

Modelos Formales de Diálogo para Sistemas Multi-Agente

Julieta Marcos **Marcelo A. Falappa** **Guillermo R. Simari**
mjm@cs.uns.edu.ar mfalappa@cs.uns.edu.ar grs@cs.uns.edu.ar

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA)
Departamento de Ciencias e Ingeniería de la Computación (DCIC)
Universidad Nacional del Sur (UNS)

Resumen

En un sistema multi-agente los agentes necesitan comunicarse, por diferentes motivos, dando lugar a diferentes tipos de *diálogos*. Existe interés por automatizar estas interacciones, sin embargo la mayoría de las propuestas existentes en la literatura son *ad hoc* y carecen de un fundamento teórico sólido. Nuestro trabajo de investigación está orientado principalmente al estudio abstracto de estas interacciones con el fin de brindar metodologías que faciliten el diseño de sistemas de diálogo.

Contexto

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

- “Programación en lógica distribuida con primitivas de sincronización, comunicación y ejecución en paralelo de agentes, para desarrollo de sistemas multi-agente en ambientes dinámicos” Código: PIP 5050. Director: Alejandro Javier García. Financiamiento: CONICET (Proyecto de Investigación Plurianual).
- “Formalismos y Tecnologías de Dinámica de Conocimiento aplicadas a Robótica Cognitiva, Sistemas Multi-agente y Web Semántica”. Código: PGI 24/ZN11. Director: Dr. Marcelo A. Falappa. Co-director: Dr. Alejandro J. García. Financiamiento: Universidad Nacional del sur.

1. Introducción

En un sistema multi-agente los agentes necesitan comunicarse, por diferentes motivos: resolver diferencias de opinión o intereses en conflicto, cooperar para resolver dilemas o encontrar pruebas, o simplemente informarse uno a otro sobre hechos pertinentes. En muchos casos no alcanza con intercambiar mensajes aislados, sino que los agentes necesitan entablar *diálogos* (secuencias de mensajes sobre el mismo tema) [10]. Existe una gran variedad de interacciones con características diferentes que podrían quererse modelar. Una posible tipología, teniendo

en cuenta el objetivo común del diálogo y las metas particulares de cada participante, es la siguiente [13]:

Diálogo de Búsqueda de Información. Un agente busca la respuesta a una pregunta en el conocimiento de otro agente. Se supone que este último conoce la respuesta.

Diálogo de Investigación. Todos los agentes colaboran para encontrar la respuesta a una pregunta. Se supone que ninguno de ellos conoce la respuesta.

Diálogo Persuasivo. Un agente trata de convencer a otro para que se adhiera a cierta creencia o punto de vista.

Negociación. Los agentes tratan de llegar a un acuerdo aceptable sobre la división de recursos escasos. Cada uno trata de maximizar su ganancia.

Diálogo Deliberativo. Los agentes colaboran para decidir que acción realizar en cierta situación.

Se puede hacer una distinción entre diálogos *colaborativos* y diálogos *no-colaborativos*. En un diálogo colaborativo los agentes no tienen metas individuales más allá de la meta común del diálogo; por lo tanto, todos colaboran en aras del mismo fin. En un diálogo no-colaborativo, en cambio, los agentes tienen metas individuales que podrían estar en conflicto. Por ejemplo, un diálogo de investigación es colaborativo (el único objetivo compartido por todos los agentes investigadores es descubrir la verdad) pero una negociación no lo es (cada agente negociador tiene como meta maximizar su propia ganancia). Un diálogo deliberativo podría ser colaborativo o no (dependiendo de si los agentes tienen algún interés particular por tomar cierto curso de acción). Un diálogo persuasivo puede ser visto como un diálogo *semi-colaborativo*, donde el agente que persuade tiene una meta particular, pero el persuadido podría no tenerla. Los *agentes colaborativos* expondrán toda la información que consideren relevante, mientras que los *agentes no-colaborativos* podrían ocultar información, sabiendo que es relevante, porque no favorece el cumplimiento de sus metas individuales.

Otra posible diferenciación es entre aquellos diálogos que son *sobre creencias*, y aquellos que no lo son. En un diálogo sobre creencias los participantes hablan sobre la verdad de cierta proposición. A esta categoría corresponden los primeros tres tipos de diálogo: búsqueda de información, investigación y diálogo persuasivo. Sin embargo, un diálogo persuasivo no es *necesariamente* sobre creencias (podría ser, por ejemplo, sobre acciones).

El objetivo final de este área de investigación es la automatización de los diferentes tipos de diálogo en un sistema multi-agente. A grandes rasgos, podemos identificar algunos elementos que están presentes en un sistema de diálogo:

1. Un **conjunto de agentes con estado mental**. Los elementos involucrados en el estado mental de los agentes dependerán de los distintos tipos de diálogo.
2. Un **tópico** en torno al cual se desarrolla el diálogo.
3. Una **secuencia de movimientos** realizados por los agentes, que conforman al diálogo propiamente dicho. En qué consisten exactamente los movimientos, y si éstos pueden causar modificaciones en el estado mental de los agentes o no, dependerá de los distintos tipos de diálogo.
4. Una **conclusión** sobre el tópico en cuestión, abordada al finalizar el diálogo.

Por ejemplo, el estado mental de un agente podría incluir simplemente creencias (en el caso de diálogos sobre creencias), o podría incluir también deseos, intenciones, preferencias, etc., en el caso de diálogos deliberativos o negociaciones. La forma de los movimientos también es variante según los distintos tipos de diálogo. Por ejemplo, en los diálogos de investigación los movimientos

en general consisten en hacer afirmaciones o aceptar afirmaciones hechas por otros, mientras que en una negociación tienen sentido también otros tipos de movimientos como por ejemplo: realizar/aceptar/rechazar propuestas y contra-propuestas, hacer promesas, etc.. En general, los sistemas de diálogo definen un conjunto de *locuciones* (por ejemplo: *afirmar*, *proponer*, *rechazar*, *aceptar*) y luego los movimientos especifican el tipo de locución junto con cierto contenido (el contenido corresponde a lo que se afirma, o lo que se promete, etc.). También suele especificarse un *protocolo de interacción*, que es un conjunto de reglas que gobiernan el diálogo, indicando principalmente qué tipos de movimientos son válidos en determinado momento (por ejemplo: después de que un agente hace una propuesta, otro agente puede aceptarla, rechazarla, o hacer una contra-propuesta). Si bien el protocolo de interacción establece en líneas generales la forma de los diálogos válidos, esto no es suficiente para lograr su automatización. Para tal fin, hace falta una *estrategia* que permita a cada agente determinar de manera autónoma el movimiento (locución + contenido) a realizar en cada momento.

Se han realizado varios trabajos con el objetivo de estudiar y modelar formalmente estas interacciones. Existe un grupo creciente de investigadores que consideran que se puede mejorar la calidad de los diálogos si los agentes exponen los *argumentos* que justifican lo que dicen [12], es decir, entablan *diálogos basados en argumentación*. Este tipo particular de diálogo respeta la formal general mencionada más arriba, pero se asumen en particular *agentes argumentativos* (es decir, agentes capaces de generar y evaluar argumentos) de manera que los movimientos que éstos realizan incluyen (o pueden incluir), como parte del contenido, argumentos y contra-argumentos. Habitualmente, los sistemas de diálogo basados en argumentación se diseñan sobre la base de un formalismo argumentativo ¹ existente. El trabajo realizado en [12] brinda un amplio panorama en lo que respecta a la *negociación basada en argumentación*, identificando y describiendo los principales elementos (tanto internos como externos a los agentes) necesarios para su modelamiento, y señalando también las principales motivaciones y ambiciones en la investigación.

La mayoría de las propuestas existentes en la literatura definen formalmente un sistema para un tipo particular de diálogo, e investigan propiedades de los diálogos generados por el sistema. Sin embargo, las soluciones propuestas en general son *ad hoc* y carecen de un fundamento teórico sólido. Algunos trabajos sólo especifican la representación del estado mental de los agentes, el conjunto de locuciones, y el protocolo de interacción, dejando sin especificar la estrategia de selección de movimientos. Otros trabajos incluyen también una estrategia, pero en general ésta se diseña de manera *ad hoc* y muchas veces no alcanza todas las propiedades deseables. A continuación mencionaremos brevemente algunas de las propuestas existentes. Todos los trabajos que mencionaremos estudian diálogos basados en argumentación. En [4] proponen un framework para diálogos basados en argumentación, con el objetivo de llegar a un acuerdo en la toma una decisión colectiva entre un conjunto de agentes autónomos, y estudian propiedades como *terminación* y *optimalidad del resultado*. En [2], [9] y [3] los autores se concentran en diálogos sobre creencias (búsqueda de información, investigación, y persuasión). Definen un conjunto de *locuciones* para que los agentes puedan intercambiar argumentos, un conjunto de *actitudes* que marcan una relación entre los argumentos que puede construir un agente y las locuciones que puede realizar (intuitivamente, los agentes “menos atrevidos” sólo afirman proposiciones soportadas por “buenos argumentos”), y definen también un conjunto de *protocolos* para llevar a cabo los diálogos. Luego consideran algunas propiedades de los diálogos generados bajo estos protocolos: en principio *terminación* y *complejidad*, y luego extienden el trabajo analizando propiedades del resultado e investigando en qué medida el resultado es dependiente de las tácticas utilizadas por los agentes.

¹Mayor información en el área de sistemas argumentativos puede encontrarse, por ejemplo, en [5] y [11].

2. Líneas de Investigación y Desarrollo

Dado que, desde nuestro punto de vista, hace falta una base teórica sólida que facilite el diseño y desarrollo de estos sistemas de diálogo, nuestro trabajo de investigación actual tiene por objetivo estudiar de manera abstracta y formal los requerimientos esenciales que deberían alcanzar estos sistemas, así como también identificar los principales desafíos en la persecución de los mismos, y esbozar estrategias generales para su cumplimiento. Actualmente estamos considerando una noción de diálogo restringida en ciertos aspectos, en particular: (1) el conjunto de posibles locuciones está reducido a *afirmar* conjuntos de proposiciones, (2) el lenguaje de representación de conocimiento y el mecanismo para obtener conclusiones son compartidos por todos los agentes participantes, y (3) el estado mental de los agentes no se modifica durante el diálogo. Se pretende brindar un enfoque abstracto, con el objetivo de abarcar un rango amplio de diálogos, ya sea basados en argumentación o no, sin especificar la representación del estado mental de los agentes ni el mecanismo para obtener conclusiones. De todas maneras, se pretende llevar, luego, los resultados generales a casos particulares para analizarlos en profundidad. En este sentido, estamos instanciando los conceptos y resultados obtenidos con formalismos lógicos específicos, en particular *Lógica Clásica Proposicional*, *Programación en Lógica* y *Programación en Lógica Rebatible* (DeLP) [6].

3. Resultados Obtenidos/Esperados

En [7] propusimos un algoritmo abstracto que simula una interacción entre dos o más participantes. Básicamente, vemos al diálogo como un proceso mediante el cual los agentes provocan sucesivos cambios sobre una base de conocimiento pública (que representa el *estado* del diálogo). Los agentes tienen *metas* que dictan qué conocimiento exponer en determinado momento. El uso de *Operadores de Cambio*² nos permitió abstraernos de la teoría lógica subyacente (lenguaje y mecanismo de inferencia). Por otro lado, el uso de una noción abstracta de *meta* nos permitió parametrizar las actitudes de los agentes en el diálogo, abstrayéndonos, en cierta medida, del *tipo* de diálogo. Mostramos cómo pueden diseñarse las metas de los agentes para modelar diferentes tipos de diálogos sobre creencias (*investigación*, *búsqueda de información* y *persuasión*) y brindamos ejemplos en Lógica Proposicional. En un trabajo posterior, también exploramos su aplicabilidad en diálogos de *negociación*.

En [8] nos concentramos en diálogos colaborativos específicamente, proponiendo un conjunto básico de *propiedades deseables* que éstos deberían satisfacer. Las propiedades propuestas persiguen la caracterización de diálogos finitos, con conclusiones razonables (considerando todo lo que ha sido dicho durante el diálogo), en los cuales toda la información relevante es expuesta y toda la información expuesta es realmente relevante. Para tal fin, definimos dos niveles de *relevancia* en el diálogo, *relevancia directa* y *relevancia potencial*, basados en las nociones de *inferencia* y *abducción* respectivamente. Mostramos ejemplos concretos utilizando Lógica Proposicional y Programación en Lógica Rebatible.

Como trabajo futuro, pretendemos generalizar los resultados a otros tipos de diálogos y relajar algunas de las restricciones impuestas en trabajos anteriores, en particular ampliar el conjunto de locuciones, y considerar conjuntos de agentes no homogéneos en cuanto al lenguaje de representación de conocimiento y al mecanismo de razonamiento.

²La *Teoría de Cambio de Creencias* estudia la *dinámica del conocimiento* en agentes o mundos, es decir, los cambios provocados en una base de conocimiento por el arribo de nueva información.

Referencias

- [1] Leila Amgoud, Yannis Dimopoulos, and Pavlos Moraitis. A unified and general framework for argumentation based negotiation. *In 6th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), Honolulu, Hawaii*, May 2007.
- [2] Leila Amgoud, Nicolas Maudet, and Simon Parsons. Modelling dialogues using argumentation. *In 4th International Conference on MultiAgent Systems, ICMAS'2000, Boston, USA. IEEE Press*, 2000. July.
- [3] Leila Amgoud, Simon Parsons, and Michael Wooldridge. On the formal outcomes of formal inter-agent dialogues. *In 2nd International joint conference on Autonomous Agents and Multi-Agent systems, AAMAS'03, Melbourne, Australia. ACM Press*, pages 616–623, July 2003.
- [4] Leila Amgoud, Henri Prade, and Sihem Belabbes. Towards a formal framework for the search of a consensus between autonomous agents. *In the proceedings of the 4th International joint Conference on Autonomous Agents and Multi-Agent Systems, AAMAS'2005, Utrecht. Franck Dignum, Michael Wooldridge, Sven Koenig, Sarit Kraus (Eds.), ACM Press*, pages 537–543, July 2005.
- [5] Carlos Iván Chesñevar, Ana Gabriela Maguitman, and Ronald Prescott Loui. Logical models of argument. *ACM Comput. Surv.*, 32(4):337–383, 2000.
- [6] Alejandro J. García and Guillermo R. Simari. Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming*, 4(1):95–138, 2004.
- [7] M. Julieta Marcos, Marcelo A. Falappa, and Guillermo R. Simari. Un modelo abstracto de diálogo sobre creencias para sistemas multi-agente. *Actas del XIII Congreso Argentino de Ciencias de la Computación (CACIC)*, pages 1371–1382, 2007.
- [8] M. Julieta Marcos, Marcelo A. Falappa, and Guillermo R. Simari. Properties for a formal model of collaborative dialogue. *Actas del XIV Congreso Argentino de Ciencias de la Computación (CACIC)*, 2008.
- [9] Simon Parsons, Leila Amgoud, and Michael Wooldridge. An analysis of formal inter-agent dialogues. *In 1st International Joint Conference on Autonomous Agents and Multi-Agent systems, AAMAS'2002, Bologna, Italy. ACM Press*, pages 394–401, July 2002.
- [10] Simon Parsons, Michael Wooldridge, and Leila Amgoud. Properties and complexity of some formal inter-agent dialogues. *Journal of Logic and Computation*, 13:347–376, 2003.
- [11] Henry Prakken and Gerard Vreeswijk. Logics for defeasible argumentation. In Dov M. Gabbay, editor, *Handbook of Philosophical Logic*. Kluwer Academic, 1999.
- [12] Iyad Rahwan, Sarvapali D. Ramchurn, Nicholas R. Jennings, Peter McBurney, Simon Parsons, and Liz Sonenberg. Argumentation-based negotiation. *The Knowledge Engineering Review*, 18:343–375, 2003.
- [13] Douglas Walton and Erik C. W. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press, Albany, NY, 1995.