Strategical considerations for argumentative agents (preliminary report)

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

Nicolas Maudet
IRIT, 118 route de Narbonne, 31062
Toulouse Cedex, France
maudet@irit.fr

Abstract

This paper aims at exploring some aspects of the strategical problem of move selection in the context of persuasive dialogues. The proposed heuristics are based on some human strategies issued from natural dialogues. A three-layered approach of strategy is defended, and a preliminary strategical deliberation process (illustrating the interplay of these different levels) is proposed.

Keywords: dialogue, strategies, argumentation.

1 INTRODUCTION

An argumentative agent is an agent whose reasoning capability is based on some argumentative system, *i.e.* an agent who is able to support his beliefs/intentions with some justifying reasons called *arguments*.
Let us consider the case of two argumentative agents involved in some dialogic interaction. Arguably, the normative models of dialogue introduced in the seventies by the pioneering work of Hamblin [9] and best known as *dialectical systems* or *dialogue games* are well-suited to capture such interactions[1]. A dialectical system mainly consists of:

- a set of moves (*e.g.* challenge, assertion, question);
- one commitment store ($\mathcal{CS}$) for each conversant;
- a set of commitment rules defining the effect of the moves on the commitment stores;
- a set of dialogue rules regulating the moves (*i.e.* a protocol);
- some termination rules defining when the dialogue ends.

While the model was primarily introduced to capture persuasive dialogue, it has recently been argued [18] and practicaly shown [3] that the model is adapted to other kinds of interactions (negociation, deliberation, etc.).

A protocol gives the set of moves possibly expected after a previous move. Choosing among these moves is the strategy problem. While protocol is a public notion, strategy is crucially an individualistic matter. There is of course an interesting debate about whether to put things on the normative or on the strategical side (Carlson [7] says: *"Is it a rule or just a good move to answer a question if one can?"*). This is an interesting question, but a design question if any. The protocol issue really depends on the kind of interaction studied, and all what we can say, following Loui [10], is that a fair protocol might leave room for strategy. Interaction between autonomous agents is not proof-theory: the outcome of a dialogue is expectedly non-deterministic, and agents might "lose" dialogue when they are not wise [14].

To make the problem more concrete consider the following simple piece of dialogue:

(1) A: You should ratify the Kyoto Protocol.
B: Why?
A: Because global warming is dangerous, and the Kyoto Protocol is the best means to avoid global warming.

Now imagine that B has (at least) a counter-argument against the fact that the Kyoto Protocol is indeed the

[1] Although the underlying logic of these systems is generally monotonic, see [16].

best means to avoid global warming. First of all, it is not surprising that most systems prevent the agent to rationally accept the reasons in this case (see *e.g.* [3]). But within these rational bounds (and within the normative bounds of the protocol) what should agent B prefer: continue to challenge the opponent or rather propose its counter-argument? And then: which element of the support is it best to challenge? Which counter-argument should be presented? And again would the choice be the same if the time of the debate is limited?

Strategy depends on a collection of factors uneasily captured: protocol, goals of the players, resource allowed, etc. A comprehensive and exhaustive account of this notion is of course out of the scope of the paper. However, this paper aims at exploring some aspects of this issue. In the next section, we present the kind of interaction studied here, and we exhibit evidence of strategy and heuristics inspired from human dialogues. Section 3 introduces our formal framework, and adapts the notion of acceptability of arguments which forms the backbone of our approach. This allows in turn to propose some strategies adapted to artificial agents. The paper ends with a discussion.

2 STRATEGY IN PERSUASIVE DIALOGUE

This section introduces the kind of interaction studied in this paper (thus defining the roles and goals of the agents), and presents some evidence of strategies inspired from naturally occuring dialogues.

2.1 CRITICAL DISCUSSION

In the case of persuasive dialogue [14], an agent (the *proponent*) has the role to defend its point of view while the *opponent* tries to attack it. This does not involve however that the opponent is reduced to only challenge the proponent: the opponent can propose counter-arguments, and this *may* (see [15]) put the burden to defend some facts on its side. According to [18], critical discussion can be considered as a special case of persuasive dialogue in that each protagonist has in the initial situation the burden to prove its own point of view on the (conflicting) issue.

2.2 EVIDENCE FOR STRATEGIES IN DIALECTICAL INTERACTIONS

Moore [13] has studied how natural critical discussions are constrained by the rule of the DC dialectical system [11]. Studying the resulting dialogues, he concludes that strategy is best analysed a three levels:

1. maintaining focus of the dispute;
2. building its point of view or destroying the opponent's one;
3. select method to fulfill the objective set at level 1 and 2.

This 3-level account appropriatly captures the classical distinction made between strategy and tactic. Strategy involves deliberation (*e.g.* planning the line of argumentation). Tactic is basically the mean to reach the aims fixed at the strategical level. In Moore's approach, level 1 and level 2 refer to strategy, while level 3 refers to tactic. Moore also notes that the dialogue participant only considers levels 1 and 2 if he has the initiative in the conversation (a participant has the initiative if its move is not strongly constrained by the partner's previous move: *e.g* at the beginning of the dialogue or after a statement). He also gives examples of heuristics triggered at level 3 (see Tab.1)

While this helps by giving example of concrete *natural* heuristics, further work needs to be done to computationally capture such notions. How to model "unwelcome" commitments, "good" reasons, or "carefully choosen" support? The aim of the next section is to give formal definitions of some of these notions in order to propose an account of the strategical reasoning of artificial argumentative agents.

3 STRATEGIES FOR ARGUMENTATIVE AGENTS

To start with, we make more precise the notions of protocol and strategy informally introduced in Section 1. The set of argumentative agents is denoted by $\mathcal{A}$. $\mathcal{M}$ is the language used by our agents to communicate. It specifies the kinds of acts the agents can make and the content language used to convey information. Although it may be extended to capture other types of dialogue such as negociation [5], this set at least contains claims (asserting conclusions and reasons), challenges (asking for reasons), and (yes/no) questions [14]. A dialogue protocol is the set of rules governing the high-level behavior of interacting agents. It first defines the allowed replies to each act (given the current state of the commitment stores). Since the update of the commitment stores ($\mathcal{CS}$) is related to the kind of act played by an agent, we can imagine that it is also part of a dialogue protocol. These two components of the

Table 1: Moore's heuristics for DC-like games

Previous move	Prefered following moves
question	answer in such a way to avoid unwelcome commitments give answers for which good reasons can be found answer truthfully
challenge	answer with some carefully choosen support
resolution	retract the least harmful conjunct
statement	if there is contradiction in partner's $\mathcal{CS}$, ask for resolution if there is evidence *vs* any partner's commitment, assert it seek objection in partner's $\mathcal{CS}$ challenge partner's arguments

dialogue protocol can be seen as two different functions denoted respectively by Π and $\mathcal{U}$. The function Π returns, from the last act made (and the current state of the commitment store) the set of legal next acts. Although we will use in this paper only simple protocols (*e.g.* where no backtracking replies are allowed), it is worth pointing out that a particular interest of dialogue-game protocols is to allow flexibility in dialogue structures by relying more on the commitment stores than on dialogue history. The function $\mathcal{U}$ describes how the commitment store of an agent is updated after a given act. We can now formally define a dialogue protocol as follows:

Definition 1 *A dialogue protocol is defined as a tuple* $\mathcal{P} = \langle \mathcal{M}, \Pi, \mathcal{U} \rangle$ *such that:*

- $\mathcal{M}$ *is a set of allowed acts;*
- $\Pi : \mathcal{M} \times \mathcal{CS} \mapsto 2^{\mathcal{M}}$*;*
- $\mathcal{U} : \mathcal{A} \times \mathcal{M} \times \mathcal{CS} \mapsto \mathcal{CS}$.

The set of moves that an agent might consider at his turn is constrained by some rationality rules. The function $\mathcal{R}$ specifies the preconditions ($\mathcal{C}$) that must be checked by the agent before providing a given act: it returns a boolean saying if the act can rationally be played or not by a given agent.

Definition 2 *The rationality preconditions is defined as a function* $\mathcal{R}$ *s.t.*

- $\mathcal{R} : \mathcal{M} \times \mathcal{A} \times \mathcal{C} \mapsto \{0,\ 1\}$

Thus, the dialogue protocol and the rationality preconditions specify the different possible replies after a given move. From that, the exact reply that an agent proposes is the result of the strategy. Hence, the strategy depends on the chosen protocol. Formally, we can therefore think of a given strategy for an agent as being a function from the set of moves identified by the protocol and the rationality preconditions to a single move which is then uttered by an agent:

Definition 3 *A* dialogue strategy *is defined as a function* $\mathcal{S}$*:*

$$\mathcal{S} : 2^{\mathcal{M}} \mapsto \mathcal{M}.$$

3.1 LAYERED STRATEGIES

As simple as it is however, the function hides some real difficulties: in particular, the agent must select the kind of act, but also the content to expose, and this depends on the beliefs of the agent and on the current state of the dialogue, among others. We argue in consequence that this notion is best captured at different levels, as suggested by Moore. In the context of our argumentative agents however, we leave aside the level 1 as defined by this author (the matter of maintaining the focus of the dispute) since this would involve more complex protocols than those discussed here —possibly to take into account some notions of relevance [16]. We replace this level with the different profiles of the agents involved in the interaction (a profile will influence the choice made at the other levels). The account that we propose is then layered as follows:

1. define some agent profile;
2. choose to build or destroy;
3. choose some appropriate argumentative content.

This does not indicate however, that the levels are independant. On the contrary, there is an interplay between these levels, and we propose at the end of the section a process that take into account this dependance, and selects the best move to play *w.r.t* to the strategy defined below. But let us first discuss the different levels separatly.

3.2 THE AGENT PROFILE

Amgoud and Parsons have recently [4] made some proposals towards the definition of broad classes of agent profiles:

- Agreeable: *accept* whenever possible.
- Disagreeable: only *accept* when no reason not to.
- Open-minded: only *challenge* when necessary.
- Argumentative: *challenge* whenever possible.
- Elephant Child: *question* whenever possible.

They recognize however that "although [they] have termed these "strategies" each is only a partial definition of $\mathcal{S}$—a full definition would have to take into account the nature of the previous move and hence the overall protocol for the dialogue." In other words, the agent profile is just the first level of strategy, it needs to be completed with some other levels (and this is what is tentatively done through the other levels of strategies). In this paper, the only agent profile that we consider is some fixed level p of *prudence*. Intuitively, prudent agents will be more reluctant to expose their arguments than non prudent agents.

3.3 BUILD OR DESTROY

At the next level, we will consider that the agent can adopt either a *build* or a *destroy* strategy (b-strategy vs. d-strategy). A difference with Moore's proposal is that the strategy is not necessarily selected towards the main claim of the dialogue (the controversial issue), but possibly towards any statement from the participants' commitment stores.

b-strategy consists of defending some facts in its commitment store;

d-strategy consists of attacking some commitments in the opponent's commitment store.

The question is of course: how does the agent choose to switch from one strategy to another? No satisfying answer can be made in isolation from the other levels.

3.4 THE ARGUMENTATIVE CONTENT

The last level concerns the problem of the (argumentative) content of the act. As suggested in [3], the argumention system might help to improve this level of strategy. Undoubtly, the argumention system has things to say about the notions used by Moore in his heuristics: "good" reasons, selection of the support, and so on. In the following, we will modify the notion of acceptability proposed in [1] in order to make the system really helpful for this level of strategy. We start with a brief description of the framework proposed in [1], which is based on the abstract argumentation frameworks proposed by Dung [8].

3.4.1 Basic Definitions

Definition 4 *An* argument *is a pair (H, h) where h is a formula of a propositional language $\mathcal{L}$ and H a subset of a knowledge base Σ such that i) H is consistent, ii) $H \vdash h$ and iii) H is minimal, so no subset of H satisfying both i) and ii) exists. H is called the* support *of the argument and h is its* conclusion.

In general, since Σ is inconsistent, arguments in $\mathcal{A}(\Sigma)$, the set of all arguments which can be made from Σ, will conflict, and we make this idea precise with the notion of undercutting:

Definition 5 *Let (H_1, h_1) and (H_2, h_2) be two arguments of $\mathcal{A}(\Sigma)$. (H_1, h_1)* undercuts *(H_2, h_2) iff $\exists h \in H_2$ such that $h \equiv \neg h_1$. In other words, an argument is undercut iff there exists an argument for the negation of an element of its support.*

To capture the fact that some facts are more strongly believed (or desired, or intended, depending on the nature of the facts) we assume that any set of facts has a preference order over it which derives from the stratification of the knowledge base Σ into non-overlapping sets $\Sigma_1, \ldots, \Sigma_n$ such that facts in Σ_i are all equally preferred and are more preferred than those in Σ_j where $j > i$. The preference level of a nonempty subset H of Σ, $level(H)$, is the number of the highest numbered layer which has a member in H.

Definition 6 *Let (H_1, h_1) and (H_2, h_2) be two arguments in $\mathcal{A}(\Sigma)$. (H_1, h_1) is* preferred *to (H_2, h_2) according to $Pref$ iff $level(H_1) \leq level(H_2)$.*

The preference order makes it possible to distinguish different types of relation between arguments:

Definition 7 *Let A, B be two arguments of $\mathcal{A}(\Sigma)$.*

- *B* strongly undercuts *A iff B undercuts A and it is not the case that $A \gg^{Pref} B$.*
- *If B undercuts A then A* defends itself *against B iff $A \gg^{Pref} B$.*

- *A set of arguments $\mathcal{S}$* defends *A if for every argument B which undercuts A and A cannot defend itself against, we can find an argument C in $\mathcal{S}$ which strongly undercuts B.*

Henceforth, $C_{undercut,Pref}$ will gather all non-undercut arguments and arguments defending themselves against all their undercutting arguments. In [2], it has been shown that the set $\underline{\mathcal{S}}$ of acceptable arguments of the argumentation system $\langle A(\Sigma), Undercut, Pref\rangle$ is the least fixpoint of a function $\mathcal{F}$ defined as follows:

$\mathcal{S} \subseteq A(\Sigma), \mathcal{F}(S) = \{(H,h) \in A(\Sigma) | (H,h)$ is defended by $\mathcal{S}\}$.

Definition 8 *The set of* acceptable *arguments for an argumentation system $\langle \mathcal{A}(\Sigma)$, Undercut, Pref$\rangle$ is:*

$$\begin{aligned}\underline{\mathcal{S}} &= \bigcup \mathcal{F}_{i\geq 0}(\emptyset)\\ &= C_{Undercut,Pref} \cup \left[\bigcup \mathcal{F}_{i\geq 1}(C_{Undercut,Pref})\right]\end{aligned}$$

An argument is acceptable if it is a member of the acceptable set.

(2) Let $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$ with $\Sigma_1 = \{\neg a\}, \Sigma_2 = \{a, a \rightarrow b\}$ and $\Sigma_3 = \{\neg b\}$.

(3) (follows 2)
$(\{\neg a\}, \neg a)$ and $(\{a, a \rightarrow b\}, b)$ are two arguments of $\mathcal{A}(\Sigma)$. The argument $(\{\neg a\}, \neg a)$ undercuts $(\{a, a \rightarrow b\}, b)$. The preference level of $(\{a, a \rightarrow b\}, b)$ is 2 whereas the preference level of $(\{\neg a\}, \neg a)$ is 1, and so $(\{\neg a\}, \neg a) \gg^{Pref} (\{a, a \rightarrow b\}, b)$. The argument $(\{\neg a\}, \neg a)$ is in $C_{Undercut,Pref}$ because it is preferred (according to *Pref*) to the unique undercutting argument $(\{a\}, a)$. Consequently, $(\{\neg a\}, \neg a)$ is in $\underline{\mathcal{S}}$. The argument $(\{\neg b\}, \neg b)$ is undercut by $(\{a, a \rightarrow b\}, b)$ and does not defend itself. On the contrary, $(\{\neg a\}, \neg a)$ undercuts $(\{a, a \rightarrow b\}, b)$ and $(\{\neg a\}, \neg a) \gg^{Pref} (\{a, a \rightarrow b\}, b)$. Therefore, $C_{Undercut,Pref}$ defends $(\{\neg b\}, \neg b)$ and consequently $(\{\neg b\}, \neg b) \in \underline{\mathcal{S}}$.

3.4.2 Levels Of Acceptability

According to the definition of [2], all arguments in $\underline{\mathcal{S}}$ are acceptable. However, the intuition is that they are not *equally* acceptable. The idea is simply that the more defense you need, the more you are exposed to contradiction. In this respect, unattacked arguments are considered stronger than self-defended arguments which are in turn considered stronger than jointly defended arguments. In the class of these defended arguments then, it is considered stronger to be defended with the shortest chain of arguments.

Definition 9 (leveled set of acceptable arguments) *The set of* acceptable *arguments at level l for an argumentation system $\langle \mathcal{A}(\Sigma)$, Undercut, Pref$\rangle$ is:*

$$\underline{\mathcal{S}^l} = \bigcup_{i=0..l} \mathcal{F}_{i\geq 0}(\emptyset)$$

Definition 10 (level of acceptable arguments) *The level l at which the argument is acceptable in the context of some argumentation system $\langle \mathcal{A}(\Sigma)$, Undercut, Pref$\rangle$ is given by the lacc function defined as follows:*

- *lacc: $\mathcal{A}(\Sigma) \mapsto N$ such that:*
  1. *if $A \in C_{Undercut}$ then $lacc(A) = 0$;*
  2. *if $A \in C_{Undercut,Pref}$ then $lacc(A) = 1$;*
  3. *$A \in S^l$ and $A \notin S^{l-1}$ then $lacc(A) = l+1$.*

Definition 11 (strengh of an argument) *Let A and B be two arguments of $\underline{\mathcal{S}}$. A is considered stronger than B iff $lacc(A) < lacc(B)$.*

(4) (follows 2)
The argument $(\{\neg a\}, \neg a)$ is in $C_{Undercut,Pref}$ because it is preferred (according to *Pref*) to the unique undercutting argument $(\{a\}, a)$. Consequently, $(\{\neg a\}, \neg a)$ is in $\underline{\mathcal{S}}$ and then the level of acceptability of this argument is $lacc(\{\neg a\}, \neg a) = 1$.
The argument $(\{\neg b\}, \neg b)$ is undercut by $(\{a, a \rightarrow b\}, b)$ and does not defend itself. However, $(\{\neg a\}, \neg a)$ undercuts in turn $(\{a, a \rightarrow b\}, b)$ and $(\{\neg a\}, \neg a) \gg^{Pref} (\{a, a \rightarrow b\}, b)$. Therefore, $C_{Undercut,Pref}$ defends $(\{\neg b\}, \neg b)$ and consequently $(\{\neg b\}, \neg b) \in \underline{\mathcal{S}}$. The level of acceptability of this argument $(\{\neg b\}, \neg b)$ is 2. As a consequence, $(\{\neg a\}, \neg a)$ is considered stronger than $(\{\neg b\}, \neg b)$.

The stronger is an argument, the more grounded (the "good" reasons of Moore's strategies) is the conclusion supported by this argument.

3.5 PUTTING THE LEVELS TOGETHER: THE STRATEGY DELIBERATIVE PROCESS

As correctly pointed out by Moore himself, the picture of an agent selecting some move to play w.r.t. to the strategy and then filling the content of the move with a carefully choosen argument is far too simple. We expect that the strategy deliberation is a complex process involving all three levels detailed below. As an example of this interplay, we cite the intuition behind the decision of whether choosing a b-strategy or d-strategy when the agent has the initiative: it is rational to select a b-strategy if there is some *sufficiently* acceptable arguments available to support some fact in its commitment store. The sufficiently acceptable level is defined at level 1 by the agent profile (depending on its attitude, *i.e* its level of prudence, the agent will easily or not commit to a b-strategy), and the level of acceptability of the facts is computed at level 3.

strategy deliberative process As an example of the strategical reasoning of argumentative agents involving these different levels, we propose a simple *strategy deliberative process*. Consider an argumentative agent A with a level of prudence p_a involved in a dialectical interaction with an agent B, through the commitment stores denoted respectively $\mathcal{CS}_A$ and $\mathcal{CS}_B$. We suppose that the agent A has adopted a current strategy **c-strategy** $\in$ {**b-strategy, d-strategy**}. When he has the initiative, the agent can trigger the following algorithm:

1. if **c-strategy= b-strategy** then consider first 2. then 3 else consider 3. then 2.
2. if the agent has some acceptable argument A such that:
   (a) $A = (H, p) \in \underline{\mathcal{S}}$ and $p \in \mathcal{CS}_A$;
   (b) $lacc(A) = l'$ and $l' < p_a$.

   then the agent adopts a **b-strategy**, and plays $assert(H)$[2].
3. if the agent has some acceptable argument A such that:
   (a) $A = (H, p) \in \underline{\mathcal{S}}$ and $\neg p \in \mathcal{CS}_B$;
   (b) $lacc(A) = l'$ and $l' < p_a$.

   then the agent adopts an **d-strategy**, and plays $assert(\neg p)$[3].
4. if the agent has no argument for some fact p in the opponent's commitment store, the adopts a **d-strategy** and plays $challenge(p)$.
5. else the agent selects p in the opponent's commitment store for which he has some acceptable argument A such that:
   (a) $A = (H, p) \in \underline{\mathcal{S}}$ and $p \in \mathcal{CS}_B$;
   (b) $\nexists p' : p' \in \mathcal{CS}_B \wedge$
   $A' = (H', p') \in \underline{\mathcal{S}} \wedge lacc(A) < lacc(A')$

   then the agent adopts a **d-strategy** and plays $challenge(p)$.

[2]In case of different arguments with the same level of acceptability, the agent may prefer to choose the smallest argument in order to restrict exposure to defeaters.

[3]Note that this prepares some forthcoming b-strategies for p.

The deliberation algorithm proposed above simply captures the intuition that an agent will explore the position of the partner (typically by challenging its position) in order to find a weak point in its line of argumentation. Point 1. simply stabilizes the strategical behaviour of the agent by giving preference to the current strategy (considered first).

3.6 AN EXAMPLE OF DIALOGUE

We conclude with an example dialogue that illustrates the notion presented below. We consider that agent A is a very prudent agent (*e.g.* the level p is 1) whose beliefs and preferences are described by the content of the layered knowledge base given in Tab. 2. We consider that the dialogue starts with the statement by the participants of the controversial issue. We do not give here the detail of the protocol and rationality preconditions used, but provide the useful information through the example. Note that the example is discussed from A's point of view.

> **Turns 1-2:**
> A_1: *I think newspapers can't publish this information.*
> B_2: *Of course they can.*

The commitment stores are updated in consequence (the statements are added in the $\mathcal{CS}$).

$$\mathcal{CS}_A = \{\neg publish\}$$
$$\mathcal{CS}_B = \{publish\}$$

Now A has the initiative. The set of legal moves given by the protocol is $\{assert(H), challenge(publish)\}$. Note that $(assert(\neg publish)$ is not permitted since $\neg publish$ is already in the $\mathcal{CS}$. So the agent A uses the strategy deliberation process to select its move. He first evaluates the level of acceptability of its arguments in favor of $\neg publish$. The argument

Table 2: The knowledge base for agent A

Σ	A
1	$\neg agree$ $minister, minister \rightarrow \neg private$
2	$health, private$ $private \wedge \neg agree \rightarrow \neg publish, health \rightarrow \neg(minister \rightarrow \neg private)$

$\mathcal{A}_1 = (\{private, \neg agree, private \wedge \neg agree \rightarrow \neg publish\}, \neg publish)$

has a level of acceptability of 2, since it is attacked by

$\mathcal{A}_2 = (\{minister, minister \rightarrow \neg private\}, \neg private)$

and is reinstated by

$\mathcal{A}_3 = (\{health, health \rightarrow \neg(minister \rightarrow \neg private)\}, \neg(minister \rightarrow \neg private))$

The level of acceptability of the argument in favor of $\neg publish$ (and of its arguments against $publish$) is higher than the level of prudence p. Agent A then adopts a d-strategy and choose to challenge B in order to find some others targets. The dialogue continues as follows:

Turns 3-4:
A_3: *Why can they publish this information?*
B_4: *Because the information is not private, and any public information can be published.*

The commitment stores are updated in consequence:

$\mathcal{CS}_A = \{\neg publish\}$
$\mathcal{CS}_B = \{publish, \neg private, \neg private \rightarrow publish\}$

Since the current strategy of agent A is a d-strategy, the agent compute the levels of acceptability of its arguments against $\neg private$ and $\neg private \rightarrow publish\}$. Agent A has no acceptable argument against $\neg private \rightarrow publish$ and an acceptable argument for $private$:

$\mathcal{A}_4 = (\{private\}, private)$

whose level of acceptability is 2, since attacked by $\mathcal{A}_2$ and reinstated by $\mathcal{A}_3$. Hence agent A stays in a d-strategy and challenges $\neg private$.

Turns 5-6:
A_5: *Why is the information public?*
B_6: *Because the information concerns a minister, and informations concerning a minister are public.*

The new state of the commitment stores is:

$\mathcal{CS}_A = \{\neg publish\}$
$\mathcal{CS}_B = \{publish, \neg private, \neg private \rightarrow publish, minister, minister \rightarrow \neg private\}$

Now, using again the strategy deliberation process, the agent will find an acceptable argument against one of the facts of the partner's commitment store ($minister \rightarrow \neg private$) with a level of acceptability 0, since supported by the argument

$\mathcal{A}_5 = (\{health, health \rightarrow \neg(minister \rightarrow \neg private)\}, \neg(minister \rightarrow \neg private)$

which is not attacked. This score is lower than the prudence level p of the agent, so this allows the agent to prepare a b-strategy by directly contradicting the opponent's these.

Turns 7-8:
A_7: *Informations concerning a minister are not necessarily public.*
B_8: *Why?*

$\mathcal{CS}_A = \{\neg publish, \neg(minister \rightarrow \neg private)\}$
$\mathcal{CS}_B = \{publish, \neg private, \neg private \rightarrow publish, minister, minister \rightarrow \neg private\}$

Agent A has not the initiative, so the process is not used. Agent A answers by giving the support of the argument $\mathcal{A}_5$, and we can suppose for convenience that the partner will in turn accept this argument.

Turns 9-10:
A_9: *Informations about the health of minister remain private.*
B_{10}: *I accept your argument.*

Note that a less prudent agent (*e.g.* p=5) would have choosen a b-strategy earlier in the dialogue.

4 DISCUSSION

The paper has presented an exploration of strategies for move selection in persuasive dialogue conducted by argumentative agents. We believe that this area of research will soon become very important, considering the recent developments of commitment-based or dialogue game based protocols for inter-agent communication —see [12] for a survey. The issue is also important when trying to make a computational argumentative agent interact with a human, for instance in educational debate. But it is worth noting that the issue slightly differs in this case: a computational agent too strategically wise might lead to difficulties (as noticed by Moore in p.c.).

At the best of our knowledge however, very few works (besides those reported in this paper) tackle the problem of strategy. A notable exception is a paper of Bench-Capon [6]. Having defined a *Toulmin Dialogue Game Machine*, the author proposes some heuristics for move selection (in a kind of interaction more cooperative than critical discussion). Note first that although the *TDG* machine has an interesting flexibility at the level of dialogue structure[4], the computational player never deliberates at this level (Bench-Capon argues that this is a "matter of personal taste" and "not essential" for strategy). So it leaves aside the level 1 as defined by Moore, as we do in this paper. But the main difference with our work lies in the underlying argumentation framework. Whereas Bench-Capon assumes that the knowledge in the base has initially the form of some defeasible Toulmin-like rules, we use a more general (extended) Dung-like framework which computes the acceptability of the arguments. And this argumentation framework is used to guide the strategical behaviour of the agent. We also defend a three-layered approach of strategy, where this level is used in combination with the agent profile and long-term strategies (build or destroy). An example of strategical deliberation —that we admit very simple as it is— illustrates the interplay of these different levels. Beside the fact that the proposed heuristics are still preliminary, we believe that some points deserve further discussion here.

First of all, the approach described in this paper is based on classical propositional logic. It was justified by our previous works in argument-based inter-agent communication [3]. As pointed out by a reviewer, this is however a serious loss of generality since it excludes many types of arguments discussed in the litterature. In further works, we plan to explore means of abstracting from this logic.

Also, the paper has only considered the case of ideal agents. Ideal agents compute (in particular, determine the acceptability of their arguments) for free. But computation takes resources, and for instance spending too much time trying to determine the acceptability of an argument may be a poor strategy. What if the agent cannot conclude within the bound of the resources? Loui suggests for instance that "players can be exhorted to advance their best arguments when time is limited" [10] The role of strategy is even more crucial when taking into account the resource-bounded nature of the agents. For instance, knowing that the partner is close to resource termination will certainly incite to try risky moves, with the hope he will not have enough resource to contradict it. Hence the level of prudence of the agent could vary through the dialogue —to avoid procrastination of attack of the agents.

Another important question that we leave aside in this paper concerns the prefered content order of the selected arguments. We consider here that the agent expose the argument as a whole. Several works has shown the importance of the content ordering of argument [17], and certainly more elaborated strategies should take into ccount this aspect of the problem.

At last, as it is now, the level of acceptability only concerns arguments and not directly the supported facts. This prevents to take into account the intuition that a fact defended by several different acceptable arguments should be prefered for exposure to a fact defended only by an acceptable argument (without considering the levels of acceptability). The question of how this might interfere with the notion of level of acceptability for the arguments defined here is an open issue for future research.

Acknowledgements

We would like to thank D. Moore and the three anonymous reviewers for the quality of their comments about this paper.

References

[1] L. Amgoud. *Contribution a l'intégration des préferences dans le raisonnement argumentatif.*

[4] Basically, a *swith focus* move is introduced, which permits to modify the stack structure of the claims under discussion.

PhD thesis, Université Paul Sabatier, Toulouse, 1999.

[2] L. Amgoud and C. Cayrol. On the acceptability of arguments in preference-based argumentation framework. In *Actes de 14th Conference on Uncertainty in AI*, pages 1–7, 1998.

[3] L. Amgoud, N. Maudet, and S. Parsons. Modelling dialogues using argumentation. In *Proceedings of the 4th International Conference on Multi Agent Systems (ICMAS00)*, Boston, 2000.

[4] L. Amgoud and S. Parsons. Agent dialogue with conflicting preferences. In *Proceedings of the International Workshop on Agent Theories, Architectures and Languages (ATAL01)*, pages 1–17, 2001.

[5] L. Amgoud, S. Parsons, and N. Maudet. Arguments, dialogue, and negociation. In *Proceedings of the 4th European Conference on Artificial Intelligence (ECAI00)*, Berlin, 2000.

[6] T. Bench-Capon. Specification and implementation of toulmin dialogue game. In *Proceedings of the JURIX eleventh Conference on Legal Knowledge-Based Systems*, pages 5–19, 1998.

[7] L. Carlson. *Dialogue games: an approach to discourse analysis.* D. Reidel Publishing company, 1983.

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

[9] C. L. Hamblin. *Fallacies.* Methuen, 1970.

[10] R. P. Loui. Process and policy: Resource-bounded non-demonstrative reasoning. *Computational Intelligence*, 14:1–38, 1998.

[11] J. MacKenzie. Question-begging in non-cumulative systems. *Journal of philosophical logic*, 8:117–133, 1979.

[12] N. Maudet and B. Chaib-draa. Commitment-based and dialogue-game based protocols: new trends in agent communication languages. *Knowledge Engineering Review*, 2002. Submitted.

[13] D. Moore. *Dialogue game theory for intelligent tutoring systems.* PhD thesis, Leeds Metropolitan University, England, 1993.

[14] H. Prakken. On dialogue systems with speech acts, arguments, and counterarguments. In *Proceedings of the 7th European Workshop on Logic for Artificial Intelligence (JELIA)*, Lecture Notes in AI 1919, pages 239–253, 2000.

[15] H. Prakken. Modelling defeasibility in law: logic or procedure? *Fundamentae Informaticae*, 2001.

[16] H. Prakken. Relating protocols for dynamic dispute with logics for defeasible argumentation. *Synthese*, 127:187–219, 2001.

[17] C. Reed and D. Long. Content ordering in the generation of persuasive discourse. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI'97)*, pages 1022–1028, 1997.

[18] D.N. Walton and E.C.W. Krabbe. *Commitment in dialogue: basic concepts of interpersonal reasoning.* State University of New York Press, Albany, NY, 1995.