

Un Modelo Lógico-Formal para el Estudio de los Argumentos Emocionales en los Procesos de Construcción de Acuerdos*

A Logical-Formal Model for the Study of Emotional Arguments in Negotiation Processes

Um Modelo Lógico-Formal para o Estudo dos Argumentos Emocionais nos Processos de Construção de Acordos

ANTONIO AGUILERA ONTIVEROS

JULIO C. CONTRERAS MANRIQUE

El Colegio de San Luis, México

Resumen

Los argumentos son parte de un proceso comunicativo con el cual se trata de incidir en la acción de otros. Gilbert (1994) identifica cuatro modos de argumentación: el modo lógico, el modo emocional, el modo visceral y el modo kisceral. Siguiendo la línea de investigación en psicología computacional marcada por Ortony, Clore y Collins (1988) y el modelo de resolución de conflictos usando negociaciones basadas en argumentos propuesto por Jung y Tambe (2001), este trabajo presenta un modelo lógico-formal para el estudio de un modo concreto de argumentos emocionales dentro del contexto de formación de consensos enmarcado en un proceso de negociación/coordinación. Se discuten sus implicaciones en los modelos cognitivos emocionales basados en el proceso de apreciación/evaluación de la emoción.

Palabras clave: teoría de la argumentación, argumentos emocionales, consenso emocional, métodos formales en sistemas multi-agente, psicología computacional.

Abstract

Arguments are part of a communicative process through which people try to influence the actions of others. Gilbert (1994) identifies four modes of argumentation: (a) logical, (b) emotional, (c) visceral, and (d) kisceral. Following the line of research in computational psychology proposed by Ortony, Clore and Collins (1988), and the model of conflict resolution using argumentation-based negotiations proposed by Jung and Tambe (2001), this paper presents a logical-formal model for studying of emotional arguments within the context of consensus building framed in negotiation and coordination processes.

Keywords: argumentation theory, emotional arguments, emotional consensus, formal methods in multi-agent systems, computational psychology.

Resumo

Os argumentos são parte de um processo comunicativo com o qual se trata de incidir na ação de outros. Gilbert (1994) identifica quatro modos de argumentação: o modo lógico, o modo emocional, o modo visceral e o modo kisceral. Seguindo a linha de investigação em psicologia computacional definida por Ortony, Clore y Collins (1988) e o modelo de resolução de conflitos usando negociações baseadas nos argumentos propostos por Jung e Tambe (2001), este trabalho apresenta um modelo lógico-formal para o estudo de um modo concreto de argumentos emocionais dentro do contexto de formação de consensos demarcado em um processo de negociação / coordenação. Discutem-se suas implicações nos modelos cognitivos emocionais baseados no processo de apreciação / avaliação da emoção.

Palavras-chave: teoria da argumentação, argumentos emocionais, consenso emocional, métodos formais em sistemas multi-agente, psicología computacional.

La correspondencia en relación con este artículo puede dirigirse a cualquiera de los autores. Antonio Aguilera Ontiveros, e-mail: aaguilera@colsan.edu.mx; Julio C. Contreras Manrique, e-mail: ccontreras@colsan.edu.mx. Departamento de Programa de Estudios Políticos e Internacionales, El Colegio de San Luis; A.C. Parque de Macul 155, Fracc, Colinas del Parque, C.P. 78299, San Luis Potosí, SLP, México.

ARTÍCULO DE REFLEXIÓN

RECIBIDO: 23 DE JULIO DEL 2010 - ACEPTADO: 03 DE OCTUBRE DEL 2011

* Los autores agradecen al Fondo Mixto de Fomento a la Investigación Científica y Tecnológica CONACYT-Gobierno del Estado de San Luis Potosí, proyecto: FMSLP-2003-C02-11121 y al Fondo Sectorial de Investigación para la Educación SEP-CONACYT: Investigación Básica n.º de apoyo: 101366, quienes financiaron esta investigación.

TODOS LOS días, en las interacciones con otros, se usan de una u otra forma argumentos, puesto que son parte de un proceso comunicativo con el cual se trata de incidir en la acción (o inacción) de los otros. Tal como mencionan Perelman y Olbrechts-Tyteca (1994) “La argumentación es una actividad que siempre trata de modificar un estado de cosas preexistente” (p. 105).

Dentro de los marcos comunicativos, la argumentación establece un proceso de gran relevancia ya que permite que dos individuos con diferentes conjuntos de metas y/o deseos puedan llegar a un acuerdo y lograr la coordinación de sus esfuerzos individuales (Wooldridge, 2002). Es más, en ciertas circunstancias, se crean estados de ánimo con la argumentación que permite pasar de la coordinación a la colaboración. Todo lo anterior hace que el estudio de los procesos argumentativos sea de gran relevancia en todas aquellas situaciones problemáticas en donde se trata de establecer la coordinación y la colaboración.

Los argumentos pueden ser vistos como el producto de una persona (Argumento 1) y como el proceso o procedimientos de interacción entre varias personas (Argumento 2) (O’Keefe, 1977; Wenzel, 1992). Estas formas de conceptualizar los argumentos llevan a establecer diferentes dominios de estudio, que van desde el estudio de los métodos y marcos conceptuales para el análisis de los argumentos hasta el establecimiento de aspectos teóricos relacionados con los estándares normativos para la evaluación de los argumentos. Siguiendo a Hamble (1992), a los Argumentos 1 y 2 se les puede añadir un tercer tipo, el Argumento 0. Dicho tipo contempla la dimensión cognitiva del argumento, es decir los procesos mentales por los cuales ocurren los argumentos dentro de las personas. Esta nueva dimensión argumentativa ampara varios modelos cognitivos, propios de la psicología computacional, que involucran las emociones como parte de los mecanismos de la argumentación (ver por ejemplo Marsella, Gratch & Petta, 2010; Thagard & Kroon, 2006).

La argumentación es un proceso complejo cuyo estudio se puede dividir en los siguientes tópicos: (a) las intenciones con las que los individuos crean los argumentos, (b) los contextos valorativos en donde se dan dichos argumentos, (c) el auditorio al que se trata de influir, y (d) el efecto producido por el argumento en la audiencia (Perelman & Olbrechts-Tyteca, 1994).

Para los autores de este estudio, y siguiendo a Perelman y Olbrechts-Tyteca (1994), la idea básica detrás de todo argumento es convencer al interlocutor de la verosimilitud de lo que dice el ponente, consiguiendo con esto una ventaja concreta o un cambio en la situación del interés del ponente. Para entender el proceso anterior se han desarrollado modelos concretos de análisis que sugieren mecanismos específicos en los cuales se basa el proceso argumentativo. Uno de dichos modelos es propuesto por Gaviria-Martínez y Corredor-Aristizábal (2011), quienes indican que el proceso argumentativo, para ser plausible, necesita de una precondition basadas en la causalidad de las partes que conforman el argumento. Estos últimos autores, siguiendo la línea de trabajo establecida por Toulmin, establecen que dicha causalidad se logra a través del uso de evidencia que respalda el argumento. Dicha evidencia tiene dos usos: como dato que respalda la afirmación, y como respaldo de una garantía de que efectivamente existe una relación entre los datos y la afirmación que está siendo defendida. Ambos casos son ejemplos de razonamiento no deductivo, en el que la verdad de las premisas no convierte en verdadera la conclusión, sino que la hace probable. Lo interesante de este trabajo es que se trata de una investigación empírica en la que se estudia la respuesta de dos grupos de participantes frente a los dos tipos de argumentos antes descritos, considerando el primer grupo con formación específica en habilidades investigativas y argumentativas en ciencia, y el otro sin ellas.

Otro tipo de modelo sobre el proceso argumentativo es el propuesto por Thagard y Kroon

(2006). Dicho modelo establece que las decisiones grupales son frecuentemente emocionales. Basados en dicha premisa proponen una teoría y un modelo computacional de decisión grupal, la cual se interpreta como el alcance de un consenso emocional. Thagard y Kroon sugieren tres mecanismos básicos que permiten lograr consenso emocional: (a) la comunicación argumental de las emociones, así como la transmisión de información relevante sobre los hechos; (b) el contagio emocional, basado en la teoría de que las personas implicadas en una discusión tienden, de forma automática, a imitar y sincronizar sus movimientos con las expresiones faciales, tonos de voz, posturas, movimientos y comportamiento general de los demás, lo que lleva a que las experiencias emocionales sean influidas, y (c) mecanismos mixtos que involucran la emisión de información tanto en forma verbal como no verbal, esto es, el altruismo, la compasión y la empatía. Thagard y Kroon construyeron un modelo computacional llamado HOTCO₃, que utiliza redes neuronales como su principal mecanismo de cálculo y el lenguaje LISP para su implementación.

En este trabajo se presenta un modelo lógico-formal para el estudio de un modo concreto de argumentos, los argumentos emocionales, dentro de un contexto de formación de consensos enmarcado en un proceso de negociación/coordinación. Este trabajo se ubica dentro de la línea de modelos cognitivos emocionales basados en el proceso de apreciación/evaluación de la emoción, tales como los desarrollados por Fridja, Lazarus y Scherer (en Marsella et al., 2010) y por Ortony, Clore y Collins (1988). En concreto, el modelo propuesto en este estudio se basa en el trabajo de Ortony et al. (1988) y de O'Rorke y Ortony (1994). La primera sección explica brevemente la problemática de la argumentación emocional y su relevancia dentro de los procesos de toma de decisiones colectivas/influencia social. En la segunda parte del trabajo se presenta el modelo propuesto, el cual está basado

en aspectos concretos del modelo cognitivo de emociones sugerido por Ortony et al. (1988). La tercera parte corresponde al establecimiento de las ideas, conceptos, constructos y relaciones que conforman dicho modelo. En la parte conclusiva de este trabajo, se adelantan algunas ideas sobre la utilidad del diseño modélico que se propone y el trabajo a seguir en el futuro.

La Argumentación Emocional y las Decisiones Emocionales

Dentro de la amplia literatura sobre el estudio de la argumentación, todo argumento tiene partes constitutivas que pueden ser abordadas usando elementos de la lógica formal. En este contexto, es de gran interés la forma en que los argumentos se articulan para crear un corpus deductivo concreto que establezca la verdad o falsedad del mensaje inmerso en la estructura del argumento. Sin embargo, también es cierto que todo argumento contiene partes que son no lógicas. Estas partes no lógicas están tanto en el contexto, como en la "carga" (e. g. el énfasis, lenguaje corporal y tono de voz) que acompañan a la transmisión de la estructura argumentativa. El reconocimiento de estas partes no lógicas ha llevado a la propuesta de Gilbert (1994) de identificar cuatro modos de argumentación:

1. El modo lógico. Este modo de argumentar se asemeja a las pruebas matemáticas. Su naturaleza es deductiva, esto es, las premisas implican la conclusión.
2. El modo emocional. Aquí la argumentación ocurre bajo un contexto en donde se recurre a los sentimientos, actitudes, filias y fobias para motivar la aceptación del mensaje transmitido en el argumento.
3. El modo visceral. Este modo de argumentación implica el uso de emociones primarias que corresponden a estados físicos instintivos muy intensos, tales como la alegría, la ira y el miedo.
4. El modo kisceral. Este modo implica cuestiones intuitivas, místicas y religiosas.

Los argumentos emocionales pueden contener aspectos de conflicto en la interrelación comunicativa entre individuos, tales como el uso de sarcasmos e ironías descalificadoras, que los argumentos lógicos no logran capturar. Estos aspectos incluyen nociones como el grado de compromiso con la postura propia, la profundidad y alcance de los sentimientos mostrados en el argumento, la sinceridad del individuo y el grado de resistencia al argumento (Gilbert, 1995). Todos estos aspectos son componentes importantes y básicos en la comunicación de una posición concreta a través de la argumentación. En este sentido, la emoción se presenta en los Argumentos 2 más que en los Argumentos 1.

Los argumentos del tipo *tu quoque* (tú también), en los que los interlocutores se achacan entre sí los mismos comportamientos, punibles o loables, son un ejemplo de argumento emocional. Obviamente, en un contexto lógico este tipo de argumentos son considerados una falacia (en este caso *ad hominem*) y prescriptos como “malos” argumentos. Sin embargo, fuera del marco de análisis formal, este tipo de argumentos son extremadamente “ricos” en su capacidad de lograr acuerdos o de motivar el rompimiento de los mismos. Esto se debe a que la emoción es esencial para romper estancamientos en los procesos de negociación/convencimiento dentro de un contexto racional tradicional.

Tradicionalmente, las emociones se han conceptualizado como una contradicción del comportamiento racional. Siguiendo esta premisa, las emociones afectarían a los procesos de negociación al menoscabar el proceso de toma de decisiones, el cual se considera como intrínsecamente racional (Howard, 1993). Sin embargo, la evidencia reciente (Damasio, 2005; Elster, 1996; Loewenstein, 2000; Mellers & McGraw, 2001; Naqvi, Shiv & Bechara, 2006) muestra que las emociones son parte del proceso de decisión. Por su parte los trabajos de Fineman (1996) y Thoits (1989) son evidencia de la variabilidad histórica y cultural de los procesos de

significar, expresar y regular las emociones. Lo anterior sugiere que las experiencias subjetivas y emocionales son creencias socialmente adquiridas y estructuradas (Thoits, 1989). Por su parte, Trapp y Payr (2002) sugieren que las emociones son un requisito previo para el comportamiento racional, mientras que Fineman (1996) enuncia que las emociones cambian la racionalidad del tomador de decisiones.

Pero ¿cómo una emoción puede afectar el comportamiento racional? Hay cuatro elementos básicos del proceso racional de toma de decisiones (Elster, 1996): (a) los deseos, (b) las creencias, (c) la información y (d) la acción. Bajo este esquema, una decisión es racional si es la mejor manera de lograr los deseos del agente en el marco del conjunto de sus creencias. Lo anterior coincide con el principio de racionalidad expuesto por Newell (1981) que establece: “Si un agente tiene conocimiento de que una de sus acciones lo llevará a alcanzar una de sus metas entonces el agente seleccionará dicha acción” (p. 8). Las creencias, dentro de un proceso de decisión, se forman bajo dos mecanismos básicos: (a) son resultado de la información disponible para el agente y (b) son resultado de creencias previas en la mente del agente (Miller, 1983). Si las creencias acerca de un problema de decisión están fundadas en un conocimiento equivocado o una suposición errada, el proceso racional de decisión lleva a una decisión no óptima (Elster, 1997). Por supuesto, la información afecta a las creencias, pero también las emociones afectan el proceso de formación de creencias. Como Frijda y Mesquita (2000) sostienen:

Las emociones influyen en las creencias mediante la creación y fortalecimiento de las convicciones, en particular haciendo a las creencias resistentes al cambio. Estas influencias son en parte el resultado directo de emociones agudas, en parte de anticipaciones emocionales, y en parte a emociones o sentimientos disponibles permanentemente. (p. 72)

Las emociones no solo afectan a las creencias. Elster (1996) sostiene que hay al menos siete maneras diferentes en que las emociones pueden encajar en el proceso de toma de decisiones, haciéndolo de esta manera un proceso complejo de modelar. En el objetivo desarrollado en este estudio se parte del cambio de la creencia como el eje de análisis central. Entonces, con este eje se puede establecer el papel que juegan las emociones en los procesos de decisión involucrados en un proceso de construcción de acuerdos. Los actores involucrados en dicho proceso tienen creencias, información y metas concretas que lograr. En el proceso de formación de acuerdos, los actores intercambian mensajes concretos sobre la forma en que el otro debe actuar. Dichos mensajes van embebidos dentro de un marco argumentativo que refiere a modos lógicos, modos emocionales y modos kicerales, que logran, o no, afectar el esquema de decisión de los actores. En el modelo propuesto, el esquema de decisión está conformado por los deseos y las creencias del actor. Las creencias establecen la forma en que los deseos se pueden concretar en la mente del actor. Si un argumento emocional logra la afectación de la estructura de decisión, entonces el actor afectado cambia sus creencias, y por ende los deseos, y con ello permite una evaluación del argumento a favor de la acción sugerida, logrando con ello el consenso y con esto la coordinación. Si todo sucede en un marco de emociones positivas tales como la alegría, el interés, la satisfacción y el amor, se dice que existe la colaboración.

En el siguiente apartado se establece, desde un punto de vista cognitivo, cómo las emociones afectan la transformación de las creencias motivando un cambio en los deseos.

El Modelo Emocional de Ortony, Clore y Collins

Ortony et al. (1988) y O'Rourke y Ortony (1994) proponen un modelo formal para explicar el comportamiento emocional. Este modelo

se basa en los aspectos cognitivos del comportamiento humano y puede formularse como un modelo computacional. Los aspectos cognitivos de las emociones implican el reconocimiento, la comparación, la clasificación, la inferencia, la atribución y la evaluación. Las emociones son vistas como reacciones de valencia hacia los objetos, los eventos, los agentes y sus acciones (Ortony et al., 1988).

De acuerdo con el modelo de Ortony et al. (1988) y de O'Rourke y Ortony (1994), los diferentes tipos de emociones conforman grupos o conjuntos de condiciones que causan comportamientos emocionales y que pueden ser codificados en algún lenguaje computacional. En este contexto hay seis grupos:

- El primer grupo implica la forma más sencilla de las emociones. Estas son las *emociones de autobienestar*, las cuales son la *alegría* y la *angustia*. Estas emociones pueden entenderse como los aspectos positivos y negativos de la retroalimentación que los individuos tienen ante posibles acontecimientos deseables o indeseables.
- El segundo grupo de emociones se refiere a la *fortuna de los demás*. Este grupo incluye cuatro subtipos de emociones: *feliz por*, *regocijo malicioso por*, *el resentimiento* y *apesadumbrado por*. Cada subtipo en este grupo es una combinación de aspectos agradables o desagradables respecto a un acontecimiento concreto, que se presume que es agradable o desagradable para otra persona.
- El tercer grupo es el de las *emociones basadas en prospectos* que incluye seis subtipos específicos de emociones: la *esperanza*, la *satisfacción*, el *alivio*, el *miedo*, los *temores confirmados* y la *decepción*. Cada subtipo representa una reacción a un evento deseable o indeseable que puede estar o no confirmado.
- El cuarto grupo se refiere a las *atribuciones* y abarca cuatro subtipos de emociones:

el *orgullo*, la *admiración*, la *vergüenza* y el *reproche*. Cada subtipo de emoción de atribución genera una reacción (positiva o negativa) a cualquier tipo de acción propia o a acciones llevadas a cabo por otros.

- El quinto grupo está relacionado con la *atracción*. Este grupo se refiere a las reacciones que las personas tienen hacia los objetos. Los dos subtipos de emociones en este grupo son los *sentimientos momentáneos de agrado* o *rechazo* hacia un objeto.
- El sexto, y último grupo, incluye ocho elementos que constituyen el resultado del producto cartesiano de conjuntos *bienestar* × *atribuciones*. Es importante notar que este tipo de emociones compuestas no se corresponde con la coocurrencia de sus componentes emocionales. La unión lógica se usa para resolver cuándo una emoción compuesta es válida en el modelo computacional (O'Rorke & Ortony, 1994).

Dentro del modelo, las condiciones de "producción" de una emoción se dan en términos de las variables que contribuyen a aumentar la actividad de las emociones. La teoría especifica las variables globales, que afectan a todas las emociones, y las variables locales, que afectan solo a subgrupos de las emociones (Ortony et al., 1988). Las variables tienen valores y pesos asociados a las emociones, y la teoría sugiere que una emoción se experimenta solo cuando ciertos umbrales que la disparan son superados.

Para desarrollar el objetivo de este estudio, se utilizó esta clasificación de las emociones como una ontología para construir los mecanismos emocionales del modelo que se propone. En la siguiente sección se explica dicho modelo con su marco formal.

El Modelo de Argumentación Emocional

En esta sección se presenta una definición formal, propuesta por los autores, del proceso de

la argumentación emocional dentro de un contexto de formación de acuerdos. Dicho formalismo utiliza el marco de la inteligencia artificial distribuida, es decir los sistemas multiagente (Wooldridge, 2000, 2002) en los que un agente es una entidad artificial con capacidades cognitivas, comunicativas y de acción. La arquitectura de agentes cognitivos será la establecida por el modelo BDI (Creencias, Deseos e Intenciones, por sus siglas en inglés) (Rao & Georgeff, 1995; Wooldridge, 2000), el cual está basado en la tradición filosófica del razonamiento práctico (Bratman, 1992).

El modelo BDI modela un agente racional que tiene las siguientes actitudes mentales: las creencias (*beliefs*), los deseos (*desires*) y las intenciones (*intentions*), los cuales corresponden respectivamente a la información, la motivación y la capacidad de deliberación de un agente. Dichas actitudes mentales determinan el comportamiento del sistema y son críticas para alcanzar un desempeño óptimo o adecuado cuando el proceso de deliberación está sujeto a recursos limitados (Rao & Georgeff, 1995).

Para el contexto de negociación/formación de acuerdos se empleó el modelo CONSA (*Collaborative Negotiation System based on Argumentation*) sugerido por Jung y Tambe (2001). Dicho modelo conceptualiza la resolución de conflictos como una meta común que explícitamente deben lograr los miembros de un equipo de trabajo. La interacción entre los miembros del equipo de trabajo que tienen conflicto entre sí es, entonces, modificada por un conjunto de argumentos vertidos por los mismos miembros del equipo. La aceptación de los argumentos depende no solo de la naturaleza del conflicto sino también de la dinámica del mismo. Así, un evento que genera un conflicto en el tiempo t_0 , puede cambiar su estado en el tiempo t_1 , inhibiendo entonces el conflicto. Esto puede ser conocido por uno de los miembros del equipo que informa a los demás de la situación, resolviendo de esta forma el conflicto. El modelo permite

que esquemas de argumentos que ayudaron a resolver conflictos en el pasado sean reciclados en situaciones similares.

En el modelo propuesto en este artículo se parte de una organización compuesta por agentes, en la que cada agente tiene un conjunto de problemas que debe resolver a través de la conformación de acuerdos con otros agentes. Como en el modelo CONSA, es explícita la meta de resolver conflictos ya que estos inhiben los acuerdos. Sin embargo, el modelo propuesto parte del supuesto de la existencia de grupos de interés dentro de los grupos de trabajo. Los agentes que pertenecen a un mismo grupo de interés son más propensos a lograr acuerdos entre sí, mientras que resisten el llegar a acuerdos con agentes de otros grupos de interés. Como restricción, los agentes solo pueden pertenecer a un solo grupo de interés. Cada problema necesita un conjunto de acuerdos específicos que lo soluciona. Esto puede ser definido como sigue:

Definición 1: Sea $\Theta = \{\alpha_1, \dots, \alpha_k\}$ un grupo de agentes. Sea G_1 y G_2 grupos de interés, tales que $\forall \alpha \in \Theta \Rightarrow \alpha \in G_1 \vee \alpha \in G_2 \wedge G_1 \cap G_2 = \emptyset$. Sea $P = \{p_1, \dots, p_m\}$ un conjunto de problemas que deben resolver los agentes. Sea $S = \{s_{\alpha\alpha}^1, \dots, s_{\alpha\alpha}^m\}$ un conjunto de acuerdos necesarios para solucionar los problemas. Los acuerdos son apareamientos de agentes del tipo $s_{ij}^k = (\alpha_i, \alpha_j)$. Se define una *estructura de solución* como $E_{\text{solución}} = \{\Theta, P, S, \Theta \cdot P\}$ en donde $\Theta \cdot P$ es una relación de equivalencia que relaciona respectivamente a los agentes con los problemas.

La estructura de solución funciona como sigue: cada agente tiene una afiliación específica a un grupo de interés. Cada agente tiene un conjunto de problemas que resolver. Cada problema requiere de un conjunto de acuerdos para ser resuelto. Los agentes pueden unirse de dos en dos para tomar acuerdos con los cuales resolver los problemas. La estructura de solución especifica las reglas de interacción/solución del sistema. Un problema se considera resuelto si el agente que lo tiene logra establecer el conjunto de acuerdos

que lo soluciona. Para entender esto considérese el siguiente ejemplo: Sea $\Theta = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ un conjunto de agentes. Sea $P = \{p_1, p_2, p_3, p_4\}$ un conjunto de problemas. Sea $S = \{s_{12}^1, s_{34}^2, s_{34}^3, s_{14}^4\}$ un conjunto de acuerdos que solucionan los problemas. Nótese que s_{12}^1 significa que el problema 1 se resuelve con el acuerdo de los Agentes 1 y 2. Sea $\Theta \cdot P = \{(\alpha_1, p_1, p_2, p_4), (\alpha_2, p_1), (\alpha_3, p_2, p_3), (\alpha_4, p_4)\}$ la relación de equivalencia entre los agentes y los problemas. Nótese que S y $\Theta \cdot P$ son congruentes. Sean $G_1 = \{\alpha_1, \alpha_4\}$ y $G_2 = \{\alpha_2, \alpha_3\}$ las filiaciones de cada agente con un grupo específico.

Como ya se estableció anteriormente, los agentes están dispuestos a crear acuerdos con los agentes de su propio grupo de interés, mientras que son reacios a generar acuerdos con agentes que no son de su mismo grupo. Nótese que el problema 1 requiere el acuerdo entre los Agentes 1 y 2, pero como estos pertenecen a grupos distintos de interés, entonces no hay acuerdo. En términos generales, solo aquellos problemas que requieran un acuerdo entre agentes que pertenecen al mismo grupo serán solucionados.

Es importante recordar que el modelo que aquí se propone usa agentes del tipo BDI, es decir que las decisiones de los agentes están en función de sus creencias, deseos e intenciones. Así, los Agentes 1 y 2 deciden no llegar a un acuerdo debido a que “creen” que miembros de distintos grupos no deben llegar a acuerdos. Por otro lado, los agentes pueden “desear” resolver sus problemas respectivos, tal y como establece la premisa del modelo CONSA, y al mismo tiempo desear ser obedientes de la disciplina del grupo. Dada la tensión entre las creencias y los deseos, los agentes están en un estado deliberativo en el que establecen su “intención”, la cual es negar el acuerdo. Nótese que la intención es una consecuencia de la creencia (i. e. miembros de distintos grupos no deben tener acuerdos) y el deseo (i. e. respetar al grupo).

Tal como establecen Jung y Tambe (2001), esta situación disfuncional (conflicto) es resuelta a través de un proceso de argumentación. El

modelo CONSA establece las representaciones y los algoritmos para evaluar el argumento. Dentro del modelo CONSA, un argumento está dentro de un patrón de argumentación de Toulmin (TAP) (Jung & Tambe, 2001; Kneupper, 1978). En un TAP, un argumento consiste en los siguientes elementos: (a) una demanda, que es la conclusión cuyo merito el agente busca establecer; (b) los datos, que son los hechos o evidencias en los cuales se fundamenta la demanda; (c) la justificación, que establece la relevancia de la evidencia sobre la demanda bajo la forma de reglas, principios y patrones, y (d) el cualificador modal, que establece el grado de fuerza que se le confiere a la demanda con base en los datos y la justificación; este suele ser del tipo: quizá, seguramente, típicamente, usualmente, algunos, pocos, algunas veces, la mayoría, probablemente y tal vez (Jung & Tambe, 2001). Nótese que el modelo CONSA se basa en los elementos del modelo original de Toulmin (2003), esto es: (a) la aserción, (b) los datos y (c) la garantía. A su vez, la aserción es modificada por (d) el respaldo, que son las evidencias que sustentan la aserción; (e) el cualificador modal, que indica el grado de fuerza o de probabilidad de la aserción, y (f) la reserva, que establece objeciones posibles que se le puedan formular a la aserción.

Lo anterior permite sintetizar un modelo formal en el que un argumento dentro de un TAP puede modificar una creencia articulada a un deseo y a una intención, con el fin de establecer la forma en que el argumento modifica la decisión de un agente. Sean entonces:

Definición 2: Sea $B^\alpha = \{b_1^\alpha, b_2^\alpha, \dots, b_n^\alpha\}$ un conjunto de creencias del agente α . Sea $D^\alpha = \{d_1(b_1^\alpha), \dots, d_n(b_n^\alpha)\}$ un conjunto de deseos del agente α . Nótese que $d(b_i^\alpha)$ significa que el agente α tiene el deseo d asociado a la creencia b_i . Sea $I = \{i_1(d_1(b_1^\alpha)), \dots, i_n(d_n(b_n^\alpha))\}$ el conjunto de intenciones i asociadas a los deseos del agente α .

En el modelo CONSA una creencia es una demanda. En la argumentación, las

creencias-demanda forman las propuestas con su TAP de soporte. Por ejemplo, retomando a los dos Agentes 1 y 2, los cuales pertenecen a diferentes facciones, la creencia-demanda puede ser “Debes llegar a acuerdos con los miembros de otros grupos”. Los datos o evidencias pueden ser también creencias-demanda, tales como: “Tú tienes el problema P, Yo tengo el problema P”, “Tu perteneces a otro grupo”, “Nuestro deseo de resolver se concreta si cooperamos”. La justificación sería: P se resuelve con el acuerdo entre los Agentes 1 y 2. El cualificador modal podría ser: la mayoría de los problemas no resueltos se deben a que el acuerdo no se da entre agentes de grupos distintos. Para formalizar lo anterior:

Definición 3: Sea $\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$ el conjunto de cualificadores modales. La estructura formal de un TAP es entonces: $TAP = \{c, E, I, \omega\}$ donde $c \in B^\alpha$, esto es, c es una creencia que pertenece al conjunto de creencias del agente α . Por su parte, $E = \{x | x \in B^\alpha\}$, es un subconjunto de creencias que pertenecen al conjunto de creencias del agente α . $J = f : \{c, E\} \rightarrow \{T, \perp\}$ es una función que asigna un valor de verdad, T (verdadero) \perp (falso) a la creencia específica c y a las creencias que son elementos del conjunto E . La expresión $\omega \equiv \exists x \in \Omega \therefore \diamond \exists \{c, x(E)\} \xrightarrow{f} \{T\}$ significa que ω se define como la existencia de un elemento x del conjunto de cualificadores modales tal que *posiblemente* (\diamond) existe una lista de creencias, en donde una de ellas es la creencia c , cuyos valores de verdad son todos T (verdadero). Nótese que la estructura define tanto las partes del argumento como la forma en que este es evaluado por un agente receptor del mismo.

Esta estructura es fácilmente modificable para considerar cualificadores emocionales en vez de los modales. Dichos cualificadores emocionales modifican las estructuras de la función $f: \{c, E\} \rightarrow \{T, \perp\}$. Para poder incorporar estos últimos al modelo TAP es necesario antes definirlos:

Definición 4: Sea \hat{o} una ontología de Ortony, esto es, $\hat{o} \neq \emptyset$ y $\hat{o} = \{\Psi_1, \dots, \Psi_n\}$, donde Ψ_k es un estado emocional de alguno de los seis

grupos de emociones definidos por Ortony et al. (1988): (a) emociones de autobienestar, (b) emociones que se refirieren a la fortuna de los demás, (c) emociones basadas en prospectos, (d) emociones referidas a las atribuciones, (e) emociones de atracción hacia objetos o situaciones, (f) emociones referidas a combinaciones establecidas por el producto cartesiano de las emociones bienestar \times atribuciones. Sea $\Omega_\Psi = \{\varphi_1, \dots, \varphi_r\}$ el dominio de valores emocionales para un estado emocional Ψ , es decir, φ_i puede tomar valores como: alegría, regocijo malicioso por, temor, esperanza, orgullo, entre otras. Sea $m = \{x|x = (\Psi(\varphi),b)\}$ un mundo posible como lo sugiere Kripke (Bell, DeVidi & Solomon, 2001), donde b es una relación binaria sobre $\Psi(\varphi)$. Sea M el conjunto de los mundos posibles para un agente. Existe una función $g: \{\top, \perp\} \rightarrow m \in M$, donde \top representa verdadero y \perp falso.

Los mundos posibles representan el *mundo emocional* de cada agente. Estos mundos posibles permiten evaluar positivamente un argumento emocional y establecer la forma en que el conjunto de creencias se modificaría. En concreto, basados en el modelo de Ortony et al. (1988), existen tres clases de emociones que son generadas en el proceso de valoración (*appraisal*) de las consecuencias de un evento en cuanto a su conveniencia: (a) las emociones basadas en prospectos, (b) las emociones basadas en la fortuna de los demás y (c) las emociones de autobienestar. Ya que el evento a evaluar es la formación de un acuerdo con un miembro de un grupo de interés distinto al que el agente pertenece, estas emociones serán las que se deberán generar en el momento de hacer la evaluación.

La estructura formal establecida en la Definición 4 permite definir la estructura de evaluar un argumento. Esta se puede escribir como:

Definición 5: Sea una TAP, entonces un cualificador emocional ω_Ψ se define como: $\omega_\Psi \stackrel{\text{def}}{=} \exists x \in \Omega_\Psi \therefore \diamond \exists \{c, x(E)\} \not\rightarrow \{\top\}$, en donde $\Omega_\Psi = \{x|x = \Psi_\gamma^\alpha(\omega)\}$ y donde

$$\Psi_\gamma^\alpha(\omega) \stackrel{\text{def}}{=} \diamond \exists m \in M \therefore c \in m \wedge m(\Psi(\varphi)) = \{\top\}.$$

Lo anterior significa que el cualificador emocional es un elemento x que pertenece al dominio de valores emocionales para un estado emocional Ψ , tal que *posiblemente* (\diamond) existe una lista de creencias, que contienen a la creencia c en donde todos los elementos toman el valor de verdad \top (verdadero). Lo anterior se logra al definir el dominio emocional como un conjunto de elementos pertenecientes a Ψ_γ^α , el cual es un estado emocional γ de un agente α que es la posibilidad (\diamond) de la existencia de un mundo emocional m , tal que la creencia c , elemento de dicho mundo emocional, y el mundo emocional m , dependiente del estado emocional, tienen valor de verdad \top (verdadero).

Usando esta estructura es entonces posible ver cómo el cualificador emocional se convierte en el evaluador del argumento, haciéndolo verdadero gracias a un mundo emocional adecuado. Esta estructura formal implica que todo el proceso argumentativo sea definido a través de los mundos emocionales posibles. Se mostrará todo este proceso por medio de un ejemplo. Una vez más se tiene a los Agentes 1 y 2 que comparten el problema P, pero que pertenecen a grupos de interés diferentes. El proceso racional indicaría que el problema P no sería resuelto debido a que existe la regla de que no se coopera con miembros distintos a los del grupo. Sin embargo, el proceso de argumentación/negociación, que se permite en el sistema, lleva a que uno de los agentes genere un argumento k , el cual en el modelo CONSA se considera una demanda. Supóngase que el argumento k es establecido por el Agente 1. El proceso cognitivo que sucede en el Agente 2 comienza con la evaluación (*appraisal*) del argumento k . El argumento k tiene como objetivo modificar la creencia c que establece la no colaboración con miembros de distinto grupo de interés. En el Agente 2 existe un mundo posible m_1 en donde se tiene que:

1. Emoción/creencia de autobienestar. Valoración del evento de colaborar: alegría si resuelvo el problema P.

2. Emoción/creencia referida a la fortuna de otros. Suponer el valor de un evento que afecta a otras personas: apenado por romper la lealtad con los miembros del grupo.
3. Emoción/creencia basada en la prospección. Valoración del evento previsto: deseo de resolver el problema P.
4. Emoción/creencia basada en la atribución. Valoración de una acción de los agentes: indulgencia con respecto a colaborar con miembros distintos a los del grupo.
5. Emoción/creencia de atracción. Valoración del objeto: amor a la resolución de problemas.
6. Emociones/creencias de bienestar/atribuciones. Emociones compuestas: Alegría + indulgencia: gratificación por resolver el problema P.

En la Tabla 1 se presentan los valores de verdad para este mundo posible.

Dicho mundo posible lleva a una negativa a la colaboración. Lo que debe hacer el argumento k es cambiar los valores de verdad de las Emociones 5 y 6 a verdaderas. Esto es posible si existe un mundo posible m_2 con los valores de verdad presentados en la Tabla 2.

En este caso, para empezar, el mundo posible m_1 debe poder alcanzar el mundo posible m_2 . Supóngase que esto es cierto. La acción de k es invocar la emoción/creencia $c \in B^1$, en

donde c es verdadero en el caso de la Emoción/Creencia 6. Al ser entonces todas las creencias verdaderas, el deseo $d(B^1)$ relacionado con la resolución del problema P, sin importar las reglas del grupo, lleva a la intención I de colaborar; y por lo tanto, a la acción de colaborar y resolver el problema P.

Discusión

Se presentó un modelo cognitivo de construcción de acuerdos usando procesos argumentativos bajo contextos emocionales. Tanto el modelo como la necesaria formalización simbólica están basados en el marco teórico de la lógica modal. El modelo está inspirado en el modelo cognitivo sobre las emociones propuesto por Ortony et al. (1988) y el modelo sobre la construcción de consensos propuesto por Jung y Tambe (2001). Su formulación está basada en la lógica modal y parte de las ideas planteadas por Wooldridge (2000) respecto a los modelos cognitivos que guían el comportamiento en los sistemas multiagente. Dicha formulación permitió estructurar el complejo proceso de la argumentación dentro de un esquema de conflicto/negociación. Adicionalmente, el modelo del proceso argumentativo sugerido por Toulmin (2003), y retomado por Jung y Tambe (2001), para proponer el modelo cognitivo CONSA, fue el marco básico en este estudio para incorporar al modelo

Tabla 1
Valores de verdad del mundo posible m_1

Emoción 1	Emoción 2	Emoción 3	Emoción 4	Emoción 5	Emoción 6
V	V	V	V	F	F

Tabla 2
Valores de verdad del mundo posible m_2

Emoción 1	Emoción 2	Emoción 3	Emoción 4	Emoción 5	Emoción 6
V	V	V	V	V	V

cognitivo de emociones sugerido por Ortony et al. (1988). El mecanismo propuesto para incorporar las emociones dentro del modelo CONSA de Jung y Tambe (2001) se basó en que el argumento/demanda puede hacer que el mundo posible de un agente, el cual inhabilita la colaboración, sea sustituido por otro mundo posible en donde sí se puede realizar la colaboración.

Si bien el modelo aquí propuesto no es el primero que sigue la línea de investigación planteada por el modelo de Ortony et al. (i. e. Becker-Asano & Wachsmuth, 2010; Gebhard, 2005), se puede afirmar, siguiendo los argumentos esgrimidos por Sun (2008), que la modesta contribución de este trabajo se circunscribe a la formulación de un modelo teórico cognitivo sobre el papel de las emociones en las situaciones de conflicto/negociación, las cuales son comunes en la psicología organizacional y del trabajo.

Por otra parte, hay que resaltar que los modelos basados en el modelo de Ortony et al. (1988) tienen diferencias significativas con respecto al propuesto en este estudio. Dichas diferencias se encuentran tanto en el objetivo del modelo como en la formalización e implementación del mecanismo de evaluación/apreciación. Por ejemplo, en el modelo ALMA (*A Layered Model of Affect*) propuesto por Gebhard (2005), diseñado para darle realismo a personajes virtuales que interactúan con seres humanos, el formalismo de la evaluación/apreciación es realizado estableciendo reglas de tipo imperativo en las que las variables son previamente establecidas con valores numéricos y la implementación se realiza en un lenguaje estilo XML. Por su parte, el modelo aquí propuesto usa para su formalización la lógica modal y puede, por lo tanto, implementarse usando la programación lógica. Esto es, el modelo se puede implementar a través de un lenguaje declarativo que sirve para especificar qué problema resolver. El problema a resolver se describe especificando qué caracteriza a las posibles soluciones de dicho problema. Después se especifican las condiciones que satisfacen las

soluciones y, por último, se deducen las soluciones a partir de las condiciones. Los lenguajes que permiten hacer esto son LISP y Prolog.

Otro ejemplo es el modelo WASABI propuesto Becker-Asano y Wachsmuth (2010). En este las capacidades cognitivas del razonamiento de un humano virtual se combinan con una personificación simulada, al estilo de un personaje de videojuego, para lograr la simulación de las emociones primarias y secundarias. La idea del modelo WASABI es usar un *espacio emocional tridimensional*, llamado espacio PAD, el cual captura la teoría de Ortony et al. (1988) sobre las emociones secundarias, mientras que usa la teoría de Ekman (1992) sobre las emociones primarias que manifiestan a través de expresiones faciales. El objetivo del modelo WASABI es indagar problemas propios de la psicología del desarrollo (*developmental psychology*), mientras que el modelo propuesto aquí está orientado hacia la psicología de las organizaciones y del trabajo. Respecto a la formulación del llamado espacio PAD, este consiste en el cálculo de funciones de probabilidad que disparan las emociones codificadas en el modelo. Por su parte, el nuevo modelo aquí presentado no permite este comportamiento estocástico, pero modificaciones específicas de dicho modelo usando la teoría de la lógica de articulaciones (*fibring logics*), desarrollada por Gabbay (1999), permitiría tener efectos semejantes a un modelo estocástico. Con respecto a la implementación, el modelo propuesto en este estudio es realmente simple ya que el modelo WASABI interactúa con un entorno de simulación de un humano virtual llamado MAX, lo que lo hace superior a la implementación posible de este primero.

Una limitación de la propuesta estriba en la forma del argumento/demanda, la cual no fue explicitada y queda como pendiente para un trabajo futuro. Asimismo, otra limitación es que los ejemplos solo han sido dados en una relación entre dos agentes y se ha omitido la situación compleja en la que el equipo de trabajo puede

estar conformado por agentes dispuestos a colaborar con los miembros de otro grupo de interés, y entonces se tendría la tarea de convencer a los demás miembros de su grupo de interés.

Se quiere hacer énfasis en que las ideas vertidas en el modelo presentado en este trabajo y la notación lógica empleada son simples. Sin embargo, al tratar de capturar la esencia de los mecanismos cognitivos que llevan a una argumentación que trata de realizar cambios en la postura colaborativa entre dos agentes, basados los argumentos en lograr generar emociones positivas hacia la colaboración, se encontró que existe una gran variedad de interrelaciones múltiples entre los esquemas emocionales y sus valores de verdad, la forma de plantear la red de mundos posibles y la forma que deben tener los argumentos para poder generar en los agentes una búsqueda interna de mundos posibles que armonice con el argumento.

Para los autores, el trabajo futuro será seguir desarrollando el modelo lógico formal que permita establecer mecanismos de aprendizaje de los actores y la modificación de las emociones de manera dinámica ante cambios en el medio ambiente, de tal forma que se pueda construir un sistema multiagente que pueda, a través de procesos de argumentación que evoquen emociones, tratar de llegar a acuerdos sobre problemas comunes, bajo contextos de rivalidades grupales, tal y como el modelo aquí presentado lo ha sugerido.

Referencias

- Becker-Asano, C. & Wachsmuth, I. (2010). Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems*, 20, 32-49.
- Bell, J. L., DeVidi, D. & Solomon. G. (2001). *Logical options: An introduction to classical and alternative logics*. Canada: Broadview Press.
- Bratman, M. E. (1992). Practical reasoning and acceptance in a context. *Mind, New Series*, 101 (401), 1-15.
- Damasio, A. (2005). *Descartes' error: Emotion, reason, and the human brain*. New York: Penguin Putnan.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6 (3-4), 169-200.
- Elster, J. (1996). Rationality and emotions. *The Economic Journal*, 106 (438), 1386-1397.
- Elster, J. (1997). *Egonomics: análisis de la interacción entre racionalidad, emoción, preferencias y normas sociales en la economía de la acción individual y sus desviaciones*. Barcelona: Gedisa.
- Fineman, S. (1996). Emotion and organizing. En S. R. Clegg, C. Hardy & W. R. Nord (Eds.). *Handbook of Organization Studies* (pp. 543-564). London: SAGE.
- Frijda, N. H. & Mesquita, B. (2000). Beliefs through emotions. En N. H. Frijda, A. S. R. Manstead & S. Bem (Eds.). *Emotions and beliefs: How feelings influence thoughts* (pp. 45-77). Cambridge, U.K.: Cambridge University Press.
- Gabbay, D. M. (1999). *Fibring logics*. Oxford: Claredon Press.
- Gaviria-Martínez, C. D. & Corredor-Aristizábal, J. (2011). Coordinación de evidencias, explicaciones y objetivos pragmáticos en la argumentación. *Revista Colombiana de Psicología*, 20 (1), 43-56
- Gebhard, P. (2005). ALMA: A layered model of affect. En M. Pechoucek, D. Steiner & S. Thompson (Eds.), *Proceedings of the fourth international joint conference on autonomous agents and multiagent systems (AAMAS'05)* (pp. 29-36). New York: ACM.
- Gilbert, M. A. (1994). Multi-modal argumentation. *Philosophy of the Social Sciences*, 24 (2), 159-177.
- Gilbert, M. A. (1995). Emotional argumentation, or, why do argumentation theorists argue with their mates? En F. H. van Eemeren, R. Grootendorst, J. A. Blair & C. A. Willard (Eds.), *Analysis and evaluation: Proceedings of the third ISSA conference on argumentation* (Vol. II). Recuperado de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.127.6403&rep=rep1&type=pdf>.
- Hample, D. (1992). A third perspective on argument. En W. L. Benoit, D. Hample & P. J. Benoit (Eds.), *Readings in argumentation* (pp. 91-116). Berlin: Foris Publications.

- Howard, N. (1993). The role of emotions in multi-organizational decision-making. *The Journal of the Operational Research Society*, 44 (6), 613-623.
- Jung, H. & Tambe, M. (2001). Conflicts in agent teams. En C. Tessier, L. Chaudron & H. J. Müller (Eds.), *Conflicting agents: Conflict management in multi-agent systems* (pp. 153-167). Norwell, MA: Kluwer.
- Kneupper, C. W. (1978). Teaching argument: An introduction to the Toulmin model. *College Composition and Communication*, 29 (3), 237-241.
- Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *The American Economic Review*, 90 (2), 426-432.
- Marsella, S., Gratch, J. & Petta, P. (2010). Computational models of emotion. En K. R. Scherer, T. Bänziger & E. Roesch (Eds.), *A blueprint for affective computing* (pp. 21-45). Oxford: Oxford University Press.
- Mellers, B. A. & McGraw, A. P. (2001). Anticipated emotions as guides to choice. *Current Directions in Psychological Science*, 10 (6), 210-214.
- Miller, A. (1983). The influence of personal biases on environmental problem-solving. *Journal of Environmental Management*, 17 (2), 133-142.
- Naqvi, N., Shiv, B. & Bechara, A. (2006). The role of emotion in decision making: A cognitive neuroscience perspective. *Current Directions in Psychological Science*, 15 (5), 260-264.
- Newell, A. (1981). The knowledge level. *Artificial Intelligence*, 18 (1), 87-127.
- O'Keefe, D. J. (1977). Two concepts of argument. *Journal of the American Forensic Association*, XIII (13), 121-128.
- O'Rourke, P. & Ortony, A. (1994). Explaining emotions. *Cognitive Science*, 18 (2), 283-323.
- Ortony, A., Clore, G. L. & Collins, A. (1988). *The Cognitive structure of emotions*. Cambridge, U.K.: Cambridge University Press.
- Perelman, C. & Olbrechts-Tyteca, L. (1994). *Tratado de la argumentación: la nueva retórica*. Madrid: Editorial Gredos.
- Rao, A. S. & Georgeff, M. P. (1995). BDI agents: From theory to practice. En V. Lesser & L. Gasser (Eds.), *Proceedings of the first international conference on multiagent systems (ICMAS-95)* (pp. 312-319). Cambridge, MA: MIT Press.
- Sun, R. (2008). *The Cambridge handbook of computational psychology*. Cambridge, U.K.: Cambridge University Press.
- Thagard, P. & Kroon, F. W. (2006). Emotional consensus in group decision making. *Mind & Society*, 5 (1), 85-104.
- Thoits, P. A. (1989). Sociology of emotions. *Annual Review of Sociology*, 15, 317-342.
- Toulmin, S. E. (2003). *The uses of argument*. New York: Cambridge University Press.
- Trappl, R. & Payr, S. (2002). Emotions: From brain research to computer game development. En R. Trappl, P. Petta & S. Payr (Eds.), *Emotions in humans and artifact* (pp. 115-148). Cambridge, MA: MIT Press.
- Wenzel, J. (1992). Perspectives on argumentation. En W. L. Benoit, D. Hample & P. J. Benoit (Eds.), *Readings in argumentation* (pp. 121-143). Berlin: Foris Publications
- Wooldridge, M. (2000). *Reasoning about rational agents*. Cambridge, MA: MIT Press.
- Wooldridge, M. (2002). *An introduction to multi-agent system*. Chichester: John Wiley & Sons.