

# Dinámica de Conocimiento en Sistemas Multi-Agentes: Plausibilidad, Revisión de Creencias y Retransmisión de Información

**Luciano H. Tamargo**

**Directores: Alejandro J. García y Marcelo A. Falappa**

Fecha de exposición: 17/03/2011

Laboratorio de Investigación y Desarrollo en Inteligencia Artificial  
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur,  
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)  
Av. Alem 1253, (B8000CPB) Bahía Blanca, Argentina  
Tel: (0291) 459-5135 / Fax: (0291) 459-5136  
e-mail: lt@cs.uns.edu.ar

## Resumen

Esta tesis propone formalismos para modelar la dinámica de conocimiento en bases de creencias de agentes en un sistema multi-agente, los cuales adquieren información de sus pares. Dicha dinámica es modelada por la teoría de cambio de creencias, la cuál busca mostrar como quedan constituidas las creencias de un agente después de recibir cierta información externa. Nuestra propuesta se enfoca en revisión de creencias con múltiples fuentes (*Multi-Source Belief Revision*), ya que los agentes pueden recibir información a través de múltiples informantes, los cuales son agentes independientes con sus propios intereses y podrían no ser completamente fiables. Por este motivo, en esta tesis se propone la utilización de un orden parcial de informantes que representa la credibilidad de los mismos para el agente que se está modelando. El orden parcial de credibilidad entre agentes no necesita permanecer estático ya que el agente puede elegir actualizar su relación de orden parcial para reflejar una nueva percepción de la credibilidad de un informante. Es por esto que, otro de los objetivos de esta tesis es el estudio y desarrollo de técnicas y formalismos para la actualización del grado de credibilidad que se le asigna a un agente por interactuar en el marco de un sistema multi-agente. Por lo tanto, se proponen operadores de cambio, tanto para revisar las creencias de un agente, como para revisar la credibilidad que un agente tiene sobre los demás agentes del sistema. De esta manera, la investigación desarrollada también se enfoca en la caracterización y desarrollo de operadores de cambio, que permitan modelar la dinámica de la confianza y reputación de agentes en un sistema. En resumen, se propone combinar formalismos de revisión de creencias y actualización de conocimiento con técnicas de mantenimiento de confianza y reputación de agentes en un ambiente distribuido, para representar la dinámica de órdenes parciales de credibilidad.

## 1. Introducción

En el contexto de sistemas multi-agente un agente puede a menudo recibir información a través de informantes. Estos informantes son agentes independientes quienes tienen sus propios

intereses y, por lo tanto, podrían no ser completamente fiables. Es natural para un agente creerle más a un informante que a otro. Para modelar este tipo de escenarios, esta tesis propone un modelo completo de cambio basado en informantes, pa lo cual utiliza un orden parcial de informantes que representa la credibilidad de los mismos.

Considere, por ejemplo, el siguiente escenario simplificado. Un agente  $A_v$  quiere viajar a un poblado que se encuentra en una montaña y tiene que tomar una decisión: realizar el viaje (que para simplificar lo representamos con  $\alpha$ ) o no realizar el viaje ( $\neg\alpha$ ). El agente  $A_v$  sabe, por el agente de la oficina de turismo ( $A_t$ ), que si nieva ( $s$ ) el camino al poblado no estará abierto ( $s \rightarrow \neg\alpha$ ).  $A_v$  también sabe que si el camino no está abierto, entonces no realizará el viaje ( $\neg\alpha \rightarrow \neg\alpha$ ) y si el camino está abierto, sí realizará el viaje ( $\alpha \rightarrow \alpha$ ). El agente  $A_v$  también sabe que puede obtener información desde otras fuentes: algún agente  $A_c$  que viene por el camino desde el poblado, un agente  $A_g$  de la estación de servicio, un agente  $A_r$  de un restaurante en el camino, o el informe climático de la radio ( $A_w$ ). Para  $A_v$ ,  $A_c$  es más creíble que  $A_w$ , que este último es más creíble que  $A_t$ , que a su vez es más creíble que  $A_g$  y que  $A_g$  es más creíble que  $A_r$  (i.e.,  $A_c > A_w > A_t > A_g > A_r$ ). Luego, supongamos que en la estación de servicio  $A_g$  informa a  $A_v$  que el camino no está abierto. El agente  $A_v$  debería poder revisar sus creencias para tomar en cuenta esta nueva pieza de información.

Existen diferentes modelos de cambio en la teoría de cambio de creencias. Uno de sus principales referentes, el modelo AGM [2], distingue claramente tres operaciones de cambio: expansiones, contracciones y revisiones. La operación de expansión es la más simple de caracterizar desde el punto de vista lógico pues consiste solamente en la adición de nuevas creencias. Sin embargo, las contracciones y revisiones no tienen esta propiedad. Las contracciones involucran la eliminación de creencias mientras que las revisiones involucran tanto la adición como la eliminación de creencias. Siguiendo con el ejemplo anterior el agente  $A_v$  puede expandir su conocimiento agregando la información ' $\neg\alpha$ ' (camino no abierto) ya que esta pieza de información no contradice su conocimiento. Si luego en el restaurant  $A_r$  informa a  $A_v$  que el camino está abierto, como  $A_r$  es menos creíble que  $A_g$ , luego  $A_v$  debería rechazar la nueva información. En cambio, si  $A_c$  informa a  $A_v$  que el camino está abierto, como  $A_c$  es más creíble que  $A_g$ , luego  $A_v$  debería cambiar sus creencias. Para poder modelar este comportamiento, en esta tesis se propone un modelo completo de cambio (expansión, contracción, revisión priorizada y revisión no priorizada) el cual utiliza el orden de credibilidad entre agentes para decidir que información prevalece.

Por otro lado, el orden parcial de credibilidad entre agentes no necesita permanecer estático, ya que el agente puede decidir actualizar su relación de orden parcial para reflejar una nueva percepción de la credibilidad de un informante. Continuando con el ejemplo anterior donde el agente  $A_v$  considera que el agente  $A_c$ , que viene del camino que lleva al poblado, es más creíble que el agente  $A_w$  del servicio climático de la radio. Supongamos ahora, que un agente le informa a  $A_v$  que el servicio climático de la radio está informando desde el camino que lleva al poblado y que, por lo tanto, según el informante, el agente  $A_w$  es más creíble que el agente  $A_c$ . Esto es, un agente le informa a  $A_v$  la siguiente relación de credibilidad  $A_c < A_w$ . En este caso, el agente  $A_v$  debería revisar su orden de credibilidad para poder reflejar esta nueva información.

En la tesis doctoral desarrollada se proponen operadores de cambio, no sólo sobre creencias (como mencionamos antes), sino que también sobre el orden parcial de credibilidad. Con estos operadores se puede modelar la dinámica de sistemas de confianza y reputación. Por lo tanto, la investigación también se enfocó en la caracterización y desarrollo de operadores de cambio, que permitan modelar la dinámica de la confianza y reputación de agentes en un sistema. De esta manera, los agentes pueden actualizar la relación de orden con la cual se representa la reputación de sus pares. Así, estos operadores son usados para alterar dinámicamente la estructura de la credibilidad de los informantes a fin de reflejar una nueva percepción de la credibilidad de un informante o la llegada de un nuevo agente al sistema.

En resumen, en la tesis desarrollada se combinan, formalismos de revisión de creencias y ac-

tualización de conocimiento con técnicas de mantenimiento de confianza y reputación de agentes en un ambiente distribuido, para representar la dinámica de órdenes parciales de credibilidad.

## 2. Contribuciones y resultados de la tesis

La tesis doctoral desarrollada presenta las siguientes contribuciones originales:

### **Modelo de cambio completo para revisión de creencias con múltiples fuentes.**

Se desarrolló un modelo epistémico para revisión de creencias con múltiples fuentes (*multi-source belief revision* - MSBR) para el cual se propuso una manera racional de comparación de creencias usando un orden de credibilidad entre agentes. Parte de los resultados de esta contribución están incluidos en los artículos *Modeling knowledge dynamics in multi-agent systems based on informants* publicado en la revista *Knowledge Engineering Review (KER)* [20], *A Belief Revision Approach to Inconsistency Handling in Multi-Agent Systems* publicado en *The IJCAI-09 Workshop on Nonmonotonic Reasoning, Action and Change (NRAC)* [19], y *Consistency Maintenance of Plausible Belief Bases Based on Agents Credibility* publicado en *12th International Workshop on Non-Monotonic Reasoning (NMR)* [18]. En este formalismo se incluye una función de plausibilidad la cual es utilizada en la definición de un criterio para comparar las creencias. Luego, definimos diferentes operadores que describen un modelo de cambio completo basado en informantes: expansión, contracción, revisión priorizada y revisión no-priorizada. Estos operadores pueden ser vistos como habilidades agregadas a los agentes para mejorar el razonamiento colectivo de un sistema multi-agente. Para cada uno de ellos damos una definición en forma constructiva y mostramos una caracterización axiomática a través de teoremas de representación. Además, formalmente mostramos que nuestro formalismo cumple con algunos principios reconocidos en la literatura: mínimo cambio, mantenimiento de consistencia y no-priorización. De esta manera, definimos un modelo completo de cambio donde introducimos revisión priorizada y no-priorizada. Algunos detalles de esta contribución serán detallados brevemente en la sección 3.

### **Actualización de órdenes parciales de credibilidad entre agentes.**

En la tesis se propone dos maneras de realizar cambios en el orden de credibilidad de un agente siguiendo algunos conceptos de los mecanismos de reputación y confianza. Uno de los formalismos desarrollados consiste en la descripción de un modelo completo de cambio para órdenes parciales de credibilidad de informantes, el cual fue publicado en [17]. El otro formalismo sugiere una revisión del orden de credibilidad basado en una creencia y su meta-información asociada. Finalmente, presentamos una técnica que brinda la posibilidad de revisar la base de creencias de un agente con el efecto colateral de modificar el orden de credibilidad cuando es necesario.

### **Criterios de retransmisión de información.**

Finalmente, se propusieron algunos criterios que mejoran las habilidades de interacción de un agente en un sistema multi-agente, combinándolos con sus habilidades de razonamiento para permitir la propagación de información creíble. Se investigó como un agente puede retransmitir información a otros agentes, la cual podría haber sido adquirida desde otros agentes. En particular, se estudió como elegir racionalmente la meta-información a ser retransmitida. Hemos presentado diferentes maneras de elegir esta meta-información y dimos una categorización de posibles enfoques, donde discutimos ventajas y desventajas. Esta discusión nos llevó a la definición de un criterio que usa una función de plausibilidad que determina la plausibilidad de una creencia basada en todas sus pruebas de acuerdo a una base. Todos estos aspectos analizados fueron publicados en el artículo *Forwarding Credible Information in Multi-agent Systems*, publicado en *the 3rd International Conference on Knowledge Science, Engineering and Management (KSEM 2009)* [11].

### 3. Dinámica del conocimiento basado en informantes

En la tesis se ha propuesto un modelo epistémico que considera que, cuando los agentes interactúan, incorporan la información recibida en sus bases de conocimiento asociando una sentencia con un identificador de agente, el cual determina una fuente o informante. Esto es, los agentes almacenan en su base de creencias objetos de información de la forma  $(c, i)$  donde  $c$  es una creencia e  $i$  un identificador de agente. Consideramos nuevamente el escenario introducido anteriormente, en el cual el agente  $A_v$  quiere viajar a un poblado en la montaña. Habíamos mencionado que  $A_v$  sabe, por el agente de la oficina de turismo ( $A_t$ ), que si nieva ( $s$ ) el camino al poblado no estará abierto ( $s \rightarrow \neg o$ ). Además,  $A_v$ , reconoce él mismo, que si el camino no está abierto, entonces no realizará el viaje ( $\neg o \rightarrow \neg \alpha$ ) y si el camino está abierto, sí realizará el viaje ( $o \rightarrow \alpha$ ). De esta manera, según el modelo epistémico sugerido en la tesis, el agente  $A_v$  cuenta con la siguiente base de creencias  $K_{A_v} = \{(s \rightarrow \neg o, A_t), (\neg o \rightarrow \neg \alpha, A_v), (o \rightarrow \alpha, A_v)\}$ .

El identificador en un objeto de información puede ser usado para evaluar la veracidad de la información. En la tesis, hemos sugerido que cada agente posea una función de valoración (*assessment*) la cual es usada para representar la credibilidad que cada agente asigna a otros agentes. Para definir esta valoración, usamos un conjunto de etiquetas de credibilidad  $\mathcal{C} = \{c_1, \dots, c_k\}$  comunes a todos los agentes, ordenadas en forma total y estricta. Un *assessment*, para un agente, es una función que asigna un valor de credibilidad a cada agente del sistema. Si bien el conjunto de etiquetas de credibilidad es el mismo para todos los agentes del sistema, cada agente tiene su propio *assessment* y diferentes agentes pueden tener diferentes *assessments*. De esta manera, cada agente puede definir un orden de credibilidad sobre los agentes del sistema basado en su propio *assessment*. En el caso de nuestro ejemplo, como  $A_v$  considera más creíble al agente  $A_c$  (que viene del poblado) que el agente  $A_g$  (de una estación de combustible), entonces la función de *assessment*  $c_{A_v}$  del agente  $A_v$  retornará  $c_{A_v}(A_g) = c_3$  y  $c_{A_v}(A_c) = c_6$  donde la etiqueta  $c_3$  es menor a la  $c_6$ . Luego, basado en el *assessment*, cada agente posee un orden de credibilidad entre agentes. En el caso de  $A_v$ , siguiendo el ejemplo planteado a lo largo de este resumen, el orden de credibilidad de acuerdo a  $A_v$  es  $A_v <_{c_{A_v}}^{A_v} A_r <_{c_{A_v}}^{A_v} A_g <_{c_{A_v}}^{A_v} A_t <_{c_{A_v}}^{A_v} A_w <_{c_{A_v}}^{A_v} A_c$ .

Es importante notar que como ventaja adicional de este formalismo, el *assessment* de un agente podría ser reemplazado o modificado de manera modular sin cambiar su base de creencias y sin afectar los *assessments* de otros agentes. Sin embargo, el cambio en el *assessment* de un agente  $A_i$  podría generar un cambio en el orden de credibilidad del agente  $A_i$  y, por lo tanto, podría generar un cambio implícito en el orden de las creencias de la base  $K_{A_i}$ .

Con esta forma de representar la información, hemos mostrado que en una base de creencias la misma sentencia puede estar en varios objetos de información (con diferentes identificadores de agentes). Esto tiene como ventaja adicional que, si el *assessment* de un agente es cambiado y la credibilidad de un agente particular es incrementada, entonces todas las sentencias asociadas a este agente automáticamente tienen más credibilidad.

Luego, basado en el orden de credibilidad entre agentes, hemos definido un criterio de comparación entre sentencias de una base. Para ello, hemos introducido una función de plausibilidad tal que, dada una sentencia, retorna un identificador de agente que representa su plausibilidad con respecto al *assessment* del agente propietario de la base. Luego, basado en esta función, introducimos un criterio de comparación entre sentencias de una base. Note que, dos agentes con la misma base de creencias, pero diferentes órdenes de credibilidad, producen diferentes órdenes de plausibilidad entre sentencias.

Note, además, que en el formalismo propuesto, la plausibilidad de las sentencias no es almacenada explícitamente, sino que es obtenida por medio de la *función de plausibilidad*. Dada una sentencia  $\alpha$ , su plausibilidad depende de sus pruebas ( $\alpha$ -*kernels*). Por lo tanto, si una de las sentencias de estas pruebas cambia, entonces la plausibilidad de  $\alpha$  podría cambiar. De esta manera, si el orden de credibilidad es reemplazado, entonces la plausibilidad de la sentencia puede cambiar sin cambiar la base de creencias. Esto no sucedería si hubiésemos decidido almacenar la

plausibilidad de una creencia como un valor fijo asociado como se puede ver en [6].

A diferencia de nuestro trabajo, hay autores que representan en forma numérica la información adicional [4]. Esto podría representar una limitación, ya que de este modo no se podrían representar órdenes parciales entre las creencias; sin embargo, como mostramos en la tesis y resumiremos en la siguiente sección, siguiendo nuestro modelo es posible hacerlo.

En base a este modelo epistémico y las *kernel contractions* de Hansson [8], hemos definido un modelo completo de cambio para MSBR. En este modelo fue incluido un operador de expansión que, a diferencia de la expansión propuesta en [9], considera objetos de información en vez de sólo sentencias. Luego, fue incluido una especialización de las *kernel contractions*. Esto es, las *kernel contractions* usan funciones de incisión, las cuales cortan en los subconjuntos mínimos que derivan la información a ser quitada. En la definición de *función de incisión* del trabajo de Hansson [8] no se especifica como la función elige las sentencias que serán descartadas de cada subconjunto. En la tesis, esto fue resuelto con la plausibilidad de las sentencias que hemos definido. De esta manera, la función de incisión elige los objetos de información menos creíbles de cada subconjunto minimal.

Finalmente, con este nuevo modelo epistémico para MSBR, y aprovechando la ventaja de contar con una forma de comparar las creencias, hemos propuesto un operador de revisión no priorizado, el cual usa la plausibilidad de las creencias tanto para evaluar la aceptación de la entrada, como para determinar qué creencias serán quitadas si la revisión tiene efecto.

Para ejemplificar los operadores definidos, considere la base de creencias

$$K_{A_v} = \{(s \rightarrow \neg o, A_t), (\neg o \rightarrow \neg \alpha, A_v), (o \rightarrow \alpha, A_v)\}$$

y el orden de credibilidad  $A_v <_{co}^{A_v} A_r <_{co}^{A_v} A_g <_{co}^{A_v} A_t <_{co}^{A_v} A_w <_{co}^{A_v} A_c$ . Suponga que  $A_v$  recibe el objeto de información  $(\neg o, A_g)$ . Como esta información no contradice el conocimiento de  $A_v$ , entonces la base de creencias de  $A_v$  se ‘*expande*’, y la base resultante es:

$$K_{A_v} = \{(s \rightarrow \neg o, A_t), (\neg o \rightarrow \neg \alpha, A_v), (o \rightarrow \alpha, A_v), (\neg o, A_g)\}$$

Suponga ahora que  $A_v$  recibe el objeto de información  $(o, A_r)$ , como esto contradice el conocimiento de  $A_v$  y  $A_r$  es menos creíble que  $A_g$ , utilizando el operador de revisión no priorizado definido en esta tesis,  $A_v$  rechaza la nueva información. En cambio, suponga que  $A_v$  recibe  $(o, A_r)$ , como  $A_c$  es más creíble que  $A_g$ ,  $A_v$  cambia sus creencias por medio del operador de revisión no priorizado, y la base de creencias resultante es:

$$K_{A_v} = \{(s \rightarrow \neg o, A_t), (\neg o \rightarrow \neg \alpha, A_v), (o \rightarrow \alpha, A_v), (o, A_c)\}$$

## 4. Dinámica sobre el orden de credibilidad

En la tesis hemos propuesto dos maneras de realizar cambios sobre el orden de credibilidad de un agente basado en la teoría de sistemas computacionales de confianza y reputación. En [15], los autores proponen un conjunto de aspectos especiales para clasificar los modelos de confianza y reputación. Estos aspectos los eligieron tomando en cuenta las características de los modelos computacionales existentes hasta ese momento. El primer aspecto que tuvimos en cuenta fue la fuente de información desde la cual se analiza el cambio (experiencia directa o información testigo). El otro aspecto que hemos tenido en cuenta en la definición de los formalismos propuestos, es la fiabilidad. Esta medida indica cuan fiable es el valor de confianza y qué relevancia merece en el proceso de toma de decisiones.

Tomando como punto de partida el trabajo presentado por Simari y Falappa en [16], fue desarrollado un modelo de cambio completo para órdenes parciales de credibilidad de los informantes. De esta manera, propusimos diferentes operadores de cambio (expansión, contracción, revisión

priorizada y revisión no priorizada) basados en una representación en forma de grafo de las relaciones de credibilidad entre agentes. En la tesis, se representa el orden de credibilidad por medio de un conjunto de tuplas, donde una tupla  $(A_i, A_j)$  (siendo  $A_i$  y  $A_j$  dos identificadores de agentes) indica que el agente  $A_i$  es menos creíble que el agente  $A_j$ . De esta manera, ampliamos el formalismo propuesto en [16] agregando fiabilidad a las tuplas de credibilidad para poder determinar que tuplas prevalecen a la hora de contraer o revisar. Con esto, hemos especificado claramente una manera racional de elegir que tuplas de credibilidad prevalecen luego de una contracción. Esta información adicional, a la cual la llamamos “fiabilidad” es representada por un identificador de agente, el cual será la fuente de la tupla de credibilidad entrante. Luego, introducimos un nuevo operador de revisión no priorizada por fiabilidad para orden parcial de credibilidad cuyo proceso es similar al operador de revisión no priorizado para creencias presentado en esta tesis.

Luego, hemos propuesto una revisión del orden de credibilidad basado en una creencia y su identificador asociado. Este nuevo operador consiste en modificar el orden de credibilidad como efecto secundario de una revisión no priorizada de ‘creencias’ que no tiene efecto por ser la entrada menos plausible que las creencias de la base del agente receptor que sustentan su contradicción.

Finalmente, ha sido presentada una técnica, la cual brinda la posibilidad de revisar la base de creencias de un agente con el efecto colateral de modificar el orden de credibilidad cuando es necesario. La misma combina el operador de revisión no priorizada para creencias y el operador de órdenes de credibilidad que se basa en las creencias.

## 5. Retransmisión de información basado en órdenes de credibilidad

Como otro de los aportes de esta tesis, hemos introducido algunos criterios que mejoran las habilidades de interacción de un agente en un sistema multi-agente, combinándolos con sus habilidades de razonamiento para permitir la propagación de información creíble. Hemos asumido, como en toda la tesis, un sistema multi-agente donde agentes deliberativos pueden recibir nueva información desde otros agentes y en el cual tienen creencias acerca de la credibilidad de sus compañeros. Al igual que hemos mencionado en las secciones previas, los agentes informantes pueden tener diferentes niveles de credibilidad.

Aquí, hemos propuesto enviar una pieza de información con un identificador de agente representando la credibilidad de la información transferida. Basado en este escenario, hemos investigado como un agente puede retransmitir información a otros agentes, la cual podría haber sido adquirida desde otros agentes. En particular, estudiamos como elegir racionalmente la meta-información a ser retransmitida. La elección del identificador de agente a ser enviado con la pieza de información es crucial, ya que influye en la decisión del receptor acerca de si aceptar la información transmitida. En la tesis, hemos presentado diferentes maneras de elegir este identificador y dimos una categorización de posibles enfoques, donde discutimos ventajas y desventajas. Esta discusión nos llevó a la definición de un criterio que usa una función de plausibilidad que determina la plausibilidad de una creencia basada en todas sus pruebas de acuerdo a una base. Es importante destacar, que en la tesis sólo se trabajó en criterios para retransmitir objetos de información. Criterios definidos y analizados:

- Identificador del emisor: consiste en enviar un objeto de información  $I = (\alpha, A_i)$  donde  $A_i$  es siempre el identificador del agente emisor y  $\alpha$  es la creencia a ser retransmitida.
- Identificador de la fuente: consiste en enviar un objeto de información  $I = (\alpha, A_i)$  donde  $A_i$  es siempre el identificador de la fuente original almacenada en la base de creencias del emisor. El escenario pensado para este criterio tiene la restricción de que la creencia  $\alpha$  a ser enviada aparece explícitamente en sólo un objeto de información en la base de creencias.

- Combinado: Como hemos fijado, puede haber diferentes tuplas conteniendo la misma sentencia. En este caso, el criterio consiste en enviar un objeto de información  $I = (\alpha, A_i)$  donde  $A_i$  puede ser el emisor o uno de los identificadores contenido en un objeto de información donde está almacenado  $\alpha$ . Podemos optar por una de dos políticas: elegir el agente más creíble, o el menos creíble.
- Función de plausibilidad: este criterio envía un objeto de información  $I = (\alpha, A_i)$  donde  $A_i$  es el identificador de agente obtenido usando la función de plausibilidad que describimos anteriormente.

Todos los criterios de retransmisión presentados fueron pensados para agentes que tiene un orden *total* de credibilidad. Sin embargo, hemos realizado un análisis de los criterios que se ven afectados cuando se consideran órdenes *parciales*. Hemos mostrado las dificultades que presentan cada uno de ellos y posibles soluciones.

Para dar información adicional a las creencias cuando son retransmitidas, hemos tomado una decisión importante, la cual consiste en retransmitir un identificador de agente con la creencia, en lugar de un etiqueta de credibilidad. Una de las razones se debe a que cada agente tiene su propio *assessment* y, por lo tanto, es más útil enviar identificadores de agente ya que de este modo el agente receptor puede evaluar la creencia recibida analizando la credibilidad que tiene el agente asociado de acuerdo a su propio *assessment*. Otra razón es porque la evaluación de la credibilidad de los identificadores es modular al uso de la función de *assessment*. Esto es, si el orden de credibilidad entre agentes cambia, entonces la plausibilidad de todas las sentencias también cambia, sin tener que modificar la base de creencias del agente.

## 6. Trabajos relacionados

En la literatura, han sido presentado diferentes formalismos para tratar con *multi-agent belief revision* (MABR) [12, 10, 13] donde se investiga la revisión de creencias global de un equipo de agentes. En contraste a éstos, nosotros nos enfocamos en *multi-source belief revision* (MSBR) el cual es uno de los componentes esenciales de MABR. Aquí, los agentes mantienen la consistencia de sus bases de creencias. Dos enfoques que hacen frente con MSBR son [6] y [5]. El modelo epistémico en estos trabajos es similar al que definimos en la Sección ??; sin embargo, nuestra teoría de cambio es diferente a ellas. Al igual que en esta tesis, ambos consideran que la fiabilidad de las fuentes afectan la credibilidad de la información entrante, y esta fiabilidad es usada para tomar decisiones. No obstante, estos dos enfoques difieren del de esta tesis en varios aspectos como se detallará a continuación.

En la tesis hemos comparado nuestro formalismo con un enfoque de revisión sobre bases de conocimiento priorizadas propuesto por Benferhat en [4]. Los operadores de revisión para bases de creencias propuestos en esta tesis, son similares al operador de revisión propuesto en [4]. Sin embargo, en [4], el estado epistémico está representado por una distribución de posibilidad la cual es un mapeo del conjunto de interpretaciones clásicas de mundos al intervalo  $[0,1]$ . Esta distribución representa el grado de compatibilidad de las interpretaciones con la información disponible y la revisión es hecha sobre la distribución de posibilidad.

Además, hemos comparado los formalismos propuesto para la dinámica de orden parcial con un proceso de revisión de creencias con dinámica de confianza y reputación realizado por Barber y Kim en [3]. Hemos mostrado que las ideas intuitivas detrás del formalismo propuesto en [3] están muy relacionadas a aquellas que motivaron el desarrollo de la técnica combinada de revisión de creencias y confianza desarrollada en la tesis. Sin embargo, a pesar de esta cercanía con el formalismo propuesto en la tesis, el enfoque desarrollado en [3] presenta varias diferencias con respecto a nuestro enfoque a nivel de modelo.

Luego, basados en el artículo [15] de Sabater y Sierra, hemos mostrado en forma resumida dos sistemas de confianza y reputación. El primero que hemos analizado fue ReGrE, propuesto por

Sabater y Sierra en [14], que a diferencia de nuestra propuesta, hace uso de información social y ontológica. También hemos mostrado como se administra información acerca de confianza y reputación en los sitios de subastas electrónicas como eBay [7]. La diferencia sustancial con nuestro enfoque radica en que en este tipo de sistemas consideran la confianza como una propiedad global. En cambio, en esta tesis la confianza es considerada una propiedad local al agente, debido al *assessment* que cada uno posee. En resumen, la principal diferencia entre las propuestas que realizamos en la tesis, con los modelos de reputación y confianza tratados en la tesis, es que nuestro objetivo está orientado hacia la actualización de órdenes de credibilidad buscando definir una teoría de cambio sobre la reputación y confianza de agentes. Esto es, combinamos formalismos de revisión de creencias y actualización de conocimiento con técnicas de mantenimiento de confianza y reputación de agentes en un ambiente distribuido. Como se puede observar, existe mucho trabajo realizado en las áreas de revisión de creencias y con fianza y reputación, sin embargo, la combinación de las mismas es algo novedoso.

## 7. Conclusiones

En esta tesis se presentaron formalismos para modelar la dinámica de conocimiento en bases de creencias de agentes que son parte de un sistema multi-agente. Para ello, hemos desarrollado un modelo epistémico para revisión de creencias realizada por un único agente que puede obtener nuevas creencias desde múltiples informantes (*multi-source belief revision* - MSBR). Estos informantes son agentes independientes quienes tienen sus propios intereses y, por lo tanto, podrían no ser completamente fiables. Es por esto, que en nuestro trabajo se propuso la organización de los informantes en un orden parcial que compara la credibilidad de los mismos. Basado en este orden, definimos una forma de pesar las creencias. Para ello, incluimos una función de plausibilidad que es utilizada en la definición de un criterio para comparar las creencias.

De esta manera, hemos investigado como la base de creencias de un agente puede ser modificada cuando un agente  $A$  recibe información desde múltiples agentes, los cuales, para  $A$ , pueden tener diferentes grados de credibilidad. Por lo tanto, una de las principales contribuciones de la tesis es la definición de diferentes operadores de cambio de creencia que usan la credibilidad de los agentes informantes para decidir que información prevalece.

A partir del modelo epistémico desarrollado, hemos definido un modelo completo de cambio para MSBR que toma como punto de partida las *kernel contractions* definidas por Hansson en [8]. En nuestra propuesta, fue incluido un operador de expansión que, a diferencia de la expansión propuesta en [9], consideramos objetos de información en vez de sólo sentencias. Luego, fue incluida una especialización de las *kernel contractions*. Esto es, las *kernel contractions* usan funciones de incisión, las cuales cortan en los subconjuntos mínimos que derivan la información a ser quitada. En la definición de *función de incisión* del trabajo de Hansson [8] no se especifica como la función elige las sentencias que serán descartadas de cada subconjunto. En el formalismo propuesto, esto fue resuelto con la plausibilidad de las sentencias que hemos definido. De esta manera, la función de incisión elige los objetos de información menos creíbles de cada subconjunto minimal. Luego, con este nuevo modelo epistémico para MSBR, y aprovechando la ventaja de contar con una forma de comparar las creencias, hemos propuesto un operador de revisión no priorizado, el cual usa la plausibilidad de las creencias tanto para evaluar la aceptación de la entrada, como para determinar qué creencias serán quitadas si la revisión tiene efecto.

Cada uno de estos operadores ha sido definido en forma constructiva y, además, para cada uno de ellos hemos introducido un conjunto de postulados. Para los operadores de cambio más importantes (contracciones y revisiones), hemos mostramos una caracterización axiomática a través de teoremas de representación.

Otra de las contribuciones fue el estudio y desarrollo de técnicas y formalismos para la actualización del grado de credibilidad que se le asigna a un agente por interactuar en el marco de un sistema multi-agente. Por lo tanto, en el desarrollo de nuestro trabajo también hemos estudiado



operadores de cambio, no sobre creencias (como mencionamos antes), sino sobre el orden parcial de credibilidad. Para ello, basado en algunos aspectos que son tenidos en cuenta en el diseño de sistemas de confianza y reputación, hemos propuesto dos maneras de realizar cambios sobre el orden de credibilidad de un agente:

- En principio hemos propuesto un modelo completo de cambio (expansión, contracción, revisión priorizada y revisión no priorizada) para órdenes parciales de credibilidad de los informantes.
- Luego, hemos definido un operador de cambio de orden parcial de credibilidad entre agentes que se basa en los objetos de información.

Como se puede apreciar, hemos introducido formalismos para que los agentes puedan adquirir creencias y actualizar sus órdenes de credibilidad. No obstante, en el desarrollo de nuestro trabajo también hemos sugerido como un agente puede retransmitir información ya sea propia u obtenida de otros agentes. Para esto, propusimos algunos criterios que mejoran las habilidades de interacción de un agente en un sistema multi-agente, combinándolos con sus habilidades de razonamiento para permitir la propagación de información creíble. De esta manera, hemos propuesto enviar una pieza de información con un identificador de agente representando la credibilidad de la información transferida. En particular, fue estudiado como elegir la meta-información a ser retransmitida. Hemos presentado diferentes maneras de elegir este identificador y dimos una categorización de posibles enfoques, donde discutimos ventajas y desventajas.

## 8. Líneas de investigación futuras

La confianza y la reputación son propiedades dependientes del contexto. Un modelo de confianza y reputación que administra un único contexto es diseñado para asociar un único valor de confianza o reputación por compañero sin tomar en consideración el contexto. Así es como hemos abordado este aspecto en la tesis. Sin embargo, un modelo de multi-contexto tiene los mecanismos para tratar con varios contextos, manteniendo diferentes valores de confianza y reputación asociados a estos contextos para un único compañero (como lo hacen Abdul-Rahman y Hailes en [1]). Como trabajo a futuro, buscaremos extender lo realizado en esta tesis hacia formalismos multi-contexto.

En esta tesis hemos sugerido diferentes criterios de retransmisión de información, teniendo en cuenta órdenes totales de credibilidad entre agentes. También, hemos realizado un análisis de cómo algunos de estos criterios se ven afectados cuando el orden es relajado a orden parcial. Por medio de este análisis, hemos arribado a posibles soluciones que requieren la modificación del modelo de cambio de creencias propuesto, para permitir la incorporación de información, considerando múltiples fuentes para la misma creencia. Como trabajo a futuro, proponemos extender el modelo de cambio de creencias de forma tal, que los objetos de información tengan asociados un conjunto de identificadores de agentes (representando todas las fuentes de la creencia), en lugar de uno sólo. Esto le daría la posibilidad al agente receptor de determinar, de manera más apropiada, la plausibilidad de la creencia entrante basado en su propio *assessment*.

Otro trabajo que consideramos interesante, consiste en hacer un análisis y desarrollo de nuevos operadores de revisión múltiple, tanto de creencias como de órdenes parciales de credibilidad. Esto nos daría la posibilidad de realizar modificaciones en las bases por medio de conjuntos de creencias o conjuntos de tuplas de credibilidad.

## Referencias

- [1] Alfarez Abdul-Rahman and Stephen Hailes. Supporting trust in virtual communities. In *Proc. of the Hawaii's International Conference on Systems Sciences*, Maui, Hawaii, 2000.

- [2] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *J. of Symbolic Logic*, 50(2):510–530, 1985.
- [3] K. S. Barber and J. Kim. Belief revision process based on trust: Simulation experiments. In *In Proceedings of Autonomous Agents '01 Workshop on Deception, Fraud, and Trust in Agent Societies*, pages 1–12, 2001.
- [4] Salem Benferhat, Didier Dubois, Henri Prade, and Mary-Anne Williams. A practical approach to revising prioritized knowledge bases. *Studia Logica*, 70(1):105–130, 2002.
- [5] John Cantwell. Resolving conflicting information. *Journal of Logic, Language and Information*, 7(2):191–220, 1998.
- [6] Aldo Dragoni, Paolo Giorgini, and Paolo Puliti. Distributed belief revision versus distributed truth maintenance. In *Proceedings of the Sixth IEEE International Conference on Tools with Artificial Intelligence (TAI 94)*, pages 499–505. IEEE Computer Society Press, 1994.
- [7] eBay. ebay. In <http://www.eBay.com>, 2002.
- [8] Sven Ove Hansson. Kernel contraction. *Journal of Symbolic Logic*, 59(3):845–859, 1994.
- [9] Sven Ove Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer Academic Publishers, 1999.
- [10] Noa E. Kfir-Dahav and Moshe Tennenholz. Multi-agent belief revision. In *Theoretical Aspects of Rationality and Knowledge: Proceeding of the Sixth Conference (TARK 1996)*, pages 175–196. Morgan Kaufmann Publishers Inc., San Francisco, 1996.
- [11] Patrick Krümpelmann, Luciano H. Tamargo, Alejandro J. García, and Marcelo A. Falappa. Forwarding credible information in multi-agent systems. *Proceedings of the 3rd International Conference on Knowledge Science, Engineering and Management (KSEM 2009)*, 5914/2009:41–53, November 2009.
- [12] Wei Liu and Mary-Anne Williams. A framework for multi-agent belief revision. *Studia Logica*, 67(2):291–312, 2001.
- [13] Benedita Malheiro, Nicholas Jennings, and Eugenio Oliveira. Belief revision in multi-agent systems. *Proceeding of the 11th European Conference on Artificial Intelligence (ECAI 94)*, pages 294–298, 1994.
- [14] Jordi Sabater and Carles Sierra. Regret: A reputation model for gregarious societies. *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, pages 61–69, 2001.
- [15] Jordi Sabater and Carles Sierra. Review on computational trust and reputation models. *Artificial Intelligence Review*, 24(1):33–60, 2005.
- [16] Patricio D. Simari and Marcelo A. Falappa. Revision of informant plausibility in multi-agent systems. *Journal of Computer Science and Technology*, 2(5), 2001.
- [17] Luciano H. Tamargo, Marcelo A. Falappa, Alejandro J. García, and Guillermo R. Simari. A change model for credibility partial order. In *Proceeding of the 5th International Conference on Scalable Uncertainty Management (SUM)*, pages 317–330, 2011.
- [18] Luciano H. Tamargo, Alejandro J. García, Marcelo A. Falappa, and Guillermo R. Simari. Consistency maintenance of plausible belief bases based on agents credibility. *12th International Workshop on Non-Monotonic Reasoning (NMR)*, pages 50–58, 2008.
- [19] Luciano H. Tamargo, Alejandro J. García, Marcelo A. Falappa, and Guillermo R. Simari. A belief revision approach to inconsistency handling in multi-agent systems. *The IJCAI-09 Workshop on Nonmonotonic Reasoning, Action and Change (NRAC)*, pages 63–70, 2009.
- [20] Luciano H. Tamargo, Alejandro J. García, Marcelo A. Falappa, and Guillermo R. Simari. Modeling knowledge dynamics in multi-agent systems based on informants (en prensa). *Proceedings of The Knowledge Engineering Review (KER)*, 27(1):87–114, 2012.