Sets Attacking Sets in Abstract Argumentation

Yannis Dimopoulos[1], Wolfgang Dvořák[2], Matthias König[2,*], Anna Rapberger[2], Markus Ulbricht[3] and Stefan Woltran[2]

[1]University of Cyprus, Department of Computer Science
[2]TU Wien, Institute of Logic and Computation
[3]Leipzig University, Department of Computer Science, Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI)

Abstract
In abstract argumentation, arguments jointly attacking single arguments is a well-understood concept, captured by the established notion of SETAFs—argumentation frameworks with collective attacks. In contrast, the idea of sets attacking other sets of arguments has not received much attention so far. In this work, we contribute to the development of set-to-set defeat in formal argumentation. To this end, we introduce so called hyper argumentation frameworks (HYPAFs), a new formalism that extends SETAFs by allowing for set-to-set attacks. We investigate this notion by interpreting these novel attacks in terms of universal, indeterministic, and collective defeat. We will see that universal defeat can be naturally captured by the already existing SETAFs. While this is not the case for indeterministic defeat, we show a close connection to attack-incomplete argumentation frameworks. To formalize our interpretation of collective defeat, we develop novel semantics yielding a natural generalization of attacks between arguments to set-to-set attacks. We investigate fundamental properties and identify several surprising obstacles; for instance, the well-known fundamental lemma is violated, and the grounded extension might not exist. Finally, we investigate the computational complexity of the thereby arising problems.

Keywords
Abstract Argumentation, Collective Attack, Indeterminisim

1. Introduction

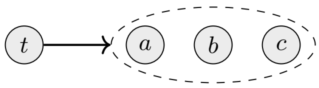
Formal argumentation is a major research area in knowledge representation and reasoning, with applications in various fields in the realm of Artificial Intelligence. The most popular formalism in the abstract setting are Argumentation Frameworks (AFs) due to Dung [1], where arguments are modeled as the nodes of a directed graph while the edges are interpreted as attacks. As oftentimes the use of *sets* instead of singular attackers comes handy, generalizations have been proposed—most notably, *collective attacks* [2]. These frameworks (referred to as SETAFs) have recently been in the focus of researchers, see e.g., [3, 4]. SETAFs, however, are restricted in the sense that a set of arguments can attack only a single argument.

The natural counter-part, namely allowing attacks between sets of arguments, has not yet been widely studied. However, some preliminary considerations were performed, e.g. in [5] where defeat is modeled not based on directed graphs but using rule-like statements; in [6] with the aim of formalizing global conflicts; Nielsen and Parsons [2] reduce these phenomena to SETAFs; and by Gabbay and Gabbay [7] who investigate (among other notions) cases where the attacking set applies conjunctively and the attacked set is understood disjunctively.

In this work, we provide the first thorough analysis of this setting. Naturally, the question arises of how to interpret an attack from a set A of arguments to another set B of arguments? We investigate three natural notions that capture different motivations: (i) *universal defeat*, i.e., accepting each $a \in A$ defeats all $b \in B$: we argue that this amounts to merely a simplification of the representation of SETAFs. (ii) *indeterministic defeat*, i.e., we model the situation where it is unknown which subset of B is attacked by A. Hence, the main motivation for this concept is to model incomplete information of an agent's knowledge base. Consequently, we show a close connection to attack-incomplete frameworks [8, 9]. (iii) *collective defeat*, i.e., we consistently generalize Dung's notions of attack and defense to be applicable to sets of arguments and investigate the emerging properties. The study of collective defeat is motivated by corresponding phenomena in structured argumentation [10, 11], where it is conceivable that a set A of arguments contradicts the conjunction of the supports of a set B of arguments, but not necessarily the support of each $b \in B$. This would result in the attack (A, B), which is not natively featured in classical (Dung-style) frameworks.

21st International Workshop on Nonmonotonic Reasoning, September 2–4, 2023, Rhodes, Greece
*Corresponding author.
yannis@ucy.ac.cy (Y. Dimopoulos); wolfgang.dvorak@tuwien.ac.at (W. Dvořák); matthias.koenig@tuwien.ac.at (M. König); anna.rapberger@tuwien.ac.at (A. Rapberger); mulbricht@informatik.uni-leipzig.de (M. Ulbricht); stefan.woltran@tuwien.ac.at (S. Woltran)
0000-0001-9583-9754 (Y. Dimopoulos); 0000-0002-2269-8193 (W. Dvořák); 0000-0003-0205-0039 (M. König); 0000-0003-0355-3535 (A. Rapberger); 0000-0002-0773-7510 (M. Ulbricht); 0000-0003-1594-8972 (S. Woltran)

Example 1. *To illustrate the concept of collective defeat, let us consider a situation in which three agents, Alice, Bob, and Carol, plan a tandem-trip. At most two of them can join the tandem, but not all of them at once. It could either be that* Alice rides tandem (a), Bob rides tandem (b), *or* Carol rides tandem (c). *However, we know that the bicycle is a* two-person tandem (t). *Utilizing collective defeat, we can depict the conflict between these statements as follows:*



The set $\{a, b, c\}$ *is* collectively defeated *by* t*. The intuition is that none of the subsets of* $\{a, b, c\}$ *is affected by the attack from* t*, but only the collection of the arguments is attacked. We can safely accept each proper subset of the set of all arguments* $\{a, b, c, t\}$*; the* $\subseteq$*-maximal acceptable sets* $\{a, b, t\}$*,* $\{a, c, t\}$*, and* $\{b, c, t\}$ *model the outcome in which exactly two of our agents enjoy their tandem-ride.*

After briefly recalling the basic notions of SETAFs and formally introducing our HYPAFs (Section 2) we discuss the three defeat-modes, namely the simple case of universal defeat (Section 3), indeterministic defeat (Section 4), and collective defeat (Section 5). Finally, we conclude in Section 7.

2. Argumentation and Set Attacks

In this section we briefly recall the definitions relevant to SETAFs (argumentation frameworks with collective attacks) and introduce our hyperframeworks (HYPAFs).

2.1. Collective Attacks (SETAFs)

Argumentation Frameworks with Collective Attacks (SETAFs) were introduced by Nielsen and Parsons [2] as a generalization of Dung's AFs [1].

Definition 2 (SETAFs)**.** *A SETAF is a pair* $SF = (A, R)$ *where* A *is a finite set of* arguments, *and* $R \subseteq 2^A \times A$ *is the* attack relation[1].

SETAFs $SF = (A, R)$, where for all $(T, h) \in R$ it holds that $|T| = 1$, amount to (standard Dung) AFs. We usually write (t, h) to denote the set-attack $(\{t\}, h)$.

An attack $(\{a_1, \ldots, a_n\}, b)$ is interpreted as follows: if we accept *all of* $a_1, \ldots, a_n$ then b is defeated. In order to defend b against this attack, it thus suffices to defeat *one of* $a_1, \ldots, a_n$. Based on this intuition, the classical Dung-semantics generalize as follows (for a recent overview, see e.g. [3, 12]).

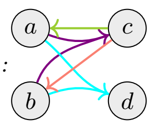
[1]While the original definition of SETAFs from Nielsen and Parsons [2] does not allow attacks of the form $(\emptyset, a)$, these attacks are often included for convenience.

Definition 3. *Let* $SF = (A, R)$ *be a SETAF and* $E \subseteq A$ *a set of arguments. Then* E *is* conflict-free *if for all* $(T, h) \in R$ *it holds* $T \subseteq E \Rightarrow h \notin E$*. An argument* $a \in A$ *is* defended *(in* SF*) by a set* $S \subseteq A$ *if for each* $B \subseteq A$*, such that* B *attacks* a*, also some* $S' \subseteq S$ *attacks some* $b \in B$*. A set* $T \subseteq A$ *is defended (in* SF*) by* S *if each* $a \in T$ *is defended by* S *(in* SF*). Let* S *be conflict-free in* SF*, then:*

- $S \in adm(SF)$*, if* S *defends itself in* SF*,*
- $S \in com(SF)$*, if* $S \in adm(SF)$ *and* $a \in S$ *for all* $a \in A$ *defended by* S*,*
- $S \in grd(SF)$*, if* $S = \bigcap_{T \in com(SF)} T$*,*
- $S \in pref(SF)$*, if* $S \in adm(SF)$ *and* $\nexists T \in adm(SF)$ *s.t.* $T \supset S$*, and*
- $S \in stb(SF)$*, if* S *attacks* a *for all* $a \in A \setminus S$*.*

Example 4. *Consider the SETAF* SF_4 *and its extensions.*

SF_4:



$adm(SF_4) = \{\emptyset, \{a, b\}, \{c\}, \{c, d\}\}$
$com(SF_4) = \{\emptyset, \{a, b\}, \{c, d\}\}$
$grd(SF_4) = \{\emptyset\}$
$pref(SF_4) = \{\{a, b\}, \{c, d\}\}$
$stb(SF_4) = \{\{a, b\}, \{c, d\}\}$

2.2. Towards HYPAFs

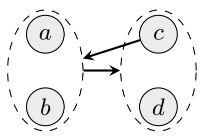
We proceed by defining HYPAFs as the faithful generalization of SETAFs where we allow sets of arguments in the second position of the attack relation.

Definition 5 (HYPAFs)**.** *A HYPAF is a pair* $HF = (A, R)$ *where* A *is a finite set of* arguments, *and* $R \subseteq 2^A \times (2^A \setminus \emptyset)$ *is the* attack relation.

For an illustration see Example 6. HYPAFs (A, R), where for all $(T, H) \in R$ it holds that $|H| = 1$, amount to SETAFs. Note that we allow for the empty set in the first position of an attack (i.e., $(\emptyset, H)$), as to be in line with our notion of SETAFs. The empty set in the second position of an attack however (i.e., $(T, \emptyset)$) we exclude. This is due to the fact that an attack towards an empty set of arguments is nonsensical and has no corresponding counter-part in any argumentation scenario.

Example 6. *Let the set* $\{a, b\}$ *attack the set* $\{c, d\}$ *and* $\{c\}$ *attack* $\{a, b\}$*, which we will illustrate as follows.*

HF_6:



Since there are different ways to interpret attacks in a HYPAFs, we will introduce different viewpoints on the matter in the next sections and illustrate their usefulness.

3. Universal Defeat

Intuition. In this section we interpret an attack T to H in the way that T defeats each element in H individually. In accordance with [5] we will call this notion *universal defeat.* Given a HYPAF $HF = (A, R)$ and an attack $(T, H) \in R$, this interpretation of a hyper-attack would be captured if the following implication holds:

> If all arguments in T are accepted, then each $h \in H$ is defeated.

However, as already observed by Nielsen and Parsons [2][2], this is mere syntactic sugar compared to usual SETAFs: Since a collective attack (T, h) from T to a single argument h encodes that h is defeated whenever all arguments in T are accepted, the above requirement can be captured by introducing the set

$$\{(T, h) \mid (T, H) \in R, h \in H\}$$

of collective attacks. In the following definition we formalize this reduction of [2] in our terminology.

Definition 7. *Let $HF = (A, R)$ be a HYPAF and $S \subseteq A$ a set of arguments. Then we say that S is a σ-extension of HF iff S is a σ-extension of the SETAF $SF = (A, \{(T, h) \mid (T, H) \in R, h \in H\})$.*

Example 6 (ctd). *Revisiting HF_6, when we interpret the attacks in the mode of universal defeat, the hyperframework HF_6 is equivalent to the SETAF SF_4 from Example 4.*

4. Indeterministic Defeat

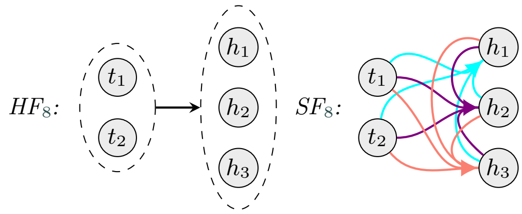
Intuition. This section aims at formalizing the intuition that for an attack $(T, H) \in R$ it is not clear (i.e. "non-deterministic") which of the arguments in H are actually attacked by T. That is, sets attacking sets are interpreted as a form of incomplete information. Formally, if we accept T, then for each $H' \subseteq H$ there shall be a possible world where H' is the precise set of arguments which is defeated due to this attack.

Indeterministic attack has been discussed by Nielsen and Parsons [2]. Here the underlying idea is as follows. If (T, H) is an attack, then the set $T \cup H$ is certainly not conflict-free. Since it is not clear how to draw more information when interpreting (T, H) as an indeterministic attack towards H, [2] refrain from encoding more than the definite conflicts we can be sure of. Thus they propose the following[3]: an attack $(\{t_1, \dots, t_n\}, \{h_1, \dots, h_m\})$ is mapped to the collective attacks

$$(\{t_1, \dots, t_n, h_1, \dots, h_{i-1}, h_{i+1}, \dots, h_m\}, h_i)$$

for each $1 \le i \le m$. Let us illustrate this approach.

Example 8. *In the construction from [2], the HYPAF HF_8 corresponds to the SETAF SF_8.*



While we indeed note that (I) $\{t_1, t_2\} \cup \{h_1, h_2, h_3\}$ is now conflicting, we want to point out some issues regarding this reduction, violating the intuition of indeterministic attacks.

(II) In order to accept any argument h_i, either an argument t_j or h_k with $k \neq i$ has to be defeated. Thus whether or not h_i is defended depends on the other h_k, but their connection is an in-coming set-attack, not an internal conflict.

(III) Any admissible set that contains t_1 and t_2 and at least one h_i argument has to contain exactly 2 arguments h_i, h_j. However, why should adding h_2 to $\{t_1, t_2, h_1\}$ render the set admissible, although the only attack involving h_2 is an in-coming one?

(IV) The arguments t_i are necessarily involved in attacks towards each h_i, although by our interpretation of indeterminism there should be a possible scenario where the arguments t_i are not part of any attack towards a single argument h_i.

From this illustrating example, we can extract the following desired properties for indeterministic HYPAFs corresponding to the observations (I)-(IV) from Example 8.

Property I Whenever we have an attack (T, H) and a jointly acceptable set of arguments S, we have $T \subseteq S \Rightarrow H \nsubseteq S$.

Property II Given an attack (T, H) and two arguments $h_1, h_2 \in H, h_1 \neq h_2$, whether h_1 is defended against the attack (T, H) does not depend on whether h_2 is accepted or not.

Property III For an attack (T, H) for each H' s.t. $\emptyset \subseteq H' \subsetneq H$ we model a situation where (T, H) does not cause a conflict in $T \cup H'$.

Property IV An attack (T, H) should not be interpreted as an attack where for two $h_1, h_2 \in H$ the argument h_1 is part of an attack towards h_2.

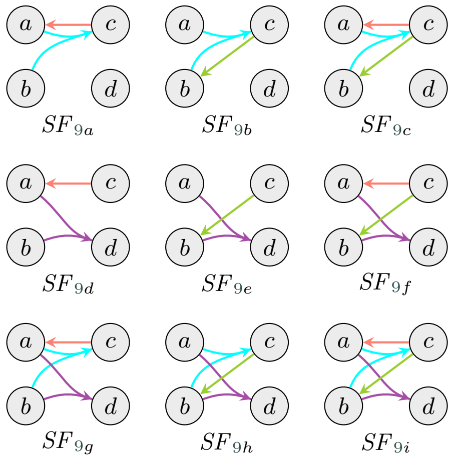
In the following section, we will introduce a different notion of indeterministic defeat that indeed satisfies all desired properties (I)-(IV).

[2]Note that in [2] this mode is called "collective defeat".
[3]While [2] only explicitly mentions the attack $(\{t_1, \dots, t_n, h_2, \dots, h_m\}, h_1)$, we assume the whole construction includes the symmetric cases towards each h_i.

4.1. Indeterministic HYPAFs

The underlying idea for our approach is to interpret the set-attacks of the form $(T, H) \in R$ as a blueprint to construct several SETAFs which represent *possible worlds.* Given a SETAF, we can rely on the rich body of research on this matter in order to assess acceptance of arguments. The following example shall illustrate our proposal.

Example 9. *Let us revisit Example 6. We can interpret this as: either $\{a, b\}$ defeats only c, $\{a, b\}$ defeats only d, or $\{a, b\}$ defeats both c and d. We do not know which is actually true, but we want to consider each possible scenario. Likewise, $\{c\}$ could defeat a, b, or both a and b. Each combination of these scenarios corresponds to a SETAF as illustrated below.*



In order to define semantics for indeterministic defeat we first introduce interpretations of HYPAFs in order the capture the possible worlds. Note that our HYPAFs with indeterministic defeat semantically coincide with conjunctive-disjunctive argumentation networks [7]. Our approach extends this notion by defining all standard (Dung-style) extension-based semantics, and that our approach is syntactically close to the established SETAFs.

Definition 10. *Let $HF = (A, R)$ be a HYPAF. For each attack $(T, H) \in R$ we choose a set of* interpreted collective attacks $R^I_{(T,H)}$ *s.t.*

$$\emptyset \subset R^I_{(T,H)} \subseteq \{(T, h) \mid h \in H\}.$$

An interpretation *of HF is any SETAF $SF^I = (A, R^I)$ s.t.*

$$R^I = \bigcup_{(T,H)\in R} R^I_{(T,H)}.$$

Example 9 (ctd)**.** *The SETAFs SF_{9a} to SF_{9i} depicted above correspond to the interpretations of HF_6 from Example 6. Each of these SETAFs realizes a possible world underlying the HYPAF HF_6 in question.*

For example, to construct SF_{9h} we let

$$R^I_{(\{a,b\},\{c,d\})} = \{(\{a, b\}, c), (\{a, b\}, d)\}$$
$$R^I_{(\{c\},\{a,b\})} = \{(c, b)\}$$

Next, we turn to the semantics of HYPAFs when interpreting attacks indeterministically. We define argument acceptance in hyper-argumentation frameworks with the following intuition: a set of arguments is *possibly* accepted (w.r.t. semantics σ) if it is accepted in at least *one* of the "instantiated" SETAFs. This leads us to the following definition of extensions in indeterministic HYPAFs.

Definition 11. *Let $HF = (A, R)$ be a HYPAF. A set $E \subseteq A$ is a (possible) σ-extension of HF iff E is an σ-extension for some interpretation of HF.*

We omit "possible" and simply speak of extensions if there is no risk of confusion.

Example 9 (ctd)**.** *In Example 9 we have that $\{a, b\}$ is a stable extension; this is witnessed by the SETAF SF_{9h}.*

We will now illustrate the adequacy of this definition by showing that our indeterministic HYPAFs indeed satisfy the desired properties (I)-(IV). Again, let $HF = (A, R)$ be a HYPAF with $(T, H) \in R$.

1. This is satisfied by the definition of conflict-freeness.
2. As desired, the attack (T, H) never maps to any scenario where two arguments $h_i, h_j \in H$ appear in the same collective attack (nor is h_i relevant for the defense of h_j).
3. In the interpretation where T defeats $H' \subseteq H$, the attack causes no conflict in $T \cup (H \setminus H')$.
4. In the interpretation where T defeats $H' \subseteq H$, there is no (partial) conflict between T and $H \setminus H'$.

It is clear however that our properties (I)-(IV) can only serve to cover a small subset of conceivable desiderata. Hence, to better put our proposal in context and demonstrate how it can be naturally captured by concepts from the literature, in the following section we characterize indeterministic HYPAFs by showing a semantic relation to attack-incomplete frameworks with correlations.

4.2. Relation to Attack-Incomplete SETAFs

As due to their similar construction, our indeterministic interpretation of HYPAFs is close to *attack-incomplete frameworks (iAFs)* [8]. In iAFs a subset of the attack relation is *uncertain*, i.e., the reasoning agent is not sure whether this attack exists or not. For their semantics each possible scenario of taking or omitting an uncertain attack is considered. In our setting, an attack

$(T, \{h_1, \ldots, h_n\})$ can be seen as the set of uncertain attacks $\{(T, h_1), \ldots, (T, h_n)\}$. However, indeterministic HYPAFs face an additional constraint, namely we require *at least one* of these attacks to be present in each scenario. [13] generalized iAFs by the addition of *correlations*. OR-Correlations pose the additional constraint that of a set $R' \subseteq R$, at least one attack of R' is present, albeit unknown which one. The semantics are defined in terms of *completions*, which correspond to our interpretations. We can straightforwardly generalize iAFs with OR-correlations to feature set attacks.

Definition 12. *A* iSETAF with correlations *is a tuple* $iSF = (A, R, R^?, \Delta)$*, where* A *is a finite set of arguments,* $R, R^? \subseteq 2^A \times A$ *are sets of certain/uncertain attacks, and* $\Delta \subseteq 2^{R^?} \setminus \emptyset$ *is a set of OR-correlations.* A valid completion *of* iSF *is a SETAF* $SF = (A, R')$*, where* $R \subseteq R' \subseteq R^?$ *such that for each* $D \in \Delta$ *it holds* $R' \cap D \neq \emptyset$*. A set* $S \subseteq A$ *is a possible* σ*-extension of* iSF *if for at least one completion of* iSF *the set* S *is a* σ*-extension.*

The following equivalence follows directly from the respective definitions (cf. Definition 10, 12).

Theorem 13. *Let* $HF = (A, R)$ *be a HYPAF.* $S \subseteq A$ *is a* σ*-extension of* HF *iff* S *is a possible* σ*-extension of the iSETAF* $iSF = (A, \emptyset, R^?, \Delta)$ *with*

$$R^? = \{(T, h) \mid (T, H) \in R, h \in H\},$$
$$\Delta = \{\{(T, h) \mid h \in H\} \mid (T, H) \in R\}.$$

Proof. The statement follows from Definition 10 and Definition 12: there is a 1-to-1 correspondence between the interpretations of HF and the valid completions of iSF. An indeterministic attack

$$(\{a_1, \ldots, a_m\}, \{b_1, \ldots, b_n\})$$

corresponds to the attacks

$$(\{a_1, \ldots, a_m\}, b_1), \ldots, (\{a_1, \ldots, a_m\}, b_n),$$

together with a corresponding OR-correlation that includes all of these attacks. □

Even though the iSETAFs we construct in Theorem 13 have no certain attacks, they are still effectively present as an attack $(T, h) \in R^?$ where $\{(T, h)\} \in \Delta$ is semantically equivalent to $(T, h) \in R$. This is not surprising, as these attacks correspond to (T, H) with $|H| = 1$ in the original HYPAF.

For the reverse direction, i.e., mapping iSETAFs as HYPAFs, we have to pose a restriction on the iSETAFs, namely that all uncertain attacks appear in at least one OR-constraint, and all attacks that appear in an OR-constraint together have the same tail, i.e., the following properties hold:

$$\bigcup_{D \in \Delta} D = R^? \quad (1)$$

$$\{(T_1, h_1), \ldots, (T_n, h_n)\} \in \Delta \Rightarrow T_1 = \cdots = T_n \quad (2)$$

Clearly, the construction in Theorem 13 maps precisely to those iSETAFs that satisfy both (1) and (2). Conversely, we show next that every iSETAF adhering to (1) and (2) can be seen as an equivalent HYPAF, i.e., the mapping is bijective.

Theorem 14. *Let* $iSF = (A, \emptyset, R^?, \Delta)$ *be a iSETAF adhering to (1), (2). A set* $S \subseteq A$ *is a possible* σ*-extension of* iSF *iff* S *is a* σ*-extension of the HYPAF* HF *given as the tuple*

$$(A, \{(T, \{h_1, \ldots, h_n\}) \mid \{(T, h_1), \ldots, (T, h_n)\} \in \Delta\}).$$

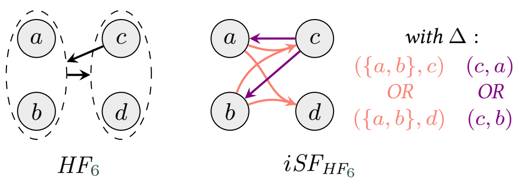
Proof. As in Theorem 13, there is a 1-to-1 correspondence between the interpretations of HF and the valid completions of iSF that satisfies properties (1), (2). An OR-correlation

$$\{(\{a_1, \ldots, a_m\}, b_1), \ldots, (\{a_1, \ldots, a_m\}, b_n)\}$$

corresponds to the indeterministic attack

$$(\{a_1, \ldots, a_m\}, \{b_1, \ldots, b_n\}). \quad \square$$

Example 6 (ctd)**.** *Let us revisit* HF_6 *(left). The corresponding iSETAF with OR-correlations is depicted below (right). Note that its valid completions coincide with the interpretations of our HYPAF (see Example 9).*

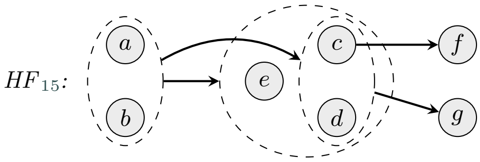


Theorems 13 and 14 provide an exact characterization of the relation between indeterministic HYPAFs and iSETAFs with OR-correlations.

5. Collective Defeat

Intuition. In this section, we interpret a set-attack (T, H) as a *collective* attack in the sense that whenever T is acceptable, then the set H of arguments (and thus, each superset of H) is not. As demonstrated in our introductory Tandem Example 15, we want to formalize the situation in which the set of arguments H is attacked as a whole but the attack does not affect any proper subset of H. We will see that with this new notion, the preferred extensions of the HYPAF in Example 1 are $\{t, a, b\}, \{t, a, c\}, \{t, b, c\}$ and thereby correspond to the intuitive outcome. We emphasize that we do not aim to reduce this attack to any conceivable notion of an attack towards (some of) the individual arguments in H.

Example 15. *Let us consider the following HYPAF HF_{15}.*



i) Our first observation here is that the attack $(\{a, b\}, \{c, d, e\})$ should be redundant as it states that c, d, e cannot be collectively accepted since $\{a, b\}$ are (they are unattacked). However, the attack $(\{a, b\}, \{c, d\})$ states already that c, d cannot be collectively accepted, which is a strictly stronger condition. ii) Secondly, both $\{a, b, c\}$ and $\{a, b, d\}$ should be acceptable because the attack $(\{a, b\}, \{c, d\})$ only forbids collective acceptance of c and d. iii) Moreover, g should be acceptable w.r.t. $\{a, b\}$ because in order to defeat g, $\{c, d\}$ are required which in turn should be interpreted defeated. iv) Finally, defending f is harder than defending g since defeating $\{c, d\}$ collectively is easier than defeating c specifically.

5.1. Semantics of Collective Defeat

Let us now define the standard concept required to generalize the usual AF semantics to capture the interaction of sets of arguments.

Definition 16. *Let $HF = (A, R)$ be a HYPAF and let $S, T \subseteq A$. We say that*

- *S attacks T iff there are $S' \subseteq S$ and $T' \subseteq T$ such that $(S', T') \in R$; we call S an attacker of T;*
- *S is conflict-free, $S \in cf(HF)$, iff it does not attack itself;*
- *S defends T iff S attacks all attacker of T, i.e. for all $(U, T') \in R$ with $T' \subseteq T$, there are $S' \subseteq S$ and $U' \subseteq U$ such that $(S', U') \in R$.*

We abuse notation and write S defends a whenever we mean that S defends $\{a\}$.

Example 15 (ctd). *Recall HF_{15}. We have that both $S_1 = \{a, b, c\}$ and $S_2 = \{a, b, d\}$ are conflict-free since $\{a, b\}$ only attacks $\{c, d\}$, but none of them individually. Moreover, $\{a, b\}$ defends g, but it does not defend f. We also want to mention that the conflict-free and defended sets in HF_{15} do not alter after removing $(\{a, b\}, \{c, d, e\})$, i.e. the attack is indeed redundant.*

We observe that if a set T is defended by some set S, then all individual arguments of T are defended as well.

Lemma 17. *Let (A, R) be a HYPAF and let $S, T \subseteq A$. If S defends T then S defends $\{a\}$ for each $a \in T$.*

Proof. The statement follows from the observation that each subset of a defended set is defended. Let $S \subseteq A$ defend T, let $T' \subseteq T$, and consider some attacker H of T'. By definition of attacks, H attacks T as well. By assumption, S attacks H and therefore also defends T' against H. Since H was an arbitrary attacker, the claim follows. □

Using these underlying notions, the definitions of the semantics naturally generalize to hyperframeworks.

Definition 18. *Let $HF = (A, R)$ be a HYPAF and let $S \in cf(HF)$. Then*

- *S is admissible, $S \in adm(H)$, iff S defends itself;*
- *S is complete, $S \in com(H)$, iff $S \in adm(HF)$ and S contains every set $T \subseteq A$ it defends;*
- *S is grounded, $S \in grd(HF)$, iff S is $\subseteq$-minimal in $com(H)$;*
- *S is preferred, $S \in pref(HF)$, iff S is $\subseteq$-maximal in $adm(H)$;*
- *S is stable, $S \in stb(HF)$, iff S attacks each $T \subseteq A \setminus S$.*

Example 15 (ctd). *Consider again HF_{15}. We have that $S = \{a, b, g\}$ is admissible ($\{a, b\}$ defends g). It is not maximal though since $S' = \{a, b, c, g, e\} \in adm(HF)$ as well. The latter is preferred. Note that f is not in any admissible set since defending f would require defeating c; no set of arguments is capable though.*

Interestingly, we can simplify our definitions for complete and stable semantics.

Lemma 19. *Let $HF = (A, R)$ be a HYPAF. Then*

- *$S \in stb(H)$ iff $S \in cf(HF)$ and S attacks each set $\{x\}$ for all $x \in A \setminus S$;*
- *$S \in com(H)$ iff $S \in adm(HF)$ and S contains each argument it defends.*

Proof. A stable set S attacks each singleton not contained in S by definition. In case each singleton is attacked, then each superset is attacked as well.

First assume S is complete. Then it is admissible and contains each set T it defends. Hence it contains each singleton it defends. Now assume S is admissible and contains each defended argument. Let T be defended by S. By Lemma 17, each argument $a \in T$ is defended by S as well. Hence $T \subseteq S$, as desired. □

Having defined our HYPAF semantics formally, we are now interested in their properties. Thereby, we pay special attention to the behavior of Dung's semantics in AFs, because they are well-behaved w.r.t. several aspects. We mention here the most common ones.

While stable extensions do not necessarily exist, we expect each HYPAF to possess admissible, complete, grounded, and preferred extensions. Moreover, the grounded extension should be unique since it intuitively formalizes the set of arguments one is willing to accept, even if the reasoning is cautious. The most important technical tool in order to ensure these properties is Dung's fundamental lemma [1].

Lemma 20 (Fundamental Lemma)**.** *Let $F = (A, R)$ be an AF, $S \in adm(F)$ and $T, T' \subseteq A$ be sets of arguments that are defended by S. Then*

1. *$S' = S \cup T$ is admissible, and*
2. *T' is defended by S'.*

Moreover, we typically expect preferred extensions (i.e. maximal admissible sets) to be complete. In summary, we get the following properties which are typically desirable for any generalization of Dung's setting.

1. (Some version of) the fundamental lemma holds.
2. There is always at least one admissible, complete, grounded, and preferred extension.
3. Every preferred extension is complete, and every stable extension is preferred.

Let us first discuss the positive news:

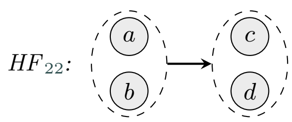
Observation 21. *Let HF be a HYPAF. Then*

1. *admissible and preferred extensions always exist; and*
2. *each stable extension is preferred.*

Both follows directly from the definitions: indeed, the empty set as well as each stable extension defends itself, moreover, each stable extension is $\subseteq$-maximal admissible.

However, the following simple example already illustrates that the semantics we defined so far violate all of the remaining properties.

Example 22. *Let us consider the following HYPAF HF_{22} in which the set $\{a, b\}$ attacks the set $\{c, d\}$.*



In HF_{22}, both sets $\{a, b, c\}$ and $\{a, b, d\}$ are conflict-free. Moreover, they are also admissible since they defend themselves (in fact, no subset of them is attacked). Moreover, they are preferred as $\{a, b, c, d\} \notin cf(HF_{22})$.

What are the complete extensions of the HYPAF? Coming from the well-behaving SETAFs, we would expect that the two preferred sets $\{a, b, c\}$ and $\{a, b, d\}$ are complete as well. However, it turns out that our HYPAF HF_{22} has no complete extension at all. Let us consider the set $\{a, b, c\}$: By definition, the set is admissible, moreover, the argument d is unattacked, hence $\{d\}$ is defended by $\{a, b, c\}$. However, we cannot extend $\{a, b, c\}$ with $\{d\}$ since the resulting set $\{a, b, c, d\}$ is not conflict-free anymore.

This shows not only that $com(HF_{22}) = \emptyset$, but also the fact that the fundamental lemma is violated.

From $com(HF_{22}) = \emptyset$ we also deduce $grd(HF_{22}) = \emptyset$.

We therefore conclude that the natural generalization of the semantics admits unexpected behavior. In summary, the previous example illustrates the following observation regarding our HYPAF semantics.

Observation 23. *Let HF be a HYPAF. Then*

1. *the fundamental lemma is in general violated;*
2. *complete and grounded extensions do not always exist;*
3. *not every preferred extension is complete.*

5.2. HYPAF Properties

In this section, we discuss complete and grounded semantics in more depth. For this, we define the *characteristic function* for HYPAFs as it is defined for (SET)AFs: Γ_{HF} applied to some set S of arguments returns all arguments which are defended by S. Due to Lemma 17 this also captures our intuition of defending sets of arguments.

Definition 24. *Let $HF = (A, R)$ be a HYPAF and let $S \subseteq A$. We define the* characteristic function *as*

$$\Gamma_{HF}(S) = \{a \in A \mid S \text{ defends } \{a\}\}.$$

Example 22 (ctd)**.** *We revisit HF_{22}. Then $\Gamma_{HF_{22}}(\emptyset) = \Gamma_{HF_{22}}(\{a, b\}) = \{a, b, c, d\}$ as each singleton is unattacked.*

We mention that our characteristic function is monotonic.

Lemma 25. *Let $HF = (A, R)$ be a HYPAF and let $S \subseteq T \subseteq A$. Then $\Gamma_{HF}(S) \subseteq \Gamma_{HF}(T)$.*

5.2.1. Complete Semantics

As for AFs and SETAFs, complete semantics can be alternatively defined via the characteristic function. By definition, the complete extensions are the conflict-free fixed points of Γ_{HF}.

Lemma 26. *Let $HF = (A, R)$ be a HYPAF. Then $S \in com(HF)$ iff $S \in cf(HF)$ and $S = \Gamma_{HF}(S)$.*

However, in contrast to Dung AFs and SETAFs, the characteristic function might have no conflict-free fixed points, as Example 22 demonstrates: $\{a,b,c,d\}$ is the only fixed point of $\Gamma_{HF_{22}}$. We can attribute the non-existence of complete extensions to set-attacks in the following sense: If the head of each attack contains at least two arguments, then complete extensions do not exist.

Lemma 27. *Let $HF = (A,R)$ be a HYPAF, $R \neq \emptyset$. If $|H| > 1$ for all $(T,H) \in R$ then com$(HF) = \emptyset$.*

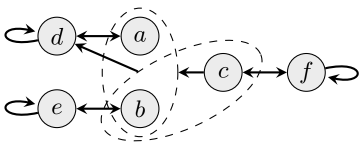
Proof. Let $a \in A$. We show that the singleton $\{a\}$ is not attacked by any set S of arguments. Suppose S attacks a. Then there is a subset $T' \subseteq \{a\}$ s.t. $(S', T') \in R$ for some $S' \subseteq S$. This contradicts our assumption $|H| > 1$ for each $(T,H) \in R$. □

This result formalizes that our notion of attacks between sets is not suitably tailored to assess the status of a *single* argument, but focuses on sets of arguments instead.

Even in the special (somewhat well-behaving) case where we do have fixed points for the characteristic function (i.e., there are complete extensions), we are not guaranteed to have the usual relations between the semantics that we know from Dung's notions.

Example 28. *Note that even in case com$(HF) \neq \emptyset$ holds, it is still not ensured that pref$(HF) \subseteq$ com(HF) holds, as HF_{28} illustrates: the set $\{b,c\}$ is preferred, but not complete (because $\{b,c\}$ defends a, but $\{a,b,c\}$ is not conflict-free). However, the empty set of arguments is complete in HF_{28}.*

HF_{28}:



5.2.2. Grounded Semantics

Now let us turn our attention towards the grounded extension. In AFs and SETAFs, the grounded extension is the least fixed point of the characteristic function. As is folklore in the argumentation community, the grounded extension can be computed by applying the characteristic function to the empty set until a fixed point is attained, i.e. we have *grd*$(F) = \{\Gamma^{\infty}_{F}(\emptyset)\}$ whenever F is a Dung-AF. An analogous result holds true in SETAFs. Since a HYPAF might have no conflict-free fixed points (cf. Example 22), this does not hold for our HYPAFs, i.e, *grd*$(HF) = \{\Gamma^{\infty}_{HF}(\emptyset)\}$ is not true anymore (in Example 22, $\Gamma_{HF_{22}}(\emptyset)$ contains all arguments and is not conflict-free).

We do however obtain the following positive result. While on the one hand we cannot guarantee that $\Gamma^{\infty}_{HF}(\emptyset)$ is conflict-free, we can on the other hand be certain that it is the *only* candidate for the grounded extension.

Proposition 29. *Let $HF = (A,R)$ be a HYPAF and let $S \subseteq A$. It holds that*

1. *grd*$(HF) = \{\Gamma^{\infty}_{HF}(\emptyset)\}$ *iff* $\Gamma^{\infty}_{HF}(\emptyset) \in$ *cf*(HF);
2. *grd*$(HF) = \emptyset$ *iff* $\Gamma^{\infty}_{HF}(\emptyset) \notin$ *cf*(HF).

Proof. By monotonicity of the characteristic function, it holds that $\Gamma^{\infty}_{HF}(\emptyset)$ is conflicting iff Γ_{HF} has no conflict-free fixed points. By definition of grounded semantics, we obtain the desired results. □

Hence, Γ_{HF} behaves as expected in case it admits a conflict-free fixed point. As an immediate corollary, we obtain that the grounded extension is unique whenever it exists.

Corollary 30. *For any HYPAF HF, $|$grd$(HF)| \leq 1$.*

Since there are never two grounded extensions, we sometimes abuse notation and write *grd*(HF) to denote the unique grounded extension of HF. Another corollary of Proposition 29 is that *grd*(HF) is a subset of each complete extension; i.e. *grd*(HF) is the least set in *com*(HF).

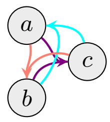
Corollary 31. *If com$(HF) \neq \emptyset$, then grd(HF) is the least complete set.*

5.3. Undirected Conflicts

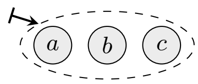
We want to point out that with our new notion of collective attack it is possible to model *undirected conflicts* natively within our framework.

Example 32. *We consider again the Tandem Example 15 from the introduction, but this time, we use only the arguments a, b, and c to model the conflict. Again, we have a conflict if we accept all of them but no conflict if we only accept a subset. In SETAFs we would model this scenario with symmetric attacks towards each argument (see SF_{32}). However, we see that this does not capture the intuition, as in this case for example a singleton set cannot be accepted. Our collective defeat allows for the attack $(\emptyset, \{a,b,c\})$ (see HF_{32}), which is an intuitive way to model the conflict and gives the desired behavior.*

SF_{32}:



HF_{32}:



The following result illustrates that if we omit the direction of attacks of a HYPAF, we retain the admissible sets.

Proposition 33. *Let $HF = (A, R)$ be a HYPAF and $S \subseteq A$ a set of arguments. If $S \in adm(HF)$, then $S \in adm(HF')$ where $HF' = (A, \{(\emptyset, T \cup H) \mid (T, H) \in R\})$.*

Proof. First note that every conflict-free set in HF' is admissible. Moreover, every conflict-free set in HF is also conflict-free in HF', as an attack $(\emptyset, T \cup H)$ in could only cause a conflict in a set S if $T \cup H \subseteq S$, which also causes a conflict in HF. □

Note however that a version of Proposition 33 where instead of admissible sets we consider semantics that maximize the extensions (like *grd*, *com*, *pref*, and *stb*) does *not* necessarily hold, as Example 34 illustrates.

Example 34. *Consider the following (HYP)AF HF_{34} and its counterpart with undirected conflicts HF'_{34}. While $\emptyset$ is the only extension of HF_{34} we have $grd(HF'_{34}) = com(HF'_{34}) = pref(HF'_{34}) = \{\{b\}\}$.*

Regarding stable semantics, if in HF_{34} we omit the self-attack of argument a, the set $\{a\}$ is stable, but in the corresponding undirected pendant there is no stable extension.

6. Computational Complexity

In this section, we briefly investigate the computational complexity of decision problems regarding our different notions of HYPAFs. We assume the reader to be familiar with the required notions; see e.g. [14] for an introduction to complexity analysis in the context of argumentation.

We focus on verifying extensions, however other computational problems are closely related and in most cases the complexity can be obtained as a corollary.

First we have that HYPAFs with universal defeat can be directly reduced to SETAFs (and vice versa). Thus the complexity coincides with the respective results of SETAFs [15], and can be found in the first line of Table 1. Second, for verifying extensions in HYPAFs with indeterministic defeat we observe a higher complexity. Intuitively, the raised complexity (as depicted in Table 1, second line) is due to the fact that in addition to the standard computational costs arising from the respective problems on SETAFs, a witnessing interpretation SETAF has to be guessed as well. Formally, we obtain the lower bounds by carefully inspecting the hardness proofs for attack-incomplete AFs [9] (for *pref*) and their more general form allowing OR-correlations [13] (for *com*, *grd*). We obtain the corresponding upper bounds by generalizing the algorithms of [13] from the Attack-Incomplete AFs to the attack-incomplete SETAFs setting. The idea is to iteratively remove conflicting attacks and attacks that it is impossible to defend against. Finally, for collective defeat we obtain the same computational properties as for SETAFs. The lower bound for preferred semantics carries over directly from the SETAF case, upper bounds can straightforwardly be obtained by the fact that conflict-freeness, defense, and the closure function can be computed in polynomial time.

Theorem 35. *Let $HF = (A, R)$ be a HYPAF and $S \subseteq A$ a set of arguments. For the problem of verifying whether S is a σ-extension of HF w.r.t. universal/indeterministic/collective defeat, the complexity results in Table 1 hold.*

Table 1
Complexity of verifying an extension

	grd	*adm*	*com*	*pref*	*stb*
universal defeat	in P	in P	in P	coNP-c	in P
indetermministic defeat	NP-c	in P	NP-c	Σ_2^P-c	in P
collective defeat	in P	in P	in P	coNP-c	in P

7. Discussion

In this paper, we provided three different defeat-modes for hyperattacks: universal, indeterministic, and collective. It turns out that *universal* defeat simply amounts to SETAF semantics. Our *indeterministic* defeat on the other hand generalizes attack-incomplete SETAFs, which in turn conservatively generalize attack-incomplete AFs [9]. Finally, we introduced *collective* defeat, which naturally generalizes Dung's original semantics to consider sets of arguments. However, we observe undesirable behavior of the characteristic function, i.e., the known well-behaved properties of Dung's framework are not preserved. In future work, we want to address these issues. HYPAFs with collective defeat resemble the semantics of ABA$^+$ [11]. Future work comprises of a deeper analysis of this connection. Moreover, we want to investigate the expressiveness of indeterministic and collective defeat and relate the obtained results to SETAFs [16] and ADFs [17]. We also plan to investigate alternative semantics for indeterministic defeat that do not heavily rely on the notion of interpretations. Finally, our undirected conflicts are similar in spirit to ideas due to [18]. In the future we want to explore this connection.

Acknowledgments

This work was funded by the Austrian Science Fund (FWF) under grant P32830 and the Vienna Science and

Technology Fund (WWTF) under grant ICT19-065 and by the Federal Ministry of Education and Research of Germany and by Sächsische Staatsministerium für Wissenschaft, Kultur und Tourismus in the programme Center of Excellence for AI-research "Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig", project identification number: ScaDS.AI.

References

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

[2] S. H. Nielsen, S. Parsons, A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments, in: Proceedings of ArgMAS 2006, Springer, 2006, pp. 54–73. doi:10.1007/978-3-540-75526-5_4.

[3] G. Flouris, A. Bikakis, A comprehensive study of argumentation frameworks with sets of attacking arguments, Int. J. Approx. Reason. 109 (2019) 55–86. doi:10.1016/j.ijar.2019.03.006.

[4] W. Dvořák, M. König, M. Ulbricht, S. Woltran, Rediscovering argumentation principles utilizing collective attacks, in: Proceedings of the 19th International Conference on Principles of Knowledge Representation and Reasoning KR22, 2022, pp. 122–131. URL: https://proceedings.kr.org/2022/13/.

[5] B. Verheij, Rules, reasons, arguments : formal studies of argumentation and defeat, Ph.D. thesis, Maastricht University, Netherlands, 1996. doi:10.26481/dis.19961205hv.

[6] A. Bochman, Collective argumentation and disjunctive logic programming, J. Log. Comput. 13 (2003) 405–428. doi:10.1093/logcom/13.3.405.

[7] D. M. Gabbay, M. Gabbay, Theory of disjunctive attacks, part I, Log. J. IGPL 24 (2016) 186–218. doi:10.1093/jigpal/jzv032.

[8] S. Coste-Marquis, C. Devred, S. Konieczny, M. Lagasquie-Schiex, P. Marquis, On the merging of Dung's argumentation systems, Artif. Intell. 171 (2007) 730–753. doi:10.1016/j.artint.2007.04.012.

[9] D. Baumeister, D. Neugebauer, J. Rothe, H. Schadrack, Verification in incomplete argumentation frameworks, Artif. Intell. 264 (2018) 1–26. doi:10.1016/j.artint.2018.08.001.

[10] A. Bondarenko, P. M. Dung, R. A. Kowalski, F. Toni, An abstract, argumentation-theoretic approach to default reasoning, Artif. Intell. 93 (1997) 63–101. doi:10.1016/S0004-3702(97)00015-5.

[11] K. Cyras, F. Toni, ABA+: assumption-based argumentation with preferences, in: C. Baral, J. P. Delgrande, F. Wolter (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016, AAAI Press, 2016, pp. 553–556. URL: http://www.aaai.org/ocs/index.php/KR/KR16/paper/view/12877.

[12] A. Bikakis, A. Cohen, W. Dvořák, G. Flouris, S. Parsons, Joint attacks and accrual in argumentation frameworks, in: D. Gabbay, M. Giacomin, G. R. Simari, M. Thimm (Eds.), Handbook of Formal Argumentation, volume 2, College Publications, 2021.

[13] B. Fazzinga, S. Flesca, F. Furfaro, Reasoning over attack-incomplete aafs in the presence of correlations, in: M. Bienvenu, G. Lakemeyer, E. Erdem (Eds.), Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning, KR 2021, Online event, November 3-12, 2021, 2021, pp. 301–311.

[14] W. Dvořák, P. E. Dunne, Computational problems in formal argumentation and their complexity, in: P. Baroni, D. Gabbay, M. Giacomin, L. van der Torre (Eds.), Handbook of Formal Argumentation, College Publications, 2018, pp. 631–687. Also appears in IfCoLog Journal of Logics and their Applications 4(8):2557–2622.

[15] W. Dvořák, A. Greßler, S. Woltran, Evaluating SETAFs via answer-set programming, in: Proceedings of SAFA 2018, volume 2171 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2018, pp. 10–21.

[16] W. Dvořák, J. Fandinno, S. Woltran, On the expressive power of collective attacks, Argument Comput. 10 (2019) 191–230. doi:10.3233/AAC-190457.

[17] J. Pührer, Realizability of three-valued semantics for abstract dialectical frameworks, Artif. Intell. 278 (2020). doi:10.1016/j.artint.2019.103198.

[18] A. Vassiliades, G. Flouris, T. Patkos, A. Bikakis, N. Bassiliades, D. Plexousakis, A multi attack argumentation framework, in: P. Baroni, C. Benzmüller, Y. N. Wáng (Eds.), Logic and Argumentation - 4th International Conference, CLAR 2021, Hangzhou, China, October 20-22, 2021, Proceedings, volume 13040 of *Lecture Notes in Computer Science*, Springer, 2021, pp. 417–436. doi:10.1007/978-3-030-89391-0_23.