

Tecnologías de Acuerdo en Modelos de Razonamiento Práctico

Marcelo Errecalde, Guillermo Aguirre, Edgardo Ferretti, Cecilia Sosa

Laboratorio de Investigación y Desarrollo en Inteligencia Computacional
Departamento de Informática, Universidad Nacional de San Luis
Ejército de los Andes 950 - (D5700HHW) San Luis - Argentina
e-mails: {merreca, gaguirre, ferretti}@unsl.edu.ar,
ceciliasosatoranzo@gmail.com

Resumen

En este artículo se describen, en forma resumida, los trabajos de investigación y desarrollo que se están llevando a cabo en la línea de investigación “Sistemas Inteligentes” en la problemática vinculada a la integración de tecnologías de acuerdo dentro de arquitecturas de razonamiento práctico, y se abordan temas vinculados a arquitecturas de agentes inteligentes, arquitecturas BDI y Procesos de Decisión Markov, mecanismos de votación, argumentación y negociación.

Palabras clave: sistemas multi-agente, tecnologías de acuerdo, toma de decisión.

Contexto

La línea de investigación “Sistemas Inteligentes” se centra en la formalización, diseño y desarrollo de agentes computacionales con capacidades cognitivas de alto nivel. Los enfoques utilizados en nuestra línea de trabajo, buscan dotar a los agentes inteligentes con las capacidades necesarias para enfrentar problemas complejos de la vida real, que involucran ambientes dinámicos, inciertos, limitados en recursos y que son parcialmente observables. Esta línea de investigación for-

ma parte del proyecto “Nuevas tecnologías para el tratamiento integral de datos multimedia”, Proyecto de Investigación consolidado de la Universidad Nacional de San Luis (UNSL), que se centra en la incorporación de información no estructurada (texto, audio, imágenes y video) en la resolución de problemas y la toma de decisiones. Este proyecto recibe financiamiento de la UNSL y de la Comisión Europea de Investigación e Innovación (Marie Curie Actions: FP7-People-2010-IRSES).

Introducción

Los sistemas computacionales distribuidos han evolucionado desde arquitecturas tradicionales basadas en el modelo cliente/servidor, a enfoques donde las componentes del sistema exhiben patrones de interacción más flexibles y elaborados, que incluyen comportamiento pro-activo (basado en objetivos), mecanismos de negociación, de argumentación, etc. A partir de esta tendencia, gran parte de las técnicas pensadas como mecanismos de interacción social de humanos, han pasado a convertirse en el fundamento de muchos sistemas computacionales donde las componentes interactúan utilizando protocolos de argumentación [12, 2], votación [8, 15, 23],

negociación [18, 14] y mercado [20]. Si bien estos mecanismos de interacción han sido frecuentemente estudiados por separado, buena parte de la comunidad científica vinculada al área, ha comenzado a reconocer la importancia de darle un tratamiento integral a estos enfoques y estudiarlos bajo el “paraguas conceptual” denominado *tecnologías de acuerdo* (en inglés, *agreement technologies (AT)*). Esta tendencia ha quedado de manifiesto en el surgimiento de consorcios de investigación europeos especializados en la temática¹ como así también en la inclusión de Workshops sobre el tema en las principales conferencias internacionales.² Como se reconoce en estos círculos científicos, “el *acuerdo* es uno de los conceptos sociales que ayuda a los humanos a desempeñarse en sus ambientes sociales y está presente en todas las interacciones humanas. En efecto, sin acuerdos no hay cooperación y finalmente los sistemas sociales no pueden emerger. Por lo tanto, el acuerdo es una componente fundamental y necesaria en nuestras actividades diarias”. Sin lugar a dudas, los *agentes artificiales inteligentes (AAI)* [19] y los *sistemas multi-agente (SMA)* [25, 24] han constituido el marco de trabajo computacional por excelencia, donde la mayoría de las AT se han investigado y utilizado en problemas concretos de la vida real.

Un agente es una entidad computacional autónoma, virtual (programa) o física (por ejemplo un robot), que puede percibir directamente su entorno a través de algún tipo de sensor y que es capaz de interactuar con ese entorno utilizando efectores. La principal fortaleza de los agentes reside en su capacidad para exhibir un comportamiento flexible, que en este contexto, refiere a la capacidad de:

- Percibir directamente un ambiente dinámico y reaccionar oportunamente a eventos y condiciones cambiantes (reactividad).
- Tomar la iniciativa cuando sea nece-

sario e iniciar comportamientos dirigidos por un objetivo (pro-actividad).

- Interactuar y comunicarse, cuando es apropiado, con otros agentes artificiales o humanos (sociabilidad).

Por otra parte, en un SMA un conjunto de agentes interactúa para conseguir algún objetivo. Usualmente, en este tipo de sistemas, cada agente posee información incompleta sobre su entorno y sobre sí mismo, teniendo además capacidades limitadas. El entorno es dinámico y cambiante, viéndose afectado por las acciones de todos los agentes. Por esta razón, no puede predecirse con certeza el estado futuro de ese entorno. El control del sistema es distribuido, los datos están descentralizados, y la computación es asincrónica. Las características mencionadas previamente, que aumentan la complejidad del sistema, proveen también las siguientes ventajas: eficiencia (debido al paralelismo que se logra al disponer de múltiples agentes), robustez y confiabilidad (gracias a la redundancia), escalabilidad y reusabilidad.

Proveer a un agente de las capacidades necesarias para exhibir un comportamiento flexible, no es una tarea trivial. De hecho, una de las áreas de investigación más activa en el ámbito de agentes ha sido la definición de modelos y arquitecturas de agentes que intentan dar una respuesta a este problema. Existen arquitecturas que se han concentrado en el aspecto de la reactividad [1, 6] y otras en cambio que han privilegiado los mecanismos de deliberación y planificación necesarios para proveer un comportamiento pro-activo [16]. Sin embargo, existe actualmente un consenso generalizado de que cualquier arquitectura realista de agente, debería proveer un soporte adecuado para todas estas capacidades. Las arquitecturas híbridas [10, 11] y las basadas en comportamientos [15, 17] han intentado lograr un balance adecuado entre reactividad y pro-actividad. Sin embargo, las arquitecturas que mayor atención han recibido para este propósito son las denominadas arquitecturas BDI [5]. De igual

¹<http://www.agreement-technologies.org>

²<http://users.dsic.upv.es/~vinglada/wat2010/wat-2010/Main.html>

manera, aquellas personas que se han orientado a los enfoques basados en utilidad, han tomado a los Procesos de Decisión Markov [3] y los juegos estocásticos [4], como los modelos principales para diseñar e implementar los agentes inteligentes.

En todas las arquitecturas mencionadas previamente, el fin último de los procesos internos y de razonamiento de los agentes, es seleccionar las acciones que éste realizará sobre su ambiente. En este contexto, estos procesos orientados hacia las acciones suelen ser referenciados como de razonamiento práctico [25], para diferenciarlo del razonamiento teórico, que sólo afecta las creencias del agente sobre el mundo. Sin embargo, desde un punto de vista más general, estas propuestas también pueden ser consideradas como modelos y arquitecturas para la toma de decisiones de agentes inteligentes.

Si bien las tecnologías de acuerdo han recibido una atención considerable en el área de SMA, su aplicación dentro de las arquitecturas y modelos de agentes individuales para la toma de decisiones racional es un área que no ha recibido, hasta el momento, la necesaria atención. Más allá de trabajos aislados sobre el uso de votación en sistemas basados en comportamiento [15, 17] y algunos esfuerzos por incorporar argumentación en arquitecturas BDI [21, 22] existe una falencia notoria de propuestas tendientes a sistematizar el uso de estas tecnologías en modelos de toma de decisión ya establecidos, como las arquitecturas BDI de razonamiento práctico y las arquitecturas basadas en utilidades.

Este no es un dato secundario si consideramos que el diseño modular de los agentes, o bien la existencia de múltiples criterios o preferencias personales que pueden entrar en conflicto, conducen naturalmente a situaciones que podrán ser modeladas como un SMA que funciona “dentro de la cabeza del agente”. Esta concepción “societaria” de los procesos mentales internos del agente, ya fue planteada por Minsky [13] en 1986. Pirjanian retoma esta visión en [15] al considerar que dentro de esta jerarquía los agentes cumplen dos roles:

como *acciones* y como *mecanismos de selección de acción*. Con respecto a sus subordinados, un agente es un mecanismo de selección de acción, y con respecto a su superior, es una acción. En [9], Ferber plantea un enfoque similar, donde tanto a los agentes como a los SMA se los considera como *organizaciones*, y un agente se constituye en *componente* de una organización de nivel más alto (el SMA), y a su vez esto se puede generalizar tomando a los agentes de nivel n como las componentes del agente de nivel $n + 1$, para un n arbitrario.

Recientemente, ideas similares a las planteadas previamente se han retomado en el Massachusetts Institute of Technology, al considerar las distintas causas que han hecho que la Inteligencia Artificial (IA) se haya demorado, o directamente detenido, en cumplir algunos de los objetivos inicialmente planteados para la disciplina [7]. En esta discusión, se plantea que “*lo que está faltando es una **ecología de modelos**, un sistema que pueda resolver problemas de distintas maneras*”, como lo hace la mente. La crítica a la investigación en IA, es que hay muchísimas piezas que trabajan bien para resolver **problemas particulares**, pero la gente ha tratado de ajustar *todas las cosas* en alguna de ellas. Lo que se necesita son maneras de “*hacer sistemas compuestos de muchas piezas*” que trabajan conjuntamente como los distintos elementos de la mente. “*En lugar de buscar la bala de plata para resolver el problema, estamos considerando un amplio rango de modelos, y probar integrarlos y agregarlos*”.

Líneas de Investigación y desarrollo

En este contexto, la idea de usar AT pensadas inicialmente para SMA, dentro de los principales modelos de toma de decisión para agentes individuales, surge naturalmente y constituye la idea generadora a partir del cual se derivan los objetivos planteados en la siguiente sección.

Respecto a los dos ejes de investigación planteados previamente: AT, y modelos y arquitecturas de razonamiento práctico, se puede decir que si bien cada uno de ellos ha sido investigado por separado con gran profundidad en los últimos tiempos, la *integración* de mecanismos de acuerdos (votación, argumentación, etc.) *dentro* de modelos clásicos de toma de decisiones (como los modelos BDI o los Procesos Markov) no ha recibido a nuestro entender, la merecida atención en la investigación de AAI y ha dado origen, al área donde se pretende generar las principales contribuciones de esta línea de investigación.

Resultados y objetivos

Las principales motivaciones del grupo surgen de la intención de dar respuestas a las falencias descritas previamente, y plantear contribuciones en el área a partir del cumplimiento de los siguientes objetivos:

1. Lograr un entendimiento más profundo de los distintos enfoques basados en AT, para su uso en distintas componentes y procesos de modelos de razonamiento práctico como las arquitecturas BDI y algoritmos basados en Procesos de Decisión Markov. Este aprendizaje debería incluir la determinación y comparación de las propiedades formales y los aspectos computacionales que involucran estas técnicas.
2. Obtener, mediante un proceso adecuado de generalización, una arquitectura abstracta para la incorporación de este tipo de tecnologías de acuerdos en modelos de toma de decisiones de agentes inteligentes. Asimismo, mediante la instanciación e implementación de esta posible arquitectura, sería conveniente determinar en qué medida esta propuesta se adecua a la resolución de problemas concretos del mundo real que involucren agentes de software y hardware. Como potenciales dominios de aplicación se plantean trabajos con robots

o dispositivos de comunicación móviles (agentes de hardware) y problemas vinculados al área de Web Intelligence [26].

En este contexto, ya se han logrado algunos avances en estos temas, relacionados al uso de votación en los procesos internos del agente [8, 23], mientras que en [21, 22] se desarrolló la primera arquitectura concreta de agente que integra al modelo BDI con servicios Web y razonamiento argumentativo en un mismo framework.

Formación de Recursos Humanos

Trabajos de tesis vinculados con las temáticas descritas previamente:

- 1 tesis Doctoral en ejecución en co-dirección con investigador de la UNS.
- 1 tesis Doctoral y 1 de Maestría finalizadas.
- 1 tesis de Maestría en trámite de inscripción.
- 3 tesis de Licenciatura aprobadas.

Referencias

- [1] R. C. Arkin. *Behaviour-Based Robotics*. The MIT Press, 1998.
- [2] J. Bentahar, R. Alam, and Z. Maaamar. An argumentation-based protocol for conflict resolution. In *Workshop on Knowledge Representation for Agents and Multi-Agent Systems*, 2008.
- [3] C. Boutilier, T. Dean, and S. Hanks. Decision-theoretic planning: Structural assumptions and computational leverage. *J. Artif. Intell. Res. (JAIR)*, 1999.
- [4] M. Bowling and M. Veloso. An analysis of stochastic game theory for multiagent reinforcement learning. Technical Report 00-165, Computer Science Department, Carnegie Mellon University, 2000.

- [5] M. Bratman, D. Israel, and M. Pollack. Plans and resource bounded reasoning. *Computational Intelligence*, 1988.
- [6] R. A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 1986.
- [7] D. Chandler. *Rethinking artificial intelligence*. MIT News Office, 2009.
- [8] M. Errecalde, G. Aguirre, and F. González. Agentes y mecanismos de votación. In *X Congreso Argentino de Ciencias de la Computación*, 2004.
- [9] J. Ferber. *Multi-Agent Systems - An Introduction to Distributed Artificial Intelligence*. Addison-Wesley, 1999.
- [10] E. Gat. Integrating planning and reacting in a heterogeneous asynchronous architecture for mobile robots. *SIGART Bulletin*, 1991.
- [11] E. Gat. On three-layer architectures. In *Artificial Intelligence and Mobile Robots*, 1998.
- [12] A. Kakas, N. Maudet, and P. Moraitis. Layered strategies and protocols for argumentation-based agent interaction. In *Argumentation in Multi-Agent Systems*, pages 64–77, 2005.
- [13] M. Minsky. *The Society of Mind*. Simon and Schuster, 1986.
- [14] H. J. Müller. Negotiation principles. In *Foundations of distributed artificial intelligence*. John Wiley & Sons, 1996.
- [15] P. Pirjanian. *Multiple Objective Action Selection and Behaviour Fusion using Voting*. PhD thesis, Department of Medical Informatics and Image Analysis, Institute of Electronic Systems, Aalborg University, 1998.
- [16] M. E. Pollack. Planning technology for intelligent cognitive orthotics. In *Proceedings of 6th International Conference on AI Planning and Scheduling*, 2002.
- [17] J. Rosenblatt. Damn: A distributed architecture for mobile navigation - thesis summary. In *Journal of Experimental and Theoretical Artificial Intelligence*, 1995.
- [18] J. S. Rosenschein and G. Zlotkin. *Rules of Encounter - Designing Conventions for Automated Negotiation among Computers*. The MIT Press, 1998.
- [19] S. Russel and P. Norvig. *Artificial Intelligence - A Modern Approach*. Prentice Hall, 2010.
- [20] T. Sandholm. Distributed rational decision making. In *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, chapter 5. The MIT Press, 1999.
- [21] F. Schlesinger, M. Errecalde, and G. Aguirre. An approach to integrate web services and argumentation into a BDI system. In *9th Intl. Conf. on Autonomous Agents and Multiagent Systems*, 2010.
- [22] F. Schlesinger, E. Ferretti, M. Errecalde, and G. Aguirre. An Argumentation-based BDI Personal Assistant. In *23rd IEA-AIE*, volume 6096 of *LNAI*. Springer, 2010.
- [23] C. Sosa-Toranzo, F. Schlesinger, E. Ferretti, and M. Errecalde. Integrating a voting protocol within an argumentation-based BDI system. In *XVI Congreso Argentino de Ciencias de la Computación*, 2010.
- [24] G. Weiss, editor. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. The MIT Press, 1999.
- [25] Michael Wooldridge. *An Introduction to MultiAgent Systems*. John Wiley & Sons, Chichester, England, 2002.
- [26] U. Zeeshan-Ul-Hassan, editor. *Web Intelligence and Intelligent Agents*. In-Tech, 2010.