

Mecanismos de confianza y reputación combinados con formalismos de revisión de creencias

Luciano H. Tamargo †
Alejandro J. García †

Marcelo A. Falappa †
Guillermo R. Simari

Laboratorio de Investigación y Desarrollo en Inteligencia Artificial
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur,
† Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Av. Alem 1253, (B8000CPB) Bahía Blanca, Argentina
Tel: (0291) 459-5135 / Fax: (0291) 459-5136
e-mail: {lt, maf, ajg, grs}@cs.uns.edu.ar

Contexto

Esta línea de investigación se realizará dentro del ámbito del Laboratorio de Investigación y Desarrollo en Inteligencia Artificial, y está asociada a los siguientes proyectos de investigación:

“Formalismos Argumentativos aplicados a Sistemas Inteligentes para Toma de Decisiones”. Código: PGI 24/ZN18. Director: Alejandro Javier García. Co-director: Marcelo Alejandro Falappa. Acreditado con evaluación externa para el período 1/1/2009 - 31/12/2011. Financiamiento: Universidad Nacional del Sur.

“Sistemas de Apoyo a la Decisión Basados en Argumentación: formalización y aplicaciones”. Código: PIP 112-200801-02798. Director: Dr. Carlos I. Chesñevar. Período: 2009 - 2011. Financiamiento: CONICET.

Resumen

El objetivo principal de esta línea de investigación es el estudio y desarrollo de técnicas y formalismos para la actualización del grado de confianza asignado a un agente y reputación que estos han obtenido al interactuar en el marco de un sistema multi-agente. Si bien el interés en esta temática es reciente (impulsado por la aparición de sistemas complejos de

información social, como *social networks*), su importancia general para las aplicaciones de sistemas multi-agente es creciente, siendo un ejemplo significativo el comercio electrónico en todas sus formas. Una característica sobresaliente de estos sistemas es la dinámica que debe existir en la confianza y/o reputación asignada a un agente, *i.e.*, la reputación de un agente que no cumple con sus compromisos asumidos debería decrementarse.

La investigación se enfoca, particularmente, en la caracterización y desarrollo de operadores de cambio, que permitan modelar la dinámica de la confianza y reputación de agentes en un sistema. Disponer de un formalismo como el propuesto significará un avance concreto en las áreas antes mencionadas, impulsando la consideración de nuevas aplicaciones que utilicen la tecnología de sistemas multi-agente.

Palabras claves: Revisión de Creencias, Sistemas Multi-Agente, Confianza y Reputación.

1. Introducción

Los agentes podrán actualizar la relación de orden con la cual se representa la reputación de sus pares. Así, estos operadores podrán ser usados para alterar dinámicamente la estructura de la credibilidad de los informantes a fin de reflejar una nueva percepción de la plausibi-

lidad de un informante o la llegada de un nuevo agente al sistema.

Los agentes son entidades computacionales autónomas, ya sea programas o robots, con la capacidad de percibir el entorno en que se desenvuelven y actuar para llevar a cabo alguna tarea. La percepción del entorno por parte de un agente es usualmente limitada, y diferentes agentes dentro de un sistema multi-agente pueden tener una percepción diferente del entorno. Por lo tanto, es importante que los agentes puedan cooperar compartiendo su conocimiento acerca del entorno. Además, si cada agente almacena conocimiento, producto de su experiencia, es importante que pueda compartir este conocimiento con otros agentes con los cuales coopera. De esta manera, agentes especialistas en cierto aspecto podrán intercambiar conocimiento con otros que tienen experiencia en otras áreas. Por lo tanto, en el contexto de sistemas multi-agente, un agente puede a menudo recibir información a través de otro que por lo general llamamos informante. Estos informantes son agentes independientes quienes tienen sus propios intereses y, por lo tanto, no son completamente fiables. Es natural para un agente estar inclinado a creerle más a un informante sobre otro. Es por esto que en algunos trabajos se ha propuesto la organización de los informantes en un orden parcial que compara la plausibilidad de los mismos.

En esta línea de investigación se pretende combinar formalismos de revisión de creencias y actualización de conocimiento con técnicas de mantenimiento de confianza y reputación de agentes en un ambiente distribuido. De esta manera, los agentes podrán actualizar la relación de orden con la cual se representa la reputación de sus pares. Así, estos operadores podrán ser usados para alterar dinámicamente la estructura de la credibilidad de los informantes a fin de reflejar una nueva percepción de la plausibilidad de un informante o la llegada de un nuevo agente al sistema. El estudio y desarrollo de agentes inteligentes, y la investigación en robótica cognitiva han demandado en los últimos años la investigación y desarrollo de formalismos de repre-

sentación y mantenimiento de conocimiento. En particular esto ha dado un fuerte impulso al área de revisión de creencias (belief revision) [1, 11, 12, 17, 13, 10] donde se estudia formalmente parte de estos temas. Además, en los últimos años se han producido importantes avances en el área de revisión de creencias en sistemas multi-agente [8, 7, 5, 15, 16, 14] y en el área de confianza y reputación [20, 3, 6, 18, 19]. Si bien en la literatura existen muchos trabajos realizados en ambas áreas, la combinación de las mismas es algo novedoso.

2. Líneas de investigación y desarrollo

Esta línea de investigación busca combinar resultados teóricos obtenidos en las áreas de sistemas multi-agentes, mecanismos de confianza y reputación y revisión de creencias. El objetivo es actualizar en forma dinámica bases de conocimiento de diferentes agentes en un ambiente distribuido. Disponer de un formalismo como el propuesto permitirá lograr un avance en las áreas antes mencionadas, así como también el desarrollo de nuevas aplicaciones que utilicen la tecnología de sistemas multi-agente. Esto facilitará el desarrollo de aplicaciones basadas en agentes para resolver problemas complejos. Los resultados obtenidos serán utilizados para su aplicación inicialmente en el campo de agentes inteligentes y sistemas multi-agente y posteriormente en robótica cognitiva, sistemas de información así como también sistemas de recomendación.

2.1. Revisión de creencias en sistemas multi-agentes

La meta principal de la teoría del cambio es intentar modelar la dinámica del conocimiento. No hay una única manera de realizar un “cambio” en el conocimiento de un agente, existen numerosas propuestas para ello. El modelo AGM [1] es uno de los principales referentes (expansión, contracción y revisión). La revisión es una de las operaciones más comu-

nes que desarrolla un agente que se desenvuelve en un entorno dinámico. Si un agente cree en un conjunto K , y se produce cierta entrada epistémica X , el mismo deberá hacer lo siguiente: revisar el contenido de K para determinar cuáles de sus creencias están en desacuerdo con X , eliminar algunas de ellas, luego incorporar X y producir un nuevo conjunto K' consistente.

Los primeros modelos de revisión de creencias han sido pensados teniendo en cuenta un único agente. Esto es, consideraban únicamente como se producen cambios en la base de conocimiento de un agente cuando recibe nueva información. Usualmente, la Revisión de Creencias Individual (*IBR: Individual Belief Revision*) en un ambiente de único agente es logrado satisfaciendo o adaptando los postulados AGM.

Con el tiempo los modelos de revisión de creencias evolucionaron hacia sistemas multi-agentes. En muchos dominios y aplicaciones multi-agentes, cada agente tiene sus propias creencias iniciales también como creencias adquiridas desde otros agentes informantes. De esta manera, un agente puede recibir información desde otros agentes que es contradictoria con sus propias creencias actuales. Por lo tanto, Revisión de Creencias Individual (*IBR: Individual Belief Revision*) necesita ser extendida hacia ambientes multi-agentes.

En las áreas de revisión de creencias y sistemas multi-agente se distingue revisión de creencias en sistemas multi-agentes (*MABR: Multi-Agent Belief Revision*), revisión de creencias con múltiples fuentes (*MSBR: Multi-Source Belief Revision*) y revisión de creencias con un único agente (*SBR: Single agent Belief Revision*). MABR investiga el comportamiento de la revisión de creencias global de un equipo o sociedad de agentes en la cual, para alcanzar una meta mutua, los agentes involucrados necesitan comunicarse, cooperar, coordinar, y negociar con otros. Un sistema MABR es un SMA cuya meta mutua involucra revisión de creencias. Diferentes formalismos han sido presentados para tratar con MABR [15, 16, 14]. En cambio, en MSBR, se

lleva a cabo un proceso de revisión de creencias individual en un ambiente multi-agente donde la nueva información puede venir desde múltiples informantes [8, 7, 5].

2.2. Sistemas de mantenimiento de confianza y reputación

La investigación científica en el área de mecanismos computacionales de confianza y reputación en sociedades virtuales, es una disciplina reciente orientada a incrementar la fiabilidad y performance de comunidades electrónicas. En artículos recientes [20, 3, 6, 18, 19] se desprende que el paradigma de agentes autónomos y sistemas multi-agentes junto con la aparición creciente de las tecnologías de información social (especialmente reflejado por la popularidad del comercio electrónico) son los responsables del creciente interés sobre mecanismos de confianza y reputación aplicados a sociedades electrónicas.

En [20], Sabater y Sierra sostienen que la importancia de la confianza y reputación en sociedades humanas está fuera de discusión, por lo tanto, no es sorpresa que varias disciplinas, cada una desde una perspectiva diferente, haya estudiado y utilizado ambos conceptos. En ciencias de la computación hay dos elementos que han contribuido sustancialmente en incrementar el interés en confianza y reputación: el paradigma de sistemas multi-agentes y la evolución creciente del e-commerce. Estos sistemas han sido reconocidos como factores claves para el éxito de la adopción del comercio electrónico. Los mismos son usados por agentes de software inteligentes como un mecanismo para buscar compañeros confiables y como un incentivo en toma de decisiones acerca de si se tiene en cuenta un contrato. La reputación es usada en el mercado electrónico como un mecanismo para evitar fraudes y estafas [2, 9, 6]. Los e-markets no son el único campo de aplicación, por ejemplo, en [3] usan la confianza para mejorar la performance de mecanismos de revisión de creencias.

3. Resultados obtenidos y esperados

Anteriormente en esta línea se ha desarrollado un modelo epistémico para MSBR [21, 22, 23] para el cual hemos propuesto una manera racional de pesar las creencias usando un orden de credibilidad entre agentes. Para ello se definió una función de plausibilidad que es utilizada en la definición de un criterio para comparar las creencias. Además, se han definido diferentes operadores que describen un modelo de cambio completo basado en informantes: expansión, contracción, revisión priorizada y revisión no-priorizada. Para cada uno de ellos hemos dado una definición en forma constructiva y hemos mostrado una caracterización axiomática a través de teoremas de representación. En este formalismo los operadores están basados en un ordenamiento de fuentes (como es sugerido en [8, 7, 5, 4]) y en la técnica de contracción propuesta por Sven Ove Hansson en [11].

El objetivo de esta línea de investigación es proponer un modelo completo de cambio para órdenes parciales de credibilidad de informantes, donde se combinen formalismos de revisión de creencias con técnicas de mantenimiento de confianza y reputación. En base a este nuevo formalismo de revisión de órdenes de credibilidad en sistemas multi-agente, se definirá una caracterización axiomática de los operadores proveyendo de esta manera teoremas de representación que le dará a la propuesta un sólido marco teórico.

La importancia de esta línea de investigación radica en el estudio y desarrollo de nuevos formalismos para su aplicación en el área de sistemas multi-agente. Estos sistemas se utilizan en la actualidad para la resolución de problemas complejos en ciencias de la computación y en otras disciplinas. Su campo de aplicación ha ido en aumento en los últimos años, fundamentalmente por los desarrollos teóricos y prácticos producidos en las áreas de inteligencia artificial distribuida y robótica cognitiva. Se prevee que el campo de aplicación de estos sistemas seguirá en aumento, lo cual requerirá el

desarrollo de nuevos formalismos teóricos para su futura aplicación.

Referencias

- [1] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *J. of Symbolic Logic*, 50(2):510–530, 1985.
- [2] Amazon. Amazon auctions. In <http://auctions.amazon.com>, 2002.
- [3] K. S. Barber and J. Kim. Belief revision process based on trust: Simulation experiments. In *In Proceedings of Autonomous Agents '01 Workshop on Deception, Fraud, and Trust in Agent Societies*, pages 1–12, 2001.
- [4] S. Benferhat, D. Dubois, H. Prade, and M. A. Williams. A practical approach to revising prioritized knowledge bases. *Studia Logica*, 70(1):105–130, 2002.
- [5] J. Cantwell. Resolving conflicting information. *Journal of Logic, Language and Information*, 7(2):191–220, 1998.
- [6] C. Dellarocas. The digitalization of word-of-mouth: Promise and challenges of online reputation mechanisms. In *Management Science*, 2003.
- [7] A. Dragoni, P. Giorgini, and M. Baffetti. Distributed belief revision vs. belief revision in a multi-agent environment: First results of a simulation experiment. In *MAAMAW*, pages 45–62, 1997.
- [8] A. Dragoni, P. Giorgini, and P. Puliti. Distributed belief revision versus distributed truth maintenance, 1994.
- [9] eBay. ebay. In <http://www.eBay.com>, 2002.
- [10] M. A. Falappa, G. Kern-Isberner, and G. R. Simari. Explanations, belief revision and defeasible reasoning. *Artificial Intelligence*, 141(1):1–28, 2002.

- [11] S. O. Hansson. Kernel contraction. *Journal of Symbolic Logic*, 59(3):845–859, 1994.
- [12] S. O. Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer Academic Publishers, march 1999.
- [13] Sven Ove Hansson, Eduardo L. Fermé, John Cantwell, and Marcelo A. Falappa. Credibility limited revision. *J. Symb. Log.*, 66(4):1581–1596, 2001.
- [14] N. E. Kfir-Dahav and M. Tennenholz. Multi-agent belief revision. In Yoav Shoham, editor, *Theoretical Aspects of Rationality and Knowledge: Pro. of the Sixth Conf. (TARK 1996)*, pages 175–196. Morgan Kaufmann, San Francisco, 1996.
- [15] W. Liu and M. A. Williams. A framework for multi-agent belief revision, part i: The role of ontology. In *Australian Joint Conference on A. I.*, pages 168–179, 1999.
- [16] W. Liu and M. A. Williams. A framework for multi-agent belief revision. *Studia Logica*, 67(2):291–312, 2001.
- [17] David Makinson. Screened revision. *Theoria: Special Issue on Non-Prioritized Belief Revision*, 1997.
- [18] A. Mui, L. Halberstadt and M Mohtasheми. Notions of reputation in multi-agent systems: A review. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02)*, pages 280–287, Bologna, Italy, 2002.
- [19] Jordi Sabater and Carles Sierra. Regret: A reputation model for gregarious societies. *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, pages 61–69, 2001.
- [20] Jordi Sabater and Carles Sierra. Review on computational trust and reputation models. *Artificial Intelligence Review*, 24(1):33–60, 2005.
- [21] L. H. Tamargo, A. J. García, M. A. Falappa, and G. R. Simari. Consistency maintenance of plausible belief bases based on agents credibility. *12th International Workshop on Non-Monotonic Reasoning (NMR)*, pages 50–58, 2008.
- [22] Luciano H. Tamargo, Alejandro J. García, Marcelo A. Falappa, and Guillermo R. Simari. A belief revision approach to inconsistency handling in multi-agent systems. *The IJCAI-09 Workshop on Non-monotonic Reasoning, Action and Change (NRAC)*, pages 63–70, 2009.
- [23] Luciano H. Tamargo, Alejandro J. García, Marcelo A. Falappa, and Guillermo R. Simari. Modeling knowledge dynamics in multi-agent systems based on informants (en prensa). *Proceedings of The Knowledge Engineering Review (KER)*, 2010.