UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

LUCAS GONÇALVES DE MOURA LEITE

RACIOCÍNIO SOBRE CRENÇAS UTILIZANDO A LÓGICA DE CONFIANÇA EM TÓPICOS

FORTALEZA
2016

LUCAS GONÇALVES DE MOURA LEITE

RACIOCÍNIO SOBRE CRENÇAS UTILIZANDO A LÓGICA DE CONFIANÇA EM TÓPICOS

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Departamento de Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Ciência da Computação.

Área de concentração: Lógica e Inteligência Artificial

Orientador: Prof. Dr. João Fernando Lima Alcântara

FORTALEZA

2016

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

L554r Leite, Lucas Gonçalves de Moura.
Raciocínio sobre crenças utilizando a lógica de confiança em tópicos / Lucas Gonçalves de Moura Leite. – 2016.
66 f. : il.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2016.
Orientação: Prof. Dr. João Fernando Lima Alcântara.

1. Representação do conhecimento. 2. Argumentação. 3. Confiança e desconfiança. I. Título.
CDD 005

LUCAS GONÇALVES DE MOURA LEITE

RACIOCÍNIO SOBRE CRENÇAS UTILIZANDO A LÓGICA DE CONFIANÇA EM TÓPICOS

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Departamento de Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Ciência da Computação.

Aprovada em: 30 de novembro de 2016

BANCA EXAMINADORA

Prof. Dr. João Fernando Lima Alcântara
Universidade Federal do Ceará (UFC)
Orientador

Prof. Dr. Pedro Porfirio Muniz Farias
Universidade de Fortaleza - UNIFOR

Profa. Dra. Ana Teresa de Castro Martins
Universidade Federal do Ceará - UFC

Prof. Dr. Davi Romero de Vasconcelos
Universidade Federal do Ceará - UFC

A todos aqueles que contribuíram direta e indiretamente para que eu superasse mais esse desafio.

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter-me dado a força necessária para concluir mais este desafio pessoal e profissional;

Agradeço a Thayane por todo amor e apoio que me concedeu durante todo este trabalho, sofrendo comigo nos momentos difíceis e se alegrando a cada vitória;

Agradeço aos meus pais, Osvaldo e Maria José, por todo amor e incentivo durante toda a elaboração deste trabalho;

Agradeço em especial ao meu orientador e amigo, Prof. João Fernando, pela confiança, dedicação, orientação, e principalmente pela paciência diante de minhas dificuldades técnicas e pessoais;

Agradeço aos professores Pedro Porfirio, Davi Romero e Ana Teresa pela participação na minha banca de defesa de mestrado;

Agradeço aos amigos do LogIA por toda ajuda e companheirismo durante a maior parte deste trabalho;

Agradeço ao MDCC e ao LSBD por ter fornecido uma estrutura bastante satisfatória para minha pesquisa, como também pelo apoio financeiro para participação em congressos científicos;

Agradeço a todos os amigos do LSBD pelo apoio e experiências trocadas;

Agradeço a todos os meus amigos pela compreensão em muitas vezes não poder estar presente nos momentos importantes, devido ao tempo dedicado a este trabalho;

Agradeço aos meus colegas de mestrado e doutorado da UFC por compartilharem conhecimento e experiências na vida acadêmica;

Por fim, agradeço ainda àquelas pessoas que contribuíram de uma forma indireta para a realização deste trabalho.

*"Eu sei que o meu trabalho é uma gota no oceano,
mas sem ele, o oceano seria menor."*
(*Santa Teresa de Calcutá*)

RESUMO

A confiança é considerada um dos conceitos vitais de sistemas multiagentes, em que agentes autônomos possuem interesses próprios, que podem divergir dos interesses de um grupo ao qual pertencem. Para que um grupo seja capaz de tomar uma decisão, os agentes envolvidos devem ser capazes de trocar informações, de modo que o ponto de vista do grupo esteja de acordo com as informações levantadas por todos os seus participantes. Eventualmente, agentes podem informar inverdades, seja por ignorância ou com o intuito de obter alguma vantagem sobre o grupo ao qual pertencem. Com base nessa situação, os agentes devem decidir se uma informação gerada por um outro agente é confiável ou não, de modo que o ponto de vista do grupo seja suporte de uma decisão que reflita suas necessidades. Neste trabalho, propusemos uma extensão de uma lógica que permite formalizar e raciocinar sobre crenças que envolvem o conceito de confiança entre agentes com relação a uma informação. Em nossa extensão introduzimos dois novos recursos na lógica: (i) tratar a confiança entre agentes sobre tópicos e (ii) tratar a desconfiança entre agentes sobre fórmulas e tópicos.

Keywords: Representação do conhecimento. Argumentação. Confiança e desconfiança.

ABSTRACT

Trust is considered one of the vital concepts of multi-agent systems, in which autonomous agents have their own interests, which may differ from the interests of a group which they belong. In order to make a decision, the parties involved must be able to exchange information, so that the group's point of view is in accordance with the information collected by all participants. Eventually, agents can tell untruths, either through ignorance or in order to get an advantage on the group to which they belong. Based on this situation, agents must decide whether information generated by another agent is reliable or not, so that the group's point of view supports a decision that reflects its needs. In this work, we proposed an extension of a logic that allows to formalize and reason about beliefs that involve the concept of trust between agents with respect to information. In our extension, we introduced two new features in this logic: (i) the trust between agents on topics and (ii) the distrust between agents on formulas and topics.

Keywords: Knowledge representation. Argumentation. Trust and distrust.

LISTA DE ILUSTRAÇÕES

SUMÁRIO

1 INTRODUÇÃO

A confiança é considerada um dos conceitos vitais de sistemas multiagentes, em que agentes autônomos possuem interesses próprios, que podem divergir dos interesses de um grupo ao qual pertencem (SABATER; SIERRA, 2005). Para que um grupo seja capaz de tomar uma decisão, os agentes envolvidos devem ser capazes de trocar informações, de modo que o ponto de vista do grupo esteja de acordo com as informações levantadas por todos os seus participantes. Eventualmente, agentes podem informar inverdades, seja por ignorância ou com o intuito de obter alguma vantagem sobre o grupo ao qual pertencem. Com base nessa situação, os agentes devem decidir se uma informação gerada por um outro agente é confiável ou não, de modo que o ponto de vista do grupo seja suporte de uma decisão que reflita suas necessidades.

1.1 Motivação e Formulação do Problema

A situação retratada no exemplo abaixo mostra como a confiança entre os agentes pode influenciar uma decisão.

Uma equipe médica, composta por Alexandre, Beatriz e Carlos é escalada para diagnosticar Patrícia. Em uma reunião, os médicos decidem apresentar as informações que já obtiveram sobre a paciente a fim de chegarem a um consenso sobre qual doença a afeta e, consequentemente, decidir qual tratamento irão administrar. Alexandre afirma que a paciente sofre de meningite, pois possui os sintomas de febre alta, forte dor de cabeça, fotossensibilidade e rachaduras na pele. Prontamente, Carlos discorda argumentando que a paciente não possui rachaduras na pele, de acordo com os primeiros exames realizados nela. A afirmação de Carlos coloca em xeque o diagnóstico de Alexandre, pois a ausência de fissuras na pele poderia indicar que a paciente estaria sofrendo de outras patologias, visto que febre, dor de cabeça e fotossensibilidade também são sintomas de várias outras doenças. Porém, antes da reunião, Alexandre e Beatriz haviam conversado e Alexandre a informou que nos últimos exames a paciente havia apresentado rachaduras características de quem possui meningite. Beatriz confia na informação de Alexandre, pois sabe que, por ser um dermatologista, ele tem conhecimento sobre as diversas patologias que afetam a pele. Por conseguinte, ela acredita na informação que seu colega a passou. Por outro lado, Beatriz também desconfia da autenticidade da informação de Carlos, visto que, Patrícia foi casada com ele e Beatriz sabe que ele está buscando se reaproximar dela. Com essas informações, Beatriz também decide dizer que Patrícia tem meningite por apresentar febre alta, forte dor de cabeça, fotossensibilidade e rachaduras na pele.

No cotidiano, temos que tomar decisões baseadas em informações disponíveis. Em Ciência da Computação, buscamos construir aplicações capazes de utilizar essas informações para tomar decisões de maneira autônoma. Diversas aplicações podem ser encontradas na literatura sobre esse assunto, tais como gerenciamento de energia (ZHAO et al., 2013), certificação de construções sustentáveis (MEDINECKIENE et al., 2015), diagnósticos de câncer de mama (WU et al., 1993), etc. No exemplo apresentado no início do capítulo, vemos que um grupo de médicos deseja descobrir qual doença afeta uma paciente a partir de um conjunto de informações disponíveis. A partir do diagnóstico, uma decisão será tomada em relação ao tratamento que será conduzido pela equipe médica. Repare que para concluir sobre qual enfermidade afeta a paciente, não podemos simplesmente considerar todas as informações do conjunto, uma vez que existem informações conflitantes como o fato de Alexandre afirmar que Patrícia possui fissuras na pele e Carlos falar exatamente o contrário. Também não podemos desconsiderar essas informações contraditórias, visto que poderíamos perder informações importantes para uma tomada de decisão condizente com a realidade do grupo.

Na literatura, encontramos o trabalho (AMGOUD; DEMOLOMBE, 2014), que apresenta um sistema capaz de dado um conjunto de agentes com informações potencialmente contraditórias, encontrar um ponto de vista desse grupo a fim de que uma decisão possa ser tomada. Para tanto, esse sistema utiliza-se de uma instanciação do Framework de Argumentação de Dung (DUNG, 1995) para raciocinar sobre essas informações. O primeiro diferencial de (AMGOUD; DEMOLOMBE, 2014) para os demais trabalhos presentes na literatura, tais como, (TANG et al., 2012; STRANDERS et al., 2007; LIAU, 2003), é que ele instância os argumentos utilizando uma lógica modal capaz de modelar as crenças e as trocas de informações entre os agentes e, a partir disso, define o que seria uma relação de confiança ternária entre dois agentes sobre uma fórmula, proporcionando uma modelagem de confiança mais próxima da realidade, visto que a confiança entre agentes nem sempre é total, ou seja, um agente pode ser confiável em relação a alguns assuntos enquanto não merece confiança em outros. O segundo diferencial vem por meio dessa relação de confiança. A maioria dos trabalhos que tratam da confiança entre agentes utiliza uma noção de confiança absoluta, tais como (TANG et al., 2012; STRANDERS et al., 2007). Nesse trabalho, são considerados seis tipos de confiança: *Sinceridade, Validade, Completude, Cooperatividade, Competência e Vigilância*. Cada um desses tipos de confiança possui suas próprias características, conforme definido inicialmente em (DEMOLOMBE, 2001; DEMOLOMBE, 2004). Essa divisão da confiança em seis tipos torna a lógica mais expressiva, uma vez que podemos diferenciar um agente competente, que está à par das verdades do mundo, de um agente sincero, que gera informações conforme suas crenças, por exemplo.

1.2 Contribuições

Apesar de (AMGOUD; DEMOLOMBE, 2014) modelar a confiança entre agentes com relação a uma fórmula, iremos estender esse trabalho nesta Dissertação com duas contribuições. No exemplo mostrado no início do capítulo, Beatriz toma uma posição no diagnóstico de Patrícia baseando-se em sua confiança em Alexandre, que é um dermatologista, e em sua desconfiança em Carlos, acreditando que esse possui segundas intenções sobre a avaliação da paciente. A lógica apresentada em (AMGOUD; DEMOLOMBE, 2014) não modela a confiança entre agentes em relação a um a um tópico de fórmulas, como é o caso da confiança de Beatriz em Alexandre sobre a dermatologia. Dessa maneira, não poderíamos expressar a confiança de um agente em alguém que ele considera como um especialista, isto é, alguém que merece confiança em relação a um conjunto infinito de fórmulas associadas ao tópico de sua especialidade. Além disso, a lógica considerada não modela a desconfiança entre agentes com relação a fórmulas e tópicos, como é o caso da desconfiança de Beatriz em Carlos sobre a validade de sua informação.

Nossas contribuições nesta Dissertação de Mestrado consistem em

- Estender (AMGOUD; DEMOLOMBE, 2014) utilizando a lógica definida em (DEMOLOMBE; JONES, 1995), que trata da Teoria dos Tópicos, que relaciona infinitas fórmulas a um determinado contexto.

- Introduzir a Desconfiança entre agentes com relação a fórmulas e tópicos, utilizando a noção de desconfiança apresentada em (MARSH; DIBBEN, 2005).

Com a primeira contribuição, podemos representar a confiança entre agentes em um conjunto de fórmulas relacionadas a um tópico. Com relação à segunda contribuição, (MARSH; DIBBEN, 2005) afirma que a desconfiança não deve ser tratada apenas como a ausência de confiança, mas como a medida na qual um indivíduo acredita que outro desconhece as verdades do mundo ou simplesmente não revela o que suas reais crenças pregam, buscando obter algum ganho com isso. Com essas duas contribuições, construímos a Lógica de Confiança em Tópicos, que será utilizada para modelar as crenças de um conjunto de agentes, levando em consideração a confiança e a desconfiança entre agentes em relação a fórmulas e tópicos, com o intuito de encontrar as fórmulas que melhor representam o conjunto de crenças desses agentes, possibilitando uma tomada de decisão condizente com o ponto de vista do grupo.

1.3 Estrutura da Dissertação

O restante deste trabalho é dividido da seguinte maneira: No Capítulo 2 apresentamos os conceitos e definições fundamentais de Confiança, Teoria da Argumentação e Teoria dos Tópicos. No Capítulo 3 mostraremos a principal contribuição desta Dissertação de Mestrado, que é a Lógica de Confiança em Tópicos, que será utilizada no lugar da lógica definida em (AMGOUD; DEMOLOMBE, 2014) para o raciocínio sobre as crenças de agentes, levando em consideração suas relações de confiança e desconfiança. Além disso, iremos utilizar como caso de estudo para a nossa contribuição a obra *Otelo, o Mouro de Veneza*. Por fim, no Capítulo 4, concluímos esta Dissertação apresentando quais as contribuições que faltam para a conclusão deste trabalho e também um conjunto de ideias que podem ser utilizadas para estender esse trabalho.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo, iremos abordar brevemente os conceitos fundamentais para a construção deste trabalho. Tais conceitos envolvem a Modelagem Formal de Relações de Confiança (DEMOLOMBE, 2001; DEMOLOMBE, 2004), Teoria da Argumentação (DUNG, 1995) e Teoria dos Tópicos (DEMOLOMBE; JONES, 1995). Além disso, iremos também discorrer sobre alguns dos trabalhos relacionados a esses temas. A seguir, iremos apresentar uma variação do exemplo mostrado na seção 1.1 enfatizando a relação de confiança entre agentes e suas consequências. Mais adiante iremos modificar este exemplo para incluir a relação entre confiança e tópicos.

Exemplo 2.1. *Alexandre, Beatriz e Carlos são médicos que estão buscando chegar a uma conclusão sobre a enfermidade de Patrícia. Em uma reunião, os médicos decidem apresentar suas conclusões sobre qual seria a doença que afeta Patrícia. Alexandre afirma que a paciente sofre de meningite, pois possui os sintomas de febre alta, forte dor de cabeça, fotossensibilidade e rachaduras na pele. Prontamente, Carlos discorda argumentando que a paciente não possui rachaduras na pele, de acordo com os primeiros exames realizados nela. A afirmação de Carlos coloca em xeque o diagnóstico de Alexandre, pois a ausência de fissuras na pele poderia indicar que a paciente estaria sofrendo de outras patologias, visto que febre, dor de cabeça e fotossensibilidade também são sintomas de várias outras doenças. Porém, antes da reunião, Alexandre e Beatriz haviam conversado e Alexandre a informou que nos últimos exames a paciente havia apresentado rachaduras na pele. Beatriz confia na sinceridade de Alexandre com relação ao seu diagnóstico, pois o havia visto examinando a paciente. Por conseguinte, ela acredita na informação que seu colega a passou. Ela, então, informa na reunião que segundo os últimos exames, a paciente tinha fissuras na pele. Assim, o grupo decide que a paciente realmente sofre de meningite.*

O exemplo acima mostra um grupo chegando a uma conclusão a partir de informações contraditórias. Para tal, note a discordância entre Alexandre e Carlos sobre a presença ou não de rachaduras na pele de Patrícia. Nesse contexto, percebemos que o ponto chave para a conclusão do diagnóstico da paciente foi a confiança de Beatriz em Alexandre, levando-a a crer na presença de rachaduras. A área que estuda esse tipo de problema chama-se Modelagem de Confiança, que vem obtendo destaque desde a última década.

Na literatura, boa parte dos trabalhos assumem que os envolvidos em um diálogo são honestos e sempre irão relatar suas crenças de maneira fidedigna, como em (KARUNATILLAKE et al., 2005; DEVEREUX; REED, 2009). Ocorre que em aplicações mais realistas, um dos envolvidos pode buscar sobressair-se sobre o outro utilizando-se de uma informação inverídica para obter alguma vantagem. Por exemplo, em uma guerra, um comandante deve confiar em seu aliado para que um ataque pelos flancos seja eficaz. Para que os aliados possam realizar seu ataque, ambos devem decidir se acreditam ou não na competência do outro em relação ao *timing* da investida. Tendo em vista que uma das maneiras de acreditar em uma informação recebida é confiar em sua fonte sobre essa informação, os aliados devem ter confiança mútua em relação aos movimentos realizados. A confiança pode ter diversas origens, como bom retrospecto em diálogos anteriores, documentos ou indicações.

Neste capítulo iremos apresentar a Lógica de Confiança (AMGOUD; DEMOLOMBE, 2014), que se diferencia de outras modelagens que utilizam a confiança para tomada de decisão por trabalhar com diferentes tipos de confiança, conforme (DEMOLOMBE, 2001; DEMOLOMBE, 2004), e ser uma relação de confiança ternária, envolvendo dois indivíduos e uma fórmula representando uma informação a qual a confiança se refere, pois geralmente a confiança não é total, ou seja, pode existir somente em relação a fatos isolados. Além disso, Lógica de Confiança utiliza os Frameworks de Argumentação de Dung (DUNG, 1995) para definir o ponto de vista do grupo a partir das crenças de cada um dos membros desse grupo. Na Seção 2.4, iremos também explorar a lógica definida em (DEMOLOMBE; JONES, 1995) que trata de tópicos de fórmulas. As definições apresentadas neste capítulo servirão de alicerce no Capítulo 3 para a construção de uma lógica geral o suficiente para lidar com formas mais complexas de confiança.

2.1 Confiança entre agentes

Nesta seção, iremos definir a sintaxe e a axiomática da Lógica de Confiança $\mathcal{LC}$. Além disso, iremos também apresentar as definições dos seis tipos de confiança introduzidos por Demolombe em (DEMOLOMBE; JONES, 1995).

2.1.1 Sintaxe de $\mathcal{LC}$

Esta subseção irá abordar o formalismo utilizado no restante do capítulo para modelar crenças e fontes de informação. As primitivas sintáticas da linguagem $\mathcal{L}$ da $\mathcal{LC}$ são definidas a seguir:

Definição 2.2 (Primitivas Sintáticas de $\mathcal{LC}$)**.** *(AMGOUD; DEMOLOMBE, 2014) As primitivas sintáticas de $\mathcal{L}$ são os seguintes conjuntos:*

- *$\mathcal{AT}$: conjunto não-vazio de proposições atômicas denotadas por $p_1, p_2, p_3, \ldots$*
- *$\mathcal{AG}$: conjunto finito não-vazio de agentes denotados por $i, j, k, \ldots$*

Agora seguimos para a definição da linguagem propriamente dita:

Definição 2.3 (Linguagem $\mathcal{L}$)**.** *(AMGOUD; DEMOLOMBE, 2014) Uma linguagem $\mathcal{L}$ definida sobre as Primitivas Sintáticas da definição 2.2 é o conjunto de fórmulas geradas pela seguinte BNF:*

$$\phi ::= p \mid \neg\phi \mid \phi \vee \phi \mid Bel_i\phi \mid Inf_{i,j}\phi$$

em que $p \in \mathcal{AT}$ e $i, j \in \mathcal{AG}$. Os conectivos $\neg$ e $\vee$ são definidos como usual. O significado intuitivo dos operadores modais Bel e Inf são

- *$Bel_i\phi$: o agente i acredita na fórmula ϕ.*
- *$Inf_{i,j}\phi$: o agente i informa ao agente j a fórmula ϕ.*

A seguir, iremos mostrar algumas definições que serão utilizadas para descrever a modelagem de uma situação. A primeira dessas definições é a de Base de Crenças de um agente, que, como o nome sugere, é um conjunto com todas crenças de um agente i.

Definição 2.4 (Base de Crenças)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG} \rangle$ as primitivas sintáticas de $\mathcal{LC}$ e $\mathcal{L}$ a linguagem definida sobre essas primitivas. Dizemos que a base de crenças de um agente $i \in \mathcal{AG}$ em $\mathcal{L}$ é um conjunto do tipo*

$$\mathcal{K}_i \subseteq \{Bel_i\phi \mid \phi \in \mathcal{L}\}.$$

A partir da Base de Crenças de um agente, iremos apresentar a Descrição das Crenças de um conjunto de agentes, que é um conjunto no qual seus elementos são as Bases de Crenças dos agentes envolvidos. Sua formalização é dada a seguir:

Definição 2.5 (Descrição das Crenças)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG} \rangle$ as primitivas sintáticas de $\mathcal{LC}$ e $\mathcal{L}$ a linguagem definida sobre essas primitivas. Uma descrição das crenças dos agentes $i, j, k \ldots \in \mathcal{AG}$ é uma tupla denotada por $\langle \mathcal{K}_i, \mathcal{K}_j, \mathcal{K}_k \ldots \rangle$ em que $\mathcal{K}_i, \mathcal{K}_j, \mathcal{K}_k \ldots$ são, respectivamente, as bases de crenças dos agentes $i, j, k \ldots$ em $\mathcal{L}$.*

Por fim, definimos a Descrição de Mundo de um conjunto de agentes, que é composta pela Descrição das Crenças desses agentes e por fórmulas que representam trocas de informações, denotadas pelo operador modal Inf.

Definição 2.6 (Descrição de Mundo)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG} \rangle$ as primitivas sintáticas de $\mathcal{LC}$ e $\mathcal{L}$ a linguagem definida sobre essas primitivas. Uma descrição de mundo dos agentes $i, j, k \ldots \in \mathcal{AG}$ em $\mathcal{L}$ é uma dupla denotada por $\langle \mathcal{K}, \mathcal{AC} \rangle$, em que $\mathcal{K}$ é uma descrição das crenças dos agentes $i, j, k \ldots \in \mathcal{AG}$ em $\mathcal{L}$ e $\mathcal{AC} \subseteq \{Inf_{i,j}\phi \mid \phi \in \mathcal{L}\}$.*

A partir deste ponto, quando não houver confusão, iremos nos referir a uma descrição de mundo $\langle \mathcal{K}, \mathcal{AC} \rangle$ dos agentes $i, j, k \ldots \in \mathcal{AG}$ em $\mathcal{L}$ como a descrição definida sobre a Descrição das Crenças $\langle \mathcal{K}_i, \mathcal{K}_j, \mathcal{K}_k \ldots \rangle$ e sobre o conjunto $\mathcal{AC} \subseteq \{Inf_{i,j}\phi \mid \phi \in \mathcal{L}\}$, que tem sua base nas primitivas sintáticas de $\mathcal{L}$, representadas pela tupla $\langle \mathcal{AT}, \mathcal{AG} \rangle$.

Exemplo 2.7. *Sejam a, b e c agentes. A Descrição de Mundo $DM = \langle\langle\{Bel_a f,\ Bel_a s,\ Bel_a d,\ Bel_a m\},\ \{\ \},\ \{Bel_c\neg f,\ Bel_c s,\ Bel_c d,\ Bel_c\neg m\}\rangle,\ \{Inf_{a,b} f,\ Inf_{c,b}\neg f\}\rangle$ retrata o exemplo 2.1, sendo a representando Alexandre; b, Beatriz; c, Carlos; f representando a frase "apresenta fissuras na pele"; s, "apresenta fotossensibilidade"; d, "apresenta fortes dores de cabeça" e m, "a paciente tem meningite".*

Repare que a princípio, as crenças de Beatriz são vazias, pois a lógica ainda não permite a modelagem da confiança entre agentes sobre fórmulas.

É válido ressaltar que $\mathcal{LC}$ não se propõe a avaliar uma fórmula quanto a sua veracidade, mas apenas representar as crenças dos agentes.

Tendo definido a sintaxe da $\mathcal{LC}$, iremos apresentar na seção seguinte a sua axiomática.

2.1.2 Axiomática de $\mathcal{LC}$

Nesta subseção, iremos mostrar os axiomas e regras de inferência de $\mathcal{LC}$, que são os mesmos da Lógica Multimodal apresentada em (CHELLAS, 1980) em conjunto com os do Cálculo Proposicional Clássico.

Antes de apresentar os axiomas e regras de inferência de $\mathcal{LC}$, iremos expor a construção dos conectivos $\rightarrow$ e $\wedge$, que representam respectivamente a implicação e a conjunção, e a definição de sequente. Note que podemos inferir os conectivos $\rightarrow$ e $\wedge$ a partir dos conectivos $\vee$ e $\neg$, sendo $(\phi \rightarrow \psi)$ equivalente a $(\neg\phi \vee \psi)$ e $(\phi \wedge \psi)$ equivalente a $\neg(\neg\phi \vee \neg\psi)$. Um sequente, denotado por $\Gamma \vdash \phi$, é formado por um conjunto de fórmulas $\Gamma \subseteq \mathcal{L}$, chamado de conjunto de premissas, uma fórmula $\phi \in \mathcal{L}$, chamada de conclusão, e um símbolo $\vdash$, chamado de consequência.

Definição 2.8 (Consequência Sintática)**.** *(MENDELSON, 2009) Dizemos que ϕ é consequência de Γ se e somente se existe uma sequência finita de fórmulas $\langle \phi_1, \phi_2, \phi_3, \ldots, \phi_k \rangle$ tal que $\phi = \phi_k$ e cada ϕ_i ou é um axioma, ou é membro de Γ ou é resultado de uma regra de inferência sobre alguma outra fórmula que ocorreu anteriormente na sequência.*

Com a definição de Consequência Sintática, iremos mostrar a axiomática de $\mathcal{LC}$.

Axiomas e Regra de Inferência da Lógica Proposicional: Sistema Axiomático de Frege

- (P1) $A \rightarrow (B \rightarrow A)$
- (P2) $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$
- (P3) $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$
- (P4) $\neg\neg A \rightarrow A$
- (P5) $A \rightarrow \neg\neg A$
- (PR) $A, (A \rightarrow B) \vdash B$

Axiomas e Regra de Inferência para o Operador Modal Bel

- (K) $Bel_i(\phi \rightarrow \psi) \rightarrow (Bel_i\phi \rightarrow Bel_i\psi)$
- (D) $\neg(Bel_i\phi \wedge Bel_i\neg\phi)$
- (Nec) Se $\vdash \phi$, então $\vdash Bel_i\phi$

Axiomas e Regra de Inferência para o Operador Modal Inf

- (EQV) Se $\vdash \phi \leftrightarrow \psi$, então $\vdash Inf_{j,i}\phi \leftrightarrow Inf_{j,i}\psi$
- (CONJ) $Inf_{j,i}\phi \wedge Inf_{j,i}\psi \rightarrow Inf_{j,i}(\phi \wedge \psi)$
- (OBS) $Inf_{j,i}\phi \rightarrow Bel_i Inf_{j,i}\phi$
- (OBS') $\neg Inf_{j,i}\phi \rightarrow Bel_i \neg Inf_{j,i}\phi$

Os axiomas (P1), (P2), (P3), (P4) e a regra de inferência (PR) pertencem à lógica proposicional, como apresentado por Russel-Bernays.

O axioma (K) garante que o agente i pode aplicar o *modus ponens* para inferir consequências a partir de suas crenças em implicações, o axioma (D) garante que as crenças do agente i não são inconsistentes e (Nec) garante que i acredita nas verdades lógicas.

O axioma (EQV) garante que dadas duas informações ϕ e ψ equivalentes, se um agente j informa i sobre ϕ, então j também informou i sobre ψ. Por exemplo, se as informações que João está em sua sala e que João está trabalhando tem o mesmo significado, informar que João está em sua sala é o mesmo que informar que João está trabalhando. O axioma (CONJ) tem como objetivo garantir que, se um agente j informa i sobre ϕ e também informa sobre ψ, o agente j informou i sobre $\phi \wedge \psi$. Por exemplo, se João informa a Maria que o almoço está na mesa e também informa que o suco está na mesa então João informou a Maria que o almoço e o suco estão na mesa. O axioma (OBS) define que a comunicação entre os agentes é perfeita. Se um agente j informa a i a mensagem ϕ, i tem conhecimento que recebeu essa mensagem. No caso do axioma (OBS'), temos a mesma situação de (OBS), porém tratando o não envio das mensagens. Se j não informou a i sobre ϕ, i tem conhecimento que não recebeu esta mensagem.

2.1.3 *Confiança em fontes de informação*

Em uma tomada de decisão, um agente pertencente a um grupo deve decidir se considera relevante ou não uma informação obtida através de outro membro do grupo. Nesta subseção iremos tratar a questão de como um agente reage a informações recebidas de agentes nos quais ele confia. A confiança a ser considerada aqui é de natureza ternária, ou seja, é a confiança de um agente em outro com relação a uma fórmula, pois um agente pode confiar em outro com relação a algumas fórmulas e não em outras. Por exemplo, um agente pode ocultar informações que o prejudiquem, mas pode ser sincero em outras que o tragam algum tipo de vantagem. Para tanto, consideramos o trabalho desenvolvido por Demolombe em (DEMOLOMBE, 2001; DEMOLOMBE, 2004), que definiu seis diferentes tipos de confiança. Com base nesses diferentes tipos de confiança, serão mostradas abaixo as reações do agente receptor em relação às fontes nas quais ele confia.

- **Sinceridade:** A sinceridade é o tipo de confiança que define a crença de alguém passar uma informação em que realmente acredita. Por exemplo, o fato de João confiar na sinceridade de Maria sobre ela estar com fome implica que, ele crê que se se ela informar isso, João acreditará que ela está com fome. De modo intuitivo, se um agente i confia na sinceridade de um agente j sobre uma fórmula ϕ, o agente i irá crer que se ele for informado por j sobre ϕ, j também crê em ϕ.

$$ConfSinc(i, j, \phi) ::= Bel_i(Inf_{j,i}\phi \rightarrow Bel_j\phi)$$

- **Validade:** Validade é o tipo de confiança que define a crença de alguém passar uma informação que é de fato verdadeira. Por exemplo, se Maria confia na validade de João sobre o mercado não abrir no feriado implica que, ela crê que se João informar isso, o mercado realmente estará fechado. De modo intuitivo, se um agente i confia na validade de um agente j sobre uma fórmula ϕ, o agente i irá crer que se ele for informado por j sobre ϕ, ϕ é verdadeiro.

$$ConfVal(i, j, \phi) ::= Bel_i(Inf_{j,i}\phi \rightarrow \phi)$$

- **Completetude:** Completetude é o tipo de confiança que define a crença que relaciona a verdade com o que o confiado fala. Por exemplo, se João confia na completude de Maria sobre o controle remoto da televisão estar sobre a mesa implica que, ele crê que se o controle está sobre a mesa, Maria irá informar João sobre isso. De modo intuitivo, se um agente i confia na completude de um agente j sobre uma fórmula ϕ, o agente i irá acreditar que se ϕ for verdadeiro, j irá informá-lo sobre isso.

$$ConfComple(i, j, \phi) ::= Bel_i(\phi \rightarrow Inf_{j,i}\phi)$$

- **Cooperatividade:** Cooperatividade é o tipo de confiança que define a crença de alguém informar algo que é de seu conhecimento. Por exemplo, se Maria confia na cooperatividade de João sobre a chave do carro estar no porta-chaves implica que, ela crê que se João souber que as chaves do carro estão no porta-chaves, ele irá informar isso à Maria. De modo intuitivo, se um agente i confia na cooperatividade de um agente j sobre uma fórmula ϕ, o agente i irá acreditar que se j possui crença em ϕ, j irá informá-lo sobre isso.

$$ConfCoop(i, j, \phi) ::= Bel_i(Bel_j\phi \rightarrow Inf_{j,i}\phi)$$

- **Competência:** Competência é o tipo de confiança que define a crença que relaciona as crenças de um confiado com a verdade. Por exemplo, se João confia na competência de Maria sobre a alta da cotação do iene em relação à rúpia nepalesa implica que, ele crê se Maria souber da alta na cotação do iene em relação à rúpia nepalesa, essa informação será de fato verdadeira. De modo intuitivo, se um agente i confia na competência de um agente j sobre uma fórmula ϕ, o agente i irá acreditar que se j possui crença em ϕ, ϕ será verdadeiro.

$$ConfCompet(i, j, \phi) ::= Bel_i(Bel_j\phi \rightarrow \phi)$$

- **Vigilância:** Vigilância é o tipo de confiança que define a crença que relaciona a verdade com as crenças de um confiado, sendo considerada o dual da competência. Por exemplo, se Maria confia na vigilância de João sobre a baixa na cotação da rúpia nepalesa em relação ao iene implica que, ela crê que se a baixa na cotação da rúpia nepalesa em relação ao iene for verdadeira, João terá conhecimento disso. De modo intuitivo, se um agente i confia na vigilância de um agente j sobre uma fórmula ϕ, o agente i irá acreditar que se ϕ for verdadeiro, j possui crença em ϕ.

$$ConfVig(i, j, \phi) ::= Bel_i(\phi \rightarrow Bel_j\phi)$$

A essa altura, podemos perceber que os seis tipos de confiança definidos acima possuem uma relação de interdependência. Repare que a confiança na sinceridade junto com a confiança na competência implicam na confiança sobre a validade e a confiança na vigilância juntamente com a confiança na cooperatividade implica confiança na completude (AMGOUD; DEMOLOMBE, 2014). Em termos formais, temos:

- (V) $\vdash ConfSinc(i, j, \phi) \wedge ConfCompet(i, j, \phi) \rightarrow ConfVal(i, j, \phi)$
- (C) $\vdash ConfVig(i, j, \phi) \wedge ConfCoop(i, j, \phi) \rightarrow ConfComple(i, j, \phi)$

Os resultados das ações de informação sobre as fórmulas de confiança (AMGOUD; DEMOLOMBE, 2014) são mostrados abaixo:

- (E1) $\vdash ConfSinc(i, j, \phi) \rightarrow (Inf_{j,i}\phi \rightarrow Bel_iBel_j\phi)$
- (E2) $\vdash ConfVal(i, j, \phi) \rightarrow (Inf_{j,i}\phi \rightarrow Bel_i\phi)$
- (E3) $\vdash ConfCoop(i, j, \phi) \rightarrow (\neg Inf_{j,i}\phi \rightarrow Bel_i\neg Bel_j\phi)$
- (E4) $\vdash ConfComple(i, j, \phi) \rightarrow (\neg Inf_{j,i}\phi \rightarrow Bel_i\neg\phi)$

Esses resultados podem ser derivados utilizando a Definição 2.8 sobre os axiomas e regra de inferência de $\mathcal{LC}$ e as fórmulas de confiança.

Como exemplo dessas ações de informação, iremos apresentar a seguir um caso de (E2), que mostra que o resultado de j informar a i a mensagem ϕ quando i confia na validade de j com relação a ϕ é que i irá acreditar em ϕ:

Exemplo 2.9. *Sejam i e j dois agentes e $DM = \langle\langle\{ConfVal(i, j, \phi)\}, \{\}\rangle, \{Inf_{j,i}\phi\}\rangle$ uma Descrição de Mundo. A partir de DM, podemos derivar que o agente i acredita em ϕ seguindo os passos abaixo.*

1	$ConfVal(i, j, \phi)$	*por DM*
2	$Inf_{j,i}\phi$	*por DM*
3	$Bel_i Inf_{j,i}\phi$	*por (OBS) em (2)*
4	$Bel_i(Inf_{j,i}\phi \rightarrow \phi)$	*pela definição de (1)*
5	$(Bel_i(Inf_{j,i}\phi \rightarrow \phi)) \rightarrow (Bel_i Inf_{j,i}\phi \rightarrow Bel_i\phi)$	*por (K)*
6	$(Bel_i Inf_{j,i}\phi \rightarrow Bel_i\phi)$	*por (PR) sobre (4) e (5)*
7	$Bel_i\phi$	*pelo axioma (PR) sobre (3) e (6)*

2.2 Argumentação

Existem situações em que desejamos defender ou atacar um ponto de vista, objetivo ou ação a partir de um conjunto de crenças. Uma maneira de encontrar esse ponto de vista é simplesmente unir todas essas crenças e verificar qual tipo de decisão elas suportam. Porém, as crenças de um agente ou de um grupo de agentes podem ser inconsistentes entre si e a união delas podem acarretar em pontos de vista contraditórios. Por outro lado, se retirarmos pelo menos uma das crenças ditas inconsistentes, perdemos a capacidade de analisar as possibilidades envolvendo as crenças eliminadas. A maneira mais utilizada na literatura para formalizar uma argumentação é por meio de Frameworks de Argumentação, que são constituídos de um conjunto de argumentos, gerados a partir das crenças de um agente ou de um grupo de agentes, e uma relação de ataque definida sobre esses argumentos. O resultado de um Framework de Argumentação é um multiconjunto de argumentos que, unidos, suportam ou atacam uma decisão, ponto de vista, objetivo, etc.

Em (AMGOUD; DEMOLOMBE, 2014) é apresentado um sistema de argumentação envolvendo um conjunto de agentes que define quais devem ser as crenças aceitas. Esse sistema utiliza como base o Framework de Argumentação de Dung (DUNG, 1995) e uma de suas semânticas para avaliar os argumentos. Antes de apresentar esse sistema, iremos primeiro mostrar a seguir as principais definições que envolvem o Framework de Argumentação de Dung.

2.2.1 Framework de Dung

O Framework de Argumentação de Dung (DUNG, 1995) consiste de um conjunto de argumentos e uma relação binária de ataque. Os argumentos e a relação de ataque presentes nos Frameworks de Dung são entidades abstratas. Ambos não possuem informações sobre sua estrutura ou sua origem. As definições formais envolvendo o Framework de Dung são apresentadas abaixo.

Definição 2.10 (Framework de Argumentação de Dung (DUNG, 1995))**.** *Um framework de argumentação de Dung é uma tupla $AF = \langle \mathcal{A}, \mathcal{R} \rangle$, em que $\mathcal{A}$ é um conjunto de argumentos abstratos e $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ é uma relação de ataque entre dois argumentos.*

Exemplo 2.11 ((KACI, 2011))**.** *Seja $AF_1 = \langle \mathcal{A}_1, \mathcal{R}_1 \rangle$ em que $\mathcal{A}_1 = \{A,B,C,D,E,F,G,H\}$ e $\mathcal{R}_1 = \{(A, G), (D, C), (D, E), (C, F), (F, C), (E, F), (F, E)\}$. Seja $AF_2 = \langle \mathcal{A}_2, \mathcal{R}_2 \rangle$ com*

$\mathcal{A}_2 = \{A, B, C, D\}$ *e* $\mathcal{R}_2 = \{(A, B), (B, A), (A, C), (B, C), (C, D)\}$. *As Figuras* 2.1 *e* 2.2 *representam os frameworks* AF_1 *e* AF_2 *respectivamente. Uma seta de um argumento para outro indica que o primeiro ataca o segundo.*

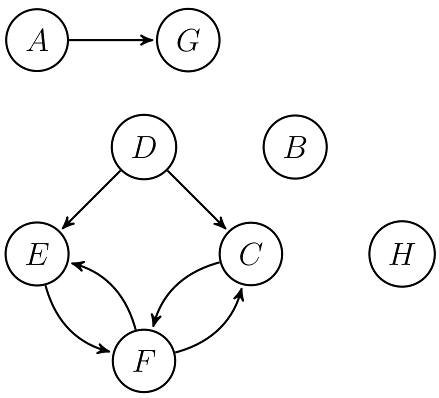

Figura 1 – AF_1

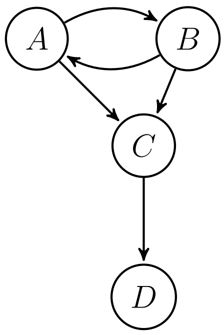

Figura 2 – AF_2

Um framework de argumentação pode ser visto como um grafo cujos vértices representam os argumentos e suas arestas representam a relação de ataque entre os argumentos. Para avaliar os argumentos, precisamos recorrer a uma semântica. Algumas semânticas foram propostas em (DUNG, 1995) e uma parte dessas foi refinada em outros trabalhos, como por exemplo em (BARONI et al., 2005; DUNG et al., 2007). Antes de mostrá-las, iremos introduzir algumas definições básicas.

Como dito anteriormente, argumentação é utilizada em prol de estabelecer ou defender um ponto de vista, objetivo ou ação. Nesse âmbito, um conjunto de argumentos que defendem um mesmo ponto de vista não deve possuir conflitos entre si. A próxima definição formaliza a noção de um conjunto livre de conflitos.

Definição 2.12 (Conjunto de argumentos livre de conflito)**.** *(DUNG, 1995) Seja* $AF = \langle \mathcal{A}, \mathcal{R} \rangle$ *um framework de argumentação de Dung. Um conjunto* $S \subseteq \mathcal{A}$ *é dito livre de conflito se e somente se,* $\forall a, b \in S$, $(a, b) \notin \mathcal{R}$.

Além de ser livre de conflitos, é desejável que um conjunto de argumentos que defendem o mesmo ponto de vista ou objetivo também seja capaz se defender de ataques de argumentos externos.

Definição 2.13 (Defesa de um conjunto de argumentos)**.** *(DUNG, 1995) Seja* $AF = \langle \mathcal{A}, \mathcal{R} \rangle$ *um framework de argumentação de Dung. Um conjunto* $S \subseteq \mathcal{A}$ *defende um conjunto* $S' \subseteq A$ *se e somente se,* $\forall b \in \mathcal{A}$ *e* $\forall a \in S'$ *com* $(b, a) \in \mathcal{R}$, $\exists c \in S$ *tal que* $(c, b) \in \mathcal{R}$.

Com as definições de defesa de um conjunto de argumentos e de conjunto de argumentos livre de conflito, podemos definir o que é um conjunto admissível de argumentos, que é exatamente um conjunto de argumentos que é simultaneamente livre de conflito, conforme a Definição 2.12, e se defende de ataques externos, conforme a Definição 2.13. Essas duas propriedades são o mínimo necessário para que um conjunto de argumentos possa sustentar uma opinião, visto que o conjunto não iria possuir uma contradição interna e seria capaz de atacar qualquer argumento do framework que pudesse apresentar algum risco ao ponto de vista representado pelo conjunto. A seguir iremos formalizar a definição de conjunto admissível de argumentos.

Definição 2.14 (Conjunto admissível de argumentos)**.** *(DUNG, 1995) Seja $AF = \langle \mathcal{A}, \mathcal{R} \rangle$ um framework de argumentação de Dung. Dizemos que um conjunto $S \subseteq \mathcal{A}$ é um conjunto admissível de argumentos se e somente se*

- *S é livre de conflito e*
- *S defende-se de ataques externos.*

Exemplo 2.15 (Continuação do Exemplo 2.11)**.** *Considere o framework de argumentação AF_1 mostrado na figura 2.1. Os conjuntos $\{A\}$, $\{D\}$, $\{F\}$, $\{A, B, D, F, H\}$ são conjuntos admissíveis de argumentos de AF_1, pois $\{A\}$ e $\{D\}$ não sofrem nenhum ataque, $\{F\}$ se defende de E e C e $\{A, B, D, F, H\}$ se defende de C e E.*

Considere agora o framework de argumentação AF_2 mostrado na figura 2.2. Os conjuntos $\{A\}$, $\{B\}$, $\{A, D\}$ e $\{B, D\}$ são conjuntos admissíveis de argumentos de AF_2, pois $\{A\}$ se defende de B, $\{B\}$ se defende de A, em $\{A, D\}$, A defende D de C e se defende de B e em $\{B, D\}$, B defende D de C e B se defende de A.

Tendo definido um conjunto admissível, podemos definir as semânticas de um framework de argumentação. As semânticas são construídas a partir da definição de extensões aceitáveis, que são conjuntos admissíveis que satisfazem algumas propriedades adicionais. Embora existam diversos tipos de extensões, como podemos ver em (DUNG, 1995; BARONI et al., 2005; DUNG et al., 2007), iremos apresentar as mais utilizadas na literatura:

Definição 2.16 (Semânticas de Frameworks de Argumentação)**.** *(DUNG, 1995) Seja $AF = \langle \mathcal{A}, \mathcal{R} \rangle$ um framework de argumentação de Dung.*

- *Um conjunto $S_p \subseteq \mathcal{A}$ é uma extensão preferida se e somente se é uma extensão admissível maximal (com respeito à inclusão de conjuntos).*
- *Um conjunto $S_e \subseteq \mathcal{A}$ é uma extensão estável se e somente se $\forall A \in \mathcal{A} \setminus S_e$, $\exists B \in S_e$ tal que $(B, A) \in \mathcal{R}$.*
- *Um conjunto $S_c \subseteq \mathcal{A}$ é uma extensão completa se e somente se é livre de conflito e $S = \{A \mid A \in \mathcal{A}, S \text{ defende } A\}$.*
- *Um conjunto $S_g \subseteq \mathcal{A}$ é uma extensão grounded se e somente se é a menor (com respeito à inclusão de conjuntos) extensão completa de AF.*

Cada uma das semânticas apresentadas possui características diferentes, podendo ser utilizadas em diferentes contextos. A extensão preferida S_p possui uma intuição simples, buscando incluir o maior número de argumentos, respeitando as restrições da extensão admissível. A extensão estável S_e também possui uma intuição bem simples, que se baseia em atacar todos os argumentos que não pertencem a S_e. A extensão completa S_c tem como característica ser um conjunto de argumentos que é capaz de se defender de ataques externos e, para todo conjunto de argumentos A, tal que se S_c defende A, A está contido em S_c, ou seja, S_c contém todos os argumentos que são defendidos por ele mesmo. A extensão grounded S_g é construída a partir dos argumentos que não são atacados por nenhum outro argumento. A partir desse primeiro passo, removemos os argumentos que são atacados pelos argumentos presentes no primeiro passo e incluímos em S_g os novos argumentos que não são atacados por ninguém. Esse processo se repete até que não seja possível incluir mais ninguém em S_g. Essa construção de S_g garante que só existe no máximo uma extensão grounded por Framework de Argumentação.

Exemplo 2.17 (Continuação do Exemplo 2.11)**.** *O conjunto $\{A, B, D, F, H\}$ é a extensão grounded e a única extensão preferida ou estável de AF_1. A extensão grounded de AF_2 é vazia. Os conjuntos $\{A, D\}$ e $\{B, D\}$ são as extensões preferidas ou estáveis de AF_2.*

Note que tanto a extensão estável como a extensão preferida são também extensões completas. O conceito de incluir todos os argumentos que podem ser defendidos pela extensão acarreta em atacar todos os que não estão presentes dentro dela, além de também maximizar a extensão com respeito a seu tamanho.

Com as definições de extensões aceitáveis encerramos as definições necessárias para o entendimento dos Frameworks de Argumentação de Dung. Na próxima seção iremos integrar a $\mathcal{LC}$ ao Framework de Dung com o objetivo de poder raciocinar sobre argumentos que defendem pontos de vista que envolvem a confiança entre agentes em relação a fórmulas da linguagem de $\mathcal{LC}$.

2.3 Confiança e Frameworks de Argumentação

Nesta seção iremos apresentar o sistema de argumentação definido em (AMGOUD; DEMOLOMBE, 2014). Esse sistema segue três passos: i) construção dos argumentos a partir de uma Descrição de Mundo, ii) estabelecer quais conjuntos de argumentos serão considerados aceitos, e iii) especificar quais fórmulas podem ser concluídas a partir de uma base de crenças.

Nesse sistema, iremos instanciar os argumentos abstratos do Framework de Argumentação de Dung a partir de $\mathcal{LC}$. O sistema contará com um conjunto de argumentos, uma relação de ataque e uma avaliação dos argumentos a partir de uma das semânticas de Dung apresentadas anteriormente. Os argumentos são construídos a partir de $\mathcal{LC}$ seguindo os princípios da consistência, não possuindo contradições entre si, e da minimalidade, evitando incluir informações desnecessárias ou redundantes. A seguir iremos mostrar a definição da construção de um argumento a partir de uma Descrição de Mundo.

Definição 2.18 (Construção de argumentos)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $DM = \langle \mathcal{K}, \mathcal{AC} \rangle$ a Descrição de Mundo construída sobre uma linguagem $\mathcal{L}$ e $K_i \in \mathcal{K}$ a base de crenças de um agente i. Uma tupla $\langle H, h \rangle$, em que H é chamado de suporte e h é chamado de conclusão, é um argumento se e somente se ela satisfaz as seguintes propriedades:*

- *$H \subseteq K_i$ e $h \in \mathcal{L}$;*
- *$H \nvdash \bot$;*
- *$H \vdash h$;*
- *$\nexists H' \subset H$ tal que $H' \vdash h$*

O símbolo $\bot$ utilizado na definição acima representa o absurdo. Definimos que um conjunto de fórmulas S deriva o absurdo ($S \vdash \bot$) se e somente se para uma fórmula qualquer ϕ, ($S \vdash \phi$) e ($S \vdash \neg\phi$). De modo análogo, dizemos que S não deriva o absurdo ($S \nvdash \bot$) se e somente se para uma fórmula qualquer ϕ, $\neg(S \vdash \phi)$ e $\neg(S \vdash \neg\phi)$.

Definição 2.19 (Conjunto dos argumentos de uma Descrição de Mundo)**.** *Seja $DM = \langle \mathcal{K}, \mathcal{AC} \rangle$ uma Descrição de Mundo construída sobre $\mathcal{L}$. Iremos chamar de $Arg(DM)$ o conjunto de argumentos construídos a partir de todas as Bases de Crenças $K_i \in \mathcal{K}$.*

A definições acima mostram a construção de um argumento a partir de uma base de crenças. O suporte de um argumento contém o mínimo necessário para derivar a fórmula que faz o papel da conclusão. Além disso, não devem existir inconsistências entre as fórmulas do suporte. O exemplo abaixo relata a criação de um conjunto de argumentos gerados a partir de uma Base de Crenças.

Exemplo 2.20. *Sejam a descrição de mundo $DM = \langle \{K_i, K_j\}, \{Inf_{j,i}\psi\} \rangle$ e as bases de crenças $K_i = \{Bel_i\phi,\ Bel_i(Inf_{j,i}\psi),\ Bel_i(\neg Inf_{k,i}\delta),\ ConfSinc(i, j, \psi),\ ConfCoop(i, k, \delta)\}$ do agente i e $K_j = \{\}$ do agente j. A partir de DM podemos construir infinitos argumentos, incluindo os seguintes:*

- $\langle \{Bel_i(\neg Inf_{k,i}\delta)\}, Bel_i(\neg Inf_{k,i}\delta) \rangle$
- $\langle \{Bel_i Inf_{j,i}\psi,\ ConfSinc(i, j, \psi)\}, Bel_i Bel_j \psi \rangle$
- $\langle \{Bel_i(\neg Inf_{k,i}\delta),\ ConfCoop(i, k, \delta)\}, Bel_i \neg Bel_j \delta\} \rangle$

Tendo definido a construção dos argumentos a partir de uma Descrição de Mundo, iremos tratar diferentes tipos de ataque entre argumentos. Primeiramente iremos definir um ataque gerado a partir de um conflito entre a conclusão de um argumento e a premissa de outro argumento.

Definição 2.21 (Ataque sobre premissa)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. Dizemos que A realiza um ataque sobre premissa em B se e somente se $\exists h'' \in H'$ tal que $h = Bel_i\phi$ e $h'' = Bel_i\neg\phi$.*

A seguir, iremos introduzir um meio de atacar argumentos que tem em seu suporte fórmulas relacionadas à confiança, a partir de situações em que um agente a quem a confiança é dirigida agiu de forma contrária ao esperado. Por exemplo, para atacar um argumento que possui em seu suporte uma fórmula que represente a confiança de um agente i na sinceridade de um agente j em relação a uma fórmula ϕ, podemos, além de utilizar um argumento que possua em sua conclusão a negação dessa confiança, utilizar um argumento que mostre uma situação na qual j falhou em ser sincero em uma outra situação, como informar uma fórmula ψ sem possuir crença na mesma. Com isso, além dos ataques sobre premissa, também devem existir meios de atacar os argumentos em relação aos tipos de confiança utilizados nos suportes dos argumentos. A seguir, serão definidos os ataques sobre a confiança dos agentes, iniciando pelo ataque sobre sinceridade:

Definição 2.22 (Ataque sobre sinceridade)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. Dizemos que A realiza um ataque sobre sinceridade em B se e somente se $h = Bel_i(Inf_{j,i}\phi \wedge \neg Bel_j\phi)$ e $ConfSinc(i, j, \psi) \in H'$.*

Em outras palavras, o ataque sobre a confiança de um agente i na sinceridade de um agente j sobre uma fórmula ϕ acontece quando um argumento conclui que i acredita que j o informou sobre uma fórmula ψ mesmo crendo em sua negação.

O ataque à confiança de um agente i na validade de um agente j sobre uma fórmula ϕ acontece quando um argumento conclui que i acredita que j o informou sobre uma fórmula ψ e ψ não é considerada verdade. Sua definição formal vem a seguir:

Definição 2.23 (Ataque sobre validade)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. Dizemos que A realiza um ataque sobre validade em B se e somente se $h = Bel_i(Inf_{j,i}\phi \wedge \neg\phi)$ e $ConfVal(i, j, \psi) \in H'$.*

Um argumento que possui fórmula de confiança sobre a completude de um agente em seu suporte também pode ser atacado. O ataque à confiança de um agente i na completude de um agente j sobre uma fórmula ϕ acontece quando um argumento conclui que i acredita que j não o informou sobre uma fórmula ψ e ψ é considerada verdade:

Definição 2.24 (Ataque sobre completude)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. Dizemos que A realiza um ataque sobre completude em B se e somente se $h = Bel_i(\phi \wedge \neg Inf_{j,i}\phi)$ e $ConfComple(i, j, \psi) \in H'$.*

O ataque à confiança de um agente i na cooperatividade de um agente j sobre uma fórmula ϕ acontece quando um argumento conclui que i acredita que j não o informou sobre uma fórmula ψ mesmo crendo nela:

Definição 2.25 (Ataque sobre cooperatividade)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. Dizemos que A realiza um ataque sobre cooperatividade em B se e somente se $h = Bel_i(Bel_j\phi \wedge \neg Inf_{j,i}\phi)$ e $ConfCoop(i, j, \psi) \in H'$.*

O ataque à confiança de um agente i na competência de um agente j sobre uma fórmula ϕ acontece quando um argumento conclui que i acredita que j crê em uma fórmula ψ e a mesma não é considerada verdadeira. A definição desse tipo de ataque é mostrada abaixo:

Definição 2.26 (Ataque sobre competência)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. A realiza um ataque sobre competência em B se e somente se $h = Bel_i(Bel_j\phi \wedge \neg\phi)$ e $ConfCompet(i, j, \psi) \in H'$.*

Por fim, o ataque à confiança de um agente i na vigilância de um agente j sobre uma fórmula ϕ acontece quando um argumento conclui que i acredita que j não crê em uma fórmula ψ que é verdadeira:

Definição 2.27 (Ataque sobre vigilância)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. A realiza um ataque sobre vigilância em B se e somente se $h = Bel_i(\phi \wedge \neg Bel_j\phi)$ e $ConfVig(i, j, \psi) \in H'$.*

Exemplo 2.28. *Seja $DM = \langle\ \langle\{ConfSinc(i, j, \phi),\ Bel_i(\neg Bel_j\psi\},\ ConfVal(i, j,\ Bel_j\psi)\rangle$, $\{Inf_{j,i}\ \psi, Inf_{j,i}Bel_j\psi\}\rangle$ uma Descrição de Mundo dos agentes i e j. A partir de DM, podemos construir os seguintes argumentos:*

- $A_1 = (\{Bel_i(Inf_{j,i}\psi), Bel_i(\neg Bel_j\psi)\}, Bel_i(Inf_{j,i}\psi) \wedge Bel_i(\neg Bel_j\psi))$
- $A_2 = (\{ConfSinc(i, j, \phi)\},\ Bel_i(Inf_{j,i}\phi \rightarrow Bel_j\phi))$
- $A_3 = (\{ConfVal(i, j, Bel_j\psi),\ Bel_iInf_{j,i}Bel_j\psi\},\ Bel_iBel_j\psi)$

Podemos notar que o argumento A_1 realiza um ataque sobre a sinceridade do argumento A_2 e o argumento A_3 realiza um ataque sobre premissa em A_1.

O exemplo acima ilustra um ataque sobre a sinceridade entre dois argumentos. Podemos notar que o argumento A_1 conclui $Bel_i(Inf_{j,i}\psi) \land Bel_i(\neg Bel_j\psi)$ e A_2 possui em seu suporte $ConfSinc(i, j, \phi)$, caracterizando um ataque de A_1 sobre a sinceridade em A_2. Além disso, o argumento A_3, com conclusão $Bel_iBel_j\psi$ efetua um ataque sobre premissa no suporte de A_1.

A seguir, iremos definir uma relação para capturar essas sete novas formas de ataques.

Definição 2.29 (Relação $\Re$ (AMGOUD; DEMOLOMBE, 2014))**.** *Seja $DM = \langle \mathcal{K}, \mathcal{AC} \rangle$ uma Descrição de Mundo. Considere (H, h) e (H', h') dois argumentos construídos a partir de DM. Dizemos que $((H, h), (H', h')) \in \Re$ se e somente se:*

- *(H, h) ataca sobre premissa (H', h'), ou*
- *(H, h) ataca sobre sinceridade (H', h'), ou*
- *(H, h) ataca sobre validade (H', h'), ou*
- *(H, h) ataca sobre completude (H', h'), ou*
- *(H, h) ataca sobre cooperatividade (H', h'), ou*
- *(H, h) ataca sobre competência (H', h'), ou*
- *(H, h) ataca sobre vigilância (H', h').*

Com a definição da relação $\Re$, podemos então utilizar um Framework de Argumentação que permite raciocinar sobre as crenças de um conjunto de agentes. A subseção a seguir irá mostrar como podemos raciocinar sobre um conjunto de argumentos formados a partir de um conjunto de crenças de vários agentes.

2.3.1 Raciocinando com Frameworks de Argumentação

Como levantamos no início da Seção 2.2, podemos utilizar Frameworks de Argumentação quando desejamos defender ou atacar um ponto de vista, objetivo ou ação a partir de um conjunto de crenças. Aqui queremos construir um Framework de Argumentação a partir das crenças presentes em uma Descrição de Mundo D. Dada essa descrição D, podemos construir um conjunto de argumentos conforme a Definição 2.19 e a Definição da relação de ataque 2.29. A formalização desse framework é descrita abaixo.

Definição 2.30 (Framework de Argumentação sobre crenças)**.** *(AMGOUD; DEMOLOMBE, 2014) Seja $DM = \langle \mathcal{K}, \mathcal{AC} \rangle$ a Descrição de Mundo construída sobre $\mathcal{L}$. Um framework de argumentação construído sobre DM é um par $\mathcal{T} = (S, \Re)$ em que $S \subseteq Arg(DM)$ é uma conjunto finito de argumentos com $Arg(DM)$ o conjunto introduzido na Definição 2.19 e $\Re \subseteq S \times S$ é a relação da Definição 2.29.*

Uma vez que os argumentos podem ser conflitantes, devemos encontrar um meio de avalia-los e decidir quais seriam os argumentos aceitáveis. Como vimos anteriormente, as semânticas dos Frameworks de argumentação de Dung podem ser utilizadas para decidir quais são esses argumentos aceitáveis dentro de um Framework de Argumentação. Nesse sistema, iremos utilizar as extensões estáveis por conta de elas permitirem particionar os subconjuntos do conjunto de argumentos em dois conjuntos: as extensões estáveis e as não extensões. As extensões extraídas a partir do Framework de Argumentação irão permitir definir quais fórmulas poderão ser inferidas a partir de uma Descrição de Mundo. Essas fórmulas inferidas indicam as crenças do grupo a partir da Descrição de Mundo. Basicamente, para uma fórmula ser inferida, ela deve ser a conclusão de pelo menos um argumento de cada uma das extensões estáveis do framework de argumentação. Vale ressaltar que o argumento não precisa ser o mesmo em todas as extensões.

Definição 2.31 (Resultado de um Framework de Argumentação $\mathcal{T}$)**.** *(AMGOUD; DEMOLOMBE, 2014) Sejam $\mathcal{T} = (S, \Re)$ um framework de argumentação definido conforme 2.30 construído sobre uma Descrição de Mundo DM, $S \subseteq Arg(DM)$ um conjunto finito de argumentos e $Ext(\mathcal{T})$ o conjunto de todas as extensões estáveis de $\mathcal{T}$. Uma fórmula ϕ é inferida de DM se e somente se para todo $\mathcal{E} \in Ext(\mathcal{T})$, existe um $\langle H, \phi \rangle \in \mathcal{E}$. O resultado de um framework de argumentação $\mathcal{T}$ é o conjunto de todas as fórmulas ϕ inferidas de $\mathcal{T}$.*

Exemplo 2.32. *Seja $DM = \langle \langle K_a, K_b, K_c \rangle, \{Inf_{a,b}\phi\} \rangle$ uma Descrição de Mundo em que:*

- $K_a = \{\langle Bel_a \alpha \wedge \beta \wedge \gamma, Bel_a \alpha \wedge \beta \wedge \gamma \wedge \psi \rightarrow \epsilon, Bel_a \phi, Bel_a \phi \rightarrow \psi, Bel_a \neg\psi \rightarrow \neg\epsilon\}$
- $K_b = \{Bel_b \alpha \wedge \beta \wedge \gamma, Bel_b \alpha \wedge \beta \wedge \gamma \wedge \psi \rightarrow \epsilon, ConfVal(b, a, \phi), Bel_b \phi \rightarrow \psi, Bel_b \neg\psi \rightarrow \neg\epsilon\}$
- $K_c = \{Bel_c \alpha \wedge \beta \wedge \gamma, Bel_c \alpha \wedge \beta \wedge \gamma \wedge \psi \rightarrow \epsilon, Bel_c \neg\psi, Bel_c \phi \rightarrow \psi, Bel_c \neg\psi \rightarrow \neg\epsilon\}$

Considere os seguintes argumentos construídos a partir de DM:

- $A_1 = (\{\phi \rightarrow \psi, \phi, \alpha \wedge \beta \wedge \gamma, \alpha \wedge \beta \wedge \gamma \wedge \psi \rightarrow \epsilon\}, \epsilon)$
- $A_2 = (\{\neg\psi, \neg\psi \rightarrow \neg\epsilon\}, \neg\epsilon)$
- $A_3 = (\{Inf_{a,b}\phi, ConfVal(b, a, \phi), \phi \rightarrow \psi\}, \psi)$

O resultado do Framework de Argumentação $\mathcal{T} = (S, \Re)$ é $\{\psi, \epsilon\}$, pois a única extensão estável $\{A_1, A_3\}$ possui essas fórmulas como conclusão de seus argumentos.

Considerando no exemplo acima pode ser visto como a formalização da situação de motivação vista no início do capítulo, desde que ϕ seja interpretado por “A paciente foi examinada”; ψ, por “A paciente possui fissuras na pele”; γ, por “A paciente apresenta febre alta”; β, por “A paciente apresenta dor de cabeça”; γ, por “A paciente apresenta fotossensibilidade” e ϵ, por “A paciente está com meningite”.

2.3.2 Trabalhos relacionados

Nesta seção iremos explorar alguns dos trabalhos relacionados, mostrando as diferenças entre esses e (AMGOUD; DEMOLOMBE, 2014).

O trabalho de Liau em (LIAU, 2003) consiste em uma modelagem da confiança entre dois agentes utilizando uma Lógica Modal. Liau defende que se um agente a acredita que um agente b lhe informa sobre uma fórmula ϕ e a confia no julgamento de b em relação a ϕ, a também irá acreditar em ϕ. Em (AMGOUD; DEMOLOMBE, 2014), a relação de confiança também envolve dois agentes e uma fórmula, porém são tratados seis tipos de confiança, possibilitando expressar outras consequências para a confiança entre agentes.

Em (STRANDERS et al., 2007), Stranders modela os motivos por trás da confiança de uma agente i em um agente j em um ambiente no qual cada agente possui interesses próprios. Para representar o comportamento dos demais agentes, Stranders utiliza uma Lógica Possibilística. Para encontrar o motivo da confiança entre dois agentes, são utilizados argumentos e dados estatísticos, tais como o grau de confiança no suporte de um argumento e quanto a conclusão de um argumento pode auxiliar na conquista de um objetivo. Em (AMGOUD; DEMOLOMBE, 2014), não são discutidas as origens das relações de confiança. Além disso, a relação de confiança de (AMGOUD; DEMOLOMBE, 2014) é qualitativa, ou seja, não utiliza números nem métodos estatísticos.

Em (TANG et al., 2012), Tang define um sistema formal para raciocinar sobre a confiança entre dois agentes utilizando argumentos que são comparados a partir de uma relação binária que expressa preferências. O sistema toma como entrada um grafo no qual os vértices representam agentes e as arestas representam a relação de confiança, tal que se existe uma aresta iniciando em um agente i e terminando em um agente j, então i confia em j. As arestas do grafo de confiança possuem pesos representando um grau de confiança. De maneira similar a (AMGOUD; DEMOLOMBE, 2014), os argumentos são criados a partir das bases de crenças dos agentes, porém cada crença possui um grau de certeza proveniente do grau de confiança entre o agente e a sua fonte de informação. Estes graus de certeza medem a confiabilidade de cada crença. Os graus de certeza são então combinados para calcular um nível do suporte de um argumento que, por sua vez, irá ser utilizado para estabelecer uma relação de preferência entre os argumentos. O diferencial de (AMGOUD; DEMOLOMBE, 2014) está em trabalhar com seis tipos de confiança ao invés de uma confiança absoluta, que no trabalho de Tang é modelada como a confiança na validade, e, por tratar a confiança em si como uma crença, permitir raciocinar sobre a própria relação de confiança, podendo adiciona-la ao suporte e conclusão de argumentos.

2.4 Tópicos

Na seção anterior, revisamos as definições de Framework de Argumentação de Dung, como estabelecer uma relação de confiança entre dois agentes sobre fórmulas e a interação desses dois assuntos. Nesta seção, iremos abordar o trabalho de Demolombe em (DEMOLOMBE; JONES, 1995), que trata de Tópicos de Fórmulas.

Existem situações as quais desejamos delimitar ou especificar um conjunto de fórmulas. Uma das maneiras de se fazer isso é definir todas as fórmulas que estarão presentes no conjunto. Embora esse método consiga especificar todas as fórmulas presentes no conjunto, ele pode se tornar inviável se tratarmos vários conjuntos diferentes, cada um podendo ter infinitos elementos. Outro meio de caracterizar um conjunto de fórmulas é utilizar o conceito de tópicos. O exemplo abaixo irá mostrar uma situação em que tópicos podem ser utilizados.

Exemplo 2.33. *Considere as seguintes sentenças: A = "Brasília é a capital do Brasil". B = "Bogotá é a capital da Colômbia". C = "Buenos Aires é a capital da Argentina". D = "A fórmula de Bhaskara calcula as raízes de uma equação de segundo grau". Repare que as três primeiras sentenças se referem a Geografia e a sentença D se refere a Matemática. Nessa situação, podemos dizer que as sentenças A, B e C pertencem ao tópico Geografia e D pertence ao tópico Matemática. Repare que a sentença E = "Brasília é a capital do Brasil e Bogotá é a capital da Colômbia" pertence ao tópico Geografia, por ser formada a partir de sentenças desse mesmo tópico. De modo similar, a sentença F = "Buenos Aires é a capital da Argentina* *e a fórmula de Bhaskara calcula as raízes de uma equação de segundo grau" pertence aos tópicos de Geografia e de Matemática, por ser formada por sentenças pertencentes aos tópicos de Geografia e Matemática.*

Tópicos podem ser utilizados de diversas maneiras em Representação de Conhecimento (DEMOLOMBE, 1994). Dentre elas, podemos utilizar tópicos para definir quais conjuntos de fórmulas pertencentes a uma base de conhecimento podem ser acessadas por um usuário (CUPPENS; DEMOLOMBE, 1996); para prover informações que não foram explicitamente requeridas, como no contexto de tópicos de interesse (CUPPENS; DEMOLOMBE, 1989); ou para caracterizar o conhecimento de um agente, a fim de não informá-lo sobre algo que já possui conhecimento (LAKEMEYER, 1993).

Escolhemos utilizar no Capítulo 3 a Lógica de Tópicos (DEMOLOMBE; JONES, 1995) para estender a $\mathcal{LC}$ por conta de sua linguagem simples e funcional. Na subseção seguinte, iremos mostrar essa Lógica de Tópicos e seus axiomas.

2.4.1 Linguagem

Nesta subseção, iremos apresentar a linguagem da lógica definida em (DEMOLOMBE; JONES, 1995). Para evitar confusão com a linguagem apresentada no capítulo anterior, iremos chamar essa lógica de *Lógica de Tópicos*, ou simplesmente de $\mathcal{LT}$. A exemplo da linguagem da lógica de confiança ($\mathcal{LC}$) começamos pelas primitivas sintáticas de $\mathcal{LT}$.

Definição 2.34 (Primitivas sintáticas de $\mathcal{LT}$)**.** *(DEMOLOMBE; JONES, 1995) As primitivas sintáticas de $\mathcal{LT}$ são os seguintes conjuntos:*

- $\mathcal{AT}$*: conjunto não-vazio de proposições atômicas denotadas por* $p_1, p_2, p_3, \ldots$
- $\mathcal{T}$*: conjunto não-vazio de tópicos denotados por* $t_1, t_2, t_3, \ldots$

A seguir, iremos definir a sintaxe de $\mathcal{LT}$ a partir de suas Primitivas Sintáticas.

Definição 2.35 (Sintaxe $\mathcal{LT}$)**.** *(DEMOLOMBE; JONES, 1995) A linguagem de $\mathcal{LT}$ é o conjunto de fórmulas definidas sobre as primitivas sintáticas $\langle \mathcal{AT}, \mathcal{T} \rangle$ pela seguinte BNF:*

$$\phi ::= p \mid \mathfrak{A}(t, \psi) \mid \neg\phi \mid \phi \vee \phi$$
$$\psi ::= p \mid \neg\psi \mid \psi \vee \psi$$

em que $p \in \mathcal{AT}$ e $t \in \mathcal{T}$. Os conectivos $\neg$ e $\vee$ são definidos como usual. O significado intuitivo do operador modal $\mathfrak{A}$ é

- $\mathfrak{A}(t, \psi)$*: a fórmula ψ da lógica proposicional pertence ao tópico t.*

Tendo definido a sintaxe da lógica $\mathcal{LT}$, iremos mostrar a axiomática dessa lógica.

2.4.2 Axiomática

Nas subseção anterior, apresentamos a sintaxe de $\mathcal{LT}$. Nesta, iremos apresentar os axiomas que caracterizam a lógica $\mathcal{LT}$.

Axiomas para o Operador Modal $\mathfrak{A}$

- (A1) $(\phi \leftrightarrow \psi) \wedge (\phi$ e ψ possuem os mesmos átomos$) \vdash \mathfrak{A}(t, \phi) \leftrightarrow \mathfrak{A}(t, \psi)$

- (A2) $(\mathfrak{A}(t, \phi) \wedge \mathfrak{A}(t, \psi)) \rightarrow \mathfrak{A}(t, \phi \wedge \psi)$
- (A3) $\mathfrak{A}(t, \phi) \rightarrow \mathfrak{A}(t, \neg\phi)$

O axioma (A1) segue do resultado apresentado em (DEMOLOMBE, 1994) ao mostrar que para toda sentença p e q da lógica proposicional, nós temos que $T(p) = T(q)$ e $F(p) = F(q)$ para todo modelo M se e somente se nós tivermos que $\vdash p \leftrightarrow q$ e p e q forem formados pelos mesmos átomos. O axioma (A2) indica que se duas sentenças p e q pertencem a um mesmo tópico t, então a conjunção delas também pertence a t. O axioma (A3) indica que a negação de uma sentença p pertencente a um tópico t também pertence a t, por exemplo, se a afirmação "Beltrano é honesto" se refere ao caráter de Beltrano, a afirmação "Beltrano é desonesto" também se refere ao caráter de Beltrano.

3 LÓGICA DE CONFIANÇA EM TÓPICOS

Este capítulo tem por objetivo apresentar duas das contribuições deste trabalho. Para tanto, iremos utilizar definições apresentadas no capítulo anterior como as de frameworks de argumentação, de confiança entre agentes em relação a fórmulas e de tópicos.

A primeira contribuição a ser apresentada neste capítulo será a definição da Lógica de Confiança em Tópicos ($\mathcal{LCT}$), uma extensão da lógica $\mathcal{LC}$ retratada no Capítulo 2, talhada para modelar a confiança entre agentes em relação a fórmulas e tópicos. Além disso, iremos definir a relação de ataque entre argumentos que possuem em seu suporte definições de confiança em tópicos.

A segunda contribuição presente neste capítulo envolve a introdução da noção de desconfiança em $\mathcal{LCT}$. Em situações do mundo real, decisões de um grupo podem ser tomadas a partir de informações falsas. Essas informações podem ser geradas a partir da ignorância, ingenuidade ou até mesmo desonestidade de um agente que busca seus próprios interesses acima dos interesses do grupo. Para tratar esse tipo de situação, iremos formalizar a relação de desconfiança entre agentes sobre fórmulas e tópicos.

Ao fim deste capítulo, iremos utilizar a obra *Otelo, o Mouro de Veneza* de *Shakespeare* para mostrar como as contribuições apresentadas podem modelar as decisões presentes na trama.

3.1 Tópicos e confiança

Nesta Seção, iremos definir a Lógica de Confiança em Tópicos, $\mathcal{LCT}$, a partir da Lógica de Confiança, $\mathcal{LC}$, e da Lógica de Tópicos, $\mathcal{LT}$, apresentadas no capítulo anterior. Antes de partirmos para a formalização de $\mathcal{LCT}$, adaptaremos o Exemplo 2.1, introduzindo uma situação em que a confiança entre agentes em um tópico pode ser decisiva na tomada de uma decisão.

Exemplo 3.1. *Alexandre, Beatriz e Carlos são médicos que estão buscando chegar a uma conclusão sobre a enfermidade de Patrícia. Em uma reunião, os médicos decidem apresentar suas conclusões sobre qual seria a doença que afeta Patrícia. Alexandre afirma que a paciente sofre de meningite, pois possui os sintomas de febre alta, forte dor de cabeça, fotossensibilidade e rachaduras na pele. Prontamente, Carlos discorda argumentando que a paciente não possui rachaduras na pele, de acordo com os primeiros exames realizados nela. A afirmação de Carlos coloca em xeque o diagnóstico de Alexandre, pois a ausência de fissuras na pele poderia indicar que a paciente estaria sofrendo de outras patologias, visto que febre, dor de cabeça e fotossensibilidade também são sintomas de várias outras doenças. Porém, antes da reunião, Alexandre e Beatriz haviam conversado e Alexandre a informou que nos últimos exames a paciente havia apresentado rachaduras na pele. Beatriz passa a acreditar na informação de Alexandre, pois confia em qualquer informação que ele passe envolvendo anomalias na pele, uma vez que Alexandre é um dermatologista. Por conseguinte, ela acredita na informação que seu colega a passou. Ela, então, informa na reunião que segundo os últimos exames, a paciente tinha fissuras na pele. Assim, o grupo decide que a paciente realmente sofre de meningite.*

O exemplo do parágrafo anterior ilustra a situação encontrada no mundo real de agentes que possuem especialização em alguma área. Por exemplo, Alexandre teria conhecimento não somente em rachaduras na pele, mas sobre qualquer assunto relacionado à dermatologia. Para tratar desse tipo de situação, iremos incorporar e adaptar à linguagem de $\mathcal{LC}$ (AMGOUD; DEMOLOMBE, 2014) as definições da linguagem de $\mathcal{LT}$ (DEMOLOMBE; JONES, 1995). Dessa maneira, podemos estender as definições de confiança entre agentes em relação a uma fórmula para um tópico, possibilitando utilizar um conjunto infinito de fórmulas as quais um agente possui confiança em outro.

3.1.1 Linguagem de $\mathcal{LCT}$

Nesta subseção iremos definir a linguagem $\mathcal{L}$ da Lógica de Confiança e Tópicos ($\mathcal{LCT}$), que é a união das lógicas $\mathcal{LC}$ e $\mathcal{LT}$ apresentadas no capítulo anterior. As primitivas sintáticas de $\mathcal{LCT}$ são apresentadas a seguir.

Definição 3.2 (Primitivas sintáticas de $\mathcal{LCT}$)**.** *As primitivas sintáticas de $\mathcal{LCT}$ são uma tripla $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$, em que*

- *$\mathcal{AT}$: conjunto não-vazio de proposições atômicas denotadas por $p_1, p_2, p_3, \ldots$*
- *$\mathcal{AG}$: conjunto finito não-vazio de agentes denotados por i, j, k, $\ldots$*
- *$\mathcal{T}$: conjunto finito não-vazio de tópicos denotados por $t_1, t_2, t_3, \ldots$*

Tendo definido as primitivas sintáticas, seguimos para a definição da linguagem:

Definição 3.3 (Linguagem $\mathcal{LCT}$)**.** *A linguagem de $\mathcal{LCT}$ é o conjunto de fórmulas definidas sobre as primitivas sintáticas $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$ pela seguinte BNF:*

$$\phi ::= p \mid \neg\phi \mid \phi \vee \phi \mid Bel_i\phi \mid Inf_{i,j}\phi \mid \mathfrak{A}(t, \psi)$$
$$\psi ::= p \mid \neg\psi \mid \psi \vee \psi$$

em que $p \in \mathcal{AT}$, $i, j \in \mathcal{AG}$ e $t \in \mathcal{T}$. Os conectivos $\neg$, $\vee$ e $\rightarrow$ são definidos como usual. O significado intuitivo dos operadores modais Bel, Inf e $\mathfrak{A}$ são

- *$Bel_i\phi$: o agente i acredita na fórmula ϕ.*
- *$Inf_{i,j}\phi$: o agente i informou ao agente j a fórmula ϕ.*
- *$\mathfrak{A}(t, \phi)$: a fórmula ϕ pertence ao tópico t.*

De maneira similar a $\mathcal{LC}$, iremos apresentar agora algumas definições que serão utilizadas no decorrer do capítulo.

Definição 3.4 (Base de Crenças)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$ as primitivas sintáticas de $\mathcal{LCT}$ e $\mathcal{L}$ a linguagem definida sobre essas primitivas. Dizemos que a base de crenças de um agente $i \in \mathcal{AG}$ em $\mathcal{L}$ é um conjunto*

$$\mathcal{K}_i \subseteq \{Bel_i\phi \mid \phi \in \mathcal{L}\}.$$

Definição 3.5 (Descrição das Crenças)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$ as primitivas sintáticas de $\mathcal{LCT}$ e $\mathcal{L}$ a linguagem definida sobre essas primitivas. Uma descrição das crenças dos agentes $i, j, k \ldots \in \mathcal{AG}$ é uma tupla denotada por $\mathcal{K} = \langle \mathcal{K}_i, \mathcal{K}_j, \mathcal{K}_k \ldots \rangle$ em que $\mathcal{K}_i, \mathcal{K}_j, \mathcal{K}_k \ldots$ são, respectivamente, as bases de crenças dos agentes $i, j, k \ldots$ em $\mathcal{L}$.*

Definição 3.6 (Descrição de Mundo)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$ as primitivas sintáticas de $\mathcal{LCT}$ e $\mathcal{L}$ a linguagem definida sobre essas primitivas. Uma descrição de mundo dos agentes $i, j, k \ldots \in \mathcal{AG}$ em $\mathcal{L}$ é uma tripla denotada por $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$, em que $\mathcal{K}$ é uma descrição das crenças dos agentes $i, j, k \ldots \in \mathcal{AG}$ em $\mathcal{L}$, $\mathcal{AC} \subseteq \{Inf_{i,j}\phi \mid \phi \in \mathcal{L}$ e $i, j \in \mathcal{AG}\}$ é o conjunto de ações de informação e $\mathcal{TP} \subseteq \{\mathfrak{A}(t, \phi) \mid t \in \mathcal{T}$ e $\phi \in \mathcal{L}\}$ é o conjunto das relações entre fórmulas e os tópicos aos quais pertencem.*

A partir deste ponto, quando não houver confusão, iremos nos referir a uma Descrição de Mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$ dos agentes $i, j, k \ldots \in \mathcal{AG}$ em $\mathcal{L}$ como a descrição definida sobre a Descrição das Crenças $\mathcal{K} = \langle \mathcal{K}_i, \mathcal{K}_j, \mathcal{K}_k \ldots \rangle$ e sobre os conjuntos de ações de informação $\mathcal{AC}$ e das relações entre fórmulas e tópicos $\mathcal{TP} \subseteq \{\mathfrak{A}(t, \phi) \mid t \in \mathcal{T}$ e $\phi \in \mathcal{L}\}$.

Tendo definido a sintaxe de $\mathcal{LCT}$, iremos apresentar na seção seguinte a sua axiomática.

3.1.2 Axiomática de $\mathcal{LCT}$

De maneira similar ao Capítulo 2, iremos expor a construção dos conectivos $\rightarrow$ e $\wedge$, que representam respectivamente a implicação e a conjunção, e a definição de sequente.

Antes de apresentar os axiomas e regras de inferência de $\mathcal{LC}$, iremos expor a construção dos conectivos $\rightarrow$ e $\wedge$, que representam respectivamente a implicação e a conjunção, e a definição de sequente. Note que podemos inferir os conectivos $\rightarrow$ e $\wedge$ a partir dos conectivos $\vee$ e $\neg$, sendo $(\phi \rightarrow \psi)$ equivalente a $(\neg\phi \vee \psi)$ e $(\phi \wedge \psi)$ equivalente a $\neg(\neg\phi \vee \neg\psi)$. Um sequente, denotado por $\Gamma \vdash \phi$, é formado por um conjunto de fórmulas $\Gamma \subseteq \mathcal{L}$, chamado de conjunto de premissas, uma fórmula $\phi \in \mathcal{L}$, chamada de conclusão, e um símbolo $\vdash$, chamado de consequência. Utilizando a definição de consequência sintática 2.8, mostramos abaixo os axiomas de $\mathcal{LCT}$:

Axiomas e Regra de Inferência da Lógica Proposicional: Sistema Axiomático de Frege

- (P1) $A \rightarrow (B \rightarrow A)$
- (P2) $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$
- (P3) $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$
- (P4) $\neg\neg A \rightarrow A$
- (P5) $A \rightarrow \neg\neg A$
- (PR) $A, (A \rightarrow B) \vdash B$

Axiomas e Regra de Inferência para o Operador Modal Bel

- (K) $Bel_i(\phi \rightarrow \psi) \rightarrow (Bel_i\phi \rightarrow Bel_i\psi)$
- (D) $\neg(Bel_i\phi \wedge Bel_i\neg\phi)$
- (Nec) Se $\vdash \phi$, então $\vdash Bel_i\phi$

Axiomas e Regra de Inferência para o Operador Modal Inf

- (EQV) Se $\vdash \phi \leftrightarrow \psi$, então $\vdash Inf_{j,i}\phi \leftrightarrow Inf_{j,i}\psi$
- (CONJ) $Inf_{j,i}\phi \wedge Inf_{j,i}\psi \rightarrow Inf_{j,i}(\phi \wedge \psi)$
- (OBS) $Inf_{j,i}\phi \rightarrow Bel_i Inf_{j,i}\phi$
- (OBS') $\neg Inf_{j,i}\phi \rightarrow Bel_i\neg Inf_{j,i}\phi$

Axiomas para o Operador Modal $\mathfrak{A}$

- (A1) $(\phi \leftrightarrow \psi) \wedge (\phi$ e ψ possuem os mesmos átomos$) \vdash \mathfrak{A}(t, \phi) \leftrightarrow \mathfrak{A}(t, \psi)$
- (A2) $(\mathfrak{A}(t, \phi) \wedge \mathfrak{A}(t, \psi)) \rightarrow \mathfrak{A}(t, \phi \wedge \psi)$

- (A3) $\mathfrak{A}(t, \phi) \rightarrow \mathfrak{A}(t, \neg\phi)$

A intuição sobre os axiomas apresentados acima é apresentada nas subseções 2.1.2 e 2.4.2, que se referem respectivamente às seções de axiomática da $\mathcal{LC}$ e $\mathcal{LT}$.

3.1.3 Confiança Ternária em $\mathcal{LCT}$

Nesta subseção, iremos apresentar a relação de confiança entre agentes em sobre tópicos utilizando a linguagem de $\mathcal{LCT}$ definida a partir das linguagens de $\mathcal{LC}$ e $\mathcal{LT}$ introduzidas no capítulo anterior. Antes, não obstante, iremos relembrar de maneira sucinta a relação de confiança mostrada anteriormente, que se limitava a relacionar agentes com respeito a uma fórmula.

Definição 3.7 (Confiança sobre uma fórmula)**.**
Confiança sobre Sinceridade:
$ConfSinc(i, j, \phi) ::= Bel_i(Inf_{j,i}\phi \rightarrow Bel_j\phi)$
Confiança sobre Validade:
$ConfVal(i, j, \phi) ::= Bel_i(Inf_{j,i}\phi \rightarrow \phi)$
Confiança sobre Completude:
$ConfComple(i, j, \phi) ::= Bel_i(\phi \rightarrow Inf_{j,i}\phi)$
Confiança sobre Cooperatividade:
$ConfCoop(i, j, \phi) ::= Bel_i(Bel_j\phi \rightarrow Inf_{j,i}\phi)$
Confiança sobre Competência:
$ConfCompet(i, j, \phi) ::= Bel_i(Bel_j\phi \rightarrow \phi)$
Confiança sobre Vigilância:
$ConfVig(i, j, \phi) ::= Bel_i(\phi \rightarrow Bel_j\phi)$

Como dito, a definição acima apenas relaciona dois agentes e uma fórmula. Nosso objetivo agora é definir a confiança entre agentes sobre tópicos. Assim iremos dizer que um agente i confia num agente j sobre um tópico t se e somente se para toda fórmula ϕ associada ao tópico t, temos que i confia em j sobre ϕ. Mais formalmente:

Definição 3.8 (Confiança sobre tópicos)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$ as primitivas sintáticas de $\mathcal{LCT}$ e $\mathcal{L}$ a linguagem definida sobre essas primitivas. Considere a descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$ sobre $\mathcal{L}$. Para $i, j \in \mathcal{AG}$ e $t \in \mathcal{T}$, definimos $\chi(i, j, t)$ como segue:*

$$\chi(i, j, t) ::= \chi(i, j, \phi) \text{ para todo } \phi \in \mathcal{L} \text{ tal que } \mathfrak{A}(t, \phi) \in \mathcal{TP}$$

em que $\chi \in \{ConfSinc, ConfVal, ConfComple, ConfCoop, ConfCompet, ConfVig\}$.

Exemplo 3.9. *Seja $DM = \langle \langle \{ConfSinc(i, j, t)\}, \{Bel_j\phi\} \rangle, \{Inf_{j,i}\phi\}, \{\mathfrak{A}(t, \phi)\} \rangle$ uma Descrição de Mundo dos agentes i e j. Tomando DM, podemos inferir que i irá acreditar que j acredita em ϕ pelo seguinte raciocínio:*

1	$ConfSinc(i, j, t)$	*por DM*
2	$\mathfrak{A}(t, \phi)$	*por DM*
3	$Inf_{j,i}\phi$	*por DM*
4	$ConfSinc(i, j, \phi)$	*por 3.8 sobre (1) e (2)*
5	$Bel_i(Inf_{j,i}\phi \rightarrow Bel_j\phi)$	*pela definição 3.7 sobre (4)*
6	$Bel_iInf_{j,i}\phi \rightarrow Bel_iBel_j\phi$	*pelo axioma (K) sobre (5)*
7	$Inf_{j,i}\phi \rightarrow Bel_iInf_{j,i}\phi$	*pelo axioma (OBS) sobre (3)*
8	$Bel_iInf_{j,i}\phi$	*por (PR) sobre (3) e (7)*
9	$Bel_iBel_j\phi$	*pelo axioma (PR) sobre (6) e (8)*

Com essa definição, introduzimos $\mathcal{LCT}$ a partir de $\mathcal{LC}$ e de $\mathcal{LT}$. Essa lógica permite que fórmulas que representam a confiança entre agentes sobre tópicos estejam presentes nas crenças de agentes e, com isso, possam ser utilizadas na construção de argumentos que serão utilizados para raciocinar sobre as crenças de um grupo de agentes. O próximo passo consiste em definir como se darão os ataques a argumentos que possuem em seu suporte fórmulas que descrevem a confiança entre agentes com relação a um tópico, uma vez que só estão definidos ataques à confiança entre agentes com relação a fórmulas.

3.1.4 Ataque a tópicos

No capítulo anterior, definimos a relação de ataque entre argumentos levando em conta as fórmulas de confiança entre agentes com relação a uma fórmula. Já na subseção anterior, construímos a $\mathcal{LCT}$ com o objetivo de introduzir fórmulas que representam a confiança entre agentes não apenas em uma fórmula, mas um tópico, que pode ser compreendido como um conjunto infinito de fórmulas. Nesta subseção, temos o objetivo de definir novas relações de ataque entre argumentos, uma vez que fórmulas envolvendo tópicos possuem um comportamento diferente das demais fórmulas.

Antes de apresentarmos as relações de ataque entre os argumentos, iremos, de maneira similar a $\mathcal{LC}$, mostrar como construiremos um argumento a partir de um conjunto de crenças de $\mathcal{LCT}$. A seguir temos a definição de construção de um argumento a partir de uma Descrição de Mundo.

Definição 3.10 (Argumentos)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$ as primitivas sintáticas de $\mathcal{LCT}$, $\mathcal{L}$ a linguagem definida sobre essas primitivas e $D = \langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$ a Descrição de Mundo construída sobre $\mathcal{L}$. Seja $K_i \subseteq \{Bel_i\phi \mid \phi \in \mathcal{L}\} \in \mathcal{K}$ a base de crenças de um agente $i \in \mathcal{AG}$. Uma tupla $\langle H, h \rangle$ é um argumento se e somente se ela satisfaz as seguintes propriedades:*

- *$H \subseteq K_i$ e $h \in \mathcal{L}$;*
- *$H \nvdash \perp$;*
- *$H \vdash h$;*
- *$\nexists H' \subset H$ tal que $H' \vdash h$*

Na definição acima, o conjunto H é chamado de suporte do argumento e a fórmula h é chamada conclusão do argumento. Além disso, o símbolo $\perp$ utilizado acima representa o absurdo. Definimos que um conjunto de fórmulas S deriva o absurdo ($S \vdash \perp$) se e somente se para uma fórmula qualquer ϕ, $(S \vdash \phi)$ e $(S \vdash \neg\phi)$. De modo análogo, dizemos que S não deriva o absurdo ($S \nvdash \perp$) se e somente se para uma fórmula qualquer ϕ, $\neg(S \vdash \phi)$ e $\neg(S \vdash \neg\phi)$.

Exemplo 3.11. *Considere a linguagem $\mathcal{L}$ de $\mathcal{LCT}$. Seja a descrição de mundo $DM = \langle\{K_i, K_j, K_k\}, \{Inf_{j,i}\ \psi\}, \{\mathfrak{A}(t, \delta)\}\rangle$ e as bases de crenças $K_i = \{Bel_i\phi,\ Bel_i(Inf_{j,i}\psi),\ Bel_i(\neg Inf_{k,i}\delta), ConfSinc(i, j, \psi),\ ConfCoop(i, k, t)\}$ do agente $i \in \mathcal{AG}$, $K_j = \{\}$ do agente $j \in \mathcal{AG}$ e $K_k = \{\}$ do agente $k \in \mathcal{AG}$, definidas sobre $\mathcal{L}$. A partir de K_i podemos construir infinitos argumentos, incluindo os seguintes:*

- $\langle\{Bel_i(\neg Inf_{k,i}\delta)\}, Bel_i(\neg Inf_{k,i}\delta)\rangle$
- $\langle\{Bel_i Inf_{j,i}\psi,\ ConfSinc(i, j, \psi)\}, Bel_i Bel_j\psi\rangle$
- $\langle\{Bel_i(\neg Inf_{k,i}\delta),\ ConfCoop(i, k, t)\}, Bel_i\neg Bel_j\delta\}\rangle$

Definição 3.12 (Conjunto dos argumentos de uma Descrição de Mundo)**.** *Sejam $\langle\mathcal{AT}, \mathcal{AG}, \mathcal{T}\rangle$ as primitivas sintáticas de $\mathcal{LCT}$, $\mathcal{L}$ a linguagem definida sobre essas primitivas e $DM = \langle\mathcal{K}, \mathcal{AC}, \mathcal{TP}\rangle$ a Descrição de Mundo construída sobre $\mathcal{L}$. Iremos chamar de $Arg(D)$ o conjunto de argumentos construídos de todas as Bases de Crenças $K_i \in \mathcal{K}$ da Descrição de Mundo D.*

Iremos definir a relação de ataque entre os argumentos a partir de seus suportes e suas conclusões, que agora podem envolver fórmulas que representam a confiança entre agentes sobre tópicos. Primeiramente iremos relembrar a definição de ataque entre os argumentos baseado em uma inconsistência por parte de suas conclusões.

Definição 3.13 (Ataque sobre premissa)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. Dizemos que A realiza um ataque sobre premissa em B se e somente se $\exists h'' \in H'$ tal que $h = Bel_i\phi$ e $h'' = Bel_i\neg\phi$.*

A seguir iremos apresentar as definições de ataque entre argumentos levando em consideração fórmulas de confiança sobre tópicos:

Definição 3.14 (Ataque sobre sinceridade em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K}, \mathcal{AC}, \mathcal{TP}\rangle$. Um argumento A ataca B em sua sinceridade sobre um tópico t se e somente se $h = Bel_i(Inf_{j,i}\phi \wedge \neg Bel_j\phi)$, $ConfSinc(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.15 (Ataque sobre validade em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K}, \mathcal{AC}, \mathcal{TP}\rangle$. Um argumento A ataca B em sua validade sobre um tópico t se e somente se $h = Bel_i(Inf_{j,i}\phi \wedge \neg\phi)$, $ConfVal(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.16 (Ataque sobre completude em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K}, \mathcal{AC}, \mathcal{TP}\rangle$. Um argumento A ataca B em sua completude sobre um tópico t se e somente se $h = Bel_i(\phi \wedge \neg Inf_{j,i}\phi)$, $ConfComple(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.17 (Ataque sobre cooperatividade em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$. Um argumento A ataca B em sua cooperatividade sobre um tópico t se e somente se $h = Bel_i(Bel_j\phi \wedge \neg Inf_{j,i}\phi)$, $ConfCoop(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.18 (Ataque sobre competência em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$. Um argumento A ataca B em sua competência sobre um tópico t se e somente se $h = Bel_i(Bel_j\phi \wedge \neg\phi)$, $ConfCompet(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.19 (Ataque sobre vigilância em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$. Um argumento A ataca B em sua vigilância sobre um tópico t se e somente se $h = Bel_i(\phi \wedge \neg Bel_j\phi)$, $ConfVig(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Exemplo 3.20. *Seja $DM = \langle\, \langle \{ConfSinc(i, j, t), Bel_i(\neg Bel_j\psi\}, ConfVal(i, j, Bel_j\psi) \rangle, \{Inf_{j,i}\, \psi, Inf_{j,i}Bel_j\psi\}, \{\mathfrak{A}(t, \phi), \mathfrak{A}(t, \psi)\} \rangle$ uma Descrição de Mundo dos agentes i e j. A partir de DM, podemos construir os seguintes argumentos:*

- $A_1 = (\{Bel_i(Inf_{j,i}\psi), Bel_i(\neg Bel_j\psi)\}, Bel_i(Inf_{j,i}\psi) \wedge Bel_i(\neg Bel_j\psi)\})$
- $A_2 = (\{ConfSinc(i, j, t)\},\ ConfSinc(i, j, \phi))$
- $A_3 = (\{ConfVal(i, j, Bel_j\psi),\ Bel_iInf_{j,i}Bel_j\psi\},\ Bel_iBel_j\psi))$

Podemos notar que o argumento A_1 realiza um ataque sobre a sinceridade do argumento A_2 e o argumento A_3 realiza um ataque sobre premissa em A_1.

O exemplo acima ilustra um ataque sobre a sinceridade entre dois argumentos. Podemos notar que o argumento A_1 conclui $Bel_i(Inf_{j,i}\psi) \wedge Bel_i(\neg Bel_j\psi)$ e A_2 possui em seu suporte $ConfSinc(i, j, t)$, caracterizando um ataque de A_1 sobre a sinceridade em A_2, uma vez que o argumento A_1 conclui uma situação que mostra j falhando na sinceridade para com i em relação a ψ, que é uma fórmula pertencente ao tópico t. Além disso, o argumento A_3, com conclusão $Bel_iBel_j\psi$ efetua um ataque sobre premissa no suporte de A_1.

Repare que as relações de ataque entre argumentos envolvendo as fórmulas de confiança em tópicos diferem da relação de ataque entre argumentos sobre as premissas. Na relação de ataque entre as premissas, a conclusão do argumento que ataca nega um elemento do suporte do argumento que sofre o ataque. Por outro lado, na relação de ataque entre as fórmulas de confiança, a conclusão do argumento que ataca é uma crença que nega uma confiança pertencente ao suporte do argumento que sofre o ataque a partir de uma fórmula que pertença ao mesmo tópico da confiança no suporte.

As definições de ataque apresentadas acima serão utilizadas na construção de uma relação entre argumentos que também captura a relação apresentada na Definição 2.29. Essa nova relação irá ser utilizada mais adiante para que possamos estender o sistema de argumentação apresentado no capítulo anterior.

Definição 3.21 (Relação $\mathcal{R}$)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG} \rangle$ as primitivas sintáticas de $\mathcal{LCT}$, $\mathcal{L}$ a linguagem definida sobre essas primitivas e $D = \langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$ a Descrição de Mundo construída sobre $\mathcal{L}$. Considere (H, h) e (H', h') dois argumentos construídos a partir da Descrição de Mundo D. Dizemos que $((H, h), (H', h')) \in \mathcal{R}$ se e somente se:*

- $((H,h),(H',h')) \in \Re$ *segundo a Definição 2.29, ou*
- (H,h) *ataca sobre sinceridade em tópicos* (H',h')*, ou*
- (H,h) *ataca sobre validade em tópicos* (H',h')*, ou*
- (H,h) *ataca sobre completude em tópicos* (H',h')*, ou*
- (H,h) *ataca sobre cooperatividade em tópicos* (H',h')*, ou*
- (H,h) *ataca sobre competência em tópicos* (H',h')*, ou*
- (H,h) *ataca sobre vigilância em tópicos* (H',h')*.*

Com isso, tratamos da inclusão de tópicos à linguagem da confiança, permitindo a modelagem e raciocínio sobre essas fórmulas por meio de um framework de argumentação. Na próxima seção, fecharemos o capítulo tratando a desconfiança de um agente em outro em relação a fórmulas e a tópicos.

3.2 Desconfiança

Nesta subseção, iremos trabalhar sobre a segunda contribuição presente neste capítulo, que envolve a introdução da noção de desconfiança em $\mathcal{LCT}$. Em situações do mundo real, decisões de um grupo são passíveis de serem tomadas a partir de informações falsas. Essas informações podem ser geradas a partir da ignorância, ingenuidade ou até mesmo desonestidade de um agente que busca seus próprios interesses acima dos interesses do grupo. Para tratar esse tipo de situação, iremos formalizar a relação de desconfiança entre agentes sobre fórmulas e tópicos.

Segundo (MARSH; DIBBEN, 2005), a Desconfiança é caracterizada pela crença que um agente possui de que um outro agente, aquele a quem a desconfiança é dirigida, irá trabalhar ativamente contra ele em uma dada situação. Formalizar a Desconfiança entre agentes com relação a uma fórmula e a um tópico permitirá o raciocínio sobre esse tipo de situação, tornando $\mathcal{LCT}$ mais expressiva.

Nesta seção, iremos introduzir as fórmulas que representam a Desconfiança entre agentes em fórmulas e em tópicos na Lógica de Confiança em Tópicos, $\mathcal{LCT}$. Antes de partirmos para a formalização desta extensão de $\mathcal{LCT}$, adaptaremos mais uma vez o Exemplo 2.1, introduzindo uma situação em que a desconfiança entre agentes em um tópico pode ser decisiva na tomada de uma decisão.

Exemplo 3.22. *Alexandre, Beatriz e Carlos são médicos que estão buscando chegar a uma conclusão sobre a enfermidade de Patrícia. Em uma reunião, os médicos decidem apresentar suas conclusões sobre qual seria a doença que afeta Patrícia. Alexandre afirma que a paciente sofre de meningite, pois possui os sintomas de febre alta, forte dor de cabeça, fotossensibilidade e rachaduras na pele. Prontamente, Carlos discorda argumentando que a paciente não possui rachaduras na pele, de acordo com os primeiros exames realizados nela. A afirmação de Carlos coloca em xeque o diagnóstico de Alexandre, pois a ausência de fissuras na pele poderia indicar que a paciente estaria sofrendo de outras patologias, visto que febre, dor de cabeça e fotossensibilidade também são sintomas de várias outras doenças. Porém, antes da reunião, Alexandre e Beatriz haviam conversado e Alexandre a informou que nos últimos exames a paciente havia apresentado rachaduras na pele. Beatriz desconfia da veracidade da informação de Alexandre, pois, uma vez que Alexandre é um médico infectologista, ele pode estar usando o* *caso de Patrícia para angariar fundos para o seu departamento. Por conseguinte, ela acredita no oposto da informação que seu colega a passou. Ela, então, afirma que a paciente não possui fissuras na pele. Como as colocações de Beatriz e Carlos contradizem o diagnóstico de Alexandre, o grupo decide que a paciente não sofre de meningite.*

O exemplo acima ilustra uma situação em que a não confiança somente não seria capaz de produzir informação suficiente para uma tomada de decisão, visto que Beatriz não poderia tirar nenhuma conclusão apenas com a ausência de confiança em Alexandre. Dentro dessa situação, iremos introduzir à linguagem uma relação para representar a noção de desconfiança, que não pode ser tratada meramente como a ausência de confiança, uma vez que não acrescenta nenhuma crença enquanto, segundo (MARSH; DIBBEN, 2005), a desconfiança possibilita que a acredite que um indivíduo b irá trabalhar ativamente contra a.

Neste trabalho, iremos adaptar essa noção definindo que a desconfiança é a medida na qual um indivíduo acredita que outro desconhece as verdades do mundo ou simplesmente não revela o que suas reais crenças pregam, buscando obter algum ganho com isso. Por exemplo, se um agente i desconfia da sinceridade de um agente j com relação a uma fórmula ϕ, i irá acreditar que j crê na negação de ϕ se j informar i sobre ϕ. Levando em consideração a confiança e a desconfiança entre agentes em relação a fórmulas e tópicos, podemos encontrar o ponto de vista do grupo, utilizando fórmulas que melhor representam o conjunto de crenças dos agentes.

A seguir, iremos mostrar como podemos definir os tipos de desconfiança de maneira similar às definições dos tipos de confiança apresentadas na seção anterior.

3.2.1 Categorias de desconfiança

Nesta subseção iremos introduzir as definições de desconfiança conforme a intuição mostrada na seção anterior.

- **Desconfiança na Sinceridade:** É o tipo de desconfiança que define a crença de um agente i em um agente j informar i sobre uma fórmula em que j não acredita. Por exemplo, suponha que João desconfia da sinceridade de Maria sobre ela estar com fome; logo, se ela informar isso, João acreditará que ela não está com fome. De modo intuitivo, se um agente i desconfia da sinceridade de um agente j sobre uma fórmula ϕ, o agente i irá acreditar que se ele for informado por j sobre ϕ, j não acredita em ϕ:

$$DesConfSinc(i, j, \phi) ::= Bel_i(Inf_{j,i}\phi \rightarrow \neg Bel_j\phi)$$

- **Desconfiança na Validade:** É o tipo de desconfiança que define a crença de um agente i em um agente j informar i sobre uma fórmula que não é verdadeira. Por exemplo, se Maria desconfia da validade de João sobre o mercado não abrir no feriado; então, se João a informar isso, Maria irá acreditar que o mercado estará aberto. De modo intuitivo, se um agente i desconfia da validade de um agente j sobre uma fórmula ϕ, o agente i irá acreditar que se ele for informado por j sobre ϕ, ϕ é falsa:

$$DesConfVal(i, j, \phi) ::= Bel_i(Inf_{j,i}\phi \rightarrow \neg\phi)$$

- **Desconfiança na Completude:** É o tipo de desconfiança que define a crença de um agente i em um agente j não informar i sobre uma fórmula que é verdadeira. Por exemplo, se João desconfiar da completude de Maria sobre o controle remoto da televisão não estar sobre a mesa, então se o controle estiver sobre a mesa, João acreditará que Maria irá informá-lo sobre isso. De modo intuitivo, se um agente i desconfia da completude de um agente j sobre uma fórmula ϕ, o agente i irá acreditar que se ϕ for falsa, j irá informá-lo sobre ϕ ser verdadeira:

$$DesConfComple(i, j, \phi) ::= Bel_i(\neg\phi \rightarrow Inf_{j,i}\phi)$$

- **Desconfiança na Cooperatividade:** É o tipo de desconfiança que define a crença de um agente i em um agente j não irá informar i sobre uma fórmula em que j acredita. Por exemplo, se Maria desconfiar da cooperatividade de João sobre a chave do carro estar no porta-chaves, então se João não souber que as chaves do carro estão no porta-chaves, Maria irá acreditar ele irá informá-la que a chave está lá. De modo intuitivo, se um agente i desconfia da cooperatividade de um agente j, o agente i irá acreditar que se j não possui crença em uma fórmula ϕ, j irá informá-lo sobre ϕ:

$$DesConfCoop(i, j, \phi) ::= Bel_i(\neg Bel_j\phi \rightarrow Inf_{j,i}\phi)$$

- **Desconfiança na Competência:** É o tipo de desconfiança que define a crença de um agente i em um agente j sobre j crer em uma fórmula, ela deve ser considerada falsa. Por exemplo, se João desconfia da competência de Maria sobre a alta da cotação do iene em relação à rúpia nepalesa, então se Maria não souber da alta na cotação do iene em relação à rúpia nepalesa, João acreditará que esta informação é verdadeira. De modo intuitivo, se um agente i desconfia da competência de um agente j em relação a uma fórmula ϕ, o agente i irá acreditar que se j possui crença em ϕ, então ϕ não é verdadeira:

$$DesConfCompet(i, j, \phi) ::= Bel_i(Bel_j\phi \rightarrow \neg\phi)$$

- **Desconfiança na Vigilância:** É o tipo de desconfiança que define a crença de um agente i em um agente j sobre j crer em uma fórmula, ela deve ser considerada falsa. Por exemplo, se Maria desconfiar da vigilância de João sobre a baixa na cotação da rúpia nepalesa em relação ao iene, então se a baixa na cotação da rúpia nepalesa em relação ao iene for verdadeira, Maria acreditará que João não tem conhecimento disso. De modo intuitivo, se um agente i confia na vigilância de um agente j sobre uma fórmula ϕ, o agente i irá acreditar que se ϕ for verdadeira, j não acredita em ϕ:

$$DesConfVig(i, j, \phi) ::= Bel_i(\phi \rightarrow \neg Bel_j\phi)$$

Repare que as fórmulas que representam a desconfiança na competência e na vigilância são equivalentes. Além disso, de acordo com a intuição de desconfiança, sua formalização não é equivalente à ausência de confiança. Por exemplo, a negação da confiança de i na sinceridade de j em relação a ϕ é equivalente a $Bel_iInf_{j,i}\phi \wedge \neg Bel_iBel_j\phi$, que não equivale à fórmula que representa a desconfiança na sinceridade, $\neg(Bel_iInf_{j,i}\phi \wedge Bel_iBel_j\phi)$. Mais resultados serão demonstrados formalmente mais adiante.

Inspirados pelas definições de ataque à confiança exploradas no capítulo anterior, iremos considerar as definições de ataque à desconfiança entre agentes em relação a fórmulas. Um argumento $A = \langle H_a, h_a\rangle$ irá atacar um argumento $B = \langle H_b, h_b\rangle$ por meio de uma desconfiança em seu suporte se e somente se existir uma fórmula de desconfiança equivalente a

$\chi(i,j,\phi) \in H_b$ e h_a for equivalente a $\neg\chi(i,j,\psi)$, em que $\chi \in \{DesConfSinc, DesConfVal, DesConfComple, DesConfCoop, DesConfCompet, DesConfVig\}$. As definições de ataque à desconfiança para cada um dos tipos de confiança serão apresentadas a seguir:

Definição 3.23 (Ataque sobre desconfiança na sinceridade)**.** *Sejam $A = (H,h)$ e $B = (H',h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K},\mathcal{AC},\mathcal{TP}\rangle$. Um argumento A ataca B em sua desconfiança na sinceridade sobre uma fórmula ϕ se e somente se $h = Bel_i(Inf_{j,i}\psi \wedge Bel_j\psi)$ e $DesConfSinc(i,j,\phi) \in H'$.*

Definição 3.24 (Ataque sobre desconfiança na validade)**.** *Sejam $A = (H,h)$ e $B = (H',h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K},\mathcal{AC},\mathcal{TP}\rangle$. Um argumento A ataca B em sua desconfiança na validade sobre uma fórmula ϕ se e somente se $h = Bel_i(Inf_{j,i}\psi \wedge \psi)$ e $DesConfVal(i,j,\phi) \in H'$.*

Definição 3.25 (Ataque sobre desconfiança na completude)**.** *Sejam $A = (H,h)$ e $B = (H',h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K},\mathcal{AC},\mathcal{TP}\rangle$. Um argumento A ataca B em sua desconfiança na completude sobre uma fórmula ϕ se e somente se $h = Bel_i(\neg\psi \wedge \neg Inf_{j,i}\psi)$ e $DesConfComple(i,j,\phi) \in H'$.*

Definição 3.26 (Ataque sobre desconfiança na cooperatividade)**.** *Sejam $A = (H,h)$ e $B = (H',h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K},\mathcal{AC},\mathcal{TP}\rangle$. Um argumento A ataca B em sua desconfiança na cooperatividade sobre uma fórmula ϕ se e somente se $h = Bel_i(\neg Bel_j\psi \wedge \neg Inf_{j,i}\psi)$ e $DesConfCoop\ (i,j,\phi) \in H'$.*

Definição 3.27 (Ataque sobre desconfiança na competência)**.** *Sejam $A = (H,h)$ e $B = (H',h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K},\mathcal{AC},\mathcal{TP}\rangle$. Um argumento A ataca B em sua desconfiança na competência sobre uma fórmula ϕ se e somente se $h = Bel_i(Bel_j\psi \wedge \psi)$ e $DesConfCompet(i,j,\phi) \in H'$.*

Definição 3.28 (Ataque sobre desconfiança na vigilância)**.** *Sejam $A = (H,h)$ e $B = (H',h')$ dois argumentos construídos da descrição de mundo $\langle\mathcal{K},\mathcal{AC},\mathcal{TP}\rangle$. Um argumento A ataca B em sua desconfiança na vigilância sobre uma fórmula ϕ se e somente se $h = Bel_i(\psi \wedge Bel_j\psi)$ e $DesConfVig(i,j,\phi) \in H'$.*

A intuição por trás desse tipo de ataque é que a desconfiança em um determinado agente sobre uma fórmula é posta em cheque quando é apresentada uma situação a qual esse agente agiu de forma sincera.

Seguindo as definições de ataque anteriores, as definições apresentadas acima serão utilizadas na construção de uma relação entre argumentos que também captura a relação apresentada na Definição 3.21.

Definição 3.29 (Relação $\mathcal{RD}$)**.** *Sejam $\langle\mathcal{AT},\mathcal{AG}\rangle$ as primitivas sintáticas de $\mathcal{LCT}$, $\mathcal{L}$ a linguagem definida sobre essas primitivas e $D = \langle\mathcal{K},\mathcal{AC},\mathcal{TP}\rangle$ a Descrição de Mundo construída sobre $\mathcal{L}$. Considere (H,h) e (H',h') dois argumentos construídos a partir da Descrição de Mundo D. Dizemos que $((H,h),(H',h')) \in \mathcal{RD}$ se e somente se:*

- *$((H,h),(H',h')) \in \mathcal{R}$ segundo a Definição 3.21, ou*
- *(H,h) ataca sobre desconfiança na sinceridade (H',h'), ou*

- (H, h) *ataca sobre desconfiança na validade* (H', h')*, ou*
- (H, h) *ataca sobre desconfiança na completude* (H', h')*, ou*
- (H, h) *ataca sobre desconfiança na cooperatividade* (H', h')*, ou*
- (H, h) *ataca sobre desconfiança na competência* (H', h')*, ou*
- (H, h) *ataca sobre desconfiança na vigilância* (H', h')*.*

Da mesma maneira que estendemos a Lógica da Confiança para introduzir as fórmulas que representam a confiança sobre tópicos, podemos também incluir as definições de desconfiança sobre tópicos na Lógica de Confiança em Tópicos. Com isso poderemos também construir argumentos que possuam em seu suporte e em sua conclusão fórmulas que descrevam a desconfiança entre agentes com relação a tópicos.

3.2.2 Desconfiança sobre tópicos

Na subseção anterior, definimos os tipos de desconfiança de maneira parecida com os tipos de confiança apresentados em (AMGOUD; DEMOLOMBE, 2014). A partir dessas definições, podemos também incorporar a noção de desconfiança entre agentes com relação a tópicos, permitindo expressar a desconfiança em um conjunto de fórmulas sem por em xeque a confiança em outras fórmulas ou tópicos. Por exemplo, é mais provável que um professor de matemática conheça mais teoremas e fórmulas trigonométricas que o professor de geografia, logo, o fato de o professor de geografia se equivocar a respeito de uma fórmula matemática não deve colocar à prova seu conhecimento em geografia, assim como uma desconfiança em seus conhecimentos matemáticos não devem por à prova seus conhecimentos em geografia. Iremos definir formalmente abaixo a desconfiança em um tópico.

Definição 3.30 (Desconfiança em tópicos)**.** *Sejam* $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$ *as primitivas sintáticas de* $\mathcal{LCT}$ *e* $\mathcal{L}$ *a linguagem definida sobre essas primitivas. Considere a descrição de mundo* $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$ *sobre* $\mathcal{L}$*. Para* $i, j \in \mathcal{AG}$ *e* $t \in \mathcal{T}$*, definimos* $\chi(i, j, t)$ *como segue:*

$$\chi(i, j, t) ::= \chi(i, j, \phi) \text{ para todo } \phi \in \mathcal{L} \text{ tal que } \mathfrak{A}(t, \phi) \in \mathcal{TP}$$

em que $\chi \in \{DesConfSinc, DesConfVal, DesConfComple, DesConfCoop, DesConfCompet, DesConfVig\}$.

A construção de argumentos ocorre da mesma maneira apresentada na Definição 3.10, visto que as fórmulas que representam a desconfiança entre agentes sobre tópicos irão estar nas crenças dos agentes e a definição supracitada cobre a construção de argumentos a partir de uma Descrição de Mundo.

Seguindo a metodologia utilizada nas definições relacionadas à confiança em tópicos, iremos definir a seguir as relações de ataque entre argumentos que possuem a desconfiança em tópicos em seu suporte.

Definição 3.31 (Ataque sobre desconfiança na sinceridade em tópicos)**.** *Sejam* $A = (H, h)$ *e* $B = (H', h')$ *dois argumentos construídos da descrição de mundo* $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$*. Um argumento* A *ataca* B *em sua desconfiança na sinceridade sobre um tópico* t *se e somente se* $h = Bel_i(Inf_{j,i}\phi \wedge Bel_j\phi)$*,* $DesConfSinc(i, j, t) \in H'$ *e* $\mathfrak{A}(t, \phi) \in \mathcal{TP}$*.*

Definição 3.32 (Ataque sobre desconfiança na validade em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$. Um argumento A ataca B em sua desconfiança na validade sobre um tópico t se e somente se $h = Bel_i(Inf_{j,i}\phi \wedge \phi)$, $DesConfVal(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.33 (Ataque sobre desconfiança na completude em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$. Um argumento A ataca B em sua desconfiança na completude sobre um tópico t se e somente se $h = Bel_i(\neg\phi \wedge \neg Inf_{j,i}\phi)$, $DesConfComple(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.34 (Ataque sobre desconfiança na cooperatividade em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$. Um argumento A ataca B em sua desconfiança na cooperatividade sobre um tópico t se e somente se $h = Bel_i(\neg Bel_j\phi \wedge \neg Inf_{j,i}\phi)$, $DesConfCoop(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.35 (Ataque sobre desconfiança na competência em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$. Um argumento A ataca B em sua desconfiança na competência sobre um tópico t se e somente se $h = Bel_i(Bel_j\phi \wedge \phi)$, $DesConfCompet(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Definição 3.36 (Ataque sobre desconfiança na vigilância em tópicos)**.** *Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos construídos da descrição de mundo $\langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$. Um argumento A ataca B em sua desconfiança na vigilância sobre um tópico t se e somente se $h = Bel_i(\phi \wedge Bel_j\phi)$, $DesConfVig(i, j, t) \in H'$ e $\mathfrak{A}(t, \phi) \in \mathcal{TP}$.*

Exemplo 3.37. *Seja $DM = \langle\ \langle \{DesConfSinc(i, j, t),\ Bel_i(Bel_j\psi\},\ ConfVal\ (i, j, \neg Bel_j\psi)\rangle$, $\{Inf_{j,i}\psi,\ Inf_{j,i}\neg Bel_j\psi\}$, $\{\mathfrak{A}(t, \phi),\ \mathfrak{A}(t, \psi)\}\rangle$ uma Descrição de Mundo dos agentes i e j. A partir de DM, podemos construir os seguintes argumentos:*

- $A_1 = (\{Bel_i(Inf_{j,i}\psi), Bel_i(Bel_j\psi)\}, Bel_i(Inf_{j,i}\psi) \wedge Bel_i(Bel_j\psi))$
- $A_2 = (\{DesConfSinc(i, j, t)\},\ DesConfSinc(i, j, \phi))$
- $A_3 = (\{ConfVal(i, j, \neg Bel_j\psi),\ Bel_i Inf_{j,i}\neg Bel_j\psi\},\ Bel_i\neg Bel_j\psi))$

Podemos notar que o argumento A_1 realiza um ataque sobre a desconfiança na sinceridade do argumento A_2 e o argumento A_3 realiza um ataque sobre premissa em A_1.

O exemplo acima ilustra um ataque sobre a desconfiança na sinceridade entre dois argumentos. Podemos notar que o argumento A_1 conclui $Bel_i(Inf_{j,i}\psi) \wedge Bel_i(Bel_j\psi)$ e A_2 possui em seu suporte $DesConfSinc(i, j, t)$, caracterizando um ataque de A_1 sobre a desconfiança sinceridade em A_2, uma vez que o argumento A_1 conclui uma situação que mostra j sendo sincero para com i em relação a ψ, que é uma fórmula pertencente ao tópico t. Além disso, o argumento A_3, com conclusão $Bel_i\neg Bel_j\psi$ efetua um ataque sobre premissa no suporte de A_1.

As definições de ataque sobre a desconfiança apresentadas acima serão utilizadas na construção de uma relação de ataque entre argumentos que também captura a relação de ataque apresentada na Definição 3.29. De maneira similar ao sistema de argumentação do capítulo anterior, a relação que será mostrada a seguir será utilizada para construir um framework de argumentação.

Definição 3.38 (Relação $\Re$)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG} \rangle$ as primitivas sintáticas de $\mathcal{LC}$, $\mathcal{L}$ a linguagem definida sobre essas primitivas e $DM = \langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$ a Descrição de Mundo construída sobre $\mathcal{L}$. Considere (H, h) e (H', h') dois argumentos construídos a partir da Descrição de Mundo DM. Dizemos que $((H, h), (H', h')) \in \Re$ se e somente se:*

- *$((H, h), (H', h')) \in \mathcal{R}$ segundo a definição 3.29, ou*
- *(H, h) ataca sobre desconfiança na sinceridade em tópicos (H', h'), ou*
- *(H, h) ataca sobre desconfiança na validade em tópicos (H', h'), ou*
- *(H, h) ataca sobre desconfiança na completude em tópicos (H', h'), ou*
- *(H, h) ataca sobre desconfiança na cooperatividade em tópicos (H', h'), ou*
- *(H, h) ataca sobre desconfiança na competência em tópicos (H', h'), ou*
- *(H, h) ataca sobre desconfiança na vigilância em tópicos (H', h').*

A exemplo do capítulo anterior, iremos encerrar com a definição de um Framework de Argumentação que será empregado para expressar o ponto de vista de um conjunto de agentes.

3.2.3 Raciocinando com Frameworks de Argumentação

Podemos utilizar Frameworks de Argumentação quando desejamos defender ou atacar um ponto de vista, objetivo ou ação a partir de um conjunto de crenças. Aqui queremos construir um Framework de Argumentação a partir das crenças presentes em uma Descrição de Mundo D. Abaixo, iremos definir o Framework com o conjunto de argumentos conforme a Definição 3.10 e a relação de ataque definida em 3.38.

Definição 3.39 (Framework de Argumentação sobre crenças)**.** *Sejam $\langle \mathcal{AT}, \mathcal{AG}, \mathcal{T} \rangle$ as primitivas sintáticas de $\mathcal{LC}$, $\mathcal{L}$ a linguagem definida sobre essas primitivas e $DM = \langle \mathcal{K}, \mathcal{AC}, \mathcal{TP} \rangle$ a Descrição de Mundo construída sobre $\mathcal{L}$. Um framework de argumentação construído sobre DM é um par $\mathcal{T} = (S, \Re)$ em que $S \subseteq Arg(DM)$ com $Arg(DM)$ o conjunto definido em 3.10 e $\Re \subseteq S \times S$ é a relação da definição 3.38.*

Uma vez que os argumentos podem ser conflitantes, devemos encontrar um meio de avalia-los e decidir quais seriam os argumentos aceitáveis. Como vimos anteriormente, as semânticas de Dung podem ser utilizadas para decidir quais são esses argumentos aceitáveis dentro de um Framework de Argumentação. Nesse sistema, iremos utilizar as extensões estáveis por conta de elas permitirem particionar os subconjuntos do conjunto de argumentos em dois conjuntos: as extensões estáveis e as não extensões. As extensões extraídas a partir do Framework de Argumentação irão permitir definir quais fórmulas poderão ser inferidas a partir de uma Descrição de Mundo. Essas fórmulas inferidas indicam as crenças do grupo a partir da Descrição de Mundo. Basicamente, para uma fórmula ser inferida, ela deve ser a conclusão de pelo menos um argumento de cada uma das extensões estáveis do framework de argumentação. Vale ressaltar que o argumento não precisa ser o mesmo em todas as extensões.

Definição 3.40 (Resultado de um Framework de Argumentação $\mathcal{T}$ (AMGOUD; DEMOLOMBE, 2014))**.** *Sejam $\mathcal{T} = (S, \Re)$ um framework de argumentação conforme a Definição 3.39 construído sobre uma Descrição de Mundo DM, $S \subseteq Arg(DM)$ um conjunto finito de argumentos e $Ext(\mathcal{T})$ o conjunto de todas as extensões estáveis de $\mathcal{T}$. Uma fórmula ϕ é inferida de DM se* *e somente se para todo $\mathcal{E} \in Ext(\mathcal{T})$, existe um $\langle H, \phi \rangle \in \mathcal{E}$. O resultado de um framework de argumentação $\mathcal{T}$ é o conjunto de todas as fórmulas ϕ inferidas de $\mathcal{T}$.*

Exemplo 3.41. *Seja $DM = \langle \langle K_a, K_b, K_c \rangle, \{Inf_{a,b}\phi\}, \{\mathfrak{A}(t, \phi)\} \rangle$ uma Descrição de Mundo em que:*

- $K_a = \{Bel_a\alpha \wedge \beta \wedge \gamma, Bel_a\alpha \wedge \beta \wedge \gamma \wedge \psi \rightarrow \epsilon, Bel_a\phi, Bel_a\phi \rightarrow \psi, Bel_a\neg\psi \rightarrow \neg\epsilon\}$
- $K_b = \{Bel_b\alpha \wedge \beta \wedge \gamma, Bel_b\alpha \wedge \beta \wedge \gamma \wedge \psi \rightarrow \epsilon, DesConfVal(b, a, \phi), Bel_b\phi \rightarrow \psi, Bel_b\neg\psi \rightarrow \neg\epsilon\}$
- $K_c = \{Bel_c\alpha \wedge \beta \wedge \gamma, Bel_c\alpha \wedge \beta \wedge \gamma \wedge \psi \rightarrow \epsilon, Bel_c\neg\psi, Bel_c\phi \rightarrow \psi, Bel_c\neg\psi \rightarrow \neg\epsilon\}$

Considere os seguintes argumentos construídos a partir de DM:

- $A_1 = (\{\phi \rightarrow \psi, \phi, \alpha \wedge \beta \wedge \gamma, \alpha \wedge \beta \wedge \gamma \wedge \psi \rightarrow \epsilon\}, \epsilon)$
- $A_2 = (\{\neg\psi, \neg\psi \rightarrow \neg\epsilon\}, \neg\epsilon)$
- $A_3 = (\{Inf_{a,b}\phi, ConfVal(b, a, t)\}, \neg\phi)$

O resultado do Framework de Argumentação $\mathcal{T} = (S, \mathfrak{R})$ é $\{\neg\phi, \neg\epsilon\}$, pois a única extensão estável $\{A_2, A_3\}$ possui essas fórmulas como conclusão de seus argumentos.

Considerando no exemplo acima pode ser visto como a formalização da situação de motivação vista no início do capítulo, desde que ϕ seja interpretado por "A paciente foi examinada"; ψ, por "A paciente possui fissuras na pele"; γ, por "A paciente apresenta febre alta"; β, por "A paciente apresenta dor de cabeça"; γ, por "A paciente apresenta fotossensibilidade" e ϵ, por "A paciente está com meningite".

3.3 Exemplo

De maneira similar a (**??**), iremos nos basear na tragédia de *Shakespeare*, *Otelo, o Mouro de Veneza*, para ilustrar o sistema de argumentação proposto nesta Dissertação de Mestrado uma vez que as personagens desenvolvem confiança e desconfiança uns pelos outros em todo o decorrer da trama. A obra conta a história de Otelo, um general mouro que serve o reino de Veneza, sua esposa Desdêmona, seu tenente Cássio e seu sub-oficial Iago, abordando diferentes temas tais como ciúme, inveja, amor, racismo e traição.

A trama começa com Rodrigo, um cavalheiro rico e dissoluto, queixando-se de Iago, um sub-oficial de Otelo, que Iago não contou a ele sobre o casamento secreto entre Desdêmona, a filha de um senador chamado Brabâncio, e Otelo. Rodrigo está furioso com essa união, porque ele ama Desdêmona e anteriormente havia pedido a Brabâncio a mão dela em casamento. Iago está chateado com Otelo por promover Cássio, um homem mais jovem, a um cargo superior e diz a Rodrigo que planeja usar Otelo para sua própria vantagem. Iago tenta de muitas maneiras ter Cássio despojado de sua posição, mas ele também quer arruinar Otelo por ter preferido Cássio a ele. Iago convence Otelo a ter suspeita de um romance entre Cássio e Desdêmona colocando um lenço que foi o primeiro presente de Otelo para Desdêmona nos pertences de Cássio, com a ajuda de sua esposa Emília. A persuasão de Iago leva Otelo a matar Desdêmona em um ataque de fúria. No epílogo, Emília revela que a suposta traição de Desdêmona foi inventada por Iago, que imediatamente a mata. Otelo, sentindo-se culpado por matar a inocente Desdêmona, se mata. Iago é preso e Cássio fica com o posto de general que pertencia a Otelo.

A história é dividida em cinco atos e em cada um deles, com exceção dos atos III e IV, que serão unificados, existe uma decisão central a ser tomada por alguma das personagens.

A sinopse a seguir foi retirada de (**??**). Iremos modificar ligeiramente a trama a fim adaptar as decisões tomadas ao sistema de argumentação proposto.

3.3.1 Sinopse - Ato I

O Ato I se inicia com Iago, sub-oficial de Otelo, tramando com Rodrigo uma forma de contar a Brabâncio, rico senador de Veneza, que sua filha, a gentil Desdêmona, tinha relações íntimas com Otelo. Iago queria vingar-se do general Otelo porque ele promoveu Cássio, jovem soldado florentino e grande intermediário nas relações entre Otelo e Desdêmona, ao posto de tenente. Esse ato deixou Iago muito ofendido, uma vez que acreditava que as promoções deveriam ser obtidas "pelos velhos meios em que herdava sempre o segundo o posto do primeiro" e não por amizades.

Brabâncio, que deixara a filha livre para escolher o marido que mais lhe agradasse, acreditava que ela escolheria, para seu cônjuge, um homem da classe senatorial ou de semelhante. Ao tomar ciência que sua filha havia fugido para se casar com o Otelo, foi à procura dele para matá-lo. No momento em que se encontraram, chegou um comunicado do Duque de Veneza, convocando-os para uma reunião de caráter urgente no senado.

Durante a reunião, Brabâncio, sem provas, acusou o Otelo de ter induzido Desdêmona a casar-se com ele por meio de bruxarias. Otelo, que era general do reino de Veneza e gozava da estima e da confiança do Estado por ser leal, muito corajoso e ter atitudes nobres, fez, em sua defesa, um simples relato da sua história de amor que foi confirmado pela própria Desdêmona. Por isso, e por ser o único capaz de conduzir um exército no contra-ataque a uma esquadra turca que se dirigia à ilha de Chipre, Otelo foi inocentado e o casal seguiu para Chipre, em barcos separados, na manhã seguinte.

3.3.2 Modelagem - Ato I

A principal decisão presente no ato I é tomada pelo Duque de Veneza. Estando no papel de julgar Otelo diante das acusações de Brabâncio, o Duque deve confiar na validade das informações dadas por ambas as partes conforme a Descrição de Mundo $DM_1 = \langle \mathcal{K}^1, \mathcal{AC}^1, \mathcal{TP}^1 \rangle$ construída sobre as seguintes primitivas sintáticas: du representando o Duque; de, Desdêmona; o, Otelo; b, Brabâncio; $a_{de,o}$, Desdêmona ama Otelo; $e_{o,de}$, Otelo enfeitiçou Desdêmona e t_1 é um tópico referente à relação entre Otelo e Desdêmona. Assim, temos que

- $\mathcal{K}^1 = \{K^1_{du}, K^1_b, K^1_{de}, K^1_o\}$
- $K^1_{du} = \{ConfVal(du, b, t_1), ConfVal(du, de, t_1), ConfVal(du, o, t_1)\}$
- $K^1_b = \{Bel_b(\neg a_{de,o} \to e_{o,de}), Bel_b(\neg a_{de,o})\}$
- $K^1_{de} = \{Bel_{de} a_{de,o}\}$
- $K^1_o = \{Bel_o a_{de,o}\}$
- $\mathcal{AC}^1 = \{Inf_{b,du}(\neg a_{de,o} \to e_{o,de}), Inf_{b,du}(\neg a_{de,o}), Inf_{de,du}((Inf_{b,du} \neg a_{de,o}) \wedge a_{de,o}), Inf_{o,du}$ $((Inf_{b,du} \neg a_{de,o}) \wedge a_{de,o})\}$
- $\mathcal{TP}^1 = \{\mathfrak{A}(t_1, a_{de,o})\}$

A partir de DM_1, os seguintes argumentos podem ser construídos:

- $A_1 = (\{ConfVal(du, b, t_1), Inf_{b,du}(\neg a_{de,o} \to e_{o,de}), Inf_{b,du}(\neg a_{de,o})\}, e_{o,de})$
- $A_2 = (\{ConfVal(du, de, t_1), Inf_{de,du}(Inf_{b,du} \neg a_{de,o} \wedge a_{de,o})\}), Inf_{b,du} \neg a_{de,o} \wedge a_{de,o}))$

- $A_3 = (\{ConfVal(du, o, t_1), Inf_{o,du}(Inf_{b,du} \neg a_{de,o} \wedge a_{de,o})\}), Inf_{b,du} \neg a_{de,o} \wedge a_{de,o}))$

Logo, podemos construir um framework de argumentação $AF_1 = (S_1, \Re_1)$, tal que $S_1 = \{A_1, A_2, A_3\}$ e $\Re_1 = \{(A_2, A_1), (A_3, A_1)\}$, com a extensão estável $\mathcal{E} = \{A_2, A_3\}$. O resultado desse framework de argumentação é $R_1 = \{Inf_{b,du} \neg a_{de,o} \wedge a_{de,o}\}$, refletindo que Brabâncio informou ao duque que Desdêmona não ama Otelo, mas, de fato, Desdêmona ama Otelo. Com isso, o Duque pôde inocentar Otelo por falta de acusação consistente, já que o argumento de Brabâncio fora derrotado.

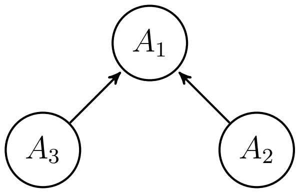

Figura 3 – AF_1

3.3.3 Sinopse - Ato II

Durante a viagem, uma tempestade separou as embarcações e, devido a isso, Desdêmona chegou primeiro à ilha. Algum tempo depois, Otelo desembarca com a novidade que a guerra tinha acabado, porque a esquadra turca fora destruída pela fúria das águas. No entanto, o que o Mouro não sabia é que na ilha ele enfrentaria um inimigo mais fatal do que os turcos.

Em Chipre, Iago, que odiava Otelo e Cássio, começou a semear a discórdia entre os dois, dando início um terrível plano de vingança que tinha como objetivo arruinar seus inimigos. Hábil e profundo conhecedor da natureza humana, e sempre fazendo reflexões sobre a humanidade, Iago sabia que de todos os tormentos que afligem a alma, o ciúme é o mais intolerável e incontrolável.

Ele sabia que Cássio, entre os amigos de Otelo, era o que mais possuía a confiança do general. Sabia também que devido a sua beleza e eloquência, qualidades que agradam às mulheres, ele era exatamente o tipo de homem capaz de despertar o ciúme de um homem de idade avançada, como era Otelo, casado com uma jovem e bela mulher. Por isso, começou a realizar seu plano.

Sob pretexto de lealdade e estima ao general, Iago induziu Cássio, responsável por manter a ordem e a paz, a se embriagar e envolver-se em uma briga com Rodrigo, durante uma festa que os habitantes da ilha ofereceram a Otelo. Quando o mouro soube do acontecido, destituiu Cássio de seu posto.

3.3.4 Modelagem - Ato II

A decisão central do ato II envolve a destituição da patente de Cássio. Por estar embriagado, Cássio se deixou levar por Iago e atacou Rodrigo acreditando que estaria mantendo a ordem no local. Otelo possuía confiança em Cássio, mas após saber por meio de Iago da briga com Rodrigo no meio da festa, vai deixar de confiar na competência do mesmo, conforme a Descrição de Mundo $DM_2 = \langle \mathcal{K}^2, \mathcal{AC}^2, \mathcal{TP}^2 \rangle$ construída sobre as seguintes primitivas sintáticas:

o representando Otelo; c, Cássio; b, Cássio brigou com Rodrigo; ct Cássio é tenente; mo, manter a ordem e t_2 é um tópico referente à capacidade de Cássio em administrar a ordem pública. Assim, temos que

- $\mathcal{K}^2 = \{K_o^2, K_c\} \cup K^1$
- $K_o^2 = \{Bel_o(ConfCompet(o, c, t_2)), Bel_o(ConfCompet(o, c, t_2) \rightarrow ct, Bel_o\neg(b \rightarrow mo)\} \cup K_o^1$
- $K_i^2 = \{Bel_c(b \rightarrow mo)\} \cup K_i^1$
- $\mathcal{AC}^2 = \{Inf_{i,o}(Bel_c(b \rightarrow mo))\} \cup \mathcal{AC}^1$
- $\mathcal{TP}^2 = \{\mathfrak{A}(t_2, b \rightarrow mo)\} \cup \mathcal{TP}^1$

A partir de DM_2, os seguintes argumentos podem ser construídos:

- $A_4 = (\{ConfCompet(o, c, t_2), ConfCompet(o, c, t_2) \rightarrow ct\}, ct)$
- $A_5 = (\{ConfVal(i, o, t_2), Inf_{i,o}(Bel_c(b \rightarrow mo)), Bel_o\neg(b \rightarrow mo)\}, Bel_c(b \rightarrow mo) \wedge \neg(b \rightarrow mo))$

Logo, podemos construir um framework de argumentação $AF_2 = (S_2, \Re_2)$, tal que $S_2 = \{A_4, A_5\} \cup S_1$ e $\Re_2 = \{(A_5, A_4)\} \cup \Re_1$, com a extensão estável $\mathcal{E} = \{A_2, A_3, A_5\}$, acrescentando ao resultado do framework AF_1 que a confiança de Otelo em Cássio é atacada e, consequentemente, o argumento que conclui que Cássio tem qualificação para ser tentente também é atacado. Logo, $R_2 = \{Bel_c(b \rightarrow mo) \wedge \neg(b \rightarrow mo)\} \cup R_1$ é resultado de AF_2.

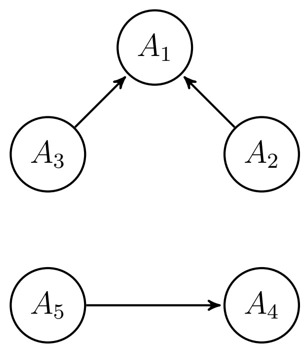

Figura 4 – AF_2

3.3.5 Sinopse - Ato III/IV

Nessa mesma noite, Iago começou a jogar Cássio contra Otelo. Ele falava, dissimulando um certo repúdio a atitude do general, que a sua decisão tinha sido muito dura e que Cassio deveria pedir a Desdêmona que convencesse Otelo a devolver-lhe o posto de tenente. Cássio, abalado emocionalmente, não se deu conta do plano traçado por Iago e aceitou a sugestão.

Dando continuidade a seu plano, Iago insinuou a Otelo que Cássio e sua esposa poderiam estar tendo um caso. Esse plano foi tão bem traçado que Otelo começou a desconfiar de Desdêmona.

Iago sabia que Otelo havia presenteado sua mulher com um velho lenço de linho bordado com morangos, o qual tinha herdado de sua mãe. Iago induz Emília, sua esposa, a roubar o lenço de Desdêmona e diz a Otelo que sua mulher havia presenteado o seu amante com ele. Otelo, já enciumado, pergunta a sua esposa sobre o lenço e ela, ignorando que o lenço estava com Iago, não soube explicar o desaparecimento dele. Nesse ínterim, Iago colocou o lenço dentro do quarto de Cássio para que ele o encontrasse. Depois, Iago fez com que Otelo se escondesse e ouvisse uma conversa sua com Cássio. Eles falaram sobre Bianca, meretriz amante de Cássio, mas como Otelo só ouviu partes da conversa, ficou com a impressão de que eles estavam falando a respeito de Desdêmona. Um pouco depois Bianca chegou e Cassio deu a ela o lenço que encontrara em seu quarto.

3.3.6 Modelagem - Ato III/IV

Nos atos III e IV, Iago arquiteta sua vingança de Otelo colocando-o contra Cássio e Desdêmona, por meio de uma manipulação complexa dos fatos ocorridos em Chipre. Com isso, Otelo passa a crer que Desdêmona foi infiel e decide matá-la, conforme a Descrição de Mundo $DM_3 = \langle \mathcal{K}^3, \mathcal{AC}^3, \mathcal{TP}^3 \rangle$ construída sobre as seguintes primitivas sintáticas: o representando Otelo; c, Cássio; de, Desdêmona; i, Iago; l, Desdêmona sabe que o lenço está com Cássio; $a_{de,o}$, Desdêmona ama Otelo; $tr_{de,c}$, Desdêmona traiu Otelo com Cássio, $m_{o,de}$, Otelo deve matar Desdêmona e t_3 é um tópico referente à relação de Desdêmona e Otelo. Assim, temos que

- $\mathcal{K}^3 = \{K_o^3, K_de^3, K_i^3\} \cup K^2$
- $K_o^3 = \{ConfVal(o, de, t_3), ConfVal(o, i, t_3), l \to DesConfVal(o, de, t_3), \neg a_{de,o} \to tr_{de,c}, tr_{de,c} \to m_{o,de}\} \cup K_o^2$
- $K_{de}^3 = \{Bel_{de} a_{de,o}, Bel_{de} \neg l\}$
- $K_i^3 = \{Bel_i \neg l\}$
- $\mathcal{AC}^3 = \{Inf_{de,o} \neg l, Inf_{i,o} l, Inf_{de,o} a_{de,o}\} \cup \mathcal{AC}^2$
- $\mathcal{TP}^3 = \{\mathfrak{A}(t_3, a_{de,o}), \mathfrak{A}(t_3, l)\} \cup \mathcal{TP}^2$

A partir de DM_3, os seguintes argumentos podem ser construídos:

- $A_6 = (\{ConfVal(o, de, t_3), Inf_{de,o} \neg l\}, \neg l)$
- $A_7 = (\{Inf_{de,o} \neg l, Inf_{i,o} l, ConfVal(o, i, t_3)\}, Inf_{de,o} \neg l \wedge l)$
- $A_8 = (\{ConfVal(o, i, t_3), Inf_{i,o} l, l \to DesConfVal(o, de, t_3), Inf_{de,o} a_{de,o}, \neg a_{de,o} \to tr_{de,c}, tr_{de,c} \to m_{o,de}\}, m_{o,de})$

Logo, podemos construir um framework de argumentação $AF_3 = (S_3, \Re_3)$, tal que $S_3 = \{A_6, A_7, A_8\} \cup S_2$ e $\Re_3 = \{(A_7, A_6)\} \cup \Re_2$, com a extensão estável $\mathcal{E} = \{A_2, A_3, A_5, A_7, A_8\}$, refletindo que Otelo passou a acreditar que deve matar Desdêmona, por tê-lo traído com Cássio. Logo, $R_3 = \{Inf_{de,o} \neg l \wedge l, m_{o,de}\} \cup R_2$ é o resultado de AF_3.

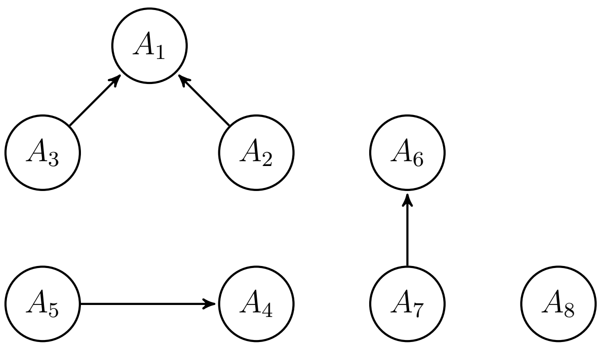

Figura 5 – AF_3

3.3.7 Sinopse - Ato V

Conforme combinado com Iago, Bianca entrega de volta o lenço para Cássio na frente de Otelo fingindo que havia descoberto que o lenço era de Desdêmona. As consequências disso foram terríveis: primeiro Iago, jurando lealdade a seu general, disse que para vingá-lo, mataria Cássio, mas sua real intenção era também matar Rodrigo, porque ele poderia estragar seus planos dizendo a verdade sobre quem estaria por trás de todo o desentendimento. No entanto, isso não ocorreu conforme suas intenções: Rodrigo acaba morrendo e Cássio ferido com uma perna mutilada. Depois Otelo, totalmente descontrolado por ter visto o lenço de Desdêmona com Cássio, foi à procura de sua esposa acreditando que ela o havia traído e a asfixia em seu quarto. Após isso, Emília, esposa de Iago, sabendo que sua senhora fora assassinada, revelou a Otelo, Ludovico (parente de Brabâncio) e Montano (governador de Chipre antes de Otelo) que tudo isso foi tramado por seu marido e que Desdêmona jamais fora infiel.

Iago então mata Emília e foge, mas logo é capturado. Otelo, desesperado ao saber que matara sua amada esposa injustamente, apunhalou-se, caindo sobre o corpo de sua mulher e morreu beijando a quem tanto amara.

Ao finalizar a tragédia Cássio passou a ocupar o lugar de Otelo, Iago foi entregue às autoridades para ser julgado e Ludovico, uma vez que seu irmão Brabâncio morrera, ficou com os bens do mouro.

3.3.8 Modelagem - Ato V

No último ato da tragédia, Otelo opta por tirar a própria vida por ter percebido que matou sua esposa injustamente. Como a confiança que possia em Iago foi quebrada, todos os argumentos que o fizeram tomar a decisão de matar Desdêmona são atacados, conforme a Descrição de Mundo $DM_4 = \langle \mathcal{K}^4, \mathcal{AC}^4, \mathcal{TP}^4 \rangle$ construída sobre as seguintes primitivas sintáticas: o representando Otelo; e, Emília; i, Iago; de, Desdêmona; l, Desdêmona sabe que o lenço está com Cássio; $a_{de,o}$, Desdêmona ama Otelo; $tr_{de,c}$, Desdêmona traiu Otelo com Cássio; $m_{o,de}$, Otelo deve matar Desdêmona, s_o, Otelo deve cometer suicídio e t_3 é um tópico referente à relação de Desdêmona e Otelo. e, portanto, $\mathfrak{A}(t_3, a_{de,o})$, $\mathfrak{A}(t_3, l)$ e $\mathfrak{A}(t_3, a_{de,o})$. Assim, temos que

- $\mathcal{K}^4 = \{K_o^4, K_{de}^4, K_i^4\} \cup K^3$
- $K_o^4 = \{ConfVal(o, de, t_3), ConfVal(o, i, t_3), l \to DesConfVal(o, de, t_3), \neg a_{de,o} \to tr_{de,c}, tr_{de,c} \to m_{o,de}, ConfVal(o, e, t_3), (Inf_{i,o} l \wedge \neg l) \to \neg tr_{de,c}, m_{o,de}, \neg tr_{de,c} \wedge m_{o,de} \to s_o\} \cup K_o^3$
- $K_{de}^4 = K_{de}^3$
- $K_i^4 = K_i^3$
- $K_e^4 = \{K_e \neg l\}$
- $\mathcal{AC}^4 = \{Inf_{e,o} \neg l\} \cup \mathcal{AC}^3$
- $\mathcal{TP}^4 = \mathcal{TP}^3$

A partir de DM_4, os seguintes argumentos podem ser construídos:

- $A_9 = (\{Inf_{i,o} l, ConfVal(o, e, t_3), Inf_{e,o} \neg l\}, Inf_{i,o} l \wedge \neg l)$
- $A_{10} = (\{Inf_{i,o} l, ConfVal(o, e, t_3), Inf_{e,o} \neg l, (Inf_{i,o} l \wedge \neg l) \to \neg tr_{de,c}, m_{o,de}, \neg tr_{de,c} \wedge m_{o,de} \to s_o\}, s_o)$

Logo, podemos construir um framework de argumentação $AF_4 = (S_4, \Re_4)$, tal que $S_4 =$

$\{A_9, A_{10}\} \cup S_3$ e $\Re_4 = \{(A_9, A_8), (A_9, A_7)\} \cup \Re_3$, com a extensão estável $\mathcal{E} = \{A_2, A_3, A_5, A_9, A_{10}\}$, refletindo que a confiança de Otelo em Iago foi atacada e que Otelo irá cometer suicídio por ter matado Desdêmona injustamente. Logo, $R_4 = \{Inf_{i,o}l \wedge \neg l, s_o\} \cup R_3$ é o resultado de AF_4.

Por questão de simplicidade, modelamos o exemplo apresentado nesta subseção com Frameworks de Argumentação com apenas uma extensão estável. Embora tenhamos formalizado dessa forma, é possível calcular o Resultado de um Framework de Argumentação para Frameworks com mais de uma extensão estável.

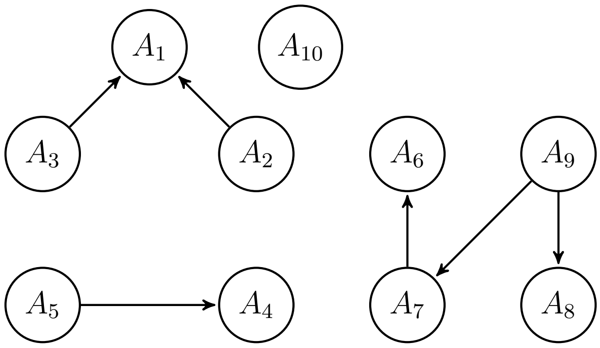

Figura 6 – AF_4

3.4 Propriedades

Nesta seção, iremos abordar algumas propriedades da relação entre as fórmulas de confiança e desconfiança e mostrar a complexidade do sistema de argumentação proposto nesta Dissertação de Mestrado.

3.4.1 Relação Confiança e Desconfiança

O resultado abaixo mostra qual crença precisa ser adicionada para que as fórmulas que representam a confiança e a desconfiança de um mesmo tipo derivem um absurdo:

Teorema 3.42. *Os seguintes conjuntos de fórmulas são inconsistentes:*

- $\{ConfSinc(i, j, \phi),\ DesConfSinc(i, j, \phi),\ Bel_i(Inf_{j,i}\phi)\}$
- $\{ConfVal(i, j, \phi),\ DesConfVal(i, j, \phi),\ Bel_i(Inf_{j,i}\phi)\}$
- $\{ConfComple(i, j, \phi),\ DesConfComple(i, j, \phi),\ Bel_i\neg(Inf_{j,i}\phi)\}$
- $\{ConfCoop(i, j, \phi),\ DesConfCoop(i, j, \phi),\ Bel_i\neg(Inf_{j,i}\phi)\}$
- $\{ConfCompet(i, j, \phi),\ DesConfCompet(i, j, \phi),\ Bel_i(Bel_j\phi)\}$
- $\{ConfVig(i, j, \phi),\ DesConfVig(i, j, \phi),\ Bel_i(\phi)\}$

Demonstração.

- $\{ConfSinc(i, j, \phi),\ DesConfSinc(i, j, \phi),\ Bel_i(Inf_{j,i}\phi)\} \vdash \bot$
  1. $ConfSinc(i, j, \phi)$ premissa
  2. $DesConfSinc(i, j, \phi)$ premissa
  3. $Bel_i(Inf_{j,i}\phi)$ premissa

4. $ConfSinc(i, j, \phi) \rightarrow (Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i(Bel_j\phi))$ — por (K)
5. $Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i(Bel_j\phi)$ — por (PR) sobre (1) e (4)
6. $DesConfSinc(i, j, \phi) \rightarrow (Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i(Bel_j\neg\phi))$ — por (K)
7. $Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\neg(Bel_j\phi)$ — por (PR) sobre (2) e (6)
8. $Bel_i(Bel_j\phi)$ — por (PR) sobre (5) e (3)
9. $Bel_i\neg(Bel_j\phi)$ — por (PR) sobre (7) e (3)
10. $Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi)$ — pela introdução do $\wedge$ sobre (8) e (9)
11. $\neg(Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi))$ — por (D)
12. $\perp$ — pela definição de $\perp$ sobre (10) e (11)

- $\{ConfVal(i, j, \phi), DesConfVal(i, j, \phi), Bel_i(Inf_{j,i}\phi)\} \vdash \perp$

1. $ConfVal(i, j, \phi)$ — premissa
2. $DesConfVal(i, j, \phi)$ — premissa
3. $Bel_i(Inf_{j,i}\phi)$ — premissa
4. $ConfVal(i, j, \phi) \rightarrow Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\phi$ — por (K)
5. $Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\phi$ — por (PR) sobre (1) em (4)
6. $DesConfVal(i, j, \phi) \rightarrow (Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\neg\phi)$ — por (K)
7. $Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\neg\phi$ — por (K) sobre (2) em (6)
8. $Bel_i\phi$ — por (PR) sobre (5) e (3)
9. $Bel_i\neg\phi$ — por (PR) sobre (7) e (3)
10. $Bel_i\phi \wedge Bel_i\neg\phi$ — pela introdução do $\wedge$ sobre (8) e (9)
11. $\neg(Bel_i\phi \wedge Bel_i\neg\phi)$ — por (D)
12. $\perp$ — pela definição de $\perp$ sobre (10) e (11)

- $\{ConfComple(i, j, \phi), DesConfComple(i, j, \phi), Bel_i\neg(Inf_{j,i}\phi)\} \vdash \perp$

1. $ConfComple(i, j, \phi)$ — premissa
2. $DesConfComple(i, j, \phi)$ — premissa
3. $Bel_i\neg(Inf_{j,i}\phi)$ — premissa
4. $Bel_i(\phi \rightarrow Inf_{j,i}\phi)$ — pela definição de (1)
5. $Bel_i(\neg\phi \rightarrow Inf_{j,i}\phi)$ — pela definição de (2)
6. $Bel_i(\neg(Inf_{j,i}\phi) \rightarrow \neg\phi)$ — por (P3) sobre (4)
7. $Bel_i(\neg(Inf_{j,i}\phi) \rightarrow \neg\phi) \rightarrow (Bel_i(\neg(Inf_{j,i}\phi)) \rightarrow Bel_i(\neg\phi))$ — por (K)
8. $Bel_i(\neg(Inf_{j,i}\phi) \rightarrow \phi)$ — por (P3) sobre (5)
9. $Bel_i(\neg(Inf_{j,i}\phi) \rightarrow \phi) \rightarrow (Bel_i(\neg(Inf_{j,i}\phi)) \rightarrow (Bel_i\phi))$ — por (K)
10. $Bel_i\neg(Inf_{j,i}\phi) \rightarrow Bel_i\neg\phi$ — por (PR) sobre (6) e (7)
11. $Bel_i\neg(Inf_{j,i}\phi) \rightarrow Bel_i\phi$ — por (PR) sobre (8) e (9)
12. $Bel_i\neg\phi$ — por (PR) sobre (10) e (3)
13. $Bel_i\phi$ — por (PR) sobre (11) e (3)
14. $Bel_i\phi \wedge Bel_i\neg\phi$ — pela introdução do $\wedge$ sobre (12) e (13)
15. $\neg(Bel_i\phi \wedge Bel_i\neg\phi)$ — por (D)
16. $\perp$ — pela definição de $\perp$ sobre (14) e (15)

- $\{ConfCoop(i, j, \phi), DesConfCoop(i, j, \phi), Bel_i\neg(Inf_{j,i}\phi)\} \vdash \perp$

1. $ConfCoop(i, j, \phi)$ — premissa

2.	$DesConfCoop(i, j, \phi)$	premissa
3.	$Bel_i\neg(Inf_{j,i}\phi)$	premissa
4.	$Bel_i((Bel_j\phi) \rightarrow (Inf_{j,i}\phi))$	pela definição de (1)
5.	$Bel_i((\neg Bel_j\phi) \rightarrow (Inf_{j,i}\phi))$	pela definição de (2)
6.	$Bel_i(\neg(Inf_{j,i}\phi) \rightarrow \neg(Bel_j\phi))$	por (P3) sobre (4)
7.	$Bel_i(\neg(Inf_{j,i}\phi) \rightarrow \neg(Bel_j\phi)) \rightarrow (Bel_i(\neg(Inf_{j,i}\phi)) \rightarrow Bel_i(\neg(Bel_j\phi)))$	por (K)
8.	$Bel_i(\neg(Inf_{j,i}\phi) \rightarrow (Bel_j\phi))$	por (P3) sobre (5)
9.	$Bel_i(\neg(Inf_{j,i}\phi) \rightarrow (Bel_j\phi)) \rightarrow (Bel_i(\neg(Inf_{j,i}\phi)) \rightarrow Bel_i(Bel_j\phi))$	por (K)
10.	$Bel_i\neg(Inf_{j,i}\phi) \rightarrow Bel_i\neg(Bel_j\phi)$	por (PR) sobre (6) e (7)
11.	$Bel_i\neg(Inf_{j,i}\phi) \rightarrow Bel_i(Bel_j\phi)$	por (PR) sobre (8) e (9)
12.	$Bel_i\neg(Bel_j\phi)$	por (PR) sobre (10) e (3)
13.	$Bel_i(Bel_j\phi)$	por (PR) sobre (11) e (3)
14.	$Bel_i\neg(Bel_j\phi) \wedge Bel_i(Bel_j\phi)$	pela introdução do $\wedge$ sobre (12) e (13)
15.	$\neg(Bel_i\neg(Bel_j\phi) \wedge Bel_i(Bel_j\phi))$	por (D)
16.	$\perp$	pela definição de $\perp$ sobre (14) e (15)

- $\{ConfCompet(i, j, \phi),\ DesConfCompet(i, j, \phi),\ Bel_i(Bel_j\phi)\} \vdash \perp$

1.	$ConfCompet(i, j, \phi)$	premissa
2.	$DesConfCompet(i, j, \phi)$	premissa
3.	$Bel_i(Bel_j\phi)$	premissa
4.	$Bel_i(Bel_j\phi \rightarrow \phi)$	pela definição de (1)
5.	$Bel_i(Bel_j\phi \rightarrow \neg\phi)$	pela definição de (2)
6.	$Bel_i(Bel_j\phi \rightarrow \phi) \rightarrow (Bel_i(Bel_j\phi) \rightarrow Bel_i\phi)$	por (K)
7.	$Bel_i(Bel_j\phi \rightarrow \neg\phi) \rightarrow (Bel_i(Bel_j\phi) \rightarrow Bel_i\neg\phi)$	por (K)
8.	$(Bel_i(Bel_j\phi) \rightarrow Bel_i\phi)$	por (PR) sobre (4) e (6)
9.	$(Bel_i(Bel_j\phi) \rightarrow Bel_i\neg\phi)$	por (PR) sobre (5) e (7)
10.	$Bel_i\phi$	por (PR) sobre (8) e (3)
11.	$Bel_i\neg\phi$	por (PR) sobre (9) e (3)
12.	$Bel_i\phi \wedge Bel_i\neg\phi$	pela introdução do $\wedge$ sobre (10) e (11)
13.	$\neg(Bel_i\phi \wedge Bel_i\neg\phi)$	por (D)
14.	$\perp$	pela definição de $\perp$ sobre (12) e (13)

- $\{ConfVig(i, j, \phi),\ DesConfVig(i, j, \phi),\ Bel_i\phi\} \vdash \perp$

1.	$ConfVig(i, j, \phi)$	premissa
2.	$DesConfVig(i, j, \phi)$	premissa
3.	$Bel_i\phi$	premissa
4.	$ConfVig(i, j, \phi) \rightarrow ((Bel_i\phi) \rightarrow Bel_i(Bel_j\phi))$	por (K)
5.	$Bel_i(\phi) \rightarrow Bel_i(Bel_j\phi)$	por (PR) sobre (1) e (4)
6.	$DesConfVig(i, j, \phi) \rightarrow (Bel_i(\phi) \rightarrow Bel_i(\neg Bel_j\phi))$	por (K)
7.	$(Bel_i\phi) \rightarrow Bel_i(\neg Bel_j\phi)$	por (PR) sobre (2) e (6)
8.	$Bel_i(Bel_j\phi)$	por (PR) sobre (5) e (3)
9.	$Bel_i\neg(Bel_j\phi)$	por (PR) sobre (7) e (3)
10.	$Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi)$	pela introdução do $\wedge$ sobre (8) e (9)
11.	$\neg(Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi))$	por (D)

12. $\perp$ pela definição de $\perp$ sobre (10) e (11)

$\square$

Além dos resultados supracitados envolvendo a confiança e desconfiança de um mesmo tipo, existem combinações de tipos que podem causar inconsistências. Veremos a seguir alguns resultados envolvendo confiança e desconfiança de tipos diferentes sobre uma mesma fórmula.

Teorema 3.43. *Os seguintes conjuntos de fórmulas são inconsistentes:*

- $\{ConfSinc(i,j,\phi), DesConfCoop(i,j,\phi), Bel_i\neg(Bel_j\phi)\}$
- $\{ConfVal(i,j,\phi), DesConfComple(i,j,\phi), Bel_i\neg\phi\}$
- $\{ConfComple(i,j,\phi), DesConfVal(i,j,\phi), Bel_i\phi\}$
- $\{ConfCoop(i,j,\phi), DesConfSinc(i,j,\phi), Bel_i(Bel_j\phi)\}$
- $\{ConfCompet(i,j,\phi), DesConfVig(i,j,\phi), Bel_i(Bel_j\phi)\}$
- $\{ConfVig(i,j,\phi), DesConfCompet(i,j,\phi), Bel_i\phi\}$

Demonstração.

- $\{ConfSinc(i,j,\phi), DesConfCoop(i,j,\phi), Bel_i\neg(Bel_j\phi)\} \vdash \perp$
  1. $ConfSinc(i,j,\phi)$ premissa
  2. $DesConfCoop(i,j,\phi)$ premissa
  3. $Bel_i(\neg Bel_j\phi)$ premissa
  4. $Bel_i((\neg Bel_j\phi) \to (Inf_{j,i}\phi))$ pela definição de (2)
  5. $Bel_i((Inf_{j,i}\phi) \to (Bel_j\phi))$ pela definição de (1)
  6. $Bel_i((\neg Bel_j\phi) \to (Inf_{j,i}\phi)) \to (Bel_i(\neg Bel_j\phi) \to Bel_i(Inf_{j,i}\phi))$ por (K)
  7. $Bel_i((Inf_{j,i}\phi) \to (Bel_j\phi)) \to (Bel_i(Inf_{j,i}\phi) \to Bel_i(Bel_j\phi))$ por (K)
  8. $Bel_i(\neg Bel_j\phi) \to Bel_i(Inf_{j,i}\phi)$ por (PR) sobre (4) e (6)
  9. $Bel_i(Inf_{j,i}\phi) \to Bel_i(Bel_j\phi)$ por (PR) sobre (5) e (7)
  10. $Bel_i(Inf_{j,i}\phi)$ por (PR) sobre (8) e (3)
  11. $Bel_i(Bel_j\phi)$ por (PR) sobre (9) e (3)
  12. $Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi)$ pela introdução do $\wedge$ sobre (3) e (11)
  13. $\neg(Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi))$ por (D)
  14. $\perp$ pela definição de $\perp$ sobre (12) e (13)

- $\{ConfVal(i,j,\phi), DesConfComple(i,j,\phi), Bel_i\neg\phi\} \vdash \perp$
  1. $ConfVal(i,j,\phi)$ premissa
  2. $DesConfComple(i,j,\phi)$ premissa
  3. $Bel_i\neg\phi$ premissa
  4. $Bel_i(\neg\phi \to (Inf_{j,i}\phi))$ pela definição de (2)
  5. $Bel_i((Inf_{j,i}\phi) \to \phi)$ pela definição de (1)
  6. $Bel_i(\neg\phi \to (Inf_{j,i}\phi)) \to (Bel_i\neg\phi \to Bel_i(Inf_{j,i}\phi))$ por (K)/
  7. $Bel_i((Inf_{j,i}\phi) \to \phi) \to (Bel_i(Inf_{j,i}\phi) \to Bel_i\phi)$ por (K)
  8. $Bel_i\neg\phi \to Bel_i(Inf_{j,i}\phi)$ por (PR) sobre (4) e (6)
  9. $Bel_i(Inf_{j,i}\phi) \to Bel_i\phi$ por (PR) sobre (5) e (7)
  10. $Bel_i(Inf_{j,i}\phi)$ por (PR) sobre (8) e (3)
  11. $Bel_i\phi$ por (PR) sobre (9) e (10)

12. $Bel_i\phi \wedge Bel_i\neg\phi$ — pela introdução do $\wedge$ sobre (3) e (11)
13. $\neg(Bel_i\phi \wedge Bel_i\neg\phi)$ — por (D)
14. $\perp$ — pela definição de $\perp$ sobre (12) e (13)

- $\{ConfComple(i,j,\phi), DesConfVal(i,j,\phi), Bel_i\phi\} \vdash \perp$
  1. $ConfComple(i,j,\phi)$ — premissa
  2. $DesConfVal(i,j,\phi)$ — premissa
  3. $Bel_i\phi$ — premissa
  4. $Bel_i(\phi \rightarrow (Inf_{j,i}\phi))$ — pela definição de (1)
  5. $Bel_i((Inf_{j,i}\phi) \rightarrow \neg\phi)$ — pela definição de (2)
  6. $Bel_i(\phi \rightarrow (Inf_{j,i}\phi)) \rightarrow (Bel_i\phi \rightarrow Bel_i(Inf_{j,i}\phi))$ — por (K)
  7. $Bel_i((Inf_{j,i}\phi) \rightarrow \neg\phi) \rightarrow (Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\neg\phi)$ — por (K)
  8. $Bel_i\phi \rightarrow Bel_i(Inf_{j,i}\phi)$ — por (PR) sobre (4) e (6)
  9. $Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\neg\phi$ — por (PR) sobre (5) e (7)
  10. $Bel_i(Inf_{j,i}\phi)$ — por (PR) sobre (8) e (3)
  11. $Bel_i\neg\phi$ — por (PR) sobre (9) e (10)
  12. $Bel_i\phi \wedge Bel_i\neg\phi$ — pela introdução do $\wedge$ sobre (3) e (11)
  13. $\neg(Bel_i\phi \wedge Bel_i\neg\phi)$ — por (D)
  14. $\perp$ — pela definição de $\perp$ sobre (12) e (13)

- $\{ConfCoop(i,j,\phi), DesConfSinc(i,j,\phi), Bel_i(Bel_j\phi)\} \vdash \perp$
  1. $ConfCoop(i,j,\phi)$ — premissa
  2. $DesConfSinc(i,j,\phi)$ — premissa
  3. $Bel_i(Bel_j\phi)$ — premissa
  4. $Bel_i((Bel_j\phi) \rightarrow (Inf_{j,i}\phi))$ — pela definição de (1)
  5. $Bel_i((Inf_{j,i}\phi) \rightarrow \neg(Bel_j\phi))$ — pela definição de (2)
  6. $Bel_i((Bel_j\phi) \rightarrow (Inf_{j,i}\phi)) \rightarrow (Bel_i(Bel_j\phi) \rightarrow Bel_i(Inf_{j,i}\phi))$ — por (K)
  7. $Bel_i((Inf_{j,i}\phi) \rightarrow \neg(Bel_j\phi)) \rightarrow (Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\neg(Bel_j\phi))$ — por (K)
  8. $Bel_i(Bel_j\phi) \rightarrow Bel_i(Inf_{j,i}\phi)$ — por (PR) sobre (4) e (6)
  9. $Bel_i(Inf_{j,i}\phi) \rightarrow Bel_i\neg(Bel_j\phi)$ — por (PR) sobre (5) e (7)
  10. $Bel_i(Inf_{j,i}\phi)$ — por (PR) sobre (8) e (3)
  11. $Bel_i\neg(Bel_j\phi)$ — por (PR) sobre (9) e (10)
  12. $Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi)$ — pela introdução do $\wedge$ sobre (3) e (11)
  13. $\neg(Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi))$ — por (D)
  14. $\perp$ — pela definição de $\perp$ sobre (12) e (13)

- $\{ConfCompet(i,j,\phi), DesConfVig(i,j,\phi), Bel_i(Bel_j\phi)\} \vdash \perp$
  1. $ConfCompet(i,j,\phi)$ — premissa
  2. $DesConfVig(i,j,\phi)$ — premissa
  3. $Bel_i(Bel_j\phi)$ — premissa
  4. $Bel_i((Bel_j\phi) \rightarrow \phi)$ — pela definição de (1)
  5. $Bel_i(\phi \rightarrow \neg(Bel_j\phi))$ — pela definição de (2)
  6. $Bel_i((Bel_j\phi) \rightarrow \phi) \rightarrow (Bel_i(Bel_j\phi) \rightarrow Bel_i\phi)$ — por (K)
  7. $Bel_i(\phi \rightarrow \neg(Bel_j\phi)) \rightarrow (Bel_i\phi \rightarrow Bel_i\neg(Bel_j\phi))$ — por (K)

8. $Bel_i(Bel_j\phi) \rightarrow Bel_i\phi$ por (PR) sobre (4) e (6)
9. $Bel_i\phi \rightarrow Bel_i\neg(Bel_j\phi)$ por (PR) sobre (5) e (7)
10. $Bel_i\phi$ por (PR) sobre (8) e (3)
11. $Bel_i(\neg Bel_j\phi)$ por (PR) sobre (9) e (10)
12. $Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi)$ pela introdução do $\wedge$ sobre (3) e (11)
13. $\neg(Bel_i(Bel_j\phi) \wedge Bel_i\neg(Bel_j\phi))$ por (D)
14. $\perp$ pela definição de $\perp$ sobre (12) e (13)

- $\{ConfVig(i,j,\phi),\ DesConfCompet(i,j,\phi),\ Bel_i\phi\} \vdash \perp$

1. $ConfVig(i,j,\phi)$ premissa
2. $DesConfCompet(i,j,\phi)$ premissa
3. $Bel_i\phi$ premissa
4. $Bel_i(\phi \rightarrow (Bel_j\phi))$ pela definição de (1)
5. $Bel_i((Bel_j\phi) \rightarrow \neg\phi)$ pela definição de (2)
6. $Bel_i(\phi \rightarrow (Bel_j\phi)) \rightarrow (Bel_i\phi \rightarrow Bel_i(Bel_j\phi))$ por (K)
7. $Bel_i(Bel_j\phi) \rightarrow \neg\phi) \rightarrow (Bel_i(Bel_j\phi) \rightarrow Bel_i\neg\phi)$ por (K)
8. $Bel_i\phi \rightarrow Bel_i(Bel_j\phi)$ por (PR) sobre (4) e (6)
9. $Bel_i(Bel_j\phi) \rightarrow Bel_i\neg\phi$ por (PR) sobre (5) e (7)
10. $Bel_i(Bel_j\phi)$ por (PR) sobre (8) e (3)
11. $Bel_i\neg\phi$ por (PR) sobre (9) e (10)
12. $Bel_i\phi \wedge Bel_i\neg\phi$ pela introdução do $\wedge$ sobre (3) e (11)
13. $\neg(Bel_i\phi \wedge Bel_i\neg\phi)$ por (D)
14. $\perp$ pela definição de $\perp$ sobre (12) e (13)

□

Como o suporte de um argumento não deve derivar contradições, os conjuntos de fórmulas dos resultados provados acima não podem estar contidos no suporte de um argumento.

Corolário 3.44. *Os seguintes conjuntos de fórmulas não podem estar contidos no suporte de um argumento:*

- $\{ConfSinc(i,j,\phi),\ DesConfSinc(i,j,\phi),\ Bel_i(Inf_{j,i}\phi)\}$
- $\{ConfVal(i,j,\phi),\ DesConfVal(i,j,\phi),\ Bel_i(Inf_{j,i}\phi)\}$
- $\{ConfComple(i,j,\phi),\ DesConfComple(i,j,\phi),\ Bel_i\neg(Inf_{j,i}\phi)\}$
- $\{ConfCoop(i,j,\phi),\ DesConfCoop(i,j,\phi),\ Bel_i\neg(Inf_{j,i}\phi)\}$
- $\{ConfCompet(i,j,\phi),\ DesConfCompet(i,j,\phi),\ Bel_i(Bel_j\phi)\}$
- $\{ConfVig(i,j,\phi),\ DesConfVig(i,j,\phi),\ Bel_i\phi\}$
- $\{ConfSinc(i,j,\phi),\ DesConfCoop(i,j,\phi),\ Bel_i\neg(Bel_j\phi)\}$
- $\{ConfVal(i,j,\phi),\ DesConfComple(i,j,\phi),\ Bel_i\neg\phi\}$
- $\{ConfComple(i,j,\phi),\ DesConfVal(i,j,\phi),\ Bel_i\phi\}$
- $\{ConfCoop(i,j,\phi),\ DesConfSinc(i,j,\phi),\ Bel_i(Bel_j\phi)\}$
- $\{ConfCompet(i,j,\phi),\ DesConfVig(i,j,\phi),\ Bel_i(Bel_j\phi)\}$
- $\{ConfVig(i,j,\phi),\ DesConfCompet(i,j,\phi),\ Bel_i\phi\}$

Demonstração. Como visto nos Teoremas 3.42 e 3.43, cada um desses conjuntos é inconsistente e pela Definição 3.10, o suporte de um argumento não deve derivar inconsistências. □

Com esses resultados, fechamos as propriedades que relacionam a confiança e desconfiança entre agentes. Embora essas propriedades tenham sido provadas para a confiança e desconfiança entre agentes com relação a fórmulas, elas podem ser facilmente estendidas para tópicos, alterando a fórmula ϕ utilizada por um tópico t nas fórmulas de confiança e desconfiança e adicionando a fórmula $\mathfrak{A}(t, \phi)$.

Em (AMGOUD; DEMOLOMBE, 2014) são mostrados alguns resultados que permanecem válidos para o sistema de argumentação apresentado nessa dissertação. A seguir, iremos mostrar esses resultados para a $\mathcal{LCT}$:

O primeiro resultado garante que o conjunto dos suportes dos argumentos presentes em uma extensão estável de um framework de argumentação é consistente:

Teorema 3.45. *Sejam um Framework de Argumentação $\mathcal{T} = (Arg(DM), \Re)$ construído a partir de uma descrição de mundo DM e $Ext(\mathcal{T})$ o conjunto das extensões estáveis de $\mathcal{T}$. Para toda $\mathcal{E} \in Ext(\mathcal{T})$, as seguintes propriedades valem:*

- *O conjunto $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ é consistente.*
- *O conjunto $\{h \mid \exists(H, h) \in \mathcal{E}\}$ é consistente.*

Demonstração. Seja $\mathcal{E}$ uma extensão estável de $\mathcal{T} = (Arg(DM), \Re)$. Assuma que o conjunto $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ é inconsistente. Logo, $\exists X \subseteq \bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ tal que X é mínimo (com respeito à inclusão de conjuntos) inconsistente. Uma vez que cada H_k é consistente, $|X| > 1$. Logo, para todo $Bel_i(x) \in X$, $X \setminus \{Bel_i(x)\}$ é um conjunto tal que $X \setminus \{Bel_i(x)\} \vdash Bel_i(\neg x)$. Então, $A = (X \setminus \{Bel_i(x)\}, Bel_i(\neg x))$ e $B = (\{Bel_i(x)\}, Bel_i(x))$ são dois argumentos. Consequentemente, A realiza um ataque sobre premissa em B. Assuma que, $\exists(H, h) \in \mathcal{E}$ tal que $Bel_i(x) \in H$. Então, A realiza um ataque sobre premissa em (H, h). Uma vez que $\mathcal{E}$ deve ser livre de conflito, $A \notin \mathcal{E}$ e $\exists(H', h') \in \mathcal{E}$ tal que $((H', h'), A) \in \Re$.

1. Suponha que (H', h') realiza um ataque sobre premissa em $(X \setminus \{Bel_i(x)\}, Bel_i(\neg x))$. Então, $\exists Bel_i(x') \in X \setminus \{Bel_i(x)\}$ tal que $H' \vdash Bel_i(\neg x')$. Porém, $Bel_i x' \in H''$ para algum $(H'', h'') \in \mathcal{E}$, pois $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ é inconsistente. Logo, (H', h') realiza um ataque sobre a premissa em (H'', h''), contradizendo que $\mathcal{E}$ é livre de conflito.
2. Suponha agora que (H', h') realiza um ataque na confiança na sinceridade em $(X \setminus \{Bel_i(x)\}, Bel_i(\neg x))$. Então $h' = Bel_i(Inf_{j,i}\phi \wedge \neg Bel_j\phi)$ e $ConfSinc(i, j, \phi) \in X \setminus \{Bel_i(x)\}$. Então, $\exists(H'', h'')$ tal que $ConfSinc(i, j, \phi) \in H''$, pois $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ é inconsistente. Logo, (H', h') realiza um ataque na confiança na sinceridade em (H'', h''), contradizendo que $\mathcal{E}$ é livre de conflito.
3. Suponha agora que (H', h') realiza um ataque na desconfiança na sinceridade em $(X \setminus \{Bel_i(x)\}, Bel_i(\neg x))$. Então $h' = Bel_i(Inf_{j,i}\phi \wedge Bel_j\phi)$ e $DesConfSinc(i, j, \phi) \in X \setminus \{Bel_i(x)\}$. Então, $\exists(H'', h'')$ tal que $DesConfSinc(i, j, \phi) \in H''$, pois $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ é inconsistente. Logo, (H', h') realiza um ataque na confiança na sinceridade em (H'', h''), contradizendo que $\mathcal{E}$ é livre de conflito.
4. Suponha agora que (H', h') realiza um ataque na confiança na sinceridade em tópicos em $(X \setminus \{Bel_i(x)\}, Bel_i(\neg x))$ e $\mathfrak{A}(t, \phi)$. Então $h' = Bel_i(Inf_{j,i}\phi \wedge \neg Bel_j\phi)$ e $ConfSinc(i, j, t) \in X \setminus \{Bel_i(x)\}$. Então, $\exists(H'', h'')$ tal que $ConfSinc(i, j, t) \in H''$, pois $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ é inconsistente. Logo, (H', h') realiza um ataque na confiança na sinceridade em (H'', h''), contradizendo que $\mathcal{E}$ é livre de conflito.

5. Suponha agora que (H', h') realiza um ataque na desconfiança na sinceridade em tópicos em $(X \setminus \{Bel_i(x)\}, Bel_i(\neg x))$ e $\mathfrak{A}(t, \phi)$. Então $h' = Bel_i(Inf_{j,i}\phi \wedge Bel_j\phi)$ e $DesConfSinc(i, j, t) \in X \setminus \{Bel_i(x)\}$. Então, $\exists(H'', h'')$ tal que $DesConfSinc(i, j, t) \in H''$, pois $\bigcup_{(H_k, h_k) \in \mathcal{E}} H_k$ é inconsistente. Logo, (H', h') realiza um ataque na desconfiança na sinceridade em (H'', h''), contradizendo que $\mathcal{E}$ é livre de conflito.

O mesmo tipo de raciocínio pode ser utilizado para os demais ataques. A partir do resultado acima, segue que o conjunto $\{h \mid \exists(H, h) \in \mathcal{E}\}$ também é consistente, uma vez que $\bigcup_{(H_k, h_k) \in \mathcal{E}} H_k \vdash \{h \mid \exists(H, h) \in \mathcal{E}\}$ e $\bigcup_{(H_k, h_k) \in \mathcal{E}} H_k$ é consistente. □

O Teorema acima prova que uma extensão não possui conflitos entre seus argumentos. Como as fórmulas presentes no Resultado de um Framework de Argumentação são as conclusões dos argumentos presentes nas suas extensões estáveis, se torna direta a demonstração de que o resultado de um framework de argumentação é consistente:

Teorema 3.46. *Sejam um Framework de Argumentação $\mathcal{T} = (Arg(DM), \Re)$ construído a partir de uma descrição de mundo DM, $Ext(\mathcal{T})$ o conjunto de todas as extensões estáveis de $\mathcal{T}$ e $R(\mathcal{T})$ o Resultado de $\mathcal{T}$. Para todo Framework de Argumentação $\mathcal{T}$, duas fórmulas ϕ e $\neg\phi$ não pertencem a $R(\mathcal{T})$ ao mesmo tempo.*

Demonstração. Pela Definição 3.40, sabemos que $R(\mathcal{T}) \subseteq \{h \mid \exists(H, h) \in \mathcal{E}\}$ para toda extensão $\mathcal{E} \in Ext(\mathcal{T})$. Como $\{h \mid \exists(H, h) \in \mathcal{E}\}$ é consistente, conforme o Teorema 3.45, não é possível que $R(\mathcal{T})$ seja inconsistente. □

Como o resultado de um Framework de Argumentação é consistente, podemos inferir que uma decisão tomada a partir do ponto de vista de um grupo de agentes gerado pelo sistema de argumentação proposto não irá conter contradições.

A definição a seguir será utilizada para definir o próximo resutado.

O próximo resultado mostra que se um argumento A está presente em uma extensão, os argumentos cujas conclusões são derivadas de um suporte X, tal que X está contido em A, também estão:

Teorema 3.47. *Seja um Framework de Argumentação $\mathcal{T} = (Arg(DM), \Re)$ construído a partir de uma descrição de mundo DM. Para toda extensão estável $\mathcal{E}$ de $\mathcal{T}$, se $(H, h) \in \mathcal{E}$ então para todo $(H', h') \in Arg(D)$ tal que $H' \subseteq H$, $(H', h') \in \mathcal{E}$.*

Demonstração. Seja $\mathcal{E}$ uma extensão estável de $\mathcal{T} = (Arg(DM), \Re)$. Suponha, por absurdo, que $(H', h') \in Arg(DM)$ tal que $(H' \subseteq H)$ e $(H', h') \notin \mathcal{E}$. Por conseguinte, existe um argumento $(H'', h'') \in \mathcal{E}$ tal que $((H'', h''),(H', h')) \in \Re$.

- Se (H'', h'') realiza um ataque sobre a premissa de (H', h'), existe então $Bel_i x \in H'$ tal que $Bel_i \neg x = h''$. Porém, como $H' \subseteq H$, então $((H'', h''),(H, h)) \in \Re$, contradizendo o fato que $\mathcal{E}$ é livre de conflito.
- Se (H'', h'') realiza um ataque na confiança na sinceridade sobre (H', h'), existe então $ConfSinc(i, j, \phi) \in H'$ tal que $Bel_i(Inf_{j,i}\phi \wedge \neg Bel_j\phi) = h''$. Porém, como $H' \subseteq H$, então $((H'', h''),(H, h)) \in \Re$, contradizendo o fato que $\mathcal{E}$ é livre de conflito.
- Se (H'', h'') realiza um ataque na desconfiança na sinceridade sobre (H', h'), existe então $DesConfSinc(i, j, \phi) \in H'$ tal que $Bel_i(Inf_{j,i}\phi \wedge Bel_j\phi) = h''$. Porém, como $H' \subseteq H$, então $((H'', h''),(H, h)) \in \Re$, contradizendo o fato que $\mathcal{E}$ é livre de conflito.

- Se (H'', h'') realiza um ataque na confiança na sinceridade em tópicos sobre (H', h'), existe então $\mathit{ConfSinc}(i, j, t) \in H'$ tal que $\mathit{Bel}_i(\mathit{Inf}_{j,i}\phi \wedge \neg \mathit{Bel}_j\phi) = h''$ com $\mathfrak{A}(t, \phi)$. Porém, como $H' \subseteq H$, então $((H'', h''),(H, h)) \in \Re$, contradizendo o fato que $\mathcal{E}$ é livre de conflito.
- Se (H'', h'') realiza um ataque na confiança na desconfiança sinceridade em tópicos sobre (H', h'), existe então $\mathit{DesConfSinc}(i, j, t) \in H'$ tal que $\mathit{Bel}_i(\mathit{Inf}_{j,i}\phi \wedge \mathit{Bel}_j\phi) = h''$ com $\mathfrak{A}(t, \phi)$. Porém, como $H' \subseteq H$, então $((H'', h''),(H, h)) \in \Re$, contradizendo o fato que $\mathcal{E}$ é livre de conflito.

O mesmo tipo de raciocínio pode ser utilizado para os demais ataques. $\square$

Esse resultado mostra que se uma fórmula ϕ pertence ao ponto de vista de um grupo, as fórmulas que derivam das mesmas premissas de ϕ também pertencem ao ponto de vista do grupo.

A próxima propriedade garante que as extensões estáveis são fechadas sob o operador de consequência sintática $\vdash$. Com isso, pode-se dizer que o sistema de argumentação não ignora as consequências das fórmulas presentes nas extensões:

Teorema 3.48. *Sejam $\mathcal{T} = (Arg(DM), \Re)$ um framework de argumentação construído a partir da descrição de mundo DM, $\mathit{Ext}(\mathcal{T})$ o conjunto das extensões estáveis de $\mathcal{T}$ e $Con(X) = \{\phi \in \mathcal{L} \mid X \vdash \phi\}$ o conjunto das fórmlas que são consequência de um conjunto de fórmulas X. Para todo $\mathcal{E} \in \mathit{Ext}(\mathcal{T})$, $\{h \mid \exists(H, h) \in \mathcal{E}\} = Con(\{h \mid \exists(H, h) \in \mathcal{E}\})$*

Demonstração. Seja $\mathcal{E}$ uma extensão estável de $\mathcal{T}$. Seja $X = \{h \mid \exists(H, h) \in \mathcal{E}\}$ e assuma por absurdo que $X \neq Con(X)$. Portanto, $\exists h \in Con(X)$ e $h \notin X$. Além disso, $X \subseteq \bigcup_{(H_k,h_k)\in\mathcal{E}} Con(H_k) \subseteq Con(\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k)$. Também segue que $Con(X) \subseteq Con(\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k)$ e, portanto, $h \in Con(\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k)$. Dois casos possíveis são

1. $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k = \emptyset$: $(\emptyset, h) \in Arg(DM)$ mas $(\emptyset, h) \notin \mathcal{E}$. Isso significa que existe um argumento (H'_k, h'_k) tal que $((H'_k, h'_k), (\emptyset, h)) \in \Re$. Porém, para esse ataque acontecer, faz-se necessário que exista um ataque sobre a premissa, confiança ou desconfiança, o que é impossível, uma vez que o suporte de $(\emptyset, h)$ é vazio.
2. $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k \neq \emptyset$: $\exists S \subseteq \bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$, de modo que $(S, h) \in Arg(DM)$, uma vez que $\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ é consistente (segundo o Teorema 3.45) e $h \in Con(\bigcup_{(H_k,h_k)\in\mathcal{E}} H_k)$. Além disso, $(S, h) \notin \mathcal{E}$, de acordo com a premissa que $h \notin X$. Portanto, $\exists(H', h') \in \mathcal{E}$ tal que $((H', h'), (S, h)) \in \Re$.
    - Assuma que $((H', h'), (S, h)) \in \Re$ é um ataque sobre premissa. Logo, $h' = \mathit{Bel}_i(\neg x)$ e $\mathit{Bel}_i(x) \in S$. Como $S \subseteq \bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ então $\exists(H'', h'') \in \mathcal{E}$ tal que $\mathit{Bel}_i(x) \in H''$, de modo que $((H', h'), (H'', h'')) \in \Re$, contradizendo que $\mathcal{E}$ é livre de conflito.
    - Assuma que $((H', h'), (S, h)) \in \Re$ é um ataque sobre a confiança na sinceridade. Logo, $h' = \mathit{Bel}_i(\mathit{Inf}_{j,i}\phi \wedge \neg \mathit{Bel}_j\phi)$ e $\mathit{ConfSinc}(i, j, \phi) \in S$. Como $S \subseteq \bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ então $\exists(H'', h'') \in \mathcal{E}$ tal que $\mathit{ConfSinc}(i, j, \phi) \in H''$, de modo que $((H', h'), (H'', h'')) \in \Re$, contradizendo que $\mathcal{E}$ é livre de conflito.
    - Assuma que $((H', h'), (S, h)) \in \Re$ é um ataque sobre a desconfiança na sinceridade. Logo, $h' = \mathit{Bel}_i(\mathit{Inf}_{j,i}\phi \wedge \mathit{Bel}_j\phi)$ e $\mathit{DesConfSinc}(i, j, \phi) \in S$. Como $S \subseteq \bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ então $\exists(H'', h'') \in \mathcal{E}$ tal que $\mathit{DesConfSinc}(i, j, \phi) \in H''$, de modo que $((H', h'), (H'', h'')) \in \Re$, contradizendo que $\mathcal{E}$ é livre de conflito.

- Assuma que $((H', h'), (S, h)) \in \Re$ é um ataque sobre confiança sinceridade em tópicos e $\mathfrak{A}(t, phi)$. Logo, $h' = Bel_i(Inf_{j,i}\phi \wedge \neg Bel_j\phi)$ e $ConfSinc(i, j, t) \in S$. Como $S \subseteq \bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ então $\exists(H'', h'') \in \mathcal{E}$ tal que $ConfSinc(i, j, t) \in H''$, de modo que $((H', h'), (H'', h'')) \in \Re$, contradizendo que $\mathcal{E}$ é livre de conflito.
- Assuma que $((H', h'), (S, h)) \in \Re$ é um ataque sobre a desconfiança na sinceridade em tópicos e $\mathfrak{A}(t, phi)$. Logo, $h' = Bel_i(Inf_{j,i}\phi \wedge Bel_j\phi)$ e $DesConfSinc(i, j, t) \in S$. Como $S \subseteq \bigcup_{(H_k,h_k)\in\mathcal{E}} H_k$ então $\exists(H'', h'') \in \mathcal{E}$ tal que $DesConfSinc(i, j, t) \in H''$, de modo que $((H', h'), (H'', h'')) \in \Re$, contradizendo que $\mathcal{E}$ é livre de conflito.

O mesmo raciocínio pode ser utilizado para os demais ataques.

□

Esse resultado mostra que se uma fórmula ϕ pertence ao ponto de vista de um grupo, as fórmulas derivadas de ϕ também pertencem ao ponto de vista do grupo.

Para finalizar esta subseção, iremos mostrar que o resultado de um framework de argumentação é fechado em relação a $\vdash$.

Teorema 3.49. *Sejam $\mathcal{T} = (Arg(DM), \Re)$ um framework de argumentação construído a partir da descrição de mundo DM e $Ext(\mathcal{T}) \neq \emptyset$ o conjunto das extensões estáveis de $\mathcal{T}$. É válido que o resultado de $\mathcal{T}$, $R(\mathcal{T})$, é igual a $Con(R(\mathcal{T}))$.*

Demonstração. Sejam $\mathcal{T} = (Arg(DM), \Re)$ um framework de argumentação construído a partir da descrição de mundo DM tal que $Ext(\mathcal{T}) \neq \emptyset$. É trivial que $R(\mathcal{T}) \subseteq Con(R(\mathcal{T}))$, sendo $Con(X) = \{\phi \in \mathcal{L} \mid X \vdash \phi\}$.

Assuma agora por absurdo que $h \in Con(R(\mathcal{T}))$ e $h \notin R(\mathcal{T})$. Então, $\exists h_1, h_2, \ldots h_n \in R(\mathcal{T})$ tal que $h \in Con(\{h_1, h_2, \ldots, h_n\})$. Além de que, $h_1, h_2, \ldots, h_n \in \bigcap_{\mathcal{E}_k\in Ext(\mathcal{T})}\{\phi \mid \exists(H, \phi) \in \mathcal{E}_k\}$. Como $h_1, h_2, \ldots h_n \in R(\mathcal{T})$, segue que $Con(h_1, h_2, \ldots, h_n) \subseteq Con(\bigcap_{\mathcal{E}_k\in Ext(\mathcal{T})} \{\phi \mid \exists(H, \phi) \in \mathcal{E}_k\})$ pela Definição 3.40 e, portanto, $h \in Con(\{\phi \mid \exists(H, \phi) \in \mathcal{E}_1\}) \cap \ldots \cap Con(\{\phi \mid \exists(H, \phi) \in \mathcal{E}_m\})$. Pelo Teorema 3.48, $h \in \{\phi \mid \exists(H, \phi) \in \mathcal{E}_1\} \cap \ldots \cap \{\phi \mid \exists(H, \phi) \in \mathcal{E}_m\}$. Consequentemente, $h \in R(\mathcal{T})$. □

Esse resultado garante, por exemplo, que se $ConfSinc(i, j, \phi) \in R(\mathcal{T})$ e $Bel_iInf_{j,i}\phi \in R(\mathcal{T})$ então $Bel_iBel_j\phi \in R(\mathcal{T})$, mostrando que se uma fórmula ϕ é consequência do resultado do ponto de vista de um grupo, as fórmulas derivadas de ϕ também pertencem ao resultado do ponto de vista do grupo.

3.4.2 Complexidade

Nesta subseção iremos calcular a complexidade de gerar o Resultado de um Framework de Argumentação e suas extensões estáveis. Para tanto, iremos dispor do algoritmo proposto em (**??**), que verifica se um conjunto de argumentos é uma extensão estável de um Framework de Argumentação. A abordagem desse algoritmo envolve representar um Framework de Argumentação como um grafo. Nesse grafo, os argumentos são representados pelos vértices e os ataques entre argumentos são representados pelas arestas. Desse modo, os autores puderam utilizar operações de multiplicação de matrizes sobre a matriz de adjacência do grafo e um vetor representando um conjunto de vértices para calcular se esse conjunto representa uma extensão estável ou não.

Consideraremos $\mathcal{T} = (S, \Re)$ um framework de argumentação construído a partir da Descrição de Mundo DM e $S \subseteq Arg(DM)$ um conjunto finito de argumentos. Para calcular o resultado de $\mathcal{T}$, devemos primeiramente encontrar suas extensões estáveis. Em (**??**) é apresentado um algoritmo que calcula se um conjunto de argumentos é uma extensão estável utilizando operações de multiplicação de matrizes com complexidade em $O(n^3)$, sendo $|Arg(S)|$ = n. Para encontrar todas as extensões estáveis de $\mathcal{T}$, devemos executar esse algoritmo para todos os subconjuntos de S. Portanto, seriam necessárias 2^n execuções do algoritmo para encontrar todas as extensões estáveis de $\mathcal{T}$. O último passo consiste em comparar quais conclusões de argumentos estão presentes em cada uma das extensões de $\mathcal{T}$ conforme a Definição 3.40. Esse passo pode ser dividido em dois passos menores: União das conclusões de argumentos estão presentes nas extensões de $\mathcal{T}$ e verificação de quais dessas conclusões estão presentes em cada uma das extensões de $\mathcal{T}$. O primeiro passo pode ser feito com complexidade $O(m * n)$, considerando no pior caso m extensões com n argumentos cada. O segundo passo pode ser feito com complexidade $O(m * n^2)$, considerando que no pior caso n conclusões de argumentos devem ser verificados em m extensões com no máximo n argumentos cada.

Assim, como a complexidade de um algoritmo é definida por sua etapa mais lenta, a complexidade do cálculo do Resultado de um Framework de Argumentação é $O(2^n)$.

3.5 Conclusão

Com as definições apresentadas neste capítulo, pudemos estender o sistema de argumentação do capítulo anterior para lidar com fórmulas que representam a confiança e a desconfiança em fórmulas e tópicos na estrutura de um argumento. Além disso, a relação de ataque entre argumentos nessa nova abordagem também passou a conter as definições de ataque apresentadas neste capítulo. Com esse sistema de argumentação repaginado, podemos extrair as entre fórmulas consideradas como resultado do framework de argumentação aquelas que representam a confiança entre agentes sobre tópicos ou a desconfiança entre agentes sobre fórmulas ou tópicos. Além disso, apresentamos um estudo de caso envolvendo a tragédia *Otelo, o Mouro de Veneza*, modelando Frameworks de Argumentação que dão suporte às decisões centrais em cada ato da obra, com exceção dos atos III e IV, que foram unificados. Além disso, destacamos diversos teoremas provados sobre a construção de argumentos, Frameworks de Argumentação e Resultados de Frameworks de Argumentação.

4 CONCLUSÃO

Neste trabalho, propusemos uma extensão da lógica $\mathcal{LC}$ apresentada em (AMGOUD; DEMOLOMBE, 2014) por Amgoud e Demolombe, que permite raciocinar e formalizar o conceito de confiança entre agentes com relação a uma informação baseada na construção e avaliação de argumentos. O trabalho de Amgoud e Demolombe tem um diferencial em relação a outros presentes na literatura, como mostrado na Subseção 2.3.2, por conta de especializar a confiança em seis diferentes tipos: a confiança na sinceridade, validade, completude, cooperatividade, competência e vigilância. Em nossa extensão, $\mathcal{LCT}$, introduzimos dois novos recursos na lógica: (i) tratar de tópicos de informações, possibilitando manipular conjuntos de informações nas fórmulas de confiança, além de termos definido a relação de ataque entre argumentos que possuem em sua estrutura fórmulas que representam tópicos e (ii) tratar a desconfiança entre agentes, que não consiste apenas da ausência de confiança, mas sim pela crença que um agente possui de que outro agente , aquele a quem a desconfiança é dirigida, irá agir ativamente contra ele em uma dada situação.

A inclusão das fórmulas de desconfiança não altera a linguagem da $\mathcal{LC}$, uma vez que ela também permitia a construção dessas fórmulas. A vantagem em definir a desconfiança pode ser constatada nas relações de ataque envolvendo fórmulas de desconfiança no suporte de um argumento. Seguindo o padrão dos ataques sobre confiança, para que um ataque sobre desconfiança seja realizado com sucesso, basta encontrar uma situação em que o agente tenha mostrado não agir de forma que mereça a desconfiança. Por sua vez, a inclusão dos tópicos altera a linguagem, o conjunto de axiomas e as relações de ataque de $\mathcal{LC}$. A linguagem é alterada com o acréscimo das fórmulas e a modalidade referentes aos tópicos. O conjunto de axiomas passa a incluir os axiomas referentes à Lógica de Tópicos. Por fim, também foram definidas relações de ataque contemplando a confiança e desconfiança em tópicos, de modo que apenas ataques envolvendo fórmulas que pertençam ao mesmo tópico da confiança ou desconfiança sejam permitidos.

A análise da obra *Otelo, o Mouro de Veneza* mostrou que o Sistema de Argumentação proposto neste trabalho é capaz de modelar situações em que decisões são tomadas a partir da confiança e desconfiança entre agentes. Cada ato foi analisado separadamente (com exceção dos atos III e IV, que foram unificados por conta de os dois contribuírem para uma única decisão), de modo que pudemos modelar as decisões centrais de cada um deles em um Framework de Argumentação gerado de uma Descrição das Crenças do agente que toma a decisão referente ao ato.

Este trabalho tem como objetivo facilitar a implementação de sistemas de diálogo e tomada de decisão em que os agentes são capazes de argumentar e tomar decisões baseadas em seus pontos de vista, modelados pelas extensões dos Frameworks de Argumentação. Tornamos este trabalho mais próximo de um ambiente real, permitindo expressar situações em que tanto a confiança quanto a desconfiança podem ser utilizadas como suporte para uma decisão. Note que o resultado do sistema apresentado neste trabalho leva em consideração as informações disponíveis no ambiente, podendo acontecer de a decisão tomada não ser suficiente para resolver um certo problema. Não está no escopo deste trabalho avaliar se o ponto de vista representado define uma decisão correta para uma determinada situação. No exemplo motivacional apresentado na Subseção 1.1, vemos que a conclusão do grupo de médicos é baseada nas relações de confiança e desconfiança entre eles e por meio de troca de informações referentes ao estado da paciente. Embora tenham chegado a uma conclusão, nada garante que as informações utilizadas seriam suficientes.

4.1 Trabalhos Futuros

Iremos mostrar abaixo algumas das possibilidades de trabalhos que podem ser desenvolvidos a partir desta Dissertação de Mestrado:

4.1.1 Propagação de Confiança e Desconfiança

Em muitos casos, uma informação não possui suporte em observações diretas, mas em informações geradas por outras fontes de informação (DEMOLOMBE, 2011). Uma informação pode ser propagada por uma sequência de agentes, de modo que nenhum deles passe necessariamente a acreditar nessa informação, apenas a propague para outros que eventualmente podem aceitar essa informação a partir de sua confiança naquele que está propagando a mensagem ou usá-la no suporte de um argumento. A partir dessa propagação, podem ser definidos diferentes graus de força entre argumentos. Como exemplo, podemos citar uma informação propagada por x agentes possui mais força que outra que foi propagada por y agentes, desde que $x > y$, utilizando a intuição de que uma informação que foi propagada mais vezes sem contestação possui mais credibilidade que uma informação menos divulgada.

4.1.2 Recomendação de Confiança e Desconfiança

Segundo (PARSONS et al., 2014), é plausível que um agente a passe a confiar em um agente b a partir de sua confiança em outros agentes que também confiem em b. Como trabalho futuro, queremos definir quais as condições e propriedades necessárias para que um agente seja capaz de aceitar a recomendação de outro em relação a confiar em um terceiro. Para que essa contribuição seja possível, iremos estender a $\mathcal{LCT}$ com novos axiomas e definições. O exemplo abaixo mostra de maneira intuitiva a recomendação de confiança.

Seja $ConfCompet(b, c, m)$ a confiança de um agente b na competência de um agente c sobre um tópico m e $ConfRecom(a, b, m)$ a confiança de a na capacidade de b recomendar um terceiro em relação ao tópico m. A partir dessas duas condições, é possível inferir que $ConfCompet(a, c, m)$.

4.1.3 Mistrust e Untrust

Em (MARSH; DIBBEN, 2005), os autores mostram a diferença entre diversas relações de confiança entre agentes. Neste trabalho, tratamos apenas dos conceitos de confiança (*trust*) e desconfiança (*distrust*), restando os conceitos de *mistrust* e *untrust*, que podem ser vistos respectivamente como a consequência de uma confiança que não deveria ser depositada e não-confiança. O conceito de *mistrust* pode ser modelado por meio de uma Lógica Dinâmica, modificando as crenças de um agente após sua confiança em um outro ser atacada. O *untrust* poderia ser modelado como um nível mais fraco de confiança, tornando um argumento que a use como suporte incapaz de atacar outros argumentos. Intuitivamente, podemos enxergar o *untrust* como “eu confio em você, mas acredito que essa informação não seja suficiente para vencer uma discussão”. Enquanto o *mistrust* é capaz de alterar as relações de confiança entre os agentes, o *untrust* pode ser utilizado para caracterizar uma informação que não seja capaz de mudar o curso de um diálogo.

4.1.4 Força de um ataque

No contexto desta Dissertação de Mestrado, os ataques entre argumentos não possuem uma diferenciação de impacto entre si. Por exemplo, considerando um framework de argumentação que possua apenas os argumentos A e B com A atacando B e B atacando A, temos duas extensões estáveis, cada uma com um dos argumentos. Como trabalho futuro, é possível diferenciar os ataques entre argumentos com relação à sua força, de modo que um argumento atacado possa participar de uma extensão, mesmo que ele não ataque o argumento que o ataca ou algum outro o defenda.

Uma maneira de modelar essa ideia é verificar qual dos argumentos seria mais específico, o que poderia indicar um argumento mais elaborado. Um argumento $A = (H_a, h_a)$ é mais específico que um argumento $B = (H_b, h_b)$ quando $H_b \subseteq H_a$.

Outro modo de capturar a força de um argumento poderia estar relacionado a quão forte seria o ataque a uma confiança ou desconfiança. Por exemplo, para que um ataque à confiança na sinceridade de i em j em relação a ϕ aconteça é suficiente que j informe i sobre ψ, mas não acredite nessa fórmula. Uma diferença de força entre os ataques poderia ser modelada de modo que um ataque sobre premissa ocorra sobre uma fórmula de confiança. Dados três argumentos

- $C = (\{\mathit{ConfSinc}(i, j, \phi), \mathit{Inf}_{j,i,\phi}\}, \mathit{Bel}_i \mathit{Bel}_j \phi)$
- $D = (\{\mathit{Bel}_i(\mathit{Inf}_{j,i}\psi \wedge \neg \mathit{Bel}_j \psi)\}, \mathit{Bel}_i(\mathit{Inf}_{j,i}\psi \wedge \neg \mathit{Bel}_j \psi))$
- $E = (\{\mathit{Bel}_i(\mathit{Inf}_{j,i}\phi \wedge \neg \mathit{Bel}_j \phi)\}, \mathit{Bel}_i(\mathit{Inf}_{j,i}\phi \wedge \neg \mathit{Bel}_j \phi))$

podemos dizer que o ataque (E, C) é mais forte que (D, C), pois i confia na sinceridade de j em relação a ϕ e D atacou essa confiança usando a fórmula ψ e E usou uma situação na qual j falhou em ser sincero com relação à própria fórmlula ϕ.

Por fim, podemos diferenciar os ataques por meio de preferências entre argumentos. Em (**??**), os autores expoem que o Framework de Argumentação de Dung não permite considerar preferências sobre argumentos. Em um Framework de Argumentação que considera preferências, se um argumento A ataca um argumento B, mas B é preferível em relação a A, então o ataque falha. O exemplo a seguir ilustra essa ideia.

Exemplo 4.1. *(Adaptado de (**??**)) Um agente quer comprar um violino. Um especialista diz que o violino é da marca Stradivari e, por isso, tem valor elevado. Baseado nessa informação, o agente possui um argumento A que conclui que o violino é caro, por ser da marca Stradivari. O agente possui um filho de três anos de idade que diz que o violino é caro e, por isso, não é da marca Stradivari. Logo, o agente possui um argumento B que conclui que o violino não é da marca Stradivari, por ser caro. Como a conclusão de B contradiz a premissa de A e vice versa, temos que um Framework de Argumentação $AF = (\{A, B\}, \{(A, B), (B, A)\})$ que tem como extensões estáveis os conjuntos $\{A\}$ e $\{B\}$.*

Note que no exemplo acima, o argumento gerado pela informação dada pela criança possui força equivalente ao argumento gerado pelo especialista. Se levarmos em conta que o argumento do especialista deve ser preferível em relação ao argumento da criança, o ataque de B sobre A não irá ser exitoso e, portanto, teremos como extensão estável apenas o conjunto $\{A\}$.

REFERÊNCIAS

AMGOUD, L.; DEMOLOMBE, R. An argumentation-based approach for reasoning about trust in information sources. **Argument & Computation**, Taylor & Francis, v. 5, n. 2-3, p. 191–215, 2014.

BARONI, P.; GIACOMIN, M.; GUIDA, G. Scc-recursiveness: a general schema for argumentation semantics. **Artificial Intelligence**, Elsevier, v. 168, n. 1, p. 162–210, 2005.

CHELLAS, B. F. **Modal logic: an introduction**. [S.l.]: Cambridge Univ Press, 1980. v. 316.

CUPPENS, F.; DEMOLOMBE, R. How to recognize interesting topics to provide cooperative answering. **Information Systems**, Elsevier, v. 14, n. 2, p. 163–173, 1989.

CUPPENS, F.; DEMOLOMBE, R. A deontic logic for reasoning about confidentiality. In: **Deontic Logic, Agency and Normative Systems**. [S.l.]: Springer, 1996. p. 66–79.

DEMOLOMBE, R. Multivalued logics and topics. **QUEEN MARY**, p. 1, 1994.

DEMOLOMBE, R. To trust information sources: a proposal for a modal logical framework. In: **Trust and deception in virtual societies**. [S.l.]: Springer, 2001. p. 111–124.

DEMOLOMBE, R. Reasoning about trust: A formal logical framework. In: **Trust Management**. [S.l.]: Springer, 2004. p. 291–303.

DEMOLOMBE, R. Transitivity and propagation of trust in information sources: An analysis in modal logic. In: **Computational logic in multi-agent systems**. [S.l.]: Springer, 2011. p. 13–28.

DEMOLOMBE, R.; JONES, A. Reasoning about topics: towards a formal theory. In: **American Association for Artificial Intelligence Fall Symposium**. [S.l.: s.n.], 1995.

DEVEREUX, J.; REED, C. Strategic argumentation in rigorous persuasion dialogue. In: **Argumentation in Multi-Agent Systems**. [S.l.]: Springer, 2009. p. 94–113.

DUNG, P. M. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. **Artificial intelligence**, Elsevier, v. 77, n. 2, p. 321–357, 1995.

DUNG, P. M.; MANCARELLA, P.; TONI, F. Computing ideal sceptical argumentation. **Artificial Intelligence**, Elsevier, v. 171, n. 10, p. 642–674, 2007.

KACI, S. **Working with preferences: Less is more**. [S.l.]: Springer Science & Business Media, 2011.

KARUNATILLAKE, N. C.; JENNINGS, N. R.; RAHWAN, I.; NORMAN, T. J. Argument-based negotiation in a social context. In: **Argumentation in Multi-Agent Systems**. [S.l.]: Springer, 2005. p. 104–121.

LAKEMEYER, G. All they know about. In: **AAAI**. [S.l.: s.n.], 1993. p. 662–667.

LIAU, C.-J. Belief, information acquisition, and trust in multi-agent systems—a modal logic formulation. **Artificial Intelligence**, Elsevier, v. 149, n. 1, p. 31–60, 2003.

MARSH, S.; DIBBEN, M. R. Trust, untrust, distrust and mistrust–an exploration of the dark (er) side. In: **Trust management**. [S.l.]: Springer, 2005. p. 17–33.

MEDINECKIENE, M.; ZAVADSKAS, E.; BJÖRK, F.; TURSKIS, Z. Multi-criteria decision-making system for sustainable building assessment/certification. **Archives of Civil and Mechanical Engineering**, Elsevier, v. 15, n. 1, p. 11–18, 2015.

MENDELSON, E. **Introduction to mathematical logic**. [S.l.]: CRC press, 2009.

PARSONS, S.; ATKINSON, K.; LI, Z.; MCBURNEY, P.; SKLAR, E.; SINGH, M.; HAIGH, K.; LEVITT, K.; ROWE, J. Argument schemes for reasoning about trust. **Argument & Computation**, Taylor & Francis, v. 5, n. 2-3, p. 160–190, 2014.

SABATER, J.; SIERRA, C. Review on computational trust and reputation models. **Artificial intelligence review**, Springer, v. 24, n. 1, p. 33–60, 2005.

STRANDERS, R.; WEERDT, M. de; WITTEVEEN, C. **Fuzzy argumentation for trust**. [S.l.]: Springer, 2007.

TANG, Y.; CAI, K.; MCBURNEY, P.; SKLAR, E.; PARSONS, S. Using argumentation to reason about trust and belief. **Journal of Logic and Computation**, Oxford Univ Press, v. 22, n. 5, p. 979–1018, 2012.

WU, Y.; GIGER, M. L.; DOI, K.; VYBORNY, C. J.; SCHMIDT, R. A.; METZ, C. E. Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer. **Radiology**, v. 187, n. 1, p. 81–87, 1993.

ZHAO, P.; SURYANARAYANAN, S.; SIMOES, M. G. An energy management system for building structures using a multi-agent decision-making control methodology. **Industry Applications, IEEE Transactions on**, IEEE, v. 49, n. 1, p. 322–330, 2013.