

SISTEMAS BASADOS EN CONOCIMIENTO: UNA ALTERNATIVA POCO EXPLORADA EN PSICOLOGÍA

KNOWLEDGE-BASED SYSTEMS: AN UNEXPLORED ALTERNATIVE IN PSYCHOLOGY

Óscar García-Leal

CENTRO DE ESTUDIOS E INVESTIGACIONES EN
COMPORTAMIENTO, UNIVERSIDAD DE GUADALAJARA,
GUADALAJARA, JAL., MÉXICO

La contribución de los autores al escrito fue equiparable. Dirigir la correspondencia al primer autor: Óscar García Leal, oscargl@cencar.udg.mx, Calle Francisco de Quevedo #180, Col. Arcos Vallarta, 44130, Guadalajara, Jalisco. México.

Francisco Pedroza-Cabrera

FACULTAD DE PSICOLOGÍA, UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES, AGUASCALIENTES, MÉXICO

RECIBIDO: 21 DE OCTUBRE, 2010

ACEPTADO: 27 DE ABRIL, 2011

Resumen

Desde la aparición de los primeros Sistemas Expertos o Sistemas Basados en Conocimiento, aproximadamente en la década de 1960, el uso de estas herramientas se ha extendido de manera extraordinaria. Más allá de las aplicaciones que estas herramientas, derivadas de la Inteligencia Artificial presentan, el análisis del conocimiento disponible acerca del tema sobre las que las mismas son expertas resulta altamente interesante. Su utilidad para la formación de nuevos profesionales es también uno de los elementos que no deben ser desdeñados. El objetivo de este artículo es, en primera instancia, caracterizar lo que son los Sistemas Expertos, para posteriormente analizar las bondades que presentan en tanto formalismos para la estructuración del conocimiento disponible. Por último, se presentan algunas de sus potenciales aportaciones al campo de la Psicología y al diagnóstico clínico en general.

Palabras clave: Sistema experto, sistema basado en conocimiento, psicología.

Abstract

From the earliest appearances of Expert Systems, also called Knowledge-Based Systems, about 1960's, the use of these tools has expanded dramatically. Beyond the applied utility of these tools, the analysis of knowledge about the topic on which they are experts is of great theoretical interest. Also the practical consequences of training new professionals should be considered, too. The aim of this paper is to characterize what is an expert system, to analyze their role in structuring available knowledge about a particular domain, and finally a discussion about some of the contributions of expert systems to the field of Psychology and to clinical diagnosis in general is presented.

Key words: Expert system, knowledge-based system, psychology.

Introducción

Una primera caracterización global de la noción de Sistema Experto afirmarí­a que son herramientas informáticas cuyo propósito fundamental es reproducir la actuación de un experto humano en un dominio de conocimiento altamente especializado (Adarraga & Zaccagnini, 1988; Durkin, 1990). El Comité de Sistemas Expertos de la British Computer Society los definió de la siguiente manera:

“Se considera que un sistema experto es la incorporación en una computadora de un componente basado en el conocimiento que se obtiene a partir de la habilidad de un experto, de forma tal que el sistema pueda dar consejos inteligentes o tomar decisiones inteligentes... Una característica adicional deseable, y que para muchos es fundamental, es que el sistema sea capaz, bajo demanda, de justificar su propia línea de razonamiento de una forma inmediatamente inteligible para el que lo usa” (Naylor, 1983).

La definición anterior arroja dos puntos que a nuestro juicio merecen ser destacados. En primer lugar, un Sistema Experto tiene como objetivo proporcionar “consejos inteligentes o tomar decisiones inteligentes”. En este ámbito, se considera que un consejo o decisión inteligente es aquella que sería proporcionada por un experto en el dominio de conocimiento del Sistema Experto tras haber analizado un problema particular. En segundo lugar, es una característica deseable, si no fundamental, que el propio sistema sea capaz de explicitar su propia línea de razonamiento. Es en este sentido, que el uso de Sistemas Expertos puede resultar útil para los profesionales de la Psicología, en la medida en que una característica básica de estas herramientas

tecnológicas es reproducir el modo en el que los expertos en un conocimiento particular consideran la información disponible en su proceso de toma de decisiones. En este punto se abundará posteriormente.

Estos intereses remarcan la importancia que este tipo de herramientas han tenido para la Ciencia Cognitiva (Gardner, 1988) y, específicamente, para dos de sus disciplinas fundamentales, ambas interesadas en el estudio del comportamiento inteligente: la Inteligencia Artificial, cuyo propósito último es la construcción de máquinas cuyo comportamiento sea considerado inteligente, al margen de los algoritmos o heurísticos particulares que posibilitan tal comportamiento; y la Psicología Cognitiva, que a través de las técnicas de simulación psicológica ha tenido como uno de sus objetivos la construcción de programas de computadora que reproduzcan los componentes postulados por una teoría o modelo psicológico de una función cognitiva particular.

Típicamente, los Sistemas Expertos se han destinado a la solución de problemas complejos o “mal estructurados” (Jonassen 1997; Ormerod, 2005; Simon, 1973; Voss & Post, 1988), a los que habitualmente se enfrentan expertos que poseen un profundo conocimiento de los mismos (Pantic & Rothkrantz, 2000; Scherer, 1993). Se trata de problemas para los que, por sus características, no existen procedimientos de solución claramente definidos y, en este sentido, no resultan reducibles a procedimientos algorítmicos, esto es, a secuencias de acción previamente conocidas y consensuadas por la comunidad de expertos en el ámbito de intervención del sistema. Los Sistemas Expertos se han aplicado a la resolución de problemas que exigen la toma de decisiones en condiciones de información insuficiente e imprecisa, de tal modo que los datos que sustentarían las decisiones ciertas no siempre son conocidos y, si lo son, presentan

distintos grados de certeza. Ejemplos típicos de tareas de estas características son: la detección de averías en mecanismos complejos, la predicción meteorológica, la determinación de la estructura de las moléculas, el análisis de datos espectrográficos, la evaluación psicológica, el diseño de programas de intervención en salud o incluso la investigación en sus propios dominios de intervención (i.e., investigación en el ámbito del razonamiento diagnóstico). Véase Liao (2005) para una revisión de algunos de las principales áreas de aplicación.

Pero sin duda, el ámbito de intervención en el que los Sistemas Expertos se han mostrado más fecundos es el diagnóstico. De manera concreta, se han aplicado habitualmente al diagnóstico médico o psicológico. Algunos de los sistemas más sobresalientes en el ámbito del diagnóstico médico son INTERNIST-I (Myers, 1990), modelo de razonamiento diagnóstico desarrollado durante la década de 1970 en el área de la medicina interna aplicado a la formación de profesionales, que derivó, a mediados de 1980, en el sistema QMR (*Quick Medical Reference*) o MYCIN (Buchanan y Shortliffe, 1984), orientado a la evaluación, diagnóstico y tratamiento de enfermedades infecciosas. Algunos sistemas desarrollados para el diagnóstico psicológico son DAI (Adarraga, 1991), sistema experto especializado en el diagnóstico del autismo infantil, o SYSDO, especializado en el ámbito de la Terapia Ocupacional. Es por esto que analizaremos principalmente las aportaciones que este tipo de herramientas han realizado al ámbito del diagnóstico, remarcando su utilidad para la propia investigación en sus dominios de intervención, no solamente desde un punto de vista estrictamente pragmático, como herramientas de apoyo al diagnóstico, sino fundamentalmente desde una perspectiva teórica, como estrategias útiles para la investigación en el ámbito del razonamiento diagnóstico e incluso para la docencia y formación de profesionales.

El diagnóstico clínico: noción general

La literatura existente relativa a la estrategia mediante la cual los clínicos realizan las tareas diagnósticas señala la existencia de un ciclo en el cual los expertos en el diagnóstico proceden generando hipótesis y sometiendo éstas a contrastación (Barrows & Bennett, 1972; Simon, 1975; Eddy & Clanton, 1982; Elstein, Shulman & Sprafka, 1978; Feltovich, Johnson, Moller & Swanson, 1984; Groen & Patel, 1985; Johnson,

Duran, Hassebrock, Moller, Prietule, Feltovich & Swanson, 1981; Pople, 1977), aplicando por tanto una estrategia hipotético-deductiva.

Este modo de proceder permite simplificar el problema abordado, de manera que un problema abierto y mal definido, como es siempre la necesidad de elaborar un diagnóstico, se transforma en un conjunto de problemas mucho mejor definidos que permiten abordar el proceso de búsqueda de información y posterior contrastación de las hipótesis diagnósticas propuestas. Típicamente, el proceso de contrastación se realiza “hacia atrás, bien buscando manifestaciones que cabría esperar encontrar si la hipótesis fuese correcta, bien tratando de hallar indicadores incompatibles que la falseen” (Adarraga, 1991).

Elstein y Bordage (1979) destacaron la existencia de un ciclo de cuatro etapas que caracteriza el proceso de diagnóstico. Durante una primera fase, de obtención de indicios, se aplican diversos métodos que permiten recabar datos relevantes del paciente (análisis del historial clínico, administración de diferentes pruebas de evaluación, etc.). Una vez obtenida esta información inicial, típicamente se inicia un proceso de generación de hipótesis, relativas a la posible existencia de un número más o menos pequeño de diagnósticos alternativos, que podrían dar cuenta de la sintomatología previamente observada. El paso siguiente, la interpretación de indicios, tiene como objetivo principal analizar los datos recopilados en la primera fase en términos de su significación de acuerdo con las distintas hipótesis bajo consideración. Finalmente, se evalúan las hipótesis propuestas en un inicio, tratando de constatar la veracidad de alguna de ellas. En caso de que tal constatación no sea posible, normalmente se procede a la “reestructuración del problema”, reiniciándose el ciclo descrito y buscando nuevas hipótesis.

El proceso de categorización que supone la tarea diagnóstica resulta enormemente complejo. No es éste el marco adecuado para efectuar un análisis minucioso del proceso diagnóstico, pero sí resulta imprescindible realizar una breve mención de algunos de los aspectos que hacen de esta actividad una de las manifestaciones del comportamiento humano más difíciles de simular en un ordenador.

Por un lado, en el juicio diagnóstico se halla implicada, necesariamente, una enorme cantidad de información. Dicha información ha de ser elaborada

por el experto, de manera tal que el resultado de este proceso sea la generación inicial de un conjunto, más o menos amplio, de hipótesis alternativas, las cuales deberán ser sometidas a contrastación. Sin embargo, el experto no siempre dispone de la totalidad de la información que idealmente requiere para tomar una decisión plenamente justificada. Algo esencial, por tanto, será abordar el proceso de obtención de la información necesaria. Determinar qué información se requiere recabar dependerá, de manera importante, de los conocimientos relacionados con el problema abordado que posea el experto, así como de su modelo de conocimiento, entendido como el marco teórico asumido. Además, los expertos no siempre tienen absoluta seguridad o certeza acerca de la veracidad de la información presentada u obtenida, incluso sus decisiones se pueden ver afectadas por aspectos ajenos al propio proceso diagnóstico (habitualmente la necesidad inmediata de intervención clínica impide efectuar un proceso completo de evaluación previo a la actuación de los profesionales). A pesar de ello, en todos estos casos los expertos en diagnóstico efectúan un juicio clínico, y en un importante porcentaje de los casos los juicios emitidos son correctos.

Estos aspectos mencionados quedaron claramente puestos de manifiesto en la elaboración de MYCIN, uno de los primeros sistemas expertos en diagnóstico médico (Buchanan & Shortliffe, 1984). Las investigaciones realizadas al respecto pusieron de manifiesto cómo habitualmente los expertos tomaban sus decisiones diagnósticas y terapéuticas basándose en datos sumamente escasos y a todas luces insuficientes; además, en ocasiones se registraba una importante carencia de información que, sin embargo, no impedía que tomaran decisiones correctas en gran parte de los casos. Los expertos fundamentaban sus decisiones en “procesos de inferencia inexactos”, aplicaban conocimientos inseguros a premisas que les garantizaban una relativa seguridad. Parecían, de este modo, emplear mecanismos de razonamiento mal definidos, que aplicaban a pesar de contar con información incierta o incluso la carencia de una parte importante de información relativa a la interrelación existente entre el conjunto de variables consideradas, al objeto de tomar las decisiones pertinentes (Buchanan & Shortliffe, 1984).

En contextos de estas características, los expertos humanos, y por tanto, los Sistemas Expertos,

han de ser capaces de enfrentarse a situaciones en las que se requiere tomar una decisión a partir de la consideración de información conocida con diferentes grados de certeza, o incluso en situaciones de carencia de información relevante para la solución de la tarea abordada. Este hecho añade una complejidad considerable al desarrollo de sistemas computacionales de estas características. Por un lado, es completamente imposible determinar en el momento de su desarrollo de qué información se va a disponer y qué información estará ausente en cada caso (la ausencia de información será variable en los diferentes casos abordados por el sistema); por otro lado, la información disponible será conocida con un grado variable de certeza, es decir, la seguridad que el usuario tenga acerca de la verdad de un hecho relevante para la solución del problema abordado no siempre será la misma, y no siempre se podrá determinar la completa verdad del dato proporcionado al sistema. En todos estos casos (i.e., ausencia variable de información, desconocimiento de datos relevantes para la decisión buscada o conocimiento de los mismos con una seguridad relativa, etc.), el sistema deberá ser capaz de tomar una decisión o, en su defecto, de informar de la escasez de información que le permita llegar a una decisión mínimamente plausible y válida para el usuario.

Aunado a lo anterior, una característica deseable de un Sistema Experto es que pueda enfrentarse a demandas variables (Adarraga & Zaccagnini, 1988). Por un lado la necesidad de “comportarse” ante situaciones como las descritas genera, por sí misma, un gran número de problemas computacionalmente diferentes (i.e., el proceso de evaluación necesario para alcanzar el diagnóstico de un determinado caso variará necesariamente en virtud de los datos clínicos conocidos). Desde un punto de vista formal, un mismo caso clínico, procesado con distintas cantidades de información y distintos grados de certeza, no es el “mismo problema”, por más que el objetivo sea siempre el diagnóstico y el tratamiento (Adarraga & Zaccagnini, 1988). Por otra parte, la demanda realizada al sistema no necesariamente es equivalente en todos los casos. El sistema debe permitir al usuario matizar el tipo de información que desea que el sistema le proporcione. De este modo, un Sistema Experto ha de permitir, si se desea, la determinación de un único diagnóstico altamente probable, con lo que el sistema habrá de buscar información que le

permita alcanzar la decisión buscada con un alto grado de seguridad, o, si por el contrario, es preferible que el resultado de la consulta efectuada sea un amplio conjunto de diagnósticos convincentes en virtud de la información proporcionada, ordenados en un índice que ponga de manifiesto su convincentes, considerando la información obtenida a lo largo de la sesión.

Sistemas Expertos:

Caracterización estructural y funcional

En la definición inicial proporcionada de Sistema Experto nada se dice acerca de la estructura que estas herramientas deben presentar para alcanzar su objetivo. A saber, dar consejos inteligentes o tomar decisiones inteligentes en su ámbito concreto de trabajo. Esto es relativamente irrelevante a efectos de implementar un sistema de estas características. Sin embargo, sí se menciona la necesidad de que el sistema presente un importante corpus de conocimiento. De hecho, lo que hace que un sistema sea experto en un área de conocimiento concreto es el conocimiento que en él se halle implementado, junto con un conjunto de procedimientos (heurísticos en la mayoría de los casos) de manipulación y gestión del conocimiento instanciado en el sistema. Es por ello que a sistemas de estas características se les denomina en ocasiones Sistemas Basados en Conocimiento (SBC), concepto que adoptaremos a partir de este momento.

El objetivo fundamental de un SBC no es emular el modo de proceder de los expertos, esto es, tomar decisiones equivalentes independientemente de cómo se alcancen tales decisiones, sino simularlo, de manera que el contenido de un SBC sea un modelo teórico, tanto de los modos en los que los expertos en un ámbito de tarea concreto representan el problema, como de las estrategias concretas que siguen para resolverlo, así como del conocimiento implicado en dicha resolución.

Hayes-Roth (1984) identificó las funciones mínimas que un SBC debe satisfacer, señalando las siguientes: a) resolver problemas muy difíciles, tan bien o mejor que un experto humano; b) razonar heurísticamente, utilizando reglas que los expertos humanos consideran eficaces; c) interactuar eficazmente y en lenguaje natural con las personas; d) manipular descripciones simbólicas y razonar sobre ellas; e) funcionar con datos erróneos y reglas imprecisas; f) contemplar

simultáneamente múltiples hipótesis alternativas; g) explicar por qué plantean sus preguntas; y h) justificar sus conclusiones.

Desde el punto de vista de su aplicación práctica, Waterman (1985) señaló las siguientes propiedades que un SBC debe satisfacer: a) debe poseer un elevado grado de eficacia en lo que hace referencia a la resolución efectiva de los problemas propios de su dominio de intervención –se considera que dicha eficacia ha de ser equivalente a la mostrada por los expertos humanos en la materia–; b) debe ser flexible, en el sentido de permitir la incorporación inmediata de nuevos conocimientos adquiridos por la comunidad científica; c) debe servir como modelo de resolución de problemas de su ámbito de conocimiento, de tal modo que no proporcione únicamente soluciones adecuadas y válidas a problemas conocidos, sino que sea capaz de abordar con éxito la resolución de nuevos problemas a los cuales nunca se haya enfrentado, y, finalmente, d) debe permitir el entrenamiento de personal especializado.

En los primeros SBC desarrollados aparecía mezclado, de manera más o menos desordenada, el conocimiento concreto necesario para abordar los problemas objeto de representación con los procedimientos que, actuando sobre esos conocimientos, permitían resolverlos. Ello tenía consecuencias desagradables, que dificultaban la obtención de los objetivos mencionados anteriormente. Entre otras: a) la actualización de la base de conocimientos del sistema, con objetivos de desarrollo, habitualmente implicaba la necesidad de modificar gran parte de la estructura del sistema; b) las explicaciones que el sistema podía ofrecer acerca del razonamiento seguido y las conclusiones tomadas, habitualmente carecían de la transparencia deseable por el usuario final del producto; c) los heurísticos desarrollados para la manipulación del conocimiento, potenciales simulaciones de funciones de respuesta (Newell, 1990) del sistema cognitivo humano, eran difícilmente susceptibles de ser sometidos a validación. La obtención de un sistema con las especificaciones funcionales que permitiera satisfacer los requerimientos enunciados anteriormente implicó hacer independientes los conocimientos y los procedimientos de manipulación (Fernández, 1985).

Con este objetivo se desarrolló una nueva caracterización estructural de los sistemas, que se mantiene aún en la gran mayoría de los proyectos actuales

(Dhaliwal & Benbasat, 1996; Gregor & Benbasat, 1999; Liao, 2004). La estructura propuesta permite que el conocimiento específico que los SBC poseen sobre su ámbito de actuación se encuentre representado explícitamente, y visiblemente separado del resto del sistema. El conocimiento es representado típicamente mediante el uso de reglas de la forma SI-ENTONCES. Este conocimiento es independiente de los procesos particulares de razonamiento implementados en el sistema, que permiten operar con estas reglas a efectos de alcanzar una solución o conjunto de soluciones posibles al problema abordado. No se discutirá ampliamente sobre los diversos formatos de representación del conocimiento que es posible adoptar (la bibliografía existente al respecto es muy amplia –véase, por ejemplo, para una introducción a estos algoritmos, Adarraga & Zaccagnini, 1994; Cuenca, 1985; Davis & King, 1984; Liao, 2004; Pazos Sierra, 1987; Waterman, 1985), pero sí se realizará una básica caracterización estructural de un SBC prototípico, así como se describirá brevemente su funcionamiento.

La mayor parte de los SBC están constituidos fundamentalmente por tres componentes (Davis, Buchanan & Shortliffe, 1985; Fernández, 1985; Waterman, 1985; Adarraga & Zaccagnini, 1994):

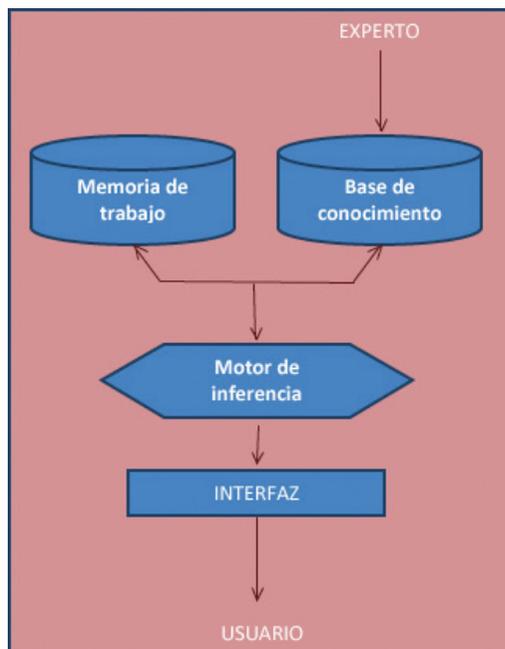


Figura 1. Caracterización estructural típica de un sistema basado en conocimiento.

Memoria de trabajo

Almacena información o conocimiento declarativo sobre el problema particular que en un momento dado se intenta resolver y sobre el estado de conocimiento del sistema a lo largo de la sesión en curso. Es, por tanto, una estructura de datos que almacena la información correspondiente al caso en curso, tanto los datos que inicialmente le han sido proporcionados al sistema como las conclusiones, provisionales o definitivas, que el sistema haya ido estableciendo a lo largo de la sesión.

Base de conocimiento

Está constituida por el conocimiento específico y procedimental acerca de la clase de problemas en los que el sistema es experto. Almacena una representación del conocimiento que el sistema posee sobre los conceptos y relaciones propios de la tarea. Característicamente (i.e., en los sistemas de producción) consta de un conjunto de objetos o parámetros, junto con una serie de reglas que especifican las relaciones existentes entre dichos parámetros. Citando a Adarraga y Zaccagnini (1994), *“su característica más destacable es que la organización interna de sus contenidos resulta significativa desde el punto de vista de los conocimientos representados. Dicho de otro modo, la base de conocimientos puede interpretarse como un “modelo” del dominio de tareas a que el sistema está dirigido”*.

Motor de inferencia

Contiene información relativa al funcionamiento del sistema en sus funciones deductivas. Así, analiza el conocimiento almacenado tanto en la base de conocimiento como en la memoria de trabajo del sistema, con el objetivo de obtener conclusiones relevantes para la solución del problema o determinar el curso de acción a seguir. Si se prefiere, posee el conocimiento necesario para que el sistema pueda razonar con el conocimiento experto que posee y la información específica del estado del problema en curso, con el fin de alcanzar conclusiones semejantes a las que tomaría un experto humano.

Como se ha comentado, la principal aportación de esta arquitectura es que permite hacer independiente la lógica del procedimiento de interpretación para resolver problemas, del conocimiento específico del dominio de tarea abordado. Este hecho, que en principio parecería estar fundamentado en razones

meramente pragmáticas, posee, sin embargo, extensas implicaciones teóricas, de gran utilidad para la investigación en los dominios de especialización de los sistemas. Su utilidad reside en la posibilidad de diferenciar el conocimiento específico de los sistemas (base de conocimiento) de los procedimientos o funciones de respuesta que permiten su utilización (motor de inferencia) (Buchanan & Smith, 1988). Ambos aspectos se constituyen, por tanto, como componentes potencialmente analizables de manera independiente, a pesar de que puedan existir relaciones o dependencias entre ellos.

Uno de los métodos procedimentales más característicos de representación del conocimiento en los SBC son las denominadas reglas de producción (Liao, 2005). Este formato se ha empleado en algunos de los principales sistemas diagnósticos citados anteriormente, MYCIN o DAL, y es uno de los formalismos que se ha mostrado más exitoso en la construcción de SBC. Típicamente, una regla de producción tiene una estructura ANTECEDENTE-CONSECUENTE, tal que el antecedente contiene las cláusulas que deben cumplirse para que la regla pueda evaluarse o ejecutarse, y el consecuente indica las conclusiones que se reducen de las premisas o las acciones que el sistema debe realizar cuando ejecuta la regla.

La unidad básica de representación del conocimiento en una regla de producción es una terna asociativa que presenta la siguiente estructura: OBJETO (contexto)-ATRIBUTO-VALOR. Cada una de ellas representa el conocimiento de que una determinada característica (*atributo*) de una cierta entidad conceptual (*objeto*) posee un valor concreto, de manera tal que para cada objeto es posible definir un conjunto de parámetros descriptivos de atributos de la entidad representada, mediante la asignación de un valor o un conjunto de valores con distintos grados de certeza (Cuena, 1985). Esta estructura de datos ternaria ha mostrado un poder expresivo suficiente para representar cualquier hecho relevante para la resolución de un problema diagnóstico (Adarraga & Zaccagnini, 1994), entre otras áreas de aplicación.

Como se ha mencionado, los elementos o módulos de representación que componen la base de conocimiento están formados por pares antecedente-consecuente. El antecedente está constituido por una combinación booleana de naturaleza conjuntiva de condiciones relevantes que toman como argumentos proposiciones sobre el área de problema objeto

de representación. Cada una de esas condiciones es, una función predicativa o una disyunción de éstas. El consecuente o acción de cada una de ellas consta de una única expresión resultante de la aplicación de una función predicativa que, habitualmente, supone la incorporación de una proposición inferida en memoria de trabajo, acompañada de una medida de fuerza de creencia asociada, que es interpretada como la verosimilitud que el sistema le atribuye (Adams, 1984; Adarraga & Zaccagnini, 1994). Nótese que este formalismo de representación es independiente del contenido semántico en él representado, en virtud de su naturaleza formal. El conocimiento de este modo representado es gestionado por un conjunto de funciones de respuesta o procedimientos que permiten transformar el conocimiento implicado en las reglas utilizadas o bien inferir nuevos conocimientos, con el objetivo de alcanzar la meta o metas activas en el sistema (Pylyshyn, 1988).

La mayor parte de los SBC desarrollan un patrón de inferencia o modelo de razonamiento orientado por objetivos, que supone un modo específico de equiparación y catalogación de las reglas constituyentes de la base de conocimiento. Partiendo de un objetivo inicial –en general, proporcionar un diagnóstico o conjunto posible de diagnósticos–, los SBC actúan de dos maneras, ya sea solicitando al usuario información necesaria para sostener el proceso de inferencia, u obteniendo la misma del análisis de la base de conocimiento en virtud del contenido proposicional de las reglas constitutivas o de la memoria de trabajo. La información requerida para avanzar en el proceso diagnóstico estará determinada por cada uno de los antecedentes constituyentes de la regla de producción que sea activada en la memoria de trabajo. En caso de que dicha información no esté disponible en la propia memoria de trabajo, la obtención de esa información se constituirá en una nueva meta a ser alcanzada por el sistema mediante el procedimiento descrito.

Buchanan y Shortliffe (1984), señalaron una serie de ventajas derivadas de la utilización conjunta de un modelo capaz de simular el razonamiento mostrado por los expertos en tareas diagnósticas, junto con el algoritmo de representación del conocimiento expuesto: a) permite emplear conocimiento global del dominio de tarea al objeto de tratar el caso específico objeto de estudio; b) permite incorporar conocimiento experto existente al objeto de abordar

el diagnóstico y tratamiento de problemas específicos, sobre los cuales no existen o no se encuentran fácilmente disponibles datos estadísticos; c) permite abordar la modificación y/o actualización de la base de conocimiento del sistema de manera sencilla; d) facilita la búsqueda de potenciales contradicciones e inconsistencias de la información representada en la base de conocimiento; e) facilita la modelización de un sistema capaz de explicitar el proceso de razonamiento seguido por el sistema; y f) potencia la posibilidad de emplear el sistema en labores de instrucción.

Conclusiones

Si bien los SBC suponen una importante aportación para el estudio de los procesos diagnósticos, la elaboración de sistemas de estas características no resulta posible para cualquier dominio de conocimiento, ni siquiera para cualquier ámbito relativo al diagnóstico. Se requiere que los dominios objeto de intervención presenten una serie de características (Cuenca, 1985). Por un lado, es necesario que los procedimientos de resolución empleados por los especialistas en la tarea incluyan un importante componente de razonamiento. Además, estos procedimientos deben ser suficientemente conocidos por la comunidad científica especializada, de tal modo que exista un mínimo grado de estructuración en el dominio del conocimiento. A estos dos criterios, relativos a las características propias del ámbito abordado, es necesario añadir dos requisitos más de naturaleza computacional. Por un lado, es necesario que el conocimiento experto sea susceptible de ser manipulado en elementos relativamente independientes, de manera que pueda introducirse en el sistema de manera incremental. Por otro, de naturaleza pragmática, la elaboración de un SBC ha de resultar más ventajosa que abordar el proceso de diseño de un sistema característico de la informática convencional. Si la tarea puede resolverse por métodos algorítmicos en un tiempo razonable, la creación de un SBC –siempre costosa– no aportaría ninguna ventaja apreciable sobre la aplicación de un programa convencional.

Al abordar el proceso de diseño de una herramienta de estas características se hace imprescindible realizar un importante trabajo de obtención y estructuración del conocimiento existente, potestad de los

expertos en el dominio de tarea abordado. Además, se requiere revisar las fuentes documentales existentes, a efectos de reunir el conocimiento necesario para construir un sistema competente. Finalmente, es necesario estructurar y elaborar el conocimiento de un modo tal que permita su implementación en la base de conocimiento del SBC y su explotación. Dicha estructuración y elaboración será el factor determinante de que el SBC resulte competente en su ámbito de aplicación. Los SBC requieren, por tanto, de un importante proceso de elaboración de conocimiento. Es necesario insistir en ello: para poder implementar un Sistema Experto es requisito indispensable que el conocimiento existente al respecto presente un mínimo grado de estructuración; pero aún más importante es el hecho de que el proceso de análisis del conocimiento experto permita obtener ese nivel de estructuración, dado que supone siempre un ejercicio de interpretación, evaluación y elaboración del conocimiento existente relativo al ámbito de trabajo objeto de estudio.

El grado de formalización requerido implica descomponer el dominio de conocimiento en unidades que tengan sentido en sí mismas. Cada una de estas unidades debe ser delimitada en términos de subproblemas perfectamente acotados y relativamente independientes entre sí, susceptibles de ser analizados y descompuestos de manera minuciosa. Este análisis debe incluir el análisis y caracterización de las relaciones existentes entre los diferentes subproblemas identificados.

Estructurar el conocimiento existente no solamente resulta útil en el sentido descrito. Para que un SBC se pueda desarrollar, el conocimiento ha de poder ser representado en unidades relativamente independientes –típicamente, reglas de producción– que permitan una aproximación incremental al problema. Junto a estos formalismos de representación es necesario proporcionar un conjunto de estrategias algorítmicas y/o heurísticas que permitan la realización de inferencias en última instancia equivalentes a las realizadas por los expertos.

Un SBC especializado en el diagnóstico clínico trata de simular el modo de proceder de los expertos en las tareas típicamente diagnósticas (Hoffman, 1987). Por tanto, dado que un SBC es, típicamente, la simulación de un proceso de toma de decisiones en un ámbito de conocimiento concreto, y puesto que ese proceso de toma de decisiones se sustenta

en la consideración de conocimiento necesariamente formalizado y estructurado en el sistema (de otro modo no estaremos en condiciones de instanciar en la Base de Conocimiento el propio conocimiento necesario que permita al sistema tomar decisiones útiles), un SBC puede ser concebido como una representación del conocimiento existente al respecto; es decir, como la "instanciación formal" de un corpus de conocimiento previamente existente en los propios expertos en el dominio de tarea abordado.

Las ventajas más inmediatas que nos ofrece poseer una representación física de un modelo de conocimiento existente acerca de un ámbito concreto son obvias. En primer lugar, nos obliga a hacer un análisis minucioso de dicho conocimiento, esto es, del grado de estructuración existente, de posibles incoherencias o redundancias existentes en los procesos de toma de decisiones observados, nos permite darnos cuenta de posibles lagunas de conocimiento sobre ámbitos concretos del dominio de tarea, etc. Una vez determinadas las posibles incoherencias existentes en el corpus de conocimiento, orienta la investigación básica necesaria para cubrir las lagunas existentes. Más allá de esto, permite determinar los procesos de decisiones implicados (Henrion, Breese & Horvitz, 1991) y caracterizar, por tanto, el curso de acción a seguir dada una situación de intervención concreta. En definitiva, posibilita tener una representación explícita del propio modelo teórico existente de los expertos en un área y, por tanto, permite abordar el proceso de formación de nuevos especialistas desde un conocimiento objetivado y a la luz de los resultados aportados por la investigación básica realizada acerca del modelo de conocimiento utilizado por los expertos.

Los SBC pueden ser entendidos como métodos de investigación básica del conocimiento en un dominio particular y, del mismo modo, como modelos de partida para investigaciones sobre el razonamiento humano.

Junto a la utilidad descrita, de naturaleza netamente teórica, no queremos dejar de señalar una importante utilidad práctica o metodológica, sobre la que es necesario insistir. Los SBC aparecen como herramientas de ayuda al diagnóstico (nunca como herramientas sustitutivas del especialista) y, en este sentido, permiten el entrenamiento de personal especializado, así como la instrucción de nuevos especialistas.

Una característica deseable de todo SBC es que sea posible explicitar el proceso de razonamiento seguido en su actuar diagnóstico. En sentido estricto, esto implica que una herramienta de estas características pueda ser interrogada acerca de los procesos que en un determinado momento está ejecutando, así como de la finalidad de los mismos.

Todo proceso de solución de problemas (i.e., tarea diagnóstica) implica el establecimiento de un conjunto de metas progresivas que se han de satisfacer para alcanzar el objetivo buscado. Un SBC debe permitir explicitar en todo momento cuáles son las metas que en un momento particular han sido establecidas a efectos de solucionar el problema enfrentado. Además, en todo momento, debe permitir al usuario ver su estado de conocimiento, es decir, presentarle un listado tanto de la información introducida al sistema por parte del usuario (típicamente, se realiza mediante sistemas sencillos a través del teclado o mediante cualquier otro dispositivo de entrada), como del conjunto de conocimiento adquirido resultado de la realización de inferencias a la luz de las reglas de producción instanciadas en el sistema. El formato en el que esta información se presente al usuario es irrelevante en lo que al diseño del sistema se refiere. En algunos SBC la explicitación de estos procesos es simplemente un listado de las reglas de producción que en un determinado momento el sistema está tratando de ejecutar, en tanto otros presentan de modo gráfico el proceso de razonamiento seguido por el sistema. Son estos últimos SBC los que presentan una mayor utilidad a efectos de abordar el proceso de formación de nuevos especialistas.

Otra característica deseable de todo SBC es que permita almacenar casos para su posterior análisis, o incluso modificación de los datos proporcionados al sistema. Este hecho se muestra de gran utilidad para la aplicación de los SBC a la instrucción (Schenk, Vitalari, Shannon & Davis, 1998), dado que permite enfrentar a sujetos en formación a situaciones clínicas completamente descritas, analizar el comportamiento que los sujetos expertos mostrarían ante un caso concreto bajo distintos supuestos sin necesidad de que estos se encuentren presentes, evaluar el proceso de aprendizaje a través de la contrastación de los diagnósticos efectuados por los sujetos con los elaborados por el sistema, etcétera.

En resumen, es posible señalar la existencia de al menos una doble aplicación de los SBC al ámbito diagnóstico:

En primer lugar, presentan una utilidad teórica. El proceso de estructuración, análisis e interpretación del conocimiento existente acerca del dominio de actuación del sistema objeto de diseño aporta información esencial para el propio estudio del área de conocimiento. Por ejemplo, señalando la necesidad de enfatizar la investigación en sub-problemas particulares detectados, o de reestructurar los conocimientos existentes. Asimismo, su elaboración obliga a formalizar el estado actual de conocimiento respecto a la tarea diagnóstica abordada y, en consecuencia, implica la formulación de un modelo teórico acerca de la estructura conceptual y relaciones que permiten adoptar decisiones inteligentes en ese ámbito.

Finalmente, los SBC presentan una utilidad práctica manifiesta. No solamente aportan información esencial de apoyo a la tarea diagnóstica aportada por los especialistas, sino que, además, permiten abordar el proceso de formación y/o instrucción de nuevos profesionales a la luz de modelos de conocimiento previamente delimitados y caracterizados de manera precisa.

Referencias

- Adams, J. B. (1984). Probabilistic reasoning and certainty factors. En: B.G. Buchanan & E.H. Shortliffe (Eds.), *Rule-based expert systems: The MYCIN experiments of the Stanford heuristic programming project*. Reading, MA: Addison-Wesley.
- Adarraga, P. (1991). "DAI": Sistema basado en conocimiento para diagnóstico de autismo. Una aproximación al razonamiento diagnóstico. Tesis doctoral no publicada. Universidad Autónoma de Madrid.
- Adarraga, P., & Zaccagnini, J. L. (1988). Sistemas expertos y psicología cognitiva: una visión general. *Estudios de Psicología*, 36, 97-127.
- Adarraga, P., & Zaccagnini, J. L. (1994). *Psicología e Inteligencia Artificial*. Madrid: Trotta.
- Barrows, H. S. & Bennett, K. (1972). Experimental Studies on the Diagnostic Problem Solving Skill of the Neurologist. *Archives of Neurology*, 26, 273-277.
- Buchanan, B. G., & Shortliffe, E. H. (1984). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley.
- Buchanan, B. G., & Smith, R. G. (1988). Fundamentals of expert systems. *Annual Review of Computer Science*, 3, 23-58, available via: <http://dx.doi.org/10.1146%2Fannurev.cs.03.060188.000323>
- Cuena, J. (1985). El razonamiento aproximado en sistemas expertos. En J. Cuena (Ed.) *Lógica informática* (pp. 495-545). Madrid: Alianza Editorial.
- Dhaliwal, J. S., & Benbasat, I. (1996). The use and effects of knowledge-based system explanations: Theoretical foundations and a framework for empirical evaluation. *Information Systems Research*, 7, 342-362, available via <http://dx.doi.org/10.1287%2Fisre.7.3.342>
- Davis, R., & King, J.J. (1984). The Origin of Rule-Based Systems in AI. En B. Buchanan & E.H. Shortliffe (Eds.), *Rule-based expert systems: The MYCIN Experiments of the Stanford heuristic programming project* (pp. 20-52). Reading, MA: Addison-Wesley.
- Davis, R., Buchanan, B., & Shortliffe, E. (1985). Production Rules as a Representation for a Knowledge-Based Consultation Program. En R.J. Brachman & H.J. Levesque (Eds.), *Readings in Knowledge Representation* (pp. 371-388). Los Altos: Morgan Kaufmann Publishers, Inc.
- Durkin, J. (1990). Introducing Students to Expert Systems. *Expert Systems: The International Journal of Knowledge Engineering*, 7 (2), 70-80, available via: <http://dx.doi.org/10.1111%2Fj.1468-0394.1990.tb00169.x>
- Eddy, D. M., & Clanton, C. H. (1982). The art of diagnosis: solving the clinipathological exercise. *The New England Journal of Medicine*, 306, 1263-1268, available via: <http://dx.doi.org/10.1056%2FNEJM198205273062104>
- Elstein, A. S., & Bordage, G. (1979). Psychology of Clinical Reasoning. En G. Stone & N. Alder (Eds.), *Health Psychology*. San Francisco: Jossey-Bass.
- Elstein, A. S., Shulman, L.S., & Sprafka, S.A. (1978). *Medical Problem Solving: An Analysis of Clinical Reasoning*. Cambridge, MA: Harvard University Press.
- Feltovich, P., Johnson, P. E., Moller, J. H., & Swanson, D. B. (1984). LCS: The Role and Development of Medical Knowledge in Diagnostic Expertise. En W.J. Clancey & E.H. Shortliffe (Eds.), *Readings in Medical Artificial Intelligence: The First Decade*. Reading, MA: Addison-Wesley.

- Fernández, G. (1985). Panorama de los sistemas expertos. En J. Cuenca, *Inteligencia Artificial: Sistemas Expertos* (pp. 23-53). Madrid: Alianza Editorial.
- Gardner, H. (1988). *La nueva ciencia de la mente: historia de la revolución cognitiva*. Barcelona: Paidós (traducción al español).
- Gregor, S., & Benbasat, I. (1999). Explanations from intelligent systems: Theoretical foundations and implications for practice. *MIS Quarterly*, 23, 497-530, available via: <http://dx.doi.org/10.2307/2F249487>
- Groen, G. J., & Patel, V.L. (1985). Medical Problem Solving and Cognitive Psychology: Some Questionable Assumptions. *Medical Education*, 19, 95-100, available via: <http://dx.doi.org/10.1111/2Fj.1365-2923.1985.tb01148.x>
- Hayes-Roth, F. (1984). The Machine as a Partner of the New Professional. *IEEE Spectrum*, 21, 28-31, available via: <http://dx.doi.org/10.1109/2FPMC.1984.1659242>
- Henrion, M., Breese, J. S., & Horvitz, E.J. (1991). Decision Analysis and Expert Systems. *AI Magazine*, 12, 64-91.
- Hoffman, R. R. (1987). The problem of extracting the knowledge of experts from the perspective of experimental psychology. *AI Magazine*, 8(2), 53-67.
- Johnson, R. E., Duran, A. S., Hassebrock, R., Moller, J., Prietule, M., Feltoovich, P. J., & Swanson, D. (1981). Expertise and error in diagnostic reasoning. *Cognitive Science*, 5, 235-283, available via: <http://dx.doi.org/10.1016/2FS0364-0213/2881/2980014-6>
- Jonassen, D.H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology: Research and Development*, 45, 65-95, available via: <http://dx.doi.org/10.1007/2FBF02299613>
- Liao, S.-H. (2004). Expert system methodologies and applications: A decade review from 1995-2004. *Expert Systems with Applications*, 1-11, available via: <http://dx.doi.org/10.1016/2Fj.eswa.2004.02.006>
- Liao, S.-H. (2005). Expert system, methodologies and applications –a decade review from 1995 to 2004. *Expert Systems with Applications*, 28, 93-103, available via: <http://dx.doi.org/10.1016/2Fj.eswa.2004.02.006>
- Myers, J.D. (1990). The background of INTERNIST-I & QMR. En: B.I. Blum & K. Duncan (Eds.), *A history of medical informatics*. Nueva York: ACM Press.
- Naylor, C. (1983). *Build your own Expert System*. Cheshire: Sigma Technical Press.
- Newell, A. (1990). Foundations of cognitive science. En A. Newell (Ed.), *Unified Theories of Cognition* (pp. 42-111). Harvard: Harvard University Press.
- Ormerod, T. C. (2005). Planning and ill-defined problems. En R. Morris & G. Ward (Eds.), *The cognitive Psychology of planning*. London: Psychology Press.
- Pantic, M., & Rothkrantz, L.J.M. (2000). Expert system for automatic analysis of facial expressions. *Image and Vision Computing*, 18, 881-905, available via: <http://dx.doi.org/10.1016/2FS0262-8856/2800/2900034-2>
- Pazos Sierra, J. (1987). Sistemas de inferencia dirigidos por patrones: Sistemas de producción. En J. Pazos Sierra (Ed.), *Inteligencia Artificial: Programación heurística* (pp. 172-260). Madrid: Paraninfo.
- Pople, H. E. (1977). The formation of composite hypothesis in diagnostic problem solving. An exercise in synthetic reasoning. *International Joint Conference on Artificial Intelligence*, 5, 1030-1037.
- Pylyshyn, Z. W. (1988). *Computation and cognition: Towards a foundation for cognitive science*. Boston, MA: The MIT Press.
- Schenk, K. D., Vitalari, N. P., & Shannon Davis, K. (1998). Differences between novice and expert systems analysts: what do we know and what do we do? *Journal of Management Information Systems*, 15, 9-50.
- Scherer, K. R. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition & Emotion*, 7, 325-355, available via: <http://dx.doi.org/10.1080/2F02699939308409192>
- Simon, H. A. (1973). The structure of ill structured problems. *Artificial Intelligence*, 4, 181-204, available via: <http://dx.doi.org/10.1016/2F0004-3702/2873/2990011-8>
- Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7, 268-288, available via: <http://dx.doi.org/10.1016/2F0010-0285/2875/2990012-2>
- Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. En: M.T.H. Chi, R. Glaser & M.J. Farr (Eds.), *The nature of expertise* (pp.

261-285). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Waterman, D. A. (1985). How are expert systems organized? En D.A. Waterman (Ed.), *A Guide to Expert Systems* (pp. 16-24). Reading, MA: Addison-Wesley.