

Desiderata for dialogue-game protocols for agent interactions

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

Abstract. Designers of agent communications protocols are increasingly using formal dialogue games, taken from the philosophy of argumentation. Based on our experiences in designing these, we propose a set of desiderata for such protocols.

KEYWORDS: Argumentation, Agent Communication Languages, Dialogue Games.

Formal dialogue games are games in which participants “move” by uttering locutions, and they have been studied by philosophers since the Middle Ages [6]. In recent years, such games have found application in Artificial Intelligence, for example as the basis for the design of protocols for communications between autonomous software agents. In the typology of dialogues of Walton and Krabbe [23], dialogue-game models have been developed for: negotiation dialogues, where participants seek to agree a division of some scarce resource [2, 3, 20]; persuasion dialogues, where one party seeks to persuade another to accept some claim [3, 16]; information-seeking dialogues, where one party seeks the answer to some question from another party [10]; and deliberation dialogues, where participants seek to jointly agree a course of action in some situation [8].

Much discussion in argumentation theory and in the philosophy of logic has concerned the meaning, usefulness and appropriate design of dialogue games, as seen recently in [9, 12]. They have also received attention in computational linguistics, as the basis for models of natural language generation and processing [22]. However, despite the numbers of dialogue-game protocols recently proposed for agent interactions, we know of no discussion of appropriate design principles for this domain. In this paper, we propose the first list of desiderata to govern their design and assessment.

We assume that the specification of a dialogue game protocol consists of: (a) a set of topics of discussion (which may be represented in some logical language); (b) the syntax for a set of defined locutions concerning these topics; (c) a set of rules which govern the utterance of these locutions; and (d) a set of rules which establish what commitments, if any, participants create by the utterance of each locution. We refer to such a specification as a *Dialogue Game* or a *Dialectical System*; our model is presented formally in [15]. We further assume that agents participating in such dialogues are autonomous, willing and free participants, able to enter and withdraw a particular dialogue as they see fit.

Our proposed desiderata are informed by the criteria proposed for assessment of automated auction and negotiation mechanisms in, e.g. [11, 17, 21], by theories of deliberative decision-making, e.g. [1, 5, 7], and by our previous work proposing criteria for assessment of argumentation-based decision-support systems [19]. Because of space limitations, we simply list our proposed desiderata, with little in the way of justification:

- Dialogue Purpose:** A dialectical system should have one or more stated objectives, and its locutions and rules should facilitate the achievement of these. For example, the stated purpose of a system for negotiation may be an agreement on the division of a scarce resource.
- Individual Purposes:** A dialectical system should permit participating agents to achieve their own individual purposes consistent with the overall purpose of the dialogue. These individual purposes may conflict, as when parties to a negotiation each seek to maximize their individual utility in any outcome.
- Formal Inclusiveness:** A dialectical system should not preclude participation by any potential agent which is qualified and willing to participate.
- Transparency:** Participants to a dialogue should know the rules and structure of the dialectical system prior to commencement of the dialogue. In particular, any reference from dialogues in a dialectical system to an external reality should be explicitly stated, and known to the participants before commencement, e.g. when commitments incurred inside a negotiation dialogue imply subsequent real-world commitments to execute a particular transaction.
- Fairness:** A dialectical system should either treat all participants equally, or, if not, make explicit any asymmetries in their treatment.
- Soundness of Argumentation:** A dialectical system should adhere, at least at the outset, to a specific theory of argument, for example Hitchcock's Principles for Rational Mutual Inquiry [7], or Roberts' Rules of Order [18]. If the participants wish to change the dialogical rules of the system in the course of using it for a particular dialogue, being free agents, they should be enabled to do so.
- Externalization:** A dialectical system should be defined entirely in terms of verifiable linguistic behaviour; in particular, there should be no requirement for participants to exhibit particular mental states before, during or after a dialogue.¹
- Encouragement of Resolution:** Resolution of each dialogue (normal termination) should be facilitated, and not precluded, by the locutions and rules of a dialectical system.
- Rule-Consistency:** The locutions and rules of a dialogue system should be internally consistent; in particular, they should not lead to deadlocks (where no participant may utter a legal locution), nor cycles of repeated locutions.
- Discouragement of Disruption:** Under normal circumstances, the rules of a dialectical system should discourage or preclude disruptive linguistic behaviour, such as uttering the same locution repeatedly. As Krabbe [13] notes with regard to retraction, achieving a balance between outlawing disruptive behaviour and permitting freedom of expression is not necessarily straightforward, and will differ by application.

¹ For example, the dialogue-game models for persuasion dialogues of [14] and negotiation dialogues of [2] do not satisfy this criterion, as they includes semantic conditions in the definition of locutions.

Self-Transformation: A dialectical system should permit participants to undergo self-transformation [5] in the course of a dialogue; e.g. participants to a negotiation may change their preferences or their valuations of utility as a result of information they receive in the dialogue. In particular, participants should have the right to retract prior commitments, although not necessarily always unconditionally.

System Simplicity: The locutions and rules of a dialectical system should be as simple as possible, consistent with these other criteria. In particular, each locution should serve a specific and explicit function in the dialogue, and the rules should lead to an efficient transfer of information between participants.

Computational Simplicity: A dialectical system should be designed to minimize any computational demands on its participants, consistent with these criteria.

In addition, there may be further desiderata appropriate for specific types of dialogue. For instance, for dialogues undertaken to negotiate a division of scarce resource, it may be considered desirable that outcomes are Pareto optimal, i.e. that any other outcome leaves at least one participant worse off. Because we assume agents are free and willing participants in a dialogue, acting under no duress, then any agreed outcome to a negotiation dialogue will satisfy this particular criterion, if the desiderata above are met.

It is important to note two criteria we have not included here. We have not specified that dialectical systems should be realistic representations of some human dialogue, as we see no reason why agent interactions should necessarily adopt human models of interaction. Indeed, dialectical systems may be applied to agent dialogues which humans do not, or, even, could never undertake, such as simultaneous negotiations with hundreds or thousands of participants. Moreover, there are many common types of human dialogue, such as command dialogues, information-provision dialogues and dialogues over the degrees of belief to be assigned to uncertain propositions, for which no formal representations yet exist.

In addition, we have not specified that the rules of a dialectical system should require that the participants satisfy some criteria of rationality or that they should adopt a particular decision-making procedure, such as the maximum-expected-utility rule. Firstly, such conditions can never be completely verified, since a sufficiently-clever agent can always insincerely simulate any desired mental state. Secondly, insisting that participants to a dialogue adhere to some normative code of correct thought is contrary to our *Weltanschauung*: as Feyerabend demonstrated [4], the progress of science has occurred despite, not because of, such rules for reasoning.

References

1. R. Alexy. A theory of practical discourse. In S. Benhabib and F. Dallmayr, editors, *The Communicative Ethics Controversy*, pages 151–190. MIT Press, Cambridge, MA, USA, 1990. (Originally published in German in 1978.)
2. L. Amgoud, N. Maudet, and S. Parsons. Modelling dialogues using argumentation. In E. Durfee, editor, *Proc. 4th ICMAS*, pages 31–38, Boston, 2000. IEEE.
3. F. Dignum, B. Dunin-Kępicz, and R. Verbrugge. Agent theory for team formation by dialogue. In C. Castelfranchi and Y. Lespérance, editors, *Proc. 7th Intern. Workshop on Agent Theories, Architectures, and Languages (ATAL-2000)*, pages 141–156, Boston, 2000.

4. P. Feyerabend. *Against Method*. Verso, London, UK, third edition, 1993.
5. J. Forester. *The Deliberative Practitioner: Encouraging Participatory Planning Processes*. MIT Press, Cambridge, MA, USA, 1999.
6. C. L. Hamblin. *Fallacies*. Methuen and Co Ltd, London, UK, 1970.
7. D. Hitchcock. Some principles of rational mutual inquiry. In F. van Eemeren et al., editor, *Proc. 2nd Intern. Conf. on Argumentation*, pages 236–243, Amsterdam, 1991. SICSAT.
8. 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, *Proc. 4th Biennial Conf. Ontario Soc. Study of Argumentation*, Windsor, Canada, 2001.
9. W. Hodges. Dialogue foundations: I – a sceptical look. In *Supplementary Volume LXXV: Joint Session of the Aristotelian Society and the Mind Association, University of York*, pages 17–32, London, UK, 2001. Aristotelian Society.
10. J. Hulstijn. *Dialogue Models for Inquiry and Transaction*. PhD thesis, Universiteit Twente, Enschede, The Netherlands, 2000.
11. N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, M. Wooldridge, and C. Sierra. Automated negotiation: prospects, methods and challenges. *International Journal of Group Decision and Negotiation*, 10(2):199–215, 2001.
12. E. C. W. Krabbe. Dialogue foundations: II – dialogue logic revisited. In *Supplementary Volume LXXV: Joint Session of the Aristotelian Society and the Mind Association, University of York*, pages 33–49, London, UK, 2001. Aristotelian Society.
13. E. C. W. Krabbe. The problem of retraction in critical discussion. *Synthese*, 127(1-2):141–159, 2001.
14. N. Maudet and F. Evrard. A generic framework for dialogue game implementation. In *Proc. 2nd Workshop on Formal Semantics and Pragmatics of Dialog*, 13th Intern. Twente Workshop on Language Technology (TWLT 13), The Netherlands, 1998. Universiteit Twente.
15. P. McBurney and S. Parsons. Agent ludens: games for agent dialogues. In S. Parsons and P. Gmytrasiewicz, editors, *Game-Theoretic and Decision-Theoretic Agents: Proc. 2001 AAAI Spring Symposium*, pages 70–77, Menlo Park, 2001. AAAI Press.
16. P. McBurney and S. Parsons. Representing epistemic uncertainty by means of dialectical argumentation. *Annals of Mathematics and Artificial Intelligence*, (in press), 2001.
17. A. Pekec. Multiround combinatorial auction design. Presentation to Workshop on Electronic Market Design, July 2001. Infonomics Group, University of Maastricht, Maastricht, The Netherlands.
18. H. Prakken and T. F. Gordon. Rules of order for electronic group decision making — a formalization methodology. In J. A. Padget, editor, *Collaboration between Human and Artificial Societies: Coordination and Agent-Based Distributed Computing*, LNAI 1624, Berlin, 1999. Springer.
19. W. Rehg, P. McBurney, and S. Parsons. Computer decision-support systems for public argumentation: Criteria for assessment. In H. V. Hansen, C. W. Tindale, J. A. Blair, and R. H. Johnson, editors, *Proc. 4th Biennial Conf. Ontario Soc. Study of Argumentation*, Windsor, Canada, 2001.
20. F. Sadri, F. Toni, and P. Torroni. Logic agents, dialogues and negotiation: an abductive approach. In M. Schroeder and K. Stathis, editors, *Proc. Sym. Info Agents for E-Commerce, AISB-2001*, York, UK, 2001. AISB.
21. 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, 1999.
22. D. R. Traum and J. F. Allen. A speech acts approach to grounding in conversation. In *Proc. 2nd Intern. Conf. Spoken Language Processing*, pages 137–140, 1992.
23. D. N. Walton and E. C. W. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press, Albany, 1995.