Inferring attack relations for gradual semantics

Nir Oren [*] and Bruno Yun
Department of Computing Science, University of Aberdeen, Scotland
E-mails: n.oren@abdn.ac.uk, bruno.yun@abdn.ac.uk

Abstract. A gradual semantics takes a weighted argumentation framework as input and outputs a final acceptability degree for each argument, with different semantics performing the computation in different manners. In this work, we consider the problem of attack inference. That is, given a gradual semantics, a set of arguments with associated initial weights, and the final desirable acceptability degrees associated with each argument, we seek to determine whether there is a set of attacks on those arguments such that we can obtain these acceptability degrees. The main contribution of our work is to demonstrate that the associated decision problem, i.e., whether a set of attacks can exist which allows the final acceptability degrees to occur for given initial weights, is NP-complete for the weighted h-categoriser and card-based semantics, and is polynomial for the weighted max-based semantics, even for the complete version of the problem (where all initial weights and final acceptability degrees are known). We then briefly discuss how this decision problem can be modified to find the attacks themselves and conclude by examining the partial problem where not all initial weights or final acceptability degrees may be known.

Keywords: Gradual semantics, argumentation, complexity

1. Introduction

Abstract argumentation semantics associate a justification status with a set of arguments based on interactions between arguments. Such interactions can include inter-argument attacks [17], or preference-based defeats [5,6,21,29,41], or may consider both of these together with some supportive relationship [2,3,12,30]. Given such a framework, an argument may be considered justified if it appears in one (resp. many or all) *extensions*, where such extensions – sets of non-conflicting arguments accepted together – are computed according to the argumentation semantics, or based on the label assigned to the argument (we refer the reader to [7] for an introduction to the most common argumentation semantics).

Arguments often have associated weights, and different semantics have been proposed which consider the weight associated with an argument [2,14,30,34]. While some semantics developed in this context return whether an argument is, or is not justified [14], many other *ranking-based* semantics either return a preference ordering over arguments. Such ranking-based semantics have been used to, for example, identify irrationality in reasoning [20] by examining whether the initial weights associated with an argument affect the argument's final acceptability degree in an appropriate manner (i.e., consistent with the ranking-based semantics being used); refining the results obtained by extension semantics [11,42]; and applied to multi-agent settings [16]. In this paper, we focus on *gradual* semantics, a subclass of ranking-based semantics which associates a numeric *final acceptability degree* with each argument to compute a preference ordering over arguments.

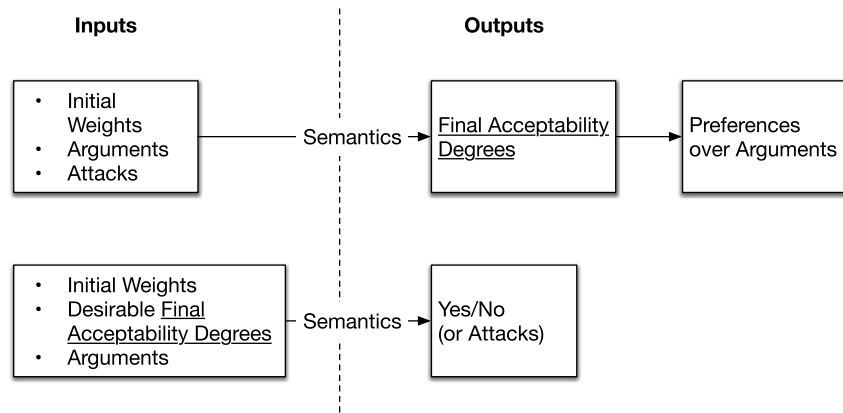[*]Corresponding author. E-mail: n.oren@abdn.ac.uk.

Fig. 1. The process (top) by which a gradual semantics is applied to compute an acceptability preference between arguments and (bottom) the inverse problem considered in this paper. The "final acceptability degrees" are now used as input. That is, given arguments, initial weights, and desirable final acceptability degrees, can we find a suitable set of attacks?

The top portion of Fig. 1 shows the reasoning process used when reasoning using gradual semantics. Here, a weighted argumentation framework consisting of arguments, attacks and initial weights associated with arguments is provided as input. A gradual semantics is then used to compute a *final acceptability degree* for each argument. In many semantics, these final acceptability degrees are then used to compute a preference ordering over the arguments.

More recent work has considered different combinations of inputs and outputs to the problem. For example, [34] seeks to identify a set of initial weights for arguments based on the final argument preference ordering and argumentation semantics, while [27] determines the preferences between arguments given argument justification status, semantics and argumentation framework. In this paper, and as illustrated in the bottom portion of Fig. 1, we consider the problems of determining whether a set of attacks between arguments can be identified given specific argumentation semantics, the final acceptability degrees, and the initial weights. As discussed further in Section 5, we leave the problem of using preferences over arguments as input for our problem as future work. While it is true that the problem we consider is somewhat abstract and has limited real-world applications, it serves as a departure point for potentially important applications of argumentation to opponent modelling [35,37] and preference elicitation [27].

While other ranking-based semantics are described in the literature, we consider three popular weighted gradual semantics (the weighted h-categoriser, the weighted card-based, and the weighted max-based semantics [4]). These three semantics were chosen due to their similarity to one another, and due to the way in which a similar approach can be used to solve the problem we consider when these semantics are used. We show that when given all final acceptability degrees and initial weights, the problems for the weighted h-categoriser and the weighted card-based semantics are both NP-complete, while the problem can be solved in polynomial time for the weighted max-based semantics.

The remainder of this paper is structured as follows. In Section 2, we provide the necessary background to understand the remainder of the paper. Section 3 formalises the decision problem of whether attacks can be found for gradual semantics and then demonstrates that it is NP-complete for some semantics. In Section 4, we discuss solvers for our problem, and consider related and future work (including a partial version of the problem) in Section 5.

2. Background

As discussed above, we situate our work in the context of *weighted* abstract argumentation frameworks (WAFs), which can be encoded as graphs with weighted nodes. Each argument has an initial weight (also called "basic score" in [36]) drawn from the range [0, 1]. The smaller the initial weight of an argument, the weaker the argument. The initial weight of an argument may represent different concepts such as the likelihood degree of its premises [9], the degree of trust in its source [15], or an aggregation of votes provided by users [22] among others. In this paper, the origin of the weights and arguments is left unspecified. Similarly, arguments and attacks are considered abstract notions.

Definition 1. A (abstract) *weighted argumentation framework* is a triple $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$ where $\mathcal{A}$ is a *finite* set of arguments, $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ is a binary attack relation, and $w : \mathcal{A} \to [0, 1]$ is a total weighting function which associates an *initial weight* between 0 and 1 to each argument.

Given an argument $a \in \mathcal{A}$, we refer to $w(a)$ as a's *initial weight*. For any argument $a \in \mathcal{A}$, the set $\{b \in \mathcal{A} | (b, a) \in \mathcal{D}\}$, denoted by $\text{Att}(a)$, contains all attackers of a. Similarly, we define $\text{Att}^*(a) = \{b \in \mathcal{A} | (b, a) \in \mathcal{D} \text{ and } w(b) > 0\}$, i.e., the set of attackers of a with strictly positive weights. When the current argumentation framework $\mathcal{F}$ is not clear from the context, we will use the notation $\text{Att}_{\mathcal{F}}$ and $\text{Att}^*_{\mathcal{F}}$ respectively.

A gradual semantics σ takes as input a weighted argumentation framework and outputs a function that associates a *final acceptability degree* for each argument.[1] Several such semantics have been described in the literature. In this paper, we focus on three widely used gradual semantics introduced by Amgoud et al., i.e. the weighted h-categoriser, the weighted max-based, and the weighted card-based semantics, denoted σ_{HC}, σ_{MB} and σ_{CB} respectively [4].

The weighted h-categoriser considers the initial weight of an argument as well as the sum of the acceptability degrees of all its attackers to determine the argument's final acceptability degree. In this semantics, all (non-zero) attackers will be considered when determining the acceptability degree of an argument and the number of attackers is not used.

Definition 2. The weighted h-categoriser semantics, denoted σ_{HC}, is the function that takes as input any weighted argumentation framework $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$ and returns the function $\sigma_{HC}^{\mathcal{F}} : \mathcal{A} \to [0, 1]$ such that for all $a \in \mathcal{A}$, $\sigma_{HC}^{\mathcal{F}}(a) = HC_{\infty}(a)$, where:

$$HC_i(a) = \begin{cases} w(a) & \text{if } i = 0 \\ \frac{w(a)}{1 + \sum_{b \in \text{Att}(a)} HC_{i-1}(b)} & \text{otherwise} \end{cases}$$

Next, we consider the weighted max-based semantics which utilises the initial weight of an argument as well as the highest acceptability degrees of its attackers to determine the argument's final acceptability degree. In this semantics, only the strongest attacker is considered and the number of attackers is not used.

Definition 3. The weighted max-based semantics, denoted σ_{MB}, is the function that takes as input any weighted argumentation framework $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$ and returns the function $\sigma_{MB}^{\mathcal{F}} : \mathcal{A} \to [0, 1]$ such that

---

[1]Often, the final step in using a gradual semantics involves using this final acceptability degree to compute a preference ordering over arguments.

for all $a \in \mathcal{A}$, $\sigma_{MB}^{\mathcal{F}}(a) = MB_\infty(a)$, where:

$$MB_i(a) = \begin{cases} w(a) & \text{if } i = 0 \\ \frac{w(a)}{1 + \max_{b \in \mathtt{Att}(a)} MB_{i-1}(b)} & \text{otherwise} \end{cases}$$

Finally, the weighted card-based semantics considers the initial weight of an argument, the number of its attackers as well as the sum of the acceptability degrees of its attackers to determine the acceptability degree of the considered argument. In this semantics, the number of attackers is the most important factor to determine the acceptability degree of an argument.

Definition 4. The weighted card-based semantics, denoted σ_{CB}, is the function that takes as input any weighted argumentation framework $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$ and returns the function $\sigma_{CB}^{\mathcal{F}} : \mathcal{A} \to [0, 1]$ such that for all $a \in \mathcal{A}$, $\sigma_{CB}^{\mathcal{F}}(a) = CB_\infty(a)$, where:

$$CB_i(a) = \begin{cases} w(a) & \text{if } i = 0 \text{ or } |\mathtt{Att}^*(a)| = 0 \\ \frac{w(a)}{1 + |\mathtt{Att}^*(a)| + \frac{\sum_{b \in \mathtt{Att}^*(a)} CB_{i-1}(b)}{|\mathtt{Att}^*(a)|}} & \text{otherwise} \end{cases}$$

Here, $|\mathtt{Att}^*(a)|$ denotes the number of arguments attacking a with non-zero final acceptability degree.

We note in passing that the properties for these semantics have been researched in depth by Amgoud et al. [4]. Perhaps the most important property, in the context of this paper, is the convergence of the σ_{HC}, σ_{MB} and σ_{CB} semantics for finite frameworks (see Theorem 7, 12, 17 in [4]). A proof of convergence for a broader class of semantics (including these three) was proved in the work of Oren et al. [33]. This means that for any given argumentation framework, semantics and initial weights, the final acceptability degrees of all arguments can always be computed.

Example 1. An agent may believe the following arguments.

- a_0: Tomatoes older than a week can go rotten; these tomatoes are a week and a half old.
- a_1: Tomatoes kept in the fridge (like these ones) can last longer than a week, and so the tomatoes are not rotten.
- a_2: My friend ate one of the tomatoes this morning, and it tasted fine, therefore they are not rotten.
- a_3: My friend is not very good at discriminating whether something is rotten by taste, so the tomatoes might be rotten.

Here, a_0 is attacked by a_1 and a_2, while the latter is attacked by a_3. The agent ascribes each argument with an initial weight encoding the agent's belief in the applicability or strength of the argument, i.e., $w(a_0) = 0.9$, $w(a_1) = 0.7$, $w(a_2) = 0.7$, and $w(a_3) = 0.6$. Fig. 2 illustrates this weighted argumentation framework $\mathcal{F} = \langle \{a_0, a_1, a_2, a_3\}, \{(a_1, a_0), (a_2, a_0), (a_3, a_2)\}, w \rangle$.

We consider three different semantics in this paper (defined above); yielding the final acceptability degrees in Table 1. In turn, a reasoner using the weighted h-categoriser semantics σ_{HC} or σ_{CB} semantics would have preferences $a_1 \succ a_3 \succ a_2 \succ a_0$, while if it were to use σ_{MB} preferences would be $a_1 \succ a_3 \succ a_0 \succ a_2$. Note that while the σ_{HC} and σ_{CB} semantics yield similar preference orderings, the agent would be less sure of its conclusions in the latter case (due to the smaller difference in weights).
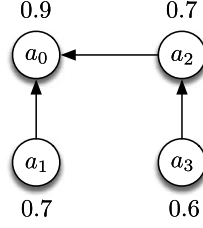
Fig. 2. The weighted argumentation framework from Example 1.

Table 1

Final acceptability degrees for each argument of Example 1 (Fig. 2) for σ_{HC}, σ_{MB}, and σ_{CB}

Argument	$\sigma_{HC}^{\mathcal{F}}(a_i)$	$\sigma_{MB}^{\mathcal{F}}(a_i)$	$\sigma_{CB}^{\mathcal{F}}(a_i)$
a_0	0.421	0.529	0.258
a_1	0.7	0.7	0.7
a_2	0.438	0.438	0.269
a_3	0.6	0.6	0.6

3. Inferring attacks

In this paper, we focus on the problem of identifying whether we can infer suitable attacks between arguments given their initial weights and desirable final acceptability degrees (as shown at the bottom of Fig. 1). However, we ground most of our discussion in the following, closely-related decision problem: *given a gradual argumentation semantics, a set of arguments, and associated information about these arguments (i.e., initial weights and some or all desirable final acceptability degrees), is there an attack relation over the arguments that, according to the semantics, lead to the desirable final acceptability degrees from the initial weights?*

In this section, we consider the *complete problem*, where the initial weights and desirable final acceptability degrees for *all* arguments are known. We formalise this problem as follows.

Problem 1. The decision problem $\mathcal{DEC}_c^x$, where $x \in \{MB, HC, CB\}$ is: *Given a set of arguments $\mathcal{A}$, a gradual semantics σ_x, a weighting function w, and a total desired final acceptability degree function $S : \mathcal{A} \to [0, 1]$, is there a set of attacks $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ such that for all $a \in \mathcal{A}$, $\sigma_x^{\mathcal{F}}(a) = S(a)$, where $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$?*

We investigate this decision problem for the three gradual semantics σ_x, for $x \in \{HC, CB, MB\}$.

Proposition 1. $\mathcal{DEC}_c^{MB}$ *is polynomial and can be decided in $O(n)$ time and space.*

Proof. We create the set $L = \{1 + S(b) | b \in \mathcal{A}\}$, for which an $O(1)$ lookup can be performed. We can then use Algorithm 1 to solve $\mathcal{DEC}_c^{MB}$:

In the Algorithm 1, line 2 checks whether arguments with zero (resp. non-zero) initial weights and non-zero (resp. zero) desired final acceptability degrees exist; the presence of such arguments would mean that the weighted-max-based semantics cannot be satisfied (line 3). Line 5 then checks the existence of an attacking argument with a suitable acceptability degree (via a lookup in the set L) to satisfy the weighted-max-based semantics in $O(1)$ time. Since lines 1–8 iterate over all arguments, the time complexity is $O(n)$; similarly, storing L takes $O(n)$ space.

Algorithm 1 Procedure to solve the decision problem $\mathcal{DEC}_c^{MB}$

1: **for all** $a \in \mathcal{A}$ **do**
2:      **if** ($w(a) = 0$ and $S(a) \neq 0$) or ($w(a) \neq 0$ and $S(a) = 0$) **then**
3:          **return** False $\triangleright$ A zero initial weight cannot lead to a non-zero final acceptability degree (and vice-versa).
4:      **end if**
5:      **if** $S(a) \neq 0$ and $w(a)/S(a) \notin L$ **then**
6:          **return** False $\triangleright$ Contradiction by definition of the σ_{MB} semantics.
7:      **end if**
8: **end for**
9: **return** True

Proposition 2. *$\mathcal{DEC}_c^{HC}$ is NP-complete.*

Proof. We show that $\mathcal{DEC}_c^{HC}$ is NP-complete using the subset-sum problem ($\mathcal{SSP}$). Formally, $\mathcal{SSP}$ is defined as answering the following question. Consider a multiset of positive numbers M and a number $T \in \mathbb{R}^+$, is there a subset $M' \subseteq M$ such that $\sum_{m' \in M'} m' = T$?

To prove $\mathcal{DEC}_c^{HC}$ is NP-complete we must demonstrate that it is in NP and identify a polynomial time reduction from $\mathcal{SSP}$ to $\mathcal{DEC}_c^{HC}$.

To demonstrate that $\mathcal{DEC}_c^{HC}$ belongs to NP we observe that the certificate of the problem is a set of attacks. Given a set of attacks, we can check in polynomial time whether $S(a) = w(a)/(1 + \sum_{b \in \text{Att}(a)} S(b))$ for all $a \in \mathcal{A}$ (as we have all initial weights and the desired final acceptability degrees), therefore the problem is in NP.

Turning to the reduction, we begin by reducing $\mathcal{SSP}$ to $\mathcal{DEC}_c^{HC}$. Let us assume we have a multiset of positive numbers $M = \{m_1, \ldots, m_n\}$ and $T \in \mathbb{R}^+$. We denote $\sum_i m_i$ by m^*. We then create a set of $n+1$ arguments $\mathcal{A} = \{a_0, a_1, a_2, \ldots, a_n\}$ such that for all $i \in \{1, \ldots, n\}$, $S(a_i) = w(a_i) = (0.4m_i)/(nm^*)$, and we set $w(a_0) = 1$ and $S(a_0) = nm^*/(0.4T + nm^*)$. This transformation is represented in Fig. 3. $\blacksquare$

Example 2. Consider the $\mathcal{SSP}$ instance where $T = 100$ and $M = \{23, 94, 1, 37, 40\}$. Here, $n = 5$ and $m^* = 195$. Using our transformation, we obtain the arguments, initial weights, and final acceptability degrees shown in Table 2.
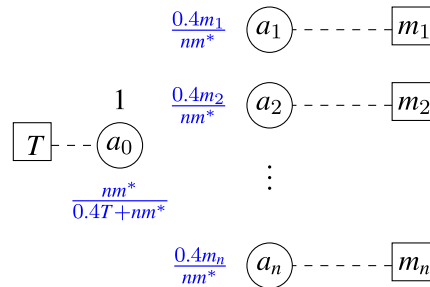

Fig. 3. Graphical representation of the polynomial transformation (dashed lines) of an $\mathcal{SSP}$ instance (rectangular nodes) into an instance of $\mathcal{DEC}_c^{HC}$ (circular nodes). The initial weight of a_0 is 1 (shown in black above a_0), and of a_1 to a_n, is equal to their final acceptability degree (where final acceptability degrees are shown in blue). The initial weights of the latter arguments have thus been omitted.

Table 2

The arguments created from Example 2 using the reduction described in Proposition 2. Note that for $a_1 \ldots a_5$, the initial weight equals the desired final acceptability degree

Argument	Initial weight	Desired final acceptability degree
a_0	1	0.96059
a_1		0.00944
a_2		0.03856
a_3		0.00041
a_4		0.01518
a_5		0.01641

- We now demonstrate that using the above reduction, denoted g, if there exists a solution to an instance x of $\mathcal{SSP}$, then the instance $g(x)$ of $\mathcal{DEC}_c^{HC}$ also has a solution. If there exists an $M' \subseteq M$ such that $\sum_{m' \in M'} m' = T$, then $\mathcal{D} = \{(f(m), a_0) | m \in M'\}$ is a set of attacks such that for all $a \in \mathcal{A}$, $\sigma^{\mathcal{F}}(a) = S(a)$, where f associates the corresponding argument and $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$. Indeed, we have for all $i \in \{1, \ldots, n\}$, $\sigma_{HC}^{\mathcal{F}}(a_i) = S(a_i)$ since a_i is not attacked and:

$$\sigma_{HC}^{\mathcal{F}}(a_0) = \frac{1}{1 + \sum_{m \in M'} S(f(m))}$$
$$= \frac{1}{1 + \frac{0.4T}{nm^*}}$$
$$= \frac{nm^*}{nm^* + 0.4T}$$
$$= S(a_0)$$

- We now show the reduction in the other direction – if an instance $g(x)$ of $\mathcal{DEC}_c^{HC}$ has a solution then the instance x of $\mathcal{SSP}$ has a solution. If there exists $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ such that for all $b \in \mathcal{A}$, $\sigma^{\mathcal{F}}(b) = S(b)$, then $M' = \{(S(b)nm^*)/0.4 | (b, a_0) \in \mathcal{D}\} \subseteq M$ is such that $\sum_{m' \in M'} m' = T$, where $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$. Indeed:

$$\sum_{m' \in M'} m' = \left(\sum_{(b,a_0) \in \mathcal{D}} S(b) \right) \frac{nm^*}{0.4}$$
$$= \left(\sum_{(b,a_0) \in \mathcal{D}} \sigma_{HC}^{\mathcal{F}}(b) \right) \frac{nm^*}{0.4}$$
$$= \left(\frac{1}{\frac{nm^*}{0.4T + nm^*}} - 1 \right) \frac{nm^*}{0.4}$$
$$= \left(\frac{0.4T}{nm^*} \right) \frac{nm^*}{0.4}$$
$$= T$$

Note that $(a_0, a_0) \notin \mathcal{D}$. Indeed, if $(a_0, a_0) \in \mathcal{D}$ then we have that $\sigma_{HC}^{\mathcal{F}}(a_0) = 1/(1 + \sigma_{HC}^{\mathcal{F}}(a_0) + Y)$, where $Y = \sum_{b \in \text{Att}(a_0) \setminus \{a_0\}} \sigma_{HC}^{\mathcal{F}}(b)$. Hence, $(\sigma_{HC}^{\mathcal{F}}(a_0)^2 + \sigma_{HC}^{\mathcal{F}}(a_0) - 1)/(-\sigma_{HC}^{\mathcal{F}}(a_0)) = Y$ and since $Y \geqslant 0$, we conclude that $0 \leqslant \sigma_{HC}^{\mathcal{F}}(a_0) \leqslant \frac{-1+\sqrt{5}}{2} \simeq 0.618$. However, we get a contradiction as $S(a_0) = \sigma_{HC}^{\mathcal{F}}(a_0)$ and:

$$\frac{T}{nm^*} \leqslant 1$$

$$\frac{0.4T}{nm^*} \leqslant 0.4$$

$$\frac{0.4T + nm^*}{nm^*} \leqslant 1.4$$

$$S(a_0) \geqslant 0.714$$

We have proved that $\mathcal{DEC}_c^{HC}$ is in NP and that $\mathcal{SSP}$ is polynomial time reducible to $\mathcal{DEC}_c^{HC}$ (the size of the argumentation framework produced is polynomial with respect to the size of the $\mathcal{SSP}$ instance), therefore $\mathcal{DEC}_c^{HC}$ is NP-complete. $\square$

Example 3 (Cont'd Example 2). We see that the subset sum problem has a solution (using values 23,37 and 40). Analogously, a solution to $\mathcal{DEC}_c^{HC}$ exists by having arguments a_1, a_4 and a_5 attack argument a_0, as shown in Fig. 4.

Proposition 3. $\mathcal{DEC}_c^{CB}$ is NP-complete.

Proof. Similar to the previous proof, we show that $\mathcal{DEC}_c^{CB}$ is NP-complete using the k-subset-sum problem ($k\mathcal{SSP}$). Formally, $k\mathcal{SSP}$ is defined as answering the following question. Consider a multiset of positive numbers M and a number $T \in \mathbb{R}^+$, is there a subset $M' \subseteq M$ such that $\sum_{m' \in M'} m' = T$ and $|M'| = k$?
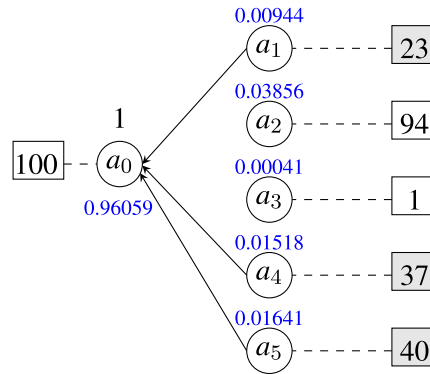

Fig. 4. Representation of how a solution to an $\mathcal{SSP}$ instance (represented with the gray rectangles) can be obtained from a solution to a corresponding instance from $\mathcal{DEC}_c^{HC}$ (the attacks drawn). Blue values for $a_1 \ldots a_5$ indicate both inital weights and final acceptability degrees, and denote final acceptability degrees for a_0, which has an initial weight of 1.

We first observe that the certificate of the problem $\mathcal{DEC}_c^{CB}$ is a set of attacks. Given a set of attacks, we can check in polynomial time whether $S(a) = w(a)/(1 + |\text{Att}^*(a)| + \sum_{b \in \text{Att}^*(a)} S(b)/|\text{Att}^*(a)|)$ for all $a \in \mathcal{A}$, meaning that the problem is in NP.

We now reduce $k\mathcal{SSP}$ to $\mathcal{DEC}_c^{CB}$. Let us assume we have a multiset of positive numbers $M = \{m_1, \ldots, m_n\}$, $1 \leqslant k \leqslant n$, and $T \in \mathbb{R}^+$. We denote by $m^* = \max_{m_i \in M} m_i$ and $u = \frac{\sqrt{k^2 + 2k + 4/k + 1} - k - 1}{3}$ (note that $u \in]0, (2\sqrt{2} - 2)/3]$ when $k \geqslant 1$). We create $n + 1$ arguments $\mathcal{A} = \{a_0, a_1, a_2, \ldots, a_n\}$ such that for all $i \in \{1, \ldots, n\}$, $S(a_i) = w(a_i) = um_i/m^*$ and we set $w(a_0) = 1$ and $S(a_0) = 1/(1 + k + Tu/(km^*))$.

- If there exists an $M' \subseteq M$ such that $\sum_{m' \in M'} m' = T$ and $|M'| = k$, then $\mathcal{D} = \{(f(m), a_0)|m \in M'\}$ is a set of k attacks such that for all $a \in \mathcal{A}$, $\sigma^{\mathcal{F}}(a) = S(a)$, where f associates the corresponding argument with m and $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$. Indeed, we have for all $i \in \{1, \ldots, n\}$, $\sigma_{CB}^{\mathcal{F}}(a_i) = S(a_i)$ since a_i is not attacked and:

$$
\begin{aligned}
\sigma_{CB}^{\mathcal{F}}(a_0) &= \frac{1}{1 + k + \frac{\sum_{m \in M'} S(f(m))}{k}} \\
&= \frac{1}{1 + k + \frac{\frac{Tu}{m^*}}{k}} \\
&= \frac{1}{1 + k + \frac{Tu}{km^*}}
\end{aligned}
$$

- Assume we have a solution to the associated $\mathcal{DEC}_c^{CB}$ instance, i.e. we have $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ such that for all $b \in \mathcal{A}$, $\sigma^{\mathcal{F}}(b) = S(b)$. We know that a_0 is attacked by k arguments in $\mathcal{D}$ as its value lies between $[1/(k + 2), 1/(k + 1)]$ and that arguments a_1 to a_n are not attacked (except if they or their attackers have initial weights of 0). We can build $M' = \{S(b) * m^*/u|(b, a_0) \in \mathcal{D}\} \subseteq M$ such that $|M'| = k$ and $\sum_{m' \in M'} m' = T$. Indeed:

$$
\begin{aligned}
\sum_{m' \in M'} m' &= \left(\sum_{(b, a_0) \in \mathcal{D}} S(b) \right) \frac{m^*}{u} \\
&= \left(\frac{Tu}{m^*} \right) \frac{m^*}{u} \\
&= T
\end{aligned}
$$

We show that $(a_0, a_0) \notin \mathcal{D}$ by contradiction. Assume we have this self-attack, then the maximum value with self attack possible for $\sigma_{CB}^{\mathcal{F}}(a_0)$ is determined by computing the unique fixed-point of the function $z(x) = 1/(1 + k + x/k)$ which is:

$$
0 < \frac{-k^2 + \sqrt{k(k^3 + 2k^2 + k + 4)} - k}{2} \leqslant 1
$$

Moreover, it holds that:

$$\frac{-k^2 + \sqrt{k(k^3 + 2k^2 + k + 4)} - k}{2} < \frac{1}{1 + k + u}$$

$$< \frac{1}{1 + k + \frac{Tu}{km^*}}$$

$$< S(a_0)$$

This is a contradiction with $S(a_0) = \sigma_{CB}^{\mathcal{F}}(a_0)$. □

To recap, our results show that $\mathcal{DEC}_c^{MB}$ is polynomial (and indeed, can be solved in linear time), while $\mathcal{DEC}_c^{HB}$ and $\mathcal{DEC}_c^{CB}$ are both NP-complete. This latter result was obtained by a reduction to a variant of the subset-sum problem.

4. Identifying solutions

Rather than simply considering the decision problem, it is useful – if possible – to be able to infer the attacks induced by some set of initial weights and final acceptability degrees. We consider each semantics individually.

4.1. The weighted max-based semantics

For the weighted max-based semantics, it is trivial to modify Algorithm 1 from Proposition 1 to infer a suitable set of attacks (when possible). Rather than simply returning True or False, we return, by modifying lines 5 and 6, any set of attacks X that satisfy the condition that for all $a \in \mathcal{A}$, there exists $(x, a) \in X$ such that $w(a)/S(a) = (1 + S(x))$. Thus, we note that this semantics does little to constrain the attacks in the framework. If we discover that there is an attack from argument a_i to argument a_0 in a solution, then under the max-based semantics, we can expand this solution by adding additional attacks from any argument a_j to a_0, where $S(a_j) \leqslant S(a_i)$, without affecting the result.

Proposition 4 (Solution Expansion). *Given a set of arguments $\mathcal{A}$, a weighting function $w : \mathcal{A} \to [0, 1]$, and a desirable final acceptability degree function $S : \mathcal{A} \to [0, 1]$. If $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ is such that for all $a \in \mathcal{A}$, $\sigma_{MB}^{\mathcal{F}}(a) = S(a)$, where $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$ then for all $a \in \mathcal{A}$, $\sigma_{MB}^{\mathcal{F}'}(a) = S(a)$, where $\mathcal{F}' = \langle \mathcal{A}, \mathcal{D}', w \rangle$, $\mathcal{D}' = \mathcal{D} \cup X$ and $X \subseteq \{(a_i, a_j) | \exists (a_k, a_j) \in \mathcal{D} \text{ and } S(a_i) \leqslant S(a_k)\} \cup \{(a_i, a_j) | w(a_j) = 0\}$. We say that $\mathcal{D}'$ is an expansion of $\mathcal{D}$.*

Similarly, if we have a set of attacks that achieve the desirable final acceptability degrees for all arguments and there is an attack from a_i to a_0, then under the max-based semantics, we can contract this solution by removing attacks from any other argument a_j to a_0, where $S(a_j) \leqslant S(a_i)$.

Proposition 5 (Solution Contraction). *Given a set of arguments $\mathcal{A}$, a weighting function $w : \mathcal{A} \to [0, 1]$, and a desirable final acceptability degree function $S : \mathcal{A} \to [0, 1]$. If $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ is such that for all $a \in \mathcal{A}$, $\sigma_{MB}^{\mathcal{F}}(a) = S(a)$, where $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$ then for all $a \in \mathcal{A}$, $\sigma_{MB}^{\mathcal{F}'}(a) = S(a)$, where $(a_i, a_0) \in \mathcal{D}$, $\mathcal{F}' = \langle \mathcal{A}, \mathcal{D}', w \rangle$, $\mathcal{D}' = \mathcal{D} \setminus X$ and $X \subseteq \{(a_j, a_0) \in \mathcal{D} | a_j \neq a_i \text{ and } S(a_j) \leqslant S(a_i)\} \cup \{(a_j, a_0) | w(a_0) = 0\}$. We say that $\mathcal{D}'$ is a contraction of $\mathcal{D}$ (on $(a_i, a_0) \in \mathcal{D}$).*
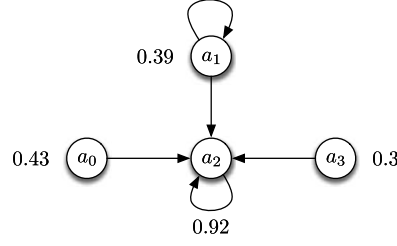
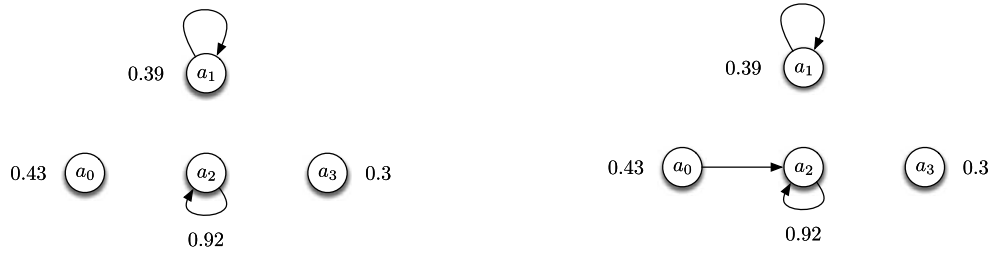Fig. 5. The argumentation framework of Example 4.

Fig. 6. Representation of $\mathcal{F}' = \langle \mathcal{A}, \mathcal{D}', w \rangle$ (left), where $\mathcal{D}'$ is a contraction of $\mathcal{D}$ on (a_2, a_2) and $\mathcal{F}'' = \langle \mathcal{A}, \mathcal{D}'', w \rangle$ (right), where $\mathcal{D}''$ is an expansion of $\mathcal{D}'$.

Example 4. Let us consider $\mathcal{A}$, w, $\mathcal{D}$ as shown in Fig. 5 and the desirable final acceptability degree function S such that $S(a_0) = 0.43$, $S(a_1) = 0.30$, $S(a_2) = 0.58$, and $S(a_3) = 0.30$. A contraction $\mathcal{D}'$ of $\mathcal{D}$ on (a_2, a_2) will remove the attacks (a_0, a_2), (a_1, a_2), and (a_3, a_2) from $\mathcal{D}$ without changing the acceptability degrees under the max-based semantics. Likewise, an expansion $\mathcal{D}''$ of $\mathcal{D}'$ can be obtained by adding the attack (a_0, a_2) as this does not change the acceptability degrees. This is represented in Fig. 6.

Once a solution for the max-based semantics is found, we can reach all the other solutions by expanding this initial solution once and then successively contracting it.

Proposition 6. *Given a set of arguments $\mathcal{A}$, a weighting function $w : \mathcal{A} \to [0, 1]$, and a desirable final acceptability degree function $S : \mathcal{A} \to [0, 1]$. If $\mathcal{D}, \mathcal{D}' \subseteq \mathcal{A} \times \mathcal{A}$ are such that for all $a \in \mathcal{A}$, $\sigma_{MB}^{\mathcal{F}}(a) = S(a)$, $\sigma_{MB}^{\mathcal{F}'}(a) = S(a)$ where $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$ and $\mathcal{F}' = \langle \mathcal{A}, \mathcal{D}', w \rangle$ then there exists $\mathcal{D}^* \subseteq \mathcal{A} \times \mathcal{A}$, such that $\mathcal{D}^*$ is an expansion of $\mathcal{D}$, and a sequence of sets of attacks $(\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n)$, where $\mathcal{D}_1 = \mathcal{D}^*$, $\mathcal{D}_n = \mathcal{D}'$ and for all $1 \leqslant i \leqslant n - 1$, $\mathcal{D}_{i+1}$ is a contraction of $\mathcal{D}_i$.*

Proof. Let us consider $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ and two arbitrary solutions $\mathcal{D}, \mathcal{D}' \subseteq \mathcal{A} \times \mathcal{A}$ for $\mathcal{DEC}_c^{MB}$. We prove this proposition by construction. By definition, for all $a \in \mathcal{A}$, $\sigma_{MB}^{\mathcal{F}}(a) = \sigma_{MB}^{\mathcal{F}'}(a) = S(a)$, where $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$ and $\mathcal{F}' = \langle \mathcal{A}, \mathcal{D}', w \rangle$. This means that for all $a \in \mathcal{A}$, such that $w(a) \neq 0$, $\max_{b \in \mathtt{Att}_{\mathcal{F}}(a)} S(b) = \max_{b \in \mathtt{Att}_{\mathcal{F}'}(a)} S(b)$. Thus, $\mathcal{D}^* = \mathcal{D} \cup \mathcal{D}'$ is an expansion of $\mathcal{D}$. Then, for all $1 \leqslant i \leqslant n$, it holds that $\mathcal{D}_{i+1} = \mathcal{D}_i \setminus \{(a_j, a_i) \in \mathcal{D} | (a_j, a_i) \notin \mathcal{D}'\}$ is a contraction of $\mathcal{D}_i$, where $\mathcal{D}_1 = \mathcal{D}^*$. It is clear that $\mathcal{D}_{n+1} = \mathcal{D}'$ as all attacks in $\mathcal{D} \setminus \mathcal{D}'$ have been removed from $\mathcal{D}^*$. $\quad\square$

4.2. The weighted h-categoriser semantics

Let us consider the case of the weighted h-categoriser semantics. Given an input $\mathcal{A} = \{a_1, \ldots, a_n\}$, an initial weight function w, and a desirable final acceptability degree function S; we seek a set of attacks $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ such that for all $a \in \mathcal{A}$, $\sigma_{HC}^{\mathcal{F}}(a) = S(a)$, where $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$.

Recall from the definition of the weighted h-categoriser semantics that any solution must mean that the following equation holds.

$$S(a_i) = \frac{w(a_i)}{1 + \sum_{b \in \mathtt{Att}(a_i)} S(b)}$$

Moreover, if $w(a_i) \neq 0$, we can rearrange the previous equation to obtain:

$$\sum_{b \in \mathtt{Att}(a_i)} S(b) = \frac{w(a_i) - S(a_i)}{S(a_i)}$$

For the complete problem, we are given $w(a_i)$ and $S(a_i)$ for all $a_i \in \mathcal{A}$, allowing us to compute $\sum_{b \in \mathtt{Att}(a_i)} S(b)$ via the above equation. In other words, for every argument a_i, we can compute a target value $T = \sum_{b \in \mathtt{Att}(a_i)} S(b)$ for which the final acceptability degrees of all its attackers must sum up to. By considering the multiset $M = \{S(a_j) | a_j \in \mathcal{A}\}$, we must solve a single subset-sum problem to identify the attackers for a single argument a_i. Extending this to all $n = |\mathcal{A}|$ arguments in our framework, we can therefore identify all attacks in the framework by solving n versions of the subset-sum problem. Algorithm 2 provides the pseudo-code for this approach.

We note in passing that the standard version of the subset-sum problem assumes that T and elements of M are all integers. If we assume that all initial weights and final acceptability degrees are rational, we can easily transform our computed T and M values to integers. The most common dynamic programming based approach to solving standard subset-sum can run in pseudo-polynomial time. However, the integers we obtain become very large very quickly, making this approach impractical. Future work will consider using state-of-the-art techniques for solving subset-sum over real numbers instead, e.g. the recent FPTAS of Costandin [13]. In addition, having transformed our problem to an instance of the subset-sum problem opens up the possibility of transforming the problem to other NP-complete problems for which efficient solvers exist, such as satisfiability.

Moreover, once a set of attacks is found to be a solution, we can obtain other solutions by replacing the attacks to a single argument x by other attacks to this argument so that the sum of the degree of the attackers remains the same. Given the simplicity of this proposition, we do not provide a proof for it here.

Proposition 7. *Consider two weighted argumentation frameworks $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$, $\mathcal{F}' = \langle \mathcal{A}, \mathcal{D}', w \rangle$, and a desirable final acceptability degree function $S : \mathcal{A} \to [0, 1]$ such that $\sigma_{HC}^{\mathcal{F}}(a) = S(a)$ for any $a \in \mathcal{A}$. If all of the following hold*

- $x \in \mathcal{A}$
- $Z \subseteq \mathcal{A} \times \{x\}$
- $\sum_{(z,x) \in Z} S(z) = (w(x) - S(x))/S(x)$ *if* $w(x) \neq 0$
- $\mathcal{D}' = (\mathcal{D} \cap (\mathcal{A} \times (a \setminus \{x\}))) \cup Z$

then for all $a \in \mathcal{A}$, $\sigma_{HC}^{\mathcal{F}'}(a) = S(a)$.

Algorithm 2 Procedure to solve $\mathcal{DEC}_c^{HC}$. SSP (lines 10, 13) calls a subset-sum solver which returns the elements of M that sum up to T, or false if no such elements can be found

---

Require: $\mathcal{A}$ a set of arguments
Require: $w : \mathcal{A} \rightarrow [0, 1]$ the initial weights for each argument
Require: $S : \mathcal{A} \rightarrow [0, 1]$ the desired final acceptability degrees for each argument
 1: **function** Solve$_{HC}(\mathcal{A}, w, S)$
 2:     $D = \{\}$
 3:     $M \leftarrow \{S(b) | b \in \mathcal{A}\}$
 4:     **for all** $a \in \mathcal{A}$ **do**
 5:         **if** $(w(a) = 0$ and $S(a) \neq 0)$ or $(w(a) \neq 0$ and $S(a) = 0)$ **then**
 6:             **return** False
 7:         **end if**
 8:         **if** $w(a) \neq 0$ and $w(a) \neq S(a)$ **then**
 9:             $T \leftarrow \frac{w(a) - S(a)}{S(a)}$
10:             **if** SSP$(T, M) =$ False **then**
11:                 **return** False
12:             **end if**
13:             $D = D \cup \{(b, a) | S(b) \in \text{SSP}(T, M)\}$
14:         **end if**
15:     **end for**
16:     **return** D
17: **end function**

4.3. The weighted card-based semantics

We can adopt a similar approach for σ_{CB} as was done for σ_{HC}. We begin by assuming that there are k attacks from arguments with non-zero initial weight against an argument a_i. Any solution must mean that the following equation holds:

$$S(a_i) = \frac{w(a_i)}{1 + k + \frac{1}{k} \sum_{b \in \text{Att}(a_i)} S(b)}$$

If $w(a_i) \neq 0$, rearranging gives us the equation:

$$\sum_{b \in \text{Att}(a_i)} S(b) = -\frac{k(S(a_i)k + S(a_i) - w(a_i))}{S(a_i)}$$

Again, all values on the right-hand side of the equation are given for the complete problem given, allowing us to compute the value of the left-hand side for each argument. Assuming a non-zero number of attacks, while we could (naively) iterate over all $k = 1 \ldots n$ to identify the number of attacks against an argument, and thereby determine k, we observe that if $S(b) = \epsilon > 0$, $\lim \epsilon \rightarrow 0$, for all $b \in \text{Att}(a_i)$ the semantic equation reduces to the following.

$$S(a_i) = \frac{w(a_i)}{1 + k}$$

Solving for k, we see that:

$$k = \frac{w(a_i) - S(a_i)}{S(a_i)}$$

Now assume instead that $S(b) = 1$ for all $b \in \mathtt{Att}(a_i)$. Performing the same operations means that:

$$k = \frac{w(a_i) - 2S(a_i)}{S(a_i)}$$

These two equations serve as an upper and lower bound for k. Since k must be an integer, we can avoid the need to iterate over the number of attacks by computing k as follows.

$$k = \left\lfloor \frac{w(a_i) - S(a_i)}{S(a_i)} \right\rfloor = \left\lceil \frac{w(a_i) - 2S(a_i)}{S(a_i)} \right\rceil$$

The pseudo-code for this approach is described in Algorithm 3.

Since k is known, the $k\mathcal{SSP}$ solver called in Algorithm 3 could potentially be more efficient than that used in the h-categoriser semantics.

Analogous to the result from Proposition 7, under the weighted card-based semantics, once a set of attacks is found to be a solution, we can obtain other solutions by replacing the k attacks on a single

---

Algorithm 3 Procedure to solve $\mathcal{DEC}_c^{CB}$. kSSP (lines 11, 14) calls a subset-sum solver which returns k elements of M which sum up to T

Require: $\mathcal{A}$ a set of arguments
Require: $w : \mathcal{A} \to \mathbb{Q}$ the initial weights for each argument
Require: $S : \mathcal{A} \to \mathbb{Q}$ the final acceptability degree for each argument
1: **function** $\text{Solve}_{CB}(\mathcal{A}, w, S)$
2:     $D = \{\}$
3:     $M \leftarrow \{S(b) | b \in \mathcal{A}\}$
4:     **for all** $a \in \mathcal{A}$ **do**
5:         **if** $(w(a) = 0$ and $S(a) \neq 0)$ or $(w(a) \neq 0$ and $S(a) = 0)$ **then**
6:             **return** False
7:         **end if**
8:         **if** $w(a) \neq 0$ and $w(a) \neq S(a)$ **then**
9:             $k \leftarrow \lfloor \frac{w(a) - S(a)}{S(a)} \rfloor$
10:            $T \leftarrow -\frac{k(S(a)k + S(a) - w(a))}{S(a)}$
11:            **if** $\text{KSSP}(T, M, k) =$ False **then**
12:                **return** False
13:            **end if**
14:            $D \leftarrow D \cup \{(b, a) | S(b) \in \text{KSSP}(T, M, k)\}$
15:        **end if**
16:    **end for**
17:    **return** D
18: **end function**

argument x by another set of k attacks to this argument so that the sum of the final acceptability degree of the attackers remains the same. Given the simplicity of this proposition, we do not provide a proof for it here.

Proposition 8. *Given two weighted argumentation frameworks $\mathcal{F} = \langle \mathcal{A}, \mathcal{D}, w \rangle$, $\mathcal{F}' = \langle \mathcal{A}, \mathcal{D}', w \rangle$, and a desirable final acceptability degree function $S : \mathcal{A} \to [0, 1]$ such that for all $a \in \mathcal{A}$, $\sigma_{CB}^{\mathcal{F}}(a) = S(a)$. If all the following hold:*

- $x \in \mathcal{A}$
- $Z \subseteq \mathcal{A} \times \{x\}$, *and*
- $\mathcal{D}' = (\mathcal{D} \cap (\mathcal{A} \times (\mathcal{A} \setminus \{x\}))) \cup Z$,
- $|Z| = \lfloor \frac{w(a) - S(a)}{S(a)} \rfloor$ *and* $\sum_{(z,x) \in Z} S(z) = -\frac{k(S(x)k + S(x) - w(x))}{S(x)}$ *if $w(x) \neq 0$.*

then for all $a \in \mathcal{A}$, it is the case that $\sigma_{CB}^{\mathcal{F}'}(a) = S(a)$

Propositions 4, 5, and 6 demonstrate that there are typically a large number of non-minimal argumentation frameworks for the weighted max-based semantics. For the weighted h-categoriser semantics, attacks can be substituted as long as the sum of the final acceptability degrees of the attacking arguments match (see Proposition 7), while for weighted card-based semantics, attacks can only be substituted by a set of attackers of the same size which final acceptability degree sums to the same value (see Proposition 8). We also note that, for all semantics, any argument with an initial weight of 0 can clearly have any number of attacks against it.

5. Discussion and future work

To this point, we have considered only the complete problem of attack inference in gradual argumentation frameworks, and we now briefly discuss the partial problem. Here, we are given a semantics, a set of arguments, a partial mapping between arguments and initial weights, and a partial mapping between arguments and final acceptability degrees, and we must determine whether attacks (and initial weights which were not provided) can be identified which make the given final acceptability degrees consistent with the semantics.

Since the complete problem is a special case of this partial problem, and since we can still verify correctness in polynomial time, the NP-completeness results of Section 3 still apply to the weighted card-based and weighted h-categoriser semantics. The behaviour of the weighted max-based version of the partial problem, which was polynomial in the complete problem, is more challenging to characterise. While it is easy to show that there are polynomial instances of the partial problem (e.g., when all initial weights are given and all but one final acceptability degree is provided), we strongly believe that the general case is NP-complete. We leave proof of this result as an element of future work. Other directions for research include providing a mapping from subset-sum to other problems for which solvers exist (e.g., satisfiability) and evaluating the performance of such solvers.[2] Argumentation researchers have found that this approach has often yielded excellent results [40]. In addition, it may be interesting to

---

[2]An implementation of our algorithms using an optimised depth-first-search SSP solver can be found at https://github.com/jhudsy/Gradual_Attack_Inference. We do not provide experimental data as the performance of our solver demonstrates the exponential growth of the underlying subset-sum problem. Systems with more than ∼13 arguments can only rarely be solved in reasonable time using our implementation.

evaluate such solvers on different (given) underlying graph topologies. Finally, we focused on three gradual semantics in this paper. However, many other weighted semantics have been proposed (e.g., [10,15,18,22,28]), and extending the current work to these, and in particular, to probabilistic semantics [19,23] would be interesting.

Another variant of the problem under consideration would involve identifying whether a subset of attacks from some given possible attacks (rather than all attacks) can be used to make the final acceptability degrees consistent with the initial weights. While this may be more realistic for some applications, from a complexity point of view, the results from the current work directly translate to this version of the problem.

We note in passing that we can consider another inverse problem, namely the computation of semantics for a given weighted argumentation framework and set of final acceptability degrees or preferences over arguments. This problem is trivial to solve in polynomial time, as one simply needs to check for consistency between the inputs (weighted argumentation framework) and outputs (final acceptability degrees or preferences) for each semantics.

Turning to related work, researchers have examined the notion of *argument realisability* in abstract argumentation frameworks [25], seeking to identify an attack relation that yields a specific labelling or extension. Recent work by Mumford et al. [31] has shown that for complete semantics and IN/OUT labellings the problem is NP-complete, whereas it can be solved in polynomial time if UNDEC labellings are also allowed. The work of Skiba et al. [38] focuses on whether, for a given ranking and ranking-based semantics, we can find an unweighted argumentation framework such that the selected ranking-based semantics induces the ranking when applied to the argumentation framework. They show that the above problem is true for a number of ranking-based semantics, including burden-based and discussion-based semantics [1], the simple product semantics on social abstract frameworks [10,22], and the probabilistic graded semantics [24,39]. Another related area to the current work is *argument synthesis* [32], where the attack relation must satisfy some positive and negative constraints, but where the problem is not fully constrained. There are still several open questions regarding the time complexity of these types of problems, even for the case of abstract argumentation semantics [31].

More generally, our work can be seen as a type of inverse argumentation problem. [33,34] has examined one such inverse problem in the context of the gradual semantics, seeking to identify a set of initial weights given a semantics, a set of arguments, attacks and final acceptability degrees. In the context of abstract argumentation, such inverse problems have examined inferring preferences from justified arguments [26,27], and has applications in the context of belief revision [8].

The current paper focuses on the complexity of the underlying decision problem and inferring attacks given the semantics, initial weights, and final acceptability degrees. This is arguably unrealistic and limits the applications of the current work. In most uses of gradual semantics, one considers the preferences obtained from final acceptability degrees rather than the final acceptability degrees themselves, as the latter are not (normally) exposed within an application. Thus, as future work, we intend to consider the problem of inferring attacks from a semantics, initial weights, and preferences over arguments. We believe that the complexity of this problem for a given semantics will be similar to that obtained in the current work. Given this more general problem one could, for example, determine whether an agent is rational given the initial weights they assign to a set of arguments, a semantics, and a resulting preference ordering over the arguments. Other applications, as discussed above, include opponent modelling [37] in the context of dialogue and preference elicitation.

6. Conclusions

We have considered the problem of inferring an attack relation given a gradual semantics and all initial weights and final acceptability degrees for an argumentation framework. We have shown that for the weighted max-based semantics, this problem can be solved in linear time, but that the obtained solution is (typically) not unique. For the weighted h-categoriser and weighted card-based semantics, where solutions have no redundancies, the problem becomes NP-complete. Our proofs are based on a reduction to variants of the subset-sum problem, and efficient solvers for this problem can be applied to our work, facilitating its application in domains such as opponent modelling. Finally, we have focused on the complete problem, and there are exciting avenues for future research dealing with its partial form.

References

[1] L. Amgoud and J. Ben-Naim, Ranking-based semantics for argumentation frameworks, in: *Proc. Scalable Uncertainty Management*, 2013, pp. 134–147. doi:10.1007/978-3-642-40381-1_11.

[2] L. Amgoud and J. Ben-Naim, Weighted bipolar argumentation graphs: Axioms and semantics, in: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018*, Stockholm, Sweden, July 13–19, 2018, J. Lang, ed., ijcai.org, 2018, pp. 5194–5198.

[3] L. Amgoud, C. Cayrol, M.-C. Lagasquie-Schiex and P. Livet, On bipolarity in argumentation frameworks, *Int. J. Intell. Syst.* **23**(10) (2008), 1062–1093. doi:10.1002/int.20307.

[4] L. Amgoud, D. Doder and S. Vesic, Evaluation of argument strength in attack graphs: Foundations and semantics, *Artificial Intelligence* **302** (2022), 103607. doi:10.1016/j.artint.2021.103607.

[5] L. Amgoud and S. Vesic, Two roles of preferences in argumentation frameworks, in: *Symbolic and Quantitative Approaches to Reasoning with Uncertainty – 11th European Conference, ECSQARU 2011*, Belfast, UK, June 29–July 1, 2011, Proceedings, 2011, pp. 86–97.

[6] L. Amgoud and S. Vesic, Rich preference-based argumentation frameworks, *Int. J. Approx. Reasoning* **55**(2) (2014), 585–606. doi:10.1016/j.ijar.2013.10.010.

[7] P. Baroni, M. Caminada and M. Giacomin, An introduction to argumentation semantics, *Knowledge Eng. Review* **26**(4) (2011), 365–410. doi:10.1017/S0269888911000166.

[8] R. Baumann and G. Brewka, Agm meets abstract argumentation: Expansion and revision for dung frameworks, in: *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[9] S. Benferhat, D. Dubois and H. Prade, Argumentative inference in uncertain and inconsistent knowledge bases, in: *Uncertainty in Artificial Intelligence*, D. Heckerman and A. Mamdani, eds, Morgan Kaufmann, 1993, pp. 411–419. doi:10.1016/B978-1-4832-1451-1.50054-8.

[10] E. Bonzon, J. Delobelle, S. Konieczny and N. Maudet, A comparative study of ranking-based semantics for abstract argumentation, in: *Proc. AAAI*, 2016, pp. 914–920.

[11] E. Bonzon, J. Delobelle, S. Konieczny and N. Maudet, Combining extension-based semantics and ranking-based semantics for abstract argumentation, in: *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018*, Tempe, Arizona, 30 October–2 November, 2018, 2018, pp. 118–127.

[12] C. Cayrol and M.-C. Lagasquie-Schiex, On the acceptability of arguments in bipolar argumentation frameworks, in: *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 8th European Conference, ECSQARU 2005*, Barcelona, Spain, July 6–8, 2005, Proceedings, 2005, pp. 378–389.

[13] M. Costandin, A FPTAS for the subset sum problem with real numbers, arXiv preprint, 2021. arXiv:2110.09901.

[14] S. Coste-Marquis, S. Konieczny, P. Marquis and M.A. Ouali, Selecting extensions in weighted argumentation frameworks, in: *Computational Models of Argument – Proceedings of COMMA 2012*, Vienna, Austria, September 10–12, 2012, 2012, pp. 342–349.

[15] C. da Costa Pereira, A. Tettamanzi and S. Villata, Changing one's mind: Erase or rewind? in: *ICJAI*, T. Walsh, ed., 2011, pp. 164–171.

[16] L.D. de Tarlé, E. Bonzon and N. Maudet, Multiagent dynamics of gradual argumentation semantics, in: *21st International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2022*, Auckland, New Zealand, May 9–13, 2022, P. Faliszewski, V. Mascardi, C. Pelachaud and M.E. Taylor, eds, International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2022, pp. 363–371.

[17] P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, *Artif. Intell.* **77**(2) (1995), 321–358. doi:10.1016/0004-3702(94)00041-X.

[18] D.M. Gabbay and O. Rodrigues, Equilibrium states in numerical argumentation networks, *Logica Universalis* **9**(4) (2015), 411–473. doi:10.1007/s11787-015-0119-7.

[19] A. Hunter and M. Thimm, Probabilistic reasoning with abstract argumentation frameworks, *J. Artif. Intell. Res.* **59** (2017), 565–611. doi:10.1613/jair.5393.

[20] B. Irwin, A. Rago and F. Toni, Argumentative forecasting, in: *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems, AAMAS '22*, Richland, SC, International Foundation for Autonomous Agents and Multiagent Systems, 2022, pp. 1636–1638.

[21] S. Kaci, *Working with Preferences: Less Is More*, Cognitive Technologies, Springer, 2011.

[22] J. Leite and J. Martins, Social abstract argumentation, in: *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, Barcelona, Catalonia, Spain, July 16–22, 2011, 2011, pp. 2287–2292.

[23] H. Li, N. Oren and T.J. Norman, Probabilistic argumentation frameworks, in: *TAFA*, Lecture Notes in Computer Science, Vol. 7132, 2011, pp. 1–16.

[24] H. Li, N. Oren and T.J. Norman, Probabilistic argumentation frameworks, in: *Theorie and Applications of Formal Argumentation – First International Workshop, TAFA 2011*, Barcelona, Spain, July 16–17, 2011, Revised Selected Papers, S. Modgil, N. Oren and F. Toni, eds, Lecture Notes in Computer Science, Vol. 7132, Springer, 2011, pp. 1–16.

[25] T. Linsbichler, J. Puehrer and H. Strass, Characterizing realizability in abstract argumentation, in: *Proceedings of the 16th International Workshop on Non-monotonic Reasoning*, 2020.

[26] Q.-A. Mahesar, N. Oren and W.W. Vasconcelos, Computing preferences in abstract argumentation, in: *Proceedings*, PRIMA 2018: Principles and Practice of Multi-Agent Systems – 21st International Conference, Tokyo, Japan, October 29–November 2, 2018, pp. 387–402.

[27] Q. Mahesar, N. Oren and W.W. Vasconcelos, Preference elicitation in assumption-based argumentation, in: *PRIMA 2020: Principles and Practice of Multi-Agent Systems – 23rd International Conference*, Nagoya, Japan, November 18–20, 2020, T. Uchiya, Q. Bai and I. Marsá-Maestre, eds, Lecture Notes in Computer Science, Vol. 12568, Springer, 2020, pp. 199–214. doi:10.1007/978-3-030-69322-0_13.

[28] P.-A. Matt and F. Toni, A game-theoretic measure of argument strength for abstract argumentation, in: *Logics in Artificial Intelligence, 11th European Conference, JELIA 2008*, Dresden, Germany, September 28–October 1, 2008, Proceedings, 2008, pp. 285–297.

[29] S. Modgil and H. Prakken, A general account of argumentation with preferences, *Artif. Intell.* **195** (2013), 361–397. doi:10.1016/j.artint.2012.10.008.

[30] T. Mossakowski and F. Neuhaus, Modular semantics and characteristics for bipolar weighted argumentation graphs, *CoRR*, 2018. arXiv:1807.06685.

[31] J. Mumford, I. Sassoon, E. Black and S. Parsons, On the complexity of determining defeat relations consistent with abstract argumentation semantics, in: *Proceedings of COMMA*, 2022.

[32] A. Niskanen, J. Wallner and M. Järvisalo, Synthesizing argumentation frameworks from examples, *Journal of Artificial Intelligence Research* **66** (2019), 503–554. doi:10.1613/jair.1.11758.

[33] N. Oren, B. Yun, A. Libman and M.S. Baptista, Analytical solutions for the inverse problem within gradual semantics, *CoRR*, 2022. arXiv:2203.01201.

[34] N. Oren, B. Yun, S. Vesic and M. Baptista, Inverse problems for gradual semantics, in: *Proceedings of the 31st International Joint Conference on Artificial Intelligence and the 25th European Conference on Artificial Intelligence*, 2022.

[35] S. Polberg and A. Hunter, Empirical evaluation of abstract argumentation: Supporting the need for bipolar and probabilistic approaches, *International Journal of Approximate Reasoning* **93** (2018), 487–543. doi:10.1016/j.ijar.2017.11.009.

[36] A. Rago, F. Toni, M. Aurisicchio and P. Baroni, Discontinuity-free decision support with quantitative argumentation debates, in: *KR 2016*, 2016, pp. 63–73.

[37] T. Rienstra, M. Thimm and N. Oren, Opponent models with uncertainty for strategic argumentation, in: *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, 2013, pp. 332–338.

[38] K. Skiba, M. Thimm, T. Rienstra, J. Heyninck and G. Kern-Isberner, Realisability of rankings-based semantics, in: *Proceedings of the Fourth International Workshop on Systems and Algorithms for Formal Argumentation Co-Located with the 9th International Conference on Computational Models of Argument (COMMA 2022)*, Cardiff, Wales, United Kingdom, September 13, 2022, S.A. Gaggl, J. Mailly, M. Thimm and J.P. Wallner, eds, CEUR Workshop Proceedings, Vol. 3236, CEUR-WS.org, 2022, pp. 73–85.

[39] M. Thimm, F. Cerutti and T. Rienstra, Probabilistic graded semantics, in: *Computational Models of Argument – Proceedings of COMMA 2018*, Warsaw, Poland, 12–14 September 2018, S. Modgil, K. Budzynska and J. Lawrence, eds, Frontiers in Artificial Intelligence and Applications, Vol. 305, IOS Press, 2018, pp. 369–380.

[40] M. Thimm and S. Villata, The first international competition on computational models of argumentation: Results and analysis, *Artif. Intell.* **252** (2017), 267–294. doi:10.1016/j.artint.2017.08.006.

[41] B. Yun, P. Bisquert, P. Buche and M. Croitoru, Arguing about end-of-life of packagings: Preferences to the rescue, in: *Metadata and Semantics Research – 10th International Conference, MTSR 2016*, Göttingen, Germany, November 22–25, 2016, Proceedings, 2016, pp. 119–131.

[42] B. Yun, S. Vesic, M. Croitoru and P. Bisquert, Viewpoints using ranking-based argumentation semantics, in: *Proceedings of the 7th International Conference on Computational Models of Argument, COMMA 2018*, Warsaw, Poland, 11th–14th September, 2018, 2018.