Forbidden Sets in Argumentation Semantics

Paul E. DUNNE
Dept. of Computer Science, University of Liverpool, Liverpool, L69 7ZF, UK

Abstract. We consider an alternative interpretation of classical Dung argumentation framework (AF) semantics by introducing the concept of "*forbidden sets*". In informal terms, such sets are well-defined with respect to any extension-based semantics and reflect those subsets of argument that collectively can never form part of an acceptable solution. The forbidden set paradigm thus provides a parametric treatment of extension-based semantics. We present some general properties of forbidden set structures and describe the interaction between forbidden sets for a number of classical semantics. Finally we establish some initial complexity results in the arena of forbidden set decision problems.

Keywords. abstract argumentation frameworks; extension-based semantics; computational complexity

Introduction

Among the many developments arising from the seminal treatment of argumentation within the abstract graph-theoretic model of argumentation frameworks (AFs) from Dung [7], one of the most prolific areas has been the formulation of alternative "argumentation semantics": that is the conditions on subsets of a framework's atomic arguments characterising which such sets present collectively "justified" arguments from which fail to do so. In addition to those presented in [7] one finds ideas such as semi-stable in Caminada [5], ideal from Dung *et al.* [8], together with CF2 semantics arising in Baroni *et al.* [3], the parametric concept of resolution-based semantics described by Baroni and Giacomin [2] together with the analysis of one specific instantiation of this by Baroni *et al.* [1].

Our aim in the present article is *not* to offer yet another semantics of abstract argumentation derived from graph-theoretic considerations within the supportng AF, but rather to examine an alternative view of such that have already been posited and, indeed, may be offered subsequently.

The central conceit underpinning our treatment stems from the property that all such extension-based semantics (as these have come to be generally known) conceptually prescribe solutions via a "*positive*" *enumeration* of "acceptable" subsets of arguments within a framework, e.g. the so-called "conflict-free" solutions are those subsets, S, in which no attack is present between any members of S. Thus, in order to validate a set as acceptable it suffices to find it among the list of allowed solution sets.

Here we examine an alternative view: examining conditions on sets, S, which suffice to eliminate any possibility that S is an acceptable position. These conditions, in a similar style to classical extension bases, may be thought of as described through an enumeration of sets, which we will call the *forbidden sets* (with respect to a given AF and argumentation semantics). In this way *if* S *is* a forbidden set with respect to an AF, $\mathcal{H}$ and semantics σ this indicates that *no* σ-extension of $\mathcal{H}$ contains S as a subset.

We provide some basic background in Section 1, proceeding to define formally the concept of forbidden set in Section 2 and prove some generic properties of these. In Section 3 we then consider comparative aspects of forbidden sets defined for some standard semantics and review some questions concerning computational complexity matters within Section 4. Conclusions and open issues are presented in Section 5.

1. Preliminaries

We begin by recalling the concept of abstract argumentation framework and terminology from Dung [7] and outline the main computational problems that have been of interest within this.

Definition 1 *We use* $\mathcal{X}$ *to denote a* finite *set of arguments with* $\mathcal{A} \subseteq \mathcal{X} \times \mathcal{X}$ *the so-called* attack *relationship over these. An* argumentation framework *(*AF*) is a pair* $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$*. A pair* $\langle x, y \rangle \in \mathcal{A}$ *is referred to as '*y is attacked by x*' or '*x attacks y*'. Using* S *to denote an arbitrary subset of arguments for* $S \subseteq \mathcal{X}$*,*

$$\begin{array}{lll} S^- & =_{\text{def}} & \{\, p \;:\; \exists\, q \in S \;\; \textit{such that} \; \langle p, q \rangle \in \mathcal{A} \} \\ S^+ & =_{\text{def}} & \{\, p \;:\; \exists\, q \in S \;\; \textit{such that} \; \langle q, p \rangle \in \mathcal{A} \} \end{array}$$

We say that: $x \in \mathcal{X}$ *is* acceptable with respect to S *if for every* $y \in \mathcal{X}$ *that attacks* x *there is some* $z \in S$ *that attacks* y*. Given* $S \subseteq \mathcal{X}$*,* $\mathcal{F}(S) \subseteq \mathcal{X}$ *is the set of all arguments that are acceptable with respect to* S*, i.e.*

$$\mathcal{F}(S) \;=\; \{x \in \mathcal{X} \;:\; \forall\, y \textit{ such that } \langle y, x \rangle \in \mathcal{A} \;,\; \exists\, z \in S \textit{ s.t. } \langle z, y \rangle \in \mathcal{A}\}$$

A subset, S*, is* conflict-free *if no argument in* S *is attacked by any other argument in* S*. with* $\subseteq$*-maximal conflict-free set referred to as* naive *extensions. A conflict-free set* S *is* admissible *if every* $y \in S$ *is acceptable w.r.t* S*.* S *is a* complete *extension if* S *is conflict-free and should* $x \in \mathcal{F}(S)$ *then* $x \in S$*, i.e. every argument that is acceptable to* S *is a member of* S*, so that* $\mathcal{F}(S) = S$*. The set of* $\subseteq$*-maximal complete extensions coincide with with the set of* $\subseteq$*-maximal admissible sets these being termed* preferred extensions*. The set* S *is a* stable *extension if* S *is conflict free and* $S^+ = \mathcal{X} \setminus S$*. and is a* semi-stable *extension (Caminada [5]) if admissible and has* $S \cup S^+$ $\subseteq$*-maximal among all admissible sets.*

The grounded extension *of* $\langle \mathcal{X}, \mathcal{A} \rangle$ *is defined as the* $\subseteq$*-minimal complete extension.*

We use σ to denote an arbitrary semantics from

$$\{\text{CF}, \text{NVE}, \text{ADM}, \text{PR}, \text{ST}, \text{COM}, \text{SST}, \text{GR}\}$$

corresponding to conflict-free, naive, admissible, preferred, stable, complete, semi-stable and grounded instances.

For a given semantics σ and AF, $\mathcal{H}(\mathcal{X}, \mathcal{A})$ we use $\mathcal{E}_\sigma(\mathcal{H})$ to denote the set of all subsets of $\mathcal{X}$ that satisfy the conditions specified by σ. We say that σ is a *unique status* semantics if $|\mathcal{E}_\sigma(\mathcal{H})| = 1$ for every AF, $\mathcal{H}$, denoting the unique extension by $E_\sigma(\mathcal{H})$.

We complete this, brief, overview by describing the three canonical decision problems that may be instantiated for a given semantics: *Verification* (VER), *Credulous Acceptance* (CA) and *Sceptical Acceptance* (SA). Formal definitions of these problems for AFs are presented in Table 1.

Table 1. Decision Problems in AFs

Problem Name	Instance	Question
Verification (VER_σ)	$\mathcal{H}(\mathcal{X}, \mathcal{A})$; $S \subseteq \mathcal{X}$	Is $S \in \mathcal{E}_\sigma(\mathcal{H})$?
Credulous Acceptance (CA_σ)	$\mathcal{H}(\mathcal{X}, \mathcal{A})$; $x \in \mathcal{X}$	$\exists\ S \in \mathcal{E}_\sigma(\mathcal{H})$ for which $x \in S$?
Sceptical Acceptance (SA_σ)	$\mathcal{H}(\mathcal{X}, \mathcal{A})$; $x \in \mathcal{X}$	$\forall\ T \in \mathcal{E}_\sigma(\mathcal{H})$ is $x \in T$?

2. Forbidden Sets and Related Structures

In this paper we introduce and explore the properties of a "parametric" operator – the *forbidden set* constructor – and its relationship with the extension-based semantics outlined in the preceding section.

Definition 2 *Let $\mathbb{S} \subseteq 2^{\mathcal{X}}$. A set $T \subseteq \mathcal{X}$ is said to be a* forbidden set *for $\mathbb{S}$ if for every set $S \in \mathbb{S}$, it is* not *the case that $T \subseteq S$.*

A set, $T \subseteq \mathcal{X}$ is a minimal *forbidden set for $\mathbb{S}$ if it is both a forbidden set for $\mathbb{S}$ but no* strict *subset of T describes a forbidden set for $\mathbb{S}$. Given $\mathbb{S}$, the notation $\kappa(\mathbb{S})$ and $\mu(\mathbb{S})$ describe those subsets of $\mathcal{X}$ for which*

$$\kappa(\mathbb{S}) \quad = \quad \{\ T \subseteq \mathcal{X}\ :\ T \textit{ is a forbidden set for } \mathbb{S}\ \}$$

$$\mu(\mathbb{S}) \quad = \quad \{\ T \subseteq \mathcal{X}\ :\ T \textit{ is a minimal forbidden set for } \mathbb{S}\ \} \quad \subseteq \quad \kappa(\mathbb{S})$$

For k, with $0 \leq k \leq |\mathcal{X}|$, the k-section of $\mathbb{S}$*, denoted $\chi^{(k)}(\mathbb{S})$, is*

$$\chi^{(k)}(\mathbb{S}) \quad = \quad \{\ P \subseteq \mathcal{X}\ :\ |P| = k \textit{ and } P \in \kappa(\mathbb{S})\ \}$$

In the special case $k = 1$, $\chi^{(1)}$ are those members of $\mathcal{X}$ that do not occur in any set of $\mathbb{S}$; while the subsets $\chi^{(2)}$ play an important role in the characterization considered in Dunne *et al.* [10] where these are referred to as "unpaired elements".

We note that $\kappa(\mathbb{S})$ and, potentially, $\mu(\mathbb{S})$, contains sets which are strict *supersets* of elements in $\mathbb{S}$. A simple example of such behaviour is given with $\mathcal{X} = \{x_1, x_2\}$ and $\mathbb{S} = \{\{x_1\}, \{x_2\}\}$: in this case $\mu(\mathbb{S}) = \kappa(\mathbb{S}) = \{\{x_1, x_2\}\}$.

Some properties of these operations are exploited in later results such as Lemma 2.

Lemma 1 *Given $\mathbb{S} \subseteq 2^{\mathcal{X}}$ and $\kappa(\mathbb{S})$ as defined in Defn. 2, the set systems $\kappa(\mathbb{S})$, $\mu(\mathbb{S})$ and $\chi^{(k)}(\mathbb{S})$, satisfy*

a. If $Q \in \kappa(\mathbb{S})$ there is (at least one) $R \subseteq Q$ with $R \in \mu(\mathbb{S})$.
b. The conditions $\emptyset \in \kappa(\mathbb{S})$, $\mu(\mathbb{S}) = \{\emptyset\}$ and $\mathbb{S} = \emptyset$ are equivalent.
c. $\kappa(\mathbb{S}) = \emptyset$ if and only if $\{x_1, \ldots, x_n\} \in \mathbb{S}$, that is to say $\mathbb{S}$ contains the set which comprises all of the arguments in $\mathcal{X}$.

Proof: Recall that we assume $\langle \mathcal{X}, \mathcal{A} \rangle$ is a *finite* structure.

For (a) suppose that $Q \in \kappa(\mathbb{S})$. If it is the case that no strict subset of Q is a forbidden set for $\mathbb{S}$ then, by definition, we have $Q \in \mu(\mathbb{S})$. Otherwise we find some $T \subset Q$ for which $T \in \kappa(\mathbb{S})$. Repeating the argument either T is a minimal forbidden set for $\mathbb{S}$ or has some subset which is a forbidden set. Eventually we find some $R \subseteq Q$ which is both forbidden and minimally so.

For (b), that $\mu(\mathbb{S}) = \{\emptyset\}$ if and only if $\emptyset \in \kappa(\mathbb{S})$ follows directly from the definition of $\mu(\mathbb{S})$. To see that $\mu(\mathbb{S}) = \{\emptyset\}$ is only possible when $\mathbb{S} = \emptyset$, again from the definition of forbidden set, were $\emptyset$ to be a forbidden set for $\mathbb{S}$ this indicates that no $S \in \mathbb{S}$ has $\emptyset \subseteq S$. This property can only be satisfied in the degenerate case $\mathbb{S} = \emptyset$.

With (c), $\kappa(\mathbb{S}) = \emptyset$ expresses the property that $\mathbb{S}$ has no forbidden sets at all, so, in particular, the set containing all arguments of $\mathcal{X}$ must belong to $\mathbb{S}$. Conversely, should $\{x_1, \ldots, x_n\} \in \mathbb{S}$ this suffices to rule out any subset of $\mathcal{X}$ as forbidden, i.e. $\kappa(\mathbb{S}) = \emptyset$. □

3. Comparative Properties of Forbidden Sets in Divers Semantics

Let $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$ be an AF. A natural question arising with respect to the forbidden set paradigm, concerns what may be said in general regarding comparisons between distinct extension sets of $\mathcal{H}$ and their associated forbidden sets.

In order to avoid an excess of parentheses, we adopt the following notation when considering a given AF $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$.

$$\begin{array}{lll} \mathcal{E}_\sigma^{\mathcal{H}} & =_{\text{def}} & \mathcal{E}_\sigma(\mathcal{H}) \\ \kappa_\sigma^{\mathcal{H}} & =_{\text{def}} & \kappa(\mathcal{E}_\sigma(\mathcal{H})) \\ \mu_\sigma^{\mathcal{H}} & =_{\text{def}} & \mu(\mathcal{E}_\sigma(\mathcal{H})) \end{array}$$

The following results present some basic relationships between forbidden sets and underlying semantics. Noting that Case (b) of Lemma 2 indicates semantics, σ defined as $\subseteq$-maximal elements of semantics τ have identical forbidden sets determining membership of $S \in \mathcal{E}_\sigma$ cannot be established simply by arguing S has no $R \in \mu(\mathcal{E}_\sigma)$ as a subset. The relationship given in part (c), however, does provide a method by which $S \in \mathcal{E}_\sigma$ can be decided via forbidden set structures.

Lemma 2

a. *If* σ, τ *are semantics that satisfy,* $\mathcal{E}^{\mathcal{H}}_{\sigma} \subseteq \mathcal{E}^{\mathcal{H}}_{\tau}$ *then* $\kappa^{\mathcal{H}}_{\tau} \subseteq \kappa^{\mathcal{H}}_{\sigma}$.
b. *If* $\mathcal{E}^{\mathcal{H}}_{\sigma}$ *is defined to be the* ($\subseteq$)*-maximal sets within* $\mathcal{E}^{\mathcal{H}}_{\tau}$ *then* $\kappa^{\mathcal{H}}_{\tau} = \kappa^{\mathcal{H}}_{\sigma}$.
c. *If* σ, τ *satisfy the condition given in (b) then, for all* $S \subseteq \mathcal{X}$ $S \in \mathcal{E}^{\mathcal{H}}_{\sigma}$ *if and only if*

$$(\exists\ Q \in \mu(\mathcal{E}^{\mathcal{H}}_{\tau} \setminus \mathcal{E}^{\mathcal{H}}_{\sigma})\ :\ Q \subseteq S) \quad \textit{and} \quad (\forall\ Q \in \mu^{\mathcal{H}}_{\sigma}\ \neg(Q \subseteq S))$$

Proof: For (a), when σ and τ satisfy $\mathcal{E}^{\mathcal{H}}_{\sigma} \subseteq \mathcal{E}^{\mathcal{H}}_{\tau}$ no set in $\kappa^{\mathcal{H}}_{\tau}$ can be a subset of *any* set in $\mathcal{E}^{\mathcal{H}}_{\tau}$. In particular if $\mathbb{S} \subseteq \mathcal{E}^{\mathcal{H}}_{\tau}$ then a forbidden set for $\mathcal{E}^{\mathcal{H}}_{\tau}$ is perforce also a forbidden set for $\mathbb{S}$. It follows that any forbidden set for $\mathcal{E}^{\mathcal{H}}_{\tau}$ is a forbidden set for $\mathcal{E}^{\mathcal{H}}_{\sigma}$, i.e. $\kappa^{\mathcal{H}}_{\tau} \subseteq \kappa^{\mathcal{H}}_{\sigma}$.

For (b), the maximality premise already ensures $\kappa^{\mathcal{H}}_{\tau} \subseteq \kappa^{\mathcal{H}}_{\sigma}$ via part (a). Consider any $S \in \kappa^{\mathcal{H}}_{\sigma}$ and suppose, for the sake of contradiction, that $S \notin \kappa^{\mathcal{H}}_{\tau}$. From the definition of forbidden set this means we can find $T \in \mathcal{E}^{\mathcal{H}}_{\tau}$ with $S \subseteq T$. Now, however, we find $R \in \mathcal{E}^{\mathcal{H}}_{\sigma}$ with $T \subseteq R$ so that $S \subseteq T \subseteq R \in \mathcal{E}^{\mathcal{H}}_{\sigma}$ contradicting $S \in \kappa^{\mathcal{H}}_{\sigma}$.

For the relationship in (c), should it be the case that $S \in \mathcal{E}^{\mathcal{H}}_{\sigma}$ then $S \notin \mathcal{E}^{\mathcal{H}}_{\tau} \setminus \mathcal{E}^{\mathcal{H}}_{\sigma}$ so that $S \in \kappa(\mathcal{E}^{\mathcal{H}}_{\tau} \setminus \mathcal{E}^{\mathcal{H}}_{\sigma})$ and the property of there being some $Q \in \mu(\mathcal{E}^{\mathcal{H}}_{\tau} \setminus \mathcal{E}^{\mathcal{H}}_{\sigma})$ with $Q \subseteq S$ follows from Lemma 1(a). Similarly the premise $S \in \mathcal{E}^{\mathcal{H}}_{\sigma}$ indicates $S \notin \kappa^{\mathcal{H}}_{\sigma}$ thus no $Q \in \mu^{\mathcal{H}}_{\sigma}$ satisfies $Q \subseteq S$.

Conversely suppose that some $Q \in \mu(\mathcal{E}^{\mathcal{H}}_{\tau} \setminus \mathcal{E}^{\mathcal{H}}_{\sigma})$ satisfies $Q \subseteq S$ but that no $Q \in \mu^{\mathcal{H}}_{\sigma}$ has this property. Then,

$$\begin{array}{rcl} Q \in \mu(\mathcal{E}^{\mathcal{H}}_{\tau} \setminus \mathcal{E}^{\mathcal{H}}_{\sigma}) \ \text{ and } \ Q \subseteq S & \Rightarrow & S \notin \mathcal{E}^{\mathcal{H}}_{\tau} \setminus \mathcal{E}^{\mathcal{H}}_{\sigma} \\ & \Rightarrow & S \notin \mathcal{E}^{\mathcal{H}}_{\tau} \ \text{ or } \ S \in \mathcal{E}^{\mathcal{H}}_{\sigma} \end{array}$$

In addition,

$$\forall\ Q \in \mu^{\mathcal{H}}_{\sigma}\ \neg(Q \subseteq S)\ \Rightarrow\ S \in \mathcal{E}^{\mathcal{H}}_{\tau}$$

Notice that as a consequence of (b) we have $\mu^{\mathcal{H}}_{\sigma} = \mu^{\mathcal{H}}_{\tau}$ so we cannot directly deduce from $\neg(Q \subseteq S)$ for each $Q \in \mu^{\mathcal{H}}_{\sigma}$ that $S \in \mathcal{E}^{\mathcal{H}}_{\sigma}$: only $S \in \mathcal{E}^{\mathcal{H}}_{\tau}$. Combining $S \notin \mathcal{E}^{\mathcal{H}}_{\tau} \setminus \mathcal{E}^{\mathcal{H}}_{\sigma}$ and $S \in \mathcal{E}^{\mathcal{H}}_{\tau}$ we deduce that $S \in \mathcal{E}^{\mathcal{H}}_{\sigma}$ as claimed. □

Corollary 1 *For all* $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$

a. $\kappa^{\mathcal{H}}_{\mathrm{ADM}} = \kappa^{\mathcal{H}}_{\mathrm{PR}} = \kappa^{\mathcal{H}}_{\mathrm{COM}}$.
b. $\kappa^{\mathcal{H}}_{\mathrm{CF}} = \kappa^{\mathcal{H}}_{\mathrm{NVE}}$.
c. $\kappa^{\mathcal{H}}_{\mathrm{PR}} \subseteq \kappa^{\mathcal{H}}_{\mathrm{SST}} \subseteq \kappa^{\mathcal{H}}_{\mathrm{ST}}$.
d. $\kappa^{\mathcal{H}}_{\mathrm{CF}} \subseteq \kappa^{\mathcal{H}}_{\mathrm{ADM}}$.

Proof: Immediate consequence of Lemma 2 and established containment properties of the featured semantics. □

It is worth noting at this point a distinguishing aspect of the forbidden set paradigm in comparison with the extension-based semantics. It is well known in the latter formalism, that $\mathcal{E}_{co}^{\mathcal{H}} \subseteq \mathcal{E}_{adm}^{\mathcal{H}}$, i.e. every complete extension is an admissible set. The converse, however, does not hold: one may construct frameworks having $S \in \mathcal{E}_{adm}^{\mathcal{H}}$ but $S \notin \mathcal{E}_{co}^{\mathcal{H}}$.[1] The forbidden set structures for both semantics, however, are identical in consequence of $\mathcal{E}_{pr}^{\mathcal{H}}$ being formed by $\subseteq$-maximal admissible sets and $\subseteq$-maximal complete sets.

As a second point Corollary 1(a) offers an interesting point of comparison with recent work of Baumann *et al.* [4]. In this regard if we wish to distinguish $S \in \mathcal{E}_{\mathrm{ADM}}^{\mathcal{H}}$ from $S \in \mathcal{E}_{\mathrm{PR}}^{\mathcal{H}}$ in order to do so via the forbidden set paradigm the additional information required in terms of Lemma 2 (c) can be used.

We next establish that the relationships from Corollary 1(c)-(d) are exact, i.e we construct instances for which $\mu_\sigma^{\mathcal{H}} \not\subseteq \mu_\tau^{\mathcal{H}}$ although $\mu_\sigma^{\mathcal{H}} \subseteq \kappa_\tau^{\mathcal{H}}$, indicating the *minimal* forbidden sets are distinct.

Lemma 3 *There are choices of $\mathcal{H}$ with which,*

a. $\mu_{\mathrm{CF}}^{\mathcal{H}} \not\subseteq \mu_{\mathrm{ADM}}^{\mathcal{H}}$.
b. $\mu_{\mathrm{PR}}^{\mathcal{H}} \not\subseteq \mu_{\mathrm{ST}}^{\mathcal{H}}$.
c. $\mu_{\mathrm{PR}}^{\mathcal{H}} \not\subseteq \mu_{\mathrm{ST}}^{\mathcal{H}}$ and $\mathcal{E}_{st}^{\mathcal{H}} \neq \emptyset$.
d. $\mu_{\mathrm{PR}}^{\mathcal{H}} \not\subseteq \mu_{\mathrm{SST}}^{\mathcal{H}}$.
e. $\mu_{\mathrm{SST}}^{\mathcal{H}} \not\subseteq \mu_{\mathrm{ST}}^{\mathcal{H}}$.

Proof: Consider the three AFs shown in Fig. 1.

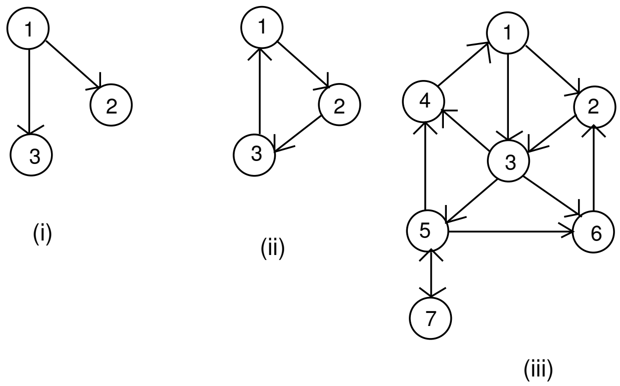

Figure 1. Non-containment properties in minimal forbidden set semantics

The AF shown in Fig. 1(i) has

$$\mu_{\mathrm{CF}}^{\mathcal{H}} = \{\{1,2\},\ \{1,3\}\}$$
$$\mu_{\mathrm{ADM}}^{\mathcal{H}} = \{\{2\},\ \{3\}\}$$

[1]Any AF for which $E_{gr}(\mathcal{H}) \neq \emptyset$ provides such an example: the empty set is always an admissible set but (in these cases) will fail to be a complete extension.

This suffices to establish (a). The AF depicted in Fig. 1(ii) has $\mathcal{E}^{\mathcal{H}}_{st} = \emptyset$ and $\mathcal{E}^{\mathcal{H}}_{pr} = \mathcal{E}^{\mathcal{H}}_{sst} = \{\emptyset\}$. In consequence,

$$\mu^{\mathcal{H}}_{\mathrm{ST}} = \{\emptyset\}$$

while

$$\mu^{\mathcal{H}}_{\mathrm{PR}} = \mu^{\mathcal{H}}_{\mathrm{SST}} = \{\{1\},\{2\},\{3\}\}$$

The relationships claimed in (b) and (e) are now immediate.

Finally in the AF shown under Fig. 1(iii) we have,

$$\begin{array}{lcl} \mathcal{E}^{\mathcal{H}}_{pr} & = & \{\{1,5\},\{7\}\} \\ \mathcal{E}^{\mathcal{H}}_{st} & = & \{\{1,5\}\} \\ \mathcal{E}^{\mathcal{H}}_{sst} & = & \{\{1,5\}\} \\ \mu^{\mathcal{H}}_{\mathrm{PR}} & = & \{\{2\},\{3\},\{4\},\{6\},\ \{1,5,7\}\} \\ \mu^{\mathcal{H}}_{\mathrm{ST}} & = & \{\{2\},\{3\},\{4\},\{6\},\{7\}\} \\ \mu^{\mathcal{H}}_{\mathrm{SST}} & = & \{\{2\},\{3\},\{4\},\{6\},\{7\}\} \end{array}$$

From which (c) and (d) are easily deduced. □

Finally we have a select number of instances where the structure of forbidden sets is characterized exactly.

Lemma 4

a. The set $\mu^{\mathcal{H}}_{\mathrm{CF}}$ is formed by the $\subseteq$-minimal sets in

$$\{\{x,y\} \ : \ \langle x,y\rangle \in \mathcal{A} \ or\ \langle y,x\rangle \in \mathcal{A}\} \ \cup \ \{\{x\} \ : \ \langle x,x\rangle \in \mathcal{A}\}$$

b. A set $S \subseteq \mathcal{X}$ is defenceless in $\mathcal{H}$ *if and only if every superset T of S satisfies*

$$T \in \mathcal{E}^{\mathcal{H}}_{cf} \ \Rightarrow \ \exists\, r \in T^- \ : \ r \notin T^+$$

The set $\mu^{\mathcal{H}}_{\mathrm{ADM}}$ is formed by the $\subseteq$-minimal defenceless sets of $\mathcal{H}$.

c. For any unique status semantics, σ,

$$\mu^{\mathcal{H}}_{\sigma} \ = \ \{\ \{x\} \ : \ x \notin E_{\sigma}(\mathcal{H})\}$$

Proof: For (a), consider any $Q \in \mu^{\mathcal{H}}_{\mathrm{CF}}$ and observe that any such Q has $|Q| \leq 2$: the property of Q being a forbidden set for conflict-free sets is easily seen to be equivalent to $(Q \times Q) \cap \mathcal{A} \neq \emptyset$ so that the corresponding minimal forbidden subsets within Q are formed by those pairs $\{x,y\} \subseteq Q$ linked by an attack in $\mathcal{A}$ together with self-attacking arguments.

For (b) if $S \subseteq \mathcal{X}$ is defenceless in $\mathcal{H}$ not only is S itself not in $\mathcal{E}^{\mathcal{H}}_{adm}$ but also S cannot be extended to an admissible set. Hence S cannot form a subset of any member of $\mathcal{E}^{\mathcal{H}}_{adm}$, i.e. $S \in \kappa^{\mathcal{H}}_{\mathrm{ADM}}$ as required.

Part (c) is trivial. □

It is easy to see that for each $0 \leq k \leq n$ ($n = |\mathcal{X}|$) one can construct AFs $\langle \mathcal{X}, \mathcal{A} \rangle$ in which there is some $S \in \mathcal{E}_{pr}(\langle \mathcal{X}, \mathcal{S} \rangle)$ for which $|S| = k$. A similar "hierarchy" is, however, not possible with respect to members of $\mu^{\mathcal{H}}_{\mathrm{PR}}$. We present a sub-optimal variant of this claim in,

Theorem 1 *For all $n \geq 4$ with $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$*

$$\begin{array}{lclcl} \forall\ \mathcal{H}(\mathcal{X}, \mathcal{A}) & : & \max_{S \in \mu^{\mathcal{H}}_{\mathrm{PR}}} |S| & \leq & n - \log_2 n \\ \exists\ \mathcal{H}(\mathcal{X}, \mathcal{A}) & : & \max_{S \in \mu^{\mathcal{H}}_{\mathrm{PR}}} |S| & = & \lfloor n/2 \rfloor \end{array}$$

Proof: Noting that for $n \in \{2, 3\}$ it is easy to form $S \in \mu^{\mathcal{H}}_{pr}$ having $|S| = 2$ (just use $\mathcal{A} = \{\langle x, y \rangle, \langle x, z \rangle, \langle y, x \rangle\}$ so that $\mu^{\mathcal{H}}_{pr} = \{\{x, y\},\ \{x, z\}\}$). The reader may easily verify by inspection that for $n = 3$, no AF having a minimal forbidden set of size 3 can be built.

Thus, assuming $n \geq 4$, we start with the upper bound claim, i.e that

$$\max_{S \in \mu^{\mathcal{H}}_{\mathrm{PR}}} |S| \leq n - \log_2 n$$

Let $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$ be any AF for which $|\mathcal{X}| = n$ and is such that no other AF, $\mathcal{G}$ of n arguments has

$$\max_{S \in \mu^{\mathcal{G}}_{\mathrm{PR}}} |S| > \max_{S \in \mu^{\mathcal{H}}_{\mathrm{PR}}} |S|$$

Consider any set S witnessing this behaviour in $\langle \mathcal{X}, \mathcal{A} \rangle$ and without loss of generality assume

$$S = \{x_1,\ x_2,\ \ldots,\ x_r\}$$

(where, trivially, $r \geq 3$). It is certainly the case that $S \in \mathcal{E}^{\mathcal{H}}_{\mathrm{CF}}$ for otherwise we find a strict subset of S which is in $\kappa^{\mathcal{H}}_{\mathrm{PR}}$ in contradiction to $S \in \mu^{\mathcal{H}}_{\mathrm{PR}}$.

For each $x_i \in S$, let S_i denote $S \setminus \{x_i\}$. By definition from $S \in \mu^{\mathcal{H}}_{\mathrm{PR}}$ we therefore have for every i, $S_i \notin \kappa^{\mathcal{H}}_{\mathrm{PR}}$ and hence we can find $T_i \subseteq \mathcal{X} \setminus S$ that satisfies

$$S_i \cup T_i \in \mathcal{E}^{\mathcal{H}}_{pr}$$

Observe that the system $\langle T_1,\ T_2, \cdots, T_r \rangle$ must consist of r *distinct* subsets of $\mathcal{X} \setminus S$, i.e. $T_i = T_j$ if and only if $i = j$. For suppose, without loss of generality, $T_1 = T_2$. Then

$$S_1 \cup T_1 \in \mathcal{E}^{\mathcal{H}}_{\mathrm{PR}} \quad \text{and} \quad S_2 \cup T_1 \in \mathcal{E}^{\mathcal{H}}_{\mathrm{PR}}$$

so that $S_1 \cup S_2 \cup T_1 = S \cup T_1 \in \mathcal{E}^{\mathcal{H}}_{\mathrm{ADM}}$ contradicting $S \in \kappa^{\mathcal{H}}_{\mathrm{PR}}$. Notice that admissibility of $S \cup T_1$ follows since the set is conflict-free and should $y \in \mathcal{X} \setminus (S \cup T_1)$ attack $S \cup T_1$ either it attacks some member of T_1 and thence is counterattacked by both S_1 and S_2 or y attacks some argument in $S = S_1 \cup S_2$ and so is defended by either T_1 or S_1 (if $y \in S_1^-$) or S_2 (should $y \in S_2^-$).

From this argument we obtain the (crude) upper bound claimed on the size of the largest possible set in $\mu^{\mathcal{H}}_{\rm PR}$: there are r arguments in S and require $n-r$ (the size of $\mathcal{X} \setminus S$) to be such that (at least) r *distinct* sets may be formed. That is we require

$$2^{n-r} \quad \geq \quad r$$

Should $r > n - \log_2 n$ then, $2^{n-r} \; = \; 2^{\log_2 n - \varepsilon}$ for some $\varepsilon > 0$, giving $2^{n-r} = n/2^{\varepsilon}$ and since $2^{\varepsilon} > 1$ (via $\varepsilon > 0$) it follows that

$$\frac{n}{2^{\varepsilon}} \quad < \quad n - \log_2 n + \varepsilon$$

as required for the upper bound.

To show that there are AFs, $\mathcal{H}(\mathcal{X},\mathcal{A})$ for which $\max_{S \in \mu^{\mathcal{H}}_{\rm PR}} \; |S|$ is at least $\lfloor |\mathcal{X}|/2 \rfloor$, let $m \geq 2$ and define

$$\mathcal{X} \quad = \quad \begin{cases} \{y_1, y_2, \ldots, y_m, z_1, z_2, \ldots, z_m\} & \text{if } n = 2m \\ \{y_1, y_2, \ldots, y_m, z_1, z_2, \ldots, z_m, \; u\} & \text{if } n = 2m+1 \end{cases}$$

We construct an AF, $\langle \mathcal{X}, \mathcal{A} \rangle$, for which

$$\{y_1, \; y_2, \ldots, \; y_m\} \quad \in \quad \mu^{\mathcal{H}}_{\rm PR}$$

We concentrate on the case $n = 2m$, since the construction for $n = 2m+1$ is identical. For the arguments, $\{z_1, z_2, \ldots, z_m\}$ all of the $m(m-1)$ attacks,

$$\{ \; \langle z_i, z_j \rangle \; : \; 1 \leq i \neq j \leq m \}$$

are added, so that *at most* one z_k can appear in any $P \in \mathcal{E}^{\mathcal{H}}_{adm}$. The set of attacks is completed with

$$\{ \; \langle z_i, y_i \rangle \; : \; 1 \leq i \leq m \; \}$$

Consider the set $S = \{y_1, y_2, \ldots, y_m\}$. Certainly $S \notin \mathcal{E}^{\mathcal{H}}_{adm}$ since, although conflict-free, there is no way of defending the attack on y_i arising from z_i. In addition, it is not possible to find a subset T of $\mathcal{X}$ for which $S \cup T \in \mathcal{E}^{\mathcal{H}}_{adm}$, since the only arguments available to form such a set are with $\{z_1, \ldots, z_m\}$ and the resulting $S \cup T$ would fail to be conflict-free.

In total these establish $S \in \kappa^{\mathcal{H}}_{\rm ADM}$. It is, however, also a *minimal* such set. To see this, let $S_i \; = \; S \setminus \{y_i\}$. It is not hard to see that for each i, $S_i \notin \kappa^{\mathcal{H}}_{\rm ADM}$: the set $S_i \cup \{z_i\}$ being in $\mathcal{E}^{\mathcal{H}}_{\rm ADM}$ (in fact it is a prefered extension). The argument z_i defends itself from attacks stemming from z_j ($j \neq i$) and, furthermore defends $y_j \in S_i$ from the attack on it by z_j. Thus $S_i \cup \{z_i\}$ is both conflict-free and defensive, i.e. in $\mathcal{E}^{\mathcal{H}}_{adm}$. This establishes that no strict subset of S belongs to $\kappa^{\mathcal{H}}_{\text{ADM}}$ while S itself is in $\kappa^{\mathcal{H}}_{\text{ADM}}$. It follows that $S \in \mu^{\mathcal{H}}_{\text{ADM}}$ with $|S| = m = \lfloor\, n/2 \,\rfloor$. □

By developing consequences of the idea of "*conflict-sensitivity*" introduced in [10] we can, in fact, show that this lower bound is optimal, i.e. for every $\mathcal{H}(\mathcal{X}, \mathcal{A})$, $\max_{S\in\mu^{\mathcal{H}}_{\text{PR}}} |S| \leq \lfloor |\mathcal{X}|/2 \rfloor$. We omit the details on account of limited space.

4. Computational Complexity of Forbidden Set Problems

Given the formal definition of forbidden set it is easy to classify the complexity of membership in $\kappa^{\mathcal{H}}_{\sigma}$ on the basis of results from [6,9,11]. Thus,

Fact 1 *Given $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$ and $S \subseteq \mathcal{X}$ deciding if $S \in \kappa^{\mathcal{H}}_{\sigma}$ is*

$$\begin{array}{ll} \textit{in } \text{P} & \textit{if } \sigma \in \{\text{CF}, \text{NVE}, \text{GR}\} \\ \textit{co}\text{NP}\textit{–complete} & \textit{if } \sigma \in \{\text{ADM}, \text{PR}, \text{ST}\} \\ \Pi^p_2\textit{–complete} & \textit{if } \sigma \in \{\text{SST}\} \end{array}$$

The last two cases holding even if S contains just a single argument.

Proof: Polynomial time methods for $\sigma \in \{\text{CF}, \text{NVE}\}$ cases simply involve checking if $S \times S$ has a non-empty intersection with $\mathcal{A}$, i.e. some attack involves arguments in S. Similarly for grounded semantics $S \in \kappa^{\mathcal{H}}_{\text{GR}}$ if and only if S contains an argument not belonging to the grounded extension. This being efficiently computable deciding $S \in \kappa^{\mathcal{H}}_{\text{GR}}$ is also so. When $S = \{x\}$ (i.e. a single argument) the decision $S \in \kappa^{\mathcal{H}}_{\sigma}$ is simply a rephrasing of $\neg\text{CA}_{\sigma}(\mathcal{H}, x)$. The complexity classifcation for $\{\text{ADM}, \text{PR}, \text{ST}\}$ is now immediate from Dimopoulos and Torres [6] while that of $\{\text{SST}\}$ follows from Dvorak and Woltran [11]. □

While obtaining exact complexity results for deciding membership of $\kappa^{\mathcal{H}}_{\sigma}$ is straightforward using well-known results, the question of membership of the *minimal* forbidden sets turns out to be rather less so. Although the *single* argument instance $\{x\} \in \mu^{\mathcal{H}}_{\sigma}$ has identical complexity to its general counterpart $\{x\} \in \kappa^{\mathcal{H}}_{\sigma}$ for $\sigma \in \{\text{ADM}, \text{PR}, \text{SST}\}$ the reason for this is that $\mathcal{E}^{\mathcal{H}}_{\sigma} \neq \emptyset$ for these semantics. From which it follows that

$$(\{x\} \in \mu^{\mathcal{H}}_{\sigma}) \quad \Leftrightarrow \quad (\{x\} \in \kappa^{\mathcal{H}}_{\sigma}) \quad \Leftrightarrow \; \neg\text{CA}_{\sigma}(\mathcal{H}, x)$$

This argument, however, fails to apply whenever S contains at least two arguments. We can observe, however, that

$$S \in \mu^{\mathcal{H}}_{\sigma} \;\Leftrightarrow\; (S \in \kappa^{\mathcal{H}}_{\sigma}) \;\wedge\; \left(\bigwedge_{y\in S} S \setminus \{y\} \not\in \kappa^{\mathcal{H}}_{\sigma} \right)$$

That is we do not need to test *every* subset of S in order to confirm its membership of $\mu^{\mathcal{H}}_{\sigma}$.

Recalling that the complexity class D^p is defined by those decision problems, Q whose positive instance are both positive instances of some decision problem,

L_1 belonging to NP *and* positive instances of some decision problem, L_2, in coNP, the following holds for verifying membership of a given set S in $\mu_\sigma^{\mathcal{H}}$.

Theorem 2

a. For $\sigma \in \{$PR,ADM,COM$\}$, given $\langle S, \langle \mathcal{X}, \mathcal{A} \rangle \rangle$ deciding if $S \in \mu_\sigma^{\mathcal{H}}$ for the AF $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$ is D^p–complete, even for instances $\langle S, \langle \mathcal{X}, \mathcal{A} \rangle \rangle$ in which $|S| = 2$.

b. For stable semantics deciding $S \in \mu_{\mathrm{ST}}^{\mathcal{H}}$ is D^p–complete even with instances having $|S| = 1$.

Proof: (Outline) In the case of (a), we recall from Corollary 1(a) that $\kappa_{\mathrm{ADM}}^{\mathcal{H}} = \kappa_{\mathrm{PR}}^{\mathcal{H}} = \kappa_{\mathrm{COM}}^{\mathcal{H}}$ so it suffices to demonstrate the upper bound for $\sigma =$ ADM. Given $\langle S, \langle \mathcal{X}, \mathcal{A} \rangle \rangle$ with $S \subseteq \mathcal{X}$, $S \in \mu_{\mathrm{ADM}}^{\mathcal{H}}$ requires,

$$\exists \langle T_1, T_2, \ldots, T_k \rangle \ : \ T_i \subseteq \mathcal{X} \setminus S \ \text{ and } T_i \cup S \setminus \{y_i\} \in \mathcal{E}_{\mathrm{ADM}}^{\mathcal{H}}$$

capturing the condition that every strict subset of S can be extended to an admissible set. In addition, S itself must be a forbidden set, i.e.

$$\forall \ U \subseteq \mathcal{X} \setminus S \ S \cup U \notin \mathcal{E}_{\mathrm{ADM}}^{\mathcal{H}}$$

And now defining

$$\begin{array}{lcl} L_1 & = & \{ \ \langle \mathcal{X}, \mathcal{A}, S \rangle \ : \ \exists \langle T_1, \ldots, T_{|S|} \rangle \ T_i \cup S \setminus \{y_i\} \in \mathcal{E}_{\mathrm{ADM}}^{\mathcal{H}} \} \\ L_2 & = & \{ \ \langle \mathcal{X}, \mathcal{A}, S \rangle \ : \ \forall \ U \supseteq S \ U \notin \mathcal{E}_{\mathrm{ADM}}^{\mathcal{H}} \} \end{array}$$

we see that $S \in \mu_{adm}^{\mathcal{H}}$ if and only if $\langle \mathcal{H}, S \rangle \in L_1 \cap L_2$. Since $L_1 \in$ NP and $L_2 \in$ coNP we deduce $S \in \mu_{adm}^{\mathcal{H}}$ can be decided in D^p.

To establish D^p–hardness we present a reduction to instances $\langle \langle \mathcal{X}, \mathcal{A} \rangle, S \rangle$ from instances $\langle \varphi_1, \varphi_2 \rangle$ of the canonical D^p–complete problem SAT-UNSAT in which these are accepted if and only if the CNF, φ_1 is satisfiable and the CNF φ_2 is unsatisfiable. Given an instance $\langle \varphi_1, \varphi_2 \rangle$ of SAT-UNSAT $\mathcal{H}$ is formed by combining three copies of the "standard translation" of CNF formulae to AFs: two of these with designated arguments φ_1^1 and φ_1^2 capturing the structure of φ_1; the other, tied with the argument φ_2, linked with the structure of φ_2. The framework uses four additional arguments, $\{p_1, p_2, q_1, q_2\}$ which are configured in a directed cycle

$$\varphi_1^1 \ \rightarrow \ p_1 \ \rightarrow \ q_1 \ \rightarrow \ \varphi_1^2 \ \rightarrow \ q_2 \ \rightarrow \ p_2 \ \rightarrow \ \varphi_1^1$$

Finally the arguments $\{q_1, p_2\}$ are attacked by φ_2.

It can be shown that $\langle \varphi_1, \varphi_2 \rangle$ is accepted as an instance of SAT-UNSAT if and only if $\{\varphi_1^1, \varphi_1^2\} \in \mu_{\mathrm{PR}}^{\mathcal{H}}$, i.e. there are admissible sets, S_1 and S_2 for which $\varphi_1^1 \in S_1$ and $\varphi_1^2 \in S_2$, however no admissible set, S, with $\{\varphi_1^1, \varphi_1^2\} \subseteq S$.

We omit the proof of (b) due to space limitations. □

5. Conclusions

We have presented an alternative view of extension-based semantics within Dung's AF model: rather than describing solutions in terms of (positive) membership of a set we focus on capturing semantics by describing those sets which cannot form part of a solution. We have demonstrated the containment relationships between extension sets determine containments between the corresponding forbidden set structures and derived some preliminary complexity results on verification. To conclude we briefly mention some further directions. In addition to analogues of generic studies of extension based semantics within the forbidden set paradigm (e.g. realizability in the style of Dunne *et. al.* [10]) one has directions specific to the operations κ and μ defined earlier. In particular since $\kappa(\mathbb{S})$ and $\mu(\mathbb{S})$ are themeselves sets of subsets, in principle these operations could be iterated. While the structure of $\kappa(\kappa(\mathbb{S}))$ is uninteresting (being either $\emptyset$ or $2^{\mathcal{X}}$) that of $\mu(\mu(\mathbb{S}))$ appears non-trivial.

References

[1] P. Baroni, P. E. Dunne, and M. Giacomin. On the resolution-based family of abstract argumentation semantics and its grounded instance. *Artificial Intelligence*, 175:791–813, 2011.

[2] P. Baroni and M. Giacomin. Resolution-based argumentation semantics. In *Proc. 2nd COMMA*, volume 172 of *FAIA*, pages 25–36. IOS Press, 2008.

[3] P. Baroni, M. Giacomin, and G. Guida. SCC-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence*, 168(1–2):162–210, 2005.

[4] R. Baumann, T. Linsbichler, and S. Woltran. Verifiability of argumentation semantics. arXiv preprint arXiv:1603.09502, 2016.

[5] M. Caminada. Semi-stable semantics. In *Proc. 1st COMMA*, volume 144 of *FAIA*, pages 121–130. IOS Press, 2006.

[6] Y. Dimopoulos and A. Torres. Graph theoretical structures in logic programs and default theories. *Th. Comp. Sci.*, 170:209–244, 1996.

[7] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and N-person games. *Artificial Intelligence*, 77:321–357, 1995.

[8] P. M. Dung, P. Mancarella, and F. Toni. Computing ideal sceptical argumentation. *Artificial Intelligence*, 171:642–674, 2007.

[9] P. E. Dunne. The computational complexity of ideal semantics. *Artificial Intelligence*, 173(18):1559–1591, 2009.

[10] P. E. Dunne, W. Dvořák, T. Linsbichler, and S. Woltran. Characteristics of multiple viewpoints in abstract argumentation. *Artificial Intelligence*, 228:153–178, 2015.

[11] W. Dvořák and S. Woltran. Complexity of semi-stable and stage semantics in argumentation frameworks. *Inf. Process. Lett.*, 110(11):425–430, 2010.