

Un modelo cognitivo para un agente BDI con compromiso social

Sonia V. Rueda
svr@cs.uns.edu.ar

Guillermo R. Simari
grs@cs.uns.edu.ar

Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR
8000 Bahía Blanca, Argentina

Abstract

Este trabajo propone un modelo para representar el conocimiento y el comportamiento de un agente BDI que actúa e interactúa condicionado por las normas y la estructura social de la organización a la que pertenece. Parte del conocimiento está compartido globalmente y algunos elementos cognitivos son específicos de cada individuo en particular. Cuando el conocimiento de un agente es insuficiente para alcanzar su meta comprometida inicia un diálogo colaborativo. La autonomía y reactividad le permiten adaptarse a las características dinámicas del entorno. La estructura normativa y la asignación de roles brindan patrones para anticipar la evolución de los diálogos.

Keywords: Sistemas Multiagente, Programación en Lógica, Argumentación, Diálogos Colaborativos, Roles, Normas

1 Introducción

Un gran número de aplicaciones desarrolladas siguiendo el modelo de agentes demandan altos niveles de colaboración. La colaboración ofrece un ambiente cooperativo que aumenta las posibilidades para que cada individuo alcance sus metas, pero también provoca interferencias y dependencias entre las tareas, exigiendo que las acciones sean coordinadas adecuadamente.

Una manera efectiva de lograr coordinación es imponer sobre los sistemas de agentes artificiales características propias de las organizaciones humanas. Estas organizaciones cuentan con patrones de comportamiento que regulan su funcionamiento. En particular, la especificación de normas, roles y relaciones restringe el comportamiento de los individuos, favoreciendo la concreción de las metas globales.

Nuestro trabajo parte del modelo de comportamiento para un agente BDI propuesto en [9] y lo extiende para incluir interacción a través de *diálogos colaborativos*. Cada individuo actúa restringido por las normas de la organización a la que pertenece e interactúa condicionado por la relación entre su rol y el de sus interlocutores.

La sección que sigue presenta alternativas para la especificación y asignación de roles. A continuación se introduce el concepto de norma social como un mecanismo para regular el comportamiento de un individuo responsable. La cuarta sección describe brevemente las características abstractas de un agente BDI dotado de razonamiento práctico y propone una alternativa basada en argumentación para representar el conocimiento y el mecanismo de planificación. La quinta sección extiende la alternativa propuesta para incluir roles, relaciones y una estructura normativa que regula la interacción entre los agentes. El algoritmo que modela el comportamiento de cada individuo en particular refleja su intervención en un diálogo colaborativo, tanto cuando actúa como iniciador como cuando interviene como colaborador. Por último se presentan algunas conclusiones y el trabajo futuro.

2 Roles y relaciones

En un contexto social el rol de un individuo determina el comportamiento esperado durante la interacción con otros miembros del grupo. En [11] se indica que los roles agrupan tipos de comportamiento en unidades significativas que contribuyen al logro de los objetivos del grupo. Las relaciones establecen vínculos entre los roles y permiten coordinar acciones, anticipar la evolución de la interacción y reducir el nivel de incertidumbre que provoca el dinamismo en el entorno. Una relación típica es la autoridad vertical o jerárquica.

La *especificación de los roles* en una organización artificial puede ser implícita o explícita. En el primer caso los roles se usan en forma intuitiva para analizar las cualidades de un sistema multiagente. En el segundo caso el concepto de rol toma mayor trascendencia porque brinda una estructura formal para construir sistemas multiagente para contextos dinámicos e impredecibles.

La *asignación de roles* puede ser fija o dinámica. El primer caso es el más simple y el rol de cada agente puede quedar establecido en el diseño de la aplicación. Cada miembro del sistema puede conocer su propio rol y el de los demás.

Una alternativa más flexible es que el rol de cada individuo pueda cambiar y deba ser considerado en el contexto de cada meta específica. Algunos miembros pueden estar dotados de la facultad de asignar roles determinando la estructura organizacional del sistema. La capacidad deliberativa de estos miembros probablemente sea más compleja que la de los demás, ya que requiere conocer las capacidades específicas de cada uno.

Una posición intermedia es que a cada agente se le asigne un rol en el momento en el que ingresa en el sistema. De este modo, el rol no queda fijado en el diseño, pero una vez que se establece, no cambia. Aun en este caso existen diferentes maneras de asignar un rol, entre ellas:

- *Imposición*: todo agente recibe un rol sin participar en la decisión. El objetivo puede ser, por ejemplo, que todos los miembros ocupen roles diferentes o por el contrario mantener varios agentes dentro de un mismo rol.
- *Aceptación sistemática*: todo agente es admitido en el sistema con el rol con el que desea ingresar.
- *Condicionada a un diálogo*: un agente que ingresa a un sistema expresa de alguna manera el rol que desea ocupar y si no es aceptado puede negociar hasta alcanzar un acuerdo que convenga a todos.
- *En función de las competencias*: en este esquema las habilidades y competencias de un agente van a determinar el rol o los roles que puede cumplir. Se requiere que la especificación de cada rol defina el conjunto de aptitudes que requiere y que la especificación de cada agente establezca el conjunto de habilidades y competencias que dispone.

Una propiedad fundamental es la *cardinalidad de los roles*, establece el número de agentes que deberían ocupar cada rol, como así también el número de roles que puede cumplir un individuo dado. La alternativa más estricta es que un agente ocupe exactamente un rol y un rol corresponda a un único agente. La más flexible es que cada rol pueda ser ocupado por un número no restringido de agentes, que a su vez pueden ocupar varios roles. Una posición intermedia, más compleja de modelar, es que algunos roles pueden tener restricciones específicas, por ejemplo puede haber roles de los cuales se exige que haya al menos un agente que ocupe esa función en cada momento dado o que haya a lo sumo uno asignado al rol. En particular, puede exigirse que exista un rol especializado, el administrador, entre cuyas funciones estará la admisión de nuevos miembros o el manejar los requerimientos de colaboración.

3 Las normas sociales

Las normas son fenómenos sociales que prescriben y proscriben patrones de comportamiento [12]. Un mismo individuo puede formar parte de diferentes grupos sociales, de modo que en cada uno deberá adaptarse a la estructura normativa correspondiente.

Las normas, así como también sucede con los roles, son elementos más o menos *estables* dentro de la vida social, por lo cual muchas veces pueden considerarse entidades estáticas. Sin embargo, las sociedades sufren transformaciones que las afectan. Estas influyen en el comportamiento social pero además son construidas a través de procesos sociales, basados en el comportamiento social. Esto es las normas son condiciones y resultados de procesos sociales [13].

En el diseño de la estructura normativa de una organización es importante considerar dos aspectos: la *perspectiva individual*, considerando como afectan las normas al mecanismo de razonamiento de los actores, y la *perspectiva institucional*, analizando el impacto de las normas sobre los procesos sociales que ocurren dentro de una organización.

Desde la perspectiva individual las normas pueden ser *restrictivas* o *impositivas*. Las normas restrictivas prohíben la ejecución de ciertas acciones en situaciones determinadas. Las normas impositivas exigen la ejecución de acciones en circunstancias establecidas.

Desde un enfoque institucional, algunas normas están codificadas en forma de *leyes* mientras que otras constituyen *convenciones*. La principal diferencia reside en que toda ley va a estar acompañada de algún mecanismo para detectar comportamiento ilegal y probablemente penalizarlo, en el caso de que se produzca. Las convenciones son lineamientos más débiles, pautas o guías que influyen en el comportamiento, sin restringirlo severamente.

Las acciones de los individuos, grupos y sociedades van a estar también determinados por *principios* y *valores* que guiarán el comportamiento. Mientras que las normas son restricciones concretas que se aplican en circunstancias específicas, los principios y valores son lineamientos más generales que afectan a un espectro más amplio de situaciones. Los principios de una sociedad pueden ser brindar protección a todos sus miembros y en particular a los más vulnerables. Los valores podrían ser libertad y solidaridad. Los principios y valores pueden afectar la construcción de la estructura normativa.

Las normas, principios y valores guían el comportamiento a través de la generación de *obligaciones*. El término *responsabilidades* se utiliza normalmente como hiperónimo de principios, valores y normas. Una persona responsable es fiel a sus principios, respeta los valores del grupo al que pertenece, cumple con las normas que afectan a cada situación específica y se compromete con sus obligaciones.

En el diseño de una organización artificial las responsabilidades pueden quedar establecidas en forma implícita o explícita. En el primer caso las pueden estar embebidas en el *protocolo de interacción*, restringiendo las acciones y las intervenciones que cada agente puede realizar o en la *arquitectura* de cada agente individual, estableciendo qué prioridad asigna a cumplir con sus obligaciones, alcanzar sus metas y brindar colaboración.

Las normas pueden vincularse a la estructura de roles de maneras diferentes. Por un lado una norma implícita en el protocolo regula las interacciones posibles de acuerdo a la jerarquía de roles. Según la relación entre los roles que ocupan dos agentes en un momento determinado uno de ellos *podrá* exigir un comportamiento determinado. Como contrapartida el otro agente *deberá* actuar de acuerdo a lo requerido. Además a cada rol podrán asociarse normas explícitas de manera tal que los agentes que ocupan ese rol deben cumplirlas. Las normas impositivas pueden ser controladas por un agente con un rol específico que monitorea la ocurrencia de los eventos o pueden ser controladas por diferentes agentes.

4 Conocimiento y Acción

Un agente es una entidad computacional capaz de percibir, reaccionar y actuar en el entorno en el que se haya situado. Un agente racional está dotado además de algún nivel de *razonamiento práctico*, que le permite seleccionar una acción adecuada a partir del conocimiento de las metas y del entorno [7].

~~El razonamiento práctico involucra dos procesos fundamentales: *deliberación* y *razonamiento*~~

sobre medios y fines. La deliberación permite decidir **qué** metas van a perseguirse y requiere considerar un espectro de opciones alternativas, seleccionar algunas de ellas y comprometerse a cumplirlas. El razonamiento de medios y fines determina **cómo** van a alcanzarse las metas comprometidas, esto es permite construir un plan que le permita lograr sus intenciones.

El modelo de creencias, deseos e intenciones (BDI) basado en el razonamiento práctico, 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 [8]. El mecanismo deliberativo debe ser de algún modo *oportunist*a en el sentido de que debe permitir seleccionar intenciones en función del beneficio que provoca, pero también considerando que el proceso de planificación pueda llegar a ser exitoso. Para ello es necesario considerar las características del entorno y de las habilidades del agente.

El *estado* de un agente BDI en un momento determinado se define en [9] como una terna $\langle \mathcal{C}, \mathcal{D}, I \rangle$ tal que $\mathcal{C} \subseteq Cre$, $\mathcal{D} \subseteq Des$ e $I \subseteq Int$, donde Cre es el conjunto de todas las creencias posibles, Des es el conjunto de todos los deseos posibles e Int es el conjunto de todas las intenciones posibles. En el trabajo citado la función de generación de opciones mapea los conjuntos de creencias e intenciones previas en un nuevo conjunto de deseos.

Este trabajo presenta una alternativa a esta caracterización, modificando el proceso de revisión de deseos y el mecanismo deliberativo. El algoritmo que sigue presenta una variación para el modelo de comportamiento de un agente BDI propuesto en el trabajo citado antes:

```

 $\mathcal{C} \leftarrow \mathcal{C}_0$ 
 $\mathcal{D} \leftarrow \mathcal{D}_0$ 
while true do
  get next percept  $p$ 
   $\mathcal{C} \leftarrow brf(\mathcal{C}, p)$ 
   $\mathcal{D} \leftarrow drf(p)$ 
   $iota \leftarrow deliber(\mathcal{C}, \mathcal{D})$ 
   $\Pi \leftarrow plan(\mathcal{C}, iota)$ 
  execute( $\Pi, \mathcal{C}$ )
end while

```

Los conjuntos \mathcal{C}_0 e \mathcal{D}_0 corresponden a las creencias y deseos iniciales. La función de revisión de creencias determina el conjunto de creencias a partir de la percepción actual y las creencias previas. En nuestra propuesta los deseos se actualizan también en cada ciclo a partir de la percepción actual. El mecanismo deliberativo genera una intención específica considerando exclusivamente los deseos y las creencias actuales.

La función *plan* retorna una secuencia de acciones a partir del conjunto de creencias y la intención comprometida, considerando por supuesto las habilidades del agente. La operación *execute* no abarca al plan completo, sino que luego de ejecutar cada acción individual se percibe y analiza la necesidad de reconsiderar intenciones o reformular el plan. La reconsideración de intenciones permite tomar ventaja de los cambios y aprovechar nuevas oportunidades.

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 una intención 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 este trabajo 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. Cuando la ejecución de un plan termina exitosamente, la intención ha sido alcanzada y se transforma en una creencia.

A continuación se describe brevemente el formalismo adoptado para representar el conocimiento de un agente BDI y soportar planificación. Más adelante se propone una variación sobre el algoritmo que modela el comportamiento de un agente BDI con compromiso social.

4.1 Representación del conocimiento de un agente 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.

Definition 4.1 [Agente Argumentativo BDI]

Un agente argumentativo BDI puede denotarse como una tupla $\mathbf{a} = \langle \mathcal{C}, \mathcal{D}, \iota, \Gamma \rangle$, donde \mathcal{C} representa el conjunto de creencias del agente, \mathcal{D} sus deseos, ι la intención comprometida y Γ es el conjunto de acciones que es capaz de ejecutar.

En este trabajo las creencias \mathcal{C} están representadas por un Programa en Lógica Rebatible (Φ, Δ) , donde Φ es un conjunto consistente de hechos, y Δ un conjunto de reglas rebatibles. La descripción detallada de DeLP puede obtenerse en [2]. El conjunto de acciones Γ le permiten al agente cambiar el mundo de acuerdo a la propuesta presentada en [1]. 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.

Definition 4.2 [Acción] Una acción A es una terna ordenada $\langle \mathbf{P}, \mathbf{X}, \mathbf{C} \rangle$, donde \mathbf{P} es un conjunto de literales que constituyen las precondiciones de A , \mathbf{X} es un conjunto consistente de literales que representan las consecuencias de ejecutar A , y \mathbf{C} es un conjunto de restricciones de la forma $\text{not } L$, donde L es un literal. Denotaremos a las acciones como:

$$\{X_1, \dots, X_n\} \xleftarrow{A} \{P_1, \dots, P_m\}, \text{not } \{C_1, \dots, C_k\}$$

Notemos que $\text{not } \{C_1, \dots, C_k\}$ representa $\{\text{not } C_1, \dots, \text{not } C_k\}$.

Definition 4.3 [Acción aplicable] Sea $\mathcal{C} = (\Phi, \Delta)$ las creencias de un agente y Γ el conjunto de acciones disponibles. Una acción A en Γ , es aplicable si toda precondición P_i en \mathbf{P} tiene una garantía construida a partir de (Φ, Δ) y toda restricción C_i in \mathbf{C} falla al intentar garantizarse.

En este trabajo las normas restrictivas se modelan como precondiciones para la ejecución de las acciones. Así, una acción no será aplicable si en el estado actual del entorno la ejecución de dicha acción provoca la violación de una norma restrictiva.

Definition 4.4 [Efecto de una acción] Sea $\mathcal{C} = (\Phi, \Delta)$ el conjunto de creencias de un agente Γ el conjunto de acciones disponibles para el agente. Sea A una acción aplicable en Γ . El efecto de ejecutar A es la revisión de Φ por \mathbf{X} , i.e. $\Phi^{*\mathbf{X}} = \Phi^{*\{X_1, \dots, X_n\}}$. La revisión consistirá de remover cualquier literal de Φ que es el complemento de cualquier literal en \mathbf{X} y luego agregar \mathbf{X} al conjunto resultante, esto es, $\Phi^{*\mathbf{X}} = \Phi^{*\{X_1, \dots, X_n\}} = (\Phi - \bar{\mathbf{X}}) \cup \mathbf{X}$, donde $\bar{\mathbf{X}}$ representa el conjunto de los complementos de los miembros de \mathbf{X} .

En [1] se describe la interacción entre acciones y el formalismo de argumentación rebatible. Por un lado la argumentación rebatible se usa para testear las precondiciones y restricciones a través de la noción de garantía. Por otra parte, los agentes ejecutan acciones para cambiar el mundo de manera tal que sea posible obtener nuevas garantías.

4.2 Planificación basada en argumentación rebatible

Cuando un agente \mathbf{a} adopta una intención ι , y ι no está garantizada por (Φ, Δ) el agente \mathbf{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, \mathbf{a} elabora una secuencia de acciones que le permitan establecerlas. La secuencia completa, incluyendo a A , conformará un plan para ι . Cada acción del plan puede modificar las creencias agregando literales que permitan construir nuevos argumentos. Al terminar la ejecución del plan la intención ι se agrega al conjunto de creencias \mathcal{C} .

La selección de acciones no es una tarea trivial y los problemas clásicos de planificación se reflejan en el esquema argumentativo. En una secuencia de acciones $[A_1, A_2]$, A_2 puede ser aplicable de acuerdo al estado inicial Φ , pero no luego de que A_1 se ejecute. En efecto, las consecuencias de A_1 pueden modificar Φ de modo tal que las precondiciones de A_2 no estén garantizadas o una restricción si lo esté.

La ejecución de A_1 puede agregar literales que permitan construir nuevos derrotadores para las precondiciones de A_2 y eliminar literales que permiten construir garantías para las precondiciones de A_2 . El criterio propuesto en [1] para seleccionar acciones es minimizar los cambios no deseados en las precondiciones y las restricciones.

Hemos indicado que en este trabajo un agente sólo adopta intenciones para las cuales puede construir al menos un plan. Esto es, existe al menos una acción entre cuyas poscondiciones está incluido el literal que corresponde a la intención. Sin embargo, es posible que las precondiciones de esta acción no se satisfagan considerando el conjunto de creencias del agente, ni este pueda elaborar un plan que le permita generarlas. En este caso el plan estará incompleto. Si el agente está aislado un plan incompleto no puede ejecutarse.

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.

5 Conocimiento e Interacción

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 individuos 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 este trabajo un sistema multiagente se concibe como organización artificial, formada por conjunto de agentes BDI colaborativos que interactúan de acuerdo a los roles que tienen asignados y restringidos por ciertas normas.

Definition 5.1 [Organización Artificial]

Una organización artificial se define como una tupla $\mathcal{V} = \langle \mathcal{A}, \mathcal{K}, \mathcal{G}, \mathcal{N}, \mathcal{R}, \mathcal{E} \rangle$, donde \mathcal{A} es el conjunto de identificadores de agentes que lo integran, \mathcal{K} el conjunto de creencias globales, \mathcal{G} el conjunto de deseos globales, \mathcal{N} es el conjunto de normas globales, \mathcal{R} es el conjunto de roles que van a asignarse a los miembros del grupo y \mathcal{E} es una estructura jerárquica que especifica las relaciones entre los roles.

Las creencias globales son aquellas que están compartidas por todos los miembros de la organización y para que uno de ellos pueda modificarlo se requiere autorización. Los deseos globales se representan a través de un conjunto de literales que constituyen las metas de la organización.

Durante el proceso perceptivo individual cada agente actualiza sus deseos individuales seleccionando entre los deseos globales aquellos para los cuales es capaz de elaborar planes. Los deseos y las creencias globales se modifican cuando se percibe el mundo y como consecuencia de la actividad de los agentes.

Una vez que un agente transforma un deseo individual en una intención, remueve el literal del conjunto de deseos global, intentando evitar que otro miembro del grupo adopte la misma meta. Evidentemente existe la posibilidad de que dos agentes incorporen el mismo deseo a su conjunto de deseos específico, lo seleccionen como una intención y elaboren un plan para alcanzarla. Si ambos llegan a tener éxito y ejecutan su plan, se habrá desperdiciado esfuerzo porque hubiera bastado con que uno de ellos lo hiciera.

5.1 La estructura de roles y relaciones

En este trabajo los roles existen dentro de una organización, independientemente de que estén instanciados o no. En un momento dado, puede producirse la situación de que ninguno de los miembros de un grupo pueda asignarse a un rol determinado y no por eso deja de existir dicho rol.

Definition 5.2 [Roles] *El conjunto de roles se define como $\mathcal{R} = \{\rho_1, \rho_2, \dots, \rho_n\}$. Cada rol ρ_i dentro del conjunto \mathcal{R} contiene al conjunto de aptitudes requeridas, donde cada aptitud va a estar representada a través de un literal.*

El rol de cada agente cambia dinámicamente dependiendo de las competencias que va adquiriendo. Las competencias están integradas por sus creencias actuales y todos aquellos planes que ha elaborado y ejecutado previamente. Un rol puede ser asignado a un agente si el conjunto de aptitudes que conforman el rol forma parte de sus competencias. Esto es, este está incluido en sus creencias o existen planes ya elaborados que le permiten cumplir con las funciones inherentes al rol.

Definition 5.3 [Relaciones entre Roles] *La estructura de relaciones entre los roles se define como $\mathcal{E} \subseteq \mathcal{R} \times \mathcal{R}$ y constituye un orden parcial satisfaciendo las propiedades reflexiva, anti-simétrica y transitiva. Además, supondremos que \mathcal{E} verifica la siguiente propiedad: si $\rho_1, \rho_2 \in \mathcal{R}$ entonces $\exists \rho_3 \in \mathcal{R}$ tal que $\rho_1 \mathcal{E} \rho_3$ y $\rho_2 \mathcal{E} \rho_3$, es decir, siempre existe el supremo entre dos roles.*

En el momento en que se inicia un diálogo colaborativo se asigna a cada participante el rol que corresponde y así queda determinada también la relación que unirá a los interlocutores durante todo el proceso. Esta relación va a restringir el tipo de locuciones que puede realizar cada uno.

Bajo esta concepción, en un grupo particular cada agente tendrá en un momento determinado un rol específico y entre cada par de agentes puede existir una relación establecida por sus roles. Esta relación condicionará la evolución del diálogo entre ambos.

5.2 La estructura normativa

En este trabajo, los valores y principios van a estar embebidos en el comportamiento de los agentes. La representación de normas restrictivas e impositivas es en cambio explícita, algunas en forma global afectando a toda la organización, otras más específicas en el modelo cognitivo de cada individuo.

Los *valores* fundamentales serán la sinceridad, honestidad y cooperación. La honestidad implica que los agentes van a respetar las normas de la organización y los roles asignados. Es decir, el comportamiento está regulado por las normas y los roles dentro de la estructura. Así, un agente honesto cumple con las normas salvo en aquellas situaciones en las cuales un superior lo autoriza a no hacerlo. Aun así solo algunas restricciones van a relajarse, las normas globales no van a ser violadas, lo cual implica que eventualmente el sistema puede bloquearse requiriendo intervención externa. La honestidad también provoca que ningún miembro de la organización modifique el conocimiento compartido sin autorización.

La sinceridad implica que las acciones se seleccionan a partir de las metas, pero se consideran solo aquellas cuyas precondiciones se satisfacen en el estado actual del mundo. Por lo tanto los agentes no actúan asumiendo creencias que no tienen.

La cooperación implica que todos los individuos estarán dispuestos a colaborar cuando reciban un requerimiento, intentando balancear sus metas propias con su intención de colaborar. El nivel de compromiso depende del rol.

El *principio* será la coordinación de acciones individuales a través de la interacción, regulada por la estructura de roles. Aunque los agentes están dotados de diferentes repertorios de acciones, todos responden a un mismo modelo de comportamiento colaborativo.

Las normas globales están asignadas a la organización en su conjunto, expresan condiciones que deben satisfacerse en todo momento y actúan simultáneamente como restrictivas e impositivas. En principio el conjunto de normas globales está incluido en el conjunto de creencias globales. Las normas son fijas, pero las creencias cambian como consecuencia de las acciones de los agentes y de los cambios externos en el entorno.

Las normas globales son restrictivas porque limitan el comportamiento de los agentes. Ningún agente puede ejecutar en condiciones normales una acción que viole una norma global, provocando que el conjunto de normas deje de estar incluido en el conjunto de creencias. Así, las acciones de un agente cuyas poscondiciones entran en conflicto con las normas globales, constituyen aptitudes que sólo se utilizarán en circunstancias excepcionales.

Las normas globales son además imposiciones porque si un cambio externo en el entorno afecta al conjunto de creencias, de manera tal que el conjunto de normas deja de estar incluido en él, los agentes quedan obligados a seleccionar esta norma como un deseo con máxima prioridad y elaborar un plan para volver a una situación normal.

5.3 El conocimiento de un agente BDI con compromiso social

Hemos considerado hasta el momento que el conocimiento de agente BDI incluye las creencias acerca del dominio en el que actúa, los deseos, su intención comprometida y el conjunto de acciones que conforman su capacidad efectora. Un agente con compromiso social respeta las normas globales y también puede tener asignadas algunas normas impositivas específicas que consideren sus habilidades:

Definition 5.4 [Norma Impositiva] *Una norma impositiva \mathcal{I} es un par ordenado $\langle H, S \rangle$, donde H es el conjunto de literales que caracterizan al estado en el cual corresponde disparar a \mathcal{I} , y S es un plan, esto es la secuencia de acciones que impone ejecutar la norma.*

Las normas impositivas disparan la ejecución de una secuencia de acciones cuando H está contenido en las creencias actuales del agente. Notemos que la especificación de las normas impositivas es fija, lo que cambian son las creencias del agente como producto de la percepción de cambios externos en el entorno. Las normas impositivas caracterizan a situaciones no deseadas que en general deben superarse tan pronto como sea posible. De modo que la secuencia debería contener pocas acciones y al menos la primera no debería tener precondiciones.

La situación más simple es que el primer elemento del par sea un conjunto con un único literal y el segundo elemento del par sea una acción simple. En este caso la norma prescribe una acción específica ante un hecho determinado. Por ejemplo, si se detecta humo, activar la alarma.

Una alternativa a este modelo es que el segundo elemento del par no sea directamente una secuencia de acciones sino un conjunto de literales que caracterizan el estado que debería alcanzarse. En este caso el agente es el responsable de elaborar un plan que le permita alcanzar este estado. Para el ejemplo anterior, si se detecta humo se debe alcanzar un estado en el cual la alarma esté activa, pero no se indica si esto se logra con una acción simple o con una secuencia de acciones.

Extendemos ahora la definición propuesta antes para incluir *normas impositivas específicas*, *requerimientos de colaboración* y *competencias*.

Definition 5.5 [Agente Argumentativo BDI con capacidad Social]

Definimos el conocimiento de un agente argumentativo BDI con capacidad social, como una tupla $\mathcal{K} = \langle \mathcal{C}, \mathcal{D}, g, \Gamma, \mathcal{P}, \mathcal{I}_a \rangle$, donde \mathcal{C} representa las creencias del agente, \mathcal{D} sus deseos, g es una intención, Γ el conjunto de acciones que es capaz de ejecutar, \mathcal{P} el conjunto de competencias adquiridas y \mathcal{I}_a el conjunto de imposiciones específicas.

El conjunto de deseos está formado ahora por dos subconjuntos \mathcal{F} y \mathcal{Q} , representando respectivamente las metas percibidas en los deseos globales y los requerimientos de colaboración de otros agentes.

El conjunto \mathcal{P} está conformado por los planes que el agente ha construido y ejecutado previamente y aumentan su conocimiento. De este modo, cada plan ejecutado tiene una doble funcionalidad, por un lado le permite al agente alcanzar una meta, pero además aumenta su experiencia.

Cada nueva competencia que se agrega al conjunto puede pensarse como una acción compuesta y esto resulta particularmente útil cuando un agente enfrenta las mismas metas reiteradamente. La existencia de un plan previo no implica necesariamente que siempre se evite construir otro nuevo, porque en cada oportunidad las creencias pueden ser diferentes y entonces las precondiciones de un plan conocido pueden satisfacerse o no.

Es importante notar que las acciones involucradas en un plan ejecutado pueden tener como precondiciones a creencias que fueron obtenidas gracias a solicitudes de colaboración, que probablemente provocaron procesos de negociación entre dos o más agentes. La representación explícita de las competencias podría distinguir las acciones que el agente pudo realizar por sí mismo, de aquellas que provocaron la interacción con otros agentes. No consideramos significativo incluir dentro de las competencias aquellos planes que resultaron infructuosos y no llegaron a ejecutarse.

El conjunto \mathcal{I}_a de *normas impositivas específicas* indica lo que se **debe** hacer bajo ciertas circunstancias. Por ejemplo, si hay fuego cortar la luz y activar la alarma.

Las *normas restrictivas* indican lo que **no** se puede hacer, expresan prohibiciones. Como vimos algunas normas restrictivas se representan a través de un conjunto global que no debería ser modificado en circunstancia normales por ninguna acción. Cada agente respeta las normas globales. En circunstancias excepcionales, un agente con un rol superior puede autorizar a otro a violar una de estas restricciones. Otras restricciones son *específicas* de un agente en particular y se representan asignándole a cada acción de ese individuo un conjunto de precondiciones que representan restricciones para su ejecución. Notemos que si no existe un mecanismo para factorizar acciones, una misma restricción va a aparecer en cada agente que tenga la habilidad que corresponde a la acción.

Una norma global se transforma en imposición cuando un cambio externo provoca una modificación en las creencias inconsistente con la norma. Por ejemplo si una norma global establece que el piso tiene que estar seco y en determinado momento las creencias se modifican indicando que el piso no está seco, inmediatamente la meta *piso seco* se transforma en un deseo global con máxima prioridad. El agente que se comprometa a esta intención deberá elaborar un plan para alcanzarla.

Notemos que existe una diferencia importante entre las normas globales que se transforman en imposiciones y las normas impositivas específicas de un agente. Para estas últimas el individuo tiene un plan, *sabe* como actuar en determinadas circunstancias. En el caso de que una norma global genere una imposición, el individuo que la adopta tiene que buscar un modo de restaurar el estado en el cual se cumple con la restricción.

5.4 El comportamiento de un agente BDI con compromiso social

La participación en un contexto social modifica el comportamiento de cada individuo de una organización. Cada agente puede solicitar colaboración para alcanzar sus metas, pero también está comprometido a brindar ayuda cuando recibe un requerimiento.

La introducción de normas y roles afecta al comportamiento de cada miembro de la organización e implica modificar el algoritmo que modela este comportamiento. La percepción de deseos y la selección de intenciones se realiza considerando también las normas globales y específicas. Además aumenta la complejidad de la interacción porque debe considerar las restricciones impuestas por los roles.

Aunque en este modelo todos los agentes respetan un mismo algoritmo general, el proceso de selección de una intención y la planificación va a variar según el rol.

```

 $\mathcal{C} \leftarrow \mathcal{C}_0$ 
 $\mathcal{D} \leftarrow \mathcal{D}_0$ 
while true do
  Determinar el rol del agente
  Obtener una percepción  $p$ 
   $\mathcal{C} \leftarrow brf(\mathcal{C}, p)$  {revisión de creencias}
   $\mathcal{D} \leftarrow grf(\mathcal{D}, p)$  {revisión de deseos y requerimientos}
   $iota \leftarrow deliber(\mathcal{C}, \mathcal{D})$ 
   $P \leftarrow plan(\Pi, iota)$ 
  if  $iota \in \mathcal{Q}$  then
     $OK \leftarrow dialogue()$ 
  else
    if  $incomplete(\Pi)$  then
       $OK \leftarrow request()$ 
    end if
  end if
  if  $OK$  then
     $execute(\Pi)$ 
  end if
end while

```

Los roles están asociados a aptitudes y en cada ciclo se asigna al agente el máximo rol dentro de la estructura de roles considerando que sus competencias puedan cubrir las aptitudes requeridas por el rol. Una vez que el rol ha sido asignado el agente percibe el entorno para actualizar creencias, deseos y requerimientos de colaboración.

Un agente solo percibe como metas literales que corresponden a las poscondiciones de sus acciones. Si una norma global se activa, un agente puede adoptarla como un deseo con máxima prioridad, en tanto su repertorio de acciones resulte adecuado para cumplir con ella. Cada individuo puede recibir requerimientos de colaboración aun cuando no se correspondan con sus capacidades, en este caso se declarará incompetente para atender la solicitud.

En este trabajo, la incorporación de roles provoca que los requerimiento de colaboración de los superiores tengan prioridad sobre los deseos percibidos en el mundo, dependiendo del nivel de exigencia del requerimiento. Notemos que en nuestro modelo el agente no recibe la orden de ejecutar una acción concreta, sino de llegar a un estado específico. Es el agente quien decide cómo hacerlo.

La elaboración de un plan es ahora más compleja porque es necesario considerar los conflictos que pueden producirse con las normas restrictivas. Esto es, si el agente está comprometido a respetar una norma, tendrá en principio inhibida cualquier acción cuya poscondición la viole.

Esta restricción tiene vigencia salvo en algunos casos excepcionales. Por ejemplo, si un agente elabora un plan para cumplir con una norma impositiva, puede violar una restrictiva, si no existe un plan alternativo que pueda realizarse sin violar las normas. Del mismo modo puede hacerlo, si recibe un requerimiento con un nivel alto de exigencia de un agente con un rol mayor que el de él mismo.

Como antes, la estrategia para elaborar un plan específico escapa a los alcances de este trabajo. Una alternativa es en primera instancia intentar buscar entre los planes elaborados previamente alguno cuyas precondiciones se satisfagan considerando las creencias actuales. Si existe podría seleccionarse inmediatamente. Si no existe un plan cuyas precondiciones se satisfagan a partir de las creencias individuales, es posible elaborar uno nuevo o completar el anterior. En ambos casos, el agente puede requerir colaboración. El mecanismo de selección podría aumentarse asignando un costo a cada plan, de modo que entre varios planes en condiciones similares, se considere el menor costo.

Una vez que el plan ha sido elaborado el agente puede interactuar con otros miembros de la organización si no dispone de todo el conocimiento necesario como para ejecutarlo como se propone en [5]. La interacción se produce tanto cuando el plan ha sido elaborado para atender un requerimiento de colaboración, como cuando se trata de un deseo propio pero se necesita colaboración. En el primer caso el agente ya está enfrascado en un diálogo, de modo que continúa interactuando en este caso para realizar una contrapropuesta que le permita obtener el conocimiento necesario para ejecutar su plan. En el segundo caso el agente inicia un diálogo solicitando colaboración.

6 Conclusiones y trabajo futuro

Uno de los mayores desafíos en el desarrollo de un sistema multiagente es lograr un comportamiento previsible ante situaciones imprevisibles en el entorno. Las normas de una organización permiten controlar de alguna manera la autonomía de los individuos en circunstancias convencionales y evitar desviaciones. La posibilidad de violar las normas en casos de emergencia brinda la flexibilidad necesaria para adaptarse a cambios bruscos en el entorno. En este trabajo las normas solo son vulneradas considerando la estructura de roles y el énfasis en el requerimiento, de modo que el comportamiento sigue siendo controlado.

El lenguaje de interacción debe permitir expresar el nivel de exigencia en un requerimiento ofreciendo un repertorio de performativas adecuado. La estructura de roles restringe las locuciones posibles, de modo que en cada diálogo el agente con un rol superior podrá usar un conjunto más amplio de performativas. El protocolo de interacción que modela las posibles formas en las que puede evolucionar un diálogo colaborativo, está entonces fuertemente ligado a la relación entre los roles. Nuestro trabajo futuro se concentrará en reformular el lenguaje de interacción y el protocolo de interacción descrito en [4] considerando la introducción de roles y normas presentada en este trabajo.

Una extensión posible sobre este trabajo es asignar normas específicas a los roles. Así la asignación de un rol se realiza considerando las aptitudes requeridas y determina las responsabilidades inherentes.

References

- [1] *Actions, Planning, and Defeasible Reasoning*. Simari, G.R., García, A.J., Capobianco, M.. In Proc. of the 10th Int. Workshop on Non-Monotonic Reasoning, Canada, pp. 377-384, 2004.
- [2] *Defeasible Logic Programming: An Argumentative Approach*. A.J., Simari, G.R., Theory and Practice of Logic Programming, 4(1), pp 95-138, 2004.
- [3] Capobianco, M., Chesñevar, C., Simari, G. *Argumentation and the Dynamics of Warranted Beliefs in Changing Environments.*, Journal of Autonomous Agents and Multi-Agent Systems, 11, 127-151,

- [4] Rueda S. V. , Martinez M.V. *Interaction among BDI Argumentative Agents: A dialogue games approach*. XI Congreso Argentino de Ciencias de la Computación (CACIC), pp. 955-966.
- [5] Rueda S. V. , García A., Simari G. *Collaboration through Deliberative Dialogues*. JCST Vol. 7. No. 2., pp. 125–133, 2007
- [6] Fagin, R., Halpern J., Moses, Y., Vardi, M. *Reasoning About Knowledge*. The MIT Press, Cambridge, MA 1995.
- [7] Bratman, M.E., Israel D.J., Pollack, M.E. *Plans and Resource-Bounded Practical Reasoning*. Computational Intelligence 4:349-355, 1988.
- [8] Rao, A. S., and Georgeff M. P. *BDI Agents: From Theory to Practice*. Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95), pp. 312-319, USA, 1995.
- [9] Wooldridge, M. *Intelligent Agents*, The MIT Press, 1999.
- [10] Huhns, M., Stephens, L. *Multiagent Systems and Societies of Agents*, The MIT Press, 1999.
- [11] Partsakoulakis I., Vouros G. *Roles in Mas: Managing the Complexity of Tasks and Environments*. In *An Application Science for Multi-Agent Systems*, T. Wagner Ed., Springer-Verlag, pp. 133-154, 2004.
- [12] Vázquez -Salceda J., Aldewereld H., Dignum F. *Norms in Multiagent Systems: Some Implementation Guidelines*. Second European Workshop on Multi-Agent Systems, pp. 737–748, Barcelona, 2004.
- [13] Köhler, M. , Moldt, D., and Rölke, H. *A Discussion of Social Norms with Respect to the Micro-Macro Link*. In G. Lindemann, D. Moldt, and M. Paolucci, Eds., 2nd Int. Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA'03), 2003. In connection with the Ninth ICAIL.
- [14] Lopez y Lopez, F., Luck, M., and, d'Inverno, M., *A Framework for Norm-Based Inter-Agent Dependence*. In 3rd Mexican International Conference on Computer Science, pp. 31-40, 2001.