

Un modelo de comportamiento para Agentes BDI colaborativos *

Sonia V. Rueda

Guillermo R. Simari

Departamento de Ciencias e Ingeniería de la Computación – Universidad Nacional del Sur
Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (L.I.D.I.A)
Instituto de Ciencias e Ingeniería de la Computación
Av. Alem 1253 – B8000CPB Bahía Blanca, ARGENTINA
Email: {svr, grs}@cs.uns.edu.ar

Resumen

El objetivo general de nuestra línea de investigación es elaborar un modelo de comportamiento para agentes BDI que participan en un sistema multi-agente cooperativo. Los miembros del sistema tienen capacidades limitadas y cuando estas resultan insuficientes para alcanzar sus metas comprometidas, interactúan buscando potenciar sus habilidades. La interacción genera un diálogo deliberativo en el cual uno de los interlocutores es el iniciador y otro el colaborador. Si se produce un conflicto el diálogo se transforma en negociación y su evolución depende de los roles asignados a los interlocutores. El modelo propone un formalismo para representar el conocimiento individual y social e incluye un protocolo de interacción y un algoritmo que especifica el mecanismo de razonamiento, acción e interacción de cada agente BDI.

Keywords:

Sistemas Argumentativos, BDI, Modelos de Comportamiento, Protocolo de Interacción, Juegos Dialógicos.

1. Introducción

Un sistema multiagente puede concebirse como una organización artificial formada por individuos dotados de algunas habilidades. Los individuos pueden compartir objetivos comunes o cada uno de ellos puede tener asignadas metas diferentes. Si la actitud es cooperativa pueden interactuar para brindar y recibir colaboración. La colaboración resulta particularmente útil cuando los agentes tienen habilidades diferentes, porque permite resolver problemas que ninguno de los miembros de la organización podría resolver individualmente. Como contrapartida, exige resolver los conflictos que puedan surgir como consecuencia de la interacción.

En nuestra línea de investigación consideramos que un agente es una entidad artificial con un propósito asignado, algunas habilidades y un comportamiento que combina reactividad, proactividad, autonomía

*Financiado parcialmente por Secretaría de Ciencia y Tecnología, Universidad Nacional del Sur, y por la Agencia Nacional de Promoción Científica y Tecnológica

y racionalidad. La autonomía le permite tomar decisiones referidas a cómo actuar considerando su propósito y su conocimiento acerca del entorno. La racionalidad favorece seleccionar *buenas* decisiones.

Los modelos basados en creencias, deseos e intenciones (BDI) intentan capturar la actitud mental de un agente racional que selecciona en cada momento una acción a realizar de acuerdo a la meta perseguida. En un ambiente cooperativo el razonamiento de cada agente BDI involucra deliberación y planificación, pero se agrega además cierto *compromiso social*. En este trabajo el compromiso social permite que cada miembro del sistema pueda brindar y solicitar colaboración participando en diálogos deliberativos. Cuando se produce un conflicto el diálogo se transforma en negociación y los interlocutores intentan alcanzar una solución adecuada para ambos.

La autonomía individual y el compromiso social son dos propiedades que se contraponen. La primera permite que cada agente *actúe* guiado por sus propias metas, de manera independiente al entorno al que pertenece. El compromiso social restringe el modo en que los agentes *interactúan* e influye en el comportamiento individual limitando el nivel de autonomía.

Nuestro trabajo está orientado al diseño de un algoritmo que permita modelar el comportamiento de un agente BDI colaborativo, tanto cuando actúa como iniciador de la interacción y como cuando interviene como colaborador. Estamos abocados además a la elaboración de un protocolo de interacción que permita especificar las diferentes alternativas a través de las cuales puede evolucionar un diálogo, comenzando siempre como un proceso deliberativo y transformándose en ocasiones en una negociación.

La introducción de normas, roles y relaciones condiciona las formas de intervención de cada agente en el diálogo y restringe el comportamiento individual. La estructura que relaciona a los roles es fija, pero la asignación de roles es dinámica y depende de los planes ejecutados hasta el momento por cada uno de los miembros de la organización.

2. El modelo BDI

El modelo BDI brinda los elementos esenciales para representar la actitud mental de un agente racional que actúa en un ambiente dinámico, sujeto a cambios bruscos y frecuentes. Las creencias conforman el conocimiento del agente acerca del mundo. Sus deseos e intenciones se refieren al estado que desea alcanzar y representan sus motivaciones y compromisos. La capacidad efectora está constituida por el conjunto de acciones que es capaz de ejecutar.

En nuestro trabajo las creencias van a ser representadas a través de un Programa en Lógica Rebatible, (Φ, Δ) , en el cual Φ es un conjunto consistente de hechos y Δ un conjunto de reglas rebatibles. La descripción detallada de DeLP puede obtenerse en [2]. Los deseos de un agente se representan como un conjunto de literales. Una intención es un literal específico, seleccionado dentro del conjunto de deseos como una meta comprometida.

Las habilidades de un agente quedan establecidas por un conjunto de acciones Γ que le permiten cambiar el mundo. Como se establece formalmente en [1] cada acción tiene asociado un conjunto de precondiciones, un conjunto de restricciones y un conjunto de consecuencias. El efecto de ejecutar una acción provoca un cambio en las creencias acerca del mundo de acuerdo a las consecuencias establecidas.

Un aspecto central para modelar racionalidad es entonces la vinculación entre conocimiento y acción. Un agente BDI va a estar dotado de cierto nivel de *razonamiento práctico* que le permite seleccionar la acción más adecuada a partir de su conocimiento respecto a sus metas y al estado del entorno en el que

se haya situado [8]. Es decir, la capacidad cognitiva del agente es la que le permitirá actuar adecuadamente en cada contexto particular. El razonamiento práctico involucra dos procesos fundamentales: *deliberación* y *razonamiento sobre medios y fines*. La deliberación permite decidir **qué** metas van a perseguirse, el razonamiento sobre medios y fines determina **cómo** van a alcanzarse.

En nuestro trabajo partimos de la formalización propuesta en [5] para modelar el *estado* y el *comportamiento* de un agente BDI en un momento determinado. Cada agente determina un nuevo conjunto de creencias a partir de una función de revisión de creencias, *brf*, que se evalúa sobre las percepciones *P* y creencias actuales. La función de generación de opciones, *options*, mapea los conjuntos de creencias y de intenciones en el conjunto de deseos. La función de deliberación, *deliber*, genera el conjunto de intenciones a partir de las creencias, deseos e intenciones previas. La función *plan* retorna una secuencia de acciones a partir del conjunto de creencias y de la intención comprometida. En aplicaciones en las cuales el entorno sufre cambios frecuentes, el algoritmo puede ser refinado progresivamente para incorporar reactividad y reconsideración de intenciones. En este caso, la operación *execute* no abarca al plan completo, sino que luego de ejecutar cada acción individual se percibe y analiza la necesidad de ajustarlo. La reconsideración de intenciones permite tomar ventaja de los cambios y aprovechar nuevas oportunidades.

Nuestra línea de investigación propone una alternativa a este modelo en la cual los deseos de un agente no se determinan a partir de las creencias actuales y las intenciones previas, sino que se perciben como necesidades en el entorno. Esto es, la función $options(B, I)$ se elimina y en su lugar los deseos se actualizan a partir de una función de revisión de deseos que considera la percepción *P*.

Consideramos que el conjunto de metas que un agente percibe e incorpora a sus deseos se corresponde con las acciones que es capaz de ejecutar. Esto no garantiza que se pueda construir un plan ejecutable, el conjunto de creencias puede resultar insuficiente. Decimos entonces que el plan está incompleto. La operación de deliberación se modifica también porque no considera las intenciones alcanzadas previamente sino que selecciona una nueva intención considerando exclusivamente las metas asignadas y las creencias actuales.

El modelo BDI es abstracto, no especifica cómo se representa el conocimiento de un agente, ni propone un mecanismo deliberativo concreto o una estrategia de planificación específica. La selección de intenciones puede realizarse de diferentes maneras, una de ellas es que todos los deseos tengan la misma importancia, otra, asociarle a cada una de ellos una *prioridad*.

En nuestro trabajo, adoptamos un mecanismo de planificación basado en argumentación rebatible, tal que, cuando un agente *a* adopta una intención ι , y ι no está garantizada por (Φ, Δ) *a* busca una acción *A* en Γ que modifique las creencias de forma tal que pueda obtenerse un argumento sin derrotadores que soporte a ι . Si *A* no puede ejecutarse porque sus precondiciones no están garantizadas, *a* elabora una secuencia de acciones que le permitan establecerlas. La secuencia completa, incluyendo a *A*, conformará un plan para ι .

Cuando un agente forma parte de un sistema y sus capacidades resultan insuficientes para completar un plan, puede solicitar colaboración. Como contrapartida, sus acciones pueden interferir con los planes de los demás miembros. Si su actitud es cooperativa, brindará colaboración cuando reciba requerimientos y solicitará autorización para provocar cambios que puedan afectar a otros.

Así, en un sistema multiagente la representación de la capacidad cognitiva no sólo determina relación entre conocimiento y acción, sino que también influye en la interacción con otros agentes. Nuestro trabajo adopta la propuesta presentada en [?] para representar el conocimiento de un agente social. Las creencias individuales de cada agente están integradas por sus creencias específicas, las creencias compartidas con cada uno de los otros miembros del grupo y el conjunto de creencias compartido global. Cada conjunto de creencias individual es consistente.

3. Colaboración entre agentes BDI

El plan de un agente individual se construye considerando únicamente sus creencias, intenciones y habilidades. Cuando un agente forma parte de un sistema, puede interactuar con otros miembros del sistema en el que participa y aprovechar sus creencias y sus habilidades.

La interacción aumenta las posibilidades de alcanzar las metas comprometidas, pero exige como contrapartida considerar el impacto que el comportamiento social provoca sobre la representación del conocimiento, el mecanismo deliberativo, el proceso de planificación y el modelo de comportamiento de cada agente individual.

En nuestro trabajo la interacción se estructura en *diálogos* entre agentes que buscan colaborar entre sí para aumentar sus capacidades individuales. Un agente *iniciador* realiza una *convocatoria al diálogo* a través de un mensaje que envía a los demás miembros del sistema. Cada receptor responde al requerimiento manifestando su disponibilidad para colaborar, su incapacidad para hacerlo o rechazándolo porque provoca conflicto. El iniciador entablará un diálogo con uno de los agentes que se manifestaron disponibles para colaborar.

El comportamiento de los interlocutores sigue siendo racional. Esto es, como antes cada individuo debe balancear el esfuerzo que dedica a elaborar un plan efectivo, en relación al tiempo que emplea en ejecutarlo. En un diálogo, además, cada agente destina recursos para solicitar y brindar colaboración, nuevamente es importante que mantenga un equilibrio entre la interacción y la ejecución de sus planes. Un diálogo *deliberativo* comienza con un problema abierto, de naturaleza práctica y la meta es una decisión acerca de cómo actuar [9]. La interacción puede fracasar, transformarse en *negociación* o concluir en forma exitosa. La negociación comienza cuando se produce un enfrentamiento entre las motivaciones individuales. El proceso dialéctico puede pensarse como una búsqueda que intenta encontrar una solución adecuada para ambos. La búsqueda también puede fracasar o tener éxito. Cada forma de diálogo, deliberación y negociación, va a estar modelada a través de un diagrama de secuencia en UML que describa las diferentes alternativas de acuerdo a las cuales puede evolucionar la interacción.

Un diálogo termina exitosamente si se encuentra un plan conformado por acciones que los interlocutores están dispuestos a ejecutar y que les permiten alcanzar un acuerdo compartido. En este trabajo los agentes intercambian propuestas y contrapropuestas intentando hacer efectiva la colaboración.

Como en una organización humana, la especificación de normas, roles y relaciones permite balancear la autonomía con el compromiso social que impone formar parte de un grupo. El comportamiento individual y la interacción queda determinado por la relación entre los roles que ocupan los interlocutores durante el diálogo. Esto es, la estructura en la cual se relacionan los roles dentro de la organización, restringe las locuciones posibles en determinados momentos del diálogo. La autonomía sigue siendo una cualidad importante, pero restringida al comportamiento esperado para el rol en el contexto que determina la estructura de relaciones.

Nuestro trabajo se orienta también a la introducción de normas, roles y relaciones en la representación del sistema de agentes, la modificación del comportamiento individual de cada uno y del protocolo de interacción. Las normas y la estructura que relaciona a los roles es fija adoptando la propuesta presentada en [4]. Los roles se asignan dinámicamente considerando las aptitudes de los interlocutores. La evolución de cada diálogo particular va a quedar determinada por el protocolo de interacción y la relación entre los roles específicos que ocupan los interlocutores. La iniciación de un diálogo depende del comportamiento individual del agente que solicita colaboración y del que la brinda.

4. Conclusiones

En un sistema multiagentes un grupo de entidades autónomas pueden cooperar en la resolución de un problema que escapa a las capacidades individuales de cada uno. En un ambiente cooperativo el razonamiento de cada agente BDI involucra deliberación y planificación, pero se agrega además cierto *compromiso social*. En este trabajo el compromiso social permite que cada miembro del sistema pueda brindar y solicitar colaboración participando en diálogos deliberativos. Cuando se produce un conflicto el diálogo se transforma en negociación.

La autonomía individual y el compromiso social son dos propiedades que se contraponen. La primera permite que cada agente *actúe* guiado por sus propias metas, de manera independiente al entorno al que pertenece. El compromiso social restringe el modo en que los agentes *interactúan* e influye en el comportamiento individual limitando el nivel de autonomía.

El algoritmo que modela el comportamiento tiene que considerar las dos formas de acuerdo a las cuales puede interactuar. Esto es, como iniciador de un diálogo o como colaborador. Cuando un agente adopta una intención que corresponde a una meta propia y elabora un plan que no está completo, realiza una convocatoria al diálogo que puede concluir en forma exitosa o no.

En el caso de que la intención comprometida corresponda a un requerimiento de colaboración el agente adopta la posición opuesta en el diálogo. Esto es, manifiesta su disposición a colaborar y queda a la espera de que se entable un diálogo deliberativo.

Referencias

- [1] Simari, G., García, A., Capobianco, M. *Actions, Planning and Defeasible Reasoning*, Proceedings of the 10th International Workshop on Non-Monotonic Reasoning (NMR 2004), pp. 377–384, Whistler, Canada, June 6-8, 2004.
- [2] García, A., Simari, G. *Defeasible Logic Programming: An Argumentative Approach*, Theory Practice of Logic Programming, vol 4, 1, pp. 95–138, 2004.
- [3] Rueda S., García A., Simari, G. *Alternativas para la representacion del conocimiento compartido entre agentes colaborativos*, II Workshop en Agentes y Sistemas Inteligentes (WASI), VIII CACIC, Argentina, pp. 1508–1919, 2004.
- [4] Rueda S., Simari, G. *Roles y Relaciones en una Organización de agentes BDI*, III Workshop en Agentes y Sistemas Inteligentes (WASI), IX CACIC, Argentina, pp. 1370–1381, 2005.
- [5] Wooldridge, M., *Intelligent Agents*, in Multiagent Systems A Modern Approach to Distributed Artificial Intelligence. Editor G. Weiss. The MIT Press, 2000.
- [6] Fagin, R., Halpern J., Moses, Y., Vardi, M. *Reasoning About Knowledge*, The MIT Press, Cambridge, MA 1995.
- [7] Rao, A. S., and Georgeff M. P., *BDI Agents: From Theory to Practice*, In Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95) pp.312-319 San Francisco, USA, 1995.
- [8] Bratman, M.E., Israel D.J., Pollack, M.E., *Plans and resource-bounded practical reasoning*, Computational Intelligence 4: 349,355 1988.
- [9] Walton, D. N., Krabbe E.C., *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, State University of New York Press, Albany, NY, 1995.