Modelling dialogues using argumentation

Leila Amgoud
Department of Electronic Engineering,
Queen Mary and Westfield College,
University of London,
London E1 4NS,
United Kingdom

Nicolas Maudet
IRIT/ENSEEIHT,
2 rue C. Camichel,
31071 Toulouse Cedex,
France

Simon Parsons
Department of Computer Science,
University of Liverpool,
Chadwick Building,
Liverpool L69 7ZF,
United Kingdom

Abstract

A number of authors have suggested the use of argumentation techniques as the basis for negotiation dialogues between agents. In this paper we augment this work by investigating the use of argumentation as the basis for a wider range of types of dialogue. The approach we take is based upon MacKenzie's dialogue game DC, and we show that a translation of this into our system of argumentation can support a subset of the types of dialogue identified by Walton and Krabbe.

1 Introduction

When building multi-agent systems, we take for granted the fact that the agents which make up the system will need to communicate. They need to communicate in order to resolve differences of opinion and conflicts of interest, work together to resolve dilemmas or find proofs, or simply to inform each other of pertinent facts. Many of these communicate requirements cannot be fulfilled by the exchange of single messages. Instead, the agents concerned need to be able to exchange a sequence of messages which all bear upon the same subject. In other words they need the ability to engage in dialogues.

As a result of this requirement, there has been much work on providing agents with the ability to hold such dialogues. Typically these focus on one type of dialogue, often negotiation [14, 19, 20]. Recently, Reed [17] has suggested a general approach which can capture a range of dialogue types, grounded on work in argumentation theory. Here we build on Reed's work by showing in detail how it is possible to carry out this range of types of dialogues. We take a system of argumentation developed to handle inconsistent information, use this to build a formal model of dialogue, and then show that the latter is general enough to capture the types of dialogue discussed by Reed.

2 The argumentation model

In this section we briefly introduce the system of argumentation which forms the backbone of our approach. This is inspired by the work of Dung [7] but goes further in dealing with preferences between arguments. Further details are available in [1]. We start with a possibly inconsistent knowledge base Σ with no deductive closure. We assume Σ contains formulas of a propositional langage $\mathcal{L}$. $\vdash$ stands for classical inference and $\equiv$ for logical equivalence.

Definition 1 *An* argument *is a pair (H, h) where h is a formula of $\mathcal{L}$ and H a subset of 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 2 *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. We suppose that this ordering derives from the fact that the knowledge base Σ is stratified 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 3 *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).

Example 1 *Let* $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$ *with* $\Sigma_1 = \{\neg a\}, \Sigma_2 = \{a, a \to b\}$ *and* $\Sigma_3 = \{\neg b\}$. *Now,* $(\{\neg a\}, \neg a)$ *and* $(\{a, a \to b\}, b)$ *are two arguments of* $\mathcal{A}(\Sigma)$. *The argument* $(\{\neg a\}, \neg a)$ *undercuts* $(\{a, a \to b\}, b)$. *The preference level of* $\{a, a \to b\}$ *is 2 whereas the preference level of* $\{\neg a\}$ *is 1, and so* $(\{\neg a\}, \neg a\}) \gg^{Pref} (\{a, a \to b\}, b)$.

We can now define the argumentation system we will use:

Definition 4 *An* argumentation system *(AS) is a triple* $\langle \mathcal{A}(\Sigma), Undercut, Pref \rangle$ *such that* $\mathcal{A}(\Sigma)$ *is a set of the arguments built from* Σ, *Undercut is a binary relation representing defeat relationship between arguments, Undercut* $\subseteq \mathcal{A}(\Sigma) \times \mathcal{A}(\Sigma)$, *and Pref is a (partial or complete) preordering on* $\mathcal{A}(\Sigma) \times \mathcal{A}(\Sigma)$. $\gg^{Pref}$ *stands for the strict pre-order associated with Pref.*

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

Definition 5 *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.*
>
> $\mathcal{S}$ defends *A if there is some argument in* $\mathcal{S}$ *which strongly undercuts every argument B where B undercuts A and A cannot defend itself against B.*

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

$$\begin{aligned} \mathcal{S} &\subseteq \mathcal{A}(\Sigma) \\ \mathcal{F}(\mathcal{S}) &= \{(H, h) \in \mathcal{A}(\Sigma) | (H, h) \text{ is defended by } \mathcal{S}\} \end{aligned}$$

Definition 6 *The set of* acceptable *arguments for an argumentation system* $\langle \mathcal{A}(\Sigma), Undercut, Pref \rangle$ *is:*

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

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

Example 2 *(follows Example 1)*
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* $(\{a, a \to b\}, b)$ *and does not defend itself. On the contrary,* $(\{\neg a\}, \neg a)$ *undercuts* $(\{a, a \to b\}, b)$ *and* $(\{\neg a\}, \neg a) \gg^{Pref} (\{a, a \to b\}, b)$. *Therefore,* $C_{Undercut,Pref}$ *defends* $(\{\neg b\}, \neg b)$ *and consequently* $(\{\neg b\}, \neg b) \in \underline{\mathcal{S}}$.

The set of acceptable arguments mutually defend one another:

Definition 7 *Let A, B be two arguments of* $\mathcal{A}(\Sigma)$ *and* $\mathcal{S} \subseteq \mathcal{A}(\Sigma)$, *then A* disqualifies *B iff A strongly undercuts B and B does not strongly undercut A.* $\mathcal{S}$ strictly defends *A iff for all B such that B strongly undercuts A, then there is a* $C \in \mathcal{S}$ *such that C disqualifies B.*

Theorem 1 $\forall (H, h) \in \underline{\mathcal{S}}, \underline{\mathcal{S}}$ *strictly defends* (H, h).

The proof of this theorem can be found in [1].

3 Proof Theory

In practice we don't need to calculate all the acceptable arguments from some Σ in order to know the status of a given argument. This can be exploited to give a proof theory for the system [1]. The proof theory takes the form of a game between two players P_0 and P_1. Player P_0 makes the argument we are interested in and its defenders, and player P_1 makes the counter-arguments or defeaters.

The basic idea behind the proof theory is to traverse the sequence $\mathcal{F}_1, ..., \mathcal{F}_n$ in reverse. Consider that A occurs for the first time in $\mathcal{F}_n$. We start with A, and then for any argument B_i which strongly undercuts A, we find an argument A_i in $\mathcal{F}_{n-1}$ which defends A. Now, because of Theorem 1, we are only interested in the strict defenders of an argument, and the strict defenders of A will disqualify the B_i. The same process is repeated for each strict defender until there is no strict defender or defeater. This leads to the idea of the argument dialogue:

Definition 8 *An* argument dialogue[1] *is a nonempty sequence of moves,* $move_i = (Player_i, Arg_i)$, $i \geq 0$, *such that:*

1. $Player_i = P_0$ *iff i is even,* $Player_i = P_1$ *iff i is odd.*
2. $Player_0 = P_0$ *and* $Arg_0 = A$.
3. *If* $Player_i = Player_j = P_0$ *and* $i \neq j$ *then* $Arg_i \neq Arg_j$.
4. *If* $Player_i = P_0$, $i > 1$, *then* Arg_i *disqualifies* Arg_{i-1}.

[1] In [1] this is called simply a "dialogue"; here we use the term "argument dialogue" to dinstinguish these dialogues from those discussed later in the paper. We will omit the term "argument" when it is clear that we mean an argument dialogue.

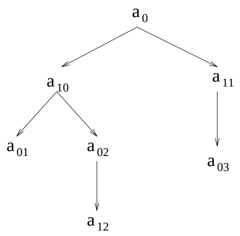

Figure 1. An argument dialogue tree

5. *If* $Player_i = P_1$ *then* Arg_i *attacks* Arg_{i-1}.

An argument dialogue tree *is a finite tree where each branch is an argument dialogue.*

Example 3 *Let* $\langle \mathcal{A}, R, Pref \rangle$ *be an AS such that* $\mathcal{A} = \{a_0, a_{01}, a_{02}, a_{10}, a_{11}, a_{12}\}$, *and Undercut* $= \{(a_{10}, a_0), (a_{01}, a_{10}), (a_{12}, a_{02}), (a_{02}, a_{10}), (a_{03}, a_{11}), (a_{11}, a_0)\}$.
Let's suppose

$$a_{03} \gg^{Pref} a_{11} \gg^{Pref} a_0$$

$$a_{01} \gg^{Pref} a_{10} \gg^{Pref} a_0$$

and

$$a_{12} \gg^{Pref} a_{02} \gg^{Pref} a_{10}$$

We are interested in the status of the argument a_0. *The corresponding argument dialogue tree is presented in Figure 1.*

The dialogue tree can be considered as an AND/OR tree. A node corresponding to player P_0 is an AND node, and a node corresponding to player P_1 is an OR node. That distinction between nodes is due to the fact that an argument is acceptable if it is defended against all its defeaters. The edges of a node containing an argument of P_0 represent defeaters so they all must be defeated. In contrast, the edges of a node containing an argument from P_1 represent defenders of P_0 so it is sufficient that one of them defeats the argument of P_1.

Definition 9 *A player* wins *an argument dialogue iff he ends the dialogue (he makes the last argument).*

A player who wins a dialogue does not necessarily win in all the sub-trees of the dialogue tree. To formalize the winning of a dialogue tree we use the concept of a solution sub-tree.

Definition 10 *A* candidate sub-tree *is a sub-tree of an argument dialogue tree containing all the edges of each AND node and exactly one edge of each OR node. A* solution sub-tree *is a candidate sub-tree whose branches are all won by* P_0.

Thus the dialogue represented in Example 3 has exactly two candidate sub-trees: S_1 and S_2 in Figure 2.

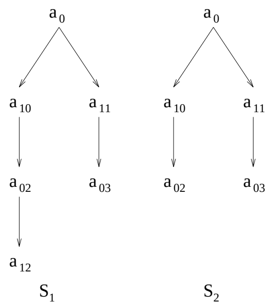

Figure 2. Sub-trees

Definition 11 P_0 wins *an argument dialogue iff the corresponding dialogue tree has a solution sub-tree.*

Thus P_0 wins the dialogue presented in Figure 1 because S_2 is a solution sub-tree.

Definition 12 *An argument A is* justified *iff there is an argument dialogue tree whose root is A, and which is won by player* P_0.

Thus the argument a_0 is justified because the player P_0 won the dialogue tree. The main result from the proof theory is:

Theorem 2 *Let* $\langle \mathcal{A}, R, Pref \rangle$ *be an argumentation system.*

- $\forall x \in \mathcal{A}$, *if x is justified then each argument made by* P_0 *which belongs to the solution sub-tree is in* $\underline{\mathcal{S}}$, *in particular x.*
- $\forall x \in \underline{\mathcal{S}}$, *x is justified.*

The proof may be found in [1].

In other words, the dialogue process constructs all acceptable arguments, and only constructs acceptable arguments. Thus Theorem 2 is a form of soundness and completeness result for the proof theory.

4 Arguments and dialogue games

Dialogue games are way of formally analysing discourse [6]. A dialogue between two individuals is seen as a game in which each individual has objectives and a set of legal moves which can be used to obtain those objectives. The moves are illocutions, and the objectives are matters such as persuading the other player of the truth of a proposition.

4.1 Dialogues in DC

One rather influential dialogue game is DC, proposed by MacKenzie [9][2] in the course of analysing the fallacy

[2]Though he describes is as a "dialectical system" rather than a dialogue game.

of question-begging. DC provides a set of rules for arguing about the truth of a proposition. Each player has the goal of convincing the other, and can assert or retract facts, challenge the other player's assertions, ask whether something is true or not, and demand that inconsistencies be resolved. Associated with each player is a "commitment store", which holds the statements players have made and the challenges they have issued. There are then rules which define how the commitment stores are updated and whether particular illocutions can be uttered at a particular time.

While DC is interesting in its own right, what is more interesting from the point of view of this paper is that DC ties in very neatly with the system of argumentation descibed above. As detailed in [3], it is possible to formulate the dialogue rules in terms of the arguments that each player can construct. This, gives an operational semantics to the system and, given an implementation of the argumentation system, makes it possible to build systems that can carry out DC-type dialogues. One reason we might want to do this, is to use these DC-style dialogues as the basis of communications between agents in a multi-agent system, and the rest of this paper explores some of the issues in doing so.

Dialogues are assumed to take place between two agents, P and C, where P is arguing "pro" some proposition, and C argues "con". Each player has a knowledge base, Σ_P and Σ_C respectively, containing their beliefs. In addition, and as in DC, we suppose that each player has a further knowledge base, accessible to both players, containing commitments made in the dialogue. These commitment stores are denoted $CS(P)$ and $CS(C)$ respectively. Note that the union of the commitment stores can be viewed as the state of the dialogue at turn t. All the bases described contain propositional formulas and are not closed under deduction, and all are stratified according to degree of belief as discussed above. Here we assume that these degrees of belief are static and that both the players agree on them, though it is possible [4] to combine different sets of degrees of belief, and it is also possible to have agents modify their beliefs on the basis of the reliability of their acquaintances [13].

Both players are equipped with an argumentation system of the kind discussed above. Each has access to their own private knowledge base and both commitment stores. Thus Player P plays with $\langle \mathcal{A}(\Sigma_P \cup CS(P) \cup CS(C)), Undercut, Pref \rangle$ and player C with $\langle \mathcal{A}(\Sigma_C \cup CS(C) \cup CS(P), Undercut, Pref \rangle$. The two argumentation systems are then used to help players to maintain the coherence of their beliefs[3], and thus to avoid asserting things which are defeated by other knowledge from $CS(P) \cup CS(C)$, and which could thus easily be undercut by the other player. In this sense the argumentation systems help to ensure that players are *rational*.

[3]Note that the players' data is allowed to be inconsistent in the classical sense, since the argumentation system can handle inconsistency.

4.2 The dialogue moves

With this background, we can present the set of dialogue moves that we will use. These are not an exact facsimile of the moves in DC—we omit those illocutions we have not found useful (in particular the retraction rules) and add a move for explicit acknowledgment. For each move, we give what we call rationality rules, dialogue rules, and update rules. These are based on the rules suggested by [10]. The rationality rules specify the preconditions for playing the move. The update rules specify how commitment stores are modified by the move. The dialogue rules specify the moves the other player can make next, and so specify the *protocol* under which the dialogue takes place.

In the following, player P adresses the move to player C. We start with the assertion of facts:

assert(p) where p is a propositional formula.

rationality the player uses its AS to check if there is an acceptable argument for the fact p.

dialogue the other player can respond with:

1. $accept(p)$,
2. $assert(\neg p)$,
3. $challenge(p)$.

update $CS_i(P) = CS_{i-1}(P) \cup \{p\}$ and $CS_i(C) = CS_{i-1}(C)$

assert(S) where S is a set of formulas representing the support of an argument. Note that in DC, players can only assert one propositional formula.

rationality the player uses the AS to check if the related argument is acceptable.

dialogue the other player can play:

1. $accept(S)$,
2. $assert(\neg p)$,
3. $challenge(p)$ where $p \in S$.

Informally, this means that the player can accept the whole support or challenge/deny an element of the support.

update $CS_i(P) = CS_{i-1} \cup S$ and $CS_i(C) = CS_{i-1}(C)$

The conditions on assertion only allow agents to assert facts which it believes, on the basis of its private knowledge and all the public knowledge, cannot be challenged. The counterpart of these moves are the acceptance moves:

accept(p) p is a propositional formula. This move has no equivalent in DC where acceptance is implicit.

rationality the player uses the AS to check if there is an acceptable argument for p.

dialogue the other player can play any allowed moves.

update $CS_i(P) = CS_{i-1}(P) \cup \{p\}$ and $CS_i(C) = CS_{i-1}(C)$

accept(S) S is a set of propositional formulas.

rationality the player uses his AS to check if each element of S is supported by an acceptable argument.

dialogue the other player can play any allowed move.

update $CS_i(P) = CS_{i-1}(P) \cup S$ and $CS_i(C) = CS_{i-1}(C)$

Thus a player can only accept something if it is not possible to build a stronger argument against it.

challenge(p) where p is a propositional formula.

rationality $\emptyset$

dialogue the other player can only $assert(S)$ where S is an argument supportting p.

update $CS_i(P) = CS_{i-1}(P)$ and $CS_i(C) = CS_{i-1}(C)$

A challenge is a means of making the other player explicitly state the argument supporting a proposition. In contrast, a question can be used to query the other player about any proposition.

question(p) where p is a propositional formula.

rationality $\emptyset$

dialogue the other player can:

1. $assert(p)$,
2. $assert(\neg p)$,
3. $question(q)$.

update $CS_i(P) = CS_{i-1}(P)$ and $CS_i(C) = CS_{i-1}(C)$

We refer to this set of moves as the set $\mathcal{M}_{\mathcal{DC}}$. These are similar to those discussed in legal reasoning [8, 16].

4.3 How the moves make a dialogue

The basic kind of dialogue which this system allows us to model is that in which one agent believes p and the other believes $\neg p$. The private aim of each agent is then simply to offer arguments in support of their thesis. In our model, this means that one agent has an acceptable argument in favour of p in its argumentation system and the other agent has an acceptable one in favour of $\neg p$. The agents then exchange arguments (defeaters) in an attempt to find an acceptable argument for whichever of p and $\neg p$ they started believing.

In this dialogue, P has an acceptable argument (S, p) in favour of p, which it advances. This argument is built using private information, and, because C does have access to it, C does not have an acceptable argument for p and so cannot accept the proposition. Instead C has an acceptable argument $(S', \neg p)$ in favour of $\neg p$, and it asserts $\neg p$. P responds to this assertion by challenging $\neg p$, and C rises to the challenge by asserting its support S' for $\neg p$. This dialogue process is continued, for example by P challenging one of the steps in S', until there are no further arguments relating to p, $\neg p$, S, and S'. This is the case when the two argumentation frameworks:

$$\langle \mathcal{A}(\Sigma_P \cup CS(P) \cup CS(C)), Undercut, Pref \rangle$$

and

$$\langle \mathcal{A}(\Sigma_C \cup CS(P) \cup CS(C)), Undercut, Pref \rangle$$

provide the same acceptable argument and the agents, perforce, agree. Thus the agents have resolved their initial conflict of opinions about the acceptability of p. If the agents do not agree, then they are able to make further arguments and keep the dialogue going.

As mentioned above, the rationality rules and dialogue rules together provide a form of protocol for carrying out these DC-style dialogues. The rationality rules tie the illocutions which can be made quite closely to the information that the agents have at their disposal. The fact that it is only possible for an agent to assert things for which it has an acceptable argument, and to only accept things for which it has an acceptable argument means that Theorem 2 can be carried forward to ensure the soundness of the dialogue. In particular we can prove:

Theorem 3 *Given two players P, with $AS_P = \langle \mathcal{A}(\Sigma_P \cup CS(P) \cup CS(C)), Undercut, Pref \rangle$, and C with $AS_C = \langle \mathcal{A}(\Sigma_P \cup CS(P) \cup CS(C)), Undercut, Pref \rangle$, which play a dialogue game using moves $\mathcal{M}_{\mathcal{DC}}$, then if S is the set of all arguments which the game can possibly generate,*

- *$\forall x \in S$, x is a justified argument of either AS_P or AS_C;*
- *If $x \in S$ is a justified argument of AS_P, it is a justfied argument of AS_C*

Type of Dialogue	Initial Situation	Participant's Goal	Goal of Dialogue
Persuasion	Conflict of opinions	Persuade other party	Resolve or clarify issue
Inquiry	Need to have proof	Find and verify evidence	Prove (disprove) hypothesis
Negotiation	Conflict of interests	Get what you most want	Reasonable settlement
Information seeking	One party lacks information	Acquire or give information	Exchange information
Deliberation	Dilemma or practical choice	Co-ordinate goals or actions	Decide best course of action

Table 1. Walton and Krabbe's classification of dialogues

Note that this result does mean that agents have to tell the truth; they can lie as long as their mechanism for generating untruths allows suitable acceptable arguments to be generated.

Despite this result, the protocol defined by the dialogue rules will only work under certain assumptions. In fact, without these assumptions it can easily descend into meaningless babble. For example, there is no reason why agents should not, under the protocol, carry out infinite dialogues which consist entirely of "accept" illocutions or repeated challenges of the same proposition. The assumptions that the protocol works under are, broadly speaking, equivalent to the usual Gricean maxim of relevance. We assume that agents have a focus to their dialogue and restrict their illocutions to those which address that focus, engage in one for a concrete reason (for example to persuade one to change its mind about giving up a resource) not simply for the sake of it, and both recognise when they have reached the useful end of a dialogue (even if it is a case where neither has persuaded the other) and stop when this point has been reached. In practice, we see these assumptions being enforced by the fact that agents engaging in these kind of dialogues will be part of some kind of electronic institution [12] which commits them to certain rules of behaviour.

For now we leave the formalisation of these points for future work. Instead we consider how the basic dialogue moves can be used to capture different types of dialogue.

5 Different types of dialogue

As mentioned by Parsons *et al.* [15] and discussed in detail by Reed [17], Walton and Krabbe [21] have identified a set of types of dialogue, distinguished by initial situation, goal of participants, and goal of dialogue. These are summarised in Table 1. In this section we discuss how these types of dialogue may be captured in our model of DC. We deal first with persuasion, inquiry and information seeking, which can be captured by our model of DC directly.

A *persuasion* dialogue, according to Walton and Krabbe, is 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. The dialogue continues until the dispute is resolved. This is clearly the kind of dialogue discussed in the previous section, and can therefore easily be captured in our system.

An *inquiry* dialogue does not start from conflict but from a lack of knowledge. The two agents will try to establish the truth or falsity of some proposition p and the dialogue will end when either this has been acheived or they realise they cannot find a proof. In our model this corresponds to the situation in which neither agent has an acceptable argument for p. In other words, the initial situation is that neither:

$$\langle \mathcal{A}(\Sigma_P \cup CS(P) \cup CS(C)), \mathit{Undercut}, \mathit{Pref} \rangle$$

nor

$$\langle \mathcal{A}(\Sigma_C \cup CS(P) \cup CS(C)), \mathit{Undercut}, \mathit{Pref} \rangle$$

can provide an acceptable argument for p. The agents then engage in a DC dialogue with the aim of determining whether the system

$$\langle \mathcal{A}(\Sigma_P \cup \Sigma_C \cup CS(P) \cup CS(C)), \mathit{Undercut}, \mathit{Pref} \rangle$$

can provide an acceptable argument for p. The dialogue, during which both agents reveal information by asserting it into their commitment store, will continue until either an acceptable argument for p is found, or it is not possible to make any further arguments which are related to p. This achieves the aim of the inquiry dialogue. Thus, although DC has its roots in adversarial dialogue, the moves $\mathcal{M}_{\mathcal{DC}}$ can be used in a non-adversarial way.

Dialogue type	Agent P	Agent C
Persuasion	$\exists(S, p) \in \underline{S}_P$	$\exists(S, \neg p) \in \underline{S}_C$
Inquiry	$\nexists(S, p) \in \underline{S}_P$	$\nexists(S, p) \in \underline{S}_C$
Info-seeking	$\nexists(S, p) \in A(\Sigma_P \cup CS(P) \cup CS(C))$	$\exists(S, p) \in \underline{S}_C$

Table 2. Initial conditions for different types of dialogue about p

Information seeking dialogues are similar to inquiries, but differ in their initial conditions. An information seeking dialogue is initiated when there is an asymmetry between the agents in the sense that one is thought by the other to have more information in regard to *p*, for instance because one agent is a recognised authority on the subject. In our model this kind of dialogue is initiated with a *question* move, asking if it is the case that *p* holds. If the other agent has an argument for or against *p* it will assert this, and the agents will then argue about its acceptability. However the argument is resolved, this exchange of information achieves the aim of the information seeking dialogue. Table 2 summarises the initial conditions for these three types of dialogue in our framework, and highlights the differences between them. This table, and our previous discussion allow us to prove:

Theorem 4 *The set of moves $\mathcal{M}_{\mathcal{DC}}$ is sufficient to model persuasion, inquiry and information seeking dialogues.*

The two remaining dialogues are slightly different. As discussed in [15], negotiation centres around conflicts between the intentions of agents—it is concerned with what agents intend to achieve. Thus to have a negotiation dialogue in our framework it is first necessary to extend the language in which the dialogue is carried out to distinguish between the beliefs which are the subject of persuasion, information seeking and inquiry dialogues and the intentions which are the subject of negotiations. The change to the language then necessitates some minor changes to the argumentation system (essentially defining which arguments undercut which others). This may be done in exactly the same was as in [15], and afterwards negotiation dialogues have exactly the same form as persuasion dialogues but concern intentions rather than beliefs. In other words, they arise from a conflict in intentions, proceed through agents making arguments about intentions, and terminate when the agents agree on an acceptable argument for the intention in question.

A deliberation dialogue represents the process of forming a plan of action, and thus is also concerned with intentions (since an agent's intentions precisely deal with its plans of action). The joint aim of a deliberation is to reach an agreement on a plan (a joint intention or a pair of individual intentions), and the individual aims are to influence this agreement to their benefit (to ensure that the intentions are in their favour). Thus the joint aims can be considered to be establishing an acceptable argument for the joint or individual intentions[4], and the individual aims are satisfied by the fact that this is acceptable to both agents. Note that deliberation starts not from a point of conflict but simply from a need for action, in exactly the same way that an inquiry dialogue starts, and, once the underlying language is extended to deal with intentions, precisely the same process can be followed as outlined above for the inquiry dialogue.

This discussion leads us to believe that $\mathcal{M}_{\mathcal{DC}}$ will be suitable as a basis for negotiation and deliberation dialogues. However to properly capture such dialogues $\mathcal{M}_{\mathcal{DC}}$ needs to be augmented, in particular to model compromises, and in [5] we discuss how to do this.

6 Conclusion

In this paper we have presented a model for inter-agent dialogues based on argumentation. We believe that this dialogue model goes further than previous attempts based on argumentation. In particular, it is a more general model than that presented in [15], and has a precisely defined protocol for the exchange of arguments which the former lacks (since [15] concentrates more on the interaction between beliefs and intentions in a specific negotiation/deliberation form of dialogue). Our model also extends the suggestion made by Reed [17] by making Reed's suggestion more concrete—providing an underlying argumentation system and the illocutions necessary to carry out the kinds of dialogues Reed discusses. Thus this work can be seen as an attempt to bridge the gap between the low level detail of handling beliefs and intentions described in [15] and the general approach of [17]. This was certainly our motivation in undertaking this work and, though it should be acknowledged that more work is required to combine the approaches, we feel it makes significant progress towards achieving this aim.

The most obvious thing we need to do now is to extend the underlying argumentation system to handle mental attitudes, in particular beliefs, desires and intentions. As argued above, this is an important step towards capturing negotiation and deliberation dialogues, though taking the approach of [15] it should be straightforward. Another obvious thing is to extend the mechanism to deal with multiparty dialogues possibly through the use of agoras [11].

[4]Thus we agree with Reed [17] that the dialogue in [15] has elements of deliberation as well as negotiation.

Other important points relate to how this work on dialogue connects with wider issues of agent communication. One such issue is what strategy an agent uses to pick which argument to put forward. One simple idea would be to choose the smallest argument in order to restrict the exposure to defeaters, but there are other aspects which might equally play a part. Another issue is how the argumentation protocol we have discussed connects with the protocol for the communication exchange in which the argumentation is embedded—if we assume that agents use argumentation for part of their communication (as for example in [18]) then agents need to know when to engage in argumentation and when to stop. Finally, we need to define the kinds of electronic institutions within which agents can engage in argumentative dialogues.

References

[1] L. Amgoud. *Contribution a l'integration des préferences dans le raisonnement argumentatif*. PhD thesis, Université Paul Sabatier, Toulouse, July 1999.

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

[3] L. Amgoud and N. Maudet. Towards a dialogue system based on argumentation. Technical report, Department of Electronic Engineering, Queen Mary and Westfield College, 1999.

[4] L. Amgoud and S. Parsons. An argumentation framework for merging conflicting knowledge bases. Technical report, Department of Electronic Engineering, Queen Mary and Westfield College, 2000.

[5] L. Amgoud, S. Parsons, and N. Maudet. Arguments, dialogue, and negotiation. In *Proceedings of the 14th European Conference on Artificial Intelligence*, 2000.

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

[7] 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.

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

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

[10] N. Maudet and F. Evrard. A generic framework for dialogue game implementation. In *Proceedings of the 2nd Workshop on Formal Semantics and Pragmatics of Dialogue*, University of Twente, The Netherlands, May 1998.

[11] P. McBurney and S. Parsons. Risk agoras: dialectical argumentation for scientific reasoning. In *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, 2000.

[12] P. Noriega. *Agent mediated auctions: the fishmarket metaphor*. PhD thesis, Universitat Autònoma de Barcelona, Bellaterra, December 1997.

[13] S. Parsons and P. Giorgini. An approach to using degrees of belief in BDI agents. In B. Bouchon-Meunier, R. R. Yager, and L. A. Zadeh, editors, *Information, Uncertainty, Fusion*. Kluwer, Dordrecht, 1999.

[14] S. Parsons and N. R. Jennings. Negotiation through argumentation—a preliminary report. In *Proceedings of the 2nd International Conference on Multi Agent Systems*, pages 267–274, 1996.

[15] 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.

[16] H. Prakken. Relating protocols for dynamic dispute with logics for defeasible argumentation. *Synthese*, (to appear).

[17] C. Reed. Dialogue frames in agent communication. In *Proceedings of the 3rd International Conference on Multi Agent Systems*, pages 246–253, 1998.

[18] C. Sierra, N. R. Jennings, P. Noriega, and S. Parsons. A framework for argumentation-based negotiation. In *Proceedings of the 4th International Workshop on Agent Theories, Architectures and Languages*, pages 167–182, 1997.

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

[20] F. Tohmé. Negotiation and defeasible reasons for choice. In *Proceedings of the Stanford Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning*, pages 95–102, 1997.

[21] 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.