

Agentes creíbles basados en argumentación en sistemas multiagente

Santiago E. Fulladoza Dalibón, Diego C. Martinez, Guillermo R. Simari

Laboratorio de Investigación y Desarrollo en Inteligencia Artificial
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur,
Av. Alem 1253, Bahía Blanca, Argentina
Tel: (0291) 459-5135
e-mail: santiagofulladoza@gmail.com, dcm@cs.uns.edu.ar, grs@cs.uns.edu.ar

Resumen

Esta investigación está centrada en la integración de formalismos de argumentación rebatible en agentes inteligentes creíbles. Los agentes creíbles son agentes autónomos ricos en aspectos de personalidad propia, en los que las emociones juegan un rol preponderante para definir su comportamiento.

Actualmente cobran vital importancia en áreas de crecimiento vertiginoso donde la interactividad es protagonista, como los juegos interactivos digitales de entretenimiento o educativos, la búsqueda personalizada de información, los sistemas de recomendaciones o la simulación computarizada. En todos estos escenarios es importante la percepción que se tiene del agente según su comportamiento.

Hasta el momento, no ha sido explorado completamente el uso de formalismos de argumentación en el modelado de aspectos emocionales en agentes inteligentes, lo que constituye el eje central del presente plan de trabajo.

Creemos que el enriquecimiento de los sistemas argumentativos con características tales como factores de personalidad estática y cambiante, características emocionales, habilidades sociales y consistencia de expresión rep-

resentará un avance significativo dentro del área de sistemas argumentativos en Inteligencia Artificial y Ciencias de la Computación.

Palabras clave: Agentes creíbles, Argumentación, Emociones

1. Contexto

Esta línea de investigación se llevará a cabo dentro del ámbito de colaboración entre el Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA) del Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur y el Laboratorio de Investigación y Desarrollo en Inteligencia Artificial Concordia (LIDIA Concordia) de la Facultad de Ciencias de la Administración, Universidad Nacional de Entre Ríos.

2. Introducción

Actualmente es predominante la utilización global de sistemas computacionales con un alto grado de interactividad. La innovación en tecnologías interactivas es una demanda actual en las tecnologías de información y co-

municaciones. Esta interactividad existe en los sistemas educativos, de entretenimiento y en las diversas formalizaciones de realidad virtual, ya sea en simulaciones computacionales o en las actuales redes sociales. Es precisamente en este aspecto donde los agentes inteligentes necesitan claras habilidades sociales y emocionales. En estos escenarios surge la noción de agente creíble, una particularidad de agentes fuertes en los cuales el énfasis es puesto en la percepción que del agente tienen quienes interactúan con él [Loy97, RCS⁺96]. El uso masivo actual de entornos virtuales de inmersión, tales como los juegos de entretenimiento o las simulaciones de navegación (aérea, terrestre o marítima) hacen de los agentes creíbles un elemento de especial interés técnico y científico. Para que la experiencia de inmersión sea completa, es necesario que los agentes virtuales que la pueblan sean consistentes con la realidad emulada.

Si bien tradicionalmente se ha considerado a la emoción como algo irracional y que resta valor al razonamiento humano, es en la actualidad universalmente aceptado que las emociones son parte esencial de la inteligencia humana, y desempeñan un papel crucial en cuanto a la percepción, el aprendizaje, la atención, la memoria, la toma de decisiones racional y otras habilidades sociales [Adolphs94, Ledoux98]. El estudio de la emoción es esencial para comprender y representar la cognición humana. Neurológicamente hablando, no existe una división clara entre razonamiento y emociones [Pic95].

Por el otro lado, la argumentación rebatible es un mecanismo de razonamiento no monótono en donde la aceptación o el rechazo de una proposición dependen de un análisis entre argumentos a favor y en contra de esa proposición. Usualmente es utilizada bajo diferentes formalizaciones para capturar aspectos del razonamiento del sentido común y la representación de información incompleta y potencialmente inconsistente.

Hay dos aspectos importantes de los agentes creíbles que alientan la vinculación con la argumentación rebatible. Primero, los agentes creíbles se desenvuelven en escenarios con un alto grado de interactividad, ya sea con humanos o con otros agentes inteligentes. Este entorno de agentes con alta actividad social presenta características naturalmente argumentativas, lo que motiva el estudio y desarrollo de formalismos argumentativos propicios. Segundo, la propiedad intrínseca de los agentes creíbles de poseer y desarrollar personalidad distintiva requiere procesos de introspección, sopesando creencias, deseos e intenciones propias. Esta actividad racional del agente también puede interpretarse como de naturaleza argumentativa.

3. Líneas de investigación

Ésta línea de investigación se centrará en el problema de desarrollar una arquitectura de agente basada en sistemas de argumentación que considere, además de elementos racionales, aspectos emocionales para definir su comportamiento; de modo que proporcionen un mayor grado de credibilidad a los agentes que la implementen.

3.1. Argumentación

La argumentación es una forma de razonamiento en la cual, para una afirmación dada, se presta atención explícita a las justificaciones presentadas y a la resolución de los posibles conflictos entre ellas; de modo que la afirmación es aceptada o no según el análisis de los argumentos a su favor y en su contra. La forma en que los argumentos y las justificaciones son considerados permite definir un tipo de razonamiento automático, en el cual puede haber información contradictoria, incompleta, e incierta. El estudio de la argumentación ha sido abordado desde diferentes enfoques en años re-

cientes. A nivel lógico puede considerarse como una forma de modelar inferencia rebatible, y a nivel dialógico como una forma de interacción entre agentes. Actualmente es un atractivo paradigma para conceptualizar el razonamiento de sentido común [CML00, PV98]. Esto produjo como resultado la formalización de diferentes frameworks de argumentación abstracta como [Dun95, BGG05, NBCD08], entre otros; y de Sistemas de Argumentación Basados en Reglas (SABR) como [AK07, DKT06, GS04, PS97]; lo que permitió el desarrollo de diversas aplicaciones del mundo real basadas en argumentación. En el último tiempo, el campo de aplicación de la argumentación se ha expandido velozmente, en gran parte debido a los avances teóricos, pero también gracias a la demostración exitosa de su uso práctico en un gran número de dominios de aplicación, tales como el razonamiento legal [PS02], la ingeniería del conocimiento [CRL00], los sistemas multi-agentes [PSJ98, AMP02], y el e-government [ABcM05], entre muchos otros.

3.2. Agentes inteligentes creíbles

Un agente inteligente es una entidad autónoma que es capaz de percibir su entorno y de actuar en él acordeamente, posiblemente alterándolo en diversos grados. El estudio y desarrollo de agentes inteligentes ha sido abordado desde diferentes perspectivas; pero esta línea de investigación particularmente está relacionada con (y parte de) el concepto fuerte de agencia (en inglés, *strong agency*), según el cual se atribuye a los agentes, además de las características propias de agente, propiedades antropomórficas que podrían ir desde actitudes mentales, como creencias, deseos e intenciones, hasta capacidades y actitudes emocionales y caracterizaciones visuales de los agentes. Así llegamos al concepto de agente creíble, cuya diferencia fundamental con el espectro tradicional de agentes inteligentes es que el énfasis es puesto en la creación de agentes que poseen personalidades propias y

distintivas [Loy97].

4. Resultados y Objetivos

Esta línea de investigación tiene por objetivo desarrollar una arquitectura de agente que integre los sistemas argumentativos rebatibles con características emocionales y de personalidad, a modo de producir un aumento tanto en la credibilidad de los agentes inteligentes artificiales como en la eficiencia con la que operan.

Por otro lado, el enriquecimiento de los sistemas argumentativos con características tales como factores de personalidad estática y cambiante, características emocionales, habilidades sociales y consistencia de expresión representará un avance significativo dentro del área de sistemas argumentativos en Inteligencia Artificial y Ciencias de la Computación.

5. Formación de Recursos Humanos

Esta investigación se lleva a cabo en el contexto de una beca de postgrado otorgada por el CONICET; por lo que tiene previsto la formación de un Doctor en Ciencias de la Computación.

Referencias

- [ABcM05] Katie Atkinson, Trevor Benchcapon, and Peter Mcburney. Multi-agent argumentation for edemocracy. In *In Proceedings of the Third European Workshop on Multi-Agent Systems*, 2005.
- [AK07] Leila Amgoud and Souhila Kaci. An argumentation framework for merging conflicting knowledge bases. *Int. J. Approx. Reasoning*, 45:321–340, July 2007.
- [AMP02] L. Amgoud, N. Maudet, and S. Parsons. An argumentation-based semantics for agent communication languages. In F. Van Harmelen, editor, *Proceedings of the European Conference on Artificial Intelligence (ECAI-2002)*, pages 38–42, Lyon, France, July 2002. IOS Press.
- [BGG05] Pietro Baroni, Massimiliano Giacomin, and Giovanni Guida. Sc-recursiveness: a general schema for argumentation semantics. *Artif. Intell.*, 168:162–210, October 2005.
- [CML00] Carlos Ivan Chesnevar, Ana Gabriela Maguitman, and Ronald Prescott Loui. Logical models of argument. *ACM COMPUTING SURVEYS*, 32:337–383, 2000.
- [CRL00] Daniela V. Carbogim, David Robertson, and John Lee. Argument-based applications to knowledge engineering. *Knowl. Eng. Rev.*, 15:119–149, June 2000.
- [DKT06] P. M. Dung, R. A. Kowalski, and F. Toni. Dialectic proof procedures for assumption-based, admissible argumentation. *Artif. Intell.*, 170:114–159, February 2006.
- [Dun95] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.
- [GS04] Alejandro J. García and Guillermo R. Simari. Defeasible logic programming: an argumentative approach. *Theory Pract. Log. Program.*, 4:95–138, January 2004.
- [Loy97] Aaron B. Loyall. *Believable agents: building interactive personalities*. PhD thesis, 1997.
- [NBCD08] Fahd Saud Nawwab, Trevor Bench-Capon, and Paul E. Dunne. A methodology for action-selection using value-based argumentation. In *Proceeding of the 2008 conference on Computational Models of Argument: Proceedings of COMMA 2008*, pages 264–275, Amsterdam, The Netherlands, The Netherlands, 2008. IOS Press.
- [Pic95] R. Picard. Affective computing. Technical Report 321, MIT Media Laboratory, Perceptual Computing Section, nov 1995.
- [PS97] Henry Prakken and Giovanni Sartor. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics*, pages –1–1, 1997.
- [PS02] Henry Prakken and Giovanni Sartor. The role of logic in computational models of legal argument: A critical survey. In *Computational Logic: Logic Programming and Beyond, Essays in Honour of Robert A. Kowalski, Part II*, pages 342–381, London, UK, 2002. Springer-Verlag.

- [PSJ98] Simon Parsons, Carles Sierra, and Nicholas Jennings. Agents that reason and negotiate by arguing. *JOURNAL OF LOGIC AND COMPUTATION*, 8:261–292, 1998.
- [PV98] Henry Prakken and G. Vreeswijk. Logical systems for defeasible argumentation. In *Handbook of Philosophical Logic*. D. Gabbay, second edition, 1998.
- [RCS⁺96] W. Scott Reilly, Jaime Carbonell, Reid Simmons, W. Scott, W. Scott, Neal Reilly, and Neal Reilly. Believable social and emotional agents, 1996.