Computational Representation of Persuasive Argument

Katie Atkinson, Trevor Bench-Capon and Peter McBurney
Department of Computer Science
University of Liverpool
Liverpool L69 3BX
UK
{k.m.atkinson,tbc,p.j.mcburney}@csc.liv.ac.uk

Abstract

In this report we discuss the typology of dialogues given by Walton and Krabbe and offer a precise interpretation of them. We go on to discuss one of these dialogue types - persuasion - in the context of practical reasoning and the problems associated with such reasoning. We propose a perspective on practical reasoning as presumptive justification and critical questions, giving an extension to the account proposed by Walton [24]. This provides us with a foundation for a protocol, named PARMA, for a dialogue game based on this theory. We go on to give an axiomatic and denotational semantics for PARMA and discuss two implementations based on PARMA.

1 Introduction

We have previously offered an account of persuasion in practical reasoning, based upon the intention of one party to persuade another that an action is justified [6]. We now go on to present a development of this work in the form of the specification, implementation and evaluation of a dialogue game in which participants aim to persuade each other, according to our theory. Our approach takes as its starting point an influential typology of human dialogues, due to Walton and Krabbe [23], which we summarise in section 2. Section 3 focuses on the topic of practical reasoning and why this is of particular importance in persuasion dialogues. It offers an account of practical reasoning which builds on that of [24]. Section 4 presents our general theory of persuasion over action, which comprises the statement of a position and a list of subsequent attacks which can be made on this position. Section 5 presents a dialogue game protocol named PARMA. We have specified a syntax, an axiomatic and a denotational semantics to allow a multi-party dialogue game to be conducted, in accordance with our theory. Section 6 describes an implementation in Java of the dialogue game, in accordance with the PARMA protocol. Section 7 presents an example application of the dialogue game in use. Section 8 provides an evaluation and summary of issues which arose from the

Java implementation. Section 9 goes on to describe an alternative implementation of the argument scheme which addresses the issues occurring in the Java implementation. Finally, section 10 offers a summary of the main points of the paper.

2 Background

Dialogue exchanges can be categorised as falling into a number of distinct dialogue types. A well known typology of distinct human dialogue types has been identified by Walton and Krabbe in [23]. Here we will recapitulate the work of Walton and Krabbe [23] on dialogue types, and present some additional analysis of our own which is intended to make these notions more precise and readily applicable. In particular we view one of these types, persuasion, as being of significant importance for use in multi agent dialogue games and we will go on to focus on this particular dialogue type in more detail in later sections of this paper.

2.1 Types of Dialogue

In [23], Walton and Krabbe have identified a number of distinct dialogue types used in human communication: Persuasion, Negotiation, Inquiry, Information-Seeking, Deliberation, and Eristic Dialogues. These types are characterised by their initial positions, main goal and the aims of the participants. These dialogue types are summarised in Table 1.

Table 1: **Types of Dialogue**

Type	Initial Situation	Main Goal	Participants Aims
Persuasion	Conflicting points of view	Resolution of such conflicts by verbal means	Persuade the other(s)
Negotiation	Conflict of interests and need for cooperation	Making a deal	Get the best out of it for oneself
Inquiry	General ignorance	Growth of knowledge and agreement	Find a proof or destroy one
Info-seeking	Personal ignorance	Spreading knowledge and revealing positions	Gain, pass on, show or hide personal knowledge
Deliberation	Need for action	Reach a decision	Influence the outcome
Eristic Dialogue	Conflict and antagonism	Reaching an accommodation in a relationship	Strike the other party and win in the eyes of onlookers

We summarize the Walton and Krabbe descriptions as follows (in the order of [23]):

- A **Persuasion** dialogue involves an attempt by one participant to have another participant endorse some proposition or statement. The statement at issue may

concern the beliefs of the participants or proposals for action, and the dialogue may or may not involve conflict between the participants. If the participants are guided only by the force of argument, then whichever participant has the more convincing argument, taking into account the burden of proof, should be able to persuade the other to endorse the statement at issue, or to give up the attempt.

- A **Negotiation** dialogue occurs when two or more parties attempt to jointly divide some resource (which may include the participants' own time or their respective capabilities to act), where the competing claims of the participants potentially cannot all be satisfied simultaneously. Here, co-operation is required by both parties in order to engage in the negotiation dialogue, but, at the same time, each participant is assumed to be seeking to achieve the best possible deal for him or herself.

- An **Inquiry** dialogue occurs when two or more participants, each being ignorant of the answer to some question, and each believing the others to be ignorant also, jointly seek to determine the answer. These dialogues do not start from a position of conflict, as no participant has taken a particular position on the question at issue; they are trying to find out some knowledge, and no one need resile from their existing beliefs. Aircraft disaster investigations may be seen as examples of Inquiry dialogues.

- An **Information-seeking** dialogue occurs when one party does not know the answer to some question, and believes (perhaps erroneously) that another party does so. The first party seeks to elicit the answer from the second by means of the dialogue. Expert consultation is a common important subtype of this type of dialogue. When the information sought concerns an action or course of action, we call this type of dialogue, a **plan-seeking** dialogue.

- A **Deliberation** dialogue occurs when two or more parties attempt to agree on an action, or a course of action, in some situation. The action may be performed by one or more the parties in the dialogue or by others not present. Here the participants share a responsibility to decide the action(s) to be undertaken in the circumstances, or, at least, they share a willingness to discuss whether they have such a shared responsibility.

- An **Eristic** dialogue is one where the participants vent perceived grievances, as in a quarrel, and the dialogue may act as a substitute for physical fighting. We do not consider this dialogue type further in this paper as we see it being beyond rational discourse.

Most human dialogues are in fact mixtures or combinations of these ideal types. For example a debate may contain persuasion, information-seeking and antagonism all at once, each embedded in the larger interaction. Moreover a dialogue may shift between types as it proceeds. With the exception of eristic dialogues, we have taken the above dialogue types as a starting point, and given a more precise characterisation to them. This is done using the initial beliefs and aims of the participants and the ways in which

these can change in the course of the dialogue. This allows us to identify any shifts in the dialogue type, and the changes which the parties can make to reach agreement.

Tables 2, 3 and 4 show our analysis for three typical situations. Table 2 shows the possibilities where two parties discuss their beliefs regarding a single proposition. Table 3 shows the possibilities when two parties discuss whether a particular action should be performed or not. Table 4 shows the situation where two parties discuss the performance of either, both or neither of two actions, which may be performed.

Table 2: **Model of a Discussion Over Beliefs**

A/B	B believes p	B believes ¬ p	B believes neither p nor ¬ p
A believes p	Agreement	Disagreement or Persuasion	B info seeks A
A believes ¬ p	Disagreement or Persuasion	Agreement	B info seeks A
A believes neither p nor ¬ p	A info seeks B	A info seeks B	Inquiry

Table 3: **Model of a Discussion Over Actions**

A wants/B wants	B does p	B does ¬ p	B does p or ¬ p
A does p	Agreement	Disagreement or Persuasion	B plan seeks A
A does ¬ p	Disagreement or Persuasion	Agreement	B plan seeks A
A does p or ¬ p	A plan seeks B	A plan seeks B	Deliberation

These tables model the space of all possible initial positions for the participants appropriate to these situations, and indicate which types of dialogue can arise from them. Representing the dialogues in this way leads to a number of observations relating to reaching agreement:

- we can see the space of possible moves available to the participants;
- we can see how agreement can be reached;
- we can see how many changes are needed if agreement is to be reached;
- we can see which participant must change if agreement is to be reached.

We can now apply this classification to some example queries.

1. In (1), the inquirer (B) does not know a piece of information, and the recipient (A) does. Thus, we are at the top right of table 2 and have an info-seeking dialogue.

Table 4: **Model of a Discussion Over Multiple Actions**

A wants/B wants	A does p and q	A does p and ¬ q	A does q and ¬ p	A does ¬ p and ¬ q	no opinion
A does p and q	Agreement	Conflict or Persuasion	Conflict or Persuasion	Conflict or Negotiation	Plan seeking
A does p and ¬ q	Conflict or Persuasion	Agreement	Conflict or Persuasion	Conflict or Persuasion	Plan seeking
A does q and ¬ p	Conflict or Persuasion	Conflict or Persuasion	Agreement	Conflict or Persuasion	Plan seeking
A does ¬ p and ¬ q	Conflict or Negotiation	Conflict or Persuasion	Conflict or Persuasion	Agreement	Plan seeking
no opinion	Plan seeking	Plan seeking	Plan seeking	Plan seeking	Deliberation

2. In (2), the inquirer(A) wishes to know whether or not to perform an action, and (B) will have the answer. This puts us in one of the first two cells at the bottom of table 3, as here we have plan seeking, a sub-type of info-seeking.

3. (3) is similar to (1), even though it is of a general nature.

4. In (4), the recipient (A) did p, but the inquirer (B) believes that ¬ p should have been done, giving rise to a situation of disagreement, requiring persuasion, as in Table 3.

5. In (5), we assume that the recipient's (A) policy is ¬ p and the inquirer (B) wants p to be done so again, this gives rise to persuasion, again as in Table 3.

6. Finally, in (6), the recipient (A) is currently performing an action p, where the inquirer (B) wishes for ¬ p to be performed instead so, we are in the third cell of the second row in table 4, again giving rise to a persuasion dialogue.

We believe that providing the information in the form of these matrices gives a more structured and precise characterisation of the dialogue types than the informal descriptions of [23]. When the participants have a clear understanding of the gaps between their positions the task of deciding what shifts in position they should try to induce, or may need to make, is facilitated. Of course, whether a party is willing to change their position will depend on their other beliefs, and the utility they ascribe to actions and the states resulting from action. The structures, however, do provide a basis for forming strategies and heuristics to inform the conduct of the various types of dialogue.

3 Practical Reasoning

3.1 Introduction to Practical Reasoning

Although many of the arguments that are deployed in everyday life are concerned with what it is sensible or practical to do, the topic of practical reasoning has been rather neglected by philosophers. Practical reasoning has, of course, been addressed (see, e.g.. [18] for a collection of essays and [21] for a recent monograph), but it has received nothing like the attention that has been paid to reasoning about beliefs. When action has been considered, it has most often been in the context of ethics, considering what is morally right or wrong, rather than what is prudentially or practically useful. It has been similarly neglected in Computer Science, where practical reasoning has been treated as little different from deduction, standard backward chaining techniques being applied to rules with goals as consequents and preconditions and actions as antecedents. In this section we first discuss some of the differences between reasoning about belief and reasoning about action which cause problems with approaches based on the practical syllogism. We then discuss the treatment of practical reasoning in [24] which makes use of an argumentation scheme and associated critical questions. We elaborate this scheme and extend the critical questions, before discussing how these critical questions appear in natural argument.

3.2 Problems With The Practical Syllogism

Practical reasoning in computer science can predominately be seen as based on a form of the practical syllogism. An example from [11] is:

K1 I'm to be in London at 4.15.
If I catch the 2.30, I'll be in London at 4.15
So, I'll catch the 2.30

This, however, cannot be quite right. It may well be possible to accept both the premises and deny the conclusion. There are at least three bases for criticism:

C1 K1 represents a species of abduction, and so there may be alternative ways of achieving the goal.

C2 Performing an action typically excludes the performance of other actions, which might have other desirable results; these may be more desirable than the stated goal.

C3 Performing an action typically has a number of consequences. If some of these are undesirable, they may be sufficiently bad to lead us to abandon the goal.

In order to act on the basis of an argument such as K1 therefore, we need to consider alternative actions, alternative goals and any additional consequences, and then choose the best of these alternatives. Note the element of choice: we can choose our goals and actions in a way in which we cannot choose our beliefs, and different people may rationally make different choices, as Searle puts it:

”Assume universally valid and accepted standards of rationality, assume perfectly rational agents operating with perfect information, and you will find that rational disagreement will still occur; because, for example, the rational agents are likely to have different and inconsistent values and interests, each of which may be rationally acceptable.” (Searle 2001, xv)

In a sense therefore any practical argument is directed to a specific person at a specific time, to encourage them towards a particular choice and the objectivity that we can find in factual matters cannot in general be attained in practical reasoning. An attempt to modify K1, similar to one put forward by Searle in [21] (although not regarded by him as satisfactory) is:

S1 I want, all things considered, to achieve E
The best way, all things considered, to achieve E is to do M
So, I will do M.

There are problems with this: we cannot in general consider all things, because we have limited reasoning resources and imperfect information. Nor is it easy to say what is meant by ”best” here. In computer science there are often attempts to define best using some kind of utility function but, as Searle points out, any preference ordering is more often the *product* of practical reasoning than an input to it. Coming to understand what we think is best is part of what we do in practical reasoning.

One way of addressing these problems is to regard practical reasoning as a species of presumptive argument. Given an argument like K1, we have a presumptive reason for performing the action. This presumption can, however, be challenged and withdrawn. Subjecting our argument to appropriate challenges is how we hope to identify and consider the alternatives that require consideration, and determine the best choice for us, in the particular context.

One account of presumptive reasoning is in terms of argument schemes and critical questions, as given in [24]. The idea here is that an argument scheme gives a presumption in favour of its conclusion. Whether this presumption stands or falls depends on satisfactory answers being given to the critical questions associated with the scheme.

3.3 Argument Schemes for Practical Reasoning

In [24] Walton gives two schemes for practical reasoning: the necessary condition scheme:

W1 G is a goal for a
Doing A is necessary for a to carry out G
Therefore a ought to do A.

and the sufficient condition scheme:

W2 G is a goal for a
Doing A is sufficient for a to carry out G

Therefore a ought to do A.

He associates with them four critical questions:

CQ1 Are there alternative ways of realising G?
CQ2 Is it possible to do A?
CQ3 Does a have goals other than G which should be taken into account?
CQ4 Are there other consequences of doing A which should be taken into account?

Here we will consider only W2: W1 is a special case in which CQ1 is answered in the negative. CQ1, CQ3 and CQ4 relate respectively to the criticisms C1, C2 and C3 identified above.

We believe, however, that the argument scheme, and the critical questions need elaboration. Firstly, we believe that the notion of a goal is ambiguous.

Consider the following situation. I am in Liverpool. My friend X is about to go to Australia indefinitely, and I am eager to say farewell to him. To catch him before he leaves London, it is necessary that I arrive in London before 4.30. So I may say:

AS1 I want to be in London before 4.30
The 2.30 train arrives in London at 4.15
So, I shall catch the 2.30 train.

Here I am justifying my action in terms of one of its consequences. Alternatively I may say:

AS2 I want to see X before he leaves London
The 2.30 train arrives in London at 4.15
So, I shall catch the 2.30 train.

Here the action is not justified by its direct consequences, but by something else that follows from it. I do not really desire to be in London at all, except in so far as it is a means to the end of seeing X. Alternatively there is a third justification:

AS3 Friendship requires that I see X before he leaves London
The 2.30 train arrives in London at 4.15
So, I shall catch the 2.30 train.

Now I justify my action not in terms of its direct consequences, nor in terms of a state of affairs which will result from the action, but in terms of the underlying social value which I hope to promote by the action.

In general we may write instead of

W1a G is a goal for a

P1 a wishes to achieve S so as to bring about G which promotes a value V

Note that the answers to CQ1 are different in the cases AS1-3:

- in the case of AS1, I must propose other ways of arriving in London on time, perhaps by driving;
- in the case of AS2 I need not go to London at all; for example I could drive to Heathrow and say goodbye at the airport;
- in the case of AS3 I need not meet with X at all; perhaps a telephone call and an apology will be enough to promote friendship.

Given this more refined notion of a goal we can extend CQ1 to

CQ1a Are there alternative ways of realising the same consequences?
CQ1b Are there alternative ways of realising the same goal?
CQ1c Are there alternative ways of promoting the same values?

We can also elaborate CQ3, in that it may be that doing A realises some other goal which promotes some other value, or it may be that doing A prevents some other goal from being realised:

CQ3a Would doing A promote some other value?
CQ3b Does doing A preclude some other action which would promote some other value?

Also CQ4 has two aspects:

CQ4a Does doing A have a side effect which demotes the value V?
CQ4b Does doing A have a side effect which demotes some other value?

Secondly, apart from the possibility of the action, Walton does not consider other problems with soundness of W2, presupposing that the second premise is to be understood in terms of what a knows or reasonably believes. In [6] we proposed an argument scheme which incorporates P1 and makes the factual context explicit:

G1 In the Circumstances R
we should perform action A
to achieve New Circumstances S.
which will realise some goal G
which will promote some value V.

It could be that:

- A is not sufficient to bring about G; either because the current circumstances are not as presupposed, or because, although the beliefs about the current situation are correct, A does not have the believed effects.

- G is not a goal for a; either because there is some problem with the link between the circumstances brought about by doing A with the value a assumes them to promote, or because G is not in fact a possible state of affairs.

We can therefore add the critical questions:

CQ5 Are the circumstances such that doing A will bring about G?
CQ6 Does G promote V?
CQ7 Is G possible?

Note that an answer to CQ5 needs to address four points:

a) whether the believed circumstances are possible
b) whether the believed circumstances are true
c) assuming both of these, whether the action has the stated consequences
d) assuming all of these, whether the action will bring about the desired goal

Similarly, if we take the more articulated view of G expressed as P1, CQ6 needs to address both

a) whether G does realise the value intended; and
b) whether the value proposed is indeed a legitimate value

Also, taking G in terms of P1, CQ7 needs to address both

a) whether the situation S believed by a to result from doing A is a possible state of affairs
b) whether the particular aspects of S represented by G are possible

We thus have an elaborated set of critical questions: four variants of CQ5; three variants of CQ1; two variants of each of CQ3, CQ4, CQ6 and CQ7; and CQ2, making sixteen in all. We will use these sixteen questions as the basis for the development of our general theory of persuasion over action in section 4.

3.4 Natural Arguments And Practical Reasoning

When we examine natural arguments about practical reasoning we find that in practice many of them are elliptical, omitting one or more elements of the justification. For example, in the case of discussion of the proposed invasion of Iraq in 2003, one argument was

IW1 We should invade Iraq to depose Saddam

Here we have an action and the resulting situation, but the current situation is presupposed and the reasons for deposing Saddam and the value promoted by so doing are left implicit. There is a pragmatic reason for this: because persuasion relies on the value

preferences of the audience rather than the speaker (see [3] for a fuller account of this), the speaker can hope that the audience will recognise some desirable consequences and acceptable values for themselves. In this way there may be agreement on the action, without the need to agree on the specific goals achieved and values promoted. There is also a problem: many of the critical questions depend on these components of the justification. Thus only critical questions CQ5 and CQ7 can be posed without making some assumption about the goals and values of the proponent. If we reply, for example,

IW2 Saddam has no weapons of mass destruction,

assuming the removal of the weapons to be the desired goal, we may be drawing the argument into a irrelevant discussion: the proponent may have had quite other reasons for deposing Saddam. Thus, in addition to posing critical questions, it must be possible to make the proponent elaborate his position by supplying the missing pieces. Also it must be possible to state alternative positions, such as

IW3 We should not invade Iraq as it would breach international law.

Because so many presentations and criticisms of justifications for action are elliptical in naturally occurring arguments, any system which represents such dialogues must allow, in addition to stating positions and posing critical questions, the capability of seeking further clarification through asking additional questions to elucidate further information about the justification, and through supplying this supplementary information. The theory presented in section 4 will address these points.

4 General Theory of Persuasion Over Action

In this section we will attempt to make the sixteen critical questions identified in the last section more precise, by giving relatively formal definitions of them. The specific situation that we consider is where one agent is attempting to persuade another to adopt a course of action, and that other agent is arguing against this. Because we see this situation as one of conflict, we will refer to the various critical questions as ”attacks”. Persuasion is intended to be rational, and so reasons are advanced, and attacked, by each side. Moreover, the persuasion is intended to lead to action, so the debates are examples of practical reasoning.

We will also consider a number of variants on the basic attacks. When an element of a position is disputed, the attacker may simply disagree, or may additionally offer extra information indicating the source of the disagreement, or making the point of contention more precise. Thus for example if there is a disagreement as to the current fact situation, the opponent may simply deny what his opponent said, or may also add what he thinks is the case.

4.1 Stating a Position

In section 3 we gave the following as the general schema for a position motivating an action:

> (G1) In the Current Circumstances R
> we should perform Action A
> to achieve New Circumstances S
> which will realize some goal G
> which will promote some value V.

We need recognize no difference between resolving on a future action and justifying a past action. Moreover, an action may achieve multiple goals, and each goal may promote multiple values. For simplicity, we assume that the proponent of an action articulates an argument in the form of schema 1 for each goal realized and value promoted. We may then formalize the schemas as follows. We assume the existence of:

- A finite set of distinct actions, denoted *Acts*, with elements, A, B, C, etc.
- A finite set of propositions, denoted *Props*, with elements, p, q, r, etc.
- A finite set of states, denoted *States*, with elements, R, S, T, etc. Each element of *States* is an assignment of truth values $\{T, F\}$ to every element of *Props*.
- A finite set of propositional formulae, *Goals*, called goals, with elements G, H, etc.
- A finite set of values *Values*, with elements v, w, etc.
- A function *value* mapping each element of *Goals* to a pair $< v, sign >$, where $v \in$ *Values* and $sign \in \{+, =, -\}$.
- A ternary relation *apply* on *Acts* $\times$ *States* $\times$ *States*, with *apply(A, R, S)* to be read as: "*Performing action A in state R results in state S.*"[1]

The argument schema G1 contains a number of problematic notions which are not readily formalized in classical logic. We can, however, see that there are four classical statements which must hold if the argument represented by schema G1 is to be valid:

Statement 1: R is the case.

Statement 2: *apply(A, R, S)* $\in$ *apply*.

Statement 3: S $\models$ G (G is true in state S).

Statement 4: *value(G)* $=< v, + >$.

[1] We remark that formalisms of actions and their effects have received a great deal of attention in AI, e.g., the situation calculus [16]. In future work, we intend to explore the connections between these formalisms and our approach.

4.2 Attacking a Position

In this subsection we will describe the attacks corresponding to the critical questions of section 3 in terms of the elements identified in 4.1. We will group them in a slightly different manner, in order to emphasise different connections between the attacks, but we will in each case relate them to the source critical question.

4.2.1 Denial of Premises

A proposal for a particular action A can first be attacked by denying one of the four statements which must obtain for the proposal to be valid. Three of these premises relate to the the action realising the goal, and so relate to CQ5, whereas as one concerns the realisation of the claimed value and so relates to CQ6.

Attack 1 (CQ5b): R is not the case.

Attack 2 (CQ5c): It is not the case that *apply(A, R, S)* $\in$ *apply*.

Attack 3 (CQ5d): It is not the case that S $\models$ G.

Attack 4 (CQ6a): It is not the case that *value(G)* $=< v, + >$.

Each of these attacks may be executed with differing degrees of force, depending on whether positive information accompanies the attack, and the severity of the consequences of disagreement, and so we are able to distinguish variants of the main attack.

We can identify two variant attacks for **Attack 1**.

Attack 1a: R is not the case.

Attack 1b: R is not the case, and there is a circumstance Q $\in$ States, where R $\neq$ Q, such that Q is the case.

We can identify seven variant attacks for **Attack 2**.

Attack 2a: It is not the case that *apply(A, R, S)* $\in$ *apply*.

Attack 2b: It is not the case that *apply(A, R, S)* $\in$ *apply*, and it is the case that *apply(A, R, T)* $\in$ *apply*, where T $\neq$ S.

Attack 2c: It is not the case that *apply(A, R, S)* $\in$ *apply*, and it is the case that *apply(A, R, T)* $\in$ *apply*, where T $\neq$ S, but it is not the case that T $\models$ G.

Attack 2d: It is not the case that *apply(A, R, S)* $\in$ *apply*, and it is the case that *apply(A, R, T)* $\in$ *apply*, where T $\neq$ S, and it is the case that T $\models$ G, but it is not the case that *value(G)* $=< v, + >$.

Attack 2e: It is not the case that *apply(A, R, S)* $\in$ *apply*, and it is the case that *apply(A, R, T)* $\in$ *apply*, where T $\neq$ S, and it is the case that T $\models$ G, but *value(G)* $=< v, - >$.

Attack 2f: It is not the case that *apply(A, R, S)* $\in$ *apply*, and it is the case that *apply(A, R, T)* $\in$ *apply*, where T $\neq$ S, and it is the case that T $\models$ G, but *value(G)* $=< w, + >$, where w $\neq$ v.

Attack 2g: It is not the case that *apply(A, R, S)* $\in$ *apply*, and it is the case that *apply(A, R, T)* $\in$ *apply*, where T $\neq$ S, and it is the case that T $\models$ G, but *value(G)* $=< w, - >$, where w $\neq$ v.

Similarly, we may distinguish six variants of **Attack 3**:

Attack 3a: It is not the case that S $\models$ G.

Attack 3b: It is not the case that S $\models$ G and there is a goal H $\in$ *Goals*, H $\neq$ G, such that S $\models$ H.

Attack 3c: It is not the case that S $\models$ G and there is a goal H $\in$ *Goals*, H $\neq$ G, such that S $\models$ H and with *value(H)* $\neq< v, + >$.

Attack 3d: It is not the case that S $\models$ G and there is a goal H $\in$ *Goals*, H $\neq$ G, such that S $\models$ H and with *value(H)* $=< v, - >$.

Attack 3e: It is not the case that S $\models$ G and there is a goal H $\in$ *Goals*, H $\neq$ G, and a value w $\in$ *Values*, w $\neq$ v, such that S $\models$ H and with *value(H)* $=< w, + >$.

Attack 3f: It is not the case that S $\models$ G and there is a goal H $\in$ *Goals*, H $\neq$ G, and a value w $\in$ *Values*, w $\neq$ v, such that S $\models$ H and with *value(H)* $=< w, - >$.

Likewise, we may distinguish four variants of **Attack 4**:

Attack 4a: It is not the case that *value(G)* $=< v, + >$.

Attack 4b: It is not the case that *value(G)* $=< v, + >$ and *value(G)* $=< v, - >$.

Attack 4c: It is not the case that *value(G)* $=< v, + >$ and there is a value w $\in$ *Values*, w $\neq$ v, such that *value(G)* $=< w, + >$.

Attack 4d: It is not the case that *value(G)* $=< v, + >$ and there is a value w $\in$ *Values*, w $\neq$ v, such that *value(G)* $=< w, - >$.

4.2.2 Alternative Action for Same Effect

These three attacks all relate to CQ1, in that they propose an alternative way of achieveing the same desired results.

Attack 5 (CQ1a): There exists an action B $\in$ *Acts*, with B $\neq$ A, and *apply(B,R,S)* $\in$ *apply*.

Attack 6 (CQ1b): There exists an action B $\in$ *Acts*, with B $\neq$ A, and *apply(B,R,T)* $\in$ *apply*, with T $\models$ G.

Attack 7 (CQ1c): There exists an action B $\in$ *Acts*, with B $\neq$ A, and *apply(B,R,T)* $\in$ *apply*, with T $\models$ H, and *value(H)* $=< v, + >$.

4.2.3 Side Effects of the Action

Two of these attacks relate to unconsidered consequences of the action, raised by CQ4. The third offers a different justification for the action, and so relates to other goals that need to be considered, CQ3. This third attack also has a variant that does not correspond to any critical question (10a). Here the dispute concerns which features of the new situation should be identified as those promoting the value. This attack does not correspond to a critical question, because it does not dispute that the action should be performed, nor that the value will be promoted. Its significance comes when the discussion concerns the justification of a past action which is taken as a precedent supporting some future action. This becomes important in, for example, legal applications, as discussed in [6].

Attack 8 (CQ4a): There is a goal H $\in$ *Goals*, with H $\neq$ G, such that *apply(A,R,S)* $\in$ *apply* with S $\models$ H, and with *value(H)* $=< v, - >$.

Attack 9 (CQ4b): There is a goal H $\in$ *Goals*, with H $\neq$ G, and there is a value w $\in$ *values*, with w $\neq$ v, such that *apply(A,R,S)* $\in$ *apply* with S $\models$ H, and with *value(H)* $=< w, - >$.

Attack 10a: There is a goal H $\in$ *Goals*, with H $\neq$ G, such that *apply(A,R,S)* $\in$ *apply* with S $\models$ H, and with *value(H)* $=< v, + >$.

Attack 10b (CQ3a): There is a goal H $\in$ *Goals*, with H $\neq$ G, and there is a value w $\in$ *values*, with w $\neq$ v, such that *apply(A,R,S)* $\in$ *apply* with S $\models$ H, and with *value(H)* $=< w, + >$.

4.2.4 Interference with Other Actions

This group of attacks all relate to the promotion of some other value, and so derive from CQ3b. The three variants arise out of whether the action is incompatible with some other action, or whether it realises a state of affairs incompatible with the goal of another action, or whether the state of affairs realised is incompatible with *all* ways of promoting some other value.

Attack 11a: It is the case that *apply(A,R,S)* $\in$ *apply*. There is a value w $\in$ *values* with w $\neq$ v. There is an action B $\in$ *Acts* with B $\neq$ A, such that *apply(B,R,T)* $\in$ *apply*, with T $\models$ H, and *value(H)* $=< w, + >$. However, there is no state X $\in$ *States* such that *apply(A&B,R,X)* $\in$ *apply*.

Attack 11b: It is the case that *apply(A,R,S)* $\in$ *apply*. There is a value w $\in$ *values* with w $\neq$ v. There is a goal H $\in$ *Goals*, such that *value(H)* $=< w, + >$. However, S $\models \neg$H.

Attack 11c: It is the case that *apply(A,R,S)* $\in$ *apply*. There is a value w $\in$ *values* with w $\neq$ v. However, if there is a goal J $\in$ *Goals*, with *value(J)* $=< w, + >$, then S $\models \neg$J.

4.2.5 Disagreements Relating to Impossibility

The final group of attacks all relate to whether an element of the position is possible or not. In the critical questions we considered possibility together with the other questions relating to the element under dispute. Therefore these attacks relate to a number of different critical questions, as shown below.

Attack 12 (CQ2): It is not the case that A $\in$ *Acts*.

Attack 13a (CQ5a): It is not the case that R $\in$ *States*.

Attack 13b (CQ7a): It is not the case that S $\in$ *States*.

Attack 14 (CQ7b): It is not the case that G $\in$ *Goals*.

Attack 15 (CQ6b): It is not the case that v $\in$ *Values*.

We can summarise our attacks and their relation to the critical questions in Table 5. The last column will be discussed in section 4.3.

4.3 Responding to an Attack and Resolution

How a proponent of a proposal for action responds to an attack depends upon the nature of the attack. For those attacks which explicitly state an alternative position, the original proponent is able to counter-attack with some subset of the attacks listed in Table 5. For example, if a proponent argues for an action on the grounds that this will promote some value v, and an attacker argues in response that the proposed action will also demote some other value w, then the proponent may respond to this attack by arguing that the action does not have this effect on w (Attack 4), or that an alternative action can promote w, or that w is not worth promoting (Attack 15), etc.

Whether or not two participants may ultimately reach agreement on a proposed action will depend on the participants and on the precise nature of the disagreement. A basis for any resolution between participants for each type of attack is shown in the fourth column of Table 5. If the disagreement concerns the nature of the current world-state (Attack 1), for example, then some process of agreed empirical investigation may resolve this difference between the participants. Alternatively, if the participants disagree over which value should be promoted by the action (Attacks 9 or 15), then resolution will require agreement between them on a preference ordering over values. Such resolution may require other types of dialogue, and some of these interactions have received considerable attention from philosophers, for example [7, 17, 19]. We leave this topic for another occasion.

5 The *PARMA Protocol*

In this section we present the syntax of the *PARMA* (for Persuasive ARgument for Multiple Agents) *Action Persuasion Protocol*. In section 5.2 we give an outline of an axiomatic semantics for the Protocol. We assume, as in recent work in agent communications languages [12], that the language syntax comprises two layers: an inner layer

Table 5: **Attacks in Previous Work and Critical Questions**

Attack	Characterisation	Critical Question	Dispute to be resolved
1	Disagree with the description of the current situation	CQ5b	what is true
2	Disagree with the consequences of the proposed action	CQ5c	what is true
3	Disagree that the desired features are part of the consequences	CQ5d	representation
4	Disagree that these features promote the desired value	CQ6a	what is true
5	Believe that the consequences can be realized by some alternative action	CQ1a	what is best
6	Believe that the desired features can be realized through some alternative action	CQ1b	what is best
7	Believe that an alternative action realizes the desired value	CQ1c	what is best
8	Believe that the action has undesirable side effects which demote the desired value	CQ4a	what is best
9	Believe that the action has undesirable side effects which demote the desired value	CQ4b	what is best
10	Agree that the action should be performed, but for different reasons	CQ3a	what is best
11	Believe that the action will preclude some more desirable action	CQ3b	what is best
12	Believe that the action is impossible	CQ2	what is true
13	Believe that the circumstances or consequences as described are not possible	CQ5a, CQ7a	representation
14	Believe that the desired features cannot be realized	CQ7b	representation
15	Disagree that the desired value is worth promoting	CQ6b	representation

in which the topics of conversation are represented formally, and an outer, wrapper, layer comprising locutions which express the illocutionary force of the inner content. In our presentation of the axiomatic semantics we will assume propositional logic as the formal representation of the inner layer, but this restriction is for simplicity only. In section 5.3 we present a denotational semantics for the Protocol.

5.1 Axiomatic Semantics of the Protocol

The locutions of the *PARMA Protocol* are shown in the left-most columns of Tables 6–10. These tables also present the pre-conditions necessary for the legal utterance of each locution under the Protocol, and any post-conditions arising from their legal utterance. Thus, Tables 6–10 present an outline of an axiomatic semantics for the PARMA Protocol [22], and imply the rules governing the combination of locutions under the

Table 6: **Locutions to Control the Dialogue**

Locution	Pre-conditions	Post-conditions
Enter dialogue	Speaker has not already uttered enter dialogue	Speaker has entered dialogue
Leave dialogue	Speaker has uttered enter dialogue	Speaker has left dialogue
Turn finished	Speaker has finished making their move	Speaker and hearer switch roles so new speaker can now make a move
Accept denial	Hearer has made an attack on an element of speaker's position	Speaker committed to the negation of the element that was denied by the hearer
Reject denial	Hearer has made an attack on an element of speaker's position	Disagreement reached

protocol [14]. We further assume, following [8] and in accordance with most work on dialogue games (e.g.[13]) and recent work in agent communications, that a *Commitment Store* is associated with each participant, which stores, in a manner which all participants may read, the commitments made by that participant in the course of a dialogue. The post-conditions of utterances shown in Tables 6–10 include any commitments incurred by the speaker of each utterance while the pre-conditions indicate any prior commitments required before an utterance can be legally made. Commitments in this protocol are dialogical — ie, statements which an agent must defend if attacked, and may bear no relation to the agent's real beliefs or intentions [8]. We thus make no assumptions about the private mental states of the agents involved in the dialogue.

5.1.1 Locutions for the Attacks

The set of attacks presented in section 4 can now be individually described by combining the previously defined locutions of the dialogue game. The attacks are made up of a mixture of the primitive locutions and the order in which the primitive locutions are presented as part of an attack is of no relevance.

Attack 1a: deny circumstances(R).

Attack 1b: deny circumstances(R) AND state circumstances(Q).

Attack 2a: deny consequences(A,R,S).

Attack 2b: state consequences(A,R,T) AND deny consequences(A,R,S).

Attack 2c: state consequences(A,R,T) AND deny consequences(A,R,S) AND deny logical consequences(T,G).

Attack 2d: state consequences(A,R,T) AND state logical consequences(T,G) AND deny purpose(G,V,D+) AND deny consequences(A,R,S).

Attack 2e: state consequences(A,R,T) AND state logical consequences(T,G) AND state purpose(G,V,D-) AND deny consequences(A,R,S).

Attack 2f: state consequences(A,R,T) AND state logical consequences(T,G) AND state purpose(G,W,D+) AND deny consequences(A,R,S).

Attack 2d: state consequences(A,R,T) AND state logical consequences(T,G) AND state purpose(G,W,D-) AND deny consequences(A,R,S).

Attack 3a: deny logical consequences(S,G).

Attack 3b: state logical consequences(S,H) AND deny logical consequences(S,G).

Attack 3c: state logical consequences(S,H) AND state purpose(H,V,D+) AND deny logical consequences(S,G).

Attack 3d: state logical consequences(S,H) AND state purpose(H,V,D-) AND deny logical consequences(S,G).

Attack 3e: state logical consequences(S,H) AND state purpose(H,W,D+) AND deny logical consequences(S,G).

Attack 3f: state logical consequences(S,H) AND state purpose(H,W,D-) AND deny logical consequences(S,G).

Attack 4a: deny purpose(G,V,D+).

Attack 4b: state purpose(G,V,D-) AND deny purpose(G,V,D+).

Attack 4c: state purpose(G,W,D+) AND deny purpose(G,V,D+).

Attack 4d: state purpose(G,W,D-) AND deny purpose(G,V,D+).

Attack 5: state action(B) AND state consequences(B,R,S).

Attack 6: state action(B) AND state consequences(B,R,T) AND state logical consequences(T,G).

Attack 7: state action(B) AND state consequences(B,R,T) AND state logical consequences(T,H) AND state purpose(H,V,D+).

Attack 8: state consequences(A,R,S) AND state logical consequences(S,H) AND state purpose(H,V,D-).

Attack 9: state consequences(A,R,S) AND state logical consequences(S,H) AND state purpose(H,W,D-).

Attack 10a: state consequences(A,R,S) AND state logical consequences(S,H) AND state purpose(H,V,D+).

Attack 10b: state consequences(A,R,S) AND state logical consequences(S,H) AND state purpose(H,W,D+).

Attack 11a: state consequences(A,R,S) AND state action(B) AND state consequences(B,R,T) AND state logical consequences(T,H) AND state purpose(H,W,D+) AND deny consequences(A&B,R,X).

Attack 11b: state consequences(A,R,S) AND state purpose(H,W,D+) state logical consequences(S,¬ H).

Attack 11c: state consequences(A,R,S) AND (IF state purpose(J,W,D+) THEN state logical consequences(S,¬ J).

Attack 12: deny action exists(A).

Attack 13a: deny initial state exists(R).

Attack 13b: deny resultant state exists(S).

Attack 14: deny goal exists(G).

Attack 15: deny value exists(R).

Once a move has legally been performed by a player, the turn can be passed, where the next player then has a set of moves from which they can choose their next utterance. These next available moves are entirely defined by the pre-conditions of the locutions. This means that checking the pre-conditions for the legality of moves ensures that the dialogue is sensibly structured and that irrelevant or inappropriate utterances cannot be made during the course of the dialogue.

5.2 Denotational Semantics of the Protocol

We now outline a denotational semantics for the *PARMA* protocol, that is a semantics which maps statements in the syntax to mathematical entities [22]. Our approach draws on the semantics proposed by Charles Hamblin for imperative statements [9], which itself may be viewed as a process theory of causality. The main proponent of such

theories has been Wesley Salmon, whose theory of causal processes *"identifies causal connections with physical processes that transmit causal influence from one spacetime location to another"* [20, p. 191]. Our approach draws on elements of category theory, namely topos theory. Our reason for using this, rather than (say) a Kripkean possible worlds framework or a labelled transition system, is that topos theory enables a natural representation of logical consequence ($S \models G$) in the same formalism as mappings between spaces ($R \stackrel{A}{\rightarrow} S$ and $G \uparrow v$). To our knowledge, no other non-categorical denotational semantics currently proposed for action formalisms permits this.

We begin by representing proposals for action. We assume, as in Section 2.1, finite sets of Acts, Propositions, States, Goals, and Values, and various mappings. For simplicity, we assume there are n propositions. Each State may be considered as being equivalent to the set of propositions which are true in that State, and so there are 2^n States. We consider the space $\mathcal{C}$ of these States, with some additional structure to enable the representation of actions and truth-values. We consider values as mappings from Goals to some space of evaluations, called $\mathcal{S}$. This need not be the three-valued set *Sign* $= \{+, =, -\}$ we assumed in Section 2.1, although we assume that $\mathcal{S}$ admits at least one partial order. The structures we assume on $\mathcal{C}$, $\mathcal{S}$ and between them is intended to enable us to demonstrate that these are categorical entities [4]. We begin by listing the mathematical entities, along with informal definitions.

- The space $\mathcal{C}$ comprises a finite collection $\mathcal{C}_0$ of objects and a finite collection $\mathcal{C}_1$ of arrows between objects.
- $\mathcal{C}_0$ includes 2^n objects, each of which may be considered as representing a State. We denote these objects by the lower-case Greek letters, $\alpha, \beta, \gamma, \ldots$, and refer to them collectively as *state objects* or *states*. We may consider each state to be equivalent (in some sense) to the set of propositions which are true in the state.
- $\mathcal{C}_1$ includes arrows between state objects, denoted by lower case Roman letters, $f, g, h, \ldots$. If f is an arrow from object α to object $beta$, we also write $f : \alpha \rightarrow \beta$. Some arrows between the state objects may be considered as representing actions leading from one state to another, while other arrows are causal processes (not actions of the dialogue participants) which take the world from one state to another. There may be any number of arrows between the same two objects: zero, one, or more than one.
- Associated with every object $\alpha \in \mathcal{C}_0$, there is an arrow $1_\alpha \in \mathcal{C}_1$ from α to α, called the identity at α. In the case where α is a state object, this arrow may be considered as that action (or possibly inaction) which preserves the status quo at a state α.
- If $f : \alpha \rightarrow \beta$ and $g : \beta \rightarrow \gamma$ are both arrows in $\mathcal{C}_1$, then we assume there is an arrow $h : \alpha \rightarrow \gamma$. We denote this arrow h by $g \circ f$ (*"g composed with f"*). In other words, actions and causal processes may be concatenated.
- We assume that $\mathcal{C}_0$ includes a special object *Prop*, which represents the finite set of all propositions. We further assume that for every object $\alpha \in \mathcal{C}_0$ there

is a monic arrow $f_\alpha : \alpha \rightarrow Prop$. Essentially, a monic arrow is an injective (one-to-one) mapping.

- We assume that $\mathcal{C}_0$ has a terminal object, **1**, ie, an object such that for every object $\alpha \in \mathcal{C}_0$, there is precisely one arrow $\alpha \rightarrow \mathbf{1}$.
- We assume that $\mathcal{C}$ has a special object Ω, and an arrow $true : \mathbf{1} \rightarrow \Omega$, called a *sub-object classifier*. The object Ω may be understood as the set comprising $\{True, False\}$.
- We assume that $\mathcal{S}$ is space of objects over which there is a partial order $<_i$ corresponding to each participant in the dialogue. Such a space may be viewed as a category, with an arrow between two objects α and β whenever $\alpha <_i \beta$. For each participant, we further assume the existence of one or more mappings v between $\mathcal{C}$ and $\mathcal{S}$, which takes objects to objects, and arrows to arrows. We denote the collection of all these mappings by $\mathcal{V}$.

The assumptions we have made here enable us to show that $\mathcal{C}$ is a category [4], and we can thus represent the statement $R \xrightarrow{A} S$, for states R and S, and action A. Moreover, the presence of a sub-object classifier structure enables us to represent statements of the form $S \models G$, for state S and goal G, inside the same category $\mathcal{C}$. This structure we have defined for $\mathcal{C}$ creates some of the properties needed for $\mathcal{C}$ to be a topos [4]. Finally, each space $\mathcal{S}$ with partial order $<_i$ is also a category, and the mappings v are functors (structure-preserving mappings) between $\mathcal{C}$ and $\mathcal{S}$. This then permits us to represent statements of the form $G \uparrow v$, for goal G and value v.

We define a denotational semantics for the *PARMA* Protocol by associating dialogues conducted according to the Protocol with mathematical structures of the type defined above. Thus, the statement of a proposal for action by a participant in a dialogue

$$R \xrightarrow{A} S \models G \uparrow v$$

is understood semantically as the assertion of the existence of objects representing R and S in $\mathcal{C}$, the existence of an arrow representing A between them, the existence of an arrow with certain properties[2] between *Prop* and Ω, and the existence of a functor $v \in \mathcal{V}$ from $\mathcal{C}$ to $\mathcal{S}$. Attacks on this position then may be understood semantically as denials of the existence of one or more of these elements, and possibly also, if the attack is sufficiently strong, the assertion of the existence of other objects, arrows or functors.

Thus, our denotational semantics for a dialogue conducted according to the *PARMA* Protocol is defined as a countable sequence of triples,

$$\langle \mathcal{C}_1, \mathcal{S}_1, \mathcal{V}_1 \rangle, \langle \mathcal{C}_2, \mathcal{S}_2, \mathcal{V}_2 \rangle, \langle \mathcal{C}_3, \mathcal{S}_3, \mathcal{V}_3 \rangle, \ldots,$$

where the k-th triple is created from the k-th utterance in the dialogue according to the representation rules just described. Then, our denotational semantics for the *PARMA*

[2]This arrow is the characteristic function for the object representing G, and the properties are that a certain diagram commutes in $\mathcal{C}$.

Protocol itself is defined as the collection of all such countable sequences of triples for valid dialogues conducted under *PARMA*. This approach views the semantics of the protocol as a space of mathematical objects, which are created incrementally and jointly by the participants in the course of their dialogue together. The approach derives from the constructive view of human language semantics of Discourse Representation Theory [10], and is similar in spirit to the denotational semantics, called a *trace semantics*, defined for deliberation dialogues in [15] and the *dialectical graph* recording the statements of the participants in the Pleadings Game of Thomas Gordon [5]. We are currently engaged in specifying formally the semantics in accordance with the outline presented here.

5.3 State Transition Diagram

Figure 1 below shows a simple state transition diagram for the Protocol. It shows the types of moves that the players can make and the choice of move which is then available in the new state. It also shows the moves that lead to the roles of speaker and hearer being switched and how the game can terminate. The diagram does not show the specific details of all moves that can be made, only the types of moves. For example a 'state' move can be any of the moves given in the axiomatic semantics for proposing an action e.g. 'state circumstances', 'state action', 'state consequences' etc. A 'deny' move can be any of the moves given in the axiomatic semantics for attacking a position e.g. 'deny circumstances', 'deny consequences', 'deny action exists' etc. An 'ask' move can be any of the moves given in the axiomatic semantics for asking about an agent's position e.g. 'ask circumstances', 'ask action', 'ask consequences' etc.

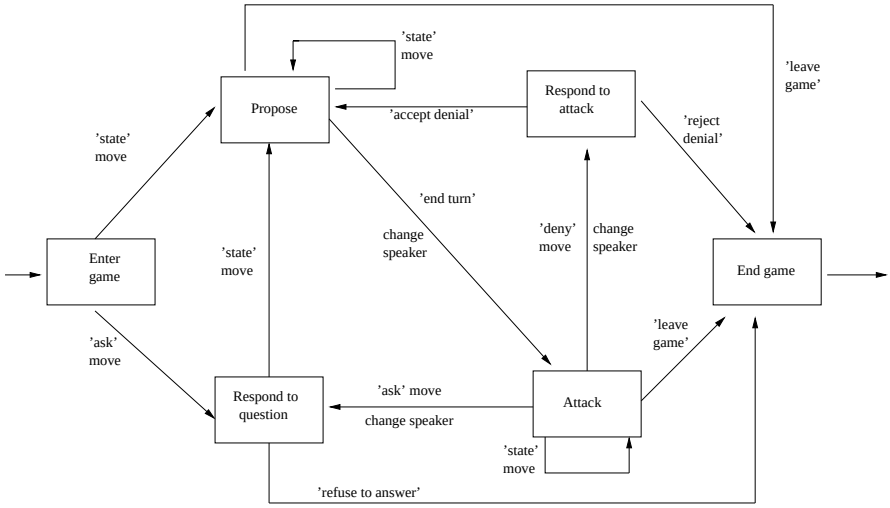

Figure 1: State Transition Diagram for *PARMA* Dialogue Game Protocol.

6 Implementation of Dialogue Game

We have implemented the *PARMA Action Persuasion Protocol* specified above in the form of a Java program. The program implements a version of the protocol so that dialogues between two human participants can be undertaken under the protocol, with each participant taking turns to propose and attack positions uttering the locutions specified above. The program checks the legality of the participants' chosen moves by verifying that all pre-conditions for the move hold. Thus, the participants are able to state and attack each other's positions with the program verifying that the dialogue always complies with the protocol. If a participant attempts to make an illegal move then they are informed about this and given the opportunity to chose an alternative move. After a move has been legally uttered, the commitment store of the participant who made the move is updated to contain any new commitments created by the utterance. All moves, whether legal or illegal, are entered into the history, which records which moves were made by which participant and the legality of the move chosen. After a move has been legally made, the commitment store of the player who made the move is printed to the screen to show all previous commitments and any new ones that have consequently been added. By publicly displaying the commitment stores in this way each participant is able to see their own and each other's commitments. Thus, participants can determine which of their commitments overlap with those of the other participant, and thereby identify points of agreement. Conversely, this also allows each participant to identify any commitments of the other participant in conflict with their own, and thus which commitments are susceptible to an attack.

Dialogues undertaken via the program can terminate in a number of ways (see Figure 1). A participant can decide to leave the game by exiting at any time, thereby terminating the dialogue. A dialogue can also terminate if disagreement about a position is reached. This occurs when a participant states an element of a position which is is consequently attacked by the other participant, and the first participant disagrees with the attack. If the first participant refuses to accept the reasons for the attack then disagreement has been identified and the dialogue terminates. Dialogues may also reach a natural end with agreement between the two participants on a course of action. If this occurs, both players may choose to exit the dialogue.

When a dialogue terminates, whether in agreement or disagreement, the history and commitment stores of both players are printed on screen and also to a file. The dialogue may then be analyzed, for example to see which attacks occurred, or how often or how successful they were. Such analysis may be useful for a study of appropriate strategies for dialogue conducted under the protocol. Further details of the implementation can be found in [1].

7 Example Use of the Dialogue Game

Table 11 below gives an example transcript of a dialogue game being conducted in accordance with the PARMA protocol. Two parties, A and B, are engaged in a discussion about where and what type of holiday to go on together. Persuasion occurs through use of the attacks from the protocol, as both parties have different preferences. The dispute

arises due to party A wanting to go on a beach holiday and party B wanting to go on a skiing holiday.[3]

The above transcript represents an example of a dialogue being conducted between two parties but it does not show any details of the internal recordings made by the program concerning the commitments incurred by the players. The program mediates the dialogue exchange by checking that the pre-conditions for chosen moves hold and if they do hold, then the program executes the post-conditions of the move by updating the appropriate player's commitment store to include the addition of any newly incurred commitments. The program also maintains a history of all moves chosen throughout the course of the dialogue, even when the move chosen is an illegal one.

The history holds details of the player who is making the move, the name of their chosen locution, the status of the locution, which can either be 1 to show that the chosen move is a legal move or -1 to denote that an illegal move was chosen and has not incurred any new commitment, and finally the history contains the content of the locution chosen. Thus, the history documents all moves attempted and made.

The commitment stores of each player contain the name of the moves they have made, the content of each move and the status of each commitment, with 1 being positive commitment to the content, -1 being commitment to the negation of the content and 0 being no commitment at all to the content.

Both the commitment stores and the history are updated and displayed on screen whenever a new move/commitment is added to either. Tables 13-15 below give example snapshots of the history and each player's commitment store showing what they contain after moves 1-7 from the example dialogue have been made[4]:

8 Issues Arising from the Implementation

Implementing the dialogue game has proved to be a very useful task as we have shown that our general theory of persuasion can be conducted via computer mediated dialogues of this form. This implementation has also raised a number of interesting issues in relation to our underlying argumentation scheme. Below we summarise the three main general insights which have arisen through our evaluation of the implemented dialogue game protocol:

1. The referee cannot use pre-conditions based on mental states of the participants: he infers these from the moves the players make. This means that the pre-conditions to allow a move are different from those to sincerely make a move.

[3]Note: we use the word goal in a loose sense to denote some state of affairs which a player wishes to achieve. Thus, the example transcript shows that although a player can commit to a goal which may be the negation of one of the opposing player's goals this does not constitute an attack, just a statement of opposing preferences.

[4]Note: as a natural language dialogue is being modelled in this example the commitment stores include words that are conversation fillers, such as "well", "so", etc, as these are naturally used in everyday conversation. However, if the game were to be used by computer agents rather than human agents such words would not be included. In such a case the content of moves would contain purely propositional statements based on the representation of the knowledge embodied in the computer agents' knowledge bases.

2. Natural dialogue is very flexible. Giving support requires constraints and what constraints are appropriate depends on context and purpose. The protocol may impose too few constraints to allow scope for useful computer support.

3. Goodwill and some co-operation is required to make sensible progress and this is again due to the fact that natural dialogue is so flexible. Thus, uncooperative players can abuse the protocol to stultify the purpose of the dispute.

For a more detailed description of the issues encountered through the implementation of the Java program see [1]

9 Alternative Use of the Argument Scheme

After reflecting on the issues raised in the previous section regarding the Java implementation of the Protocol, we have concluded that the implementation poses many problems for casual users of the system. In order to correctly follow the Protocol the users must have prior knowledge of the underlying theory of persuasion. If they do not have prior knowledge of the theory then they will be unable to recognise which locutions need to be chosen in order to realise the correct attack, in a given situation. The users must also be familiar with the names and meanings of the locutions used to represent the statement and denial of a position. As well as these usability problems, we mentioned in the previous section that the dialogue game does rely somewhat on the goodwill of the players to use the Protocol sensibly, as legal moves may well be unhelpful and unconstructive.

Some of these problems have arisen due to the amount of freedom of expression afforded by the program and this leaves the users with an overwhelming variety of options to select between. All these points related to problems with the usability of the program are obviously undesirable. Therefore, we have addressed these issues by going on to implement our theory of persuasion in an entirely different format.

We have developed an online discussion forum, named PARMENIDES (Persuasive ArguMENt In DEmocracieS) which allows a much simpler form of interaction to take place. The user is guided through a series of web pages in order to elicit their views on a particular topic, in accordance with our theory. The user interaction occurs through a simple web based interface which guides them in a structured fashion through a justification of an action, giving opportunities to disagree at selected points. Each of these disagreements represents one of the attacks from our theory of persuasion, so the exact nature of the disagreement can be unambiguously identified. By constraining the choice of the user in such a way, the need for them to understand the underlying argumentation scheme and thus select the correct moves is removed. The responses of the users are written to a database so we are able to gather and analyse the information in order to identify what points of the argument are more strongly supported than others.

This system has been successfully implemented and we are satisfied that it is, given the particular situation of intended use, an improved alternative implementation to the Java program, as it overcomes many of the usability problems presented by the Java program, which we highlighted above. We now intend to focus on this system to extend our theory and implementation further. We hope to further explore other elements, such

as counter attacks and allow the construction of positive alternative arguments, as the system currently focuses on the negative criticism of arguments. We will then consider how this approach might be adapted to different use situations, including a different selection of attacks. Details of the PARMENIDES online discussion forum can be found in [2].

10 Summary

In this report we have:

- Discussed the typology of dialogues of Walton and Krabbe and offered a precise interpretation of them.
- Discussed the problems associated with practical reasoning and proposed a perspective on practical reasoning in presumptive justification and critical questions. Our account represents an extension of Walton [24].
- Given a precise definition of the main critical questions and variants of them which could provide the foundation for a protocol for a dialogue game based on this theory.
- Given an axiomatic and denotational semantics for this protocol.
- Discussed two implementations based on this protocol.

This work has drawn our attention to the importance of the context in which the Protocol is useful. One line of future work will be to explore the Protocol in different specific contexts. A second line of investigation will be to extend the Protocol to accommodate responses to the critical questions which form the basis of our dialogue moves.

References

[1] K.M. Atkinson, T.J.M Bench-Capon, and P.J. McBurney. Implementation of a dialogue game for persuasion over action. Technical Report ULCS-04-005, Department of Computer Science, University of Liverpool, UK, 2003.

[2] K.M. Atkinson, T.J.M. Bench-Capon, and P.J. McBurney. Parmenides: Facilitating democratic debate. In R. Traunmüller, editor, *Electronic Government 2004*, LNCS. Springer, Berlin, 2004. *To appear*.

[3] T.J.M. Bench-Capon. Agreeing to differ: Modelling persuasive dialogue between parties without a consensus about values. *Informal Logic*, 22(3), 2003.

[4] R. Goldblatt. *Topoi: The Categorial Analysis of Logic*. North-Holland, Amsterdam, The Netherlands, 1979.

[5] T. F. Gordon. The Pleadings Game: An exercise in computational dialectics. *Artificial Intelligence and Law*, 2:239–292, 1994.

[6] K.M. Greenwood, T.J.M. Bench-Capon, and P.J. McBurney. Towards a computational account of persuasion in law. In *Proc. Ninth Intern. Conf. AI and Law (ICAIL-2003)*, pages 22–31, New York, NY, USA, 2003. ACM Press.

[7] J. Habermas. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. MIT Press, Cambridge, MA, USA, 1996. (Translation by W. Rehg).

[8] C. L. Hamblin. *Fallacies*. Methuen, London, UK, 1970.

[9] C. L. Hamblin. *Imperatives*. Basil Blackwell, Oxford, UK, 1987.

[10] H. Kamp and U. Reyle. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer Academic, Dordrecht, The Netherlands, 1993. Two Volumes.

[11] A.J.P. Kenny. *Practical Reasoning and Rational Appetite*. 1975. Reprinted in Reprinted in Raz, J., (ed) (1978). Practical Reasoning. Oxford University Press,Oxford,UK.

[12] Y. Labrou, T. Finin, and Y. Peng. Agent communication languages: The current landscape. *IEEE Intelligent Systems*, 14(2):45–52, 1999.

[13] J. D. MacKenzie. Question-begging in non-cumulative systems. *Journal of Philosophical Logic*, 8:117–133, 1979.

[14] P. McBurney and S. Parsons. Games that agents play: A formal framework for dialogues between autonomous agents. *J. Logic, Language and Information*, 11(3):315–334, 2002.

[15] P. McBurney and S. Parsons. A denotational semantics for deliberation dialogues. In N. R. Jennings, C. Sierra, E. Sonenberg, and M. Tambe, editors, *Proc.Third Intern. Joint Conf. Autonomous Agents and Multi-Agent Systems (AAMAS 2004)*, 2004. *In press*.

[16] J. McCarthy and P. J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In B. Melzer and D. Michie, editors, *Machine Intelligence 4*, pages 463–502. Edinburgh University Press, 1969.

[17] C. Perelman and L. Olbrechts-Tyteca. *The New Rhetoric: A Treatise on Argumentation*. University of Notre Dame Press, Notre Dame, IN, USA, 1969.

[18] J. Raz(editor). *Practical Reasoning*. Oxford University Press, Oxford, UK, 1978.

[19] H. S. Richardson. *Practical Reasoning about Final Ends*. Cambridge University Press, Cambridge, UK, 1994.

[20] W. C. Salmon. *Causality and Explanation*. Oxford University Press, New York, NY, USA, 1998.

[21] J.R. Searle. *Rationality in Action*. MIT Press, Cambridge, MA, USA, 2001.

[22] R. D. Tennent. *Semantics of Programming Languages*. Prentice-Hall, Hemel Hempstead, UK, 1991.

[23] D. N. Walton and E. C. W. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. SUNY Press, Albany, NY, USA, 1995.

[24] D.N. Walton. *Argument Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1996.

Table 7: **Locutions to Propose an Action**

Locution	Pre-conditions	Post-conditions
State circumstances(R)	Speaker has uttered enter dialogue	Speaker committed to R Speaker committed to R $\in$ States
State action(A)	Speaker has uttered enter dialogue Speaker committed to R Speaker committed to R $\in$ States	Speaker committed to A Speaker committed to A $\in$ Acts
State consequences(A,R,S)	Speaker has uttered enter dialogue Speaker committed to R Speaker committed to R $\in$ States Speaker committed to A Speaker committed to A $\in$ Acts	Speaker committed to apply(A,R,S) $\in$ apply Speaker committed to S $\in$ States
State logical consequences(S,G)	Speaker has uttered enter dialogue Speaker committed to R Speaker committed to R $\in$ States Speaker committed to A Speaker committed to A $\in$ Acts Speaker committed to apply(A,R,S) $\in$ apply Speaker committed to S $\in$ States	Speaker committed to S $\models$ G Speaker committed to G $\in$ Goals
State purpose(G,V,D)	Speaker has uttered enter dialogue Speaker committed to R Speaker committed to R $\in$ States Speaker committed to A Speaker committed to A $\in$ Acts Speaker committed to apply(A,R,S) $\in$ apply Speaker committed to S $\in$ States Speaker committed to S $\models$ G Speaker committed to G $\in$ Goals	Speaker committed to (G,V,D) Speaker committed to V $\in$ Values

Table 8: **Locutions to ask about an Agent's Position**

Locution	Pre-conditions	Post-conditions
Ask circumstances(R)	Hearer has uttered enter dialogue Speaker has uttered enter dialogue Speaker not committed to circumstances(R) about topic in question	Hearer must reply with state circumstances(R) or don't know(R)
Ask action(A)	Hearer has uttered enter dialogue Speaker has uttered enter dialogue Speaker not committed to action(A) about topic in question	Hearer must reply with state action(A) or don't know(A)
Ask consequences(A,R,S)	Hearer has uttered enter dialogue Speaker has uttered enter dialogue Speaker not committed to consequences(A,R,S) about topic in question	Hearer must reply with state consequences(A,R,S) or don't know(A,R,S)
Ask logical consequences(S,G)	Hearer has uttered enter dialogue Speaker has uttered enter dialogue Speaker not committed to logical consequences(S,G) about topic in question	Hearer must reply with state logical consequences(S,G) or don't know(S,G)
Ask purpose(G,V,D)	Hearer has uttered enter dialogue Speaker has uttered enter dialogue Speaker not committed to purpose(G,V,D) about topic in question	Hearer must reply with state purpose(G,V,D) or don't know(G,V,D)

Table 9: **Locutions to Attack Elements of a Position**

Locution	Pre-conditions	Post-conditions
Deny circumstances(R)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R Hearer committed to R $\in$ States	Speaker committed to deny circumstances(R)
Deny consequences(A,R,S)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R Hearer committed to R $\in$ States Hearer committed to A Hearer committed to A $\in$ Acts Hearer committed to apply(A,R,S) $\in$ apply Hearer committed to S $\in$ States	Speaker committed to deny consequences(A,R,S) $\in$ apply
Deny logical consequences(S,G)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R Hearer committed to R $\in$ States Hearer committed to A Hearer committed to A $\in$ Acts Hearer committed to apply(A,R,S) $\in$ apply Hearer committed to S $\in$ States Hearer committed to S $\models$ G Hearer committed to G $\in$ Goals	Speaker committed to deny logical consequences S $\models$ G
Deny purpose(G,V,D)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R Hearer committed to R $\in$ States Hearer committed to A Hearer committed to A $\in$ Acts Hearer committed to apply(A,R,S) $\in$ apply Hearer committed to S $\in$ States Hearer committed to S $\models$ G Hearer committed to G $\in$ Goals Hearer committed to (G,V,D) Hearer committed to V $\in$ Values	Speaker committed to deny purpose(G,V,D)

Table 10: **Locutions to Attack Validity of Elements**

Locution	Pre-conditions	Post-conditions
Deny initial circumstances exist(R)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R $\in$ States	Speaker committed to deny initial circumstances exist(R)
Deny action exists(A)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R Hearer committed to R $\in$ States Hearer committed to A $\in$ Acts	Speaker committed to deny action exists(A)
Deny resultant state exists(S)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R Hearer committed to R $\in$ States Hearer committed to A $\in$ Acts Hearer committed to S $\in$ States	Speaker committed to deny resultant state exists(S)
Deny goal exists(S)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R Hearer committed to R $\in$ States Hearer committed to A $\in$ Acts Hearer committed to S $\in$ States Hearer committed to G $\in$ Goals	Speaker committed to deny goal exists(G)
Deny value exists(S)	Speaker has uttered enter dialogue Hearer has uttered enter dialogue Hearer committed to R Hearer committed to R $\in$ States Hearer committed to A $\in$ Acts Hearer committed to S $\in$ States Hearer committed to G $\in$ Goals Hearer committed to V $\in$ Values	Speaker committed to deny value exists(V)

Table 11: **Example Dialogue**

Move No.	Player	Locutions	Content
1	A	state circs(R_1)	I would like to book our summer holiday this weekend.
2	A	ask circs(R)	Would you like to book it this weekend?
3	B	state circs(R_1)	Yes, I would like to do that too.
4	B	ask action(A)	So, where do you think we should go?
5	A	state action(A_1)	Well, I was thinking about one of the Greek islands.
6	A	ask action(A)	Where did you have in mind?
7	B	state action(A_2)	I fancied a skiing holiday somewhere
8	B	state conseq(A_2,R_1,S_1)	so we'd get to go to a ski resort
9	B	state log conseq(S_1,G_1)	where there would be lots of things to do during the day as well as at night
10	B	state purpose(G_1,V_1,D+)	So I think we'd have a really good time.
11	A	*attack 6*: state action(A_1) state conseq(A_1,R_1,S_2) state log conseq(S_2,G_1)	 But going to the Greek islands means we'd get to go to a beach resort where there's also lots to do during the day and at night too.
12	A	state conseq(A_1,R_1,S_2)	And, the Greek islands are nice and hot during the summer months
13	A	state log conseq(S_2,G_2)	and you know how I like to spend my holidays in the sun
14	A	state purpose(G_2,V_2,D+)	as it helps me to relax.
15	B	*attack 9*: state conseq(A_1,R_1,S_3) state log conseq(S_3,G_3) state purpose(G_3,V_3,D-)	 But we went on a similar holiday last year and it'll be just the same and I'd like to do something different this year.
16	B	*attack 9*: state conseq(A_2,R_1,S_4) state log conseq(S_4,G_4) state purpose(G_4,V_2,D-)	 But going skiing means going on holiday to somewhere that's in a cold climate and I want a holiday in the sun.
17	B	*attack 9*: state conseq(A_2,R_1,S_5) state log conseq(S_5,G_5) state purpose(G_5,V_4,D+)	 But it will be an activity holiday which is different from what we're used to and that will make it more exciting.

Table 12: **Example Dialogue cont'd**

18	B	*attack 6*: state action(A_3) state conseq(A_3,R_1,S_6) state log conseq(S_6,G_2)	 And, we can go on a beach holiday in winter instead of our usual city break so we'll get to go on two holidays one of which will be your beach holiday.
19	A	state purpose(G_2,V_2,D+)	I suppose that would be nice.
20	A	*attack 9*: state conseq(A_2,R_1,S_2) state log conseq(S_2,G_6) state purpose(G_6,V_5,D-)	 But going skiing means going on holiday to somewhere that's very expensive and we can't afford to spend a lot of money.
21	B	*attack 7*: state action(A_4) state conseq(A_4,R_1,S_4) state log conseq(S_4,G_7) state purpose(G_7,V_5,D+)	 Not if you go to somewhere in Eastern Europe as there are ski resorts there which are very cheap at the moment and it really won't cost a lot of money.
22	A	state purpose(G_7,V_5,D+)	Yes, that's true. A lot of friends have told me the same thing.
23	B	ask action(A)	Great. So you agree to a skiing holiday then?
24	A	state action(A_2)	Yes, I'll give it a go.
25	A	state action(A_5)	We can go to the travel agent's tomorrow and look at some destinations and prices.

Table 13: **History after move 7**

Player	Move	Status	Content
A	state circs	1	I would like to book our summer holiday this weekend.
A	ask circs	1	Would you like to book it this weekend?
B	state circs	1	Yes, I would like to do that too.
B	ask action	1	So, where do you think we should go?
A	state action	1	Well, I was thinking about one of the Greek islands.
A	ask action	1	Where did you have in mind?
B	state action	1	I fancied a skiing holiday somewhere

Table 14: **Player A's commitment store after move 5**

Move	Status	Content
enter dialogue	1	enter dialogue
state circs	1	I would like to book our summer holiday this weekend
circ exist	1	I would like to book our summer holiday this weekend exists in set of possible circs
state action	1	Well, I was thinking about one of the Greek islands
act exist	1	Well, I was thinking about one of the Greek islands exists in the set of possible actions

Table 15: **Player B's commitment store after move 7**

Move	Status	Content
enter dialogue	1	enter dialogue
state circs	1	Yes, I would like to do that too
circ exist	1	Yes, I would like to do that too exists in set of possible circs
state action	1	I fancied a skiing holiday somewhere
act exist	1	I fancied a skiing holiday somewhere exists in the set of possible actions