ABA: Argumentation Based Agents

A. Kakas[1] and **L. Amgoud**[2] and **G.Kern-Isberner**[3] and **N. Maudet**[4] and **P. Moraitis**[5]

Abstract. Many works have identified the potential benefits of using argumentation to address a large variety of multiagent problems. In this paper we take this idea one step further and develop the concept of a fully integrated argumentation-based agent architecture that allows us to develop agents that are coherently designed on an underlying argumentation based foundation. Under this architecture, an agent is composed of a collection of modules each of which is equipped with a local argumentation theory. Similarly, the intra-agent control of the agent is governed by local argumentation theories that are sensitive to the current situation of the agent through dynamically enabled feasibility arguments.

1 Introduction

In recent years, many authors have promoted argumentation as a means to deal with specific multi-agent problems, for instance negotiation or communication with other agents. Indeed, recently argumentation has seen its scope greatly extended, so that it now covers many of the features usually associated to the theories of agency [5]. The benefits of argumentation are well established: a high-level of flexibility and expressiveness, allowing powerful and diverse reasoning tasks to be performed. In this work we take the use of argumentation one step further by developing the concept of a fully integrated argumentation-based agent (ABA) architecture. This can be seen as a global framework where all these separated features could be glued together, both in terms of abstract design and technical specification. We lay the foundations of such an approach to agency, present an abstract agent architecture based on argumentation and indicate the type of properties that we could expect from ABA agents.

It is important to note that the ABA architecture does not depend on any specific argumentation framework but only requires some quite general properties such as including some notion of (relative) strength of arguments and some notion of dynamically enabled arguments, as found for example in [1, 2]. Irrespective of the framework used, the argumentation-based foundation of ABA agents provides various advantages, including that of its rational and explainable decisions that facilitate the focus of purpose by the agent.

This work grew out of the initiative of the 2008 Dagstuhl meeting on Argumentation to ask groups of researchers to propose ways of consolidating the work on several main themes of argumentation in Computer Science, such as the theme of argumentation in agents, which is the concern of this paper.

2 ABA Architecture

The ABA architecture's basic principle is to build an agent from a loosely coupled set of modules that are to a large extent independent from each other with no or minimal central control. Each module is based on an argumentation theory, concerning a certain internal capability of the agent, that provides a policy of how to take decisions (preferred choices) for this type of task. A module contains also another argumentation theory responsible for its involvement in the intra-agent control (IAC) of the agent. Together these local IAC theories give an argumentation-based communication protocol between the modules, which effects the internal operation of the agent.

Definition 1 (ABA Agent Module) *An ABA agent module is a tuple* $M = \langle IAC, T, R\rangle$ *where:*

- IAC *is an argumentation theory for intra-agent control,*
- T *is an argumentation theory for the task of the module,*
- $R = \langle P, C\rangle$ *where* P *and* C *are sets of names of other modules, the parent and child modules of* M *respectively.*

The argumentation theory, T, of each module is an expert (preference) policy comprising of basic arguments for the different decision choices together with priority arguments on the relative strength of arguments. The priority arguments can be build based on a parametrization of the arguments and the relative importance of parametric criteria. The sets P and C of a module give a dependence between the modules that captures a request-server relationship where the decisions taken by a parent module form part of the problem task of a child module. For example, a PLANNING module will be a child of a GOAL DECISION module since PLANNING decides on plans to achieve the goals decided by GOAL DECISION.

Definition 2 (ABA Agent) *An ABA agent is a tuple,* $\langle Ms, Mot, WV\rangle$*, where*

- $Ms = \{M_1, ..., M_n\}$ *is a set of ABA modules for the different internal capabilities of the agent,*
- Mot *is an argumentation theory for its motivations and needs,*
- WV *is a theory that captures the world view that the agent has about external its environment.*

The number of modules and the capability they each provide to the agent is not fixed but can vary according to the type of application that the agent is built for. However, the MOTIVATIONS AND NEEDS (Mot) and the WORLD VIEW (WV) modules play a central role and are arguably required to design any ABA agent.

Motivations and needs. The Mot module governs the high-level Motivations and Needs of an ABA agent (c.f. [2] for capturing the agent's personality using Maslow's motivations theory). The argumentation theory in Mot decides on the current high-level Needs of the agent, its *Desires*, that drive its behaviour. Needs act as parameters for the arguments in many of the other modules, thus allowing motivations to shape the overall behavior of the agent.

---

[1] University of Cyprus, antonis@ucy.ac.cy
[2] IRIT-CNRS at Toulouse, amgoud@irit.fr
[3] University of Dortmund, gabriele.kern-isberner@cs.uni-dortmund.de
[4] University Paris 9 Dauphine, maudet@lamsade.dauphine.fr
[5] Paris Descartes University, pavlos@mi.parisdescartes.fr

World view. The WORLD VIEW module, WV, contains the world beliefs of the agent providing a common view of the current state of the world to all other modules. The basic and priority arguments of the agent depend on the world view, thus making them context dependent and adaptable to changes in the external environment of the agent. The WV module itself can also be based, if the designer so wishes, on an argumentation theory.

Definition 3 (Agent State) *A state of an ABA agent, $\langle Ms, Mot, WV \rangle$, is a tuple $\langle V, \mathcal{D} \rangle$ where:*

- *V represents the current view of the world as given by WV,*
- *$\mathcal{D} = \{CS_{M_1}, ..., CS_{M_n}\}$ where each, CS_{M_i}, is a tuple $\langle D, L, S \rangle$, representing the current state of the module M_i, where D is its current decision, as given by its argumentation theory, T_i, L is the level of commitment on D and $S \in \{keep, abandon\}$ is the current status of the decision D.*

The level of commitment and status of a module's decision are maintained by the intra-agent control, IAC theory of the module.

Definition 4 (IAC Argumentation Theory) *The intra agent control theory of a module, M, is a tuple $\langle T_L, P_{Status} \rangle$ where:*

- *T_L is theory for defining the commitment level, L, for the (object-level) decisions in M,*
- *P_{Status} is an ABA argumentation theory for the options $Keep(D)$ or $Abandon(D)$, with D a decision in M.*

The arguments in P_{Status} for keeping or not a current decision can be annotated (or expressed) in terms of relative changes in the level of commitment as time passes and new information from the external environment is acquired. Although this can be specified in different ways, the argumentative basis of an ABA agent suggests the following natural form of commitment:

Definition 5 *Let D be a decision of a module and $T(V)$ denote the module's argumentation theory T grounded on the current world view V. Then the current* commitment level *for D is given as follows:*

- *Level 4, iff D is uniquely (sceptically) preferred by $T(V)$*
- *Level 3, iff D is credulously preferred by $T(V)$*
- *Level 2, iff D is not preferred by $T(V)$, but there exists a basic argument for D*
- *Level 1, iff D does not have a basic argument in $T(V)$*

Hence the commitment level reflects the degree of preference of the decision with respect to the agent's subjective optimality arguments in its module that are enabled by the current world view. Changes in this view then affect the commitment level of the current decision, which can change its status and in turn this can change the commitment and status of decisions of other parent or child modules.

Feasibility arguments. When deciding the status of a decision it is useful to distinguish between *feasibility* arguments and *optimality* arguments. Feasibility arguments refer to the feasibility of a given decision based on some external feedback pertaining to its validity in the real world, while *optimality* arguments are situation independent arguments for the value of a given decision.

When to reconsider? The reconsideration of the commitment level and status of the current decision in a module can be computationally non-effective. Hence to make the operation of P_{Status} more practical we can layer its decision process into two stages. In the first stage we apply a lightweight *Decision Reconsideration* policy that efficiently tells us whether we indeed need to reconsider the current decision. Only if the result from this is affirmative we continue to consider the full P_{Status} reasoning for deciding the fate of the current decision. Otherwise, we keep the current decision.

3 Properties of ABA Agents

ABA agents are designed so that their operation is based on informed decisions. The preferred choices in any module are meant to capture the best solutions available at the time. Hence an ABA agent's operation should follow these choices as intended in its design.

Property 1 *An ABA agent such that for any of its states, $\langle V, \mathcal{D} \rangle$, every decision $D \in \mathcal{D}$ is preferred by the argumentation policy in its module, is called a* **strongly sound agent.**

A strongly sound agent is therefore one whose decisions are not only optimal at the time that they are taken but remain optimal at any subsequent situation where its view of the world may have changed. In practice though in some applications this may be too strong to require as it may mean that decisions are abandoned too often. This can be mitigated, *e.g.*, by taking the cost induced by discarding this decision into account, or by requiring a weaker form of soundness where only some of the decisions are optimal throughout the operation of the agent. In particular, the higher level decisions in the "hierarchy" of modules, such as the goal decisions should remain optimal. The individual module decisions need to be coherent with each other and give some overall sense to the agent's operation. This is the role of the Motivations and Needs policy of the agent: the agent must operate in accordance to its current high-level desires and needs. We can then (re)formulate properties of a **soundly motivated** agent where its decisions always remain preferred with respect to its Motivations and Needs policy.

4 Conclusions

We have proposed an agent architecture uniformly based on argumentation with a highly modular structure. The focus is on a high-level architecture mainly concerned with managing the currently available best options for the agent's constituent tasks in a way that provides a coherent behaviour of the agent. Our work shares similarities with other argumentation based agent approaches, when it comes to addressing specific issues and features of agents, e.g. in the KGP model of agency [3] goal decision and cycle theories for internal control are also captured through argumentation. The closest connection is with the work in [4] which also proposes an Agent Argumentation Architecture (called AAA). As in our case, argumentation is used to arbitrate between conflicting motivations and goals.

An important distinguishing characteristic of an ABA agent is that its argumentation based decisions are not rigid but rather they are defeasible decisions for currently preferred options that can be different under a different view of the world. This means that the agent is flexible and versatile in a changing environment, able to adapt graciously to changes in the agent's current situation, without the need for an explicit mechanism of adaptation.

REFERENCES

[1] Trevor J. M. Bench-Capon, 'Value based argumentation frameworks', *CoRR*, **cs.AI/0207059**, (2002).

[2] A. C. Kakas and P. Moraitis, 'Argumentation based decision making for autonomous agents', in *AAMAS '03*, pp. 883–890, (2003).

[3] Antonis C. Kakas, Paolo Mancarella, Fariba Sadri, Kostas Stathis, and Francesca Toni, 'Computational logic foundations of kgp agents', *J. Artif. Intell. Res. (JAIR)*, **33**, 285–348, (2008).

[4] Mark Witkowski and Kostas Stathis, 'A dialectic architecture for computational autonomy', in *Agents and Computational Autonomy*, pp. 261–274, (2003).

[5] M. J. Wooldridge and A. Rao, *Foundations of rational agency*, Kluwer, 1999.