Games that agents play: A formal framework for dialogues between autonomous agents

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

24 September 2001

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 is computational and its modular nature enables different types of dialogues to be represented.

Keywords: Autonomous Agents, Computational Dialectics, Dialogue Games.

1. Introduction

Autonomous intelligent software agents have become an important new paradigm in computer science [16]. 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, seek information from them, or collaborate with them to achieve some common task. Recently, argumentation theory, the formal study of human argument and dialogue, has been proposed for modeling agent interactions, for example by Parsons and Jennings [27, 28], and Reed [32].

Reed's work built on a typology of human dialogues due to Walton and Krabbe [36], and we start from the same typology, which we summarize in Section 2. Several of these atomic dialogue types have been modeled by means of formal dialogue games, adopted from the philosophy of argumentation, and Section 2 continues with a presentation of the generic elements of such games. Our ultimate objective in this work is to represent complex dialogues occurrences which may involve more than one atomic type, e.g. dialogues which may contain sub-dialogues embedded within them. We are drawn, as was Reed [32], to an hierarchical representation. Our formalism is presented in Section 3, and it integrates a dialogue game model of atomic types with a for-



© 2002 Kluwer Academic Publishers. Printed in the Netherlands.

malism for combining multiple dialogues of potentially different types. Our formalism also enables representation of dialogues *about* dialogues. In Section 4, we illustrate our formalism with an example involving a dialogue occurrence of multiple types between a potential buyer and a potential seller of second-hand cars. Finally, Section 5 discusses related and future research.

2. Dialogues and Dialogue Games

2.1. Types of dialogues

An influential model of human dialogues is the typology of primary dialogue types of argumentation theorists Doug Walton and Erik Krabbe [36]. This categorization is based upon the information the participants have at the commencement of a dialogue (of relevance to the topic of discussion), their individual goals for the dialogue, and the goals they share. **Information-Seeking Dialogues** are those where one participant seeks the answer to some question(s) from another participant, who is believed by the first to know the answer(s). In Inquiry Dialogues the participants collaborate to answer some question or questions whose answers are not known to any one participant. Persuasion Dialogues involve one participant seeking to persuade another to accept a proposition he or she does not currently endorse. In **Negotiation Dialogues**, the participants bargain over the division of some scarce resource. Here, the goal of the dialogue — a division of the resource acceptable to all — may be in conflict with the individual goals of the participants. Participants of Deliberation Dialogues collaborate to decide what action or course of action should be adopted in some situation. Here, participants share a responsibility to decide the course of action, or, at least, they share a willingness to discuss whether they have such a shared responsibility. Note that the best course of action for a group may conflict with the preferences or intentions of each individual member of the group; moreover, no one participant may have all the information required to decide what is best for the group. In Eristic Dialogues, participants quarrel verbally as a substitute for physical fighting, aiming to vent perceived grievances. Since these dialogues are not generally rule-governed, we do not discuss them further in this paper.

Most actual dialogue occurrences — both human and agent — involve mixtures of these dialogue types. 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. Instances of individual dialogue types contained entirely within other dialogue types are said to be *embedded* [36].

2.2. Dialogue games

Formal dialogue games are interactions between two or more players, where each player "moves" by making utterances, according to a defined set of rules. Although they date from the time of Aristotle [11], they have found recent application in philosophy, in computational linguistics and in Artificial Intelligence (AI). In philosophy, dialogue games have been used to study fallacious reasoning [11, 20] and to develop a game-theoretic semantics for intuitionistic and classical logic [19]. In linguistics, they have been used to explain sequences of human utterances [17], with subsequent application to machine-based natural language processing and generation [15], and to human-computer interaction [2]. Within AI, they have been applied to modeling complex human reasoning, for example in legal domains [31], and as the basis for systems to support public argumentation [9]. More recently, dialogue games have found application in AI as the basis for protocols for interactions between autonomous software agents, e.g. [1, 5, 33].

Formal dialogue-game models have been presented for several of the atomic dialogue types in the typology of Walton and Krabbe [36]. These include: persuasion dialogues [22, 30, 36]; information-seeking dialogues [15]; negotiations [1, 33]; and deliberations [14]. In addition, game formalisms have been proposed for certain combinations of atomic types, such as for the formation of teams [5] and of collective intentions [6], both of which involve combinations of persuasion and negotiation dialogues. Building on these particular game models and on attempts to abstract general games [3, 21, 30], we can identify several types of dialogue game rules. We first assume that the topics 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. The components of a dialogue game are then:

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. 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.¹

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.

Commitments: Rules which define the circumstances under which participants express commitment to a proposition. Typically, the 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. Since [11], formal dialogue systems typically establish and maintain public sets of commitments, called commitment stores, for each participant; these stores are usually non-monotonic, in the sense that participants can also retract committed claims, although possibly only under defined circumstances.

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

Two comments are appropriate on this generic model of dialogue games. Firstly, in the computational linguistics tradition, dialogue games have been used to explain and to generate sequences of (human or machine) utterances [15, 17]. Here, the intentions of the speaker are important, and so play a key role in an influential model of dialogue [10]. Similar considerations have led designers of multi-agent systems to link utterances in a dialogue with the mental states of the participants, as in the operational semantics for agent speech acts of [4]. This has led some designers of dialogue game protocols, e.g. [1], to impose conditions on utterances, allowing agents to assert a statement only when they themselves believe it. However, such rationality conditions

For example, propositions with implicitly different levels of commitment may be presented in the dialogue games of [36]; degrees of commitment are expressed explicitly in the system of [22].

can never be completely verified, as it will always be possible to design a sufficiently-clever agent able to simulate insincerely any required mental state [37]. One response to this problem is the social semantics of [35], in which agents declare publicly their mental states, for example their beliefs and intentions, in the dialogue, and these declarations are used as a semantics for the speech acts uttered. However, even these declarations may be insincere. In contrast, here we suggest that agent dialogue protocols should be defined in purely syntactical terms, so that conformance with the protocol may always be verified by observing actual agent utterances.² Indeed, because our intended application domains are computational systems, non-conforming utterances can be precluded from being broadcast to other participants.

A second comment relates to commitments. In the philosophical tradition of dialogue games [11, 20, 36], commitments are understood as solely dialogical, e.g. speakers who utter an assertion locution are required to defend these assertions if subsequently questioned or challenged. The statements uttered may bear no relationship with any reality external to the dialogue, so, for example, speakers may not necessarily believe statements they endorse in the dialogue. However, the use of dialogue games to model human discourses or as protocols for agent interactions, leads to a different understanding of a commitment, namely as a statement with some external referant. In a negotiation dialogue, for instance, the utterance of an offer may express a willingness to undertake a subsequent transaction on the terms contained in the offer. For this reason, we distinguish between dialogical commitments, which incur burdens on the speaker only inside the dialogue, and semantic commitments, which incur burdens on the speaker in the world beyond the dialogue. This distinction will be useful later.

As mentioned above, dialogue game models have been articulated for most of the atomic types of dialogues in the typology of Walton and Krabbe. Given these, how may we then represent dialogue occurrences which consist of combinations of different types? The only proposal known to us is that of Reed [32], who defines a formalism called a Dialogue Frame. This is a 4-tuple, where the first element 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 participants to the dialogue during its actual occurrence. Utterances are statements assumed taken from some dictionary agreed between the participants, along with arguments for these statements. Utterances

² This protocol property is termed *externalization* in [13].

can also include requests to switch to a different dialogue type, and, if agreed by the participants, the new dialogue type then continues until completed or until a switch to another type occurs. Hence, this formalism permits both the functional embedding of different dialogue types and sequential combinations of different dialogues or dialogue types.

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

3. Formal Dialogue Frameworks

We now present a three-level hierarchical formalism for agent dialogues. At the lowest level are the topics which are the subjects of dialogues. At the next level are the dialogues themselves — instantiations of persuasions, inquiries, etc, and combinations of these — which we represent by means of formal dialogue games. At the highest level are control dialogues, where agents decide which dialogues to enter, if any. Our motivation for this structure is the Game Logic of Rohit Parikh [26, 29], which was a dynamic logic formalism [12] developed for representing and studying the formal properties of games in multigame contexts. However, while Game Logic has provided the starting point, our formalism differs from it in several aspects.

We assume throughout this Section that dialogues are being undertaken by finite set of distinct software agents, 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. One implication of this assumption is 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 unilaterally 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 incur a commitment.

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} . Topics are denoted by the lower-case Roman letters p, q, r, etc. We assume that all the matters of interest to the participating agents can be represented in this logical language. Topics may refer either to real-world objects or to states of affairs, and the formalism presented below can accommodate either interpretation. Note that \mathcal{L} may be a modal language, containing for instance temporal or deontic modalities.

3.2. Dialogue layer

At the next level in the hierarchy we model particular types of dialogues, using the generic theory of formal dialogue games presented in Section 2. We examine the components of this theory in turn. Firstly, consider Commencement Rules. Because our agents are consenting participants, a particular dialogue occurrence 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; this is described with the Control Layer in the next subsection. As explained there, agreement reached during a commencement dialogue concerning a particular type of dialogue on a specified topic leads to the immediate commencement of that dialogue-type on that topic, which is then said to be open.

Next, 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. We assume that all dialogue games contain a rule which asserts that participants to a dialogue may only utter locutions in the dialogue while the dialogue is open. 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 to the participants that they interrupt the current dialogue and return to the Control Layer. This locution can be made by any participant at any

time.³ We denote this locution 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 a generic dialogue not part of any one dialogue type.

Thirdly, 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 $1, 2, 3, \ldots$ We could think, therefore, of a dialogue of k rounds as a sequence (of length k) from the cross-product set Θ^k . We only consider finite dialogues, but they may be arbitrarily long. 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. The utterances valid at round k will depend upon what utterances were made in prior rounds, so M^k will be a function of all of M^1, M^2, \dots, M^{k-1} . We may view the combination rules at each round as (possibly multiple-valued) functions which define the valid utterances at round k on the basis of those utterances valid in previous rounds. The valid utterances at any round k will be the intersection of the images of the set of combination rules at that round. In other words, each combination rule, R_i^k , at round k can be considered as a function R_i^k from Θ^k to 2^{Θ} , such that

$$M^k = \bigcap_i Image(R_i^k(M^1 imes M^2 imes M^3 \ldots imes M^{k-1}))$$

In addition, some combination rules may specify for each locution what other locutions, if any, must have preceded 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 function which maps Θ to 2^{Θ} , and whose image is M^1 . For any dialogue game G, we denote the set of combination functions by \mathcal{R}^G .

We can readily see how the representation described here captures different types of combination rules. For instance, many dialogue games (e.g. [22]) 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 functions which map $M^1 \times M^2 \times M^3 \dots \times M^{k-2} \times M^{k-2}$

This is an example of a metalinguistic utterance called a *Point of Order* in [11, p. 284].

 $\{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 the first such contestation subsequent to the execution of $assert_a(p)$ in round i, or that multiple utterances of contestations of a proposition are not legal.

Next, we may also model rules which define Commitments, this time by means of functions similar to truth-valuation functions in logic. 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 sequences in $M^1 \times M^2 \times M^3 \ldots \times M^k \times \ldots$ 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 CF_a^k is the function CF_a restricted to the domain $M^1 \times M^2 \times M^3 \ldots \times M^k$, and its image is PCS_a^k . We denote the set of commitment functions for dialogue G by \mathcal{CF}^G .

Finally, we consider Termination Rules. These are rules which allow or require the dialogue to end upon achieving certain conditions. For example, a Persuasion Dialogue may end when all the agents involved accept the proposition at issue. We can therefore 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 sequences in $M^1 \times M^2 \times M^3 \dots \times M^k$ to $\{0,1\}$, for arbitrary k. For any dialogue game G, we denote the set of termination functions 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 with 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 functions, \mathcal{T}^G the set of termination functions, 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 Game Logic [26], 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 between the agents concerned. Our formalism therefore needs to represent such debate. 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 modeled by formal dialogue games using the same structure as for the dialogues presented above.

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, non-Eristic, 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), also called an instantiation of dialogue-type G by topic p. Sometimes we write simply G. We denote the set of atomic dialogue-types by Π_{Atom} .

We next define Control Dialogues, which are dialogues that have as their discussion subjects not topics, but other dialogues, and we 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. Another control dialogue is the dialogue in which the participating agents discuss whether to terminate the Control Layer itself, a dialogue we denote by END(CONTROL). We denote the set of control dialogues by $\Pi_{Control}$. The BEGIN(G(p)) dialogue commences with an utterance by one agent which seeks the consent of the other participating agents to commence a dialogue of type G over proposition p. To achieve this consent, may require embedded dialogues of other types, for instance, persuasions, information-seeking dialogues, and negotiations. 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 also a dialogue, being that dialogue which consists of the n-fold repetition of G, each occurrence being undertaken until closure, and then being followed immediately by the next occurrence.

Sequencing: If G and H are both dialogues, then G; H is also a dialogue, representing that dialogue which consists of undertaking G until its closure and then immediately undertaking H.

Parallelization: If G and H are both dialogues, then $G \cap H$ is also a dialogue, representing that dialogue which consists of undertaking both G and H simultaneously, until each are closed. ⁴

Embedding: If G and H are both dialogues, and $\Phi \subseteq M^1 \times M^2 \dots \subseteq \Theta^G \times \Theta^G \dots$ is a finite set of legal locution sequences in G, then $G[H|\Phi]$ is also a dialogue, representing that dialogue which consists of undertaking G until a sequence in Φ has been executed, and then switching immediately to dialogue H which is undertaken until its closure, whereupon dialogue G resumes from immediately after the point where it was interrupted and continues until closure. Dialogue G is said to be embedded in G, at one level lower than G. In the time between when G0 pens and closes, dialogue G1 remains open, no matter how many embedded dialogues G3 itself may contain.

Testing: If p is a wff in \mathcal{L} , then $\langle p \rangle$ is a dialogue to assess the truthstatus of p. We assume such a dialogue returns a truth-value for p to whichever was the lowest-level dialogue open at the time of commencement of the testing dialogue. Typically, p will be some proposition which has become the subject of contention in a dialogue, and which makes reference to the world external to that dialogue. For example, in dialogue systems associated with a particular database, the testing of p may involve an interrogation of the database. In dialogues in scientific domains, testing may involve the conduct of a scientific experiment, the design,

⁴ As an example of parallel dialogues, complex human inquiries such as air-crash investigations are often divided into simpler, parallel sub-inquiries. Similarly, an intending purchaser of some product may engage in simultaneous bilateral negotiations with potential suppliers.

analysis and interpretation of which may themselves involve much discussion.⁵

We denote by Π the closure of the set $\Pi_{Atom} \cup \Pi_{Control}$ under the dialogue combination operations defined here. We next consider the impact of dialogue combinations on the combination and termination rules of dialogues, and on the commitments incurred in particular dialogue occurrences.

3.4. Interaction of Rules and Commitments

The previous sub-section presented a formalism for combinations of atomic dialogue-types and derived dialogues. For application of this formalism, we need to define interaction rules for the dialogue rules and commitments arising in the dialogues under the different combinations; these interaction rules may be seen as analogous to the bridge rules between logics in multi-context systems [8].

Firstly, we consider combination and termination rules under each of the possible Dialogue Combinations. For **Iterated**, **Sequenced** and **Parallel** dialogues, these rules for each dialogue apply only to that dialogue and do not interact, so no interaction rules are necessary. For **Embedded** dialogues, we assume that the rules of the lowest (most-deeply embedded) open dialogue and only these apply while this dialogue is in process; thus, any conflicting rules from higher dialogues are over-ridden by those of the lowest open dialogue. If a participant (or participants) does not wish to conduct an embedded dialogue on these terms, that participant can always refuse to participate; in other words, such a participant would not accede to a request to have an embedded dialogue in a $BEGIN(G[H|\Phi])$ Commencement dialogue if he or she did not accept the rules of H should over-ride those of G while H is in progress.

Next, for commitment rules, it is useful to distinguish dialogic from semantic commitments. For dialogic commitments, the interaction rules are defined exactly as for the interaction of combination and termination rules just presented. This is possible because dialogic commitments have no impact beyond the dialogue they occur in. For semantic commitments, a different approach is required, since these commitments refer to and may impact upon external reality. We assume that semantic commitments arising within a single dialogue occurrence do not conflict, either because such conflicts are disallowed by the combination rules of the dialogue (e.g. preclusion of commitment to an action and

⁵ We have proposed formal dialogue-game models for some of these scientific dialogues in [22].

to its negation) or because the participants to the dialogue, knowing that semantic commitments refer to an external reality, do not allow the dialogue to close until all conflicting commitments are resolved. For conflicts between semantic commitments from different dialogue occurrences, the dialogue participants may have different opinions on the appropriate form of resolution. For example, for dialogues conducted sequentially, one opinion may be that the commitments from earlier dialogues should take precedence over those from later ones, as is usually the case in legal and contractual domains. Alternatively, another opinion may be that the commitments from later dialogues should take precedence, as is usually the case with party political promises, or edicts issued by religious authorities.

The appropriate place for the resolution of such differences of opinion is in the Control dialogue before the dialogues in question commence. We therefore attach a suffix to our syntax for the various Dialogue Combinations, which indicates the proposed interaction rule for the resolution of any conflicts between semantic commitments; this is of the form SC(G(p)) > SC(H(q)), indicating that the semantic commitments arising in dialogue G(p) take precedence over those arising in dialogue G(p). Thus, a Commencement dialogue for two dialogues G(p) and G(p) to be conducted in parallel, with commitments from the earlier dialogue to over-ride those from the later, could be denoted:

$$BEGIN(G(p) \cap H(q) \mid SC(G(p)) > SC(H(q)))$$

If all potential participants are willing to accept such a prioritization constraint on commitments then the commencement dialogue can terminate, and the parallel dialogues $G(p) \cap H(q)$ can start, subject to the commitment interaction rule expressed by the constraint SC(G(p)) > SC(H(q)). If not all participants to the commencement dialogue agree with this constraint, then the parallel dialogues $G(p) \cap H(q)$ will not open.

This particular syntax permits the representation of constraints on commitments in terms of named dialogues, e.g. G and H. Because the formalism of Section 3.3 names dialogues and indicates their order of commencement, a constraint syntax which names dialogues enables the representation of constraints in terms of dialogue commencement order. For example, for a sequential dialogue G; H, we may express both possible prioritizations of constraints on semantic commitments based on the commencement order of the two dialogues: SC(G) > SC(H) and SC(G) < SC(H). But this does not permit the expression of constraints on commitments in terms of the order of termination of the dialogues. We therefore introduce an expression to enable this. If $\pi(G_1, G_2, \ldots, G_n)$ is some legal dialogue combination of the n di-

alogues G_1, G_2, \ldots, G_n , and G'_1, G'_2, \ldots, G'_n is some (possibly identical) re-ordering and/or renaming of these dialogues, then the expression:

$$BEGIN(\pi(G_1, G_2, ..., G_n) | SC(G'_1) > SC(G'_2) > ... > SC(G'_n))$$

denotes a Commencement dialogue for the dialogue $\pi(G_1, G_2, \ldots, G_n)$ subject to the constraint on semantic commitments expressed in the suffix, $SC(G'_1) > SC(G'_2) > \ldots > SC(G'_n)$. We permit renaming of the n dialogues instantiating π so as to allow for different time-orderings, e.g. $G_{Last_to_finish}$, $G_{Second_Last_to_finish}$, \ldots , $G_{First_to_finish}$ is a renaming which presents the dialogues in reverse order of termination. In effect, this construction permits the participants to a dialogue to jointly provide their own interpretation — their own semantics — to the commitments expressed in dialogues.

3.5. Agent dialogue frameworks

We mentioned earlier that the Control Layer is the level at which dialogues about dialogues are conducted. It is useful to consider a fixed set of agents engaged in discussions on a fixed set of topics, according to rules from a fixed set of atomic dialogue games. In this context, we assume that the Control Layer has a specific starting time and continues indefinitely into the future, unless and until participating agents agree to terminate it, through an execution of the END(CONTROL) dialogue. Depending on the rules of association adopted by the participating agents (rules we leave unspecified here), termination of the Control Layer may occur whenever one participant wishes to withdraw, or only when a majority wish to do so, or when all but one wish to do so.

Drawing on the definitions given earlier, we now define an Agent Dialogue Framework (ADF) as a 5-tuple $(\mathcal{A}, \mathcal{L}, \Pi_{Atom}, \Pi_{Control}, \Pi)$, where \mathcal{A} is a set of agents, \mathcal{L} is a logical language for representation of discussion topics, Π_{Atom} is a set of atomic dialogue-types, $\Pi_{Control}$ a set of Control dialogues and Π the closure of $\Pi_{Atom} \cup \Pi_{Control}$ under the combination rules presented in the previous subsection. To reprise, each formal dialogue in $\Pi_{Atom} \cup \Pi_{Control}$ 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 functions, \mathcal{T}^G the set of termination functions, and \mathcal{CF}^G the set of commitment functions of the dialogue-type G.

The framework we have presented is defined in terms of rules of dialogue games and is potentially generative. For it to be so, we would

⁶ Such rules of association reveal a connection between our work and research defining *institutions* in electronic negotiation [24, 34].

need to 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 [36] typology.

For **Information-Seeking Dialogues** an agent a may be preprogrammed 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, for example, because it can construct a proof of por of $\neg p$, can be programmed to agree to such a dialogue and, within it, to respond with the appropriate truth-value. For defeasible reasoning, we may have undefeated arguments supporting p or $\neg p$ rather than deductive proofs. If questioned further, b can present the proof of or the argument for p to a. Thus, the ADF formalism can generate information-seeking 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). For Persuasion Dia**logues** we can imagine that agents a and b are pre-programmed as follows: If a accepts the truth of some proposition p and requires that balso 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 ais rational, a should have such a proof of (or argument for) p before it believed p to be true. Thus, for rational agents, we are able to generate persuasion dialogues.

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 for each agent are partially ordered, then a cake-cutting algorithm, such as that described in [26, 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. An alternative generative mechanism for negotiation dialogues over the purchase of consumer durables is proposed in [23], drawing on marketing models of consumer decision making. For deliberation dialogues, there is no obvious generative mechanism. These dialogues can be initiated automatically whenever an agent believes that the group of agents

needs to jointly decide on a course of action. 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 each 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. However, it is not clear how the initial proposal may be automatically generated when the deliberation dialogue commences. One approach could be for the agent which requested the dialogue to propose an action at random. However, this method may take a long time to converge to an agreed solution, if indeed it ever terminates.⁷

If we had generative mechanisms for each of the atomic dialoguetypes, then we would have them for all dialogue-types, by simple inspection of the Dialogue Combination Rules presented in Section 3.3.

4. Example

We illustrate the framework with a dialogue occurrence 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 dialogue-types may be embedded in one another. Because our formalism has been designed for any dialogue game, it does not specify legal locutions within games. For ease of understanding therefore, the example is given in a pseudo-narrative form, with dialogue moves annotated as sub-dialogues open and close; we also ignore interaction constraints. 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 \boldsymbol{B} requests commencement of an information-seeking dialogue regarding purchase of a second-hand car.

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 Layer.

⁷ The modal logic formalism of [25], for example, deals with this problem by assuming each agent first develops its own proposal for action and then seeks to persuade the others to adopt it; thus in Walton and Kabbe's terminology, the dialogue is modeled as a multi-way persuasion rather than as a deliberation.

- B: AGREE(RETURN_CONTROL)
- S: BEGIN(INFOSEEK(Budget))

 ${m S}$ 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 Layer.

S: BEGIN(INFOSEEK((Purchase_Criteria))

 $m{S}$ requests Information-Seeking dialogue over $m{B}$'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 Layer.

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" as a surrogate 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 Layer.

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 Layer.

B: BEGIN(INFOSEEK(Ratings_of_cars))

 ${\it B}$ requests an Information-Seeking Dialogue concerning ratings of cars against the agreed purchase criteria.

S: AGREE(INFOSEEK(Ratings_of_cars))

INFORMATION-SEEKING Dialogue 4 opens.

B: REQUEST(Cars, Models, Price, Mileage, Number_of_Owners)

S: Provides this information.

INFORMATION-SEEKING Dialogue 4 closes. Return to CONTROL Layer.

This example demonstrates a Negotiation Dialogue embedded in a Persuasion Dialogue, an embedding not permitted in [36]. In our example, the persuasion dialogues concern purchase decision criteria, which may be viewed as constraints on the space of the buyer's possible intentions. It seems reasonable that negotiations may occur over intentions and the constraints upon them, a view shared by [6] in modeling joint problem solving between agents. However, whether or not a particular type of sub-dialogue is appropriate at a specific place in a larger dialogue should be a matter for the participants to the dialogues to decide at the time. The formalism we have presented here enables such decisions to be made mutually and contextually.

⁸ Purchase decision criteria are important for the buyer's decision because they may determine which subset of available products receive detailed evaluation by the buyer [18, Chap. 2]. In this example, the particular criteria used by **B** may result in a subset containing cars which are not those which **S** most desires to sell to **B** at this time, and so **S** may well seek to influence **B**'s selection of criteria. The criteria therefore become part of the broad purchase negotiation dialogue between the two.

5. Discussion

The major contribution of this paper has been to develop a formal and potentially-generative language for dialogues between autonomous agents which admits combinations of different types of dialogues. Our formalism extends previous work in formalizing generic dialogue game protocols. For example, Prakken [30] seeks to develop a generic formalism for dialogue game protocols, but only those involving persuasion dialogues. In addition, his formalism does not represent combinations of dialogues. Bench-Capon et al. [3] take a different approach, presenting a generic method for syntactical specification of a dialogue game in terms of the dialogic pre-conditions and post-conditions for each legal locution. Such a specification may be considered as an operational semantics for the game protocol, linking utterances with states of the dialogue (although not with mental states of the participants). However, this formalism, while applicable to any dialogue game type, similarly does not represent combinations of dialogues or dialogue types.

Another approach is the dialogue game framework proposed by Maudet and Evrard [21], in the tradition of computational linguistics. Because the explanation of human dialogues and the generation of artificial dialogues are key concerns, their formalism incorporates semantic elements, such as rationality conditions, in the definition of the dialogue game syntax. As explained earlier, we are concerned to ensure protocols are defined purely syntactically, even when we have specific applications in mind. Moreover, their framework appears to be focused on information-seeking and persuasion dialogues and, as with those mentioned above, has no means to represent combinations of dialogues or dialogue types. Reed's formalism [32] avoids these problems, in that it permits any type or combination of types of dialogues to be modeled, with the exception of parallel dialogues. However, as explained earlier, this formalism is descriptive rather than generative, in that it does not specify the forms of utterances, nor the rules which govern their formation, issuance and effects.

In contrast, the ADF formalism we have proposed permits the representation of any type of dialogue and a wide diversity of combinations of dialogues and dialogue types; it does so in a manner which is purely syntactic, and which is potentially generative as well as descriptive. The formalism provides a single, unifying framework for representing disparate types of dialogue, including those in the typology of [36]. In addition, our formalism is modular, so that other dialogue types may be

inserted readily into the framework, directly or as embedded dialogues. Because the ADF structure is a logical formalism, it may also facilitate the study of the formal properties of dialogue game protocols, e.g. their computational complexity. Recent work by one of us has looked at the satisfiability and complexity of using logical languages for agent negotiation [38]. The issue of participant strategies in dialogue games is another area amenable to formal analysis through a logical formalism.

Acknowledgements

We are grateful for partial funding from the EU IST Programme through the Sustainable Lifecycles in Information Ecosystems Project (IST-1999-10948) and from the British EPSRC through a PhD studentship. We are also grateful to Mike Wooldridge and to the anonymous referees for their comments. An earlier version of this paper was presented at the AAAI Spring Symposium on Game-Theoretic and Decision-Theoretic Agents, held at Stanford University in March 2001. We thank the audience that day for their comments, particularly Barbara Grosz.

References

- Amgoud, L., N. Maudet, and S. Parsons: 2000, 'Modelling dialogues using argumentation'. In: E. Durfee (ed.): Proceedings of the Fourth International Conference on Multi-Agent Systems (ICMAS 2000). Boston, MA, USA, pp. 31–38.
- Bench-Capon, T. J. M., P. E. Dunne, and P. H. Leng: 1991, 'Interacting with knowledge-based systems through dialogue games'. In: Proceedings of the Eleventh International Conference on Expert Systems and Applications. Avignon, France, pp. 123-140.
- Bench-Capon, T. J. M., T. Geldard, and P. H. Leng: 2000, 'A method for the computational modelling of dialectical argument with dialogue games'. Artificial Intelligence and Law 8, 233-254.
- 4. Cohen, P. R. and C. R. Perrault: 1979, 'Elements of a plan-based theory of speech acts'. *Cognitive Science* 3, 177–212.
- Dignum, F., B. Dunin-Keplicz, and R. Verbrugge: 2000, 'Agent theory for team formation by dialogue'. In: C. Castelfranchi and Y. Lespérance (eds.): Pre-Proceedings of the Seventh International Workshop on Agent Theories, Architectures, and Languages (ATAL-2000). Boston, USA, pp. 141-156.
- Dignum, F., B. Dunin-Keplicz, and R. Verbrugge: 2001, 'Creating collective intention through dialogue'. Logic Journal of the IGPL 9(2), 305-319.

⁹ For example, Rod Girle has described command dialogues, where one party orders another to do something, and the second party challenges or questions this order [7].

- 7. Girle, R.: 1996, 'Commands in Dialogue Logic'. In: D. M. Gabbay and H. J. Ohlbach (eds.): *Practical Reasoning: Proceedings of the First International Conference (FAPR 1996)*. Berlin, Germany, pp. 246–260.
- Giunchiglia, F. and L. Serafini: 1994, 'Multilanguage hierarchical logics (or: how we can do without modal logics)'. Artificial Intelligence 65(1), 29-70.
- Gordon, T. F. and N. Karacapilidis: 1997, 'The Zeno argumentation framework'. In: Proceedings of the Sixth International Conference on AI and Law. pp. 10-18.
- Grosz, B. J. and C. L. Sidner: 1986, 'Attention, Intentions, and the structure of discourse'. Computational Linguistics 12(3), 175-204.
- 11. Hamblin, C. L.: 1970, Fallacies. London, UK: Methuen and Co Ltd.
- Harel, D.: 1984, 'Dynamic Logic'. In: D. Gabbay and F. Guenther (eds.): Handbook of Philosophical Logic, Volume II: Extensions of Classical Logic. Dordrecht, The Netherlands: D. Reidel, pp. 497-604.
- 13. Hitchcock, D.: 1991, 'Some principles of rational mutual inquiry'. In: F. Eemeren, R. Grootendorst, J. A. Blair, and C. A. Willard (eds.): *Proceedings of the Second International Conference on Argumentation*. Amsterdam, The Netherlands, pp. 236–243.
- Hitchcock, D., P. McBurney, and S. Parsons: 2001, 'A Framework for Deliberation Dialogues'. In: H. V. Hansen, C. W. Tindale, J. A. Blair, and R. H. Johnson (eds.): Proceedings of the Fourth Biennial Conference of the Ontario Society for the Study of Argumentation (OSSA 2001). Windsor, Ontario, Canada.
- 15. Hulstijn, J.: 2000, 'Dialogue Models for Inquiry and Transaction'. Ph.D. thesis, Universiteit Twente, Enschede, The Netherlands.
- Jennings, N. R., K. P. Sycara, and M. Wooldridge: 1998, 'A Roadmap of Agent Research and Development'. Journal of Autonomous Agents and Multi-Agent Systems 1(1), 7-36.
- Levin, J. A. and J. A. Moore: 1978, 'Dialogue-Games: metacommunications structures for natural language interaction'. Cognitive Science 1(4), 395–420.
- 18. Lilien, G. L., P. Kotler, and K. S. Moorthy: 1992, *Marketing Models*. Englewood Cliffs, NJ, USA: Prentice-Hall.
- 19. Lorenzen, P. and K. Lorenz: 1978, *Dialogische Logik*. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.
- MacKenzie, J. D.: 1979, 'Question-begging in non-cumulative systems'. Journal of Philosophical Logic 8, 117-133.
- 21. Maudet, N. and F. Evrard: 1998, 'A generic framework for dialogue game implementation'. In: *Proceedings of the Second Workshop on Formal Semantics and Pragmatics of Dialog.* Universite Twente, The Netherlands.
- McBurney, P. and S. Parsons: 2001, 'Representing epistemic uncertainty by means of dialectical argumentation'. Annals of Mathematics and Artificial Intelligence 32(1-4), 125-169.
- 23. McBurney, P., R. van Eijk, S. Parsons, and L. Amgoud: 2002, 'A dialogue-game protocol for agent purchase negotiations'. *Journal of Autonomous Agents and Multi-Agent Systems*. (In press).
- Noriega, P. and C. Sierra: 1997, 'Towards layered dialogical agents'. In: J. P. Muller, M. J. Wooldridge, and N. R. Jennings (eds.): Intelligent Agents III: Agent Theories, Architectures, and Languages. Berlin, Germany, pp. 173-188.
- Panzarasa, P., N. R. Jennings, and T. J. Norman: 2002, 'Formalizing collaborative decision-making and practical reasoning in multi-agent systems'. *Journal of Logic and Computation* 12(1). (In press).

- Parikh, R.: 1985, 'The logic of games and its applications'. Annals of Discrete Mathematics 24, 111–140.
- Parsons, S. and N. R. Jennings: 1996, 'Negotiation through argumentation —
 a preliminary report'. In: Proceedings of the Second International Conference on Multi-Agent Systems (ICMAS 1996). pp. 267-274.
- Parsons, S., C. Sierra, and N. R. Jennings: 1998, 'Agents that reason and negotiate by arguing'. *Journal of Logic and Computation* 8(3), 261—292.
- Pauly, M.: 1999, 'An introduction to Game Logic'. In: M. Faller, S. Kaufmann, and M. Pauly (eds.): Proceedings of the Seventh CSLI Workshop on Logic, Language and Computation.
- Prakken, H.: 2000, 'On dialogue systems with speech acts, arguments, and counterarguments'. In: Proceedings of the Seventh European Workshop on Logic in Artificial Intelligence (JELIA 2000). Berlin, Germany.
- 31. Prakken, H. and G. Sartor: 1998, 'Modelling reasoning with precedents in a formal dialogue game'. Artificial Intelligence and Law 6, 231-287.
- 32. Reed, C.: 1998, 'Dialogue frames in agent communications'. In: Y. Demazeau (ed.): Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS-98). pp. 246-253.
- Sadri, F., F. Toni, and P. Torroni: 2001, 'Logic agents, dialogues and negotiation: an abductive approach'. In: M. Schroeder and K. Stathis (eds.):
 Proceedings of the Symposium on Information Agents for E-Commerce, Artificial Intelligence and the Simulation of Behaviour Conference (AISB-2001).
 York, UK.
- Sierra, C., N. R. Jennings, P. Noriega, and S. Parsons: 1998, 'A framework for argumentation-based negotiations'. In: M. P. Singh, A. Rao, and M. J. Wooldridge (eds.): Intelligent Agents IV: Agent Theories, Architectures, and Languages. Berlin, Germany, pp. 177-192.
- Singh, M. P.: 2000, 'A social semantics for agent communications languages'.
 In: F. Dignum, B. Chaib-draa, and H. Weigand (eds.): Proceedings of the IJCAI-99 Workshop on Agent Communication Languages. Berlin, Germany.
- 36. Walton, D. N. and E. C. W. Krabbe: 1995, Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning. Albany, NY, USA: SUNY Press.
- 37. Wooldridge, M.: 2000, 'Semantic Issues in the Verification of Agent Communication Languages'. *Journal of Autonomous Agents and Multi-Agent Systems* 3(1), 9-31.
- 38. Wooldridge, M. and S. Parsons: 2000, 'Languages for negotiation'. In: W. H. (Editor) (ed.): Proceedings of the Fourteenth European Conference on Artificial Intelligence (ECAI 2000). Berlin, Germany, pp. 393–397.