

Las redes bayesianas como herramientas de modelado en psicología*

Jorge López Puga^{1**}, Juan García García^{1**}, Leticia de la Fuente Sánchez¹ y Emilia Inmaculada de la Fuente Solana²

¹Universidad de Almería, ²Universidad de Granada

Resumen: Cada vez son más numerosas las investigaciones que trabajan con un amplio número de variables donde existen relaciones complejas entre ellas. Las redes bayesianas son herramientas estadísticas surgidas en el campo de la Inteligencia Artificial que nos permiten afrontar situaciones de investigación de estas características. Una red bayesiana es un grafo dirigido acíclico que codifica relaciones probabilísticas de dependencia e independencia condicional y que actualiza el modelo con base en las evidencias muestrales mediante la regla de Bayes. Este artículo describirá los principios matemático-estadísticos esenciales de las redes bayesianas y las ventajas que tienen frente a otras herramientas multivariantes. Finalmente, revisaremos las aplicaciones, que desde la Psicología, se han aportado, así como se describen sus usos potenciales.

Palabras clave: Redes bayesianas; sistemas expertos; estadística bayesiana y modelado estadístico.

Title: Bayesian Nets as Modelling Tools in Psychology.

Abstract: There is more and more research projects which have a lot of variables with complex relations between them. Bayes nets are statistical tools emerged from Artificial Intelligence field that allow us to face up these research situations. A Bayes net is a directed acyclic graph that represents conditional dependence and independence probabilistic relations and it uses the Bayes rule to update the model when some evidences are took into account. The paper describes Bayes nets mathematical essentials and the advantages as opposed to others multivariate tools. Finally, we review applications and potential applications of Bayes nets in Psychology.

Key words: Bayes nets; expert systems; bayesian statistic and psychological modelling.

El término *Inteligencia Artificial* (IA) fue acuñado en 1956 durante un congreso celebrado en los Estados Unidos. Sin embargo, la idea de desarrollar máquinas inteligentes tiene un bagaje más largo. Por ejemplo, los árabes de Al-Andalus ya experimentaban durante la Edad Media con artilugios derivados del astrolabio (aparato diseñado para la navegación naval teniendo en cuenta la disposición de las estrellas) que les permitían razonar teológica y científicamente (Zaccagnini, 1994).

Los inicios de la IA podrían hallarse en los trabajos de Allan M. Turing, representados por el artículo *Computing machinery and intelligence*, publicado en *Mind* durante 1950; y a los de Claude Shannon, en su tesis *Análisis simbólico de los circuitos de relé y conmutación* (Rodríguez, 1996). Los objetivos de la IA se concretaban en el desarrollo de sistemas artificiales que emularan los procesos psicológicos inteligentes que llevan a cabo las personas de forma natural. Este deseo de modelado psicológico se hizo realidad gracias al gran desarrollo que estaban sufriendo los mecanismos de cómputo artificiales. Aunque de esta fusión inicial de las ciencias de la computación y la Psicología surgió lo que se ha venido a denominar como *paradigma cognitivo* en Psicología (Adarraga y Zaccagnini, 1988), es posible que la mala estructuración de los problemas investigados por la Psicología y el conocimiento difuso que implican ha estado produciendo una separación paulatina de las disciplinas (Adarraga y Zaccagnini, 1992).

Podríamos distinguir otra rama parcialmente separada de la anterior donde la IA y la Psicología evolucionaron conjuntamente. En este caso tendríamos que aludir a lo que po-

dríamos denominar *filosofía de red* en el campo de los sistemas expertos. La filosofía de red, surgida en el contexto del conexionismo y el asociacionismo, se inicia en los años 50, cuando el neurobiólogo Warren McCulloch y el estadístico Walter Pitts desarrollaron la primera neurona artificial (Quinlan, 1991). Esta iniciativa fue evolucionando hasta que, coincidiendo con el auge conexionista en Psicología (Hollnagel, 1995/1997), se forjó el paradigma de las redes neurales artificiales (RNA). Desde esta perspectiva se trató de desarrollar sistemas inteligentes que modelizasen los procesos de cómputo que tienen lugar en el sistema neural humano. La base principal de las RNA es el *principio de Hebb*, *principio de estabilización selectiva* o *principio de convergencia sincrónica* (Carlson, 1993/2000), que viene a ser una generalización del aprendizaje clásico aplicado a la estructura química y anatómica de las sinapsis.

El paradigma de las RNA comenzó a perder importancia en los alrededores de 1970 cuando se demostró que, como mecanismos de cómputo artificiales, se veían desbordadas cuando se las desafiaba con problemas lineales simples (Quinlan, 1991; SPSS Inc. y Recognition Systems Inc., 1997). Sin embargo, pese a las críticas y a los problemas que experimentó la investigación en RNA, en los años 80 del siglo pasado se revitalizó el interés por este tipo de herramientas estadísticas. Como hitos destacados de este nuevo auge podríamos señalar el desarrollo de los Mapas Autoorganizativos de Kohonen, que simulan procesos de aprendizaje neural basados en la reestructuración topográfica de las conexiones sinápticas; y el diseño de las redes tipo Hopfield, que se valen de principios termodinámicos y de funciones de energía para entender el estado global de una red neural en un instante concreto del tiempo. Hoy en día, pese a los grandes avances que se produjeron a finales del siglo XX, el paradigma de las RNA se encuentra relativamente estancado. No han surgido nuevos algoritmos de propagación de la información que hayan tenido un efecto considerable sobre la comunidad científica y muchos de los métodos de estima-

* Este trabajo ha sido realizado en el marco del proyecto del Plan Nacional de I+D+I del MCYT (BSO2002-03347).

** **Dirección para correspondencia** [Correspondence address]: Jorge López Puga (jpuga@ual.es) y Juan García García (jgarcia@ual.es), Área de Metodología de las Ciencias del Comportamiento. Departamento de Ciencias Humanas y Sociales. Facultad de Humanidades y Ciencias de la Educación, Universidad de Almería. Carretera de Sacramento s/n. 04120 Almería (España).

ción que utilizan las RNA han sido equiparados a técnicas estadísticas clásicas (Pérez y Martín, 2003).

Las redes bayesianas surgieron hace aproximadamente tres décadas como alternativas a los sistemas expertos clásicos orientados a la toma de decisiones y la predicción bajo incertidumbre en términos probabilísticos (Cowell, Dawid, Lauritzen, y Spiegelhalter, 1999; Edwards, 1998); sin embargo, la idea de construir redes que representasen información probabilística se remonta al primer cuarto del siglo XX (Pearl, 2001). Una red bayesiana (también denominada como *red causal probabilística*, *sistema experto bayesiano*, *sistema experto probabilístico*, *red causal*, *red de creencia* o *diagrama de influencia*) es una herramienta estadística que representa un conjunto de incertidumbres asociadas teniendo en cuenta las relaciones de independencia condicional que se establecen entre ellas. Forman parte de la familia de los *Sistemas Estocásticos Altamente Estructurados* y pertenecen al conjunto de técnicas dirigidas a la modelización gráfica (Cowell et al., 1999; Martínez y Rodríguez, 2003). Siguiendo a Kadie, Hovel, y Horvitz, (2001) podríamos definir una red bayesiana como un conjunto de variables codificadas probabilísticamente, una estructura gráfica que conecta esas variables en términos de relaciones de independencia condicional y un conjunto de distribuciones de probabilidad condicional susceptibles de ser modificadas con base en evidencias por medio del teorema de Bayes.

Nuestro objetivo es popularizar el formalismo de las redes bayesianas en el contexto de la investigación psicológica. En primer lugar, describiremos los principios matemático-estadísticos en los que se sustentan las redes bayesianas. A continuación, señalaremos las ventajas que presentan estas técnicas estadísticas sobre otras herramientas multivariantes de análisis de datos. Y, por último, revisaremos algunas aportaciones que se han hecho a la Psicología a partir del uso de redes bayesianas, así como los usos potenciales que se podrían dar a estas técnicas en la investigación psicológica.

Fundamentos de las Redes Bayesianas

Un criterio que se ha utilizado para describir la estructura de una red bayesiana atiende al grado en que sus elementos son "visibles" para el usuario (Edwards, 1998; Edwards y Fasolo, 2001). En este sentido una red bayesiana sería una estructura compuesta por cuatro niveles. En el nivel superior, una red bayesiana sería un conjunto de variables representadas por nodos y un conjunto de flechas que relacionan estas variables en términos de influencia. En un nivel inferior estarían los niveles o estados, también conocidos como *espacio de estados* (Nadkarni y Shenoy, 2001, 2004), que pueden asumir cada una de las variables del modelo. En tercer lugar, tendríamos un conjunto de funciones de probabilidad condicional, una para cada nodo, donde se representaría la probabilidad de ocurrencia de cada estado de la variable condicionado a los posibles valores de las variables que determinan el valor de la variable. Por último, en el nivel más subordinado

estarían un conjunto de algoritmos que permitirían que la red recalculase las probabilidades asignadas a cada uno de sus niveles cuando conocemos alguna evidencia sobre el modelo.

Sin embargo, la descripción más frecuente de una red bayesiana se basa en dos elementos, una dimensión cualitativa y otra cuantitativa (p. e. Cowell et al., 1999; Garbolino y Taroni, 2002; Nadkarni y Shenoy, 2001, 2004; Martínez y Rodríguez, 2003). Los fundamentos teóricos de estos dos elementos o dimensiones se sustentan en dos grandes pilares de la modelización matemática, la teoría de grafos y la teoría de la probabilidad (Ríos, 1995). A continuación describiremos en que consiste la dimensión cualitativa y cuantitativa de las redes bayesianas.

Dimensión Cualitativa

Una de las ventajas más importantes de las redes bayesianas es que pueden representar tanto el aspecto cuantitativo de un problema como su aspecto cualitativo. El soporte teórico de la dimensión cualitativa en las redes bayesianas lo aporta la teoría de grafos. La teoría de grafos trata de crear modelos gráficos (*grafos*) que representen los elementos del problema en un sentido holista y fue introducida por Euler para dar solución al problema de los puentes de Königsberg (Harary, 1969; Ríos, 1995). La teoría de grafos se ha utilizado, por ejemplo, para la solución de problemas en la búsqueda de moléculas isomórficas, en el diseño de circuitos electrónicos e incluso en la representación de estructuras empresariales (Harary, 1969; Ríos, 1995; Ronald, 1988); en el campo de la Psicología fue Lewin quien usó la teoría de grafos para construir sus teorías sobre el espacio vital y los vectores psicológicos (Harary, 1969; Hothersall, 1995/1997).

Una red bayesiana es un *grafo* y aunque la definición de grafo, y la terminología que lo acompaña, varía en función de los autores (p. e. Harary, 1969; Ronald, 1988; Spirtes, Glymour, y Scheines, 2000; Tutte, 1984; Xiang, 2002); podemos definirlo como un par $G = (V, E)$, donde V es un conjunto finito de vértices, nodos o variables y E es un subconjunto del conjunto $V \times V$ de pares ordenados de vértices llamados enlaces o aristas. Además, una red bayesiana es un tipo particular de gráfico que se denomina *grafo dirigido acíclico*. Dirigido hace referencia a que los enlaces entre los vértices del grafo están orientados. Así, si los dos pares ordenados (A, B) y (B, A) pertenecen a G diremos que existe un enlace *no dirigido* entre A y B y se representa como $A \sim B$. También podemos decir que A y B son *vecinos* o que son *adyacentes*. El conjunto de vecinos del vértice A es denotado por $Ady(A)$. Sin embargo, 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 $A \rightarrow B$. También podremos decir que A es *padre* o la variable *madre* de B y que B es *hijo* o la variable *madre* de A . Acíclico se refiere a que no pueden existir ciclos o *loops* en el grafo, esto es; que si comenzamos a recorrer un camino dirigido desde un nodo, nunca podríamos regresar al punto de partida.

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 valiosísima información ya que representa independencia condicional (Edwards, 1998). El *principio de independencia condicional* puede enunciarse del siguiente modo: *sean tres conjuntos X , Y y Z de variables; X e Y son (condicionalmente) independientes dado Z si $P(x|z) = P(x|yz)$. Otro modo de enunciar el principio de independencia condicional es el siguiente: *dos variables, X e Y , son independientes en probabilidad de una tercera Z si y solamente si $P(x,y|z) = P(x|z) \times P(y|z)$. La consecuencia fundamental de este principio es que la probabilidad de X es la misma condicionándola a Z que condicionándola a Z e Y . O lo que es lo mismo, que si conocemos el estado de Z , el conocimiento del estado de Y es irrelevante para saber algo de X . Por último, una propiedad importante del principio de independencia condicional es que es simétrico; esto es, cuando X es independiente de Y , Y es independiente de X .**

En una red bayesiana pueden existir tres tipos de conexiones básicas (seriales, convergentes y divergentes), cada una con propiedades cualitativas diferentes y que favorecen la propagación de probabilidades ante una nueva evidencia sobre el modelo. Las *conexiones en serie* representan a un conjunto de variables asociadas linealmente que denota dependencia entre las variables (por ejemplo $A \rightarrow B \rightarrow C$). En este caso, B depende de A y C de B . En términos causales diríamos que A es causa de B y que B es causa de C . En este caso, dada la dependencia entre las variables, cuando conocemos información sobre A podemos modificar nuestra certeza sobre el estado de C ; y a la inversa, cuando sabemos algo sobre el estado de C la creencia sobre el estado de A se altera. Sin embargo, si el estado de B es conocido; conocer alguna información sobre A o C no modificará nuestra creencia sobre el estado de estas variables. Podríamos decir que la propagación de la información se bloquea y se dice que A y C se tornan *condicionalmente independientes* dado B .

En las *conexiones divergentes* tenemos un nodo padre que proyecta sus arcos hacia varios hijos; o lo que es lo mismo, las flechas que salen de él divergen hacia sus hijos (por ejemplo, $A \leftarrow B \rightarrow C$). Este tipo de conexión es el más apropiado para representar enfermedades que son detectables por sus síntomas (Herskovits y Dagher, 1997) y en términos causales se interpretaría que una causa (p. e. un resfriado) genera varios síntomas (p. e. tos, congestión, dolor de cabeza, etc.). Cuando no conocemos el estado de la variable madre existe una dependencia entre las variables análoga a la que se da en la conexión serial. Sin embargo, cuando el estado de la variable madre se conoce, las variables hijo se tornan independientes y la información no se propaga si añadimos evidencias sobre los nodos hijo. En nuestro ejemplo, A y C son independientes dado B .

Por último, en las *conexiones convergentes* varias variables apuntan con sus arcos hacia una variable objetivo (por ejemplo, $A \rightarrow B \leftarrow C$), esto es, convergen en un mismo nodo hijo. En este tipo de conexiones las variables madre son independientes entre sí, pero la variable de convergencia

puede influir sobre sus nodos padre. La propiedad importante de este tipo de conexión a la hora de propagar la información es que cuando se tiene evidencia sobre la variable de convergencia, los nodos padre se vuelven dependientes y la evidencia sobre el estado de uno de ellos se propaga por los demás. Se diría, en nuestro ejemplo, que A y C son dependientes dado B .

Dimensión cuantitativa

Hay 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, el teorema de Bayes como heurístico actualizador de creencias y un conjunto de funciones de probabilidad condicionada.

Existen, al menos, cuatro formas de entender la probabilidad: la clásica, empírica, axiomática y la subjetiva. Desde la concepción clásica, introducida por Laplace, la probabilidad de que ocurra un evento de un espacio muestral viene dado por la razón que se establece entre el número de casos favorables asociados al suceso y el número de casos posibles. Sin embargo, este principio conduce a conclusiones falaces en determinadas situaciones (p. e., cuando la muestra de sucesos es muy pequeña). Para solventar este problema se propuso otra definición de probabilidad, compatible con la clásica, donde se entiende que la probabilidad de un evento es entendida como una función del número de ensayos cuando estos tienden a infinito. Pero esta definición de probabilidad también plantea problemas porque, por ejemplo, hay situaciones en las que no es posible repetir un experimento infinitas veces. Por lo tanto se hizo necesario desarrollar otra forma de entender la probabilidad. Fue Kolmogoroff, entre 1930 y 1950, quien propuso un conjunto de axiomas que hicieran matemáticamente tratable el concepto de probabilidad; a saber: que la probabilidad está comprendida entre 0 y 1, que la probabilidad de un suceso seguro es 1 y que la probabilidad de la unión de sucesos incompatibles es la suma de sus probabilidades.

Por último, la probabilidad puede ser entendida como algo subjetivo, como el grado de creencia que tenemos sobre la ocurrencia de un suceso. La probabilidad sería entendida como "algo que está en nuestra mente y no en el mundo físico real" (Dixon, 1967/1970, p. 5.2). A esta concepción de la probabilidad, derivada del *principio de la razón insuficiente* o *principio de incertidumbre*, se le denomina *bayesiana* y para clarificarla se suele aludir a la dicotomía *frecuentista/bayesiana* (p. e. Alonso y Tubau, 2002; Cowell *et al.*, 1999; De la Fuente, García, y De la Fuente, 2002; Heckerman, 1995; Serrano, 2003). Por otro lado, la concepción bayesiana de la probabilidad está lejos de ser unitaria; más bien, como señalan Neapolitan y Morris (2004) pueden identificarse tres grupos de perspectivas bayesianas que difieren entre sí en la concepción de *probabilidad física*: a) una rama que la considera inexistente, b) otra que la rechaza y c) un tercer grupo que la usa como si existiese aunque no la justifica filosóficamente. De

una manera u otra, la probabilidad es una forma de cuantificar la incertidumbre asociada a la ocurrencia de eventos y en el caso de las redes bayesianas la probabilidad es entendida en sentido subjetivo, es decir, como la creencia subjetiva que tenemos sobre la ocurrencia de un suceso.

El teorema de Bayes es la regla básica de actualización de creencias en una red bayesiana. Se deduce del axioma que relaciona la probabilidad de la intersección de sucesos y la probabilidad condicional [$p(A \cap B) = p(A|B) \times p(B) = p(B|A) \times p(A)$], lo que favorece que podamos trabajar de manera eficiente con la propagación de probabilidades en modelos gráficos en términos de dependencia e independencia condicional (Cowell *et al.*, 1999). La regla de Bayes es una herramienta de cálculo útil cuando tenemos que actualizar probabilidades asociadas a un modelo del que recibimos evidencias sobre alguna de las variables implicadas. Por ejemplo, si sabemos que en nuestra ciudad llueve el 3% de los días y que el 30% de los días está nublado (algunas veces está nublado y no llueve); usando el teorema de Bayes podríamos concluir que la probabilidad de que hoy llueva ya que está nublado es de 0.1 (asumiendo que siempre que llueve está nublado). En el teorema de Bayes la probabilidad que conocemos gracias a nuestra experiencia se le denomina *previa* o *a priori* (en nuestro ejemplo que el 3% de los días llueve en nuestra ciudad) mientras que la probabilidad que se obtiene tras encontrar alguna evidencia (que hoy está nublado) sobre el modelo se denomina *posterior* o *a posteriori* (en nuestro ejemplo sería la probabilidad de que llueva dado que está nublado). Lo que hace una red bayesiana, a grandes rasgos, es actualizar probabilidades dentro de un grafo dirigido acíclico teniendo en cuenta los principios de independencia condicional en el momento que incorporamos una evidencia al modelo.

Por último, una red bayesiana necesita un conjunto de funciones de probabilidad condicional, una por cada variable o nodo en la red, sobre los que ha de aplicarse la regla de Bayes. Más específicamente, cada variable de la red está caracterizada por una *tabla de probabilidad condicional* donde se representan los valores que puede tomar esa variable en función de los valores que toman el conjunto de variables de las que depende. En este sentido siguiendo a Cowell *et al.* (1999), y teniendo en cuenta que trabajamos con un grafo dirigido acíclico G con un conjunto de nodos V , para cada $v \in V$ tenemos que especificar las distribuciones condicionales de X_v dado sus "padres" $X_{pa(v)}$. Si suponemos que esta densidad es $p(x_v | x_{pa(v)})$, la función de densidad conjunta sería:

$$p(x) = \prod_{v \in V} p(x_v | x_{pa(v)})$$

Ventajas de las Redes Bayesianas

La ventaja más importante que ofrecen las redes bayesianas respecto a otros métodos de análisis multivariante es que

permiten representar al unísono la dimensión cualitativa y cuantitativa de un problema en un entorno gráfico inteligible (Edwards, 1998; Heckerman, 1995; López, García, y De la Fuente, 2005).

Otra ventaja importante de las redes bayesianas es que pueden trabajar con datos perdidos de una manera eficiente, algo que en la práctica es deseable (p. e. destacado por Heckerman, 1995; Jansen, *et al.*, 2003; Nadkarni y Shenoy, 2004). Sin embargo, una desventaja de los casos perdidos es que si un estado de la variable no aparece la probabilidad para este es cero. También permiten reducir el sobre ajuste de los datos (Heckerman, 1995) y, como técnica estadística bayesiana, combinar el conocimiento previo que tenemos respecto al problema de estudio con datos experimentales (Nadkarni y Shenoy, 2004).

Las redes bayesianas también permiten descubrir la estructura causal subyacente en un conjunto de datos (p. e., Cowell *et al.*, 1999; Heckerman, 1995; Neapolitan y Morris, 2004; Pearl, 2001; Spirtes *et al.*, 2000). Esto significa que podemos construir un modelo gráfico probabilístico a partir de una base de datos que contenga un conjunto de observaciones sobre un conjunto de variables.

Las redes bayesianas presentan ventajas frente a los sistemas expertos clásicos basados en reglas cuando se van a utilizar, por ejemplo, para la toma de decisiones. En primer lugar, las redes bayesianas representan toda la información en un único formato (probabilístico y gráfico) lo que hace sencillas las interpretaciones, permiten retractarse de conclusiones obtenidas con anterioridad que ya no son razonables a la luz de nuevas evidencias, nos proporcionan una visión general del problema, generan un conjunto de alternativas ordenadas y facilita la explicación de las conclusiones (Huete, 1998). Por otro lado, cuando construimos una red bayesiana a partir del conocimiento de un experto para usarla en la orientación ante la toma de decisiones, la asignación de probabilidades (o su generación) es más sencilla (Edwards, 1998; Martínez y Rodríguez, 2003). Además, las redes bayesianas permiten trabajar con conceptos de la teoría de la decisión como *valor* o *valor esperado* frente a problemas de decisión (Edwards, 1998; Edwards y Fasolo, 2001; Maciá, Barbero, Pérez-Llantada, y Vila, 1990).

En el contexto de la inferencia, las redes bayesianas permiten realizar inferencias bidireccionales; esto es, desde los efectos a las causas y desde las causas a los efectos. Y lo que es más interesante, permiten llevar a cabo *inferencias abductivas*; o sea, encontrar la mejor explicación para un conjunto de datos (Gámez, 1998; Huete, 1998). Sin embargo, la ventaja más importante de una red bayesiana en el ámbito de la inferencia estadística se deriva de su habilidad para realizar *computaciones locales* (Pearl, 2001; Xiang, 2002). Esta propiedad permite que la actualización de probabilidades se lleve a cabo eficientemente sin tener que calcular todas las posibles combinaciones entre todos los niveles de las variables (que supondría un incremento exponencial de los cálculos a medida que aumenta el número de variables del modelo o el número de estados por variable) cuando se incorpora deter-

minado conocimiento a la estructura gráfica. Gracias a esta propiedad derivada de los principios de dependencia e independencia condicional, las redes bayesianas son herramientas especialmente indicadas para tareas que requieren una actualización rápida y continua en los procesos de control, por ejemplo, industriales.

Aplicaciones y Perspectivas de las Redes Bayesianas en Psicología

El uso de redes bayesianas como herramienta de análisis de datos no está extendido en el contexto de la Psicología. Sin embargo, podríamos encontrar fácilmente los beneficios de la utilización de estas herramientas en todas las grandes áreas de aplicación de la Psicología. Edwards (1998) destaca que las redes bayesianas son importantes para la Psicología en dos sentidos: económica y científicamente. A nivel económico las redes bayesianas podrían generar un mercado de “elicitación de probabilidades” orientado al desarrollo de sistemas expertos donde la Psicología podría jugar un papel fundamental. La tarea a realizar por la Psicología en este sentido sería orientar a los expertos sobre como hacer juicios de probabilidad apropiados intentando sortear los sesgos en que solemos incurrir las personas cuando realizamos tareas de este tipo. En el plano científico las redes bayesianas no pueden perderse de vista si la Psicología pretende conocer los mecanismos por los cuales las personas evaluamos, decidimos y realizamos inferencias; ya que pueden servir de referencia analítica y teórica en el desarrollo de modelos de razonamiento, aprendizaje y percepción de la incertidumbre.

La propuesta más llamativa del uso de las redes bayesianas en Psicología proviene de la mano de Clark Glymour (2001) en su libro *The mind's arrow. Bayes nets and graphical causal models in psychology*. La obra de Glymour gira en torno al fenómeno de la *causalidad* y al proceso de construcción de conocimiento por medio de la observación y la manipulación de nuestro entorno. El aprendizaje causal es el punto arquimédico con el que inicia su trabajo, proponiendo que en nuestra interacción con el entorno aplicamos una serie de procesos (análogos a los algoritmos desarrollados para estudiar las relaciones causales estadísticas) que conducen a una representación interna de la estructura causal. Para Glymour, las personas llevan a cabo procesos de aprendizaje y descubrimiento causal que serían compatibles con el formalismo de las redes bayesianas. Sin embargo, su propuesta no se limita al desarrollo ontogenético del aprendizaje causal humano, sino que traslada esta filosofía al campo de la Neuropsicología (reentendiendo los diagramas de flujo de la Psicología Cognitiva clásica en un sentido bayesiano) o a la Psicometría (criticando el análisis factorial y el análisis de regresión como métodos fiables para extraer el número correcto de variables latentes y para representar la estructura causal subyacente).

Pese a que la propuesta teórica de Glymour en relación al aprendizaje causal ha mostrado evidencias empíricas y ha

dado lugar a una teoría del aprendizaje causal en niños pequeños (Gopnik, *et al.*, 2004; Gopnik y Schulz, 2004; Gopnik, Sobel, Schulz, y Glymour, 2001; Sobel, Tenenbaum, y Gopnik, 2004), ha sido criticada por ser demasiado especulativa y por carecer de una sólida evidencia experimental (Borsboom, 2002). En primer lugar, los pocos experimentos usados para justificar sus razonamientos parecen seleccionados “intencionadamente” para soportar su teoría y los resultados son interpretados sobre consideraciones *post hoc*. Por otro lado, Glymour parece no compensar equitativamente los *pros* y los *contras* de su propuesta. Por ejemplo, si su argumentación es correcta y cada uno de nosotros llevamos implementados en nuestro cerebro un sistema que nos permite “desenmascarar” la estructura causal del ambiente en términos de relaciones de dependencia e independencia probabilística, ¿cómo es posible que el aprendizaje formal de estos conceptos sea tan difícil? Estas críticas serían consistentes con los trabajos clásicos que informan de dificultades en la estimación y/o comprensión de probabilidades condicionales en términos causales (p. e. Kahneman, 2003; Kahneman y Tversky, 1973, 1996; Tversky y Kahneman, 1974, 1982). Sin embargo, este hecho podría deberse a los formatos de presentación de la información o a “artefactos metodológicos”. Así pues, la presentación de información probabilística de una manera más comprensiva para los participantes experimentales y enfatizando que los datos de aprendizaje son extraídos aleatoriamente de una muestra con distribución normal podría mejorar su aprendizaje/razonamiento y acercarlo más a lo establecido por la estadística (p. e. Gigerenzer, 1996; Seldmeier y Gigerenzer, 2001).

Psicología educativa

En el contexto educativo el modelo teórico propuesto por Glymour cobra importancia cuando se pretendan desarrollar programas formativos orientados al aprendizaje de conceptos estadísticos o científicos que impliquen causalidad. Por otro lado, el uso de redes bayesianas en educación ha aparecido bajo la insignia de los Sistemas de Tutores Inteligentes (STI). Según Xiang (2002), los STI son un tipo de *agente* inteligente caracterizado por la toma de decisiones y la realización de acciones sin la intervención de ninguna persona que monitoriza el funcionamiento del sistema. Esto supondría que el sistema tiene que, en primer lugar, evaluar el conocimiento que el aprendiz tiene sobre la materia. A continuación, ha de presentar información acorde a ese nivel (p. e. definiciones, material gráfico o auditivo, etc.). Después ha de evaluar si la persona ha asimilado este nuevo conocimiento y presentar nueva información acorde con lo que ha aprendido. En general, un STI genera un modelo del conocimiento que tiene el alumno y le presenta la información acorde a él (p. e., Conejo, Millán, Perez de la Cruz, y Trella, 2001; Martín y VanLehn, 1995; Mislevy y Gitomer, 1996; Padilla, Lara, y Márquez, 2001).

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 (*Off-Line Assessment of Expertise*) fue un sistema que evaluaba diferentes aspectos clave en la resolución de problemas de física. El sistema OLAE permitía recoger información en cinco tipos de actividades por medio de una computadora. Aquí se describe la actividad de *resolución de problemas*. Constaba de una interface 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, una representación gráfica donde el aprendiz tenía que señalar los fenómenos físicos implicados en el problema y una sección de respuestas donde el aprendiz tenía que expresar las fórmulas a utilizar y donde tenía que dar su respuesta. El sistema iba registrando las respuestas de los alumnos y generaba una red bayesiana que representaba las probabilidades asociadas a la utilización de cada fórmula física. El sistema OLAE, sin embargo, no funcionaba como un tutor inteligente ya que se limitaba a estimar la probabilidad relativa de utilización correcta de fórmulas y principios físicos. Además, no realizaba estos cálculos dinámicamente, o sea, a medida que el aprendiz daba una respuesta; si no que lo hacía a *posteriori*. OLAE supuso un avance en el campo de los STI porque mostró que el conocimiento que tiene una persona sobre un dominio puede evaluarse usando redes bayesianas de un modo estructurado.

Mislevy y Gitomer (1996) trabajaron con el programa HYDRIVE desarrollado por los Laboratorios Armstrong de las Fuerzas Aéreas de los Estados Unidos. HYDRIVE es un STI desarrollado para simular el funcionamiento de un avión de combate F-15. Un problema comienza con una secuencia de video donde un piloto describe algunas deficiencias en el funcionamiento de un aparato que está a punto de aterrizar o que ya ha aterrizado (por ejemplo, el chequeo rutinario del timón de aterrizaje no responde correctamente). La interface gráfica permite al estudiante llevar a cabo una tarea de solución de problemas revisando videos del aparato y permitiéndole actuar 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 estaría la ejecución global. En un nivel inferior estarían los tres tipos de conocimientos evaluados: conocimiento de sistema, conocimiento estratégico y conocimiento procedimental. En la tercera capa estarían los subcomponentes de cada tipo de conocimiento. La tercera capa sería una capa oculta donde se integra la información de las capas precedentes. La cuarta capa es donde se interpretan las acciones del estudiante en relación a los posibles escenarios que plantea el sistema. Todos los arcos se orientan desde los niveles jerárquicos más altos hacia los más bajos.

Uno de los avances más importantes que se han realizado en este campo ha sido el desarrollo e implementación de STI en internet. Conejo *et al.*, (2001) han desarrollado un sistema (SIETTE) que lleva a cabo procesos de diagnóstico y evaluación del aprendizaje por medio de internet. Aunque este sistema se basa en la Teoría de Respuesta al Ítem para

escalar el material didáctico y a los usuarios, el motor de inferencia que diagnostica el estado de conocimiento del aprendiz lo lleva a cabo una red bayesiana.

Psicología forense

El uso de las redes bayesianas en el ámbito forense cobra especial relevancia en la investigación policial (p. e., Garbolino y Taroni, 2002; Oatley y Ewart, 2002). Por ejemplo, Garbolino y Taroni (2002) proponen que las redes bayesianas podrían ser usadas en el ámbito forense como herramientas que ayudasen a sopesar evidencias (pruebas) de cara a la resolución de los casos. Una ilustración de esta utilidad de las redes bayesianas lo encontramos en el proyecto OVER. Este proyecto fue fruto de la colaboración entre la *West Midlands Police*, el *Centro de Sistemas Adaptativos, Computación y Tecnología*; y la *División de Psicología* de la Universidad de Sunderland. Como resultado de este trabajo conjunto, Oatley y Ewart, (2002, 2003) han informado de la construcción de un sistema informático que predice la probabilidad de robos en casas de una región metropolitana de Birmingham. Este programa está integrado por varios módulos que funcionan controlados por diferentes herramientas estadísticas. Sin embargo, el motor de inferencia principal es una red bayesiana que estima la probabilidad de asaltos a casas en función de un conjunto de variables (p. e. *modus operandi*, número de asaltos en una zona, etc.). Todo el trabajo está monitorizado por un plano de la ciudad donde se representa el riesgo de asalto con diferentes colores. A su vez, todos los análisis pueden ser exportados a formato Excel para que la policía pueda elaborar informes fácilmente.

Psicodiagnóstico

Los sistemas expertos han estado relacionados con las Ciencias de la Salud desde sus inicios; de hecho, el surgimiento de las redes neurales artificiales (que pueden considerarse como un tipo particular de sistema experto) estuvo asociado al diagnóstico médico (Quinlan, 1991). MICYN es considerado como el primer sistema experto utilizado en el diagnóstico médico de enfermedades infecciosas. Sin embargo, MICYN fue un sistema experto basado en reglas.

La primera red bayesiana orientada al diagnóstico médico surgió, al igual que MICYN, en la Universidad de Stranford y constaba de cinco nodos. La red ASIA (Cowell *et al.*, 1999) está compuesta por ocho nodos y está dirigida al diagnóstico diferencial de tuberculosis, bronquitis y cáncer de pulmón. Sin embargo, estas redes "primigenias" no estaban orientadas a resolver problemas reales, sino que, más bien pretendían ser un ejemplo de la potencialidad de las redes bayesianas en el diagnóstico médico (Díez, 1998). Las primeras redes bayesianas con aplicación clínica real se desarrollaron en la Universidad de Aalborg y tuvieron como objetivo el diagnóstico de enfermedades musculares (Hugin Expert, 2004) y la monitorización de insulina en pacientes diabéticos.

En Psicología, al igual que en medicina, el diagnóstico es un proceso de toma de decisiones donde el profesional, o profesionales, tiene que poner en marcha una búsqueda activa de evidencias y actuar acorde a ellas (Godoy, 1996). Por ello, se han utilizado numerosas herramientas estadísticas para minimizar las posibilidades de error en el diagnóstico. Por ejemplo, la teoría de la utilidad esperada junto a la maximización de la utilidad multiatributo y el uso de la regla de Bayes para calcular probabilidades condicionales (Esteve, 1996a, 1996b) han sido usadas con la idea de apoyar a los profesionales en su toma de decisiones.

MENTOR (Mani, McDermott y Valtorta, 1997) fue construido en la Universidad de California con la colaboración de la Fundación Kaiser y su objetivo era la predicción y el diagnóstico precoz de retraso mental en bebés. Para la estimación de los parámetros del modelo se usó una base de datos obtenida en un estudio sobre mujeres embarazadas y sus hijos que contenía datos desde 1959 hasta la década de 1980. El resultado de este estudio fue una base de datos que contenía registros de, aproximadamente, 6000 niños con edades comprendidas entre los 5 y los 9 años y unas 3000 madres. No obstante, los autores encontraron algunos problemas en el desarrollo de su trabajo. En primer lugar, observaron que las redes generadas por diferentes conjuntos de datos no eran consistentes con las reglas cronológicas. Por otro lado, los parámetros obtenidos con una muestra de datos recogidos desde 1959 hasta 1974 no son representativos, o generalizables, de lo que sucede en la actualidad ya que el cuidado a recién nacidos ha evolucionado mucho durante los últimos años.

Por su parte, el Instituto de Investigación en Salud Mental de Victoria (Melbourne, Australia, www.mhri.edu.au) es una organización que está trabajando con la tecnología de las redes bayesianas para comprender, tratar y prevenir trastornos psicológicos como la esquizofrenia, los trastornos del estado de ánimo o el Alzheimer. La idea de Glymour (2001) de reinterpretar todos los modelos neuropsicológicos clásicos bajo la idea de las redes bayesianas podría ser un paso importante de cara al diagnóstico y tratamiento de pacientes con dolencias de esta índole.

No obstante, todas las iniciativas orientadas a desarrollar modelos, sistemas o herramientas que pretendan ser de utilidad en la toma de decisiones clínica han de tener en cuenta los problemas que acarrear. Por ejemplo, ¿qué habría que modelar, la toma de decisiones que lleva a cabo un experto clínico (lo *descriptivo*); o lo que tendría que hacer el experto clínico si funcionase como un mecanismo estadístico ausente de sesgos (lo *normativo*)? Algunos autores (p. e., Godoy, 1996; Raiffa, 1994) proponen que los mecanismos, como las redes bayesianas, de optimización ante la toma de decisiones clínica tendrían que ser meros *apoyos* y no sustitutos a estos profesionales.

Investigación básica

Como se ha sugerido anteriormente, el área de trabajo más prolífica en el uso de las redes bayesianas en los últimos años ha sido la del aprendizaje causal (p. e., Gopnik *et al.*, 2004; Gopnik y Schulz, 2004; Gopnik *et al.*, 2001; Lagnado y Sloman, 2004; Sobel *et al.*, 2004; Waldman y Hagmayer, 2005). A este nivel, se están desarrollando estudios que ponen a prueba la hipótesis de que las personas representamos el conocimiento causal de manera análoga a como lo hace una red bayesiana. En este sentido, se pretenden desarrollar modelos del aprendizaje causal que permitan integrar la dimensión observacional e instrumental del aprendizaje en un único modelo. Esto es, obtener un modelo que integre los principios del aprendizaje clásico y del aprendizaje instrumental en la percepción de la causalidad.

Por otro lado, la investigación básica en Psicología del lenguaje podría dar lugar a un mejor entendimiento de los procesos psicológicos implicados en el aprendizaje, desarrollo y funcionamiento de las habilidades lingüísticas. Daniel Jurafsky (1996) propuso una conexión entre la estructura del lenguaje en términos de probabilidades condicionadas y ciertos aspectos léxicos. Aunque en este trabajo tan sólo se hace una ligera referencia a la utilidad de las redes bayesianas en el estudio del lenguaje a nivel psicológico, en estudios posteriores (Narayan y Jurafsky, 1998, 2002) se han replicado algunos resultados experimentales usando modelos de red bayesiana. Estos hallazgos también han tenido utilidad práctica. Por ejemplo, en un estudio orientado a comparar la habilidad en la recuperación de información electrónica entre redes neurales artificiales y redes bayesianas, las redes bayesianas ofrecieron mejores resultados (Martinez, Díaz, Rivas y Ureña, 2003). En este contexto y como sugiere Glymour (2001), las redes bayesianas podrían servir como una herramienta normativa para el desarrollo/actualización de los modelos clásicos de procesamiento de lenguaje (Ellis y Young, 1988).

Una de los campos donde la potencialidad de las redes bayesianas no ha sido explotada es la neurociencia cognitiva. El uso de este tipo de herramientas estadísticas podría encuadrarse dentro de uno de los grandes flancos que constituyen a la Neurociencia Cognitiva: la simulación computacional (Kandel, Schwartz y Jessell, 1995/1997; Parasuraman, 2000). Podríamos destacar tres áreas en el contexto de la neurociencia cognitiva donde las redes bayesianas podrían ser de utilidad como herramientas orientadas al modelado: en los mecanismos neuroquímicos de la incertidumbre, en la caracterización de los circuitos neurales encargados de la toma de decisiones y en la modelización de los fenómenos psicológicos básicos.

El estudio del sustrato neuroquímico de la incertidumbre ha comenzado a usar la estadística bayesiana y el uso de redes bayesianas podría servir para aumentar el conocimiento que tenemos de este tipo de fenómenos (Cohen, y Aston-Jones, 2005). Por ejemplo, Yu y Dayan (2005a) han propuesto un modelo basado en la estadística bayesiana capaz de

modelar el comportamiento de nuestra atención ante diferentes tipos de incertidumbre en relación con dos neurotransmisores: la noradrenalina y la acetilcolina. A grandes rasgos, el modelo propone que existen dos sistemas de incertidumbre que interactúan entre sí: la *incertidumbre esperada*, mediada por la acetilcolina (ACh), y la *incertidumbre no esperada*, donde el neurotransmisor implicado es la noradrenalina (NA) (Cohen y Aston-Jones, 2005; Yu y Dayan, 2003, 2005a, 2005b). La incertidumbre esperada se refiere al conocimiento relativo que tenemos a la ocurrencia de eventos en nuestro entorno natural; esto es, a la probabilidad que asociamos a la ocurrencia de los eventos ambientales. Por su parte, la incertidumbre no esperada se refiere a cambios bruscos, repentinos e inesperados que se producen sobre nuestra percepción incierta del entorno. El modelo de Yu y Dayan, que es consistente con los datos fisiológicos, farmacológicos y conductuales reportados hasta el momento; tiene la virtud de proponer dos sistemas de incertidumbre mediados por diferentes neurotransmisores.

Sin embargo, no tendríamos un conocimiento ajustado de los procesos neurocognitivos implicados en la toma de decisiones si solamente nos centrásemos en estudiar los neurotransmisores implicados en la elección de alternativas. Tendríamos que complementar esta información con datos relativos a los circuitos neurales que llevan a cabo tal tarea. Recientemente se están llevando a cabo un vasto conjunto de investigaciones que pretenden identificar los circuitos neurales implicados en la toma de decisiones. Esto ha dado lugar a una nueva línea de investigación que bajo la etiqueta de *neuroeconomía* está integrando las ideas de las Neurociencias, la Psicología y las Ciencias Económicas con el objetivo de caracterizar los sustratos anatómico-funcionales de la toma de decisiones (p. e., Lafuente y Romo, 2003; Lee, 2005; Platt, 2004).

Por último, cabe destacar que las redes bayesianas podrían ser una herramienta normativa de referencia a la hora de desarrollar modelos neuropsicológicos de los procesos psicológicos básicos. En este sentido, la *hipótesis de la codificación bayesiana* podría considerarse como un punto de arranque ante el modelado de la percepción (Knill y Pouget, 2004) que podría trasvasarse a otros procesos cognitivos básicos como la memoria o la atención.

Psicología de las Organizaciones

Como herramienta estadística, las redes bayesianas podrían usarse en cualquier tópico investigado por la Psicología Social. Sin embargo, los estudios que se han realizado por el momento se han centrado en aplicaciones orientadas al entorno empresarial. En el mundo empresarial la toma de decisiones en términos estadísticos juega un papel fundamental. El interés por la estadística aplicada a la toma de decisiones en el contexto empresarial se ha acentuado gracias a la evolución de los estilos de dirección acaecida en los últimos años (p. e., León, 1994). Por ello, se ha pasado de un estilo

de dirección centrado en la experiencia del empresario a un modo de dirección basado en indicadores estadísticos.

Una de las facetas donde las redes bayesianas han mostrado ser herramientas útiles ha sido en los estudios de mercado (Nadkarni y Shenoy, 2001, 2004; Concejero, 2004). Por ejemplo, Baesens *et al.* (2004) encontraron que una red bayesiana con cuatro nodos y seis arcos (frente a otros modelos más complejos y menos parsimoniosos que implicaban más variables) clasificaba a los clientes de una empresa en relación a la perspectiva de compras a largo plazo. Encontraron que la evolución de las compras de los clientes de una empresa cliente se relacionaba (positivamente) con el porcentaje máximo de productos que adquiriría de una misma categoría; y (negativamente) con la contribución inicial y el número total de artículos comprados.

El vínculo más claro del mundo empresarial y la Psicología aplicada podríamos encontrarlo en la gestión y dirección de recursos humanos. HumRRO (Human Resources Research Organization, www.humrro.org) se fundó en 1951 y es un ejemplo de la utilización de las redes bayesianas en la consultoría de recursos humanos. De hecho, como se ha hecho notar por Maciá *et al.* (1990), la selección de personal es un buen ejemplo de un proceso de toma de decisiones secuencial dentro de una organización. En este sentido los diagramas de influencia serían apropiados para ir guiando el proceso de selección de personal en función de ciertas restricciones.

Conclusiones

Las redes bayesianas no son una herramienta conocida en Psicología, y menos aún en nuestro contexto. Retomando la doble perspectiva de Edwards (1998), se pueden plantear diferentes cuestiones que incidan tanto en el ámbito más de mercado o profesional de la Psicología y por otro lado los aspectos de investigación.

Desde el ámbito profesional, un interés primordial lo tiene el diagnóstico y el tratamiento psicológico. El desarrollo de sistemas expertos probabilísticos podría mejorar, en casos particulares, el diagnóstico y el tratamiento psicológico; y, en términos generales, podrían desarrollarse mejores modelos de los trastornos psicológicos que conllevarían a una comprensión más profunda de los mismos. Por otro lado, la ya comentada elicitación de probabilidades para la creación de sistemas expertos orientados a la toma de decisiones empresarial lleva implícito el *handicap* de los sesgos en el juicio bajo incertidumbre que la Psicología lleva estudiando desde hace tres décadas. Por tanto, si este mercado evolucionase, la Psicología tendría a su disposición otro nicho laboral donde ubicarse. Finalmente, y como se ha contado a lo largo del artículo, otros campos aún emergentes, como la Psicología Forense o determinados aspectos de la Psicología de la Intervención Social y Educativa, pueden ser campos de aplicación profesional de estas herramientas.

En el plano investigador, podemos esperar que vayan cobrando protagonismo en los próximos años. Ya sea como

herramientas de análisis de datos o como modelos normativos comparables a datos experimentales, las redes bayesianas aportan una gran ventaja para la Psicología: integrar en un único paradigma la dimensión cualitativa y cuantitativa de un problema.

Es previsible que la comprensión de los procesos de aprendizaje causal se optimicen gracias a la referencia que aportan las redes bayesianas. Por otro lado, creemos que las redes bayesianas pueden ayudar a mejorar los modelos teóricos que se han planteado en todas las áreas de la Psicología

ya que el formalismo de la teoría de probabilidades y de grafos aportan un marco de referencia que hace posible el planteamiento y el contraste de hipótesis.

Por último, nuevamente, gracias a la potencia modeladora que tienen estas herramientas, hay que destacar que las redes bayesianas podrían usarse para generar más y mejores modelos del funcionamiento de las organizaciones, los grupos o cualquier agregado social que sea objeto de estudio para la Psicología.

Referencias

- Adarraga, P., y Zaccagnini, J. L. (1988). Sistemas expertos y Psicología cognitiva: una visión general. *Estudios de Psicología*, 36, 97-127.
- Adarraga, P., y Zaccagnini, J. L. (1992). DAI: a knowledge-based system for diagnosing autism. A case study on the application of artificial intelligence to psychology. *European Journal of Psychological Assessment*, 8, 25-46.
- Alonso, D., y Tubau, E. (2002). Inferencias bayesianas: una revisión. *Anuario de Psicología*, 33, 25-47.
- Baesens, B., Verstraeten, G., Van den Poel, D., Egmont-Petersen, M., Van Kenhove, P., y Vanthienen, J. (2004). Bayesian classifiers for identifying the slope of the customer life cycle of long-life customer. *European Journal of Operational Research*, 156, 508-523.
- Borsboom, D. (2002). [Revisión del libro *The mind's arrows. Bayes nets and graphical causal models in psychology*]. *Acta Psychologica*, 111, 355-357.
- Carlson, N. R. (2000). *Fisiología de la conducta* (3ª ed.). Barcelona: Ariel. (Trabajo original publicado en 1993).
- Cohen, J. D., y Aston-Jones, G. (2005, Julio 28). Decision amid uncertainty. *Nature*, 436, 471-472.
- Concejero, P. (2004). *Comparación de los modelos de curvas ROC para la evaluación de procedimientos estadísticos de predicción en investigación de mercados*. Tesis doctoral no publicada, Facultad de Psicología, Universidad Complutense de Madrid, Madrid.
- 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.
- 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.
- Diez, J. F. (1998). Aplicaciones de los modelos gráficos probabilistas en medicina. En A. Gámez y J. M. Puerta (Eds.), *Sistemas expertos probabilísticos* (pp. 239-263). Cuenca: Servicio de Publicaciones de la Universidad de Castilla-La Mancha.
- Dixon, J. R. (1970). *Introducción a la probabilidad. Texto programado*. México: Limusa-Wiley. (Trabajo original publicado en 1967).
- Edwards, W. (1998). Hailfinder. Tools for and experiences with bayesian normative modelling. *American Psychologist*, 53, 416-428.
- Edwards, W., y Fasolo, B. (2001). Decision technology. *Annual Review of Psychology*, 52, 581-606.
- Ellis, A., y Young, A. W. (1988). *Human cognitive neuropsychology*. Hove: Lawrence Erlbaum.
- Esteve, M. R. (1996a). Teoría de la decisión (II): el teorema de Bayes. En A. Godoy (Ed.), *Toma de decisiones y juicio clínico. Una aproximación psicológica* (pp. 191-247). Madrid: Pirámide.
- Esteve, M. R. (1996b). Teoría de la decisión (I): teoría de la utilidad esperada. En A. Godoy (Ed.), *Toma de decisiones y juicio clínico. Una aproximación psicológica* (pp. 175-190). Madrid: Pirámide.
- Gámez, J. A. (1998). Abducción en modelos gráficos. En J. A. Gámez y J. M. Puerta (Eds.), *Sistemas expertos probabilísticos* (pp. 79-111). Cuenca: Servicio de Publicaciones de la Universidad de Castilla-La Mancha.
- Garbolino, P., y Taroni, F. (2002). Evaluation of scientific evidence using bayesian networks. *Forensic Science International*, 125, 149-155.
- Gigerenzer, G. (1996). On narrow norms and vague heuristic: a reply to Kahneman and Tversky. *Psychological Review*, 103, 592-596.
- Glymour, C. (2001). *The mind's arrows. Bayes nets and graphical models in psychology*. Cambridge, MA: MIT Press.
- Godoy, A. (1996). *Toma de decisiones y juicio clínico. Una aproximación psicológica*. Madrid: Pirámide.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., y Danks, D. (2004). A theory of causal learning in children: causal and bayes nets. *Psychological Review*, 111, 3-32.
- Gopnik, A., Sobel, D. M., Schulz, L., y Glymour, C. (2001). Causal learning mechanisms in very young children: two, three and four-years-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, 37, 620-629.
- Gopnik, A., y Schulz, L. (2004). Mechanisms of theory formation in young children. *Trends in Cognitive Sciences*, 8, 371-377.
- Harary, F. (1969). *Graph theory*. Philippines: Westview Press.
- Heckerman, D. (1995). *A tutorial with bayesian networks* (Tech. Rep. MS-TR-95-06). Redmon, WA: Microsoft Research.
- Herskovits, E. H., y Dagher, A. P. (1997). *Applications of bayesian networks to health care* (Tech. Rep. NSI-TR-1997-02). Baltimore, MD: Noetic Systems.
- Hothersall, D. (1997). *Historia de la Psicología* (3ª ed.). México: McGraw-Hill. (Trabajo original publicado en 1995)
- Huete, J. F. (1998). Sistemas expertos probabilísticos: modelos gráficos. En J. A. Gámez y J. M. Puerta (Eds.), *Sistemas expertos probabilísticos* (pp. 1-40). Cuenca: Servicio de Publicaciones de la Universidad de Castilla-La Mancha.
- Hugin Expert. (2004). *Tutorials. Using learning facilities*. Descargado el 21 de Septiembre, 2005, desde <http://www.hugin.com/tutorials/>
- Jansen, R., Yu, H., Greenbaum, D., Kluger, Y., Krogan, N. J., Chung, S., et al. (2003, Octubre 17). A bayesian network approach for predicting protein-protein interactions from genomic data. *Science*, 302, 449-453.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137-194.
- Kadie, C. M., Hovel, D., y Horvitz, E. (2001). *MSBNx: A component-centric toolkit for modelling and inference with bayesian networks* (Tech. Rep. MSR-TR-2001-67). Redmond, WA: Microsoft Research.
- Kahneman, D. (2003). Experiences of collaborative research. *American Psychologist*, 58, 723-730.
- Kahneman, D., y Tversky, A. (1973). On the psychology of prediction. *American Psychologist*, 58, 723-730.
- Kahneman, D., y Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103, 582-591.
- Kandel, E. R., Schwartz, J. H., y Jessell, T. M. (1997). *Neurociencia y conducta*. Madrid: Prentice Hall. (Trabajo original publicado en 1995).
- Knill, D. C., y Pouget, A. (2004). The bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27, 712-719.
- Lafuente, V., y Romo, R. (2003). Decisions arising from opposing views. *Nature Neuroscience*, 6, 792-793.
- Lagnado, D. A., y Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 856-876.
- Lee, D. (2005). Neuroeconomics: making risky choices in the brain. *Nature Neuroscience*, 8, 1129-1130.
- León, O. G. (1994). *Análisis de decisiones. Técnicas situacionales aplicables a directivos y profesionales*. Madrid: McGraw-Hill.

- López, J., García, J., y De la Fuente, E. I. (2005, Septiembre). Las redes bayesianas en Psicología. Comunicación presentada en el IX Congreso de Metodología de las Ciencias Sociales y de la Salud. Granada
- Maciá, M. A., Barbero, M. I., Pérez-Llantada, M. C., y Vila, E. (1990). *Psicología y teoría de la decisión: aplicaciones*. Madrid: Universidad Nacional de Educación a Distancia.
- Mani, S., McDermott, S., y Valtorta, M. (1997). MENTOR: a bayesian model for prediction of mental retardation in newborns. *Research in Developmental Disabilities, 18*, 303-318.
- Martin, J., y VanLehn, K. (1995). Student assessment using bayesian nets. *International Journal of Human-Computer Studies, 42*, 575-591.
- Martínez, F., Díaz, M. C., Rivas, V. M., y Ureña, L. A. (2003, Septiembre). *Aplicación de redes neuronales y redes bayesianas en la detección de multipalabras para tareas IR*. Artículo presentado en las II Jornadas de Tratamiento y Recuperación de la Información. Madrid.
- 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.
- Mislevy, R. J., y Gitomer, D. H. (1996). The role of probability-based inference in an intelligent tutoring system. *User-Mediated and User-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.
- Narayan, S., y Jurafsky, D. (1998, Agosto). *Bayesian models of human sentence processing*. Artículo presentado en la XX Annual Meeting of the Cognitive Science Society. Madison, WI.
- Narayan, S., y Jurafsky, D. (2002). A bayesian model predicts human parse preference and reading times in sentence processing. En T. G. Dietterich y Z. Ghahramani (Eds.), *Advances in Neural Information Processing 14* (pp. 59-65). Cambridge, MA: MIT Press.
- 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.
- Oatley, G. C., y Ewart, B. W. (2002, Septiembre). *Constructing a bayesian belief network to determine likelihood of burglary*. Poster presentado en la V International Conference on Forensic Statistics. Venecia.
- Oatley, G. C., y Ewart, B. W. (2003). Crimes analysis software: 'pins in maps', clustering and bayes net prediction. *Expert Systems with Applications, 25*, 569-588.
- Padilla, J. H., Lara, A., y Márquez, P. R. (2001, Octubre). Sistema inteligente para la enseñanza de las matemáticas. Artículo presentado en el XXIII Congreso Internacional de Ingeniería Electrónica ELECTRO. México
- Parasuraman, R. (2000). The attentive brain. Issues and prospects. En R. Parasuraman (Ed.), *The attentive brain* (pp. 3-15). Cambridge, MA: MIT Press.
- Pearl, J. (2001). *Bayesian networks, causal inference and knowledge discovery* (Tech. Rep. R-281). Los Angeles: University of California.
- Pérez, M. L., y Martín, Q. (2003). *Aplicaciones de las redes neuronales artificiales a la estadística*. Madrid: Muralla-Hespérides.
- Platt, M. L. (2004). Unpredictable primates and prefrontal cortex. *Nature Neuroscience, 7*, 319-320.
- Quinlan, P. (1991). *Connectionism and psychology: a psychological perspective on new connectionist research*. Hertfordshire: Cambridge University Press.
- Raiffa, H. (1994). The prescriptive orientation of decision making: a synthesis of decision analysis, behavioral decision making, and game theory. En S. Ríos (Ed.), *Decision theory and decision analysis* (pp. 3-13). Norwell, MA: Kluwer Academic Press.
- Ríos, S. (1995). *Modelización*. Madrid: Alianza Universidad.
- Rodríguez, R. J. (1996). Modelos cognoscitivos para una filosofía contemporánea de la mente. *Revista de Filosofía de la Universidad de Costa Rica, 34*, 423-432.
- Ronald, G. (1988). *Graph theory*. Menlo Park, CA: Benjamin/Cummings.
- Seldmeier, P., y Gigerenzer, G. (2001). Teaching bayesian reasoning in less than two hours. *Journal of Experimental Psychology, 130*, 380-400.
- Serrano, J. (2003). *Iniciación a la estadística bayesiana*. Madrid: Muralla-Hespérides.
- Sobel, D. M., Tenenbaum, J. B., y Gopnik, A. (2004). Children's causal inferences from indirect evidence: backwards blocking and bayesian reasoning in preschoolers. *Journal of Experimental Psychology, 130*, 380-400.
- Spirtes, P., Glymour, C., y Scheines, R. (2000). *Causation, prediction and search* (2ª ed.). Cambridge, MA: MIT Press.
- SPSS Inc. y Recognition Systems Inc. (1997). *Neural connection 2.0. User's guide*. Chicago: Autores.
- Tutte, W. T. (1984). *Graph theory*. Cambridge: Cambridge University Press.
- Tversky, A., y Kahneman, D. (1974, Septiembre 27). Judgement under uncertainty: heuristic and biases. *Science, 185*, 1124-1131.
- Tversky, A., y Kahneman, D. (1982). Causal schemas in judgements under uncertainty. En D. Kahneman, P. Slovic y A. Tversky (Eds.), *Judgement under uncertainty: heuristic and biases* (pp. 117-128). New York: Cambridge University Press.
- Waldman, M. R., y Hagmayer, Y. (2005). Seeing versus doing: two models of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 216-227.
- Xiang, Y. (2002). *Probabilistic reasoning in multiagent systems. A graphical models approach*. Cambridge: Cambridge University Press.
- Yu, A., y Dayan, P. (2003). Expected and unexpected uncertainty: ACh and NE in the neocortex. En S. T. S. Becker y K. Obermayer (Eds.), *Advances in Neural Information Processing Systems 15* (pp. 157-164). Cambridge, MA: MIT Press.
- Yu, A., y Dayan, P. (2005a). Uncertainty, neuromodulation, and attention. *Neuron, 46*, 681-692.
- Yu, A., y Dayan, P. (2005b). Inference, attention and decision in a bayesian neural architecture. En L. K. Saul y L. Bottou (Ed.), *Advances in Neural Information Processing Systems 17* (pp. 1577-1584). Cambridge, MA: MIT Press.
- Zaccagnini, J. L. (1994). Introducción al campo de la inteligencia artificial. En P. Adarraga y J. L. Zaccagnini (Ed.), *Psicología e inteligencia artificial* (pp. 13-36). Madrid: Trotta.

(Artículo recibido: 4-9-2006; aceptado: 3-10-2007)