Measures for persuasion dialogs: A preliminary investigation

Leila AMGOUD [a,1], Florence DUPIN DE SAINT CYR [a]
[a] *Institut de Recherche en Informatique de Toulouse (IRIT), France*

Abstract. Persuasion is one of the main types of dialogs encountered in everyday life. The basic idea behind a persuasion is that two (or more) agents disagree on a state of affairs, and each one tries to persuade the other to change his mind. For that purpose, agents exchange arguments of different strengths.

Several systems, grounded on argumentation theory, have been proposed in the literature for modeling persuasion dialogs. These systems have studied more or less deeply the different protocols required for this kind of dialogs, and have investigated different termination criteria. However, nothing is said about the *properties* of the generated dialogs, nor on the behavior of the interacting agents. Besides, analyzing dialogs is a usual task in everyday life. For instance, political debates are generally deeply dissected.

In this paper we define *measures* for analyzing dialogs from the point of view of an external agent. In particular, three kinds of measures are proposed: i) measures of the quality of the exchanged arguments in terms of their strengths, ii) measures of the behavior of each participating agent in terms of its *coherence*, its *aggressiveness* in the dialog, and finally in terms of the *novelty* of its arguments, iii) measures of the quality of the dialog itself in terms of the *relevance* and *usefulness* of its moves.

Keywords. Argumentation, Persuasion dialogs, Quality measures

1. Introduction

Since the seminal work by Walton and Krabbe [11] on the role of argumentation in dialog, and on the classification of different types of dialogs, there is an increasing interest on modeling those dialog types using argumentation techniques. Indeed, several *dialog systems* have been proposed in the literature for modeling *information seeking* dialogs (e.g. [8]), *inquiry* dialogs (e.g. [4]), *negotiation* (e.g. [10]), and finally *persuasion* dialogs (e.g. [3,6,9,13]). Persuasion dialogs are initiated from a position of conflict in which one agent believes p and the other believes $\neg p$, and both try to persuade the other to change its mind by presenting arguments in support of their thesis.

It is worth noticing that in all these disparate works, a dialog system is built around three main components: i) a *communication language* specifying the locutions that will be used by agents during a dialog for exchanging information, arguments, offers, etc., ii) a *protocol* specifying the set of rules governing the well-definition of dialogs, and iii)

[1]Corresponding Author: IRIT–CNRS, 118, route de Narbonne, 31062, Toulouse Cedex, France; E-mail: amgoud@irit.fr.

agents' strategies which are the different tactics used by agents for selecting their moves at each step in a dialog.

All the above systems allow agents to engage in dialogs that obey of course to the rules of the protocol. Thus, the only properties that are guaranteed for a generated dialog are those related to the protocol. For instance, one can show that a dialog terminates, the turn shifts equally between agents in that dialog (if such rule is specified by the protocol), agents can backtrack to an early move in the dialog, etc. Note that the properties inherited from a protocol concern the way the dialog is generated. However, they don't say anything about the *properties* of that dialog.
Judging the properties of a dialog may be seen as a subjective issue. Two people listening to the same political debate may disagree, for instance, on the "winner" of the debate, and more generally on their feeling about the dialog itself. Nevertheless, it is possible to define more objective criteria, for instance, the aggressiveness of each participant, the way agents may borrow ideas from each others, the self-contradiction of agents, the relevance of the exchanged information, etc.

Focusing only on persuasion dialogs, in this paper, we are concerned by analyzing already generated dialogs whatever the protocol used is and whatever the strategies of the agents are. We place ourselves in the role of an external observer that tries to evaluate the dialog. For this purpose, three kinds of measures are proposed: 1) Measures of the quality of the exchanged arguments in terms of their *weights*, 2) Measures of the behavior of each participating agent in terms of its *coherence*, its *aggressiveness* in the dialog, and finally in terms of the *source* of its arguments, 3) Measures of the properties of the dialog itself in terms of the relevance and usefulness of its moves. These measures are of great importance since they can be used as guidelines for a protocol in order to generate the "best" dialogs. They can also serve as a basis for analyzing dialogs that hold between agents.

The rest of the paper is organized as follows: Section 2 recalls the basics of argumentation theory. In Section 3, we present the basic concepts of a persuasion dialog. Section 4 details our dialog measures as well as their properties. Section 5 is devoted to some concluding remarks and conclusions.

2. Basics of argumentation systems

Argumentation is a reasoning model based on the construction and the comparison of arguments whose definition will be given in Section 3. In [5], an argumentation system is defined as follows:

Definition 1 (Argumentation system) *An* argumentation system *(AS) is a pair* $T = \langle \mathcal{A}, \mathcal{R} \rangle$*, where* $\mathcal{A}$ *is a set of arguments and* $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ *is an attack relation. We say that an argument* α_i *attacks an argument* α_j *iff* $(\alpha_i, \alpha_j) \in \mathcal{R}$ *(or* $\alpha_i \; \mathcal{R} \; \alpha_j$*).*

Note that to each argumentation system is associated an oriented graph whose nodes are the different arguments, and the edges represent the attack relation between them. Let $\mathcal{G}_T$ denote the graph associated to the argumentation system $T = \langle \mathcal{A}, \mathcal{R} \rangle$.

Since arguments are conflicting, it is important to know which arguments are acceptable. For that purpose, in [5], different acceptability semantics have been proposed. Let us recall them here.

Definition 2 (Conflict-free, Defence) *Let $\mathcal{B} \subseteq \mathcal{A}$.*

- $\mathcal{B}$ *is* conflict-free *iff* $\nexists \alpha_i, \alpha_j \in \mathcal{B}$ *such that* $(\alpha_i, \alpha_j) \in \mathcal{R}$.
- $\mathcal{B}$ defends *an argument* α_i *iff for each argument* $\alpha_j \in \mathcal{A}$*, if* $(\alpha_j, \alpha_i) \in \mathcal{R}$*, then* $\exists \alpha_k \in \mathcal{B}$ *such that* $(\alpha_k, \alpha_j) \in \mathcal{R}$.

Definition 3 (Acceptability semantics) *Let $\mathcal{B}$ be a conflict-free set of arguments of $\mathcal{A}$.*

- $\mathcal{B}$ *is an* admissible extension *iff $\mathcal{B}$ defends all its elements;*
- $\mathcal{B}$ *is a* preferred extension *iff it is a maximal (w.r.t. set-$\subseteq$) admissible extension;*
- $\mathcal{B}$ *is a* stable extension *iff it is a preferred extension that attacks w.r.t. the relation $\mathcal{R}$ all arguments in $\mathcal{A} \backslash \mathcal{B}$.*

Let $\mathcal{E}_1, \ldots, \mathcal{E}_n$ denote the possible extensions under a given semantics.

Now that the acceptability semantics are defined, we can define the status of any argument. As we will see, an argument may have one among three possible statuses: *skeptically accepted, credulously accepted* and *rejected*.

Definition 4 (Argument status) *Let $\langle \mathcal{A}, \mathcal{R} \rangle$ be an argumentation system, and $\mathcal{E}_1, \ldots, \mathcal{E}_n$ its extensions under stable (resp. preferred) semantics. Let $\alpha \in \mathcal{A}$.*

- α *is* skeptically accepted *iff* $\alpha \in \mathcal{E}_i, \forall \mathcal{E}_i \neq \emptyset$ *with* $i = 1, \ldots, n$*;*
- α *is* credulously accepted *iff α is in some extensions and not in others.*
- α *is* rejected *iff* $\nexists \mathcal{E}_j$ *such that* $\alpha \in \mathcal{E}_j$.

3. Persuasion dialogs

Let $\mathcal{L}$ be a logical language from which arguments may be built. In our application, arguments are reasons of believing something. Throughout the paper, the structure and the origin of such arguments are supposed to be unknown. However, an argument is assumed to have at least two parts: a *support* (representing the set of premises or formulas used to build the argument) and a *conclusion* (representing the belief one wants to justify through the argument). Arguments will be denoted by lowercase Greek letters.

Definition 5 (Argument) *An* argument α *is a pair* $\alpha = \langle H, h \rangle$ *where* $h \in \mathcal{L}$ *and* $H \subseteq \mathcal{L}$. H *is the* support *of the argument returned by the function* $H = \texttt{Support}(\alpha)$*, and h is its* conclusion *returned by the function* $h = \texttt{Conc}(\alpha)$.

In what follows, $\arg$ denotes a function that returns for a given set $S \subseteq \mathcal{L}$ all the arguments that may be built from formulas of S. Thus, $\arg(\mathcal{L})$ is the set of arguments that may be built from the whole logical language $\mathcal{L}$.

As well established in the literature and already said, arguments may be conflicting since, for instance, they may support contradictory conclusions. In what follows, $\mathcal{R}_\mathcal{L}$ is a binary relation that captures all the conflicts that may exist among arguments of $\arg(\mathcal{L})$.

Thus, $\mathcal{R}_\mathcal{L} \subseteq \arg(\mathcal{L}) \times \arg(\mathcal{L})$. For two arguments $\alpha, \beta \in \arg(\mathcal{L})$, the pair $(\alpha, \beta) \in \mathcal{R}_\mathcal{L}$ means that the argument α attacks the argument β.

Let $\mathtt{Ag} = \{a_1, \ldots, a_n\}$ be a set of symbols representing agents that may be involved in a persuasion dialog. Each agent is supposed to be able to recognize each argument of $\arg(\mathcal{L})$ and each conflict in $\mathcal{R}_\mathcal{L}$. Note that this does not mean at all that an agent is aware of all the arguments. This assumption means that agents use the same logical language and the same definition of arguments.

A persuasion dialog consists mainly of an exchange of arguments. Of course other kinds of moves can be exchanged like questions and assertions. However, arguments play the key role in determining the outcome of the dialog. Thus, throughout the paper, we are only interested by the arguments exchanged in a dialog. The subject of such a dialog is an argument, and its aim is to compute the status of that argument. If at the end of the dialog, the argument is "skeptically accepted" or "rejected", then we say that the dialog has *succeeded*, otherwise the dialog has *failed*.

Definition 6 (Moves) *A* move $m \in \mathcal{M}$ *is a triple* $\langle S, H, x\rangle$ *such that:*

- $S \in \mathtt{Ag}$ *is the agent that utters the move,* $\mathtt{Speaker}(m) = S$
- $H \subseteq \mathtt{Ag}$ *is the set of agents to which the move is addressed,* $\mathtt{Hearer}(m) = H$
- $x \in \arg(\mathcal{L})$ *is the content of the move,* $\mathtt{Content}(m) = x$.

During a dialog several moves may be uttered. Those moves constitute a sequence denoted by $\langle m_0, \ldots, m_n\rangle$, where m_0 is the initial move whereas m_n is the final one. The empty sequence is denoted by $\langle\rangle$. For any integer n, the set of sequences of length n is denoted by $\mathcal{M}^n$. These sequences are built under a given protocol. A protocol amounts to define a function that associates to each sequence of moves, a set of valid moves. Several protocols have been proposed in the literature, like for instance [3,9]. In what follows, we don't focus on particular protocols.

Definition 7 (Persuasion dialog) *A* persuasion dialog D *is a non-empty and finite sequence of moves* $\langle m_0, \ldots, m_n\rangle$.
The subject *of* D *is* $\mathtt{Subject}(D) = \mathtt{Content}(m_0)$, *and the* length *of* D, *denoted* $|D|$, *is the number of moves* $n+1$. *Each sub-sequence* $\langle m_0, \ldots, m_i\rangle$ *(with* $i < n$*) is a* sub-dialog D^i *of* D. *We will write also* $D^i \sqsubset D$.

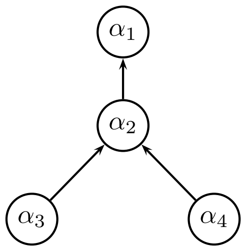
It is worth noticing that to each persuasion dialog D, one may associate an argumentation system that will be used to evaluate the status of each argument uttered during the dialog. This argumentation system is also used to compute the output of the dialog.

Definition 8 (AS of a persuasion dialog) *Let* $D = \langle m_0, \ldots, m_n\rangle$ *be a persuasion dialog. The* argumentation system *of* D *is the pair* $\mathsf{AS}_D = \langle \mathtt{Args}(D), \mathtt{Confs}(D)\rangle$ *s.t:*

1. $\mathtt{Args}(D) = \{\mathtt{Content}(m_i) \mid i = 0, \ldots, n\}$
2. $\mathtt{Confs}(D) = \{(\alpha, \beta)$ *such that* $\alpha, \beta \in \mathtt{Args}(D)$ *and* $(\alpha, \beta) \in \mathcal{R}_\mathcal{L}\}$

In other words, $\mathtt{Args}(D)$ and $\mathtt{Confs}(D)$ return respectively, the set of arguments exchanged during the dialog and the different conflicts among those arguments.

Example 1 *Let D be the following persuasion dialog between two agents a_1 and a_2. $D = \langle\langle a_1, \{a_2\}, \alpha_1\rangle, \langle a_2, \{a_1\}, \alpha_2\rangle, \langle a_1, \{a_2\}, \alpha_3\rangle, \langle a_1, \{a_2\}, \alpha_4\rangle, \langle a_2, \{a_1\}, \alpha_1\rangle\rangle$. Let us assume that there exist conflicts in $\mathcal{R}_\mathcal{L}$ among some of these arguments. Those conflicts are summarized in the figure below.*



In this case, $\texttt{Args}(D) = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ *and* $\texttt{Confs}(D) = \{(\alpha_2, \alpha_1), (\alpha_3, \alpha_2), (\alpha_4, \alpha_2)\}$.

Property 1 *Let D be a persuasion dialog. $\forall D^j$ such that $D^j \sqsubset D$,* $\texttt{Args}(D^j) \subseteq \texttt{Args}(D)$, *and* $\texttt{Confs}(D^j) \subseteq \texttt{Confs}(D)$.

Any dialog has an output. In case of a persuasion, the output of a dialog is either the status of the argument under discussion (i.e. the subject) when that status is "skeptically accepted" or "rejected", or failure in case the status of the subject is "credulously accepted". The idea is that a dialog succeeds as soon as the status of the subject is determined, and thus a winner agent is known. However, when an argument is credulously accepted, this means that each agent keeps its position w.r.t. the subject, and the dialog fails to meet its objective.

Definition 9 (Output of a persuasion dialog) *Let D be a persuasion dialog. The* output *of D, denoted by* $\texttt{Output}(D)$ *is*

$$\begin{cases} A & \textit{iff } \texttt{Subject}(D) \textit{ is } \text{skeptically accepted } \textit{in } \mathsf{AS}_D \\ R & \textit{iff } \texttt{Subject}(D) \textit{ is rejected in } \mathsf{AS}_D \\ \texttt{Fail} & \textit{iff } \texttt{Subject}(D) \textit{ is } \text{credulously accepted } \textit{in } \mathsf{AS}_D \end{cases}$$

It may be the case that from the set of formulas involved in a set E of arguments, it is possible to build new arguments that do not belong to E. Let

$$\texttt{Formulas}(E) = \cup_{\alpha \in E} \texttt{Support}(\alpha)$$

be that set of formulas. Due to the monotonic construction of arguments, $E \subseteq \arg(\texttt{Formulas}(E))$ but the reverse is not necessarily true. Indeed, an argument remains always an argument even when new attackers are received. However, its status is non-monotonic, and may change. As a consequence of the previous inclusion, new conflicts may also appear among those new arguments, and even between new arguments and elements of E. This shows clearly that the argumentation system associated with a dialog is not necessarily "complete". In what follows, we define the complete version of an argumentation system associated with a given dialog D.

Definition 10 (Complete AS) *Let D be a persuasion dialog, and* $\mathsf{AS}_D = \langle \texttt{Args}(D), \texttt{Confs}(D)\rangle$ *its associated AS.*
The complete AS *is* $\mathsf{CAS}_D = \langle \arg(\texttt{Formulas}(\texttt{Args}(D))), \mathcal{R}_c\rangle$ *where* $\mathcal{R}_c = \{(\alpha, \beta)$ *s.t.* $\alpha, \beta \in \arg(\texttt{Formulas}(\texttt{Args}(D)))$ *and* $(\alpha, \beta) \in \mathcal{R}_\mathcal{L}\}$.

Recall that $\texttt{Args}(D) \subseteq \arg(\texttt{Formulas}(\texttt{Args}(D)))$ and $\texttt{Confs}(D) \subseteq \mathcal{R}_c \subseteq \mathcal{R}_{\mathcal{L}}$. Note that the status of an argument α in the system AS_D is not necessarily the same as in the complete system CAS_D.

4. Measuring persuasion dialogs

In this section we discuss different measures of persuasion dialogs. Three aspects can be analyzed: 1) the quality of the exchanged arguments during a persuasion dialog, 2) the agent's behavior, and 3) the properties of the dialog itself.

4.1. Measuring the quality of arguments

During a dialog, agents utter arguments that may have different *weights*. A weight may highlight the quality of information involved in the argument in terms for instance of its certainty degree. It may also be related to the cost of revealing that information. In [1], several definitions of such arguments' weights have been proposed, and their use for comparing arguments has been made explicit. It is worth noticing that the same argument may not have the same weight from one agent to another. In what follows, a weight in terms of a numerical value is associated to each argument. The greater this value is, the better the argument.

$$\texttt{weight} : \arg(\mathcal{L}) \longrightarrow \mathbb{N}^*$$

The function `weight` is given by the agent which wants to analyze the dialog. Thus, it may be given by an agent that is involved in the dialog, or by an external one. On the basis of arguments' weights, it is possible to compute the weight of a dialog as follows:

Definition 11 (Measure of dialog weight) *Let D be a persuasion dialog. The* weight *of D is* $\texttt{Weight}(D) = \sum_{\alpha_i \in \texttt{Args}(D)} \texttt{weight}(\alpha_i)$

It is clear that this measure is monotonic. Formally:

Property 2 *Let D be a persuasion dialog. $\forall D^j \sqsubset D$ then $\texttt{Weight}(D^j) \leq \texttt{Weight}(D)$.*

This measure allows to compare pairs of persuasion dialogs only on the basis of the exchanged arguments. It is even more relevant when the two dialogs have the same subject and got the same output.

It is also possible to compute the weight of arguments uttered by each agent in a given dialog. For that purpose, one needs to know what has been said by each agent. This can be computed by a simple projection on the dialog given that agent.

Definition 12 (Dialog projection) *Let $D = \langle m_0, \ldots, m_n \rangle$ be a persuasion dialog, and $a_i \in \texttt{Ag}$. The* projection *of D on agent a_i is $D^{a_i} = \langle m_{i_1}, \ldots, m_{i_k} \rangle$ such that $0 \leq i_1 \leq \ldots \leq i_k \leq n$ and $\forall l \in [1, k]$, $m_{i_l} \in D$ and $\texttt{Speaker}(m_{i_l}) = a_i$.*

The contribution of each agent is defined as follows:

Definition 13 (Measure of agent's contribution) *Let $D = \langle m_0, \ldots, m_n \rangle$ be a persuasion dialog, and $a_i \in$ Ag. The* contribution *of agent a_i in D is*

$$\texttt{Contr}(a_i, D) = \frac{\sum \texttt{weight}(\alpha_i) \text{ s.t. } \alpha_i \in \texttt{Args}(D^{a_i})}{\texttt{Weight}(D)}.$$

Example 2 *Let us consider the persuasion dialog D presented in Example 1. Recall that $\texttt{Args}(D) = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, $D^{a_1} = \{\alpha_1, \alpha_3, \alpha_4\}$ and $D^{a_2} = \{\alpha_1, \alpha_2\}$. Suppose now that an external agent wants to analyze this dialog. The function "$weight$" of this agent is as follow: $\texttt{weight}(\alpha_1)$ = 1, $\texttt{weight}(\alpha_2)$ = 4, $\texttt{weight}(\alpha_3)$ = 2 and $\texttt{weight}(\alpha_4)$ = 3. It is then clear from the definitions that the overall weight of the dialog is $\texttt{Weight}(D)$ = 10. The contributions of the two agents are respectively $\texttt{Contr}(a_1, D)$ = 6/10 and $\texttt{Contr}(a_2, D)$ = 5/10.*

Consider now an example in which an agent sends several times the same argument.

Example 3 *Let us consider a persuasion dialog D such that $\texttt{Args}(D) = \{\alpha, \beta\}$. $D^{a_1} = \{\alpha\}$ and $D^{a_2} = \{\beta\}$. Let us assume that there are 50 moves in this dialog of which 49 moves are uttered by agent a_1 and one move uttered by agent a_2. Suppose now that an external agent wants to analyze this dialog. The function "$weight$" of this agent is as follow: $\texttt{weight}(\alpha)$ = 1 and $\texttt{weight}(\beta)$ = 30. The overall weight of the dialog is $\texttt{Weight}(D)$ = 31. The contributions of the two agents are respectively $\texttt{Contr}(a_1, D)$ = 1/31 and $\texttt{Contr}(a_2, D)$ = 30/31.*

It is easy to check that when the protocol under which a dialog is generated does not allow an agent to repeat an argument already given by another agent, then the sum of the contributions of the different agents is equal to 1.

Proposition 1 *Let $D = \langle m_0, \ldots, m_n \rangle$ be a persuasion dialog and $a_1, \ldots, a_n$ the agents involved in it. $\sum_{i=1,\ldots,n} \texttt{Contr}(a_i, D)$ = 1 iff $\nexists m_i, m_j$ with $0 \leq i, j \leq n$, such that $\texttt{Speaker}(m_i) \neq \texttt{Speaker}(m_j)$, and $\texttt{Content}(m_i) = \texttt{Content}(m_j)$.*

It is worth noticing that the measure Contr is not monotonic. Indeed, the contribution of an agent may change during the dialog. It may increase then decreases in the dialog. However, at a given step of a dialog, the contribution of the agent that will present the next move will for sure increase, whereas the contributions of the other agents may decrease. Formally:

Proposition 2 *Let $D = \langle m_0, \ldots, m_n \rangle$ be a persuasion dialog and $a_i \in$ Ag. Let $m \in \mathcal{M}$ such that $\texttt{Speaker}(m) = a_i$. Then, $\texttt{Contr}(a_i, D \oplus m) \geq \texttt{Contr}(a_i, D)$, and $\forall a_j \neq a_i$, $\texttt{Contr}(a_j, D \oplus m) \leq \texttt{Contr}(a_j, D)$, with $D \oplus m = \langle m_0, \ldots, m_n, m \rangle$.*

4.2. Analyzing the behavior of agents

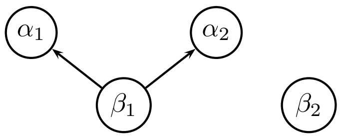
The behavior of an agent in a given persuasion dialog may be analyzed on the basis of three main criteria: i) its degree of *aggressiveness* in the dialog, ii) the source of its arguments, i.e. whether it builds arguments using its own formulas, or rather the ones revealed by other agents, and finally iii) its degree of *coherence* in the dialog.

The first criterion, i.e. the aggressiveness of an agent in a dialog, amounts to compute to what extent an agent was attacking arguments sent by other agents. Such agents prefer to destroy arguments presented by other parties rather than presenting arguments supporting their own point of view. Formally, the *aggressiveness degree* of an agent a_i towards an agent a_j during a persuasion dialog is equal to the number of its arguments that attack the other agent's arguments over the number of arguments it has uttered in that dialog.

Definition 14 (Measure of aggressiveness) *Let* $D = \langle m_0, \ldots, m_n \rangle$ *be a persuasion dialog, and* $a_i, a_j \in \mathtt{Ag}$*. The* aggressiveness degree *of agent* a_i *towards* a_j *in* D *is*

$$\mathtt{Agr}(a_i, a_j, D) = \frac{|\{\alpha \in \mathtt{Args}(D^{a_i}) \text{ s.t. } \exists \beta \in \mathtt{Args}(D^{a_j}) \text{ and } (\alpha, \beta) \in \mathtt{Confs}(D)\}|}{|\mathtt{Args}(D^{a_i})|}.$$ [2]

Example 4 *Let* D *be a persuasion dialog between two agents* a_1 *and* a_2*. Let us assume that* $\mathtt{Args}(D) = \{\alpha_1, \alpha_2, \beta_1, \beta_2\}$, $D^{a_1} = \{\alpha_1, \alpha_2\}$ *and* $D^{a_2} = \{\beta_1, \beta_2\}$*. The conflicts among the four arguments are depicted in the figure below.*



The aggressiveness degrees of the two agents are respectively $\mathtt{Agr}(a_1, a_2, D)$ *= 0/2 = 0, and* $\mathtt{Agr}(a_2, a_1, D)$ *= 1/2.*

The aggressiveness degree of an agent changes as soon as a new argument is uttered by that agent. It decreases when that argument does not attack any argument of the other agent, and increases otherwise. Formally:

Proposition 3 *Let* $D = \langle m_0, \ldots, m_n \rangle$ *be a persuasion dialog and* $a_i, a_j \in \mathtt{Ag}$*. Let* $m \in \mathcal{M}$ *such that* $\mathtt{Speaker}(m) = a_i$*, and let* $D \oplus m = \langle m_0, \ldots, m_n, m \rangle$*.*
$\mathtt{Agr}(a_i, a_j, D \oplus m) \geq \mathtt{Agr}(a_i, a_j, D)$ *iff* $\exists \alpha \in \mathtt{Args}(D^{a_j})$ *such that* $(\mathtt{Content}(m), \alpha) \in \mathcal{R}_{\mathcal{L}}$*, and* $\mathtt{Agr}(a_i, a_j, D \oplus m) < \mathtt{Agr}(a_i, a_j, D)$ *otherwise.*

The second criterion concerns the source of arguments. An agent can build its arguments either from its own knowledge base, thus using its own formulas, or using formulas revealed by other agents in the dialog. In [2], this idea of borrowing formulas from other agents has been presented as one of the tactics used by agents for selecting the argument to utter at a given step of a dialog. The authors argue that by doing so, an agent may minimize the risk of being attacked subsequently.
Let us now check to what extent an agent borrows information from other agents. Before that, let us first determine which formulas are owned by each agent according to what has been said in a dialog. Informally, a formula is owned by an agent, if this formula is revealed for the first time by that agent. Note that a formula revealed for the first time by agent a_i may also pertain to the base of another agent a_j. Here, we are interested by who reveals first that formula.

Definition 15 (Agent's formulas) *Let* $D = \langle m_0, \ldots, m_n \rangle$ *be a persuasion dialog, and* $a_i \in \mathtt{Ag}$*. The* formulas owned *by agent* a_i *are:* $\mathtt{OwnF}(a_i, D) = \{x \in \mathcal{L}$ *such that* $\exists m_j$ *with* $\mathtt{Speaker}(m_j) = a_i$ *and* $x \in \mathtt{Support}(\mathtt{Content}(m_j))$ *and* $\nexists m_k$ *s.t* $k < j$ *and* $\mathtt{Speaker}(m_k) \neq a_i$ *and* $x \in \mathtt{Support}(\mathtt{Content}(m_k))\}$.

[2] The expression $|E|$ denotes the cardinal of the set E.

Now that we know which formulas are owned by each agent, we can compute the *degree of loan* for each participating agent. It may be tactically useful to turn an agents' arguments against him since they should be immune from challenge. This loan degree can thus help for evaluating the strategical behavior of an agent.

Definition 16 (Measure of loan) *Let $D = \langle m_0, \ldots, m_n \rangle$ be a persuasion dialog, and $a_i, a_j \in \texttt{Ag}$. The* loan degree *of agent a_i from agent a_j in D is:*

$$\texttt{Loan}(a_i, a_j, D) = \frac{|\texttt{Formulas}(\texttt{Args}(D^{a_i})) \cap \texttt{OwnF}(D, a_j)|}{|\texttt{Formulas}(\texttt{Args}(D^{a_i}))|}$$

The third criterion concerns the coherence of an agent. Indeed, in a persuasion dialog where an agent a_i defends its point of view, it is important that this agent does not contradict itself. In fact, there are two kinds of self contradiction:

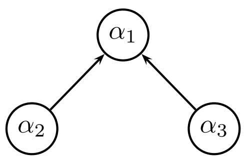
1. an *explicit* contradiction in which an agent presents at a given step of a dialog an argument, and later it attacks that argument. Such conflicts already appear in the argumentation system $\mathsf{AS}_{D^{a_i}} = \langle \texttt{Args}(D^{a_i}), \texttt{Confs}(D^{a_i}) \rangle$ associated to the moves uttered by agent a_i. In other words, the set $\texttt{Confs}(D^{a_i})$ is not empty.
2. an *implicit* contradiction that appears only in the complete version of that system, i.e. in $\mathsf{CAS}_{D^{a_i}}$.

In what follows, we will define a measure, a *degree of incoherence*, for evaluating to what extent an agent was incoherent in a dialog.

Definition 17 (Measure of incoherence) *Let D be a persuasion dialog, $a_i \in \texttt{Ag}$, and $\mathsf{CAS}_{D^{a_i}} = \langle \mathcal{A}_c^{a_i}, \mathcal{R}_c^{a_i} \rangle$ its complete system. The* incoherence degree *of agent a_i in D is*

$$\texttt{Inc}(a_i, D) = \frac{|\mathcal{R}_c^{a_i}|}{|\mathcal{A}_c^{a_i} \times \mathcal{A}_c^{a_i}|}$$

Example 5 *Let D be a persuasion dialog in which agent a_1 has uttered two arguments α_1 and α_2. Let us assume that from the formulas of those arguments a third argument, say α_3, is built. The figure below depicts the conflicts among the three arguments. The incoherence degree of agent a_1 is equal to 2/9.*



Note that, the above definition is general enough to capture both explicit and implicit contradictions. Moreover, this measure is more precise than the one defined on the basis of attacked arguments, i.e. $\texttt{Inc}(a_i, D) = \frac{|\{\beta \in \mathcal{A}_c^{a_i} \text{ such that } \exists (\alpha,\beta) \in \mathcal{R}_c^{a_i}\}|}{|\mathcal{A}_c^{a_i}|}$. Using this measure, the incoherence degree of the agent in the above example is 1/3. Indeed, even if the argument α_1 is attacked by two arguments, only one conflict is considered.

It is easy to check that if an agent is aggressive towards itself, then it is incoherent.

Property 3 *Let D be a persuasion dialog, and $a_i \in$ $\texttt{Ag}$. If $\texttt{Agr}(a_i, a_i, D) > 0$ then $\texttt{Inc}(a_i, D) > 0$.*

The following example shows that the reverse is not always true.

Example 6 *Let D be a persuasion dialog, $a_i \in$ $\texttt{Ag}$. Let us assume that $\texttt{Args}(D^{a_i})$ = $\{\alpha_1, \alpha_2\}$, and $\texttt{Confs}(D^{a_i}) = \emptyset$. This means that $\texttt{Agr}(a_i, a_i, D) = 0$. Let $\mathsf{CAS}_{D^{a_i}}$ = $\langle \mathcal{A}_c^{a_i}, \mathcal{R}_c^{a_i} \rangle$ be the complete version of the previous system with $\mathcal{A}_c^{a_i}$ = $\{\alpha_1, \alpha_2, \alpha_3\}$ and $\mathcal{R}_c^{a_i}$ = $\{(\alpha_3, \alpha_1), (\alpha_3, \alpha_2)\}$. It is clear that $\texttt{Inc}(a_i, D)$ = 2/9.*

Similarly, it can be shown that if all the formulas of an agent a_i are borrowed from another agent a_j and that a_i is aggressive towards a_j, then a_j is for sure incoherent.

Proposition 4 *Let D be a persuasion dialog, and $a_i, a_j \in$ $\texttt{Ag}$. If $\texttt{Loan}(a_i, a_j, D) = 1$ and $\texttt{Agr}(a_i, a_j, D) > 0$, then $\texttt{Inc}(a_j, D) > 0$.*

4.3. Measuring the dialog itself

In the previous sections, we have defined measures for evaluating the quality of arguments uttered in a persuasion dialog, and others for analyzing the behavior of agents involved in such a dialog. In this section, we define two other measures for evaluating the quality of the dialog itself. The first measure checks to what extent moves uttered in a given dialog are in relation with the subject of that dialog. It is very common in everyday life, that agents deviate from the subject of the dialog. Before introducing the measure, let us first define formally the notion of relevance.

Definition 18 (Relevant move) *Let $D = \langle m_0, \ldots, m_n \rangle$ be a persuasion dialog. A move $m_{i=0,\ldots,n}$ is* relevant *to the subject of D iff there exists a path (not necessarily directed) from $\texttt{Subject}(D)$ to $\texttt{Content}(m_i)$ in the directed graph associated with AS_D.*

Example 4 (continued): *Let us consider the persuasion dialog given in Example 4. Suppose that $\texttt{Subject}(D) = \alpha_1$. It is clear that the arguments α_2, β_1 are relevant, whereas the argument β_2 is not.*

On the basis of this notion of relevance, one can define a measure for knowing the percentage of moves that are relevant in a dialog.

Definition 19 (Measure of relevance) *Let D = $\langle m_0, \ldots, m_n \rangle$ be a persuasion dialog. The* relevance degree *of D is*

$$\texttt{Relevance}(D) = \frac{|\{m_{i=0,\ldots,n} \textit{ such that } m_i \textit{ is relevant }\}|}{|D|}$$

Example 4 (continued): *In the previous example, $\texttt{Relevance}(D)$ = 3/4.*

It is clear that the greater this degree is, the better the dialog. When the relevance degree of a dialog is equal to 1, this means that agents did not deviate from the subject of the dialog. However, this does not mean at all that all the moves have an impact on the result of the dialog, i.e. on the status of the subject. Another measure is then needed to compute the percentage of useful moves. Before introducing this measure, let us first define what is a useful move. The following definition is similar to the one used in [9].

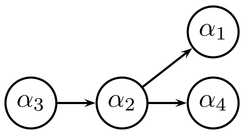
Definition 20 (Useful move) *Let $D = \langle m_0, \ldots, m_n \rangle$ be a persuasion dialog. A move $m_{i=0,\ldots,n}$ is* useful *in D iff $\texttt{Output}(D) \neq \texttt{Output}(D \setminus m_i)$ where $D \setminus m_i$ is the dialog obtained by removing the move m_i*[3].

It can be checked that the two notions of usefulness and relevance are closely related.

Proposition 5 *Let D be a persuasion dialog, and $m \in \mathcal{M}$. If m is useful in D, then m is relevant to the subject of D.*

Note that the converse is not true as shown in the following example.

Example 7 *Let us assume a dialog D whose subject is α_1, and whose graph is the one presented below.*



The set $\{\alpha_1, \alpha_3, \alpha_4\}$ is the only preferred extension of AS_D. It is clear that the argument α_4 is relevant to α_1, but it is not useful for D. Indeed, the removal of α_4 will not change the status of α_1 which is skeptically accepted.

On the basis of the above notion of usefulness of moves, it is possible to compute the degree of usefulness of the dialog as a whole.

Definition 21 (Measure of usefulness) *Let $D = \langle m_0, \ldots, m_n \rangle$ be a persuasion dialog. The* usefulness degree *of D is*

$$\texttt{Usefulness}(D) = \frac{|\{m_{i=0,\ldots,n} \text{ such that } m_i \text{ is useful }\}|}{|D|}$$

It is worth noticing that according to the structure of the graph associated to the argumentation system of a dialog, it is possible to know whether the degrees of relevance and usefulness of that dialog are less that 1 or not. Formally:

Proposition 6 *Let D be a persuasion dialog, and $\mathcal{G}$ be the graph associated with AS_D. If $\mathcal{G}$ is not connected then $\texttt{Usefulness}(D) < 1$ and $\texttt{Relevance}(D) < 1$.*

5. Conclusion

Several systems have been proposed in the literature for allowing agents to engage in persuasion dialogs. Different dialog protocols have then been discussed. These latter are the high level rules that govern a dialog. Examples of such rules are "how the turn shifts between agents", and "how moves are chained in a dialog". All these rules should ensure "correct" dialogs, i.e. dialogs that terminate and reach their goals. However, they

[3] $D \setminus m_i = \begin{cases} \langle m_0, \ldots m_{i-1}, m_{i+1}, \ldots, m_n \rangle & \text{if } i \neq 0 \text{ and } i \neq n \\ \langle m_1, \ldots, m_n \rangle & \text{if } i = 0 \\ \langle m_0, \ldots, m_{n-1} \rangle & \text{if } i = n \end{cases}$

don't say anything on the *quality* of the dialogs. One even wonders whether there are criteria for measuring the quality of a dialog. In this paper we argue that the answer to this question is yes. Indeed, under the same protocol, different dialogs on the same subject may be generated, and some of them may be judged better than others. There are three kinds of reasons, each of them is translated into quality measures: i) the arguments exchanged are stronger, ii) the generated dialogs are more *concise* (i.e. all the uttered arguments have an impact on the result of the dialog), iii) the behavior of agents was "ideal". In the paper, the behavior of an agent is analyzed on the basis of three main criteria: its *degree of aggressiveness*, its *degree of loan*, and its *degree of coherence*. In sum, different measures have been proposed in this paper for the quality of dialogs. To the best of our knowledge, this is the first work on such measures in dialogs. Exceptions may be the works by Hunter [7] and by Yuan et col. [12] on defining dialog strategies. For instance, Hunter has defined a strategy for selecting arguments in a dialog. The basic idea is that an agent selects the ones that will satisfy the goals of the audience. The agent is thus assumed to maintain two bases: a base containing its own beliefs, and another base containing what the agent thinks are the goals of the audience. These works are thus more concerned with proposing dialog strategies than with analyzing dialogs.

An extension of this work would be the study of the general properties of protocols generating good dialogs w.r.t. the measures presented in this paper. Another future work consists of applying these measures to other types of dialogs, especially negotiation.

References

[1] L. Amgoud and C. Cayrol. Inferring from inconsistency in preference-based argumentation frameworks. *Int. Journal of Automated Reasoning*, Volume 29 (2):125–169, 2002.

[2] L. Amgoud and N. Maudet. Strategical considerations for argumentative agents (preliminary report). In *Proceedings of the 10th International Workshop on Non-Monotonic Reasoning NMR'2002 (Collocated with KR'2002), session Argument, Dialogue, Decision*, pages 409–417, 2002.

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

[4] L. Black and A. Hunter. A generative inquiry dialogue system. In *International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 963–970. ACM Press, 2007.

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

[6] T. F. Gordon. The pleadings game. *Artificial Intelligence and Law*, 2:239–292, 1993.

[7] A. Hunter. Towards higher impact argumentation. In *Proceedings of the National Conference on Artificial Intelligence*, pages 963–970. ACM Press, 2004.

[8] S. Parsons, M. Wooldridge, and L. Amgoud. Properties and complexity of some formal inter-agent dialogues. *Journal of Logic and Computation*, 13(3):347–376, 2003.

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

[10] K. Sycara. Persuasive argumentation in negotiation. *Theory and Decision*, 28(3):203–242, 1990.

[11] D. N. Walton and E. C. W. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press, Albany, 1995.

[12] T. Yuan, D. Moore, and A. Grierson. An assessment of dialogue strategies for a human computer debating system, via computational agents. In *Proc. of ECAI'2004 Workshop on Computational Models of Natural Argument*, pages 17–24, 2004.

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