Argumentation-Based Reasoning Using Preferences over Sources of Information

(Extended Abstract)

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

ABSTRACT

Argumentation-based reasoning plays an important role in agent reasoning and communication. In this work, we extend an argumentation-based reasoning mechanism to take into account preferences over arguments supporting contrary conclusions. 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. 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.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems, intelligent agents

Keywords

Argumentation-Based Reasoning, Trust

1. INTRODUCTION

An important topic of research in argumentation for autonomous agents is argumentation-based reasoning, and recent work has brought this to the context of agent-oriented programming languages [4]. In that context, agents need to reason about possible arguments in order to make decisions and to communicate. Further, argumentation-based reasoning includes the capability of agents to construct arguments in the face of uncertainty (i.e., incomplete and incorrect information). Therefore, it is important that the agents are able to construct arguments using the most precise pieces of information available to them, based on the most trustworthy sources, in order to avoid as much as possible elements of doubt in the arguments used, hence improving their decisions and therefore their actions (in particular interactions). With these issues in mind, we propose an approach to combine argumentation-based reasoning and preferences over sources of information, for example using information about trust on the agents who provided the information used to construct arguments. This combination allows the agents to make decisions in situations where they would not have been able to make a clear choice, for example because of undecided conflicts in argumentation-based reasoning mechanisms without such preferences.

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 [1], where beliefs are annotated with all known sources of that information. Furthermore, elaborate trust systems have been studied [6] in the context of multi-agent systems which could provide reliable trust information about each such source.

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

2. REASONING MECHANISM

We here use the work presented in [4] as argumentation-based reasoning mechanism, so arguments are constructed using strict and defeasible inferences rules. In [4], conflict between arguments are of two types[1]: (i) an argument $\langle S_1, c_1\rangle$ *rebuts* the argument $\langle S_2, c_2\rangle$ if $c_1 \equiv \overline{c_2}$, and (ii) an argument $\langle S_1, c_1\rangle$ *undercuts* $\langle S_2, c_2\rangle$ if $c_1 \equiv \overline{c_3}$ with $c_3 \in S_2$.

An argument attacking another does not necessarily mean that it defeats the other. Defeat is a "successful" attack, and it considers the set of arguments that defend each other, including preferences between the conflicting arguments [8]. In [4], the set of acceptable arguments from an agent's knowledge base is defined in terms of the *defeasible semantics* introduced in [2], which considers some basic preferences over arguments. Clearly, strict arguments are stronger than defeasible arguments, so they always have priority, i.e., when arguments are in conflict, strict arguments always defeat defeasible ones. Considering only defeasible arguments, the work in [4] considers two types of priority[2]: (i) priority by specificity, where more specific conclusions are preferred over more general ones, and (ii) the explicit declaration of priority between defeasible rules, using a special predicate `sup(Rule1,Rule2)`, indicating that `Rule1` has priority over (i.e., is superior to) `Rule2`.

[1] We use '$\neg$' for strong negation and an overline as a *contrary* operation, thus $\overline{\varphi} \equiv \neg\varphi$ and $\overline{\neg\varphi} \equiv \varphi$.
[2] Originally from Nute's defeasible logic [3].

Although the approach presented in [4] provides some means for dealing with conflicting information (arguments in particular), when conflicts cannot be resolved considering only the set of arguments, that approach proves limited. Therefore, we propose that when an agent has conflicting arguments that cannot be decided upon by specificity or explicit declaration of preference, the agent can use meta-information such as which of the arguments was constructed using more trustworthy information.

3. TRUST ON BELIEFS

There are many approaches to models of trust in the agents literature [5, 6, 7]. Often a function $tr(Ag_i, Ag_j)$ returning a value between 0 and 1 is used to represent how much agent Ag_i trusts agent Ag_j[3]. However, an agent can obtain information from different sources (e.g., from other agents and from perceiving the environment). Therefore, we expand the trust function to have as input an agent and the possible sources of information, so function $tr(Ag_i, Ag_j)$ is generalised to $tr(Ag_i, s_j)$, where s_j represents one of the sources of information for agent Ag_i.

In order to define how much Ag_i trusts in some information φ (denoted by $trb_i(\varphi)$), we consider the tr value associated with each source of φ for Ag_i. To this end, we introduce two agent profiles for calculating trust values over beliefs: (i) **credulous agents** consider only the most trustworthy source of information, i.e., $trb_i(\varphi) = max\{tr(Ag_i, s_1), ..., tr(Ag_i, s_n)\}$, where $\{s_1, ..., s_n\}$ is the set of sources that informed φ to Ag_i, and (ii) **sceptical agents** consider all the sources from which they received the information as well as the trust value of each such source in order to have some form of social trust value, i.e., $trb_i(\varphi) = \frac{\sum_{s \in S_\varphi^+} tr(Ag_i, s)}{|S_\varphi^+|+|S_\varphi^-|}$, 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 $\overline{\varphi}$.

4. TRUST ON ARGUMENTS

In order to enable agents to make decisions over those still undecided conflicts between arguments, we combine the argumentation-based reasoning mechanism from [4] and the trust value for beliefs introduced above. The approach presented here is applicable to both premises and inference rules as used in [4]. 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.

Definition 1 (Trust on Arguments). *The trust on an argument is based on the trust value of its support. For an argument $\langle S, c\rangle$, its trust value is given by the trust on its support S, i.e., $tra(\langle S, c\rangle) = trb(\varphi_1) \otimes^{tra} \ldots \otimes^{tra} trb(\varphi_n)$ where $S = \{\varphi_1, \ldots, \varphi_n\}$ is the support set for that argument.*

Considering the agent profiles we introduced, the generic operator $\otimes^{tra}$ can be defined as follows: (i) **credulous agents** use $\otimes^{tra}$ as the maximum trust value (the highest trust value present in the argument's support), i.e., $tra(\langle S, c\rangle) = \max\{trb(\varphi_1), \ldots, trb(\varphi_n)\}$; and (ii) **sceptical agents** use $\otimes^{tra}$ as the minimum trust value, i.e., $tra(\langle S, c\rangle) = \min\{trb(\varphi_1), \ldots, trb(\varphi_n)\}$. When agent Ag_i has multiple arguments for a conclusion c, it can opt for the argument that has the highest trust value, i.e., $tra(\langle S, c\rangle) = \max\{tra(\langle S_1, c\rangle), \ldots, tra(\langle S_n, c\rangle)\}$. Therefore, when it has an undecided conflict between two arguments, it can resolve the conflict by looking at the trust values.

Definition 2 (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 3 (Undercutting Defeat using Trust). *Let $\langle S_1, c_1\rangle$ and $\langle S_2, c_2\rangle$ be two conflicting arguments, where $c_1 \equiv \overline{c_3}$ with $c_3 \in S_2$. We say that $\langle S_1, c_1\rangle$* undercuts *$\langle S_2, c_2\rangle$ iff $tra(\langle S_1, c_1\rangle) > tra(\langle S_2, c_2\rangle)$.*

Although we introduced only two simple agent profiles above, clearly other profiles and instantiations for the generic operators could be used, as suggested in [5, 7].

5. FINAL REMARKS

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 such sources is used to generate trust values for beliefs by combining the trust on the multiple sources for the same piece of information. Our approach allows agents to have a *social* perspective on the information they use to construct arguments. 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. As the trust values for beliefs are then used to calculate trust values for arguments, the agents can decide upon conflicting arguments based on such values. Further, we have introduced some agent profiles that can be used to model different attitudes towards aggregating the trust on multiple sources of information. The profile choice for an agent is arguably domain dependent.

REFERENCES

[1] 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.

[2] G. Governatori, M. J. Maher, G. Antoniou, and D. Billington. Argumentation semantics for defeasible logic. *J. Log. Comp.*, 14(5):675–702, 2004.

[3] D. Nute. Defeasible logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 353–395. Oxford University Press, 2001.

[4] 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.

[5] S. Parsons, Y. Tang, E. Sklar, P. McBurney, and K. Cai. Argumentation-based reasoning in agents with varying degrees of trust. In *AAMAS*, 2011.

[6] I. Pinyol and J. Sabater-Mir. Computational trust and reputation models for open multi-agent systems: a review. *Artificial Intelligence Review*, 40(1):1–25, 2013.

[7] Y. Tang, K. Cai, P. McBurney, E. Sklar, and S. Parsons. Using argumentation to reason about trust and belief. *J. Log. Comp.*, page 38, 2011.

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

[3] We refer to [5] for further details about the trust framework.