

UN SISTEMA DE APRENDIZAJE MULTIESTRATÉGICO BASADO EN UN MODELO DE ARQUITECTURA FLUIDA

M.D. del Castillo, M.P. Sesmero, P. Bustos, L. Barrios, S. Ros
Instituto de Automática Industrial, CSIC
Carretera de Campo Real Km 0.200, La Poveda
28500 Arganda del Rey, España.
e-mail: lola@iai.csic.es

Palabras clave: aprendizaje multiestratégico, representación, procesos de inferencia

Resumen. En este artículo se presenta un sistema de aprendizaje multiestratégico para abordar la adquisición de conocimiento a partir de ejemplos. Este sistema integra los procesos de inferencia comunes a diversas estrategias de aprendizaje, permitiendo que compartan un espacio de trabajo con conocimiento de diferente grado de complejidad que debe satisfacer los objetivos de aprendizaje del sistema. La estrategia de control del sistema se encuentra distribuida en los tres espacios que definen su arquitectura. La dinámica de esta arquitectura se basa en el ciclo percepción-representación-operación.

1. Introducción

Desde un punto de vista formal, un proceso de aprendizaje se puede interpretar como un procedimiento de búsqueda a través de un espacio de conocimiento con una determinada representación. Los estados de dicho espacio se generan aplicando operadores o reglas de inferencia. El elemento que aprende, o aprendiz, debe ser capaz de razonar o inferir a partir de la información que se le presenta y del conocimiento que posee y de conservar el resultado de su razonamiento. Por tanto, el tipo de inferencia que lleva a cabo (deducción, inducción, analogía,...) conforma su capacidad para razonar y ha de constituir parte de su conocimiento básico. Este depósito de conocimiento también puede incluir toda la información necesaria para restringir el proceso de búsqueda y hacerlo eficiente.

Un aspecto de destacable interés cuando el ser humano aprende es su capacidad para aplicar diferentes estrategias de aprendizaje, en función de la situación presente, y para adaptarse dinámicamente a los cambios existentes. A partir de una mayor comprensión de las ventajas y las limitaciones de los métodos de aprendizaje automático de una sola estrategia ha surgido un creciente interés por los sistemas de aprendizaje multiestratégico que integran diferentes tipos de inferencias. Las características fundamentales en el diseño de estos sistemas son el abanico de estrategias disponibles y los métodos de control empleados para seleccionarlas y combinarlas adecuadamente. Su versatilidad y potencia es mayor. Sin embargo, también es mayor su complejidad, por lo que la realización de un sistema integral real supone un enorme desafío para la comunidad científica. Puesto que el aprendizaje en el ser humano es claramente multiestratégico, la investigación en este tipo de sistemas tiene una relevancia significativa para el entendimiento del mismo.

En este artículo se presenta un sistema de aprendizaje multiestratégico, que se sustenta en la idea de que toda estrategia de aprendizaje es el resultado de determinados procesos de inferencia elementales que pueden complementarse, trabajando de modo recurrente sobre un espacio con una adecuada representación, para aprender nuevo conocimiento o mejorar el ya existente. Los procesos de inferencia son procesos de ensamblado y de transformación de las estructuras presentes en ese espacio.

2. Proceso de aprendizaje

Un proceso o tarea de aprendizaje se define en función de la información de entrada, del conocimiento presente, del objetivo u objetivos a alcanzar y de la estrategia empleada para ello.

El proceso de aprendizaje que se presenta en este artículo recibe como entrada un conjunto de ejemplos y contraejemplos de un concepto. Tiene como objetivo aprender la descripción del concepto a partir de los ejemplos suministrados. Los ejemplos del concepto constituyen ejemplos positivos y los contraejemplos son ejemplos negativos. Dicho de otro modo, el objetivo, independiente del dominio de aplicación, es el de encontrar una definición intensional de un concepto partiendo de una definición extensional del mismo.

El lenguaje de representación para ejemplos y conocimiento es la lógica proposicional, mediante la que se pueden aprender descripciones basadas en atributos. Este tipo de descripciones son las más empleadas, ya que la mayoría de las bases de datos más conocidas caracterizan los ejemplos mediante atributos. Las descripciones finales sólo contienen atributos presentes en los ejemplos de entrada. Puesto que se parte de ejemplos de diferentes clases, el sistema pretende obtener descripciones discriminantes de un concepto que distingan la clase de ejemplos que satisfacen el concepto de un número limitado de otras clases. Estas descripciones cumplen la condición de ser completas y consistentes, es decir, deben cubrir todos los ejemplos de una clase y ningún ejemplo de las restantes. Las descripciones están formadas por atributos unidos mediante conectivas lógicas (conjunción, disyunción e implicación). Un ejemplo de una clase viene dado como una conjunción de pares atributo-valor, denominados selectores (Michalski 98). Una descripción de un concepto estará formada por la disyunción de expresiones conjuntivas. Con esta descripción se construye una regla de decisión que asocia la descripción del concepto con el nombre del mismo a través de la implicación.

El sistema de aprendizaje propuesto cuenta con conocimiento genérico, que puede ser tanto dependiente como independiente del dominio al que pertenecen los ejemplos.

Para lograr el objetivo de aprendizaje, las estrategias de aprendizaje del sistema se han descompuesto en procesos de inferencia básicos (Michalski 93) que han de integrarse de un modo adecuado para resolver el objetivo propuesto: encontrar la descripción intensional de un concepto.

2.1. Un pequeño problema de aprendizaje supervisado

La operativa habitual de un ser humano ante un problema de aprendizaje supervisado, de magnitud razonable como para ser abordado manualmente, comienza detectando los distintos valores de los atributos que intervienen en los ejemplos positivos. La Figura 1 muestra un problema de este tipo.

Una vez reconocidos los selectores de interés, se pasa a comprobar si alguno de ellos aparece sólo en los ejemplos positivos -se trataría de encontrar, en el menor número de pasos posibles, algún selector consistente-. Si así fuera, el siguiente paso es observar si ese selector aparece en todos los ejemplos positivos -se tendría de ese modo una expresión consistente y completa y, por tanto, una solución al problema planteado-. Cuando el selector no es completo, se busca otro selector que de nuevo sólo cubra ejemplos positivos y se realiza la “o lógica” entre ambos.

Si no existe ningún selector consistente, se elige el de mayor consistencia y se realiza la “y lógica” con otro selector de alta consistencia. Esto se repite hasta lograr una expresión consistente. Por último, la expresión consistente obtenida se intenta completar.

Resumiendo, cuando se afronta un proceso de aprendizaje de este tipo, la secuencia de operaciones llevada a cabo es:

1. Se consideran exclusivamente los selectores involucrados en la descripción de los ejemplos positivos.
2. Se busca la mínima expresión consistente.

3. Si se encuentra, se generaliza para que sea completa. Con la expresión consistente y completa se acaba el proceso.
4. Si no existe a simple vista una expresión consistente, se especializa la más consistente.
5. Vuelta al paso 2.

Atributo 1	Atributo 2	Atributo 3	Clase
y	n	r	+
x	m	r	+
y	n	s	+
x	n	r	+
x	m	s	-
y	m	t	-
y	n	t	-
z	n	t	-
z	n	r	-
x	n	s	-

Descripción: $r \wedge x \vee r \wedge y \vee y \wedge s \equiv r \wedge (x \vee y) \vee (y \wedge s)$

Figura 1. Definición extensional e intensional del concepto o clase +

3. Arquitectura fluida

Uno de los aspectos esenciales cuando se aborda un problema de aprendizaje automático es el de elegir una representación adecuada. Esta representación no se refiere sólo al formalismo de representación propiamente dicho (reglas, árboles de decisión, o redes neuronales), sino a los atributos empleados para describir los ejemplos, el conocimiento genérico o el resultado del aprendizaje y a las características de los procesos de inferencia empleados. Se trata encontrar una representación efectiva de todo el proceso de aprendizaje.

En este sentido, la arquitectura en la que se basa el sistema propuesto pretende adaptarse a la operativa de un ser humano, desde la percepción de una situación dada hasta la comprensión o experiencia que obtiene de la misma, cuando se enfrenta al problema de encontrar una regla que describa una clase de ejemplos.

La representación de un problema es una consecuencia de la percepción del mismo. Los datos (ejemplos, hechos, conocimiento) que se perciben pasan a través de un proceso de filtrado y organización dando lugar a una representación estructurada que puede tener múltiples usos (Hosftadter 95).

La mayor parte del trabajo en inteligencia artificial ha conducido a modelos sobre los procesos de aprendizaje independientes de los procesos de percepción, estando bastante alejados del aprendizaje humano y aquejados de serias restricciones. Uno de los desafíos de esta disciplina es el de entender cómo darle significado al mundo: qué información es relevante y de qué forma hay que representarla. La representación del mundo debe variar en función del contexto y de los conceptos que se ponen en juego al percibirlo.

La percepción es un proceso que pone en marcha conceptos de diferente grado de abstracción. Esto garantiza su flexibilidad, ya que el mismo conjunto de datos de entrada se puede percibir de diferentes formas dependiendo del contexto y del estado del que lo percibe. Esta flexibilidad impide asociar una representación fija a una situación percibida.

La arquitectura que se propone adopta un modelo en el que la representación, en un proceso cognitivo como es el aprendizaje conceptual, no está fijada de antemano, sino que depende del

contexto y de los conceptos o ideas involucrados en el proceso de aprendizaje. El calificativo fluida simboliza la flexibilidad que se pretende de la representación.

Para construir un sistema en el que la representación de una situación, como la descrita en la Figura 1, se adapte a lo que de ella se percibe, se ha diseñado una arquitectura formada por tres bloques o espacios relacionados: espacio conceptual, espacio de trabajo y espacio de operadores. El diseño y el desarrollo de todos los elementos componentes de esta arquitectura está orientado a objeto.

La dinámica de la arquitectura se ajusta al ciclo percepción-representación-operación, siendo cada uno de los espacios responsable de una de las actividades del ciclo. La información que un espacio puede requerir o transferir a otro se materializa mediante unas estructuras de trasvase. Los elementos de cada espacio pueden operar en paralelo y canalizar sus peticiones o sus respuestas a otros a través de dichas estructuras.

Los siguientes apartados describen cada uno de estos espacios ejemplificándolos para el problema de la Figura 1.

3.1. Espacio conceptual

El espacio conceptual reúne las ideas que, a priori, posee el ser humano a la hora de abordar un problema de aprendizaje supervisado. Está formado por conceptos que incluyen todas las posibles descripciones de los objetos percibidos, el tipo de descripciones, así como otros que son propiedades o que relacionan entre sí conceptos. En otras palabras, los conceptos expresan todo aquello que se pueda decir de los objetos que constituyen el problema que se estudia y se intenta resolver.

El espacio conceptual del sistema de aprendizaje se asienta en la filosofía percepción-representación empleada para resolver problemas por analogía en un determinado microdominio (Mitchell 93). De acuerdo a ella, los conceptos se caracterizan por la profundidad conceptual, el grado de activación y los enlaces con otros conceptos.

La extrapolación de esta filosofía a un problema como el del aprendizaje supervisado implica la comprensión del contenido de cada uno de los términos y la adaptación de los mismos. Las ideas que subyacen a un problema del aprendizaje supervisado pueden resumirse en los dos siguientes silogismos:

- Si una expresión es demasiado general hay que especificarla
- Si una expresión es demasiado específica hay que generalizarla

El espacio conceptual debe tenerlos en cuenta estableciendo el conjunto de conceptos y relaciones que los representen. Los objetos que se van a percibir son las expresiones iniciales así como las que se generan durante la evolución del sistema intentando satisfacer el objetivo propuesto. Los conceptos también han de representar las características de una expresión, que pueden ser: consistente, completa, general y específica.

Los conceptos General y Específica están unidos mediante enlaces unidireccionales etiquetados con los conceptos Completa y Consistente.

La profundidad conceptual revela la facilidad para reconocer un concepto en el espacio de trabajo. Los conceptos Completa y Consistente son más fáciles de percibir que Específica o General, con lo cual su profundidad conceptual sería menor.

La activación de los conceptos de este espacio refleja la percepción del contenido del espacio de trabajo. El concepto Específica se activa siempre que existan expresiones susceptibles de ser especializadas y, por tanto, de llegar a ser expresiones consistentes. Lo mismo sucede con el concepto General. El grado de activación de estos conceptos representa el número de operaciones de su categoría que pueden efectuarse sobre el espacio de trabajo.

La activación de un concepto provoca la actuación de los procesos u operadores de inferencia para transformar el espacio de trabajo. El concepto Específica tiene asociados operadores de

especialización deductiva e inductiva y el concepto General tiene asociados operadores de generalización deductiva e inductiva. La activación de los conceptos Completa y Consistente simboliza la existencia de expresiones de esas categorías. Su función es la de permitir el traspaso de activación entre los conceptos General y Específica.

Los conceptos pierden activación siempre que sobre el espacio de trabajo se han efectuado correctamente las operaciones asociadas a ellos.

El espacio conceptual contendrá también el conocimiento dependiente del dominio. La representación de todo el conocimiento es uniforme: conceptos y enlaces que reflejan las relaciones entre ellos.

3.2. Espacio de trabajo

El espacio de trabajo se define como un recipiente en el que las diferentes piezas de información están representadas como fórmulas de lógica proposicional. La aplicación de los operadores de inferencia (especialización, generalización) sobre este espacio permite transformar las mismas generando nuevas expresiones o actualizando las existentes. La evolución del sistema determina que el espacio de trabajo pueda estar constituido por estructuras de información con diferente complejidad. Los operadores de inferencia trabajan sobre dichas estructuras, independientemente de la complejidad de las mismas, sin sufrir ninguna remodelación.

Las expresiones iniciales del espacio de trabajo son los selectores de los ejemplos positivos de la clase. A medida que el sistema evoluciona se crean estructuras derivadas de los selectores iniciales que conviven con ellos.

El estado del espacio de trabajo en cada momento refleja la evolución del problema. Este espacio se ha representado como un grafo formado por nodos que contienen las expresiones generadas unidos a otros nodos mediante enlaces etiquetados con el tipo de operador que los vincula. Una expresión puede formar parte de más de una expresión de mayor orden de generalidad. La Figura 2 muestra el estado final del espacio de trabajo para el problema planteado en la Figura 1.

El espacio de trabajo alberga dos tipos de expresiones: virtuales y reales. Las expresiones virtuales son las expresiones candidatas a ser reales. Son expresiones generadas por la actuación de los operadores inferenciales. El grafo que simboliza el estado del espacio de trabajo está formado sólo por expresiones reales.

Cada expresión tiene asociado un parámetro que indica la importancia que tiene la misma en el contexto de trabajo y con relación a los conceptos del espacio conceptual. Cuando un operador propone al espacio de trabajo una nueva expresión o expresión virtual, el espacio de trabajo decide en función de la importancia de la misma si ésta se va a construir. Por ejemplo, una expresión consistente es más importante que una expresión sólo específica: involucra un concepto de baja profundidad conceptual (Consistente) y cede el trabajo de operadores de especialización a otras expresiones menos importantes, ya que no es posible especializarla más. Además, el espacio de trabajo considera el grado de generalidad de una expresión en relación con las existentes. Si los ejemplos que cubre una expresión virtual son un subconjunto de los ejemplos cubiertos por una expresión real, la primera no se construirá. Si, por el contrario, la expresión virtual es más general que otra expresión real, el espacio de trabajo construirá la primera y destruirá la segunda.

Estos parámetros juegan un papel fundamental intentando impedir un crecimiento desordenado del conocimiento que los procesos de inferencia generan. De esta manera se establece un mecanismo de ponderación dinámica de las expresiones que permite valorar cada una de acuerdo a los objetivos del sistema.

La construcción de expresiones en el espacio de trabajo provoca la actualización de los conceptos del espacio conceptual.

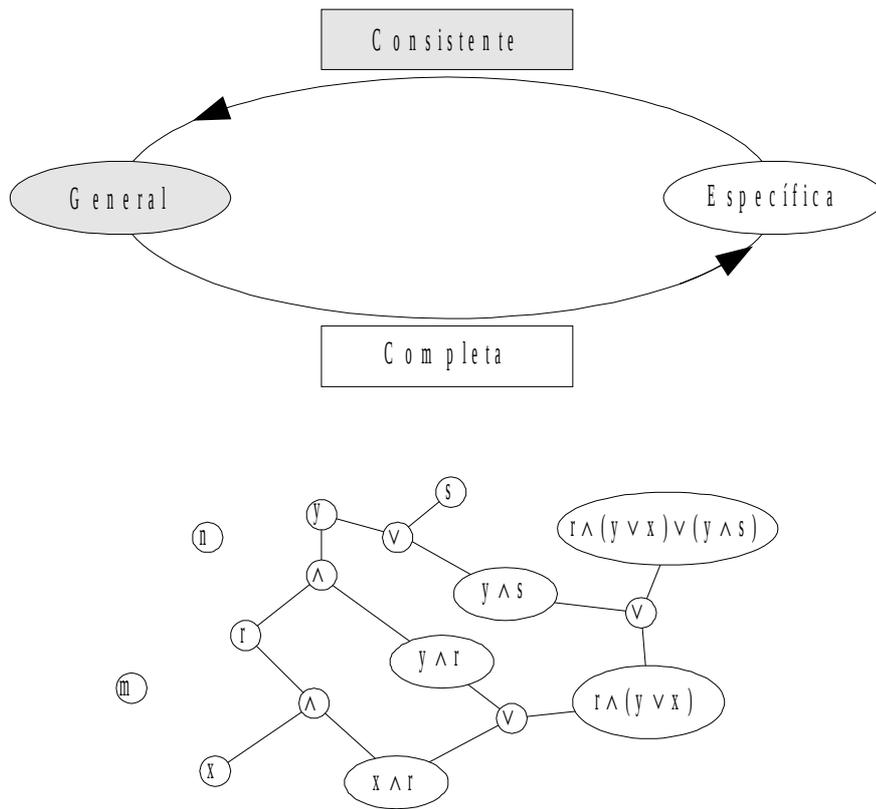


Figura 2. Parte del espacio conceptual y estado final del espacio de trabajo para el problema propuesto

El algoritmo de búsqueda de la descripción del concepto que cubre un conjunto de ejemplos es de cobertura simultánea (Mitchell 97): elige las descripciones más generales sin eliminar los ejemplos cubiertos. Este enfoque es válido cuando existen muestras con un gran número de ejemplos ya que el número de preguntas independientes a realizar es menor.

La búsqueda comienza con la descripción más general: de la descripción vacía se obtienen descripciones sucesivas por especialización. Es una búsqueda en rayo, que mantiene un número determinado de candidatos en cada paso: los de mejor valor de consistencia y más completos.

La generación de las descripciones no está guiada por los ejemplos de la muestra, sino que se producen todas las descripciones sintácticamente válidas. Esto permite minimizar el efecto del ruido en los ejemplos de entrada, ya que cada descripción se evalúa sobre muchos ejemplos.

3.3. Espacio de operadores

En este sistema se propone, en lugar de integrar las estrategias de aprendizaje a un nivel macroscópico, la integración de los diferentes procesos de inferencia que constituyen las estrategias individuales, que muchas de ellas comparten. Un proceso de inferencia es cualquier tipo de razonamiento o transformación del conocimiento llevada a cabo por un agente.

Los procesos u operadores de inferencia se ocupan de la parte de operación del ciclo tripartita que realiza el sistema. Su función es la de satisfacer los objetivos primarios de los conceptos del espacio conceptual que, en el caso que se presenta, consiste en buscar expresiones completas y consistentes.

El objetivo de cada operador será el de razonar a partir de la entrada suministrada y del conocimiento de que dispone.

Con el objetivo planteado, este espacio consta de cuatro clases de operadores: especialización inductiva, especialización deductiva, generalización inductiva y generalización deductiva. De forma esquemática, el proceso llevado por cada uno de ellos es:

- especialización inductiva: $at1 = \{x, y, z\}$ $at2 = \{r, s\}$: $D_i : at1 \wedge r \Rightarrow D_{i+1} = x \wedge r$
- especialización deductiva: $at1 = \{x\}$ $at2 = \{r, s\}$: $D_i : r \Rightarrow D_{i+1} = x \wedge r$
- generalización inductiva: $at1 = \{x, y, z\}$: $D_i : x \Rightarrow D_{i+1} = x \vee r$
- generalización deductiva: $at1 = \{x, y\}$ $x \subset y$: $D_i : x \Rightarrow D_{i+1} = y$

Los operadores trabajan del mismo modo en diferentes iteraciones del ciclo de aprendizaje. Es decir, no es necesario reestructurar su comportamiento sobre las expresiones del espacio de trabajo, a pesar del aumento de complejidad de las mismas. Cada operador toma dos expresiones del espacio de trabajo y le propone a éste la expresión resultante o expresión virtual. La actuación de cada operador tiene la pretensión de mejorar el estado del espacio de trabajo, obteniendo descripciones que sean solución al problema propuesto. Sólo propone una expresión si la importancia de la misma mejora con relación a la importancia de las dos expresiones que recibe.

El número de ejemplares de una clase de operadores depende de la activación de los conceptos que los ponen en marcha.

4. Estrategia de control

La integración de los diferentes procesos de inferencia esencialmente persigue construir unidades de conocimiento mayores (descripciones) a partir de unidades más pequeñas (ejemplos), generando estructuras transitorias o provisionales (descripciones virtuales) que intenten satisfacer categorías prefijadas (conceptos).

Cada proceso podrá operar independientemente del resto, bastará con que se cumplan sus condiciones de aplicación. De este modo, es posible llevar a cabo la búsqueda de descripciones en paralelo. Mientras que un proceso de inferencia inductiva generaliza un conjunto de ejemplos produciendo diversas descripciones, un proceso de inferencia deductiva puede encontrar que alguna de ellas es la premisa de una de las reglas deductivas del espacio de conocimiento.

No obstante, ha sido necesario imponer restricciones en el orden de aplicación de los mismos para asegurar un comportamiento óptimo del sistema en lo que respecta al uso de sus recursos. Estos criterios pretenden evitar la intervención simultánea de procesos de búsqueda innecesarios en favor de otros que son más eficientes en la generación de conocimiento o que un proceso pueda destruir las descripciones inmediatamente producidas por otro.

Con todo ello se pone de manifiesto nuevamente la semejanza del sistema con la manera en la que el ser humano resuelve los problemas: emplea diversas líneas de razonamiento y toma decisiones en función de las conclusiones que se obtengan de ellas. Si todas las líneas conducen a la misma conclusión, es muy probable que ésta sea acertada. Si, por el contrario, conducen a conclusiones contradictorias y todas las líneas tienen la misma fuerza, será muy difícil adoptar una conclusión final. Sin embargo, aunque uno podría poner en marcha en principio todas sus facultades de razonamiento en paralelo, las condiciones del problema rápidamente determinan qué líneas son más prometedoras y cuáles de ellas pueden postergarse. El proceso de aprendizaje finaliza cuando no existe activación en ningún nodo de la red conceptual.

5. Integración de estrategias

La integración de técnicas diferentes exige la caracterización de cada una de ellas en términos de parámetros comunes con el fin de establecer una representación única de los ejemplos de aprendizaje y del resultado de los procesos de inferencia.

Puesto que cada estrategia individual de aprendizaje es, en última instancia, la consecuencia de determinados procesos más elementales de inferencia, la integración de las estrategias de aprendizaje se plantea al nivel de los diferentes procesos de inferencia que las componen.

La mayoría de los sistemas multiestratégicos, hasta la actualidad, abordan la integración de diversas estrategias mediante dos tipos de solución: competitiva y cooperativa. El primer tipo está orientado a la selección del mejor aprendiz, mientras que el segundo pretende obtener el mejor resultado global utilizando las aportaciones que cada uno puede hacer a los demás. Las desventajas de cualquiera de estas soluciones radica en la complejidad de los métodos de control empleados para seleccionar las estrategias, la dificultad para construir estructuras de conocimiento válidas para todas ellas y, por tanto, la dificultad para compartir realmente el conocimiento. El sistema desarrollado pretende superar estas desventajas.

Es cooperativo en el sentido de que todas las estrategias contribuyen a un espacio de trabajo común y unas estrategias reutilizan las estructuras generadas por otras. Es competitivo en el sentido de que el espacio de trabajo sólo incorpora una estructura propuesta por un operador si ésta es mejor que otras ya creadas o que las que simultáneamente estén proponiendo otros operadores inferenciales. Como se ha descrito en el espacio de trabajo, la calidad de una estructura está determinada tanto por el espacio conceptual como por el espacio de trabajo.

6. Conclusiones

En este artículo se han expuesto las características de un sistema de aprendizaje multiestratégico con un objetivo como el de aprender la descripción del concepto que satisfacen un conjunto de ejemplos. La arquitectura en la que descansa el sistema propone la integración cooperativa y competitiva de diferentes procesos de inferencia.

La estrategia de control seguida por el sistema se encuentra distribuida en los tres bloques que lo componen. No existe ningún módulo central que gobierne el funcionamiento del sistema. Las comunicaciones entre los diferentes bloques se ha resuelto de una forma sencilla.

Se propone un bucle de ejecución infinito que ponga en marcha secuencialmente el espacio conceptual, el espacio de trabajo y el espacio de operadores. Si la actuación de cada una de los bloques es pequeña se obtiene una buena simulación de la concurrencia de todos ellos y una analogía mayor con el proceso de percepción-representación del ser humano.

Las descripciones obtenidas por el sistema residen en el espacio de trabajo. Actualmente se está estudiando la incorporación de este conocimiento al espacio conceptual y todos los aspectos a considerar en los tres espacios con relación a ello.

Agradecimientos

El trabajo presentado en este artículo ha sido posible gracias al proyecto de investigación financiado por la Comunidad de Madrid 07T/0028/97.

Referencias

- Mitchell, M. (1993): *Analogy-Making as Perception*. MIT Press, 1993.
- Michalski, R.S. (1993): Inferential Theory of Learning as a Conceptual Basis for Multistrategy Learning. *Machine Learning*, 11, 1993.
- Hosftadter, D. (1995): *Fluid Concepts and Creative Analogies*. Basic Books, 1995.
- Mitchell, T.M. (1997): *Machine Learning*. McGraw-Hill, 1997.
- Michalski, R.S., Bratko, I. y Kubat M. (1998): *Machine Learning and Data Mining, Methods and Applications*. Wiley, 1998.