Dialogue Games in Multi-Agent Systems

Peter McBurney
Department of Computer Science
University of Liverpool
Liverpool, L69 7ZF, UK
p.j.mcburney@csc.liv.ac.uk

Simon Parsons
Department of Computer and Information Science
Brooklyn College, City University of New York
Brooklyn, NY 11210, USA
parsons@sci.brooklyn.cuny.edu

October 23, 2002

Abstract

Formal dialogue games have been studied in philosophy since at least the time of Aristotle. Recently they have been applied in various contexts in computer science and artificial intelligence, particularly as the basis for interaction between autonomous software agents. We review these applications and discuss the many open research questions and challenges at this exciting interface between philosophy and computer science.

1 Introduction

Formal dialogue games are interactions between two or more players, where each player "moves" by making utterances, according to a defined set of rules. Their history in philosophy dates at least to Aristotle [5] and they were widely studied by philosophers in medieval times [24, 61]. In modern times, philosophers have used formal dialogue games to study non-deductive forms of reasoning, such as the classical fallacies [24, 33], and to provide a game-theoretic semantics for intuitionistic and classical logic [32] and for quantum logic [44]. These games have also found application in computational linguistics and Artificial Intelligence (AI). In linguistics, they have been used to explain sequences of human utterances [30], with subsequent application to machine-based natural language processing and generation [27], and to human-computer interaction [7]. Within artificial intelligence, they have been applied to modeling complex human reasoning, for example in legal domains [8, 51], and to requirements specification for complex software projects [20]. Dialogue games differ from the games of economic game theory [45] in that payoffs for winning or losing a game are not considered, and because there is no use of uncertainty measures, such as probabilities, to model the possible moves of opponents. They also differ from the abstract games recently used as a semantics for interactive computation [1], in the tradition of Jaako Hintikka's game-theoretic semantics [25], since these abstract games do not share the rich rule structure of dialogue games, and are not intended to have themselves a semantic interpretation involving the beliefs or actions of multiple agents.

The main application of dialogue games in Artificial Intelligence has been to the design of protocols for interaction between autonomous software entities, called *agents*, and our focus in this paper will be on this application. We begin, in Section 2, with a brief introduction to the subject of autonomous agents and multi-agent systems. This is followed, in Section 3, with an overview of an influential typology of human dialogues, which will be used to classify agent interactions. Section 4 then presents a model of a formal dialogue game protocol, following which we consider, in Section 5, several recent proposals for agent interaction protocols based on dialogue games. In Section 6, we discuss some of the many open issues current in this domain, whose resolution will require contributions from computer science, philosophy and mathematics.

2 Agents and Agent Interactions

Over the last decade, the concept of autonomous software agents has become increasingly important in Computer Science, simultaneously with the rise of large-scale distributed computer systems, such as the Internet.[1] There are many distinct definitions, and as yet no consensus among computer scientists, of what comprises an autonomous software agent. However, for our purposes, it will suffice to use the formulation of Michael Wooldridge and Nicholas Jennings [68]. They described an autonomous software agent as being situated in a specific environment and exhibiting there four primary characteristics: autonomy of decision-making; social awareness; reactivity to events in its environment; and proactivity, as it seeks to actively achieve some goals within its environment. A multi-agent system is a community of two or more autonomous computational agents interacting. An example could be a network of water data monitors, each responsible for water and flood control in part of some geographic region. In this example, each monitor may be able to collect local information, send and receive information to and from other monitors, and take local action, such as opening a dam sluice to release excess water. Actions may be taken independently or in collaboration with other monitors, in which case dialogue between the agents may be necessary to agree the need for, and nature of, any such actions. Similarly, some global actions may require each agent to act locally for their desired effects to be realized.

There are two sets of design issues which pre-occupy researchers in this field: The design and engineering of the individual agents in a multi-agent system,

[1] As so often happens in Computer Science, practice here has preceded theory [60].

and the design and engineering of the system itself. For the design of individual agents, the philosophy of intention, particularly the work of Michael Bratman [9] and John Pollock [50], has been influential, along with ideas from economic theory and cognitive psychology. Formalization of these approaches has often drawn heavily on modal logic, e.g., [65].

For the design of multi-agent systems, attention has focused on the design of communications and interaction protocols. Here, several approaches have been adopted. One widespread approach has been the use of economic interaction mechanisms, such as auctions [54]. Typically, these have advantages of simplicity and analytical tractability; for example, one can often determine the extent to which the outcomes of an auction have desirable properties, such as Pareto-optimality [56]. However, the information which can be transmitted in these interactions is limited: participants in an auction, for example, cannot normally provide the reasons for their bids or the reasons for their acceptance or rejection of the bids of others. It is reasonable to suppose that the ability to present such reasons should increase the likelihood that agents participating in an interaction would reach successful resolution [46].

A second approach to communications protocol engineering, permitting greater expressiveness than do auction mechanisms, has been the design of generic agent communications languages. Here, there are two main proposals: the Knowledge Query and Manipulation Language (KQML) [34], which arose from work sponsored by the U.S. Government's Defense Advanced Research Projects Agency (DARPA); and the Foundation for Intelligent Physical Agents' Agent Communications Language [21], the FIPA ACL, which arose from attempts to develop an industry standard for agent communications. In both languages there is a two-level hierarchy for communications messages. At the lower, or "inner" level, the content of messages — the topics of discussion — may be expressed in any logical language agreed by the participants. Such content is then wrapped, at the next level, in one of a number of defined locutions. In the FIPA ACL, for example, there are 22 of these; the *inform* locution, for example, has the following syntax [21, p. 11]:

```
(inform
    :sender (agent-identifier :name j)
    :receiver (set (agent-identifier :name i))
    :content
        "weather (today, raining)"
    :language Prolog)
```

In this example the content of the message is *"weather (today, raining)"*, while the locution in which this content is wrapped is *inform*. The syntax of the locution indicates that the computational language in which the message content is expressed is Prolog. The syntax also indicates that the sender of the message has agent-identifier j, while the intended recipient has agent-identifier i.

The FIPA ACL has been given an axiomatic semantics, called *SL* (for *Seman-*

tic Language), based on speech act theory [21, Informative Annex A].[2] Speech act theory, due originally to philosophers John Austin [6] and John Searle [57], considers spoken human utterances as actions, in so far as they may change the state of the world. Some utterances do this demonstrably as when countries issue *declarations of war* on other countries. Other utterances change the world unobservably, as when one person informs another of something he or she does not already know, and thereby causes a revision to the beliefs of the listener. Speech act theory classifies utterances according to their pre-conditions and effects, and has been used to provide a semantics for agent utterances, particularly in work by Philip Cohen, Hector Levesque and Raymond Perrault [10, 11]. Drawing on this work, the FIPA ACL Semantic Language *SL* is a modal logic formalism which defines pre- and post-conditions for the 22 locutions of the language in terms of the beliefs, desires and uncertain beliefs of the agents participating in a dialogue using the language. For example, the *inform* locution may only be uttered if the first agent believes the proposition to be true, which is thus a semantic condition for the locution. Further information regarding FIPA ACL and KQML may be found in [29, 34]; a recent introduction to the subject of autonomous agents and multi-agent systems can be found in [67].

3 Types of dialogues

An influential model of human dialogues is the typology of primary dialogue types of argumentation theorists Doug Walton and Erik Krabbe [63]. 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.

[2] For a description of different types of computer program semantics, see, e.g., [62].

This classification has been influential in argumentation theory and in the application of dialogue game protocols to multi-agent systems. However, 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* [63].

4 Formal Dialogue Games

We now present a model of a generic formal dialogue game in terms of the components of its specification, taken from [39]. 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. A dialogue game specification then consists of the following elements:

Commencement Rules: Rules which define the circumstances under which a 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.

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 the work of philosopher Charles Hamblin [24], 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 a dialogue ends.

It is worth noting here that more than one notion of *commitment* is present in the literature on dialogue games. For example, Hamblin treats commitments in a purely dialogical sense: *"A speaker who is obliged to maintain consistency needs to keep a store of statements representing his previous commitments, and require of each new statement he makes that it may be added without inconsistency to this store. The store represents a kind of persona of beliefs; it need not correspond with his real beliefs . . ."* [24, p. 257]. In contrast, Walton and Krabbe [63, Chapter 1] treat commitments as obligations to (execute, incur or maintain) a course of action, which they term action commitments. These actions may be utterances in a dialogue, as when a speaker is forced to defend a proposition he has asserted against attack from others; so Walton and Krabbe also consider propositional commitment as a special case of action commitment [63, p. 23]. As with Hamblin's treatment, such dialogical commitments to propositions may not necessarily represent a participant's true beliefs. In contrast, Munindar Singh's social semantics [59], requires participants in an interaction to express publicly their beliefs and intentions; these expressions are called *social commitments.* These include both expressions of belief in propositions and expressions of intent to execute or incur future actions.[3] Our primary motivation is the use of dialogue games as the basis for interaction protocols between autonomous agents. Because such agents will typically enter into these interactions in order to achieve some wider objectives, and not just for the enjoyment of the interaction itself, we believe it is reasonable to define commitments in terms of future actions or propositions external to the dialogue. In a commercial negotiation dialogue, for instance, the utterance of an offer may express a willingness by the speaker to undertake a subsequent transaction on the terms contained in the offer. For this reason, we can view commitments as semantic mappings between locutions and subsets of some set of statements expressing actions or beliefs external to the dialogue.

Dialogue game protocols have been articulated for each of the rule-governed primary types of dialogues in the typology of Walton and Krabbe: information-seeking dialogues [27, 63]; inquiries [38]; persuasion dialogues [2, 12]; negotiation dialogues [4, 27, 36, 55]; and deliberations [26]. Some of these proposals are discussed in the next section. There has even been work on non-co-operation in

[3]It is worth noting that all these notions of *commitment* differ from that commonly used in discussion of the internal states of software agents, namely the idea of the persistence of a belief or an intention [67, p. 205]. As Singh [58] argues, there is a qualitative difference between social commitments of the kind discussed here, and personal commitments of the kind encoded in beliefs, desires, and intentions. He further argues that one kind of commitment cannot be derived from another.

dialogues [23, 22], work which may yield dialogue game protocols for eristic dialogues. However, as mentioned earlier, most real-world dialogues (whether human or agent) involve aspects of more than one of these primary types. Two formalisms have been suggested for computational representation of combinations of dialogue: the *Dialogue Frames* of Chris Reed [52], which enable iterated, sequential and embedded dialogues to be represented; and our own *Agent Dialogue Frameworks* [39], which permit iterated, sequential, parallel and embedded dialogues to be represented. Both these formalisms are neutral with regard to the modeling of the primary dialogue types themselves, allowing the primary types to be represented in any convenient form, and allowing for types other than the six of the Walton and Krabbe typology to be included.

5 Examples of dialogue game protocols

We now briefly review several of the proposals for dialogue game protocols published in the autonomous agent literature. The first of these is the protocol of Leila Amgoud, Nicolas Maudet and Simon Parsons [2]. This protocol is based on James MacKenzie's philosophical dialogue game DC [33], a game for two players, both subject to the same rules. DC enables the participants to argue about the truth of a proposition and was designed to study the fallacy of begging the question (*petitio principii*, or circular reasoning). The agent interaction protocol of [2] based on DC allows four distinct locutions: *assert, accept, question* and *challenge*; these can be instantiated with a single proposition, and also, for the locutions *assert* and *accept*, with a set of propositions which together constitute an argument for a proposition. Thus the participants may communicate both propositional statements and arguments about these statements, where arguments may be considered as tentative proofs (i.e., logical inferences from assumptions which may not all be confirmed). The locutions of this protocol are similar to those of DC except that they do not include a locution for retraction of assertions (called *withdrawal* in DC). As with MacKenzie's game, the protocol of [2] establishes commitment stores which record, in full public view, the statements each participant has asserted. The syntax for this protocol has only been provided for dialogues between two participants, but could be readily extended to more agents, as the same authors did subsequently in [3].

Amgoud *et al.* demonstrate that their system enables persuasion, inquiry and information-seeking dialogues [2]. However, as the authors note, to permit negotiation dialogues, the protocol requires additional locutions. These are proposed in a subsequent paper by the same authors [4], in which three additional locutions are suggested: *request, promise* and *refuse*. In addition to instantiation with propositions and with arguments for propositions, several of these locution can also be instantiated with a two-valued function expressing a relationship between two resources. For example, the locution $promise(p \Rightarrow q)$ indicates a promise by the speaker to provide resource q in return for receiving resource p.

Building on the protocol of Amgoud *et al.* [4], Fariba Sadri, Francesca Toni and Paolo Torroni [55] propose a similar protocol but with fewer locutions. The legal locutions proposed here are: *request, promise, accept, refuse, challenge,* and *justify*. The contents of the locutions *request* and *promise* are resources, while the contents for the other four locutions are any of the six locutions. In addition, the locution *justify* allows the utterance of some support for a previous locution.

The dialogue-game protocols presented in the work of Frank Dignum, Barbara Dunin-Kęplicz and Rineke Verbrugge [12, 13] are intended to enable agents to form teams and to agree joint intentions, respectively. For both protocols, the authors assume that one agent, an *initiator* or *proponent*, seeks to persuade others (*opponents*) to join a team, and that another initiator (possibly the same agent) seeks to persuade team members to adopt a group belief or intention. The team-formation dialogue is modeled as an information-seeking dialogue followed by a persuasion, while the joint-intentions-formation dialogue is modeled as a persuasion dialogue, which may include embedded negotiation dialogues. For the persuasion dialogues, the authors adapt the rigorous persuasion dialogue game of Walton and Krabbe [63]. This game is a formalization of a critical discussion — i.e., a rigorous persuasion — in philosophy. Such dialogues involve two parties, one seeking to prove a proposition, and one seeking to disprove it.[4] Unfortunately, the authors do not specify their team-formation dialogue game models completely, for example, nowhere stating the set of locutions available to the participating agents. The protocol for joint intention formation dialogues [13] includes seven locutions: *statement, question, challenge, challenge-with-statement, question-with-statement* and *final remarks*; these last include: *"quit"* and *"won."* The statements associated with challenges and questions may be concessions made by the speaker. The protocol for team formation dialogues [12] may also use the same set of locutions, although this is not absolutely clear.

Finally, we briefly mention some of our own proposals for dialogue game agent interaction protocols. Firstly, in joint work with Rogier van Eijk and Leila Amgoud [36], we have articulated a dialogue game protocol for negotiation dialogues between potential buyers and sellers of consumer durables; this work drew on a standard model of consumer decision-making taken from marketing theory [31]. Secondly, together with David Hitchcock [26], we have presented a dialogue game protocol for deliberation dialogues, drawing on a theory of deliberative argument from the philosophy of argumentation [64]. Thirdly, in [38], we articulated a dialogue game protocol for agents engaged in an inquiry dialogue. This protocol enables the participants to express uncertain beliefs about claims and to resolve these on the basis of the arguments for and against the claims presented in the dialogue, and was designed in accordance with the philosophy of science of Paul Feyerabend [19] and Marcello Pera [49]. Inquiries involve a disinterested search for the truthful answer to some question. In many instances, however, we may desire *particular* answers to a question, such as when we seek to identify the possible risks of a new technology. In these cases

[4] Note, however, that the persuasion dialogues of Walton and Krabbe [63] deal only with beliefs, not intentions.

our search is overlaid with values we impose on the answer-space; in [37], we proposed a dialogue game protocol for agents engaged in such an inquiry.

6 Issues and challenges

The use of formal dialogue games as the basis for agent interaction protocols has only just begun, and there are many challenging issues still open. In this section we consider some of these.

6.1 Protocol semantics

One of the reasons for the popularity of the FIPA ACL is the fact that it has been given a well-defined semantics [29]. This semantics, as we mentioned above, is based on speech act theory and is defined in terms of the certain beliefs, uncertain beliefs and intentions of the participating agents. Having such a defined semantics means that participants know precisely what other speakers mean and intend by their utterances, *assuming those others are conforming to the semantics.* However, verifying conformance to the semantics is a conceptually challenging task [66], since it is always possible for a sufficiently-clever agent to simulate insincerely any required semantic state.

The development of appropriate semantics for dialogue game protocols is still very immature, although several different types of semantics have been defined for these protocols. The first of these involves defining locutions in terms of the pre-conditions necessary for, and the post-conditions arising from, their utterance, i.e., what is termed an *axiomatic* semantics in computer programming language theory [62]. We can distinguish, as in [35], between two types of axiomatic semantics. In a *public* axiomatic semantics, the pre-conditions and post-conditions of each locution are given only in terms of observable linguistic behaviour, and so conformity with the protocol can be verified by anyone observing the dialogue. The protocols in [26, 36, 37, 38] have been given a public axiomatic semantics. A *private* axiomatic semantics, on the other hand, is one in which some or all locutions have pre- or post-conditions defined in terms of the internal states of the participants. The protocols of Amgoud *et al.* [2, 4], of Sadri *et al.* [55] and of Dignum *et al.* [13] have been given such a semantics, as has the FIPA ACL [21]. For the protocols of Amgoud *et al.* [2, 4], each participating agent is assumed to be vested with a private argumentation-based reasoning mechanism, which generates defeasible arguments from premises according to defined procedures, as in [14]. The mechanism also permits a preference ordering over the arguments. An agent can only utter a locution $assert(p)$, for p a proposition, if that agent has an acceptable argument for p in its own knowledge base, or in its knowledge base combined with the public commitment stores of the other participants. Here, *acceptable arguments* are those which survive a specified process of attack and defence within an argumentation framework [14].

In the case of the protocol of Sadri *et al.* [55], which is designed for dialogues over scarce resources (i.e, negotiation dialogues), utterances are linked to a first-order logic describing resources. In this semantics, the knowledge of an agent is described [55, Section 3.1] as an abductive logic program consisting of if-then rules and of statements regarding the resources owned by the agent. Integrity constraints are placed on this knowledge in the form of rules[5] which provide an agent with possible responses to utterances it receives. An example of such a rule is: *Accept a request.* The abducibles of this logic program are then the possible locutions which the agent may utter in response to a message it receives. For the protocol of Dignum *et al.* [13], the authors assume that the participating agents have a Belief-Desire-Intention (BDI) mental architecture [65] and vest the locutions with a private axiomatic semantics: as with the speech-act semantics of the FIPA ACL [21], locutions are defined in terms of their impacts on, and pre-conditions of, these private mental states.

One may also view a dialogue under a given protocol as a sequence of commands in a programming language which operates on the states of a virtual computer comprising the internal states of the agents participating in the dialogue. This view leads to the definition of an *operational* semantics for the protocol [62] in which locutions are defined in terms of their state transitions. Operational semantics have recently been defined for various agent communications languages, e.g., [18]. To our knowledge, the only dialogue game protocol provided with an operational semantics is the consumer negotiation protocol of [36].

Programming language theory also entertains the concept of a *denotational* semantics [62], in which a translation is given between the commands in a programming language and the objects of some mathematical domain. In [40], we defined a denotational semantics for protocols in terms of sub-spaces of n-dimensional euclidean space; in this semantics, dialogues conducted according to a given protocol are mapped to directed paths in the associated sub-space. Another denotational approach arises from viewing agents engaged in dialogue as jointly constructing meaning as the dialogue proceeds, in the same way that humans using natural language may be thought to do. Thus, there may be value in defining a denotational semantics which is constructed incrementally in the course of a dialogue, in a manner analogous to the Discourse Representation Structures of natural language semantics in linguistics [28]. This is a line of research we are pursuing.

6.2 Formal properties

Why define a semantics for a protocol? One reason, mentioned above, is so that the participants share common meanings for utterances. Another reason is to enable a better understanding of the formal properties of a protocol and of the dialogues conducted under it. The study of the formal properties of dialogues and protocols is, like the development of formal semantics, still very immature, and considerable scope exists for further research in this area.

One property of great interest is *termination*: under what circumstances will

[5] Confusingly called *dialogue protocols* by these authors.

dialogues conducted under a given protocol terminate? For instance, Sadri and her colleagues [55] demonstrate that a dialogue under their protocol for resource negotiations will always terminate in a finite number of steps, if conducted between agents vested with the abductive logic program mechanisms described in the paper.[6] Similarly, Parsons *et al.* [47, 48] consider the termination properties of the protocols of [2, 4] for agents vested with argumentation-based decision architectures.

A set of related questions concern the *computational complexity* of dialogues using dialogue game protocols: How many algorithm steps are required, for the most efficient algorithm, for a participant to decide what to utter in a dialogue under a given protocol? How many algorithm steps are required, for the most efficient algorithm, to determine whether the locution a participant proposes to utter actually conforms to the protocol under which the dialogue is conducted? And, how many dialogue utterances are required for normal termination of a dialogue under a given protocol? There has been some preliminary work by computer scientists addressing these questions. In the papers just cited [47, 48], Parsons, Wooldridge and Amgoud consider the first and third questions for dialogue game protocols for information-seeking, inquiry and persuasion between agents with different degrees of scepticism regarding the information they receive. Dunne and McBurney [17] consider the first question for general persuasion dialogue protocols, while Dunne and Bench-Capon [15] consider the third question for a specific two-party, persuasion dialogue protocol. In addition, there has been related work by some of these authors, looking at the computational complexity of termination of general negotiation mechanisms (not only those involving dialogue game protocols) [69] and the computational complexity of deciding the status of particular arguments in computational argumentation systems [16].[7]

A third property of importance in practical applications concerns the degree to which a dialogue protocol may support automated dialogues between suitably-equipped software agents, what may be called *automatability.* In [36] we showed, using an operational semantics, that agents vested with appropriate, domain-specific, internal decision mechanisms could undertake fully automated consumer negotiation dialogues using the dialogue game protocol we defined. The internal decision mechanisms were derived from standard models of buyer and seller decision-making taken from marketing theory [31]. No other dialogue game protocol so far proposed has been shown to be automatable.

6.3 Protocol design and assessment

One reason a more thorough study of protocol properties is needed is to provide assistance to designers and users of agent interaction protocols. At present,

[6] To some extent this result is not surprising since these mechanisms require the agents to be co-operative.

[7] It is interesting that all the work mentioned here on the computational complexity of argumentation and dialogue has been undertaken by people who have, or do currently, work at the University of Liverpool, UK.

designers of dialogue game protocols currently have no guidance for issues such as:

- How many locutions should there be?
- What types of locutions should be included? For example: assertions, questions, etc.
- What are the appropriate rules for the combination of locutions?
- When should behavior be forbidden, e.g., repeated utterance of one locution?
- Under what conditions should dialogues be made to terminate?
- When are dialogues conducted according to a particular protocol guaranteed to terminate?

Similarly, the absence of formal studies of dialogue game protocols means that agents (or their human principals) intending to use such protocols have no guidance for issues such as:

- When are two protocols the same? With the proliferation of protocols, it will become important for agents to decide which protocol should be used for particular interactions.
- How may different protocols be compared and how may differences be measured?
- Which protocols are to be preferred under which circumstances? In other words, what are their advantages and disadvantages?
- What is the formal relationship between dialogue game protocols and general agent communications languages, e.g., FIPA ACL.
- When are dialogue game protocols preferable to other forms of agent interaction, such as auction mechanisms or general agent communications languages?

As part of a longer-term project to develop a formal theory of interaction protocols, we have taken initial steps towards answering some of these questions. In work with Michael Wooldridge [42], we proposed thirteen desirable properties of interaction protocols using dialogue games. These properties included: *Separation of syntax and semantics*, so that conformance with the protocol could be assessed on the basis only of observable behaviour; *Clarity of argumentation theory*, so that participants would know in advance their dialectical obligations when making assertions, contesting others' assertions, etc; and *Enablement of self-transformation*, so that participants would be able to change their beliefs, preferences or intentions, and readily express such changes, in the course of an interaction. We applied these 13 properties to assess several existing dialogue game protocols and to the FIPA ACL; all were found wanting, to a greater or lesser extent.

In another step towards a formal theory of interaction protocols we have considered when two protocols may be considered the same. Not only is an answer to this question necessary for choosing between protocols, but it is also essential in order to assess if a protocol is new (in the sense of providing a different functionality to an existing protocol rather than just having equivalent locutions with different names) or to assess if a protocol conforms to some specification (such as that laid down as the standard for interacting within some electronic institution [53]). In work with Mark Johnson [41], we have recently defined several reasonable notions of protocol equivalence and shown that these lead to distinct classes of protocols. Some of these notions were derived from ideas of process equivalence [43], and the links between protocols and processes are worth further exploration.

7 Conclusion

In this paper, we have provided a brief review of the application of formal dialogue games from the philosophy of argumentation to the problem of the design of protocols for interaction between autonomous software agents. These games have almost 2500 years of study in philosophy behind them, and show great promise for agent communications. This area of Artificial Intelligence still has many open questions, research challenges and implementation issues. We believe that dialogue game protocols have great potential for multi-agent systems applications because they represent an effective compromise between the strict rule-governed nature of economic auction mechanisms and the greater expressiveness of generic agent communications languages, such as the FIPA ACL.

References

[1] S. Abramsky. Semantics of interaction: an introduction to game semantics. In A. M. Pitts and P. Dybjer, editors, *Semantics and Logics of Computation*, pages 1–31. Cambridge University Press, Cambridge, UK, 1997.

[2] 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 (ICMAS 2000)*, pages 31–38, Boston, MA, USA, 2000. IEEE Press.

[3] L. Amgoud and S. Parsons. Agent dialogues with conflicting preferences. In J-J. Meyer and M. Tambe, editors, *Pre-Proceedings of the Eighth International Workshop on Agent Theories, Architectures, and Languages (ATAL 2001)*, pages 1–14, Seattle, WA, USA, 2001.

[4] L. Amgoud, S. Parsons, and N. Maudet. Arguments, dialogue, and negotiation. In W. Horn, editor, *Proceedings of the Fourteenth European Conference on Artificial Intelligence (ECAI 2000)*, pages 338–342, Berlin, Germany, 2000. IOS Press.

[5] Aristotle. *Topics.* Clarendon Press, Oxford, UK, 1928. W. D. Ross, Editor.

[6] J. L. Austin. *How To Do Things with Words.* Oxford University Press, Oxford, UK, 1962.

[7] T. J. M. Bench-Capon, P. E. Dunne, and P. H. Leng. Interacting with knowledge-based systems through dialogue games. In *Proceedings of the Eleventh International Conference on Expert Systems and Applications*, pages 123–140, Avignon, France, 1991.

[8] T. J. M. Bench-Capon, T. Geldard, and P. H. Leng. A method for the computational modelling of dialectical argument with dialogue games. *Artificial Intelligence and Law*, 8:233–254, 2000.

[9] M. E. Bratman. *Intention, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA, USA, 1987.

[10] P. R. Cohen and H. J. Levesque. Rational interaction as the basis for communication. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 221–255. MIT Press, Cambridge, MA, USA, 1990.

[11] P. R. Cohen and C. R. Perrault. Elements of a plan-based theory of speech acts. *Cognitive Science*, 3:177–212, 1979.

[12] F. Dignum, B. Dunin-Kęplicz, and R. Verbrugge. Agent theory for team formation by dialogue. In C. Castelfranchi and Y. Lespérance, editors, *Intelligent Agents VII: Proceedings of the Seventh International Workshop on Agent Theories, Architectures, and Languages (ATAL 2000)*, Lecture Notes in Artificial Intelligence 1986, pages 150–166, Berlin, Germany, 2000. Springer.

[13] F. Dignum, B. Dunin-Kęplicz, and R. Verbrugge. Creating collective intention through dialogue. *Logic Journal of the IGPL*, 9(2):305–319, 2001.

[14] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-persons games. *Artificial Intelligence*, 77:321–357, 1995.

[15] P. E. Dunne and T. J. M. Bench-Capon. Two party immediate response disputes: properties and efficiency. Technical Report ULCS-01-005, Department of Computer Science, University of Liverpool, Liverpool, UK, 2001.

[16] P. E. Dunne and T. J. M. Bench-Capon. Coherence in finite argument systems. *Artificial Intelligence*, 141(1–2):187–203, 2002.

[17] P. E. Dunne and P. McBurney. Optimal utterances in dialogue protocols. Technical report, Department of Computer Science, University of Liverpool, Liverpool, U.K., 2002.

[18] R. van Eijk, F. S. de Boer, W. van der Hoek, and J-J. Ch. Meyer. A fully abstract model for the exchange of information in multi-agent systems. *Theoretical Computer Science*, 2002. *In press.*

[19] P. Feyerabend. *Against Method.* Verso, London, UK, third edition, 1993. First edition published 1971.

[20] A. Finkelstein and H. Fuks. Multi-party specification. In *Proceedings of the Fifth International Workshop on Software Specification and Design*, Pittsburgh, PA, USA, 1989. ACM Sigsoft Engineering Notes.

[21] FIPA. Communicative Act Library Specification. Technical Report XC00037H, Foundation for Intelligent Physical Agents, 10 August 2001. Available from: `www.fipa.org`.

[22] D. M. Gabbay and J. Woods. More on non-cooperation in Dialogue Logic. *Logic Journal of the IGPL*, 9(2):321–339, 2001.

[23] D. M. Gabbay and J. Woods. Non-cooperation in dialogue logic. *Synthese*, 127(1-2):161–186, 2001.

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

[25] J. Hintikka. *The Game of Language: Studies in Game-Theoretical Semantics and Its Applications*, volume 22 of *Synthese Language Library.* D. Reidel, Dordrecht, The Netherlands, 1983.

[26] D. Hitchcock, P. McBurney, and S. Parsons. A framework for deliberation dialogues. In H. V. Hansen, C. W. Tindale, J. A. Blair, and R. H. Johnson, editors, *Proceedings of the Fourth Biennial Conference of the Ontario Society for the Study of Argumentation (OSSA 2001)*, Windsor, Ontario, Canada, 2001.

[27] J. Hulstijn. *Dialogue Models for Inquiry and Transaction.* PhD thesis, Universiteit Twente, Enschede, The Netherlands, 2000.

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

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

[30] J. A. Levin and J. A. Moore. Dialogue-games: metacommunications structures for natural language interaction. *Cognitive Science*, 1(4):395–420, 1978.

[31] G. L. Lilien, P. Kotler, and K. S. Moorthy. *Marketing Models*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1992.

[32] P. Lorenzen and K. Lorenz. *Dialogische Logik*. Wissenschaftliche Buchgesellschaft, Darmstadt, Germany, 1978.

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

[34] J. Mayfield, Y. Labrou, and T. Finin. Evaluating KQML as an agent communication language. In M. J. Wooldridge, J. P. Müller, and M. Tambe, editors, *Intelligent Agents II*, Lecture Notes in Artificial Intelligence 1039, pages 347–360, Berlin, Germany, 1996. Springer.

[35] P. McBurney. *Rational Interaction*. PhD thesis, Department of Computer Science, University of Liverpool, Liverpool, UK, 2002.

[36] P. McBurney, R. M. van Eijk, S. Parsons, and L. Amgoud. A dialogue-game protocol for agent purchase negotiations. *Journal of Autonomous Agents and Multi-Agent Systems*, 2002. *In press.*

[37] P. McBurney and S. Parsons. Chance discovery using dialectical argumentation. In T. Terano, T. Nishida, A. Namatame, S. Tsumoto, Y. Ohsawa, and T. Washio, editors, *New Frontiers in Artificial Intelligence: Joint JSAI 2001 Workshop Post Proceedings*, Lecture Notes in Artificial Intelligence 2253, pages 414–424. Springer, Berlin, Germany, 2001.

[38] P. McBurney and S. Parsons. Representing epistemic uncertainty by means of dialectical argumentation. *Annals of Mathematics and Artificial Intelligence*, 32(1–4):125–169, 2001.

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

[40] P. McBurney and S. Parsons. A geometric semantics for dialogue-game protocols for autonomous agent interactions. *Electronic Notes in Theoretical Computer Science*, 52(2), 2002.

[41] P. McBurney, S. Parsons, and M. W. Johnson. When are two protocols the same? In M. P. Huget, F. Dignum, and J. L. Koning, editors, *Agent Communications Languages and Conversation Policies. Proceedings of a Workshop at the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2002)*, Bologna, Italy, 2002.

[42] P. McBurney, S. Parsons, and M. Wooldridge. Desiderata for agent argumentation protocols. In C. Castelfranchi and W. L. Johnson, editors, *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2002)*, Bologna, Italy, 2002.

[43] R. Milner. *Communication and Concurrency.* Prentice Hall, Hemel Hempstead, UK, 1989.

[44] P. Mittelstaedt. *Quantum Logic.* D. Reidel, Dordrecht, The Netherlands, 1979.

[45] M. J. Osborne and A. Rubinstein. *A Course in Game Theory.* MIT Press, Cambridge, MA, USA, 1994.

[46] S. Parsons, C. Sierra, and N. R. Jennings. Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3):261–292, 1998.

[47] S. Parsons, M. Wooldridge, and L. Amgoud. An analysis of formal inter-agent dialogues. In C. Castelfranchi and W. L. Johnson, editors, *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2002)*, pages 394–401, New York, NY, USA, 2002. ACM Press.

[48] S. Parsons, M. Wooldridge, and L. Amgoud. Properties and complexity of some formal inter-agent dialogues. *Journal of Logic and Computation*, 2003. *In press.*

[49] M. Pera. *The Discourses of Science.* University of Chicago Press, Chicago, IL, USA, 1994.

[50] J. L. Pollock. *Cognitive Carpentry: A Blueprint for How to Build a Person.* MIT Press, Cambridge, MA, USA, 1995.

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

[52] C. Reed. Dialogue frames in agent communications. In Y. Demazeau, editor, *Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS-98)*, pages 246–253. IEEE Press, 1998.

[53] J. A. Rodríguez, F. J. Martin, P. Noriega, P. Garcia, and C. Sierra. Towards a test-bed for trading agents in electronic auction markets. *AI Communications*, 11(1):5–19, 1998.

[54] J. S. Rosenschein and G. Zlotkin. *Rules of Encounter: Designing Conventions for Automated Negotiation among Computers.* MIT Press, Cambridge, MA, USA, 1994.

[55] F. Sadri, F. Toni, and P. Torroni. Logic agents, dialogues and negotiation: an abductive approach. In M. Schroeder and K. Stathis, editors, *Proceedings of the Symposium on Information Agents for E-Commerce, Artificial Intelligence and the Simulation of Behaviour Conference (AISB-2001)*, York, UK, 2001. AISB.

[56] T. W. Sandholm. Distributed rational decision making. In G. Weiss, editor, *Multiagent Systems: A Modern Introduction to Distributed Artificial Intelligence*, pages 201–258. MIT Press, Cambridge, MA, USA, 1999.

[57] J. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, UK, 1969.

[58] M. P. Singh. A conceptual analysis of commitments in multiagent systems. Technical Report 96-09, Department of Computer Science, North Carolina State University, Raleigh, NC, USA, 1996.

[59] M. P. Singh. A social semantics for agent communications languages. In F. Dignum, B. Chaib-draa, and H. Weigand, editors, *Proceedings of the Workshop on Agent Communication Languages, International Joint Conference on Artificial Intelligence (IJCAI-99)*, Berlin, Germany, 2000. Springer.

[60] B. C. Smith. The foundations of computing. In M. Scheutz, editor, *Computationalism: New Directions*, pages 23–58. MIT Press, Cambridge, MA, USA, 2002.

[61] P. V. Spade. Recent research on medieval logic. *Synthese*, 40:3–18, 1979.

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

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

[64] H. Wohlrapp. A new light on non-deductive argumentation schemes. *Argumentation*, 12:341–350, 1998.

[65] M. J. Wooldridge. *Reasoning about Rational Agents*. MIT Press, Cambridge, MA, USA, 2000.

[66] M. J. Wooldridge. Semantic issues in the verification of agent communication languages. *Journal of Autonomous Agents and Multi-Agent Systems*, 3(1):9–31, 2000.

[67] M. J. Wooldridge. *Introduction to Multiagent Systems*. John Wiley and Sons, New York, NY, USA, 2002.

[68] M. J. Wooldridge and N. R. Jennings. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2):115–152, 1995.

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