A note on the semi-stable semantics of abstract argumentation systems

Bart Verheij
Artificial Intelligence, University of Groningen

Abstract

The paper adds to the formal analysis of argumentation, following up on Dung's abstract approach. The focus is on a specific kind of semantics, the so-called semi-stable extensions, the properties of which are reviewed. Special attention is paid to the existence of semi-stable extensions. An example is provided showing that not all argumentation frameworks have a semi-stable extension. The counterexample is rather involved, but the complexity is warranted by the property that, if an attack graph has no semi-stable extension, then there is an infinite sequence of preferred extensions with strictly increasing ranges.

1 Introduction

The formal study of argumentation is flourishing (e.g., Pollock 1994, Nute 1994, Dung 1995, Prakken & Sartor 1996, Bondarenko et al. 1997, Besnard & Hunter 2001, Verheij 2003, García & Simari 2004, Amgoud et al. 2008). Dung's abstract approach (1995) has been especially influential. In Dung's work, the focus is on the mathematical properties of one aspect of argumentation, namely the attack relation between arguments. Dung's analysis of the attack relation uses sets as a central tool. He proposed four kinds of extensions of an argumentation framework: stable, preferred, grounded and complete extensions. Verheij (1996) continued the analysis using labelings. He defined labeling analogues of stable and preferred extensions, and added two new kinds of extensions, arising naturally in the setting of labelings: stage extensions and admissible stage extensions. Instead of maximizing the set of arguments, the set of labeled arguments was maximized. In a sense, this meant that the set of arguments taken into account was maximized (whether defeated or not), instead of the set of undefeated arguments. Verheij (2003) continued the labeling analysis of argumentation, but in a more expressive setting, namely one in which both support and attack can be analyzed (cf. also Verheij 1999, Amgoud et al. 2008). Recently, Caminada (2006b) has resumed the analysis of argumentation frameworks in terms of labelings. In Caminada's work, Verheij's admissible stage extensions (1996) occur by the elegant name of semi-stable extensions.

The present paper focuses on these semi-stable extensions, emphasizing the problem of their existence. Section 2 provides the core definitions of the set and labeling approaches. Section 3 reviews results on the semi-stable semantics in different publications and gives connections with other argumentation semantics. In section 4, the focus is on the existence of semi-stable extensions. An argumentation framework is provided that has no semi-stable extension. The counterexample to the existence of semi-stable extensions is rather complex, which is warranted by the theorem that, if an attack graph has no semi-stable extension, then there is an infinite sequence of preferred extensions with strictly increasing ranges.

2 Analyzing the attack relation in terms of sets and in terms of labelings

The starting point of Dung's (1995) work is an argumentation framework, which is essentially a directed graph expressing the attack relations between arguments:

Definition (1). An *argumentation framework* is a pair (*Arguments*, *Attacks*), where *Arguments* is any set, and *Attacks* is a subset of *Arguments* × *Arguments*. The elements of *Arguments* are the arguments of the theory, the elements of *Attacks* the attacks.

When (*Arg*, *Arg*') is an attack, the argument *Arg* is said to *attack* the argument *Arg*'. A set of arguments *Args* is said to *attack* an argument *Arg* if and only if there is an element of *Args* that attacks *Arg*.

Some of Dung's central notions are the following:

Definition (2). 1. A set of arguments *Args* is *conflict-free* if it contains no arguments *Arg* and *Arg*', such that *Arg* attacks *Arg*'.
2. An argument *Arg* is *acceptable* with respect to a set of arguments *Args* if for all arguments *Arg*' in the argumentation framework the following holds:

> If *Arg*' attacks *Arg*, then there is an argument *Arg*'' in *Args*, such that *Arg*'' attacks *Arg*'.

Technical report, January 2009; minor revisions June 2010.

3. A set of arguments *Args* is *admissible* if it is conflict-free and all arguments in *Args* are acceptable with respect to *Args*.
4. An admissible set of arguments *Args* is a *complete extension* if each argument that is acceptable with respect to *Args* is an element of *Args*.
5. A *preferred extension* of an argumentation framework is an admissible set of arguments, that is maximal with respect to set inclusion.
6. A conflict-free set of arguments *Args* is a *stable extension* of an argumentation framework if for any argument *Arg* of the framework that is not in *Args*, there is an argument *Arg'* in *Args*, such that *Arg'* attacks *Arg*.

The complete extension that is minimal with respect to set inclusion (which exists and is unique; see Dung 1995) is called the *grounded extension*.

Central definitions of the labeling approach are as follows (Verheij 1996, 2007):

Definition (3). A pair (J, D) is a *labeling* if J and D are disjoint subsets of the set *Arguments* of the argumentation framework. The elements of J and D are the *justified* and *defeated* arguments, respectively. The elements of $J \cup D$ are *labeled*, other elements of *Arguments unlabeled*. The set $J \cup D$ is the *range* of the labeling.

The following definition contains the main notions of the labeling approach.

Definition (4). 1. A labeling (J, D) is *conflict-free* if the set J is conflict-free.
2. A labeling (J, D) has *justified defeat* if for all elements *Arg* of D there is an element in J that attacks *Arg*.
3. A labeling (J, D) is *closed* if all arguments that are attacked by an argument in J are in D.
4. A conflict-free labeling (J, D) is *attack-complete* if all attackers of arguments in J are in D.
5. A conflict-free labeling (J, D) is *defense-complete* if all arguments of which all attackers are in D are in J.
6. A conflict-free labeling (J, D) is *complete* if it is both attack-complete and defense-complete.
7. A labeling (J, D) is a *stage* if it is conflict-free and has justified defeat.

Caminada's (2006b) reinstatement labelings are closed complete labelings with justified defeat. The set of labelings of an argumentation framework AF is denoted as $\text{Labelings}_{\text{AF}}$.

The following properties summarize the relations between the set and labeling approach.

Properties (5). Let J be a set of arguments and D be the set of arguments attacked by the arguments in J. Then the following properties obtain:
1. J is conflict-free if and only if (J, D) is a conflict-free labeling.
2. J is admissible if and only if (J, D) is an attack-complete stage.
3. J is a complete extension if and only if (J, D) is a complete stage.
4. J is a preferred extension if and only if (J, D) is an attack-complete stage with maximal set of justified arguments.
5. J is a stable extension if and only if (J, D) is a labeling with no unlabeled arguments.

Proof. 1 follows by checking the definitions.
2. If J is admissible, it is conflict-free and attacks all arguments attacking it. Hence (J, D) is conflict-free, has justified defeat and is attack-complete. If (J, D) is an attack-complete stage, J is conflict-free and attacks all arguments attacking it. Hence J is admissible.
3. If J is a complete extension, it is admissible, hence (J, D) is an attack-complete stage. Moreover, all arguments of which all attackers are attacked by J are already in J. This is another way of saying that (J, D) is defense-complete. The other way around: If (J, D) is a complete stage, J is admissible (by part 2). J also contains all arguments acceptable with respect to J: Let *Arg* be acceptable with respect to J. Then all attackers of *Arg* are attacked by J. Since (J, D) is attack-complete, these attackers are all in D. The defense-completeness of (J, D) then implies that *Arg* is in J.
4 and 5 follows from the parts 2 and 3 and the definitions.

3 The semi-stable semantics and its connections to other argumentation semantics

Semi-stable extensions (an elegant term coined by Caminada 2006b) are admissible sets of arguments, for which the union of the set with the set of arguments attacked by it is maximal.[1] They have been introduced by Verheij (1996), in an analysis of Dung-style attack graphs (Dung 1995) in terms of - what are now referred to as - labelings.[2] Verheij (1996) uses the term "admissible stage extensions" for semi-stable extensions. Amongst other things, the following four central connections with Dung's stable and preferred extensions are shown:

1. Stable extensions are semi-stable.
2. Semi-stable extensions are preferred.
3. Preferred extensions are not always semi-stable (example in section 4.4 of Verheij 1996).
4. Semi-stable extensions are not always stable (example in section 4.3 of Verheij 1996).

Since preferred extensions exist for all attack graphs, while there exist attack graphs without a stable extension (Dung

[1] Caminada (2006b) shows that a semi-stable extension can also be defined as a complete extension, for which the union of the set with the set of arguments attacked by it is maximal.

[2] See Caminada (2006a, 2007) and Verheij (2007) for recent uses of the labeling approach. Other work on labelings, but four-valued, was performed by Jakobovits & Vermeir 1999 and Jakobovits 2000.

Dung (1995)	Verheij (1996)	Verheij (2000, 2003)	Caminada (2006b)	Encompassing proposal
stable extension	complete stage extension	extension, dialectical interpretation	stable extension	stable extension
preferred extension	preferred stage	dialectically preferred stage	preferred extension	preferred extension
grounded extension	-	-	grounded extension	grounded extension
complete extension	-	-	complete extension	complete extension
-	admissible stage extensions	maximal dialectically preferred stages	semi-stable extension	semi-stable extension
-	stage extension	maximal stage	-	stage extension
-	-	compatibility class (in Verheij (2000): satisfiability class)	-	conflict-free extension

Table 1: Comparison of terminology

1995), it is natural to consider the question whether all attack graphs have a semi-stable extension. This question was answered negatively by Verheij (2000, 2003). The attack graph of example 5.8 (Verheij 2003, p. 338)[3] has no semi-stable extension.[4] The result is obtained using the DefLog language, a straightforward generalization of Dung's attack graphs. DefLog[5] is a logical language in which attack is interpreted as a kind of conditional relation. The language adds support, nested conditionals and - what might be called - negation-as-defeat[6] to the expressiveness of Dung's attack graphs. Analogues of Dung's stable and preferred extensions are defined, and shown to be faithful generalizations (in the sense that translating an attack graph into DefLog does not affect its stable and preferred extensions). Next to the semi-stable semantics, Verheij (1996, 2003) adds a second kind of semantics that is new with respect to Dung's definitions, namely the stage semantics. A stage extension is a conflict-free set of arguments, for which the union of the set with the set of arguments attacked by it is maximal (Verheij 1996).[7] For the sake of completeness of the analysis, Verheij (2003) adds maximal conflict-free sets to the comparative analysis (using the term "compatibility class"). Table 1 contains an overview of the different uses of terminology.

4 An attack graph without semi-stable extension

Here are some elementary facts about the existence of semi-stable extensions:

1. There exist attack graphs without a semi-stable extension.
2. Finite attack graphs always have a semi-stable extension.
3. An attack graph with a finite number of preferred extensions has a semi-stable extension.
4. An attack graph with a stable extension has a semi-stable extension.

Verheij's example (2003, example 5.8, p. 338) showing that semi-stable extensions do not exist for all attack graphs is not the simplest possible. Here another example is provided, that is perhaps somewhat more transparent. The key idea is that - by the following theorem - we must look for an infinite series of preferred extensions with strictly increasing ranges.[8]

Theorem. If an attack graph has no semi-stable extension, then there is an infinite sequence of preferred extensions with strictly increasing ranges.

Proof. Pick a preferred extension P_0 of the attack graph. It is not semi-stable, so there is an admissible set A_1 with larger range (i.e., the range of A_1 is a proper superset of the range of P_0). There exists a preferred extension $P_1 \supseteq$ A1. P_1 has larger range than P_0. P_1 is not semi-stable, so (using the same reasoning) there is a preferred extension P_2 with larger range. Repeating this process gives (by induction) an infinite sequence of preferred extensions with strictly increasing ranges.

Example 5.8 given by Verheij (2003, p. 338) uses this criterion for the non-existence of semi-stable extensions. The following is a perhaps somewhat more transparent example.

[3] Example 7.12 in Verheij (2000).

[4] Somewhat confusingly, Verheij (2003) refers to semi-stable extensions using a different term than the 1996 term "admissible stage extensions". In 2003, semi-stable extensions are called "maximal dialectically preferred stages".

[5] Verheij (2003) is based on a technical report containing extensive additional material (Verheij 2000). The first publication on DefLog is Verheij (2002).

[6] Verheij (2003) speaks of 'dialectical negation'.

[7] Verheij (2003) refers to stage extensions as "maximal stages".

[8] The range of a conflict-free set of arguments is defined as the union of the set with the set of arguments attacked by it (Verheij 1996).

Example: An attack graph without semi-stable extension. Consider the following attack graph:

$p_0, p_1, p_2, p_3, \ldots$
$q_0, q_1, q_2, q_3, \ldots$
$r_0, r_1, r_2, r_3, \ldots$
q_i ~> xp_i
q_i ~> xq_i
r_i ~> xq_j $(i \geq j)$
r_i ~> xr_j $(i \neq j)$

In words: Each p is attacked by one q, the one with corresponding index. Each q attacks itself. Each r attacks all qs with equal or lower index. Each r is attacked by all other rs.

Let P be a preferred extension. Then:

1. All conditional sentences are in P as they cannot be attacked.
2. If P contains r_i for some i, then no r_j $(j \neq i)$ is also in P, for otherwise P would not be conflict-free.
3. If P contains r_i for some i, then P is the set consisting of r_i, all p_j with $j \leq i$ and all conditional sentences. Proof: Step I. P is admissible since it defends all its elements against their attackers: r_i defends itself against all its attackers (the other r_j), p_i is defended by r_i against its only attacker q_i, while the conditional sentences need no defense since they are not attacked. Step II. P is maximal since no other r_j than r_i are in P (see 2 above). No p_j with $j > i$ can be in P for such a p_j would need defense against q_j, which can only be provided by r_k with $k > j$ and such an r_k is not in P. Since the q_j are self-attacking they cannot be in P.
4. P contains an r_i for some i. Proof: Assume that P contains no r_i. Then P would contain all conditional sentences since they need no defense. No p_i is in P as that would require defense by r_i. No q_i is in P as they are self-attacking. In other words, P would consist of the conditional sentences, but that is not a maximal admissible set since it is properly contained in the preferred extensions described under 3.

By 3 and 4, we find that the sets P_i consisting of r_i, all p_j with $j \leq i$ and all conditional sentences (as they occur in 3) are the preferred extensions of the attack graph. The range of P_i is the set consisting of all r_j, all p_k and q_k with $k \leq i$ and all conditional sentences. As a result, the range of P_i is properly contained in the range of P_j when $i < j$. As the P_i are all preferred extensions and none of these has a range containing the range of all others, we find that the attack graph has no semi-stable extension. QED

References

Amgoud, L., Cayrol, C., Lagasquie-Schiex, M.C., & Livet, P. (2008). On Bipolarity in Argumentation Frameworks. *International Journal of Intelligent Systems* 23 (10), 1062-1093.

Besnard, P., & Hunter, A. (2001). A Logic-Based Theory of Deductive Arguments. *Artificial Intelligence* 128 (1-2), 203-235.

Bondarenko, A., Dung, P.M., Kowalski, R.A., & Toni, F. (1997). An Abstract, Argumentation-Theoretic Approach to Default Reasoning. *Artificial Intelligence* 93, 63-101.

Caminada, M. (2006a). On the Issue of Reinstatement in Argumentation. *Logics in Artificial Intelligence, 10th European Conference, Jelia 2006, Liverpool, Uk, September 13-15, 2006, Proceedings. Lecture Notes in Computer Science, Vol. 4160* (eds. M. Fisher, W. van der Hoek, B. Konev & A. Lisitsa), 111-123. Berlin: Springer.

Caminada, M. (2006b). Semi-Stable Semantics. *Computational Models of Argument: Proceedings of Comma 2006* (eds. P.E. Dunne & T.J.M. Bench-Capon), 121-130. Amsterdam: IOS Press.

Caminada, M. (2007). An Algorithm for Computing Semi-Stable Semantics. *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 9th European Conference, Ecsqaru 2007, Hammamet, Tunisia, October 31 - November 2, 2007, Proceedings. Lecture Notes in Computer Science, Vol. 4724* (ed. K. Mellouli), 222-234. Berlin: Springer.

Dung, P.M. (1995). On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and N-Person Games. *Artificial Intelligence* 77, 321-357.

García, A.J., & Simari, G.R. (2004). Defeasible Logic Programming: An Argumentative Approach. *Theory and Practice of Logic Programming* 4 (1-2), 95-138.

Jakobovits, H. (2000). *On the Theory of Argumentation Frameworks*. Brussel: Dissertation Vrije Universiteit.

Jakobovits, H., & Vermeir, D. (1999). Robust Semantics for Argumentation Frameworks. *Journal of Logic and Computation* 9 (2), 215-261.

Nute, D. (1994). Defeasible Logic. *Handbook of Logic in Artificial Intelligence and Logic Programming. Volume 3. Nonmonotonic Reasoning and Uncertain Reasoning* (eds. D.M. Gabbay, C.J. Hogger & J.A. Robinson), 353-395. Oxford: Clarendon Press.

Pollock, J.L. (1994). Justification and Defeat. *Artificial Intelligence* 67, 377-407.

Prakken, H., & Sartor, G. (1996). A Dialectical Model of Assessing Conflicting Arguments in Legal Reasoning. *Artificial Intelligence and Law* 4, 331-368.

Verheij, B. (1996). Two Approaches to Dialectical Argumentation: Admissible Sets and Argumentation Stages. *Naic'96. Proceedings of the Eighth Dutch Conference on Artificial Intelligence* (eds. J.-J.C. Meyer & L.C. van der Gaag), 357-368. Utrecht: Universiteit Utrecht. A preliminary version was presented at the Computational Dialectics Workshop at FAPR-96. June 3-7, 1996, Bonn. .

Verheij, B. (1999). Automated Argument Assistance for Lawyers. *The Seventh International Conference on Artificial Intelligence and Law. Proceedings of the Conference*, 43-52. New York (New York): ACM.

Verheij, B. (2000). DefLog - a Logic of Dialectical Justification and Defeat. Technical Report.

Verheij, B. (2002). On the Existence and the Multiplicity of Extensions in Dialectical Argumentation. *Proceedings of the 9th International Workshop on Non-Monotonic Reasoning (Nmr'2002)* (eds. S. Benferhat & E. Giunchiglia), 416-425. Toulouse.

Verheij, B. (2003). DefLog: On the Logical Interpretation of Prima Facie Justified Assumptions. *Journal of Logic and Computation* 13 (3), 319-346.

Verheij, B. (2007). A Labeling Approach to the Computation of Credulous Acceptance in Argumentation. *Ijcai 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007* (ed. M.M. Veloso), 623-628. http://www.ijcai.org/proceedings07.php.