On the Maximal and Average Numbers of Stable Extensions

Ringo Baumann and Hannes Strass

Institute of Computer Science, Leipzig University

Abstract We present an analytical and empirical study of the maximal and average numbers of stable extensions in abstract argumentation frameworks. As one of the analytical main results, we prove a tight upper bound on the maximal number of stable extensions that depends only on the number of arguments in the framework. More interestingly, our empirical results indicate that the distribution of stable extensions as a function of the number of attacks in the framework seems to follow a universal pattern that is independent of the number of arguments.

1 Motivation

Stable extensions constitute one of the most important and well-researched semantics for abstract argumentation frameworks (AFs). Dung used the stable extension semantics in his original paper to relate AFs to Reiter's default logic, different forms of logic programming, and to solve the stable marriage problem, among others [1]. Alas, there are some fundamental questions to be asked about stable extension semantics which have yet remained unanswered.

Given an abstract argumentation framework for which the only thing we know is that it has n arguments and x attacks, how many stable extensions does it have at most? How many on average?

For $x = 0$, without attacks, the case is quite clear – there will be exactly one stable extension, the set of all arguments. For $x = n^2$, the AF contains all possible attacks, in particular all self-attacks, and there will be no stable extension. But what happens in between, when $0 < x < n^2$?

This paper takes a step towards analytical and empirical answers to these questions. In particular, we develop predictions on the maximal and average number of stable extensions when only the number of arguments and attacks are known (and finite).

In the considerable zoo of semantics for abstract argumentation, stable extension semantics is the only one for which extension existence is not guaranteed for finite AFs. While this is usually regarded as a weakness, there is an obvious benefit to it when AFs are used to model NP-complete problems, that do not necessarily possess a solution. In this setting, the fact that an NP problem instance encoded as an AF has no stable extension elegantly reflects the fact that the problem instance has no solution. Using other semantics, unsolvability would have to be represented by introducing new (meta-)language constructs.

NP problems typically have elements that are generating (that is, generate possible solution candidates) and elements that are constraining (that is, eliminate possible solution candidates). The classical example of an NP-complete problem is of course deciding the satisfiability of a given propositional formula in conjunctive normal form, the SAT problem. There, the propositional variables are the generating elements (since solution candidates are among all interpretations for the variables) while the disjunctive clauses are the constraining elements (they remove those interpretations not satisfying some clause).

Can the same be said about arguments and attacks? Surely, arguments are generating, since extension candidates are sets of arguments. But are attacks always constraining?

This is unlikely, since classical propositional logic is a monotonic formalism, while argumentation frameworks are nonmonotonic. So while adding a clause to a CNF may never increase the number of models, adding attacks to an AF may in general both increase or decrease the number of stable extensions. This leads to the conjecture that there are numerical parameters of AFs (for example the numbers of attacks in relation to the number of arguments) where the average number of stable extensions is locally maximal and locally minimal, respectively.

Can we already hypothesize where such extrema can be found? Roughly, to be a stable extension, a set has to satisfy two properties: it has to be conflict-free, and has to attack all arguments not in the set. Intuitively, the number of attacks in an AF correlates negatively with the number of conflict-free sets – the more attacks (that is, conflicts) there are, the less conflict-free sets are found. At the same time, the number of attacks correlates positively with the number of sets which attack all outsiders. So how will these two interleaved and counteracting forces come to terms in general?

The paper is structured as follows. We next introduce the necessary background in graph theory and Dung's abstract argumentation frameworks. Then Section 3 presents our analytical results; Section 4 describes the results we obtained empirically. We conclude with a discussion of the results and give some perspectives on future work.

2 Background

Throughout the paper we assume some familiarity with standard analysis, combinatorics and statistics. For a set X, a *(binary) relation* over X is any set $R \subseteq X \times X$. Special among these relations is the *identity* $\mathrm{id}_X = \{(x, x) \mid x \in X\}$. A relation R over X is *irreflexive* iff $R \cap \mathrm{id}_X = \emptyset$, that is, for each $x \in X$ we have $(x, x) \notin R$. It is *symmetric* iff for each $(x, y) \in R$ we have $(y, x) \in R$. The *inverse* of a relation R is given by $R^{-1} = \{(y, x) \mid (x, y) \in R\}$.

2.1 Graph Theory

A *directed graph* is a pair (V, E) where V is a finite set and E a binary relation over V. The elements of V are called *nodes* and those of E are called *edges*.

A directed graph is *symmetric* iff its edge relation E is symmetric. For a directed graph $G = (V, E)$, we denote by $sym(G) = (V, E \cup E^{-1})$ its symmetric version. Similarly, the irreflexive version of a graph $G = (V, E)$ is defined as $irr(G) = (V, E \setminus \mathrm{id}_V)$.

An *undirected graph* is a pair (V, F) where V is as above and $F \subseteq \binom{V}{2} \cup \binom{V}{1}$ is a set of 2- and 1-element subsets of V, which represent the undirected edges. For a directed graph $G = (V, E)$, we denote by $und(G) = (V, \{\{u, v\} \mid (u, v) \in E\})$ its associated undirected graph. An undirected graph (V, F) is *simple* iff for no $v \in V$ we find $\{v\} \in F$. We denote by $\mathcal{G}_n$ the set of all simple graphs with n nodes.

For a simple graph $G = (V, F)$, a set $M \subseteq V$ is *independent* iff for all $u, v \in M$ we have $\{u, v\} \notin F$. A set $M \subseteq V$ is *maximal independent* iff it is independent and there is no proper superset of M which is independent. The set of all maximal independent sets of a simple graph G is denoted by $MIS(G)$.

2.2 Abstract Argumentation

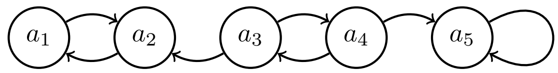
An *argumentation framework (AF)* $\mathcal{F} = (A, R)$ is a directed graph; the elements of A are also called *arguments* and the elements of R are also called *attacks*. All other graph theoretic notions carry over to AFs. A *full AF* is of the form $(A, A \times A)$ for some set A.

For the purposes of this paper, we denote by $\mathcal{A}_n$ the set of all AFs with n arguments, and by $\mathcal{A}_{n,x}$ the set of all AFs with n arguments and x attacks. There, not the precise arguments are of interest to us but only the *number* of arguments; we will implicitly assume that the n arguments can be numbered by $1, \ldots, n$. Once the arguments are fixed, however, we consider two AFs the same if and only if they have the same attack relation. So the AF with two arguments $1, 2$ where 1 attacks 2 is different from the AF with two arguments $1, 2$ where 2 attacks 1, although the two are isomorphic in a graph theoretic sense. This guarentees that all possible scenarios, that is, any arrangement of attacks for fixed numbers of arguments and attacks is considered.

The semantics of AFs is defined by determining those subsets $S \subseteq A$ which are acceptable according to specific criteria, so-called *extensions*. Among the various semantics from the literature, we are only interested in the stable semantics: a set $S \subseteq A$ is a *stable extension* for (A, R) iff (1) there are no $a, b \in A$ with $(a, b) \in R$, and (2) for all $a \in S \setminus A$, there is a $b \in S$ with $(b, a) \in R$. For an AF $\mathcal{F}$, the set of its stable extensions is denoted by $\mathcal{E}_{st}(\mathcal{F})$.

Interpreting the attack relation as denoting some kind of directed conflict between arguments, a stable extension can be seen as a set of arguments that is without internal conflict and attacks all arguments not contained in it. We call an argumentation framework a *y-AF* iff it has exactly y stable extensions. For the purpose of illustration consider the following example.

Example 1. Consider the following AF $\mathcal{F}$:



$\mathcal{F}$ has two stable extensions – $\mathcal{E}_{st}(\mathcal{F}) = \{\{a_1, a_4\}, \{a_2, a_4\}\}$ – thus $\mathcal{F}$ is a 2-AF.

3 Analytical Results

Baroni et al. [2] showed that counting the number of stable extensions is a computationally hard problem. The analysis of counting techniques may yield upper bounds for algorithms computing extensions. Furthermore, a fast counting algorithm gives a first advice on how controversial the information represented in an AF is. In this section, we contribute some analytical results to this direction of research.

For a fixed number n of arguments there are $|\mathcal{A}_n| = 2^{n^2}$ different AFs, since any attack relation whatsoever is possible and significant. Furthermore, if we additionally know that the AF in question possesses x attacks, then the total number of possibilities equals $|\mathcal{A}_{n,x}| = \binom{n^2}{x}$, the number of x-element subsets of an n^2-element set. This means that in principle, one may obtain numerically precise results by brute force for classes of AFs possessing a certain number of arguments and attacks. For example, specific classes of AFs could be enumerated and each element analyzed separately. But obviously, such an approach cannot provide a solution which is parametric in the numbers of arguments and attacks.

3.1 Maximal Number of Stable Extensions

What is the maximal number of stable extensions given an AF $\mathcal{F} = (A, R)$ with $|A| = n$ arguments? Since argumentation semantics choose their extensions from the set of subsets of A, we have $\mathcal{E}_{st}(\mathcal{F}) \subseteq 2^A$. This yields an immediate upper bound on the number of extensions for any semantics, namely $|\mathcal{E}_{st}(\mathcal{F})| \leq \left|2^A\right| = 2^n$. Can this quite naive bound be improved? In case of semantics satisfying I-maximality the answer is "yes." For short, I-maximality is fulfilled if no extension can be a proper subset of another [3]. In other words, the cardinality of one of the largest $\subseteq$-antichains S being a subset of an n-element set gives a further upper bound on the number of extensions. The maximal cardinality of such antichains is given by Sperner's theorem [4], namely $|S| = \binom{n}{\lfloor \frac{n}{2} \rfloor}$. By a straightforward calculation one may show that $\binom{n}{\lfloor \frac{n}{2} \rfloor} \leq \frac{2^n}{n}$. Without any further knowledge about the considered semantics it is impossible to find better bounds.

Let us turn to stable semantics. In any case, we can achieve a high number of stable extensions by grouping. For instance, the maximal number of stable extensions for an AF possessing an even number $n = 2m$ of arguments is at least $2^m = 2^{\frac{n}{2}}$. Such a framework is given by grouping the arguments in pairs that mutually attack each other:

$$\mathcal{F} = (\{a_i, b_i \mid 1 \leq i \leq m\}, \{(a_i, b_i), (b_i, a_i) \mid 1 \leq i \leq m\})$$

Is grouping in pairs the best we can do?

Assume we group not in pairs but in groups of arbitrary size k such that all members of a single group attack each other. Then for n arguments the number of stable extensions is given by the following function:

$$f : \mathbb{N} \to \mathbb{N} \text{ where } f(k) = k^{\lfloor \frac{n}{k} \rfloor}$$

To approximate the maximum of $f(k)$ we calculate the extrema of the associated real-valued function

$$g : \mathbb{R} \to \mathbb{R} \text{ where } g(k) = k^{\frac{n}{k}} = e^{\frac{n}{k} \cdot \ln(k)}$$

For that, we have to solve the following equation:

$$k^{\frac{n}{k}} \left(-\frac{n}{k^2} \cdot \ln(k) + \frac{n}{k^2} \right) = k^{\frac{n}{k}} \cdot \frac{n}{k^2} \cdot (1 - \ln(k)) = 0$$

The only solution for this equation is that k equals Euler's number e. Of course, it is very difficult to arrange in groups of e when dealing with arguments. Nevertheless, the obtained result provides an upper bound for the initial problem – namely the value $g(e) = e^{\frac{n}{e}}$ – assuming that grouping is the best. We will see that the exact value is not far away.

On the path to the main theorem we start with two simple observations which hardly need a proof. Being aware of this fact, we still present them in the form of a proposition to be able to refer to them later on. For one, whenever a set E is a stable extension of $\mathcal{F}$, then E is also a stable extension in the symmetric and self-loop free version of $\mathcal{F}$. Observe that the converse is not true in general.

Proposition 1. *For any argumentation framework $\mathcal{F} = (A, R)$ and any $E \in \mathcal{E}_{st}(\mathcal{F})$ we have $E \in \mathcal{E}_{st}(sym(irr(\mathcal{F})))$.*

For another, the second proposition establishes a simple relationship between stable extensions in symmetric AFs and maximal independent sets in undirected graphs.

Proposition 2. *For any symmetric argumentation framework $\mathcal{F} = (A, R)$ we have: $E \in \mathcal{E}_{st}(\mathcal{F})$ iff $E \in MIS(und(\mathcal{F}))$.*

Now we turn to the main theorem which is mainly based on a graph theoretical result by J.W. Moon and L. Moser from 1965 [5]. The theorem establishes a tight upper bound for the number of stable extensions of an AF with n arguments. The upper bound is obtained as a function $\sigma_{\max}$ of n.

Theorem 1. *For any natural number n, it holds that*

$$\max_{\mathcal{F} \in \mathcal{A}_n} |\mathcal{E}_{st}(\mathcal{F})| = \sigma_{\max}(n)$$

where the function $\sigma_{\max} : \mathbb{N} \to \mathbb{N}$ is defined by

$$\sigma_{\max}(n) = \begin{cases} 1, & \text{if } n = 0 \text{ or } n = 1, \\ 3^s, & \text{if } n \geq 2 \text{ and } n = 3s, \\ 4 \cdot 3^{s-1}, & \text{if } n \geq 2 \text{ and } n = 3s + 1, \\ 2 \cdot 3^s, & \text{if } n \geq 2 \text{ and } n = 3s + 2. \end{cases}$$

Proof. *The cases $n = 0$ and $n = 1$ are obvious. For $n \geq 2$ we show first that $\sigma_{\max}(n)$ does not exceed the specified values. We already observed that for any AF $\mathcal{F}$ we have $\mathcal{E}_{st}(\mathcal{F}) \subseteq \mathcal{E}_{st}(sym(irr(\mathcal{F})))$ (Proposition 1). Consequently, $|\mathcal{E}_{st}(\mathcal{F})| \leq |\mathcal{E}_{st}(sym(irr(\mathcal{F}))|$ follows and hence*

$$\max_{\mathcal{G} \in \mathcal{A}_n} |\mathcal{E}_{st}(\mathcal{G})| \leq \max_{\mathcal{G} \in \mathcal{A}_n} |\mathcal{E}_{st}(sym(irr(\mathcal{G}))|$$

In the light of Proposition 2 we get

$$\max_{\mathcal{G} \in \mathcal{A}_n} |\mathcal{E}_{st}(sym(irr(\mathcal{G})))| = \max_{\mathcal{G} \in \mathcal{A}_n} |MIS(und(sym(irr(\mathcal{G}))))|$$

Observe that the functions $irr(\cdot)$, $sym(\cdot)$ and $und(\cdot)$ do not change the number of nodes (respectively arguments). Consequently, we may estimate thus:

$$\max_{\mathcal{G} \in \mathcal{A}_n} |MIS(und(sym(irr(\mathcal{G}))))| \leq \max_{\mathcal{U} \in \mathcal{G}_n} |MIS(\mathcal{U})| .$$

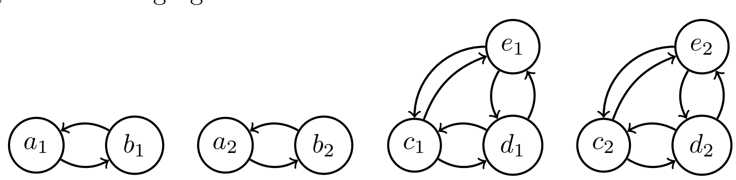
This means, the value $\sigma_{\max}(n)$ does not exceed the maximal number of maximal independent sets of simple undirected graphs of order n. Due to Theorem 1 in [5] these values are exactly given by the last three lines of the claimed value range of $\sigma_{\max}(n)$.

For completeness we describe examples proving that the bounds provided by $\sigma_{\max}(n)$ are tight. To this end we define:

- *$A_2(i) = \{a_i, b_i\}$ and $A_3(i) = \{c_i, d_i, e_i\}$,*
- *$\mathcal{F}_2(i) = irr(A_2(i), A_2(i) \times A_2(i))$ and $\mathcal{F}_3(i) = irr(A_3(i), A_3(i) \times A_3(i))$.*
- *For $n = 3s$ consider $\mathcal{F}_{3s} = \bigcup_{i=1}^{s} \mathcal{F}_3(i)$.*
- *For $n = 3s + 1$ consider $\mathcal{F}_{3s+1} = (\bigcup_{i=1}^{2} \mathcal{F}_2(i)) \cup (\bigcup_{i=1}^{s-1} \mathcal{F}_3(i))$.*
- *Finally, in case of $n = 3s + 2$ consider $\mathcal{F}_{3s+2} = \mathcal{F}_2(1) \cup (\bigcup_{i=1}^{s} \mathcal{F}_3(i))$.*

It is straightforward to verify that $|\mathcal{E}_{st}(\mathcal{F}_{3s})| = 3^s$, $|\mathcal{E}_{st}(\mathcal{F}_{3s+1})| = 4 \cdot 3^{s-1}$ and $|\mathcal{E}_{st}(\mathcal{F}_{3s+2})| = 2 \cdot 3^s$. □

For illustration we present here an instantiation of the presented prototypes, namely $\mathcal{F}_{10} = \mathcal{F}_{3 \cdot 3+1} = (\bigcup_{i=1}^{2} \mathcal{F}_2(i)) \cup (\bigcup_{i=1}^{2} \mathcal{F}_3(i))$ which is graphically represented by the following figure:



Observe that $|\mathcal{E}_{st}(\mathcal{F}_{10})| = |\mathcal{E}_{st}(\mathcal{F}_{3 \cdot 3+1})| = 4 \cdot 3^2$. In general, the function $\sigma_{\max}$ looks more complicated than it is, because the numbers are slightly different depending on the remainder of n on division by 3. Here is a much simpler version.

Corollary 1 (Upper bound short cut). *For any natural number n, we find:*

$$\sigma_{\max}(n) \leq 3^{\frac{n}{3}} \leq 1,4423^n .$$

As a final note we want to mention that it does not make much sense to ask for the minimal number of stable extensions, since for any $n > 0$ and $0 < x \leq n^2$ there are always AFs without stable extensions.

3.2 Average Number of Stable Extensions

What is the average number of stable extensions of argumentation frameworks with n arguments and x attacks?

As in the case of the maximal number of stable extensions, the precise value is computable in principle. This is immediate from its formal definition:

Definition 1. *The function $\bar{\sigma}(n,x)$ returns the average number of stable extensions of all AFs with n arguments and x attacks, and is defined thus:*

$$\bar{\sigma} : \mathbb{N} \times \mathbb{N} \to \mathbb{R} \text{ where } \bar{\sigma}(n,x) = \frac{\sum_{\mathcal{F} \in \mathcal{A}_{n,x}} |\mathcal{E}_{st}(\mathcal{F})|}{\binom{n^2}{x}}$$

While this definition makes it precise what we mean by "average number of stable extensions," computing this number for a given AF still remains as hard as computing all stable extensions of that AF.

But we are looking for a way to compute the number $\bar{\sigma}(n,x)$ *without actually inspecting at the AF* except for determining the parameters n and x. This would be useful since the number n of arguments and the number x of attacks can be determined in linear time, and knowing $\bar{\sigma}(n,x)$ gives some guidance on how many extensions a given AF $\mathcal{F} \in \mathcal{A}_{n,x}$ will have.

The best-case scenario would be the specification of a closed-form function that returns the exact values of $\bar{\sigma}(n,x)$. Unfortunately, the combinatorial blowup even in case of small numbers of attacks turns this endeavor into a challenging task. Nevertheless, we were able to specify certain values. The following proposition presents some exact values of $\bar{\sigma}(n,x)$ given that the number of attacks x is close to 0 or close to n^2.

Proposition 3. *For any $n \in \mathbb{N}$, we have*

$$\bar{\sigma}(n,0) = 1 \qquad \bar{\sigma}(n,n^2-3) = \begin{cases} \frac{3\cdot(n^2-n-1)}{(n+1)\cdot(n^2-2)}, & \text{if } n \geq 3, \\ 1 - \frac{1}{n}, & \text{if } n = 2 \\ 0, & \text{otherwise} \end{cases}$$

$$\bar{\sigma}(n,1) = \begin{cases} 1 - \frac{1}{n}, & \text{if } n \geq 1, \\ 0, & \text{otherwise} \end{cases} \qquad \bar{\sigma}(n,n^2-2) = \begin{cases} \frac{2}{n+1}, & \text{if } n \geq 2, \\ 0, & \text{otherwise} \end{cases}$$

$$\bar{\sigma}(n,2) = \begin{cases} 1 - \frac{2n-2}{n^2+n}, & \text{if } n \geq 2, \\ 0, & \text{otherwise} \end{cases} \qquad \bar{\sigma}(n,n^2-1) = \begin{cases} \frac{1}{n}, & \text{if } n \geq 1, \\ 0, & \text{otherwise} \end{cases}$$

$$\bar{\sigma}(n,n^2) = \begin{cases} 1, & \text{if } n = 0, \\ 0, & \text{otherwise} \end{cases}$$

Proof. *The values of $\bar{\sigma}(n,0)$ and $\bar{\sigma}(n,n^2)$ are obvious. Consider $\bar{\sigma}(n,1) = 1 - \frac{1}{n}$. This can be seen as follows: If the belonging attack is a self-loop, then we have no extensions. If it is not, then we have exactly one extension which is the union of all unattacked arguments. Obviously, we have $|\mathcal{A}_{n,1}| = \binom{n^2}{1} = n^2$ and* *furthermore, there are n different AFs in $\mathcal{A}_{n,1}$ possessing exactly one loop. Thus $\bar{\sigma}(n,1) = \frac{n^2-n}{n^2} = 1 - \frac{1}{n}$. Analogously one may prove $\bar{\sigma}(n, n^2-1) = \frac{1}{n}$.*

We want to emphasize that the other values are non-trivial. To get an idea of the complextity of the remaining proofs we consider the value $\bar{\sigma}(n, n^2-3)$. W.l.o.g. we may assume $n \geq 2$ since the number of attacks has to be non-negative. Furthermore we may even assume that $n \geq 3$ because if $n = 2$, then $\bar{\sigma}(n, n^2-3) = \bar{\sigma}(n,1)$ which is already solved. An AF $\mathcal{F} \in \mathcal{A}_{n,n^2-3}$ can be seen as the result of the following process: One starts with a full AF with n arguments. We then stepwise delete 3 attacks which are either loops or non-loops. We list now the probabilities to end up in an AF where k loops are deleted.

$$\begin{aligned}
P(k=3) &= 1 \cdot \frac{n}{n^2} \cdot \frac{n-1}{n^2-1} \cdot \frac{n-2}{n^2-2} \\
P(k=2) &= 3 \cdot \frac{n}{n^2} \cdot \frac{n-1}{n^2-1} \cdot \frac{n^2-n}{n^2-2} \\
P(k=1) &= 3 \cdot \frac{n}{n^2} \cdot \frac{n^2-n}{n^2-1} \cdot \frac{n^2-n-1}{n^2-2}
\end{aligned}$$

We omit the consideration of $P(k=0)$ since such kind of frameworks do not possess an extension and thus does not contribute anything to $\bar{\sigma}(n, n^2-3)$. We list now the average number of extensions of AFs in $\mathcal{A}_{n,n^2-3}$ where k loops are deleted.

$$\begin{aligned}
av(k=3) &= 3 \\
av(k=2) &= 1 \cdot \frac{2(n-1)}{n^2-n} + 2 \cdot \frac{(n^2-n)-2(n-1)}{n^2-n} \\
&= 2 \cdot \left(1 - \frac{1}{n}\right) \\
av(k=1) &= 1 - \left(\frac{n-1}{n^2-n} + \frac{(n^2-n)-(n-1)}{n^2-n} \cdot \frac{n-1}{n^2-n-1}\right) \\
&= \frac{n^2-3n+2}{n^2-n-1}
\end{aligned}$$

The average numbers can be seen as follows. If we delete exactly three loops we end up in an AF with 3 stable extensions, namely the singletons of the non-looping arguments. Consequently, $av(k=3) = 3$. If we delete 2 loops and 1 non-loop we either end up with 1 extension, namely if the deleted non-loop starts by an self-loop free argument or 2 extensions otherwise. The probability of the former is $\frac{2(n-1)}{n^2-n}$. Since both cases are mutual exclusive and exhaustive we derive a probability of $\frac{(n^2-n)-2(n-1)}{n^2-n}$ for the latter case proving the claimed value of $av(k=2)$.

Consider now $av(k=1)$. Observe that the maximal number of extensions equals 1 because only 1 self-loop is deleted. In the following we call this argument arg. We specify now the probability that we end up in AF with zero stable *extension. This is the case if at least one deleted non-loop starts by arg. The probability for the "first" non-loop is $\frac{n-1}{n^2-n}$. Furthermore, the probability for the "second" deleted non-loop to start by arg providing that the first one does not started by arg is given by $\frac{(n^2-n)-(n-1)}{n^2-n} \cdot \frac{n-1}{n^2-n-1}$. Thus, the claimed value for $av(k=1)$ follows. Finally, we have to sum up, that is,*

$$\bar{\sigma}(n, n^2 - 3) = \sum_{i=1}^{3} P(k=i) \cdot av(k=i) = 3 \cdot \frac{n^2 - n - 1}{(n+1)(n^2-2)}$$

We omit the consideration of $\bar{\sigma}(n,2)$ and $\bar{\sigma}(n,n^2-2)$ since their treatment is similar in style to the above proof. □

It can be seen that the values of $\bar{\sigma}(n,1)$ and $\bar{\sigma}(n,2)$ do not give any indication on how $\bar{\sigma}(n,3)$ could look like, not even qualitatively. The same holds for $\bar{\sigma}(n,n^2-2)$ and $\bar{\sigma}(n,n^2-3)$, and potential informed guesses about $\bar{\sigma}(n,n^2-4)$.

But having these exact values at hand we may consider the limit values for AFs with an increasing number of arguments. We have

$$\lim_{n\to\infty} \bar{\sigma}(n,0) = \lim_{n\to\infty} \bar{\sigma}(n,1) = \lim_{n\to\infty} \bar{\sigma}(n,2) = 1$$

On the other hand, we obtain

$$\lim_{n\to\infty} \bar{\sigma}(n,n^2) = \lim_{n\to\infty} \bar{\sigma}(n,n^2-1) = \lim_{n\to\infty} \bar{\sigma}(n,n^2-2) = \lim_{n\to\infty} \bar{\sigma}(n,n^2-3) = 0$$

This means that for increasing numbers of arguments, the average number of stable extensions in the case of very small numbers of attacks approaches from below to 1. In the case of very large numbers of attacks we have a convergence to 0 from above. So far, so good; but it is still unclear how many extensions there usually are in between. With an increasing number of attacks, does the average number of stable extensions just decrease in a monotone fashion? It turns out that this is a really hard problem.[1]

Of course, we can look at simple special cases. For example, for $n = 2$, Proposition 3 yields the precise values for all possible numbers of attacks $0 \leq x \leq n^2 = 4$: an AF with 2 arguments and $0, 1, 2, 3, 4$ attacks will have an average number of $1, \frac{1}{2}, \frac{2}{3}, \frac{1}{2}, 0$ stable extensions, respectively. So while the number of attacks linearly increases, the average number of extensions first decreases, then increases and then decreases again. Qualitatively speaking, this means that for a fixed number of arguments, there are certain numbers of attacks where the average number of extensions is locally maximal or minimal, respectively.

We have seen in the proofs of the results above that already the closed-form solutions for values of $\bar{\sigma}(n,2)$ and $\bar{\sigma}(n,n^2-3)$ are quite hard to obtain. To nevertheless get an inkling of the characteristic distribution of stable extensions, we have set out to study the problem in an empirical way.

[1] We therefore introduce the "average-number-stable-challenge" which is: present a closed-form function for $\bar{\sigma}(n,x)$ or at least specific values like $\bar{\sigma}(n,n^2-n)$ or $\bar{\sigma}(n,2n)$. The prize is a hot or cold drink with the authors.

4 Empirical Results

As we have seen, combinatorial explosion stood in our way of mathematically analyzing the average number of stable extensions. While the same combinatorial explosions prevent us from an exhaustive empirical analysis of the average number of stable extensions, we can still use methods from descriptive statistics to draw some meaningful conclusions.

The basic idea is simple: instead of computing the average number of stable extensions for *all* AFs in some class such as $\mathcal{A}_{n,x}$, we only analyse a uniformly drawn random sample $S \subseteq \mathcal{A}_{n,x}$ of a fixed size $|S|$. We thereby obtain a point estimation of the actual (hidden) parameter $\bar{\sigma}(n, x)$.

4.1 Experimental Setup

We wrote a program that randomly samples AFs with specific parameters and determines how many stable extensions they have. To create a random AF, we first set $A = \{1, \dots, n\}$. To create attacks we then randomly select x elements from the set $A \times A$ with equal probability for each pair. Thus we obtain an AF $\mathcal{F} = (A, R) \in \mathcal{A}_{n,x}$. For a given n, this process is repeated for all $0 \leq x \leq n^2$. Now for each AF thus created, we determine the number of stable extensions as follows: We use the translation of Dung [1, Section 5] to transform the AF into a logic program. By [1, Theorem 62], the stable models of this logic program and the stable extensions of the AF are in one-to-one-correspondence. Using the answer set solver `clingo` [6], we determine the number of stable models of the program and thus the number of stable extensions of the AF. So for a given n, we can empirically estimate the average number of stable extensions in each sample set of AFs with n arguments and x attacks for all $0 \leq x \leq n^2$.

4.2 Average Number of Stable Extensions

To check the experimental setup, we first ran the experiment with $n = 2$ and observed that the empirical results agreed with the predictions of Section 3.2. The results for $n = 20$ are depicted in a scatter plot, in Figure 1 on page 11; the results for $n = 50$ are plotted likewise in Figure 2, page 12.

The empirical data clearly vindicate our analytical predictions for very small and very large numbers of attacks. In between, the data furthermore confirm our predictions about the emergence of local minima and maxima. In addition to the experiments that are graphically depicted, we present the positions of these empirically obtained minima and maxima for several additional small n in Table 1.

For the local minimum and for small n, an approximation of the position $x_{\min}$ of the local minima from below is given by $n^2 - n \cdot \sqrt{n}$. More precisely – and astonishingly –, the position of the local maximum *always* coincides with $n^2 - n$. On an intuitive level, this suggests that removing n attacks from a full AF with n arguments quite probably leads to AFs for which *both adding and removing* attacks leads to a *decrease* in the number of stable extensions. To investigate

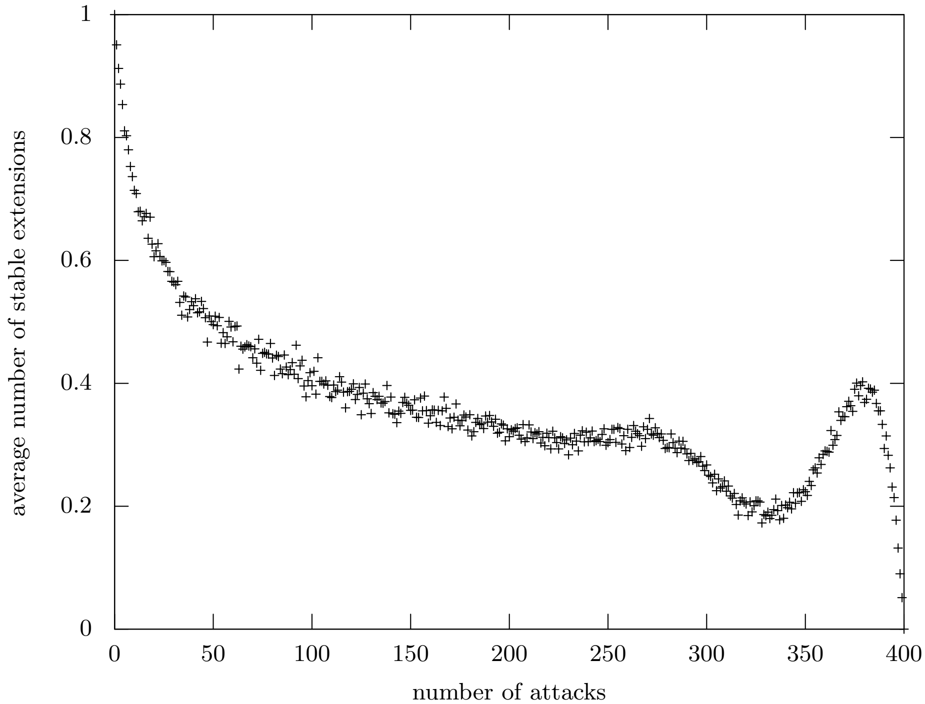


Figure 1: Average number of stable extensions of AFs with $n = 20$ arguments. The values have been obtained from a random sample of size 2500 for each possible number $0 \leq x \leq 400$ of attacks. (So the total sample size is 1 002 500.) We can see that there is a significant local minimum at $x_{\min} \approx 330$ and a local maximum at $x_{\max} \approx 380$.

this issue somewhat deeper, we next analysed how the average number of stable extensions came about.

4.3 Number of AFs with at most one stable extension

The point estimator *sample mean* we used for approximating $\bar{\sigma}(n, x)$ does not per se tell us anything about the distribution of 0-AFs, 1-AFs, ..., y-AFs among the AFs sampled.[2] In principle, an average number of 0.5 stable extensions could be obtained by a 50/50-ratio of 0-AFs to 1-AFs, or likewise by a 75/25-ratio of 0-AFs to 2-AFs. To find out what is the case, we extracted the absolute frequency of 0-AFs and 1-AFs from our results for $n = 50$ and plotted them in the stacked histogram on page 14.

[2] Recall that a y-AF is an AF with exactly y stable extensions.

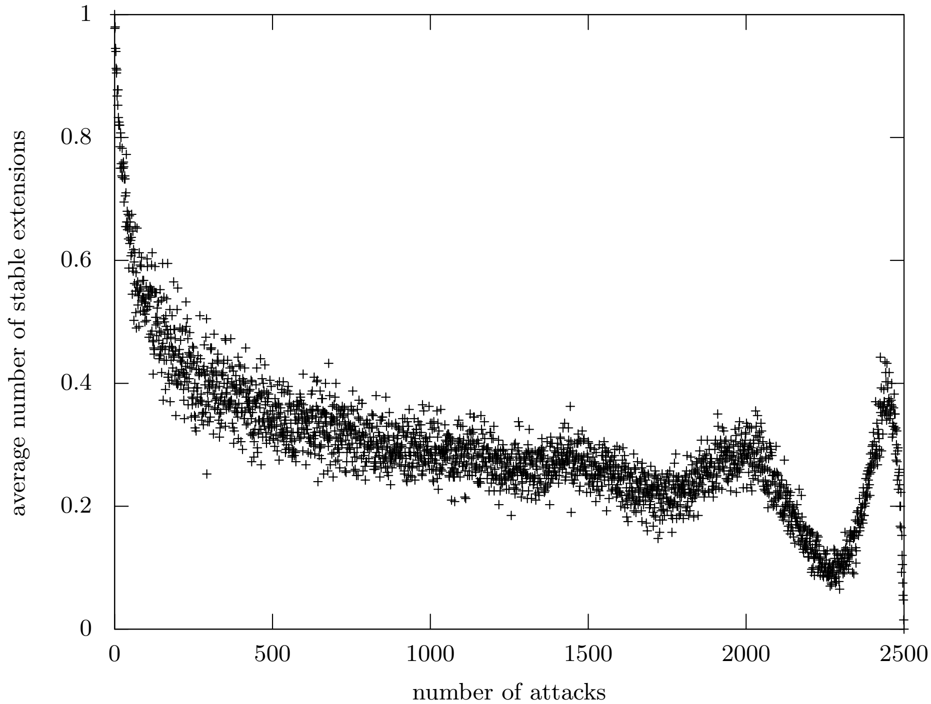


Figure 2: Average number of stable extensions of AFs with $n = 50$ arguments and sample size 400 for each $0 \leq x \leq 2500$. Again, there are significant extrema: a local minimum at $x_{\min} \approx 2250$ and a local maximum at $x_{\max} \approx 2450$. It even seems that there is another local maximum at $x'_{\max} \approx 2000$ and another local minimum before that, but the data are unreliable. (Recall that for $x = 2000$ the number of AFs to sample from is $|\mathcal{A}_{50,2000}| = \binom{2500}{2000} \geq \left(\frac{2500}{2000}\right)^{2000} \approx 6.6 \cdot 10^{193}$.)

The stacked histogram for $n = 20$ looks alike, indeed as much as the scatterplots in Figures 1 and 2 do. This suggests that there are certain recurring features in this distribution that are independent of the number n of arguments.

It cannot be seen in the histogram, but we also observed that for any set of sampled AFs from $\mathcal{A}_{50,x}$ with $0 \leq x \leq 50^2$, there are typically more 1-AFs than 2-AFs, more 2-AFs than 3-AFs, and so on. This gives some hints about the sizes of the subclasses of 1-AFs, 2-AFs, ... in a given class $\mathcal{A}_{n,x}$.

We close the empirical section by presenting two conjectures supported by the obtained results. The first one is concerned with the cardinality of y-AFs for a fixed number n of arguments.

Conjecture 1. For any natural numbers n, k and l with $0 < k < l \leq n$ we have:

$$|\{\mathcal{F} \mid \mathcal{F} \in \mathcal{A}_n,\ \mathcal{F} \text{ is a } k\text{-AF}\}| \geq |\{\mathcal{G} \mid \mathcal{G} \in \mathcal{A}_n,\ \mathcal{G} \text{ is an } l\text{-AF}\}|\,.$$

n	2	3	4	5	6	7	8	9	10
$x_{\min}$	1	4	9	15	23	32	45	57	73
$n^2 - n \cdot \sqrt{n}$	1.17	3.80	8	13.82	21.30	30.48	41.37	54	68.38
e_{abs}	0.17	0.2	1	1.18	1.7	1.52	3.63	3	4.62
e_{rel}	0.17	0.04	0.11	0.08	0.07	0.05	0.08	0.05	0.06
$x_{\max}$	2	6	12	20	30	42	56	72	90
$n^2 - n$	2	6	12	20	30	42	56	72	90
e_{abs}	0	0	0	0	0	0	0	0	0
e_{rel}	0	0	0	0	0	0	0	0	0

Table 1: Positions (at a specific number x of attacks) of empirically observed local minima (denoted by $x_{\min}$) and maxima ($x_{\max}$) of the average number of stable extensions of AFs with n arguments. We additionally present the values of our analytical estimations. To approximate the position of the minima, we devised the function $n^2 - n \cdot \sqrt{n}$; for the maxima we obtained $n^2 - n$. The rows labelled by e_{abs} and e_{rel} show the absolute and relative error of these estimates.

The second conjecture claims that the average number of stable extensions of AFs is always located in between 0 and 1. Here is the precise formulation.

Conjecture 2. For any natural numbers n and x with $0 < x < n^2$ we have:

$$0 < \bar{\sigma}(n, x) < 1.$$

5 Discussion

We have conducted a detailed analytical and empirical study on the maximal and average numbers of stable extensions in abstract argumentation frameworks. First of all, we have proven a tight upper bound on the maximal number of stable extensions. For specific numbers of attacks, we have also given the precise average number of stable extensions in terms of closed-form expressions. As the calculation of these analytical values tends to be quite complex, we turned to

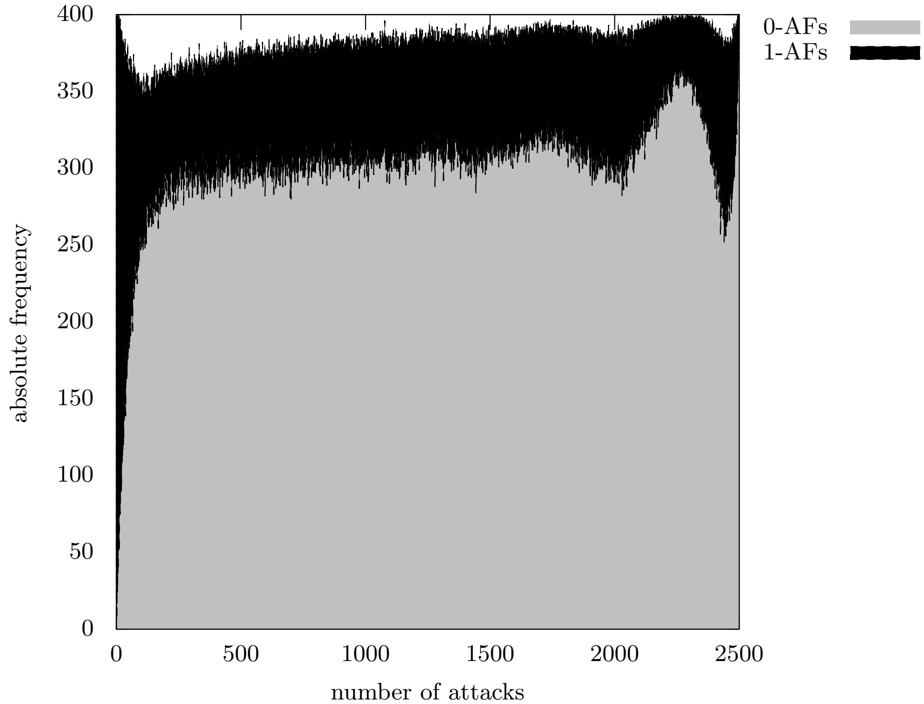


Figure 3: Absolute frequencies of 0*-AFs (grey) and* 1*-AFs (black) among all AFs with* $n = 50$ *arguments and* x *attacks for* $0 \leq x \leq n^2 = 2500$ *with a total sample size of* $1\,000\,400$*. It is obvious from the histogram that the majority (at least two thirds) of all sampled AFs have no stable extension. Additionally, almost all AFs have at most one stable extension. The white area at the top consequently depicts the* y*-AFs for* $y \geq 2$*. For* $x \approx 100 = 2n$*, there is a meaningful number of such* y*-AFs, which however decreases with increasing* x*. (Note that the extremal graphs defined in Theorem 1 have* n *arguments and* $2n$ *attacks.) At* $x \approx 2250$*, where the average number of stable extensions has a local minimum, the absolute frequency of* 0*-AFs has a local maximum; furthermore at this position there are almost no* y*-AFs for* $y \geq 2$*. Conversely, at* $x \approx 2450$ *where the average number of stable extensions has a local maximum, the absolute frequency of* 0*-AFs has a local minimum; furthermore there are yet again* y*-AFs for* $y \geq 2$*.*

studying the problem empirically. There, we obtained data about the distribution of stable extensions in samples of AFs which were randomly drawn with a uniform probability. Our empirical results offer new insights into the average number and also the distribution of stable extensions for AFs, given only the parameters n (number of arguments) and x (number of attacks).

We could not provide exhaustive theoretical explanations for the many empirical observations we have made, and consider this as one of the major future directions of this research. First and foremost we consider it important to work on proving or disproving the conjectures we explicitly formulated at the end of the previous section. Also the conjectured local maximum of the average number of stable extensions at $n^2 - n$ attacks deserves some attention. A possible way to tackle these conjectures may be to look at subclasses of AFs with special structural properties, such as having no self-loops, or more generally no cycles, those being symmetric, or the ones with a specific average connectivity. Finally, it is clear that many of the questions we asked about stable extension semantics can be asked about the other standard semantics.

Note that our results are not only of interest to the argumentation community: We have seen in the proof of Theorem 1 that there is a close relationship between stable extensions of AFs and maximal independent sets of undirected graphs.[3] In a sense, stable extensions represent a directed generalization of maximal independent sets, where the $\subseteq$-maximality condition has been replaced by the condition that all nodes not in the set must be reached by a directed edge from the set. So there is also a graph theoretical significance to our results.

For abstract argumentation, our results show that – in the context of stable semantics – attacks cannot simply be thought of as constraining: adding an attack may sometimes increase and sometimes decrease the number of stable extensions. Although this might be obvious in general to argumentation researchers (AFs are, after all, a nonmonotonic formalism), for the first time we were able to present some precise numerical figures around this phenomenon.

References

1. Dung, P.M.: On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. Artificial Intelligence **77** (1995) 321–358
2. Baroni, P., Dunne, P.E., Giacomin, M.: On extension counting problems in argumentation frameworks. In: Proceedings of COMMA. (2010) 63–74
3. Baroni, P., Giacomin, M.: On principle-based evaluation of extension-based argumentation semantics. Artificial Intelligence **171**(10-15) (2007) 675–700
4. Sperner, E.: Ein Satz über Untermengen einer endlichen Menge. Mathematische Zeitschrift **27**(1) (1928) 544–548
5. Moon, J., Moser, L.: On cliques in graphs. Israel Journal of Mathematics **3**(1) (1965) 23–28
6. Gebser, M., Kaminski, R., Kaufmann, B., Ostrowski, M., Schaub, T., Schneider, M.: Potassco: The Potsdam Answer Set Solving Collection. AI Communications **24**(2) (2011) 105–124 Available at `http://potassco.sourceforge.net`.

[3] Indeed, maximal independent sets are sometimes called "stable sets" in the graph theory literature.