Formal Argumentation in Symbolic AI

Markus Ulbricht
ScaDS.AI Dresden/Leipzig, Leipzig Universtiy, Germany
mulbricht@informatik.uni-leipzig.de

Abstract

In the area of symbolic AI, researchers strive to develop techniques to teach machines (commonsense) reasoning. Human reasoning is often argumentative in its nature,and consequently, computational models of argumentation constitute a vibrant research area in symbolic AI.In this paper I describe my most significant contributions to the field spanning from general non-monotonic logics to formal argumentation.

1 Inconsistency in Non-Monotonic Logics

Statistical machine learning, in particular the development of artificial neural networks, has produced a plethora of scientific breakthroughs within the last years, with the most recent ones being generative models like GPT. However, it is commonly agreed in the AI community that up to now, those models lack proper reasoning capabilities which leads to nonsensical outputs of language models or images of physically impossible situations produced by e.g. diffusion models.

While statistical machine learning techniques are specialized in finding patterns in data, the research area of knowledge representation and reasoning (KR) is driven by the goal to teach machines common sense reasoning and deriving conclusions on their own. A vital observation in this context is that classical propositional logics are too restrictive in their nature since in many real world scenarios, one has to draw conclusions from unknown, fuzzy, or even inconsistent knowledge. Moreover, commonsense reasoning is oftentimes *non-monotonic*, that is, one might believe something to be true, but withdraws this conclusion as soon as some novel information is learned. In propositional logic, this behavior cannot be modeled.

In order to overcome these issues, researchers propose and study various logics and (non-monotonic) formalisms, tailored to model different real world scenarios [Van Harmelen *et al.*, 2008]. Thereby, the trade-off underlying any KR formalism is the expressive power vs. the computational complexity of reasoning. The goal is to find formalisms that are capable of expressing involved reasoning tasks, while ensuring that nonetheless, drawing conclusions in an automated way can then be performed efficiently.

Independent of the formalism under consideration, an important issue is the handling of *inconsistent* information. Suppose we are given a knowledge base $\mathcal{K}$ (for instance, $\mathcal{K}$ could be a set of inference rules). Assuming $\mathcal{K}$ models some real world knowledge, has undergone updates, or represents the aggregated beliefs of different agents, $\mathcal{K}$ might contain conflicting information, i.e., it entails both an atom p and its contrary $\neg p$. In classical propositional logic, this would cause $\mathcal{K}$ to collapse entirely, i.e., by the *principle of explosion*, any atom can now be deduced from $\mathcal{K}$. This is, however, not feasible in commonsense reasoning; a conflict in a subset of the knowledge base is usually no reason to neglect *any* information encoded in $\mathcal{K}$. For instance, if a database contains two different phone numbers of Alice, would we then assume that all information stored about Alice is wrong?

Striving to tackle this issue, in 1987 Raymond Reiter published his seminal article entitled *A theory of diagnosis from first principles* [Reiter, 1987]. Here, Reiter studies how to identify reasons for misbehavior in logical systems. One of the main results of this article is his so-called *hitting set duality*. It states that any $\subseteq$-maximal consistent subset of a knowledge base $\mathcal{K}$ can be computed by removing a $\subseteq$-minimal hitting set of the $\subseteq$-minimal inconsistent subsets of $\mathcal{K}$[1]. Reiter studied this issue in a quite specific setting, but it is folklore that this result generalizes to arbitrary monotonic logics.

The main subject of investigation during my PhD studies was handling inconsistency in non-monotonic logics. One of the central observations was that Reiter's hitting set duality can be extended to this more general setting when using the notion of *strong inconsistency* [Brewka *et al.*, 2019]. We call a subset $\mathcal{H} \subseteq \mathcal{K}$ strongly inconsistent if not just $\mathcal{H}$ itself is inconsistent, but also each superset up until the entire knowledge base $\mathcal{K}$; intuitively, strong inconsistency adapts ordinary inconsistency by taking the non-monotonic behavior of $\mathcal{K}$ into account. We studied this notion and its consequences thoroughly [Brewka *et al.*, 2019; Ulbricht, 2019]. Moreover, we generalize research on inconsistency measurement [Thimm, 2019] to non-monotonic logics [Ulbricht *et al.*, 2016; Ulbricht *et al.*, 2020]. In this area, the goal is to quantitatively assess the severity of inconsistency in a knowledge base which provides deeper insights into the nature of and culprits for the semantical collapse.

[1] $\mathcal{H}$ is a *hitting set* of a set $\mathcal{S}$ of sets if $\mathcal{H} \cap S \neq \emptyset$ for each $S \in \mathcal{S}$.

2 Abstract Argumentation

Human reasoning, in particular resolving conflicts among agents with different beliefs can oftentimes be seen as *argumentative* [Antaki and Leudar, 1992]. Researchers thus study computational models of argumentation in order to explicate, resolve, and reason in the presence of conflicting information. Formal argumentation [Baroni *et al.*, 2018b; Bench-Capon and Dunne, 2007] is a flourishing research area and can nowadays be seen as one of the classical fields in AI.

2.1 Dung's Abstract Argumentation

A main booster for research on abstract argumentation was Phan Minh Dung's seminal 1995 paper [Dung, 1995] where he introduces *abstract argumentation frameworks* (AFs). Dung abstracts away certain parts of the argumentation procedure, such as the internal structure of arguments derived from a debate or premises required to infer a certain conclusion: In Dung's AFs, arguments are modeled as atomic entities $a \in A$ and the conflicts between them as a binary relation R, amounting to the representation of an AF as a directed graph $F = (A, R)$.

Definition 2.1. An (AF) [Dung, 1995] is a directed graph $F = (A, R)$ where A is a set (of *arguments*) and $R \subseteq A \times A$ models *attacks* between them.

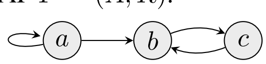
In order to evaluate a discussion using AFs, researchers investigate so-called *semantics* σ whose purpose it is to formalize jointly acceptable sets of arguments. Intuitively, each set $E \in \sigma(F)$ of arguments is considered a reasonable point of view w.r.t. the debate represented by F. The vast majority of semantics which are agreed upon by the AF community are based on the concept of *admissibility* (adm). This notion formalizes that arguments within an acceptable set $E \in adm(F)$ should not have conflicts among them and be able to refute objections against E.

Definition 2.2. Let $F = (A, R)$ be an AF. A set $E \subseteq A$ of arguments is called *admissible*, $E \in adm(F)$, whenever the following two conditions hold.

- If $a, b \in E$, then $(a, b) \notin R$. E is conflict-free
- If $(a, b) \in R$ for some $b \in E$, then there is some $c \in E$ with $(c, a) \in R$. E defends itself

Let us illustrate this notion in the following example.

Example 2.3. Consider the directed graph depicted below which represents an AF $F = (A, R)$.

F : 

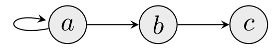
Take the singleton $E = \{c\}$. The edges in F represent attacks among arguments, so we see that E has no internal conflict (c does not attack itself). The only attacker of E is b, and as c counter-attacks this threat, we infer $E \in adm(F)$.

Dung's semantics [Dung, 1995], which are nowadays seen as the classical ones, augment admissibility with further requirements, for instance so-called *preferred* semantics maximize the admissible sets. AFs have been thoroughly investigated ever since [Baroni *et al.*, 2018b] and therefore provide a solid formal groundwork for argumentative reasoning approaches to build upon them.

2.2 Weak Admissibility

Perhaps our most significant contribution to abstract argumentation research was our proposal of the semantics family based on *weak admissibility*. Towards illustrating this notion, lead us recall our Example 2.3 in a slightly modified version.

Example 2.4. Consider the AF G as depicted below.

G : 

In G, c does not refute the attack from b anymore, so it is not admissible this time. So let us consider a: it attacks itself and can thus not be accepted, either. Since no other argument in G attacks it, a also blocks acceptance of b. In summary, the only admissible set here is the empty set, i.e., $adm(G) = \{\emptyset\}$.

In our paper on repairing AFs we investigated under which circumstances a semantical collapse as in the above AF G can be resolved [Baumann and Ulbricht, 2019]. Applied to this particular graph, our study amounts to the intuitive result that removing the argument a from the graph resolves the aforementioned issue. However, this requires us to alter the debate which is represented by G, because we have to manually delete an argument. A better solution would be to utilize semantics that can handle such situations on their own.

Let us therefore inspect the situation a bit closer. The only reason for b to be non-acceptable is the paradoxical a, yet arguably, a self-attacker like a should not be able to refute reasonable arguments. Interestingly, Dung already made a similar observation in his seminal 1995 paper [Dung, 1995]:

> "An interesting topic of research is the problem of self-defeating arguments as illustrated in the following example. Consider the argumentation framework $\big(\{A, B\}, \{(A, A), (A, B)\}\big)$. The only preferred extension here is empty though one can argue that since A defeats itself, B should be acceptable."

Although a plethora of different AF semantics has been proposed by the AF community within the last decades [Baroni *et al.*, 2018a], for the most time, no commonly agreed solution to this problem was available. Striving to tackle this issue, in 2020 we proposed a novel family of semantics based on so-called *weak admissibility* and *weak defense* [Baumann *et al.*, 2020]. Based on a recursive definition, our weak admissibility notion distinguishes between "serious" arguments whose attacks need to be refuted and "non-serious" ones that we can neglect. This induces a liberalization of Dung's classical defense avoiding the aforementioned problematic behavior; one of the core features of weak admissibility is that paradoxical self-attacking arguments can be ignored.

Formally, weak admissibility is based on the *reduct* of an AF $F = (A, R)$: Given a set $E \subseteq A$ of arguments, the reduct F^E is a tool to partially evaluate F. It restricts F to those arguments that are neither contained in nor attacked by E. Intuitively, F^E partially evaluates F by setting the arguments in E to "true", those attacked by E to "false" and then focuses on the remaining ones only. Based on this, weak admissibility recursively decides which arguments are serious and which are not, as explained above.

Definition 2.5. Let $F = (A, R)$ be an AF. A set $E \subseteq A$ of arguments is called *weakly admissible*, $E \in adm^w(F)$, whenever the following two conditions hold.

- If $a, b \in E$, then $(a, b) \notin R$. E is conflict-free
- If $(a, b) \in R$ for some $b \in E$, then $a \notin \bigcup adm^w(F^E)$. Attackers of E are not serious enough to be a threat

As the reader may verify, in the above G, the singleton $E' = \{b\}$ is indeed weakly admissible, as we desired.

Let us report two core features of weak admissibility. First of all, weak admissibility faithfully generalizes Dung's classical admissibility.

Proposition 2.6. *[Baumann* et al.*, 2022a] For any AF F, it holds that $adm(F) \subseteq adm^w(F)$.*

We recall that the motivation underlying weak admissibility is that we strive to neglect the impact of paradoxical arguments. Indeed, adm^w ignores self-attackers entirely.

Proposition 2.7. *[Baumann* et al.*, 2022a] Let F be an AF and F° be F after each self-attacker is removed. Then it holds that $adm^w(F) = adm^w(F^\circ)$.*

We discuss further fundamental properties of weak admissibility including a comprehensive comparison to related work in [Baumann *et al.*, 2022a; Blümel and Ulbricht, 2022].

Perhaps the most pressing future work direction consists in finding feasible techniques to implement solvers that compute weakly acceptable arguments. In our study [Dvorák *et al.*, 2022] we could show that almost all problems in this context are PSPACE-complete, i.e., reasoning with weak admissibility is highly intractable. Finding efficient implementations is thus a challenging endeavor for future research.

3 Structured Argumentation

In the research area of *structured argumentation*, arguments are systematically derived from a given knowledge base [Baroni *et al.*, 2018b]. In this so-called *instantiation* procedure, different conclusions as well as their conflicts within the knowledge base are made explicit by constructing an associated AF. This way, many well-established KR formalisms can be captured by AFs and thus enjoy the considerable research which is already available for them out of the box. Structured argumentation formalisms have been studied extensively [Baroni *et al.*, 2018b] with prominent examples being ASPIC$^+$ [Modgil and Prakken, 2013], defeasible logic programming (DeLP) [García and Simari, 2004], and deductive argumentation [Besnard and Hunter, 2001]. Most research we conduct on structured argumentation focuses on assumption-based argumentation (ABA) [Bondarenko *et al.*, 1997], a popular and well-investigated formalism.

3.1 Assumption-Based Argumentation

An ABA framework (ABAF) is a tuple $\mathcal{K} = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$ where $(\mathcal{L}, \mathcal{R})$ is a deductive system, that is, $\mathcal{L}$ is a language and $\mathcal{R}$ a set of inference rules of the form $p_0 \leftarrow p_1, \ldots, p_n$ where all p_i are in $\mathcal{L}$. The intuitive meaning of a rule is that the *head* p_0 is inferred from the conjunction $p_1, \ldots, p_n$. Moreover, $\mathcal{A} \subseteq \mathcal{L}$ is a set of defeasible assumptions, and the mapping $^- : \mathcal{A} \to \mathcal{L}$ formalizes their *contraries*.

A commonly studied ABA fragment are so-called *flat* ABAFs where assumptions cannot be entailed, only assumed to hold or not. However, researchers also investigate more expressive ABA classes like non-flat ABA [Bondarenko *et al.*, 1997] or so-called ABA$^+$ in order to model preferences among the given assumptions [Cyras and Toni, 2016].

The semantics of flat ABAFs resemble those of AFs in considering the interaction of sets of assumptions. For $S \subseteq \mathcal{A}$ we write $S \vdash p$ if we can successively infer p from the assumptions in S via the rules in $\mathcal{R}$. If it holds that $S \vdash p$ and p is the contrary of some assumption in $a \in T$, i.e., $p = \overline{a}$, then S we say *attacks* T. As in the case of AFs, S is admissible if S does not attack itself and counter-attacks any set T of assumptions which in turn attacks S.

Definition 3.1. Let $\mathcal{K} = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$ be an ABAF. A set $S \subseteq \mathcal{A}$ of assumptions is called *admissible*, $E \in adm(\mathcal{K})$, whenever the following two conditions hold.

- If $S \vdash p$ and $p = \overline{a}$, then $a \notin S$. E is conflict-free
- If T attacks S, then S also attacks T. E defends itself

Example 3.2. Consider the ABAF $\mathcal{K} = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$ where $\mathcal{L} = \{a, b, c, b_c, c_c\}$, $\mathcal{A} = \{a, b, c\}$, $\overline{a} = a$, $\overline{b} = b_c$, $\overline{c} = c_c$, and $\mathcal{R}$ is the following set of rules:

$$r_1 \ : b_c \leftarrow a. \qquad r_2 : \ c_c \leftarrow b. \qquad r_3 : \ b_c \leftarrow c.$$

This ABAF is similar in spirit to the AF F from before and yields e.g. $\{c\} \in adm(\mathcal{K})$ as the reader might verify.

ABA and AFs are similar in their spirit, so each flat ABAF $\mathcal{K}$ can be translated into a semantics-preserving AF $F_\mathcal{K}$, a procedure which is referred to as *instantiation* in the argumentation community [Baroni *et al.*, 2018b]. Not just ABA, but many KR formalisms can be captured by AFs this way.

3.2 Expressive Instantiations and Implementations

Instantiations of this kind are a popular technique in argumentation research: the constructed graph provides the user with a comprehensible graphical depiction of the conflicts arising in the given knowledge base. Moreover, the plethora of research which is available for AFs can then be applied to the knowledge base $\mathcal{K}$ that has been translated into the AF $F_\mathcal{K}$.

There are, however, two issues which are common for such instantiations:

- Not all structured argumentation formalisms can be captured by means of an AF, because they have features AFs cannot model. The aforementioned advantages are thus not available in those cases.
- Even if the knowledge base can be captured by an AF, the instantiated graph is oftentimes unacceptably large. This diminishes the advantages of the whole procedure due to the sheer number of constructed arguments.

Our most recent research strongly focuses on these issues. Striving to overcome them, we first observe that different types of ABA frameworks have different modeling capabilities, as we studied in [Berthold *et al.*, 2023b]. Awareness of their expressiveness helps us in identifying the features necessary to capture such knowledge bases. As a result, we utilized the following expressive abstract argumentation formalisms.

Collective Argumentation: An important tool are argumentation formalisms that allow for attacks between sets of arguments. For instance, *argumentation frameworks with collective attacks* (SETAFs) [Nielsen and Parsons, 2006] can model situations where several arguments are necessary to defeat another one. Moreover, so-called HYPAFs can even model set-to-set attacks among arguments [Gabbay and Gabbay, 2016].

Supporting Arguments: Another extension of AFs are so-called *bipolar argumentation frameworks* (BAFs) [Cayrol and Lagasquie-Schiex, 2005]. In addition to attacks, BAFs can also model arguments supporting each other.

As it turns out, these features – collective attacks and supports between arguments – are the missing building blocks in order to capture popular, expressive ABA formalisms.

Theorem 3.3. *There is a semantics-preserving translation*

- *from non-flat ABA to BAFs [Ulbricht* et al.*, 2024];*
- *from ABA+ to HYPAFs [Dimopoulos* et al.*, 2024].*

Moreover, even existing instantiation techniques can be enhanced. While the known translation from flat ABA to AFs yields infinitely many arguments in general, we can capture any flat ABA knowledge base $\mathcal{K}$ with a SETAF that has only linearly many arguments (in $|\mathcal{K}|$).

Theorem 3.4. *There is a semantics-preserving translation*

- *from flat ABA to SETAFs [König* et al.*, 2022]*

admitting linearly many arguments.

Perhaps the most surprising observation is that in many of these cases, the constructed argumentation graph admits a lower computational complexity compared to the initial knowledge base. We study this thoroughly in the respective works [Ulbricht *et al.*, 2024; Dimopoulos *et al.*, 2024]. Hence, there is a computational cost for constructing the graph (BAF resp. HYPAF), but once it has been paid, many reasoning problems can be solved more efficiently. We thus implement instantiation-based solvers for non-flat ABA [Lethonen *et al.*, 2024] and indeed, the resulting solver is competitive with state-of-the-art approaches which compute accepted assumptions on the knowledge base $\mathcal{K}$ directly.

While these results promote the usage of expressive argumentation formalisms to capture ABAFs, there is also a significant potential in enhancing the efficiently of AF-based instantiations. We showed that any flat ABAF $\mathcal{K}$ can be translated into an AF $F_{\mathcal{K}}$ of polynomial size (in $\mathcal{K}$) if the knowledge base $\mathcal{K}$ is suitably preprocessed [Lehtonen *et al.*, 2023]. This approach also resulted in an instantiation-based solver [Lehtonen *et al.*, 2021]. We also want to mention our approach [Anh and Ulbricht, 2024] which operates on an auxiliary graph in order to compute acceptable assumptions.

These results are encouraging steps towards our overall goal of establishing a stronger connection between expressive structured and abstract argumentation. Nonetheless, we still need better means to compute argumentation graphs more efficiently, in order to enhance instantiation-based solvers even further. Many of the aforementioned formalisms need to be studied more thoroughly in terms of their computational properties, similar in spirit to our study of SETAFs [Dvořák *et al.*, 2024]. This is necessary to lay the foundations for more advanced instantiation techniques.

3.3 Dynamics and Structured Argumentation

Instantiations as discussed in the previous subsection are tailored for static reasoning environments, i.e., situations where the given knowledge base $\mathcal{K}$ persists. However, exchanging arguments is a continually evolving process, which inspired the investigation of dynamics, i.e., knowledge bases that change over time [Gabbay *et al.*, 2021]. For instance, researchers study situations where further arguments are brought forward in a debate or additional rules are added.

However, the connection between an ABAF $\mathcal{K}$ and its constructed AF $F_{\mathcal{K}}$ is not close enough to handle situations of this kind [Prakken, 2023]. Indeed, in our studies dealing with dynamic reasoning tasks in ABA, we realized that we had to produce results from scratch [Rapberger and Ulbricht, 2023; Berthold *et al.*, 2023a] despite the close correspondence between ABA and AFs as well as the fact that these problems had been studied for AFs before.

Driven by this observation, our goal was to come up with an instantiation technique that is also suitable for dynamic reasoning, i.e., we strive to equip arguments with additional information in order to better anticipate their role after the underlying knowledge base is extended. The appropriate tool are so-called *semi-structured* argumentation formalisms [Baumann *et al.*, 2022b; Rapberger and Ulbricht, 2023]. For instance, our idea in [Rapberger and Ulbricht, 2023] is to consider an extension of AFs, called *claim and vulnerability augmented AFs* (cvAFs), where each argument is augmented with the conclusion it represents as well as the premises which are necessary to entail it. In this study, we could show that this instantiation technique provides sufficient information in order to capture a significant subclass of ABA in both static and dynamic reasoning environments. In addition to its conceptual significance, our study provides insights towards computational improvements for solving dynamic tasks in ABA. Moreover, we could showcase that our methodology also scales to other structured formalisms.

The results we established thus far are, however, quite theoretical and the actual implementation of efficient solvers for dynamic reasoning building upon our techniques is still open. Moreover, our cvAFs extend AFs as proposed by Dung, but as we saw in the previous subsection, expressive argumentation formalisms like SETAFs or BAFs provide us with significant improvements when connecting abstract and structured argumentation. This potential can only be fully utilized when studying more expressive semi-structured formalisms.

4 Conclusion

Symbolic AI techniques provide transparency, robustness, and interpretability in decision-making processes, and thus complement statistical machine learning approaches. Formal argumentation in particular plays a central role as it can capture various KR formalisms, thereby providing a descriptive, visual, and argumentative representation of the encoded knowledge. The foundational work outlined in this paper establishes vital theoretical underpinnings for harnessing the full potential of formal argumentation. This way, we push the developments of argumentation, which promises advances for future challenges within its rich application areas.

References

[Anh and Ulbricht, 2024] Kiet Ngyuen Anh and Markus Ulbricht. Preferred reasoning in aba by cycle-breaking. In *Proc. IJCAI*, 2024. To appear.

[Antaki and Leudar, 1992] Charles Antaki and Ivan Leudar. Explaining in conversation: Towards an argument model. *European Journal of Social Psychology*, 22(2):181–194, 1992.

[Baroni *et al.*, 2018a] Pietro Baroni, Martin Caminada, and Massimiliano Giacomin. Abstract argumentation frameworks and their semantics. In *Handbook of Formal Argumentation*. College Publications, 2018.

[Baroni *et al.*, 2018b] Pietro Baroni, Dov Gabbay, Massimiliano Giacomin, and Leendert van der Torre, editors. *Handbook of Formal Argumentation*. College Publications, 2018.

[Baumann and Ulbricht, 2019] Ringo Baumann and Markus Ulbricht. If nothing is accepted - repairing argumentation frameworks. *J. Artif. Intell. Res.*, 66:1099–1145, 2019.

[Baumann *et al.*, 2020] Ringo Baumann, Gerhard Brewka, and Markus Ulbricht. Revisiting the foundations of abstract argumentation - semantics based on weak admissibility and weak defense. In *Proc. AAAI*, pages 2742–2749. AAAI Press, 2020.

[Baumann *et al.*, 2022a] Ringo Baumann, Gerhard Brewka, and Markus Ulbricht. Shedding new light on the foundations of abstract argumentation: Modularization and weak admissibility. *Artif. Intell.*, 310:103742, 2022.

[Baumann *et al.*, 2022b] Ringo Baumann, Anna Rapberger, and Markus Ulbricht. Equivalence in argumentation frameworks with a claim-centric view - classical results with novel ingredients. In *Proc. AAAI*, pages 5479–5486, 2022.

[Bench-Capon and Dunne, 2007] Trevor J. M. Bench-Capon and Paul E. Dunne. Argumentation in artificial intelligence. *Artif. Intell.*, 171(10-15):619–641, 2007.

[Berthold *et al.*, 2023a] Matti Berthold, Anna Rapberger, and Markus Ulbricht. Forgetting aspects in assumption-based argumentation. In *Proc. KR*, pages 86–96, 2023.

[Berthold *et al.*, 2023b] Matti Berthold, Anna Rapberger, and Markus Ulbricht. On the expressive power of assumption-based argumentation. In *Proc. JELIA*, volume 14281 of *Lecture Notes in Computer Science*, pages 145–160. Springer, 2023.

[Besnard and Hunter, 2001] Philippe Besnard and Anthony Hunter. A logic-based theory of deductive arguments. *Artif. Intell.*, 128(1-2):203–235, 2001.

[Blümel and Ulbricht, 2022] Lydia Blümel and Markus Ulbricht. Defining defense and defeat in abstract argumentation from scratch - A generalizing approach. In *Proc. KR*, 2022.

[Bondarenko *et al.*, 1997] Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artif. Intell.*, 93:63–101, 1997.

[Brewka *et al.*, 2019] Gerhard Brewka, Matthias Thimm, and Markus Ulbricht. Strong inconsistency. *Artif. Intell.*, 267:78–117, 2019.

[Cayrol and Lagasquie-Schiex, 2005] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *Proc. ECSQARU*, pages 378–389, 2005.

[Cyras and Toni, 2016] Kristijonas Cyras and Francesca Toni. Proc. kr. pages 553–556, 2016.

[Dimopoulos *et al.*, 2024] Yannis Dimopoulos, Wolfgang Dvorák, Matthias König, Anna Rapberger, Markus Ulbricht, and Stefan Woltran. Redefining ABA+ semantics via abstract set-to-set attacks. In *Proc. AAAI*, pages 10493–10500. AAAI Press, 2024.

[Dung, 1995] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, 77(2):321–358, 1995.

[Dvorák *et al.*, 2022] Wolfgang Dvorák, Markus Ulbricht, and Stefan Woltran. Recursion in abstract argumentation is hard - on the complexity of semantics based on weak admissibility. *J. Artif. Intell. Res.*, 74:1403–1447, 2022.

[Dvorák *et al.*, 2024] Wolfgang Dvorák, Matthias König, Markus Ulbricht, and Stefan Woltran. Principles and their computational consequences for argumentation frameworks with collective attacks. *J. Artif. Intell. Res.*, 79:69–136, 2024.

[Gabbay and Gabbay, 2016] Dov M. Gabbay and Michael Gabbay. Theory of disjunctive attacks, part I. *Log. J. IGPL*, 24(2):186–218, 2016.

[Gabbay *et al.*, 2021] Dov Gabbay, Massimiliano Giacomin, Guillermo R. Simari, and Matthias Thimm, editors. *Handbook of Formal Argumentation*, volume 2. College Publications, 2021.

[García and Simari, 2004] Alejandro Javier García and Guillermo Ricardo Simari. Defeasible logic programming: An argumentative approach. *Theory Pract. Log. Program.*, 4(1-2):95–138, 2004.

[König *et al.*, 2022] Matthias König, Anna Rapberger, and Markus Ulbricht. Just a matter of perspective. In *Proc. COMMA*, pages 212–223. IOS Press, 2022.

[Lehtonen *et al.*, 2021] Tuomo Lehtonen, Anna Rapberger, Markus Ulbricht, and Johannes P Wallner. Acbar–atomic-based argumentation solver. *ICCMA 2023*, 42(3):16, 2021.

[Lehtonen *et al.*, 2023] Tuomo Lehtonen, Anna Rapberger, Markus Ulbricht, and Johannes P. Wallner. Argumentation frameworks induced by assumption-based argumentation: Relating size and complexity. In *Proc. KR*, pages 440–450. IJCAI, 2023.

[Lethonen *et al.*, 2024] Tuomo Lethonen, Anna Rapberger, Francesca Toni, Markus Ulbricht, and Johannes P. Wallner. Instantiations and computational aspects of non-flat assumption-based argumentation. In *Proc. IJCAI*, 2024. To appear.

[Modgil and Prakken, 2013] Sanjay Modgil and Henry Prakken. A general account of argumentation with preferences. *Artif. Intell.*, 195:361–397, 2013.

[Nielsen and Parsons, 2006] Søren Holbech Nielsen and Simon Parsons. A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *Proc. ArgMAS 2006*, volume 4766 of *Lecture Notes in Computer Science*, pages 54–73. Springer, 2006.

[Prakken, 2023] Henry Prakken. Relating abstract and structured accounts of argumentation dynamics: the case of expansions. In *Proc. KR*, pages 562–571, 2023.

[Rapberger and Ulbricht, 2023] Anna Rapberger and Markus Ulbricht. On dynamics in structured argumentation formalisms. *J. Artif. Intell. Res.*, 77:563–643, 2023.

[Reiter, 1987] Raymond Reiter. A theory of diagnosis from first principles. *Artif. Intell.*, 32(1):57–95, 1987.

[Thimm, 2019] Matthias Thimm. Inconsistency measurement. In *Proc. SUM*, volume 11940 of *Lecture Notes in Computer Science*, pages 9–23. Springer, 2019.

[Ulbricht *et al.*, 2016] Markus Ulbricht, Matthias Thimm, and Gerhard Brewka. Measuring inconsistency in answer set programs. In *Proc. JELIA*, pages 577–583, 2016.

[Ulbricht *et al.*, 2020] Markus Ulbricht, Matthias Thimm, and Gerhard Brewka. Handling and measuring inconsistency in non-monotonic logics. *Artif. Intell.*, 286:103344, 2020.

[Ulbricht *et al.*, 2024] Markus Ulbricht, Nico Potyka, Anna Rapberger, and Francesca Toni. Non-flat ABA is an instance of bipolar argumentation. In *Proc. AAAI*, pages 10723–10731. AAAI Press, 2024.

[Ulbricht, 2019] Markus Ulbricht. Repairing non-monotonic knowledge bases. In *Proc. JELIA*, pages 151–167, 2019.

[Van Harmelen *et al.*, 2008] Frank Van Harmelen, Vladimir Lifschitz, and Bruce Porter. *Handbook of knowledge representation*. Elsevier, 2008.