# Corrupt Strategic Argumentation: The Ideal and the Naive

Michael J. Maher(✉)

School of Engineering and Information Technology,
University of New South Wales, Canberra, ACT 2600, Australia
michael.maher@unsw.edu.au

**Abstract.** Previous work introduced a model of corruption within strategic argumentation, and showed that some forms of strategic argumentation are resistant to two forms of corruption: collusion and espionage. Such a result provides a (limited) basis on which to trust agents acting on our behalf. That work addressed several argumentation semantics, all built on the notion of admissibility. Here we continue this work to three other well-motivated semantics: the ideal, naive, and stage semantics. The latter two are not admissibility-based. We show that the naive semantics does not support strategic argumentation, in the sense that the outcome of the game is determined by the initial state, if the players are not corrupt. As a result, the semantics is corruption-proof. We show that the ideal semantics is resistant to both collusion and espionage. The stage semantics is resistant to espionage, but its resistance to collusion depends on the strategic aims of the players.

## 1 Introduction

Strategic argumentation provides a simple model of disputation and negotiation among agents. Agents are intended to act on our behalf but – whether they are human or software – we cannot be sure that they are acting in our best interests. Social structures, including criminal sanctions, are used to discourage corruption by human agents. For computational agents, [2] introduced another way corruption can be discouraged: if the computational requirements to take advantage of the corruption are too great then there is no incentive to act corruptly. [13–15] adapted this idea to strategic argumentation. That work formulated a notion of resistance to corruption, where corruption may be collusion of the two nominally opposed agents, or espionage by one agent in gaining illicit knowledge of her opponent's arsenal of arguments.

[14,15] addressed the problem in terms of Dung's abstract argumentation [6]. Abstract argumentation abstracts away from the structure of arguments and the conflicts between them; the resulting argumentation framework can be interpreted by many different semantics. Each semantics expresses compatibility criteria that a set of arguments must satisfy to be jointly accepted. The different semantics reflect different intuitions and principles of how conflicting arguments should be resolved.


B.H. Kang and Q. Bai (Eds.): AI 2016, LNAI 9992, pp. 17–28, 2016.
DOI: 10.1007/978-3-319-50127-7_2

While [14,15] addressed many semantics for argumentation, attention was focussed on semantics based on admissibility: that for a set of arguments to be compatible it must defend itself against any attack. In this paper we address one admissibility-based semantics (the ideal semantics) that was not addressed in previous work, and two semantics that are not based on admissibility (the naive and stage semantics). It has been argued ([1,17], among others) that admissibility-based semantics do not accord with intuitions, for some argumentation frameworks.

The ideal semantics [7] offers a form of scepticism that is not as severe as that provided by the grounded semantics. Starting from the preferred semantics, it provides a single maximal coherent set of arguments that are accepted in every preferred extension, and defends itself against attack from any other arguments. It is claimed to express "justifiably accepted skeptical belief" [9].

The other semantics define a collection of sets of arguments, rather than a single set, and are not based on self-defence. The naive semantics [4] is quite natural: it consists of maximal conflict-free sets of arguments, where conflict-free entails that no argument in the set is incompatible with another. It reflects a desire to accept as many arguments as possible that are consistent with each other.

The stage semantics [17] is, like the naive semantics, based on conflict-free sets. It consists of those conflict-free sets that maximize the arguments that are decided – accepted or rejected. It reflects the intuition that as many arguments as possible should have their status decided. It has a close relation to some admissible semantics [17].

This paper is structured as follows. The next section provides necessary background on abstract argumentation and computational complexity. The following sections introduce strategic argumentation, and the computation problems arising in playing strategic argumentation games, including problems that arise when exploiting corruption. Then these problem are examined for, respectively, the ideal, naive, and stage semantics. Their computational complexity is established, which provides the basis for identifying resistance to corruption.

## 2 Background

### 2.1 Abstract Argumentation

This work is based on abstract argumentation in the sense of [6], which addresses the evaluation of a static set of arguments. An *argumentation framework* $\mathcal{A} = (S, \gg)$ consists of a finite set of arguments $S$ and a binary relation $\gg$ over $S$, called the attack (sometimes, defeat) relation. If $(a, b) \in \gg$ we write $a \gg b$ and say that $a$ attacks $b$. The semantics of an argumentation framework is given in terms of *extensions*, which are subsets of $S$.

Given an argumentation framework, an argument $a$ is said to be *accepted* in an extension $E$ if $a \in E$, and said to be *rejected* in $E$ if some $b \in E$ attacks $a$. The set of rejected arguments in $E$ is denoted by $E^-$. An argument that is neither accepted nor rejected in $E$ is said to be *undecided* in $E$. We say an argument $a$ is *self-defeating* if it attacks itself (that is, $a \gg a$). We say there is a *conflict* between arguments $a_1$ and $a_2$ if either $a_1 \gg a_2$ or $a_2 \gg a_1$. An extension $E$ is *conflict-free* if the restriction of $\gg$ to $E$ is empty. The *naive* extensions are the maximal conflict-free extensions. The *stage* extensions are the conflict-free extensions that are maximal under the containment ordering of $E \cup E^-$. Alternatively, the stage extensions are the conflict-free extensions that minimize the set of undecided arguments.

An argument $a$ is *defended* by $E$ if every argument that attacks $a$ is attacked by some argument in $E$. An extension $E$ of $\mathcal{A}$ is *admissible* if it is conflict-free and, for every argument $a \in E$ is defended by $E$. An extension $E$ of $\mathcal{A}$ is *complete* if it is conflict-free and, $a \in E$ iff $a$ is defended by $E$. The least complete extension under the containment ordering exists and is called the *grounded* extension. It reflects a strongly sceptical attitude towards accepting arguments. The *preferred* extensions are the maximal admissible extensions under the containment ordering. The *ideal* extension is the maximal admissible extension contained in all preferred extensions. A semantics is *unitary* if every argumentation framework has a single extension under the semantics. The grounded and ideal semantics are unitary; the naive and stage semantics are not.

A semantics is defined to be a set of extensions: the ideal semantics consists only of the ideal extension, the naive semantics is the set of naive extensions, the stage semantics is the set of stage extensions, etc. Each semantics expresses a criterion for what arguments can coherently be accepted together, given an argumentation framework. Each extension in the semantics represents a "reasonable" adjudication, according to that criterion, of the arguments in the argumentation framework.

## 2.2 Computational Complexity

We can view a complexity class as a set of decision problems. We assume the reader has knowledge of the polynomial complexity hierarchy (see, for example, [12]). We use PTIME to refer to the class of problems solvable in polynomial time. PSPACE is the class of decision problems solvable in polynomial space. It contains the entire polynomial hierarchy $PH$. As usual, the notation $\mathcal{C}^{\mathcal{D}}$, where $\mathcal{C}$ and $\mathcal{D}$ are complexity classes, refers to the class of problems that can be decided by an algorithm of complexity $\mathcal{C}$ with calls to a $\mathcal{D}$ oracle.

There are some additional complexity classes within the hierarchy that we need. $D_2^p$ is the class of problems that can be expressed as the conjunction of a problem in $\Sigma_2^p$ and a problem in $\Pi_2^p$. $\Theta_2^p$ is the class of decision problems solvable by a deterministic polynomial algorithm with $\mathrm{O}(log\, n)$ calls to an NP oracle. It is equal to $\mathrm{PTIME}_{||}^{\mathrm{NP}}$, the class of problems solvable by a deterministic polynomial algorithm with non-adaptive calls to an NP oracle. Non-adaptive refers to the restriction that oracle calls cannot depend on the outcome of previous calls. We have

$$NP, coNP \subseteq D^p \subseteq \Theta_2^p \subseteq \Delta_2^p \subseteq \Sigma_2^p, \Pi_2^p \subseteq D_2^p \subseteq \Sigma_3^p, \Pi_3^p$$

with $\mathrm{NP}^{\Theta_2^p} = \Sigma_2^p$ and $\mathrm{NP}^{D_2^p} = \Sigma_3^p$. Also, $\mathrm{PTIME}^{\Sigma_i^p} = \Delta_{i+1}^p$.

**Table 1.** Complexity of several argumentation reasoning problems under selected semantics

| | Credulous acceptance | Sceptical acceptance | Verification | Non-emptiness |
|---|---|---|---|---|
| Ideal | in $\Theta_2^p$ | in $\Theta_2^p$ | in $\Theta_2^p$ | in $\Theta_2^p$ |
| Naive | in PTIME | in PTIME | in PTIME | in PTIME |
| Stage | $\Sigma_2^p$-c | $\Pi_2^p$-c | coNP-c | in PTIME |

There are several prominent decision problems in argumentation. For any semantics $\sigma$:

- The *Verification problem* asks, given an argumentation framework $\mathcal{A}$ and a set of arguments $S$, is $S$ a $\sigma$-extension?
- The *Credulous Acceptance problem* asks, given $\mathcal{A}$ and an argument $a$, is there a $\sigma$-extension containing $a$?
- The *Sceptical Acceptance problem* asks, given $\mathcal{A}$ and an argument $a$, do all $\sigma$-extensions contain $a$?
- The *Non-emptiness problem* asks, is there a $\sigma$-extension of $\mathcal{A}$ that is non-empty?

Table 1 summarizes complexity results for these problems under the semantics of interest in this paper, drawn from [8–10]. For a complexity class $\mathcal{C}$, $\mathcal{C}$-c denotes $\mathcal{C}$-completeness.

## 3 Strategic Argumentation

Strategic argumentation provides a simple model of dynamic argumentation. Originally [11] it was formulated for a concrete argumentation system based in a defeasible logic, but we will use the model of [15] which is defined in terms of abstract argumentation. In strategic abstract argumentation, players take turns to add arguments to an argumentation framework. At each turn, the player adds arguments so that the argumentation framework is in a desired state. We refer to such states interchangeably as *desired outcomes* or *strategic aims* of the player. A player loses the strategic argumentation game when she is unable to achieve her desired outcome. In general, both players can win if the argumentation reaches a state that is desired by both players, but in this paper we consider an adversarial setting where the players' aims are mutually exclusive.

Strategic abstract argumentation is formalized as follows [15]. We assume there are two players, a proponent $P$ and her opponent $O$. A *split argumentation framework* $(\mathcal{A}_{Com}, \mathcal{A}_P, \mathcal{A}_O, \gg)$ consists of three sets of arguments: $\mathcal{A}_{Com}$ the arguments that are common knowledge to $P$ and $O$; $\mathcal{A}_P$ the arguments available to $P$, and $\mathcal{A}_O$ the arguments available to $O$; and an attack relation $\gg$ over $\mathcal{A}_{Com} \cup \mathcal{A}_P \cup \mathcal{A}_O$. $\mathcal{A}_P$ is assumed to be unknown to $O$, and $\mathcal{A}_O$ is unknown to $P$. Each player is aware of $\gg$ restricted to the arguments they know. We assume that $P$'s desired outcome is that a distinguished argument $a$ is accepted, in some sense, while $O$'s aim is to prevent this. Starting with $P$, the players take turns in adding sets of arguments to $\mathcal{A}_{Com}$ from their available arguments, ensuring that their desired outcome is a consequence of the resulting argumentation framework[1]. As play continues, the set of arguments that are common knowledge $\mathcal{A}_{Com}$ becomes larger. When a player is unable to achieve her aim when it is her turn to play, she loses. We say that a player is *honest* if she plays rationally in trying to win. In particular, an honest player does not abandon a game when she has a play that achieves her aim, and does not play arguments that are unnecessary to achieve her aim.

[15] identifies several plausible strategic aims that the proponent $P$ might have under an argumentation semantics $\sigma$. In this paper we focus on the following four:

1. **Existential:** $a$ is accepted in at least one $\sigma$-extension
2. **Universal:** $a$ is accepted in all $\sigma$-extensions
3. **Unrejected:** $a$ is not rejected in any $\sigma$-extension
4. **Uncontested:** $a$ is accepted in at least one $\sigma$-extension and is not rejected in any $\sigma$-extension

The existential and universal aims are credulous and sceptical acceptance. [15] also identifies some "counting aims". They will not be addressed in this paper.

In addition to these aims, a player may wish to "spoil" or prevent such aims from being achieved. Such aims are the negation of the above aims. For example, the negation of the uncontested aim aims to have $a$ not accepted in any extension or have $a$ rejected in some extension. In this paper, player $O$'s aim is to prevent $P$'s desired outcome; thus $O$'s aim is the negation of $P$'s aim.

In general, all these aims are distinct. However, for a unitary semantics $\sigma$ (such as the ideal semantics) this variety of aims collapses: all the above aims – except the unrejected aim – collapse into one, that $a$ is accepted in the $\sigma$-extension. For a unitary semantics $\sigma$ there are six possible aims: (1) $a$ is accepted in the $\sigma$-extension; (2) $a$ is rejected; (3) $a$ is undecided; (4) $a$ is not accepted; (5) $a$ is not rejected; and (6) $a$ is not undecided. Each of these aims can be expressed as a disjunction of (some of) the three properties: $a$ is accepted; $a$ is rejected; and $a$ is undecided. There are, in theory, two other aims. One is the empty disjunction, which represents an aim that can never be satisfied in a unitary semantics[2]. The other is the disjunction of all three properties, but this is always satisfied in any non-empty semantics. Of the six possible aims, the second three are the negations of the first three.

[1] Each player's move is a normal expansion [3].

[2] This possibility is not so outré in general: the stable semantics can be empty, and this is a sensible aim when a player wants to sabotage the game (that is, prevent any conclusion about the status of $a$).

## 4 Corruption in Strategic Argumentation

[13–15] presents a model of corruption within strategic argumentation. Two corrupt behaviours are defined: espionage, where one player (say $P$) violates the privacy of $\mathcal{A}_O$, and collusion, where $P$ and $O$ arrange for one of them to win, in violation of the best interests of the other's client.

Resistance to corruption under this model, adapting the notion of resistance of voting systems to manipulation [2], arises when the computational problems that arise when exploiting corruption and hiding the corruption from view require greater computational resources than playing the game honestly. This greater computational cost can act as a disincentive to corruption, since a player might be unable to exploit the results of corruption. The computational problems are formulated as decision problems, rather than functional problems, to avoid less familiar complexity classes.

The problem of verifying that an aim is satisfied by some state of strategic argumentation is a fundamental part of each move in a game, and of the exploitation of corrupt behaviour. However, its main interest is as a component of other problems.

**The Aim Verification Problem**

**Instance** An argumentation framework $(\mathcal{A}_{Com}, \gg)$, an argumentation semantics, and an aim.
**Question** Is the aim satisfied under the given semantics by the given argumentation framework?

The Desired Outcome problem [15] is the problem that a player must solve at each step of a strategic abstract argumentation game. It involves identifying that the player has a legal move.

**The Desired Outcome Problem for $P$**

**Instance** A split argumentation framework $(\mathcal{A}_{Com}, \mathcal{A}_P, \mathcal{A}_O, \gg)$ and a desired outcome for $P$.
**Question** Is there a set $I \subseteq \mathcal{A}_P$ such that $P$'s desired outcome is achieved in the argumentation framework $(\mathcal{A}_{Com} \cup I, \gg)$?

It is not difficult to see that this problem can be solved by a non-deterministic algorithm with an oracle for the Aim Verification problem.

Playing strategic argumentation involves solving the desired outcome problem at each turn. We can formulate this as a deterministic polynomially bounded algorithm with an oracle for the player's desired outcome problem. Consequently, we can identify the complexity of playing strategic argumentation as $\mathrm{PTIME}^{DO}$, where $DO$ is the complexity of the desired outcome problem.

We now turn to corruption, and the computational problems that must be solved to exploit corruption.

In the case of collusion between $P$ and $O$ to ensure that (say) $P$ wins, the players must arrange a sequence of moves that satisfy the rules of the game and leads to $P$ winning. This sequence must give the appearance of being normal play. In particular, $O$ cannot simply "give up" and fail to make a move – such behaviour would open her to charges of incompetence or corruption. Instead, she must exhaust her possible moves.

**The Winning Sequence Problem for** $P$

**Instance** A split argumentation framework $(\mathcal{A}_{Com}, \mathcal{A}_P, \mathcal{A}_O, \gg)$ and a desired outcome for $P$.
**Question** Is there a sequence of moves such that $P$ wins?

This problem can be solved by a non-deterministic algorithm that guesses moves for $P$ and $O$ and uses oracles for the aim verification problem for $P$ and $O$ and the (complement of) the desired outcome problem for $O$.

In the case of espionage, one player, say $P$, illicitly learns her opponent's arguments $\mathcal{A}_O$ and desires a strategy that will ensure $P$ wins, no matter what moves $O$ makes. A *strategy* for $P$ in a split argumentation framework $(\mathcal{A}_{Com}, \mathcal{A}_P, \mathcal{A}_O, \gg)$ is a function $s_P$ from a set of common arguments and a set of playable arguments to the set of arguments to be played in the next move. A sequence of moves $S_1, T_1, S_2, T_2, \ldots$ resulting in common arguments $\mathcal{A}_{Com}^{P,1}, \mathcal{A}_{Com}^{O,1}, \mathcal{A}_{Com}^{P,2}, \mathcal{A}_{Com}^{O,2}, \ldots$ is *consistent with* a strategy $s$ for $P$ if, for every $j$, $S_{j+1} = s_P(\mathcal{A}_{Com}^{O,j}, \mathcal{A}_P)$. A strategy for $P$ is *winning* if every valid sequence of moves consistent with the strategy is won by $P$.

**The Winning Strategy Problem for** $P$

**Instance** A split argumentation framework $(\mathcal{A}_{Com}, \mathcal{A}_P, \mathcal{A}_O, \gg)$ and a desired outcome for $P$.
**Question** Is there a winning strategy for $P$?

Strategic argumentation is said to be *resistant to collusion (espionage)* if the complexity of the Winning Sequence (Winning Strategy) problem is greater than the complexity of playing the strategic argumentation game, under the widely-believed complexity-theoretic assumption that the polynomial hierarchy does not collapse. In that case, the computational work needed to exploit the corrupt behaviour is greater than that required to simply play the argumentation game.

In the next sections we investigate the complexity of the problems defined above for the three semantics of abstract argumentation under investigation.

## 5 Strategic Argumentation Under the Ideal Semantics

Building on the work of [8], and the previous analysis of the aims under a unitary semantics, we have an upper bound on the complexity of aim verification under the ideal semantics.

**Theorem 1.** *The Aim Verification problem for* $P$ *with any of the six strategic aims under the ideal semantics is in* $\Theta_2^p$. *The corresponding aim verification problem for* $O$ *is also in* $\Theta_2^p$.

On the other hand, for the remaining problems we can give a tight analysis of their complexity. The ideal semantics is amenable to the techniques developed in [14] for other admissibility-based semantics. By combining constructions in [8,14] we obtain the following results.

**Theorem 2.** *The Desired Outcome problem for $P$ with any of the aims under the ideal semantics is $\Sigma_2^p$-complete. The same complexity holds for $O$.*

**Theorem 3.** *The Winning Sequence problem for $P$ with any of the aims under the ideal semantics is $\Sigma_3^p$-complete.*

**Theorem 4.** *The Winning Strategy problem for $P$ with any of the aims under the ideal semantics is PSPACE-complete.*

The complexity of honestly playing strategic argumentation is $\text{PTIME}^{\Sigma_2^p} = \Delta_3^p$, using Theorem 2. Consequently, we see that strategic argumentation under the ideal semantics is resistant to both collusion and espionage.

## 6 Argumentation Under the Naive Semantics

We can characterize the aims under the naive semantics.

**Lemma 1.** *Consider an argument $a$ in an argumentation framework and the naive semantics.*

1. *$a$ is in at least one naive extension iff $a$ is not self-defeating*
2. *$a$ is in every naive extension iff the only arguments that attack or are attacked by $a$ are self-defeating*
3. *$a$ is unrejected in every naive extension iff the only arguments that attack $a$ are self-defeating*
4. *$a$ is uncontested iff $a$ is not self-defeating, and the only arguments that attack $a$ are self-defeating*

It follows from this lemma that the aim verification problem under the naive semantics can be solved in polynomial time for each of the aims.

However, using the above characterization we find a more surprising result: under the naive semantics, if the players are honest, the outcome of the strategic argumentation game is determined by the initial split argumentation framework $\mathcal{A}$. There is no strategy involved.

For example, for the unrejected aim, if $\mathcal{A}_O$ contains an argument $b$ that attacks $a$ and is not self-defeating then $O$ simply has to play $b$ in order to win. Furthermore, additional arguments do not affect the existence of $b$, so it is sufficient for $O$ to play her entire set of arguments $\mathcal{A}_O$ at her first move. She does not even need to know what the focal argument is!

**Theorem 5.** *Consider strategic argumentation under the naive semantics where $P$ and $O$ are honest. Suppose that $P$ can make an initial move that includes the focal argument $a$.*

1. *If $P$ has the existential aim, then $P$ wins.*
2. *If $P$ has the universal aim, then $P$ wins iff all the arguments in $\mathcal{A}_O$ that conflict with $a$ are self-defeating*
3. *If $P$ has the unrejected or uncontested aim, then $P$ wins iff all the arguments in $\mathcal{A}_O$ that attack $a$ are self-defeating*

*In each case, if $P$ does not win then $O$ wins, and if $P$ cannot make an initial play then $O$ wins.*

The characterization for the uncontested aim is the same as the characterization for the unrejected aim because the assumption that $P$ can make an initial move ensures that there is a naive extension where $a$ is accepted.

Consequently, for each of these aims, the outcome is totally predictable. Any deviation from that result is a sign that one of the players is corrupt. Thus strategic argumentation under naive semantics is more than resistant to corruption: any collusive behaviour cannot be hidden, and the results of espionage cannot be used to affect the outcome. We say that strategic argumentation under the naive semantics is *proof against corruption* or *corruption-proof.*

This result contrasts markedly with a result of [16]. That work characterized strategy-proof games in similar argumentation games under the grounded semantics. That characterization suggests that only rarely is a game under the grounded semantics not strategic.

## 7 Strategic Argumentation Under the Stage Semantics

Unlike the ideal semantics, but like many other semantics [14], the complexity of the problems under the stage semantics varies with the players' strategic aim.

**Proposition 1.** *Consider the Aim Verification problem for $P$ under the stage semantics.*

1. *The complexity for the existential aim is $\Sigma_2^p$-complete*
2. *The complexity for the universal aim is $\Pi_2^p$-complete*
3. *The complexity for the unrejected aim is $\Pi_2^p$-complete*
4. *The complexity for the uncontested aim is $D_2^p$-complete*

The complexity of Aim Verification for $O$, assuming $O$'s aim is to prevent $P$ from achieving her aim, is the complement of the complexity of Aim Verification for $P$.

As mentioned earlier, the Desired Outcome problem is in $\mathrm{NP}^{AV}$, where $AV$ is the Aim Verification problem. Building on constructions of [10,15], we can establish hardness results.

**Theorem 6.** *Consider the Desired Outcome problem for $P$ under the stage semantics.*

1. *The problem with the existential aim is $\Sigma_2^p$-complete*
2. *The problem with the universal aim is $\Sigma_3^p$-complete*
3. *The problem with the unrejected aim is $\Sigma_3^p$-complete*
4. *The problem with the uncontested aim is $\Sigma_3^p$-complete*

Hence the complexity of $P$ honestly playing strategic argumentation is $\Delta_3^p$ for the existential aim, and $\Delta_4^p$ for the other aims.

**Theorem 7.** *Consider the Winning Sequence problem for $P$ under the stage semantics.*

1. *The problem with the existential aim is $\Sigma_4^p$-complete*
2. *The problem with the universal aim is $\Sigma_3^p$-complete*
3. *The problem with the unrejected aim is $\Sigma_3^p$-complete*
4. *The problem with the uncontested aim is $\Sigma_4^p$-complete*

Like all the admissibility-based semantics, and unlike the naive semantics, under the stage semantics the Winning Strategy problem is PSPACE-complete.

**Theorem 8.** *The Winning Strategy problem under the stage semantics is PSPACE-complete for each of the aims addressed in this paper.*

Thus we see that there is resistance to espionage under the stage semantics, but resistant to collusion only for the existential and uncontested aims.

**Table 2.** Resistance to collusion to ensure $P$ wins, for several aims and semantics.

| | Grounded | Preferred | Ideal | Naive | Stage |
|---|---|---|---|---|---|
| Existential | Resistant | Resistant | Resistant | Proof | Resistant |
| Universal | Resistant | | Resistant | Proof | |
| Unrejected | Resistant | | Resistant | Proof | |
| Uncontested | Resistant | Resistant | Resistant | Proof | Resistant |

## 8 Conclusion

We have investigated the resistance to corruption of strategic argumentation under three semantics for argumentation. The naive semantics is proof against corruption, but only because there is no significant strategy involved. The ideal semantics is resistant to corruption, for each of the aims we studied. The resistance to collusion of the stage semantics varies according to the aim: It is resistant to collusion for the existential and uncontested aims, but not for the other aims. The stage semantics is also resistant to espionage for each of the aims. A summary of the results of this paper on resistance to collusion appears in Table 2.

There remain several argumentation semantics, including other extension-based semantics and ranking-based semantics [5], for which resistance to corruption has not been determined.

## References


1. Baroni, P., Giacomin, M., Guida, G.: SCC-recursiveness: a general schema for argumentation semantics. Artif. Intell. **168**(1–2), 162–210 (2005)
2. Bartholdi, J.J., Tovey, C.A., Trick, M.A.: The computational difficulty of manipulating an election. Soc. Choice Welf. **6**(3), 227–241 (1989)
3. Baumann, R., Brewka, G.: Expanding argumentation frameworks: enforcing and monotonicity results. In: COMMA, pp. 75–86 (2010)
4. Bondarenko, A., Dung, P.M., Kowalski, R.A., Toni, F.: An abstract, argumentation-theoretic approach to default reasoning. Artif. Intell. **93**, 63–101 (1997)
5. Bonzon, E., Delobelle, J., Konieczny, S., Maudet, N.: A comparative study of ranking-based semantics for abstract argumentation. In: AAAI Conference on Artificial Intelligence, pp. 914–920 (2016)
6. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artif. Intell. **77**(2), 321–358 (1995)
7. Dung, P.M., Mancarella, P., Toni, F.: Computing ideal sceptical argumentation. Artif. Intell. **171**(10–15), 642–674 (2007). http://dx.doi.org/10.1016/j.artint.2007.05.003
8. Dunne, P.E.: The computational complexity of ideal semantics. Artif. Intell. **173**(18), 1559–1591 (2009). http://dx.doi.org/10.1016/j.artint.2009.09.001
9. Dunne, P.E., Dvořák, W., Woltran, S.: Parametric properties of ideal semantics. Artif. Intell. **202**, 1–28 (2013). http://dx.doi.org/10.1016/j.artint.2013.06.004
10. Dvořák, W., Woltran, S.: Complexity of semi-stable and stage semantics in argumentation frameworks. Inf. Process. Lett. **110**(11), 425–430 (2010). http://dx.doi.org/10.1016/j.ipl.2010.04.005
11. Governatori, G., Olivieri, F., Scannapieco, S., Rotolo, A., Cristani, M.: Strategic argumentation is NP-complete. In: Proceedings of European Conference on Artificial Intelligence, pp. 399–404 (2014)
12. Johnson, D.S.: A catalog of complexity classes. In: van Leeuwen, J. (ed.) Handbook of Theoretical Computer Science. Algorithms and Complexity, vol. A, pp. 67–161. Elsevier, Amsterdam (1990)
13. Maher, M.J.: Complexity of exploiting privacy violations in strategic argumentation. In: Proc. Pacific Rim International Conf. on Artificial Intelligence. pp. 523–535 (2014)
14. Maher, M.J.: Resistance to corruption of general strategic argumentation. In: Proceedings of International Conference Principles and Practice of Multi-Agent Systems, pp. 61–75 (2016)
15. Maher, M.J.: Resistance to corruption of strategic argumentation. In: AAAI Conference on Artificial Intelligence, pp. 1030–1036 (2016)

16. Rahwan, I., Larson, K., Tohmé, F.A.: A characterisation of strategy-proofness for grounded argumentation semantics. In: Boutilier, C. (ed.) IJCAI, pp. 251–256 (2009)
17. Verheij, B.: Two approaches to dialectical argumentation: admissible sets and argumentation stages. In: Proceedings of the 8th Dutch Conference on Artificial Intelligence, pp. 357–368 (1996)

http://www.springer.com/978-3-319-50126-0

AI 2016: Advances in Artificial Intelligence
29th Australasian Joint Conference, Hobart, TAS, Australia, December 5-8, 2016, Proceedings
Kang, B.H.; Bai, Q. (Eds.)
2016, XVI, 729 p. 433 illus., 120 illus. in color., Softcover
ISBN: 978-3-319-50126-0