An axiomatic approach for persuasion dialogs

Leila Amgoud, Florence Dupin de Saint-Cyr

To link to this article : DOI :10.1109/ICTAI.2013.97
URL : http://dx.doi.org/10.1109/ICTAI.2013.97

To cite this version : Amgoud, Leila and Dupin De Saint Cyr - Bannay, Florence *An axiomatic approach for persuasion dialogs*. (2013) In: IEEE International Conference on Tools with Artificial Intelligence - ICTAI 2013, 4 November 2013 - 6 November 2013 (Washington, United States).

An axiomatic approach for persuasion dialogs

Leila Amgoud
IRIT - CNRS
118, route de Narbonne,
31062 Toulouse Cedex 9 France
amgoud@irit.fr

Florence Dupin de Saint-Cyr
IRIT - Université de Toulouse
118, route de Narbonne,
31062 Toulouse Cedex 9 France
bannay@irit.fr

***Abstract*—Several systems were developed for supporting *public persuasion dialogs* where two agents with conflicting opinions try to convince an audience. For computing the outcomes of dialogs, these systems use (abstract or structured) argumentation systems that were initially developed for non-monotonic reasoning.**

Despite the increasing number of such systems, there are almost no work on high level properties they should satisfy. This paper is a first attempt for defining *postulates* that guide the well-definition of dialog systems and that allow their comparison. We propose six basic postulates (including e.g. the finiteness of generated dialogs). We then show that this set of postulates is incompatible with those proposed for argumentation systems devoted for nonmonotonic reasoning. This incompatibility confirms the differences between persuading and reasoning. It also suggests that reasoning systems are not suitable for computing the outcomes of dialogs.

***Keywords*-Argumentation; Dialog; Postulates;**

I. INTRODUCTION

Argumentation theory has become a hot topic in Artificial Intelligence. It is studied for modeling agent's internal reasoning, namely for handling inconsistent/incomplete and uncertain information (e.g., [6], [13]) and for making decisions (e.g., [4], [12]). There is also an extensive literature devoted to modeling agent's interactions, namely dialogs in which agents may exchange arguments with each other, like persuasion (e.g., [3], [20]), negotiation (e.g., [15], [18]) and deliberation (e.g., [17]). In all these disparate applications, an argumentation theory [10] consists of a set of arguments justifying claims, attacks among those arguments and a semantics. This latter is a set of criteria describing which arguments are acceptable together.

Persuasion is a type of dialog in which two agents having conflicting opinions about an issue try to convince each other either in *public*, i.e., in presence of an audience (e.g. [7], [8], [16]) or in *private*, i.e., in absence of any audience (e.g. [3], [20], [24]). In both cases persuasion is done by exchanging arguments. An important feature of private persuasion is the evolution of the argumentation systems of both agents by adding the arguments received from the other party. An agent is persuaded if the subject of the dialog is supported by its system at a given step of the dialog. Thus, this kind of dialogs is more concerned with the dynamics of the argumentation systems of the agents.

In this paper, we focus on public persuasion in which agents try more to convince an audience rather than the other party. In the debate between Holland and Sarkozy before the presidential election, both candidates tried to convince the voters. Systems that support this type of dialogs have three main components: i) a protocol which is a set of rules that define coherent dialogs, ii) a set of reasoning systems of agents involved in dialog and iii) a system for computing the outcomes of dialogs. In existing literature, the reasoning system of an agent is either a Dung style abstract argumentation system [10] or one of its logic-based instantiations (e.g., [9], [2]). The system that is used for computing the outcomes of dialogs is exactly of the same nature as those of the agents. However, its arguments come from the exchanges made by the agents during the dialog and the attacks are the conflicts among them.

Despite the increasing number of works on modeling public persuasion, there are almost no work on high level requirements expected from dialog systems. Consequently, apart from the termination of their dialogs, it is not clear what other properties they satisfy. This makes their proper evaluation difficult if not impossible. This paper provides a first attempt for defining *postulates* that guide the well-definition of dialog systems. We focus on systems that use Dung style argumentation systems both for modeling the reasoning of agents and for computing the outcomes of dialogs. We propose six basic postulates that any such dialog system should satisfy. Some of them (non-triviality, natural-attacks-allowanceand dissimulation) ensure that a dialog system captures natural language dialogs. The three other postulates (finiteness, consistency and non-determinism) are more about the quality of dialogs. Since one cannot speak about a dialog without referring to its subject thus to its content, then in what follows we consider dialog systems whose various argumentation systems are logic-based. Our postulates hold for any instantiation of Dung's framework. However, for illustration purposes, we have chosen those based on deductive logics [2]. The second main contribution of the paper consists of comparing the six postulates with those proposed for argumentation systems devoted for nonmonotonic reasoning, i.e., the systems used by the agents. We show that the two sets are incompatible. This incompatibility confirms the differences between persuading and reasoning. It also suggests that reasoning systems are not suitable for computing the outcomes of dialogs.

The paper is organized as follows: Section II recalls both the argumentation system proposed in [2] for reasoning about inconsistent information and the set of postulates it should enjoy. Section III defines a public persuasion system and Section IV proposes a set of postulates the system should satisfy. Section V compares the two sets of postulates. The last section is devoted to some concluding remarks.

II. ARGUMENTATION FOR REASONING

This section recalls the argumentation system proposed in [2] for reasoning about inconsistent information. It is a logic-based instantiation of Dung's framework [10]. Note that the same kind of results could also be obtained for rule-based systems like ASPIC [9].

A. Basic definitions

In [2], the argumentation system is grounded on Tarski's logics [23]: i.e., pairs $(\mathcal{L}, \texttt{CN})$ where $\mathcal{L}$ is a set of well-formed *formulas* and CN is a *consequence operator* that satisfies the following basic properties: For $X \subseteq \mathcal{L}$,

- *Expansion:* $X \subseteq \texttt{CN}(X)$
- *Idempotence:* $\texttt{CN}(\texttt{CN}(X)) = \texttt{CN}(X)$
- *Absurdity:* $\texttt{CN}(\{x\}) = \mathcal{L}$ for some $x \in \mathcal{L}$

The notion of *consistency* is defined as follows:

A set $X \subseteq \mathcal{L}$ is *consistent* wrt a logic $(\mathcal{L}, \texttt{CN})$ iff $\texttt{CN}(X) \neq \mathcal{L}$. It is *inconsistent* otherwise.

Arguments are built from a *knowledge base* $\Sigma \subseteq \mathcal{L}$.

Definition 1 (Argument)**:** *An* argument *built from a knowledge base Σ is a pair (X, x) s.t.*

- $X \subseteq \Sigma$
- *X is consistent*
- $x \in \texttt{CN}(X)$
- *$\nexists X' \subset X$ such that $x \in \texttt{CN}(X')$*

An argument (X, x) is atomic *iff $X = \{y\}$ and $\texttt{CN}(\{x\}) = \texttt{CN}(\{y\})$. It is a* sub-argument *of (X', x') iff $X \subseteq X'$.*

Example 1: *Let $\Sigma = \{\neg wh, \neg wh \to fe, vs \to \neg fe, vs\}$ be a knowledge base representing the following information: Mary does not work hard ($\neg wh$), if somebody does not work hard then he will fail his exams (fe), if somebody is very smart (vs) then he will not fail his exams, Mary is very smart. The following arguments may be built from Σ: $a_0 = (\emptyset, fe \vee \neg fe)$, $a_1 = (\{\neg wh\}, \neg wh)$, $a_2 = (\{\neg wh, \neg wh \to fe\}, fe)$, $a_3 = (\{vs, vs \to \neg fe\}, \neg fe)$.*

The following proposition shows that it is possible to build an atomic argument from any formula that is neither a tautology nor a contradiction.

Property 1: *Let Σ be a knowledge base. For all $x \in \Sigma$, if $x \notin \texttt{CN}(\emptyset)$ and $\texttt{CN}(\{x\}) \neq \mathcal{L}$ then $(\{x\}, x)$ is an (atomic) argument.*

Proof: For all $x \in \Sigma$, it holds that $\{x\} \subseteq \Sigma$, moreover if $\texttt{CN}(\{x\}) \neq \mathcal{L}$ then $\{x\}$ is consistent. Due to expansion, $x \in \texttt{CN}(\{x\})$. Since $\emptyset$ is the only strict subset of $\{x\}$, if $x \notin \texttt{CN}(\emptyset)$ then $\nexists X \subset \{x\}$ such that $x \in \texttt{CN}(X)$. ■

Note that for most classical logics (instances of Tarski's ones, e.g. propositional logic, first order logic ...), the set of all arguments that may be built from a (finite) knowledge base is *infinite*.

Notation 1: Supp *and* Conc *are two functions that return respectively the* support X *and the* conclusion x *of an argument* (X, x). Sub *is a function that returns all the sub-arguments of a given argument.*

An argumentation system is defined as follows.

Definition 2 (Argumentation system)**:** *An* argumentation system *(AS) over a knowledge base Σ is a pair $\mathcal{T} = (\mathcal{A}, \mathcal{R})$ such that $\mathcal{A}$ is a set of arguments built from Σ using Definition 1, $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ is an* attack *relation. For $a, b \in \mathcal{A}$, $(a, b) \in \mathcal{R}$ (or $a\mathcal{R}b$) means that a attacks b.*

In the next sections, we will show that, for a reasoning system and for a dialog system, the set $\mathcal{A}$ of arguments should not be chosen in an arbitrary way. There are some constraints that should be fulfilled. The attack relation $\mathcal{R}$ is left *unspecified* in the sequel since our analysis is independent from its exact definition. Finally, arguments are evaluated using any Dung's semantics [10]. The following definition recalls some of them.

Definition 3 (Semantics)**:** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R})$ be an AS over a base Σ and $\mathcal{E} \subseteq \mathcal{A}$ s.t. $\nexists a, b \in \mathcal{E}$ s.t. $a\mathcal{R}b$.*

- *$\mathcal{E}$ is an* admissible *set iff $\mathcal{E}$ defends all its elements (i.e., it attacks any attacker of its arguments).*
- *$\mathcal{E}$ is a* preferred *extension iff $\mathcal{E}$ is a maximal (for set inclusion) admissible set.*
- *$\mathcal{E}$ is a* stable *extension iff $\forall a \in \mathcal{A} \setminus \mathcal{E}$, $\exists b \in \mathcal{E}$ s.t. $b\mathcal{R}a$.*

$\texttt{Ext}(\mathcal{T})$ *returns the set of extensions of the AS $\mathcal{T}$ under a given semantics.*

The extensions are used in order to define the plausible conclusions to be drawn from Σ. The idea is to infer a formula x from Σ iff x is the conclusion of an argument in each extension. $\texttt{Output}(\mathcal{T})$ is the set of all such formulas.

Definition 4 (Output)**:** *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R})$ be an AS over a knowledge base Σ. $\texttt{Output}(\mathcal{T}) = \{x \in \mathcal{L} \mid \forall \mathcal{E} \in \texttt{Ext}(\mathcal{T}), \exists a \in \mathcal{E}$ s.t. $\texttt{Conc}(a) = x\}$.*

B. Basic postulates for reasoning systems

A set of desirable properties that the previous systems should satisfy was proposed in [1]. The postulates are compatible (i.e., can be satisfied all together). The first one ensures that each extension supports consistent conclusions. The second postulate concerns the closure of its output under the consequence operator CN. The third postulate concerns *sub-arguments*. It ensures that the acceptance of an argument should imply also the acceptance of all its sub-parts.

Consistency: *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R})$ be an AS over a base Σ. For all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\{\texttt{Conc}(a) \mid a \in \mathcal{E}\}$ is consistent.*

Closure under $\texttt{CN}$: *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R})$ be an AS over a base Σ. For all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, $\{\texttt{Conc}(a) \mid a \in \mathcal{E}\} = \texttt{CN}(\{\texttt{Conc}(a) \mid a \in \mathcal{E}\})$.*

Closure under sub-arguments: *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R})$ be an AS over a base Σ. For all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, if $a \in \mathcal{E}$, then $\texttt{Sub}(a) \subseteq \mathcal{E}$.*

The following simple property is useful for our discussion in next sections. It shows that an argumentation system which is closed under sub-arguments and which admits non empty extensions contains a set of atomic arguments. Moreover, if the system is closed under $\texttt{CN}$ then it should contain all arguments supporting tautologies.

Proposition 1: *Let $\mathcal{T} = (\mathcal{A}, \mathcal{R})$ be an AS. For all $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$, if $\mathcal{E} \neq \emptyset$ then:*

- *If $\mathcal{T}$ is* closed under sub-arguments, *then $\forall x \in \cup_{a \in \mathcal{E}} \texttt{Supp}(a)$, if $x \notin \texttt{CN}(\emptyset)$, then $(\{x\}, x) \in \mathcal{E}$.*
- *If $\mathcal{T}$ is* closed under $\texttt{CN}$, *then $\forall x \in \texttt{CN}(\emptyset)$ $(\emptyset, x) \in \mathcal{E}$.*

Proof: Let $\mathcal{T} = (\mathcal{A}, \mathcal{R})$ be an AS and $\texttt{Ext}(\mathcal{T})$ its set of extensions under a given semantics. Assume that $\mathcal{E} \in \texttt{Ext}(\mathcal{T})$ and $\mathcal{E} \neq \emptyset$.

Let $a \in \mathcal{E}$. Assume that $\texttt{Supp}(a) \neq \emptyset$. For all $x \in \texttt{Supp}(a)$, $\{x\}$ is consistent (since $\texttt{Supp}(a)$ is consistent and due to Property 2 in [2]). If $x \notin \texttt{CN}(\emptyset)$, then $(\{x\}, x)$ is an argument. Moreover, it is a sub-argument of a. $\mathcal{T}$ being closed under sub-arguments, this means that $(\{x\}, x)$ is also an argument of $\mathcal{E}$.

By monotonicity of $\texttt{CN}$, since $\emptyset \subseteq \{\texttt{Conc}(a) | a \in \mathcal{E}\}$ then $\texttt{CN}(\emptyset) \subseteq \texttt{CN}(\{\texttt{Conc}(a) | a \in \mathcal{E}\})$. If $\mathcal{T}$ is closed under $\texttt{CN}$ then $\texttt{CN}(\emptyset) \subseteq \{\texttt{Conc}(a) | a \in \mathcal{E}\}$. This means that $\forall x \in \texttt{CN}(\emptyset)$, $\exists a \in \mathcal{E}$ such that $\texttt{Conc}(a) = x$. Due to Definition 1, $\texttt{Supp}(a) = \emptyset$. Hence, $(\emptyset, x) \in \mathcal{E}$. ■

III. Public persuasion dialog systems

This section defines an abstract public persuasion system that may be used, for instance, in online debate platforms. The system is abstract since it keeps one of its main components, protocol, unspecified. Moreover, the notion of audience is not explicitly represented.

A dialog system has three main components: a set of agents represented by their reasoning models, a protocol, and a rule for computing the outcome of any dialog that takes place between the agents. A protocol specifies the set of rules governing the well-definition of dialogs (e.g., who is allowed to say what and when?). In the sequel, we leave this component unspecified. Thus, our system can be instantiated by any protocol.

For the purpose of our paper and without loss of generality, we focus on persuasion dialogs between two agents P and C. Each of them is equipped with a knowledge base Σ_k (with $k \in \{P, C\}$) and an argumentation system (in the sense of Definition 2) $\mathcal{T}_k = (\mathcal{A}_k, \mathcal{R}_k)$. It is worth mentioning that the two agents may use two distinct attack relations (for instance, P may use the undercut relation [19] whereas C assumption attack [11] (see section IV)). They may also choose distinct semantics for the evaluation of arguments. However, they use the same underlying monotonic logic $(\mathcal{L}, \texttt{CN})$. Indeed, in order to be able to understand each other they should at least share the same language.

Before defining the third component of a dialog system, i.e., its rule for computing the outcomes of dialogs, let us first define what is a persuasion dialog. The notion of move is the backbone of a dialog. It consists of two agents (a speaker and a hearer) and a speech act together with a content. The speech act is taken from a set $\mathcal{S}$[1]. The only restriction on $\mathcal{S}$ is that it should contain at least two kinds of speech acts: "Argue" for exchanging arguments and "Assert" for making claims.

Definition 5 (Move)**:** *Let $\mathcal{S}$ be a set of speech acts symbols containing at least "Argue" and "Assert" symbols. A* move *m is a triple $\langle s, h, a \rangle$ s.t.*

- *$s \in \{P, C\}$ is the agent that utters m.*
- *$h \in \{P, C\}$ is the agent to whom the move is addressed.*
- *$a = act : content$ s.t. $act \in \mathcal{S}$ and $content \in \mathcal{L} \cup \mathcal{A}_P \cup \mathcal{A}_C$. If $act =$ Argue (respectively $act =$Assert) then $content \in \mathcal{A}_P \cup \mathcal{A}_C$ (resp. $content \in \mathcal{L}$). $\texttt{Act}$ and $\texttt{Content}$ are two functions s.t. $\texttt{Act}(m) = act$ and $\texttt{Content}(m) = content$.*

A persuasion dialog is a "valid" sequence of moves, i.e., a sequence that satisfies all the rules of the protocol. Since we do not focus on particular protocols, then we use the term 'valid' without defining it formally. Besides, the subject of a persuasion dialog is a claim made via an Assert move by one of the agents. Arguments are exchanged in order to increase or decrease its acceptability.

Definition 6 (Persuasion dialog)**:** *A* persuasion dialog *$\mathcal{D}$ generated by a dialog system DS[2] is a non-empty (finite or infinite) valid sequence of moves (m_i) s.t. $\texttt{Act}(m_1) =$ Assert. The* subject *of $\mathcal{D}$ is $\texttt{Subject}(\mathcal{D}) = \texttt{Content}(m_1)$.*

For computing the outcome of a persuasion dialog, an argumentation system in the sense of Definition 2 is used. Its arguments are those exchanged in the dialog in addition to the atomic arguments built from the assertions made in the dialog. The idea is to consider all the different kinds of claims (either in form of assertions or arguments) made by the agents. Defining the attack relation of this system is more tricky since the agents may use different relations. In what follows, we assume the existence of a third relation denoted $\mathcal{R}$ which results from a merging of the two relations $\mathcal{R}_P$ and $\mathcal{R}_C$ using an operator $\oplus$ not specified in this paper. Thus, $\mathcal{R} = \mathcal{R}_P \oplus \mathcal{R}_C$. An example of a merging operator is the *union* which considers all the attacks which hold either in $\mathcal{R}_P$ or in $\mathcal{R}_C$.

[1] In the literature the following set of basic speech acts is often used $\mathcal{S} = \{$Assert, Argue, Declare, Question, Request, Challenge, Promise$\}$.

[2] Throughout the paper we refer to a dialog system by DS without specifying its components.

Definition 7 (AS of a persuasion dialog)**:** *Let $\mathcal{D}$ be a persuasion dialog generated by a dialog system* DS*. The argumentation system associated with $\mathcal{D}$ is the pair $\mathsf{AS}_{\mathcal{D}}$ = $(\texttt{Args}(\mathcal{D}), \texttt{Confs}(\mathcal{D}))$ s.t.*

- $\texttt{Args}(\mathcal{D}) = \{\texttt{Content}(m) \mid m \in \mathcal{D}$ *and* $\texttt{Act}(m) =$ *Argue*$\} \cup \{(\{\texttt{Content}(m)\}, \texttt{Content}(m)) \mid m \in \mathcal{D}$ *and* $\texttt{Act}(m) = Assert\}$
- $\texttt{Confs}(\mathcal{D}) = \{(a, b) \mid a, b \in \texttt{Args}(\mathcal{D})$ *and* $(a, b) \in \mathcal{R}\}$

The outcome of a persuasion dialog $\mathcal{D}$ is the status of its subject wrt $\mathsf{AS}_{\mathcal{D}}$.

Definition 8 (Dialog output)**:** *Let $\mathcal{D}$ be a persuasion dialog generated by a dialog system* DS*.* $\texttt{Subject}(\mathcal{D})$ *is true iff* $\texttt{Subject}(\mathcal{D}) \in \texttt{Output}(\mathsf{AS}_{\mathcal{D}})$.

We can go further by checking wether the agent who asserted the subject wins or not the dialog.

Definition 9 (Dialog winner)**:** *Let $\mathcal{D}$ be a persuasion dialog generated by a dialog system* DS *with $m_1 = \langle s, h, a\rangle$. If* $\texttt{Subject}(\mathcal{D}) \in \texttt{Output}(\mathsf{AS}_{\mathcal{D}})$ *then s* wins the dialog *$\mathcal{D}$ and h* looses *it. Otherwise, h wins and s looses the dialog.*
Let Winner *be a function such that* $\texttt{Winner}(\mathcal{D})$ *returns the agent that wins the dialog $\mathcal{D}$.*

IV. Basic postulates for dialog systems

In the previous section, we have defined what a dialog system is. It generates non-empty persuasion dialogs. In what follows, we propose some key features, called also postulates, that should be satisfied by the system and the dialogs it generates.

The first postulate concerns the finiteness of the generated dialogs. This requirement is already known in the literature. In [14], protocols should ensure termination. Here, we require finiteness not only for the number of moves but also for the content of each move. For instance, it is not allowed to assert $x \wedge x \wedge \ldots$.

Finiteness: *Finiteness holds for a dialog system* DS *iff for all persuasion dialog $\mathcal{D}$ generated by* DS*, $size(\mathcal{D}) \in \mathbb{N}$ where $size(\mathcal{D}) = \sum_{m \in \mathcal{D}} sizemove(m)$ with $sizemove(m)$ is the number of occurrences of atoms and operators used in the content of m.*

As said before, a protocol guides the well-definition of dialogs and is common to all agents. However, the outcome of a dialog depends on the strategies of the agents. This is captured by the next postulate which constrains the dialog system to be able to generate at least one dialog in which a subject is accepted and one dialog in which it is not.

Non-determinism: *A dialog system* DS *is non-determinist iff for all formula $x \in \mathcal{L}$, s.t. $x \notin \texttt{CN}(\emptyset)$ and $\texttt{CN}(\{x\}) \neq \mathcal{L}$, there exist at least two dialogs $\mathcal{D}_1$ and $\mathcal{D}_2$ generated by* DS*, such that* $\texttt{Subject}(\mathcal{D}_1) = \texttt{Subject}(\mathcal{D}_2) = x$ *and* $\texttt{Output}(\mathcal{D}_1) \neq \texttt{Output}(\mathcal{D}_2)$.

Note that if the set of non-trivial formulas of $\mathcal{L}$ (i.e., without considering tautologies and contradictions) is infinite then any dialog system satisfying non-determinism can generate an infinite number of dialogs.

The third important postulate concerns the formalism that is used for computing the outcomes of dialogs. In our context, Dung's system should ensure sound results. Namely, extensions (under any semantics) represent various positions in a dialog. Thus, each of them should be coherent. This leads to a consistency postulate similar to the one presented for reasoning systems in [1].

Consistency: *A dialog system* DS *ensures consistency iff for all persuasion dialog $\mathcal{D}$ generated by* DS*, for all $\mathcal{E} \in \texttt{Ext}(\mathsf{AS}_{\mathcal{D}})$, $\{\texttt{Conc}(a) \mid a \in \mathcal{E}\}$ is consistent.*

The aim behind building systems for persuasion dialogs is to automate such dialogs and to conduct efficient ones. However, these systems should capture as much as possible natural dialogs. Works by linguists [21], [22] have emphasized the main forms of counter-argumentation that may take place in everyday life dialogs. The first one, known as "rebuttal" in [11], consists of undermining the conclusion of another argument. The second form, known as "assumption attack" in [11], consists of undermining a premise in the support of another argument.

- An argument a *rebuts* an argument b iff the set $\{\texttt{Conc}(a), \texttt{Conc}(b)\}$ is inconsistent.
- An argument a *assumption-attacks* an argument b iff $\exists x \in \texttt{Supp}(b)$ s.t. the set $\{\texttt{Conc}(a), x\}$ is inconsistent.

It is thus important for a dialog system to capture these two forms of attacks. The following postulate ensures this by constraining the attack relation $\mathcal{R}$.

Natural-attacks-allowance: *A dialog system* DS *allows for natural attacks iff for all persuasion dialog $\mathcal{D}$ generated by* DS*, for all $a, b \in \texttt{Args}(\mathcal{D})$,*

- *if a rebuts b then $(a, b) \in \mathcal{R}$, and*
- *if a assumption-attacks b then $(a, b) \in \mathcal{R}$.*

It is well-known that in public persuasion dialogs, agents try to convince others about a given claim even if they think that this latter does not hold. They then hide arguments and information in order to reach their objectives. A dialog system should thus allow dissimulation of information. More formally a dialog system allows dissimulation if it can generate some dialogs in which the winner would have change if one of the agents had uttered (had not concealed) some argument.

Dissimulation: *A dialog system* DS *allows dissimulation iff there exists a persuasion dialog $\mathcal{D}$ between two agents (say k and l) generated by* DS *such that $\exists a \in \mathcal{A}_k$, such that* $\texttt{Winner}(D) \neq \texttt{Winner}(D; \langle k, l, a\rangle)$

The last postulate is about the efficiency of persuasion dialogs. Recall that in such dialogs, agents try to convince other parties to accept some assertion by putting forward arguments. These latter are intended to justify the assertion by *new evidences*. Thus an argument in which an assertion is justified by the assertion itself fails to meet the objective of arguing. Assume a politician who tries to convince a population that taxes should be increased. Nobody will accept an argument of the form: "taxes should be increased because they should be increased". This does not mean that nobody will accept the idea of increasing taxes especially people who have good reasons in favor of tax increase. Thus, for a persuasion to be efficient, atomic arguments should be avoided. Similarly, tautologies are not allowed in dialogs since they are not informative. To put it differently, they do not bring new information and this is certainly not suitable in persuasion dialogs.

Non-triviality: *A dialog system* DS *ensures non-triviality iff for all persuasion dialog* $\mathcal{D}$ *generated by* DS*, for all* $a \in \{\texttt{Content}(m) \mid m \in \mathcal{D}, \texttt{Act}(m) = \textit{Argue}\}$*,* a *is not atomic and* $\texttt{Conc}(a)$ *is not a tautology (i.e.,* $\texttt{Conc}(a) \notin \texttt{CN}(\emptyset)$*).*

The postulates are compatible (i.e., can be satisfied all together by a dialog system). There is however a problem for ensuring consistency together with natural-attacks-allowance. This is due to Dung's framework, when it is instantiated by symmetric attack relations the system is not able to guarantee to obtain consistent conclusions. Indeed, in [2], an example of violation of the consistency postulate in the context of symmetric attack is described. Nevertheless, the non compatibility of consistency and natural-attacks-allowance in Dung's framework does not mean that the two postulates are not required for dialog systems.

Proposition 2: *Finiteness, consistency, natural-attacks-allowance, non-triviality, non-determinism and dissimulation are compatible.*

Proof: (Sketch) Let $(\mathcal{L}_0, \vdash)$ be a propositional language whose vocabulary contains at least two propositional variables Let $\mathcal{A}_0$ be the arguments built from $\mathcal{L}_0$ by using Definition 1 and let $\mathcal{R}_0 \subset \mathcal{A}_0 \times \mathcal{A}_0$ an arbitrary non reflexive attack relation containing rebut and assumption-attack.

Let DS based on $(\mathcal{L}_0, \vdash)$, with $\{P, C\}$ the set of agents and s.t. a sequence of moves $\mathcal{D} = (m_i)$ is a valid dialog wrt to the protocol iff $\mathcal{D} = (m1)$ or $\mathcal{D} = (m1, m2)$ such that $m1 = \langle s, h, Assert : \varphi \rangle$ and $m2 = \langle s, h, Argue : a \rangle$ where $s \in \{P, C\}$, $h \in \{P, C\}$ and $\varphi \in \mathcal{L}_0$ and φ is finite and $a \in \mathcal{A}_0$ such that a is not atomic and not supporting a tautology and $\texttt{Conc}(a)$ is finite. Let the output of $\mathcal{D}$ be computed under stable semantics. Finiteness, non-triviality, natural-attack allowance hold by construction. Consistency is ensured by the fact that there is at most two arguments, they could be together in the basic extension only if their conclusions are consistent (due to the presence of rebuttal attack). Non determinism and dissimulation can be shown by using a dialog with only the first assert move compared to a dialog in wich an argument against it is added. ■

Proposition 3: *Finiteness, consistency, natural-attacks-allowance, non-triviality, non-determinism and dissimulation are independant.*

Proof: (sketch) For each postulate, we may provide a dialog system in which every postulate holds except the one considered. ■

V. Dialog systems postulates vs. reasoning systems postulates

A dialog system, with two participating agents, uses three argumentation systems of the same kind. They are all logical instantiations of the abstract framework of Dung [10]. Two of the systems are used for modeling the *nonmonotonic reasoning* of the agents and should thus obey to postulates like those recalled in Section II-B. The third argumentation system is devoted to a completely different purpose which is computing the outcomes of persuasion dialogs. The difference of tasks raises the question of the suitability of the postulates of the two reasoning systems for the one that computes the outcomes of dialogs. More generally, are those postulates compatible with the ones proposed previously for dialog systems? In this section we show that the two sets of postulates are incompatible.

A. Reasoning postulates in a dialog context

The first postulate that a reasoning system should satisfy concerns the consistency of its extensions. It ensures that the system returns sound results. A similar postulate is required for a dialog system, namely for its argumentation system that computes the outcomes of dialogs. While this postulate is compatible with the two closure ones, in case of dialogs this is unfortunately not guaranteed. Indeed, we have shown in the previous section that consistency is not compatible with natural-attacks-allowance, namely when symmetric attack relations are used. It is worth mentioning that in reasoning, symmetric attack relations can be avoided. Indeed, there exist non-symmetric attack relations that ensure the consistency postulate (see [2]). However, things are not so simple in dialogs. Getting rid of rebuttals in dialogs would constrain the kind of moves agents may utter. This would also mean that it is not possible to design dialog systems (based on Dung's system) that capture everyday life dialogs in which rebuttals are very common.

Closure under $\texttt{CN}$ is a suitable postulate for reasoning systems since it guarantees a form of "completeness" of their outputs. Tautologies are among the plausible conclusions that are ensured. These formulas, even if they are trivial, may serve as a basis for testing the quality of those systems. However, in a dialog context, they are not suitable since they are not "informative". Thus, closure under $\texttt{CN}$ is not a required postulate for dialog systems. We can even show that it is incompatible with the non-triviality postulate.

Proposition 4: *Let* DS *be a dialog system. If* $\texttt{CN}(\emptyset) \neq \emptyset$ *and* DS *satisfies non-triviality, then for all dialog* $\mathcal{D}$ *generated by* DS*, if the argumentation system* $\mathsf{AS}_{\mathcal{D}}$ *admits at least one extension then it violates closure under* $\texttt{CN}$*.*

Proof: Let DS be a dialog system and $\texttt{CN}(\emptyset) \neq \emptyset$. Since DS satisfies non-triviality then for any dialog $\mathcal{D}$ generated under DS, $\nexists a \in \texttt{Args}(\mathcal{D})$ such that $\texttt{Conc}(a) \in \texttt{CN}(\emptyset)$.

Let $\mathcal{E}$ be an extension of the argumentation system $\mathsf{AS}_{\mathcal{D}}$ under a given semantics and let $\varphi \in \mathtt{CN}(\emptyset)$. Then, $\varphi \notin \{\mathtt{Conc}(a), a \in \mathcal{E}\}$. However, $\varphi \in \mathtt{CN}(\{\mathtt{Conc}(a), a \in \mathcal{E}\})$ since $\emptyset \subseteq \{\mathtt{Conc}(a), a \in \mathcal{E}\}$ and from the monotonicity of $\mathtt{CN}$, it follows that $\mathtt{CN}(\emptyset) \subseteq \mathtt{CN}(\{\mathtt{Conc}(a), a \in \mathcal{E}\})$. Hence, $\mathsf{AS}_{\mathcal{D}}$ violates closure under $\mathtt{CN}$. ■

Some dialog systems may even miss some non trivial conclusions. Let us consider a dialog in which only two arguments $(\{x \wedge y\}, x)$[3] and $(\{z \wedge t\}, z)$ are exchanged. The argumentation system associated with this dialog has only one stable/preferred extension which contains only the two arguments. It is easy to check that this extension is not closed under $\mathtt{CN}$ since, for instance, y, t and $x \wedge z$ are not supported by arguments in the extension.

Closure under sub-arguments is another postulate which makes sense for reasoning systems but not for dialog ones. Indeed, in a dialog context, this postulate is ensured in case agents utter all the sub-arguments of their arguments. This is certainly not realistic. Let us consider the following dialog between Carla and Peter.

> **Carla:** Mary will miss her exams. She did not work hard.
> **Peter:** She worked hard. Her eyes are encircled and she is very tired.

The corresponding argumentation system contains the two exchanged arguments and one attack from the argument (say b) of Peter to that of Carla (say a). This system has one stable/preferred extension: $\{b\}$. This extension is not closed under sub-arguments since b has at least two sub-arguments (one for "Mary's eyes are encircled" and one for "Mary is very tired") which are not in the extension. In order to satisfy the postulate, Peter should utter two additional arguments for the two statements. The following result shows that this postulate is even not compatible with the non-triviality one.

Proposition 5: *Let DS be a dialog system. If DS satisfies non-triviality, then for all dialog $\mathcal{D}$ generated by DS whose argumentation system $\mathsf{AS}_{\mathcal{D}}$ admits non-empty extensions, $\mathsf{AS}_{\mathcal{D}}$ violates closure under sub-arguments.*

Proof: Let $\mathcal{D}$ be a dialog generated by a dialog system DS. Let $\mathsf{AS}_{\mathcal{D}}$ be its argumentation system and $\mathcal{E}$ be a non-empty extension of $\mathsf{AS}_{\mathcal{D}}$ under a given semantics. Thus, $\exists a \in \mathcal{E}$. Since $\mathcal{E} \subseteq \mathtt{Args}(\mathcal{D})$ and DS satisfies non-triviality, then $\mathtt{Supp}(a) \neq \emptyset$. Consequently, $\exists \varphi \in \mathtt{Supp}(a)$. There are two possible cases:

- $\varphi \in \mathtt{CN}(\emptyset)$. Then, $(\emptyset, \varphi)$ is a tautological argument and $(\emptyset, \varphi) \notin \mathtt{Args}(\mathcal{D})$ since DS satisfies non-triviality. Thus, $(\emptyset, \varphi) \notin \mathcal{E}$.
- $\varphi \notin \mathtt{CN}(\emptyset)$. Then, $(\{\varphi\}, \varphi)$ is an atomic argument and $(\{\varphi\}, \varphi) \notin \mathtt{Args}(\mathcal{D})$ since DS satisfies non-triviality. So $(\{\varphi\}, \varphi) \notin \mathcal{E}$.

Both arguments $(\emptyset, \varphi)$ and $(\{\varphi\}, \varphi)$ are sub-arguments of a and do not belong to $\mathcal{E}$. Thus, $\mathsf{AS}_{\mathcal{D}}$ is not closed under sub-arguments. ■

B. Dialog postulates in a reasoning context

This section discusses the suitability of the postulates of dialog systems in a reasoning context. We start with the *finiteness* postulate which ensures finite dialogs. An important question is: do argumentation systems for reasoning need to be finite (i.e., have a finite number of arguments)? From a computational perspective, finiteness is certainly a desirable property since the computation of the extensions of infinite systems would be hard if not impossible. However, in practice the finiteness property depends broadly on the logic $(\mathcal{L}, \mathtt{CN})$ underlying the argumentation system. For a broad class of logics, the set of *all* arguments that may be built from a knowledge base is infinite. This is particularly the case for classical logics. Nevertheless, it was shown in [5] that for some logics it is possible to consider only a *subset* of the whole set of arguments. The corresponding argumentation system, called *core*, returns exactly the plausible conclusions of the argumentation system that takes as input all the arguments built from the base. For some logics, as shown below, the core is finite. Before presenting the formal result, let us first introduce some useful notations.

Notation 2: *For $X \subseteq \mathcal{L}$, $\mathtt{Cncs}(X) = \{x \in \mathcal{L} \mid \exists Y \subseteq X$ s.t. $\mathtt{CN}(Y) \neq \mathcal{L}$ and $x \in \mathtt{CN}(Y)\}$ is the set of formulae that are drawn from consistent subsets of X, and $(\mathtt{Cncs}(X)/\equiv) = \{[x] \mid x \in X\}$ with $[x] = \{x' \in \mathcal{L} \mid \mathtt{CN}(\{x'\}) = \mathtt{CN}(\{x\})\}$ is the quotient set of $\mathtt{Cncs}(X)$ wrt logical equivalence.*

Proposition 6: *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation system built over a knowledge base Σ. If $(\mathtt{Cncs}(\Sigma)/\equiv)$ is finite, then there exists an argumentation system $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ s.t. $\mathcal{A}' \subseteq \mathcal{A}$, $\mathcal{R}' \subseteq \mathcal{R}$, $\mathtt{Output}(\mathcal{F}) = \mathtt{Output}(\mathcal{F}')$ and $\mathcal{A}'$ is finite.*

Proof: Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an argumentation system built over a knowledge base Σ. In [5], it is shown that each argumentation system has a core. The idea is to take exactly one argument from each equivalence class of arguments of $\mathcal{A}$. Let $\mathcal{F}' = (\mathcal{A}', \mathcal{R}')$ s.t. $\mathcal{A}' \subseteq \mathcal{A}$, $\mathcal{R}' \subseteq \mathcal{R}$ be that system. Theorem 4 of [5] shows that $\mathcal{F}$ and $\mathcal{F}'$ have equivalent extensions. Thus, $\mathtt{Output}(\mathcal{F}) = \mathtt{Output}(\mathcal{F}')$. Besides, Theorem 5 of [5] shows that in this case $\mathcal{A}'$ is finite. ■

However, the core may be infinite for some other logics. To sum up, two cases can be distinguished:

1) Argumentation systems that have finite cores, which, then, can be replaced by systems that satisfy the finiteness postulate.
2) Argumentation systems that have infinite cores which, thus, violate the finiteness postulate.

Note that the notion of core is crucial in reasoning since it gathers the key arguments that are necessary and sufficient to

[3]We assume here that $(\mathcal{L}, \mathtt{CN})$ is propositional logic.

define the plausible conclusions of an argumentation system. Recall that in such systems, arguments are generated from a knowledge base without discrimination. However, in dialogs, agents choose the arguments to utter and may hide some of them. Thus, the argumentation system associated with a dialog does not necessarily contain its core. This will impact the outcome of the dialog since, as shown in [5], when a system does not contain its core, then the status of its arguments are not final and may evolve.

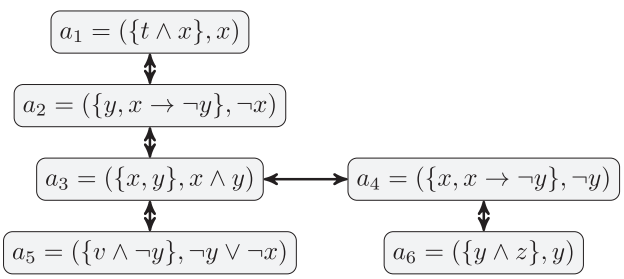
Proposition 7: *Let* DS *be a dialog system satisfying non-triviality. If* $\mathcal{L}$ *contains tautologies and* $\mathcal{R}$ *does not contain attacks against tautoligical arguments, then for all non-empty dialog* $\mathcal{D}$ *generated by* DS, $\mathtt{Output}(\mathsf{AS}_{\mathcal{D}}) \neq \mathtt{Output}(\mathsf{CAS}_{\mathcal{D}})$

where $\mathsf{CAS}_{\mathcal{D}} = (\mathtt{Arg}(\bigcup_{a \in \mathtt{Args}(\mathcal{D})} \mathtt{Supp}(a)), \mathcal{R})$ *and for* $\mathcal{S} \subseteq \mathcal{L}$, $\mathtt{Arg}(\mathcal{S})$ *is the set of all arguments that may be built from* $\mathcal{S}$ *using Definition 1.*

Proof: Indeed, if $\mathcal{D}$ is non-empty then $\mathtt{Args}(\mathcal{D})$ is not empty, thus $\mathsf{CAS}_{\mathcal{D}}$ contains tautological arguments. These arguments should belong to $\mathtt{Ext}(\mathsf{CAS}_{\mathcal{D}})$. Thus, tautologies should belong to $\mathtt{Output}(\mathsf{CAS}_{\mathcal{D}})$ but due to non-triviality they cannot appear in $\mathtt{Output}(\mathsf{AS}_{\mathcal{D}})$. ■

As said before, *consistency* may be violated by the argumentation system that computes the outcomes of dialogs due to the use of the rebutting relation (imposed by the natural-attacks-allowancepostulate). This is particularly the case when the knowledge base contains a ternary or more minimal inconsistent subset as shown below:

Example 2: *Assume a dialog* $\mathcal{D}$ *whose AS is as follows:*



The set $\{a_1, a_5, a_6\}$ *is a preferred extension of this system. However, its set* $\{x, y, \neg x \vee \neg y\}$ *of conclusions is clearly inconsistent.*

The inconsistency problem can be avoided by getting rid of the natural-attacks-allowancepostulate, reducing thus the kind of natural language dialogs that may be conducted. This is certainly not a desirable solution.

Non-triviality postulate (which consists of avoiding atomic arguments and arguments supporting tautologies) is violated by reasoning systems as shown by Proposition 1.

Non-determinism postulate expresses that for any non-trivial formula, it should be possible to generate two dialogs with opposite conclusions. This is very important in dialogs to ensure strategic debates. However, this postulate is not suitable for reasoning systems since these latter should determine in an objective way whether a formula holds or not.

Dissimulation postulate is clearly not compatible with a reasoning system. It highlights a main difference between reasoning and persuading. While in reasoning one looks for the truth of formulas and considers thus all available information, in persuading one looks for convincing another agent about an issue. This may be done by hiding crucial information which run counter the issue.

VI. Conclusion

Since early nineties, there is an increasing number of works trying to formalize dialogs in which agents may exchange arguments. Persuasion and negotiation dialogs have received particular attention from AI community. Several systems were developed for each of them. In those systems arguments are exchanged in order to support *claims* in persuasion dialogs and *offers* in a negotiation context. The arguments are then evaluated using argumentation systems that were originally developed for nonmonotonic reasoning or for reasoning about inconsistent information.

In this paper, we focused on persuasion dialogs, and more precisely on public persuasions where two agents with conflicting opinions try to persuade each other in presence of an audience. Note that the aim here is rather to persuade the audience. We studied whether the approach followed in the literature for defining dialog systems is sound or not. For that purpose, we considered a recent argumentation system proposed in [2] for reasoning about inconsistent information. Note that our study holds for any other logic-based instantiation of the abstract framework of Dung [10], like ASPIC system [9]. We then proposed a *general* persuasion dialog system. This persuasion system is general since one of its basic components (the protocol) is left unspecified. For each of the reasoning and persuading systems, we propose a set of postulates that should hold. For reasoning systems, we considered the three basic postulates defined in [1], namely the consistency of the conclusions supported by the extensions, their closure under the consequence operator, and finally the closure under sub-arguments of the extensions. Regarding dialog systems, we proposed six postulates: the finiteness postulate ensures termination of the dialog. Non-determinism postulate imposes that the system may generate dialogs with different outcomes. The third postulate imposes consistency of the outcomes of dialogs. This postulate ensures that the system that is used for evaluating the arguments exchanged in a dialog is sound. The fourth postulate aims to capture as much as possible everyday life use of counter-argumentation. The dissimulation postulate ensures that an agent may hide some information. The last postulate ensures that agents do not utter trivial arguments during a dialog. An important contribution of this paper consists of investigating the compatibility of the two sets of postulates. We have shown that the three postulates of reasoning systems cannot be satisfied by a dialog system since in this latter the set of exchanged arguments is not complete (due to the finiteness of dialogs and also to the fact that in dialogs, some arguments are considered as trivial and thus do not need to be exchanged). Similarly, we have shown that four postulates of the dialog system cannot be satisfied by the argumentation system. Moreover, we have established that the outcome of a dialog system can be different from the outcome that should be obtained by a reasoning system that would use a knowledge base containing all the formulas exchanged during the dialog. To sum up, the study has revealed that a dialog system needs particular argumentation systems for evaluating its outcomes. Those systems should obey the nature of dialog.

This work can be extended in different ways. The first one consists of defining argumentation systems that are more suitable for public persuasion dialogs and that ensure the postulates discussed in this paper. Another future work consists of defining new postulates for dialogs, namely for capturing manipulation in dialogs. Finally, we are planning to undertake a similar study in the context of negotiation dialogs. Recall that in those systems, the outcome of a negotiation is evaluated by argumentation systems developed for making decisions.

REFERENCES

[1] L. Amgoud. Postulates for logic-based argumentation systems. *International Journal of Approximate Reasoning*, 2013.

[2] L. Amgoud and P. Besnard. Logical limits of abstract argumentation frameworks. *Journal of Applied Non-Classical Logics*, 2013.

[3] L. Amgoud, N. Maudet, and S. Parsons. Modelling dialogues using argumentation. In *Proc. of the Int. Conference on Multi-Agent Systems*, pages 31–38, 2000.

[4] L. Amgoud and H. Prade. Using arguments for making and explaining decisions. *Artificial Intelligence*, 173(3-4):413–436, 2009.

[5] L. Amgoud and S. Vesic. On the equivalence of logic-based argumentation systems. In *SUM'11*, pages 123–136, 2011.

[6] P. Besnard and A. Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128(1-2):203–235, 2001.

[7] E. Bonzon and N. Maudet. On the outcomes of multiparty persuasion. In *10th Int. Conf. on Autonomous Agents and Multiagent Systems*, pages 47–54, 2011.

[8] E. Cabrio and S. Villata. Natural language arguments: A combined approach. In *20th European Conference on Artificial Intelligence*, pages 205–210, 2012.

[9] D2.2. Towards a consensual formal model: inference part. *Deliverable of ASPIC project*, 2004.

[10] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence Journal*, 77:321–357, 1995.

[11] M. Elvang-Gøransson, J. Fox, and P. Krause. Acceptability of arguments as logical uncertainty. In *ECSQARU'93*, pages 85–90, 1993.

[12] J. Fox and S. Parsons. Arguing about beliefs and actions. In *Applications of Uncertainty Formalisms*, pages 266–302, 1998.

[13] A. García and G. Simari. Defeasible logic programming: an argumentative approach. *Theory and Practice of Logic Programming*, 4:95–138, 2004.

[14] M. Johnson, P. McBurney, and S. Parsons. When are two protocols the same? In *Communication in Multiagent Systems*, pages 253–268, 2003.

[15] A. Kakas and P. Moraitis. Adaptive agent negotiation via argumentation. In *Proc. 5th Int. J. Conf. on Auton. Agents and Multi-Agents syst.*, pages 384–391, 2006.

[16] D. Kontarinis, E. Bonzon, N. Maudet, and P. Moraitis. Picking the right expert to make a debate uncontroversial. In *Comp. Models of Arg.*, pages 486–497, 2012.

[17] P. McBurney, D. Hitchcock, and S. Parsons. The eightfold way of deliberation dialogue. *Int. J. Intell. Syst.*, 22(1):95–132, 2007.

[18] S. Parsons, C. Sierra, and N. R. Jennings. Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3):261—292, 1998.

[19] J. L. Pollock. How to reason defeasibly. *Artif. Intelligence Journal*, 57:1–42, 1992.

[20] H. Prakken. Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation*, 15:1009–1040, 2005.

[21] G. Quiroz, D. Apothéloz, and P. Brandt. How counter-argumentation works. *Argumentation Illuminated*, pages 172–177, 1992.

[22] C. Salavastru. *Logique, Argumentation, Interprétation*. L'Harmattan, Paris, 2007.

[23] A. Tarski. *Logic, Semantics, Metamathematics (E. H. Woodger, editor))*, chapter On Some Fundamental Concepts of Metamathematics. Oxford Uni. Press, 1956.

[24] S. Zabala, I. Lara, and H. Geffner. Beliefs, reasons and moves in a model for argumentative dialogues. In *25th Latino-Amer. Conf. on Comp. Science*, 1999.