Abstracting Assumptions in Structured Argumentation

Extended Abstract

Iosif Apostolakis
Institute of Software Technology
TU Graz
Graz, Austria
apostolakis@ist.tugraz.at

Zeynep G. Saribatur
Institute of Logic and Computation
TU Wien
Vienna, Austria
zeynep@kr.tuwien.ac.at

Johannes P. Wallner
Institute of Software Technology
TU Graz
Graz, Austria
wallner@ist.tugraz.at

ABSTRACT

In this work we apply a form of existential abstraction on the prominent structured approach of assumption-based argumentation (ABA) via clustering assumptions, leading to simplified argumentation scenarios supporting explainability. We present ways of interpreting clustered ABA frameworks, look at use cases, and provide an interactive automated tool that obtains faithful clusterings.

KEYWORDS

Structured Argumentation; Existential Abstraction

ACM Reference Format:
Iosif Apostolakis, Zeynep G. Saribatur, and Johannes P. Wallner. 2024. Abstracting Assumptions in Structured Argumentation: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

1 INTRODUCTION

Computational argumentation is a well-established area within Artificial Intelligence (AI) [3, 15]. Specifically within multi-agent systems, argumentation was applied in several directions, e.g., in dialogues and negotiation [1, 10], persuasion [22], decision making [16], debates [13], and interaction of agents [7, 21].

This field can be categorized into two central approaches, namely structured argumentation [5, 8, 17–20] and abstract argumentation [3, 11]. The former provides principled approaches of how to reason argumentatively on given, possibly conflicting, knowledge bases. This is achieved by prescribing how to instantiate argument structures and their relationships from given knowledge. Abstract argumentation, on the other hand, provides approaches how to find acceptable (sets of) arguments when arguments are seen as abstract entities. For instance, in argumentation frameworks (AFs) [11] arguments are represented as vertices and a directed counter-argument (attack) relation is represented as directed edges.

Among the methods supporting explainability, we in particular find approaches to simplify given argumentation scenarios [2, 12, 14, 26, 28], with a recent approach utilizing abstraction on AFs via clustering of arguments [27]. This method is of use for focusing on relevant details and abstracting away redundant or undesired parts, providing at the same time an automated way of doing so in an interactive fashion with users querying for more or less abstraction.

This approach was presented for AFs, and, as witnessed by many works in the field, a lifting or adaptation to structured argumentation is both not direct [23–25, 29] and brings such approaches closer to applications. In this paper we apply existential abstraction by clustering (parts of) given knowledge bases and thus lift clustering to the level of structured argumentation. We focus on the prominent structured approach of assumption-based argumentation (ABA) [8], with applications in, e.g., medical reasoning [9] and multi-agent systems [16]. We focus on flat ABA frameworks in the prominent logic-programming fragment [6].

2 ASSUMPTION-BASED ARGUMENTATION

An assumption-based argumentation framework (ABAF) $D = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{})$, consists of a set of inference rules $\mathcal{R}$ over the formal language $\mathcal{L}$. From the language $\mathcal{L}$ we distinguish the set of assumptions $\mathcal{A} \subseteq \mathcal{L}$, that represent the pieces of information that are defeasible. Each assumption can have an atom as its contrary, and conflicts (derivations of contraries) are resolved using argumentation semantics. A rule r has the form $a_0 \leftarrow a_1, \ldots, a_n$, with $cl(r) = a_0 \in \mathcal{L} \setminus \mathcal{A}$ as the head of the rule representing the derived atom, while the $body(r) = \{a_1, \ldots, a_n\} \subseteq \mathcal{L}$ constitutes the rule body representing its premises. In this context, an argument is a derivation using rules in $\mathcal{R}$ to derive a claim from a given set of assumptions $S \subseteq \mathcal{A}$. For a set $S \subseteq \mathcal{A}$ we define derivability in an ABAF D via the set $Th_D(S)$ that contains all claims s with arguments based on subsets of S.

A set A of assumptions attacks a set B of assumptions if one can derive the contrary of some assumption in B from A. If a set can derive none of its own contraries, that is if the set does not attack itself, then this set is said to be conflict-free in D. Additionally, if A is conflict-free, and for each set $C \subseteq \mathcal{A}$ that attacks an assumption $a \in A$, it holds that A attacks C, then A is called admissible, and we also say that A defends a. The sets $cf(D)$ and $adm(D)$ are called semantics and correspond to the sets of all conflict-free sets and admissible sets respectively. Stable semantics is composed of conflict-free sets A that attack every assumption in $\mathcal{A} \setminus A$.

Example 1. *As an example we consider the ABAF D with assumption set $\mathcal{A} = \{a, b, c, d, e\}$, $\mathcal{L} = \{a, b, c, d, e, x, y, z, w, t\}$, and contraries $\overline{a} = z$, $\overline{b} = t$, $\overline{c} = x$, $\overline{d} = w$, and $\overline{e} = t$. Consider also the rules to be the set $\mathcal{R} = \{x \leftarrow a, b, w \leftarrow e, y \leftarrow c, z \leftarrow d\}$. The set $\{b, c, e\}$ is conflict-free, and this set is also stable and admissible. On the contrary the sets $\{a, b, c\}$ and $\{a, b, d\}$ are not conflict-free, and thus they are neither admissible nor stable.*

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

3 CLUSTERING ASSUMPTIONS

Let us start by defining the notion of existential abstraction in ABAFs. In this work, existential abstraction refers to clustering the set of assumptions of an ABAF, obtained by a mapping m, intuitively mapping assumptions to clusters.

Definition 1. *Given an ABAF* $D = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{})$*, let* $\hat{\mathcal{A}}$ *be a partition of* $\mathcal{A}$ *and m be the surjective mapping from* $\mathcal{A}$ *to* $\hat{\mathcal{A}}$*. Then, the clustered ABAF (cABAF) of D according to m is* $\hat{D} = m(D) = (\hat{\mathcal{L}}, \hat{\mathcal{R}}, \hat{\mathcal{A}}, \hat{\overline{}})$ *where*

- $\hat{\mathcal{L}} \setminus \hat{\mathcal{A}} = \mathcal{L} \setminus \mathcal{A}$ *and* $\hat{\mathcal{A}} = m(\mathcal{A})$,
- *each rule* $r \in \mathcal{R}$ *is mapped to the rule* $\hat{r}$ *of the form* $cl(r) \leftarrow m(body(r))$*, hence obtaining* $\hat{\mathcal{R}}$,
- $\hat{\overline{}}$ *is a total mapping from* $\hat{\mathcal{A}}$ *to* $2^{\hat{\mathcal{L}}}$ *such that for all* $\hat{a} \in \hat{\mathcal{A}}$ *we have* $\overline{\hat{a}} = \{\hat{b} \in \hat{\mathcal{A}} \mid \exists b \in \hat{b}, a \in \hat{a} \text{ s.t. } b = \overline{a}\} \cup \{\hat{x} \in \hat{\mathcal{L}} \setminus \hat{\mathcal{A}} \mid \exists a \in \hat{a} \text{ s.t. } x = \overline{a}\}$.

Example 2. *In the example above, let m be the mapping that maps assumptions d and c into the same cluster* $\hat{c}$*, and the assumptions a, b, e to themselves. This mapping leads to a cABAF* $\hat{D}$ *with an assumption set* $\hat{\mathcal{A}} = \{a, b, \hat{c}, e\}$*. It is evident that a cABAF does not strictly fit in the definition of a classical ABAF, since the contrary of the cluster* $\hat{c}$ *is not an element of the set* $\hat{\mathcal{L}}$*. In fact,* $\overline{\hat{c}} = \{x, w\} \notin \hat{\mathcal{L}}$*.*

Regarding the rules, as Definition 1 states, for each rule in $\mathcal{R}$*, we get a clustered rule by applying m to its body. Thus, the clustered rule set is* $\hat{\mathcal{R}} = \{x \leftarrow a, b, w \leftarrow e, y \leftarrow \hat{c}, z \leftarrow \hat{c}\}$*.*

4 SEMANTICS OF CLUSTERINGS

We interpret cABAFs by defining (abstract) semantics that aim to approximate classical semantics on ABAFs. More formally, for a given classical semantics σ, we define abstract semantics $\hat{\sigma}$ s.t. $m(\sigma(D)) \subseteq \hat{\sigma}(\hat{D})$ for any frameworks D with $\hat{D} = m(D)$. For such abstract semantics, we say that $\hat{\sigma}$ abstracts σ (thus over-approximating the classical counterparts).

It turns out that it is useful to distinguish three kinds of attacks on cABAFs, each representing different "strengths" of attacks, when considering the clustering. With $Single(\hat{A})$ we denote all singleton clusters in $\hat{A}$ (non-abstracted assumptions).

Definition 2. *Let* $\hat{D} = (\hat{\mathcal{L}}, \hat{\mathcal{R}}, \hat{\mathcal{A}}, \hat{\overline{}})$ *be a cABAF, and* $\hat{A}, \hat{B} \subseteq \hat{\mathcal{A}}$*.*

- $\hat{A}$ *(normally) attacks* $\hat{B}$ *if* $\exists \hat{b} \in \hat{B}$ *and* $\overline{\hat{b}} \cap Th_{\hat{D}}(\hat{A}) \neq \emptyset$,
- $\hat{A}$ *fully attacks* $\hat{B}$ *if* $\exists \hat{b} \in \hat{B}$ *and* $\overline{\hat{b}} \subseteq Th_{\hat{D}}(\hat{A})$*, and*
- $\hat{A}$ *truly attacks* $\hat{B}$ *if* $\exists \hat{b} \in \hat{B}$ *and* $\overline{\hat{b}} \subseteq Th_{\hat{D}}(Single(\hat{A}))$*.*
- $\hat{A}$ *defends a cluster* $\hat{a}$*, if* $\exists x \in \overline{\hat{a}}$ *s.t.* $\forall \hat{C}$ *with* $x \in Th_{\hat{D}}(\hat{C})$*, it holds that* $\hat{A}$ *attacks* $\hat{C}$*.*

Now we are ready to define our abstract semantics.

Definition 3. *Let* $\hat{D} = (\hat{\mathcal{L}}, \hat{\mathcal{R}}, \hat{\mathcal{A}}, \hat{\overline{}})$ *be a cABAF. A set of clusters* $\hat{A} \subseteq \hat{\mathcal{A}}$ *is*

- *conflict-free in* $\hat{D}$ *if it does not attack itself truly,*
- *admissible, if it is conflict-free and it defends all of its clusters, and*
- *stable iff it is conflict-free,* $\forall \hat{a} \notin \hat{A}$ *there must be a full attack from* $\hat{A}$ *to* $\hat{a}$*, and if* $\hat{a} \in \hat{A}$*, then if* $\hat{S} \subseteq \hat{A}$ *is a set of clusters that fully attack* $\hat{a}$*, then* $\hat{A}$ *must attack some cluster in* $\hat{S}$*.*

As is the case in classical ABAFs, a stable set is also admissible.

Example 3. *The set of singletons* $\{a, b, e\}$ *truly attacks the cluster* $\hat{c}$*. Hence, the set* $\{a, b, e, \hat{c}\}$ *is not conflict-free. Cluster* $\hat{c}$ *attacks fully the singleton a. This implies that* $\{a, \hat{c}\} \notin \hat{adm}(\hat{D})$*. However, the singleton e defends a from atom z, and consequently* $\{a, e\} \in \hat{adm}(\hat{D})$*. The set* $\{b, e, a\}$ *and* $\{b, e, \hat{c}\}$ *are both stable under the abstract semantics. The former set is a set of singletons, and the same set is also stable in the classical framework.*

We have shown that the abstract semantics defined above are all abstracting their classical counterparts. This is a key property of the abstract semantics as it guarantees that through abstracting we do not lose any essential information regarding assumption sets. However, we still have to avoid adding extra information that does not exist in the classical framework. A set that lies in the abstract semantics but has no preimage through m in the classical semantics is called spurious. A cABAF without spurious sets is deemed faithful. Spuriousness cannot be avoided, but we can aim to minimize spuriousness. We prove that our definition of abstract conflict-free semantics is optimal in a formally defined sense. For admissible semantics, we prove that defining an optimal semantics that abstracts *adm* requires conditions that are NP-hard to verify. Deciding whether a set is spurious under $\hat{adm}$ is coNP-hard.

5 OBTAINING CLUSTERINGS AND USE CASES

We present two methods of constructing faithful clusterings: one by automatically refining initially coarse clusterings and the other by iteratively abstracting more and more assumptions into clusters. We present a tool for the former approach, and use cases on ABAFs that abstract large parts of the given set of assumptions to provide, e.g., small reasons for no stable assumptions to exist or for interactions between assumptions for showing that an assumption set deriving a queried atom exists.

6 CONCLUSION

In this work we introduced existential abstraction to assumption-based argumentation (ABA), by clustering assumptions, lifting a recent work on clustering on AFs [27].

We believe that our approach can be beneficial for supporting explainability, a key area of formal argumentation, by providing foundational work towards abstracting certain parts of argumentative reasoning in a faithful manner. Interactive tools that give users the ability to "zoom in" or "zoom out" can be useful to improve understanding and employment of formal argumentation.

We think more research is needed to make argumentation more accessible and to help users to digest argumentative results. Among interesting avenues for future works are, e.g., extending our approach to other formal approaches to structured argumentation [4]. To facilitate interaction, visualization tools are a fruitful direction to present abstractions, and more generally argumentative reasoning.

ACKNOWLEDGEMENTS

This research was funded in whole or in part by the Austrian Science Fund (FWF) by grants P35632 and T-1315.

REFERENCES

[1] Leila Amgoud, Yannis Dimopoulos, and Pavlos Moraitis. 2007. A unified and general framework for argumentation-based negotiation. In *Proc. AAMAS*, Edmund H. Durfee, Makoto Yokoo, Michael N. Huhns, and Onn Shehory (Eds.). IFAAMAS, 967–974.

[2] Pietro Baroni, Guido Boella, Federico Cerutti, Massimiliano Giacomin, Leendert W. N. van der Torre, and Serena Villata. 2014. On the Input/Output behavior of argumentation frameworks. *Artif. Intell.* 217 (2014), 144–197.

[3] Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre (Eds.). 2018. *Handbook of Formal Argumentation.* College Publications.

[4] Philippe Besnard, Alejandro Javier García, Anthony Hunter, Sanjay Modgil, Henry Prakken, Guillermo Ricardo Simari, and Francesca Toni. 2014. Introduction to structured argumentation. *Argument Comput.* 5, 1 (2014), 1–4.

[5] Philippe Besnard and Anthony Hunter. 2018. A Review of Argumentation Based on Deductive Arguments. In *Handbook of Formal Argumentation*, Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre (Eds.). College Publications, Chapter 9, 437–484.

[6] Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. 1997. An Abstract, Argumentation-Theoretic Approach to Default Reasoning. *Artif. Intell.* 93 (1997), 63–101.

[7] Álvaro Carrera and Carlos Angel Iglesias. 2015. A systematic review of argumentation techniques for multi-agent systems research. *Artif. Intell. Rev.* 44, 4 (2015), 509–535.

[8] Kristijonas Čyras, Xiuyi Fan, Claudia Schulz, and Francesca Toni. 2018. Assumption-Based Argumentation: Disputes, Explanations, Preferences. In *Handbook of Formal Argumentation*, Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre (Eds.). College Publications, Chapter 7, 365–408.

[9] Kristijonas Cyras, Tiago Oliveira, Amin Karamlou, and Francesca Toni. 2021. Assumption-based argumentation with preferences and goals for patient-centric reasoning with interacting clinical guidelines. *Argument Comput.* 12, 2 (2021), 149–189.

[10] Yannis Dimopoulos, Jean-Guy Mailly, and Pavlos Moraitis. 2019. Argumentation-based Negotiation with Incomplete Opponent Profiles. In *Proc. AAMAS*, Edith Elkind, Manuela Veloso, Noa Agmon, and Matthew E. Taylor (Eds.). IFAAMAS, 1252–1260.

[11] Phan Minh Dung. 1995. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. *Artif. Intell.* 77, 2 (1995), 321–358.

[12] Wolfgang Dvořák, Matti Järvisalo, Thomas Linsbichler, Andreas Niskanen, and Stefan Woltran. 2019. Preprocessing Argumentation Frameworks via Replacement Patterns. In *Proc. JELIA (Lecture Notes in Computer Science, Vol. 11468)*, Francesco Calimeri, Nicola Leone, and Marco Manna (Eds.). Springer, 116–132.

[13] Xiuyi Fan and Francesca Toni. 2012. Agent Strategies for ABA-based Information-seeking and Inquiry Dialogues. In *Proc. ECAI (Frontiers in Artificial Intelligence and Applications, Vol. 242)*, Luc De Raedt, Christian Bessiere, Didier Dubois, Patrick Doherty, Paolo Frasconi, Fredrik Heintz, and Peter J. F. Lucas (Eds.). IOS Press, 324–329.

[14] Xiuyi Fan and Francesca Toni. 2015. On Computing Explanations in Argumentation. In *Proc. AAAI*, Blai Bonet and Sven Koenig (Eds.). AAAI Press, 1496–1502.

[15] Dov Gabbay, Massimiliano Giacomin, Guillermo R. Simari, and Matthias Thimm (Eds.). 2021. *Handbook of Formal Argumentation.* Vol. 2. College Publications.

[16] Yang Gao, Francesca Toni, Hao Wang, and Fanjiang Xu. 2016. Argumentation-Based Multi-Agent Decision Making with Privacy Preserved. In *Proc. AAMAS*, Catholijn M. Jonker, Stacy Marsella, John Thangarajah, and Karl Tuyls (Eds.). ACM, 1153–1161.

[17] Alejandro Javier García and Guillermo Ricardo Simari. 2018. Argumentation Based on Logic Programming. In *Handbook of Formal Argumentation*, Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre (Eds.). College Publications, Chapter 8, 409–435.

[18] Thomas F. Gordon, Henry Prakken, and Douglas Walton. 2007. The Carneades model of argument and burden of proof. *Artif. Intell.* 171, 10-15 (2007), 875–896.

[19] Antonis C. Kakas, Pavlos Moraitis, and Nikolaos I. Spanoudakis. 2019. *GORGIAS*: Applying argumentation. *Argument Comput.* 10, 1 (2019), 55–81.

[20] Sanjay Modgil and Henry Prakken. 2018. Abstract Rule-Based Argumentation. In *Handbook of Formal Argumentation*, Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre (Eds.). College Publications, Chapter 6, 287–364.

[21] Simon Parsons, Carles Sierra, and Nicholas R. Jennings. 1998. Agents That Reason and Negotiate by Arguing. *J. Log. Comput.* 8, 3 (1998), 261–292.

[22] Henry Prakken. 2009. Models of Persuasion Dialogue. In *Argumentation in Artificial Intelligence*, Guillermo Ricardo Simari and Iyad Rahwan (Eds.). Springer, 281–300.

[23] Henry Prakken. 2023. Relating Abstract and Structured Accounts of Argumentation Dynamics: the Case of Expansions. In *Proc. KR*, Pierre Marquis, Tran Cao Son, and Gabriele Kern-Isberner (Eds.). ijcai.org, 562–571.

[24] Henry Prakken and Michiel De Winter. 2018. Abstraction in Argumentation: Necessary but Dangerous. In *Proc. COMMA (Frontiers in Artificial Intelligence and Applications, Vol. 305)*, Sanjay Modgil, Katarzyna Budzynska, and John Lawrence (Eds.). IOS Press, 85–96.

[25] Anna Rapberger and Markus Ulbricht. 2023. On Dynamics in Structured Argumentation Formalisms. *J. Artif. Intell. Res.* 77 (2023), 563–643.

[26] Chiaki Sakama. 2018. Abduction in argumentation frameworks. *J. Appl. Non Class. Logics* 28, 2-3 (2018), 218–239.

[27] Zeynep G. Saribatur and Johannes P. Wallner. 2021. Existential Abstraction on Argumentation Frameworks via Clustering. In *Proc. KR*, Meghyn Bienvenu, Gerhard Lakemeyer, and Esra Erdem (Eds.). ijcai.org, 549–559.

[28] Markus Ulbricht and Johannes P. Wallner. 2022. Strongly Accepting Subframeworks: Connecting Abstract and Structured Argumentation. In *Proc. COMMA (Frontiers in Artificial Intelligence and Applications, Vol. 353)*, Francesca Toni, Sylwia Polberg, Richard Booth, Martin Caminada, and Hiroyuki Kido (Eds.). IOS Press, 320–331.

[29] Johannes P. Wallner. 2020. Structural constraints for dynamic operators in abstract argumentation. *Argument Comput.* 11, 1-2 (2020), 151–190.