Inteligencia Artificial 49(2012), 13-30

INTELIGENCIA ARTIFICIAL

http://erevista.aepia.org/

Deliberative DeLP agents with multiple informants

Luciano H. Tamargo, Sebastian Gottifredi, Alejandro J. García, Marcelo A. Falappa and Guillermo R. Simari

Department of Computer Science and Engineering
Artificial Intelligence Research and Development Laboratory
Universidad Nacional del Sur
Consejo Nacional de Investigaciones Científicas y Técnicas
Bahía Blanca, Argentina
E-mail: {lt,sg,ajg,mfalappa,grs}@cs.uns.edu.ar

Abstract In this paper we define a trust-based argumentative reasoning formalism where the source of the received information is used to decide the warranted conclusions. In the proposed formalism, the agent's tentative conclusions are supported by arguments, and these conclusions can in turn be attacked by other arguments, referred to as counter-arguments. The inference mechanism compares arguments and counter-arguments to decide which conclusion prevails. We propose a novel development of different argument comparison criteria based on trust. Thus, given a particular application domain, the more appropriate criterion can be used. Finally, since a realistic scenario should consider that trust on informants is usually subject to change, a complete change theory over the agents' trust is introduced.

Keywords: DeLP, Trust, Argumentation, Preference.

1 Introduction

In this paper, we propose a trust-based argumentative formalism for deliberative agents that participate in a Multi-agent system. We consider that each agent can act as an informant for other agents in the system, and that agents collaborate by sharing information. Therefore, in our proposal each agent can obtain information from multiple sources, and the agent's attribution of trust to a particular informant can be higher than the trust attributed to others. Hence, when the information provided for different agents is in conflict, trust can be used in the decision process leading to a prevailing conclusion. Since in many application domains the agent's assessment of trust of the agents in the system varies, we have developed change operators to handle this dynamics.

The importance of trust models have been widely emphasized in the literature. In [12], the authors mention that two particular elements have contributed to substantially increase the interest on trust in this area: the multi-agent system paradigm and the spectacular evolution of e-commerce. The results coming from the study of trust has many applications in Information and Communication Technologies; *e.g.*, it has been recognized as a key factor for successful electronic commerce adoption. These results are used by intelligent software agents as a mechanism to search for trustworthy exchange partners, and also as an important factor in the decision-making process related to whether or not enter into a contract or even honor them. However, it is clear that trust models are needed in most problems where critical decisions depend on the perceived value of the information received from other agents. The work that

ISSN: 1988-3064(on-line)

we propose here can be applied to multi-agent systems requiring that the trust or credibility assigned to informants be considered as part of a decision process.

As a simple example, consider an agent that has to decide whether to buy a digital camera; this agent receives advice from two informant agents $\mathtt{I}_1$ and $\mathtt{I}_2$. Lets assume also that the agent regards $\mathtt{I}_2$ as less credible than $\mathtt{I}_1$; now, $\mathtt{I}_1$ informs that "the camera xq is a good option", and $\mathtt{I}_2$ informs that "xq is not a good option". Clearly, the conclusion "xq is a good option" should prevail. As time passes, the agent changes its assessment of the credibility relation of its informants: $\mathtt{I}_1$ is now less credible than $\mathtt{I}_2$; therefore the conclusion "xq is not a good option" should prevail.

As we will describe below, in our proposal we will consider logic-based deliberative agents with an argumentative inference mechanism. Hence, conclusions will be supported by arguments built from information coming from different sources. As an agent could receive contradictory information from different informants, conflictive arguments can appear. To solve these conflicts, a decision mechanism based on trust will be provided.

Consider for instance an scenario where an agent V is visiting a town and an agent $\mathtt{I}_A$ informs V that "it was snowing during the night". Other informant $\mathtt{I}_B$ tells V that "roads are usually closed when snows during the night". Then V tentatively should conclude that "roads are closed". Another agent $\mathtt{I}_C$ tells V that "if temperature is high, roads will not be closed", and the weather service $\mathtt{I}_W$ has reported that "now temperature is high". From this information, there are also reasons for conclude that "roads are not closed"; that is, the information obtained from $\mathtt{I}_A$ and $\mathtt{I}_B$ allows to conclude that "roads are closed", whereas the information obtained from $\mathtt{I}_C$ and $\mathtt{I}_W$ allows for the conclusion that "roads are not closed". In this scenario, the trust on $\mathtt{I}_A$,$\mathtt{I}_B$, $\mathtt{I}_C$ and $\mathtt{I}_W$ can be used for deciding which conclusion prevails. For instance, if $\mathtt{I}_C$ and $\mathtt{I}_W$ are both more credible than $\mathtt{I}_A$ and $\mathtt{I}_B$, it is clear that the conclusion should be "roads are not closed".

As we mentioned, in the formalism proposed here agent's tentative conclusions will be supported by arguments; but counter-arguments against these arguments are considered in a dialectical process. When contradictory information appears, a trust-based argument comparison criterion will be used to decide the predominant arguments. Recently, some approaches that combine argumentation and trust have been developed [11, 16, 15]; in general, these formalisms are focused on using argumentation to reason about trust. Here we will follow a different route, our proposal uses trust assigned to informants in the decision process leading to the prevailing information in the context of conflictive knowledge.

Trust is a concept that is both complex and rather difficult to define with precision; for that reason, the literature is not in agreement regarding this elusive notion. In our approach, we consider trust as representing the credibility attached to information sources. This description is closely related to the following definitions: Barber and Kim in [2] suggest that trust is the "confidence in the ability and intention of an information source to deliver correct information"; Gambetta in [4] states that "trust is the subjective (assessment) by which an individual, A, expects that another individual, B, performs a given action on which its welfare depends"; and Mcknight and Chervany in [10] propose that "trust is the extent to which one party is willing to depend on something or somebody in a given situation with a feeling of relative security, even though negative consequences are possible."

Next, we present a scenario that will be used as a running example in the rest of the paper. Assume Eve is considering whether to travel next week. She can obtain information from three different informants: her boss, her colleague Paul, and her assigned client. Eve knows that: "she can travel if she has no work to do", "Paul is ill", and "if Paul is ill he cannot replace her at work". Her boss has informed her that: "if she has work to do, then she cannot travel", and "he thinks that she has some work to do". After checking with her assigned client, Eve knows that "she has no work to do". With the information obtained from her boss, Eve can build an argument supporting that "she cannot travel"; also, using the information received from her client, she can conclude that "she can travel". Since Eve puts more trust on her boss than in Paul, and she trusts more her assigned client than her boss, she can use that trust-based preference to reach a decision on which argument prevails.

The main contribution of this paper is the definition of a trust-based argumentative reasoning formalism where the source of the received information is used to decide the warranted conclusions. In our formalism, the agent's tentative conclusions are supported by arguments, and these conclusions can in turn be attacked by other arguments, referred to as counter-arguments. The inference mechanism compares arguments and counter-arguments to decide which conclusion prevails. Another novel development is the definition of different argument comparison criteria based on trust. Thus, given a particular application domain, the more appropriate criterion can be used. Finally, since a realistic scenario should consider that trust on informants is usually subject to change, a revision operator for the credibility order affecting the informants is introduced.

The rest of the paper is organized as follows: in Section 2 we introduce how agents represent the trust assigned to their informants as a credibility order among them, and how agents represent knowledge and store the received information. Section 3 shows how and agent can build arguments and counter-arguments using our formalism. Next, Section 4 develops different concrete argument comparison criteria that are based on trust. In Section 5 we will show how to determine in which literals the agent believes using the arguments and the relation presented in the previous sections. Section 6 introduces a complete change theory over the agents' trust. Finally, in Section 7 conclusions and related work are commented.

2 Knowledge Representation

The main intuitions in our approach call for the consideration of the sources of information in a multi-agent system, and the realization that, in general, they are not equally credible. That is, the trust that an agent has in each of its informants can be different and, as importantly, that trust may vary over time. To reflect these observations, we will start by considering a finite set of agent identifiers $\mathbb{I}$ enough to assign a different identifier to each member of the system. Also, we will assume that an order is defined over $\mathbb{I}$. For instance, $\mathbb{I} = \{\mathtt{I}_1, \mathtt{I}_2, \mathtt{I}_3\}$ or $\mathbb{I} = \{\mathtt{I}_a, \mathtt{I}_b, \mathtt{I}_c, \mathtt{I}_d\}$ or $\mathbb{I} = \{A, B, C, D, E, F\}$ are all valid sets of agent identifiers. To represent a credibility order among the set $\mathbb{I}$, and to have the ability of changing that order dynamically, the formalism proposed in [14] will be used. For the purpose of clarity, we will recall the definition of *generator set* that allows for the representation of a partial order among informants.

Definition 1 (Generator set [14]) *Given a finite set of informants* $\mathbb{I}$*, a* generator set *over* $\mathbb{I}$ *is a binary relation on* $\mathbb{I}$ *called G-set, (G-set* $\subseteq \mathbb{I} \times \mathbb{I}$*). An informant* $\mathtt{I}_1 \in \mathbb{I}$ *is less credible than an informant* $\mathtt{I}_2 \in \mathbb{I}$*, according to G-set, if* $(\mathtt{I}_1, \mathtt{I}_2) \in$ *G-set***, where G-set** *represents the reflexive transitive closure of G-set. We will say that G-set is* sound *if G-set** *results in a partial order.*

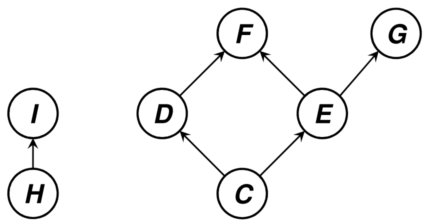

Figure 1: A graph representation of a generator set.

Graphically, a generator set *G-set* is represented as a directed graph, where the nodes are labeled with identifiers in $\mathbb{I}$, and the arcs represent the tuples in *G-set*, *i.e.*, $(\mathtt{I}_1, \mathtt{I}_2) \in$ *G-set* iff there is an arc from node $\mathtt{I}_1$ to node $\mathtt{I}_2$. For example, given the set $\mathbb{I} = \{B, C, D, E, F, G, H, I\}$ the digraph in Figure 1 shows the representation of the generator set $G\text{-}set_1 = \{(C, D), (C, E), (D, F), (E, F), (E, G), (H, I)\}$.

The relation *G-set**, *i.e.*, the reflexive transitive closure of *G-set*, is the smallest preorder containing *G-set*. Any pair $(\mathtt{I}_1, \mathtt{I}_2) \in$ *G-set** is referred to as a *credibility tuple*, and it represents the relation that $\mathtt{I}_1$ is less credible than $\mathtt{I}_2$. For instance, in the generator set $G\text{-}set_1$ introduced above, $(C, G) \in G\text{-}set_1^*$ and $(C, F) \in G\text{-}set_1^*$. That is, in this example, C is less credible than D, C is less credible than F, C is less credible than G, and C and H are not related.

As in [14] we assume that a generator set *G-set** is a partial order, that is, a generator set must be *sound*. Since *G-set** is the reflexive and transitive closure of *G-set*, for soundness to fail on *G-set**, it is necessary that the antisymmetry property fails, *i.e.*, there is at least a couple of pairs (C, D) and (D, C) in *G-set**. This will mean that both C is less credible than D and that D is less credible than C. For that reason, to avoid this undesirable situation, we require for the generator set to be a partial order, *i.e.*, it must obey reflexivity, antisymmetry and transitivity. For example, the generator set $G\text{-}set_1$ showed in

Figure 1 is sound. Nevertheless, $G\text{-}set_2 = G\text{-}set_1 \cup \{(F, C)\}$ is *not* sound because $(C, F) \in G\text{-}set_2^*$ and $(F, C) \in G\text{-}set_2^*$, violating the antisymmetry condition for partial orders.

Example 1 *Consider the scenario presented in the introduction, where: Eve* ($\mathtt{I}_e$) *is less credible than her assigned client* $\mathtt{I}_c$*, Paul* $\mathtt{I}_p$ *is less credible than Eve, Paul is less credible than Eve's boss* ($\mathtt{I}_b$)*, and also that Eve's boss is less credible than Eve's assigned client. This information can be represented by* $\mathbb{I} = \{\mathtt{I}_e, \mathtt{I}_p, \mathtt{I}_c, \mathtt{I}_b\}$ *and the generator set:* $G\text{-}set_{Eve} = \{ (\mathtt{I}_p, \mathtt{I}_b), (\mathtt{I}_b, \mathtt{I}_c), (\mathtt{I}_e, \mathtt{I}_c), (\mathtt{I}_p, \mathtt{I}_e) \}$*. Figure 2 shows the graph representation for* $G\text{-}set_{Eve}$

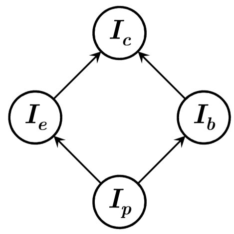

Figure 2: Graph representation for $G\text{-}set_{Eve}$ of Example 1.

In the scenario described, agents will receive information from different sources that are not equally credible, therefore the information, that can be contradictory, will be considered as tentative. Therefore, there is a need for a formalism capable of deciding what to believe by performing reasoning in the presence of tentative and contradictory information; the reasoning mechanism will consider the trust of reporting agents to decide among contradictory conclusions. In our approach, knowledge representation and reasoning will be based on Defeasible Logic Programming (DeLP). Our interest in this particular computational tool is that its language provides the declarative capability of representing weak information in the form of *defeasible rules* and *presumptions*, and its defeasible argumentation inference mechanism allows warranting conclusions in the presence of contradictory information. Below, we will summarily introduce the elements in DeLP and we will refer the reader to [5] for a full account of Defeasible Logic Programming and to [3] to experiment with the fully implemented system. DeLP is an extension of Logic Programming, and for that reason we will freely use its language and terminology (see [9] for more details).

Defeasible rules are used to represent tentative information that may be used when nothing could be posed against it. They are denoted as an ordered pair "$L_0 \prec \{L_1, \ldots, L_n\}$" where L_0 is a ground literal called the "*Head*" of the rule, and $\{L_1, \ldots, L_n\}$ is a set of ground literals called the "*Body*" of the rule. The symbols "{" and "}" on the body of a rule will be omitted when no confusion could arise. Presumption are a special type of defeasible rules with an empty body, denoted "$L_0 \prec$ ". Following the terminology of Logic Programming, literals represent atomic information or the negation of atomic information using strong negation. In DeLP, the symbol "$\sim$" is used for representing strong negation. A ground literal is a literal that has no variables. Two ground literals *disagree* if they are complementary with respect to strong negation (*e.g.*, $\sim L$ and L).

A defeasible rule "*Head* $\prec$ *Body*" expresses that "reasons to believe in the antecedent *Body* give reasons to believe in the consequent *Head*." For instance, the defeasible rule "*road_closed* $\prec$ *snowing*" represents that "*reasons to believe that is snowing provides reasons to believe that roads are closed*", whereas "$\sim$*road_closed* $\prec$ *snowplows*" represents that "*reasons to believe that snowplows are working provides reasons to believe that roads are not closed.*" The rule "*snowplows* $\prec$ " represents the presumption that *snowplows* are working.

In our approach, agents will incorporate the received information in the form of *defeasible information objects*. A defeasible information object will associate a defeasible rule (or presumption) with an informant agent.

Definition 2 (Defeasible information object) *Let* $\mathbb{I}$ *be the set of informants. A* defeasible information object *is a tuple* $(I,\ R)$*, where* $I \in \mathbb{I}$ *and* R *is a defeasible rule.*

Definition 3 (Informant-based DeLP program) *An* informant-based DeLP program *(or* IBDP *for short) is a finite set of defeasible information objects denoted as* Δ^+.

Example 2 *Consider the set of identifiers of Example 1:* $\mathbb{I} = \{\mathtt{I}_e, \mathtt{I}_p, \mathtt{I}_c, \mathtt{I}_b\}$ *where* $\mathtt{I}_e$ *is Eve,* $\mathtt{I}_p$ *is her colleague Paul,* $\mathtt{I}_b$ *Eve's boss, and* $\mathtt{I}_c$ *is Eve's assigned client. Consider the literals* t, w, r, i, p *and* i *with the following informal interpretations.*

t : *travel*
w : *work to do*
r : *Paul can replace Eve*
i : *Paul is ill*
p : *Paul has work to do*

Then, the following information-based DeLP program represents Eve's knowledge:

$$\Delta^+{}_{Eve} = \left\{ \begin{array}{llll} (\mathtt{I}_b,\ w \prec\ ), & (\mathtt{I}_p,\ r \prec {\sim}p), & (\mathtt{I}_c,\ {\sim}w \prec\ ), & (\mathtt{I}_e,\ t \prec {\sim}w), \\ (\mathtt{I}_b,\ {\sim}t \prec w), & (\mathtt{I}_p,\ t \prec r), & & (\mathtt{I}_e,\ {\sim}r \prec i), \\ (\mathtt{I}_b,\ {\sim}r \prec i), & (\mathtt{I}_p,\ {\sim}p \prec\ ), & & (\mathtt{I}_e,\ i \prec\ ) \end{array} \right\}$$

Note that $\Delta^+{}_{Eve}$ *has ten defeasible information objects that have been arranged in columns just for reader's convenience. The first column contains the information obtained from Eve's boss: the object "*$(\mathtt{I}_b,\ w \prec\ )$*" represents that "according to her boss Eve has work to do"; the object "*$(\mathtt{I}_b,\ {\sim}t \prec w)$*" states that "according to her boss if she has work to do, then she can not travel"; whereas the object* $(\mathtt{I}_b,\ {\sim}r \prec i)$ *represents that "her boss has informed her that if Paul is ill then he can not replace her". The second column contains the information obtained from Paul: "if he has no work to do then he may replace her", "if he replaces her then she may travel", and "Paul has no work to do". The third column shows the information received from Eve's client: "there is no work to do". The last column has information that Eve has from herself.*

Note that in a IBDP there can be two or more informant objects that have the same defeasible rule R but a different informant (*e.g.*, $(\mathtt{I}_e,\ {\sim}r \prec i)$ and $(\mathtt{I}_b,\ {\sim}r \prec i)$); this does not mean that there is redundancy in the agent knowledge base, since it represents that the same piece of knowledge was received from different sources. This feature can be considered as an advantage of our representation since the credibility order is dynamic, at any moment the more credible informant of R can be consider. Note also that the IBDP of an agent A can contain information objects with the agent identifier A.

In our approach, each agent will integrate four elements: its own *agent identifier*, an *informant-based DeLP program* used to store its knowledge, a *generator set* that represents the credibility order among informants, and a *trust-based argument comparison criterion* that will be used in the warranting process for deciding which information prevails when arguments for and against a conclusion exist. The first three elements have been introduced above, the trust-based argument comparison criterion will be introduced in Section 4. The definition of agent is included next.

Definition 4 (Agent) *Let* $\mathbb{I}$ *be a finite set of agent identifiers. An* agent *is a tuple* $(\mathrm{I}, \Delta^+, G\text{-}set, \succ)$ *where* $\mathrm{I} \in \mathbb{I}$, Δ^+ *is an informant-based DeLP program,* $G\text{-}set$ *is a generator set over* $\mathbb{I}$*, and "* $\succ$*" is a trust-based argument comparison criterion.*

As we will show below, two different agents $(\mathtt{I}_1, \Delta^+, G\text{-}set_1, \succ)$ and $(\mathtt{I}_2, \Delta^+, G\text{-}set_2, \succ)$ could have the same IBDP and the same comparison criterion but different generator set, and for that reason their inferred conclusions could be different. Similarly, two agents $(\mathtt{I}_3, \Delta^+, G\text{-}set, \succ_3)$ and $(\mathtt{I}_4, \Delta^+, G\text{-}set, \succ_4)$ could have the same IBDP and the same generator set but a different comparison criterion and their conclusions could also differ. Also note that since agent identifiers are unique, having the agent identifier as a component allows to have two different agents $(\mathtt{I}_5, \Delta^+, G\text{-}set, \succ)$ and $(\mathtt{I}_6, \Delta^+, G\text{-}set, \succ)$ that have the same IBDP, the same generator set and the same comparison criterion. In this last case, they will obtain the same conclusions.

Example 3 *Consider the scenario presented in the introduction, the agent Eve can be represented as* $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)$ *where* $G\text{-}set_{Eve}$ *is the generator set of Example 1 and the* IBDP $\Delta^+{}_{Eve}$ *of Example 2.*

In the following sections a trust-based argumentative reasoning formalism will be introduced. Using this formalism, the knowledge base Δ^+ the agent $(\mathrm{I}, \Delta^+, G\text{-}set, \succ)$ will use for warranting its derived beliefs contains its own information and the information objects received from other agents. A belief B will be warranted if a non-defeated argument for B exists. Intuitively, the trust-based argument comparison criterion $\succ$ will be used to decide when an argument is preferred to other contradicting argument. Hence, warranted beliefs depend on the credibility order *G-set*.

In Section 6, we will introduce the change operators of expansion, contraction and revision. The agent I can use these operator for changing the order in its *G-set*, reflecting the dynamics in the credibility assigned to its informants. Since warrant relies on the credibility attached to informants, a change in *G-set* could mean that previously warranted belief are no longer warranted, and new beliefs can become warranted. This process will reflect the dynamics of the belief base of an agent.

3 Arguments and counter-arguments

A central piece of this formalism that will allow the agent to handle contradictory information is the notion of argument; arguments will be built as support for the tentative beliefs of the agent. Intuitively, a concrete argument is a structure where a claim is supported from a set premises through the use of a reasoning mechanism; in here, the claims of the arguments will be the tentative beliefs of the agent. When considering the argument support for a particular (tentative) belief, the agent can find arguments, referred to as counterarguments, that are in conflict with it because they support contradictory claims or the claim supported by one contradicts some part of the structure the other, *i.e.*, opposes the reasoning used or is against the premises of the original argument. In this situation, it is necessary to have a way of comparing the arguments in conflict to decide which one is "better". The analysis of the situation follows a dialectical process seeking to validate the arguments in conflict. The arguments that survive all possible attack by counterarguments are said to *warrant* their conclusion or claim.

In this section we will show how an agent $(\mathrm{I}, \Delta^+, G\text{-}set, \succ)$ can build arguments using the defeasible information objects stored in its informant-based DeLP program Δ^+, we will discuss the comparison of arguments in Section 4, and the dialectical process will be described in Section 5.

As a preliminary notion before introducing argument we need to define defeasible derivation.

Definition 5 (Defeasible Derivation) *Let Δ^+ be a* IBDP *and L a ground literal. A* defeasible derivation *of L from Δ^+, denoted $\Delta^+ \mid\!\sim L$, consists of a finite sequence $L_1, L_2, \ldots, L_n = L$ of ground literals, where each literal L_i $(1 \leq n)$ is in the sequence because:*

(a) *there exists a defeasible information object $(\mathrm{I},\ L_i \prec\ )$ in Δ^+, or*

(b) *there exists a defeasible information object $(\mathrm{I},\ R)$ in Δ^+, such that R has head L_i and body $B_1, B_2, \ldots, B_k$ and every literal of the body is an element L_j of the sequence appearing before L_i $(j < i.)$*

A derivation for a literal L is called "defeasible" because as we will show next, there may exist information in contradiction with L, or any of the literals appearing in the sequence, and in certain conditions this could prevent the acceptance of L as a warranted belief. Note that rules from different informants can be combined together to derive a literal; the source of the information is not relevant for the derivation, however, as we will show when we develop our preference criterion in the following section, the source of the information will be used to appraise the strength of the conclusion's support. It is important to observe that from the same IBDP there could be several, distinct, defeasible derivations for a given literal. The following example shows how, from a given IBDP, it is possible to obtain defeasible derivations for contradictory literals.

Example 4 *Consider the* IBDP $\Delta^+{}_{Eve}$ *of example 2. From that program there exist two defeasible derivations for the literal 't'. The sequence '$\sim p, r, t$' is a defeasible derivation for 't', obtained from the defeasible information objects:* $(\mathtt{I}_p,\ \sim p \prec\ ), (\mathtt{I}_p,\ r \prec \sim p), (\mathtt{I}_p,\ t \prec r)$. *The sequence '$\sim w, t$' is another defeasible derivation for 't', obtained from* $(\mathtt{I}_c,\ \sim w \prec\ )$ *and* $(\mathtt{I}_e,\ t \prec \sim w)$. *Note that the first derivation for 't' was built using information from the same source* $(\mathtt{I}_p)$, *whereas the other uses information from different sources. From the same program $\Delta^+{}_{Eve}$ there is also a defeasible derivation for '$\sim t$', that uses the objects*

$(\mathtt{I}_b,\ w \prec\ )$ *and* $(\mathtt{I}_b,\ \sim t \prec w)$*. That is, from a* IBDP *is possible to have defeasible derivations for contradictory literals. Finally, note that the following literals have also a defeasible derivation from* $\Delta^+{}_{Eve}$*: '*w*', '*$\sim r$*', '*r*', '*$\sim p$*', '*$\sim w$*' and '*i*'.*

Definition 6 (Contradictory set of rules) *Let* Δ^+ *be an* IBDP *and* $\mathcal{A} \subseteq \Delta^+$*. We say that the set of information objects* $\mathcal{A}$ *is* contradictory *if and only if from* $\mathcal{A}$ *is possible to defeasible derive a literal* L *and its complement* $\sim L$ *.*

Example 5 *Consider the informant-based defeasible program* $\Delta^+{}_{Eve}$ *introduced in example 2. The subset* $\mathcal{A}_1 \subseteq \Delta^+{}_{Eve}$ $\mathcal{A}_1 = \{(\mathtt{I}_b,\ w \prec\ ), (\mathtt{I}_b,\ \sim t \prec w), (\mathtt{I}_p,\ r \prec \sim p), (\mathtt{I}_p,\ t \prec r)\}$ *is contradictory whereas the subset* $\mathcal{A}_2 \subseteq \Delta^+{}_{Eve}$ $\mathcal{A}_2 = \{(\mathtt{I}_b,\ w \prec\ ), (\mathtt{I}_b,\ \sim t \prec w)\}$ *is not. It is clear that if a set* $\mathcal{A}_1$ *is contradictory, then any superset of* $\mathcal{A}_1$ *is contradictory.*

As we have mentioned, the notion of *argument* plays a fundamental role in this system. Now, we will introduce the definition of argument, and the related notion of *subargument*, both are adapted from similar definitions given in [13, 5].

Definition 7 (Argument - Subargument) *Let* L *be a literal, and* Δ^+ *an* IBDP*. We say that* $\mathcal{A} \subseteq \Delta^+$ *is an* argument *for* L*, denoted* $\langle \mathcal{A}, L\rangle$*, if:*

1. *there exists a defeasible derivation of* L *from* $\mathcal{A}$*,*
2. *the set* $\mathcal{A}$ *is non-contradictory,*
3. $\mathcal{A}$ *is minimal: there is no proper subset* $\mathcal{A}'$ *of* $\mathcal{A}$ *such that* $\mathcal{A}'$ *satisfies conditions (1) and (2).*

Given two arguments $\langle \mathcal{A}_1, L_1\rangle$ *and* $\langle \mathcal{A}_2, L_2\rangle$*, we say that* $\langle \mathcal{A}_2, L_2\rangle$ *is a* subargument *of* $\langle \mathcal{A}_1, L_1\rangle$ *if and only if* $\mathcal{A}_2 \subseteq \mathcal{A}_1$*.*

The definition above characterizes an argument as a minimal and non-contradictory set of rules that allows to defeasibly derive a conclusion. Therefore, although an agent could maintain a contradictory IBDP Δ^+, agent's beliefs will be supported by arguments which are, by definition, non-contradictory.

Example 6 *Consider again the* IBDP $\Delta^+{}_{Eve}$ *of Example 2. Then the agent* $\mathtt{I}_e$ *will be able to build the following arguments:*

$$
\begin{array}{rlrl}
\langle \mathcal{A}_1, w\rangle = & \langle\{(\mathtt{I}_b,\ w \prec\ )\}, w\rangle & \langle \mathcal{A}_2, \sim t\rangle = & \langle\{(\mathtt{I}_b,\ \sim t \prec w), (\mathtt{I}_b,\ w \prec\ )\}, \sim t\rangle \\
\langle \mathcal{A}_3, \sim r\rangle = & \langle\{(\mathtt{I}_b,\ \sim r \prec i), (\mathtt{I}_e,\ i \prec\ )\}, \sim r\rangle & \langle \mathcal{A}_4, r\rangle = & \langle\{(\mathtt{I}_p,\ r \prec \sim p), (\mathtt{I}_p,\ \sim p \prec\ )\}, r\rangle \\
\langle \mathcal{A}_5, t\rangle = & \langle\{(\mathtt{I}_p,\ t \prec r), (\mathtt{I}_p,\ r \prec \sim p), (\mathtt{I}_p,\ \sim p \prec\ )\}, t\rangle & \langle \mathcal{A}_6, \sim p\rangle = & \langle\{(\mathtt{I}_p,\ \sim p \prec\ )\}, \sim p\rangle \\
\langle \mathcal{A}_7, \sim w\rangle = & \langle\{(\mathtt{I}_c,\ \sim w \prec\ )\}, \sim w\rangle & \langle \mathcal{A}_8, t\rangle = & \langle\{(\mathtt{I}_e,\ t \prec \sim w), (\mathtt{I}_c,\ \sim w \prec\ )\}, t\rangle \\
\langle \mathcal{A}_9, \sim r\rangle = & \langle\{(\mathtt{I}_e,\ \sim r \prec i), (\mathtt{I}_e,\ i \prec\ )\}, \sim r\rangle & \langle \mathcal{A}_{10}, i\rangle = & \langle\{(\mathtt{I}_e,\ i \prec\ )\}, i\rangle
\end{array}
$$

Observe that $\langle \mathcal{A}_4, r\rangle$ *is a subargument of* $\langle \mathcal{A}_5, t\rangle$*. Also note that although arguments* $\langle \mathcal{A}_3, \sim r\rangle$ *and* $\langle \mathcal{A}_9, \sim r\rangle$ *share the defeasible rule* $\sim r \prec i$*, they differ in the informant recorded as part of the information object in each case:* $\mathtt{I}_b$ *for* $\mathcal{A}_3$*, and* $\mathtt{I}_e$ *for* $\mathcal{A}_9$*. As we develop the system, the importance of maintaining this information will become clear. The main reason for this is the credibility order among informants may change and, therefore, the argument* $\mathcal{A}_3$ *for "*$\sim r$*" may became stronger than the argument* $\mathcal{A}_9$ *for "*$\sim r$*" which uses the same defeasible rule but different informants.*

From the examples above, it is clear that from an IBDP there can be arguments that are in conflict: for instance, $\langle \mathcal{A}_3, \sim r\rangle$ and $\langle \mathcal{A}_4, r\rangle$ or $\langle \mathcal{A}_3, \sim r\rangle$ and $\langle \mathcal{A}_5, t\rangle$. This situation is captured by the notion of *counterargument* or *attack*.

Definition 8 (Attack) *Let* $\langle \mathcal{A}_1, L_1\rangle$ *and* $\langle \mathcal{A}_2, L_2\rangle$ *be two arguments. We say that* $\langle \mathcal{A}_1, L_1\rangle$ counterargues*, or* attacks $\langle \mathcal{A}_2, L_2\rangle$ *at literal* L*, if and only if there exists a subargument* $\langle \mathcal{A}, L\rangle$ *of* $\langle \mathcal{A}_2, L_2\rangle$ *such that* L *and* L_1 disagree, i.e., L *and* L_1 *are complementary literals regarding "*$\sim$*".*

In Example 6 for instance, the argument $\langle \mathcal{A}_3, {\sim}r\rangle$ attacks the argument $\langle \mathcal{A}_5, t\rangle$ because $\langle \mathcal{A}_5, t\rangle$ has $\langle \mathcal{A}_4, r\rangle$ as a subargument and "${\sim}r$" disagrees with "r". Also note that the argument $\langle \mathcal{A}_5, t\rangle$ attacks the argument $\langle \mathcal{A}_2, {\sim}t\rangle$ since "t" disagrees with "${\sim}t$" (an argument is a subargument of itself). Observe that it also is the case that $\langle \mathcal{A}_2, {\sim}t\rangle$ attacks $\langle \mathcal{A}_5, t\rangle$. Usually attacks to intermediate conclusions (like the one from $\langle \mathcal{A}_3, {\sim}r\rangle$ to $\langle \mathcal{A}_5, t\rangle$) are known as "inner attacks", whereas attacks directly to the conclusion are known as "direct attacks" (like the one from $\langle \mathcal{A}_5, t\rangle$ to $\langle \mathcal{A}_2, {\sim}t\rangle$).

Remark 1 *If an argument $\langle \mathcal{A}_1, L_1\rangle$ attacks $\langle \mathcal{A}, L\rangle$ at literal L, then $\langle \mathcal{A}_1, L_1\rangle$ also attacks any other $\langle \mathcal{A}_2, L_2\rangle$ such that $\langle \mathcal{A}, L\rangle$ is a subargument of $\langle \mathcal{A}_2, L_2\rangle$. Observe also that the attack notion has the property that every time an argument $\langle \mathcal{A}_1, L_1\rangle$ attacks other argument $\langle \mathcal{A}_2, L_2\rangle$, then, there always exist other argument that attacks $\langle \mathcal{A}_1, L_1\rangle$,* i.e., *the corresponding subargument of $\langle \mathcal{A}_2, L_2\rangle$ whose conclusion disagrees with L_1.*

To establish whether $\langle \mathcal{A}, L\rangle$ is a non-defeated argument, it is necessary to analyze all of its associated counterarguments; let $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_k$ be all the counterarguments of $\mathcal{A}$, each one of them potentially being a (defeasible) reason for rejecting $\mathcal{A}$. If any $\mathcal{B}_i$ is (somehow) "better" than, or unrelated to, $\mathcal{A}$, then $\mathcal{B}_i$ is a candidate for defeating $\mathcal{A}$. However, if the argument $\mathcal{A}$ is "better" than the argument $\mathcal{B}_i$, then $\mathcal{B}_i$ will not be taken in consideration as a defeater $\mathcal{A}$. Therefore, to decide on the result of an attack we need to introduce a preference relation that will settle the question. We will do that next with the help of the trust relation among agents.

4 Trust-based argument comparison criteria

Until now the dialectical process considers the support of piece of information in form of an argument. We have progressed to the point we can consider the potential attacks of one argument over other. In this section we will introduce comparison criteria that will establish if an attack is successful; in that case, the counterargument involved will be successful becoming a defeater and the argument under attack will become defeated.

Consider the following arguments from Example 6 as an illustration of a conflictive situation:

$$\begin{array}{lcl}\langle \mathcal{A}_2, {\sim}t\rangle & = & \langle \{(\mathtt{I}_b,\ {\sim}t \prec w), (\mathtt{I}_b,\ w \prec\ )\}, {\sim}t\rangle \\ \langle \mathcal{A}_5, t\rangle & = & \langle \{(\mathtt{I}_p,\ t \prec r), (\mathtt{I}_p,\ r \prec {\sim}p), (\mathtt{I}_p,\ {\sim}p \prec\ )\}, t\rangle\end{array}$$

In this case, both arguments are in conflict and they are in a situation of mutual attack. To decide whether one of them *defeats* the other, we need a way of comparing them. The following definition introduces this idea in an abstract manner.

Definition 9 (Argument comparison criterion) *Let Args be the set of arguments that can be obtained from a* IBDP. *An argument comparison criterion is a relation $\succ \subseteq \mathsf{Args} \times \mathsf{Args}$, that is irreflexive,* i.e., *if $\mathcal{A} \in \mathsf{Args}$ then $\mathcal{A} \not\succ \mathcal{A}$; and asymmetric,* i.e., *given $\mathcal{A}, \mathcal{B} \in \mathsf{Args}$ when $\mathcal{A} \succ \mathcal{B}$ it is the case that $\mathcal{B} \not\succ \mathcal{A}$.*

Notice that the "$\succ$" relation on Args is not assumed to be transitive.

In this section we will develop three different concrete argument comparison criteria that are based on informants credibility. However, first we need to introduce a more atomic preference relation among defeasible information objects. This notion formalizes how trust is used by agents for compare knowledge that it is explicitly stored in their knowledge bases (Δ^+) and will be used below by concrete argument comparison criteria. Note that the following definition captures the idea that information from a more credible agent should be preferred over the information of a less credible agent.

Definition 10 (Defeasible information object preference) *Let G-set* be a generator set and let $(\mathtt{I}_1,\ R)$, $(\mathtt{I}_2,\ K)$ be two defeasible information objects (dio). We say that $(\mathtt{I}_1,\ R)$ is preferred over $(\mathtt{I}_2,\ K)$, noted by $(\mathtt{I}_1,\ R) >_{dio} (\mathtt{I}_2,\ K)$, if and only if $(\mathtt{I}_2,\ \mathtt{I}_1) \in$ G-set**, i.e., *$\mathtt{I}_2$ is less credible than $\mathtt{I}_1$.*

Example 7 *Consider the program $\Delta^+{}_{Eve}$ of Example 2 and the generator set G-set$_{Eve}$ of Example 1, G-set$_{Eve}$ = $\{(\mathtt{I}_p, \mathtt{I}_e), (\mathtt{I}_p, \mathtt{I}_b), (\mathtt{I}_b, \mathtt{I}_c), (\mathtt{I}_e, \mathtt{I}_c)\}$. In this case, we have that $(\mathtt{I}_c\ , {\sim}w \prec\ )$ is preferred to $(\mathtt{I}_p\ , r \prec {\sim}p)$ since $(\mathtt{I}_p, \mathtt{I}_c) \in$ G-set*. Also note that $(\mathtt{I}_e\ , t \prec {\sim}w)$ is not preferred to $(\mathtt{I}_b\ , {\sim}t \prec w)$ (or viceversa) because $\mathtt{I}_e$ and $\mathtt{I}_b$ are incomparable according to this agent.*

As we will show below, based on the defeasible information object preference ($>_{dio}$) defined above, several trust-based argument comparison criteria can be defined. This notion is modular, *i.e.*, the argument comparison criterion can be replaced without affecting the rest of the formalism; therefore, it is possible to select the more appropriate criterion for a particular application domain. Observe that the definition of an agent as $(\mathrm{I}, \Delta^+, G\text{-}set, \succ)$ contains an argument comparison criterion "$\succ$ ", in such a way that it can be conveniently changed.

The first criterion is based on the credibility attached to the informants and will be called *single rule credibility criterion*. This criterion will prefer an argument $\mathcal{A}$ over an argument $\mathcal{B}$ if there is at least one defeasible information object in $\mathcal{A}$ with an informant that is more credible than an informant in some defeasible information object in $\mathcal{B}$, and no defeasible information object in $\mathcal{B}$ is preferred to some other in $\mathcal{A}$. This intuition is formalized in the following definition.

Definition 11 (Single rule credibility criterion) *Let $\langle\mathcal{A}, L\rangle$ and $\langle\mathcal{B}, Q\rangle$ be two arguments built from an* IBDP Δ^+. *We will say that $\langle\mathcal{A}, L\rangle$ is s-preferred to $\langle\mathcal{B}, Q\rangle$, noted $\langle\mathcal{A}, L\rangle \succ_S \langle\mathcal{B}, Q\rangle$ iff:*

- *There are $(\mathtt{I}_i,\ R) \in \mathcal{A}$ and $(\mathtt{I}_j,\ W) \in \mathcal{B}$ such that $(\mathtt{I}_i,\ R)>_{dio}(\mathtt{I}_j,\ W)$, and*
- *There is no $(\mathtt{I}_k,\ V) \in \mathcal{A}$ and $(\mathtt{I}_l,\ U) \in \mathcal{B}$ such that $(\mathtt{I}_l,\ U)>_{dio}(\mathtt{I}_k,\ V)$.*

Proposition 1 *The relation $\succ_S$ is an argument comparison criterion.*
Proof: *We have to prove that for all $\mathcal{A}, \mathcal{B} \in$ Args:*

1. *$\mathcal{A}\succ_S\mathcal{A}$ does not hold (irreflexivity).*

   We will use reductio ad absurdum. Suppose $\mathcal{A} \succ_S\mathcal{A}$. Then, there are $(\mathtt{I}_i,\ R), (\mathtt{I}_j,\ W) \in \mathcal{A}$ such that $(\mathtt{I}_i,\ R)>_{dio}(\mathtt{I}_j,\ W)$, which immediately contradicts the second condition of the Definition 11.

2. *if $\mathcal{A}\succ_S\mathcal{B}$ then $\mathcal{B}\not\succ_S\mathcal{A}$ (asymmetry).*

   Suppose $\mathcal{A}\succ_S\mathcal{B}$. Then, there are $(\mathtt{I}_i,\ R) \in \mathcal{A}$ and $(\mathtt{I}_j,\ W) \in \mathcal{B}$ such that $(\mathtt{I}_i,\ R)>_{dio}(\mathtt{I}_j,\ W)$, and there is no $(\mathtt{I}_k,\ V) \in \mathcal{A}$ and $(\mathtt{I}_l,\ U) \in \mathcal{B}$ such that $(\mathtt{I}_l,\ U)>_{dio}(\mathtt{I}_k,\ V)$. Thus, there is no $(\mathtt{I}_h,\ Y) \in \mathcal{B}$ and $(\mathtt{I}_g,\ Z) \in \mathcal{A}$ such that $(\mathtt{I}_h,\ Y)>_{dio}(\mathtt{I}_g,\ Z)$ violating the first condition of the Definition 11. Hence, $\mathcal{B}\not\succ_S\mathcal{A}$.

Example 8 *Consider the agent $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)$ of Example 3 and its arguments presented in Example 6. Consider now that the selected comparison criteria $\succ$ is $\succ_S$. Note that the argument $\langle\mathcal{A}_8, t\rangle = \langle\{(\mathtt{I}_e,\ t \prec {\sim}w), (\mathtt{I}_c,\ {\sim}w \prec\ )\}, t\rangle$ is s-preferred to $\langle\mathcal{A}_2, {\sim}t\rangle = \langle\{(\mathtt{I}_b,\ {\sim}t \prec w), (\mathtt{I}_b,\ w \prec\ )\}, {\sim}t\rangle$ since the defeasible information object $(\mathtt{I}_c,\ {\sim}w \prec\ )$ from $\langle\mathcal{A}_8, t\rangle$ is preferred to $(\mathtt{I}_b,\ w \prec\ )$ from $\langle\mathcal{A}_2, {\sim}t\rangle$, and there is no defeasible information object from $\langle\mathcal{A}_2, {\sim}t\rangle$ preferred to some from $\langle\mathcal{A}_8, t\rangle$.*

The single rule criterion is not discriminative enough since, in the general case, will leave many comparisons of arguments undecided. Next, we will consider the criterion of *least credible rules* that takes a more refined approach. To realize the comparison this criterion considers just those defeasible information objects of each argument with the least credible informant. Previously to the introduction of this criterion, we will first define which are the least credible defeasible objects in an argument.

Definition 12 (Least credible defeasible information objects) *Let $\langle\mathcal{A}, L\rangle$ an argument of an agent Ag. The least credible defeasible information objects $\langle\mathcal{A}, L\rangle$, noted* $\mathsf{Min}(\langle\mathcal{A}, L\rangle)$*, is a set such that $(\mathtt{I}_1,\ R) \in \mathsf{Min}(\langle\mathcal{A}, L\rangle)$ if and only if $(\mathtt{I}_1,\ R) \in \mathcal{A}$ and there is no $(\mathtt{I}_2,\ K) \in \mathcal{A}$ such that $(\mathtt{I}_1,\ R)>_{dio}(\mathtt{I}_2,\ K)$.*

For instance, the argument $\langle\mathcal{A}_8, t\rangle = \langle\{(\mathtt{I}_e,\ t \prec {\sim}w), (\mathtt{I}_c,\ {\sim}w \prec\ )\}, t\rangle$ of the agent $\mathtt{I}_e$ from Example 6 will be such that $\mathit{Min}(\langle\mathcal{A}_8, t\rangle) = \{(\mathtt{I}_e,\ t \prec {\sim}w)\}$. Note that the *Min* set of an argument can contain several defeasible information objects. Next we will formalize the intuition introduced above for the *least credibility criterion*.

Definition 13 (Least credible rules criterion) *Let $\langle\mathcal{A}_1, L_1\rangle$ and $\langle\mathcal{A}_2, L_2\rangle$ be two arguments of an* IBDP Δ^+. *We will say that $\langle\mathcal{A}_1, L_1\rangle$ is l-preferred to $\langle\mathcal{A}_2, L_2\rangle$, noted $\langle\mathcal{A}_1, L_1\rangle \succ_L \langle\mathcal{A}_2, L_2\rangle$ iff for every $(\mathtt{I}_i,\ R) \in$* $\mathsf{Min}(\langle\mathcal{A}_1, L_1\rangle)$ *and for every $(\mathtt{I}_j,\ K) \in$* $\mathsf{Min}(\langle\mathcal{A}_2, L_2\rangle)$ *it holds that $(\mathtt{I}_i,\ R)>_{dio}(\mathtt{I}_j,\ K)$.*

Proposition 2 *The relation $\succ_L$ is a an argument comparison criterion.*
Proof: *We have to prove that for all $\mathcal{A}, \mathcal{B} \in$ Args:*

1. *$\mathcal{A} \succ_L \mathcal{A}$ does not hold (irreflexivity).*
   Straightforward by Definition 13 and Definition 12.
2. *if $\mathcal{A} \succ_L \mathcal{B}$ then $\mathcal{B} \not\succ_L \mathcal{A}$ (asymmetry).*
   Suppose $\mathcal{A} \succ_L \mathcal{B}$. Then for every $(\mathtt{I}_i,\ R) \in \mathsf{Min}(\langle \mathcal{A}_1, L_1\rangle)$ and for every $(\mathtt{I}_j,\ K) \in \mathsf{Min}(\langle \mathcal{A}_2, L_2\rangle)$ it holds that $(\mathtt{I}_i,\ R) >_{dio} (\mathtt{I}_j,\ K)$. Thus, there is no $(\mathtt{I}_k,\ Y) \in \mathsf{Min}(\langle \mathcal{A}_2, L_2\rangle)$ and $(\mathtt{I}_h,\ Z) \in \mathsf{Min}(\langle \mathcal{A}_1, L_1\rangle)$ it holds that $(\mathtt{I}_k,\ Y) >_{dio} (\mathtt{I}_h,\ Z)$. Hence, $\mathcal{B} \not\succ_L \mathcal{A}$.

Example 9 *Consider the agent $(I_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)$ of Example 3 and its arguments presented in Example 6. Consider now that the selected comparison criteria $\succ$ is $\succ_L$. Observe that the argument $\langle \mathcal{A}_7, \sim w\rangle = \langle \{(\mathtt{I}_c,\ \sim w \prec)\}, \sim w\rangle$ is l-preferred to $\langle \mathcal{A}_8, t\rangle = \langle \{(\mathtt{I}_e,\ t \prec \sim w), (\mathtt{I}_c,\ \sim w \prec)\}, t\rangle$ because the element in $\mathsf{Min}(\langle \mathcal{A}_7, \sim w\rangle) = \{(\mathtt{I}_c,\ \sim w \prec)\}$ is preferred to the only element in $\mathsf{Min}(\langle \mathcal{A}_8, t\rangle) = \{(\mathtt{I}_e,\ t \prec \sim w)\}$. Note that $\langle \mathcal{A}_8, t\rangle$ is not l-preferred to $\langle \mathcal{A}_2, \sim t\rangle = \langle \{(\mathtt{I}_b,\ \sim t \prec w), (\mathtt{I}_b,\ w \prec)\}, \sim t\rangle$ since the element in $\mathsf{Min}(\langle \mathcal{A}_8, t\rangle)$ is not preferred to the ones in $\mathsf{Min}(\langle \mathcal{A}_2, \sim t\rangle) = \{(\mathtt{I}_b,\ \sim t \prec w), (\mathtt{I}_b,\ w \prec)\}$.*

The last criterion we will present is the *multiple rule credibility criterion*, which also takes a more careful approach than our previously introduced single credibility criterion. This criterion will prefer an argument $\mathcal{A}$ over another $\mathcal{B}$ when every defeasible information object from $\mathcal{A}$ is preferred to some defeasible information object from $\mathcal{B}$, and there is no defeasible information object in $\mathcal{B}$ that is preferred to one in $\mathcal{A}$.

Definition 14 (Multiple rule credibility criterion) *Let $\langle \mathcal{A}, L\rangle$ and $\langle \mathcal{B}, Q\rangle$ be two arguments of an* IBDP *Δ^+. We will say that $\langle \mathcal{A}, L\rangle$ is m-preferred to $\langle \mathcal{B}, Q\rangle$, noted $\langle \mathcal{A}, L\rangle \succ_M \langle \mathcal{B}, Q\rangle$ iff:*

- *For all $(I_i,\ U) \in \mathcal{A}$ and there exists $(I_j,\ V) \in \mathcal{B}$ such that $(I_i,\ U) >_{dio} (I_j,\ V)$, and*
- *There is no $(I_k,\ X) \in \mathcal{A}$ such that there exists $(I_l,\ Y) \in \mathcal{B}$ such that $(I_l,\ Y) >_{dio} (I_k,\ X)$*

Proposition 3 *The relation $\succ_M$ is a an argument comparison criterion.*
Proof: *We have to prove that for all $\mathcal{A}, \mathcal{B} \in$ Args:*

1. *$\mathcal{A} \succ_M \mathcal{A}$ does not hold (irreflexivity).*
   If for all $(I_i,\ U) \in \mathcal{A}_1$ and there is $(I_j,\ V) \in \mathcal{A}$ such that $(I_i,\ U) >_{dio} (I_j,\ V)$, then it does not hold that there is no $(I_k,\ X) \in \mathcal{A}$ such that there is a $(I_l,\ Y) \in \mathcal{A}$ holding that $(I_l,\ Y) >_{dio} (I_k,\ X)$ which contradicts the second condition of Definition 14. Hence, $\mathcal{A} \succ_M \mathcal{A}$ does not hold.
2. *if $\mathcal{A} \succ_M \mathcal{B}$ then $\mathcal{B} \not\succ_M \mathcal{A}$ (asymmetry)*
   Suppose $\mathcal{A} \succ_M \mathcal{B}$. Then, for all $(I_i,\ U) \in \mathcal{A}_1$ and there is $(I_j,\ V) \in \mathcal{B}$ such that $(I_i,\ U) >_{dio} (I_j,\ V)$, and there is no $(I_k,\ X) \in \mathcal{A}$ such that there is a $(I_l,\ Y) \in \mathcal{B}$ holding that $(I_l,\ Y) >_{dio} (I_k,\ X)$. Thus, there is no $(I_h,\ Y) \in \mathcal{B}$ and there is $(I_g,\ Z) \in \mathcal{A}$ such that $(I_h,\ Y) >_{dio} (I_g,\ Z)$ violating the first condition of the Definition 14. Hence, $\mathcal{B} \not\succ_M \mathcal{A}$.

Example 10 *Consider the agent $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)$ in Example 3 and its arguments presented in Example 6. Consider now that the selected comparison criteria $\succ$ is $\succ_M$. Observe that the argument $\langle \mathcal{A}_8, t\rangle = \langle \{(\mathtt{I}_e,\ t \prec \sim w), (\mathtt{I}_c,\ \sim w \prec)\}, t\rangle$ is preferred to $\langle \mathcal{A}_4, r\rangle = \langle \{(\mathtt{I}_p,\ r \prec \sim p), (\mathtt{I}_p,\ \sim p \prec)\}, r\rangle$ under the multiple rule credibility criterion because every element in $\mathcal{A}_8$ is preferred to an element in $\mathcal{A}_4$, and no element in $\mathcal{A}_4$ is preferred to one in $\mathcal{A}_8$.*

Below, in Figure 3 we present a table showing the comparison for the arguments of Example 6 using the three criterions based on informants credibility previously defined. In this table we compare pairs of arguments that directly attack each other. Each row corresponds to a pair of arguments, and the columns to the criterion used to compare them. The special symbol "≈" is used to denote that no argument in the pair is preferred in using the criterion identified in column.

Arguments	$\succ_S$	$\succ_L$	$\succ_M$
$\langle\mathcal{A}_1, w\rangle, \langle\mathcal{A}_7, \sim w\rangle$	$\langle\mathcal{A}_7, \sim w\rangle \succ_S \langle\mathcal{A}_1, w\rangle$	$\langle\mathcal{A}_7, \sim w\rangle \succ_L \langle\mathcal{A}_1, w\rangle$	$\langle\mathcal{A}_7, \sim w\rangle \succ_M \langle\mathcal{A}_1, w\rangle$
$\langle\mathcal{A}_2, \sim t\rangle, \langle\mathcal{A}_5, t\rangle$	$\langle\mathcal{A}_2, \sim t\rangle \succ_S \langle\mathcal{A}_5, t\rangle$	$\langle\mathcal{A}_2, \sim t\rangle \succ_L \langle\mathcal{A}_5, t\rangle$	$\langle\mathcal{A}_2, \sim t\rangle \succ_M \langle\mathcal{A}_5, t\rangle$
$\langle\mathcal{A}_2, \sim t\rangle, \langle\mathcal{A}_8, t\rangle$	$\langle\mathcal{A}_8, t\rangle \succ_S \langle\mathcal{A}_2, \sim t\rangle$	$\langle\mathcal{A}_2, \sim t\rangle \approx \langle\mathcal{A}_8, t\rangle$	$\langle\mathcal{A}_2, \sim t\rangle \approx \langle\mathcal{A}_8, t\rangle$
$\langle\mathcal{A}_3, \sim r\rangle, \langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_3, \sim r\rangle \succ_S \langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_3, \sim r\rangle \succ_L \langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_3, \sim r\rangle \succ_M \langle\mathcal{A}_4, r\rangle$
$\langle\mathcal{A}_4, r\rangle, \langle\mathcal{A}_9, \sim r\rangle$	$\langle\mathcal{A}_9, \sim r\rangle \succ_S \langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_9, \sim r\rangle \succ_L \langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_9, \sim r\rangle \succ_M \langle\mathcal{A}_4, r\rangle$

Figure 3: Table comparing agent $\mathtt{I}_e$ arguments using the credibility based criterions

5 Defeat and Warrant

Agents build arguments from their knowledge bases to support their beliefs; but, the possibility of conflict among them exists bringing the focus on the notion of attack that was analyzed in Section 3. Attacks could succeed or fail and to decide what happens a comparison criterion becomes necessary. In Section 4, we have characterized preference criteria that can be applied to arguments built in the context of an IBDP Δ^+ to decide which one is better. We will now put these things together to introduce the inference mechanism that will obtain the beliefs of an agent.

Informally, the notion of defeat can be considered as an attack that is effective; an attack will be deemed effective when the argument that attacks is better than the argument that receives the attack. The formal definition, adapted from [5], follows.

Definition 15 (Defeater) *Let $\langle\mathcal{A}_1, L_1\rangle$ and $\langle\mathcal{A}_2, L_2\rangle$ be two arguments build from an* IBDP Δ^+*, and let $\succ$ be an argument comparison criterion. The argument $\langle\mathcal{A}_1, L_1\rangle$ is a* defeater *for $\langle\mathcal{A}_2, L_2\rangle$ at literal L under $\succ$, if and only if there exists a subargument $\langle\mathcal{A}, L\rangle$ of $\langle\mathcal{A}_2, L_2\rangle$ such that $\langle\mathcal{A}_1, L_1\rangle$ counterargues $\langle\mathcal{A}_2, L_2\rangle$ at L, and $\langle\mathcal{A}, L\rangle$ is not preferred to $\langle\mathcal{A}_1, L_1\rangle$ under $\succ$.*

When the counterargument $\langle\mathcal{A}_1, L_1\rangle$ is better than $\langle\mathcal{A}, L\rangle$ with respect to the comparison criterion used, then $\langle\mathcal{A}_1, L_1\rangle$ will be called a *proper defeater* for $\langle\mathcal{A}_2, L_2\rangle$. If the comparison is undecided, *i.e.*, neither argument is better, nor worse, than the other, a blocking situation occurs, and we will say that $\langle\mathcal{A}_1, L_1\rangle$ is a *blocking defeater* for $\langle\mathcal{A}_2, L_2\rangle$. If $\langle\mathcal{A}_2, L_2\rangle$ is better than its attacker $\langle\mathcal{A}_1, L_1\rangle$, then $\langle\mathcal{A}_1, L_1\rangle$ will not be considered as a defeater for $\langle\mathcal{A}_2, L_2\rangle$.

Example 11 *Consider the agent $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)$ of Example 3 and its arguments presented in Example 6. Using the $\succ_S$ comparison criterion, the argument $\langle\mathcal{A}_8, t\rangle$ is a defeater for $\langle\mathcal{A}_2, \sim t\rangle$, since $\langle\mathcal{A}_8, t\rangle$ is a counter-argument for $\langle\mathcal{A}_2, \sim t\rangle$ and as we have shown in Example 8 $\langle\mathcal{A}_8, t\rangle \succ_S \langle\mathcal{A}_2, \sim t\rangle$. Observe that $\langle\mathcal{A}_2, \sim t\rangle$ is not a defeater for $\langle\mathcal{A}_8, t\rangle$, then $\langle\mathcal{A}_8, t\rangle$ is a proper defeater for $\langle\mathcal{A}_2, \sim t\rangle$. Using the $\succ_L$ criterion the argument $\langle\mathcal{A}_8, t\rangle$ is a defeater for $\langle\mathcal{A}_2, \sim t\rangle$, but also $\langle\mathcal{A}_2, \sim t\rangle$ defeats $\langle\mathcal{A}_8, t\rangle$. That is, under $\succ_L$ the arguments $\langle\mathcal{A}_2, \sim t\rangle$ and $\langle\mathcal{A}_8, t\rangle$ are blocking defeaters.*

To determine whether a belief L can be accepted it is necessary to find out if there exists at least an undefeated argument $\mathcal{A}$ that supports that belief; naturally, this will lead to analyze all possible attackers to see if any of them can be regarded as a defeater of $\langle\mathcal{A}, L\rangle$. In case that any defeater $\langle\mathcal{B}, Q\rangle$ of $\langle\mathcal{A}, L\rangle$ is found we need to repeat the process of considering the possible defeaters of $\langle\mathcal{B}, Q\rangle$, and so on. This dialectical analysis will create a tree structure of defeaters, called *dialectical tree*, where every node below the root is a defeater of its parent, and the leaves are undefeated arguments (for full details see [5]). In a dialectical tree every node can be marked as defeated "D" or undefeated "U"; leaves are marked "U" and the inner nodes, including the root, are marked "D" when they have at least a children marked "U", or marked "U" when all their children are marked "D".

Example 12 *Consider the agent $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)$ of Example 3 and its arguments presented in Example 6. In the Figure 4 we show the dialectical trees for the arguments $\langle\mathcal{A}_5, t\rangle$, $\langle\mathcal{A}_8, t\rangle$ and $\langle\mathcal{A}_2, \sim t\rangle$. In particular, in the Figure 4(a) we show the dialectical trees T_1, T_2 and T_3 for these arguments under the*

$\succ_S$ comparison criterion and in the Figure 4(b) the trees T_4, T_5 and T_6 are built using the $\succ_L$ criterion. In both figures arguments are depicted using triangles, proper defeats using filled arrows and blocking defeats using dashed arrows. The circles beside the defeasible rule symbol in the arguments are used to illustrate the informant associated to the defeasible information objects for the respective argument. Also observe that each argument is labeled defeated or undefeated with a circle containing a **D** *or a* **U** *respectively.*

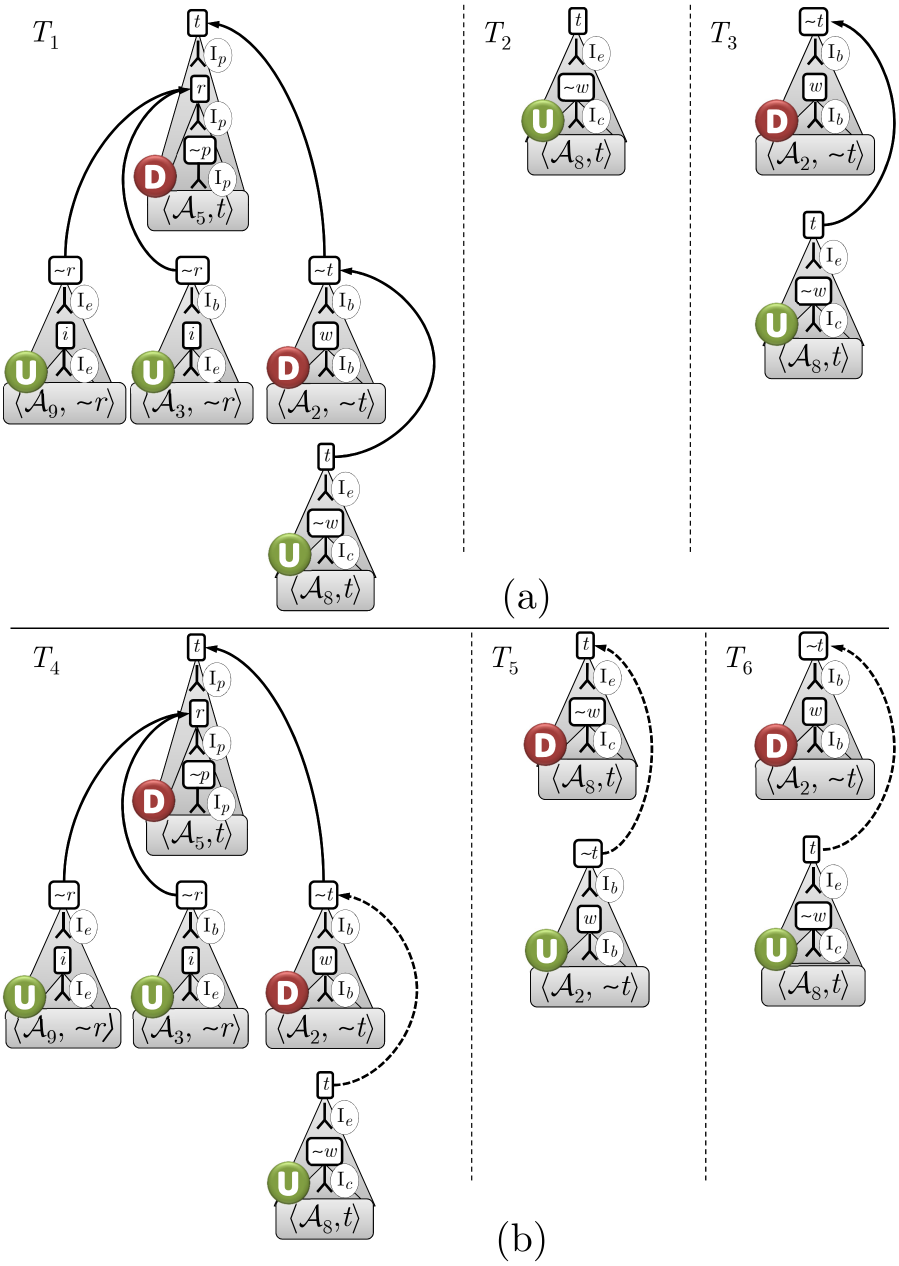

Figure 4: Dialectical trees for the arguments supporting "t" and "$\sim t$" using the criterion $\succ_S$ in (a) and $\succ_L$ in (b)

A marked dialectical tree represents a dialectical analysis considering every possible argument the agent can build for and against the argument in the root of the tree. Therefore if the the root argument is marked undefeated, it means that the conclusion of the argument is warranted and can be considered a belief for the agent; one warranted argument is enough as the support for a particular belief. Contrariwise, if every argument for a literal is marked as defeated in its own tree, then this literal will not be warranted, thus the literal cannot be a belief for the agent.

Definition 16 (Warranted literals) *Let* $(\mathrm{I}, \Delta^+, G\text{-}set, \succ)$ *be an agent and let* L *be a literal, then* $(\mathrm{I}, \Delta^+, G\text{-}set, \succ)$ *warrants* L *iff there is an argument* $\langle \mathcal{A}, L\rangle$ *from* Δ^+ *such that there is a dialectical tree under* $\succ$ *with* $\langle \mathcal{A}, L\rangle$ *as root argument and* $\langle \mathcal{A}, L\rangle$ *is marked defeated.*

Example 13 *Consider the agent* $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)$ *of Example 3 and the marked dialectical trees presented in Figure 4. If* $\succ$ *is the* $\succ_S$ *criterion this agent will have "t" as warranted since, as depicted in Figure 4(a), the argument* $\langle \mathcal{A}_8, t\rangle$ *is marked as "U" in its dialectical tree* T_2*. Note that using* $\succ_S$ *the literal "∼t" is not warranted since the only argument that supports it* $\langle \mathcal{A}_2, \sim t\rangle$ *is marked as defeated in its own tree* T_3*. It can be seen from* T_4*,* T_5 *and* T_6 *in Figure 4(b) that using the* $\succ_L$ *criterion our agent does not warrant neither "t" nor "∼t" because all the dialectical trees have their roots marked as "D".*

6 Change operators for the dynamics of the credibility order

In the preceding sections, we have introduced a trust-based argumentation formalism where agent's conclusions are supported by arguments built using information from different sources. When that information is in conflict, the credibility attached to the informants is used in the decision process leading to a prevailing conclusion. Since in many application domains, the credibility of informant agents changes, we will introduce the final piece in our framework to handle this common characteristic. We will introduce change through the classic operators of expansion, contraction and revision. An agent can use these operators for modifying its credibility order stored in its *G-set*; this will allow to modify the credibility of its informants dynamically. Since the warranting process relies on the credibility of informants, a change in its *G-set* can bring about changes in the set of warranted beliefs.

In [14] a form of updating the credibility order to reflect the change in the perceived agent's credibility was introduced. Their model formalizes change operators over that credibility order, and provides the capability of dynamically modifying the credibility of informants to reflect a new perception of the informant's plausibility; also provides means for extending the set of informants by admitting the arrival of new agents to the system. While the concept of trust is complex, in that work the authors have taken the position that trust can be seen as a type of credibility value the agents can assign to each other, position we will adhere here. In Section 2 we have already introduced the notion of Generator Set following [14], and now we will adapt the change operators proposed in that work to our framework. We will formalize the expansion and contraction operators for the credibility order of an agent and, using the *Levi identity* [6, 1], we will define the revision operator out of the contraction and expansion operators already defined.

6.1 Expansion operator for credibility orders

Here we will introduce an expansion operator for credibility partial orders called C-expansion, denoted $\oplus^C$, and adapted to the agent structure given in Definition 4. A C-expansion consists simply in the addition of a new credibility tuple to a generator set. Given an ordered pair of informants $(\mathtt{I}_1, \mathtt{I}_2)$ and a generator set *G-set*, this operator returns a new generator set *G-set*′ in which said agents are now related, *i.e.*, $(\mathtt{I}_1, \mathtt{I}_2) \in G\text{-}set'$. According to this new generator set we may say that the first informant is "less credible" than the second. The construction of expansions on credibility partial order is formally defined as follows.

Definition 17 (C-expansion) *Let* $\mathtt{I}_1, \mathtt{I}_2, \mathtt{I}_3 \in \mathbb{I}$ *be three agent identifiers and let* $(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)$ *an agent. The operator "*$\oplus^C$*", called C-expansion is defined as follow:*

$$(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)\oplus^C(\mathtt{I}_2,\mathtt{I}_3) = (I_1, \Delta^+{}_1, G\text{-}set_1 \cup \{(\mathtt{I}_2,\mathtt{I}_3)\}, \succ)$$

Expansion does not preserve soundness in the generator set *per se*, but is conditioned. That is, if *G-set* is a sound generator set and $(\mathtt{I}_1, \mathtt{I}_2) \notin G\text{-}set^*$ then $G\text{-}set \cup \{(\mathtt{I}_2,\mathtt{I}_1)\}$ is a sound generator set.

Example 14 *Consider agent* $\mathtt{I}_e$ *of Example 3 where* $G\text{-}set_{Eve} = \{(\mathtt{I}_p,\mathtt{I}_b), (\mathtt{I}_p,\mathtt{I}_e), (\mathtt{I}_b, \mathtt{I}_c), (\mathtt{I}_e,\mathtt{I}_c)\}$ *(see Figure 2), then,* $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ) \oplus^C (\mathtt{I}_e,\mathtt{I}_b) = (\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve} \cup \{(\mathtt{I}_e,\mathtt{I}_b)\}, \succ)$*. In the previous expansion,* $\mathtt{I}_e$ *and* $\mathtt{I}_b$ *that were not related, are now related in the resulting set. However, an expansion can introduce information that was implicit in the set. For instance, consider the following expansion:* $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ) \oplus^C (\mathtt{I}_p,\mathtt{I}_c) = (\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve} \cup \{(\mathtt{I}_p,\mathtt{I}_c)\}, \succ)$ *here the tuple* $(\mathtt{I}_p,\mathtt{I}_c)$ *is added to the G−set but the two agents were already related because* $(\mathtt{I}_p,\mathtt{I}_c) \in G\text{-}set^*_{Eve}$*.*

6.2 Contraction operator for credibility orders

We will define a contraction operator for credibility partial order among agents, called C-contraction, that will be denoted "$\ominus^C$", and will be adapted to the agent structure introduced in Definition 4. Consider the informants $\mathtt{I}_1$, $\mathtt{I}_2$ and $\mathtt{I}_3$. Suppose that, $(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)$ is an agent such that $(\mathtt{I}_2, \mathtt{I}_3) \in G\text{-}set^*_1$; the basic task of the $\ominus^C$ operator is to construct a new generator set $G\text{-}set'_1$ in which $(\mathtt{I}_2, \mathtt{I}_3) \notin G\text{-}set'^*_1$, losing as less information as possible. As we will show below, contraction does not mean simply to remove those credibility tuples containing $(\mathtt{I}_2, \mathtt{I}_3)$ from $G\text{-}set_1$. Every path from $\mathtt{I}_2$ to $\mathtt{I}_3$ has to be considered. Hence, the notions of *non-redundant path* and *paths set* are introduced.

Definition 18 (Non-redundant Path [14]) *Let* $\mathtt{I}_1, \mathtt{I}_2, \mathtt{I}_3, \mathtt{I}_4 \in \mathbb{I}$*, let* $(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)$ *be an agent and* $\mathbb{P} \subseteq G\text{-}set_1$*. We say that the set of credibility tuples* $\mathbb{P}$ *is a* path from $\mathtt{I}_2$ to $\mathtt{I}_4$ *in* $G\text{-}set_1$*, if* $(\mathtt{I}_2,\mathtt{I}_4) \in \mathbb{P}$*, or* $(\mathtt{I}_2,\mathtt{I}_3) \in \mathbb{P}$ *and there is a path from* $\mathtt{I}_3$ *to* $\mathtt{I}_4$ *in* $\mathbb{P}$*. We say that* $\mathbb{P}$ *is a* non-redundant path *from* $\mathtt{I}_2$ *to* $\mathtt{I}_4$ *if there is no other path* $\mathbb{P}'$ *from* $\mathtt{I}_2$ *to* $\mathtt{I}_4$ *in* $G\text{-}set_1$*, such that* $\mathbb{P}' \subsetneq \mathbb{P}$*.*

Definition 19 (Paths set [14]) *Let* $\mathtt{I}_1, \mathtt{I}_2, \mathtt{I}_2 \in \mathbb{I}$ *and let* $(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)$ *be an agent, we define the* paths set *from* $\mathtt{I}_2$ *to* $\mathtt{I}_3$ *in* $G\text{-}set_1$*, denoted* $G\text{-}set_1^{(\mathtt{I}_2\text{-}\mathtt{I}_3)}$*, as* $G\text{-}set_1^{(\mathtt{I}_2\text{-}\mathtt{I}_3)} = \{\mathbb{P} : \mathbb{P}$ *is a non-redundant path from* $\mathtt{I}_1$ *to* $\mathtt{I}_2$ *in* $G\text{-}set_1\}$*.*

Example 15 *Consider agent* $\mathtt{I}_e$ *of Example 3 where* $G\text{-}set_{Eve} = \{(\mathtt{I}_p,\mathtt{I}_b), (\mathtt{I}_p,\mathtt{I}_e), (\mathtt{I}_b, \mathtt{I}_c), (\mathtt{I}_e,\mathtt{I}_c)\}$ *(see Figure 2). Then, the paths set from* $\mathtt{I}_p$ *to* $\mathtt{I}_c$ *in* $G\text{-}set_{Eve}$ *is* $G\text{-}set_{Eve}^{(\mathtt{I}_p\text{-}\mathtt{I}_c)} = \{\mathbb{P}_1, \mathbb{P}_2\}$*, where:*

$$\mathbb{P}_1 = \{\ (\mathtt{I}_p,\mathtt{I}_e),\ (\mathtt{I}_e,\mathtt{I}_c)\}$$
$$\mathbb{P}_2 = \{\ (\mathtt{I}_p,\mathtt{I}_b),\ (\mathtt{I}_b,\mathtt{I}_c)\}$$

Remark 2 *Given* $G\text{-}set \subseteq \mathbb{I} \times \mathbb{I}$*, since* $G\text{-}set^*$ *is defined over the transitive closure of the credibility order among informants, then,* $(\mathtt{I}_1,\mathtt{I}_2) \in G\text{-}set^*$ *if and only if there exists at least one element in* $G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)}$*.*

It is clear from Remark 2 that given a generator set $G\text{-}set \subseteq \mathbb{I} \times \mathbb{I}$, for the contraction of $G\text{-}set$ by $(\mathtt{I}_1, \mathtt{I}_2)$ we need to eliminate at least one credibility tuple in every path of $G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)}$. In other words, we need to eliminate a set of credibility tuples from $G\text{-}set$ so that no path is left from $\mathtt{I}_1$ to $\mathtt{I}_2$ in $G\text{-}set$. In our approach this set to be eliminated will be required to be *minimal.* In [14], the contraction of a generator set by a credibility tuple $(\mathtt{I}_1, \mathtt{I}_2)$ uses a cut mechanism to decide which tuples are erased from each path from $\mathtt{I}_1$ to $\mathtt{I}_2$. Next, we give the definition of *cut function* defined there.

Definition 20 (Cut function) *Given a paths set* $G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)}$*, we say that* σ *is a* cut function *for* $G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)}$*, denoted* $\sigma(G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)})$*, if and only if:*

1. $\sigma(G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)}) \subseteq \bigcup(G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)})$*.*
2. *For each path* $\mathbb{P} \in G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)}$*,* $\mathbb{P} \cap \sigma(G\text{-}set^{(\mathtt{I}_1\text{-}\mathtt{I}_2)}) \neq \emptyset$*.*

Example 16 *Consider again agent* $\mathtt{I}_e$ *of Example 3 and the paths set* $G\text{-}set_{Eve}^{(\mathtt{I}_p\text{-}\mathtt{I}_c)} = \{\mathbb{P}_1, \mathbb{P}_2\}$ *obtained in Example 15, where* $\mathbb{P}_1 = \{(\mathtt{I}_p,\mathtt{I}_e), (\mathtt{I}_e,\mathtt{I}_c)\}$ *and* $\mathbb{P}_2 = \{(\mathtt{I}_p,\mathtt{I}_b), (\mathtt{I}_b,\mathtt{I}_c)\}$*. To obtain the set* $\sigma(G\text{-}set_{Eve}^{(\mathtt{I}_p\text{-}\mathtt{I}_c)})$*, at least one element of each path has to be selected. Then, to avoid the occurrence of* $(\mathtt{I}_p,\mathtt{I}_c)$*, the* cut function *selects the tuples that will be erased from each path. Note that in each path the* cut function *can select two tuples or either one of these, depending on its specification. Suppose that* $(\mathtt{I}_p,\mathtt{I}_e)$ *is selected by the* cut function *from* $\mathbb{P}_1$ *and* $(\mathtt{I}_b,\mathtt{I}_c)$ *is selected from* $\mathbb{P}_2$*. Therefore,* $\sigma(G\text{-}set_{Eve}^{(\mathtt{I}_p\text{-}\mathtt{I}_c)}) = \{(\mathtt{I}_p,\mathtt{I}_e), (\mathtt{I}_b,\mathtt{I}_c)\}$*.*

Next, a C-contraction operator will be defined.

Definition 21 (C-contraction) *Let $\mathtt{I}_1$, $\mathtt{I}_2$, $\mathtt{I}_3 \in \mathbb{I}$ be three agent identifiers, $(\mathtt{I}_2,\mathtt{I}_3)$ a credibility tuple, $(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)$ an agent, $G\text{-}set_1^{(\mathtt{I}_2\text{-}\mathtt{I}_3)}$ a paths set, and let σ be a cut function for $G\text{-}set_1^{(\mathtt{I}_2\text{-}\mathtt{I}_3)}$. The operator "$\ominus^C$", called C-contraction, is defined as follows:*

$$(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)\ominus^C(\mathtt{I}_2,\mathtt{I}_3) = (\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1 \setminus \sigma_\downarrow(G\text{-}set_1^{(\mathtt{I}_2\text{-}\mathtt{I}_3)}), \succ)$$

It is important to note that we have defined a family of contraction operators following other established formalisms of belief revision [1, 7]. The specification of the cut function will allow the introduction of different possibilities.

Example 17 *Consider agent $\mathtt{I}_e$ of Example 3 where $G\text{-}set_{Eve} = \{(\mathtt{I}_p,\mathtt{I}_b), (\mathtt{I}_p,\mathtt{I}_e), (\mathtt{I}_b, \mathtt{I}_c), (\mathtt{I}_e,\mathtt{I}_c)\}$. Then, suppose Eve ($\mathtt{I}_e$) wants to reflect that the client assigned ($\mathtt{I}_c$) to her is no longer more credible than her colleague Paul ($\mathtt{I}_p$). That is, $\mathtt{I}_e$ wants to contract by the credibility tuple ($\mathtt{I}_p$, $\mathtt{I}_c$) using "$\ominus^C$". As shown in Example 16, $\sigma(G\text{-}set_{Eve}^{(\mathtt{I}_p\text{-}\mathtt{I}_c)}) = \{(\mathtt{I}_p,\mathtt{I}_e), (\mathtt{I}_b,\mathtt{I}_c)\}$. Hence, $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)\ominus^C(\mathtt{I}_p,\mathtt{I}_c) = (\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve} \setminus \sigma_\downarrow(G\text{-}set_{Eve}^{(\mathtt{I}_p\text{-}\mathtt{I}_c)}), \succ) = (\mathtt{I}_e, \Delta^+{}_{Eve}, \{(\mathtt{I}_p,\mathtt{I}_b), (\mathtt{I}_e,\mathtt{I}_c)\}, \succ)$. Figure 5 shows (left) the graph for the original generator set $G\text{-}set_{Eve}$ and (right) the graph for the resulting generator set after the contraction by ($\mathtt{I}_p$, $\mathtt{I}_c$) where two arcs have been deleted.*

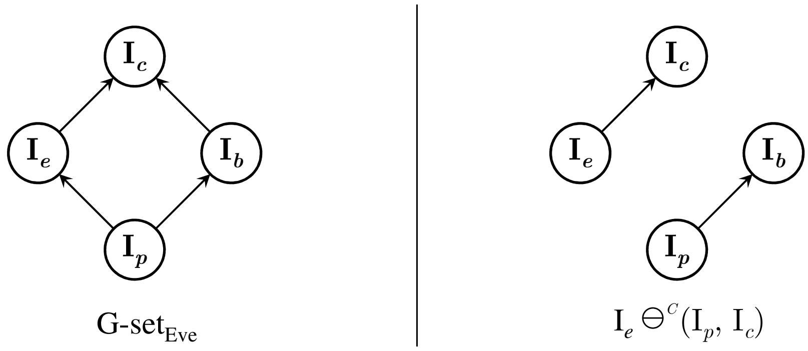

Figure 5: $\mathtt{I}_e$ contracted by ($\mathtt{I}_p$, $\mathtt{I}_c$).

Notice that here, in contrast to the case of expansion, the soundness preservation property of contraction is not conditioned. This is due to the way we define contraction. Since contraction is basically a process of elimination, it is impossible for this operation to introduce cycles if there were none with which to begin.

6.3 Prioritized revision operator for credibility orders

In this section, we will define a prioritized revision operator for credibility partial order among agents, called C-revision, denoted "$\circledast^C$", and adapted to the agent structure introduced in Definition 4. When an agent $(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)$ receives a new credibility tuple, the operator $\circledast^C$ will modify $G\text{-}set_1$ and will guarantee: soundness preservation and that the new information prevails. Consider agent $\mathtt{I}_1$ receives the information that $\mathtt{I}_2$ is less credible than $\mathtt{I}_3$. Since the new information has priority, the basic task of the C-revision operator is to construct a new generator set in which $(\mathtt{I}_2, \mathtt{I}_3) \in G\text{-}set_1^*$ but $(\mathtt{I}_3, \mathtt{I}_2) \notin G\text{-}set_1^*$. When an agent $(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)$ is revised by a credibility tuple $(\mathtt{I}_2, \mathtt{I}_3)$ there exist two tasks:

1. to maintain the soundness of $G\text{-}set_1$: if $(\mathtt{I}_3, \mathtt{I}_2) \in G\text{-}set_1^*$ (*i.e.*, $(\mathtt{I}_2, \mathtt{I}_3)$ generates a cycle in $G\text{-}set_1$), then is necessary to erase some credibility tuples from $G\text{-}set_1$ to avoid cycles.

2. to add $(\mathtt{I}_2, \mathtt{I}_3)$ to $G\text{-}set_1$: this is the most simple task to characterize from the logical point of view because it consists only in the addition of a new credibility tuple.

The first task can be accomplished contracting by $(\mathtt{I}_3, \mathtt{I}_2)$. The second task can be accomplished expanding by $(\mathtt{I}_2, \mathtt{I}_3)$. This composition is based on the *Levi identity* [6, 1], which proposes that a revision can be constructed out of two operations: a contraction and an expansion.

Definition 22 (Prioritized C-revision) *Let* $\mathtt{I}_1$, $\mathtt{I}_2$, $\mathtt{I}_3 \in \mathbb{I}$ *be three agent identifiers,* ($\mathtt{I}_2$,$\mathtt{I}_3$) *a credibility tuple,* $(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)$ *an agent,* $\ominus^C$ *the C-contraction operator and* $\oplus^C$ *the C-expansion operator. The operator "*$\circledast^C$*", called prioritized C-revision, is defined as follows:*

$$(\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)\circledast^C(\mathtt{I}_2,\mathtt{I}_3) = ((\mathtt{I}_1, \Delta^+{}_1, G\text{-}set_1, \succ)\ominus^C(\mathtt{I}_3,\mathtt{I}_2))\oplus^C(\mathtt{I}_2,\mathtt{I}_3)$$

Example 18 *Consider aging the agent Eve* $\mathtt{I}_e$*. Then, suppose* $\mathtt{I}_e$ *wants to reflect that her boss* ($\mathtt{I}_b$) *is more credible than the client assigned* ($\mathtt{I}_c$) *to her. That is,* $\mathtt{I}_e$ *wants to revise by* ($\mathtt{I}_c$, $\mathtt{I}_b$) *using "*$\circledast^C$*". Since* ($\mathtt{I}_b$,$\mathtt{I}_c$) $\in$ $G\text{-}set_{Eve}$ *then it is first necessary to contract* $(\mathtt{I}_e, \Delta^+{}_{Eve}, G\text{-}set_{Eve}, \succ)$ *by* ($\mathtt{I}_b$,$\mathtt{I}_c$) *and then it is necessary to expand by* ($\mathtt{I}_c$, $\mathtt{I}_b$)*. Figure 6 shows (left) the graph for the original generator set* $G\text{-}set_{Eve}$ *and (right) the graph for the resulting generator set after the revision by* ($\mathtt{I}_c$, $\mathtt{I}_b$) *where one arc have been deleted, and an arc have been added.*

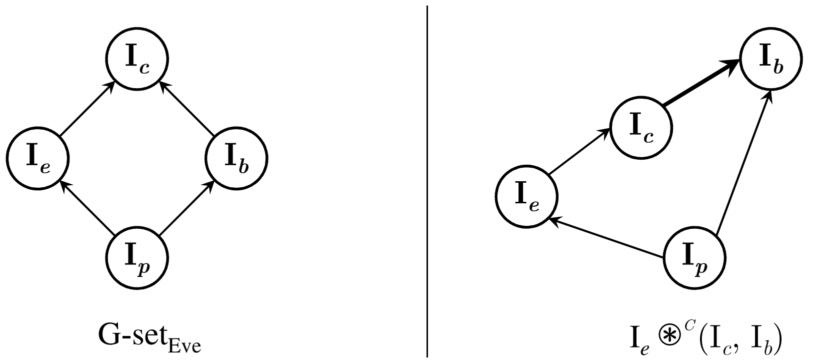

Figure 6: $\mathtt{I}_e$ revised by ($\mathtt{I}_c$, $\mathtt{I}_b$).

Note that here, when the the credibility order considered by Eve is changed, the preferences among its argument also changes. In the following table we depict the preferences among the arguments of Eve in this new situation considering the comparison criteria which were presented in Section 4.

Arguments	$\succ_S$	$\succ_L$	$\succ_M$
$\langle\mathcal{A}_1, w\rangle, \langle\mathcal{A}_7, {\sim}w\rangle$	$\langle\mathcal{A}_1, w\rangle\succ_S\langle\mathcal{A}_7, {\sim}w\rangle$	$\langle\mathcal{A}_1, w\rangle\succ_L\langle\mathcal{A}_7, {\sim}w\rangle$	$\langle\mathcal{A}_1, w\rangle\succ_L\langle\mathcal{A}_7, {\sim}w\rangle$
$\langle\mathcal{A}_2, {\sim}t\rangle, \langle\mathcal{A}_5, t\rangle$	$\langle\mathcal{A}_2, {\sim}t\rangle\succ_S\langle\mathcal{A}_5, t\rangle$	$\langle\mathcal{A}_2, {\sim}t\rangle\succ_L\langle\mathcal{A}_5, t\rangle$	$\langle\mathcal{A}_2, {\sim}t\rangle\succ_M\langle\mathcal{A}_5, t\rangle$
$\langle\mathcal{A}_2, {\sim}t\rangle, \langle\mathcal{A}_8, t\rangle$	$\langle\mathcal{A}_8, t\rangle\succ_S\langle\mathcal{A}_2, {\sim}t\rangle$	$\langle\mathcal{A}_2, {\sim}t\rangle\succ_L\langle\mathcal{A}_8, t\rangle$	$\langle\mathcal{A}_2, {\sim}t\rangle\succ_M\langle\mathcal{A}_8, t\rangle$
$\langle\mathcal{A}_3, {\sim}r\rangle, \langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_3, {\sim}r\rangle\succ_S\langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_3, {\sim}r\rangle\succ_L\langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_3, {\sim}r\rangle\succ_M\langle\mathcal{A}_4, r\rangle$
$\langle\mathcal{A}_4, r\rangle, \langle\mathcal{A}_9, {\sim}r\rangle$	$\langle\mathcal{A}_9, {\sim}r\rangle\succ_S\langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_9, {\sim}r\rangle\succ_L\langle\mathcal{A}_4, r\rangle$	$\langle\mathcal{A}_9, {\sim}r\rangle\succ_M\langle\mathcal{A}_4, r\rangle$

Also observe that this change will affect Eve's warranted literals. In this scenario $\mathtt{I}_e$ will have "${\sim}t$" as warranted under the three possible criterions whereas in Example 13 we have shown that under $\succ_S$ $\mathtt{I}_e$ had "t" warranted and neither "t" nor "${\sim}t$" under the other two criterions.

7 Conclusions and related work

We have defined an argumentative reasoning formalism where the notion of trust plays a central role. Each individual agent considers tentative conclusions obtained from its knowledge base by building arguments; often, arguments have contradictory conclusions and/or contradictory elements in their structure. Arguments conflicting with the support for tentative conclusions act as counterarguments and the need for an argument comparison criterion becomes clear in this context. In the formalism, every source of information is attributed certain credibility leading to a perception of trustworthiness. Trust in turn leads to a preference criterion over the set of arguments helping to decide which conclusions deserve to be warranted. Thus, the inference mechanism compares arguments and counter-arguments in the course to decide which argument should prevail. An agent $(I, \Delta^+, G\text{-}set, \succ)$ can use its own information and also the information previously received stored in Δ^+ to warrant its derived beliefs; a belief B is warranted if there exists a non-defeated argument for B. The trust-based argument comparison criterion $\succ$ was used for deciding when an argument defeats another; hence, warranted beliefs depend on the credibility order *G-set*. We have also introduced different argument comparison criteria based on trust that could be used accordingly with the representational needs of different domains.

Finally, the trust assigned to informants is inherently changeable and most realistic scenarios should consider this possibility; for that reason, we have introduced the change operators for *expansion*, *contraction* and *revision* following the tradition in Belief Change Theory. An agent I can use these operators to act on the credibility relation established on the set agent *G-set*; this brings the ability of dynamically changing the credibility assigned to its informants. Therefore, as the warranting process relies on the credibility of informants, a change in *G-set* opens the possibility that previously warranted beliefs become unsupported, and that new beliefs turn into accepted conclusions.

Recently, some approaches combining argumentation and trust have been developed [11, 16, 15, 8]; in general, these formalisms are focused on using argumentation to reason about trust. Here we have follow a different route that in a sense is complementary, our proposal uses the trust relation defined in the set of informants as part of the decision process which leads to obtain the prevailing information in the context of conflictive knowledge. In [8] an argumentative formalism for reasoning about the appropriateness of an argument's proponent is given; that is, the paper explores how arguments can be evaluated in terms of the proponents trustiness in relation to that particular argument. Hunter's work is related to our approach in that he also attributes importance to the source of the information; however, in contrast with ours, his proposal includes augmenting the representation and reasoning infrastructure with a meta-level system for reasoning about the object-level arguments and their proponents. The meta-level system incorporates axioms for raising the object-level argumentation to the meta-level (in particular to capture when one argument is a counterargument to another argument), and meta-level axioms that specify when proponents are appropriate for arguments.

Our work is also related with [5]. Similar to that work we use defeasible rules and presumptions for knowledge representation, and we use DeLP marked dialectical trees for computing warranted belief. However, in contrast to that approach, we propose some new argument comparison criteria that differ from those reported in the DeLP literature. Here, we define three new trust-based argument comparison criteria and we compare them. Thus, the more appropriate criterion to the application domain can be selected.

References

[1] C. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[2] K. S. Barber and J. Kim. Belief revision process based on trust: Simulation experiments. In *In Proceedings of Autonomous Agents '01 Workshop on Deception, Fraud, and Trust in Agent Societies*, pages 1–12, 2001.

[3] DeLP-home-page. http://lidia.cs.uns.edu.ar/delp. 2007.

[4] D. Gambetta. Can we trust them? In *Trust: Making and beraking cooperative relations*, pages 213–238, 1990.

[5] A. J. García and G. R. Simari. Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming*, 4(1-2):95–138, 2004.

[6] P. Gärdenfors. An epistemic approach to conditionals. *American Philosophical Quarterly*, 18(3):203–211, 1981.

[7] S. O. Hansson. Kernel contraction. *Journal of Symbolic Logic*, 59(3):845–859, 1994.

[8] A. Hunter. Reasoning about the appropriateness of proponents for arguments. In D. Fox and C. P. Gomes, editors, *Proc. of the 23rd AAAI Conference on Artificial Intelligence, AAAI*, pages 89–94. AAAI Press, 2008.

[9] Vladimir Lifschitz. Foundations of logic programs. In G. Brewka, editor, *Principles of Knowledge Representation*, pages 69–128. CSLI Pub., 1996.

[10] D. H. Mcknight and N. L. Chervany. The meanings of trust. Technical report, Carlson School of Management, University of Minnesota, 1996.

[11] S. Parsons, P. McBurney, and E. Sklar. Reasoning about trust using argumentation: A position paper. In *ArgMAS*, pages 159–170, 2010.

[12] J. Sabater and C. Sierra. Review on computational trust and reputation models. *Artificial Intelligence Review*, 24(1):33–60, 2005.

[13] G. R. Simari and R. P. Loui. *A Mathematical Treatment of Defeasible Reasoning and its Implementation. Artificial Intelligence*, 53:125–157, 1992.

[14] L. H. Tamargo, M. A. Falappa, A. J. García, and G. R. Simari. A change model for credibility partial order. In *Proceeding of the 5th International Conference on Scalable Uncertainty Management (SUM)*, pages 317–330, 2011.

[15] S. Villata, G. Boella, D. M. Gabbay, and L. Van der Torre. Arguing about trust in multiagent systems. In *Proc. of the 11th Symposium on Artificial Intelligence of the Italian Association for Artificial Intelligence (AIIA'10)*, pages 236–243, 2010.

[16] S. Villata, G. Boella, D. M. Gabbay, and L. Van der Torre. Arguing about the trustworthiness of the information sources. In *ECSQARU*, pages 74–85, 2011.