

1 PRINCIPIOS BÁSICOS DE LOS SISTEMAS MULTIAGENTE

El capítulo se dedica a la presentación de los Sistemas Multiagente. Estos sistemas se caracterizan por estar compuestos por múltiples entidades inteligentes que interactúan entre sí para la consecución de un conjunto de objetivos o la realización de tareas.

Con el fin de facilitar una primera aproximación a esta línea de investigación se explica y motiva su aparición y conceptos básicos en relación con otras ciencias. La importancia actual de los Sistemas Multiagente se hace patente en el elevado número de contribuciones que se dan tanto a nivel de investigación como de aplicación.

1.1 Algunos conceptos básicos

La evolución de los sistemas centralizados hacia la inteligencia distribuida se hace más patente y necesaria cada día. La complejidad alcanzada por sistemas industriales que engloban subsistemas heterogéneos y distribuidos espacialmente, hace de estos, entre otros casos, un claro ejemplo de la necesidad de distribuir las actividades y la inteligencia.

Las razones que llevan a este ámbito de investigación, originado en la década de los ochenta, son:

- Existencia de problemas físicamente distribuidos.
- Problemas ampliamente distribuidos y heterogéneos en cuanto a funcionalidad.
- Necesidad de un punto de vista distribuido para el estudio de sistemas de redes.
- Elevada complejidad de algunos problemas que hace necesario un estudio inicial de los mismos desde un punto de vista local.
- Necesidad de sistemas que sean capaces de adaptarse a cambios en su estructura o en el entorno al que pertenecen.
- Evolución de la ingeniería de software hacia diseños que usan conceptos de unidades autónomas que interactúan entre sí.

Actualmente, la inteligencia artificial distribuida (*Distributed Artificial Intelligence*, DAI) constituye un campo de investigación asentado y prometedor al mismo tiempo. Ha surgido de la confluencia de múltiples disciplinas tales como la inteligencia artificial, la informática, la sociología, la economía, la investigación operativa y la filosofía, lo que le confiere un amplio espectro de posibilidades.

Una rama dentro de la inteligencia artificial distribuida se dedica al estudio, la construcción y la aplicación de los denominados sistemas multiagente (*Multiagent Systems*, MAS). Se

entiende por sistema multiagente a aquel en el que varios agentes inteligentes interactúan para la consecución de ciertos objetivos o la realización de tareas (Weiss, 1999).

La interacción entre los agentes, basada en la comunicación (ya sea para transmitir información o inducir conductas en los otros), hace posible la cooperación para la consecución de determinados objetivos. La cooperación conlleva la colaboración entre agentes, coordinación de sus acciones y la resolución de los posibles conflictos que pueden aparecer en el proceso. El problema de la adaptación también es de máximo interés. Este puede tratarse desde el punto de vista individual, hablándose entonces de aprendizaje, o desde el punto de vista colectivo de la población, tratándose así de evolución.

La inteligencia del sistema se basa en su comportamiento social, en las interacciones que tienen lugar entre los agentes, a partir de las cuales, emerge el comportamiento global deseado.

La gran importancia y el interés conseguido por este ámbito de investigación se debe a las razones ya expuestas con anterioridad. Dos de ellas, fundamentalmente, constituyen las fuerzas desencadenantes del estudio de la inteligencia artificial, tal y como hoy se conoce. Por un lado, la DAI supuso una posible solución para el problema del tratamiento y manejo de grandes cantidades de datos y aplicaciones distribuidas en el ámbito de la informática; y por otro lado, los sistemas multiagente constituyen una herramienta adecuada para el análisis de modelos y teorías de interacción en la sociedad humana, tanto desde el punto de vista de la sociología como desde el de la psicología.

Si se toman los sistemas multiagente (*multiagent systems*, MAS) como objeto fundamental de análisis, es necesario establecer claramente qué se entiende por ellos.

1.1.1 Sistemas multiagente

Un sistema multiagente es aquel que se compone de los siguientes elementos (Ferber, 1999):

- Un entorno, es decir, un espacio.
- Un conjunto de objetos situados en el entorno.
- Un conjunto de agentes (objetos específicos) que representan entidades activas.
- Una serie de relaciones que conectan los objetos entre sí.
- Una batería de operaciones que permite a los agentes percibir, producir, consumir, transformar y manipular objetos del entorno.
- Un conjunto de operadores que representan el efecto de las operaciones en el entorno, las denominadas leyes del universo.

Las características, por excelencia, de un sistema multiagente (Jennings et al., 1998) son:

- Cada agente tiene capacidades limitadas y dispone de una información incompleta.
- Los datos están descentralizados.
- El control está distribuido.
- La computación es asíncrona.

1.1.2 Definición y tipos de agentes

Los agentes pueden verse como entidades inteligentes por sí mismas, o como simples seres que reaccionan a cambios en el entorno.

Un agente puede definirse como una entidad computacional que percibe el entorno en el que se encuentra, que es capaz de actuar en él y que tiene un cierto comportamiento autónomo.

Como entidad inteligente, un agente se puede comportar de forma flexible y racional frente a determinadas circunstancias, basándose para ello en sus percepciones y su capacidad de actuación. Por ser una entidad interactiva, también se ve afectada por las actividades realizadas por otros agentes o elementos en su entorno. Una posible clasificación de los mismos es la que se aprecia en la Tabla 1.

Tabla 1: Tipos de agentes (Fuente: Ferber, 1999).

Relación con el mundo		
Conducta	Agentes cognitivos	Agentes reactivos
Teleonómica	Agentes intencionales	Agentes de base guiada
Reflejos	Agentes de base modulada	Agentes tropísticos

Esta clasificación divide a los agentes según cómo sea su relación con el mundo y el comportamiento que los mismos manifiesten en este. Cuando la conducta del agente esté dirigida a la consecución de un objetivo se denomina teleonómica. Si esta conducta está regulada por las percepciones del agente, entonces se agrupa dentro de la categoría de reflejos.

En general, son agentes cognitivos aquellos que poseen una representación simbólica y específica del mundo, lo que les permite razonar sobre este. Por el contrario, los agentes reactivos interactúan con el mundo según sus capacidades sensomotoras.

Agentes intencionales son aquellos que tienen objetivos específicos propios que motivan sus acciones; los agentes de base modulada son agentes cognitivos que realizan tareas que no forman parte de objetivos propios, sino encargadas por otros.

Se denominan agentes tropísticos a agentes reactivos cuyo comportamiento está determinado por los estímulos del entorno. Los agentes reactivos cuyo comportamiento es guiado por la consecución de unos objetivos, ya sean internos o marcados por el diseñador del sistema, forman el grupo de agentes de base guiada.

Los agentes reactivos presentan un comportamiento menos flexible que los cognitivos; no obstante, disponen de un gran potencial basado en su capacidad para formar grupos. Esta característica les permite adaptarse a cambios en el entorno. Por otra parte, los agentes cognitivos ofrecen la posibilidad de adelantarse a sucesos futuros y prepararse para ellos. Para conseguir esto, usan su capacidad para representar el mundo, memorizar situaciones y analizarlas.

Salvo para análisis detallados de sistemas multiagente, en los que es necesario precisar las estructuras a un elevado nivel de detalle, se distingue básicamente, entre agentes reactivos y cognitivos únicamente. No obstante, algunas otras tendencias añaden un tercer tipo de agentes a este mismo nivel: agente social (López, 2004).

Atendiendo a la definición y caracterización de un agente que se ha presentado al inicio del apartado, es posible clasificar los tipos de agentes en función del modelado de sus capacidades. Esta clasificación resulta más apropiada que la de la Tabla 1, ya que en ella se profundizaba en un agente de determinadas características que circunstancialmente se comportaba de un modo u otro.

- Agentes reactivos: actúan (reaccionan) a cambios en el entorno según unas reglas prefijadas.
- Agentes cognitivos: poseen además de las características del tipo anterior, capacidad para razonar y decidir su actuación según sus motivaciones o necesidades.
- Agentes sociales: tienen capacidad para representar al resto de agentes de su entorno y pueden razonar teniendo en cuenta las intenciones y actuaciones de estos.

La flexibilidad en los comportamientos de los agentes, así como su racionalidad, se basan en distintos procesos. Entre estos procesos se pueden citar los de resolución de problemas, planificación, toma de decisiones y aprendizaje.

En general, un agente se caracteriza, básicamente, por su arquitectura (estructura interna) y su comportamiento (propiedades que manifiesta en el entorno). En consecuencia y desde el punto de vista del diseño, el comportamiento constituye la fuente de especificaciones externas a alcanzar mediante el adecuado diseño de la arquitectura interna del agente.

1.1.3 Arquitectura del agente

Desde los inicios de la DAI hasta ahora, distintas líneas han surgido para abordar las posibles arquitecturas de un agente. Como ya se ha dicho, la arquitectura del agente debe ser adecuada

al comportamiento que este ha de mostrar en el entorno y, además, tratará de ajustarse a los conceptos teóricos sobre los que se apoya el diseñador del sistema.

Los tipos de arquitecturas existentes son:

- Arquitectura deliberativa.
- Arquitectura reactiva.
- Arquitectura de niveles.
- Arquitectura de razonamiento práctico.

La arquitectura deliberativa se caracteriza por considerar el proceso de toma de decisiones como un proceso deductivo. En este caso, el agente dispone de una base de teoría lógica y otras bases de conocimiento, información. Para tomar una decisión, el agente aplica las reglas lógicas que conoce a los datos de los que dispone sobre el problema.

Este tipo de arquitectura es inapropiado para entornos complejos y dinámicos. No es válida tampoco para situaciones que imponen restricciones temporales a la toma de decisiones. Su inconveniente fundamental reside en asumir que no se produce ningún cambio en el entorno durante el proceso deductivo y que la decisión final será razonable.

El enfoque reactivo propone una arquitectura de control del agente que carece de representación simbólica alguna o de algún tipo de razonamiento. Una de las primeras representaciones de esta línea fue la propuesta por Brooks (1991), la denominada arquitectura subsumida (*subsumption architecture*). Este enfoque consistía, básicamente, en una colección de comportamientos para la implementación de tareas. Cada uno de los comportamientos venía dado por una máquina de estado finito que mapeaba las percepciones de entrada a acciones de salida.

La inexistencia de representación del mundo, del entorno circundante para un agente, dificulta la comprensión de la relación entre los comportamientos locales, el entorno y el comportamiento global emergente (Huhns y Stephens, 1999). Los inconvenientes de esta arquitectura hicieron que cayera en desuso a favor de arquitecturas que hibridaban las puramente reactivas y las deliberativas.

Las arquitecturas de niveles se desarrollaron tomando como base un conjunto de capas o niveles software (*software layers*), los cuales podían distribuirse horizontal o verticalmente. En este último caso, una única capa dispone de acceso al medio, tal y como se muestra en la Figura 1.

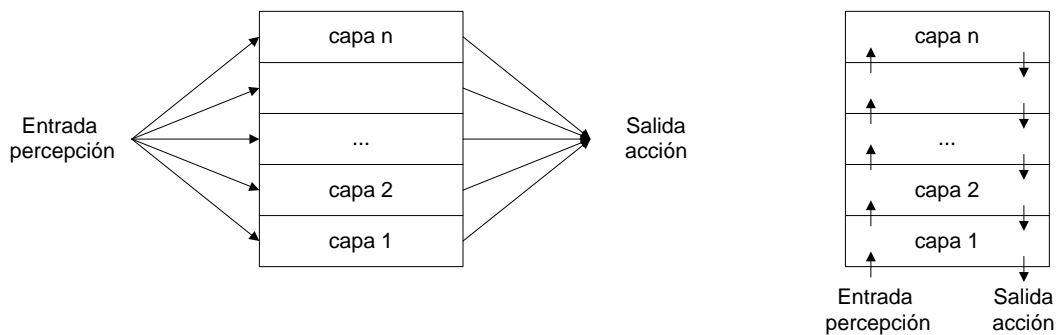


Figura 1: Capas software distribuidas horizontal y verticalmente (Fuente: Müller et al., 1995).

Las distintas capas se organizan según una estructura jerárquica para tratar la información del entorno a distintos niveles de abstracción. En la mayoría de los casos, se consideran suficientes tres niveles:

- Nivel inferior: de carácter reactivo, maneja directamente la información de entrada.
- Nivel medio: hace uso de la representación simbólica que el agente tiene.
- Nivel superior: trata aspectos sociales del entorno, como por ejemplo, la representación de otros agentes.

La principal ventaja de este enfoque es que presenta una descomposición natural de la funcionalidad, pues se distinguen claramente el comportamiento reactivo, proactivo y social del agente. No obstante, como inconvenientes presenta la falta de una semántica clara en la toma de decisiones, así como la definición de las interacciones entre los distintos niveles (Wooldridge, 1999).

Otra de las líneas tradicionales para el desarrollo de la arquitectura del agente es la marcada por los agentes de razonamiento práctico (*practical reasoning*) (Bratman et al., 1998). Este enfoque toma su inspiración de las teorías de razonamiento humano. El mayor exponente de este tipo de arquitecturas es el denominado modelo BDI, del acrónimo en inglés *Belief, Desire, Intention* (Georgeff y Lansky, 1983).

Como su propio nombre indica, los agentes BDI se caracterizan por tres componentes:

- Creencias: representa la información que el agente tiene de su entorno.
- Deseos: opciones de las que dispone el agente y con las que se compromete.
- Intenciones: estados elegidos por el agente y con los que ha comprometido sus recursos.

Este modelo ha sido y es uno de los más usados en la implementación de sistemas multiagente en distintas aplicaciones, como por ejemplo en (Atkison et al., 2005) o (Vicari et al., 2003). No obstante, como todos, presenta dificultades y/o inconvenientes. Una de las mayores dificultades de las arquitecturas BDI hace referencia a las posibilidades de escalabilidad.

1.2 Revisión histórica y evolución de los sistemas multiagente

Pese a la relativa juventud de este ámbito de investigación, las aportaciones al mismo son muy numerosas, habiéndose diferenciado a los pocos años de su aparición diversas líneas dentro de él.

El elevado número de aportaciones, así como su gran aplicabilidad hacen que hoy día esté presente en numerosos ámbitos de la ciencia y de la industria.

1.2.1 Orígenes y tendencias en los sistemas multiagente

Los sistemas multiagente nacen en la confluencia de las disciplinas de Inteligencia Artificial Distribuida (*Distributed Artificial Intelligence*, DAI) y Vida Artificial (*Artificial Life*, AL).

La primera de ellas tiene como objetivo la creación de organizaciones capaces de resolver problemas mediante el razonamiento basado en la manipulación de símbolos, de donde emana la inteligencia. La Vida Artificial, por su parte, trata de comprender y modelar los sistemas vivos, que son capaces de adaptarse, reproducirse, etc. incluso en entornos hostiles. La Vida Artificial se centra, por tanto, en el estudio del comportamiento, la autonomía y cuestiones de viabilidad.

Desde su aparición, los MAS han sido amplia y profusamente estudiados. Algunas líneas de investigación se han centrado en el estudio de los sistemas multiagente cognitivos, otras en la formalización de los agentes racionales autónomos, o bien en la relación de los sistemas multiagente con la teoría de juegos y las técnicas de aprendizaje.

La Tabla 2 resume algunas de las tendencias que se han dado en los últimos años.

Tabla 2: Líneas de Investigación (Fuente: elaboración propia).

Línea de Investigación	Referencias
Estudio de MAS cognitivos	(Lesser, 1999) (Durfee, 2002)
Formalización de agentes racionales autónomos (interacción)	(Jennings, 1998) (Wooldridge, 1999)
DAI y Teoría de Juegos	(Rosenschein, 1994)
Aprendizaje en MAS	(Weiss, Sen, 1999) (Panait y Luke, 2005)

1.2.2 Aplicaciones de los sistemas multiagente

Los dominios de aplicación de los sistemas multiagente se caracterizan por un conjunto de rasgos típicos, tales como modularidad, descentralización, complejidad, capacidad de interacción flexible, dinamismo e inconcreción inicial del problema a resolver (Weiss, 1999).

Los criterios para la clasificación de las aplicaciones de los sistemas multiagente son variados y dependen de los autores. Ferber (1999) los divide según las líneas de investigación en las que se enmarcan; Jennings et al. (1998) atiende al ámbito del sector industrial o profesional en el que se aplica; y Oliveira et al. (1999) distingue en función de la naturaleza software o hardware del agente, refiriéndose a la relación entre el agente y su entorno con el mundo real.

Los MAS pueden verse como una herramienta de resolución de problemas. Como ya se ha dicho, se componen de entidades solucionadoras de problemas, autónomas y capaces de interactuar entre sí para resolver problemas o perseguir objetivos. Esto explica lo adecuados que resultan a numerosos ámbitos de investigación y de la industria. Así, aparecen aplicaciones de MAS para simulación de sistemas sociales, en sistemas de planificación y control de la producción, gestión de la cadena de suministro, sistemas de transporte, etc. La Tabla 3, sin pretender ser exhaustiva en cuanto a la clasificación realizada se refiere, recoge numerosos ejemplos que representan la aplicación de los MAS en estos ámbitos.

Tabla 3: Ejemplos de aplicaciones de los sistemas multiagentes en distintos ámbitos (Fuente: elaboración propia).

Ámbito de aplicación	Referencias
Simulación de sistemas sociales	(Emmanouilides y Davies, 2007) (Koes et al., 2005)
Planificación y Control de la Producción	(Tchikou y Gouardres 2003), (Mahesh et al. 2007) (Lima et al., 2006).
Cadena de Suministro	(Carvalho y Custódio, 2005) (Guo et al., 2004)
Transporte y Logística	(Böcker et al., 2001) (Van Dam et al., 2007)

La simulación de sistemas sociales mediante MAS permite analizar y validar propiedades y modelos teóricos. Para ello, se observan los comportamientos y la interacción de los individuos, así como los resultados que emergen a un nivel de abstracción superior. Esta área es de gran interés para el estudio de las ciencias económicas y sociales. La simulación mediante el empleo de MAS supuso un salto cualitativo y cuantitativo con respecto a las técnicas de simulación anteriores. Esta nueva aproximación resulta más apropiada al ámbito de estudio y, además, permite englobar a los métodos anteriormente usados (López, 2004).

En cuanto a la aplicación de MAS en los sistemas de producción, esta está motivada por las características y exigencias propias de los mercados actuales. Estos mercados requieren que los nuevos sistemas de planificación y control de la producción sean flexibles, ágiles, capaces

de adaptarse a los cambios en el ambiente, al mismo tiempo que deben ser tolerantes a fallos e inteligentes, características que le son propias a los MAS (Caridi et al., 2004).

Por su parte, la aparición del concepto de cadena de suministro también exigió el uso de técnicas de análisis y estudio distribuido, que se alejaron de los tradicionales enfoques centralizados, (Moyaux et al., 2006).

La gestión de sistemas de transporte y las aplicaciones de logística de transporte tienen también características que las hacen adecuadas para el uso de MAS. En este ámbito, la forma de incursión de los MAS más habitual es como sistema de apoyo a la toma de decisiones, (*Decision Support System, DSS*) (Davidsson et al., 2004).