

## Aprendizaje implícito y explícito: ¿dos procesos diferentes o dos niveles de abstracción?

Elisabet Tubau  
Universitat de Barcelona  
Joan L. Moliner  
Universitat Autònoma de Barcelona

*En los estudios centrados en analizar las diferencias entre el aprendizaje implícito y el explícito es frecuente observar definiciones que relacionan estas dos formas de aprendizaje con dos procesos diferentes: procesamiento consciente y no consciente de la información. En este artículo abordamos la distinción explícito versus implícito desde la perspectiva de distintos niveles de abstracción de reglas. Consideramos que tanto la evidencia experimental como las aportaciones de los modelos conexionistas nos permiten pensar en la distinción implícito versus explícito relacionada con las diferencias entre representaciones específicas basadas en covariaciones o ejemplares y representaciones basadas en reglas más abstractas. Los resultados de algunos experimentos así como el análisis de simulaciones mediante redes neuronales han posibilitado la identificación de algunas características asociadas a estos niveles de abstracción.*

*Palabras clave: procesamiento implícito y explícito, inducción de reglas y redes neuronales, niveles de abstracción.*

*Regarding the differences between implicit and explicit learning reported in the literature, the definitions frequently relate these two kinds of learning with two different processes: conscious and unconscious information processing. In this paper we focus on the distinction between explicit and implicit from the point of view of different levels of rule abstraction. We think that evidence from experiments and connectionist models permits us to relate the explicit-implicit distinction to the differences between specific representations based on covariations or exemplars and those based on more abstract rules. Both experimental results and connectionist simulations have made it possible to identify certain features associated with these levels of abstraction.*

*Key words: Implicit and explicit processing, Rule induction and neural networks, Levels of abstraction.*

Es frecuente establecer una distinción entre dos procesos de aprendizaje y pensamiento. En algunos de los aprendizajes que realizamos aparentemente somos conscientes del curso de pensamiento que nos lleva a tomar ciertas decisiones; podemos describirlo mediante algún tipo de estructura lógica y, en consecuencia, las decisiones van acompañadas de un conocimiento fácilmente verbalizable. Por el contrario, en otras situaciones nuestras decisiones parecen basarse en la intuición o en el azar; si la decisión es la correcta pensamos, a menudo, que debe tratarse de una casualidad. No obstante, como muestran numerosas investigaciones también en este caso puede haber existido un tipo de aprendizaje aunque, a diferencia del anterior, no se ajusta aparentemente a ninguna lógica. No podemos explicar las razones que nos han llevado a tomar una determinada decisión y, por lo tanto, el conocimiento fruto de este tipo de aprendizaje es considerado implícito y no verbalizable.

Froufe (1996) realiza una ilustrativa síntesis del pensamiento de numerosos investigadores centrados en analizar las diferencias entre el aprendizaje implícito y explícito. Sin embargo, como este mismo autor señala, al igual que otros artículos recientes (p.e. Shanks y St John, 1994), el principal problema de esta distinción reside en el papel atribuido a la conciencia, entendida como experiencia subjetiva, como frontera entre lo explícito y lo implícito. Una de las cuestiones que emergen de estos trabajos es ¿puede existir aprendizaje sin conciencia de nuevas relaciones capaces de producir cambios en las decisiones tomadas a nivel consciente?

Al hablar de aprendizaje sin conciencia se excluyen los cambios que podemos observar (y medir) del comportamiento como consecuencia de la práctica. Mediante el entrenamiento de una habilidad, las acciones (motoras o mentales) se realizan con mayor precisión y velocidad sin que exista conciencia del verdadero efecto producido por la repetición. En este sentido, nos centraremos en analizar especialmente los procesos de adquisición de nuevas estructuras de conocimiento más que la mejora procedimental de una habilidad. Una de las cuestiones principales que sigue debatiéndose en la literatura sobre este tema es la que se refiere al continuum específico-abstracto del conocimiento adquirido en condiciones implícitas de aprendizaje (p.e. Neal y Hesketh, 1997) y si es posible diferenciar tal modo de aprendizaje del procesamiento más explícito (Shanks y St. John, 1994).

Podemos observar que esta distinción ha estado presente, de forma más o menos explícita, en numerosos tratados del conocimiento humano; desde Sócrates (Dreyfus y Dreyfus, 1986) a investigadores más recientes como Luria (1961), Skinner (1969) o Smolensky (1988). Por ejemplo, Skinner (1974/1977) fue uno de los primeros autores en proponer una clara distinción entre dos tipos de comportamiento y, al mismo tiempo, dos tipos de aprendizaje: comportamiento moldeado por las contingencias y comportamiento gobernado por reglas (o aprendizaje de contingencias versus aprendizaje de reglas verbales). Veámoslo en un ejemplo:

«La persona que está aprendiendo a conducir un automóvil responde al comportamiento verbal de la persona que está sentada a su lado; arranca, frena, hace los cam-

bios, hace señales, etc., cuando se le dice que lo haga. Estos estímulos verbales pueden ser inicialmente directrices, pero se convierten en instrucciones si la ayuda verbal sólo se da cuando se necesita. Entonces, naturalmente, las contingencias naturales, no verbales, de la conducción del automóvil se imponen en el comportamiento del conductor. Aprender a conducir únicamente por medio de la exposición a esas contingencias requeriría mucho tiempo. El aprendiz de conductor tendría que descubrir lo que sucede cuando mueve la palanca de cambios, gira el volante, presiona el acelerador, aplica el embrague, etc., y todo esto con mucho peligro para él. Cuando sigue instrucciones evita verse expuesto a muchas de estas contingencias y eventualmente se comporta como lo hace el mismo instructor.» (p. 114).

De esta manera las instrucciones, leyes o reglas verbales describen contingencias que se dan en el ambiente (o cuando se interacciona con él). En la mayoría de los casos el aprendizaje de reglas es más rápido que aprender a partir de las contingencias descritas en la regla. Sin embargo, la mejora procedimental de la habilidad requiere complementar las instrucciones verbales con la experiencia directa. El aprendiz deberá adaptar su comportamiento a las contingencias variables del entorno, con el fin de comportarse progresivamente con mayor pericia. En este sentido, el comportamiento del experto está moldeado por las contingencias relevantes para la ejecución de la tarea y por las reglas que le han sido transmitidas o que ha «descubierto».

Esta distinción implica dos niveles de actuación y, por consiguiente, dos niveles de representación. En un nivel tenemos la representación del estímulo o del ejemplar que tiene un papel activador de la representación o respuesta asociada. Tal como ha señalado recientemente Shanks (1995):

«En una situación de aprendizaje asociativo, el entorno (o el experimentador) organiza una relación de contingencia entre eventos, permitiendo a la persona predecir uno desde la presencia de otros. Los eventos predictivos o las propias acciones del sujeto serán señales externas (claves). Las relaciones predictivas pueden ser de dos tipos; causales o estructurales. La forma más obvia de relación es la causal, en la que la ocurrencia de un evento o conjunto de eventos es seguida, después de un intervalo de tiempo, por otro/s evento/s. [...] Por el contrario, podemos decir que una relación es estructural cuando un organismo aprende a predecir un rasgo o atributo de un objeto o evento a partir de la presencia de otros rasgos que regularmente coocurren con él [...] La capacidad de clasificar objetos es otro ejemplo de la predicción estructural.» (pág. 2).

El otro nivel de representación, el de las reglas, es mucho menos comprendido. Tan sólo algunos autores apuntan a señalar la existencia de distintos niveles de abstracción. ¿Cuál es el nivel de abstracción correspondiente a una regla? Según Nisbett (1993) la dificultad de responder a esta cuestión se debe al rechazo generalizado de las reglas, y la abstracción en general, que se observa en la psicología actual:

«La psicología del siglo XX ha tenido un fuerte prejuicio contra la abstracción, es decir, contra el punto de vista de que el mundo es comprendido por medio de reglas que trascienden la percepción de estímulos físicos particulares o la comprensión de un dominio de eventos relacionados.» (pág. 1).

Según este autor uno de los errores fundamentales de las últimas teorías sobre el razonamiento y la resolución de problemas es considerar que estas actividades, tradicionalmente vistas como de alto nivel, pueden explicarse a partir de mecanismos analógicos, es decir, a partir de asociaciones basadas en la semejanza entre problemas o situaciones (razonamiento basado en casos).

Es cierto. Hasta finales de la década de los ochenta ha existido un interés considerable en unificar estos dos niveles de representación: reglas y ejemplares, considerando el nivel representado por las reglas tan sólo como una forma de describir el comportamiento de un sistema capaz de relacionar ejemplares asociados entre ellos por el número de atributos compartidos. Es decir, y en términos de Schanks, conjuntos de ejemplares que mantienen ciertas relaciones estructurales. De esta forma, el mecanismo asociativo seguiría siendo un único y común nivel explicativo, tanto para explicar la organización de ejemplares en la memoria, como para explicar el funcionamiento aparentemente controlado por reglas. Este punto de vista sigue siendo defendido por algunos de los investigadores en el área del conexionismo (Cleeremans y McClelland, 1991; Cleeremans, 1993; Keele y Jennings, 1992).

Como señalábamos anteriormente, a finales de los años 80, y sobre todo en la década de los 90, adquiere fuerza una línea de investigación centrada en diferenciar dos tipos de aprendizaje: explícito o selectivo e implícito o no selectivo (Berry y Broadbent, 1988; Reber, 1989). No es casual que algunos autores centrados en esta línea hayan asociado el aprendizaje explícito con la inducción de reglas conceptuales (Mathews *et al.*, 1989), mientras que el aprendizaje implícito se explicaría en términos de percepción de covariaciones (Lewicki, 1986) o memoria de ejemplares (Mathews *et al.*, 1989). Sin entrar en el debate existente sobre la realidad de esta distinción (ver p.e. Shanks y St. John, 1994), es evidente que esta línea de investigación ha vuelto a resucitar y poner en consideración la antigua separación, ya señalada, entre reglas o representaciones abstractas y ejemplares o representaciones basadas en asociaciones.

El objetivo de este artículo es analizar la distinción entre procesos implícitos y explícitos de aprendizaje desde la perspectiva de distintos formatos de representación del conocimiento. Quizás sea difícil diferenciar ambos modos de aprendizaje en términos del nivel de consciencia que acompaña a cada uno, pero creemos que podemos encontrar algunas regularidades en relación al nivel de abstracción del conocimiento en situaciones implícitas y explícitas de aprendizaje. Para ello seguiremos dos caminos complementarios. Por un lado, analizaremos algunos de los experimentos más relevantes en el área del aprendizaje implícito versus explícito con el fin de considerar posibles interpretaciones alternativas. Por otro lado, introduciremos algunas aportaciones del enfoque conexionista que pueden ser relevantes para esta discusión, tanto al considerar el nivel de abstracción del conocimiento como algunas de las limitaciones de los mecanismos asociativos.

## Tareas de aprendizaje implícito

Como tareas de aprendizaje implícito podríamos identificar una numerosa variedad. Algunas de ellas tratan de demostrar la asociación no consciente de de-

terminados eventos como las que se dan en situaciones de condicionamiento evaluativo, en las que estímulos neutros (p.e. fotografías) acaban adquiriendo la carga afectiva asociada al estímulo apetitivo o aversivo con que han sido emparejados (ver p.e. Baeyens *et al.*, 1992).

Más directamente relacionado con el campo del aprendizaje implícito, Lewicki (1986) publicó varios experimentos sobre detección no consciente de covariaciones. En uno de ellos, por ejemplo, presentaba a los sujetos fotografías de mujeres las cuales aparentemente sólo variaban de forma sistemática en la longitud del pelo. En algunas condiciones el hecho de tener el pelo largo iba acompañado por descripciones de personalidad que enfatizaban la «amabilidad» de la persona, mientras que el tener el pelo corto se asociaba a la «eficiencia» de la persona. En otras condiciones la relación se establecía a la inversa y en otras variaban las características de personalidad.

Posteriormente, Lewicki observó que los sujetos tendían a aceptar descripciones de nuevas fotografías de acuerdo con la covariación observada y sin conocimiento consciente de tal relación. Es decir que, según estos estudios, parece posible la detección no consciente de simples covariaciones. ¿Qué es lo que demuestran las investigaciones sobre aprendizaje implícito de relaciones más complejas? En este apartado analizaremos los resultados principales de dos de estas tareas: aprendizaje de gramáticas artificiales y control de sistemas interactivos.

### *Gramáticas artificiales*

Las gramáticas artificiales más frecuentemente utilizadas se basan en diagramas que determinan las transiciones válidas entre un conjunto de estados finitos (véase p. ej. el diagrama de la Figura 1). Podemos considerar tal diagrama como el conjunto de reglas que, una vez aprendidas, nos permitirán generar nuevas cadenas y decidir si nuevas secuencias de letras cumplen o no las reglas de la gramática. El procedimiento utilizado en la mayoría de experimentos (p. ej. Reber, 1967; Reber y Allen; 1978) consiste en pedir a los sujetos que intenten memorizar una lista de cadenas gramaticales. Tras esta fase de estudio, se explica a los sujetos que todas las cadenas habían sido generadas a partir de una misma gramática (conjunto de reglas) y que a continuación observarán nuevas cadenas. En esta segunda fase, la tarea de los sujetos consiste en decidir si la nueva secuencia es o no gramatical.

Los resultados de estos experimentos muestran que los sujetos son capaces de clasificar las nuevas cadenas significativamente mejor que si las clasificaran al azar. Pero sorprendentemente son incapaces de explicitar los criterios (reglas) en los que basan sus decisiones. Estos resultados han sido cuestionados más recientemente. Por ejemplo, Perruchet y Pacteau (1990) encontraron que el comportamiento de clasificación podía ser explicado a partir del conocimiento de fragmentos de las cadenas observadas. Es decir, los sujetos eran igual de hábiles clasificando nuevas cadenas o tan sólo bigramas o trigramas de letras que podían o no formar parte de una cadena gramatical. En este sentido, y tal como

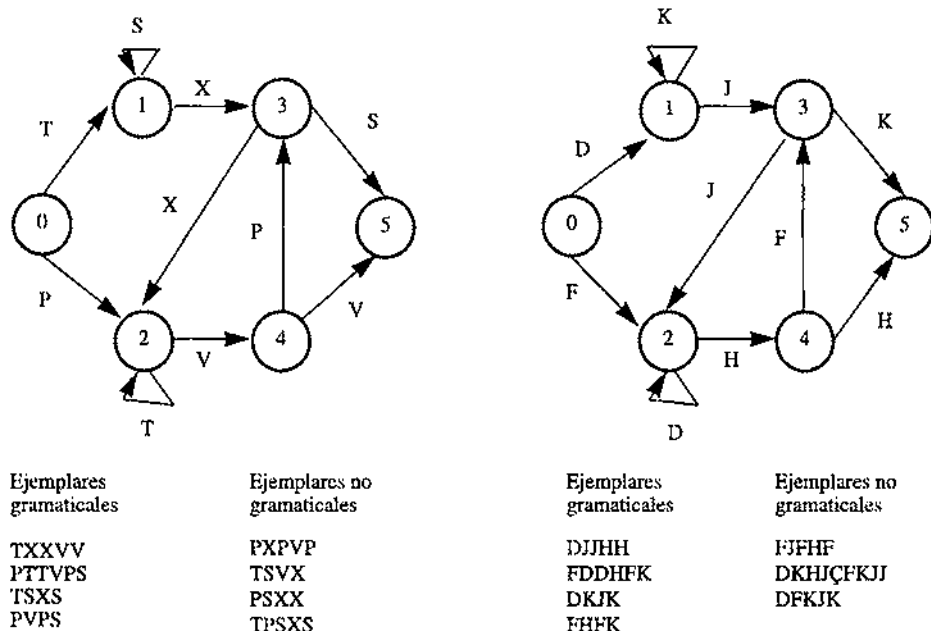


Figura 1. Dos versiones de una misma gramática de estados finitos.

proponen los autores, más que de un conocimiento implícito se trataría de un conocimiento incompleto constituido por fragmentos que funcionarían como indicadores de gramaticalidad.

También los experimentos de Mathews *et al.* (1989) mostraron que los sujetos eran capaces de verbalizar mucho más de lo que habían supuesto los primeros experimentos. Para ello pidieron a los sujetos del experimento que, una vez superada la primera fase de estudio de cadenas gramaticales, dieran instrucciones sobre lo que habían aprendido de los ejemplares estudiados a nuevos sujetos que no habían tenido experiencia alguna con las cadenas gramaticales. En una segunda fase, todos los sujetos (experimentales y los que recibían instrucciones, «instruidos») clasificaban nuevas cadenas las cuales podían ser o no gramaticales. Los resultados de este experimento mostraron que los sujetos «instruidos» pudieron realizar la tarea de clasificación a un nivel superior al azar, aunque inferior al de los sujetos experimentales. Es decir, los sujetos que habían observado ejemplares de la gramática demostraron cierta capacidad de verbalización pero parecían saber más de lo que podían explicitar.

¿Es este conocimiento residual, que parece imposible de ser verbalizado, la demostración de la existencia de conocimiento implícito? Como veremos en el apartado siguiente, las dificultades de verbalización pueden estar señalando determinadas características del conocimiento adquirido, aparte del nivel de

conciencia atribuido al aprendizaje. Es decir, algunas representaciones serían fácilmente accesibles y traducibles al código lingüístico mientras que otras representaciones no tendrían esta propiedad.

### *Control de sistemas interactivos*

Los experimentos que analizan cómo se aprende a controlar un sistema interactivo han enfatizado especialmente el papel de la complejidad de la relación en la distinción explícito versus implícito. La tarea utilizada consiste en la simulación simplificada de algún sistema. Por ejemplo el nivel de producción en una azucarera en función del número de trabajadores (Berry y Broadbent, 1984); el comportamiento de una persona «computerizada» a partir de las entradas obtenidas (Berry y Broadbent, 1988); o el comportamiento de una nave espacial controlada por un extraño robot (Tubau y Herrera, 1994). El sistema simulado siempre se comporta de acuerdo a una ecuación (regla) que relaciona la entrada del sujeto con la salida del sistema. En la mayoría de tareas utilizadas, esta relación puede ser de dos tipos: directa o «simple» cuando la salida depende de la entrada actual, o indirecta o «compleja» cuando la salida es función de la entrada anterior.

$$\begin{aligned} \text{Sistema simple: } S_t &= E_t - 2 + A (-1, 0, +1) \\ \text{Sistema complejo: } S_t &= E_{t-1} - 2 + A (-1, 0, +1) \end{aligned}$$

$S$  = Salida del sistema.  $E$  = Entrada del sujeto.  $t$  = tiempo de la interacción.  
 $A$  = número aleatorio que puede ser  $-1$ ,  $0$  o  $+1$ .

Los primeros estudios (p. ej. Berry y Broadbent, 1984; Hayes y Broadbent, 1988) llegaron a la conclusión de que la interacción con el sistema sencillo producía un aprendizaje explícito; el nivel de ejecución correlacionaba con el nivel de verbalización. Por el contrario, la interacción con el sistema complejo producía un aprendizaje implícito, observándose una disociación entre el nivel de actuación y el nivel de verbalización. Según Hayes y Broadbent (1988) el aprendizaje explícito o selectivo se basaría en estrategias analíticas típicas de las tareas de generación y comprobación de hipótesis, produciendo como resultado una regla (o la contingencia más relevante). En cambio, el aprendizaje implícito o no selectivo se basaría en estrategias no analíticas capaces de almacenar todas las contingencias del entorno (sin ninguna selección aparente).

Más específicamente, Lee (1995) concluye «parece que cuando la estructura del estímulo es compleja y el formato de codificación conduce a una discriminación perceptual, el aprendizaje se convierte en un proceso más pasivo y automático, y es difícil para los sujetos hacer que el proceso de aprendizaje sea explícito. Por otro lado, los sujetos son conscientes de su proceso de aprendizaje cuando las relaciones entre variables son simples y el aprendizaje incluye una manipulación simbólica». Lee sólo observa aprendizaje implícito en una tarea perceptual; los sujetos tenían que descubrir qué valores numéricos generaban de-

terminadas longitudes de dos líneas que aparecían en la pantalla y cuya relación la controlaban dos ecuaciones «complejas». No obstante, el criterio de complejidad, aparentemente crucial para la distinción explícito versus implícito, sigue apareciendo de forma poco clara.

En los últimos años, varios experimentos han puesto en duda esta distinción, al menos en lo referente al control de sistemas (Green y Shanks, 1993; Tubau y Herrera, 1994; Buchner, Funke y Berry, 1995). Según Green y Shanks (1993) la mayor complejidad de la regla explicaría la mayor dificultad de verbalización. Es decir, según estos autores, las dos tareas sólo se diferenciarían en el nivel de dificultad y no en procesos de aprendizaje aparentemente distintos. En todo caso la diferencia estaría en que en una situación (reglas simples) el aprendizaje sería más rápido que en la otra (reglas complejas).

Tubau y Herrera (1994) concluyen que la supuesta disociación entre el nivel de actuación y el nivel de verbalización característica del aprendizaje implícito sencillamente puede estar reflejando un conjunto de interacciones más amplio y variado y, seguramente, menos accesible. Cuando se interactúa con el sistema simple (relación directa) la entrada sólo puede llevar asociada tres posibles salidas. Pero cuando se trata de una relación compleja (indirecta), en cada interacción la entrada puede estar asociada a cualquier salida, ya que esta última depende de la interacción anterior. Como consecuencia, los dos sistemas producen bases de «ejemplares» muy diferentes. En el caso del sistema simple existiría una variabilidad limitada y, probablemente, los sujetos habrían experimentado la mayoría de situaciones. En cambio, al interactuar con el sistema complejo probablemente sólo se observarían algunas de las situaciones del conjunto total de combinaciones entrada-salida.

En la misma línea, Buchner *et al.* (1995) muestran que las características de las representaciones codificadas en la memoria del sujeto pueden explicar las diferencias al responder a las preguntas de los cuestionarios de predicción típicos en estas investigaciones. Los sujetos que han experimentado la mayoría de interacciones (condición más probable cuando se aprende el sistema simple) responden correctamente mayor número de preguntas que los que sólo pueden recordar algunas situaciones y que no han inducido la regla (más probable cuando el entorno coincide con el sistema complejo). Es decir, las disociaciones señaladas en la mayoría de experimentos podrían estar reflejando las características de la base de ejemplares. Más que procesos de aprendizaje diferenciados se trataría de distintos conjuntos de experiencias. Obviamente, la inducción de la(s) regla(s) de forma explícita conduciría a un nivel de ejecución (tanto en la tarea de control como en las pruebas de predicción) perfecta (o casi). Pero esto raramente ocurre incluso cuando el sistema se rige por ecuaciones «simples».

### **¿Qué se aprende de forma implícita?**

De momento, sólo tenemos una clara evidencia de aprendizaje implícito de covariaciones aparentemente «sin sentido» o «irrelevantes» para los sujetos



como las que se dan en las situaciones de condicionamiento evaluativo descritas anteriormente. En relación al aprendizaje de regularidades más complejas como las representadas en las gramáticas artificiales o en las tareas de control de sistemas no existe suficiente evidencia que demuestre su existencia. Podríamos encontrar una excepción en los experimentos de aprendizaje secuencial (Nissen y Bullemer, 1987; Lewiccki, Czyzewska y Hoffman; 1987) en los que los sujetos deben llevar a cabo una tarea serial de tiempo de reacción ante estímulos que siguen una secuencia estructurada encubierta. Repetidamente los autores defienden la existencia de aprendizaje implícito en este tipo de situación; el tiempo de reacción disminuye sin conciencia alguna de regularidad.

No obstante, tampoco queda claro qué han aprendido los sujetos en una tarea de aprendizaje secuencial. ¿Aprendizaje implícito de las complejas regularidades (Jiménez, 1996; Reed y Johnson, 1994) o adquisición de conocimiento procedimental como consecuencia de la práctica (Perruchet, Gallego y Savy, 1990; Shanks, Green y Colodny, 1994)? Si dentro del aprendizaje implícito se excluye aquel aprendizaje fruto de la repetición de unas mismas acciones, entonces el aprendizaje secuencial no sería un buen ejemplo de aprendizaje sin conciencia e implícito.

Dejando por el momento de lado la adquisición no consciente de simples asociaciones y la mejora procedimental fruto de la práctica, nos quedan dos alternativas sobre cuál podría ser el producto del aprendizaje implícito, en el caso de que así sea caracterizado: aprendizaje de un conjunto de episodios, experiencias concretas o «ejemplares» y aprendizaje implícito de regularidades o «reglas no conscientes».

### *Memoria de ejemplares*

En diversos experimentos, Brooks y colaboradores han aportado datos que permiten profundizar en la distinción entre aprendizaje de reglas versus aprendizaje de ejemplares. Por ejemplo, en uno de los experimentos de Brooks (1978) los sujetos observaban animales imaginarios como los de la Figura 2. Estos animales diferían en 5 dimensiones binarias: longitud del cuello, longitud de las patas, número de patas, forma del cuerpo y presencia o no de manchas. Cada animal se presentaba en un entorno distinto (véase Figura 2) el cual también variaba en algunas dimensiones como el clima (cálido o frío) y la humedad (presencia o no de agua) y podían aparecer en distintos colores: verde, azul o marrón. Los animales se presentaban en parejas y en distinta posición, con el fin de facilitar la eliminación de algunos atributos como relevantes.

Algunos de los animales utilizados habían sido clasificados dentro de una categoría ficticia: Lepton. Ser Lepton estaba determinado por la conjunción de dos atributos (p.e. forma del cuerpo y presencia de manchas) los cuales fueron variando a través de los sujetos. La dimensión clima estaba relacionada con dos de los atributos mediante una regla conjuntiva (p.e. longitud del cuello y número de patas) y la dimensión humedad estaba relacionada con un atributo del animal (p. ej. longitud de las patas).

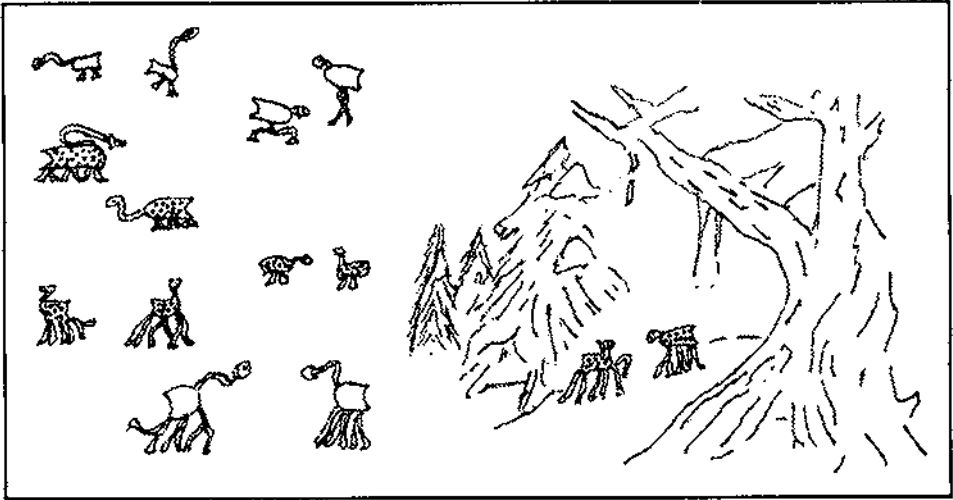


Figura 2. Ejemplos de los animales Lepton utilizados en los experimentos de Brooks (1978) y Allen y Brooks (1991).

Brooks (1978) comparó el porcentaje de clasificaciones correctas de tres grupos sometidos a distintas tareas de adquisición. El primer grupo aprendía mediante el clásico paradigma de recepción (Bruner *et al.*, 1956); generando y comprobando hipótesis tenían que descubrir los atributos definitorios de Lepton. El segundo grupo aprendía por memorización. Desconocían la existencia de una categoría relevante (Lepton) y su tarea consistía en asociar cada uno de los animales con un nombre propio. Una vez superada esta primera fase, se les indicaba el nombre de los animales Lepton. Finalmente un tercer grupo aprendía mediante una estrategia mixta (aprendían los nombres propios al mismo tiempo que se les indicaba si pertenecían o no a la categoría Lepton). Los tres grupos de sujetos desconocían la existencia de regularidades entre el entorno y las características de los animales.

Los resultados mostraron que los sujetos del primer grupo fueron los más hábiles en clasificar nuevos animales Lepton pero fueron incapaces de predecir en qué entorno podrían hallarse. Los sujetos del segundo grupo tuvieron un peor rendimiento al clasificar nuevos animales, pero fueron los que mejor podían establecer correctas relaciones con el entorno. Finalmente, los sujetos del tercer grupo obtuvieron resultados similares, y en algunas ocasiones peores, que los del segundo grupo. Los sujetos que habían aprendido regularidades relacionadas con el contexto (especialmente los del segundo grupo), lo hicieron de forma implícita, sin mostrar conciencia alguna de ellas y, como muestran los resultados, esto ocurría especialmente cuando existía una buena base de ejemplares.

Más recientemente, Allen y Brooks (1991), utilizando estímulos similares a los de Brooks (1978) (véase Figura 2) investigaron el efecto de la práctica con una regla operativa (*additive rule* o regla conjuntiva que incluía valores de tres

de las cinco dimensiones binarias). Sus resultados mostraron que, aunque la regla estuviera accesible y posibilitara realizar clasificaciones con mayor éxito, los sujetos seguían mostrando una significativa influencia de los episodios o ejemplares más familiares, tanto en la precisión de sus decisiones (porcentaje de aciertos) como en el tiempo de respuesta. En relación a los ejemplares nuevos y negativos (no formaban parte de la categoría), cuando se identificaban correctamente como «no ejemplares» se observaba un incremento del tiempo de reacción, pero sólo en el caso de que los sujetos ya conocieran la regla.

Allen y Brooks también encontraron que el efecto de la familiaridad sólo ocurría con estímulos pictóricos (dibujos donde los atributos son fácilmente integrables), pero no cuando los estímulos eran listas verbales de atributos. En este último caso los sujetos parecían basarse únicamente en la regla, no dejándose influir por los episodios o experiencias previas. Brooks, Norman y Allen (1991) confirmaron los hallazgos antes mencionados (influencia de la similitud o familiaridad cuando los estímulos se presentaban en forma de dibujos) utilizando como estímulos «material más ecológico». En concreto se trataba de diapositivas en color de lesiones dermatológicas escogidas de un catálogo de dermatólogos profesionales.

Tanto los expertos como los estudiantes de medicina mostraron una clara influencia del grado de similitud entre los nuevos ejemplares y los vistos anteriormente, independientemente de su pertenencia a la misma categoría diagnóstica. En relación al comportamiento de los expertos, esta influencia se reflejaba especialmente en el tiempo de reacción. Es decir, el conocimiento de la regla definitoria de la categoría diagnóstica les permitía corregir su decisión inicial (frecuentemente influenciada por la similitud entre ejemplares). Y, en consecuencia, al final los expertos eran capaces de dar respuestas más acertadas pero consumiendo más tiempo.

### *¿Representación implícita de las reglas?*

Si un organismo ha almacenado un conjunto de ejemplares (o un conjunto de contingencias) y su comportamiento (p.e. en una tarea de discriminación) puede describirse en forma de reglas, ¿posee reglas que emergen, de forma implícita, del conjunto de ejemplares? Y si éste es el caso, ¿de qué tipo de reglas se trata? La prueba crucial parece ser analizar la capacidad de generalización: ¿cuál es el poder de estas supuestas reglas implícitas?

Los estudios sobre aprendizaje de gramáticas artificiales aportan resultados relevantes respecto a este punto. Algunos investigadores han intentado analizar el nivel de abstracción del conocimiento adquirido realizando pruebas sobre nuevas cadenas generadas por la misma gramática pero con distintas letras (Reber y Allen, 1978; Mathews *et al.*, 1989). Los sujetos expuestos a tales variaciones siguen siendo capaces de clasificar mejor que si lo hicieran aleatoriamente. Sin embargo, para algunos autores como Brooks y Vockey (1991) o Vockey y Brooks (1992) estos resultados no estarían señalando necesariamente la inducción implícita de reglas, sino tan sólo la capacidad de establecer analogías

abstractas. Según estos investigadores, los sujetos habrían aprendido, por ejemplo, que las cadenas gramaticales no empiezan y acaban con la misma letra, que la segunda y cuarta letra pueden aparecer repetidas, etc. ¿Se trata de reglas abstractas implícitas o de conocimiento fragmentado de episodios o ejemplares?

Se considera que una regla es abstracta cuando representa un resumen de las regularidades detectadas en un conjunto de episodios separados de aprendizaje (Neal y Hesketh, 1997). Pero no todas las reglas reflejan el mismo nivel de abstracción. Por ejemplo, en relación a este punto, Shanks (1995) introduce dos niveles principales: uno más específico o de bajo nivel, donde la regla estaría directamente conectada con los ejemplares y otro más abstracto o de alto nivel, donde la regla tendría en cuenta dimensiones o atributos distintos a los observados en la fase de entrenamiento. Probablemente el tipo de reglas que parecen ser aprendidas de forma implícita formarían parte de la primera categoría; reglas que funcionarían de forma dependiente a los ejemplares experimentados. En relación a las reglas más abstractas no existe evidencia experimental de que sean inducidas de forma implícita.

Con la aparición de los modelos basados en redes neuronales artificiales (conexionismo), ha habido un considerable número de trabajos que han abordado el problema de la inducción «implícita» de reglas desde este nuevo enfoque. De forma genérica, estos estudios aportan, tal como veremos en el siguiente apartado, datos interesantes en el campo de la adquisición de conocimiento y, específicamente, en relación a la inducción de reglas. Este último punto creemos que puede ser discutido teniendo en cuenta estos niveles de abstracción introducidos.

## **La inducción en redes neuronales: reglas y ejemplares**

En la introducción mencionamos cómo algunos autores conexionistas (por ejemplo, Cleeremans, 1993; Keele y Jennings, 1992, etc.) proponen un único modo de aprendizaje de tipo asociativo que daría cuenta tanto de la organización de los diferentes ejemplares en memoria como del funcionamiento basado en reglas. Algunas simulaciones muestran que este aprendizaje modelizado por redes neuronales, y basado en mecanismos elementales, puede procesar material con un alto grado de complejidad en diferentes tareas (aprendizaje de secuencias, aprendizaje de gramáticas artificiales y de conceptos, etc.) (Hinton, 1986).

El conocimiento adquirido mediante este mecanismo toma forma de patrón de activación a través de un conjunto de unidades de procesamiento, llámense unidades representacionales, intermedias, ocultas o de conocimiento según los autores. Este conocimiento refleja asociaciones entre configuraciones específicas de estímulos, ya sean temporales o espaciales, y respuestas específicas. Existe, además, evidencia (Cleeremans, 1993) de que este conocimiento refleja las covariaciones más relevantes entre atributos en el sentido de Shanks (1995) de una forma fragmentada similar a la propuesta por autores como Perruchet y Pacteau (1990).

En muchas ocasiones, la ejecución correcta de una tarea supone la aplicación de una regla abstracta. ¿Es capaz una red neuronal de generar una representación lo suficientemente independiente de los ejemplares utilizados en el aprendizaje como para dar cuenta de un alto grado de abstracción? La naturaleza de las representaciones en un red neuronal no es un campo exento de debate y polémica. La respuesta a esta pregunta supone abordar, a nuestro entender, los límites de la inducción por parte de las redes neuronales. En este apartado debatiremos este tipo de representación a partir de cómo los modelos conexionistas organizan los ejemplares y hasta qué punto dicha representación puede dar cuenta de un comportamiento basado en reglas del nivel de abstracción más elevado que introdujimos anteriormente. Pensamos que la distinción entre estos niveles de abstracción proporciona elementos interesantes para el análisis de los límites de los modelos conexionistas en la representación de reglas.

### *Organización explícita e implícita de ejemplares*

Un aspecto importante que nos permite distinguir diferentes grados de abstracción en las representaciones en los modelos conexionistas nos lo proporciona el hecho de que un modelo conexionista puede ser diseñado para almacenar explícitamente copias de los distintos ejemplares que han sido vistos por una red neuronal. El modelo basado en ejemplares de Estes (1986), el de Medin y Schaffer (1978) o el Modelo contextual generalizado de Nosofsky (1986) son trabajos representativos de este enfoque. En estos modelos, cada ejemplar se representa en la capa de entrada por un conjunto de atributos que suelen ser binarios. Cada unidad de la entrada codifica la información sobre un atributo; si la unidad está activa denota la presencia del atributo correspondiente, mientras que si la unidad no está activa denota su ausencia. En la capa intermedia, tendremos una unidad para cada ejemplar almacenado. Es decir, hay una correspondencia uno a uno entre los ejemplares a representar y el número de unidades intermedias.

Podemos darnos cuenta con bastante facilidad de las limitaciones representacionales que supone esta vía prefijada de organización. Aunque resulta mucho más fácil el análisis de la representación interna de la red, no resultan, en modo alguno, plausibles los requerimientos de memoria para representar un gran número de ejemplares. El potencial de la red para dar cuenta de un número elevado de ejemplares es muy limitado.

La alternativa más clara a esta vía de representación es no restringir la activación de una unidad intermedia a la presencia de un único ejemplar en la capa de entrada. Podemos dejar libertad a la red neuronal para que, con los recursos disponibles (el número de unidades intermedias con que cuenta), represente y organice libremente los ejemplares, no restringiendo de manera explícita la representación de cada ejemplar a una sola unidad intermedia como hacen los modelos basados en ejemplares que hemos citado más arriba. De esta forma, el acceso a la representación obtenida por la red no es directa, pero puede ser accesible mediante métodos de análisis como puede ser el análisis de componentes principales o análisis de agrupaciones o conglomerados (para más deta-

les véase Elman, 1991 y Moliner, 1993). En general, se habla de este tipo de representación como distribuida.

Ha habido muchas discusiones sobre el tipo de representaciones descubiertas por las unidades de la capa intermedia (ver por ejemplo, Sejnowski y Rosenberg 1987; Rosenberg 1987). Estas discusiones normalmente enfatizan el hecho de que la red neuronal descubre representaciones distribuidas de forma que los estímulos de entrenamiento que tienen asociadas respuestas similares desarrollarán representaciones similares en la capa intermedia (Hinton *et al.*, 1986).

También se ha discutido sobre qué propiedades debe tener esta representación distribuida en las unidades de la capa intermedia de forma que permita la generalización a nuevos estímulos. Normalmente, un punto importante es que una red con un mínimo número de unidades representacionales o intermedias se fuerza a captar las regularidades más frecuentes que se encuentran en el conjunto de estímulos de aprendizaje o entrenamiento, pero, en este caso, siendo incapaces de retener información arbitraria de ejemplares muy específicos. No obstante, Kruschke (1992) consigue, en un sistema conexionista, retener información sobre ejemplares con propiedades muy excepcionales a partir de representaciones distribuidas.

En resumen, parece ser que hay un continuum representacional en las redes neuronales que va desde el almacenamiento explícito de ejemplares al desarrollo de representaciones que codifican propiedades compartidas por gran número de ellos posibilitando así unas bases para la generalización. En este último extremo, el criterio de agrupación va más allá de una copia de ejemplares específicos. Las modelizaciones evidencian que las redes son capaces de utilizar los atributos relevantes a través de un amplio número de ejemplares, lo cual, sin duda alguna, denota cierto grado de abstracción de la representación obtenida. En el siguiente apartado contrastaremos estos grados de abstracción con los propuestos anteriormente.

### *Representaciones en las redes neuronales: algunas limitaciones*

En la sección anterior hemos considerado dos grados de abstracción por lo que a la organización de los ejemplares en redes neuronales se refiere. Sin duda, las representaciones distribuidas permiten una capacidad de generalización ante nuevos ejemplares que no permite una representación prefijada de los ejemplares. Sin embargo, ¿hasta qué punto una representación distribuida puede dar cuenta de reglas lo suficientemente abstractas como para decir que son independientes de los ejemplares? A continuación debatiremos sobre este punto.

A lo largo de este trabajo hemos delineado una serie de dimensiones referentes a la posibilidad de explicitar o no el conocimiento adquirido y cómo estas dimensiones se pueden relacionar con los dos niveles de abstracción. En algunos trabajos sobre la modelización de tareas con redes neuronales se ha considerado cómo se puede abordar la introducción explícita de conocimiento y en qué grado éste ayuda a la realización de la tarea. Por ejemplo, Cleeremans (1993) simula la tarea de aprendizaje de secuencias de Lewicki *et al.* (1988) y el efecto de la in-

roducción de información explícita. En este caso, además de la presentación temporal de los elementos, se introduce a la red información sobre la posición serial de los mismos. Las conclusiones a las que llega apuntan a que la red no codifica la posición serial y sólo utiliza los patrones de covariación temporal para resolver la tarea. Sólomente un aprendizaje exhaustivo hace que la red se muestre sensible a la información adicional, pero en ningún caso, a diferencia de lo que puede ocurrir con humanos, la ejecución mejora significativamente. Así pues, la naturaleza de las representaciones obtenidas mediante mecanismos asociativos parece no ser sensible a la introducción de información adicional explícita.

La simulación con redes neuronales también puede aportar alguna información útil sobre el por qué representaciones implícitas son difícilmente verbalizables. El carácter distribuido y holístico de las representaciones implica que la información relevante está incrustada dentro de complejos patrones representacionales y es difícilmente analizable (Servan-Schreiber y Anderson, 1990). Aunque, como señala Cleeremans (1993), es muy costoso el análisis sobre si estos tipos de representaciones basadas en la detección de covariaciones permiten una ejecución satisfactoria de tareas a partir de reglas abstractas, pensamos que sí hay unas limitaciones que consideramos relevante comentar.

Como hemos mencionado, existen datos experimentales derivados de las simulaciones que indican que redes neuronales pueden aislar qué atributos son relevantes y cuáles no para resolver una tarea. Por ejemplo, podemos obtener representaciones abstractas de gramáticas de estados finitos, donde cada estado de la gramática se representa por un conjunto de valores de activación en las neuronas intermedias. Nadie dudaría de que existe cierto grado de abstracción en la codificación de dimensiones relevantes, ya que vamos más allá de las copias literales de ejemplares específicos.

Sin embargo, estas representaciones aún dependen de atributos superficiales que no dejan de definir los distintos elementos. En este sentido, no podemos hablar de representaciones abstractas en el sentido introducido en los apartados anteriores. Por ejemplo, la red no podría generalizar la misma gramática a un conjunto diferente de letras representadas por nuevas unidades de entrada. Esto no debe sorprendernos, ya que la red no puede inducir que las diferentes letras jueguen el mismo papel en cada gramática. No obstante, se podrían diseñar mecanismos que permitiesen a la red evaluar este tipo de similitud más abstracta (Hinton, 1986; Cleeremans, 1993). Aunque estos aspectos concernientes a las analogías abstractas han sido abordados por modelos conexionistas, no existen soluciones definitivas a este problema.

Consideremos los ejemplares de la Figura 3. Cada ejemplar se compone de tres objetos que vienen definidos por atributos como tamaño y textura. Supongamos que en la fase de aprendizaje los sujetos han aprendido que en los ejemplares positivos se da un aumento de la dimensión «forma» a través de los objetos que forman el ejemplar (Figura 3a). La tarea consiste en escoger el ejemplar de la derecha (3b o 3c) que más se parece al ejemplar positivo (3a). En 3c no sólo hay una semejanza en la dimensión «forma», sino que hay una misma relación creciente en el atributo «textura». Es decir, mientras que en 3a, se da una relación creciente en el atributo «tamaño», en 3c se da una relación creciente en la

tonalidad de gris de la textura. Este juicio implica lo que se conoce como una alineación de estructura (*structural alignment*) (Markman y Gentner, 1990). Hay una transferencia de relación entre valores de un atributo a otro atributo diferente. Difícilmente la red puede generalizar de manera sistemática la misma relación del atributo tamaño al atributo textura si durante el aprendizaje este último atributo no varía de valor. Esta restricción no se da en la emisión de juicios de similitud basados en relaciones en humanos. Aunque existen modelos conexionistas que intentan implementar la alineación de estructura (p. ej. Handler y Cooper, 1993), no hay una verdadera transferencia de relaciones a nuevos atributos.

Por lo tanto, parece ser que la red sólo podrá responder adecuadamente ante un nuevo estímulo susceptible de ser representado en el espacio de entrada utilizando los mismos atributos. Aunque la generalización se puede basar en covariaciones de atributos muy complejas, estas covariaciones siempre tienen lugar entre los mismo atributos, es decir, en el mismo espacio. Existen reglas, que por su nivel de abstracción, no están sujetas a unas dimensiones previas.

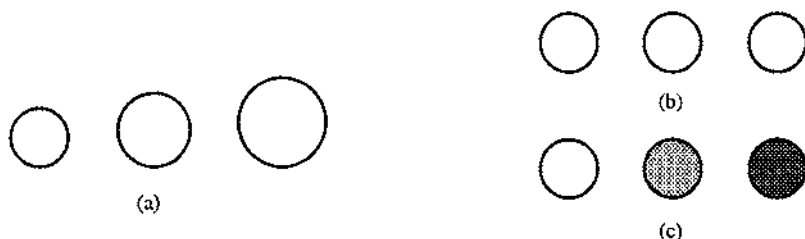


Figura 3. Estímulos utilizados en una tarea de alineación de estructura (véanse explicaciones en el texto).

Cuando emitimos juicios de similitud entre dos ejemplares, las dimensiones o atributos de evaluación pueden variar de un par de ejemplares a otro par de ejemplares. En una red neuronal existen limitaciones representacionales en este aspecto. No es lo mismo emitir un juicio de similitud en base a unas dimensiones preestablecidas (por ejemplo, decir si dos objetos tienen los mismos valores de forma, textura, etc.) que poder generalizar la misma relación a cualquier dimensión. Si en una red neuronal representamos dos objetos, cualquier juicio podrá ser emitido solamente en base a las diferencias/semelanzas entre las distintas dimensiones que se utilizan para representar los objetos en la red. Pero no se puede generalizar una respuesta más allá del espacio representacional fijo que implican las unidades de entrada de la red.

## Conclusiones

Uno de los principales objetivos que nos planteábamos era analizar la distinción entre reglas, o representaciones abstractas, y ejemplares, o representa-



ciones más específicas. En concreto, nos preguntábamos hasta qué punto el nivel de representación basado en reglas podía funcionar de forma independiente del nivel más específico de los ejemplares. ¿Cuáles son las respuestas que podemos ofrecer ahora? Analicemos, en primer lugar, las dos dimensiones que necesariamente tienen que mezclarse en esta discusión.

Tal como distinguen Shanks y St. John (1994) por un lado está la dimensión consciente-no consciente. Podemos llegar a identificar claramente la inducción de reglas abstractas, pero ¿en qué punto de la dimensión podemos hallarlas? Lo mismo sería aplicable al hablar de ejemplares o covariaciones. Con el fin de intentar disminuir la posible confusión generada por esta dimensión, tomaremos como diferencia fundamental entre los dos polos la posibilidad de traducir el conocimiento al código lingüístico. En otras palabras, el conocimiento no consciente sería aquel que sólo nos permite actuar (p. ej. tomar decisiones sin saber «por qué»), mientras que el conocimiento consciente sería aquel que, además de permitirnos realizar ciertas tareas, podemos comunicarlo a través del lenguaje u otros sistemas representacionales (p. ej. dibujos).

Por otro lado nos encontramos con la dimensión específico-abstracto, principal para este artículo. Aunque los autores antes mencionados consideran que ambas dimensiones son independientes, nosotros consideramos que no son fácilmente separables. Como hemos visto a lo largo de los distintos apartados analizados, y tal como resaltaremos a continuación, la naturaleza de algunas de las representaciones formadas debe permitir más fácilmente el ser traducidas o codificadas mediante un sistema simbólico como lo es por excelencia el lenguaje humano. Parece ser que, en general, los sujetos pueden verbalizar más fácilmente representaciones en forma de reglas, mientras que existiría una mayor dificultad cuando las representaciones están basadas en ejemplares.

Tanto los estudios centrados en la adquisición de gramáticas artificiales como los que han investigado las estrategias de control de sistemas interactivos han mostrado una fuerte limitación al intentar demostrar la existencia de inducción no consciente de reglas abstractas. Quizás uno de los problemas subyacentes sea precisamente la falta de definición operativa sobre qué es una regla abstracta. En cualquier caso todos los estudios muestran una importante influencia de los episodios o ejemplares experimentados. En el supuesto caso de que los sujetos hubieran inducido reglas, tal como defiende Reber (1989), éstas no serían capaces de evitar la influencia producida por las estructuras de semejanza que relacionan los nuevos ejemplares con los almacenados en la memoria.

¿En qué situaciones podemos observar una clara separación entre decisiones basadas en la similitud entre estímulos y decisiones basadas en reglas? Los estudios sobre clasificación (o adquisición de conceptos) que hemos visto al hablar sobre «memoria de ejemplares» ofrecen una ilustrativa ejemplificación. El uso de estrategias analíticas o estrategias de generación y comprobación de hipótesis suele conducir a la identificación de la regla definitoria de la categoría la cual puede, en algunas circunstancias, actuar de forma independiente a las contingencias estructurales. Sin embargo, estas reglas, aunque fácilmente traducibles al código lingüístico, siguen siendo poco abstractas. Es decir, en todos los estudios sobre clasificación analizados la regla definitoria coincide con algún o

algunos de los atributos compartidos por todos los ejemplares de la categoría. Formarían parte, pues, del primer nivel de abstracción antes mencionado.

Este tipo de reglas, es decir, las que corresponden al primer nivel de abstracción, se han mostrado susceptibles de ser fácilmente inducidas por modelos conexionistas. El hecho de que los modelos conexionistas tengan éxito a partir de un conocimiento fragmentado radica en la habilidad de combinar estos fragmentos de conocimiento sobre covariaciones en representaciones más abstractas. Este hecho ha permitido a las redes inducir reglas basadas en correlaciones entre atributos e ir más allá de simples reglas definidas a partir de un atributo definitorio. No obstante, quedan por responder muchas preguntas y no es tarea nada fácil abordar este tema y lo demuestra el hecho de que no existen trabajos que realicen un estudio detallado sobre este aspecto.

La observación de que cierto conocimiento abstracto pueda emerger de mecanismos asociativos contradice la concepción extendida de que este tipo de conocimiento se debe adquirir necesariamente a través de operaciones que producen representaciones explícitas basadas en reglas. Pero aunque podamos llegar a hablar de emergencia de conocimiento que, por su simplicidad, es fácilmente expresable en un código simbólico mediante procedimientos de análisis, ¿qué podemos decir sobre la inducción de regularidades más abstractas como por ejemplo el concepto de igualdad, o reglas sobre relaciones cuantitativas como las que se dan en los sistemas dinámicos?

Los pesos aprendidos por una red neuronal posibilitan el acceso a una representación a partir de la especificación de un ejemplar en la entrada. Esta representación emerge gracias a la propagación de activación a través de los pesos desde el espacio que representa a los ejemplares y no desde ningún otro. Además, esta representación puede ser sensible a interacciones complejas entre los atributos que definen un ejemplar. En caso de un gran número de atributos, difícilmente podemos hacer explícitas las interacciones o regularidades que ha observado la red. La capacidad de generalización se debe a estas regularidades inducidas y que son comunes a todos los ejemplares de entrenamiento. Sin embargo, el nivel de abstracción adquirido por la red no permite la generalización a nuevos ejemplares que no sean susceptibles de ser definidos mediante las mismas dimensiones. En otras palabras, las reglas inducidas no pueden desligarse del espacio utilizado para representar los ejemplares de aprendizaje y su aplicación tendrá sentido ante nuevos ejemplares siempre que éstos sean nuevos puntos en el mismo espacio representacional. Estas argumentaciones toman mayor sentido al considerar la evidencia experimental existente de la inducción implícita de reglas sencillas y el hecho de que no se ha demostrado el mismo tipo de inducción de reglas más complejas que implican un desmarque de los ejemplares que se han utilizado en el aprendizaje.

En páginas anteriores hemos esbozado algunas limitaciones las cuales hacían referencia a la generalización de relaciones conocidas a nuevos atributos. Actualmente, los modelos conexionistas que intentan modelizar procesos cognitivos de más alto nivel como el razonamiento, no suelen disponer de sistemas sensoriales anclados en el mundo real. Es decir, difícilmente una red neuronal puede desarrollar representaciones a partir de los datos sensoriales directos del

mundo real. Ello supone que cualquier objeto debe ser representado o codificado en la entrada de la red a partir de un código preestablecido, por ejemplo mediante un conjunto de atributos a modo de primitivos semánticos. Cualquier generalización se debe basar necesariamente en relaciones entre los atributos codificados. Sin duda, un reto muy importante para los modelos basados en redes neuronales en los próximos años será la conexión directa con información del mundo físico, de forma que pueda ser posible el desarrollo de representaciones internas a partir de la experiencia directa con objetos reales. Además queda por solucionar de qué forma se deben combinar las representaciones obtenidas a partir de este contacto con información del mundo real (Miikkulainen, 1993).

Aunque no hay respuestas definitivas sobre los límites de los mecanismos asociativos, los resultados analizados tanto de los experimentos sobre aprendizaje implícito como del tipo de conocimiento que puede ser representado mediante una red neuronal muestran una importante dificultad en relación a la inducción de reglas del nivel más alto de abstracción. Probablemente esta limitación sea una manifestación de la relación, frecuentemente señalada, entre la capacidad de abstracción y la existencia de un código representacional (normalmente el lenguaje) que permita acceder a relaciones más amplias que las similitudes perceptuales. El aprendizaje explícito, a diferencia del implícito, parece basarse en este código representacional como herramienta principal para la adquisición de conocimiento.

## REFERENCIAS

- Allen, S. W. & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, 1, 3-19.
- Baeyens, F., Eelen, P., Crombez, G. & Berg, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy*, 30, 133-142.
- Berry, D. C. & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology*, 36A, 209-231.
- Berry, D. C. & Broadbent, D. E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology*, 79, 251-272.
- Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. En E. Rosch y B. Lloyd (Eds.), *Cognition and concepts* (pp. 169-211). Hillsdale, NJ: Erlbaum.
- Brooks, L. R. & Vokey, J. R. (1991). Abstracts analogies and abstracts grammars. *Journal of Experimental Psychology: General*, 120, 316-323.
- Brooks, L. R., Norman, G. R. & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General*, 3, 178-287.
- Bruner, J. S., Goodnow, J. & Austin, G.A. (1956). *A study of thinking*. New York: Wiley.
- Buchner, A., Funke, J. & Berry, D.C. (1995). Negative correlations between control performance and verbalizable knowledge: indicators for implicit learning in process control tasks? *The Quarterly Journal of Experimental Psychology*, 1, 166-187.
- Cleermans, A. (1993). *Mechanisms of implicit learning: Connectionist models of sequence processing*. Cambridge, MA: MIT Press.
- Cleermans, A. & McClelland, J. L. (1991). Learning the structure of sequence events. *Journal of Experimental Psychology: General*, 120, 235-253.
- Dreyfus, H. L. & Dreyfus, S. E. (1986). Why skills cannot be represented by rules. En N. E. Sharkey (Ed.), *Advances in Cognitive Science* (pp. 315-335). Chichester, England: Ellis Horwood.
- Elman, J. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*.

- Estes, W. K. (1986). Memory storage and retrieval processes in category learning. *Journal of Experimental Psychology: General*, *115*, 155-174.
- Froufe, M. (1996). *El inconsciente cognitivo. Datos, teorías y aplicaciones*. Ediciones de la Universidad Autónoma de Madrid.
- Green, R. E. A. & Shanks, D. R. (1993). On the existence of independent explicit and implicit learning systems: An examination of some evidence. *Memory and Cognition*, *21*, 304-317.
- Handler, J. B. & Cooper, P. R. (1993). Analogical Similarity: Performing Structure Alignment in a Connectionist Network. In *Proceedings of the 15th. Annual Conference of the Cognitive Science Society*, pp. 522-527. Boulder, Co., Erlbaum.
- Hayes, N. A. & Broadbent, D. E. (1988). Two modes of learning for interactive tasks. *Cognition*, *28*, 249-276.
- Hinton, G. E. (1986). Learning distributed representations of concepts. En *Proceedings of the 8th. Annual Conference of the Cognitive Science Society*. Boulder, Co., Erlbaum.
- Hinton, G. E., McClelland, J. L. & Rumelhart, D. E. (1986). Distributed Representations, En D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel Distributed Processing*. Cambridge, Ma.: MIT Press.
- Jiménez, L. (1996). Aprendizaje secuencial: modelos de reglas versus modelos conexionistas. *Cognitiva*, *2*, 235-255.
- Keele, S. W. & Jennings, P. J. (1992). Attention in the representation of sequence. Experiment and theory. *Human Movements Studies*, *11*, 125-138.
- Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review*, *99*, 22-44.
- Lee, Y. S. (1995). Effects of learning contexts on implicit and explicit learning. *Memory & Cognition*, *23*, 723-734.
- Lewicki, P. (1986). *Nonconscious social information processing*. New York: Academic Press.
- Lewicki, P., Czyżewska, M. & Hoffman, H. (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*, 523-530.
- Lewicki, P., Hill, T. & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, *20*, 24-37.
- Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behavior*. New York: Live-right.
- Markman, A. B. & Gentner, D. (1990). Analogical mapping during similarity judgments. In *Proceedings of the 12th. Annual Conference of the Cognitive Science Society*, pp. 38-44. Hillsdale, NJ. Erlbaum.
- Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R. & Druhan, B. (1989). Role of explicit and implicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 1083-1100.
- Medin, D. I. & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207-238.
- Miikkulainen, R. (1993). *Subsymbolic Natural Language Processing*. Cambridge, Ma: MIT Press.
- Moliner, J. L. (1993). Inducció de regles mitjançant una xarxa neuronal seqüencial. Tesis de Licenciatura no publicada, Universitat de Barcelona.
- Neal, A. & Hesketh, B. (1997). Episodic knowledge and implicit learning. *Psychonomic Bulletin & Review*, *4* (1), 24-37.
- Nisbett, R. E. (Ed., 1993). *Rules for reasoning*. Hillsdale, NJ: LEA.
- Nissen, M. J. & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1-32.
- Nosofsky, R. M. (1986). Attention, similarity and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.
- Perruchet, P. & Pacteau, C. (1990). Synthetic grammar learning: implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, *3*, 264, 275.
- Perruchet, P., Gallego, J. & Savy, I. (1990). A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognitive Psychology*, *22*, 493-516.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *7*, 317-327.
- Reber, A. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219-235.
- Reber, A. (1993). *Implicit learning and tacit knowledge*. Oxford: Oxford University Press.
- Reber, A. S. & Allen, R. (1978). Analogic and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition*, *6*, 189-221.
- Reed, J. & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Human Learning and Memory*, *20*, 585-594.

- Rosenberg, C. R. (1987). Revealing the structure of NETalk's internal representations. En *Proceedings of the 9th Annual Conference of the Cognitive Science Society*, pp. 537-554. Hillsdale, NJ: Erlbaum.
- Sejnowsky, T. J. & Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English text. *Complex Systems*, 1, 145-168.
- Servan-Schreiber, E. & Anderson, J. R. (1990). Learning artificial grammars with competitive chunks. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 592-608.
- Shanks, D. R. (1995). *The psychology of associative learning*. New York: Cambridge University Press.
- Shanks, D. R. & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17, 367-447.
- Shanks, D. R., Green, R. E. & Colodny, J. (1994). A critical examination of the evidence for unconscious (implicit) learning. En C. Ulmìtà y M. Moscovitch (Eds.), *Attention and performance XV: Conscious and non conscious information processing*. Cambridge, MA: Bradford.
- Skinner, B. F. (1969). *Contingencies of reinforcement*. Englewood Cliffs, NJ: Prentice-Hall.
- Skinner, B. F. (1974/1977). *Sobre el conductismo*. Barcelona: Fontanella.
- Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11, 1-74.
- Tubau, E. y Herrera, P. (1994). Aprendizaje implícito y representación del conocimiento en tareas de control de sistemas interactivos. *Cognitiva*, 6, 47-65.
- Vokey, J. R. & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 328-344.

