

UN PROTOCOLO PARA CO-EDICIÓN ARGUMENTATIVA ENTRE AGENTES

Autores: Miguel Rebollo, Carlos Carrascosa, Vicente Julián
Dept. Sistemas Informáticos y Computación
Universidad Politécnica de Valencia
e-mail: {mrebollo,carrasco,vinglada}@dsic.upv.es

Resumen: En este trabajo se propone un conjunto de protocolos para la colaboración entre agentes inteligentes que permite el intercambio de ideas y la discusión. Utiliza en el estándar propuesto por FIPA para la comunicación entre agentes. Los protocolos se basan en un esquema para co-edición argumentativa, adaptado a las necesidades de un sistema multi-agente. Este esquema supone un primer paso hacia la construcción social del conocimiento en agentes autónomos.

Palabras Clave: agentes; colaboración; conocimiento distribuido; comunicación

Abstract: In this paper a set of protocols that allows collaboration between intelligent agents is proposed. This schema allows to interchange ideas and agent discussion. It uses the FIPA standard for agent communication. Protocols are based on an schema of argumentative co-edition, adapted to satisfy multi-agent system requirements. This schema supposes a first stage towards social construction of knowledge for autonomous agents.

Introducción

Los recientes avances en las Tecnologías de la Información y de las Comunicaciones han ampliado el tipo y el ámbito de uso de las actividades que involucran a personas y ordenadores. Una de las últimas tendencias es la creación de comunidades “virtuales”, entendiendo como tales un conjunto de usuarios geográficamente dispersos que comparten información, recursos y/o conocimiento a través de un medio electrónico (usualmente Internet). Estas comunidades se crean para usos profesionales y sociales, si bien las necesidades de ambas son muy semejantes desde el punto de vista del intercambio de información.

De esta forma, surge un nuevo fenómeno: la *distribución del conocimiento*. Ya no es necesario que toda la información se encuentre agrupada en un mismo lugar, sino que se puede trabajar con ella de forma remota. El problema que aparece es el de cómo compartir el conocimiento dentro de una comunidad de usuarios, para lo que se precisa:

- gestión de la **inteligencia** individual y colectiva, así como del conocimiento emergente de la propia comunidad.
- métodos de **acceso** a los recursos de información disponibles.
- **interacción** social entre los miembros de la comunidad.

El paradigma de sistemas multi-agente resulta un modelo adecuado para dar soporte a este tipo de comunidades, principalmente por dos motivos: (i) permite la construcción de sistemas inteligentes distribuidos situados en un entorno y (ii) facilita la cooperación y la colaboración a través de protocolos de comunicaciones de alto nivel, usualmente basados en la teoría de los actos comunicativos (“*speech acts*”).

En este trabajo se propone un protocolo para la colaboración entre agentes como una capa por encima de los protocolos de interacción de FIPA. Sigue un modelo basado en conversaciones, en las cuales los agentes realizan una serie de contribuciones. La estructura de la conversación define el tipo de contribuciones que se puede realizar en cada momento. Este protocolo permite el intercambio de ideas y la discusión entre varios agentes y supone un primer paso hacia la construcción social del conocimiento en agentes autónomos.

El resto del artículo se estructura de la siguiente forma: la sección 2 presenta los conceptos de agente y sistema multi-agente. En la sección 3 se muestran las necesidades detectadas para permitir el intercambio de ideas entre agentes inteligentes, exponiendo las limitaciones de algunas de las técnicas empleadas en la actualidad. En la sección 4 se exponen los protocolos de interacción que se proponen para permitir el intercambio de ideas y la discusión entre agentes autónomos. Por último, la sección 5 resume las conclusiones del presente trabajo y se indican las líneas futuras de investigación y desarrollo en las que se está trabajando.

1 Sistemas multi-agente

Los agentes inteligentes son considerados como una nueva técnica que permitirá abordar de una manera más apropiada la construcción de sistemas inteligentes complejos aplicados a muy diversos campos. La definición del concepto de agente es sumamente compleja y se puede encontrar propuestas en la literatura un gran número de definiciones del concepto de agente. La definición más aceptada de agente es posiblemente la propuesta por Wooldridge y Jennings que lo definen como “*un sistema computacional, situado en algún entorno, que es capaz de actuar de forma autónoma y flexible con tal de lograr sus objetivos de diseño*” [Wooldridge95], entendiéndose por flexible que un agente sea:

- **Reactivo:** responde al entorno en que se encuentra.
- **Proactivo:** intenta cumplir sus propios planes u objetivos.
- **Social:** se comunica con otros agentes mediante algún tipo de lenguaje.

Junto con esta definición básica, existen toda una serie de atributos opcionales que puede o no tener un agente y que están en función de las características funcionales del mismo [Franklin1996]. En la mayoría de las ocasiones, el objetivo no es disponer de un único agente capaz de resolver todo el problema, sino tener un conjunto de agentes (a este sistema se le denomina *sistema multi-agente* [Huhns, 1998]) que cooperan en la resolución del mismo, y por tanto, interactúan, se comunican y razonan sobre su propio conocimiento y sobre el conocimiento que poseen los demás.

La existencia de múltiples definiciones de agente provoca no pocas divergencias entre los investigadores. La aparición, en este caso, de un estándar que sirva de referencia parece más que necesaria y es considerado un prerequisite para la explotación con éxito (e incluso la comercialización) de la tecnología de agentes inteligentes. En esta línea, FIPA —*Foundation for Intelligent Physical Agents*— es una organización internacional dedicada a promover la aplicación comercial de la tecnología basada en agentes mediante en desarrollo abierto de aplicaciones, servicios y equipamiento basados en agentes. FIPA proporciona una serie de propuestas para la especificación de la tecnología de agentes básica que pueden utilizar los desarrolladores de sistemas basados en agentes para la construcción de sistemas complejos con un alto grado de interoperatividad.

2 Construcción social del conocimiento

Una de las características de los sistemas multi-agente es la autonomía que disfruta cada una de las entidades que forma el sistema. Cada agente es responsable de una parcela del entorno en el que se encuentra, centrándose únicamente en aquellos aspectos del problema que son relevantes para él. De esta manera se consigue simplificar la construcción de un sistema software que aborde problemas complejos.

Sin embargo, el coste que se debe pagar por esta simplificación es que, en ocasiones, un agente puede necesitar conocimiento de una parcela del problema que no le pertenece directamente, por lo que se ve obligado a solicitar la colaboración de otros agentes para resolver el problema. Esta colaboración tiene lugar básicamente de dos maneras:

- intercambiando información para que el primer agente disponga de todos los datos para resolver el problema.
- delegando tareas a otros agentes, de forma que sea otra entidad la que resuelva esa parte del problema y devuelva únicamente el resultado.

En ambos casos, es necesario un mecanismo de comunicación entre los agentes. En general, podemos ver éste intercambio de mensajes como una forma de influir en el comportamiento de los demás agentes: el agente iniciador intentará que otro agente trabaje para él (consiguiéndole información o delegando en él la responsabilidad de la realización de una subtarea), mientras

que el receptor pretende influenciar al agente iniciador para que éste varíe sus creencias. En resumen, los agentes utilizan la comunicación para alterar el estado de sus compañeros.

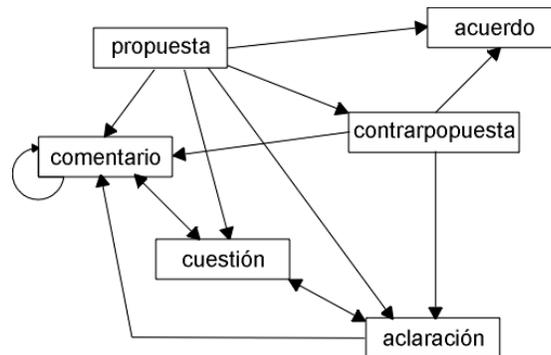


FIGURA 1. Grafo de conversación para co-edición argumentativa

Los esquemas de comunicación entre agentes disponibles hasta la fecha no contemplan la posibilidad de que un agente no esté de acuerdo con la información facilitada por otros agentes. En este artículo se presenta una adaptación de un modelo de comunicación basado en conversaciones [Bobrow 1990]: el grafo de co-edición argumentativa. Este esquema se ha aplicado con éxito en entornos de aprendizaje colaborativo a distancia para la construcción de documentos entre equipos de personas [Barros 1998]. Esa aproximación no pretende ser un modelo de comunicación entre entidades autónomas, sino una ayuda para estructurar las conversaciones y el intercambio de conocimiento entre humanos.

La principal limitación de esta aproximación para tratar de incluirla en un proceso de razonamiento automático es que no se puede asegurar que el proceso termine alguna vez, puesto que la única forma de acabar la discusión es que todos los participantes estén de acuerdo en algún punto de la misma. Para que este esquema sea válido en un sistema multi-agente siguiendo las recomendaciones de FIPA para sus protocolos de comunicaciones es necesario garantizar la terminación del protocolo, forzando a que todos los agentes involucrados en la discusión muestren su acuerdo o su desacuerdo con cada una de las propuestas que se realizan. Para garantizar la terminación del protocolo en cualquier caso debemos obligar a:

- a) un plazo máximo (*deadline*) para que los participantes evalúen las propuestas y decidan si la aceptan o no.
- b) tener la posibilidad de que el iniciador cancele una propuesta (porque no se llegue a un acuerdo o la situación que dio origen a la discusión haya cambiado).
- c) que siempre se emita un juicio sobre cada propuesta, aceptándola o rechazándola.

3 Protocolos de interacción

Lo que se propone en este artículo es una adaptación del grafo de conversación para co-edición argumentativa (ver FIGURA 1) para que pueda ser utilizado por agentes inteligentes autónomos. El primer paso es aplicarlo dentro de asistentes personales, de forma que sea un apoyo para los usuarios en entornos colaborativos. Para simplificar, asumiremos que los grupos están ya formados y que se trata de grupos cerrados en los que todos los agentes se conocen entre sí.

Un *protocolo* define una secuencia fija de mensajes que establecen un flujo de comunicación bien definido entre agentes. Los agentes pueden verse involucrados en varios diálogos simultáneamente. El término *conversación* se emplea para denotar una instancia particular de un diálogo. Así, el agente puede estar involucrado en múltiples conversaciones, con diferentes agentes, utilizando distintos protocolos.

Dentro de una conversación, una vez se ha establecido el protocolo, un agente está esperando un tipo determinado de respuestas. Si recibe un mensaje que no está en este conjunto, debe enviar un mensaje de tipo `not-understood`. Este mismo mensaje se emplea también para indicar al emisor que no se ha entendido su mensaje porque estaba mal formado o porque lo que ha llegado resulta ininteligible. Así mismo, un agente puede rehusar (`refuse`) el tratar el mensaje que acaba de recibir (por ejemplo, porque está realizando una tarea crítica que no puede interrumpir). Ambos tipos de mensaje terminan el diálogo entre los dos agentes.

Se ha preferido mantener los protocolos y mensajes propuestos por FIPA siempre que ha sido medida de lo posible, con el fin de conseguir una especificación cercana al estándar.

3.1 Propuesta

El proceso se desencadena cuando un agente emite una propuesta al grupo. En ese momento, todos los miembros del grupo reciben la propuesta y deben evaluarla, emitiendo un veredicto de acuerdo (`accept-proposal`) o desacuerdo (`reject-proposal`) con la propuesta antes de que se cumpla el plazo indicado por el agente iniciador.

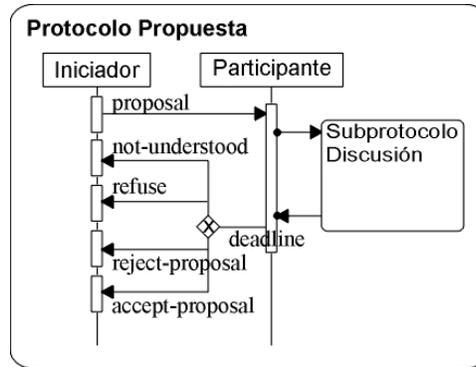


FIGURA 2. Protocolo para la creación de propuestas

El agente puede emitir el juicio directamente o puede iniciar una discusión con el resto del grupo sobre la propuesta antes de decidir su voto.

3.2 Discusión

El protocolo de discusión permite a los agentes explicar la propuesta inicial que se ha emitido o bien emitir una contrapropuesta. Una contrapropuesta es, desde el punto de vista de los mensajes, equivalente a una propuesta, por lo que se utiliza el mismo protocolo. La única diferencia es que la contrapropuesta debe etiquetarse con la propuesta inicial (para que los agentes puedan controlar el flujo de la comunicación) y el plazo máximo nunca debe ser superior al de la propuesta inicial (si no se indica nada, se asume el mismo).

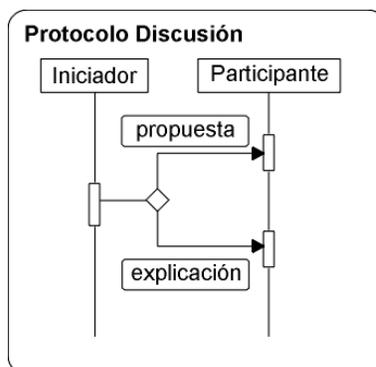


FIGURA 3. Protocolo para la discusión de propuestas

Por otra parte, el agente puede solicitar una explicación de la propuesta para decidir si le satisface o no. En ese caso, utiliza el protocolo de explicación, que le permite discutir acerca de una propuesta en concreto.

Con este esquema, todos los agentes deben emitir sus contrapropuestas y, una vez emitidas, se discuten y se "votan". No se admite la creación de una contrapropuesta dentro de la explicación de una propuesta anterior. El

protocolo no fuerza a que los agentes deban estar de acuerdo con alguna de las propuestas, pudiendo rechazarlas todas.

3.3 Explicación

Las explicaciones de las propuestas se emplean para que un agente sea capaz de decidir si la acepta o la rechaza. Se proporcionan dos mecanismos: los comentarios, que no es más que el envío de un mensaje dando una opinión, y la aclaración, que es una pregunta que requiere una contestación.

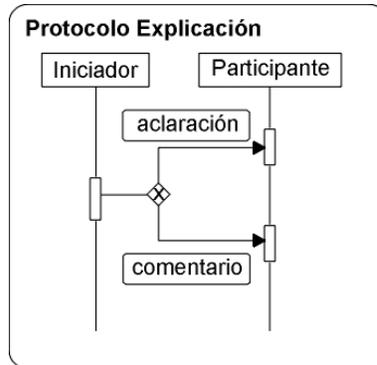


FIGURA 4. Protocolo para la explicación de propuestas

3.4 Aclaración

Una aclaración consiste en un intercambio de pregunta / respuesta entre agentes. La pregunta se lanza al grupo completo y cualquiera de ellos puede responder (incluso varios). Como puede observarse en el protocolo, si se acepta la pregunta es obligatorio responder a ella.

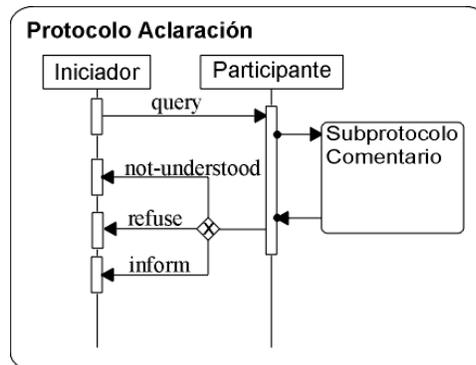


FIGURA 5. Protocolo para la aclaración de propuestas

3.5 Comentario

El mensaje más simple que se puede realizar es el emitir un comentario. Se trata simplemente de una opinión que no espera contestación por parte de otros agentes.

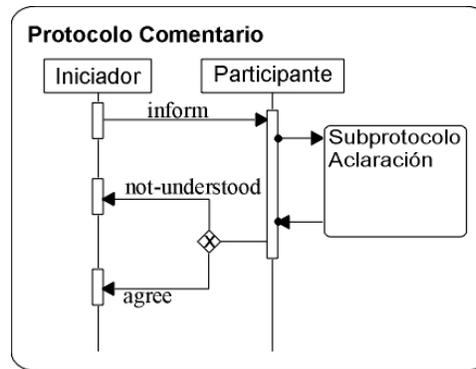


FIGURA 6. Protocolo para el comentario de propuestas

Con el fin de mantener la misma filosofía en todos los protocolos, se ha incluido un mensaje de aceptación (*agree*) dentro del protocolo de comentario. La única función de éste mensaje es que el emisor pueda confirmar que el grupo ha recibido y entendido el comentario.

4 Conclusiones

En este artículo se ha presentado un conjunto de protocolos que permiten intercambiar conocimiento entre agentes inteligentes empleando un esquema de co-edición argumentativa. La principal contribución es el asegurar que la conversación termina en algún momento, con lo que se consigue un esquema susceptible de ser tratado automáticamente por ordenadores.

Las principales aplicaciones de este protocolo de comunicación se dan en el campo del trabajo cooperativo asistido por ordenador (CSCW), la educación a distancia o el tele-trabajo. Actualmente, se está desarrollando un primer prototipo de este protocolo dentro de una herramienta para la creación de aulas virtuales. En todas estas aplicaciones los agentes son un apoyo para las personas que comparten información a través de Internet.

Si se obtienen buenos resultados, el siguiente paso es emplear estos protocolos para la comunicación directa entre agentes de forma autónoma, permitiéndoles intercambiar información sobre sus creencias.

5 Referencias

- [Wooldridge 1995] M. Wooldridge;N.R. Jennings. Intelligent agents: Theory and practice. The Knowledge Engineering Review, 10(2):115–152, 1995.
- [Franklin 1996] Franklin, S., Graesser, A. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages. Springer-Verlag 1996.
- [Barros 1998] B. Barros; M.F. Verdejo . Designing wokspaces to support collaborative learning. In Proceedings of the IEA-98-AIE Workshop. 1998

[Bobrow 1990] D. Bobrow. Dimensions of Interaction. AAAI-90 Presidential Address, AI Magazine 12(3): 64–80. 1990

[Huhns 1998] Huhns, M., Singh, M. P.: Readings in Agents. Readings in Agents. Chapter 1, 1-24. 1998