Using Preferences over Sources of Information in Argumentation-Based Reasoning

Alison R. Panisson, Victor S. Melo, and Rafael H. Bordini
Postgraduate Programme in Computer Science — School of Informatics (FACIN)
Pontifical Catholic University of Rio Grande do Sul (PUCRS) – Porto Alegre – Brazil
Emails: {alison.panisson,victor.melo.001}@acad.pucrs.br, rafael.bordini@pucrs.br

***Abstract*—Argumentation-based reasoning plays an important role in agent reasoning and communication, yet little research has been carried out on the issues in integrating argumentation techniques into practical multi-agent platforms and the various sources of information in such systems. In this work, we extend an argumentation-based reasoning mechanism to take into account preferences over arguments supporting contrary conclusions, which in practice means the agent will be able to act more informedly, being able to decide on beliefs about which it would be otherwise ambivalent. Such preferences come from elements that are present or can be more easily obtained in the context of practical multi-agent programming platforms, such as multiple sources from which the information (used to construct the arguments) was acquired, as well as varying degrees of trust on them. Further, we introduce different agent profiles by varying the way certain operators are applied over the various information sources leading to the preferences over competing arguments in our approach. Unlike previous approaches, our approach accounts for multiple sources for a single piece of information and is based on an argumentation-based reasoning mechanism implemented on a multi-agent platform so arguably more computationally grounded than those approaches.**

I. Introduction

Argumentation in multi-agent systems has received much research attention over the years, perhaps because it provides the exchange of additional information in dialogues for negotiation, deliberation, and many other important aspects of multi-agent systems [1], [2], [3]. This exchange of information allows agents to communicate and understand each other in a more informed way. It also can change the mental attitudes of the agents who receive such information. This is an important aspect of argumentation-based approaches to multi-agent systems, because of the inherent uncertainty and lack of information in these systems [4].

Argumentation-based reasoning is an important topic of research [5], and recent work has brought this to the context of agent-oriented programming languages [6], [7], [8]. In that context, agents can reason about arguments in order to make decisions and communicate. Further, argumentation-based reasoning allows agents to construct arguments in the face of uncertainty (i.e., incomplete and incorrect information). Therefore, it is important that the agents construct arguments using the most precise pieces of information available to them, based on the most trustworthy sources, avoiding as much as possible sources of doubt in the arguments used, hence improving their decisions and therefore their actions.

With these issues in mind, we propose an approach to combine argumentation-based reasoning and preferences over sources of information[1]. For example using information about trust on the agents who provided some information used in an argument. This allows agents to make decisions in situations where they would not have been able, for example because of unresolved conflicts in argumentation-based reasoning mechanisms without such preferences. This is interesting, as not resolving conflicts between arguments — in particular using defeasible arguments for contrary conclusions — can be unsatisfactory in general, specially in multi-agent systems where efficient ways of solving conflicts are typically required [9].

Differently from previous approaches, we here consider that an agent might have various different sources for the same piece of information. This is in fact often the case in multi-agent systems developed on agent-programming platforms, such as Jason [10], where beliefs are annotated with all known sources of that information. Furthermore, elaborate trust systems have been studied [11] in the context of multi-agent systems which could provide reliable trust information about each such source.

The main contributions of this paper are: (i) the argumentation-based reasoning mechanism that uses trust to support decision when there are arguments supporting contrary conclusions, (ii) accounting for multiple sources for the same information with varying degrees of trust in each, (iii) agent profiles providing different attitudes towards such multiple sources, and (iv) the implementation of such reasoning in the context of an agent programming platform.

II. Argumentation-Based Reasoning

In this work, we chose to use the argumentation-based reasoning mechanism reported in [7] (which was extended and formalised in [8]), one of the few practical approaches that implement argumentation-based reasoning in agent-oriented programming languages; in particular, that reasoning mechanism is implemented in Jason [10], a well-know multi-agent platform for the development of multi-agent systems. In [7], arguments are constructed using strict and defeasible inferences rules. Intuitively, arguments that use only *strict* rules are stronger than arguments that use defeasible rules. That is, there are two types of arguments: (i) **strict arguments** are formed

[1] A preliminary discussion about this appeared in [24], [21].

only of facts and strict rules (i.e., indisputable knowledge). It is assumed that the strict part of any knowledge base is consistent (i.e., contradictions cannot be derived); and (ii) **defeasible arguments** are created using at least one defeasible rule (indicating the points of weakness of the argument).

In order to define the acceptability of an argument we need to consider conflicting arguments. In [7], conflict between arguments are of two types[2]:

Definition 1 (Attack Between Arguments): Let $\langle S_1, c_1\rangle$ and $\langle S_2, c_2\rangle$ be two arguments. Attacks between arguments can be generalised into two types:

- The argument $\langle S_1, c_1\rangle$ rebuts the argument $\langle S_2, c_2\rangle$ if $c_1 \equiv \overline{c_2}$.
- The argument $\langle S_1, c_1\rangle$ undercuts the argument $\langle S_2, c_2\rangle$ if $c_1 \equiv \overline{c_3}$ for some $\langle S_3, c_3\rangle$ where $S_3 \subseteq S_2$.

When two arguments are in conflict, i.e., the arguments attack each other, this does not necessarily mean that an argument defeats the other argument. Defeat is a "successful" attack, and it considers the set of arguments that defend each other, including preferences between the conflicting arguments [12]. In [7], the set of acceptable arguments from an agent's knowledge base is defined in terms of the *defeasible semantics* introduced in [13]. The defeasible semantics is similar to the *grounded semantics* from Dung's work [14] and it is based on the so-called *preempting defeaters* [15]. The preempting defeaters of [15] are called ambiguity blocking (in regards to the argumentation system) in [13]. This means that defeasible rules that are rebutted by a superior rule cannot be used to rebut other rules. An example of *preempting defeaters* is the knowledge base represented by Δ below, where we use $\Rightarrow$ to refer to defeasible inferences:

$$\Delta = \left\{ \begin{array}{ccc} a & a \Rightarrow b & x \Rightarrow e \\ x & b \Rightarrow c & e \Rightarrow \neg c \\ y & c \Rightarrow d & y \Rightarrow \neg e \end{array} \right\}$$

where, in this example, we may conclude d using the inferences $\{a,\ a \Rightarrow b,\ b \Rightarrow c,\ c \Rightarrow d\}$, although there is a derivation $\{x,\ x \Rightarrow e,\ e \Rightarrow \neg c\}$ which rebuts the sub-argument for d that concludes c (undercutting the first argument); this argument (the argument that derives $\neg c$) is defeated (by undercut) by $\{y,\ y \Rightarrow \neg e\}$ which prevents the use of that argument to undercut the inference of d. Although in the example above we have an acceptable argument for d, the arguments $\{y,\ y \Rightarrow \neg e\}$ and $\{x,\ x \Rightarrow e\}$ (which is a sub-argument for c in the example) are in conflict, and the approach presented in [7][3] is not able to decide which one is acceptable, i.e., both are treated as unacceptable. A way to deal with unresolved conflicts is to use preferences over the arguments.

Clearly, strict arguments are stronger than defeasible arguments and they have priority, i.e., when arguments are involved in conflict, strict arguments always defeat defeasible ones. Considering only defeasible arguments, the work in [7] considers two types of priority: (i) priority by specificity, which is originally defined in defeasible logic [16], and (ii) the explicit declaration of priority between defeasible rules, using a special predicate. In priority by specificity, more specific conclusions have priority over more general ones. To exemplify this idea, consider the well-known Tweety example:

```
def_rule(flies(X),bird(X)).
def_rule(¬flies(X),penguin(X)).
def_rule(bird(X),penguin(X)).
penguin(tweety).
```

All clauses in the example are defeasible rules (written using the representation of defeasible rules in Jason platform [10] as in [7]). Considering the knowledge above, we have two conflicting arguments, one supporting that Tweety flies: "Tweety flies, because it is a penguin, penguins are birds, and birds fly", and one supporting that Tweety does not fly: "Tweety does not fly, because it is a penguin and penguins do not fly". The mechanism implemented in [7] (as well as the defeasible-Prolog [15]) concludes, in this case, that Tweety does not fly, because the rule for penguins is more specific than a rule for birds, given that penguin is a subclass of birds. In this manner, the argument for Tweety not flying has priority over the other, and so defeats it. Considering the explicit declaration of priority, [7] allows to declare that a `Rule1` has priority over `Rule2` (using the predicate `sup(Rule1, Rule2)`). Therefore, when two arguments are constructed using these rules, and they are in conflict, this declaration is used in order to decide which conclusion will actually be derived.

Although the approach presented by [7] has ways to deal with conflicting information (conflicting arguments) when the conflict cannot be resolved considering the set of arguments, this characteristic is limited. This limitation can be substantially circumvented when we consider preferences over the arguments, generating fewer unresolved conflicts. Such preferences may come from information typically available in multi-agent systems, such as trust values for the information perceived and received from other agents. Following this idea, we propose to extend the preference relations described above to consider the source of information in order to decide the most reasonable conclusions to reach in the argumentation-based reasoning. The extensions are built on top of the work in [7] and implemented in the Jason platform [10].

III. Trust in Multi-agent Systems

In trust-based approaches, agents can use the level of trust associated with the sources of contradictory information in order to decide about which one to believe. There are many different approaches to trust in the literature [17], [18], [19], [11], [20], but here we will build our definitions mostly based on [19], [20]. In this section, we describe trust as a relation between agents, while in Section IV we expand it, associating trust values for beliefs, which represent how much an agent trusts on some information. Afterwards, in Section V, we apply the trust values of beliefs in order to calculate the overall trust on arguments.

[2]We use "$\neg$" for strong negation and a general operator for contradictory information, where $\overline{\varphi} \equiv \neg\varphi$ and $\overline{\neg\varphi} \equiv \varphi$.

[3]This characteristic is from the original implementation of defeasible Prolog [15], and it is what gave rise for the name *ambiguity blocking* in [13].

Considering trust as a relation between agents and following [20], a *trust relation* can be formalised as: $\tau \subseteq Ags \times Ags$, where the existence of the relation indicates that an agent assigns some level of trust to another. For example, $\tau(Ag_i, Ag_j)$ means that agent Ag_i has at least some trust on Ag_j. It is important to realise that this is not a symmetric relation, so if $\tau(Ag_i, Ag_j)$ holds, this does not imply that $\tau(Ag_j, Ag_i)$ holds.

A *trust network* is a directed graph representing trust relations. It can be defined as: $\Gamma = \langle Ags, \tau \rangle$, where Ags is the set of nodes in the graph, representing the agents of the trust network, and τ is the set of edges, where each edge is a pairwise trust relation between agents of Ags. An example of a trust network can be seen in Figure 1.

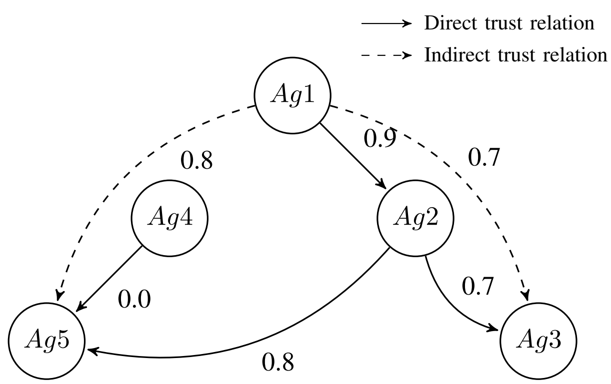

Fig. 1. Trust Network Example

In order to measure trust, we follow the definition given in [19], [20] where a function $tr : Ags \times Ags \mapsto \mathbb{R}$ is used. It returns a value between 0 and 1, representing how much an agent trusts another. However, differently from [19], [20], we define the relation between tr and τ as:

$$\begin{array}{lcl} tr(Ag_i, Ag_j) \geq 0 & \Leftrightarrow & (Ag_i, Ag_j) \in \tau \\ tr(Ag_i, Ag_j) = \texttt{null} & \Leftrightarrow & (Ag_i, Ag_j) \notin \tau \end{array}$$

where, a trust level can in fact be zero, represented by $tr(Ag_i, Ag_j) = 0$, which means that Ag_i does not trust Ag_j. This is different from cases where Ag_i has no trust value assigned to Ag_j, represented by $tr(Ag_i, Ag_j) = \texttt{null}$. Both cases can be seen in Figure 1, where there we have $tr(Ag4, Ag5) = 0$ and $tr(Ag1, Ag4) = \texttt{null}$.

As in [19], we assume trust as a transitive relation, so an agent Ag_i can trust Ag_j directly or indirectly. Direct trust occurs when agent Ag_i directly assigns a trust value to Ag_j. Indirect trust occurs when, continuing the previous example, Ag_j trusts another agent Ag_k: in this case we could say that Ag_i indirectly trusts Ag_k. We say there is a *path* between agents Ag_i and Ag_j if it is possible to create sequence of nodes $\langle Ag_0, Ag_1, Ag_2, \ldots, Ag_{n-1}, Ag_n \rangle$ so that $\tau(Ag_0, Ag_1), \tau(Ag_1, Ag_2), \ldots, \tau(Ag_{n-1}, Ag_n)$, with $Ag_0 = Ag_i$ and $Ag_n = Ag_j$. In order to measure the trust from one particular path from Ag_i to Ag_j we need to use an operator to consider all the direct trust values in that path. Following the idea proposed in [19], if have a general operator $\otimes^{tr}$ with $tr(Ag_i, Ag_j) = tr(Ag_0, Ag_1) \otimes^{tr} ... \otimes^{tr} tr(Ag_{n-1}, Ag_n)$ as the trust value that Ag_i has on Ag_j according to the path $A_0, \ldots, A_n$ from Ag_i to Ag_j. If it happens that there are m different paths between Ag_i and Ag_j, a first possible path having a trust value of $tr(Ag_i, Ag_j)^1$ and the mth having $tr(Ag_i, Ag_j)^m$, following [19] we can define a generic operator $\oplus^{tr}$, where $tr(Ag_i, Ag_j) = tr(Ag_i, Ag_j)^1 \oplus^{tr} \ldots \oplus^{tr} tr(Ag_i, Ag_j)^m$. For simplicity, in this paper we use those generic operators instantiated as: (i) the trust of a path operator $\otimes^{tr}$ is the minimum trust value along the path, i.e., $tr(Ag_i, Ag_j) = min\{tr(Ag_0, Ag_1), \ldots, tr(Ag_{n-1}, Ag_n)\}$ given a path $Ag_0, \ldots, Ag_n$ from Ag_i to Ag_j; (ii) the operator $\oplus^{tr}$ over trust paths is defined as $tr(Ag_i, Ag_j) = max\{tr(Ag_i, Ag_j)^1, \ldots, tr(Ag_i, Ag_j)^m\}$, where m is the number of different possible paths between Ag_i and Ag_j.

IV. Trust on Beliefs

In this section, we introduce trust applied to beliefs, which is based on the trust value applied to the sources of these beliefs. We consider not only other agents as sources of information, but also perception of the environment, artifacts, and "mental notes" (beliefs created by an agent itself). For trust values for information received from other agents, we assume that these values are explicitly asserted in the belief base of agents (but calculated dynamically) based on the approach presented in the previous section. For trust values of information perceived from the environment, these values depend on the application domain, where, for example, multiple sensors could have varying degrees of trustworthiness. For the purpose of a running example, we use the following trust values: $tr(ag1, 0.3)$, $tr(ag2, 0.4)$, $tr(ag3, 0.5)$, $tr(ag4, 0.8)$, $tr(self, 1.0)$, $tr(percept_1, 0.9)$, and $tr(percept_2, 0.6)$.

Therefore, we expand trust to be a relation between an agent and the possible sources of information. So function $tr(Ag_i, Ag_j)$ that returns the trust level of Ag_i on Ag_j is generalised to $tr(Ag_i, s_j)$, where s_j represents one of the sources of information for agent Ag_i. This way, an agent Ag_i has a trust level on other kinds of sources, percepts or mental notes. Therefore, as a belief φ of an agent Ag_i can come from multiple sources, in order to know how much Ag_i trusts φ, we must consider the *tr* value associated with each source of φ for Ag_i. For this, we introduce the function $trb_i : \varphi \rightarrow \mathbb{R}$, where $trb_i(\varphi)$ returns the trust value that Ag_i has on belief φ based on the trust level Ag_i has on the sources that asserted information φ. The operation that calculates $trb_i(\varphi)$ varies according to agent profiles, corresponding to different attitudes towards one's sources of information. We introduce two agent profiles for calculating trust values over beliefs. They both may be interesting in different domains, depending on whether we are interested in credulous or sceptical agents.

Definition 2 (Credulous Agent): A credulous agent considers only the most trustworthy source of information, and does not look for an overall social value.

The formula used by a credulous agent to consider the most trusted source is $trb_i(\varphi) = max\{tr(Ag_i, s_1), ..., tr(Ag_i, s_n)\}$, where $\{s_1, ..., s_n\}$ is the set of sources that informed φ to Ag_i.

Definition 3 (Sceptical Agent): A sceptical agent considers the number of sources from which it has received the information, and the trust value of each such source, in order to have some form of social trust value.

A sceptical agent considers the quantity of sources that the information φ comes from. Therefore, we use a formula that sums the trust values of sources that information φ has been received from by Ag_i, determining a social trust value as follows:

$$trb_i(\varphi) = \frac{\sum_{s \in S^+_\varphi} tr(Ag_i, s)}{|S^+_\varphi| + |S^-_\varphi|}$$

where $S^+_\varphi = \{s_1, ..., s_n\}$ is the set of n different sources of φ and S^-_φ is the set of sources for $\overline{\varphi}$.

For example, considering an agent Ag_i with the trust values we assumed, if Ag_i receives an information φ from a set of sources $S^+_\varphi = \{Ag1, Ag2, Ag3\}$ and receives $\overline{\varphi}$ from $S^-_\varphi = \{Ag4\}$, then: (i) A credulous agent will consider only the maximum trust values in S^+_φ and S^-_φ, so it will have $trb_i(\varphi) = 0.5$ and $trb_i(\overline{\varphi}) = 0.8$. (ii) A sceptical agent will consider all the various sources. In particular, it will have $trb_i(\varphi) = \frac{0.3+0.4+0.5}{4} = 0.3$ and $trb_i(\overline{\varphi}) = \frac{0.8}{4} = 0.2$.

As another example, when Ag_i receives an information φ where the sources of φ are $S^+_\varphi = \{percept_1\}$ and receives $\overline{\varphi}$ with sources $S^-_\varphi = \{Ag2, Ag3, Ag4\}$, then: (i) A credulous agent will have $trb_i(\varphi) = 0.9$, and $trb_i(\overline{\varphi}) = 0.8$, having greater trust in φ than in its negation. (ii) A sceptical agent however will have $trb_i(\varphi) = \frac{0.9}{4} = 0.225$ and $trb_i(\overline{\varphi}) = \frac{0.4+0.5+0.8}{4} = 0.425$, preferring to believe $\overline{\varphi}$ instead.

V. Trust on Arguments

In this section, we describe how trust and the argumentation-based reasoning mechanism presented in Section II are combined, focusing on the unresolved conflicting arguments. As described before, the reasoning mechanism introduced in [7] has some mechanisms to support decision on conflicting arguments, but some situations remain unresolved.

In order to have more acceptable arguments available to agents, allowing them to make decisions over these still unresolved conflicts, we propose to combine the argumentation-based reasoning mechanism from [7] and the trust value for beliefs introduced in the previous section. For this purpose, we introduce an approach to the calculation of trust for arguments, which allows us to consider the trust value of arguments in order to decide on those conflicting arguments by comparing such trust values. The approach presented here is applicable to both premises and inference rules as used in [7]. This is possible because the inference rules are represented using special predicates in the format of AgentSpeak beliefs.

The trust value on an argument depends on the values of each element in its support (in our case, premises and inference rules, both stored as beliefs), as defined below.

Definition 4 (Trust on arguments): The trust on an argument is based on the trust value of its support. Let $\langle S, c\rangle$ be an argument, its trust value is given by the trust of its support S, as follows: $tra(\langle S, c\rangle) = trb(\varphi_1) \otimes^{tra} \ldots \otimes^{tra} trb(\varphi_n)$, with $S = \{\varphi_1, ..., \varphi_n\}$ the support of the argument.

Considering again the profiles introduced in Section IV, the generic operator $\otimes^{tra}$ can be defined as: (i) **credulous agents** use $\otimes^{tra}$ as the maximum trust value, i.e., taking the highest trust value present in the argument's support set of the argument as the trust value for the argument as a whole: $tra(\langle S, c\rangle) = \max\{trb(\varphi_1), \ldots, trb(\varphi_n)\}$; and (ii) **sceptical agents** use the minimum value for $\otimes^{tra}$, considering the lowest trust value present in the argument's support set as the trust value for the argument: $tra(\langle S, c\rangle) = \min\{trb(\varphi_1), \ldots, trb(\varphi_n)\}$.

When agent Ag_i has multiple arguments for the same conclusion c, for example, the argument $\langle S_1, c\rangle$ and $\langle S_2, c\rangle$, the agent can opt for the argument that has the highest trust value: $argument(\langle S, c\rangle) = \max\{tra(\langle S_1, c\rangle), \ldots, tra(\langle S_n, c\rangle)\}$. Therefore, when we have an unresolved conflict between two arguments, we can solve the conflict by looking at the trust values, as follows.

Definition 5 (Rebutting Defeat using Trust): Let $\langle S_1, c_1\rangle$ and $\langle S_2, c_2\rangle$ be two conflicting arguments, with $c_1 \equiv \overline{c_2}$. We say that $\langle S_1, c_1\rangle$ *rebuts* $\langle S_2, c_2\rangle$ iff $tra(\langle S_1, c_1\rangle) > tra(\langle S_2, c_2\rangle)$.

Definition 6 (Undercutting Defeat using Trust): Let $\langle S_1, c_1\rangle$ and $\langle S_2, c_2\rangle$ be two conflicting arguments, with $c_1 \equiv \overline{c_2}$. We say that $\langle S_1, c_1\rangle$ *undercuts* $\langle S_3, c_3\rangle$ iff $tra(\langle S_1, c_1\rangle) > tra(\langle S_2, c_2\rangle)$ with $S_2 \in S_3$.

Although we introduced two simple agent profiles above, clearly other profiles and instantiations for the generic operators could be used, as suggested in [19], [20], [21].

VI. A Stock Market Scenario

In this paper, we use a stock market scenario to exemplify our approach of combining argumentation-based reasoning and trust on the sources of information. In our scenario, the agents are responsible for buying and selling stocks, looking for the most lucrative transactions. In this scenario, similar the real life, an agent that invests in stocks has a series of advisers who can be consulted about whether to invest or not in specific stocks given the market situation, or asking suggestions about which stocks to invest on, based on how much it can spend. We have an agent called *stockholder* (sh for short) and a series of *advisers*, we use adv_1, adv_2, adv_3 and adv_4 to name them. The agents communicate by message passing, whereby they exchange arguments.

Initially, the stockholder has \$20.000 to invest in some stocks, but it does not know in which market to invest. Therefore, the stockholder requests suggestions from the advisers about the markets in which it should invest. Assuming agent *sh* will choose only one stock to invest, it will choose the most trustworthy suggestion. However, in order to make a more confident choice, the agent can ask all advisers about the specific choice resulting from the first interaction (used to assess the trust). For example, assume the trust values $tr(adv_1, 0.3)$, $tr(adv_2, 0.4)$, $tr(adv_3, 0.5)$, and $tr(adv_4, 0.8)$ as the ones calculated in past interaction. Assume further that, in the first interaction, the advisers suggested the soybean and orange juice markets, in particular adv_1, adv_2 and adv_3

$$\Delta^{sh}_{scep} = \left\{ \begin{array}{l} \texttt{def\_rule(invest(soybean),[dollar(high),production(low)])[source(adv\_1),trustValue(0.3)].} \\ \texttt{def\_rule(}\neg\texttt{invest(soybean),[production(high),}\neg\texttt{buying(china,soybean)])[source(adv\_2),trustValue(0.4)].} \\ \texttt{def\_rule(}\neg\texttt{invest(soybean),[}\neg\texttt{buying(china,soybean),crisis(china)])[source(adv\_3),trustValue(0.5)].} \\ \texttt{def\_rule(invest(soybean),[buying(china,soybean),dollar(high)])[source(adv\_4),trustValue(0.8)].} \\ \texttt{buying(china,soybean)[source(adv\_4),trustValue(0.27)].} \\ \neg\texttt{buying(china,soybean)[source(adv\_2),source(adv\_3),trustValue(0.3)].} \\ \texttt{crisis(china)[source(adv\_3),trustValue(0.5)].} \\ \texttt{production(high)[source(adv\_2),trustValue(0.2)].} \\ \texttt{production(low)[source(adv\_1),trustValue(0.15)].} \\ \texttt{dollar(high)[source(adv\_1),source(adv\_4),trustValue(0.55)].} \\ \texttt{comp(production(high),production(low))[source(self)].} \end{array} \right\}$$

suggested the orange juice market and adv_4 suggested the soybean market. A sceptical agent would select the orange juice market, and a credulous agent would choose the soybean market instead.

In this first interaction, we will assume that the stockholder is credulous, so it concludes that to invest in soybean could be better. The stockholder agent then queries all of the advisers in order to get their opinions about that planned investment, receiving positive and negative suggestion by the arguments:

- adv_1: argument arg_1 $\langle S, \texttt{invest(soybean)}\rangle$, with S = [def_rule(invest(soybean), [dollar(high), production(low)]), dollar(high), production(low)].
- adv_2: argument arg_2 $\langle S, \neg\texttt{invest(soybean)}\rangle$, with S = [def_rule(¬invest(soybean), [production(high),¬buying(china, soybean)], production(high), ¬buying(china, soybean))].
- adv_3: argument arg_3 $\langle S, \neg\texttt{invest(soybean)}\rangle$, with S = [def_rule(¬invest(soybean), [¬buying(china, soybean), crisis(china)]), ¬buying(china, soybean), crisis(china)].
- adv_4: argument arg_4 $\langle S, \texttt{invest(soybean)}\rangle$, with S = [def_rule(invest(soybean), [buying(china, soybean), dollar(high)]), buying(china, soybean), dollar(high)].

After receiving the arguments the agent has all that information, with respective trust value (calculated by *trb*), in its knowledge base. The knowledge base of sh in represented by Δ^{sh}_{scep}. In Δ^{sh}_{scep}, we are considering that sh is a sceptical agent. Thus, agent sh now has arguments for both `invest(soybean)` and `¬invest(soybean)`. Using the reasoning mechanism in [7], both (i) `buying(china,soybean)` and `¬buying(china,soybean)`, and (ii) `production(low)` and `production(high)` are in contradiction with each other, and because they are both blocked to support to the competing arguments, even with undercutting the conflict cannot be resolved. However, with our approach presented in Section V, our extended reasoning mechanism is capable of deciding which argument is stronger, using tra. Therefore, we have the following trust values for the advisers arguments:

- arg_1: $tra(\langle S, \texttt{invest(soybean)}\rangle)$ = 0.15.
- arg_2: $tra(\langle S, \neg\texttt{invest(soybean)}\rangle)$ = 0.2.
- arg_3: $tra(\langle S, \neg\texttt{invest(soybean)}\rangle)$ = 0.3.
- arg_4: $tra(\langle S, \texttt{invest(soybean)}\rangle)$ = 0.27.

where the stockholder agent could not invest in soybean market, given that the argument for `¬invest(soybean)` is stronger and defeats the others.

In contrast, if we assume an credulous agent just to tra function, we have a different result, as showed below:

- arg_1: $tra(\langle S, \texttt{invest(soybean)}\rangle)$ = 0.55.
- arg_2: $tra(\langle S, \neg\texttt{invest(soybean)}\rangle)$ = 0.4.
- arg_3: $tra(\langle S, \neg\texttt{invest(soybean)}\rangle)$ = 0.5.
- arg_4: $tra(\langle S, \texttt{invest(soybean)}\rangle)$ = 0.8.

which is interesting depending of the application domain, and allows different attitudes towards aggregating the trust on multiple sources of information.

VII. Related Work

Tang et al., in [20], combine argumentation and trust, taking into account trust on information used in argument inferences. That work is based on work presented by Parsons et al. [19], which proposes a formal model for combining trust and argumentation, aiming to find relationships between these areas. Our work differs from [19], [20] in some points. We introduce an approach for computing trust values for beliefs that differs from [19], [20], where trust on a piece of information is assumed to be more directly available. Different from those approaches, we allow for different sources for the same information (which is often the case in Jason agents) a propose ways to combine them into a single trust value for that information. We also define some agent profiles to facilitate the development of agents that require different *social* perspectives on the trust values of multiple sources; this is not considered in [19], [20] either. Another difference, although [19], [20] also consider inference rules and the structure of arguments, is that we use an implemented argumentation-based reasoning mechanism integrated with a well-known agent-oriented language, which we argue is a strong point of our work.

Parsons et al., in [17], identify ten different patterns of argumentation, called schemes, through which an individual/agent can acquire trust on another. Using a set of critical questions, the authors show a way to capture the defeasibility inherent in argumentation schemes and are able to assess whether an argument is good or fallacious. Our approach differs from [17] in that we are not interested in agents arguing about the trust they have on each other. We are interested in using such trust values and combining them with an argumentation-based reasoning mechanism in order to use trust to resolve undecided conflicts between arguments.

Biga and Casali, in [22], present an extension of Jason, called G-Jason, to allow the creation of more flexible agents to reason about uncertainty, representing belief degrees and grades using the annotation feature provided by Jason. The authors define `degOfCert(X)`, where `X` is a value between 0 and 1, as a value associated with certainty of a belief and `planRelevance(LabelDegree)` as a value associated with plans, where the `LabelDegree` value is based on its context and its triggering event's `degOfCert` level. Our approach differs from [22] in that we use the notion of trust on agents and sensors in order to infer a level of certainty on beliefs, and in our case from belief certainty we calculate the certainty of arguments as well.

Pereira et al., in [23], present an approach for agents not to miss information that is currently incorrect about the environment. That work proposes a framework for changing the agent's mind without completely erasing previous information. The authors use possibility theory to represent uncertainty about information, using a fuzzy labeling function that sets a trust degree n to sources and arguments, where $n \in [0, 1]$. Our approach differs from [23] in some aspects. First, the authors in [23] define two agent profiles: *optmistic* and *pessimistic*. Consider an argument A and $S(A) = \{a_1, ..., a_n\}$ as the set of sources of A. An *optimistic* agent will set the trust of A according to the most reliable source $a_i \in S(A)$ and a *pessimistic* agent will set the trust of A according the least reliable source $a_j \in S(A)$. Our approach considers that the trust of an argument is defined according to the trust of its beliefs, and the trust of a belief is defined according to the trust of its sources. Our approach allows for a social perspective, as a *sceptical* agent will consider the number of sources of each belief to set its trust value. In [23], it is stated that if an agent believes in φ, it could not believe in $\neg\varphi$. Our approach differs in this aspect too, as we allow an agent to believe in φ and $\neg\varphi$. Another interesting difference is that we extend the work [7], which use defeasible logic, while [23] uses a fuzzy approach and possibility theory.

VIII. Final Remarks

In this paper, we showed how an argumentation-based reasoning mechanism, implemented in an agent-oriented programming language, can be extended to take into account trust over the sources of information. Trust on the sources is used to generate trust values for beliefs by combining the trust on the multiple sources for the same piece of information. Therefore, our approach allows agents to have a *social* perspective on the information they use to construct arguments where, for example, more trustworthy sources could have less influence over the final trust value for a belief if there are more sources asserting the contrary. Also, the trust values for beliefs are used to calculate trust values for arguments constructed over these information, allowing agents to decide conflicting arguments based on such values. Further, we have introduced some profiles for agents. These profiles can be used to model different attitudes towards aggregating the trust on multiple sources of information. These differences are domain-dependant, so in some domains a credulous/sceptical agent may facilitate obtaining the desired agent behaviour as well as the overall multi-agent system behaviour. We used a stock market scenario to exemplify the approach and show how the different agent profiles could influence the resulting decisions.

Some directions of research are to consider other criteria such as the time the beliefs were added to the belief base to calculate confidence on a belief (improving our approach in terms of reducing ambivalence even further), and devising other agent profiles based on the various criteria. We also plan to investigate the use of social choice mechanisms in order to aggregate other agents' preferences on contrary beliefs.

Acknowledgements

This research was partially funded by CNPq and CAPES.

References

[1] A. R. Panisson, A. Freitas, D. Schmidt, L. Hilgert, F. Meneguzzi, R. Vieira, and R. H. Bordini, "Arguing About Task Reallocation Using Ontological Information in Multi-Agent Systems," in *ArgMAS*, 2015.

[2] S. Parsons and P. McBurney, "Argumentation-based dialogues for agent co-ordination," *Group Decision and Negotiation*, vol. 12, no. 5, pp. 415–439, 2003.

[3] S. Parsons, M. Wooldridge, and L. Amgoud, "An analysis of formal inter-agent dialogues," in *AAMAS*. ACM Press, 2002, pp. 394–401.

[4] M. Wooldridge, *An introduction to multiagent systems*. John Wiley & Sons, 2009.

[5] N. Maudet, S. Parsons, and I. Rahwan, "Argumentation in multi-agent systems: Context and recent developments." in *ArgMAS*, 2006.

[6] T. Berariu, "An argumentation framework for bdi agents," in *Intelligent Distributed Computing VII*, Springer, 2014, vol. 511, pp. 343–354.

[7] A. R. Panisson, F. Meneguzzi, R. Vieira, and R. H. Bordini, "An Approach for Argumentation-based Reasoning Using Defeasible Logic in Multi-Agent Programming Languages," in *ArgMAS*, 2014.

[8] A. R. Panisson and R. H. Bordini, "Knowledge representation for argumentation in agent-oriented programming languages," in *BRACIS*, 2016.

[9] L. Amgoud and J. Ben-Naim, "Argumentation-based ranking logics," in *AAMAS*, 2015, pp. 1511–1519.

[10] R. H. Bordini, J. F. Hübner, and M. Wooldridge, *Programming Multi-Agent Systems in AgentSpeak using Jason (Wiley Series in Agent Technology)*. John Wiley & Sons, 2007.

[11] I. Pinyol and J. Sabater-Mir, "Computational trust and reputation models for open multi-agent systems: a review," *Artificial Intelligence Review*, vol. 40, no. 1, pp. 1–25, 2013.

[12] D. Walton, C. Reed, and F. Macagno, *Argumentation Schemes*. Cambridge University Press, 2008.

[13] G. Governatori, M. J. Maher, G. Antoniou, and D. Billington, "Argumentation semantics for defeasible logic." *Journal of Logic and Computation*, vol. 14, no. 5, pp. 675–702, 2004.

[14] P. M. Dung, "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games," *Artificial Intelligence*, vol. 77, pp. 321–357, 1995.

[15] D. Nute, *Defeasible Prolog*, ser. Research report. Artificial Intelligence Programs, University of Georgia, 1993.

[16] ——, "Defeasible logic," in *Handbook of logic in artificial intelligence and logic programming*, Oxford University Press, Inc., 1994, ch. Defeasible logic, pp. 353–395.

[17] S. Parsons, K. Atkinson, K. Haigh, K. Levitt, P. M. J. Rowe, M. P. Singh, and E. Sklar, "Argument schemes for reasoning about trust," in *COMMA*, vol. 245, p. 430, 2012.

[18] S. Parsons, E. Sklar, and P. McBurney, "Using argumentation to reason with and about trust," in *Argumentation in multi-agent systems*. Springer, 2012, pp. 194–212.

[19] S. Parsons, Y. Tang, E. Sklar, P. McBurney, and K. Cai, "Argumentation-based reasoning in agents with varying degrees of trust," in *AAMAS*, 2011, pp. 879–886.

[20] Y. Tang, K. Cai, P. McBurney, E. Sklar, and S. Parsons, "Using argumentation to reason about trust and belief," *Journal of Logic and Computation*, p. 38, 2011.

[21] V. S. Melo, A. R. Panisson, and R. H. Bordini, "Trust on beliefs: Source, time and expertise," in *8th Int. Ws. on Trust in Agent societies*, 2016.

[22] A. C. Adrián Biga, "G-jason: An extension of jason to engineer agents capable to reason under uncertainty," in *Proc. of 14th Intelligent Agent and Systems Workshop*, 2014.

[23] C. da Costa Pereira, A. G. Tettamanzi, and S. Villata, "Changing one's mind: Erase or rewind? possibilistic belief revision with fuzzy argumentation based on trust," in *IJCAI*, vol. 1, 2011, pp. 164–171.

[24] V. S. Melo, A. R. Panisson, and R. H. Bordini, "Argumentation-based reasoning using preferences over sources of information," in *AAMAS*, 2016.