



Sistemas de Tutorización Inteligente Basados en Redes Bayesianas

Jorge López Puga y Juan García García¹
Universidad de Almería

RESUMEN

La enseñanza programada ha sufrido grandes avances con el desarrollo de las ciencias de la computación. Cada vez son más los programas informáticos que nos permiten aprender cualquier materia o destreza. En este contexto surgieron los Sistemas de Tutorización Inteligente asociados a redes bayesianas. Las redes bayesianas son herramientas estadísticas surgidas en el contexto de la Inteligencia Artificial orientadas a la inferencia probabilística y en el ámbito de la tutorización electrónica se utilizan para modelar la incertidumbre asociada al estudiante y su nivel de conocimiento. En este trabajo presentamos la metodología de las redes bayesianas en relación a los Sistemas de Tutorización Inteligente, revisamos algunos ejemplos importantes de estos sistemas y proponemos ciertos aspectos que podrían mejorar el desarrollo de los sistemas de tutorización.

Palabras clave: Sistemas de Tutorización Inteligente, toma de decisiones, psicología del aprendizaje, enseñanza, redes bayesianas.

ABSTRACT

Programmed teaching has suffered major advances because of the computer science development. There are more and more software designed to learn any matter or skill. Intelligent Tutoring Systems arose in that context associated with bayesian networks. Bayesian nets are statistical tools developed in the field of Artificial Intelligence to manage probabilistic inference. In educational context, bayesian nets are used to model the uncertainty about student knowledge. We present the Bayes nets methodology related with Intelligent Tutoring Systems, we review several examples of bayesian tutoring systems and we propose how tutoring systems could be optimized.

Keywords: Intelligent Tutoring Systems, making decisions, psychology of learning, education, Bayes nets.

¹ Dirección de contacto:

Facultad de Humanidades y Ciencias de la Educación, Departamento de Ciencias Humanas y Sociales, Área de Metodología de las Ciencias del Comportamiento, Universidad de Almería, Ctra. Sacramento S/N, La Cañada de San Urbano, C.P. 04120, Almería, España.

E-mails: jpuga@ual.es y jgarciag@ual.es.



1.- Introducción

El desarrollo de las ciencias de la computación está generando un abaratamiento en la maquinaria y los programas informáticos. Este avance ha afectado al ámbito educativo produciendo una auténtica revolución en los procedimientos clásicos de enseñanza-aprendizaje (Urretavizcaya, 2001). Hoy en día existen diferentes sistemas que permiten la tutorización y el aprendizaje casi de cualquier materia académica. La plataforma WebCT, donde se puede gestionar todo un curso a través de Internet, es un claro ejemplo de ello. Sin embargo, este tipo de sistemas tienen algunas limitaciones si los comparamos con lo que haría una persona en una tutoría. El problema es que estos sistemas son una mera plataforma comunicativa que establece un proceso interactivo entre el profesorado y el alumnado. Así pues, aunque plataformas como WebCT permiten que el personal docente pueda poner información y materiales a disposición de sus estudiantes, el sistema no hace nada automáticamente para determinar qué información o materiales serían más apropiados para cada persona en función del nivel de conocimientos que cada uno tiene.

El área de la Inteligencia Artificial en el ámbito de los sistemas expertos ha desarrollado la idea de lo que se conoce como Sistemas de Tutorización Inteligente (STI). Según Xiang (2002), un STI es un tipo de *agente* inteligente caracterizado por su habilidad para tomar decisiones y la ejecución de acciones sin la intervención de ninguna persona que monitoriza el funcionamiento del sistema. Así pues, un STI tendría que, en primer lugar, evaluar el conocimiento que el aprendiz tiene sobre la materia. A continuación, ha de presentar información acorde a ese nivel (por ejemplo, definiciones, material gráfico o auditivo). Y posteriormente ha de evaluar si la persona ha asimilado el conocimiento nuevo que se pretendía que el usuario hubiese adquirido. En otras palabras, el sistema tiene que generar una representación del conocimiento que tiene el usuario y guiarlo a través de un proceso de enseñanza que consiga que adquiera algún conocimiento o competencia.

Los STI comenzaron a desarrollarse en los años setenta y experimentaron su máximo apogeo a mediados de los años ochenta coincidiendo con el auge de los sistemas expertos, aunque se popularizaron durante la década de 1990. Durante la década de 1970 se pusieron en marcha numerosas investigaciones encaminadas al desarrollo de programas informáticos orientados a la enseñanza bajo el lema lo que podríamos denominar como Instrucción Facilitada por Computadora (IFC, del inglés *Computer Aided Instruction* o CAI), Aprendizaje Asistido por Computadora (AAC, del inglés *Computer Assisted Learning* o CAL) o Entrenamiento Basado en Ordenador (EBC, del inglés *Computer Based Training* o CBT). Como señala Wu (1993), el inicio del desarrollo de los primeros STI estuvo marcado por el intento de simular el proceso físico de interacción que se establece entre la persona que tutoriza y la que aprende. En concreto, se pretendía simular el proceso comunicativo que se produce en el tutor y el aprendiz en entornos naturales.

Podríamos ubicar el origen de los sistemas informáticos inteligentes orientados a la enseñanza en el Instituto Tecnológico de Massachussets (MIT). En este sentido, SCHOLAR podría considerarse como el primer STI que, a su vez, estuvo basado una aplicación conocida como ELIZA diseñada para el estudio de la comunicación entre persona-máquina basándose en la Terapia Rogeriana. SCHOLAR se diseñó para enseñar geografía del Continente Americano y supuso un salto cualitativamente sustancial entre los sistemas clásicos *Orientados a Marcos Apriorísticos (ad hoc-frame-oriented systems* o AFO), donde el proceso



de enseñanza se basaba en series de preguntas y respuestas pre-diseñadas que hacían el sistema demasiado rígido y que constreñían la creatividad de la persona que aprende; y lo que se denominó como sistemas *Orientados a Estructuras de Información* (*information – structure-oriented systems* o ISO), donde el formato de representación de la información permitía al sistema anticipar las respuestas de la persona que aprende y adaptarse a sus necesidades (Carbonell, 1970).

El Sistema WHY es otro ejemplo de la evolución de los tutores inteligentes al auspicio del MIT (Stevens y Collins, 1977). WHY se diseñó para enseñar a interpretar los fenómenos lluviosos en términos causales. El sistema se concibió para hacer reflexionar a los usuarios del sistema en términos socráticos; esto es, desafiaba al educando con preguntas que minaban creencias erróneas reconduciéndolo a la construcción del conocimiento por medio de ejemplos *contrafactuales* y preguntando por causas generales e intermedias.

Esta idea mecanizada de enseñanza fue introducida alrededor de 1958 por el psicólogo conductista Burrhus Frederic Skinner en el concepto de *enseñanza programada*. La idea subyacente a la enseñanza programada estuvo condicionada por la filosofía conductista-asociacionista y su objetivo era dirigir el aprendizaje humano bajo condiciones controladas. Los investigadores educativos de la época trataron de implementar estas ideas en un contexto aplicado y se desarrollaron lo que se denominó *libros de texto programados*. Un libro programado presenta una materia tan estructurada que permite que el usuario vaya asimilando conceptos secuencial y autorreguladamente. Por ejemplo, Dixon (1964) diseñó un libro tras varios años de investigaciones orientado a enseñar los principios de la probabilidad en cursos universitarios de ingeniería. En una de sus secciones se presentaban preguntas (problemas) en formato de elección múltiple con cuatro alternativas de respuesta. Cada respuesta remitía a otra página y sección donde daba retroalimentación al usuario sobre la idoneidad de su respuesta. Otro formato usado en los libros programados era dividir cada página en dos secciones, una (generalmente la de la derecha) para plantear conceptos o preguntas y la otra para mostrar las respuestas (Maning y Rosenstock, 1961). El usuario tenía que leer las cuestiones mientras mantenía ocultas las respuestas. A continuación, tenía que comparar sus respuestas con las que daba el libro.

Durante la época de máximo esplendor del conductismo, las ciencias de la computación estaban en una etapa primigenia de su desarrollo por lo que el intento de implementar la idea de la enseñanza programada en un dispositivo automático pasó desapercibida. Lo que vino a denominarse como la *revolución cognitiva* en Psicología relegó a la filosofía conductista, junto a la idea de la formación programada, a un segundo plano como objeto de investigación científica (Hothersall, 1997). Hoy día, las computadoras son extremadamente más potentes que entonces lo que podría abrir paso a una nueva revolución en el campo de la enseñanza programada.

Un STI es un *sistema experto*; esto es, un mecanismo informático que simula lo que haría un experto humano en un área particular de especialización (Castillo, Gutiérrez y Hadi, 1997). En el área educativa, el sistema experto tendría que recoger información sobre el usuario, evaluar el grado de conocimiento que tiene sobre la materia y llevar a cabo un proceso de toma de decisiones sin intervención humana acerca del suministro de ayuda que habría que proporcionar para que la asimilación del contenido fuese óptima.



No obstante, el desarrollo de sistemas expertos orientados a la enseñanza no ha gozado de una gran atención frente a otros campos como los negocios, la industria, la medicina, las aplicaciones militares o la investigación espacial (Castillo et al., 1997). Por otra parte, los sistemas expertos más utilizados en el desarrollo de este tipo de herramientas han sido los sistemas basados en reglas clásicos. Sin embargo, los sistemas clásicos basados en reglas han sido criticados porque no tienen en cuenta eficientemente la naturaleza incierta de la realidad y porque tienden a generar conclusiones falaces (Cowel, Dawid, Lauritzen, y Spiegelhalter, 1999).

La investigación en Inteligencia Artificial ha desarrollado un nuevo tipo de herramientas que resuelven algunos de los problemas detectados en los sistemas expertos tradicionales. Las redes bayesianas son una generalización de los sistemas expertos clásicos que permiten modelar la realidad tanto cuantitativamente como cualitativamente. Gracias a su habilidad para representar los estados de un sistema de incertidumbres relacionadas, la utilización de redes bayesianas en el diseño de STI se ha mostrado fructífera y supone un gran avance en el campo de la enseñanza programada.

Con lo expuesto anteriormente el objetivo de este artículo es; en primer lugar, presentar algunas nociones básicas sobre las redes bayesianas. En concreto, se tratará su habilidad para representar gráficamente la estructura cualitativa de la realidad entendida como un conjunto de dependencias condicionales; por otro lado, trataremos el modo en que representan la incertidumbre asociada al modelo por medio de funciones probabilísticas. A continuación, describiremos el funcionamiento de algunos STI basados en redes bayesianas. Por último, propondremos algunos aspectos que consideramos importantes ante el desarrollo de STI con base en los avances recientes en el campo de la psicología del aprendizaje y de la medición psicológica.

2.- Redes Bayesianas

Las redes bayesianas (también conocidas como redes causales probabilísticas, redes causales, sistemas expertos bayesianos, redes de creencia, sistemas expertos probabilísticas o diagramas de influencia) son herramientas estadísticas que representan un conjunto de incertidumbres asociadas sobre la base de las relaciones de independencia condicional que se establecen entre ellas (Edwards, 1998). Siguiendo a Kadie, Hovel y Horvitz (2001) diríamos que una red bayesiana es un conjunto de variables, una estructura gráfica conectando estas variables y un conjunto de distribuciones de probabilidad condicional. Codifica incertidumbre asociada a cada variable por medio de probabilidades y, gracias al teorema de Bayes, esta incertidumbre es susceptible de ser modificada con base en observaciones (o evidencias) sobre el modelo.

Aunque se podría establecer una distinción más específica de los elementos que componen una red bayesiana (Edwards, 1998), se suele decir que una red bayesiana tiene dos dimensiones: una cualitativa y otra cuantitativa (p. e. Cowell, et al., 1999; Garbolino y Taroni, 2002; Nadkarni y Shenoy, 2001, 2004; Martínez y Rodríguez, 2003).



2.1.- Dimensión cualitativa

Una red bayesiana es un *grafo*; esto es, una representación gráfica de un problema. Aunque la definición de grafo, y la terminología que la acompaña varía en función de los autores (Harary, 1969; Ronald, 1988; Spirtes, Glymour y Scheines, 2000; Tutte, 1984); podemos definir grafo como *un par* $G = (V, E)$, donde V es un conjunto finito de vértices, nodos o variables y E es un subconjunto del producto cartesiano $V \times V$ de pares ordenados de nodos llamados enlaces o aristas.

Por otro lado, una red bayesiana es un tipo concreto de grafo que se denomina *grafo dirigido acíclico* (GDA). Es dirigido porque los enlaces entre los vértices de la estructura están orientados. Por ejemplo, si $(A, B) \in E$ pero $(B, A) \notin E$, diremos que hay un enlace dirigido o arco entre los nodos y lo representaremos como $A \rightarrow B$. Por su parte, es acíclico porque no pueden existir ciclos o bucles en el grafo; esto es, que si empezamos a recorrer un camino dirigido desde un nodo nunca podríamos regresar al punto de partida.

Una conexión tipo $A \rightarrow B$ indica dependencia o relevancia directa entre las variables. En este caso se está representando que B depende de A o que A es la causa de B y B el efecto de A . Dada esta habilidad para codificar causalidad, las redes bayesianas han sido utilizadas para el modelado o la búsqueda automática de estructuras causales en bases de datos (López, García y De la Fuente; 2006). También se dice que A es *padre* (o la variable *madre*) de B y que B es el *hijo* (o la variable *hija*) de A . Aunque la presencia de arcos entre nodos codifica información esencial sobre el modelo representado en la red, la ausencia de arcos entre nodos aporta una valiosa información ya que el grafo codifica independencia condicional.

El principio de independencia condicional en una representación gráfica. Podríamos decir que las redes bayesianas son una representación gráfica del principio de independencia condicional en términos probabilísticos. Este principio quedaría enunciado del siguiente modo: *sean tres conjuntos X , Y y Z de variables; diríamos que los conjuntos X e Y son (condicionalmente) independientes dado el conjunto Z si y solamente si*

$$P(x | z) = P(x | yz) \quad (1)$$

O dicho de otro modo, *dos variables X e Y , son independientes en términos probabilísticos de una tercera Z si y solamente si*

$$P(xy | z) = P(x | z) \times P(y | z) \quad (2)$$

La consecuencia fundamental de este principio es que la probabilidad de X es la misma condicionándola a Z o condicionándola a Z e Y .

Cualquier red bayesiana podría descomponerse en tres tipos de conexiones básicas, cada una con propiedades diferentes en el proceso de propagación de probabilidades. En primer lugar, las conexiones *seriales* o *cadena causales* representan un conjunto de variables asociadas linealmente que denota dependencia entre las variables (Figura 1). En nuestro ejemplo, la variable B depende de A y la variable C depende del valor de B . Así, cuando sabemos algo sobre A podemos modificar nuestra creencia sobre el estado de B y esta



información se propagará hasta *C*. Sin embargo, si encontramos una evidencia sobre *B*, añadir evidencias sobre *C* no alterará nuestro conocimiento sobre *A* y viceversa. En este caso decimos que *A* y *C* son condicionalmente independientes dado *B*.

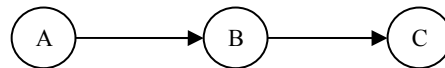


Figura 1. Conexión serial

En las conexiones *divergentes*, también conocidas como *clasificadores ingenuos de Bayes*, tenemos un nodo padre (o clase) que proyecta sus arcos sobre varios hijos (Figura 2a). Este tipo de conexión es el más apropiado para representar procesos de diagnóstico médico (Herskovits y Dagher, 1997) y como veremos más adelante son útiles para estimar el conocimiento asociado a conceptos en función del comportamiento del usuario. Cuando no conocemos el estado de la variable madre existe dependencia entre las variables hijas. Sin embargo, cuando el estado de esta variable se conoce, las evidencias sobre las variables hijas no se propagarán entre ellas. En nuestro ejemplo diremos que *A* y *C* son independientes dado *B*.

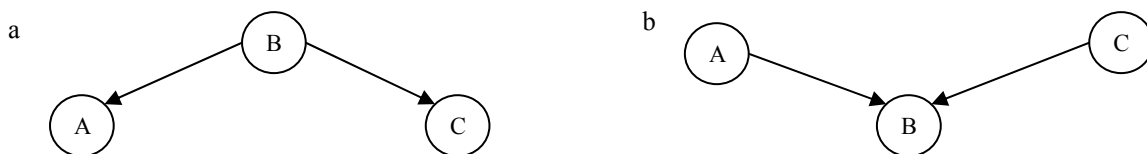


Figura 2. Conexión divergente (a) y conexión convergente (b).

Por último, en las conexiones *convergentes* (llamadas también *cabeza a cabeza*) varias variables apuntan con sus arcos hacia una variable de convergencia (Figura 2b). En este tipo de conexiones las variables madre son independientes entre si. Sin embargo, tenemos una evidencia sobre la variable hija, las variables madre se tornarán dependientes. En nuestro ejemplo diríamos que *A* y *C* son condicionalmente dependientes dado *B*.

2.2.- Dimensión cuantitativa

Existen tres elementos esenciales que caracterizan la dimensión cuantitativa de una red bayesiana: el concepto de probabilidad como un grado de creencia subjetiva relativa a la ocurrencia de un evento, un conjunto de funciones de probabilidad condicionada que definen a cada variable en el modelo y el teorema de Bayes como herramienta básica para actualizar probabilidades con base en experiencia.

Tendríamos, como mínimo, cuatro formas de entender la probabilidad: desde un punto de vista clásico, desde una perspectiva empírica, axiomáticamente y la concepción bayesiana o subjetiva. Por un lado, tenemos las teorías objetivistas dentro de las cuales se encuadra la concepción *frecuentista* de la probabilidad, y por otro tenemos las epistemológicas dentro de las cuales la *subjetiva* es la más famosa (p. e. Alonso y Tubau, 2002; Cowell et al., 1999; Cox, 1964; De la Fuente, García y De la Fuente, 2002; Neapolitan y Morris, 2004). De una manera u otra, la probabilidad es una manera de cuantificar la incertidumbre asociada a la ocurrencia



de eventos y las redes bayesianas se basan en una idea subjetiva de la probabilidad, siendo el teorema de Bayes el *motor* de actualización de probabilidades.

Las redes bayesianas utilizan el teorema de Bayes como herramienta básica para actualizar probabilidades con base en las evidencias aportadas sobre el modelo. En primer lugar hay que compilar la red para que se cree una representación interna de las probabilidades del modelo. De esta manera se estima lo que se conoce como distribución previa (o *distribución a priori*) del modelo. A continuación se van añadiendo evidencias sobre el estado de las variables del modelo y se van obteniendo sucesivamente lo que se conoce como distribuciones posteriores (o *a posteriori*).

De esta forma, en una red bayesiana cada variable es entendida como una función de probabilidad condicionada o una función de densidad condicionada a los valores que toman las variables de las que depende.

La ventaja de utilizar un GDA para construir un modelo probabilístico es que siempre tenemos la posibilidad de conocer la verosimilitud del modelo factorizándolo. Así pues, para cada variable $v \in V$, tenemos que especificar las distribuciones condicionales de X_v dados sus padres $X_{pa(v)}$. Si entendemos que esta densidad es $P(x_v / x_{pa(v)})$, entonces la densidad global conjunta se deriva de

$$P(x) = \prod_{v \in V} P(x_v | x_{pa(v)}) \cdot \quad (3)$$

3.- Ejemplos de tutores bayesianos inteligentes

Los sistemas de tutorización basados en redes bayesianas no son muy abundantes aunque la utilización de este tipo de herramientas en este sentido sería de gran utilidad práctica. No obstante, se han desarrollado algunos sistemas que han mostrado ser eficientes en la modelización del alumnado que ha trabajado con ellos. A continuación presentaremos algunos de los STI que se han desarrollado usando redes bayesianas y comentaremos sus rasgos más relevantes.

3.1.- OLAE y SIETTE

Una de las primeras aproximaciones orientadas a evaluar el aprendizaje por medio de redes bayesianas se llevó a cabo con estudiantes de secundaria en la asignatura de física (Martin y VanLehn, 1995). OLAE (acrónimo de la expresión anglosajona *Off-Line Assessment of Expertise*) fue un sistema que evaluaba diferentes aspectos clave en la resolución de problemas de física. Constaba de una interfaz gráfica dividida en cuatro secciones: una barra de tareas para seleccionar categorías de problemas, un recuadro que presentaba el enunciado del problema junto a una representación gráfica del mismo, un croquis del problema destinado a que el aprendiz señalase las fuerzas y las ecuaciones



implicadas en el problema y, por último, una sección donde se tenía que dar la respuesta. El sistema iba creando redes bayesianas a partir de las respuestas del usuario para modelar el nivel de conocimiento del estudiante.

No obstante, OLAE no era un STI. Más bien, era una herramienta de evaluación y su funcionamiento no era dinámico sino estático ya que realizaba las operaciones después de que el usuario hubiese dado sus respuestas. Por otro lado, se limitaba a estimar la probabilidad relativa a la utilización de ecuaciones y principios físicos, no a enseñarlos. OLAE supuso un avance en la estimación de habilidades con redes bayesianas y dio paso a sistemas de tutorización dinámicos más sofisticados.

El sistema SIETTE (Sistema Inteligente de Evaluación mediante Test para la TeleEducación) integrado en el proyecto TREE (*TRaining of European Environmental Trainers and Technicians*) utiliza la filosofía de los Tests Adaptativos Computerizados (TAC) basados en la Teoría de Respuesta al Ítem (TRI) como herramientas básicas para desarrollar sistemas de tutorización inteligente más eficientes (Conejo, Millán, Pérez-de-la-Cruz y Trella, 2001; Conejo et al., 2004). Una de las características más importantes del sistema SIETTE es que funciona a través de Internet. Sin embargo, el sistema no está implementado en un STI *per se*; sino que, más bien, SIETTE se dedica a evaluar el conocimiento de forma adaptativa por medio de procedimientos eficientes.

3.2.- HYDRIVE

El sistema HYDRIVE fue desarrollado por los Laboratorios Amstrong de las Fuerzas Aéreas de los Estados Unidos para simular el funcionamiento del avión de combate F-15 (Mislevy y Gitomer, 1996). Los problemas se presentan en formato de video donde el piloto describe algunas deficiencias en el funcionamiento de un aparato que está aterrizando o que ya ha aterrizado (por ejemplo, el chequeo rutinario del timón de aterrizaje no responde correctamente). La interfaz gráfica permite al estudiante llevar a cabo una tarea de resolución de problemas revisando videos del aparato y actuando sobre ellos.

La red bayesiana generada a partir de HYDRIVE consta de 22 nodos organizados jerárquicamente en cuatro capas. En la cúspide de la jerarquía aparece la ejecución global del usuario. En un nivel inferior se reflejan los tres tipos de conocimientos que el sistema evalúa: conocimiento del sistema, conocimiento estratégico y conocimiento procedimental. En la tercera capa están los subcomponentes de cada tipo de conocimiento. La cuarta capa sirve para recoger los datos de la ejecución del alumno. Todos los arcos están orientados en sentido decreciente desde las capas superiores a las inferiores.

3.3.- Andes

El rasgo que caracteriza a Andes es que no reduce la iniciativa del usuario estableciendo vías estrictas en el aprendizaje (Conati, Gertner, y VanLehn, 2002). Más bien, es un sistema abierto que permite aprender una habilidad o concepto por medio de diferentes procedimientos. Otra característica importante de Andes es que puede suministrar ayuda dinámica cuando el/la estudiante se encuentra en un callejón sin salida y no sabe como



continuar en la resolución de un problema. Se desarrolló a partir de OLAE y su objetivo es la enseñanza de física newtoniana a nivel universitario.

Andes tiene dos módulos principales, uno para resolución de problemas y otro para el estudio de ejemplos. El módulo de resolución de problemas es parecido al de OLAE pero incorporando una ayuda dinámica más eficiente. En el módulo de estudio de ejemplos los estudiantes van siguiendo la resolución de los problemas de manera análoga a como lo harían en una clase impartida por un tutor humano. Así, el desarrollo de la explicación del problema va adaptándose a lo que la persona sabe a lo largo del proceso formativo.

El sistema hace una serie de asunciones acerca del proceso de generalización del aprendizaje; esto es, el grado en el que los usuarios son capaces de trasladar lo aprendido en un contexto a otro. Esto es un problema crucial en el proceso de enseñanza-aprendizaje ya que uno de los objetivos de la enseñanza es aplicar lo que se ha aprendido en clase a problemas de la vida cotidiana. Para modelar el proceso de generalización utiliza una estructura de red bayesiana divergente donde el nodo de divergencia representa la probabilidad de uso correcto de una determinada ecuación física y los nodos hijos representan diferentes contextos.

Para llevar a cabo el proceso de evaluación Andes utiliza redes bayesianas dinámicas (Millán, Perez-de-la-Cruz, y Suárez, 2000). La estructura de las redes generadas por Andes codifican dos tipos de conocimiento: conocimiento de dominio general y conocimiento específico de la tarea. En la parte de conocimiento general existen dos tipos de nodos: *nodos de regla*, que representan la probabilidad de uso correcto de una ecuación física; y *nodos de regla contextuales*, que representan la probabilidad del uso correcto de una ecuación en el correspondiente contexto aunque no pueda aplicarse en otros contextos. Estos nodos son, como se ha comentado, los responsables de estimar el grado de generalización en el uso de ecuaciones. En la parte de conocimiento específico hay nodos de *hechos* y nodos de *objetivos*, que representan proposiciones concretas definidas en el problema o que han de alcanzarse con la resolución del mismo; y nodos de *estrategia* que indican la estrategia más probable seguida por el usuario.

4.- Perspectivas

Como se ha comentado, la idea de la enseñanza automática o programada no es nueva. Más bien, podríamos decir que fue introducida en el contexto de la psicología científica de manos del conductismo. Sin embargo, dado que la informática no estaba tan desarrollada como lo está hoy día y debido a la influencia del paradigma cognitivo, los modelos de enseñanza automática han estado relegados a un segundo plano. Por otro lado, las redes bayesianas son herramientas estadísticas que permiten representar la incertidumbre asociada a los procesos de aprendizaje y esta representación puede utilizarse en la toma de decisiones orientada a la tutorización automática. Sin embargo, creemos que hay ciertos aspectos que deberían de trabajarse con más profundidad de cara al desarrollo de sistemas de tutorización inteligente más eficientes. En este sentido, podríamos hacer apreciaciones a dos niveles: metodológica y teóricamente.



Desde el punto de vista metodológico, sería interesante profundizar en los modelos de medida. La profundización en las teorías de medida, como la TRI, podrían optimizar los sistemas de tutorización inteligente al aportar técnicas más validas y fiables para estimar el nivel de conocimiento de los estudiantes. Por ello, creemos que sería apropiado ir introduciendo paulatinamente la filosofía de los tests adaptativos computerizados en el campo de los sistemas de tutorización inteligente (Olea y Ponsoda, 1996).

Para terminar, creemos que los sistemas de tutorización se optimizarían en la medida en que estuviesen basados en el conocimiento que tenemos de cómo aprendemos. O dicho de otro modo, en la medida en que sepamos cuales son los procesos psicológicos básicos implicados en el aprendizaje, podremos desarrollar sistemas de tutorización más eficientes. Es por ello que creemos que la Psicología se encuentra en una situación privilegiada para incorporarse al estudio y uso de las Redes Bayesianas como herramientas de modelado (López, García y de la Fuente, 2005; López, García, de la Fuente y de la Fuente, 2007). A este nivel, la Teoría del Marco Relacional (Hayes, Barner-Holmes y Roche, 2002; Blackledge, 2003), dada su relación con el concepto de aprendizaje programado, podría ser de utilidad en el desarrollo de sistemas de tutorización inteligente. Por ello, sería necesario profundizar en los aspectos básicos del aprendizaje en contextos educativos relacionados con las nuevas tecnologías con el fin de mejorar el grado en que son efectivos (López y García, 2007a y 2007b).

5.- Referencias

- Alonso, D., y Tubau, E. (2002). Inferencias bayesianas: una revisión. *Anuario de Psicología*, 33, 25-47.
- Blackledge, J. T. (2003). An introduction to relational frame theory: basic and applications. *The Behavior Analyst Today*, 3, 421-433.
- Carbonell, J.R. (1970). AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction. *IEEE Transactions on Man-Machine Systems*, 11, 190-202.
- Castillo, E., Gutiérrez, J. M. y Hadi, A. S. (1997). *Expert systems and probabilistic network models*. New York: Springer-Verlag.
- Conati, C., Gertner, A. y VanLehn, K. (2002). Using bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Interaction*, 12, 371-417.
- Conejo, R., Guzmán, E., Millán, E., Trella, M., Perez-de-la-Cruz, J. L. y Ríos, A. (2004). SIETTE: A web-based tool for adaptive testing. *International Journal of Artificial Intelligence in Education*, 14, 1-33.
- Conejo, R., Millán, E., Perez-de-la-Cruz, J. L. y Trella, M. (2001). Modelado del alumno: un enfoque bayesiano. *Revista Iberoamericana de Inteligencia Artificial*, 12, 50-58.



- Cowell, R. G., Dawid, A. P., Lauritzen, S. L., y Spiegelhalter, D. J. (1999). *Probabilistic networks and expert systems*. Harrisonburg, VA: Springer.
- Cox, R. T. (1946). Probability, frequency and reasonable expectation. *American Journal of Physics*, 14, 1-13.
- De la Fuente, E. I., García, J., y De la Fuente, L. (2002). Estadística bayesiana en la investigación psicológica. *Metodología de las Ciencias del Comportamiento*, 4, 185-200.
- Dixon, J. R. (1964). *A programmed introduction to probability*. New York: John Wiley & Sons.
- Edwards, W. (1998). Hailfinder. Tools for and experiences with bayesian normative modelling. *American Psychologist*, 53, 416-428.
- Garbolino, P., y Taroni, F. (2002). Evaluation of scientific evidence using bayesian networks. *Forensic Science International*, 125, 149-155.
- Harary, F. (1969). *Graph theory*. Philippines: Westview Press.
- Hayes, C. S., Barnes-Holmes, D and Roche, B. (2002). *Relational frame theory. A post-Skinnerian account of human language and cognition*. New York: Kluwer/Plenum.
- Herskovits, E. H. y Dagher, A. P. (1997). *Applications of bayesian networks to health care (Rep. Téc. NSI-TR-1997-02)* Baltimore, MD: Noetic Systems.
- Hothersall, D. (1997). *Historia de la Psicología* (3ª ed.). México: Mc-Graw Hill. (Trabajo original publicado en 1995)
- Kadie, C. M., Hovel, D., y Horvitz, E. (2001). *MSBNx: a component-centric toolkit for modelling and inference with bayesian networks (Rep. Téc. MST-TR-2001-67)*. Redmon, WA: Microsoft.
- López, J. y García, J. (2007a). Intelligent tutoring systems with bayesian nets. *Proceedings of the International Technology, Education and Development Conference, INTED2007*. Valencia, España.
- López, J. y García, J. (2007b). A Bayes net inside an intelligent tutoring system. How does it work? *Proceedings of the International Technology, Education and Development Conference, INTED2007*. Valencia, España.
- López, J., García, J. y de la Fuente, M. L. (2006). Modelado causal con redes bayesianas. *Actas de las XVII Jornadas de Automática*, 198-202.
- López, J., García, J., de la Fuente, L. y de la Fuente, E. I. (2007). Las redes bayesianas como herramientas de modelado en psicología. *Anales de Psicología*, 23, 307-316.



- López, J., García, J., y De la Fuente, E. I. (2005). Las redes bayesianas en Psicología. *Actas del IX Congreso de Metodología de las Ciencias Sociales y de la Salud*. Granada. España.
- Manning, S. A. y Rosenstock, E. H. (1961). *Classical psychophysics and scaling*. New York: Mc-Graw Hill.
- Martínez, I., y Rodríguez, C. (2003). Modelos gráficos. En Y. del Águila et. al. (Eds.), *Técnicas estadísticas aplicadas al análisis de datos* (pp. 217-257). Almería: Servicio de Publicaciones de la Universidad de Almería.
- Martin, J. y VanLehn, K. (1995). Student assessment using bayesian nets. *International Journal of Human-Computer Studies*, 42, 575-591.
- Millán, E., Perez-de-la-Cruz, J. L. y Suárez, E. (2002). Adaptive bayesian networks for multilevel student modelling. En G. Gauthier, C. Frasson y K. VanLehn (Eds.), *Lecture notes in computer science 1839*, 534-543. Berlin: Springer-Verlag.
- Mislevy, R. J. y Gitomer, D. H. (1996). The role of probability-based inference in an intelligent tutoring system. *User-Mediated and User-Adapted Interaction*, 128, 253-282.
- Nadkarni, S., y Shenoy, P. P. (2001). A bayesian network approach to making inferences in causal maps. *European Journal of Operational Research*, 128, 479-498.
- Nadkarni, S., y Shenoy, P. P. (2004). A causal mapping approach to constructing bayesian networks. *Decision Support Systems*, 38, 259-281.
- Neapolitan, E. R., y Morris, S. (2004). Probabilistic modelling with bayesian networks. En D. Kaplan (Ed.), *The SAGE handbook of quantitative methodology for the social sciences* (pp. 371-390). Thousand Oaks, CA: Sage Publications.
- Olea, J., y Ponsoda, V. (1996) Tests adaptativos informatizados. En J. Muñiz (Ed.), *Psicometría* (pp. 729-783). Madrid: Universitas.
- Ronald, G. (1988). *Graph theory*. Menlo Park, CA: Benjamin/Cummings.
- Spirtes, P., Glymour, C., y Scheines, R. (2000). *Causation, prediction and search* (2ª). Cambridge, MA: MIT Press.
- Stevens, A. L. y Collins, A. (1977). The goal structure of a socratic tutor. *Proceedings of the 1977 Annual Conference, ACM '77*. New York.
- Tutte, W. T. (1984). *Graph theory*. Cambridge: Cambridge University Press.
- Urretavizcaya, M. (2001). Monografía: sistemas inteligentes en el ámbito de la educación. *Revista Iberoamericana de Inteligencia Artificial*, 12, 2-4.



- Wu, A. K. (1993). Paradigms for ITS (intelligent tutoring system). *Proceedings of the Conference on Computer, Communication, Control and Power Engineering, TENCON '93*. Beijing, China.
- Xiang, Y. (2002). *Probabilistic reasoning in multiagent systems. A graphical models approach*. Cambridge: Cambridge University Press.