A formal framework for inter-agent dialogues

Peter McBurney and Simon Parsons
Department of Computer Science
University of Liverpool
Liverpool L69 7ZF, United Kingdom
{p.j.mcburney,s.d.parsons}@csc.liv.ac.uk

ABSTRACT

We present a logic-based formalism for modeling of dialogues between intelligent and autonomous software agents, building on a theory of abstract dialogue games which we present. The formalism enables representation of complex dialogues as sequences of moves in a combination of dialogue games, and allows dialogues to be embedded inside one another. The formalism can be readily operationalized and its modular nature enables different types of dialogues to be represented.

Keywords

Computational Dialectics, Conversational Agents, Dialogue Games, Multi-agent Communication/Collaboration

1. INTRODUCTION

Autonomous intelligent software agents have become a powerful paradigm in modern computer science. In this paradigm, discrete software entities — autonomous agents — interact to achieve individual or group objectives, on the basis of possibly different sets of assumptions, beliefs, preferences and objectives. For instance, agents may negotiate the purchase of goods or services from other agents, or seek information from them, or collaborate with them to achieve some common task, such as management of a telecommunications network. Recently, argumentation theory, the formal study of argument and dialogue, has been proposed for modeling agent interactions, for example by Parsons and Jennings [8], Reed [10] and Sycara [12].

Reed's work built on an influential model of human dialogues due to argumentation theorists Doug Walton and Erik Krabbe [14], and we also take their dialogue typology as our starting point. Walton and Krabbe set out to analyze the concept of commitment in dialogue, so as to "provide conceptual tools for the theory of argumentation" [14, page ix]. This led to a focus on persuasion dialogues, and their work presents formal models for such dialogues. In attempting this task, they recognized the need for a characterization of dialogues, and so they present a broad typology for inter-personal dialogue. They make no claims for its comprehensiveness.

Their categorization identifies six primary types of dialogues and three mixed types. The categorization is based upon: firstly, what information the participants each have at the commencement of the dialogue (with regard to the topic of discussion); secondly, what goals the individual participants have; and, thirdly, what goals are shared by the participants, goals we may view as those of the dialogue itself. As defined by Walton and Krabbe, the six primary dialogue types are (re-ordered from [14]):

Information-Seeking Dialogues: One participants seeks the answer to some question(s) from another participant, who is believed by the first to know the answer(s).

Inquiry Dialogues: The participants collaborate to answer some question or questions whose answers are not known to any one participant.

Persuasion Dialogues: One party seeks to persuade another party to adopt a belief or point-of-view he or she does not currently hold. These dialogues begin with one party supporting a particular statement which the other party to the dialogue does not, and the first seeks to convince the second to adopt the proposition. The second party may not share this objective.

Negotiation Dialogues: The participants bargain over the division of some scarce resource in a way acceptable to all, with each individual party aiming to maximize his or her share. The goal of the dialogue may be in conflict with the individual goals of each of the participants.[1]

Deliberation Dialogues: Participants collaborate to decide what course of action to take in some situation. Participants share a responsibility to decide the course of action, and either share a common set of intentions or a willingness to discuss rationally whether they have shared intentions.

Eristic Dialogues: Participants quarrel verbally as a substitute for physical fighting, with each aiming to win

[1] Note that this definition of Negotiation is that of Walton and Krabbe. Arguably negotiation dialogues may involve other issues besides the division of scarce resources.

the exchange. We include Eristic dialogues here for completeness, but we do not discuss them further.

Most actual dialogues — both human and agent — involve mixtures of these dialogue types, rather than being pure instances. A purchase transaction, for example, may commence with a request from a potential buyer for information from a seller, proceed to a persuasion dialogue, where the seller seeks to persuade the potential buyer of the importance of some feature of the product, and then transition to a negotiation, where each party offers to give up something he or she desires in return for something else. The two parties may or may not be aware of the different nature of their discussions at each phase, or of the transitions between phases. Indeed, even this three-phase description may be an idealization, as sub-dialogues may be embedded (to use the terminology of [14]) in each different dialogue, for example when further information is requested by either party in the midst of the negotiation phase.

Our aim in this paper is to provide a formal framework, motivated by game logic [7], for representing the five kinds of dialogue identified by Walton and Krabbe, as well as dialogues *about* dialogues.

2. DIALOGUE GAMES

Recent work in the philosophy of argumentation and in Artificial Intelligence has undertaken to develop formal models of dialogues, a discipline known as computational dialectics. Walton and Krabbe follow their typology with formal models of persuasion dialogues [14] and similar models have found application in artificial intelligence [1, 9].

A standard approach to this task is the use of dialogue-games, following the work of Hamblin [3] and MacKenzie [4]. This approach defines a dialogue game between two or more players in terms of various game rules. For instance, Amgoud and her colleagues [1] provide a syntax for negotiation dialogues between two agents, based upon MacKenzie's Dialogue Game DC [4]. This syntax enables the presentation of offers and counter-offers (formulae in some logical language) between the agents, along with arguments which support or contest these various offers. The formalism defines precisely the protocol for when and how such arguments may be presented by a participant, and how they should be handled by another participant receiving them. The formalism can therefore be readily operationalized in a computer system for agent negotiations.

Abstracting from the rules for any one game — an abstraction we might refer to as the meta-theory of dialogue formalization — we can identify several types of dialogue game rules, as follows. We assume that the issues of discussion between the agents can be represented in some logical language, whose well-formed formulae are denoted by the lower-case Roman letters, p, q, r, etc.

Commencement Rules: Rules which define the circumstances under which the dialogue commences.

Locutions: Rules which indicate what utterances are permitted. Typically, legal locutions permit participants to assert propositions, permit others to question or contest prior assertions, and permit those asserting propositions which are subsequently questioned or contested to justify their assertions. Justifications may involve the presentation of a proof of the proposition or an argument for it, and such presentations may also be legal utterances. In multi-agent system applications of dialogue games (e.g. [1]), it is common to impose rationality conditions on utterances, for example allowing agents to assert statements only when they themselves have a prior argument or proof from their own knowledge base. The dialogue game rules may also permit participants to utter propositions to which they assign differing degrees of commitment, for example: one may merely *propose* a proposition, a speech act which entails less commitment than would an *assertion* of the same proposition.[2]

Combination Rules: Rules which define the dialogical contexts under which particular locutions are permitted or not, or obligatory or not. For instance, it may not be permitted for a participant to assert a proposition p and subsequently the proposition $\neg p$ in the same dialogue, without in the interim having retracted the former assertion. Similarly, assertion of a proposition by a participant may oblige that participant to defend it following contestation by other participants.

Commitments: Rules which define the circumstances under which participants express commitment to a proposition. Typically, assertion of a claim p in the debate is defined as indicating to the other participants some level of commitment to, or support for, the claim. In a negotiation dialogue, for example, assertion of an offer may express a willingness to undertake a transaction on the terms contained in the offer. However, depending on the rules of the game, commitment may express merely that the speaker has an argument for p, and this is not necessarily the same as belief or intention.

Termination Rules: Rules which define the circumstances under which the dialogue ends.

As mentioned above, Walton and Krabbe [14] presented formal models for persuasion dialogues. No formal models yet exist for the other dialogues in the typology of the previous sub-section. Although the task of formalizing these different dialogue types is incomplete, it should be possible to define a formal dialogue-game model for any rule-governed dialogue. In the next Section, we present an abstract formalism for any dialogue game, based on these elements.

Given such formal models of each dialogue type, how do we then represent conversations which consist of multiple types? The only proposal known to us is that of Chris Reed [10], who has proposed a formalism called Dialogue Frames. Building on the Walton and Krabbe typology, a Dialogue Frame is defined as a 4-tuple, where the first element of the tuple identifies the type of dialogue; the second element, the object of the dialogue (a belief, an action-plan, a sales-contract, etc); the third element, the topic of the dialogue (understood as an element of some database related to the object); and the fourth element, the sequence of utterances made by the parties to the dialogue. Utterances are assumed taken from some dictionary agreed between the participants, along with arguments for these. Utterances can also include requests to switch to a different dialogue type, and, if agreed

[2]For example, propositions with implicitly different levels of commitment may be presented in the dialogue games of [14]; degrees of commitment are expressed explicitly in the system of [5].

by the participants, the new dialogue then continues until completed or until a switch to another type occurs. Hence, this formalism permits the functional embedding of different dialogue types, as occurs in real dialogues.

However, the fourth element of Reed's Dialogue Frame tuples present records of a dialogue (real or hypothetical), in terms of legal utterances. The representation does not specify the form of such utterances, nor the rules which govern their formation and issuance; the formalism, although admirably flexible, is descriptive and not generative. Thus, Dialogue Frames are analogous to tape-recordings of human conversations, rather than to the rules of syntax and dialogue used by the speakers in the conversations recorded. We seek a formalism which can represent such rules of syntax and dialogue — in our case, the formal dialogue game rules for each type of dialogue — as well as representing the nesting of one dialogue inside another. The next section presents our formalism for this representation.

3. FORMAL DIALOGUE FRAMEWORKS

In this Section, we present a hierarchical formalism for agent dialogues which has three levels. At the lowest level are the topics which are the subjects of dialogues. At the next level are the dialogues — information-seeking, inquiry, etc — which we represent by means of dialogue games. At the highest level we represent control dialogues, where agents decide which dialogues to enter, if any. Our motivation for this structure is the Game Logic of Rohit Parikh [7], which was developed for representing and studying the formal properties of games in multi-game contexts.

We assume throughout this Section that dialogues are being undertaken by agents from a set denoted $\mathcal{A}$, whose individual members are denoted by lower-case Roman letters, a, b, c, etc. We further assume that the agents involved are (or represent) reasonable, consenting participants in the dialogues. This assumption means that no particular dialogue may commence without the consent of all those agents participating. This is an assumption not shared by Game Logic, which sometimes permits one player to choose the type of game to be played. We do assume, however, that the participating agents have agreed to join the control-level dialogue. Another implication of the assumption that the agents are consenting and reasonable is that no agent may be *forced* to agree to a proposition or statement.

3.1 Topic layer

Topics are matters under discussion by the participating agents, and we assume that they can be represented in a suitable logic $\mathcal{L}$ with defined connectives. Topics are denoted by the (possibly-indexed) lower-case Roman letters p, q, r, etc. We assume that all the matters of interest to the participating agents can be represented by well-formed formulae in this logical language. Note that $\mathcal{L}$ may be a modal language, with, for example, temporal modalities.

3.2 Dialogue layer

At the next level in the hierarchy we model particular types of dialogues, using the meta-theory of formal dialogue games presented in Section 2. We examine each of the components of this theory in turn. Firstly, we consider Commencement Rules. Because our agents are consenting participants, a dialogue of a specific type cannot commence without the agreement of all those involved. Such agreement may itself only be reached after a dialogue concerning the desirability or otherwise of conducting such a dialogue on the specified topic at that particular time. For this reason, we model the commencement rules by means of their own dialogue, the Commencement Dialogue, which we describe when presenting the Control Layer in the next subsection.

Locutions are legal dialogue moves made by dialogue participants regarding the discussion topics, within a particular dialogue game. Such moves may include assertions, contestations, justifications, etc, and we denote them by lower-case Greek letters, θ, ϕ, etc. Because in most dialogue games these moves refer to particular topics, we sometimes write $\theta(p)$ for a move θ which concerns discussion topic p. For any dialogue game G, the set of legal locutions is denoted by Θ^G, or by Θ when only one game is under consideration. We assume that every dialogue game has a legal locution which proposes that participants interrupt the current dialogue and return to the Control Layer. This locution can be made by any participant at any time, and is an example of a metalinguistic utterance called a *Point of Order* by Hamblin [3, p. 284]. We denote this by *PROPOSE_RETURN_CONTROL*. Any debate over whether or not to undertake this return to the Control Layer is assumed itself to be undertaken in the Control Layer, since it is not part of any one dialogue type.

Combination Rules define which locutions are valid in which different dialogical circumstances. Imagine a dialogue which proceeds through successive utterances, which we may call *rounds*, numbered 1, 2, 3, ... We could think, therefore, of a dialogue as a (possibly infinite) subset of the set $\Theta \times \Theta \times \ldots \times \Theta \times \ldots$. However, the Combination rules specify that not all possible utterances are valid in every round of the dialogue, or that certain utterances are required at certain rounds. Suppose then, for each round k we define the set M^k to be that subset of utterances Θ which are valid under the combination rules at round k. Then the combination rules may be thought of as relations which define the valid utterances at round k on the basis of those utterances valid in previous rounds. In other words, each combination rule can be considered as a function R from $\Theta \times \Theta \times \ldots \times \Theta \times \ldots$ to Θ, which maps $M^1 \times M^2 \times M^3 \ldots \times M^{k-1}$ to M^k. In addition, some combination rules may specify for each locution what other locutions, if any, must have preceeded it, for it to be legally uttered. Those locutions which do not have any such preconditions constitute precisely the set of valid locutions at the first round of the dialogue, and so we have a particular combination relation which maps from Θ to Θ, and whose image is M^1. For any dialogue game G, we denote the set the combination relations by $\mathcal{R}^G$.

The representation described here captures different types of combination rules. For instance, many dialogue games (e.g. [5]) require assertions, when contested, to be then justified by the agent who made the assertion. Thus, the move $assert_a(p)$ made at one round by agent a and then followed at a subsequent round by the move $contest_b(p)$ made by agent b obliges agent a to subsequently move $justify_a(p)$. Such a combination rule can be represented by a set of combination relations which map $M^1 \times M^2 \times M^3 \ldots \times M^{k-2} \times \{contest_b(p)\}$ to $M^k = \{justify_a(p)\}$, when $assert_a(p) \in M^i$, for some $i = 1, 2, \ldots, k-2$. Of course, we would also need to specify that the execution of $contest_b(p)$ in round $k-1$ was also the first such contestation subsequent to the execution of $assert_a(p)$ in round i, or that multiple utterances of contestations of the same proposition are not legal.

We may also model rules which define Commitments, this time by means of functions similar to truth-valuation functions. For each agent $a \in \mathcal{A}$ participating in the dialogue we define a's Commitment Function CF_a as a function which maps finite subsets of the set $M^1 \times M^2 \times M^3 \ldots \times M^k \times \ldots$ to subsets of $\mathcal{L}$, by associating a set of propositions with each combination of legal dialogue moves. Those subsets of $\mathcal{L}$ which are contained in the image of CF_a are called Commitment Stores for a. We denote the restriction of CF_a to the k-th round by CF_a^k, and the set of possible commitment stores of agent a at round k, by $PCS_a^k \subseteq \mathcal{P}(\mathcal{L})$. Thus PCS_a^k, is the image of CF_a^k on $M^1 \times M^2 \times M^3 \ldots \times M^k$. We denote the set of commitment functions for dialogue G by $\mathcal{CF}^G$.

Finally, we consider Termination Rules, which allow or require the dialogue to end upon achieving certain conditions. For example, a Persuasion Dialogue may end when all those involved accept the proposition at issue. Thus we can model termination rules in a similar fashion to combination rules, by means of functions T which map valid combinations of utterances to the set $\{0, 1\}$, where the symbol 1 denotes the termination of the dialogue and the symbol 0 its continuation. That is, each function T maps finite subsets of $M^1 \times M^2 \times M^3 \ldots \times M^k \times \ldots$ to $\{0, 1\}$. For any dialogue game G, we denote the set of termination relations by $\mathcal{T}^G$.

A dialogue may also terminate when all the participants agree to so terminate it. This may occur even though the dialogue may not yet have ended, for instance, when a persuasion dialogue does not result in all the participants accepting the proposition at issue. As with the Commencement Dialogue, we can model this with a specific type of control-level dialogue, which we term the Termination Dialogue. This is discussed at the Control Layer in the next subsection.

Given a set of participating agents $\mathcal{A}$, we then define a formal dialogue G as a 4-tuple $(\Theta^G, \mathcal{R}^G, \mathcal{T}^G, \mathcal{CF}^G)$, where Θ^G is the set of legal locutions, $\mathcal{R}^G$ the set of combination relations, $\mathcal{T}^G$ the set of termination relations, and $\mathcal{CF}^G$ the set of commitment functions. We omit the superscript G if this causes no confusion.

3.3 Control layer

The control layer seeks to represent the selection of specific dialogue types and transition between these types. In Parikh's Game Logic [7], this selection is undertaken by one or other of the participants deciding autonomously, and this is represented by the game sort. Because our application domain involves consenting agents, the selection of dialogue-type may itself be the subject of debate and possibly even negotiation between the agents concerned. Our formalism therefore needs to represent such dialogue. As suggested in the description of the Dialogue Layer, we do this by defining certain control dialogues, namely the Commencement Dialogue and The Termination Dialogue. These can be modelled by formal dialogue games using the same structure as for the dialogues presented in the previous subsection.

The Control Layer is defined in terms of the following components. We first define a finite set of dialogue-types, called *Atomic Dialogue-Types*, which include the five dialogues of the Walton and Krabbe typology. Atomic Dialogue-types are denoted by the (possibly indexed) upper case Roman letters G, H, J, K, etc. To denote a dialogue conducted according to dialogue-type G and concerning a specific proposition p, we write $G(p)$. When no confusion would be caused we omit the argument and write simply G. We denote the set of atomic dialogue-types by Π_0.

We next define *Control Dialogues*, which are dialogues that have as their discussion subjects not topics, but other dialogues, and we can define them formally as 4-tuples in the manner of subsection 3.2. They include the Commencement and Termination Dialogues for any dialogue $G(p)$, which we denote by $BEGIN(G(p))$ and $END(G(p))$ respectively, and the Control Dialogue itself, denoted $CONTROL$. We denote the set of control dialogues by Π_{CON}. If a $BEGIN(G(p))$ dialogue leads to agreement between the participating agents to commence a $G(p)$ dialogue, then the $BEGIN(G(p))$ dialogue immediately terminates, and the specific $G(p)$ dialogue begins. In this case, from the moment of termination of $BEGIN(G(p))$ to the moment following termination of $G(p)$, the dialogue $G(p)$ is said to be *open*. Following termination of $G(p)$, $G(p)$ is said to be *closed*.

Also defined as dialogues are the following combinations of atomic or control dialogues or any legal combination thereof, which we term *Dialogue Combinations*:

Iteration: If G is a dialogue, then G^n is that dialogue which consists of the n-fold repetition of G, each occurrence being undertaken until normal completion.

Sequencing: If G and H are dialogues, then $G; H$ is that dialogue which consists of undertaking G until its normal completion and then immediately undertaking H.

Parallelization: If G and H are dialogues, then $G \cap H$ is that dialogue which consists of undertaking G and H simultaneously, until both are completed normally.[3]

Embedding: If G and H are both dialogues, and $\Phi \subseteq M^1 \times M^2 \ldots \subseteq \Theta^G \times \Theta^G \ldots$ is a sequence of legal locutions in G, then $G[H|\Phi]$ is that dialogue which consists of undertaking G until Φ has been executed, and then switching immediately to dialogue H which is undertaken until completion, whereupon dialogue G resumes from immediately after the point where it was interrupted and continues until normal completion. In the time between G commencing and concluding, dialogue G remains open.

Testing: If p is a wff in $\mathcal{L}$, then $< p >$ is a control dialogue which consists of testing the truth of p. If p is found to be false then the current open dialogue at the lowest embedded level (or dialogues, if parallel dialogues are open at the same level) immediately ends; otherwise, the current dialogue (or dialogues) continues.

We denote by Π the closure of the set $\Pi_0 \cup \Pi_{CON}$ under the dialogue combination operations defined here.

We next define the rules for commencement of the *CONTROL* dialogue, which commences precisely when a participating agent in the set of agents $\mathcal{A}$ commences the *BEGIN* $(G(p))$ dialogue for some G and p. The $BEGIN(G(p))$ dialogue commences with a locution which seeks the consent of the other participating agents to commence a dialogue of type G over proposition p. Immediately upon execution of this consent-seeking locution, the Control Layer is said to be *open*.

[3] As an example of parallel dialogues, complex human inquiries such as air-crash investigations are often divided into simpler, parallel sub-inquiries.

An open Control dialogue terminates precisely when either[4] (i) there are no open dialogues apart from the *CONTROL* dialogue itself, or (ii) the participating agents all agree to terminate the *CONTROL* dialogue, by undertaking and completing an *END*(*CONTROL*) dialogue.

These various components at the Control level form the basis for Agent Dialogue Frameworks, which we define in the next subsection.

3.4 Agent dialogue frameworks

We define an Agent Dialogue Framework (ADF) as a 5-tuple $(\mathcal{A}, \mathcal{L}, \Pi_0, \Pi_{CON}, \Pi)$, where $\mathcal{A}$ is a set of agents, $\mathcal{L}$ is a logical language for representation of discussion topics, Π_0 is a set of atomic dialogue-types, Π_{CON} a set of Control dialogues and Π the closure of $\Pi_0 \cup \Pi_{CON}$ under the combination rules presented in the previous subsection. To reprise, each formal dialogue in $\Pi_0 \cup \Pi_{CON}$ is defined as a 4-tuple, $G = (\Theta^G, \mathcal{R}^G, \mathcal{T}^G, \mathcal{CF}^G)$, where: Θ^G is the set of legal locutions, $\mathcal{R}^G$ the set of combination relations, $\mathcal{T}^G$ the set of termination relations, and $\mathcal{CF}^G$ the set of commitment functions of the dialogue type G. In the next two subsections we explore some of the formal properties of this framework: firstly, can the framework be used for automatic generation of dialogues between autonomous agents; and secondly, the circumstances under which dialogues terminate.

3.5 Generating dialogues

The framework we have presented is defined in terms of rules of dialogue games and can be used to generate dialogues if we have procedures which could automatically generate each of the types of dialogues if and when required. We examine how this might occur for each of the five atomic dialogue types in the Walton and Krabbe typology.

Information-Seeking Dialogues: An agent a may be pre-programmed as follows: If, in the course of a dialogue, a realizes there is some proposition p for which it requires, but does not know, the truth-value, then a automatically seeks permission to commence an information-seeking dialogue concerning p. Any other agent who knows the truth-value of p can be programmed to agree to such a dialogue and, within it, to respond with the appropriate truth-value. If questioned further, b can present the proof of or the argument for p to a.

Inquiry Dialogues: A similar line of reasoning applies to inquiry dialogues, except that here agents pool their knowledge and also potentially their reasoning capabilities (if, for example, they are using logics with different rules of inference).

Persuasion Dialogues: We can imagine that agents a and b are programmed as follows: If a accepts the truth of some proposition p and requires that b also accepts its truth (for example, to support some joint goal they are collaborating on), then a may seek consent for a persuasion dialogue for p. If b already accepts the truth of p, it then says so to a and the dialogue is quickly concluded. If b does not initially accept the truth of p, then b should accept a proof (or an undefeated argument) for p when presented by a, provided b is rational and reasonable. Provided a is rational, a should have such a proof of (or argument for) p before it believed p to be true.

Negotiation Dialogues: In Walton and Krabbe's definition, negotiation dialogues arise when agents wish to divide a scarce resource between themselves. If divisions of the resource can be quantified, and if each agent has knowledge of their own utilities with regard to these possible divisions, and the utilities are partially ordered, then a cake-cutting algorithm, such as that described by Parikh [7, Section 5], could be used to generate agent locutions. Note that one agent's utilities need not be known to the other agents, and the utilities of different agents need not be commensurate.

Deliberation Dialogues: A deliberation dialogue can be initiated automatically whenever an agent believes that the group of agents needs to jointly decide on a course of action. And, if a proposal for action is presented by some agent once inside such a dialogue, this proposal could be considered rationally by each of the other agents (assuming as before they have partially-ordered utilities with respect to the features of the proposal, and assuming each agent knows its own utilities). Thus a proposal could be discussed inside the dialogue, and revisions proposed, based on the individual-agent utilities of each proposal.

Thus each of the atomic dialogue types can be generated, and as a result it is possible to generate all dialogue-types, by simple inspection of the Dialogue Combination Rules presented in Section 3.3.

3.6 Dialogue termination

Under what circumstances does a dialogue terminate? In exploring this question, we can distinguish between two types of termination. *Regular termination* occurs when a dialogue achieves its objectives, for example when all parties to a Persuasion Dialogue are persuaded to accept or commit to the proposition at issue. Amgoud and colleagues [1] defined what constitutes such termination for each of the five Walton and Krabbe dialogue-types, where these dialogues are instantiated with MacKenzies's game DC [4] and a standard argumentation model. In each case, regular termination may be considered to occur once specific formulae are contained in the Commitment Stores of some or all of the participants. In a Persuasion Dialogue, for example, a proposition p will be accepted by each of the participants with the utterance of a specific locution, such as $Assert(p)$ or $Accept(p)$; the Commitment rules will then specify the insertion of proposition p into the respective commitment stores. Thus, asking the question: *"Will a dialogue of Type T concerning a proposition p terminate?"* is equivalent to asking: *"Is there a finite integer k such that the formulae p can be found in the appropriate commitment store(s) for agents engaged in dialogue-type T after round k?"*. The question of regular termination thus becomes one of satisfiability.

However, because we are modeling dialogues in which participants may enter and or leave any dialogue at any time, a specific dialogue may terminate before its objectives have been achieved. For example, an agent may be engaged in simultaneous purchase negotiations with multiple vendors for the same product, as in the aircraft purchase example

[4] Note that we are assuming agents do not engage in non-cooperative or unreasonable behaviour.

of [13, Ch. 8]; one of these negotiations may achieve regular termination before the others do, and so the agent concerned could summarily terminate those others. We may refer to such termination as *Irregular termination.* One way to capture this would be introduce a dummy topic formula, indexed by the dialogue type and the topic, which is inserted into an agent's commitment store for a dialogue whenever that agent agrees to commence that dialogue. If the agent subsequently decides to exit the dialogue, this variable is withdrawn from the commitment store. Irregular termination can thereby also be seen as a satisfiability problem.

The more interesting question regarding termination is under what conditions are dialogues guaranteed to terminate. Here, various assumptions regarding the nature of the agents involved in the dialogue will be important, e.g. that they are rational, well-intended, non-whimsical, etc; this is the subject of further research by the authors.

4. EXAMPLE

We illustrate the framework with a dialogue between a potential buyer and a potential seller of used motor cars. The example shows how a dialogue may evolve as information is sought and obtained by one or other party, and how dialogues may be embedded in one another. For ease of understanding, we write the example in a narrative form, annotating the dialogue moves whenever a sub-dialogue opens or closes. The two participants, a Potential Buyer and a Potential Seller, are denoted by **B** and **S** respectively.

B: BEGIN(INFOSEEK(New_car_purchase))

*Potential Buyer **B** requests commencement of an information-seeking dialogue regarding purchase of a second-hand car. The CONTROL Dialogue opens.*

S: AGREE(INFOSEEK(New_car_purchase))

*Potential Seller **S** agrees. INFORMATION-SEEKING Dialogue 1 opens.*

B: REQUEST(Cars,Models)

B asks what cars and models S has available, using legal locutions in the INFORMATION-SEEKING Dialogue.

S: PROPOSE_RETURN_CONTROL

Return to CONTROL Dialogue.

B: AGREE(RETURN_CONTROL)

S: BEGIN(INFOSEEK(Budget))

***S** requests commencement of an Information-Seeking dialogue regarding the budget B has available.*

B: AGREE(INFOSEEK(Budget))

***B** agrees. INFORMATION-SEEKING Dialogue 2 opens, embedded in 1.*

S: REQUEST(Budget)

B: Budget = $ 8000.

INFORMATION-SEEKING Dialogue 2 closes. Return to INFORMATION-SEEKING Dialogue 1.

S: (Cars, Models) = {(Mazda, MX3), (Mazda, MX5), (Toyota, MR2)}

INFORMATION-SEEKING Dialogue 1 closes. Return to CONTROL Dialogue.

S: BEGIN(INFOSEEK((Purchase_Criteria))

***S** requests Information-Seeking dialogue over **B's** purchase criteria.*

B: AGREE(INFOSEEK((Purchase_Criteria))

INFORMATION-SEEKING Dialogue 3 opens.

S: REQUEST(Purchase_Criteria)

B: Purchase_Criterion_1 = Price, Purchase_Criterion_2 = Mileage, Purchase_Criterion_3 = Age

INFORMATION-SEEKING Dialogue 3 closes. Return to CONTROL Dialogue.

S: BEGIN(PERSUASION(Make);
PERSUASION(Condition_of_Engine);
PERSUASION(Number_of_Owners))

***S** requests a sequence of three Persuasion dialogues over the purchase criteria Make, Condition of the Engine, and Number of Owners.*

B: AGREE(PERSUASION(Make);
PERSUASION(Condition_of_Engine);
PERSUASION(Number_of_Owners))

PERSUASION Dialogue 1 in the sequence of three opens.

S: Argues that *"Make"* is the most important purchase criterion, within any budget, because a typical car of one Make may remain in better condition than a typical car of another Make, even though older.

B: Accepts this argument.

*PERSUASION Dialogue 1 closes upon acceptance of the proposition by **B.** PERSUASION Dialogue 2 opens.*

S: Argues that that *"Condition_of_Engine"* is the next most important purchase criterion.

B: Does not accept this. Argues that he cannot tell the engine condition of any car without pulling it apart. Only **S**, as the Seller, is able to tell this. Hence, **B** must use *"Mileage"* in place of for *"Condition_of_Engine."*

*PERSUASION Dialogue 2 closes with neither side changing their views: **B** does not accept "Condition_of_Engine" as the second criterion, and **S** does not accept "Mileage" as the second criterion. PERSUASION Dialogue 3 opens.*

S: Argues that the next most important purchase criterion is *"Number_of_Owners."*

B: Argues that *"Mileage"* and *"Age"* are more important than *"Number_of_Owners."*

S: Argues that *"Number_of_Owners"* is important because owners who keep their cars for a long time tend to care for them more than owners who change cars frequently.

B: PROPOSE_RETURN_CONTROL

Return to CONTROL Dialogue.

S: AGREE(RETURN_CONTROL)

B: BEGIN(NEGOTIATION(Purchase_criteria)

S: AGREE(NEGOTIATION(Purchase_criteria)

NEGOTIATION Dialogue 1 (embedded in PERSUASION Dialogue 3) opens.

B: Says he will accept *"Number_of_Owners"* as the third purchase criterion in place of *"Age"* if **S** accepts *"Mileage"* in place of *"Condition_of_Engine"* as the second.

S: Agrees.

NEGOTIATION Dialogue 1 closes. PERSUASION Dialogue 3 resumes and closes immediately. Return to CONTROL Dialogue.

One feature of this example is that it shows a Negotiation Dialogue embedded in a Persuasion Dialogue, an embedding not everyone considers valid (e.g. [14]). We believe that the desirability of particular combinations of dialogue-types should be a matter for the participants to the dialogues to decide at the time of the dialogue. The formalism we have presented here enables such decisions to be made.

5. A GAME-THEORETIC SEMANTICS

In this section, we present a semantics for the five different types of dialogues based on the notion of abstract games. Thus our approach is in the game-theoretic tradition associated with Jaako Hintikka, but which is increasingly applied in artificial intelligence, e.g. [9]. We assume as above an underlying logical language, whose well-formed formulae are denoted by lower-case Roman letters, p, q, etc. For each such wff, p, we associate a game between two players, V (for *Verifier*) and F (for *Falsifier*), which we label $G(p)$. We assign p the value "true" if and only if there is a winning strategy for V in the game $G(p)$. What is meant by a winning strategy may be defined differently for different types of games or for different application domains. For example, a winning strategy may be that V is able to provide a deductive proof for p in the logical language concerned. By contrast, in argumentation-based games a winning strategy may be defined as the capability of V to provide a set of arguments for p which defend themselves against all contestations possibly articulated by F, e.g. [1, 9]. The argumentation definition is analogous to the conduct of real-world legal proceedings, where claims are accepted as true if and only if they survive attempts to defeat them in validly-constituted and appropriately-conducted legal forums.

With this understanding of "truth", we provide a game-theoretic interpretation of each of the five dialogue types of Walton and Krabbe. For simplicity, we assume each dialogue is undertaken by two agents, denoted a and b; the general case is an obvious extension. We also assume that both agents accept this game-theoretic semantics.

Information-Seeking Dialogues: a asks b the truth-status of some proposition p. The proposition will be true iff V has a winning strategy in the game $G(p)$. Whether or not V has such a strategy in $G(p)$ is a fact unknown to a, but may be known to b.

Inquiry Dialogues: a and b both seek to know the truth-status of p. As for the previous dialogue, p will be true iff V has a winning strategy in the game $G(p)$. Neither agent knows at the outset of the dialogue whether V has such a strategy, but together they may be able to determine if this is the case.

Persuasion Dialogues: a seeks to persuade b of the truth of p. Here, a believes that p is true and hence that V has a winning strategy in the game $G(p)$. Agent b is not able to show this at the outset of the dialogue. If a can convince b that V does have such a strategy, then (because b accepts the game-theoretic semantics), b will then accept the truth of p. Note that a may believe that V has a winning strategy without being able to exhibit that strategy, for example if its proof of the existence of the winning strategy is non-constructive. Agent b may or may not accept such proofs.

Negotiation Dialogues: a and b seek to divide some scarce resource between them. Wooldridge and Parsons [15] propose a general framework for representation of multi-agent negotiations in logical languages, in which the two agents make successive offers and counter-offers in a sequence of n moves:

$$(p_a^1, p_b^1, p_a^2, p_b^2, \dots ,\ p_a^n, p_b^n)$$

Here, p_j^k represents the offer made by Agent j in move k.[5] Success in such a negotiation occurs when $p_a^n \Leftrightarrow p_b^n$, where $\Leftrightarrow$ denotes logical equivalence in the underlying language $\mathcal{L}$. Our game theoretic interpretation is that success is achieved after n moves when V has a winning strategy in the game $G(p_a^n \Leftrightarrow p_b^n)$.

Deliberation Dialogues: Agents a and b seek to decide a course of action in some situation. These dialogues can be represented in a similar fashion to negotiation dialogues, where the statements p_j^k denotes the proposal for action made by Agent j in round k. As with negotiation dialogues, success is achieved after n moves when V has a winning strategy in the game $G(p_a^n \Leftrightarrow p_b^n)$.

As mentioned above, this semantic interpretation of dialogues is in terms of abstract games. We have not identified the nature of the games $G(p)$, nor defined the winning strategies in these games. It is possible that both games and strategies may differ by dialogue-type and seem likely to be domain dependent. Provided the participants agree on the particular instantiations appropriate to their domain, there is no problem with this level of abstraction.

6. DISCUSSION

The major contribution of this paper has been to develop a formal and potentially-generative language for dialogues between autonomous agents which admits different types of dialogues. Abstracting from recent work in philosophy and artificial intelligence developing formal dialogue games, we have proposed a meta-theory of such dialogue games, and used this as the basis of our Agent Dialogue Framework. A second contribution of this paper has been to provide a

[5] Note that there may be other legal utterances besides offers and counter-offers, for instance, questions regarding offers, and justifications for them. Hence, the moves listed here may only be a subset of all the rounds of a negotiation dialogue.

simple game-theoretic semantics for each of these dialogue types. Note that designers of multi-agent systems which support dialogues between agents have used slightly different terminology to that presented here. Noriega and Sierra [6], for example, call formal dialogue games *institutions* and call combinations of locutions *dialogical frameworks*. Their definitions permit the legal locutions to differ from one agent to another, as in a marketplace where different participants may play different roles, e.g. buyers, sellers, auctioneers, etc. The same is true of the subsequent refinement of their model presented in [11], although here the term *dialogical framework* refers to a broader concept, closer to our use of *dialogue game*. In both cases, however, the intention of the work was to provide a formal representational framework for agents undertaking negotiation dialogues, defined broadly to include information-seeking and persuasion locutions; neither framework enables representation or embedding of different types of dialogue. The work of Dignum and colleagues [2] uses formal models of dialogues to represent team-formation by agents in co-operative problem solving contexts. This work does allow embedding of some dialogues within others, although not in a generic way.

We see a number of advantages of the Agent Dialogue Framework to represent agent dialogues. Firstly, the formalism provides a single, unifying framework for representing disparate types of dialogue, including the dialogues in the typology of Walton and Krabbe [14]. Although most work to date in agent interactions has involved some form of negotiation, other types of dialogue are arguably as important to the development of full agent societies. Indeed, as shown above, many negotiations may involve other types of dialogues. Secondly, the use of an explicit representation for the dialogue-type in the ADF means that the nature of the current dialogue being undertaken is always known to the participants. This is not always evident in human conversations, and much disagreement may be due to misunderstandings of the nature of the dialogue being undertaken [14]. Thirdly, the ADF formalism is modular, so that other dialogue types may be inserted readily into the framework. Similarly, our formalism permits incorporation of specialized sub-types of dialogues, for instance, public policy debates over environmental risk assessment; these comprise a complex combination of aspects of inquiry, persuasion, deliberation and negotiation dialogues in a specialized domain.

These three advantages are also features of Reed's Dialogue Frames [10]. However, a fourth advantage, not shared by Dialogue Frames, is that the ADF can be used to generate dialogues. Also, because it is based on a meta-theory of dialogue-games, the ADF enables us to use the recent work in computational dialectics in designing such games. A final advantage arises from the use of a logical formalism, which permits us to study the formal properties of these systems, for example, their computational complexity as in [15]. The issue of participant strategies in dialogue games is another area potentially amenable to formal analysis.

Acknowledgments

This work was partly funded by the EU SLIE Project (IST-1999-10948) and by the British EPSRC through a PhD studentship. We also thank Mike Wooldridge for comments.

7. REFERENCES

[1] L. Amgoud, N. Maudet, and S. Parsons. Modelling dialogues using argumentation. In E. Durfee, editor, *Proceedings of the Fourth International Conference on Multi-Agent Systems*, pages 31–38, Boston, MA, USA, 2000. IEEE Press.

[2] F. Dignum, B. Dunin-Kęplicz, and R. Verbrugge. Agent theory for team formation by dialogue. In C. Castelfranchi and Y. Lespérance, editors, *Seventh Workshop on Agent Theories, Architectures, and Languages*, pages 141–156, Boston, USA, 2000.

[3] C. L. Hamblin. *Fallacies*. Methuen and Co Ltd, London, UK, 1970.

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

[5] P. McBurney and S. Parsons. Risk agoras: Dialectical argumentation for scientific reasoning. In C. Boutilier and M. Goldszmidt, editors, *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, Stanford, CA, USA, 2000. UAI.

[6] P. Noriega and C. Sierra. Towards layered dialogical agents. In J. P. Muller, M. J. Wooldridge, and N. R. Jennings, editors, *Intelligent Agents III*, pages 173–188, Berlin, Germany, 1997. Springer Verlag.

[7] R. Parikh. The logic of games and its applications. *Annals of Discrete Mathematics*, 24:111–140, 1985.

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

[9] H. Prakken and G. Sartor. Modelling reasoning with precedents in a formal dialogue game. *Artificial Intelligence and Law*, 6:231–287, 1998.

[10] C. Reed. Dialogue frames in agent communications. In Y. Demazeau, editor, *Proceedings of the Third International Conference on Multi-Agent Systems*, pages 246–253. IEEE Press, 1998.

[11] C. Sierra, N. R. Jennings, P. Noriega, and S. Parsons. A framework for argumentation-based negotiations. In M. P. Singh, A. Rao, and M. J. Wooldridge, editors, *Intelligent Agents IV*, pages 177–192, Berlin, Germany, 1998. Springer Verlag.

[12] K. Sycara. Argumentation: Planning other agents' plans. In *Proceedings of the Eleventh Joint Conference on Artificial Intelligence*, pages 517–523, 1989.

[13] R. van Eijk. *Programming Languages for Agent Communications*. PhD thesis, Department of Computer Science, Utrecht University, The Netherlands, 2000.

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

[15] M. Wooldridge and S. Parsons. Languages for negotiation. In W. Horn, editor, *Proceedings of the Fourteenth European Conference on Artificial Intelligence*, pages 393– 397, Berlin, Germany, 2000. IOS Press.