Computational Models of Argument
H. Prakken et al. (Eds.)

doi:10.3233/FAIA200512

Possible Controllability of Control Argumentation Frameworks

Jean-Guy MAILLY
LIPADE, Université de Paris, jean-guy.mailly@u-paris.fr

Abstract. The recent Control Argumentation Framework (CAF) is a generalization of Dung's Argumentation Framework which handles argumentation dynamics under uncertainty; especially it can be used to model the behavior of an agent which can anticipate future changes in the environment. Here we provide new insights on this model by defining the notion of possible controllability of a CAF. We study the complexity of this new form of reasoning for the four classical semantics, and we provide a logical encoding for reasoning with this framework.

Keywords. Abstract argumentation, uncertainty, computational complexity

1. Introduction

Abstract argumentation [1] has become an important subfield of Knowledge Representation and Reasoning research in the last decades. Intuitively, an abstract argumentation framework (AF) is a directed graph where nodes are arguments and edges are relations (usually attacks) between these arguments. The outcome of such an AF is an evaluation of the arguments' acceptance (through extensions [1,2], labellings [3] or rankings [4]). The question of argumentation dynamics has arisen more recently, and many different approaches have been proposed (see *e.g.* [5,6,7,8,9,10,11,12,13,14]). Roughly speaking, the question of these works is "how to modify an AF to be consistent with a given piece of information?". Such a piece of information can be "argument a should be accepted in the outcome of the AF". A particular version of this problem is called *extension enforcement* [7,15,10,12,13]: it consists in modifying an AF s.t. a given set of arguments becomes (included in) an extension of the AF. The recently proposed *Control Argumentation Framework* (CAF) [14] is a generalization of Dung's AF which incorporates different notions of uncertainty in the structure of the framework. The *controllability* of a CAF w.r.t. a set of arguments is the fact that, whatever happens in the uncertain part of the CAF (*i.e.* whatever is the real situation of the world), the target set of arguments is accepted. This is somehow a generalization of extension enforcement, where uncertainty is taken into account.

In this paper, we study what we call *possible controllability* (and then, controllability defined in [14] can be renamed as *necessary controllability*). The idea of possible controllability w.r.t. a target set of arguments is that this target should be accepted in *at least* one of the possible completions of the uncertain part. Necessary controllability trivially implies possible controllability, while the converse is not true. This form of reasoning can be applied in different situations. Possible controllability makes sense in situations where an agent is unable to guarantee some result (the fact that some argument a is accepted), but she wants to be sure that the opposite result (a is rejected) is not necessary true. For instance, possible controllability is similar to the reasoning of the defendant's lawyer during a trial. Thanks to the principle of *presumption of innocence*, the lawyer does not have to prove that the defendant *is* innocent, but he has to prove that the defendant *may be* innocent. This means that if there is some uncertainty in the case, the lawyer wants to exhibit the fact that one possible world encompassed by this uncertainty implies that his client is innocent.[1] This means that the lawyer's knowledge about the case can be represented by a CAF, and the lawyer wants to guarantee that the argument "the defendant is innocent" is accepted in at least one completion of the CAF, *i.e.* one possible world. In this kind of scenario, possible controllability is particularly useful since it is (presumably) easier to search for one completion that accepts the target instead of checking that the target is accepted in each of the (exponentially many) completions.

The paper is organized as follows. We first recall the background notions of logic and introduce the CAF setting in Section 2. In Section 3 we define formally this new form of controllability, and we determine the complexity of this reasoning problem for the four classical semantics introduced by Dung. We also propose a QBF-based encoding which allows to determine whether a CAF is possible controllable w.r.t. a target and the stable semantics (and moreover, which allows to determine *how* to control it). We describe the related work in Section 4, and finally Section 5 concludes the paper and draws interesting future research tracks.

2. Background

2.1. Propositional Logic and Quantified Boolean Formulas

We consider a set V of Boolean variables, *i.e.* variables which can be assigned a value in $\mathbb{B} = \{0,1\}$, where 0 and 1 are associated respectively with *false* and *true*. Such variables can be combined with connectives $\{\vee, \wedge, \neg\}$ to build formulas. $x \vee y$ is true if at least one of the variables x, y is true; $x \wedge y$ is true if both x, y are true; $\neg x$ is true is x is false. Additional connectives can be defined, *e.g.* $x \Rightarrow y$ is equivalent to $\neg x \vee y$; $x \Leftrightarrow y$ is equivalent to $(x \Rightarrow y) \wedge (y \Rightarrow x)$. The definition of the connectives is straightforwardly extended from variables to formulas (*e.g.* if ϕ and ψ are formulas, then $\phi \wedge \psi$ is true when both formulas are true). A truth assignment on the set of variables $V = \{x_1, \dots, x_n\}$ is a mapping $\omega : V \to \mathbb{B}$.

Quantified Boolean Formulas (QBFs) are an extension of propositional formulas with the universal and existential quantifiers. For instance, the formula $\exists x \forall y (x \vee \neg y) \wedge (\neg x \vee y)$ is satisfied if there is a value for x such that for all values of y the proposition $(x \vee \neg y) \wedge (\neg x \vee y)$ is true. More formally, a canonical QBF is a formula $\mathcal{Q}_1 X_1 \mathcal{Q}_2 X_2 \dots \mathcal{Q}_n X_n \Phi$ where Φ is a propositional formula, $\mathcal{Q}_i \in \{\exists, \forall\}$, $\mathcal{Q}_i \neq \mathcal{Q}_{i+1}$, and $X_1, X_2, \dots, X_n$ disjoint sets of propositional variables such that $X_1 \cup X_2 \cup \dots \cup X_n = V$.[2] It is well-known that QBFs span the polynomial hierarchy. For instance, deciding whether the formula $\exists X_1 \forall X_2 \dots \mathcal{Q}_i X_i \Phi$ is true is Σ_i^p-complete. The decision problem associated

[1] On the opposite, necessary controllability [14] is close to the reasoning of the prosecutor.

[2] If some variable $x \in V$ does not explicitly belong to any X_i, *i.e.* $X_1 \cup \dots \cup X_n \subset V$, then it implicitly means that x can be existentially quantified at the rightmost level.

to QBFs of the form $\exists V, \Phi$ is equivalent to the satisfiability problem for propositional formulas (SAT), which is well-known to be NP-complete. For more details about propositional logic, QBFs and complexity theory, we refer the reader to [16,17,18].

2.2. Abstract Argumentation and Control Argumentation Frameworks

An *argumentation framework* (AF), introduced in [1], is a directed graph $\mathcal{AF} = \langle A, R \rangle$, where A is a set of *arguments*, and $R \subseteq A \times A$ is an *attack relation*. The relation *a attacks b* is denoted by $(a,b) \in R$. In this setting, we are not interested in the origin of arguments and attacks, nor in their internal structure. Only their relations are important to define the acceptance of arguments.

In [1], different acceptability semantics were introduced. They are based on two basic concepts: *conflict-freeness* and *defence*. A set $S \subseteq A$ is:

- conflict-free iff $\forall a, b \in S$, $(a,b) \notin R$;
- admissible iff it is conflict-free, and defends each $a \in S$ against its attackers.

The semantics defined by Dung are as follows. An admissible set $S \subseteq A$ is:

- a complete extension iff it contains every argument that it defends;
- a preferred extension iff it is a $\subseteq$-maximal complete extension;
- the unique grounded extension iff it is the $\subseteq$-minimal complete extension;
- a stable extension iff it attacks every argument in $A \setminus S$.

The sets of extensions of an $\mathcal{AF}$, for these four semantics, are denoted (respectively) $\text{co}(\mathcal{AF})$, $\text{pr}(\mathcal{AF})$, $\text{gr}(\mathcal{AF})$ and $\text{st}(\mathcal{AF})$.

Our approach could be adapted for any other extension semantics. Based on these semantics, we can define the status of any (set of) argument(s), namely *skeptically accepted* (belonging to each σ-extension), *credulously accepted* (belonging to some σ-extension) and *rejected* (belonging to no σ-extension). For more details about argumentation semantics, we refer the reader to [1,2].

We introduce now the notions of CAF and (necessary) controllability [14].

Definition 1. *A Control Argumentation Framework* (CAF) *is a triple $\mathcal{CAF} = \langle \mathcal{F}, \mathcal{C}, \mathcal{U} \rangle$ where $\mathcal{F}$ is the* fixed part*, $\mathcal{U}$ is the* uncertain part *and $\mathcal{C}$ is the* control part *of $\mathcal{CAF}$ with:*

- *$\mathcal{F} = \langle A_F, \rightarrow \rangle$ where A_F is a set of arguments and $\rightarrow \subseteq (A_F \cup A_U) \times (A_F \cup A_U)$ is an attack relation.*
- *$\mathcal{U} = \langle A_U, (\rightleftarrows \cup \dashrightarrow) \rangle$ where A_U is a set of arguments, $\rightleftarrows \subseteq (((A_U \cup A_F) \times (A_U \cup A_F)) \setminus \rightarrow)$ is a conflict relation and $\dashrightarrow \subseteq (((A_U \cup A_F) \times (A_U \cup A_F)) \setminus \rightarrow)$ is an attack relation, with $\rightleftarrows \cap \dashrightarrow = \emptyset$.*
- *$\mathcal{C} = \langle A_C, \Rightarrow \rangle$ where A_C is a set of arguments, and $\Rightarrow \subseteq \{(a_i, a_j) \mid a_i \in A_C,\ a_j \in A_F \cup A_C \cup A_U\}$ is an attack relation.*

A_F, A_U and A_C are disjoint subsets of arguments.

The different sets of arguments and attacks have different meanings. The fixed part $\mathcal{F}$ represents the part of the system which cannot be influenced either by the agent or by the environment. This means that if $a \in A_F$, then it is sure that a is an "active" argument (for instance, all of its premises are true, and cannot be falsified). Similarly, if $(a,b) \in \rightarrow$, the attack from a to b is actually part of the system and cannot be removed.

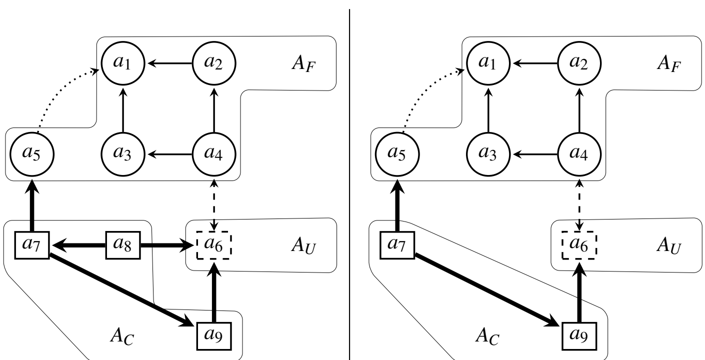
$\mathscr{U}$ is the uncertain part of the system. This means that it cannot be influenced by the agent, but it can be modified by the environment (in a wide way, this can also represent the possible actions of other agents). The uncertainty can appear in different ways. First, if $a \in A_U$, this means that there is some uncertainty about the presence of an argument (for instance, the agent is not sure whether her opponent in the debate will state argument a, or she is not sure whether the premises of a will be true at some moment). If $(a,b) \in \rightleftarrows$, then the agent is sure that there is a conflict between a and b, but she is not sure of the direction of the attack (this could be an attack (a,b), an attack (b,a), or even both at the same time). This is possible, for instance, if the agent is not sure about some preference between a and b [19]. Finally, $(a,b) \in \dashrightarrow$ means that the agent is not sure whether there is actually an attack from a to b.

The last part $\mathscr{C}$ is the *control* part. This is the part of the system which can be influenced by the agent. This means that the agent has to choose which arguments she will actually use (uttering them in the debate, or making an action to switch their premises to true). When the agent uses a subset $A_{conf} \subseteq A_C$, called a *configuration*, this defines a configured CAF where the arguments from $A_C \setminus A_{conf}$ (and the attacks concerning them) are removed. We illustrate these concepts on an example adapted from [14].

Example 1. *We define* $\mathscr{CAF} = \langle F, C, U \rangle$ *as follows:*

- $\mathscr{F} = \langle \{a_1, a_2, a_3, a_4, a_5\}, \{(a_2, a_1), (a_3, a_1), (a_4, a_2), (a_4, a_3)\} \rangle$;
- $\mathscr{U} = \langle \{a_6\}, \rightleftarrows \cup \dashrightarrow \rangle$, *with* $\rightleftarrows = \{(a_6, a_4)\}$, *and* $\dashrightarrow = \{(a_5, a_1)\}$;
- $\mathscr{C} = \langle \{a_7, a_8, a_9\}, \{(a_7, a_5), (a_7, a_9), (a_8, a_6), (a_8, a_7), (a_9, a_6)\} \rangle$.

$\mathscr{CAF}$ is given at Figure 1a. The configuration of $\mathscr{CAF}$ by $A_{conf} = \{a_7, a_9\}$ yields the configured CAF $\mathscr{CAF}'$ described at Figure 1b. On the figures, arguments from A_F, A_U and A_C are respectively represented as circle nodes, dashed square nodes and plain square nodes. Similarly, the attacks from $\rightarrow$, $\rightleftarrows$, $\dashrightarrow$ and $\Rrightarrow$ are represented (respectively) as plain, double-headed dashed, dotted and bold arrows.



(a) The CAF $\mathscr{CAF}$ (b) $\mathscr{CAF}$ configured by $A_{conf} = \{a_7, a_9\}$

Figure 1. A CAF and a configured CAF

Now we recall the notion of completion, borrowed from [20], and adapted to CAFs in [14]. Intuitively, a completion is a classical AF which describes a situation of the world coherent with the uncertain information encoded in the CAF.

Definition 2. *Given $\mathscr{CAF} = \langle F, C, U\rangle$, a completion of $\mathscr{CAF}$ is $\mathscr{AF} = \langle A, R\rangle$, s.t.*

- $A = A_F \cup A_C \cup A_{comp}$ *where* $A_{comp} \subseteq A_U$*;*
- *if* $(a,b) \in R$*, then* $(a,b) \in \rightarrow \cup \rightleftarrows \cup \dashrightarrow \cup \Rrightarrow$*;*
- *if* $(a,b) \in \rightarrow$*, then* $(a,b) \in R$*;*
- *if* $(a,b) \in \rightleftarrows$ *and* $a,b \in A$*, then* $(a,b) \in R$ *or* $(b,a) \in R$*;*
- *if* $(a,b) \in \Rrightarrow$ *and* $a,b \in A$*, then* $(a,b) \in R$*.*

Example 2 (Continuation of Example 1)**.** *We describe two possible completions of $\mathscr{CAF}'$. First, we consider a completion $\mathscr{AF}_1$ where the attack (a_5, a_1) is not included, while the argument a_6 (with the attack (a_6, a_4)) is included. Another possible completion is $\mathscr{AF}_2$, where a_6 is not included (so, neither the attacks related to it) while the attack (a_5, a_1) is included.*

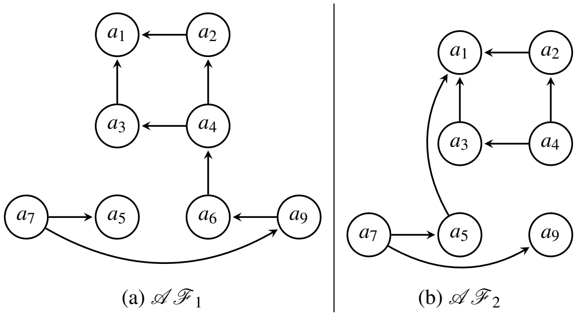


Figure 2. Two possible completions of $\mathscr{CAF}'$

Now, a CAF is necessary controllable w.r.t. a target $T \subseteq A_F$ if the agent can configure it in a way which guarantees that T is accepted in every completion of the configured CAF. This necessary controllability has two versions, depending on the kind of acceptance under consideration (skeptical or credulous).

Definition 3. *Given a set of arguments $T \subseteq A_F$ and a semantics σ, $\mathscr{CAF}$ is* necessary skeptically (resp. credulously) controllable *w.r.t. T and σ iff $\exists A_{conf} \subseteq A_C$ s.t. T is included in each (resp. some) σ-extension of each completion of $\mathscr{CAF}' = \langle F, C', U\rangle$, with $C' = \langle A_{conf}, \{(a_i, a_j) \in \Rrightarrow \mid a_i, a_j \in (A_F \cup A_U \cup A_{conf})\}\rangle$.*

[14] proposes a QBF-based method to determine whether a CAF is necessary controllable, and to obtain the corresponding configuration if it exists.

3. Possible Controllability

3.1. Formal Definition of Possible Controllability

The intuition of necessary controllability is that the agent is satisfied when its target is reached in every possible world encoded by the uncertain information in the CAF. While this is an interesting property (especially for applications like negotiation [21]), this may seem unrealistic for some applications, where the graph is built in such a way that some completions cannot accept the target. Here, we adapt the definition of controllability to consider that the agent is satisfied whether there exists at least one possible world (*i.e.* one completion) which accepts the target.

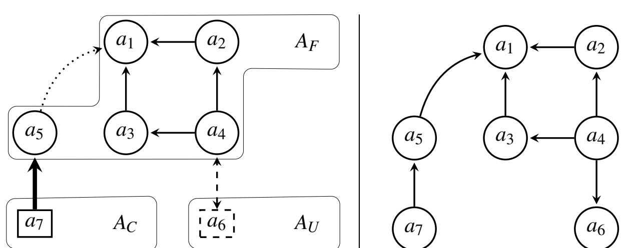
Definition 4. *Given a set of arguments $T \subseteq A_F$ and a semantics σ, $\mathscr{CAF}$ is* possibly skeptically (resp. credulously) controllable *w.r.t. T and σ iff $\exists A_{conf} \subseteq A_C$ s.t. T is included in each (resp. some) σ-extension of some completion of $\mathscr{CAF}' = \langle F, C', U \rangle$, with $C' = \langle A_{conf}, \{(a_i, a_j) \in \Rightarrow \mid a_i, a_j \in (A_F \cup A_U \cup A_{conf})\} \rangle$.*

Observation 1. Given a set of arguments $T \subseteq A_F$ and a semantics σ, if $\mathscr{CAF}$ is necessary skeptically (resp. credulously) controllable w.r.t. T and σ, then $\mathscr{CAF}$ is possibly skeptically (resp. credulously) controllable w.r.t. T and σ. The converse is false.

Example 3 (Continuation of Example 1)**.** *We observe that $\mathscr{CAF}$ from the previous example is not necessary skeptically controllable w.r.t. the target $\{a_1\}$. Indeed,*

- *if $A_{conf} = \{a_7, a_8, a_9\}$, then because of the attack (a_8, a_7), the target is not defended against the potential threat $(a_5, a_1) \in \dashrightarrow$. The same thing happens if $A_{conf} = \{a_7, a_8\}$ or $A_{conf} = \{a_8, a_9\}$.*
- *if $A_{conf} = \{a_7, a_9\}$, this time the target is not defended against the potential threat coming from a_6 (in the completions where a_6 belongs to the system, along with the attack (a_6, a_4), a_1 is not accepted).*
- *if A_{conf} is one of the three possible singletons, then again a_1 is not accepted in every completion (since either a_5 or a_6 is unattacked).*

On the opposite, it is possible to configure $\mathscr{CAF}$ is such a way that a_1 is skeptically accepted in at least one completion. For instance, Figure 3a describes such a configured CAF, with a successful completion given at Figure 3b.



(a) $\mathscr{CAF}$ configured by $A_{conf} = \{a_7\}$ (b) A successful completion of the CAF

Figure 3. A configured CAF and a successful completion

3.2. Computational Complexity of Possible Controllability

Now we focus on the computational complexity of deciding whether a CAF is possibly controllable. Formally, for $x \in \{sk, cr\}$ standing respectively for "skeptically" and "credulously", and $\sigma \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$, we study the decision problem:

$\mathsf{Control}_{\sigma,p,x}^{\mathscr{CAF},T}$ Is the CAF $\mathscr{CAF}$ possibly x-controllable w.r.t. σ and T?

Proposition 1. *The complexity of* $\mathsf{Control}_{\sigma,w,x}^{\mathscr{CAF},T}$, *for* $x \in \{sk, cr\}$ *and* $\sigma \in \{\text{co}, \text{pr}, \text{gr}, \text{st}\}$, *is given at Table 1.*

σ	sk	cr
st	Σ_2^P-complete	NP-complete
co	NP-complete	NP-complete
gr	NP-complete	NP-complete
pr	Σ_3^P-complete	NP-complete

Table 1. The complexity of $\mathsf{Control}_{\sigma,p,x}^{\mathscr{CAF},T}$, for $x \in \{sk, cr\}$

Detailled proofs are omitted for space reasons. However, we can explain lower bounds from existing results. In [22], the decision problems σ-PSA (possible skeptical acceptance) and σ-PCA (possible credulous acceptance) for Incomplete Argumentation Frameworks (IAFs) have been studied. A IAF corresponds to a CAF where $\rightleftarrows = \emptyset$ and $A_C = \emptyset$ (and obviously, $\Rrightarrow$ is empty too). Thus, an argument a is skeptically (resp. credulously) accepted in some completion of the IAF iff the corresponding CAF is skeptically (resp. credulously) controllable w.r.t. the target $\{a\}$. This means that if σ-PSA (resp. σ-PCA) is C-hard (for some class C of the polynomial hierarchy), then $\mathsf{Control}_{st,p,sk}^{\mathscr{CAF},T}$ (resp. $\mathsf{Control}_{st,p,cr}^{\mathscr{CAF},T}$) is C-hard as well.

For upper bounds, we obtain some of them from the known complexity of skeptical or credulous acceptance in Dung's AFs [23]. Indeed, a completion that skeptically (resp. credulously) accepts the target is a witness that the CAF is possibly skeptically (resp. credulously) controllable w.r.t. the target. This leads to the upper bounds for possible skeptical controllability, as well as the possible credulous controllability under grounded semantics. The possible credulous controllability for the other semantics can be reduced to SAT, so they belong to NP (the method is given in details for stable semantics in the next part of the paper).

Let us also briefly discuss the complexity of possible controllability for simplified CAFs, defined by [14] as CAFs with no uncertainty (*i.e.* $A_U = \rightleftarrows = \dashrightarrow = \emptyset$). Such a CAF has only one completion for each control configuration, thus possible and necessary controllability are equivalent in this case, and complexity remains the same as in the general case, described at Table 1.

3.3. Possible Controllability Through QBFs

Inspired by [14], we propose a QBF-based method to compute possible controllability for the stable semantics. Let us first give the meaning of the propositional variables used in the encoding.

Given $\mathscr{AF} = \langle A, R \rangle$,

- $\forall x_i \in A$, acc_{x_i} represents the acceptance status of the argument x_i;
- $\forall x_i, x_j \in A$, att_{x_i,x_j} represents the attack from x_i to x_j.

Φ_{st} is the formula $\Phi_{st} = \bigwedge_{x_i \in A}[acc_{x_i} \Leftrightarrow \bigwedge_{x_j \in A}(att_{x_j,x_i} \Rightarrow \neg acc_{x_j})]$. This modified version of the encoding from [24] describes in a generic way the relation between the structure of an AF (*i.e.* the set of attacks) and the arguments' acceptance (*i.e.* the extensions) w.r.t. stable semantics.

When the *att*-variables are assigned the truth value corresponding to the attack relation of $\mathscr{AF}$ (*i.e.* att_{x_i,x_j} is assigned 1 iff $(x_i,x_j) \in R$), the models of Φ_{st} (projected on the *acc*-variables) correspond in a bijective way to $\mathrm{st}(\mathscr{AF})$.

Given $\mathscr{AF} = \langle A,R\rangle$, we define the formula $\Phi_{st}^R = \Phi_{st} \wedge (\bigwedge_{(x_i,x_j)\in R} att_{x_i,x_j}) \wedge (\bigwedge_{(x_i,x_j)\notin R} \neg att_{x_i,x_j})$ which represents this assignment of *att*-variables corresponding to a specific AF. For any model ω of Φ_{st}^R, the set $\{x_i \mid \omega(acc_{x_i}) = 1\}$ is a stable extension of $\mathscr{AF}$. In the other direction, for any stable extension $\varepsilon \in \mathrm{st}(\mathscr{AF})$, ω s.t. $\omega(acc_{x_i}) = 1$ iff $x_i \in \varepsilon$ is a model of Φ_{st}^R.

These variables and formula are enough to encode the stable semantics of AFs. But to determine the controllability of a CAF, we need also to consider propositional variables to indicate which arguments are actually in the system:

- $\forall x_i \in A_C \cup A_U$, on_{x_i} is true iff x_i actually appears in the framework.

Now, we can recall the encoding which relates the attack relation and the arguments statuses in $\mathscr{CAF} = \langle F,C,U\rangle$ [14]:
Notation: $\mathbf{A} = A_F \cup A_C \cup A_U$, $\mathbf{R} = \rightarrow \cup \rightleftarrows \cup \dashrightarrow \cup \Rrightarrow$

$$\Phi_{st}(\mathscr{CAF}) = \bigwedge_{x_i\in A_F}[acc_{x_i} \Leftrightarrow \bigwedge_{x_j\in\mathbf{A}}(att_{x_j,x_i} \Rightarrow \neg acc_{x_j})]\wedge$$
$$\bigwedge_{x_i\in A_C\cup A_U}[acc_{x_i} \Leftrightarrow (on_{x_i} \wedge \bigwedge_{x_j\in\mathbf{A}}(att_{x_j,x_i} \Rightarrow \neg acc_{x_j}))]\wedge$$
$$\bigwedge_{(x_i,x_j)\in\rightarrow\cup\Rrightarrow} att_{x_i,x_j}) \wedge (\bigwedge_{(x_i,x_j)\in\rightleftarrows} att_{x_i,x_j} \vee att_{x_j,x_i}) \wedge (\bigwedge_{(x_i,x_j)\notin\mathbf{R}} \neg att_{x_i,x_j})$$

The first line states that an argument from A_F is accepted when all its attackers are rejected (similarly to the case of classical AFs). Then, the next line concerns arguments from A_C and A_U; since these arguments may not appear in some completions of the CAF, we add the condition that on_{x_i} is true to allow x_i to be accepted. The last line specify the case in which there is an attack in the completion: attacks from $\rightarrow$ and $\Rrightarrow$ are mandatory, and their direction is known; attacks from $\rightleftarrows$ are mandatory, but the actual direction is not known. We do not give any constraint about $\dashrightarrow$, which is equivalent to the tautological constraint $att_{x_i,x_j} \vee \neg att_{x_i,x_j}$: the attack may appear or not. Finally, we know that attacks which are not in $\mathbf{R}$ do not exist.

Given a set of arguments T, the fact that T must be included in all the stable extensions is represented by:

$$\Phi_{st}^{sk}(\mathscr{CAF},T) = \Phi_{st}(\mathscr{CAF}) \Rightarrow \bigwedge_{x_i\in T} acc_{x_i}$$

Given a set of arguments T, the fact that T must be included in at least one stable extension is represented by:

$$\Phi_{st}^{cr}(\mathscr{CAF},T) = \Phi_{st}(\mathscr{CAF}) \wedge \bigwedge_{x_i\in T} acc_{x_i}$$

Now we give the logical encodings for possible controllability for $\sigma = \mathrm{st}$.

Proposition 2. *Given $\mathscr{CAF}$ and $T \subseteq A_F$, $\mathscr{CAF}$ is possibly skeptically controllable w.r.t. T and the stable semantics iff*

$$\begin{array}{l}\exists\{on_{x_i} \mid x_i \in A_C\}\exists\{on_{x_i} \mid x_i \in A_U\}\\ \exists\{att_{x_i,x_j} \mid (x_i,x_j) \in \dashrightarrow \cup \rightleftarrows\}\forall\{acc_{x_i} \mid x_i \in \mathbf{A}\}\\ [\Phi^{sk}_{st}(\mathscr{CAF},T) \vee (\bigvee_{(x_i,x_j)\in\rightleftarrows}(\neg att_{a_i,a_j} \wedge \neg att_{a_j,a_i}))]\end{array} \quad (1)$$

is valid. In this case, each valid truth assignment of the variables $\{on_{x_i} \mid x_i \in A_C\}$ *corresponds to a configuration which reaches the target.*

This encoding is not a direct adaptation of the encoding proposed in [14]. We have to explicitly exclude the joint assignment of the variables att_{x_i,x_j} and att_{x_j,x_i} to false, when $(x_i,x_j) \in \rightleftarrows$, which would be in contradiction with the definition of this conflict relation. Another method is used in [14] to rule out these assignments, but it does not yield a QBF in prenex form. But this is the method that was proposed in [21], when necessary controllability has been applied to automated negotiation.

The following result holds for possible credulous controllability:

Proposition 3. *Given* $\mathscr{CAF}$ *and* $T \subseteq A_F$, $\mathscr{CAF}$ *is possible credulously controllable w.r.t. T and the stable semantics iff*

$$\begin{array}{l}\exists\{on_{x_i} \mid x_i \in A_C\}\exists\{on_{x_i} \mid x_i \in A_U\}\\ \exists\{att_{x_i,x_j} \mid (x_i,x_j) \in \dashrightarrow \cup \rightleftarrows\}\exists\{acc_{x_i} \mid x_i \in \mathbf{A}\}\\ [\Phi^{cr}_{st}(\mathscr{CAF},T) \vee (\bigvee_{(x_i,x_j)\in\rightleftarrows}(\neg att_{a_i,a_j} \wedge \neg att_{a_j,a_i}))]\end{array} \quad (2)$$

is valid. In this case, each valid truth assignment of the variables $\{on_{x_i} \mid x_i \in A_C\}$ *corresponds to a configuration which reaches the target.*

We notice that in the case of possible credulous controllability, the problem reduces to SAT since all the quantifiers are existential. This corresponds to the NP upper bound for possible credulous controllability under stable semantics (Proposition 1). We keep the QBF-style notation for homogeneity with Equation 1.

Example 4 (Continuation of Example 1)**.** *Let us describe the logical encoding for possible controllability with* $\mathscr{CAF}$ *as described previously and* $T = \{a_1\}$. *We give here the example for possible skeptical controllability:*

$$\begin{array}{l}\exists on_{a_7}, on_{a_8}, on_{a_9}, \exists on_{a_6}, \exists att_{a_5,a_1}, att_{a_6,a_5}, att_{a4,a_6},\\ \forall acc_{a_1}, acc_{a_2}, \ldots, acc_{a_9},\\ [\Phi^{sk}_{st}(\mathscr{CAF},T) \vee (\bigvee_{(x_i,x_j)\in\rightleftarrows}(\neg att_{a_i,a_j} \wedge \neg att_{a_j,a_i}))]\end{array}$$

Below, we give the formula $\Phi^{sk}_{st}(\mathscr{CAF},T)$. *For a matter of readability, several simplifications are made. For instance, an implication like* $att_{x_j,x_i} \Rightarrow \neg acc_{x_j}$ *can be removed when* att_{x_j,x_i} *is known to be false (because* x_j *does not attack* x_i*), and can be replaced by* $\neg acc_{x_j}$ *when* att_{x_j,x_i} *is known to be true. Only the uncertain attacks need to be kept explicit in the encoding. The first three lines give the condition for the acceptance of the fixed arguments. Then, two lines give the condition for the acceptance of the control and uncertain arguments. The other lines describe the structure of the graph (*i.e. *the attack relations), and the implication gives the target for skeptical acceptance.*

$$[[acc_{a_1} \Leftrightarrow \neg a_2 \wedge \neg a_3 \wedge (att_{a_5,a_1} \Rightarrow \neg acc_{a_5})]$$
$$\wedge$$
$$[acc_{a_2} \Leftrightarrow \neg acc_{a_4})] \wedge [acc_{a_3} \Leftrightarrow \neg acc_{a_4})]$$
$$\wedge$$
$$[acc_{a_4} \Leftrightarrow (att_{a_6,a_4} \Rightarrow \neg acc_{a_6})] \wedge [acc_{a_5} \Leftrightarrow \neg acc_{a_7})]$$
$$\wedge$$
$$[acc_{a_6} \Leftrightarrow (on_{a_6} \wedge \neg acc_{a_8} \wedge \neg acc_{a_9} \wedge (att_{a_4,a_6} \Rightarrow \neg acc_{a_4}))]$$
$$\wedge$$
$$[acc_{a_7} \Leftrightarrow (on_{a_7} \wedge \neg acc_{a_8})] \wedge [acc_{a_8} \Leftrightarrow on_{a_8}] \wedge [acc_{a_9} \Leftrightarrow (on_{a_9} \wedge \neg acc_{a_7}))]$$
$$\wedge$$
$$att_{a_2,a_1} \wedge att_{a_3,a_1} \wedge att_{a_4,a_2} \wedge att_{a_4,a_3} \wedge att_{a_7,a_5} \wedge att_{a_7,a_9} \wedge att_{a_8,a_6} \wedge att_{a_8,a_7} \wedge att_{a_9,a_6}$$
$$att_{a_4,a_6} \vee att_{a_6,a_4}$$
$$\bigwedge_{(x_i,x_j) \notin \mathbf{R}} \neg att_{x_i,x_j}$$
$$] \Rightarrow acc_{a_1}$$

4. Related Work

Qualitative uncertainty has been considered in other frameworks. Partial AFs [20] are special instances of CAFs where only $\dashrightarrow$ is considered. They are used as a tool in a process of aggregating several AFs. Then [25] studies the complexity of verifying in a PAF whether a set of arguments is an extension of some (or every) completion. [26] conducts a similar study for argument-incomplete AFs, *i.e.* there is some uncertainty about the presence of arguments (the part called A_U)in our framework). Finally, [27] combines both. Let us notice than in [25,26,27], both versions of the verification problem (existential and universal w.r.t. the set of completions) are studied. As mentioned previously, [22] gives the complexity of skeptical and credulous acceptance for IAFs. While being a quite general model of uncertainty, this Incomplete AF is strictly included in the CAF setting: [26] does not allow to express the uncertainty about the direction of a conflict (*i.e.* our $\rightleftarrows$ relation cannot be encoded in this framework). Moreover, none of these works [20,25,26,27] is concerned with argumentation dynamics.

Quantitative models of uncertainty have also been used; while being an interesting approach, they require more input information than qualitative models like ours. This approach is out of the scope of this paper and is kept for future work. In particular, probabilistic CAFs based on the constellations approach [28] are a promising research tracks.

Argumentation dynamics has received a lot of attention in the last ten years. Except the initial paper about CAFs [14], most of the existing work consider complete information about the input (*i.e.* no uncertainty of the initial AF is considered). As far as we know, the only proposal which can encompass uncertainty is the update of AFs through the YALLA language [11]. However, YALLA pays the price of its expressiveness, and we are not aware of any efficient computational approach for reasoning with it, contrary to our QBF-based approach for CAFs.

5. Conclusion

In this paper, we push forward the study of the Control Argumentation Frameworks. We define a ”weaker” version of controllability, where a target set of arguments needs to be accepted in at least one completion (instead of every completion). This kind of reasoning is related to a lawyer’s plea: at the end of a trial, the lawyer needs to pick arguments (in our setting, the configuration A_{conf}) such that the target (“the defendant is innocent”) is accepted in at least one completion. Somehow, possible controllability is to necessary controllability what credulous acceptance is to skeptical acceptance.

Many research tracks are still open. We plan to propose logical encodings and to study the complexity of controllability for other extension-based semantics. Also, other methods can be used for computing control configuration, especially SAT-based counter-example guided abstract refinement (CEGAR), that was successfully used for reasoning problems at the second level of the polynomial hierarchy [12]. An interesting other form of controllability to be studied is ”optimal” controllability, *i.e.* finding a configuration that allows to reach the target in as many completions as possible. This is useful in situations where a CAF is not necessary controllable, and possible controllability seems too weak. Techniques like CEGAR or QBF with soft variables [29] may be helpful for solving this problem. Also, as mentioned previously, we will study quantitative models of uncertainty in the context of CAFs. In particular, it would be interesting for real world applications to define a form of controllability w.r.t. the most probable completion, or w.r.t. the set of completions with a probability higher than a given threshold. Finally, we think that an important work to be done, in order to apply CAFs to real applications scenarios, is to determine how CAFs and controllability can be defined when the internal structure of arguments (*e.g.* based on logical formulas or rules) is known.

References

[1] Dung PM. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Art Intel. 1995;77:321–357.

[2] Baroni P, Caminada M, Giacomin M. Abstract Argumentation Frameworks and Their Semantics. In: Baroni P, Gabbay D, Giacomin M, van der Torre L, editors. Handbook of Formal Argumentation. College Publications; 2018. p. 159–236.

[3] Caminada M. On the Issue of Reinstatement in Argumentation. In: Proc. of JELIA’06; 2006. p. 111–123.

[4] Amgoud L, Ben-Naim J. Ranking-Based Semantics for Argumentation Frameworks. In: Proc. of SUM’13; 2013. p. 134–147.

[5] Boella G, Kaci S, van der Torre LWN. Dynamics in Argumentation with Single Extensions: Abstraction Principles and the Grounded Extension. In: Proc. of ECSQARU’09; 2009. p. 107–118.

[6] Cayrol C, de Saint-Cyr FD, Lagasquie-Schiex M. Change in Abstract Argumentation Frameworks: Adding an Argument. J Artif Intell Res (JAIR). 2010;38:49–84.

[7] Baumann R, Brewka G. Expanding Argumentation Frameworks: Enforcing and Monotonicity Results. In: Proc. of COMMA’10; 2010. p. 75–86.

[8] Coste-Marquis S, Konieczny S, Mailly J, Marquis P. On the Revision of Argumentation Systems: Minimal Change of Arguments Statuses. In: Proc. of KR’14; 2014. .

[9] Doutre S, Herzig A, Perrussel L. A Dynamic Logic Framework for Abstract Argumentation. In: Proc. of KR’14; 2014. .

[10] Coste-Marquis S, Konieczny S, Mailly J, Marquis P. Extension Enforcement in Abstract Argumentation as an Optimization Problem. In: Proc. of IJCAI’15; 2015. p. 2876–2882.

[11] de Saint-Cyr FD, Bisquert P, Cayrol C, Lagasquie-Schiex M. Argumentation update in YALLA (Yet Another Logic Language for Argumentation). Int J Approx Reasoning. 2016;75:57–92.

[12] Wallner JP, Niskanen A, Järvisalo M. Complexity Results and Algorithms for Extension Enforcement in Abstract Argumentation. J Artif Intell Res. 2017;60:1–40.

[13] Doutre S, Mailly J. Semantic Change and Extension Enforcement in Abstract Argumentation. In: Proc. of SUM'17; 2017. p. 194–207.

[14] Dimopoulos Y, Mailly JG, Moraitis P. Control Argumentation Frameworks. In: Proc. of AAAI'18; 2018. p. 4678–4685.

[15] Baumann R. What Does it Take to Enforce an Argument? Minimal Change in abstract Argumentation. In: Proc. of ECAI'12; 2012. p. 127–132.

[16] Biere A, Heule M, van Maaren H, Walsh T, editors. Handbook of Satisfiability. vol. 185 of Frontiers in Artificial Intelligence and Applications. IOS Press; 2009.

[17] Kleine Büning H, Bubeck U. Theory of Quantified Boolean Formulas. In: Handbook of Satisfiability; 2009. p. 735–760.

[18] Arora S, Barak B. Computational Complexity - A Modern Approach. Cambridge University Press; 2009.

[19] Amgoud L, Vesic S. Rich preference-based argumentation frameworks. Int J Approx Reasoning. 2014;55(2):585–606.

[20] Coste-Marquis S, Devred C, Konieczny S, Lagasquie-Schiex M, Marquis P. On the merging of Dung's argumentation systems. Artif Intell. 2007;171(10-15):730–753.

[21] Dimopoulos Y, Mailly J, Moraitis P. Argumentation-based Negotiation with Incomplete Opponent Profiles. In: Proc. of AAMAS'19; 2019. p. 1252–1260.

[22] Baumeister D, Neugebauer D, Rothe J. Credulous and Skeptical Acceptance in Incomplete Argumentation Frameworks. In: Proc. of COMMA'18; 2018. p. 181–192.

[23] Dvorák W, Dunne PE. Computational Problems in Formal Argumentation and their Complexity. In: Baroni P, Gabbay D, Giacomin M, van der Torre L, editors. Handbook of Formal Argumentation. College Publications; 2018. p. 631–688.

[24] Besnard P, Doutre S. Checking the acceptability of a set of arguments. In: Proc. of NMR'04; 2004. p. 59–64.

[25] Baumeister D, Neugebauer D, Rothe J. Verification in Attack-Incomplete Argumentation Frameworks. In: Proc. of ADT'15; 2015. p. 341–358.

[26] Baumeister D, Rothe J, Schadrack H. Verification in Argument-Incomplete Argumentation Frameworks. In: Proc. of ADT'15; 2015. p. 359–376.

[27] Baumeister D, Neugebauer D, Rothe J, Schadrack H. Verification in incomplete argumentation frameworks. Artif Intell. 2018;264:1–26.

[28] Hunter A. A probabilistic approach to modelling uncertain logical arguments. Int J Approx Reasoning. 2013;54(1):47–81.

[29] Reimer S, Sauer M, Marin P, Becker B. QBF with Soft Variables. ECEASST. 2014;70.