

Un Marco No-Monotónico para Representar Intenciones Bajo un Modelo BDI

José Martín Castro-Manzano¹, Axel Arturo Barceló-Aspeitia¹ y
Alejandro Guerra-Hernández²

¹ Instituto de Investigaciones Filosóficas
Universidad Nacional Autónoma de México
Circuito Mario de la Cueva s/n Ciudad Universitaria, México, D.F., México, 04510
jmcmanzano@hotmail.com, abarcelo@minerva.filosoficas.unam.mx

² Departamento de Inteligencia Artificial
Universidad Veracruzana
Sebastián Camacho No. 5, Xalapa, Ver., México, 91000
aguerra@uv.mx

Resumen Los avances que se han hecho sobre la propuesta filosófica de Bratman para representar intenciones son sumamente complejos e interesantes; sin embargo, aún existen detalles de su análisis que no han sido explorados completamente. Siguiendo este estudio fundacional sobre intenciones comenzamos a formalizar algunos de los resultados de dicha investigación y desarrollamos un marco no-monotónico para representar las intenciones. Usando este marco mostramos *i*) una distinción entre intenciones fuertes y débiles y *ii*) cómo es que este marco permite evitar los problemas del efecto colateral.

Key words: Revisión de intenciones, agentes BDI, no-monotonidad.

1. Introducción

Las intenciones son fenómenos cognitivos que han llamado mucho la atención desde dos perspectivas: la Filosofía y la Inteligencia Artificial. Y esto ha sido gracias a que las intenciones tienen propiedades muy interesantes: son proactivas, están íntimamente relacionadas con las creencias y los planes, al tiempo que proveen elementos para explicar la conducta de los agentes en términos de coordinación y predicción. Sin embargo, así como son muy interesantes, también presentan problemas muy interesantes, pues las intenciones no son meramente ejecutadas, sino que también son formadas, retenidas, combinadas, constreñidas, modificadas y reconsideradas [3].

En efecto, hay teorías filosóficas y formales sobre intenciones [3,6,7,14,18] pero muy pocas consideran la revisión de intenciones [13] o la no-monotonidad sobre intenciones [10] como posibilidades lógicas legítimas. Así, resulta ilustrativo notar que aunque el cambio de creencias sobre la base de nueva información ha sido estudiada con éxito durante los últimos 25 años [1,8,12,16], llevando a teorías de revisión de creencias y lógicas no-monotónicas para creencias, los

procesos dinámicos de otros estados cognitivos han recibido menos atención y, en particular, las intenciones [13].

En un sentido general, la meta de las lógicas no-monotónicas es formalizar el fenómeno del razonamiento por *default*; sin embargo, lo hacen de diferentes formas y, por tanto, existen diferentes sistemas no-monotónicos. La mayoría de ellos, no obstante, utilizan como unidad no-estándar la noción de *default* que puede ser descrita mediante expresiones como *típicamente* o *a menos que se muestre de otro modo*. Esta noción de *default*, común a muchos sistemas, no garantiza que los consecuentes se den siempre que se den los antecedentes y, por tanto, las lógicas no-monotónicas nos permiten derivar consecuencias de manera derrotable. Con todo, estos útiles sistemas generalmente se las ven con creencias en vez de otros componentes del espectro cognitivo.

Basados en un estudio clásico y fundacional acerca de las intenciones, el objetivo de este trabajo consiste en acercarnos a una formalización de dicho estudio para desarrollar un marco no-monotónico para representar intenciones. La motivación es sencilla: sabemos que la revisión de intenciones es posible dado que el análisis filosófico muestra que las intenciones poseen características como proactividad, inercia y admisibilidad [3]. Estas características, como veremos más adelante, garantizan que las intenciones necesitan, respectivamente, mecanismos de compromiso, derrotabilidad y consistencia que, a su vez, permiten la investigación sobre intenciones en términos de revisión y no-monotonicidad: así como los cambios en las creencias requieren una teoría de revisión de creencias y ciertas aproximaciones no-monotónicas, los cambios en las intenciones requieren de una teoría de revisión de intenciones.

El trabajo está organizado del siguiente modo. En la Sección 2 presentamos parte del análisis filosófico de las intenciones y el problema de la monotonicidad. En la Sección 3 desplegamos algunas de estas ideas en términos de un marco no-monotónico y discutimos algunos resultados. Finalmente, en la Sección 4 resumimos los resultados preliminares y mencionamos trabajo futuro.

2. Intenciones y no-monotonicidad

La teoría fundacional de la agencia BDI se encuentra principalmente en el trabajo de Bratman [3]. Ciertamente los avances que se han hecho sobre esta propuesta filosófica son sumamente complejos e interesantes, pero aún existen detalles de su análisis que no han sido explorados completamente. Así por ejemplo, al seguir los modelos de agencia BDI basados en las propuestas bratmanianas es usual interpretar las intenciones como un bloque unitario cuando, en contraste, Bratman ha distinguido tres tipos de intenciones [3]. Así pues, cuando un agente en t_1 intenta ϕ en t_2 como un proceso de deliberación, la intención se llama deliberativa. Si el agente tiene una intención, no por una deliberación, sino porque ha estado suspendida desde un momento previo t_0 y el agente la ha mantenido desde t_0 a t_1 sin reconsiderarla, se llama no-deliberativa. Finalmente, cuando las intenciones son generales y acerca de una circunstancia particular, se llaman intenciones basadas-en-políticas.

Estas últimas son de particular importancia tanto por su comportamiento como por su estructura. Las intenciones basadas-en-políticas se comportan como reglas porque tienen estructura de regla. Estas observaciones son importantes porque los formalismos existentes, aunque reconocen la íntima relación entre intenciones y planes, parecen olvidarse de que las intenciones tienen una estructura y conducta de reglas y que, por tanto, funcionan como planes. Ahora bien, como Bratman ha notado [3], los planes son intenciones. Esta idea nos permite pensar los planes como intenciones. En efecto, no podemos olvidar que los planes por sí solos son meramente cursos de acción que esperan por alguna forma de activación. Sabemos, también, que para que los planes se conviertan en intenciones necesitamos una forma de compromiso, ya que las intenciones son cursos de acción que el agente se ha comprometido a cumplir. Estas dos ideas –que las intenciones funcionan como planes y que los planes son intenciones– las podemos resumir al fijar las intenciones basadas-en-políticas como planes con una estructura $g : circ \leftarrow cuerpo$ [2] donde g denota la meta de la intención, $circ$ el conjunto de circunstancias particulares y $cuerpo$ el un conjunto de acciones o sub-intenciones.

Así pues, cuando vemos la descripción de las intenciones en conjunto podemos inferir razonablemente que las intenciones basadas-en-políticas son derrotables [10]. A manera de ejemplo, veamos el siguiente caso:

$$coloca(X, Y) \leftarrow acercate(X)$$

Esta fórmula nos indica que colocar X sobre Y típicamente implica que el agente se tiene que acercar a X . Si imaginamos que hay un agente moviéndose en el ambiente y adquiere la meta de colocar X sobre Y , entonces, *típicamente* el agente se acercará a X ; sin embargo, esta intención es derrotable en el siguiente sentido. El agente hará lo que tenga que hacer *a menos que* hayan evidencias –o circunstancias– que sugieran que tal cosa no es posible. Por tanto, podemos decir que el agente intenta ciertas cosas *típicamente*, pero no de manera absoluta e indiscutible. Y por ello podemos notar que todo esto constituye un requisito de derrotabilidad que difiere de las interpretaciones tradicionales de la agencia BDI que modelan la conducta intencional bajo la batuta de un sistema modal normal y, por ende, de un sistema monotónico.

2.1. Omnisciencia lógica e intenciones

Sin duda los sistemas monotónicos son útiles en sus contextos, pero fuera de ellos dan lugar a problemas como el de permitir a los agentes inferir todas las consecuencias de su base de conocimiento. El término es usualmente aplicado sobre creencias y conocimiento, en cuyo caso se suele hablar de *omnisciencia lógica*. Así, un agente es lógicamente omnisciente si sabe (o cree) todas las consecuencias de su base de conocimientos (o creencias). De esta manera, así como se habla de omnisciencia lógica para las creencias o conocimientos, se habla del problema del *efecto colateral* cuando se trata de intenciones [6].

Para modelar agentes BDI es usual seguir la estructura $B^{KD45}D^{KD}I^{KD}$. Y al enfocarnos en el fragmento I podemos ver que este funciona bajo las normas del axioma K y por ello esta estructura induce un sistema modal normal para las intenciones, y por tanto, el entendimiento de las intenciones en dicha estructura da lugar a los siguientes problemas (originalmente diseñados para explicar la omnisciencia lógica [15], aquí los seguimos para caracterizar el efecto colateral):

- *Cerradura*. Si un agente intenta un conjunto de fórmulas Γ y si Γ implica una fórmula γ , entonces el agente también intenta γ .
- *Irrelevancia*. Bajo la interpretación de mundos posibles una fórmula válida es aquella que es satisfecha en todo mundo posible. Esto implica que el agente ha de intentar todas las fórmulas válidas de su base de datos.
- *Inconsistencia*. Un agente no puede intentar γ y $\neg\gamma$ sin intentar cualquier otra fórmula.
- *Intratabilidad*. Computacionalmente, dado los problemas anteriores, el agente tiene que computar todas las consecuencias de sus acciones, lo cual, en la práctica, no parece posible dados los recursos limitados.

Estos problemas de la noción de efecto colateral pueden resumirse, formalmente, así [17]:

- 1 $\text{INT}(\phi) \wedge \phi \rightarrow \psi \vdash \text{INT}(\psi)$ (efecto colateral)
- 2 $\phi \rightarrow \psi \vdash \text{INT}(\phi) \rightarrow \text{INT}(\psi)$ (efecto colateral)
- 3 $\vdash \psi \Leftrightarrow \psi \rightarrow \vdash \text{INT}(\phi) \Leftrightarrow \text{INT}(\psi)$ (efecto colateral)
- 4 $\vdash \phi \rightarrow \vdash \text{INT}(\phi)$ (problema de transferencia)
- 5 $(\text{INT}(\phi) \wedge \text{INT}(\psi) \rightarrow \text{INT}(\psi \wedge \phi))$ (combinación irrestricta)
- 6 $\text{INT}(\phi) \rightarrow \text{INT}(\phi \vee \psi)$ (debilitamiento irrestricto)
- 7 $\neg(\text{INT}(\phi) \wedge \text{INT}(\neg\phi))$

Evidentemente, el problema aquí es que ninguna de estas propiedades es válida para modelar el comportamiento de las intenciones, excepto por la número 7 [10], la cual es una norma para tratar la inconsistencia intencional.

3. Un marco para representar intenciones

A partir del estudio filosófico de Bratman [3] dividimos las propiedades de las intenciones en tres grupos: propiedades funcionales, descriptivas y normativas. El conjunto de propiedades funcionales se encuentran en el tope de los componentes de la arquitectura BDI junto con las creencias y los deseos. Estas características funcionales tienen la virtud de que nos permiten distinguir las intenciones de las creencias y los deseos:

- *Pro-actividad*. Las intenciones son estructuras pro-activas, mueven al agente a alcanzar una meta [3]. En este sentido, las intenciones son componentes que controlan la conducta. Es importante notar, sin embargo, que las intenciones no son iguales a los deseos: tanto las primeras como los segundos son pro-actitudes, pero las intenciones implican compromiso y consistencia, los deseos

no. Como veremos más adelante, para indicar esta propiedad adjuntamos una fórmula g que denota la meta de la intención de tal manera que la pro-actividad quede implícita en la estructura sintáctica de la intención.

- *Inercia*. Una vez que una intención ha sido tomada resiste ser abandonada. Si una intención se toma y se abandona inmediatamente no tiene sentido hablar de una intención; sin embargo, si las razones por las cuales la intención fue tomada desaparecen, es racional abandonar la intención [3]. Para expresar esta propiedad necesitamos proveer al agente de un mecanismo de compromiso. Por el momento asumiremos que ese mecanismo está dado y es de tipo flexible (*single minded*) [18], pues si el agente carece de compromiso o tiene un compromiso ciego, no tiene sentido hablar de inercia [11].
- *Admisibilidad*. Las intenciones proveen un filtro de admisibilidad. Una vez que se toma una intención, esta restringe los posibles razonamientos prácticos: mientras el agente mantiene una intención particular, no considerará opciones contradictorias [3,6]. Para representar esta idea permitimos que el fragmento intencional se comporte bajo el axioma D de tal suerte que no sea el caso que $\text{INT}\phi$ y $\text{INT}\neg\phi$.

Es fácil ver, por tanto, que las intenciones son funcionalmente diferentes de las creencias y los deseos aunque comparten algunas propiedades. Así, por ejemplo, las intenciones y los deseos son proactitudes; y las intenciones y las creencias proveen filtros de admisibilidad. Lo interesante de estas propiedades es que nos indican que las intenciones requieren compromiso (dado el principio de pro-actividad), consistencia (por el principio de admisibilidad) y retractabilidad (por el principio de inercia).

El segundo grupo de propiedades están, por decirlo así, en un nivel más bajo. Estas son descriptivas y nos muestran características descriptivas de las intenciones.

- *Parcialidad*. Las intenciones son parciales e incompletas, pues no es posible tener información completa del mundo. Esta propiedad queda implícita en la definición de la intención como un plan, pues los componentes del cuerpo de la intención están parcialmente instanciados.
- *Dinamicidad*. Las intenciones no son estructuras estáticas, pues el mundo es cambiante. Podemos expresar esta característica por medio de mecanismos de aprendizaje que, por el momento, asumiremos como fijos. Estos mecanismos deben garantizar dos cosas: *i*) que los agentes pueden retraer intenciones cuando estas presentan problemas y *ii*) que puedan expandir sus intenciones cuando sea pertinente.
- *Jerarquía*. Las intenciones contienen razones medios-fines de manera ordenada. Expresamos la jerarquía entre planes por medio de una relación de orden que sea acíclica.

Finalmente llegamos al último grupo de propiedades normativas, las cuales son sumamente importantes ya que dictan cómo es que las intenciones se comportan y cómo es que los fragmentos B e I se relacionan, al menos, en principio.

- *Consistencia interna.* Las intenciones deben ser ejecutables. Requerimos que las circunstancias de la intención sean consecuencia lógica de las creencias del agente y que la meta de la intención exprese pro-actividad.
- *Consistencia fuerte.* Las intenciones deben ser consistentes con las creencias de los agentes. Esta propiedad queda implicada por la consistencia interna al requerir que las circunstancias sean consecuencia de las creencias.
- *Coherencia medios-fines.* El razonamiento medios-fines de las intenciones debe ser consistente con la red global de intenciones. Esta es preservada por el principio de admisibilidad.

Siguiendo esta caracterización podemos comenzar a definir el siguiente marco:

Definición 1 (*Marco no-monotónico intencional*) *Un marco no-monotónico intencional es una estructura $\langle R, S, F_{BI}, I \rangle$ donde:*

- $R = \langle \oplus, \ominus, \odot \rangle$ expresa la dinámica de las intenciones.
- $S = \langle \text{compromiso} \rangle$ expresa la estrategia de compromiso embebida en el agente.
- B expresa la base de creencias.
- F_{BI} representa los hechos, tanto en creencias como intenciones.
- $I = \langle I, >, \vdash, \vdash \sim \rangle$ t.q.
 - $I = \phi_1, \dots, \phi_n$
 - $> \subseteq I^2$ t.q. $>$ es acíclica.
 - \vdash es una relación de consecuencia monótona.
 - $\vdash \sim$ es una relación de consecuencia no-monótona.

La idea general del marco es sencilla y consiste en que podemos expresar los detalles del análisis filosófico anterior mediante este marco de la siguiente manera.

En el ítem R expresamos la dinamicidad de las intenciones por medio de funciones de aprendizaje: las funciones de reconsideración de intenciones –expansión \oplus , contracción \ominus y revisión \odot – indican cómo es que funciona la dinamicidad de intenciones [4]. A su vez, estas funciones nos permiten definir la admisibilidad [4]. En el ítem S expresamos la inercia al proveer un elemento que represente un mecanismo de compromiso. Por el momento asumiremos que ese mecanismo está dado y es de tipo flexible (*single minded*) [18], pues si el agente carece de compromiso o tiene un compromiso ciego, no tiene sentido hablar de inercia [11].

El componente B representa la base de creencias, las cuales pueden ser entendidas como literales. F_B , a su vez, representa las creencias de base, las creencias que se consideran indisputables. Similarmente, F_I representa las intenciones de base.

Finalmente los componentes de la estructura I nos permiten representar los restantes componentes. Cada intención $\phi \in I$ es una estructura $g : circ \leftarrow cuerpo$ donde g representa la meta de la intención –y así permitimos la proactividad–, $circ$ las circunstancias y el resto el cuerpo. En el caso de que las circunstancias o el cuerpo sean vacíos simplemente omitimos el resto del plan y así podemos representar las intenciones basadas-en-políticas como estructuras $g : circ \leftarrow$

cuero y el resto de las intenciones como un subconjunto de éstas tal que $g : \top \leftarrow \top$ o simplemente g .

La consistencia interna consiste en que $circ_\phi \in B$, e.d., que la circunstancia de la intención ϕ sea consecuencia lógica de la base de creencias y que g sea cabeza de la intención ϕ . Por tanto, la consistencia fuerte queda implicada por la consistencia interna (pues la segunda se define cuando $circ_\phi \in B$). La coherencia medios-fines queda garantizada por la admisibilidad. Y la jerarquía de las intenciones la expresamos mediante una relación de orden entre intenciones.

3.1. Intenciones fuertes y débiles

Usando este marco podemos mostrar dos cosas: *i*) una distinción entre intenciones fuertes y débiles y *ii*) cómo es que este marco permite evitar los problemas, previamente descritos, del efecto colateral. Comenzamos con el primer objetivo. Usamos $\phi[g]$ para denotar que g es la cabeza de la intención ϕ , $circ(\phi)$ denota el conjunto de creencias de la intención ϕ y $cuero(\phi)$ denota el cuerpo de la intención ϕ .

Definición 2 (Intención fuerte) *Sea I un conjunto de intenciones. Decimos que $I \vdash g$ si*

- 1) $g \in F_I$ o
- 2) $\exists \phi[g] \in F_I : circ(\phi) \in F_B \wedge \forall \psi[g'] \in cuero(\phi) \vdash g'$

Así pues, una intención es fuerte –o fuertemente derivada– si es un hecho básico e indisputable (1) o si hay otra intención fuerte ϕ con cabeza g cuyas circunstancias son creídas y cuyo cuerpo también es derivado fuertemente (2).

Similarmente, para las intenciones débiles:

Definición 3 (Intención débil) *Sea I un conjunto de intenciones. Decimos que $I \vdash \sim g$ si*

- 1) $I \not\vdash \neg g$ y
- 2) $\exists \phi[g] \in I : circ(\phi) \in B \wedge \forall \psi[g'] \in cuero(\phi) \vdash \sim g'$ y
- 3) $\forall g' \in I : I \vdash \sim g$ si $\phi[g] > \phi[g']$

Decimos que una intención es débil si su negación no es fuerte (1) y si hay una intención débil ϕ con cabeza g y cuyas circunstancias son creídas y cuyas subintenciones son derrotables (2). Y para toda meta g' , g es débilmente derivable si g tiene prioridad sobre g' (3).

Por tanto, es evidente que:

Proposición 1 *Una intención fuerte tiene prioridad sobre una débil:*

$$\vdash \phi \Rightarrow \vdash \sim \phi$$

Lo cual resulta claro dado que $F_I \subseteq I$ y $F_B \subseteq B$.

Proposición 2 *Las intenciones fuertes no son retractables.*

$$\vdash \phi \Rightarrow \neg(I \ominus \phi)$$

Lo cual resulta ser el caso al observar que si asumimos $\vdash \phi$ pero $(I \ominus \phi)$, entonces $\phi \notin I$, lo cual contradice la definición de una intención fuerte.

3.2. El efecto colateral

Con estos resultados podemos visualizar el segundo de nuestros objetivos: que nuestra representación de intenciones evita el problema del efecto colateral. Por las definiciones anteriores, las consecuencias de una intención son sub-intenciones que necesitan ser cumplidas para poder alcanzar la meta de la intención original. Si dejamos que ϕ y ψ sean intenciones, la consecuencia de una intención ϕ se caracteriza como $Cn(\phi) = \{\psi : \psi \in cuerpo(\phi)\}$. Por tanto, si consideramos un conjunto de intenciones I es fácil ver que

$$Cn(I) = \bigcup_{\phi \in I} \psi \in cuerpo(\phi)$$

Esta interpretación de las intenciones como $\phi \leftarrow \psi$ permite pensar en la noción de implicación de una manera procedimental y no veritativo-funcional. Por tanto, las consecuencias de una intención no funcionan como implicaciones del sistema K . Cuando requerimos que ψ tiene que cumplirse para alcanzar ϕ no nos referimos a una implicación de la forma $\psi \rightarrow \phi$, pues tal forma implica un problema de efecto colateral, que va en contra de las teorías intencionales fundacionales [3,6]. Por el contrario, entendemos esta como $\phi \leftarrow \psi$, como una receta de acciones, lo cual es coherente con la interpretación de las intenciones como planes. Así, un agente intenta sólo las subintenciones de una intención tomada y no todas las consecuencias de sus intenciones.

4. Conclusión

Al formalizar ciertos resultados del análisis de Bratman sobre intenciones, recuperamos la idea de que las intenciones son planes y no meramente estructuras atómicas de una arquitectura BDI. Con estas suposiciones construimos un marco para representar intenciones donde podemos aproximar una distinción entre intenciones fuertes y débiles así como una solución al problema del efecto colateral. La relevancia de esta aproximación consiste, principalmente, en la exploración de detalles que no han sido tratados completamente.

Trabajo futuro consiste en afinar las definiciones de las intenciones así como en comparar las virtudes de nuestra aproximación con otros enfoques [5,14,20]. La ventaja de nuestro acercamiento es que permite formalizar de manera directa los componentes del análisis bratmaniano que no habían sido explorados previamente así como una interpretación de las intenciones en términos procedimentales y no veritativo-funcionales.

Agradecimientos. Los autores agradecen a los revisores anónimos por sus acertados comentarios y precisas correcciones. El primer autor está apoyado por la beca CONACyT 214783.

Referencias

1. Alchourrón, C. E., Gardenfors, P., Makinson, D.: On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50, 510-530 (1985).
2. Bordini, R.H., Hübner, J.F., Wooldridge, M.: *Programming Multi-Agent Systems in AgentSpeak using Jason*. Wiley, England (2007)
3. Bratman, M.: *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge (1987)
4. Castro-Manzano, J. M.: An Introduction to Intention Revision: Issues and Problems, in: M. Osorio, C. Zepeda, I. Olmos, C. Medina and J. Arrazola (eds.), *Proceedings of the Fifth Latin American Workshop on Non-Monotonic Reasoning 2009*, CEUR, (2009).
5. Chen, X, Liu, G.: A logic of intention. In *ICJAI99*, (1999).
6. Cohen, P., Levesque, H.: Intention is choice with commitment. *Artificial Intelligence* 42(3), 213-261 (1990).
7. Dignum, F., Meyer, J.-J. Ch., Wieringa, R. J., Kuiper, R.: A modal approach to intentions, commitments and obligations: Intention plus commitment yields obligation. In M. A. Brown, J. Carmo (eds.), *Deontic logic, agency and normative systems*, pp. 80-97, Springer-Verlag (1996).
8. Fagin, R., Ullman, J. D., Vardi, M. Y.: On the semantics of updates in databases. *Proceedings of Second ACM SIGACT-SIGMOD*, Atlanta, 352-365 (1983).
9. Gardenfors, P., Makinson. D.: Revisions of knowledge systems using epistemic entrenchment. In *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, M. Vardi (ed.), Los Altos, CA: Morgan Kaufmann (1988).
10. Governatori, G., Padmanabhan, V. and Sattar, A.: A Defeasible Logic of Policy-based Intentions. In *AI 2002: Advances in Artificial Intelligence*, LNAI-2557. Springer Verlag (2002).
11. Guerra-Hernández A., Castro-Manzano, J.M., El-Fallah-Seghrouchni, A.: *CTLA-agentSpeak(L): a Specification Language for Agent Programs*. *Journal of Algorithms in Cognition, Informatics and Logic* (2009).
12. Harper, W. L.: Rational conceptual change. In *PSA 1976*, East Lansing, Mich: Philosophy of Science Association (1977).
13. Hoek, W. van der, Jamroga, W., Wooldridge, M.: *Towards a theory of intention revision*. Synthese, Springer-Verlag (2007).
14. Konolige, K., Pollack, M. E.: A representationalist theory of intentions. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-93)*, 390-395, San Mateo: Morgan Kaufmann (1993).
15. K. M. Sim: *Epistemic Logic and Logical Omniscience: A Survey*. *International Journal of Intelligent Systems*, vol. 12, 57-81, John Wiley and Sons, Inc (1997).
16. Levi, I.: *The Enterprise of Knowledge*. MIT Press, Cambridge, Massachusetts (1980).
17. Linder, B. Van: *Modal Logic for Rational Agents*. PhD thesis, Department of Computer Science, Utrecht University, 19th June 1996.

18. Rao, A.S., Georgeff, M.P.: Modelling Rational Agents within a BDI-Architecture. In: Huhns, M.N., Singh, M.P., (eds.) *Readings in Agents*, pp. 317–328. Morgan Kaufmann (1998).
19. Rao, A.S.: AgentSpeak(L): BDI agents speak out in a logical computable language. In: de Velde, W.V., Perram, J.W. (eds.) *MAAMAW. LNCS*, vol. 1038, pp. 42–55. Springer, Heidelberg (1996).
20. Sugimoto, T.: A preference-based theory of intention. In *PRICAI-2000*, Springer-Verlag, (2000).
21. Wooldridge, M.: *Reasoning about Rational Agents*. MIT Press, Cambridge (2000).