfoque clásico del diseño ..	49
<b>4. Panorama de las técnicas de construcción de Sistemas Expertos.....</b>	12	Los sistemas basados en el co- nocimiento para apoyo del di- seño.....	50
Lógica de primer orden	12	<b>3. Aplicaciones a la explo- tación.....</b>	52
Redes semánticas.....	14	Concepto.....	52
Estructuras.....	14	El enfoque actual.....	52
Procedimientos inferenciales...	19	Los sistemas basados en el co- nocimiento de ayuda a la explo- tación.....	53
Extensiones, particiones y es- cenas (frames).....	19	<b>4. Conclusiones.....</b>	53
Redes inferenciales.....	20	<b>5. Referencias.....</b>	53
Método MYCIN.....	21	<b>LOS SISTEMAS EXPERTOS EN MEDICINA</b>	55
Definiciones.....	21	Introducción.....	55
Consistencia.....	23	Características del proceso de diagnóstico médico.....	56
Acumulación de evidencia.....	23	Sistemas clásicos.....	59
Encadenamiento de niveles de certidumbre.....	24	Evolución actual.....	62
El Método PROSPECTOR.....	25	Ingeniería para el desarrollo de Sistemas IA en Medicina.....	65
Definiciones.....	25	Expectativas.....	65
Acumulación de evidencia.....	27	Recursos humanos.....	67
Consideraciones sobre los pro- blemas de consistencia.....	30	Sistemas expertos en Medicina	69
Formulación modificada de PROSPECTOR.....	33	Sistemas Expertos en áreas pa- ramédicas y de investigación biológica.....	69
Técnicas avanzadas de repre- sentación.....	34	Ayudas para el desarrollo de Sistemas Expertos en Medicina	69
El Método DEMPSTER-SHA- FER.....	34	Bibliografía.....	73
La lógica difusa. El Método PRUF.....	38	<b>CONSIDERACIONES GENERALES EN EL CAMPO DE LOS NEGOCIOS</b>	75
Semántica difusa de fórmulas..	38		
Composición de significados...	39		
Inferencia.....	41		
<b>5. La construcción de Siste- mas Expertos.....</b>	41		
Entornos generales de Logical ..	42		
El Lenguaje LISP.....	42		
El Lenguaje PROLOG.....	43		
Entornos específicos.....	44		
El Sistema OPS 5.....	45		
El Sistema EMYCIN.....	46		
Metodología general.....	46		
<b>6. Referencias.....</b>	47		



# **LOS SISTEMAS EXPERTOS: CONCEPTO, REALIZACIONES Y TECNICAS DE CONSTRUCCION**

**J. Cuena BARTOLOME**

Profesor de Lógica e Inteligencia Artificial  
Facultad de Informática  
Universidad Politécnica  
Madrid

## **0. INTRODUCCION**

Los Sistemas Expertos, producto de los desarrollos en la Inteligencia Artificial en la década de los setenta, constituyen un nuevo concepto estructural de las aplicaciones informáticas que hace posible el desarrollo de proyectos que no eran fácilmente viables con la estructuración tradicional de algoritmos y datos.

Este artículo pretende dar un panorama de los aspectos técnicos y metodológicos de este tipo de sistemas; para ello se ha organizado en:

- Una primera parte descriptiva de los conceptos básicos y la evolución histórica conducente a la introducción de los Sistemas Expertos.
- Una segunda parte dedicada al tratamiento de los aspectos técnicos que se ha estructurado en dos niveles, incluyendo los aspectos más densos con tipo distinto de letra.
- Una última parte donde se presentan unas ideas generales sobre los entornos de software y metodología de construcción, aspectos, debido a lo reciente del tema, todavía incipientes. Finalmente, se presenta una bibliografía comentada.

## I. DEFINICIONES

Un sistema experto es un sistema informático:

— que representa el conocimiento de una persona experta sobre un tema de forma accesible y modificable.

— que utiliza esta representación mediante un proceso general para:

- producir respuestas sobre los temas objeto del sistema.
- explicar estas respuestas a demanda.

Por tanto, un sistema experto no es simplemente un programa convencional que responde preguntas sobre un tema, ya que su estructura le da capacidad para justificar las respuestas y para modificar el conocimiento que utiliza para producir estas respuestas.

Así, por ejemplo, un programa convencional de diagnóstico médico podría redactarse en la forma:

---

```
if temperatura  $\neq$  36,5  
then if temperatura > 35
```

---

```
    then escribir "normal" y explicar  
    else if globulos blancos < 5000
```

---

```
        then escribir "grave" y explicar  
        else if globulos rojos > 4M
```

---

```
else escribir "leve" y explicar.
```

---

Este programa está orientado a responder una pregunta. En el caso de que los conceptos utilizados en él pudieran usarse para contestar otras preguntas, sería preciso programar la obtención de estas contestaciones en otro proceso adecuado.

También en el caso de modificar los conceptos, se requiere modificar todo el programa.

Por otro lado, la forma de razonar se apoya exclusivamente en la lógica clásica, sin tener en cuenta que en muchos casos la opción entre alternativas no es tan clara como en este tipo de lógicas.

Resulta, por tanto, de interés utilizar una estructura fija de representación del conocimiento sobre la que pueda actuarse, para hacer modificaciones, y un programa general independiente de obtención de respuestas y explicación de las mismas, que no sea preciso modificar cada vez que se quiere añadir una nueva respuesta posible o cambiar los conceptos de análisis y que, por otro lado, pueda incorporar formas no clásicas de razonamiento, más próximas a la forma de hacer conjeturas en el razonamiento natural. De acuerdo con los conceptos anteriores, puede definirse en forma más técnica, un sistema experto por su arquitectura, basada en dos elementos fundamentales:

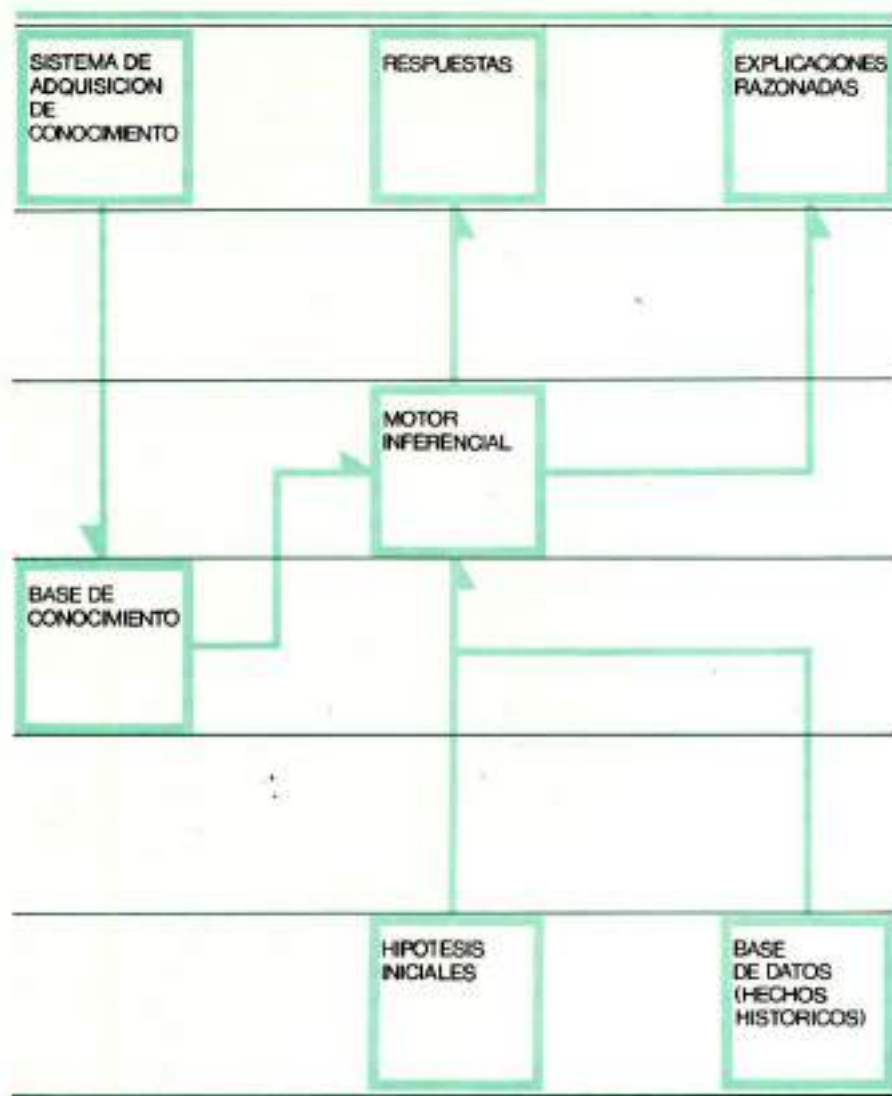
— **Base de conocimiento** que incorpora una representación de los hechos conocidos y deducibles, descriptivos del tema.

— **Motor inferencial**, procedimiento capaz de responder y explicar preguntas, apoyado en la estructura de la base de conocimiento.

Asociado con ambos elementos hay un subsistema de consulta y modificación de la base de conocimiento, habitualmente planteado con interfaz en **lenguaje natural**, de forma que sea posible a los expertos comunicarse con la base, para su permanente control y mejora.

Todo este conjunto puede articularse con una base de datos desde la que se obtienen los hechos básicos, premisa de las conjeturas razonadas del sistema. Por ejemplo, en el caso del sistema de diagnóstico, la base de datos estaría constituida por las características de los enfermos, que podrían obtenerse automáticamente de archivos o por sensores, operando en tiempo real sobre el enfermo, o bien, proporcionarse directamente por el usuario mediante un sistema convencional de consulta.

El conjunto se resume en el gráfico adjunto:



Esta arquitectura de sistema se denomina más propiamente, Sistema Basado en el Conocimiento (SBC), (KBS: Knowledge Based System), denominándose "Sistema Experto" cuando incorpora precisamente el conocimiento de una persona experimentada. Sin embargo, como arquitectura informática, puede emplearse para el tratamiento de problemas tradicionales y constituye una alternativa de metodología de programación y computación que, como se verá más adelante, constituye la base conceptual de los ordenadores de la quinta generación.

## 2. ANTECEDENTES HISTORICOS

El concepto de sistemas expertos fue formulado por primera vez por Feigenbaum en el congreso mundial de Inteligencia Artificial de 1977 (IJCAI 77); sin embargo, la producción de este tipo de sistemas arranca de los primeros tiempos de la Inteligencia Artificial, en donde se planteaban técnicas heurísticas para resolución de problemas; en ellos, el conjunto de heurísticas para resolución de una clase de problemas, constituía ya una base de conocimiento, así como los procesos generales de búsqueda pueden considerarse motores inferenciales.

Se consideran tres etapas en el desarrollo de los sistemas expertos:

— Etapa de invención, en la que la complejidad e incrementalidad de los heurísticos desarrollados para algunos sistemas empieza a darles personalidad como bases de conocimiento; tal es el caso de los sistemas DENDRAL y MACSYMA. Esta etapa puede estimarse entre 1965 y 1970.

— Etapa de prototipos, en que una vez constatado el concepto se producen desarrollos según distintos criterios de estructuración y proceso del conocimiento. Los sistemas significativos de esta etapa son CASNET, MYCIN, HEARSAY II y PROSPECTOR. El período de la misma puede fijarse entre 1970 y 1977 en que se produce la presentación de Feigenbaum y se pu-



blican los distintos sistemas (MYCIN se publica en 1975 y 76, PROSPECTOR en 1976 y 1978).

— Etapa de experimentación, en la que los conceptos de los prototipos se aplican a distintas áreas (las ideas de CASNET se utilizan en INTERNIST, las de MYCIN en PUF, F GUIDON, SACON, las de PROSPECTOR en HYDRO, las de HEARSAY-II en HEARSAY III y HASP). En esta etapa, cuyo discurso puede establecerse entre 1977 y 1981, se producen, además, los primeros intentos de formalización del proceso de construcción de sistemas expertos, mediante sistemas abstractos. Así se producen derivados de DENDRAL, el sistema METADENDRAL y de MYCIN, los sistemas TEIRESIAS, EMYCIN, RITA y ROSIE, derivados de PROSPECTOR el sistema KAS y derivado de HEARSAY II, el sistema AGE.

— Etapa de industrialización, en la que comienzan a surgir empresas dedicadas al diseño de sistemas basados en el conocimiento y se empieza a producir una industrialización de los instrumentos (así de EMYCIN se derivan K S300 y SYSTEM I). La serie de sistemas OPS, producidos en CARNEGIE MELLON para el tratamiento de reglas de producción, presentan versiones operativas a partir de OPS-5.

Esta etapa, que es la actual en los países más avanzados, puede caracterizarse por los dos sistemas: RI, de configuración de equipos DIGITAL y DRILLING ADVISOR, producido por ELFAQUITAINE para interpretación de sondeos.

Como ya se ha indicado, en esta etapa aparecen nuevas entidades de desarrollo y aplicación de estas técnicas que inicialmente se centraban en el MIT, universidades de PITTSBURG y CARNEGIE MELLON, universidad de STANFORD (Computer Science Department, Heuristic Programming Project), Stanford Research Institute (SRI) y Rand Corporation. En esta etapa, investigadores procedentes de estos centros crean grupos consultores (Teknowledge, Syntelligence, the Carnegie

group...) y otras entidades adquieren mayor fuerza (Fairchild, universidad de Rutgers, laboratorio de Xerox...) a la par que organizaciones clásicas crean secciones de inteligencia artificial, de facto dedicadas a sistemas basados en el conocimiento, de manera que en léxico profesional informático aparece una nueva especialidad: **ingeniero de conocimiento**.

Estas etapas, como puede verse, ilustran una evolución localizada en EE.UU. En Europa hasta primeros de los setenta hubo una línea de investigación en inteligencia artificial, basada fundamentalmente en la universidad de Edimburgo. Sin embargo, al llegar la crisis económica en 1973, el informe Lighthill dio lugar a que se redujeran considerablemente los fondos en este área de investigación durante la década de los setenta, lo que ha dado lugar a un retraso importante respecto de América, precisamente del centro que más podría irradiar esta tecnología en Europa. Las razones exhibidas en el informe Lighthill para desautorizar las inversiones en inteligencia artificial, se basaban en la tradición académica-científica europea, empeñada en valorar los temas de inteligencia artificial, como si fueran pasos destinados a construir una teoría universal matemática de la inteligencia, a la que referir los problemas de representación del pensamiento humano, sueño ya acariciado por Leibniz y a primeros de este siglo por Hilbert; como, desde un punto de vista científico, Godel comprobó que no era factible formalizar, en lógica de primer orden, las teorías científicas (proyecto científico de los treinta primeros años de siglo) se juzgó que la investigación en inteligencia artificial era volver a repetir, envuelto en los nuevos conceptos informáticos, una aventura ya fallida.

Sin embargo, la manera de valorar los proyectos de inteligencia artificial por los investigadores americanos, creadores del concepto en 1956, era distinta; buscaban la definición de conceptos y métodos para proyectar sis-

temas inteligentes; es decir, sistemas que en un determinado tema tuvieran un comportamiento que pudiera considerarse inteligente. No se trata, por tanto, de elaborar una teoría general de la inteligencia, sino de crear bases teóricas y técnicas para construir inteligencias concretas y esto sí que es factible, como lo demuestra la experiencia americana.

El freno a la investigación en inteligencia artificial, en el Reino Unido, hace que la única aportación europea en este campo, en la década de los setenta, proceda de la universidad de Marsella, que construyó el primer lenguaje de programación lógica, PROLOG, vehículo bastante adecuado para la formulación e interpretación de un tipo de bases de conocimiento. Este lenguaje se tratará brevemente en el apartado 5.

No obstante, la constatación a primeros de los ochenta, del avance americano (conferencia mundial de Vancouver 1981) ha relanzado la actividad en Europa, creándose sociedades nacionales de coordinación y promoción de esta actividad, ya considerada industrial por el valor añadido que puede aportar\*, así como un comité europeo de coordinación de estas sociedades nacionales (ECCA), European Coordination Committee for Artificial Intelligence). En el momento actual existen ya grupos de investigación y un proyecto europeo (ESPRIT) que da una importancia preferente a los temas de inteligencia artificial, notablemente a los sistemas basados en el conocimiento.

Asimismo, empiezan a crearse, en Europa también, empresas de servicios en este área de proyectos informáticos, algunas de ellas filiales de los grupos americanos.

\* En España se ha constituido recientemente AEPiA (Asociación Española para la Inteligencia Artificial).



### 3. LAS REALIZACIONES MAS CARACTERISTICAS

— El sistema MACSYMA, (Martín y Fateman 1971), que es uno de los sistemas especializados en cálculo integral y diferencial iniciado con SAINT (Slagle 1961) y seguido posteriormente por SIN (Moses 1967). MACSYMA incorpora un conjunto de reglas de manipulación simbólica de expresiones matemáticas, que constituyen la acumulación de toda experiencia en manejo de los anteriores sistemas y en el propio MACSYMA. El sistema identifica el tipo de problema y aplica una cadena de procedimientos de transformación. En el MIT se utiliza ampliamente por los científicos del instituto. Versiones del sistema se emplean, a efectos demostrativos, en las manifestaciones de inteligencia artificial. MACSYMA es un ejemplo de las posibilidades de acumulación del conocimiento que permite este tipo de sistema, ya que desde sus primeras versiones hasta la actual, se han producido extensiones sucesivas, fruto de los casos fallidos experimentados, que han producido un sistema actualmente comparable en su rendimiento con calculistas humanos experimentados.

— El sistema DENDRAL realiza, en el campo de la química, hipótesis de interpretación de estructuras moleculares a partir de datos de densidad espectral y de respuesta a ensayos magnéticos nucleares.

Dado su carácter pionero, su capacidad de explicación es baja y tampoco contiene facilidades de comunicación con el usuario; sin embargo, su rendimiento es comparable con expertos personales. Este sistema es ampliamente utilizado por los investigadores en Estados Unidos.

— MYCIN es el primer sistema en que una vez constatada la arquitectura base de conocimiento-motor inferencial, se emplea para construir un modelo conceptual, una red de relaciones

causa-efecto y un motor inferencial apoyado en lógica no clásica. Para ello Shortliffe y Buchanan construyeron un cálculo de factores de certidumbre elaborado "ad hoc" y que constituye un paradigma generalizado para este tipo de sistemas, como lo muestra su utilización para los sistemas PUFF, SACON y GUIDON, y la integración de su técnica inferencial en la instrumentación industrial de este tipo de sistemas (EMYCIN, ROSIE, KS300, SYSTEM I, etc.).

MYCIN es un sistema con una red formada por un número de relaciones próximo a 400, que, a partir de la información básica sobre características de un enfermo (generales y resultados de análisis) produce una serie de diagnósticos verosímiles, ordenados por grado de certidumbre, sobre enfermedades de tipo bacteriano y meningítico.

El sistema no se ha implementado para uso general, aunque han sido validadas sus conclusiones por expertos como equivalentes a las de éstos.

— PROSPECTOR, desarrollado por SRI para establecer hipótesis sobre el grado de certidumbre de existencia de yacimientos minerales a partir de la descripción geológica y geotécnica del área investigada, utiliza un esquema conceptual del tipo de MYCIN, pero con una variante sobre el motor de inferencia, basado en probabilidades subjetivas bayesianas como MYCIN, pero con otro tipo de parametrización distinta de los factores de certidumbre, pero incorporando además una red semántica integrada con la red de relaciones causales. El paradigma de inferencia de PROSPECTOR se ha revelado menos eficaz, desde el punto de vista de capacidad de representación, que el de MYCIN, como muestra su menor difusión, por su mayor dificultad de ajuste. No obstante debidamente complementado con modificaciones locales, ha producido un sistema que ha detectado un yacimiento, desapercibido por expertos humanos,

de molibdeno, valorado en cien millones de dólares y ha colaborado en otros, proporcionando respuestas aceptadas por los expertos. El sistema tiene asociada una interfaz para adquisición del conocimiento (KAS, Knowledge Acquisition System) y tiene un planteamiento muy operativo, que está en uso por el US Geological Survey. Con la estructura de PROSPECTOR, se desarrolló el sistema HYDRO destinado a realizar estimaciones de los parámetros de comportamiento de cuencas hidrográficas, a partir de sus características geológicas y morfológicas.

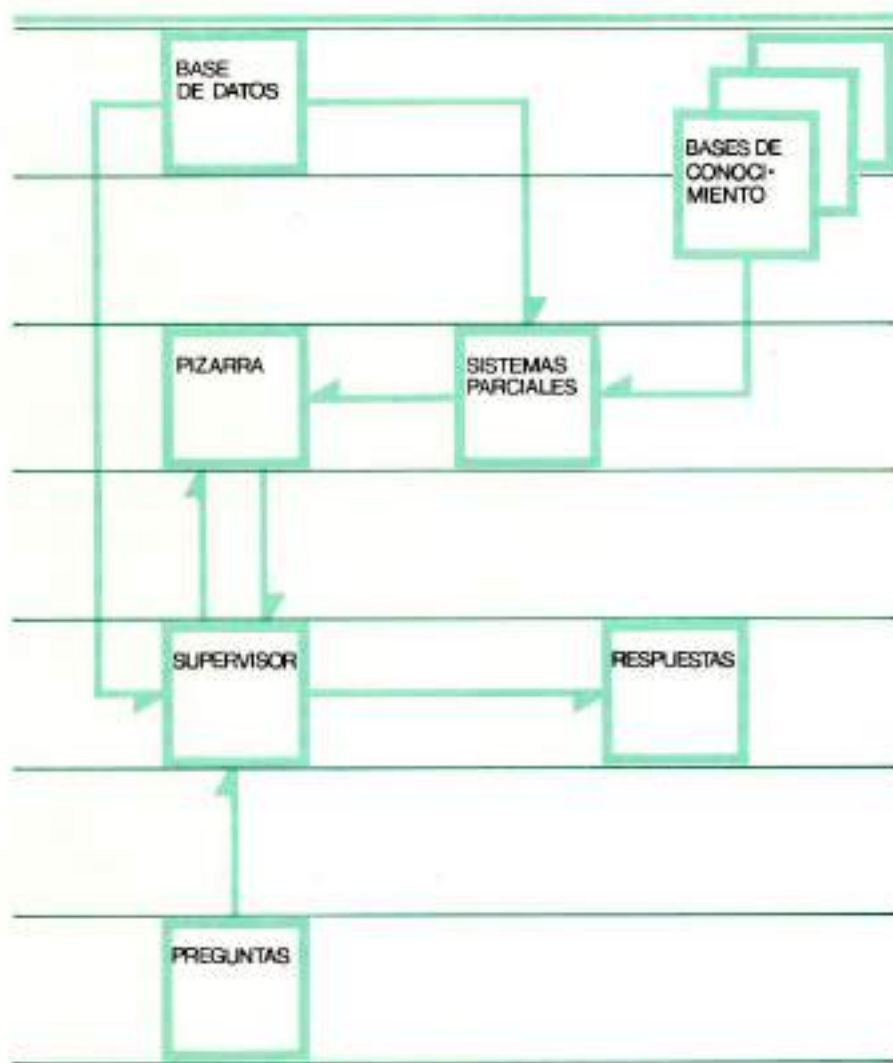
Los sistemas CASNET e INTERNIST son variaciones sobre estos esquemas, basados fundamentalmente en redes causales y tratamiento probabilístico de la incertidumbre.

— CASNET incorpora una red para diagnóstico de glaucoma, su esquema conceptual, como MYCIN, se ha empleado para un instrumento de tipo general denominado EXPERT. Tanto uno como otro tienen una arquitectura avanzada, desde el punto de vista de comunicación con el usuario y posibilidad de modificación y ampliación de la base de conocimiento.

— INTERNIST, incorpora una gama amplia de conocimiento sobre medicina interna que incluye: 350 síntomas, 500 enfermedades y cerca de 10.000 relaciones entre síntomas. El sistema, todavía en fase de desarrollo, incorpora el 90 % del conocimiento previsto y tendrá un impacto importante en la diagnosis de medicina interna.

El sistema HEARSAY-II, ofrece la innovación de una arquitectura distinta, del sistema de representación del conocimiento. HEARSAY es un proyecto de identificación de la palabra hablada, en el que ante la complejidad del problema y la variedad de formas posibles de tratamiento, todas ellas insuficientes por sí solas, se optó por un sistema basado en el conocimiento constituido por una serie de subsistemas, que realizan hipótesis que graban sobre un área denominada Pizarra

o Agenda. Sobre ésta opera otro sistema denominado planificador o supervisor, que adopta hipótesis alternativas deducidas de las propuestas, por cada sub-sistema. Este esquema, denominado de planificación oportunista, que se representa en la figura adjunta, constituye una solución ingeniosa para integrar diversas fuentes de conocimiento, aunque no resulta fácil de construir el sistema supervisor. El sistema HEARSAY, cuyos resultados en cuanto a su objetivo primordial de interpretación de la palabra hablada, no han sido muy brillantes, aporta en cambio este nuevo concepto de estructura, que también ha sido utilizado para el sistema HASP, dedicado a fines estratégicos clasificados, cuyo objetivo es producir hipótesis sobre señales de sensores. El concepto de HEARSAY se ha abstraído e instrumentado como el caso de EMYCIN en el sistema AGE.



— El sistema R 1 se comenzó a desarrollar en 1979, por acuerdo entre Digital y la universidad de Carnegie Mellon, con el objetivo ya no de experimentación sino de aprovechamiento industrial. El sistema se desarrolló a lo largo de un año, partiendo de un prototipo de 300 reglas, con el objetivo de producir las listas más adecuadas de componentes, para las demandas de equipos VAX.

La versión obtenida al cabo de un año, tenía 800 reglas y su rendimiento era 90 % aceptable por los vendedores de Digital; en junio de 1983 la base de conocimiento tenía ya 2.500 reglas y el sistema estaba plenamente operativo, siendo significativamente distintas en calidad las configuraciones producidas por el sistema de las generadas por los profesionales.

Este es un sistema planteado con un concepto actual de ingeniería de conocimiento y su diseño se ha realizado apoyado en el sistema OPS5.



#### 4. PANORAMA DE LAS TÉCNICAS DE CONSTRUCCIÓN DE SISTEMAS EXPERTOS

De acuerdo con la definición, una técnica para diseño de este tipo de sistemas se definiría por el conjunto de métodos:

- para formular bases de conocimiento.
- para utilizar la base de conocimiento.
- obteniendo respuestas a los problemas.
- explicando estas respuestas a demanda.

de forma que la operación sea suficientemente eficaz y la construcción del sistema sea factible, en forma operativa. Es importante tener en cuenta esta última consideración, ya que de nada sirve una conceptualización absolutamente perfecta, si su operación es ineficiente y su complejidad hace difícil y poco fiable la creación de la base de conocimiento, a partir de las afirmaciones de los expertos.

La mayoría de los sistemas desarrollados, se apoyan en mecanismos de razonamiento sobre relaciones de causalidad entre conceptos. En algunos casos estas relaciones están claras, pero en otros, a la hora de su formulación, los expertos no se definen en forma absoluta, de manera que es inevitable tener en cuenta en la metodología de representación e inferencia, el grado de definición de estas relaciones de causalidad.

Esta necesidad de tener en cuenta, en cierta forma, la imprecisión, hace que sea posible definir dos niveles de soporte, según se parta de la hipótesis de implicación total o parcial.

En el primer nivel cabría incluir los **métodos basados en la lógica clásica**, tanto a nivel de sistemas basados en técnicas de decisión de fórmulas de lógica de primer orden, origen de la programación lógica (lenguaje PROLOG), como a **nivel de técnicas algorí-**

**tmicas construídas a la medida de un tipo de problemas, como son las redes semánticas.**

En el segundo nivel cabe incluir los sistemas basados en lógicas **probabilísticas** (de tipo subjetivo bayesiano), como MYCIN y PROSPECTOR o los sistemas basados en lógica difusa, intento no totalmente conseguido de construir una teoría del razonamiento impreciso.

En la construcción de un sistema concreto cabe que haya elementos de ambas técnicas; notablemente cabe la representación de una red causal de tipo probabilístico como red semántica.

A continuación se describen las diversas técnicas de representación e inferencia, aplicables a los tipos de sistemas más usuales. Dado el carácter inicial de esta monografía, se han tratado de presentar en forma simplificada. Sin embargo, un mínimo de lenguaje matemático es inevitable en algunos casos. Cuando esto ocurre, se ha tratado el tema a dos niveles, incluyendo en el primero los resultados y en el segundo, con letra de menor tipo, su justificación.

##### 4.1. Lógica de primer orden

La lógica clásica se ocupó, desde sus orígenes, de estudiar el razonamiento deductivo. Los desarrollos del siglo pasado debidos a Boole para formulación del cálculo proposicional y a Frege, creador del cálculo de predicados, y los desarrollos del primer tercio del presente siglo, crearon un cuerpo de doctrina sobre el que, desde primeros de los años sesenta, se basaron los investigadores para construir técnicas de deducción automática, dando lugar a diversos métodos, de los que el más importante es el de resolución, debido a Robinson en 1965, refinado posteriormente por Luckham y Loveland.

De una manera general, el cálculo de predicados es indecidible, es decir, no



existe un proceso, que aplicado a una fórmula del cálculo de predicados termine necesariamente afirmando si la fórmula es válida o no. Por tanto, un algoritmo basado en el método de resolución aplicado a una fórmula, puede terminar afirmando esto o puede operar indefinidamente sin encontrar una solución. No obstante existen importantes subconjuntos de fórmulas, del cálculo de predicados, en los que sí existe decisión y por tanto las técnicas de deducción automática, constituyen un posible instrumento de desarrollo de aplicaciones.

De acuerdo con las consideraciones anteriores, un sistema basado en el conocimiento puede formularse en lógica de primer orden, de la forma siguiente:

— El estado del mundo sobre el que se razona, se describe por fórmulas de cálculo de predicados, redactadas a partir de las relaciones base del discurso.

— La base de conocimiento se describe por el conjunto de relaciones de implicación que caracterizan:

— las transformaciones a realizar en el mundo de discurso

— los predicados complejos definidos a partir de los básicos.

— El motor de inferencia es un procedimiento de deducción automática que para cada pregunta:

— estudia si es deducible del estado del mundo y la base de conocimiento

— caso de serlo, obtiene los valores de las variables para los que la pregunta es deducible

— el proceso de búsqueda de la demostración constituye la base para explicación de las respuestas.

A continuación se ilustra el concepto con un ejemplo sencillo:

Sea un sistema abstracto que concede licencias de acuerdo con los siguientes criterios:

- Se conceden sólo las licencias que puede tramitar el organismo concesionario por estar dentro de sus competencias.

- Una petición está dentro de las competencias si se cumplen dos requisitos:

- El primer requisito se satisface si existe algún otro licenciatario que avale la petición.

- El segundo requisito se satisface si el domicilio del peticionario está en Madrid.

- Una licencia, aceptada a trámite, no se concede si se producen impactos medioambientales indeseables o la zona de influencia de la concesión nueva coincide con las otras licencias.

- En caso contrario, se concede la licencia.

Para construir un sistema lógico capaz de decidir y explicar la adjudicación de licencias, se requiere formular el conjunto de afirmaciones anteriores en lógica de primer orden. Para ello, en primera etapa, es preciso definir un marco de discernimiento, entendido como el dominio de objetos y relaciones y propiedades en base a los que se plantean las afirmaciones anteriores. En este caso, el dominio está constituido por lugares y personas objeto de la gestión de la organización en la que se encuadra el sistema a construir; sobre las mismas, se definen las relaciones y propiedades representadas por los predicados siguientes:

**concedida** (X): X es una licencia concedida.

**dentro** (X): X está dentro de las competencias del organismo.

**tramitable** (X): X es tramitable.

**requisito 1** (X): X cumple el primer requisito.

**requisito 2** (X): X cumple el segundo requisito.

**peticionario** (X,Y): X es peticionario de licencia Y.

**domicilio** (X,Y): X está domiciliado en Y.

**impacto** (X): X produce impactos ambientales no deseables.

**coincide** (X,Y): la zona de influencia de X coincide con la de Y.

**avala** (X,Y): X avala a Y.

La base de conocimiento se formula a partir de estos elementos de discernimiento de la forma siguiente:

- $\forall X \text{ Dentro}(X) \rightarrow \text{Tramitable}(X)$

(Son tramitables las actuaciones dentro de las competencias del organismo).

- $\forall X [\text{requisito 1}(X) \wedge \text{requisito 2}(X) \rightarrow \text{dentro}(X)]$

(Una petición está dentro de las competencias del organismo si se cumplen los requisitos 1 y 2.)

- $\forall X [\text{peticionario}(X,Y) \wedge \text{EZ avala}(Z,Y) \rightarrow \text{requisito 1}(X)]$

(Si alguien avala al peticionario de X, X cumple el requisito 1.)

- $\forall X \forall Y [\text{peticionario}(X,Y) \wedge \text{domicilio}(Y,\text{Madrid}) \rightarrow \text{requisito 2}(X)]$

(Si el domicilio del peticionario es Madrid se cumple el requisito 2.)

- $\forall X [\text{tramitable}(X) \wedge \text{impacto}(X) \vee \text{EZ coincide}(X,Z) \rightarrow \sim \text{concedida}(X)]$

(Si el domicilio del peticionario es Madrid se cumple el requisito 2.)

- $\forall X [\text{tramitable}(X) \wedge \text{impacto}(X) \vee \text{EZ coincide}(X,Z) \rightarrow \sim \text{concedida}(X)]$

(Una licencia tramitable se dictamina no-concedida si produce impactos ambientales no deseables o coincide su zona de influencia con otra concesión.)

- $\forall X \forall Z [\text{tramitable}(X) \wedge \sim \text{impacto}(X) \wedge \sim \text{coincide}(X,Z) \rightarrow \text{concedida}(X)]$

(En caso contrario se dictamina concedida.)

Este conjunto de fórmulas permite, a partir de una base de datos, contestar preguntas sobre licencias.

La base de datos estará constituida por la información sobre solicitudes formuladas, en base a los predicados discernidos por la base de conocimiento.



Así, por ejemplo, si hay dos solicitudes A1, A2 se describirán por la información:

petionario (Gómez,A1)  
avala (Pérez,Gómez)  
domicilio (Gómez,Madrid)  
~ impacto (A1)  
petionario (García,A2)  
domicilio (García,Arganda)  
impacto (A2)

El motor inferencial sería un proceso de deducción automática, que contesta preguntas del tipo: ¿concedida (A1)? o bien ¿concedida (X)?

En el primer caso, el proceso estudia si concedida (A1) es deducible de los bloques de fórmulas (1) y (2). En el segundo a la par que analiza si E x concedida (X) es deducible, obtiene el valor, o valores, de X para los que concedida (X) es verdad.

Dado que el motor inferencial es un proceso genérico de deducción automática, el único elemento a definir por el diseñador, es la base de conocimiento, lo que se hace trasladando, la forma de entender el tema a nivel de lenguaje natural, a una formulación a nivel de predicados. Este acto es considerablemente más sencillo que la forma actual de plantear la programación, en la que hay que pasar de la forma de entender un tema, a la estructura de información y mecanismos de proceso de las máquinas tipo Von Neuman, más o menos acomodadas por los lenguajes procedurales habituales.

Esta idea de modelizar los conceptos en lógica de primer orden y resolver los problemas, mediante un sistema de deducción automática, constituye la base de la programación lógica, de la que el primer producto operativo es el lenguaje PROLOG.

En el momento actual, el ideal de la programación lógica no está alcanzado, porque las técnicas de deducción automática disponibles para ser operativas, dando respuesta en tiempos aceptables, requieren una programa-

ción de control que limita el ámbito de aplicabilidad y hace más compleja la formulación de procesos, que, por otra parte, cuando se plantean bases de conocimiento pueden requerir recursos importantes de tiempo y de computador (\*).

## 4.2. Redes semánticas

### 4.2.1. ESTRUCTURAS

La crítica tradicional a la formulación en lógica de las bases de conocimiento, a pesar de su aplicabilidad general y su base teórica, es su poca eficiencia, que puede dar lugar a sistemas de implementación costosa. La alternativa a estas formulaciones basadas en un marco lógico general, está en la representación a la medida del problema, mediante un sistema en el que se incorporen todos los elementos del marco de conceptos, específico del tipo de problemas, procesable en forma sencilla. Este enfoque, adolece de que los problemas de inconsistencia que en una formulación lógica se evitan por construcción, deben ser controlados por el diseñador, lo que en bases complejas no siempre es fácil.

Uno de los métodos de representación "ad hoc" más difundido es el de las redes semánticas. Este concepto fue introducido por QUILLIAM en 1968 para representar la memoria asociativa humana, y desde entonces ha sido objeto de atención y desarrollo según diferentes orientaciones.

Una red semántica constituye una representación de un universo lógico en el que el conocimiento se describe en base a relaciones en un dominio de objetos. En la versión más simple de una red semántica los vértices representan los elementos del dominio y los arcos las relaciones entre ellos. Es de señalar que, al representar los arcos relaciones concretas, van etiquetados por el nombre de la relación correspondiente, de forma que los pro-

(\*) El fundamento teórico de las técnicas de deducción automática, se apoya en los siguientes resultados teóricos:

a.—Una deducción  $A, B, C, \dots \vdash R$  es correcta si y sólo si la fórmula  $A \wedge B \wedge C \wedge \dots \wedge \sim R$  es insatisfacible.

Se dice que una deducción es correcta, si toda interpretación que satisface (hace verdad) las premisas, satisface la conclusión.

Una fórmula es insatisfacible, si no existe ninguna interpretación que la haga verdad. En el caso de la afirmación objeto de este epígrafe, si existiera una interpretación que satisficiera  $A \wedge B \wedge C \wedge \dots \wedge \sim R$ , haría verdad  $A, B, C, \dots$  y falso  $R$ , por lo que la deducción dada no sería correcta.

b.—Una fórmula es insatisfacible si y sólo si su forma clausal es insatisfacible (SKOLEM).

Una fórmula se dice que está en forma clausal si:

- está cuantificada universalmente
- el signo de negación está aplicado a los predicados simples
- la estructura de la fórmula se describe como la conjunción de disyunciones (cada una de ellas se denomina cláusula):

Por ejemplo la fórmula:

$$\forall X \forall Y \forall Z [p(X) \wedge (q(X) \vee c(X, Z)) \sim co(X)]$$

Se pone en forma clausal con la formulación:

$$\forall X \forall Y \forall Z [( \sim p(X) \vee \sim q(X) \vee \sim co(X) ) \wedge$$

$$( \sim p(X) \vee \sim c(X, Z) \vee \sim co(X) )$$

cuya anotación habitual haciendo abstracción de los cuantificadores y las conjunciones es:

$$\sim p(X) \vee \sim q(X) \vee \sim co(X)$$

$$\sim p(X) \vee \sim c(X, Z) \vee \sim co(X)$$

c.—La condición necesaria y suficiente para que una conjunción de cláusulas sea insatisfac-

tible, es que por resolución y unificación se deduzca de ellas la cláusula vacía.

Dadas dos cláusulas, por ejemplo:

$$\sim A(f(X)) \vee R(X, f(Z)) \\ A(Y) \vee S(Y, V)$$

Se obtiene a partir de ellas otra resolvente si:

- Existe en ambas un mismo átomo de predicado, afirmado en una y negado en otra.
- Los términos de ambos son unificables, es decir, existe una sustitución de las variables de ambos que unifica términos de  $\sim A(f(X))$  y  $A(V)$ .
- En el ejemplo la sustitución es  $V/f(X)$ , y aplicando resolución a ambas se deduce de ellas la cláusula:  $S(Y, f(X)) \vee R(X, f(Z))$ .
- Como la regla de resolución es una regla de deducción, si dos cláusulas con un único átomo son unificables, de ellas se deduce la cláusula vacía, lo que significa que existe una sustitución que las hace contradictorias.

Por ejemplo:

$\sim S(X, f(X, Z))$  y  $\sim S(f(Z), f(a), a)$  se unifican con la sustitución  $\langle X/f(a), \langle Z/a \rangle$  lo que significa que hay una especificación de ambas contradictoria.

• Por tanto, si de un conjunto de cláusulas se deduce por resolución la cláusula vacía, es que de ellas se deduce una contradicción y por tanto el conjunto de cláusulas es insatisfacible, puesto que si no fuera así, una interpretación que satisficiera las cláusulas, satisficiera la contradicción deducida, lo que no es posible.

• Se demuestra recíprocamente, basado en el teorema de Herbrand, que si un conjunto de cláusulas es insatisfacible, necesariamente por resolución se alcanza la cláusula vacía.

• Por tanto, la regla de resolución, salvo los casos de indecidibilidad, constituye un mecanismo que asegura la obtención de respuestas, en problemas de deducción automática.

d.—Un algoritmo de deducción automática consiste por tanto en:

- Formulación de la deducción en forma clausal
- Búsqueda de secuencias de resolución y unificación que conduzcan a la cláusula vacía. Cada sucesión de resoluciones que obtiene la cláusula vacía, incluye una especificación de las variables resultantes del proceso de unificación, que constituye una respuesta posible al problema planteado en términos deductivos.

e.—Para simplificar el proceso de búsqueda, se han planteado criterios restrictivos en cuanto a la formación de parejas de resolventes, ya que del tamaño de este universo de parejas posibles depende la eficacia del proceso. Un criterio restrictivo constituye una estrategia de resolución. Una estrategia es completa cuando el criterio restrictivo en que se basa, no mena su eficacia teórica, es decir, sigue ocurriendo que si existe una deducción de la cláusula vacía, mediante la estrategia aplicada, necesariamente se encuentra.

f.—El lenguaje PROLOG se apoya para un tipo específico de cláusulas, en la técnica de resolución lineal, asumiendo para definir parejas de resolventes un orden en las cláusulas, precisamente aquel en que se formulan.

También, con objeto de limitar la búsqueda, incorpora en su formulación órdenes de "congelación" de la búsqueda en una rama (regla "cut" o "/" y de limitación de las posibilidades de unificación, predicado "dif"). Por tanto, en PROLOG, no basta con la especificación del conocimiento mediante fórmulas lógicas; se requiere también codificar el "control" de la forma de búsqueda de la máquina lógica que va a interpretar el programa. Esto es expresado por Kowalski en 1976 mediante la ecuación:

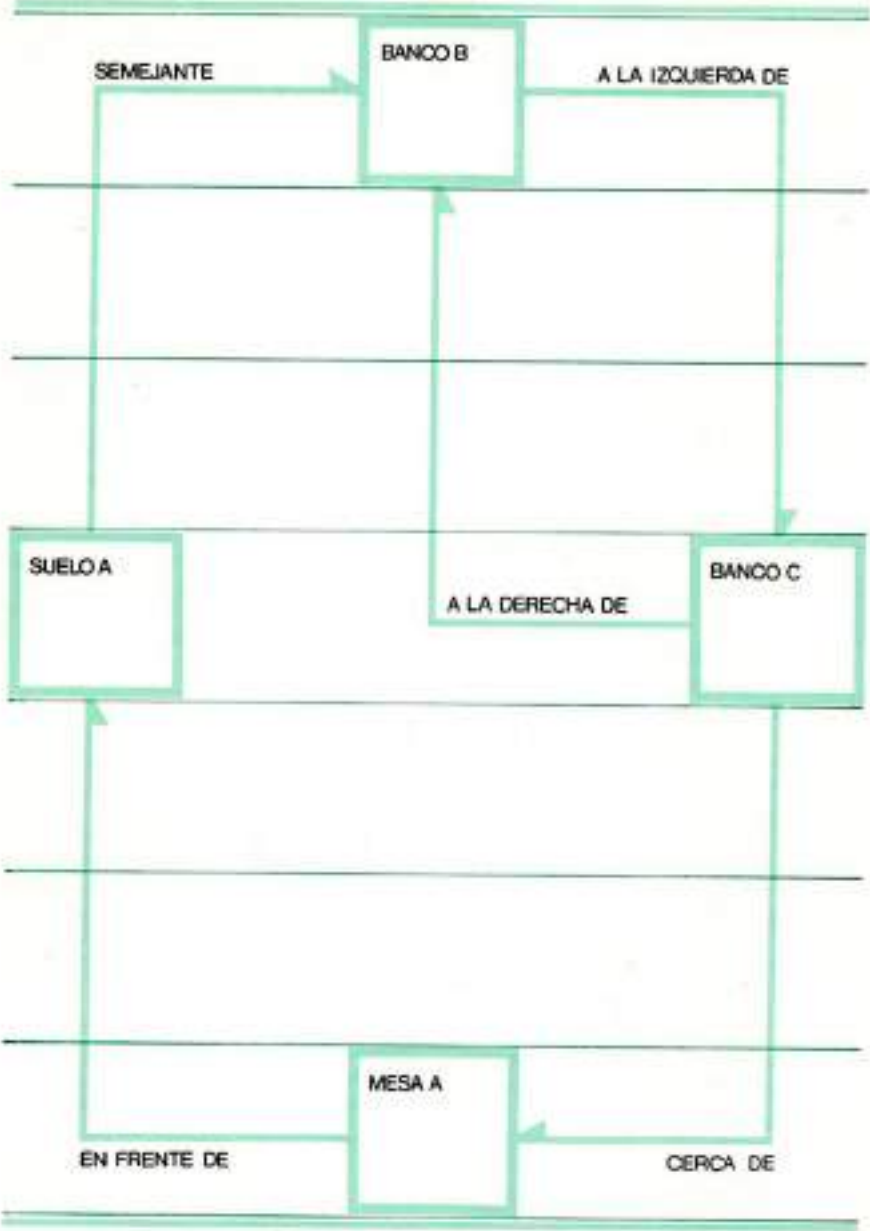
$$\text{algoritmo} = \text{lógica} + \text{control}$$

Todavía no se ha encontrado un sistema de programación lógica ideal, pero los avances actuales son significativos. En el apartado 5 se presenta PROLOG con algo más de detalle.



cedimientos de obtención de respuestas del motor de inferencia, tienen en cuenta la etiqueta del arco para resolver los problemas planteados. En la figura adjunta se incluye una red, referida en este caso a relaciones de situación o semejanza.

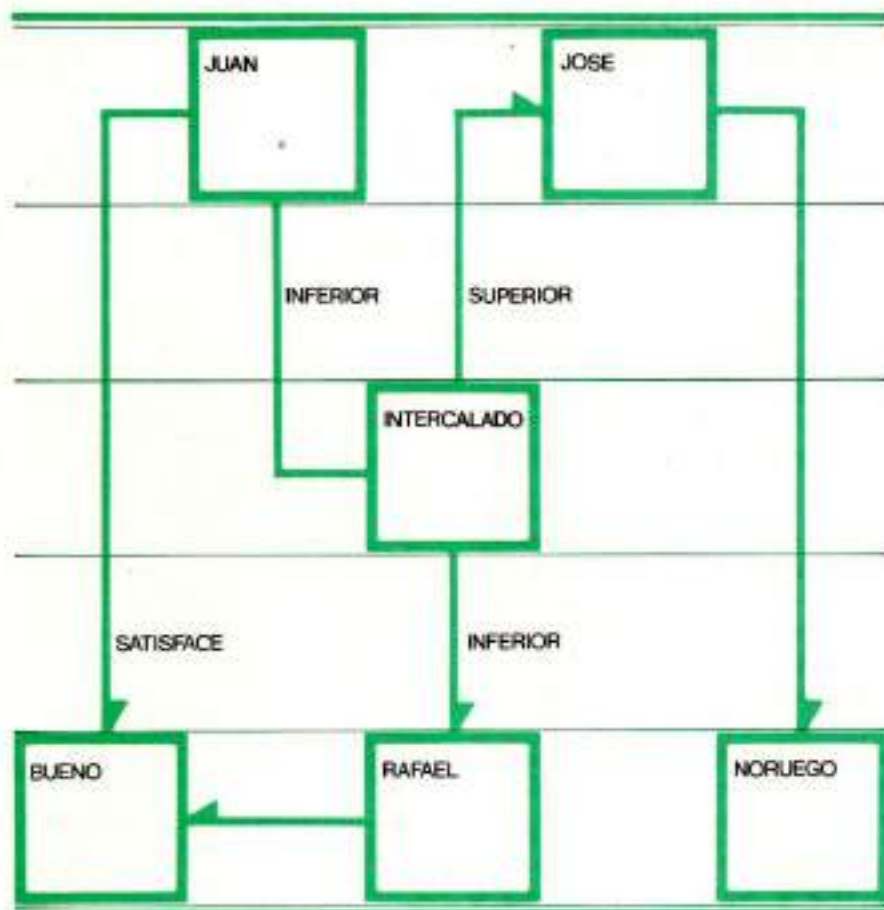
Para resolver el problema de representar predicados (relaciones) no binarias, mediante redes semánticas, ha sido preciso generalizar el concepto de nodo o vértice de la red, que de una manera general puede ser tanto un objeto del dominio como un nombre de relación. De esta forma aparece un nuevo tipo de arco que representa la relación entre el predicado y el término. Así, puede haber términos o nodos "propiedad" y nodos "objeto" que satisfacen una determinada propiedad conectados por arcos con la etiqueta "satisface".





Por ejemplo, en el caso de relaciones ternarias, la propiedad que liga el predicado con el objeto, tiene que ver con el papel del objeto en la relación representada. Así, para representar la relación, Juan está intercalado entre José y Rafael, cabe definir un nodo-relación cuyo nombre es "intercalado" y tres nodos-término Juan, José y Rafael y tres arcos relación-término cuyas etiquetas serían borde superior (José), borde inferior (Rafael) y el elemento interior (Juan).

En la figura aparece una red, que incluye los conceptos anteriores:







#### 4.2.2. PROCEDIMIENTOS INFERENCIALES

Los motores inferenciales de las redes semánticas se diseñan "ad hoc" en forma acorde con la semántica de las etiquetas de los arcos de conexión. En efecto, de acuerdo con la descripción anterior, una red semántica puede describirse sintéticamente por un conjunto de nodos (que pueden ser objetos o conceptos) y un conjunto de arcos de distintas clases (elemento de, subconjunto de, ... y relaciones específicas).

Los problemas a resolver con esta base de conocimiento, son del tipo de encontrar objetos o conjuntos de objetos que satisfagan una serie de condiciones. De acuerdo con esto, el motor inferencial se define como el conjunto de procedimientos que, apoyado en la estructura de la base y la semántica de sus arcos, proporcionan las respuestas a las preguntas posibles. Para ello se descompone la pregunta en función de sus proposiciones componentes y se lanzan procesos de búsqueda sobre la red, que seleccionen los conjuntos que satisfacen cada componente, lo que se comprueba por técnicas denominadas de "Pattern Matching", analizando si una cadena de símbolos representativa de un elemento de la red tiene la estructura del patrón (pattern) que se investiga.

Para agilizar estos procesos de búsqueda, se apela a un soporte de diccionarios asociados a la red, del tipo empleado en las bases de datos tradicionales; por ello una vía de formulación de este tipo de bases de conocimiento, pueden ser los actuales lógicos de ayuda al diseño de bases de datos relacionales. En efecto, las redes semánticas constituyen el vehículo de representación del conocimiento más integrable con las bases de datos tradicionales; de hecho en los ejemplos presentados, el nivel de asignación de propiedades a objetos es una base de datos tradicional, siendo propiamente la red de relaciones

entre conceptos la base de conocimiento. Es posible, por tanto, manejar el mismo software para resolver problemas en dos instancias: a nivel de los datos y a nivel de los conceptos y en esa línea se han hecho algunas propuestas.

#### 4.2.3. EXTENSIONES: PARTICIONES Y ESCENAS (FRAMES)

La utilización de la representación en red, tiene gran tradición en Inteligencia Artificial, principalmente por la libertad que deja al diseñador, de creación desde cero, de un sistema inteligente. Sin embargo, adolece de que en los sistemas complejos, que son la mayoría, el riesgo de inconsistencia es importante, ya que es posible que un procedimiento de obtención de respuestas encuentre, en una misma red, respuestas afirmativas y negativas a una misma cuestión, porque el sistema de conexiones, si está en forma no ordenada y el volumen de la red es importante, y si además está creado incrementalmente, puede introducir esa posibilidad no controlada, por el mecanismo inferencial que se limita a la búsqueda en la red. La forma de evitar esto, es no perder el control mental de la red para lo que, como en la programación, es necesario introducir estructura en la organización de la red, que a su vez sirva de base a una metodología sistemática. Hendrix, uno de los autores que ha construido redes complejas (para los sistemas de comprensión del lenguaje natural LIFER y LADDER, 1977) propuso en 1975 y 1979 el concepto de partición, que permitía englobar trozos de red en subredes, entre las que se definían según la clase de problema a resolver, grafos de relaciones denominados **vistas**. Esto, a la par que simplificaba el diseño e inspección de la red, hace más eficientes los procesos de resolución de problemas, ya que mediante el concepto de **vista** asociada a los distintos tipos de problema, se puede operar en cada caso con el subconjunto

pertinente de la red, con la economía subsiguiente.

Otro concepto estructurante mixto de ideas de redes semánticas y lógica de primer orden, es el de **escena** (FRAME) debido a MINSKY. El concepto de escena es una generalización del concepto de predicado. En efecto, un predicado es una abstracción de una propiedad o relación, en la que existen plazas a ser cubiertas por términos de un dominio. Cuando se cubren estas plazas por estos términos, se obtiene una afirmación o declaración. Así por ejemplo:

\_\_\_ está entre \_\_\_ y \_\_\_

es una "forma" de frase cuyas palabras pueden ser cubiertas por objetos de un dominio obteniéndose, por ejemplo, si el predicado está definido en el dominio de ciudades españolas, una frase con significado:

**Zaragoza está entre Madrid y Barcelona**

De manera análoga, una escena o cuadro es una estructura abstracta, como el predicado, pero cuyas plazas se denominan **ranuras** (slots) y pueden ser ocupadas no sólo por términos, sino por otras escenas que pueden degenerar en predicados, por funciones o procedimientos para encontrar el elemento que ocupa una **ranura**. Así por ejemplo, puede definirse la escena "oficina bancaria" con las siguientes **ranuras**:

oficina bancaria:  
nombre (Bilbao, Central, Hispano, etc...)  
por defecto: Banco Central  
tipo: por defecto: sucursal 1  
mostrador: (recto, segmentado, en forma de U) X (mármol, madera)  
defecto: recto de mármol  
horario: (completo, de mañana, de tarde)  
defecto: de mañana  
decoración: (clásica, funcional) X (con alfombras, con moquetas, etc...)  
servicio: organización bancaria  
remuneración: procedimiento (depósito, plazo, interés)



En cada una de las **ranuras** se incluye el dominio de valores que corresponde; en un caso se ha incluido como valor "organización bancaria" que a su vez es una **escena** descriptiva de la estructura de una unidad de servicio bancario en la que, por ejemplo, puede haber **ranuras** para los distintos puestos del organigrama, a su vez representables por **escenas** (esquema operativo típico de un interventor, por ejemplo).

Asimismo, en la **ranura** dedicada a "remuneración" se incluye, como concepto de relleno, un procedimiento que define la renta en función de las condiciones de la imposición.

Hayes (1980) demostró que las **escenas** pueden referirse a predicados, por lo que, aunque incorporan un mayor grado de complejidad, la lógica de las **escenas** es una lógica de predicados. Sin embargo, como vehículo de implementación, constituyen un elemento importante, ya que para cada **escena** pueden construirse una serie de funciones, para resolver problemas relativos a su estructura, que pueden aplicarse para cada instancia (especificación) de la **escena**. Asimismo, la ordenación del discurso sobre un tema, mediante **escenas**, permite una mejor estructuración de los conceptos de la base de conocimiento a diseñar.

El concepto de **escena** aparece, aunque acuñado por MINSKY (1975), en diversos autores, con diversa variante. Por ejemplo, SCHANK y ABELSON en 1977, denominaban **guiones** (Scripts), a patrones de comportamiento con estructura general preestablecida.

Otro aspecto a subrayar en el concepto de **escena**, es la posibilidad de definir la **escena** con especificación de dominios de valores y procedimientos y con valores por defecto, de manera que con ellos se plantea un cierto paradigma del concepto al que se refiere. Cuando no se especifica el concepto de la **ranura** correspondiente, las **escenas** con valores por defecto, incorporan un cierto patrón definido en

forma abstracta y concretado con los valores de lo que podría considerarse el arquetipo del concepto.

Las **escenas**, como las afirmaciones universales en cálculo de predicados, tienen reglas de inferencia propias. Así, de una definición general de la **escena** "oficina bancaria", se deduce cualquier oficina bancaria por especificación de sus **ranuras** y recíprocamente, si una entidad definida por un conjunto de datos satisface las condiciones de cada **ranura**, se deduce que es una oficina bancaria. Esta segunda forma de inferir conceptos a partir de otros, de menor nivel, utiliza la definición de la **escena** como cuadro de criterios par apasar de una categoría de conceptos a otra. Esta forma de inferir, en base a la estructura del cuadro o **escena**, es semejante a la forma con que los médicos hacen diagnósticos (se habla de un cuadro de gripe o de neumonía). En este sentido, el concepto de **escena** puede considerarse como una unidad de discurso correcta aunque persisten las limitaciones mencionadas de riesgo de inconsistencia, si bien están paliadas.

#### 4.3. Redes inferenciales

El conjunto de métodos que se describen a continuación trata de modelizar el razonamiento impreciso. La estructura de representación es una red, cuyos nodos son afirmaciones y cuyos arcos son relaciones causa-efecto, parametrizados en forma continua en uno o varios parámetros, que evalúan el grado de implicación de las **afirmaciones causa** respecto de las **afirmaciones efecto**.

Cada uno de los nodos tiene a su vez asignado uno o varios valores, medida del grado de soporte de la afirmación correspondiente.

El soporte de la afirmación en cada nodo debe tener valores en todos los nodos que deben ser consistentes con el grado de implicación de los ar-



cos de la red y los valores atribuidos a los nodos iniciales de la red. (Nodo inicial es aquel que no es efecto de ninguna causa y por tanto, su grado de soporte es un dato de partida.)

El motor inferencial evalúa los parámetros de soporte de los nodos de la red, mediante búsqueda en ella y aplicación de un método de cálculo de transmisión y acumulación de evidencia.

Es posible construir redes semánticas integradas con redes inferenciales, que describen la estructura de las afirmaciones de cada nodo de la red inferencial, en la que haya su propio mecanismo de inferencia local. En esa línea está el sistema PROSPECTOR.

Los arcos de la red inferencial se denominan **reglas** y por ello, este tipo de sistemas se denominan también **basados en reglas** (rule based). En el momento actual, este concepto constituye el soporte de la gran mayoría de Sistemas Expertos.

Como ya se ha indicado en otra parte de esta monografía, no hay una teoría unánimemente admitida de formulación del razonamiento impreciso. En principio, a un modelo de este tipo de razonamiento debe exigirse:

- Un grado de verosimilitud y consistencia suficiente.
- Un grado de eficiencia aceptable en los procesos.
- Una metodología sencilla, y por tanto, fiable para trasladar al tipo de cuantificación elegido la forma de valorar unos conceptos respecto de otros por los expertos.

Estas dos últimas condiciones son importantes ya que de nada sirve una conceptualización muy sofisticada, si los procesos de resolución de problemas son excesivamente prolijos, o no es factible o no es fiable la forma de trasladar el conocimiento de los expertos.

Se han propuesto diversos métodos, todos ellos objeto de crítica en alguno

de los aspectos anteriores, pero algunos vienen avalados porque en base a ellos se han construido sistemas que operan en forma convincente, notablemente MYCIN y PROSPECTOR. A continuación se describen los métodos empleados en estos sistemas y las variantes hacia una mejor conceptualización que en este momento son objeto de estudio y experimentación.

#### 4.3.1. METODO MYCIN

##### 4.3.1.1. Definiciones

Este procedimiento fue desarrollado por Shortliffe y Buchanan, para el sistema MYCIN cuyo objetivo era diagnosticar enfermedades bacterianas y proponer tratamientos a partir de información sobre síntomas, análisis clínicos y características somáticas del enfermo.

Se definen:

MC (h,e) = medida del incremento de credibilidad de la hipótesis h debida a la evidencia e.

MI (h,e) = medida del aumento de incredibilidad de h debida a la evidencia e.

El método MYCIN propone una forma de calcular ambas magnitudes de manera que, como consecuencia de la aplicación del proceso, a partir de la introducción de cada nueva evidencia e, se actualizan ambas medidas y se integran en un factor de certidumbre. El sistema contestará a cada pregunta con la hipótesis h que sea una respuesta correcta o bien, propondrá un conjunto de respuestas ordenadas por factor de certidumbre respecto del conjunto de evidencia de partida.

Para evaluar ambas medidas se parte de:

P(h/e): probabilidad subjetiva de que h supuesto que pasa e.

P(h) : probabilidad subjetiva previa a e, de que ocurra h.

Para evaluar MC(h,e) se hace el razonamiento siguiente:

- Si P(h) mide la probabilidad de ocurrencia de h,  $1 - P(h)$  mide el grado de desconfianza de que ocurra h.
- Por tanto, una medida relativa del grado de credibilidad que aporta e es la relación entre:
  - incremento del grado de certidumbre respecto del existente a priori:  $P(h/e) - P(h)$ .
  - grado de incredibilidad a priori  $1 - P(h)$ .

Es decir:

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

Caso de que el numerador sea negativo, e no aporta nada a la credibilidad de h sino todo lo contrario. En esta hipótesis, se incrementa el grado de incredibilidad que puede medirse por el cociente entre:

- Disminución de la credibilidad a priori P(h):

$$P(h) - P(h/e)$$

- Credibilidad a priori P(h):

Por tanto:

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

La medida adoptada para MC(h,e) evalúa la disminución del grado de incredibilidad previo, y para MI(h,e) evalúa la disminución de la credibilidad.

De acuerdo con la formulación anterior, como una nueva evidencia e sólo puede dar lugar bien a:

$$P(h/e) > P(h) \text{ ó}$$

$$P(h/e) \leq P(h)$$

Se adopta la hipótesis de que, como consecuencia de e, si  $MI(h,e) > 0$   $MC(h,e) = 0$ , y recíprocamente.

En el caso  $P(h/e) = P(h)$  obviamente:

$$MI(h,e) = MC(h,e) = 0$$

En base a las dos medidas antes descritas se define un factor de certidumbre CF que combina ambas:

$$CF(h,e) = MC(h,e) - MI(h,e)$$

El comportamiento de estos conceptos en los extremos es el siguiente:

— En el caso en que h es cierto (supuesto e)  $P(h/e) = 1$

$$\begin{aligned} MC(h,e) &= \frac{P(h/e) - P(h)}{1 - P(h)} = \\ &= \frac{1 - P(h)}{1 - P(h)} = 1 \end{aligned}$$

Como  $P(h) = 1 \geq P(h)$ :

$$MI(h,e) = 0$$

$$CF(h,e) = 1$$

— En el caso en que h es falso  $P(\sim h/e) = 1$

$$P(h/e) = 0$$

$$MI(h,e) = \frac{P(h) - P(h/e)}{P(h)} = 1$$

$$CF(h,e) = -1$$

De acuerdo con las definiciones anteriores, el rango de cada medida es:

$$0 \leq MC \leq 1 \quad 0 \leq MI \leq 1 \quad -1 \leq CF \leq 1$$

Por tanto, el factor de certidumbre sobre el nivel de implicación de h por e, oscila desde -1 en el caso de que la evidencia e incida en forma negativa totalmente, hasta +1, cuando se tiene una implicación perfecta a favor.

Los factores de certidumbre constituyen la medida en que se basará el método de razonamiento y aparte de las ventajas como vehículo teórico de modelización que se describen a continuación, cabe señalar su sencillez a efectos de traslación a ellos de la forma de entender la relación de causalidad por los expertos, ya que tienen

que pronunciarse sobre una única medida simétrica en el intervalo (-1,1). Los factores de certidumbre constituyen una medida distinta de la probabilidad, ya que los grados de certidumbre de afirmación y negación son opuestos y no suman 1 como en probabilidades, ya que:

$$\begin{aligned} CF(h,e) + CF(\sim h/e) &= MC(h,e) - \\ &- MI(h,e) + MC(\sim h,e) - MI(\sim h/e) = \\ &= \frac{P(h) - P(h/e)}{1 - P(h)} - \frac{P(h/e) - P(h)}{P(h)} + \\ &+ \frac{P(\sim h) - P(\sim h/e)}{1 - P(\sim h)} - \\ &- \frac{P(\sim h) - P(\sim h)}{P(\sim h)} = \\ &= \frac{P(h) - P(h/e)}{1 - P(h)} - \frac{P(h/e) - P(h)}{P(h)} + \\ &+ \frac{1 - P(h) - 1 + P(h/e)}{P(h)} - \\ &- \frac{1 - P(h/e) - 1 + P(h)}{1 - P(h)} = 0 \end{aligned}$$

Es importante subrayar este hecho de que los factores de certidumbre de hipótesis opuestas, no son complementarios, sino opuestos. Eso valora este concepto como vehículo adecuado de representación del conocimiento impreciso, ya que en el caso de emplearse probabilidades, si el soporte de una hipótesis es bajo, ello acarrea necesariamente que el soporte de su negación es alto (el complemento a la unidad), lo cual no tiene por qué ser real, ya que el soporte de una y otra debería ser bajo cuando existe un grado importante de ignorancia. La modelación de esta ignorancia es factible con los CF que pueden ser igualmente bajos en valor absoluto para la afirmación y negación de una hipótesis, aunque opuestos de signo.



### 4.3.1.2. Consistencia

A pesar de las ventajas, ya comentadas, en cuanto a la traslación del conocimiento de los expertos, dado un conjunto de factores, es preciso analizar y modificar los factores iniciales con objeto de que sean coherentes con el planteamiento conceptual antes descrito. Las condiciones a controlar al recibir factores por parte de los expertos son:

- Los factores de certidumbre deben estar entre  $-1$  y  $+1$ .
- Una afirmación de certidumbre total sobre una hipótesis acarrea la atribución inmediata de factores  $-1$  a todas las otras hipótesis, que son disjuntas respecto a ella.
- Los factores de certidumbre positivos atribuidos a hipótesis, mutuamente excluyentes, no pueden sumar más de 1.

Esta última propiedad puede comprobarse de la forma siguiente:

Sean  $h_1, h_2, \dots, h_i, \dots, h_n$  las hipótesis excluyentes. Al estar confirmadas por  $e$ , los factores de certidumbre  $CF(h_i/e)$  son en todos los casos iguales a  $MC(h_i/e)$ .

De acuerdo con ello, si se denomina  $S_n$  la suma de los factores:

$$S_n = \sum CF(h_i/e) = \sum_{i=1, n} \frac{P(h_i/e) - P(h_i)}{1 - P(h_i)}$$

- Evidentemente, para  $n=1$ ,  $S_1 \leq 1$
- Para  $n > 1$

$$\sum \frac{P(h_i/e) - P(h_i)}{1 - P(h_i)} \leq \sum \frac{P(h_i/e) - P(h_i)}{(1 - P(h_i/e)) \cdot \pi_i (1 - P(h_i))}$$

ya que  $\pi_i (1 - P(h_i/e)) \leq 1$  y está en el denominador. A su vez, teniendo en cuenta que  $P(h_i) \leq 1$ , el segundo miembro es menor o igual que:

$$\sum \frac{P(h_i/e) - P(h_i)}{\pi_i (1 - P(h_i))} = \frac{\sum P(h_i/e) - \sum P(h_i)}{\pi_i (1 - P(h_i))}$$

Además  $\pi_i (1 - P(h_i)) > 1 - \sum P(h_i)$

ya que el producto de  $\pi_i (1 - P(h_i))$  se puede desarrollar operando:

$$\begin{aligned} 1 - \sum x_i + \sum \sum x_i x_j - \sum \sum \sum x_i x_j x_k + \dots = \\ = 1 - \sum x_i + \sum \sum x_i x_j (1 - \sum x_k) + \\ + \sum x_i x_j x_k \dots x_l (1 - \sum x_r + \dots) > 1 - \sum x_i \end{aligned}$$

siendo  $x_i = P(h_i)$ .

Y, por tanto, como las  $h_i$  son exclusivas  $\sum x_i \leq 1$  la suma anterior es:  $> 1 - \sum x_i = 1 - \sum P(h_i)$ , y, por tanto:

$$\begin{aligned} \sum \frac{P(h_i/e) - P(h_i)}{\pi_i (1 - P(h_i))} > \\ > \frac{\sum P(h_i/e) - \sum P(h_i)}{1 - \sum P(h_i)} \end{aligned}$$

como  $\sum P(h_i/e) \leq 1$ , este valor es  $\leq 1$  y por tanto:

$$\begin{aligned} \sum CF(h_i/e) = \\ \sum \frac{P(h_i/e) - P(h_i)}{1 - P(h_i)} \leq 1 \end{aligned}$$

Los criterios antes descritos constituyen una base para plantear una verificación de los valores introducidos por los expertos, que deberán estar entre  $-1$  y  $1$  y en ningún caso, para conjuntos de hipótesis mutuamente excluyentes, la suma debe rebasar la unidad.

### 4.3.1.3. Acumulación de evidencia

En las redes inferenciales, como ya se ha indicado, se produce el soporte de una hipótesis por acumulación de sucesivas evidencias a favor y en contra; por ello, un modelo de razonamiento impreciso debe proponer una forma de acumular la evidencia y de transmisión de ésta a lo largo de la red. Se requiere, por tanto, definir la forma de evaluar  $CF(h/e_1 \wedge e_2)$  en base a  $CF(h/e_1)$  y  $CF(h/e_2)$ .

Asimismo, aunque la red inferencial relacione un conjunto de hechos básicos, el tipo de pregunta puede afectar a composiciones de ellos; para producir respuestas relativas a hechos compuestos, en base al conocimiento sobre hechos básicos, es preciso definir un método de evaluar  $CF(h_1 \wedge h_2/e)$  y  $CF(h_1 \vee h_2/e)$ .

Las propuestas de los autores de MYCIN para estos conceptos se basan en el manejo de las medidas MC y MI; son las siguientes:

$$\begin{aligned}
 MC(h, e_1 \wedge e_2) & \left\{ \begin{aligned} &= 0, \text{ si } MI(h, e_1 \wedge e_2) = 1 \\ &= MC(h, e_1) + MC(h, e_2) - \\ &\quad - MC(h, e_1) \cdot \\ &\quad \cdot MC(h, e_2) \\ &= 1 \text{ si } MC(h, e_1) = 1 \\ &\quad \text{o } MC(h, e_2) = 1 \end{aligned} \right. \\
 MI(h, e_1 \wedge e_2) & \left\{ \begin{aligned} &= 0 \text{ si } MC(h, e_1 \wedge e_2) = 1 \\ &= MI(h, e_1) + MI(h, e_2) - \\ &\quad - MI(h, e_1) \cdot \\ &\quad \cdot MI(h, e_2) \\ &= 1 \text{ si } MI(h, e_1) = 1 \\ &\quad \text{o } MI(h, e_2) = 1 \end{aligned} \right.
 \end{aligned}$$

Ambas definiciones de composición cumplen la propiedad de que sus resultados son independientes del orden en que se da la evidencia a componer:

$$MC(h, e_1 \wedge e_2) = MC(h, e_2 \wedge e_1)$$

$$MI(h, e_1 \wedge e_2) = MI(h, e_2 \wedge e_1)$$

Asimismo, la regla de composición definida es monótona, ya que cuando el nivel de credibilidad aportado por una de las evidencias es nulo, el nivel de credibilidad compuesto resultante es el correspondiente a la otra.

Los valores resultantes de la composición cumplen la propiedad de estar limitados superiormente con la unidad. En efecto,

$$\text{Si } MC(h, e_1) \leq 1 \text{ y } MC(h, e_2) \leq 1$$

el valor de

$$MC(h, e_1 \wedge e_2) \leq 1$$

Para ello basta comprobar que  $1 - MC(h, e_1 \wedge e_2) \geq 0$

$$\begin{aligned}
 1 - MC(h, e_1 \wedge e_2) &= 1 - MC(h, e_1) - \\ & - MC(h, e_2) + (1 - MC(h, e_1)) \cdot \\ & = (1 - MC(h, e_1)) \cdot (1 - MC(h, e_2)) \geq 0
 \end{aligned}$$

ya que por hipótesis  $MC(h, e_1) \leq 1$  y  $MC(h, e_2) \leq 1$ .

Estas propiedades de conmutatividad, monotonía y límite superior, muestran que esta forma de componer medidas de credibilidad produce valores numéricos coherentes con la forma cualitativa de entender la acumulación de evidencia.

Sin embargo, las formas de acumulación prejuzgan la independencia de  $e_1$  y  $e_2$ , ya que si  $MC(h, e_1) = 0.8$  y  $MC(h, e_2) = 0.6$ , pero ocurre por ejemplo que  $e_1 \rightarrow e_2$ , la acumulación de evidencia por la fórmula sería optimista.

Para acumular las medidas de credibilidad en fórmulas compuestas se define:

$$MC(h_1 \wedge h_2, e) = \min(MC[h_1, e], MC[h_2, e])$$

$$MI(h_1 \wedge h_2, e) = \max(MI[h_1, e], MI[h_2, e])$$

Se adopta, por tanto, como criterio de credibilidad para la conjunción el nivel de credibilidad mínimo de sus componentes, para el nivel de incredibilidad se opta por el criterio simétrico tomando la máxima de las incredibilidades de sus componentes.

Esta definición, que puede considerarse acorde con la forma de entender la composición, se ajusta, en el límite, con la definición semántica en lógica bivalente, ya que en ella, en el caso en que cualquiera es falso el significado de la conjunción es falso, lo que es correlativo con el valor nulo atribuido por la fórmula anterior.

• Asimismo, en línea con lo anterior, para la disyunción se define:

$$MC(h_1 \vee h_2, e) = \max(MC[h_1, e], MC[h_2, e])$$

$$MI(h_1 \vee h_2, e) = \min(MI[h_1, e], MI[h_2, e])$$

#### 4.3.1.4. Encadenamiento de niveles de certidumbre

En el proceso de obtención de factores de certidumbre, una hipótesis con-

firmada por una evidencia previa puede figurar a su vez como evidencia confirmativa de otra; sin embargo, hasta ahora las fórmulas definidas presuponen un grado de evidencia unidad en  $e_1$  y en el caso de encadenamiento puede ocurrir que la hipótesis que actúa como evidencia para otra tenga factor de certidumbre distinto de uno. El problema se plantea por tanto de forma siguiente:

• Se trata de obtener  $MC(h, s)$  y  $MI(h, s)$  supuesto conocido:

•  $CF(s, E)$  factor de certidumbre de  $s$  como consecuencia de la evidencia global  $E$ .

•  $MC'(h, s)$  y  $MI'(h, s)$  grados de credibilidad a favor y en contra supuesto que  $s$  es totalmente cierto.

• La propuesta del modelo MYCIN es aplicar a  $MC'$  y  $MI'$  los factores de certidumbre de  $s$  respecto de  $E$  siempre que sean positivos para obtener los valores correspondientes de  $h$  respecto de  $s$ .

$$MC(h, s) = MC'(h, s) \cdot \max\{0, CF[s, e]\}$$

$$MI(h, s) = MI'(h, s) \cdot \max\{0, CF[s, e]\}$$

• Este planteamiento multiplicativo de la transmisión de grados de certidumbre es consistente con el concepto de implicación total en el límite, pero implica que al proponer el grafo de relaciones antecedente-consecuente (red inferencial) no existan circuitos que hagan que una afirmación sea a la vez consecuente y antecedente de sí misma, lo que produciría una serie indefinida de valores de  $CF$  en los nodos afectados por el circuito. Sobre esta base el motor de inferencia puede construirse según distintos procedimientos:

— Hacia adelante: partiendo de la evidencia de los hechos iniciales se avanza ordenadamente en la red inferencial actualizando los factores de certidumbre. Una vez actualizada toda la red pueden obtenerse las respuestas a las preguntas por simple búsqueda.



— Hacia atrás: partiendo de las preguntas, se investiga en la red mediante un árbol de búsqueda los elementos de evidencia a favor y en contra, hasta que se acumula evidencia suficientemente estable, en cuyo caso se termina el proceso.

Tanto en un caso como en otro son necesarias estrategias de control que limiten la explosión combinatoria en los procesos de búsqueda.

#### 4.3.2. EL METODO PROSPECTOR

##### 4.3.2.1. Definiciones

El método de este sistema, propuesto en 1976 por Duda Hart y Nilsson, está en línea con las técnicas bayesianas subjetivas propuestas en MYCIN e incluye variantes conceptuales significativas en cuanto a las medidas a proporcionar por los expertos y su forma de proceso.

Como medida del grado de implicación se utiliza la disparidad (odds) de una hipótesis, definida a partir de las probabilidades bayesianas. A continuación se introduce este concepto y sus propiedades.

Por el Teorema de Bayes:

$$P(h/e) = \frac{P(h/e) \cdot P(h)}{P(e)} \quad [1]$$

Siendo como el caso de MYCIN:

$P(h/e)$  la probabilidad de que ocurra  $h$  supuesta la evidencia  $e$

$P(h)$ ,  $P(e)$  probabilidades previas de ocurrencia de  $h$  y  $e$ .

Si se denomina  $\sim h$  la hipótesis negación de  $h$ , análogamente:

$$P(\sim h/e) = \frac{P(e/\sim h) \cdot P(\sim h)}{P(e)} \quad [2]$$

Duda, Hart y Nilsson utilizan como medida análoga al factor CF en MY-

CIN la relación entre probabilidades de ocurrencia a favor y contra  $h$ . Esta relación denominada **disparidad** se define como:

$$d(h) = \frac{P(h)}{P(\sim h)} = \frac{P(h)}{1 - P(h)}$$

este valor es la disparidad a priori la disparidad a posteriori referida a los sucesos condicionados por  $e$  es:

$$d(h/e) = \frac{P(h/e)}{P(\sim h/e)} =$$

$$= \frac{P(h/e) \cdot P(h)}{P(\sim h/e) \cdot P(\sim h)} = \lambda_e d(h)$$

siendo  $\lambda_e$  una relación que mide el grado relativo de verosimilitud de  $e$  en las hipótesis de que ocurra  $h$  y  $\sim h$

$$\lambda_e = \frac{P(e/h)}{P(e/\sim h)}$$

Por tanto, supuesta conocida la disparidad de una hipótesis, su nuevo valor como consecuencia de la introducción (conocimiento) de nueva evidencia  $e$  cuya relación con  $h$  se valora por  $\lambda_e$ , viene dado por:

$$d(h/e) = \lambda_e \cdot d(h)$$

Esta fórmula aporta, por tanto, un procedimiento sistemático y sencillo de actualización de la disparidad supuesto que  $e$  se verifica; por tanto, también es sencillo actualizar las probabilidades, ya que por definición:

$$P(h/e) = \frac{d(h/e)}{1 + d(h/e)}$$

De manera similar en el caso de que  $e$  es falsa:

$$d(h/\sim e) = \bar{\lambda} d(\sim h)$$

siendo:

$$\bar{\lambda} = \frac{P(\sim e/h)}{P(\sim e/\sim h)}$$

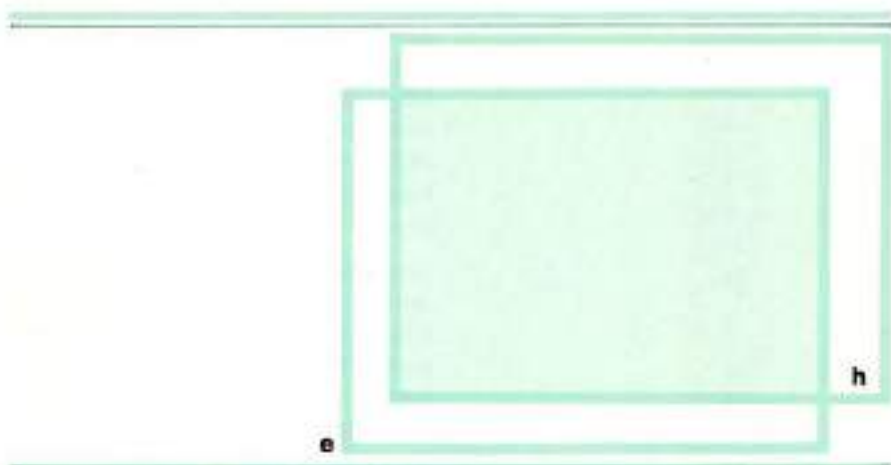
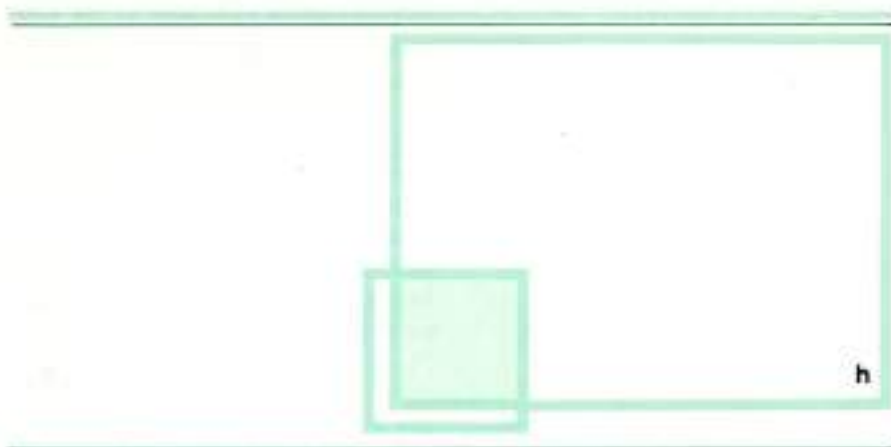
Los valores de  $\lambda_e$  y  $\bar{\lambda}_e$  son las medidas adoptadas en este sistema del grado de implicación entre e y h. En efecto:

— Si  $\lambda_e$  es muy alto, indica que la mayor parte de los casos en que ocurre e se producen cuando se verifica h, y, por tanto, hay pocos casos relativamente de ocurrencia de e sin que ocurra h, lo que indicaría que e es prácticamente suficiente para que ocurra h.

$$\approx (e \rightarrow h)$$

Esto puede ilustrarse con la figura adjunta, en donde sobre los casos posibles en el universo, se acotan mediante diagramas de Venn los subconjuntos en que se verifican e y h. Cuando se da este caso, e está prácticamente contenido en h, es decir, hay pocos casos en que e es verdad fuera de h.

— En cambio si  $\bar{\lambda}_e$  es próximo a cero, indicaría que la mayor parte de los casos en que ocurre  $\sim e$  coinciden con casos en que ocurre  $\sim h$ , pudiendo haber casos en que ocurre  $\sim e$  y h y, por tanto, hay pocos casos de h fuera de e como se indica en la figura. En este caso, prácticamente e es necesario para h:  $\approx h \rightarrow e$ .





Ambos valores deben ser aportados por los expertos, ya que uno mide el grado de implicación de e sobre h y el otro el grado de contra implicación. Sin embargo, asumiendo los grados de certidumbre como probabilidades, ambos valores de  $\lambda$  están relacionados y, por tanto, si se da un valor se da implícitamente el otro. En efecto,

$$\begin{aligned}\bar{\lambda}_e &= \frac{P(\sim e/h)}{P(\sim e/\sim h)} = \frac{1 - P(e/h)}{1 - P(e/\sim h)} = \\ &= \frac{1 - \lambda_e P(e/\sim h)}{1 - P(e/\sim h)}\end{aligned}$$

Despejando  $P(e/\sim h)$ :

$$\bar{\lambda}_e - \lambda_e \cdot P(e/\sim h) = 1 - \lambda_e P(e/\sim h)$$

$$P(e/h) \cdot (\lambda_e - \bar{\lambda}_e) = 1 - \lambda_e$$

Por tanto:

$$P(e/\sim h) = \frac{1 - \bar{\lambda}_e}{\lambda_e - \bar{\lambda}_e} \quad [5]$$

$$y \quad P(e/h) = \frac{1 - \bar{\lambda}_e}{\lambda_e - \bar{\lambda}_e} \lambda_e \quad [6]$$

Las fórmulas [5] y [6] muestran las relaciones entre cuatro datos a proporcionar por los expertos como imagen de su forma de entender e y h de los que basta proporcionar dos, ya que los otros resultan como consecuencia. Este es uno de los inconvenientes de utilizar una medida de certidumbre basada tan directamente en la probabilidad, ya que afirmar algo equivale a negar la fórmula complementaria, lo que en muchos casos del conocimiento real no es cierto. Esto es una desventaja respecto del concepto planteado por MYCIN, aunque los valores de  $\lambda$  que oscilan entre 0 e  $\infty$  pueden ser fácilmente correlativos de la forma de entender de los expertos.

#### 4.3.2.2. Acumulación de evidencia

Cuando en favor de una hipótesis h puede haber distintas evidencias  $e_1, e_2, e_3, \dots$  que incidan a favor o en contra, la manera de acumular evidencias es la siguiente:

— Si se suponen independientes respecto de h y  $\sim h$  las distintas evidencias  $e_1, e_2, \dots, e_n$  y con certidumbre total, puede formularse; siguiendo el cálculo de probabilidades:

$$P(e_1, e_2, \dots, e_n/h) = \prod_{i=1, k} P(e_i/h)$$

$$P(e_1, e_2, \dots, e_n/\sim h) = \prod_{i=1, k} P(e_i/\sim h)$$

De acuerdo con ello, la disparidad puede formularse como:

$$D(h/e_1, e_2, \dots, e_n) = \frac{P(e_1, e_2, \dots, e_n/h)}{P(e_1, e_2, \dots, e_n/\sim h)}$$

$$\cdot D(h) = \frac{P(e_1/h)}{P(e_1/\sim h)} \cdot \frac{P(e_2/h)}{P(e_2/\sim h)} \dots$$

$$\dots \frac{P(e_n/h)}{P(e_n/\sim h)} \cdot D(h) = [\prod_i \lambda_i] \cdot D(h)$$

$$\text{siendo } \lambda_i = \frac{P(e_i/h)}{P(e_i/\sim h)}$$

la disparidad aislada de cada evidencia.

Por tanto, cuando las evidencias son independientes, la manera de actualizar las  $\lambda$  tanto a favor como en contra es el producto del valor previo y el nuevo, lo que constituye un mecanismo muy operativo.

Para el caso de hipótesis compuestas, los autores de PROSPECTOR proponen combinar las probabilidades mediante las conocidas fórmulas de lógica difusa:

$$P(h_1 \wedge h_2/e) = \min [P(h_1/e), P(h_2/e)]$$

$$P(h_1 \vee h_2/e) = \max [P(h_1/e), P(h_2/e)]$$

Esta definición es criticable desde un punto de vista estrictamente probabilístico, pero no tanto si se considera que el sentido real de las probabilidades subjetivas es el de medida del grado de certidumbre y, por tanto, su combinación, de acuerdo con las fórmulas anteriores, asegura su coherencia, en el límite, con la forma de componer distribuciones bivalentes en cálculo proposicional.

Como ya se ha indicado al tratar el sistema MYCIN, la evidencia que soporta una hipótesis no tiene por qué ser completamente cierta; de hecho el nivel de información puede soportar parcialmente la hipótesis  $e$ . Se precisa, por ello, definir fórmulas para niveles de certidumbre intermedios en la evidencia de soporte. Cuando la evidencia de  $e$  es intermedia respecto de la evidencia total  $E$  hay dos grados de certidumbre extremos:

$$P(e/E) \text{ y } P(\sim e/E)$$

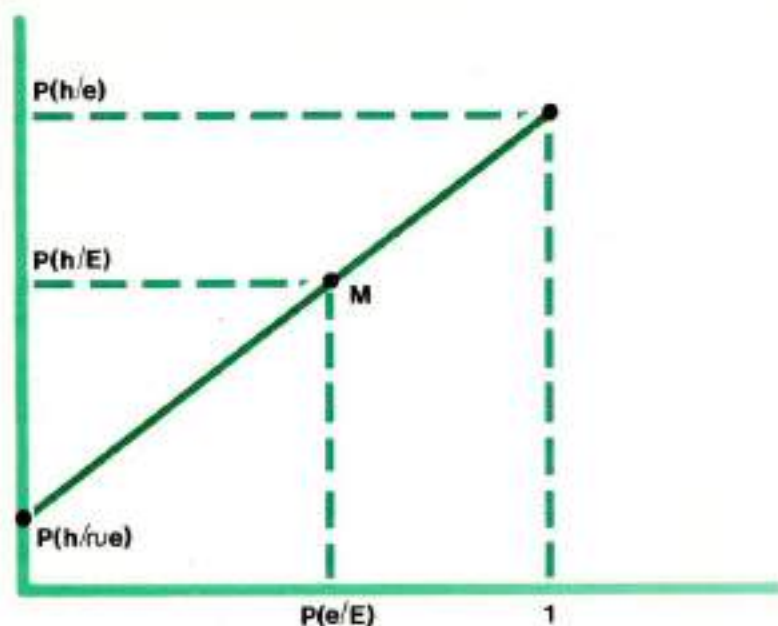
una hipótesis razonable de fijación del grado de certidumbre de  $h$  respecto de  $E$  sería aplicar la fórmula de probabilidades compuestas:

$$P(h/E) = P(h/e) \cdot P(e/E) + P(h/\sim e) \cdot P(\sim e/E)$$

Si se asume que estos grados de certidumbre son probabilidades, la fórmula anterior define la probabilidad  $P(h/E)$  como interpolada linealmente entre los valores extremos correspondientes a  $e$  verdadero o falso, de acuerdo con la figura adjunta.

Sin embargo esta relación lineal no tiene por qué cumplirse entre los valores introducidos, por estimación intuitiva de los expertos y usuarios, ya que éstos proporcionan:

- Valores a priori de  $P(e/E)$ : el usuario afirma el grado de certidumbre del hecho  $e$  entre los datos de entrada.



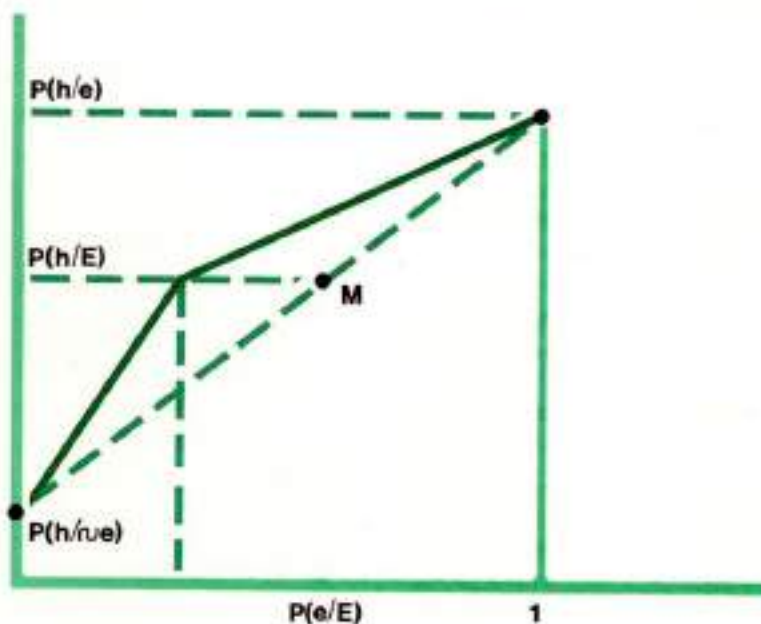
- El experto, por otra parte, aporta valores de  $\lambda$  y  $\bar{\lambda}$  para definir la regla  $e \rightarrow h$  y, por tanto,  $P(h/\sim e)$  y  $P(h/e)$ .
- El usuario, por otro lado, puede aportar  $P(h/E)$  como dato de entrada si  $h$  resulta también ser un hecho inicial.

Esto obligaría, caso de que se utilizara la fórmula lineal, a que los valores de  $P(e/E)$  y  $P(h/E)$  coincidieran en  $M$ , lo que no tiene por qué ocurrir en el caso general. Esta circunstancia de incoherencia entre probabilidades a priori y grado de fuerza de una regla, hace que sea necesario prescindir de la hipótesis de ley lineal, apelando a leyes de tipo poligonal.

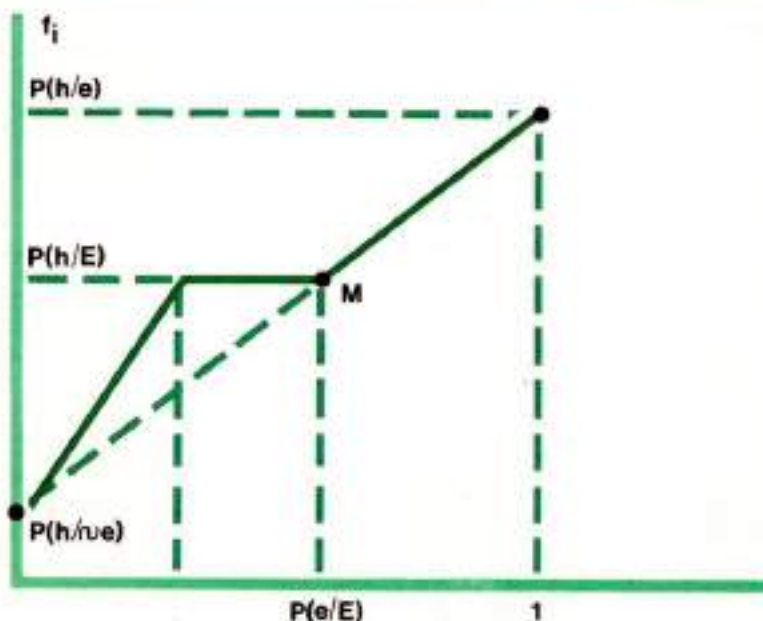


Los autores de PROSPECTOR proponen como acciones posibles de ajuste:

— Transformación de la función lineal en una quebrada, de manera que el punto de quiebro coincida con la probabilidad a priori atribuida, tanto a la evidencia como al hecho  $h$  de partida. Sin embargo, se mantienen los valores  $P(h/\sim e)$  y  $P(h/e)$ .



— Transformación de la función lineal en quebrada, pero con escalón constante, que no respeta  $P(h/\sim e)$ . Una razón en favor de esta hipótesis sería considerar que si existen discrepancias en el punto M, ello implica que el experto no está muy seguro de  $P(h/\sim e)$  y a partir de este punto se mantiene la probabilidad de  $h$  constante independiente de  $e$ .



— Transformación de la función lineal en quebrada, pero con escalón que respeta  $P(h/\sim e)$ ; en este caso, se recorta el criterio anterior admitiendo sólo una zona de valor constante de  $P(h)$ , de amplitud igual al intervalo de diferencia entre  $P(e/E)$  a priori y  $P(e/E)$  calculada de acuerdo con el valor a priori de  $P(h/E)$ . Se considera para esa zona un tramo horizontal.

Para las hipótesis no afectadas por estos problemas de inconsistencia con los valores a priori se utiliza la regla de composición lineal bayesiana. Por tanto, una vez realizado este análisis:

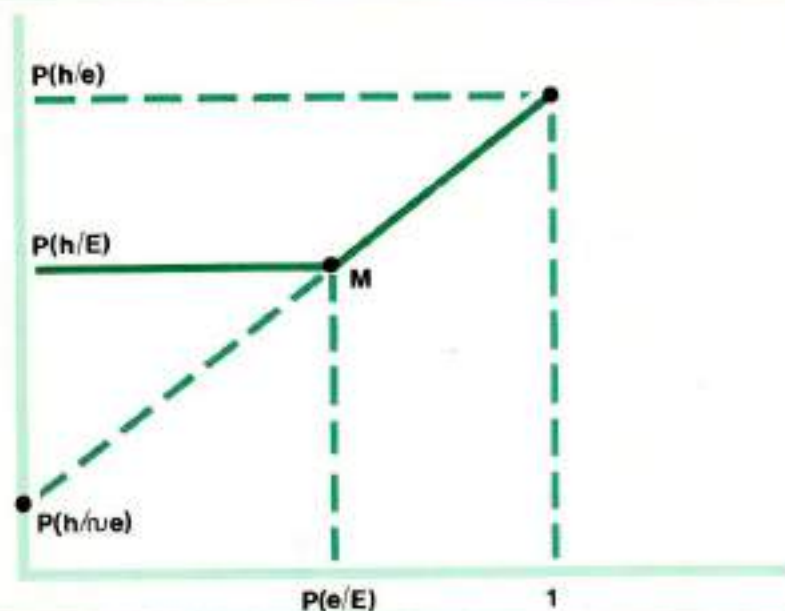
— es posible, por progreso ordenado en el grafo, obtener la evidencia acumulada a favor y en contra de una hipótesis mediante las fórmulas de 4.3.2.2.

— Es posible, mediante las correspondientes reglas de combinación (normal o modificada según los casos), obtener los valores anteriores para valores de evidencia intermedia.

En base a estos conceptos puede, por tanto, propagarse la evidencia en una red mediante el oportuno motor inferencial que tenga en cuenta las conclusiones anteriores.

Cuando la red es compleja, no es conveniente para el motor inferencial un método de propagación generalizada hacia adelante de la evidencia. Los autores de PROSPECTOR plantearon para su sistema un proceso que arrancaba de la pregunta y realizaba una búsqueda en la red hacia atrás hasta que la variación en la evidencia acumulada no era significativa. El procedimiento se basaba en el mantenimiento de una lista de preguntas pendientes de obtener su valoración del grado de certidumbre. Esta lista se inició con la pregunta de partida. Para cada elemento de la lista puede ocurrir:

- que sea de evidencia inicial asignada, en cuyo caso se quita de la lista y se propaga su efecto en la subred generada por el proceso;



- que sea preguntable; es decir, que pueda demandarse a la persona que ha planteado la pregunta inicial como si se tratara de una aclaración;

- que no esté en los casos anteriores, o bien, estando en el caso segundo, el usuario no la contesta. Entonces debe ser deducible, en cuyo caso se introduce la pregunta en la subred de propagación y se introducen en la lista los nodos antecedentes en la red. Si no tiene nodos antecedentes no existe información en la red para contestar la pregunta y termina el proceso. Si tiene antecedentes, se vuelve al inicio del procedimiento para tratar otro elemento de la lista.

El proceso termina cuando existe algún nodo imposible o cuando, como ya se ha indicado, la variación de evidencia en la pregunta no es considerable.

Este método tiene la ventaja de que se realiza la propagación de evidencia exclusivamente en la subred pertinente a la pregunta planteada, tanto por tener relación con ella como por la importancia de esa relación.

Un factor de economía en el motor de inferencia es la estrategia de control

basada en el orden en que se toman las preguntas pendientes de la lista; los autores de PROSPECTOR emplearon un heurístico que evaluaba la incidencia posible de la pregunta elegida en la respuesta a obtener.

#### 4.3.2.3. Consideraciones sobre los problemas de consistencia

Dada su condición de modelo plausible de razonamiento más que teoría totalmente rigurosa, en el planteamiento inicial de PROSPECTOR existen lagunas que pueden tener impacto desfavorable en sus resultados. La forma de resolverlo en la implementación operativa es con modificaciones locales que sustituyen el efecto del modelo estándar por métodos locales. Tal es el caso ya comentado de la inconsistencia entre los factores  $\lambda$  de implicación y los grados de certidumbre a priori. Otro aspecto criticable, ya comentado, es la composición de probabilidades con fórmulas tomadas de lógica difusa en lugar de las fórmulas conocidas de composición de probabilidades tomadas del álgebra de conjuntos.



Como ya se ha indicado al tratar MYCIN, tanto éste como PROSPECTOR incorporan formulaciones que presuponen la no existencia de circuitos en la red inferencial, ya que ambos procedimientos provocarían variaciones no verosímiles de los factores de certidumbre o índices de verosimilitud, en el caso de circuitos.

Asimismo, se presenta en PROSPECTOR el problema de que la acumulación de evidencia se apoya en la asunción de independencia respecto de la hipótesis estudiada de los elementos de evidencia.

En efecto, la fórmula de acumulación se apoya en presuponer para dos evidencias:  $e_1$  y  $e_2$ :

$$P(e_1, e_2/h) = P(e_1/h) \cdot P(e_2/h)$$

$$P(e_1, e_2/\sim h) = P(e_1/\sim h) \cdot P(e_2/\sim h)$$

Partiendo de esto,

$$P(h/e_1, e_2) = \frac{P(e_1, e_2/h) \cdot P(h)}{P(e_1, e_2)}$$

$$= \frac{P(e_1/h) \cdot P(e_2/h) \cdot P(h)}{P(e_1) \cdot P(e_2)}$$

$$\cdot \frac{P(e_1) \cdot P(e_2)}{P(e_1, e_2)} = \frac{P(e_1/h) \cdot P(h)}{P(e_1)}$$

$$\cdot \frac{P(e_2/h) \cdot P(h)}{P(e_2)} = \frac{P(e_1) \cdot P(e_2)}{P(e_1, e_2) \cdot P(h)}$$

$$= \frac{P(h/e_1) \cdot P(h/e_2)}{P(h)} \cdot \frac{P(e_1) \cdot P(e_2)}{P(e_1, e_2)}$$

$$\text{ya que } P(h/e_1) = \frac{P(e_1/h) \cdot P(h)}{P(e_1)} \quad (1)$$

Por otro lado, habiéndose cuenta de la independencia condicional:

$$P(e_1, e_2) = P(e_1, e_2/h) \cdot P(h) +$$

$$+ P(e_1, e_2/\sim h) \cdot P(\sim h) = P(e_1/h) \cdot$$

$$\cdot P(e_2/h) \cdot P(h) + P(e_1/\sim h) \cdot$$

$$\cdot P(e_2/\sim h) \cdot P(\sim h)$$

Por el Teorema de Bayes:

$$P(e_1/h) \cdot P(h) = P(h/e_1) \cdot P(e_1)$$

Sustituyendo resulta:

$$P(e_1, e_2) =$$

$$= \frac{P(h/e_1) \cdot P(e_1) \cdot P(h/e_2) \cdot P(e_2)}{P(h)} +$$

$$+ \frac{P(\sim h/e_1) \cdot P(e_1) \cdot P(\sim h/e_2) \cdot P(e_2)}{P(\sim h)}$$

$$= P(e_1) \cdot P(e_2) \left[ \frac{P(h/e_1) \cdot P(h/e_2)}{P(h)} + \frac{P(\sim h/e_1) \cdot P(\sim h/e_2)}{P(\sim h)} \right]$$

Por tanto,

$$\frac{P(e_1) \cdot P(e_2)}{P(e_1, e_2)} = \frac{1}{\frac{P(h/e_1) \cdot P(h/e_2)}{P(h)} + \frac{P(\sim h/e_1) \cdot P(\sim h/e_2)}{P(\sim h)}}$$

$$\text{ya que } \left( D(h) = \frac{P(h)}{P(\sim h)} \right) \frac{P(h)}{P(h/e_1) \cdot P(h/e_2) + D(h) \cdot P(\sim h/e_1) \cdot P(\sim h/e_2)}$$

Sustituyendo:

$$P(h/e_1, e_2) = \frac{P(h/e_1) \cdot P(h/e_2)}{P(h/e_1) \cdot P(h/e_2) + D(h) \cdot P(\sim h/e_1) \cdot P(\sim h/e_2)}$$

Konolige, en base a esta expresión, ponía de relieve los problemas de inconsistencia de la hipótesis de independencia condicional. En efecto, si el experto define  $P(h/e_1)$ ,  $P(h/e_2)$  y  $P(h)$  ha definido totalmente  $P(h/e_1, e_2)$ , si con los valores dados el valor resultante es aceptable no hay problema; sin embargo, puede ocurrir que existiendo valores excesivamente altos de  $P(h/e_1)$  y  $P(h/e_2)$  no tenga que serlo tanto  $P(h/e_1, e_2)$  por existir evidencia redundante en  $\{e_1, e_2\}$ .

Cuando se detecta este caso es preciso hacer modificaciones a la medida, no sencillas, que en algunos casos quitan verosimilitud al modelo bayesiano, aunque localmente. Asimismo, esta hipótesis de independencia condicional produce riesgos cuando se trata de hipótesis exhaustivas y mutuamente excluyentes, es decir, cuando

$$P(h_1) + P(h_2) + P(h_3) = 1$$

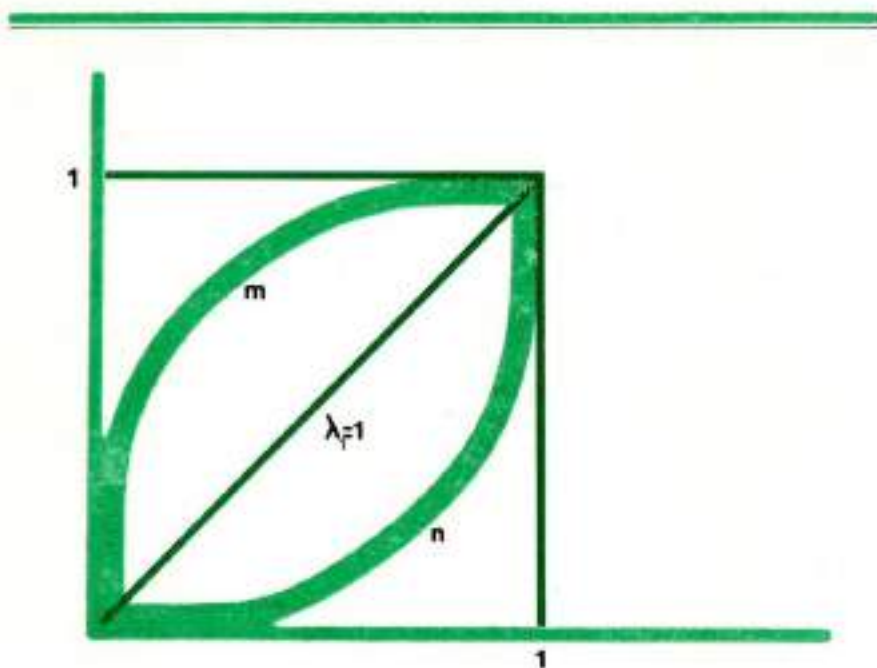
$$P(h_i \wedge h_j) = 0 \quad \forall i, j$$

Konolige demostró que las probabilidades condicionales

$$P(h_1/e_1, e_2) + P(h_2/e_1, e_2) + P(h_3/e_1, e_2) \neq 1$$

Por tanto, esta hipótesis de independencia condicional es inconsistente en este caso de exhaustividad y exclusividad. Sin embargo, hay que tener en cuenta que los Sistemas Expertos se plantean en dominios complejos y con el objetivo de producir respuestas sobre un número limitado de hipótesis, por tanto, no es fácil que este problema de inconsistencia se plantee en aplicaciones realistas, aunque podría aparecer, en pequeña medida, en forma de error en la estimación de los grados de certidumbre.

Finalmente hay otro aspecto problemático en el concepto de PROSPECTOR. Se trata del mecanismo multiplicativo de acumulación de evidencia, en donde pueden presentarse problemas locales de excesiva sensibilidad.



En efecto, según la fórmula de acumulación de evidencia la disparidad conjunta  $D(h/e_1, e_2)$  se formula:

$$D(h/e_1, e_2) = \lambda_1 - D(h)$$

$$\text{siendo } \lambda_1 = \frac{D(h/e_1)}{D(h)}$$

Poniendo en forma de probabilidades:

$$\begin{aligned} P(h/e_1, e_2) &= \frac{D(h/e_1, e_2)}{1 + D(h/e_1, e_2)} = \\ &= \frac{\lambda_1 D(h/e_2)}{1 + \lambda_1 D(h/e_2)} = \\ &= \frac{\lambda_1 P(h/e_2)}{1 + (\lambda_1 - 1) P(h/e_2)} \quad [2] \end{aligned}$$

Esta fórmula expresa la variación de la evidencia acumulada  $P(h/e_2)$  por introducción de  $e_1$ . Si se representa en la forma indicada en la figura, según el valor de  $\lambda_1$ , la curva representativa de [2] puede variar desde formas tipo  $m$  ( $\lambda_1$  alto) a tipo  $n$  ( $\lambda_1$  bajo) pasando por la recta correspondiente a  $\lambda_1 = 1$ .

Para casos extremos en las zonas indicadas se presentan variaciones excesivas del valor compuesto para pequeñas variaciones de  $P(h/e_2)$ . Cuando se da esta circunstancia, se adopta otra función de interpolación.

No obstante, las observaciones anteriores, el modelo general debidamente modificado, en aquellos vértices de la red en que las inexactitudes producen valores no admisibles, ha permitido desarrollar un modelo suficientemente complejo y representativo, como lo muestra la experiencia en el uso de este sistema.



### 4.3.3. FORMULACION MODIFICADA DE PROSPECTOR

En 1979 Konolige propuso una formulación alternativa del mecanismo inferencial de PROSPECTOR que obvia alguno de los inconvenientes detectados en la versión implementada. A continuación se describe el planteamiento de esta nueva formulación.

Se considera el universo de discurso del sistema caracterizado por un conjunto de conceptos básicos  $C_1, C_2, \dots, C_n$ , de forma que todo acontecimiento pertinente al tema se puede describir por una conjunción de afirmaciones o negaciones de estos conceptos. Existen, por tanto, los sucesos:

$$S_1: C_1 \wedge C_2 \wedge C_3 \wedge \dots \wedge C_n$$

$$S_2: \sim C_1 \wedge C_2 \wedge C_3 \wedge \dots \wedge C_n$$

$$S_3: C_1 \sim C_2 \wedge C_3 \wedge \dots \wedge C_n$$

$$S_n: \sim C_1 \wedge \sim C_2 \wedge \sim C_3 \wedge \dots \wedge \sim C_n$$

Si se conocen las probabilidades  $P(S_i)$  será posible, por tanto, por combinación de sus valores en la forma adecuada según el cálculo de probabilidades, obtener las probabilidades de sucesos compuestos (que por hipótesis podrán describirse por unión de los sucesos básicos  $S_i$ ).

El procedimiento de Konolige consiste en obtener una distribución verosímil de la distribución  $P(S_i)$ , que satisfaga las condiciones expuestas por los expertos. Los tipos de condiciones son (por razones de simplificación se denomina a partir de ahora  $P_i$  a  $P(S_i)$ ):

— generales:

$$\sum P_i = 1 \quad [1]$$

$$P_i \geq 0 \quad \forall i \quad [2]$$

— específicas:

• **probabilidades a priori impuestas:** es decir, se impone la propiedad de

que un suceso  $S$  tiene la probabilidad dada  $C$ :

Si  $S$  es un elemento  $K$  del universo de discurso:

$$P_k = C \quad [3]$$

Si  $S$  no es un elemento, por hipótesis debe poderse describir por unión de elementos, luego,

$$\sum_{k \in S} P_k = C \quad [3']$$

( $K \in S$  se refiere a elementos que forman parte del subdominio en que se verifica  $S$ ).

• **reglas de inferencia:**

Si el grado de implicación de  $A \rightarrow B$  es  $g$ , de acuerdo con la forma de representación de PROSPECTOR, ello significa:

$$P(B/A) = g = \frac{P(A \wedge B)}{P(A)}$$

$B$  y  $A$  se pueden describir a partir de los elementos del marco del discurso. Por tanto, debe ocurrir:

$$\sum_{i \in A \Delta B} P_i - g \cdot \sum_{k \in A} P_k = 0 \quad [4]$$

Como el denominador no debe ser nulo

$$\sum_{k \in A} P_k > 0 \quad [4']$$

Cualquier distribución de probabilidades que satisfaga las condiciones [1], [2], [3], [3'], [4] y [4'] es una distribución consistente, que es coherente con el conocimiento aportado por los expertos. Existen, por tanto, en caso general, varias soluciones posibles de distribución de  $P_i$ . Para elegir la más adecuada, Konolige introdujo un criterio: seleccionar aquella que minimiza la cantidad de información introducida respecto del conjunto de condiciones antes descrito, lo que significa, de

acuerdo con la teoría de la información: maximizar la entropía.

$$H(p) = -\sum P_i \cdot \log P_i \quad [6]$$

El motor inferencial será un procedimiento que halle máx.  $H(p)$  con las condiciones antes descritas.

Konolige propone un mecanismo heurístico en dos etapas:

— Obtención de una primera solución que satisfaga el conjunto de condiciones.

Esto puede hacerse mediante un algoritmo de programación lineal, ya que el conjunto de restricciones es lineal.

— Optimización de la solución factible obtenida mediante una versión diferencial de algoritmo simplex de forma que en cada etapa se maximiza la derivada de  $H(p)$ , de forma que se obtienen las variaciones de los valores de las variables que producen un incremento máximo de  $H(p)$ . Este proceso garantiza que en cada iteración  $K$ ,  $H^k(p) \geq H^{k-1}(p)$ . Sin embargo, no garantiza que se alcance el óptimo absoluto, pudiendo alcanzarse óptimos locales.

Cheeseman (1983) ha presentado una alternativa para computar la solución mediante multiplicadores de Lagrange.

El método propuesto obvia todos los inconvenientes de la formulación inicial, ya que una vez obtenida la distribución básica  $P_i$ , es sencillo obtener las probabilidades de cualquier pregunta describible en el marco del discurso, por agregación de las probabilidades básicas de los sucesos elementales que comprende el concepto de la pregunta.

El inconveniente del procedimiento es su prolijidad, que puede hacerlo inoperante cuando los conceptos básicos a manejar son superiores a 10 ó 15 ( $2^{10} = 1.024$  sucesos elementales cuyas probabilidades hay que obtener por la técnica algorítmica antes descrita).

Se ha discutido también, en este planteamiento, el criterio de maximizar la entropía, que constituye realmente una forma de introducir información. Este criterio representa que, si la única condición de subconjunto de  $n$  sucesos es  $\sum P_i = a$ , puede adjudicarse cada uno  $a/n$  como valor. Sin embargo, es sencillo durante el proceso de construcción de la base de conocimiento, obviar este inconveniente, ya que a medida que se analizan los resultados no deseables que provoca esta hipótesis, pueden neutralizarse de forma adecuada, por introducción de información en forma de condiciones o reglas.

#### 4.3.4. TECNICAS AVANZADAS DE REPRESENTACION

Los problemas de los modelos de razonamiento impreciso tratados hasta aquí, residen, en gran parte, en el intento de representar el nivel de certidumbre mediante un único valor; la forma de mejorar la calidad de la representación es aceptar un mayor número de parámetros para evaluar el grado de certidumbre de una regla. En esta línea está la teoría de la evidencia de Dempster Shafer, que utiliza dos valores que acotan un intervalo y la teoría de la posibilidad de Zadeh, que asocia una función de distribución de valores de posibilidad a toda proposición.

Asimismo, al margen de esta línea, empieza a desarrollarse una representación del razonamiento impreciso, no basado en la cuantificación del grado de certidumbre, sino en las lógicas denominadas no monótonas.

En la valoración de estos métodos no debe olvidarse que su idoneidad depende tanto de la capacidad y consistencia de su concepto de representación como de las posibilidades metodológicas de traslación al mismo, del conocimiento de los expertos. En este sentido, el método de Dempster Shafer aparece como un vehículo intere-

sante, ya que aporta un grado de complejidad no excesiva con relación a los métodos probabilísticos bayesianos y, por tanto, resulta factible de construcción, aportando en cambio un mayor nivel teórico en su concepto de representación. No ocurre lo mismo con el cálculo possibilístico de mayor nivel teórico, pero de más compleja validación, del que quizá podría hacerse una versión restringida que fuera de fácil construcción.

Hasta el momento no se han desarrollado sistemas expertos basados en esta clase de representaciones (aunque se han hecho aplicaciones del método de Dempster Shafer, no pueden considerarse estrictamente sistemas expertos). No obstante, se hace una descripción sucinta de los criterios que informan tales sistemas, porque pueden constituir una base futura interesante para su construcción.

##### 4.3.4.1. El método Dempster-Shafer

Este procedimiento parte de un marco de discernimiento  $U$ , constituido por el conjunto de descripciones posibles del mundo sobre el que se plantea la representación.

A partir de este marco puede definirse  $P(U)$ , conjunto de las partes de  $U$ , formado por todos los subconjuntos posibles de  $U$  entre los que está el conjunto vacío  $\emptyset$ . Dado que un predicado  $B$  definido en  $U$  define un subconjunto (el formado por todos aquellos elementos que hacen verdad  $B$ ), el conjunto  $P(U)$  representa, en cierta forma, el conjunto de afirmaciones que pueden hacerse sobre las descripciones incluidas en  $U$ .

Sobre esta base, una unidad de evidencia  $E$  se define por una distribución probabilística  $m(A)$  sobre los elementos de  $P(U)$ , de tal forma que:

$m(A)$  = Probabilidad de que se verifique la propiedad  $A$  (o de otra forma, probabilidad de que se presente un



suceso cuya descripción está incluida en A).

$$\text{Debe ocurrir: } \sum_{A \in \mathcal{P}(U)} m(A) = 1$$

y la probabilidad de subconjunto vacío debe ser nula:

$$m(\emptyset) = 0$$

Los elementos A tales que  $m(A) \neq 0$  se denominan puntos focales de la evidencia E. En base a la distribución m se definen dos medidas de certidumbre de cualquier afirmación en U.

La credibilidad de una afirmación A es la suma de probabilidades de todos los subconjuntos incluidos en A:

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

Dado que la propiedad de inclusión es una traducción, en términos de conjuntos, de la condición de implicación, la anterior definición puede expresarse también como:

— La credibilidad de A es la suma de probabilidades m de todas las afirmaciones que implican A.

La plausibilidad de A se define como la suma de probabilidades m de todos los subconjuntos de U, cuya intersección con A no es vacía:

$$\text{Pl}(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

El concepto de plausibilidad acumula, por tanto, todos los valores m de los elementos que sirven para obtener la credibilidad, ya que en ellos  $B \cap A = B$ , y, además, todos aquellos que tienen alguna descripción en común con A y por tanto tienen un cierto grado de implicación. De acuerdo con ello se verifica siempre:

$$\text{Pl}(A) \geq \text{Cr}(A) \quad [1]$$

Por tanto, dada una evidencia E, se valoran las distintas afirmaciones posibles por dos valores  $\text{Pl}(A/E)$  y  $\text{Cr}(A/E)$ , obtenidos de la forma antes descrita. Esta es una variación sustancial respecto de la caracterización por único valor, planteado en los métodos probabilísticos. En lo sucesivo se prescindirá de E en la notación de Pl y Cr por razones de simplificación. Además de la propiedad [1] pueden comprobarse fácilmente las siguientes propiedades:

$$\text{Pl}(\emptyset) = \text{Cr}(\emptyset) = 0 \quad [2]$$

$$\text{Cr}(U) = \sum(A) = 1 \quad [3]$$

$$\text{Pl}(U) = \sum m(A) = 1 \quad [4]$$

Además,

$$\begin{aligned} \text{Cr}(A) + \text{Cr}(\sim A) &= \\ = 1 \sum_{B \subset A} m(B) + \sum_{B \subset \sim A} m(B) &\leq 1 \quad [5] \end{aligned}$$

ya que esta suma sólo podría ser la unidad, en el caso en que no existiera ningún B tal que  $B \cap A = \emptyset$

$$\text{Pl}(A) + \text{Pl}(\sim A) = \sum_{B \cap A \neq \emptyset} m(B) + \sum_{B \cap \sim A \neq \emptyset} m(B) \geq 1 \quad [5]$$

ya que en cuanto exista algún B tal que  $B \cap A \neq \emptyset$  y  $B \cap \sim A \neq \emptyset$  con  $m(B) > 0$ , este valor se sumaría dos veces y como  $\sum m(B) = 1$ , la suma antes descrita tiene como límite inferior la unidad.

La propiedad [5] muestra cómo el concepto de credibilidad sirve para modelar el concepto de ignorancia, al propio tiempo que el de certidumbre, ya que en el caso en que la evidencia E aportara exclusivamente valores  $m(B) \neq 0$ , para elementos tales que  $B \cap A \neq \emptyset$  y  $B \cap \sim A \neq \emptyset$ , tanto la medida  $\text{Cr}(A)$  como  $\text{Cr}(\sim A)$  serían nulas; los valores de plausibilidad de ambas serían la unidad. La situación de falta de información respecto de una afirmación, aparece representada por tanto por el par (0,1). En cambio la si-

tuación de máxima información respecto de una afirmación A correspondería al caso en que la distribución  $m(A)$  incluye como puntos focales precisamente todas las descripciones contenidas en A. En este caso  $\text{Cr}(A)$  estaría valorada por la suma de probabilidades de los elementos incluidos en A y por tanto sería exactamente la probabilidad de A:  $\text{Pr}(A)$ .

Asimismo, en este caso de máxima información, el valor de plausibilidad sería también la suma de probabilidades de los elementos incluidos en A, ya que dado el carácter de la distribución  $m(A)$ , referida a conjuntos de un único elemento, no existen conjuntos cuya intersección con A y  $\sim A$  sea distinta del vacío.

Por tanto, en el caso de máxima información en la evidencia E,

$$\text{Pl}(A) = \text{Cr}(A) = \text{Pr}(A)$$

La representación propuesta por DEMPS-TER y SHAFER, aun aceptando que el grado de certidumbre de una información está medido por su probabilidad, incluye una matización que representa el grado de detalle de la evidencia de partida, de forma que como resultado se obtiene el par:

$$[\text{Cr}(A), \text{Pl}(A)]$$

cuyos valores, según el grado de precisión de la distribución de E, oscilan desde la máxima ignorancia correspondiente al par (0,1), a la máxima información definida por el intervalo con un único valor  $[\text{Pr}(A), \text{Pr}(A)]$ .

Otra propiedad de ambos conceptos es:

$$\text{Cr}(A) + \text{Pl}(\sim A) = 1 \quad [7]$$

$$\text{En efecto, } \text{Cr}(A) = \sum_{B \subset A} m(B) = \sum_{B \cap \sim A = \emptyset} m(B)$$

$$\text{Pl}(\sim A) = \sum_{B \cap \sim A \neq \emptyset} m(B)$$

$$\text{Por tanto, } \text{Cr}(A) + \text{Pl}(\sim A) = \sum_{B \subset U} m(B) = 1$$

La obtención de los grados de credibilidad y plausibilidad asociados a proposiciones compuestas para un mismo cuerpo de evidencia, se obtienen aplicando la fórmula general. En efecto:

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

$$Cr(A \wedge B) = \sum_{A, C \subset A \wedge B} m(A)$$

$$Pl(A \vee B) = \sum_{A \cap (A \vee B) \neq \emptyset} m(A, 2)$$

$$Pl(A \wedge B) = \sum_{A \cap (A \wedge B) \neq \emptyset} m(A)$$

Para negación, las fórmulas ya conocidas:

$$Cr(\sim A) = 1 - Pl(A) \quad Pl(\sim A) = 1 - Cr(A)$$

Ejemplo:

Sea el marco de discernimiento definido por las proposiciones mutuamente exclusivas  $P_1, P_2, P_3, \dots, P_8$ .

Sea una asignación de soporte  $m(A_i)$  procedente de un cuerpo de evidencia cuyos puntos focales son:

Punto focal	Descripción	$m(A_i)$
$A_1$	$P_1$	0.1
$A_2$	$P_2 \vee P_3$	0.2
$A_3$	$P_3 \vee P_4$	0.3
$A_4$	$P_5 \vee P_7 \vee P_8$	0.2
$A_5$	$P_3 \vee P_6 \vee P_7$	0.2

Para definir el soporte de esta evidencia a la afirmación  $B = P_1 \vee P_3 \vee P_6$  se calculará una medida de credibilidad:

$$Cr(B) = \sum_{A_i \subset B} m(A_i) = 0.1$$

ya que la única afirmación focal contenida en B es  $A_1$ . En cambio la medida de plausibilidad sería:

$$Pl(B) = \sum_{A_i \cap B \neq \emptyset} m(A_i) = 0.1 + 0.2 + 0.3 + 0.2 = 0.8$$

Por tanto, la afirmación B viene caracterizada por el par (0.1; 0.8).

La afirmación  $\sim B = \sim P_1 \wedge \sim P_3 \wedge \sim P_6$  por

$$Cr(\sim B) = \sum_{A_i \subset \sim B} m(A_i) = m(A_4) = 0.2$$

$$Pl(\sim B) = \sum_{A_i \subset \sim B} m(A_i) = m(A_2) + m(A_3) + m(A_4) + m(A_5) = 0.2 + 0.3 + 0.2 + 0.2 = 0.9$$

y, por tanto,  $\sim B$  (0.2; 0.9) lo que a la vez muestra las fórmulas:

$$Cr(B) = 1 - Pl(\sim B)$$

$$Pl(B) = 1 - Cr(\sim B)$$

El motor de inferencia de este tipo de representación debe ser capaz de:

- Componer la evidencia procedente de diversas fuentes E;
- en la distribución resultante evaluar de la forma indicada en el ejemplo, el intervalo de soporte para las distintas cuestiones.

Como regla para componer dos cuerpos de evidencia, Dempster (1967) propuso la formulación:

$$m(C_k) = \frac{\sum_{A_i \cap B_j = C_k} m_1(A_i) \cdot m_2(B_j)}{\sum_{A_i \cap B_j \neq \emptyset} m_1(A_i) \cdot m_2(B_j)} \quad [8]$$

siendo  $m(C_k)$  la distribución resultante y  $m_1(A_i), m_2(B_j)$  las distribuciones a componer. La fórmula anterior atribuye a cada afirmación  $C_k$  la suma de los productos de las asignaciones  $m_1$  y  $m_2$  correspondientes a las afirmaciones cuya conjunción es  $C_k$ .

La formulación anterior tiene el inconveniente de que atribuye al conjunto vacío exceso de probabilidad, cuando existen muchos casos de evidencia disjunta entre ambas asignaciones.

Dempster propone asignar valor nulo a  $m(\emptyset)$  y redistribuir la probabilidad asignada por la regla a  $\emptyset$ , proporcionalmente a los  $m(C_k)$  obtenidos. Es decir, si K es el valor asignado a  $m(\emptyset)$ , se hace  $m(\emptyset) = 0$  y se afectan a los

$m(C_k)$  de  $\frac{1}{1-k}$  lo que equivale a

sustituir [8] por  $m(C_k) = 0$

$$m(C_k) = \frac{\sum m_1(A_i) \cdot m_2(B_j)}{\sum m_1(A_i) \cdot m_2(B_j)} \quad [9]$$

extendida la sumatoria numerador a todos los  $A_i, B_j$  cuya intersección es  $C_k$  y a todos aquellos tales que  $A_i \cap B_j \neq \emptyset$  en el denominador.

Esta propuesta ha sido discutida por Zadeh, el creador de la lógica difusa, ya que constituye una forma de ocultar un aspecto disonante de la teoría. Para ello Zadeh plantea el contraejemplo siguiente:

Sea un tema de diagnóstico sobre una persona. El primer cuerpo de evidencia es propuesto por el médico A cuya distribución es:

$m(A_1)$ : 0.99, siendo  $A_1$ : el paciente tiene meningitis.

$m(A_2)$ : 0.01, siendo  $A_2$ : el paciente tiene tumor cerebral.

El segundo médico B propone otro cuerpo de evidencia cuyos puntos focales son:

$m(B_1)$ : 0.99, siendo  $B_1$ : el paciente tiene una embolia.

$m(B_2)$ : 0.01, siendo  $B_2$ : el paciente tiene tumor cerebral.

La composición de ambos cuerpos de evidencia, mediante la regla de Dempster Shafer, dado que  $A_1 \cap B_1$  tiene intersección nula, una vez normalizada daría soporte 1 a la hipótesis de tumor cerebral, siendo nulo el soporte de las otras hipótesis. Evidentemente esta crítica es correcta, sin embargo, cuando se analiza un tema es difícil que dos fuentes de conocimiento produzcan distribuciones  $m_1$  y  $m_2$  con un factor k importante, ya que no ocurre que al estudiar un tema se incluyan conceptos excesivamente próximos. Debe, sin embargo, controlarse que los coeficientes de normalización no sean excesivamente distintos de la unidad.

La distribución resultante se simboliza  $m_1 \oplus m_2$ .



Se demuestra en forma sencilla:

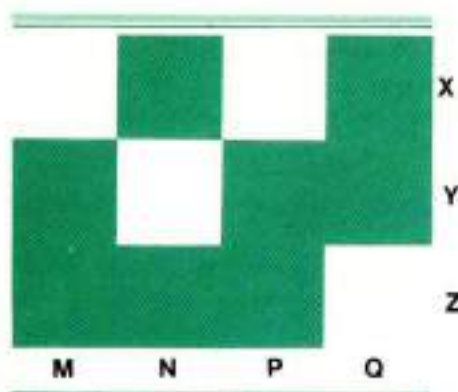
$$m_1 \oplus m_2 = m_2 \oplus m_1$$

$$m_1 \oplus (m_2 \oplus m_3) = (m_1 \oplus m_2) \oplus m_3$$

La regla de Dempster constituye una generalización de reglas de cálculo de probabilidades. En efecto, si las distribuciones de dos cuerpos de evidencia son distribuciones de probabilidad  $P_1(u)$  y  $P_2(u)$ , como en  $P$ , se atribuyen valores a cada suceso básico necesariamente existe intersección (precisamente para cada descripción elemental) de los elementos a componer por la regla de Dempster, resultando:

$$P(u_i) = \frac{P_1(u_i) \cdot P_2(u_i)}{\sum [P_1(u_k) \cdot P_2(u_k)]}$$

La regla de combinación de Dempster puede representarse mediante un diagrama en dos dimensiones, donde en cada uno de los ejes se describe cada cuerpo de evidencia elemental a combinar. Así en el ejemplo, una evidencia se distribuye sobre las afirmaciones X, Y, Z y en la otra sobre M, N, P, Q. El cuerpo de evidencia resultante se distribuye sobre  $4 \times 3 = 12$  afirmaciones elementales como máximo, cada una constituida por las posibles formas de clasificar cada afirmación en un cuerpo de evidencia según las otras condiciones, siempre que la conjunción no sea vacía; por ello podría ocurrir que de los 12 elementos hubiera sólo 8 en condiciones sobre los que se produciría la distribución del cuerpo de evidencia compuesto, teniendo en cuenta el procedimiento de normalización antes descrito. La distribución resultante serviría de base para, representada sobre un eje, combinarse con otro cuerpo de evidencia basado en los conceptos W, T, S, por ejemplo, en un nuevo diagrama. Esta forma de acumulación de evidencia es la base del procedimiento de inferencia basado en esta teoría. (Ver figura adjunta).



#### 4.4.4.2. La lógica difusa. El método PRUF

En 1965 Zadeh introdujo el concepto de conjunto difuso, definiendo una función característica  $f(x)$  asociada para evaluar el grado de pertenencia de un elemento  $x$  a un conjunto  $A$ , de manera que  $f(x)$  le asigna un valor en el intervalo cerrado  $[0,1]$  que mide su grado de pertenencia al conjunto  $A$ . Esto constituye una generalización del concepto clásico, en que la función característica sólo podría tomar valores 0 y 1.

En términos de significado de proposiciones, esto equivale a establecer una gradación continua del valor de verdad desde 0 (totalmente falso) a 1 (totalmente cierto).

Desde el año 1965, en que Zadeh introdujo el concepto, más de un millar de artículos han descrito aplicaciones del mismo a distintos ámbitos de la ciencia y la técnica. No se ha conseguido, sin embargo, una teoría unificada del razonamiento difuso. Recientemente Zadeh (1979, 1981) ha ordenado las distintas propuestas en un sistema integrado de cálculo, con significados difusos de los distintos tipos de afirmaciones en lenguaje natural denominado PRUF (Possibilistic Relational Universal Fuzzy), apoyado en las funciones de posibilidad introducidas por Gaines (1976).

A continuación se describen las características generales del método en los tres aspectos:

- Semántica de fórmulas simples.
- Semántica de fórmulas compuestas por conectivas y cuantificadores.
- Forma de representar la inferencia.

#### 4.3.4.2. Semántica difusa de fórmulas

Dada una fórmula  $F$  para definir su significado se le asocia una distribución de posibilidad en un dominio de medidas que evalúa el grado de certidumbre de la misma para distintos valores de la medida.

Por ejemplo, a la afirmación  $F$ : "Juan llega tarde", puede asociarse una distribución de posibilidad en un dominio de medidas de retraso,

$\pi_i(x)$  0/0 0.1/5 0.25/10 0.40/15  
0.50/20 0.60/25 0.70/30 1/35

$x$  se evalúa en minutos y los pares expresan en primer lugar la medida del retraso y en segundo lugar la posibilidad de que se afirme que Juan llega tarde con esa medida de retraso.

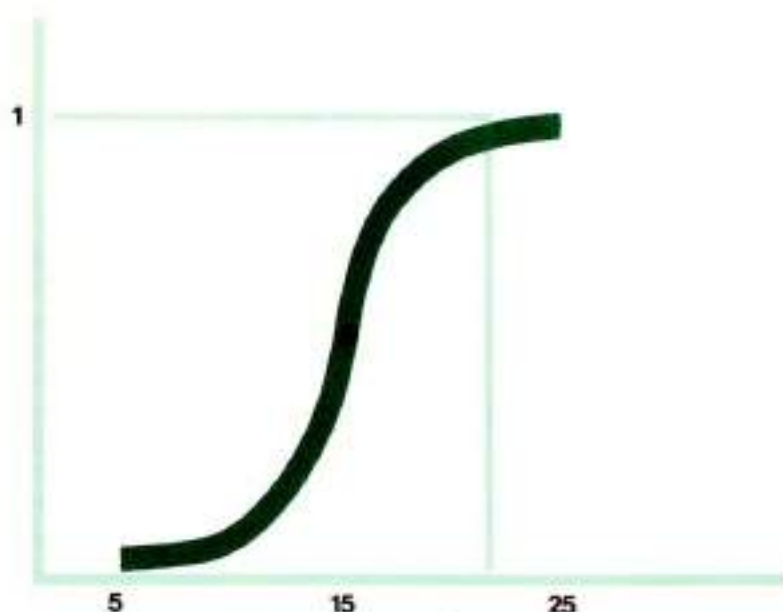
En este caso, la distribución  $\pi_i$  es de tipo discreto, pero cabe también plantear funciones continuas. Las formulaciones más frecuentes para ésta son las curvas en  $S$ :

$S(u, \alpha, \beta, \gamma)$ : distribución en  $S$  de la variable  $u$  con parámetros  $\alpha, \beta, \gamma$ , que se calcula de la forma siguiente:

$$S(u, \alpha, \beta, \gamma) \begin{cases} = 0 & \text{para } u > \alpha \\ = \frac{2(u - \alpha)^2}{(\gamma - \alpha)^2} & \alpha \leq u \leq \beta \\ = 1 - 2 \frac{(u - \gamma)^2}{(\gamma - \alpha)^2} & \text{para } \beta \leq u \leq \gamma \\ = 1 & \text{para } u > \gamma \end{cases}$$

En la figura se incluye una posible curva en  $S$  asociable a la medida de tardanza:

$S(X,5,15,25)$



La fijación del universo de medidas no tiene por qué ser correlativo de número de plazas del predicado que se modela, así, por ejemplo, la afirmación con dos términos:

$F$ : Juan llega antes que José

puede evaluarse en base a una función posibilística de la diferencia entre horas de llegada de Juan y José:

$\pi_i(x)$ : 0/-15 0.4/-10  
0.6/-5 0.8/0 1/5



#### 4.3.4.2.2. COMPOSICION DE SIGNIFICADOS

Una vez definida la forma de representar el significado de fórmulas en términos difusos, es preciso definir la forma de obtener el significado de fórmulas compuestas a partir del significado de sus componentes. De esta manera, será posible evaluar fórmulas cualesquiera compuestas de conectivas y cuantificadores.

Para obtener el significado de una fórmula compuesta por conectivas o cuantificadores, basta definir la función de posibilidad resultante a partir de las funciones de los componentes. En el caso de conectivas cabe distinguir dos casos:

— Las funciones de posibilidad de los componentes se definen en el mismo espacio de medidas. En este caso,

$$\pi_{p \wedge q}(x_1, x_2, \dots, x_n) = \min [\pi_p(x_1, x_2, \dots, x_n),$$

$$\pi_q(x_1, x_2, \dots, x_n)]$$

$$\pi_{p \vee q}(x_1, x_2, \dots, x_n) = \max [\pi_p(x_1, \dots, x_n),$$

$$\pi_q(x_1, \dots, x_n)]$$

$$\pi_{\neg p}(x_1, x_2, \dots, x_n) = 1 - \pi_p(x_1, x_2, \dots, x_n)$$

como es sencillo comprobar, estas definiciones son coherentes con la composición de significados bivalentes en lógica clásica.

— Las funciones de posibilidad de los componentes se definen en espacios distintos de medidas.

Este caso sólo afecta a la composición por conjunción y disyunción. Para tratarlo se introduce el concepto de extensión cilíndrica de una distribución de posibilidad.

Si una distribución  $\pi(x_1, \dots, x_r)$  es independiente de una variable "y" se define su extensión cilíndrica a/y en la forma

$$\pi(x_1, x_2, \dots, x_r, y) = \pi(x_1, x_2, \dots, x_r)$$

es decir, para cada valor de y se asigna la misma posibilidad a cada combinación de valores de las otras variables que hay en la función sin extender a/y.

De acuerdo con esta definición para obtener la distribución de posibilidad correspondiente a la composición de dos distribuciones, en donde no coinciden las variables de medida, se realizan las extensiones cilíndricas de cada una pertinentes, para hacer que tengan la misma base de medidas y se obtienen las distribuciones compuestas aplicando las definiciones del caso anterior.

A continuación se ilustra la definición con un ejemplo:

Si la proposición P tiene la distribución:

$$\pi_p(x): 0/a \ 0.2/b \ 0.3/c \ 0.6/d \ 0.8/e \ 1.0/f$$

(La medida es x cuyos valores posibles son a, b, c, d, e, f) y la proposición q tiene la distribución:

$$\pi_q(f): 0/m \ 0.5/n \ 1/s$$

La extensión cilíndrica de  $\pi_p$  es

$$\pi_p(x, y): 0/am \ 0/an \ 0/as \ 0.2/bm \ 0.2/bn \ 0.2/bs \ 0.3/cm \ 0.3/cn \ 0.3/cs \ 0.6/dm \ 0.6/dn \ 0.6/ds \ 0.8/em \ 0.8/en \ 0.8/es \ 1/fm \ 1/fn \ 1/fs$$

Zadeh, en lo referente a cuantificadores, propone no sólo los convencionales universal y existencial, sino una gama de ellos tomados del lenguaje usual. Así, si una fórmula cuantificada en la variable x se escribe de manera general:  $Q \times H$ , Q puede referirse a conceptos como "todos", "algunos", "varios", "muchos", "la mayor parte", "pocos", etcétera.

La modelación de este concepto se realiza entendiendo que una fórmula cuantificada puede considerarse como una proposición difusa sobre la medida de los elementos que satisfacen el predicado H.

La extensión cilíndrica de  $\pi_q$  es

$$\pi_q \wedge q: 0/am \ 0/bm \ 0/cm \ 0/dm \ 0/em \ 0/fm \ 0/an \ 0.2/bn \ 0.3/cn \ 0.5/dn \ 0.5/en \ 0.5/fn \ 0/as \ 0.2/bs \ 0.3/cs \ 0.6/ds \ 0.8/es \ 1/fs$$

Ambas distribuciones se refieren ahora, al mismo espacio de medidas y, por tanto, puede evaluarse la distribución de posibilidad de  $P \wedge q$  de acuerdo con la definición dada:

$$\pi_{p \wedge q}: 0/am \ 0/bm \ 0/cm \ 0/dm \ 0/em \ 0/fm \ 0/an \ 0.2/bn \ 0.3/cn \ 0.5/dn \ 0.5/en \ 0.5/fn \ 0/as \ 0.2/bs \ 0.3/cs \ 0.6/ds \ 0.8/es \ 1/fs$$

Se ha dejado para el final la forma de evaluar la conectiva de implicación para lo que se han propuesto distintas fórmulas, todas ellas cumpliendo la condición de ser consistentes con el concepto de esta conectiva en semántica bivalente.

Dos de las definiciones más empleadas son:

$$\pi_{p \rightarrow q}(x_1, x_2, \dots, x_n) = 1 - \pi_p(x_1, x_2, \dots, x_n) - \pi_q(x_1, x_2, \dots, x_n)$$

con la condición suplementaria de que si esta expresión rebasa el valor unidad se asigna valor unidad,

$$\pi_{p \rightarrow q}(x_1, x_2, \dots, x_n) = \pi_{(p \rightarrow q)} \vee p'$$

Siendo p' la negación de p.

De acuerdo con ello, si a cada cuantificador se le asigna una distribución posibilista,  $\pi_Q$ , el significado de una fórmula cuantificada  $Q \times H$  vendrá definida por la distribución de posibilidad  $\pi_Q(mH)$ , siendo mH la medida de los elementos que satisfacen H según la distribución posibilística  $\pi_H$ .

Como medida del número de elementos que satisfacen H, Zadeh propone utilizar la suma de los valores de posibilidad de  $\pi_H$  (integral en el caso de distribución continua) que en cierta forma evalúa la posibilidad total distri-

buida. De esta forma, de igual manera que a una proposición no cuantificada se le asignaba un conjunto difuso definido a partir de una medida (en el caso de los ejemplos iniciales, para evaluar si alguien llegaba tarde se utilizaba la medida del retraso en minutos) en el caso de las proposiciones cuantificadas, se toma como medida la posibilidad total, de forma que su valor se supone distribuido según la posibilidad asignada al cuantificador correspondiente.

La posibilidad total (potencia) de una distribución discreta, cuyos valores de medida básica considerados son  $x_1, x_2, \dots, x_n$ , a los que corresponde  $\text{pos}(x_1), \text{pos}(x_2), \dots, \text{pos}(x_n)$  se evalúa como  $\sum \text{pos}(x_i)$  extendida al conjunto de  $x_i$ , tales que  $\text{pos}(x_i) > 0$ .

Si se tiene una proposición cuantificada del tipo:

"Varios coches son grandes"

Su significado se valora en el sentido de:

— Si se tiene una medida de la "grandeza" de los coches, por ejemplo su longitud, la distribución de posibilidades definitoria del concepto "ser grande" es tal que la posibilidad total  $\sum \text{pos}(x_i)$  ( $x_i =$  longitud tipo  $i$ ) se distribuye de acuerdo con la distribución de la medida decuantificador "varios" y, por tanto, dada una distribución de "ser grande" es posible evaluar su nivel de posibilidad concreta calculando para ella  $\sum \text{pos}(x_i)$  y entrando en la distribución de "varios" para establecer la posibilidad de la fórmula cuantificada. Por ejemplo:

Si "x es grande" tiene las posibilidades siguientes (sobre un universo de tamaños de 8 elementos):

3/0.1 3.50/0.2 4/0.40 5/0.70  
6/0.90 7/1.00 8/1.00 9/1.00

y "varios" se distribuye como:

2/0.1 3/0.2 4/0.5 5/0.6 6/0.7  
7/1.0 8/1.0 9/0.8 10/0.70  
11/0.60 12/0.50 13/0.1

la posibilidad total es:

$$0.1 + 0.2 + 0.4 + 0.7 + 0.90 + 1.00 + 1.00 + 1.00 = 5.3$$

que da una posibilidad entre 0.6 y 0.90 en la fórmula cuantificada.

La medida de la extensión de un conjunto puede plantearse en forma relativa de manera que, por ejemplo, "varios" podría entenderse en fracciones del universo total en la distribución siguiente:

0.1/0.1 0.2/0.3 0.3/0.4 0.5/0.6  
0.6/1 0.7/1 0.8/0.6 0.9/0.4 1/0.1

En este caso, la medida de la posibilidad total debe hacerse también en forma relativa al máximo, que en el universo de 8 elementos sería precisamente 8 (cada elemento con posibilidad 1), en cuyo caso la potencia relativa del concepto "ser grande" sería  $5.3/8 = 0.66$ , cuya posibilidad en la distribución de "varios" es 1.

El cuantificador universal tendría una distribución relativa:

0.1/0 0.2/0 0.3/0 0.4/0 0.6/0 0.7/0  
0.8/0 0.9/0 1/1

de forma que la afirmación "todos son grandes", con la distribución dada del concepto "grande" tendría la posibilidad nula.

Las fórmulas cuantificadas se entienden en PRUF, por tanto, como afirmaciones difusas sobre las distribuciones de posibilidad de los predicados cuantificados, representadas éstas por su potencia o posibilidad total.

La potencia de una distribución continua se define de una manera análoga mediante la utilización de integrales



en el intervalo de valores de la variable de medida.

#### 4.3.4.2.3. INFERENCIA

En base a la semántica antes descrita de las afirmaciones del lenguaje usual, se utiliza una generalización del concepto de deducibilidad de la lógica clásica: Dada una estructura deductiva con premisas  $P_1, \dots, P_n$  y conclusión  $Q$  se dice que es correcta si siempre que las premisas son verdad, la conclusión es verdad.

Esta definición se matiza en lógica difusa indicando que una estructura deductiva es correcta cuando para los mismos valores de las variables de medida, la posibilidad de las premisas es menor o igual que la de la conclusión. Es decir:

Si la conjunción de las premisas  $P_i$  tiene distribución  $\pi_p(x_1, \dots, x_n)$  y la conclusión  $\pi_q(x_1, x_2, \dots, x_n)$  (una vez realizadas las correspondientes extensiones cilíndricas). Si  $Q$  es deducible de  $P_1, \dots, P_n$ , debe ocurrir para todo  $x_1, \dots, x_n$ :

$$\pi_p(x_1, \dots, x_n) \leq \pi_q(x_1, \dots, x_n)$$

Por tanto, el conjunto de premisas permite definir con esta condición un subespacio en el espacio de medidas en el que se encuentran las distribuciones de las afirmaciones deducibles de las premisas.

El planteamiento en lógica difusa resulta, por tanto, más consistente desde el punto de vista teórico; sin embargo, dada la representación de los significados mediante distribuciones de posibilidad que pueden tener excesivos parámetros, no es sencilla la traslación a este tipo de representación del criterio de expertos. Podría ser de interés implementar modelos de este tipo, con distribuciones muy restringidas en el número de parámetros.

Para ilustrar la mayor capacidad de matización en la acumulación de evi-

dencia de este tipo de modelización, se trata el problema del diagnóstico contra ejemplo de Zadeh a la regla de composición de Dempster.

Si un médico asegura un 99 % de probabilidad que es meningitis y un 1 % tumor, y otro asegura que 99 % embolia y 1 % tumor, ¿qué puede afirmarse sobre el diagnóstico de tumor?

Si se aceptan las afirmaciones de ambos médicos como premisa y las medidas de probabilidad se toman como medida de posibilidad, si se denominan:

- a: tiene meningitis
- b: tiene embolia
- c: tiene tumor

las distribuciones de posibilidad siendo  $x$  variable en  $\{a, b, c\}$  son:

$$P_1: 0.99/a \ 0/b \ 0.01/c$$
$$P_2: 0/a \ 0.99/b \ 0.01/c$$

La distribución de  $P_1 \wedge P_2$  es:  $0/a \ 0/b \ 0.01/c$  y, por tanto son deducibles afirmaciones cuya distribución sea:  $\alpha/a \ \beta/b \ \gamma/c$ , tales que  $1 \geq \alpha \geq 0$ ,  $1 \geq \beta \geq 0$  y  $0.01 \leq \gamma \leq 1.00$  y, por tanto, la afirmación "tiene tumor" tiene una posibilidad que puede oscilar entre 0.01 y 1 (con la técnica de composición de Dempster salía solamente 1).

## 5. LA CONSTRUCCION DE SISTEMAS EXPERTOS

Como proyecto informático un Sistema Experto debe definirse con los siguientes elementos:

— Un subsistema de adquisición de la información de partida para inferencia, que puede operar tomando información de una base de datos, de cuya inspección deduce el grado de veracidad de los hechos iniciales de la red inferencial; también puede plantearse sistemas de adquisición de información mediante interrogación directa al usuario, y, obviamente, puede plantearse sistemas mixtos de consulta a usuarios y bases de datos.

— Un subsistema de actualización y validación de la base de conocimiento, permitirá la inspección de la base y su modificación por eliminación, sustitución o inclusión de nuevos conceptos. Normalmente, dado que las personas que deben controlar esta base son expertos en el tipo de conocimiento y no necesariamente en informática, conviene que la consulta y modificación pueda hacerse en lenguaje próximo al natural.

— Motor inferencial que utilice alguna de las técnicas antes descritas en el apartado 4, con el oportuno trazador, de forma que sea factible tanto la obtención de respuestas como su explicación. Pueden, así mismo, en casos específicos, integrarse en este módulo elementos de presentación gráfica cuando se considere de interés (por ejemplo, cuando se trata de sistemas expertos de ayuda al diseño).

Cada uno de estos módulos constituye una unidad de proyecto informático que puede tratarse mediante las técnicas convencionales; sin embargo, tanto el subsistema de validación y actualización de la base de conocimiento, como el motor inferencial constituyen dos típicos proyectos de inteligencia artificial, cuya forma de tratarse tiene por ello una tradición en los entornos de lógica para esta técnica.

Por otro lado, la experiencia en la construcción de determinados sistemas basados en el conocimiento ha dado forma, orientada a esta clase de aplicaciones, a determinados instrumentos específicos. No cabe tampoco olvidar que, en la construcción de estos sistemas, el aparato informático es sólo un aspecto, siendo el otro la propia base de conocimiento, en donde reside gran parte del valor añadido de la inversión, por ello de que el diseño haga fácil la verificación y mejora de la base a partir del conocimiento de expertos, depende en forma importante el éxito del proyecto.

A continuación se incluyen descripciones sucintas de los entornos de lo-

gical genérico o específico, que pueden utilizarse para diseño, así como algunas consideraciones metodológicas sobre la construcción de bases de conocimiento. Estas últimas deben entenderse como sugerencias, dada la escasa experiencia en la construcción y uso de estos sistemas.

### 5.1. Entornos genéricos de logical

Los lenguajes generales de soporte de los proyectos de inteligencia artificial en el momento actual son LISP y PROLOG. Mediante ambos lenguajes se han desarrollado también entornos específicos orientados al diseño y operación de sistemas expertos.

#### 5.1.1. El lenguaje LISP

Desde que se inició la actividad en el ámbito de la inteligencia artificial en 1956, los investigadores de esta comunidad se plantearon la necesidad de contar con un lenguaje que se adaptara a su tipo de problema, haciendo abstracción de las características concretas de las máquinas en que se implementara. Ello era debido a la constatación de algunas características diferenciales entre la programación para Inteligencia Artificial y la típica de implementación de algoritmos. En efecto, se dice que existe un algoritmo para resolver un tipo de problemas cuando analizado éste, se ha encontrado un conjunto de conocimiento en base al cual se define una secuencia ordenada de operaciones (algunas de ellas estructuradas en bucles repetitivos que se sabe que terminan necesariamente) que conduce a la obtención de la solución del tipo de problemas.

En cambio, en Inteligencia Artificial se plantean sistemas basados en conocimientos incompletos que por tanto, no alcanzan para definir un algoritmo en sentido estricto, pero en base al cual se pueden investigar posibles se-

cuencias de operaciones que resuelvan el problema. Los elementos de conocimiento para plantear un proceso inteligente son por tanto, las características de las operaciones básicas para resolver el problema y un conjunto de criterios para investigar formas posibles de encadenarlas. Por ello, el resultado en un proceso inteligente es, no sólo la solución, sino el procedimiento para encontrarla, en contraposición con los procesos algorítmicos limitados a la solución, ya que el proceso está totalmente definido. Estas distintas características de los procesos específicos de inteligencia artificial hicieron que, al plantear un entorno de programación adaptado, se llegara a una solución muy peculiar como es el lenguaje LISP, debido a J. Mc Carthy, cuyas primeras versiones datan de 1958 y 1962 y que posteriormente ha sido objeto de desarrollo en distintos dialectos, hasta la época actual.

LISP es un lenguaje al propio tiempo sencillo y capaz de manejar procesos complejos; ello es debido a un planteamiento orientado fundamentalmente a la mejor definición de los conceptos necesarios para el tipo de procesos de inteligencia artificial, aun a costa de una pérdida de eficiencia. Esta idea, una de las bases del movimiento de programación estructurada, estaba ya patente en el concepto de LISP. Aunque no es objeto de estas notas hacer una presentación formal de LISP, algunas de sus características son las siguientes:

— El concepto del lenguaje es un conjunto de funciones operando sobre una estructura única: la lista entendida en su forma más general, es decir, sus componentes pueden ser a su vez listas, lo que permite realmente la formulación de gran variedad de estructuras arboriformes.

— De las funciones, unas son básicas del sistema (manejo de listas, predicados de control) y otras pueden construirse directamente por el usuario,



mediante funciones especiales para definición de funciones.

— La sintaxis define la formulación de un programa como estructura de lista.

— Una de las funciones estándar del lenguaje es el propio intérprete de LISP (función EVAL) de manera que, por programa, se pueden construir listas que, a su vez, pueden ejecutarse como programas. Esta posibilidad hace que debido a la unicidad de estructura de datos y programas, sea sencillo incluir en un mismo programa la investigación sobre secuencias de operaciones, cuyos nombres pueden manejarse como variables, elementos de una lista de funciones y ejecutarse una vez especificadas mediante las oportunas llamadas a EVAL.

— Es posible formular procesos mediante llamadas recursivas de funciones. La recursividad no es en LISP una concesión sino un concepto básico en su diseño.

— Los modernos sistemas LISP son interactivos, de forma que la construcción de programas se realiza definiendo primero las funciones básicas que se evalúan y por tanto se prueban separadamente en forma interactiva. Sucesivamente, basándose en las ya definidas, se van construyendo y probando nuevas funciones, de manera que un entorno LISP es un vehículo idóneo en el que se puede ir construyendo y verificando gradualmente un sistema.

— La unicidad de la lista como estructura de información, así como su dinamicidad, libera al diseñador de todos los problemas de manejo de memoria, tradicionales en los lenguajes clásicos.

Junto a estas ventajas, aparece el inconveniente de la ilegibilidad de los programas, debido al exceso de paréntesis y la posible pérdida de eficacia debido a las discordancias entre la máquina que supone LISP y la real, lo que en épocas de máquinas menos eficientes, hizo afirmar que LISP era

una de las formas más elegantes de desaprovechar un computador.

La generalidad de concepto de LISP ha hecho que sea factible programar en este lenguaje, tanto con enfoque funcional como con enfoque clásico imperativo, ya que es sencillo definir como funciones de tipo "macro" las formas clásicas de la programación estructurada.

En el momento actual existen dos dialectos LISP ya clásicos:

— MACLISP, procedente en sus primeras versiones del sistema MACSYMA utilizado en la Costa Este de EE.UU. (universidades de CARNEGIE MELLON y MIT) que hace énfasis en los aspectos de cálculo matemático.

— INTERLISP, desarrollado en la Costa Oeste de EE.UU., fue la primera versión interactiva de LISP; es el resultado de una evolución del LISP 1.5 de MC CARTHY debida a BBN (Bolt Beranek Newman) que implementó una versión en PDP 1 en 1966 y en PDP-10 en 1970. En 1973 se produjo la primera versión de INTERLISP, mediante una colaboración BBN y XEROX PARC. INTERLISP hace énfasis en aspectos de facilidad operativa.

El desarrollo de LISP alcanzó un punto significativo en 1979 cuando en el MIT se construyó un chip en el que estaban instrumentadas, via hardware, las funciones básicas de LISP; con ello se había creado la base para las máquinas LISP que ya aparecían comercialmente en la exposición de la conferencia mundial de Inteligencia Artificial 1981 (IJCAI 1981) en Vancouver. Actualmente hay dos marcas comerciales que construyen máquinas de este tipo; resulta curioso que en una de ellas se soportan otros lenguajes (PASCAL y FORTRAN) y en ellas sus compiladores están escritos en LISP; con ello se cierra una etapa en la evolución de la actitud ante los lenguajes de programación:

— desde los primeros lenguajes basados en la arquitectura material del procesador, a los lenguajes basados en planteamientos conceptuales orientados a soportar un tipo de procesos.

Las máquinas LISP se basan en una versión extendida de MACLISP denominada ZETALISP. Finalmente, el nivel de difusión de este lenguaje ha conseguido cotas suficientes como para crear la necesidad de su unificación y estandarización; desde 1980, los trabajos en esa línea han conducido a la definición de una versión estándar denominada COMMON LISP, de la que se están realizando compiladores e intérpretes incluso para microprocesadores.

### 5.1.2. EL LENGUAJE PROLOG

Este lenguaje, como ya se ha indicado en 4.1 procede de una etapa de mayor desarrollo de las técnicas basadas de Inteligencia Artificial; si LISP se desarrolló a la medida de los problemas de los primeros tiempos en que los programas inteligentes se planteaban como una búsqueda general, PROLOG, es una consecuencia del desarrollo de las técnicas de deducción automática y por tanto incorporan mayor cantidad de teoría que LISP, aunque éste sea más general.

Como ya se ha indicado en 4.1, un programa PROLOG consiste en:

— una especificación de hechos iniciales, formulado por predicados asumidos como ciertos.

— una especificación de hechos deducibles a partir de otros, formulados mediante reglas de definiciones del tipo:

A ← B, C, D,

que debe leerse lógicamente: A si B C y D y procedualmente: para conseguir A es necesario conseguir primeramente que se cumplan B, C y D. Con este "programa" pueden resolverse



problemas especificados como conjunción de objetivos:

### R S T

Tanto A, B, C, D como R, S, T son predicados que en las formulaciones aparecen referidos a constantes o variables.

Resolver un problema en PROLOG es:

- bien comprobar que se cumple o no el conjunto de objetivos, si en estos no aparecen variables libres.
- o bien obtener uno o todos los valores de las variables libres en la conjunción de objetivos planteada, que hacen que estos se satisfagan.

Como ya se ha indicado, la definición de estos valores solución se hace empleando una técnica de búsqueda en el conjunto de afirmaciones-programa. El procedimiento general consiste en buscar en los antecedentes de las cláusulas programa, la primera, en el orden de escritura que es unificable con el primer objetivo (es decir, aquella para la que existe una especificación de las variables libres, que hace que se obtenga la misma instancia para ambas) y se sustituye el objetivo por los predicados consecuentes de la cláusula instanciados para los valores de las variables que unifican. Este proceso se repite bien hasta que la lista de objetivos queda vacía, lo que indicaría que los valores de las variables en ese momento son una solución o bien hasta que, sin estar vacía, sus predicados no son unificables con ninguno de los antecedentes de las afirmaciones-programa en cuyo caso se trata de un camino fallido y el intérprete PROLOG hace vuelta atrás para seleccionar, en alguna de las unificaciones previas en el proceso, la cláusula siguiente a la seleccionada en la rama fallida. Si no existe ninguna el proceso termina sin encontrar más soluciones.

La máquina PROLOG interpreta una formulación en este lenguaje con un proceso de búsqueda sistemática. El control, en este tipo de programas,

consiste en la codificación de condiciones restrictivas de este proceso de búsqueda. La programación lógica, un área de investigación en la que se producirán desarrollos, ha creado distintas técnicas de este tipo, entre ellas, una de las más generalizadas es el símbolo / (indicativo de corte) que permite limitar el número de unificaciones a considerar para un objetivo.

Los sistemas PROLOG incorporan la definición axiomática de una serie de funciones y operadores lógicos y aritméticos así como reglas que producen como efecto lateral la escritura de datos o resultados; asimismo, incluyen facilidades para la manipulación de Bases de Conocimiento (denominadas "mundos") de forma que pueden formularse programas PROLOG utilizando determinados "mundos" ya catalogados a los que se añaden nuevas cláusulas.

Para la aplicación directa a la construcción de Sistemas Expertos, PROLOG ofrece una estructura muy idónea ya que los hechos iniciales y las reglas son asimilables a las bases de datos y base de conocimiento, y la máquina PROLOG de búsqueda constituye un motor inferencial. El único reparo que puede oponerse es que su forma de buscar y razonar se basa en la lógica de primer orden, siendo menos directa la formulación de los motores inferenciales que se basen en la lógica probabilística o difusa.

El concepto de PROLOG ha sido tomado como base por los diseñadores de la quinta generación de ordenadores por la simplicidad de su formulación y de su intérprete que puede ser optimizado; de hecho ya existen versiones experimentales mediante procesadores de búsqueda en paralelo. Fundamentalmente la importancia de este lenguaje reside en el cambio de estilo que produce en la formulación de procesos ya que en PROLOG basta especificar "qué hacer" ocupándose la máquina de desarrollar el "cómo hacer" aspecto en cambio que debe ser formulado en los lenguajes procedura-

les actuales incluso en los más avanzados. Esta necesidad de poner de acuerdo los objetivos de un proceso con la manera de formularlo, llevó a primeros de los setenta al desarrollo de los métodos de verificación de programas, que eran formas de razonar con elementos de formulación procedural para demostrar que un programa satisface los objetivos de su diseño. En un programa lógico ideal este proceso no sería necesario aunque todavía lo sea en PROLOG debido a las, hasta ahora, inevitables componentes de control de este lenguaje.

De todas formas la verificación de estas formulaciones puede ser mucho más sencilla, por su proximidad con el lenguaje lógico.

Hay que señalar que, aunque PROLOG sea un lenguaje válido para la formulación de procesos inteligentes y por ello de Sistemas Expertos, su concepto es aplicable a cualquier clase de problemas. En efecto, si la Base de Conocimiento formulada es completa, el intérprete realizará un proceso similar al del algoritmo y si no lo es, realizará una búsqueda con las vueltas atrás que la imperfección de la Base de Conocimiento y las hipótesis del control de búsqueda originen.

### 5.2. Entornos específicos

Como ya se ha indicado, los entornos específicos existentes proceden de sistemas ya instalados. De ellos, el que ha tenido más influencia ha sido MYCIN, del que se derivó como entorno abstracto EMYCIN (Essential MYCIN) en base al cual se construyeron, entre otros, los Sistemas Expertos ya mencionados PUFF SACON; asimismo de EMYCIN precede una línea de sistemas comercializados (KS300, S1, M1).

Basado en PROSPECTOR se desarrolló KAS, que no ha tenido el mismo grado de aceptación que EMYCIN.

Como sistema general de construcción de Bases de Conocimiento se de-



sarrolló sin abstraer de ningún sistema, OPS (Operating Production Systems); en base al mismo se formuló posteriormente el sistema R1 de configuración automática de equipos VAX.

A continuación se describen las características generales de ambos entornos, a título ilustrativo sin ánimo de exhaustividad: OPS 5 como sistema general y EMYCIN como ejemplo de sistema abstraído de un sistema experto preexistente.

### 5.2.1. EL SISTEMA OPS 5

Este sistema soporta un concepto general de sistema de producciones, sus elementos son:

— Una memoria de trabajo cuyos elementos son:

— Listas de números y átomos constituidos éstos por cadenas de símbolos encabezados por una letra.

— Listas de atributos-valores; por ejemplo, podría definirse un elemento persona cuyos atributos son nombre, profesión, edad, número de hijos, etc... que se formulan mediante pares denominación de atributo-valor (por ejemplo: edad 47); para diferenciar el atributo del valor se utiliza un signo especial precediendo el nombre del atributo.

— una Base de Conocimiento constituida por producciones de las que se codifican:

- Los elementos antecedentes (LHS)
- Los elementos consecuentes (RHS)

Los elementos LHS constituyen formas abstractas (patterns) de posibles elementos de la memoria de trabajo o condiciones definidas por conjunciones o disyunciones de siete predicados básicos de relación aplicados a elementos de la memoria de trabajo.

Los elementos RHS son secuencias de llamadas a un conjunto de funciones básicas de modificación de los elementos de la memoria de trabajo:

(MAKE, para crear nuevos elementos según diferentes funciones (GENATOM, SUBSTR, ...), MODIFY para cambiar valores, REMOVE para suprimir, etc...)

— una estrategia de control que permita decidir la regla aplicable mediante un proceso consistente en:

**Repetir** una vez leído un estado inicial de la memoria de trabajo, las dos etapas siguientes:

— Selección de reglas aplicables (pattern matching): se recorre el conjunto de reglas comprobando si el estado corriente de la memoria de trabajo verifica el LHS de cada regla. Si el conjunto de reglas obtenido es vacío, el problema no es resoluble en la rama de búsqueda que se analiza y se realiza vuelta atrás (backtracking). Si el conjunto de reglas aplicables obtenido tiene más de un elemento, es preciso ordenarlas según su idoneidad para lo que se requiere un criterio. OPS5 tiene una serie de criterios implícitos para realizar esta ordenación (según la "antigüedad" de los elementos de la memoria de trabajo que se ajustan al LHS de la regla, el número de estos elementos etc...).

— Aplicación de la regla seleccionada mediante sus funciones operando sobre los elementos de la memoria de trabajo.

**Hasta** que en la memoria de trabajo se verifican los objetivos a obtener en el proceso.

El sistema OPS5 además de proveer esta mecánica fundamental de operación dispone de facilidades.

— para manejar (consultar, modificar) la memoria de trabajo.

— para manejar y validar las reglas de producción y la estrategia de control.

— para hacer trazas de la ejecución del proceso y listar valores intermedios de elementos específicos de la memoria de trabajo, en la fase de puesta a punto.

OPS5 es un sistema debido a Forgy (1980-81) aunque ya desde los trabajos de Newell y Simon que dieron lugar a PSG a primeros de los 70, se desarrollaron versiones preliminares de OPS.

### 5.2.2. EL SISTEMA EMYCIN

Este sistema es más específico que OPS5 ya que, si bien ambos manejan un sistema de reglas, en OPS5 son genéricas (pueden utilizarse tanto para deducir como para transformar, mediante operadores, la memoria de trabajo) y en EMYCIN están orientados a la acumulación de evidencia de acuerdo con el modelo de razonamiento impreciso descrito en el apartado 4.3.1.

El conocimiento se representa por hechos y reglas. Los hechos se representan por ternas contexto-parámetro-valor:

Un contexto es una entidad conceptual sobre la que se basa el análisis, por ejemplo, un paciente en temas de medicina, un coche en temas de mecánica, etc... En cada contexto se agrupan un conjunto de parámetros descriptivos de atributos de la entidad representada, para cada uno de los cuales hay un valor (dentro de un dominio de valores posibles), o un conjunto de valores con distintos grados de credibilidad.

Las reglas se describen por tres elementos cuya estructura general es:

- premisa: conjunción de disyunciones de afirmaciones sobre valores de parámetros.
- conclusión: que modifica la base de hechos, bien por introducción de nuevos elementos o por variación de sus factores de certidumbre.
- factor de certidumbre: medida del nivel de implicación de la regla.

Sobre esta base el motor de inferencia opera mediante una búsqueda, de atrás adelante, partiendo de la pregunta cuyo grado de evidencia se quiere establecer.

— Si la pregunta está compuesta de distintas afirmaciones elementales, se obtienen por análisis de su estructura las afirmaciones elementales componentes a investigar.

— Para cada afirmación elemental se investiga si está presente en las conclusiones de las reglas según un determinado criterio (se seleccionan reglas "relevantes").

— En las reglas en que aparece se investiga su antecedente:

— Si se conoce su grado de certidumbre a partir de la información de hechos básicos, se calcula la variación del factor de certidumbre de la respuesta mediante las fórmulas de razonamiento descritas en el apartado 4.3.1

— Si no se conoce el factor de certidumbre del antecedente, se aplica el mismo proceso recursivamente para éste.

— Si se llega a obtener factor de certidumbre unidad positivo o negativo el proceso termina.

— Si para un determinado objetivo no hay reglas que lo expliquen, el sistema consulta al usuario; si éste no responde, el proceso termina por falta de información.

— Una vez obtenido el factor de certidumbre de cada afirmación elemental obtiene el de la pregunta compuesta mediante aplicación de las reglas de composición del apartado 4.3.1.3.

Sobre esta base operativa EMYCIN ofrece como ayudas principales:

— un módulo que permite explicar las respuestas a demanda del usuario

— un subsistema de consulta y modificación de la Base de Conocimiento, que controla su validez sintáctica y la coherencia entre las distintas reglas.

— elementos de traza de procesos, para puesta a punto.

Asociado a este tipo de formulación existe además el sistema TEIRESIAS

(Davis 1980) que ayuda a la construcción de Bases de Conocimiento. TEIRESIAS a su vez tiene estructura basada en el conocimiento.

### 5.3. Metodología general

El desarrollo en el área de los Sistemas Expertos, aunque ha permitido decantar el concepto y su instrumentación técnica, no ha permitido concretar todavía un procedimiento suficientemente efectivo de elaboración de estos sistemas. No obstante en 1983 Hayes-Roth, Waterman y Lenat publicaron los resultados de un *work shop* reciente con aportaciones de distintos equipos que han construido este tipo de sistemas, en donde se plantean una serie de etapas en el proceso de elaboración de un sistema de este tipo que pueden resumirse en la forma siguiente:

— Establecimiento de una relación de equipo entre los expertos humanos y el ingeniero del conocimiento que diseña la base, de forma que se produce información mutua sobre las posibilidades de la técnica y sobre las características del tema cuyo conocimiento se modeliza. Como consecuencia, se fijan claramente los objetivos del sistema (qué respuestas se desea obtener, sobre qué problemas y con qué nivel).

— Construcción de un primer prototipo a partir de la información anterior.

— Valoración del prototipo a partir de una serie de casos propuestos por los expertos.

— Perfilado inicial de una primera Base, como consecuencia de las críticas al comportamiento ante los distintos casos.

— Prueba del sistema ante usuarios.

— Modificación de la Base de Conocimientos y la interfaz con el usuario en base a las críticas de estos.

— Instalación del sistema:



- Documentación.
- Educación.
- Implementación optimizada.

En un sistema de tipo medio, este proceso, si se cuenta con instrumental adecuado puede tardar un año; no obstante no hay que olvidar que un Sistema Experto no alcanza su máxima eficiencia hasta que no está operativo, y modificado en forma acorde durante dos años. Esta, al menos, es la experiencia del desarrollo de R1, sistema de configuración automática de equipos VAX, cuya versión inicial como ya se ha indicado, se desarrolló en tres meses y tenía 300 reglas, y cuya versión final contiene más de 2.500 reglas en su Base de Conocimiento, al cabo de dos años, por ello, es importante hacer énfasis en el potencial de mejora de este tipo de sistemas, debido a la incrementalidad estructural de su Base de Conocimiento.

## 6. REFERENCIAS

El tema de Sistemas Expertos, por su novedad, es todavía objeto de tratamiento fragmentado, aunque aparecen descripciones detalladas en los textos de autores diversos:

- D. Michie (ed.) "Expert systems in the microelectronic age" Edinburgh University Press, 1979.
- A. Barr E. Feigenbaum "The Handbook of artificial intelligence". Vol. 2. W. Kauffman, 1982.
- F. Hayes Roth D. A. Waterman y D. B. Lenat (Eds.) "Building Expert Systems". Addison Wesley, 1983.

Descripciones de sistemas se han publicado en textos sobre DENDRAL, MYCIN y TEIRESIAS.

- P. Davis D. B. Lenat. "Knowledge based systems" MC Graw Hill, 1983.
- E. Shortliffe "Computer based medical consultations: Mycin" Elsevier, 1976.
- R. K. Lindsay, B. G. Buchanan, E. A. Feigenbaum, J. Lederberg. "Applications of artificial intelligence for organic chemistry: The Dendral project" MC Graw Hill 1980.

Sobre LISP se ha producido últimamente una nueva generación de textos acordes con el desarrollo actual de este lenguaje.

- B. Horn, P. H. Winston "LISP" Addison Wesley 1981.
- G. Steele "Common LISP" Digital 1984.
- D. S. Touretzky. "LISP a gentle introduction to symbolic computation". Harper Row 1984.

Sobre PROLOG están los textos:

- W. Clocksin y W. Mellish "Programming in PROLOG", Springer Verlag, 1981.
- K. L. Clark, F. G. McCabe "Micro PROLOG: Programming in logic". Prentice Hall 1984.

Sobre teoría de la evidencia puede consultarse:

- G. Shafer. "Mathematical Theory of evidence". Princeton, 1976.

El método PRUF aparece en:

- L. Zadeh "PRUF a meaning representation language for natural languages", en "Fuzzy reasoning and its applications" (B. R. Gaines, Ed.). Academic Press 1984.

Finalmente pueden consultarse dos textos generales ya clásicos aunque modernos sobre Inteligencia Artificial:

- N. J. Nilsson. "Principles of Artificial Intelligence". Springer Verlag, 1981.
- P. H. Winston. "Artificial Intelligence". Addison Wesley 1981, 1984.

# LOS SISTEMAS EXPERTOS EN INGENIERIA

J. Cuena BARTOLOME

## 1. INTRODUCCION

En el campo de la ingeniería es donde habitualmente se encuentra mayor número de personas expertas y, por ello, aparentemente, es en donde deberían haberse producido mayor número de aplicaciones. Sin embargo, la realidad muestra que no existen Sistemas Expertos importantes orientados a temas ingenieriles; de hecho, dos de los más conocidos (SACON e HYDRO) tienen un papel auxiliar en procesos de ingeniería y no se refieren a temas de excesiva complejidad. El primero, es un sistema que propone los procesos de comprobación a realizar con una estructura cuyas formas se definen y el segundo, propone valores verosímiles de los parámetros de los modelos de simulación del comportamiento de cuencas hidrográficas.

La falta de grandes aplicaciones de este tipo de sistemas puede ser debida a dos razones: 1.—los procesos involucrados en ingeniería son excesivamente complejos para ser reducidos a las estructuras de representación del conocimiento habitualmente manejadas por los sistemas actuales y 2.—los procesos ingenieriles han sido instrumentados informáticamente de manera importante siguiendo líneas de enfoque distintas de los Sistemas Expertos y no se siente la necesidad de éstos.

Las consideraciones que siguen constituyen una primera reflexión sobre las posibles orientaciones a seguir para aplicar los Sistemas Expertos a las dos líneas principales de actividad en ingeniería: el diseño y la explotación de instalaciones. En ambos se analiza, en forma general, el tipo de tratamiento actual y el posible papel de los sistemas basados en el conocimiento.

## 2. APLICACIONES AL DISEÑO

### 2.1. Concepto

En general, diseñar consiste en decidir los elementos que permiten construir una instalación, infraestructura o equipamiento, de forma que se cumplan determinadas condiciones de limitaciones de recursos y se optimicen objetivos económicos y/o sociales.

### 2.2. El enfoque clásico del diseño

El soporte del diseño son las teorías físicas descriptivas del comportamiento de los elementos que componen el objeto a diseñar, de acuerdo con ello, el enfoque se compone de las tareas:

— Formulación matemática descriptiva de la relación entre magnitud de evaluación del comportamiento del objeto y magnitudes a decidir en el proceso de diseño.

— Utilización de la formulación anterior para definir las magnitudes de diseño.

• Si las funciones formuladas admiten versiones inversas, pueden plantearse las condiciones de evaluación en función de los aspectos a diseñar y eventualmente, técnicas de obtención directa de las magnitudes de diseño que optimizan los objetivos a evaluar.

• Si, como en el caso general, la circunstancia anterior no se presenta, el procedimiento aplicado tradicionalmente es un ciclo repetido de:

- fijación de unas dimensiones.
- obtención de las magnitudes descriptivas del comportamiento.
- evaluación de los aspectos calculados.
- si son aceptables, fin, si no reconsideración de las dimensiones y vuelta a iniciar el proceso.

Las aplicaciones de informática al diseño en ingeniería civil se han orientado a la mecanización de este proceso.



Primero, programando técnicas numéricas ya existentes y en una segunda etapa, construyendo modelos de simulación del comportamiento físico que no era posible hacer operativos sin contar con el computador. De esta forma, en la década de los sesenta, la mayor parte de las aplicaciones se orientaban a la integración numérica de ecuaciones en derivadas parciales y a la resolución de sistemas de ecuaciones. Posteriormente, se produjeron métodos como el de los elementos finitos que constituyeron una aportación original hecha posible por la existencia de los computadores. En otros ámbitos como la hidrología, se produjeron técnicas de modelización analógica que representaban procesos complejos de difícil descripción precisa, como es la acción de la lluvia sobre una cuenca hidrográfica, tanto superficial como subterránea, que se asimila a la circulación según varias secuencias de depósitos debidamente parametrizada.

En el momento actual, tras veinte años de trabajo intenso, puede decirse que existe una metodología informática de diseño sedimentada de acuerdo con la segunda de las opciones de uso de las formulaciones matemáticas. Notablemente este método se ha visto reforzado por las técnicas interactivas que la introducción de datos y recepción de los resultados, vía pantalla, que además en distintas áreas de aplicación puede hacerse en forma gráfica con procesos muy avanzados de representación en dos y tres dimensiones (técnicas CAD).

El entorno informático actual de diseño puede describirse como un conjunto de herramientas de realización de procesos técnicos y de expresión gráfica y escrita, al servicio de la inteligencia y los criterios de la persona responsable. Sin embargo, estos entornos a pesar de su brillantez tienen inconvenientes:

- son complejos de manejo, exigen del operador un esfuerzo importante

de representación del problema en los conceptos del sistema.

- al ser sistemas de tanteo y verificación, el perfilado del proyecto definitivo si no se cuenta con experiencia por el diseñador en el predimensionado o éste no tiene bien asumido los conceptos con que opera el sistema, puede ser muy prolijo.

### **2.3. Los sistemas basados en el conocimiento para apoyo al diseño**

Como ya se ha indicado, existen pocas aplicaciones de Sistemas Expertos para ayuda al diseño; debido a que, por un lado, los entornos existentes por su alta tecnificación pueden dar la impresión de ser un campo ya muy instrumentado y, por otro lado, dada su complejidad y la condición de experimentales de los primeros sistemas, resultaban más adecuados campos con mayor contenido empírico.

Los sistemas ya existentes se ocupan de problemas de ayuda al planteamiento previo para los sistemas de evaluación y simulación.

En efecto, el sistema SACON proporciona consejos sobre el tipo de análisis estructural a realizar en una estructura, de la que se conocen las subestructuras componentes y en cada una de ellas, nuevas subestructuras hasta el nivel en el que se definen cargas sobre las estructuras elementales finales. El sistema incorpora una serie de reglas que permiten razonar en base a la organización y materiales de la estructura y el tipo de cargas y su colocación, sobre el tipo de análisis a realizar (líneal, no líneal, pandeo, etc.). Las conclusiones del sistema sirven de base para que el usuario de un sistema general de análisis estructural conozca la forma de utilizarlo en forma adecuada a su problema. Se trata, por tanto, de un tipo de sistema que ayuda a resolver uno de los problemas de los actuales sistemas de ayuda al diseño

que ofrecen un cúmulo de opciones sólo utilizables en forma eficaz por personas muy expertas.

El sistema HYDRO constituye en cambio, un sistema de ayuda al dimensionamiento previo; en efecto, los modelos de simulación del comportamiento de cuencas hidrográficas ante la acción de la lluvia, incorporan mecanismos analógicos formados por varios depósitos con drenaje lateral intercomunicados en vertical, mediante relaciones representativas de los fenómenos de infiltración y drenaje instantáneo y diferido. El establecimiento de las capacidades de dichos depósitos y los coeficientes de desagüe e infiltración a partir de la morfología y geología de la cuenca, así como el comportamiento observado ante las acciones de lluvia y temperatura, es un problema arduo; de manera que, cuando no se tiene experiencia, es necesario gran número de tanteos hasta encontrar el conjunto de valores de parámetros que verosimilmente representa el comportamiento observado.

El sistema HYDRO maneja, con un mecanismo inferencial tipo PROSPECTOR, un conjunto de reglas que producen como resultado de sus procesos de razonamiento, recomendaciones sobre los valores a fijar de los parámetros de un modelo del tipo antes descrito en base a una descripción de los aspectos más inmediatamente constatables sobre la estructura de la cuenca y el tipo de acciones climáticas sobre ella.

Ambos sistemas, así como los problemas generales observados en el apartado anterior, muestran las dos líneas de aplicación de los sistemas expertos al diseño de ingeniería:

— Sistemas de interfaz, tipo SACON, usuario-sistema de ayuda al diseño en todos aquellos casos en que la complejidad de éste es de manejo difícil:

- tanto a nivel de opciones de proceso posibles

- como a nivel de formulación de los problemas en el entorno (lenguaje) de datos del sistema.

— Sistemas de predimensionamiento en que deben definirse los valores iniciales del proceso de tanteo de forma que éstos estén suficientemente centrados como para que se limite el número de tanteos, en línea con el concepto de HYDRO.

La construcción de estos tipos de sistemas ofrece dificultades diversas:

— Los sistemas de interfaz pueden construirse aplicando la metodología típica de ingeniería de conocimiento para trasladar a una representación adecuada, la forma de entender el manejo del sistema de ayuda al diseño y su teoría de base por personas expertas.

— Los sistemas de predimensionamiento requieren representar el conocimiento de la forma de operar el sistema; es decir, las relaciones entre las premisas y los resultados de la operación de éste. Este planteamiento ofrece la dificultad de reducir el complejo proceso interno del sistema a un conjunto de reglas, lo que obviamente es más impreciso por ser esta estructura de menor potencial de representación, en muchos casos, que la estructura funcional modelizada en el sistema de ayuda al diseño.

No obstante, tanto en uno como en otro caso, puede conseguirse modelos suficientemente representativos. Sussman (1981) ha desarrollado un sistema de predimensionamiento de circuitos integrados en esta línea. Existen técnicas, aunque incipientes, aplicables para síntesis de resultados de explotación de modelos como sistemas de reglas.



### 3. APLICACIONES A LA EXPLOTACION

#### 3.1. Concepto

Las aplicaciones a la explotación se plantean entendiendo ésta como el proceso de definición de asignaciones de los recursos existentes a funciones en el tiempo para cumplir unos objetivos económicos y sociales de una instalación.

#### 3.2. El enfoque actual

Para llevar a cabo la explotación de instalaciones, actualmente, se plantea la definición de estrategias con dos niveles temporales:

- Definición de estrategias generales óptimas a escala anual o mensual.
- Toma de decisiones diarias de explotación ante las diversas incidencias.

Las acciones en el primer caso se adoptan cuando esto es posible, como resultado de estudios técnicos, en los que se formulan matemáticamente tanto los objetivos como las restricciones y se obtienen como resultado las variables de decisión. Tradicionalmente, las técnicas de modelación aplicadas son de dos tipos, ambas dentro de la investigación operativa:

— Técnicas de optimización directa, en las que el resultado del modelo es el conjunto de decisiones.

— Técnicas de optimización por simulación, en las que al ser poco representativa la modelación para optimización directa, se realiza una representación más certera del funcionamiento de la instalación mediante un modelo de simulación y se tantea con él una gama de alternativas posibles de decisiones de explotación, que, una vez evaluadas, permiten decidir la más adecuada.

En la primera línea, el instrumental clásico es la programación matemática para la que existen algoritmos de

optimización en un número restringido de casos, tanto de formulación de restricciones como de la función objetivo; para el caso general, los algoritmos aplicables son de enumeración implícita de soluciones factibles mediante búsqueda heurística con restricciones. Es en este ámbito, en donde existen líneas coincidentes entre técnicas de investigación operativa e inteligencia artificial, con la diferencia de que para los objetivos de aquella (búsqueda de técnicas matemáticas que permitan obtener la solución óptima y demostrarlo) la heurística es un recurso, mientras en inteligencia artificial, es prácticamente el método, ya que el conjunto de heurísticos es una base de conocimiento.

El método basado en la programación matemática adolece de cierta rigidez, al formularse con un único objetivo a optimizar cuando en la gestión real se requiere muchas veces tener en cuenta varios de ellos; también se considera de interés en los modelos de programación matemática la inclusión de variables bivalentes, dadas las dificultades de introducción de condiciones lógicas con la formulación clásica en variables reales. Esta progresiva evolución de los modelos clásicos de programación matemática hacia formulaciones de tipo lógico y a la utilización de técnicas heurísticas de resolución, ambos aspectos clásicos del enfoque de Inteligencia Artificial, muestra el mayor realismo de esta última técnica.

El enfoque mediante modelos de simulación permite una representación más afinada del comportamiento del sistema a gobernar y con un adecuado módulo de evaluación, permite tener en cuenta objetivos múltiples en la selección de alternativas, mediante las técnicas multicriterio aplicadas a los resultados de la simulación. Por ello ésta es la forma más generalizada de apoyar la justificación de decisiones.

El soporte de decisiones diarias no se realiza habitualmente mediante la utilización de modelos, ya que, en el ca-

so general, su planteamiento y operación (con una gama importante de alternativas a estudiar) requieren un tiempo del que no dispone el responsable. Por ello, la forma habitual es apoyarse en un sistema de información que aporte un entorno suficientemente descriptivo de la situación actual y su evolución reciente de manera que la persona experta responsable actúe a su juicio. Esto, cuando el sistema es complejo, puede dar lugar a una formación errónea de decisiones, es por ello deseable la aportación de ayudas en este aspecto.

### 3.3. Los sistemas basados en el conocimiento de ayuda a la explotación

Para la planificación de estrategias de explotación, los sistemas basados en el conocimiento son una alternativa a las técnicas de programación matemática con las que, como se ha indicado, coinciden en los casos en que los modelos de optimización no existen; es factible construir bases de conocimiento que incorporen los heurísticos de búsqueda, las condiciones lógicas que se manejan al modelizar por programación bivalente y los diversos criterios de evaluación que, por tanto, constituyen una representación accesible y modificable de los conceptos con que se elaboran las estrategias.

En el caso de obtención de estrategias mediante modelos de simulación, los Sistemas Expertos como en el caso de diseño, pueden jugar dos papeles:

— Sistemas de interfaz que ayuden a la formulación del problema para el modelo de simulación cuando la complejidad de éste lo justifique.

— Sistemas de predimensionamiento que incorporen la experiencia de los utilizadores del modelo y sintetizen una muestra de pasadas de éste, en donde aparecen explícitos datos y resultados de simulación.

La opción de Sistemas Expertos en predimensionamiento puede llegar a ser una alternativa a la utilización de los propios modelos de simulación si, mediante un sistema de aprendizaje, área de investigación intensiva actual en Inteligencia Artificial, se procesan los archivos de resultados de cada campaña de simulación con el modelo y se obtienen nuevos elementos de la base de conocimiento representativos de estas pasadas. Este proceso repetido a lo largo del tiempo, incorporando además los conceptos que los utilizadores personales del modelo juzguen oportunos tanto para corregir las propuestas del sistema de aprendizaje como para complementarlas, puede llegar a producir un sistema del que se obtengan directamente conclusiones suficientes para la planificación de la instalación.

En particular, la utilización de estos Sistemas Expertos, resumen de experiencia de modelos, resultaría muy útil para ayudar a la toma de decisiones en tiempo real a nivel diario ya que en muchos casos los responsables no tienen tiempo de plantear, procesar y sintetizar una campaña de simulación con el modelo, sobre todo si éste es complejo; sin embargo, sí es factible integrando el Sistema Experto con el sistema de información de la instalación, obtener por proceso del mismo, recomendaciones y diagnósticos en forma prácticamente inmediata. Este concepto de aplicación de estos sistemas ya ha producido algunas aplicaciones. Así, por ejemplo, el sistema ESTRAC-II (Araya Fukumori, 1984) resuelve los problemas de replanificación de los trenes en una red japonesa en tiempo real cuando se producen incidencias. En esta misma línea está planteado el desarrollo de un Sistema Experto para alarma en tiempo real para inundaciones (Cuena, 1983).

## 4. CONCLUSIONES

Como resumen de la panorámica de realizaciones y de posibles líneas de desarrollo, cabe concluir que, aunque en el tema de aplicaciones a ingeniería, las realizaciones en Sistemas Expertos son escasas y de un nivel no excesivamente complejo, el elevado valor añadido que pueden tener estas aplicaciones por la agilización del diseño y de la toma de decisiones que implican debería producir, sin duda, realizaciones a plazo corto, sin embargo, para ello es preciso una redistribución del trabajo en las organizaciones actuales, tanto para la realización de sus operaciones con estos nuevos sistemas, como para la transferencia y creación del conocimiento de estos sistemas.

## 5. REFERENCIAS

1. Bennet J. S. et al SACON: "A knowledge based consultant for structural analysis". *Technical Report HPP-78-23*. Computer Science Department, Universidad de Stanford.
2. R. Reboh et al "HYDRO": "A computer consultation System for solving water resources problems". *T. report SRI International*.
3. J. Cuena: "The use of simulation models and human advice to build an expert system for the defense and control of river floods". *International Joint Conference on Artificial Intelligence (IJCAI-83)*. Karlsruhe, 1983.
4. Araya S., Fukumori, K.: "ESTRAC-II an expert system for train traffic control in disturbed situations". *Proc Sexta Conferencia Europea de Inteligencia Artificial (ECAI-84)* Pisa, North Holland, 1984.



# LOS SISTEMAS EXPERTOS EN MEDICINA

José L. MONTEAGUDO

Ingeniero de Telecomunicación  
Jefe de Investigación  
del Centro Ramón y Cajal

## INTRODUCCION

Analizando la relación de Sistemas Expertos desarrollados hasta la fecha, los temas médicos aparecen como el objeto preferido de los investigadores de inteligencia artificial aplicada (1), (33).

Dentro de las aplicaciones médicas, el mayor interés tanto teórico como práctico, se ha dirigido hacia los problemas relativos al diagnóstico, quizá siguiendo la orientación de las primeras aplicaciones (MYCIN, CASNET, INTERNIST, PIP).

Otro tipo de desarrollos se han enfocado a aspectos tales como el análisis de datos, la planificación, la monitorización, la asistencia inteligente o la misma enseñanza.

También se han invertido esfuerzos en herramientas para la construcción de Sistemas Expertos en medicina y para resolver el problema concreto de la adquisición de conocimiento. Aunque no sean estrictamente médicos, están íntimamente ligados al área biomédica desarrollos como el MOLGEN y ESPEX, para genética molecular y otros dedicados a síntesis de productos químicos.

### Razones que justifican la introducción de Sistemas Expertos en Medicina

Consustancial al formidable desarrollo de las ciencias médicas y de los servicios de salud, ha sido la creciente especialización de los médicos y la multiplicación de las técnicas de análisis y diagnóstico. Esto hace que para su actuación requieran cada vez más la asistencia de otros expertos, y muy especialmente cuando la complejidad del problema presenta aspectos que caen fuera de su propia área de especialidad. Esta situación resulta cotidiana en la estructura actual de nuestros hospitales.

Es un hecho, que diferentes expertos de un campo especializado pueden

llegar a decisiones muy distintas de cómo llevar un caso y se está demostrando que los diagnósticos y las decisiones terapéuticas basadas en ellos varían mucho en exactitud y eficacia según la idiosincrasia de cada especialista. La elección entre las posibles pruebas clínicas y los tratamientos tienen también una incidencia económica importante, que en general no se tiene en cuenta por el médico, especialmente en los sistemas de Seguridad Social, aunque el mayor costo viene ligado siempre al fracaso en el diagnóstico y/o terapia.

Los problemas de diagnóstico médico requieren mucha más información que la que un médico puede almacenar y además éste debe tomar su tiempo para sopesar secuencialmente un largo número de hipótesis. En el proceso es muy fácil que se pase por alto un dato o una posibilidad alternativa. En un proceso mental se pueden manejar entre cuatro y ocho piezas de información, mientras que, por ejemplo, en el diagnóstico de tumores óseos se debe trabajar con más de veinte elementos. Parece obvio, que un sistema computerizado puede ayudar en este tipo de procesos librando al médico de una dura tarea memorística y dejándole más tiempo para dedicar al propio paciente. La elaboración de un diagnóstico médico involucra dos aspectos básicos: la recogida de datos y la toma de decisiones. Los profesionales que acceden a la práctica de la Medicina están generalmente bien entrenados para realizar una descripción completa de un paciente, con los síntomas o indicios cuidadosamente detallados. Sin embargo, parece que es mucho más difícil enseñar bien el proceso de asimilar esa información y utilizarla para emitir un diagnóstico ideal. De hecho, el proceso de diagnóstico se enseña y entiende como un arte más que como una técnica. Los médicos aprenden a tomar decisiones de una forma no estructurada que se basa fundamentalmente en la observación, tratando de captar el razona-

miento subyacente que deducen de la actuación de sus "maestros". La experiencia final se gana a través de la práctica experimental y es inevitable que en este camino se puedan producir errores.

Por lo tanto, las motivaciones para intentar formalizar y automatizar los procesos de tomas de decisiones en Medicina son numerosas y entre las razones para introducir los sistemas informáticos en este campo se pueden citar los siguientes:

**1) Para mejorar la fiabilidad y exactitud de los diagnósticos, los pronósticos y las acciones terapéuticas** mediante procedimientos sistemáticos, completos y capaces de integrar datos de distinta naturaleza procedentes de distintas fuentes, evitando polarizaciones subjetivas y haciendo explícitos los criterios en que se basan las decisiones.

**2) Para mejorar la relación costo-eficacia** dentro de un contexto, donde la proliferación de nuevas técnicas de análisis y terapia altamente sofisticadas, están empujando los costes hospitalarios bajo una demanda exhaustiva de "lo mejor y más moderno" ante la falta de criterios para seleccionar lo pertinente a cada caso sin menoscabo de la calidad de la asistencia.

Subsecuentemente y no menos importantes, aunque no tan inmediatos, son los beneficios que se pueden derivar del entendimiento y formalización de la estructura del conocimiento médico que puede ayudar al desarrollo de técnicas para identificar inconsistencias en ese conocimiento, para mejorar la formación de nuevos especialistas y para hacer que los futuros sistemas informáticos de soporte médico sean más efectivos y más identificados con las necesidades de usuario. Conviene no olvidar el mercado de software para el gran público. Tan pronto como las redes de información interactiva lleguen a los hogares, se desarrollará un nuevo mercado para el acceso del consumidor a la informa-

ción sobre temas de Salud. Existe potencialmente un incentivo económico fuerte para ofrecer bases de conocimiento lo más completas posibles explicando indicios, síntomas, tratamientos, medicamentos y los procesos fisiológicos involucrados en términos llanos, comprensibles para el consumidor general. De hecho, las previsiones del mercado para programas de medicina preventiva, para ordenadores personales muestran cifras muy significativas.

## CARACTERISTICAS DEL PROCESO DE DIAGNOSTICO MEDICO

De una forma global se puede decir que los datos médicos se obtienen o son suministrados por el paciente y son recogidos, coleccionados e interpretados por el médico. El proceso es interactivo y complejo, estando sometido a considerables distorsiones.

El proceso se muestra en el diagrama de la figura 1.

En general, los datos clínicos están definidos, son recogidos y se interpretan con un grado de variabilidad e inexactitud muy grandes.

La fuente de ambigüedad se encuentra a lo largo de todo el proceso de consulta médica, desde cómo los pacientes se definen a sí mismos como "enfermos" y se deciden a solicitar la asistencia sanitaria, hasta los propios test de laboratorio, pasando por las preguntas que el médico realiza al paciente y las contestaciones de éste. Incluso las definiciones de las enfermedades en sí mismas son a menudo ambiguas: algunas dolencias no se pueden definir por criterios objetivos y en las que sí se pueden, los criterios se pueden interpretar muy variablemente según el médico.

En la Tabla I, se indican los tipos de datos que se manejan en los actos clínicos de diagnóstico.



**TABLA I**  
**TIPOS DE DATOS MEDICOS**

**A. Datos expresados**

**1. Verbalmente**

a) "Objetivos":

- Historia médica pasada
- Descripción dolencia actual
- Respuestas al examen físico

b) "Subjetivos":

- Sensaciones

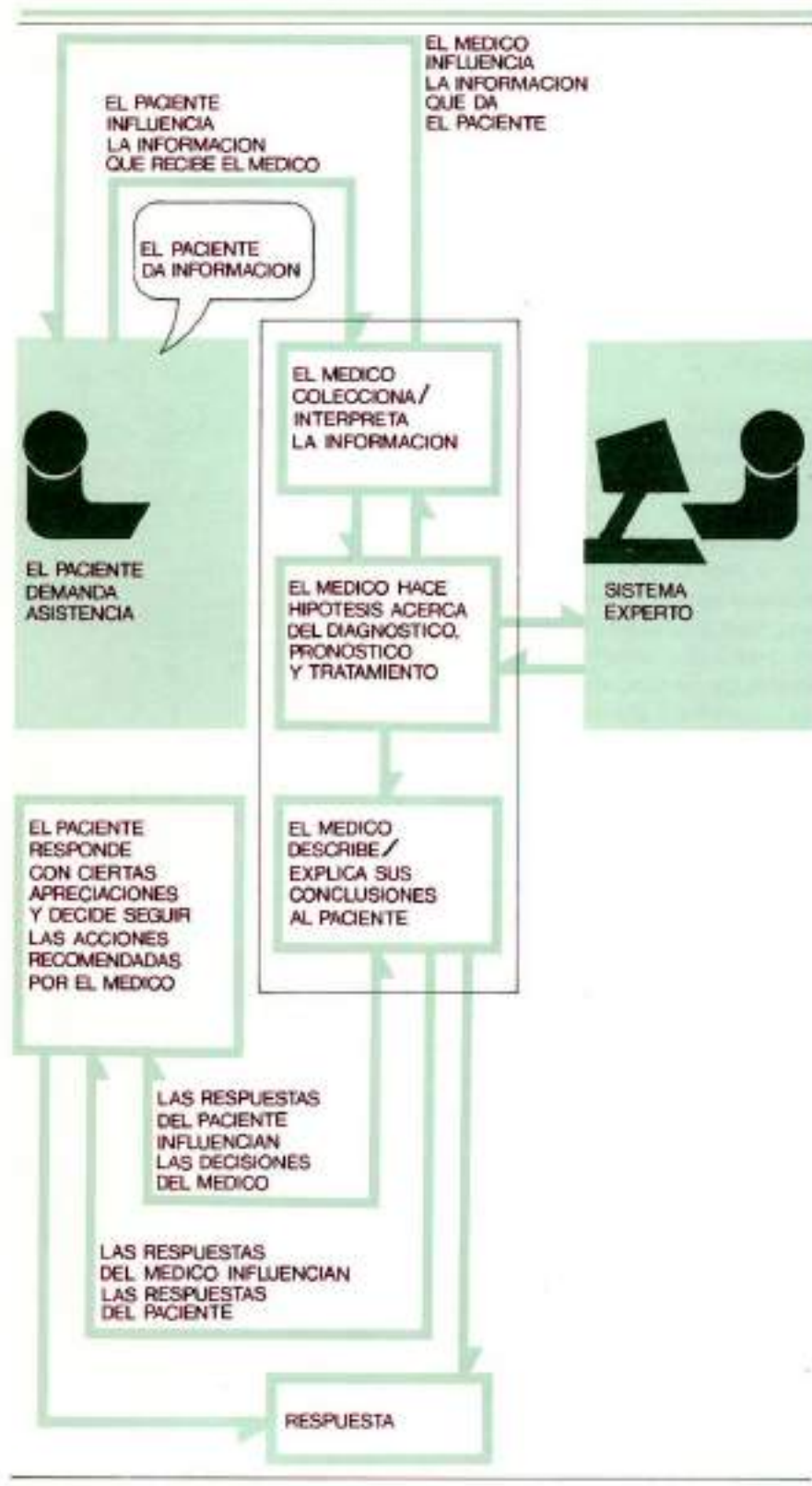
**2. No verbalmente**

**B. Datos obtenidos directamente**

**1. Elementos del examen físico**

**2. Pruebas y análisis:**

- Variables continuas
- Variables discretas
- Imágenes
- Informes clínicos

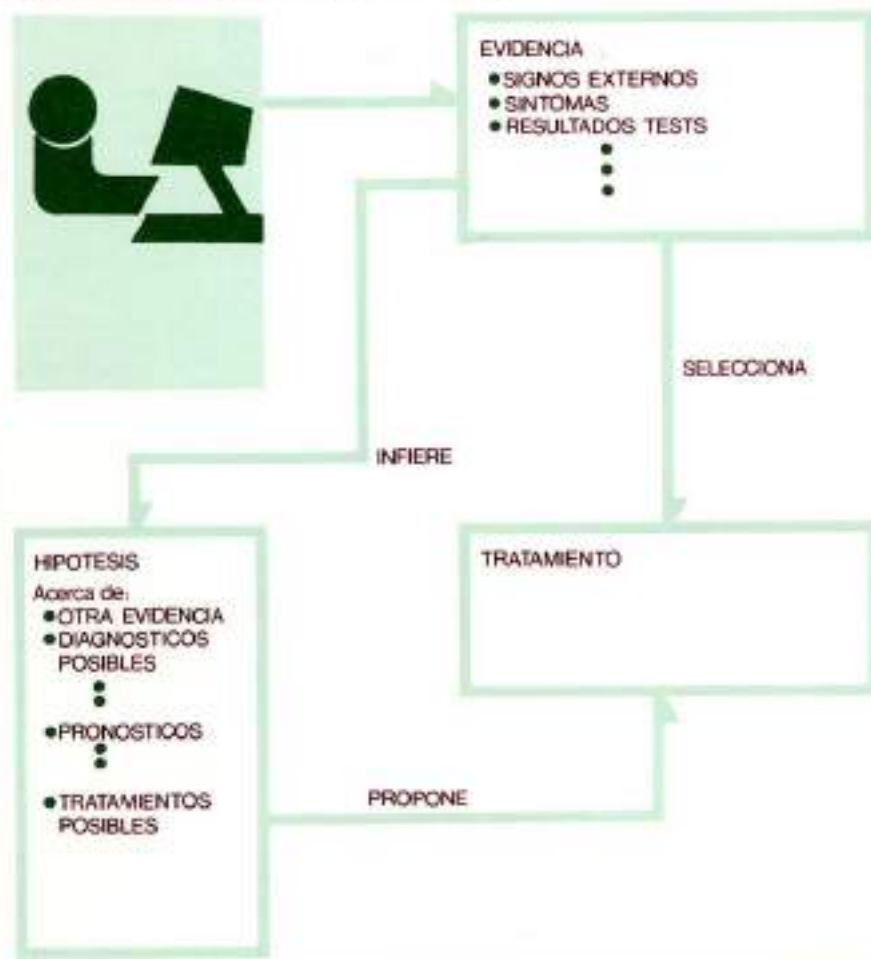


En la figura 2, se muestra un diagrama simplificado de los conceptos y actuaciones médicas típicamente involucradas en la solución de los problemas de consulta por un experto.

Los hechos médicos acerca de un paciente se pueden ver como la **evidencia** directa a partir de la cual se infieren hipótesis de diagnósticos y pronósticos posibles, y se **seleccionan** los tratamientos y planes de seguimiento.

La evidencia comprende la historia clínica y los síntomas referidos por el paciente, los signos observados por el médico durante el curso de un examen y los resultados de pruebas específicas para la detección de estados patofisiológicos. La estructura de datos utilizada para describir una evidencia debe incluir detalles acerca de su técnica de medida, el rango de valores, su fiabilidad, el momento, el costo y la relación lógica con otras medidas. Además, estará asociada a la hipótesis por reglas y vías de razonamiento. Las hipótesis requieren normalmente una estructura descriptiva diferente. Estas hipótesis descansan en los conceptos médicos utilizados en el razonamiento, tal como las categorías de diagnóstico y pronóstico aplicables al paciente, pero pueden incluir también una variedad de elementos intermedios, tales como síndromes, estados patofisiológicos y clínicos, los cursos de la enfermedad y asociaciones de datos clínicos. Estos conceptos intermedios se pueden utilizar para definir los conceptos de alto nivel. Aunque la mayoría de las hipótesis se refieren directamente a las condiciones clínicas del paciente, también es posible que sea necesario representar las hipótesis de forma explícita como aserciones acerca de contextos relativos al ambiente que rodea al paciente. Algunas hipótesis pueden ser subconceptos de otras, en cuyo caso pueden heredar propiedades del concepto parental, y otras pueden ser antecedentes causales, lo que implica que deben ser anteriores en el tiempo.

DIAGRAMA GENERAL DEL PROCESO DE UN SISTEMA EXPERTO DE CONSULTA MEDICA





Un Sistema Experto debe representar también los distintos tratamientos que potencialmente serían capaces de controlar la enfermedad del paciente. Los tratamientos estarán interrelacionados en términos de aplicabilidad y factores riesgo/beneficio tales como efectividad terapéutica, toxicidad, posibilidad de interacciones indeseables y otras limitaciones. Para tratar un paciente aquejado de una enfermedad compleja o prolongada se debe formular un plan de actuaciones que debe consistir en las posibles secuencias de tratamientos que están disponibles para controlar los cursos alternativos que puede tomar la enfermedad después de un tratamiento inicial.

En los esquemas de los Sistemas Expertos de consulta médica, es importante representar un amplio campo de alternativas y sus relaciones con las hipótesis y evidencias primarias de los pacientes. Tomando como base estas relaciones, se pueden derivar y explicar las reglas para seleccionar los tratamientos.

Puesto que una componente muy significativa del razonamiento de un especialista médico lo constituyen los juicios, cuando se diseña un sistema programado es necesario instrumentar de la mejor forma posible tales juicios si queremos simular el comportamiento experto.

Una escuela de pensamiento mantiene que sería mejor si se pudiera reemplazar dichos juicios por métodos más objetivos, normalmente del tipo de teoría estadística de decisiones (2). No obstante, incluso con esta filosofía se necesita un conocimiento de los juicios para elegir los umbrales de decisión. Otros han intentado captar la experiencia de los especialistas en la forma de reglas de razonamiento que directamente incorporan elementos de juicio (3). Independientemente de los criterios, es necesario dar un valor relativo a los resultados de razonamiento alternativos. Por lo tanto, se debe incluir una representación de es-

tos valores y la forma de utilizarlos depende de la estructura del conocimiento en el programa, las estrategias generales de la estructura del conocimiento en el programa, las estrategias generales de razonamiento y la naturaleza de los valores involucrados.

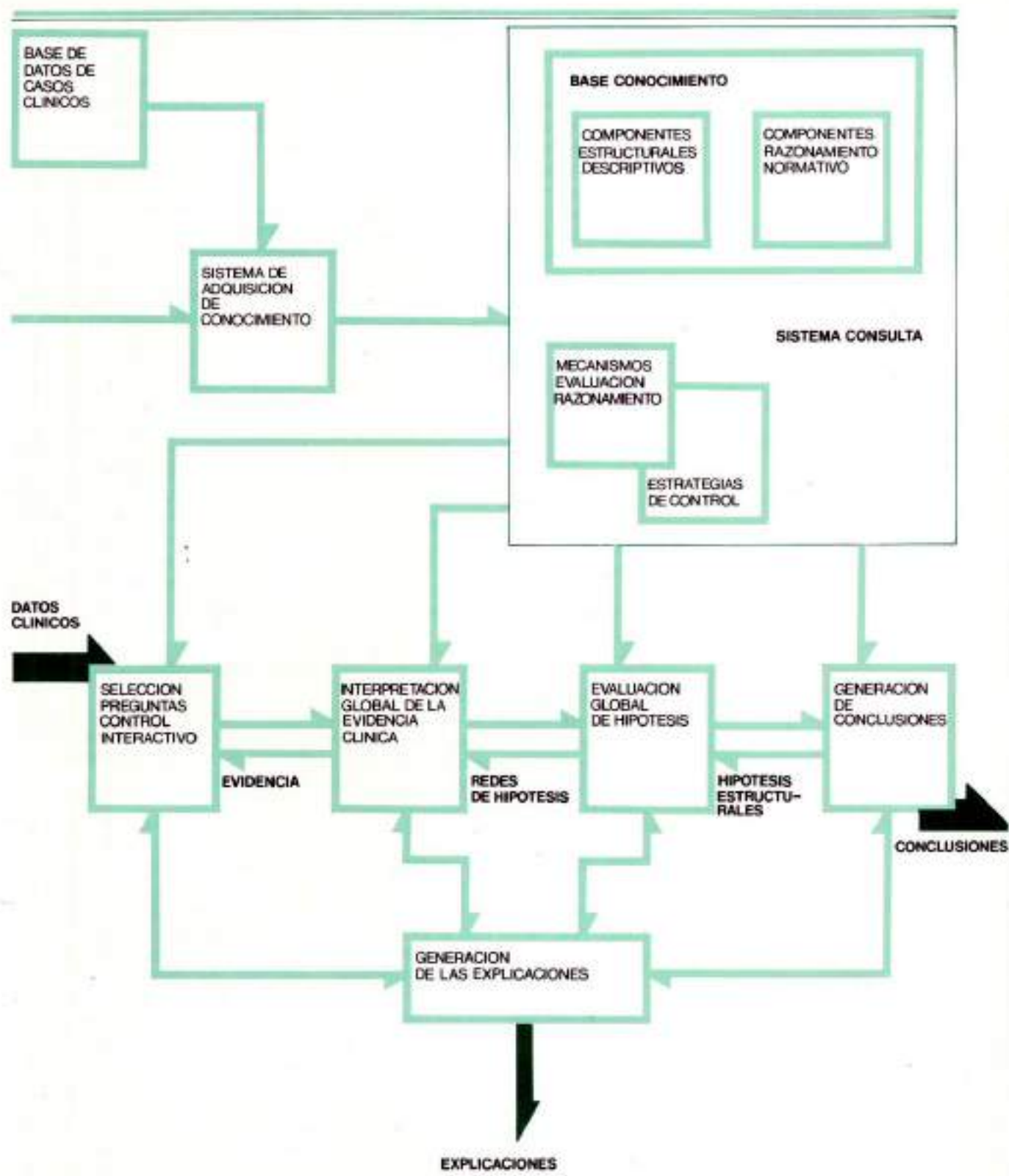
Como ya hemos indicado anteriormente, diferentes médicos pueden no estar de acuerdo en cómo tratar a un mismo paciente, dando cada uno una justificación de su punto de vista. Esta fuente de variabilidad asegura que en la mayoría de los casos no habrá un único modelo "óptimo" ó "correcto" de tratar a un paciente, y el papel de un sistema experto estaría contemplado como la presentación de alternativas, con una clara indicación de la fuente de los juicios que soportan cada decisión. Para formarse una idea de lo que pueden significar los sistemas expertos en Medicina, resulta apropiado repasar las realizaciones actuales, examinar sus características y finalmente escrutar las posibilidades de investigación y desarrollo.

## SISTEMAS CLASICOS

Los prototipos iniciales de programas expertos utilizando conceptos de Inteligencia Artificial (IA), se desarrollaron para su uso en oftalmología (CASNET), enfermedades infecciosas (MYCIN), medicina interna (INTERNIST) y enfermedades renales (PIP). Estos cuatro sistemas y todos sus sucesores comparten ciertas propiedades generales que caracterizan la estructura conceptual y funcional de los desarrollos para aplicaciones médicas de IA (15), (24), (34), (35).

La figura 3 ilustra sobre los componentes básicos y el proceso resultante.

En contraste con otros métodos formales basados en estadística o reconocimiento de formas, hay una separación deliberada de la base de conocimiento específica de un dominio de los mecanismos generales de evalua-





ción y las estrategias de control del sistema. La evaluación de razonamiento y los componentes de control reciben a veces el nombre de "máquina de inferencia". La base de conocimiento se divide claramente entre un **componente descriptivo** de estructuras de datos ligadas por relaciones específicas del dominio (categorizaciones jerárquicas, adscripción de subcomponentes, procedencia causal o antecedencia, etc.), y una **componente normativa** o reglas de razonamiento prescriptivo que operan sobre los componentes descriptivos usando los mecanismos de evaluación en una forma especificada por las estrategias de control.

Esta organización no es más que una variante especializada de la estructura utilizada en la producción generalizada de sistemas de IA.

En los sistemas CASNET, INTERNIST y PIP, el proceso de razonamiento está centrado alrededor de un componente descriptivo estructural explícito. Las redes causales y las taxonomías jerárquicas se pueden ver como casos especiales de redes semánticas (4), que han sido el primer y más usado medio de representar el conocimiento para la interpretación de lenguajes naturales.

En contraste, MYCIN centra su conocimiento alrededor del comportamiento normativo: las reglas de producción. A su componente descriptivo no se le da mayor importancia, aunque el árbol de contexto y la red para actualizar los valores de los parámetros clínicos son cruciales para que la llamada de reglas sea efectiva. Este método puede facilitar la adquisición de conocimiento inferencial estrictamente, pero deja abierta la cuestión de cómo relacionar las producciones específicas a los conceptos prototipo en el dominio médico. El árbol de contexto realiza esta tarea pero de una manera muy específica y subsidiaria.

Los métodos para cuantificar la incertidumbre varían de un sistema a otro,

pero comparten ciertas propiedades generales. En ellos se trata de confirmar o desechar hipótesis como procesos independientes, aunque se necesitan funciones para producir medidas de la confianza general para guiar al proceso de razonamiento.

El número de distintos niveles de incertidumbre estimados subjetivamente por los expertos viene a ser de 5 ó 6 y utilizan funciones de combinación lógica fuzzy para la evaluación de la incertidumbre de una combinación booleana de aseveraciones.

El flujo de información característico de un sistema IA, ilustrado en la figura 3, muestra cómo después de que un conjunto inicial de datos clínicos le hayan sido presentados al programa, las estrategias le pueden conducir a generar interpretaciones locales (tales como decidiendo sobre normalidad, anomalía o consistencia de los datos del paciente o su interpretación en términos de hipótesis directamente relacionadas); requerir más datos según las sugerencias derivadas de la interpretación inicial; proceder a una interpretación global sobre toda la base de conocimientos evaluando y comparando las interpretaciones parciales y seleccionando las agrupaciones de hipótesis más adecuadas y coherentemente estructuradas; generar conclusiones integrando las distintas hipótesis en una declaración final, y produciendo explicación para cualquiera de las etapas anteriores.

La habilidad para reciclar a través de las fases anteriores de razonamiento, permitiendo al usuario requerir explicaciones, y posibilitando cambiar el enfoque del razonamiento mediante la introducción selectiva de nuevos datos, son características que introducen un grado significativo de flexibilidad y generalidad típico de los sistemas de IA.

No obstante, resulta interesante observar que los sistemas de consulta médica que dan indicaciones de tratamiento lo hacen sin recurrir a métodos



generales de planificación. Esto no hace sino reflejar el hecho de que la mayoría de los planes de tratamiento en medicina son a corto plazo y centrados alrededor del control de un número limitado de variables clínicas o fisiológicas, haciendo posible la utilización de estrategias relativamente simples de selección sobre planes alternativos preespecificados.

En la construcción de un sistema de consulta de IA, nos apoyamos más sobre el conocimiento de los especialistas médicos que en los sistemas probabilísticos o de reconocimiento de patrones. La variedad de estructuras utilizadas por los especialistas da lugar a un proceso de adquisición de conocimiento mucho más complejo que el que se encuentran los diseñadores de sistemas tradicionales y se necesita dedicar mucho más esfuerzo a estos problemas, cuando se enfrenta el desarrollo de una nueva aplicación.

## EVOLUCION ACTUAL

Desde la introducción de los primeros sistemas expertos de aplicación médica (CASNET, MYCIN, INTERNIST y PIP), los desarrollos posteriores se pueden diferenciar en dos tipos. En uno de ellos prevalece un cierto sentido de investigación con una preocupación principal en la búsqueda de nuevos métodos de representación del conocimiento. No obstante, el grueso de los nuevos desarrollos está aprovechando la experiencia obtenida con los sistemas del primer ciclo de realizaciones, usando y ampliando las metodologías ya probadas.

En el "Digitalis Therapy Advisor", se han combinado un modelo matemático compartamental para los efectos de la digitalina (droga de uso cardíaco), con los métodos del razonamiento simbólico para la interpretación de los datos del paciente. Después de que una determinación inicial de la dosis de digitalina se introduce en el modelo matemático, el sistema utiliza

información realimentada acerca de la respuesta clínica del paciente a la dosis (incluyendo tanto aspectos cuantitativos como niveles bioquímicos y valores cardíacos como síntomas y aspectos externos) para modificar sus recomendaciones para los niveles de digitalina subsiguientes. El sistema se ha sometido a evaluaciones formales muy rigurosas que han demostrado que sus recomendaciones eran comparables en efecto a la de los expertos clínicos, sugiriendo que el sistema podría ser útil en situaciones de cuidado de la salud cuando no es posible la consulta directa con un especialista en cardiología (5), (6).

El sistema IRIS, ha incluido una generalización de las estructuras de representación utilizadas inicialmente en el CASNET, que usaba una red semántica para representar el conocimiento descriptivo de los procesos patológicos, de los principios de razonamiento y de los controles de estado. El IRIS, ha sido diseñado en realidad como una herramienta para la experimentación con diferentes estrategias de razonamiento y control, más que como un sistema completo de consulta. Esto da al usuario un mecanismo general para tratar los hechos específicos y las hipótesis, y un mecanismo para la propagación de inferencias entre ellas basadas en reglas de producción (7). Las estrategias específicas de control se podrían escribir en INTERLISP haciendo uso de la estructura de base de conocimientos y de los elementos de razonamiento de IRIS. De esta forma se pueden emular fácilmente parte de las estrategias de control de MYCIN, INTERNIST, PIP y CASNET (7).

El sistema MEDICO, que también se aplica a oftalmología, utiliza redes semánticas y de inferencia para la adquisición de conocimiento y el diseño de un sistema consultor experto (8).

El objetivo del sistema TEIRESIAS ha sido facilitar la adquisición de conocimiento de los expertos y la actuación de los modelos del tipo MYCIN, dan-

do un nuevo paso al desarrollar un esquema de meta-reglas, razonando acerca del propio razonamiento del programa MYCIN. El sistema trabaja principalmente analizando los errores del programa experto de consulta, visualizando los hechos del caso específico de consulta, las reglas utilizadas por MYCIN y la traza del razonamiento. Después se introduce en un diálogo de alto nivel dentro de un conjunto restringido de lenguaje natural, con el experto de razonamiento constructor de la base de conocimiento para tratar de descubrir los procedimientos por los que se pueden evitar los errores. Este conocimiento se interpreta por TEIRESIAS para sugerir posibles cambios en las reglas del programa de consulta. Tomado junto con el modelo de consulta, el sistema TEIRESIAS representa un ejemplo importante de un sistema que "conoce lo que el conoce" (9).

Una aplicación diferente de las técnicas MYCIN, consiste en ayudar en el análisis de casos de la base de datos que se han implementado para su uso con ARAMIS (10). (ARAMIS es un sistema de ayuda al diagnóstico desarrollado en la Universidad de Stanford para reumatología, basado en el método formal de comparación lógica del perfil del paciente con casos almacenados en una base de datos).

La necesidad de emular de forma más precisa la secuencia del razonamiento de un experto ha conducido a una nueva formulación de INTERNIST. El interés principal ha consistido en desarrollar una representación que pudiera soportar estrategias para manipular hipótesis múltiples o compuestas y rendir comportamientos que convergen más rápidamente a las conclusiones correctas.

Algunos elementos introducidos en el INTERNIST-II (11) han sido:

- "relaciones constrictoras" para describir asociaciones muy específicas entre los datos de partida y la hipótesis de alto nivel;



- Un generador de hipótesis multi-problema con un cuantificador heurístico modificado para tomar ventajas de las relaciones constrictoras;

- y estrategias de control para evaluar conjuntos de hipótesis en vez de estructuras de hipótesis individuales del sistema original.

Más recientemente se ha adaptado al INTERNIST una etapa de entrada de adquisición de conocimiento basada en el sistema ZOG permitiendo al especialista (experto) introducir sus conocimientos de una forma natural (12).

El problema de representar grupos de hipótesis relacionadas, de tal forma que estén "agregadas" de una forma natural durante la inferencia, ha sido un problema típico para todos los investigadores que tratan con grandes espacios de hipótesis. Esta cuestión cobra un papel principal en el diseño de nuevos programas para el diagnóstico del equilibrio ácido-base y su tratamiento (13).

Uno de los mayores problemas que no ha sido tratado adecuadamente en los esquemas de consulta experta actuales es el del razonamiento sobre secuencias temporales de eventos e hipótesis. Un ataque a esta cuestión se ha aplicado en el sistema VM (14) para tratamiento de ventilación pulmonar, basándose en una reevaluación de reglas en tiempo real dentro de un esquema similar al MYCIN. En esta aplicación, la estrategia dirigida por objetivos del MYCIN no se utiliza, ya que el sistema debe responder en una forma dirigida por eventos a los cambios en el estado fisiológico del paciente en el respirador. La inferencia de cambios en las hipótesis a lo largo del curso de los estados de enfermedades crónicas se ha modelado en el sistema CASNET/Glaucoma (15) mediante funciones del tiempo, y en el "Digitalis Advisor Program" se utiliza la realimentación de los valores de los parámetros fisiológicos. Estos ejemplos representan aplicaciones espe-

cializadas, pero aún se necesita un esquema general de razonamiento que tenga en cuenta la evolución temporal.

Una de las preocupaciones principales en el diseño de los sistemas de consulta tipo MYCIN, ha sido la explicación de los razonamientos, que se han ampliado para incluir información tutorial en el sistema GUIDON (16), (17), orientado a la enseñanza.

En el "Digitalis Advisory Program" se ha desarrollado un esquema de explicación que está basado en modelos fisiológicos y estructurales (18).

Un programa perenne para los diseñadores de los programas basados en el conocimiento ha sido equilibrar la mezcla de las formas de conocimiento declarativas y de procedimiento en sus representaciones. En general, esta cuestión se ha obviado mediante la combinación de redes semánticas con reglas de producción como en los programas IRIS (7), NEUREX (19) y NEUROLOGIST (20), y en los esquemas basados en conocimiento de EXPERT (21) y AGE (22).

En relación con este problema están las cuestiones de modificar las estrategias de control de forma que se aplique la clase idónea de conocimiento para cada aplicación. Un intento de esta dirección lo constituye el sistema MDX, que desarrolla una jerarquía de diferentes "expertos de procedimiento" dentro de un sistema de consulta, con transferencia estricta de los protocolos de control entre ellos. La estructura de los expertos en MDX es paralela a las relaciones entre las distintas sub-especialidades en Medicina. Se necesita más investigación para estudiar, no sólo esta cuestión, sino otras formas más flexibles en las que el control de expertos operando en forma concurrente se pueda coordinar eficazmente (23).

Resulta conveniente anotar algunos aspectos de la experiencia obtenida con estos desarrollos.

El sistema MYCIN, fue inicialmente catalogado como un modelo sencillo que podría ayudar en el desarrollo de muchos otros problemas de diagnóstico médico. Los investigadores que han intentado crear otros programas tipo MYCIN han descubierto que gran parte de su potencia se debía atribuir, no a las características de su estructura, sino a la experiencia de su creador Shortliffe (24). Los intentos para extender MYCIN han requerido, entre otras cosas, estudiar como Shortliffe elegía sus reglas. MYCIN se ha extendido de varias formas. Así, se ha desarrollado un software general para soportar sistemas de producción de reglas similares al MYCIN (EMYCIN) (25), y se han escrito programas para ayuda en el diagnóstico de otras enfermedades, tales como la obstrucción pulmonar y tratamiento de hemorragias (26).

Quizá el mayor esfuerzo se ha dirigido al desarrollo de las capacidades de explicación del sistema en la toma de decisiones (27).

En la experiencia de transferir al sistema MYCIN a nuevos dominios como en psicofarmacología (28) usando EMYCIN, ha puesto de manifiesto algunas dificultades que parecen relacionadas con propiedades más generales de los sistemas basados en reglas. Las más importantes son de dos tipos:

#### 1) Problemas en el encadenamiento hacia atrás

La estructura MYCIN realiza la inferencia por encadenamiento hacia atrás mediante un conjunto de reglas hasta que se alcanza un punto en que no se puede proseguir y se debe pedir al usuario que suministre información adicional. Una de las propiedades de esta estructura de control, es que, tanto la selección como el orden de las preguntas al usuario, se determinan dinámicamente durante el proceso de consulta. No obstante, este proceso choca con la forma en que normal-

mente un médico organiza sus datos sobre el paciente, agrupándolos típicamente atendiendo a: "dolencia principal", "historia de la enfermedad actual", e "historia clínica anterior". Con la estructura de control de cadena hacia atrás, la agrupación de las preguntas viene determinada por el orden de aparición de los parámetros o las variables en las reglas, y por lo tanto, las preguntas acerca de la dolencia y la historia actual, se pueden mezclar con otras relativas a la historia clínica anterior del paciente o de otros miembros de su familia. Este hecho desconcierta en alguna forma al usuario, cuya organización heurística de conocimiento es diferente a la que se le presenta y no acierta a comprender la intencionalidad subyacente.

#### 2) Problemas con la búsqueda paralela

Para la deducción del valor de un parámetro, el software EMYCIN aplica todas las reglas posibles incluso aunque pueda conducir a resultados contradictorios, dando como resultado no un valor único, sino una lista de posibilidades con su factor de certeza asociado que da una medida de la evidencia en favor de esa posibilidad.

El factor de certeza que se asocia a una regla que usa un parámetro en su premisa, es el producto del factor de certeza de la premisa por el de la regla. Este mecanismo funciona satisfactoriamente para la mayoría de los sistemas; por ejemplo, en psiquiatría, pero se han detectado problemas en la parte de selección de terapia.

Este problema no se debe al uso de los factores de certeza sino que también aparece si se utilizan probabilidades convencionales de la misma manera.



## INGENIERIA PARA EL DESARROLLO DE SISTEMAS IA EN MEDICINA

Algunos de los elementos generales de ayuda en la construcción de Sistemas Expertos basados en conocimientos son EMYCIN, EXPERT y AGE. EMYCIN, creado a partir de MYCIN, permite al creador de una base organizarla para que se pueda correr con una estructura de control de tipo MYCIN.

El sistema EXPERT (21), se deriva principalmente en la experiencia con CASNET, y también da un programa general de consulta que se puede rellenar con una base de conocimiento en cualquier especialidad médica que se elija. Su esquema representacional incluye una red causal jerárquica para las hipótesis y los tratamientos, un esquema estructural para los datos primarios de examen del paciente y un conjunto de reglas de producción que permite la especificación de contextos en términos de estos elementos.

EXPERT se está utilizando para desarrollar programas expertos de consulta en reumatología, neurooftalmología y endocrinología (12).

El sistema está diseñado para que médicos con alguna experiencia en informática puedan construir modelos escribiendo ellos en un archivo, con cualquier sistema editor, usando un lenguaje descriptivo sencillo. El archivo se compila después por un programa especial que chequea los errores sintácticos y producen un modelo compilado que se puede correr por el programa de consulta. También se puede actualizar la base de datos y la adquisición de conocimientos para ayudar en el proceso de verificación del modelo y se prueba con casos conocidos fiables (21).

El sistema AGE (attempt to generalize) (22), provee un conjunto general de herramientas técnicas para modelar situaciones de consulta usando el modelo de "pizarra" (29), (30), que se ha desarrollado para manipular la repre-

sentación y el procesado de información procedente de múltiples fuentes de conocimiento en el reconocimiento de expresión verbal.

La construcción de un modelo experto de consulta con AGE, requiere el conocimiento de INTERLISP, por lo que este sistema está diseñado principalmente para su uso por especialistas en informática trabajando con especialistas médicos.

### EXPECTATIVAS

En el estado actual de nuestros recursos tecnológicos existe la capacidad objetiva para diseñar y realizar sistemas expertos que funcionen de forma fiable, en un amplio rango de aplicaciones del campo médico.

Es evidente que será necesario potenciar la investigación sobre las herramientas de diseño y muy especialmente lo que respecta al manejo conversacional de nuestro idioma y otros elementos de diálogo máquina-usuario. No obstante, existen cuestiones básicas, no estrictamente técnicas, que son elementos clave en la viabilidad del desarrollo de una aplicación.

¿Con qué maestros de la Medicina se puede contar para un desarrollo de este tipo y cómo se les puede motivar para colaborar en el proceso de extraer su experiencia?

¿Cómo podemos motivar de forma efectiva el uso de Sistemas Expertos por los médicos u otros usuarios potenciales?

¿Existe una dimensión de mercado mínima fuera del ámbito de la investigación?

Para la introducción y difusión de los Sistemas Expertos en Medicina es importante considerar tres factores: sencillez de la aplicación, teoría médica y factores humanos.

La experiencia en el desarrollo del mercado informático médico muestra

que el análisis automático de electrocardiogramas (ECG), ha sido una aplicación comercialmente viable porque se trata de un problema específico y de gran volumen de demanda.

El registro de historias clínicas y el área de laboratorios son también aplicaciones idóneas para la automatización e introducción de sistemas informáticos. En ambos casos, se presentan problemas muy sencillos tales como copiar la información de donde está almacenada para enviarla donde hace falta y también presentan un gran volumen de trabajo ya que virtualmente cada paciente tiene registrada una historia y se somete a pruebas de laboratorio. Las necesidades inmediatas pasan por establecer una base de datos y aunque no exista una necesidad perentoria de establecer programas de ayuda al diagnóstico, estos sistemas se pueden añadir más tarde a la base de datos. Al principio, los Sistemas Expertos así implantados deben ser muy sencillos, y según progrese la automatización en otras áreas del hospital, compartirán más información dando la oportunidad para que el diagnóstico automático crezca en paralelo con el aumento de la base de datos y la familiaridad del usuario con los medios informáticos.

En el desarrollo de Sistemas Expertos aplicados en Medicina, se debe tener en cuenta que el mayor problema no radica tanto en las herramientas informáticas disponibles, sino, más bien, en el entendimiento del proceso de asistencia médica.

En los últimos años ha habido un gran desarrollo de técnicas de medida y análisis clínicos, de forma que ahora, el médico tiene a su disposición resultados sobre cateterización cardíaca, análisis de enzimas, tomogramas axiales computarizados y un largo etcétera de parámetros. No obstante, no se ha avanzado a la misma velocidad en la investigación de cómo interpretar estos datos o cómo usarlos adecuadamente para establecer el diagnóstico



(31). Antes de que se pueda desarrollar con éxito una aplicación de Sistema Experto se debe conocer y atender bien el proceso cognoscitivo que lo soporta. Es por ello que algunos estudios del tema (32) han expresado la necesidad del desarrollo de la "lógica de la Medicina". Más aún, se insiste en que la teoría de la Medicina debe hacer más hincapié en las acciones que en los diagnósticos, teniendo en cuenta que el proceso de cuidado de la salud es de naturaleza dinámica y las pruebas y los tratamientos realizados a través del tiempo son más importantes que los diagnósticos en sí.

Los problemas básicos en este tema son:

1) Las secuencias lógicas subyacentes a los procesos de diagnóstico clínico no se deducen fácilmente de los textos o artículos médicos y, por tanto, es muy difícil expresarlas en forma de algoritmos.

2) Existe un conocimiento incompleto relativo a las probabilidades de los parámetros clínicos para una enfermedad dada y viceversa. Los informes médicos son deficientes en el registro de síntomas y signos residuales subsiguientes al tratamiento y cura de la enfermedad.

3) No existen todavía métodos satisfactoriamente estructurados en el enfoque del diagnóstico. El método lógico, que busca parámetros específicos o intenta resolver ciertos objetivos intermedios, limita en gran manera el ámbito de las posibilidades de enfermedades "diagnosticables". El método de las probabilidades puede extraer todas las posibilidades pero adolece de problemas cuando se presentan varias enfermedades simultáneamente.

Alguien se preguntaría ¿por qué no recoger todos los textos en medicina y toda la documentación disponible, en forma automática y formar una base de conocimiento completa? Incluso en este caso, seguimos contando con obstáculos importantes:

1) Aunque se han hecho muchos progresos en el análisis de textos en algún idioma como el inglés, poco o nada se ha hecho respecto del castellano, aún es muy difícil obtener extracto sistematizado a nivel de computador, incluso de hechos simples a partir de un cuerpo de texto.

2) Sólo existen nomenclaturas normalizadas en ciertas especialidades médicas. Sólo un número muy reducido de servicios de nuestros hospitales usan vocabularios restringidos en sus informes e historias clínicas, y desde luego, los autores de libros no limitan en absoluto su léxico. No existe un "Diccionario" de la terminología médica y no se sabe bien como crearla.

3) Las herramientas de programación utilizadas para la construcción de los Sistemas Expertos son incomprensibles para la mayoría de los médicos. La labor de incorporar nuevo conocimiento resulta tediosa y frustrante para el médico que es una persona a la que le falta tiempo y más si es un especialista.

El rango de aplicaciones de los Sistemas Expertos en Medicina, depende, en gran parte, del peso relativo de los procesos cognoscitivos frente a las habilidades perceptuales utilizadas por el especialista humano en cada caso. Si en un tipo de aplicación potencial, la obtención de datos descansa en un examen físico involucrando la detección de signos aparentemente imperceptibles a través de inferencias táctiles o visuales, de forma que esta información constituye la base del dictamen, no es razonable esperar que un Sistema Experto pueda ayudar significativamente en este tipo de evaluaciones.

Si por el contrario, el papel principal del experto humano a emular consiste en emitir un diagnóstico o seleccionar un tratamiento en base de interpretaciones sobre datos de análisis y una historia clínica bien definida, entonces no es desatinado investigar estos procesos de interpretación e intentar

simularlos mediante un Sistema Experto programado.

Si además es posible construir una base de conocimiento que incorpora modelos descriptivos de mecanismos patofisiológicos y componentes normativos del razonamiento experto y si las estrategias de explicación se pueden formular tal que permitan al programa contestar problemas acerca del propio razonamiento, entonces parece razonable invocar que un sistema con estas características muestra ciertos elementos de "entendimiento" no muy diferentes de los expuestos de manifiesto por los especialistas médicos.

La Tabla II da una indicación de las aplicaciones desarrolladas hasta el momento, dando una idea de las posibilidades prácticas de este tipo de sistemas.

La experiencia en la introducción de sistemas informáticos en el hospital a lo largo de los últimos años, indica que la limitación más importante para el desarrollo de una aplicación, no radica en el hardware ni en el software sino en el "medicalware", es decir, en la parte que tiene que ser suministrada por el especialista en el tema.

Hay que tener en cuenta que la formación y modo de operar de los profesionales de la informática no siempre se acomoda con los del campo biomédico. Tradicionalmente los médicos no han tenido una buena formación matemática de forma, que el diálogo con los ingenieros de sistemas resulta a menudo frustrante para ellos.

El desarrollo y perfeccionamiento de un programa experto involucra la dedicación de bastante tiempo del especialista médico que aporta el conocimiento. Los médicos y en mayor medida los buenos especialistas, son personas con muy poco tiempo, y éste tiene un alto valor. No es de extrañar que los programas más conocidos (MYCIN, INTERNIST) se hayan debido a grandes médicos con una larga ex-



perencia en su campo y con una gran pasión por el uso de los computadores. Por ello, el éxito de un programa experto en el área biomédica depende fundamentalmente de encontrar este tipo de profesionales multidisciplinarios con motivaciones profundas para abordar este tipo de desarrollos.

Incluso eligiendo un problema relativamente sencillo, y contando con un "medicalware" idóneo, aún hace falta considerar un aspecto fundamental para que el desarrollo de un sistema tenga éxito en su aplicación: el sistema debe interaccionar con los usuarios. En los desarrollos de informática médica que se produjeron en los 60 y parte de los 70, el papel del usuario no fue tenido en cuenta y fue origen, en gran parte, de que los ordenadores no hayan jugado el papel que se esperaba en Medicina. En definitiva, el factor humano juega al final un papel primordial en la aceptación "comercial" del producto.

En las aplicaciones propias de los sistemas expertos, dentro de la complejidad del diagnóstico médico, el usuario debe estar activamente involucrado en la operación de sistemas, y debe ser capaz de manipular el equipo sin tener que llamar a un programador o aprender un nuevo lenguaje. En definitiva, se trata de que se diseñe el sistema con alto nivel de cordialidad con el usuario.

## RECURSOS HUMANOS

La captación de expertos médicos puede tener más probabilidades de éxito en los servicios hospitalarios que ya tienen una cierta experiencia en el uso de sistemas informáticos. Esto sucede en Medicina Nuclear, Radioterapia, Neurocirugía, Cardiología y fundamentalmente, en las unidades de Investigación.

Se estima que el número total de expertos en IA que existe en el mundo, no sobrepasa la cifra de 300 (33) y en

nuestro país, se pueden contar con los dedos de la mano. Aunque el trabajo de construcción de un Sistema Experto se ha reducido de 50 a 5 hombres-año, la disponibilidad de una aplicación por un usuario potencial, se sitúa en varios años a causa de la escasez de especialista en IA. Por otra parte, el costo de uno de estos "expertos en la construcción de expertos" junto con su equipamiento necesario se sitúa entre 150.000 y 300.000 \$ año. Esto supone que el desarrollo de un sistema, incluso modesto, quedará fuera del alcance de la mayoría de los laboratorios de investigación si no cuentan con apoyos institucionales fuertes.

Por mucho que hablemos acerca de las grandes cualidades de los Sistemas Expertos, éstos no tendrán impacto en el área médica hasta que al menos una aplicación esté funcionando en un gran hospital. No es probable que las aplicaciones clínicas surjan desde una iniciativa comercial en nuestro país. De hecho, la industria de Sistemas Expertos que está naciendo en USA, Europa y Japón, busca un mercado a través de la oferta de soportes generales de construcción.

Parece poco viable que se puedan comercializar sistemas para el campo médico de tipo "pret a porter" debido a las limitaciones de la base de conocimientos para cada aplicación. La cuestión de aceptabilidad impone a corto plazo los sistemas "a medida" para un usuario específico, lo que supone una barrera muy fuerte a la iniciativa comercial e incluso para los laboratorios universitarios fuera del contexto hospitalario.

En consecuencia, los ingenieros y médicos que trabajan en informática médica, se verán forzados a incorporar los conceptos de IA en su software de aplicación, formándose a sí mismos como expertos en la construcción de expertos.

La experiencia demuestra que el desarrollo de modelos que trabajen a nivel de experto necesitan de colaboracio-

nes interdisciplinarias intensas y esta forma de trabajo continuará siendo la base de los desarrollos y de investigaciones, al menos hasta que haya muchos más especialistas que combinen un entrenamiento avanzado en ambos campos.

### La cuestión de aceptabilidad

Después de revisar las publicaciones sobre los sistemas expertos aplicados en Medicina, se puede observar que prácticamente todos ellos se han limitado al uso dentro de áreas de investigación, incluso cuando sus prestaciones han sido excelentes. Este hecho sugiere que es un error concentrar sólo el esfuerzo de investigación y desarrollo en los aspectos técnicos, cuando el impacto clínico depende en gran medida, de la solución de los problemas de aceptación por el usuario.

No se puede olvidar que existen fuertes reticencias en algunos sectores de la profesión médica hacia los sistemas computerizados y este extremo cobra mayor importancia en la introducción de los Sistemas Expertos en la práctica clínica. Los argumentos contra la adopción de esta nueva tecnología abarcan desde los más radicales, que acusan de que se está intentando suplantar al médico, hasta los menos emocionales que se preguntan acerca de problemas éticos o pérdidas de control de la calidad de la asistencia médica. Varias encuestas en distintos países, muestran que las actitudes de los médicos ante nuevas tecnologías son opuestas a todas las aplicaciones que menoscaban, o parezcan hacerlo, su papel profesional ante la sociedad. Al mismo tiempo, temen quedar fuera de un nuevo campo de desarrollo que no les permita alcanzar los mismos niveles que otros colegas, y en cualquier caso, terminan aceptando las ayudas informáticas que realmente mejoran la calidad de la asistencia sanitaria. No obstante, detestan los sistemas de datos centralizados que invitan a la inter-

vención de las autoridades y en general, opinan que hay tareas y decisiones que nunca deberán salir del ámbito personal para dejarlas en manos de una máquina.

Se debe prestar una atención adecuada a la extraordinaria limitación de tiempo que presentan los médicos en general. Los sistemas computerizados no deben hacerles tomar más tiempo de lo que ellos normalmente dedican a la misma tarea y un parámetro especialmente crítico es el tiempo que necesitan invertir para familiarizarse con la nueva técnica. Aunque no pueda parecer tan importante a primera vista la capacidad de un programa acerca del auto-conocimiento, es decir, el "conocer lo que él conoce", la experiencia demuestra que es uno de los aspectos más sensibles para la aceptación de los sistemas expertos. Los profesionales que ejercen toma de decisiones, y en particular los médicos, dan más crédito a los sistemas que parecen comprender sus propias limitaciones y capacidades y que conocen cuándo admitir su ignorancia o su incapacidad para emitir alguna conclusión. Además, los médicos valoran muy positivamente la posibilidad de chequear a sus "asistentes" computerizados, lo que se facilita si los programas presentan la capacidad para explicar, no sólo la cadena de razonamientos que conducen a sus decisiones, sino también la estrategia en la solución de los problemas.

En resumen, la introducción de las técnicas de ingeniería del conocimiento, aplicadas a programas de diagnóstico médico, debe abordar dos aspectos:

- 1) asegurar unas prestaciones propias del experto,
- 2) lograr la aceptación por el usuario.

Estos elementos deben ser tenidos en cuenta desde el conocimiento del diseño ya que dictarán en cada aplicación y entorno específicos, la metodología y el dominio de tareas que en sí mismo debe satisfacer.



Una lección que se debe aprender de la experiencia en desarrollos anteriores, consiste en reconocer que todavía queda mucha tarea de investigación básica en informática médica, y que este campo requiere algo más que la simple aplicación de las técnicas de programación a los problemas médicos.

Debemos tener en cuenta las limitaciones de la tecnología actual, la complejidad de los problemas clínicos, la importancia de los factores humanos y la estructura socioeconómica en que se mueve el sistema de Salud en España. No hace falta diseñar sistemas perfectos, ni disponer la tecnología ideal, sino una elección sensata y razonada que sea el mejor compromiso para un problema específico dentro de un entorno concreto.

#### SISTEMAS EXPERTOS EN MEDICINA

**ABEL:**  
Diagnóstico desórdenes electrolíticos. MIT.

**CAA:**  
Análisis arritmias. Universidad de Toronto.

**CASNET:**  
Asociación de tratamiento glaucoma con hipótesis de diagnóstico. Universidad Rutgers.

**DIGITALIS THERAPY ADVISOR:**  
Asistencia inteligente de tratamientos con digitalina. MIT.

**GUIDON:**  
Enseñanza asistida. Stanford University.

**HODGKINS:**  
Planificación de diagnósticos. MIT.

**INTERNIST/CADUCEUS:**  
Diagnóstico en medicina interna. Univ. de Pittsburgh.

**INTERNIST II:**  
Ampliación INTERNIST.

**MDX:**  
Diagnóstico Universidad del Estado de Ohio.

**MEDAS:**  
Diagnóstico. Universidad de Southern (California).

**MEDICO:**  
Oftalmología.

**MYCIN:**  
Diagnóstico de enfermedades infecciosas. Universidad de Stanford.

**NEUREX:**

**NEUROLOGIST:**

**ONCOCIN:**  
Asistencia inteligente, tratamiento quimioterápico cáncer. Universidad de Stanford.

**PIP:**  
Diagnóstico. MIT.

**PUFF:**  
Diagnóstico enfermedades pulmonares. Universidad de Stanford.

**RAYDEX:**  
Asistencia inteligente en Radiología. Universidad Rutgers.

**RECONSIDER:**  
Diagnóstico rápido en Medicina. Universidad de California en San Francisco.

**RX:**  
Evaluación de datos estadísticos de pacientes crónicos. Universidad de Stanford.

**VM:**  
Control ventilación pulmonar, monitorizando al paciente y sugiriendo terapia respiratoria. Stanford University.

#### SISTEMAS EXPERTOS EN AREAS PARAMEDICAS Y DE INVESTIGACION BIOLÓGICA

**CRYSLIS:**  
Análisis de datos relativos a la cristalografía de proteínas. Universidad de Stanford.

**DENDRAL:**  
Determinación de la estructura molecular por espectrografía de masas. Universidad de Stanford.

**GENESIS:**  
Ingeniería Genética. Intelligenetics Inc.

**METADENDRAL:**

Ayuda a formular reglas tomando datos del espectrómetro de masas sobre fragmentos moleculares. Universidad de Stanford.

**MOLGEN:**

Planificación de experimentos de bioingeniería con DNA. Universidad de Stanford.

**REX:**

Análisis estadístico. Bell Laboratories.

**SECS:**

Síntesis de componentes orgánicos. Universidad de California en Santa Cruz.

**SPEX:**

Diseño de experimentos en genética molecular. Universidad de Stanford.

**SMP:**

Resuelve ecuaciones trascendentes, diferenciales e integrales, Inference Corp.

**TUNER:**

Basado en Pascal para ajustar los sistemas de proceso de señales - TRW Corp.

#### MYCIN

**If** 1) La infección que requiere terapia es meningitis.

**And** 2) el paciente muestra una seria infección de la piel o los tejidos blandos.

**And** 3) no se han observado organismos en el cultivo

**And** 4) el tipo de infección es bacteriana.

**Then:** Existe evidencia de que el organismo que puede ser la causa de la infección es estafilococcus - coagpos (.75) o estreptococcus (.5).

*Ejemplo de proceso deductivo en el sistema experto MYCIN utilizado en el diagnóstico de enfermedades infecciosas.*

## EJEMPLO DE OPERACION CON INTERNIST-1

**Internist 1:** Enter case name

El médico introduce el nombre del caso.

**Internist 1:** Initial positive manifestations.

**Médico:**

- + Age gtr than 55
- + Arthritis HX
- + Depression HX
- + Sex female
- + Chest pain lateral sharp
- + Liver enlarged moderate
- + Feces guaiac test positive

**Internist 1:** Initial negative findings.

**Médico:**

- Alcoholism chronic HX
- Diarrhea chronic
- Fever
- Uric acid blood increased

**Internist 1:** Disregarding: Depression HX, weight increase recent HX, chest pain lateral exarcebation with breathing, chest pain lateral sharp.

.....  
Considering: Age gtr than 55, sex female, urine dark HX, anorexia, feces light colored, jaundice, liver enlarged moderate.  
.....

**Internist 1,** enfoca su atención sobre el área del problema formado por las hipótesis de enfermedad y sus manifestaciones asociadas con puntuación más alta.

**Internist 1:** Please enter findings of lipid (s) blood.

Buscando información genérica adicional para ayudar en la selección de las hipótesis altas.

**Médico:** Go.

El médico no tiene información adicional para responder.



**Internist 1:** Cholesterol blood decreased?

Pregunta por el elemento específico que acorta la pregunta genérica que el médico no puede contestar.

**Médico:** N/A.

No aplicable.

.....

La interacción continúa hasta que el Internist 1 alcanza un diagnóstico.

**Internist 1:** Conclude: Hepatitis chronic active.

Internist 1, recicla ahora al grupo más próximo de hipótesis alternativas y continúa hasta que se explican todos los síntomas.

## BIBLIOGRAFIA

1. **Gavarrar, W. B.:** "Expert systems: limited but powerful". *IEEE Spectrum*, vol. 20, n.° 8, pp. 39-45, 1983.
2. **Gremy, F.:** "Decision-Making and medical care: Can information science help?". In: *Proc. IFIP Working Conference Decision-Making Med. Care*, F. T. de Dombal and F. Gremy (Eds.), North-Holland, pp. 32, 1978.
3. **Shortliffe, E. H., and B. Buchanan:** "A model of inexact reasoning in Medicine". *Math. Biosci.*, vol. 23, pp. 351-379, 1975.
4. **Quillian, M. R.:** "Semantic Memory". In: *Semantic Information Processing*, M. Minsky (Ed.), Cambridge, MA: MIT Press, 1968.
5. **Gorry, G. A.; H. Silverman, and S. Panker:** "Capturing clinical expertise: A computer program that considers clinical responses to digitalis". *Am. J. Med.*, vol. 64: 452-460, 1978.
6. **Silverman, H.:** "A digitalis therapy advisor". MIT, Cambridge MA, Rep. MAC-TR, 143, 1975.
7. **Trigoboff, M., and C. A. Kulikowski:** "IRIS: A system for propagation of inferences in a semantic net". In: *Proc. 5th Int. Joint Conf. Artificial Intell.*, 274, 1977.
8. **Walser, R. L., and B. H. McCormick:** "Organization of clinical knowledge in MEDICO". In: *Proc. 3rd Illinois Conf. Med. Inform. Systems*, 159, 1978.
9. **Davis, R.:** "Interactive transfer of expertise: Acquisition of new inference rules". *Artificial Intell.*, vol. 12, 121-158, 1974.
10. **Blum, R. L., and G. Wiederhold:** "Inferring knowledge from clinical data banks utilizing techniques from artificial intelligence". In: *Proc. 3rd Annu. Symp. Comp. Applic. in Med. Care*, Washington, 303-307, 1979.
11. **Pople, H.:** "The formation of composite hypotheses in diagnostic problem solving". In: *Proc. 5th Int. Joint Conf. Artificial Intell.*, Boston, MA, 1030-1037, 1977.
12. **Freiherr, G.:** "The seeds of artificial intelligence: SUMEX-AIM". *Div. Res. Resources, NIH, Pub. 80-2071*, 1979.
13. **Patil, R. S.:** "Design of a program for expert diagnosis of acid-base and electrolyte disturbances". *AIM Workshop M.I.T. Tech. Note*, 1979.
14. **Fagan, L. M.; J. C. Kunz; E. A. Feigenbaum, and J. J. Osborn:** "Representation of dynamic clinical knowledge: Measurement interpretation in the intensive care unit". In: *Proc. 6th Int. Joint Conf. Artificial Intell.*, Tokyo, 260-262, 1979.
15. **Weiss, S.; Kulikowski, C. A., and Safir, A.:** "Glaucoma consultation by computer". *Comput. Biol. Med.*, vol. 8: 24-40, 1978.
16. **Clancey, W. J.:** "Tutoring rules for guiding a case method dialogue". *Int. J. Man-Mach. Studies*, vol. 11: 25-49, 1979.
17. **Clancey, W. J.:** "Dialogue management for rule based tutorials". In: *Proc. 6th Int. Joint Conf. Artificial Intell.*, Tokyo, 155-161; 1979.
18. **Swartout, W. R.:** "Producing program explanations from multiple hierarchical models". *5th AIM Workshop MIT Tech. Note*, 1979.
19. **Reggia, S. A.:** "A production rule system for neurological localization". In: *Proc. 2nd Annu. Symp. Comput. Application in Med. Care*, Washington, 254-260; 1978.
20. **Catanarite, V. A., and Greenburg, A. G.:** "Neurologist: A computer program for diagnosis in neurology". *Proc. 3rd Annu. Symp. Comput. Application in Med. Care*, Washington, 64-71, 1979.
21. **Weiss, S., and Kulikowski, C. A.:** "EXPERT: A system for developing consultation models". In: *Proc. 6th Int. Joint Conf. Artificial Intell.*, Tokyo, 942-950, 1979.
22. **Nii, H. P., and Aiello, N.:** "AGE (attempt to generalize): A knowledge-based program for building knowledge based programs". In: *Proc. 6th Int. Joint Conf. Artificial Intell.*, Tokyo, 645-655, 1979.
23. **Chandrasekharan, B.; Gómez, F.; Mittal S., and Smith, J.:** "An approach to medical diagnosis based on conceptual schemes". In: *Proc. 6th Int. Joint Conf. Artificial Intell.*, Tokyo, 134-142, 1979.
24. **Shortliffe, E. H.:** "Computer based Medical Consultations: MYCIN". New York: Elsevier, 1976.
25. **Van Melle, W.:** "A domain-independent production rule system for consultation programs". In: *Proc. Int. Joint Conf. Artificial Intell.*, 923-925, Tokyo, 1979.
26. **Bennett, J., and Goldman:** "D CLOT A knowledge based consultation for bleeding disorders". Computer Science Dept. Stanford Univ. Memo HPP-80-7, 1980.
27. **Clancey, W.; Shortliffe, E., and Buchman, B.:** "Intelligent computer aided instruction for medical diagnosis". *Proc. 3rd Ann. Comp. Appl. Med. Care*, R. Dunn (Ed.), 117-122, IEEE Comp. Society: Los Alamitos, California, 1979.
28. **Brooks, R., and Heiser, J.:** "Some experience with transferring the MYCIN System to a new domain". *IEEE Trans. on Pattern Analysis and Medicine Intelligence*, vol. PAMI-2, n.° 5: 477-478, 1980.
29. **Lesser, V. P.:** "Organization the HEARSAY-II speech understanding system". *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-23: 11, 1975.
30. **Lesser, V. R., and Ertan, L. D.:** "A retrospective view of the HEARSAY-II architecture". In: *Proc. 5th Joint Conf. Artificial Intell.*, Boston MA, 79, 1977.
31. **Warner, H.:** "Computer-assisted medical decision making". *Academic Press*, New York, 1979.
32. **Murphy, E.:** "The logic of medicine". Johns Hopkins University Press, Baltimore, 1976.
33. **Schindler, M.:** "Artificial Intelligence begins to pay off with expert systems for engineering". *Electronic Design*, vol. 32, n.° 16: 106-146, 1984.
34. **Myers, J., and H. Pople:** "INTERNIST: A consultative diagnostic program in internal medicine". *Proc. 1st Ann. Computer Appl. Med. Care*, IEEE, New York, 1977.
35. **Pauker, S. G.; Gorry, G. A.; Kassirer, J. P., and Schwartz, W. B.:** "Towards the simulation of clinical cognition: Taking a present illness by computer". *Am. J. Med.*, vol. 60: 981-996.



## CONSIDERACIONES GENERALES SOBRE APLICACION EN EL CAMPO DE LOS NEGOCIOS

**Juán Luis VALDERRABANO**

Ingeniero Industrial  
Dtor. de Proceso de Datos  
del grupo GENERALI

"En ocasiones tuve que asesorar a empresas sobre la compra de un computador, y mi consejo fue que antes de tomar ninguna decisión o compromiso deberían analizar cuidadosamente si necesitaban o no este dispositivo y lo que iban a hacer con él. Pronto observé que mi consejo no era útil, ya que una compañía adquiere la capacidad de tomar decisiones sobre cualquier tecnología nueva a través de la experiencia con la misma." ("The new Science of Management Decision" H. A. Simon. Premio Nobel Economía 1979).

Parece pues que gran parte de la decisión sobre cómo y dónde utilizar Sistemas Expertos corresponderá a los resultados que se obtengan de la experiencia en el uso de los mismos. Pero al menos podemos conjeturar sobre las áreas de la empresa y del mercado en las que pueden ser útiles. Por lo pronto estamos seguros de que el uso de estos sistemas aportará un beneficio de tipo general a nivel de empresa, que se corresponde con una cualidad inherente a la naturaleza de los Sistemas Expertos. Nos referimos a la mejora de la calidad de las decisiones. En efecto, dentro de la Psicología de la Gestión (ver por ejemplo H. J. Leavitt), está experimentalmente comprobado que las decisiones del grupo son de mejor calidad que las del individuo. El uso de estos sistemas aproximará la dinámica de la decisión en grupo a través del contraste que se producirá cuando un utilizador del dispositivo analice la línea de razonamiento seguida por un Sistema Experto para llegar a la decisión, lo que mejorará ineludiblemente el nivel medio de las decisiones. Este es un fenómeno general independiente del área de aplicación. Entrenamiento y formación pues, son dos actividades que pueden verse beneficiadas con la llegada de estos sistemas, por lo que una de las primeras funciones al margen de su trabajo principal puede consistir en servir de instructores.

Respecto al área, diría que hay dos tipos de consideraciones a hacer, unas de tipo general y otras de índole particular. En efecto, los Sistemas Expertos modelan un proceso complejo como es el de la toma de decisiones, y básicamente, aquellas decisiones que entendemos como difíciles, difusas o sobre las que la experiencia puede ser un factor determinante. Este tipo de decisiones se caracteriza por la indefinición del problema o de las soluciones, ya que o no se conocen todas las alternativas o el propio problema está mal definido.

Por lo tanto hay que aplicar reglas provenientes de la experiencia, juicio, etc. del experto que muchas veces son difíciles de modelar precisamente por lo poco concretas, o por la sutileza que requieren los grados de aplicación. A través de un proceso cíclico donde se comparan las consecuencias de decisiones tomadas en otras ocasiones anteriores con la situación del momento presente se alcanzan niveles de decisión cada vez mejores, asociados a una o varias alternativas.

Esta es una labor compleja y difícil que normalmente realizan personas de experiencia en el terreno. Por otra parte, existe toda una jerarquía de temas en los cuales no estamos dispuestos a ceder el derecho a decidir. Por ejemplo, es claro que en temas relativos a la salud, nuestra disposición ofrezca una cierta resistencia a poner la solución de nuestro caso en manos de un artefacto, aunque en otros casos bastará que reconozcamos que es un tema demasiado complejo para un ser humano. Un ejemplo de esto último es el control global del tráfico de una ciudad o de una estación de ferrocarril, que aún estando monitorizados por un operador humano aceptamos que son lo suficientemente complejos como para estar asistido por un computador.

No es raro pues, pensar que este tipo de sistemas preveamos se utilicen en áreas como el asesoramiento fiscal o el control de cierto tipo de procesos



en forma activa, en tanto que operarían en forma pasiva en otras áreas como la Medicina. El término activo o pasivo se refiere a la sustitución o no del experto que tomaba la decisión antes de utilizar este dispositivo. Un sistema como Mycin instruye a un médico en período de formación, al exponer las reglas utilizadas para llegar al diagnóstico, pero es difícil aceptarlo en el papel de decisor solitario de cara a un enfermo, fundamentalmente por el problema de la responsabilidad. Un autor muy conocido, J. McCarthy, señala que en el caso de dirigirse al sistema explicándole que el tratamiento falló no existe la posibilidad de reconsiderar la situación en el tiempo, o que carece de sentido común, por lo que es muy difícil considerarlos en una posición activa.

Por el contrario, aquellos sistemas que se utilicen para el control de procesos dentro de la Ingeniería, la Geología, etc. una vez superado un período de prueba equivalente al de un operador humano, podrían ser utilizados de forma regular, al mismo tiempo que se reduciría el tiempo de entrenamiento y formación. Como ejemplos figuran un sistema que controla un horno de cemento o la operación automática de un tren.

Dentro de los aspectos particulares hay algunos que creo especialmente curiosos. Son aquellos relativos a decisiones que podemos calificar de irrepetibles. En efecto, hay situaciones en las que una decisión desencadena un proceso que es irreversible, o bien que la misma situación es irrepetible. La Economía, además de la Política y la Medicina ofrecen ejemplos inmejorables. Dentro de este área el Sistema Experto tendría un papel equivalente al del antagonista en la teoría de juegos o protagonizaría la simulación de situaciones en las que se analizarían previsible escenarios en función de lo que hubieran hecho uno u otro participante.

Por ejemplo, existe un Sistema Experto que permite a una empresa modelar

su estrategia dentro del mercado simulando diferentes políticas que se describen en términos análogos a los utilizados en la vida ordinaria. Se analiza la repercusión de las mismas sobre los beneficios y por lo tanto, permite establecer una serie de cursos de acción potenciales.

Vamos a sugerir lo que a nuestro modo de entender, son áreas específicas en las cuales puede ser rentable invertir dinero en el desarrollo de un Sistema Experto. Téngase en cuenta que el problema más importante, como hemos mencionado, aparece a la hora de extraer el conocimiento de un grupo de expertos y estructurarlo dentro de un programa. Referente al conocimiento, representarlo no sólo es una labor de habilidad, razón por la cual empiezan a existir los denominados ingenieros del conocimiento, sino que necesita de un vehículo para expresarlo y de una cierta dosis de creatividad para sintetizarlo. La clave reside fundamentalmente en lograr hacerlo con sencillez. Algunos autores comparan el estado actual de estos sistemas al de la ingeniería genética antes del descubrimiento del ADN, en el sentido de que aún falta una buena teoría sobre el conocimiento que lo haga manejable. Se espera que dentro de 5 ó 10 años lo que se ha hecho hoy aparezca como una aproximación muy cruda al problema real.

Probablemente desde un punto de vista comercial interesen sistemas orientados hacia los consumidores, aplicados a áreas relativas a juegos, esparcimiento, etc. Pero el sector de los servicios se encuentra con un gran volumen de gastos en concepto de personal y aunque se han automatizado desde las últimas décadas los trabajos repetitivos a nivel administrativo todavía quedan áreas complejas propicias de ser rentabilizadas a través del uso de estos sistemas. Por lo tanto nos atrevemos a sugerir las siguientes aplicaciones:

—Diseños de Sistemas Sencillos: Por ejemplo diseño de mobiliario de cocinas

(Siemens), asistencia a los aficionados en la elección del equipo fotográfico, etc. Emulan de alguna manera Sistemas Expertos más sofisticados como el utilizado por la casa de ordenadores DEC para configurar los ordenadores que vende (OPS4).

— Asistencia en toma de decisiones en sectores servicios: Consideramos que la Banca dentro del departamento de concesión de créditos y descuentos, así como el Seguro dentro del área de las cotizaciones a riesgos, pueden ser sectores ideales para este tipo de sistemas.

— Asesoramiento: Los gabinetes fiscales en las condiciones actuales pueden encontrar no sólo un instrumento muy útil para asesorar a los individuos, sino que la asesoría a nivel de sociedades es un terreno que demanda asistencia estructurada y organizada en lugar de la continua recurrencia a expertos que realicen interpretaciones particulares de la legislación en vigor.

— Formación: Sectores como la gerencia, la medicina, la abogacía, etc., se pueden beneficiar de la elaboración de Sistemas Expertos de forma de ir erradicando la necesidad, en unos casos, de un grupo de personas con conocimientos en diversas áreas, o en otros, de un tipo de entrenamiento muy largo para alcanzar prestaciones aceptables.

— Defensa: La capacidad no sólo de simular situaciones sino de incorporar nuevas experiencias, los convierte en un instrumento de análisis y entrenamiento importante. Por otro lado, tal y como hemos indicado en el primer punto, es importante considerar no sólo la formación sino la asistencia en el diseño de dispositivos lógicos que pueden ser analizados previamente por parte de uno de estos sistemas (por ejemplo formación de convoyes, etc.).

La idea de sugerir algunas áreas de aplicación como hemos hecho anteriormente, lo único que pretende es orientar y no suplir. Estimamos que la



visión del empresario reside en imaginar, examinado el entorno y las posibilidades de este medio, cuáles son aquellos usos que pueden deparar mejores resultados. En las áreas mencionadas existen ya Sistemas Expertos trabajando en forma más o menos experimental desde hace tiempo, pero hay otras oportunidades que están muy ligadas al avance de la tecnología donde hay espacio para innovaciones muy interesantes.

En efecto, una de las áreas donde el desarrollo de estos sistemas se espera que sea más espectacular, si cuentan con ordenadores suficientemente rápidos, es en el terreno de la visión y en particular acoplados a sistemas que reconozcan formas u objetos. Si a partir de aquí dejamos volar la imaginación aparecen posibilidades enormes. A modo de ejemplo piénsese que el resto de tareas administrativas que quedan por automatizar en el Sector de Servicios son las relativas al movimiento de grandes cantidades de papel donde el ver y reconocer juega un papel definitivo. Imaginemos un Sistema Experto que pudiendo "ver", de acuerdo con la información extraída y con ciertas reglas muy sencillas decidir en consecuencia, a una velocidad superior a la de un operador humano. Hoy en día un volumen importante y costoso de trabajo por automatizar se corresponde con este esquema: ver-reconocer-decidir. A nivel industrial ya está disponible un Sistema Experto asociado a un robot con visión que hace soldaduras de tal forma que le permite seguir las juntas a soldar con gran precisión.

Finalmente, otra área pendiente de un gran avance es la relativa al reconocimiento de la voz o mejor del lenguaje. El estadio es todavía muy primitivo pero ya hay dispositivos que reconocen órdenes. Por ejemplo ya se vende una "hoja electrónica" que admite órdenes verbales y es dentro de este discurso mucho más limitado en el que cabe esperar avances.

Por ejemplo, desde hace unos pocos años un Banco inglés contaba con un sistema adaptado a los operadores de compra-venta de moneda extranjera de tal forma que cerrado el trato telefónicamente el operador sólo citaba a un computador algunos datos como el cliente, importe, interés, etc., y unos segundos más tarde se le llevaba el contrato a la firma con objeto de materializar la operación. Hoy día la misma tarea puede llevarse a cabo con un ordenador de precio y tamaño mucho más reducido.

No quiero ni puedo terminar sin hacer una vez más mención de los términos en los cuales se debe plantear una incursión comercial en este terreno y para ello debo de recurrir a un ejemplo ilustrativo. Cuatro ingenieros especialistas de diferentes campos de la electrónica de una misma compañía, se reunieron un día de 1976. Su objetivo era simple: desarrollar un sistema de asistencia en el aprendizaje para niños. Había un problema: las técnicas existentes en ese momento necesitaban muchas operaciones aritméticas y grandes cantidades de memoria de computador. El producto no sería portátil y costaría miles de dólares. Unos 18 meses después el producto salió al mercado después de haber sido inicialmente financiado con 25.000 \$ (1.750.000 ptas.) y tres meses de plazo para demostrar su viabilidad técnica. Se llamaba Speak & Spell de Texas Instruments.

## AGRADECIMIENTO

El CDTI agradece el esfuerzo que ha supuesto para los autores la síntesis de ideas tan complejas en tan pocas páginas.

Asimismo, se agradece también el esfuerzo de las personas que han tenido la paciencia de corregir estos textos; en especial al Sr. D. Fernando Pena Möller, el cual ha sacrificado algunos fines de semana para lograr este cuaderno.



**CUADERNOS CDTI  
PUBLICADOS**



**La Innovación Industrial y las relaciones Industria-Universidad**



**La Innovación Industrial y su Tratamiento Fiscal**



**La conversión fotovoltaica de la energía solar**



**La ingeniería genética en la biotecnología**



**Innovación industrial y sistema educativo**



**¿Qué es la innovación tecnológica?**



**La telemática**



**Innovación Industrial y Empleo**



**Robótica Industrial**



**La financiación de la innovación industrial.**



**Pequeñas centrales hidráulicas**



**El microprocesador en la industria**



**Aprovechamiento energético de la biomasa**



**Innovación y Diseño Industrial**



**La investigación industrial farmacéutica.  
La política de patentes**



**El láser y sus aplicaciones industriales**



CDTI

Centro para el  
Desarrollo Tecnológico  
Industrial

Ministerio  
de Industria y Energía

Edificio Cuzco IV  
Pº Castellana, 141 - 3º  
28046 - Madrid  
España

Apto. de Correos: 29136  
Telf. 450 40 64\*  
Télex: 23121 CDTI E