

Comunicación y colaboración entre agentes artificiales

Guillermo Aguirre,[†] Marcelo Errecalde,[†] Guillermo Simari[‡]

[†]Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC)¹
Departamento de Informática. Universidad Nacional de San Luis
Ejército de los Andes 950 - Local 106. (D5700HHW) - San Luis - Argentina
Tel: (02652) 420823 / Fax: (02652) 430224. e-mail: {gaguirre, merreca}@unsl.edu.ar

[‡]Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA)²
Departamento de Ciencias e Ingeniería de la Computación. Universidad Nacional del Sur
Av. Alem 1253, (B8000CPB) Bahía Blanca, Argentina
Tel: (0291) 459-5135 / Fax: (0291) 459-5136. e-mail: grs@cs.uns.edu.ar

Resumen

No se puede relativizar la importancia que tiene la comunicación cuando un grupo de agentes inteligentes deben colaborar para realizar distintas tareas conjuntas. En la mayoría de estos casos, el comportamiento coordinado del grupo se logra a partir del supuesto de que los agentes poseen determinadas creencias sobre diversos aspectos del mundo que los circunda. Para lograr que estas creencias sean consistentes a nivel global, es fundamental que los agentes se comuniquen de manera efectiva. En este sentido, la investigación sobre los *actos del habla* (en inglés *speech acts*) ha establecido algunas de las condiciones fundamentales que los interlocutores y el entorno deben cumplir en un proceso de comunicación efectivo. Han surgido numerosas propuestas para sistematizar la aplicación de la teoría de los actos del habla a la comunicación entre agentes artificiales. En este trabajo se describen brevemente algunas de las principales propuestas que están siendo consideradas en la línea de “Agentes y Sistemas Multi-agente” del LIDIC, en conjunto con investigadores del LIDIA. El objetivo de este trabajo es presentar los principales aspectos de los enfoques que intentan formalizar el rol de la comunicación en las creencias de los agentes, destacando similitudes y diferencias. También se introducen los aspectos generales de un nuevo enfoque que está siendo elaborado en nuestro grupo de investigación basado en programación lógica rebatible.

1. Introducción

El presente trabajo se desarrolla en el contexto de las tareas de investigación que están llevando a cabo de manera conjunta, investigadores del LIDIC y del LIDIA, relacionadas al diseño de agentes colaborativos con capacidades cognitivas de alto nivel. En este contexto, la axiomatización del estado de creencias de los agentes necesario para lograr un comportamiento coordinado es una componente fundamental, como así también el impacto que tiene la comunicación en las creencias de los agentes.

Este trabajo tiene los siguientes *objetivos*

- Introducir las ideas más destacadas de la teoría de los *actos del habla*.
- Presentar algunas de las propuestas para su formalización.
- Destacar la importancia de la comunicación en el trabajo colaborativo.

¹Financiado por la Universidad Nacional de San Luis y la ANPCyT (PICT 2002, Nro. 12600).

²Financiado por la Universidad Nacional del Sur y la ANPCyT (PICT 2002, Nro. 13096).

En la sección 2 se presenta una breve síntesis de la teoría de los actos del habla. Las principales propuestas para especificar el efecto de los actos del habla son introducidas en la sección 3. En la sección 4 se describe en forma sintética algunos de los aspectos que deberían ser considerados para llevar a cabo esta tarea empleando programación lógica rebatible.

2. La teoría de speech acts

A partir de las ideas de Austin planteadas en [1] a comienzo de los años 60, se produce un cambio en la manera de analizar los enunciados empleados en la comunicación. Hasta entonces, solamente se rescató la valoración lógica (como *correctas* o *incorrectas*) que pudiera hacerse sobre las frases, reflejando con estas posiciones el pensamiento positivista propio de la época. Austin en cambio, resalta la importancia de ampliar el análisis del lenguaje hacia un universo mucho más amplio: el lenguaje cotidiano, es decir, rescatar el valor pragmático de las frases que la gente común usa a diario. Una pregunta, una promesa o una orden tienen sentido más allá de clasificarlas como verdaderas o falsas; son *exitosas* en la medida que consiguen las intenciones del orador. Surge así una nueva percepción del *lenguaje como acción*. Cuando un juez dicta sentencia en la corte y dice: "Culpable", se produce un cambio en el mundo real tan concreto como el producido por cualquier acción física.

Se pueden reconocer en la propuesta austiniana dos ejes fundamentales: a) la consideración del lenguaje operando sobre lo real, lo concreto (no lo simbólico) y b) una nueva percepción del éxito de un enunciado en función de llegar a satisfacer las intenciones que le dieron origen. En la teoría de los actos del habla se identifican tres actos o dimensiones para cada enunciado:

- el acto de emitir los sonidos correspondientes al vocabulario usado, el acto *locutorio*.
- el acto *ilocutorio* que se lleva a cabo **cuando** se dice algo: prometer, afirmar, amenazar, y
- el acto *perlocutorio* que se produce **como consecuencia** de decir algo: intimidar, asombrar, convencer, ofender.

La realización exitosa de un enunciado ejecutivo (performative) requiere satisfacer ciertas condiciones relativas a las circunstancias y personas involucradas, las que deben participar del procedimiento completo y de una manera sincera (sin actuar). Estas condiciones fueron refinadas por John Rogers Searle a fin de establecer el contexto más propicio para lograr los resultados esperados.

Los distintos enfoques que buscan sistematizar el empleo de los actos del habla, dan mucha relevancia al contexto y las condiciones generales que deben cumplir los participantes, particularmente las creencias y metas que tienen.

3. Enfoques para la especificación de los actos del habla

A partir del surgimiento de la idea de los actos del habla se propusieron distintos enfoques destinados a axiomatizar el efecto de los actos del habla en las creencias de los interlocutores participantes. En la sección 3.1 se describe el enfoque de Cohen y Perrault quienes proponen el uso de planes para implementar los actos del habla. En la sección 3.2 se introducen los actos del habla según el enfoque de Perrault, usando lógica default. En la sección 3.3 se considera una propuesta similar a la anterior pero basada en este caso en una lógica autoespistémica jerárquica.

3.1. Propuesta de Cohen y Perrault

Considerada la aproximación pionera [2] al tratamiento de los actos del habla con técnicas de inteligencia artificial, es una propuesta genérica que considera la elaboración de *planes* donde se integran tanto acciones físicas como acciones lingüísticas. Siguiendo una notación tipo *STRIPS*, Cohen

y Perrault, caracterizan las acciones involucradas de acuerdo a las precondiciones y poscondiciones propias de cada tipo. La combinación de estos *operadores* permite a los agentes realizar tareas más complejas en las que se deben considerar las creencias y objetivos propios y de los otros agentes. Considerando que la resolución de problemas consiste en alcanzar un estado objetivo realizando una secuencia de acciones a partir de un estado inicial, tiene sentido establecer un modelo adecuado para integrar los actos del habla en la planificación de tareas. En la vida diaria las personas suelen construir, ejecutar, simular y depurar planes como así también determinar los planes de otras personas en base a su comportamiento. Así durante una comunicación, las personas esperan que su interlocutor reconozca sus planes y facilitan este reconocimiento para que la comunicación sea efectiva. Entonces, por un lado es necesario que el orador escoja un plan adecuado a las circunstancias y por el otro, que el oyente sea capaz de identificarlo.

A fin de considerar un contexto adecuado en el que se realizan los actos del habla, los agentes deben tener una correcta representación de sus creencias y de las del interlocutor. Algo similar ocurre con las intenciones. Las condiciones previas y las posteriores deben ser descritas con precisión y para ello se usa un lenguaje lógico que represente los mundos posibles que se van sucediendo. El artículo se concentra en tres operadores: *requerir*, *informar* y *preguntar*. Se hace un refinamiento de los dos primeros para poder combinarlos adecuadamente y construir un plan para requerir información (preguntar). Los operadores resultantes son probados en situaciones complejas que llegan a considerar la participación de un tercer agente: “Pregúntale a Juan para que me diga donde están las llaves”.

3.2. Propuesta de Perrault

Para reflejar los efectos de los actos del habla en el estado mental de los interlocutores, Perrault [3] introduce la teoría de *persistencia de creencias*, que luego formaliza mediante los axiomas que integran la teoría default. Usando razonamiento default describe los cambios de actitudes en el tiempo, las transferencias entre agentes, la actualización de los estados mentales de los participantes, la consideración de las intenciones como disparadores de acciones, etc., todas cualidades requeridas en este enfoque. Las creencias previas persisten y las nuevas son adoptadas como resultado de observar los hechos externos, siempre que no estén en conflicto con los conocimientos previos. Como caso particular, se asume que un agente adopta las creencias que él cree que otro agente tiene y que no contradicen las propias.

En esta propuesta se consideran las creencias de un observador de la escena donde se hace el anuncio. El conocimiento que posee el observador es una descripción sobre las creencias de los participantes. Ese estado del mundo evolucionará según se sucedan los distintos enunciados. La descripción del mundo se realiza mediante las *extensiones* obtenidas desde la teoría default. Todas las reglas default utilizadas son normales por lo que se garantiza la existencia de al menos una extensión y, en caso de existir más de una, éstas serán mutuamente inconsistentes. No es posible establecer preferencias entre reglas default que compiten y no se puede lograr un cambio de creencia, es decir, pasar de creer p antes del acto comunicativo a creer $\neg p$ después del mismo. Esta limitación, reconocida por el mismo Perrault, es justificada por el deseo de mantener una mayor simplicidad y claridad en su propuesta.

Para probar la correctitud de la teoría default propuesta en cada caso, se detalla la verificación de la extensión correspondiente (E) para los casos de una *afirmación sincera* y para una *mentira*. Para ambos actos, se parte de la situación inicial en que se produce el anuncio (W) y mediante la aplicación sucesiva de axiomas y reglas default (D) se alcanza el conjunto de fórmulas E , que es la clausura de los defaults D sobre los hechos W . Si todo sucede de acuerdo a lo considerado, luego de una afirmación sincera p se alcanzará un conocimiento mutuo del contenido de la proposición (p) y la creencia mutua de que se mantienen las condiciones previas. En el caso que el enunciado sea una mentira, se llega al mismo conocimiento mutuo que en el caso previo pero el orador continúa creyendo la negación de la proposición, $\neg p$.

El conjunto inicial de axiomas y reglas default se restringen inicialmente a las *creencias* pero luego

se consideran las intenciones, y se agrega un nuevo conjunto de axiomas y reglas defaults, lográndose el cuerpo axiomático definitivo. Al reconsiderar los actos analizados, las nuevas extensiones ahora incluyen las intenciones puestas en juego por el orador.

Finalmente Perrault considera que un acto del habla es definido en el contexto de una teoría default como las condiciones que deben cumplir las extensiones obtenidas a partir de la teoría de acuerdo al conjunto definitivo de reglas default. Dichas extensiones deben describir el estado mental inicial de los participantes, el hecho del anuncio y las condiciones de observación.

3.3. Propuesta con lógica autoespistémica jerárquica

Aquí se presenta una propuesta alternativa para la revisión de creencias de los agentes involucrados en una acción de comunicación. La lógica autoespistémica jerárquica (HAEL por sus siglas en inglés) permite manejar de una manera elegante tanto las creencias de un agente como lo que puede ser asumido en ausencia de información más específica. En [4] se destaca la importancia de una adecuada descripción del contexto en el cual se realizan los enunciados y sus consecuencias. Es decir hay una clara dependencia entre una correcta representación de las circunstancias que rodean a la acción y la determinación de las consecuencias de la misma. En Hael, se distingue entre las creencias (o su ausencia) que son consecuencia de los hechos conocidos y las creencias que se asumen como un default porque información más específica está ausente. Esto caracteriza al proceso de revisión de creencias presente en este tipo de lógicas. Con Hael, los autores atacan “dos deficiencias de la lógica autoespistémica”: la introducción de preferencias entre inferencias default y la forma en que se determina si una proposición forma parte de la expansión de un conjunto base.

No es una única teoría uniforme, sino una estructura de varias subteorías enlazadas en una jerarquía. Cada subteoría representa una fuente de información distinta disponible para el agente y la jerarquía expresa la forma en que esta información es combinada. Las distintas subteorías mantienen un ordenamiento indexado por un operador sobre las propias creencias. Por ejemplo, al representar los defaults de una taxonomía, la información más específica tendría precedencia sobre los atributos generales. Para el caso de los actos del habla, en el nivel más bajo se encuentran las creencias del orador y del oyente previas al enunciado. En el próximo nivel, los efectos producidos en las creencias de los participantes luego de realizado el acto comunicativo. En la siguiente subteoría se mantienen las creencias de cada uno sobre las creencias del otro. En todas las subteorías se mantienen axiomas que establecen los hechos y deducciones que se producen sobre las intenciones y creencias de los agentes. Estos cambios de creencias se producen básicamente por un par de operadores: uno para las creencias de un agente y otro para describir sus intenciones.

Se definen axiomas distintos para describir las consecuencias de los enunciados en el orador y en el oyente, por ejemplo para asegurar que el hecho de producir un anuncio de una proposición no cambia el estado mental ³ de quien lo dijo. Al analizar el proceso de revisión de creencia que haría un oyente se presentan distintas alternativas. Si la creencia previa es más fuerte que el enunciado, la creencia inicial $\neg p$ debiera estar ubicada en el nivel más bajo para bloquear un cambio de idea. Caso contrario, debería quedar en un nivel más alto, para permitir que adopte el contenido de la proposición.

4. Enfoque basado en Programación lógica rebatible

Los sistemas multi-agente artificiales deben poder enfrentarse a situaciones cada día más complejas en un mundo cambiante, con acceso solamente a información parcial y muchas veces incierta. Los entornos donde se desempeñan los agentes son dinámicos y generalmente tienen una perspectiva limitada, local, obtenida mediante ciertas variables relevantes. Es importante que las creencias de un agente le permitan razonar y actuar en distintas circunstancias y concebir a sus metas como parte de

³En relación al contenido de la proposición

un objetivo más amplio. De esta manera, la acción que efectivamente se realizará puede ser escogida entre un grupo de acciones similares de acuerdo a la conveniencia del momento.

Una arquitectura de agente que provee estas características es la arquitectura BDI y corresponde al tipo de agente requerido para aplicar muchas de las propuestas presentadas en este artículo. Nuestro enfoque consiste en aplicar la teoría de speech acts con interlocutores que responden a la arquitectura BDI y que utilizan como mecanismo de representación y razonamiento un sistema de programación en lógica rebatible (*DeLP* [5]).

Diversos trabajos previos [6, 7, 8] han abordado la utilización de *DeLP* con agentes, acciones y planes. Las acciones son definidas por sus consecuencias y sus precondiciones, y la selección de la acción a ser ejecutada requiere analizar el contexto, para determinar si su uso es adecuado en una circunstancia o no. El agente representa sus metas como un conjunto de literales y a medida que realiza acciones se aproxima al estado objetivo en el cual esos literales estarán garantizados, es decir, no habrán podido ser derrotados. De esta manera se puede justificar haber cumplido con lo que se esperaba (estrictamente, su representación).

El encuentro de *DeLP* y la teoría de speech acts plantea como desafíos la representación de los estados mentales de los participantes, la elección del acto del habla adecuado según el contexto, lograr que cada acto funcione como un operador, establecer un adecuado mecanismo de revisión de creencias, integrar actos del habla en planes, poder deducir los planes del interlocutor, etc. Cada uno de estos aspectos constituyen temas de investigación de considerable complejidad que están siendo abordados en forma conjunta por nuestros grupos de investigación.

5. Conclusiones

En este trabajo, se han considerado distintos enfoques a la teoría de los actos del habla que otorgan especial importancia a la representación de los estados mentales de los participantes, al modo en que se reflejan las consecuencias de cada tipo de acto del habla, al manejo de las acciones y el contexto en el que se realiza el enunciado. En este sentido, las lógicas no monotónicas parecen tener las cualidades requeridas para abordar la teoría de los actos de habla.

Los planes para tareas que requieren colaboración entre agentes inteligentes podrían incluir acciones lüísticas. Al lograr definir en forma precisa a los actos del habla como operadores, es posible combinarlos con las acciones “físicas” de los agentes. De esta manera, una comunicación efectiva permite que agentes heterogéneos logren un estado mental conjunto y poder así colaborar.

Referencias

- [1] J. L. Austin. *How to Do Things with Words*. Harvard University Press, Cambridge, MA, Cambridge, MA, 1962.
- [2] Philip R. Cohen and C. R. Perrault. Elements of a plan-based theory of speech acts. *Cognitive science*, 3(3):177–212, 1979.
- [3] C. R. Perrault. An application of default logic to speech act theory. 1990.
- [4] Douglas E. Appelt and Kurt Konolige. A practical nonmonotonic theory for reasoning about speech acts. In *ACL*, pages 170–178, 1988.
- [5] Alejandro Javier García and Guillermo Ricardo Simari. Defeasible logic programming: an argumentative approach. *Theory and Practice of Logic Programming*, 4(2):95–138, 2004.
- [6] Guillermo Ricardo Simari, Alejandro Javier García, and Marcela Capobianco. Actions, planning and defeasible reasoning. In James P. Delgrande and Torsten Schaub, editors, *NMR*, pages 377–384, 2004.
- [7] Marcela Capobianco, Carlos Iván Chesñevar, and Guillermo Ricardo Simari. An argument-based framework to model an agent’s beliefs in a dynamic environment. In Iyad Rahwan, Pavlos Moraitis, and Chris Reed, editors, *ArgMAS*, volume 3366 of *Lecture Notes in Computer Science*, pages 95–110. Springer, 2004.
- [8] Marcela Capobianco, Carlos Iván Chesñevar, and Guillermo Ricardo Simari. An argumentative formalism for implementing rational agents. pages 1051–1062. 7mo Congreso Argentino de Ciencias de la Computacion (CACIC), 2001.