A Framework for Deliberation Dialogues

David Hitchcock (1), Peter McBurney (2) and Simon Parsons (3)

(1) Department of Philosophy, McMaster University, Hamilton, Canada L8S 4K1. Email: hitchckd@mcmaster.ca

(2) Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, United Kingdom. Email: p.j.mcburney@csc.liv.ac.uk

(3) Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, United Kingdom. Email: s.d.parsons@csc.liv.ac.uk

Abstract:

Deliberation dialogues involve reasoning about the appropriate course or courses of action for a group to undertake. No models currently exist for the conduct of such dialogues. Beginning with an analysis of the differences between deliberations and other types of dialogue (such as negotiations or information-seeking dialogues), we propose a generic framework in which to develop such models. We then consider various instantiations of our generic deliberation framework so as to illustrate its applicability.

1. Introduction

Argumentation is increasingly important in computer science, for example in the design of systems of autonomous software agents (Parsons *et al.* 1998, Jennings *et al.* 2001). Recently, the use of argumentation in such applications has focused on formal dialogue systems and in this work the typology of dialogues of Walton and Krabbe (1995) has been influential. This typology identifies several primary categories of dialogue, distinguished by their initial situations, the goals of each of their participants, and the goals of the dialogue itself (which may differ from those of its participants). The dialogue types are: *Information-seeking dialogues*, in which a participant who wishes to obtain the presently unknown answer to some question seeks to get it from another participant who does know; *Inquiries*, in which all participants collaborate to answer an open factual question to which none initially has the answer; *Persuasion dialogues*, in which a participant who endorses some proposition seeks to convince others to accept it; *Negotiations*, in which participants seek to agree on how to divide a scarce resource among themselves; *Deliberations*, in which participants discuss what action is to be taken in some situation; and *Eristic* (strife-ridden) dialogues, in which participants spar verbally, for example in an attempt to vent perceived grievances. While this typology is quite rich, Walton and Krabbe do not claim it is comprehensive. Many real-world dialogues are actually combinations of these different types; for example, human purchase negotiations may include periods of information-seeking, persuasion and deliberation interactions. Each of these types can be seen in the paradigm example of Parsons *et al.* 1998, in which agents collaborate to furnish a room.

Formal models have been developed for persuasion dialogues (Traum and Allen 1992, Walton and Krabbe 1995, Dignum *et al.* 2000a & b, Prakken 2000), for information-seeking dialogues (Hulstijn 2000) and for negotiation dialogues (Amgoud *et al.* 2000a, Hulstijn 2000). Formal models have also been proposed for combinations of dialogue-types (Reed 1998, McBurney and Parsons 2001a). Less attention has been paid to models for deliberation dialogues, despite their importance. Indeed, a major part of Artificial Intelligence (AI) research concerns the design of autonomous entities, such as robots, able to devise sequences of actions to achieve pre-determined goals. Given such a focus on action and the increasing focus in AI on collaborative decision-making, it is surprising that models for deliberation dialogue have not been a feature of AI research. To our knowledge, only one project – the TRAINS project of Allen *et al.* (1995) – ostensibly seeks to model a deliberation dialogue, and this, as we show in Section 8, is instead modelled as a two-way persuasion dialogue.

In this paper, we present a formal and implementable model for deliberation dialogues between autonomous agents. We begin in Section 2 by presenting a formal model of dialogue games, and then proceed in Section 3 to discuss the distinguishing characteristics of deliberations. Section 4 contains our eight-stage model for deliberation dialogues, while Section 5 presents the dialogue-game rules we specify to implement this model. Section 6 examines the potential of our formalism for use in automated dialogues, and presents a portfolio of mechanisms for the participating agents to enable a deliberation dialogue to be generated automatically. Our mechanisms are analogous to recent work in automated agent negotiations. We present a simple example dialogue in Section 7 and conclude with a brief discussion of future work in Section 8.

2. Dialogue Games

Formal dialogue games were first proposed in philosophy for the study of fallacies (Hamblin 1970, 1971, MacKenzie 1979) and have recently found application in Artificial Intelligence (Amgoud *et al.* 2000a, 2000b, Bench-Capon *et al.* 2000, Hulstijn *et al.* 2000, Stathis 2000). Building on (Walton & Krabbe 1995, Prakken 2000) in abstracting from the rules for any one game, we can identify five types of dialogue game rules, as follows.

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

- **Locution Rules:** Rules which indicate what utterances are permitted. Typically, legal locutions permit participants to assert propositions, permit others to question or contest prior assertions, and permit those asserting propositions which are subsequently questioned or contested to justify their assertions. Justifications may involve the presentation of a proof of the proposition or an argument for it, and such presentations may also be legal utterances.

- **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. If a dialogue has an underlying logic, then the rules of inference of this logic will be combination rules of the dialogue, which may for example permit one participant to infer a proposition from one or more propositions in the commitment store of another participant.

- **Commitment Rules:** 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. In a negotiation dialogue, for example, assertion of an offer may express a willingness to undertake a transaction on the terms contained in the offer. Since Hamblin (1970), 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 which define the circumstances under which the dialogue ends. Such rules may also define what is the upshot of a dialogue, e.g. whether the proponent of a thesis has successfully defended it.

This model has been used to define a dynamic modal logic formalism for combinations of dialogue games (McBurney and Parsons 2001a). We next examine distinguishing characteristics of deliberations, and what implications these have for dialogue game rules.

3. Deliberation Dialogues

What distinguishes deliberation dialogues from other types of dialogue? A first characteristic arises from the focus of a deliberation, which is about what is to be done in some situation by some agent, either an individual or a group of individuals. This focus on action distinguishes deliberation dialogues from inquiry, information-seeking and eristic dialogues, although not from persuasion and negotiation dialogues; these latter two may also be about action. Moreover, information-seeking and inquiry dialogues involve a search for the true answer to some factual question, either by one participant or all. In such a search for truth, appeals to value assumptions (goals, preferences, etc) would be inappropriate. However, this is not the case for deliberations, where a course of action may be selected on the basis of such factors.

A second characteristic of deliberation dialogues is the absence of a fixed initial commitment by any participant on the basic question of the dialogue. Although the participants may express individual positions about what is to be done, the discussion is a mutual one directed at reaching a joint decision over a course of action; the actions under consideration, however, need not be joint, and may indeed be enacted by others. A deliberation dialogue is not, at least not at its outset, an attempt by one participant to persuade any of the others to agree to an initially defined proposal. In this respect, deliberation dialogues differ from persuasion dialogues.

A third characteristic of deliberations relates to their mutual focus. Although the participants may evaluate proposed courses of actions according to different standards or criteria, these differences are not with respect to personal interests which they seek to accommodate in the resulting decision. In this respect, a deliberation dialogue differs from a negotiation dialogue, which must deal with reconciling competing interests. In a negotiation, for example, it may be deleterious for a participant to share her information and preferences. But a sharing strategy should behoove participants in a deliberation; to the extent that agents are unwilling to share information or preferences, we would define their discussion to be a negotiation and not a deliberation.

These last two characteristics lead to an important observation about deliberations. An action-option which is optimal for the group when considered as a whole may be seen as sub-optimal from the perspective of each of the participants to the deliberation. This could be because a demonstration of optimality requires more information than is held by any one participant at the start of the dialogue, or because individual participants do not consider all the relevant criteria for assessment. Similarly, an option for which the group has a compelling argument may be such that no one participant, on his or her own, has such an argument; only by pooling information or resources is the group able to construct a winning argument for the option. This characteristic means that the common assumption by agent designers of an individual rationality condition on agent utterances (e.g. Amgoud *et al.* 2000a) is not appropriate: if we were to impose this condition, the optimal option may never be proposed, as no one participant has, on its own, an acceptable argument for it. We might call the individual rationality condition *narrow rationality* and distinguish it from the *broader rationality* of an agent considering both its own arguments and those of the group collectively. Moreover, real-life deliberations often benefit from whimsical or apparently-random proposals, which lead participants to discuss creative (*"off-the-wall"*) alternatives.

How do dialogues commence and proceed? Information-seeking dialogues, persuasions and inquiries each commence with a question or a statement by a participant and proceed by means of responses from other participants. Likewise, negotiation dialogues arise when a resource needs to be divided, and they can commence with a proposal by a participant to divide the resource in some manner, perhaps optimally for that participant. The negotiation will then proceed via responses to this proposal, including counter-proposals, which, in the best case, converge on a mutually acceptable settlement.

A deliberation dialogue arises with a need for action in some circumstance. In general human discourse, this need may be initially expressed in governing questions which are quite open-ended, as in *Where shall we go for dinner this evening?* or *How should we respond to the prospect of global warming?* Proposals for actions to address the expressed need may only arise late in a dialogue, after discussion on the governing question, and discussion on what considerations are relevant to its resolution. When possible courses of action are proposed, they may be evaluated on a large number of attributes, including: their direct or indirect costs and benefits; their opportunity costs; their consequences; their practical feasibility; their ethical, moral or legal implications; their resourcing implications; their likelihood of realization or of success; their conformance with other goals or strategies; their timing, duration or location; etc. To achieve resolution of a deliberation dialogue, one or more participants must make a proposal for an appropriate course of action. But where do such proposals for action arise? And how do the participants know when they have identified all the possible alternatives, or at least all those alternatives worth considering? These are not easy questions, for human or machine deliberators.

Negotiations over multi-attribute outcomes share the characteristic of multi-dimensionality with deliberations. Research on agent negotiation frameworks has typically made simplifying assumptions about such attributes and about agents' reactions to them, e.g. that the attribute values can be partially-ordered and that each agent has a real-valued utility function assigning values to potential outcomes which can be used to produce a rank order of outcomes on a single scale (Jennings *et al.* 2001). We desire not to make such assumptions for our model of deliberation dialogues, at least not in its most general form.

4. A Formal Model of Deliberations

Guided by the considerations discussed in the previous section, we now present a formal, high-level model for deliberation dialogues. This builds from the work of Joris Hulstijn (2000), who presented an idealized, five-stage model for negotiation dialogues, consisting of:

- Opening the dialogue;
- Sharing information;
- Making proposals and counter-proposals;
- Confirming accepted proposals;
- Closing the dialogue.

We propose a similar structure for deliberations.

We also draw on a philosophical model for non-deductive argument termed *retroflexive argumentation*, due to Harald Wohlrapp (1998). This talks of a *matter-in-question*, equivalent to a governing question or a proposal for action, being considered from a number of different *frames* or *perspectives*; we use the latter term, to avoid confusion with Reed (1998). As mentioned above, perspectives may be factors such as moral implications, opportunity costs, etc. An argument for or against a particular option is a partial understanding of that option from one or more, but rarely all, perspectives. Having heard an argument for or against an option, Wohlrapp argues, one proceeds by re-examining the underlying assumptions or modifying the action proposal, in the light of that argument. Thus, an argument against a law permitting euthanasia may be that such practices are open to abuse of ill patients by malicious relatives. A retroflexive response to this argument is to modify the proposed law by adding restrictions which inhibit or preclude such abuses, such as a requirement that the patient be of sound mind and give prior consent to the act of euthanasia.

With Wohlrapp's model in mind, we assume that the subject-matter of dialogues can be represented in a propositional language, with propositions and propositional functions denoted by lower-case Roman letters, e.g. *"p", "q"*. We define the following types of propositions:

- **Questions:** A question is a proposition, or a propositional function with one or more free variables (possibly conjoined with the proposition that exactly one sequence of objects satisfies the function), denoted by a lower-case Roman letter followed by a question-mark, e.g. *"p?"*. A governing question is the overall issue or issues which motivated the participants to convene the particular deliberation dialogue.

- **Actions:** An action is a proposition representing a deed or an act (possibly a speech act) which may be undertaken or recommended as a result of the deliberation dialogue. The purpose of the deliberation dialogue is to decide on an answer to the governing question, which will be some (course of) action. Possible actions are also called action-options.

- **Goals:** A goal is a proposition representing a future world state (external to the dialogue), possibly arising following execution of one or more actions and desired by one or more participants. Goals express the purpose(s) for which actions are being considered in the dialogue.

- **Constraints:** A constraint is a proposition expressing some limitation on the space of possible actions.

- **Perspectives:** A perspective is a proposition representing a criterion by which a potential action may be evaluated by a participant.

- **Facts:** A fact is a proposition expressing some possible state of affairs in the world external to the dialogue.

- **Evaluations:** An evaluation is a proposition expressing an assessment of a possible action with respect to a goal, constraint or perspective.

These types are mutually exclusive. With these elements defined, we now present a formal model of the dialogue itself, which consists of the following eight stages:

- **Open:** Opening of the deliberation dialogue, and the raising of a governing question about what is to be done.

- **Inform:** Discussion of: (a) the governing question; (b) desirable goals; (c) any constraints on the possible actions which may be considered; (d) perspectives by which proposals may be evaluated; and (e) any premises (facts) relevant to this evaluation.

- **Propose:** Suggesting of possible action-options appropriate to the governing question.

- **Consider:** Commenting on proposals from various perspectives.

- **Revise:** Revising of: (a) the governing question, (b) goals, (c) constraints, (d) perspectives, and/or (e) action-options in the light of the comments presented; and the undertaking of any information-gathering or fact-checking required for resolution. (Note that other types of dialogues, such as information seeking or persuasion, may be embedded in the deliberation dialogue at this stage.)

- **Recommend:** Recommending an option for action, and acceptance or non-acceptance of this recommendation by each participant.

- **Confirm:** Confirming acceptance of a recommended option by each participant. We have assumed that all participants must confirm their acceptance of a recommended option for normal termination.

- **Close:** Closing of the deliberation dialogue.

This is a model of an ideal dialogue. The stages may occur in any order, and may be entered by participants as frequently as desired, subject only to the following constraints:

- The first stage in every dialogue is the **Open** stage. Once a second participant enters the dialogue, the dialogue is said to be *"open."*

- The **Open** stage may occur only once in any deliberation dialogue. All other stages may occur more than once.

- The only stages which must occur in every dialogue which terminates normally are **Open** and **Close**.

- At least one instance of the **Inform** stage must precede the first instance of every other stage, excepting **Open** and **Close**.

- At least one instance of the **Propose** stage must precede the first instance of the **Consider**, **Revise**, **Recommend** and **Confirm** stages.

- The **Confirm** stage can only be entered following an instance of a **Recommend** stage.

- Upon successful completion of a **Confirm** stage, the dialogue must enter the **Close** stage.

- The last stage in every dialogue which terminates normally is the **Close** stage.

- Subject only to the constraints expressed in these rules and constraints expressed in the locution-combination rules (articulated in the Appendix), participants may enter any stage from within any other stage at any time.

Some comments are appropriate on the rules constraining the order of stages. Firstly, the participants may enter a **Close** stage more than once in a particular dialogue. As the locution rules below will demonstrate, participants are required to indicate publicly that they wish to leave the dialogue. Whenever a participant does this, the dialogue enters a **Close** stage. However, the **Close** stage remains unconcluded, and the dialogue remains open, as long as there are at least two participants who wish to continue speaking. It is therefore possible for the **Close** stage, as with all the other stages except the **Open** stage, to be entered multiple times in any one dialogue.

Secondly, we have assumed for simplicity in this initial model that unanimity of the participants is required for a decision on a course of action to be made. It would be perfectly possible for the participants to adopt a different procedure for confirmation, such as majority voting or consensus procedures. We have not done this here, but it is a topic for future work. If such alternative voting procedures were to be adopted, it would be useful to announce the results of any votes formally to the participants, with a statement of the group's decision, just as the minutes of human meetings usually record these. For this reason, we have demarcated a separate stage, **Confirm**, to record final commitments to action. In addition, the requirement that participants once again assert their endorsement for a particular course of action reinforces their commitment to this course as the group's decision. Once all participants have confirmed their acceptance of a recommended action, the dialogue must end, and any further discussion relevant to the same governing question can only occur by commencement of a new deliberation dialogue.

Apart from the constraints listed here, the order of stages is not fixed and participants may return to different stages multiple times in any one dialogue. Thus, a dialogue undertaken according to this model may cycle repeatedly through these stages, just as human dialogues do. In this way, our model gives practical effect to Wohlrapp's model of retroflexive argumentation. The eightfold model is also quite general; we have not specified the nature of the governing questions, goals, constraints, facts, action-options, perspectives or evaluations. Nor have we specified here any particular mechanisms for producing, revising or accepting action-options. Wohlrapp's model of retroflexive argumentation and our formalization of it have some similarities with Imre Lakatos' theory of mathematical discovery (Lakatos 1976). According to Lakatos, mathematicians work by proposing statements they believe may be theorems and then seeking proofs for these. In doing so, a counter-example to the proposed theorem may be found, which leads the mathematician to modify the proposal. A new attempt at seeking a proof is then undertaken, with the process repeated until such a time as a theorem is identified for which a proof can be found. The theories of Lakatos and Wohlrapp may be seen as describing (in part) arguments which proceed by *precization*, in the terminology of Arne Naess (1947/1966).

5. A Deliberation Dialogue Game

We now list a set of dialogue-game locutions which, taken together, enable a deliberation dialogue to be conducted according to the eight-stage model just presented. In this section, we present only the locutions, and not also the necessary pre-conditions for, and the consequences of, their utterance; these conditions are presented in detail in the Appendix. We continue to assume that the subject-matter of dialogues can be represented in a propositional language by lower-case Roman letters. We denote participating agents by P1, P2, . . . Pi, . . . and we assume that a Commitment Store, denoted *CS(Pi),* exists for each agent Pi. This store contains the various propositions which the agent has publicly asserted or preferences he or she has declared; entries in the store are thus of two forms: (a) 2-tuples of the form (*type, t*), where *t* is a valid proposition instance of type *type*, with *type* an element of the set: {*question, goal, constraint, perspective, fact, action, evaluation*}; and (b) 3-tuples of the form (*prefer, a, b*), where *a* and *b* are proposition actions. Each store can be viewed by all participants. The permissible locutions are:

- **open_dialogue(Pi, *q?*):** Participant Pi proposes the opening of a deliberation dialogue to consider the governing question *q?*. A dialogue can only commence with this move.

- **enter_dialogue(Pj, *q?*):** Participant Pj indicates a willingness to join a deliberation dialogue to consider the governing question *q?.* All intending participants other than the mover of open_dialogue(.) must announce their participation with this move. Note that neither the open_dialogue(.) nor the enter_dialogue(.) move implies that the speaker accepts that *q?* is the most appropriate governing question, only that he or she is willing to enter into a discussion about it at this time.

- **propose(Pi, *type, t*):** Participant Pi proposes proposition *t* as a valid instance of type *type*, where *type* is an element of the set {*question, goal, constraint, perspective, fact, action, evaluation*}.

- **assert(Pi, *type, t*):** Participant Pi asserts proposition *t* as a valid instance of type *type*, where *type* is an element of the set {*question, goal, constraint, perspective, fact, action, evaluation*}. This is a stronger locution than propose(.), and results in the tuple *(type,t)* being inserted into *CS(Pi)*, the Commitment Store of Pi.

- **prefer(Pi, *a, b*):** Participant Pi indicates a preference for action-option *a* over action-option *b*. This locution can only be uttered following utterance (possibly by other participants) of assert(Pj,*evaluation,e*) locutions of at least two evaluations *e*, one of which has *a* as its first argument, and one *b*. This combination rule ensures that preferences expressed in the dialogue are grounded in an evaluation of each action-option according to some proposed goal, constraint or perspective, and thus contestable. This locution inserts (*prefer, a, b*) into *CS(Pi)*, the Commitment Store of Pi.

- **ask_justify(Pj, Pi, *type, t*):** Participant Pj asks participant Pi to provide a justification of proposition *t* of type *type*, where *t* is in *CS(Pi).*

- **move(Pi, *action, a*):** Participant Pi proposes that each participant pronounce on whether they assert proposition *a* as the action to be decided upon by the group. This locution inserts (*action,a*) into *CS(Pi).*

- **retract(Pj, *locution*):** Participant Pj expresses a retraction of a previous locution, *locution*, where *locution* is one of three possible utterances: **assert(Pj, *type, t*)** or **move(Pi, *action, a*)** or **prefer(Pi, a, b)** locution. The retraction locution deletes the entry from *CS(Pi)* which had been inserted by *locution.*

- **withdraw_dialogue(Pi,*q?*):** Participant Pi announces her withdrawal from the deliberation dialogue to consider the governing question *q?*.

The locution **ask_justify(Pj, Pi, *type, t*)** is a request from participant Pj of participant Pi, seeking justification from Pi for the assertion that proposition *t* is a valid instance of type *type.* Following this, Pi must either retract the proposition *t* or shift into an embedded persuasion dialogue in which Pi seeks to persuade Pj that proposition *t* is such a valid instance. One could model such a persuasion dialogue with a formal dialogue-game framework consistent with the deliberation framework we present here, drawing, for example, on the models proposed by Walton and Krabbe (1995) or Prakken (2000).

The **move(.)** locution requests that participants who agree with a particular action being decided upon by the group should utter an **assert(.)** locution with respect to this action. Participants who do not agree that the particular action should be the decision of the group, or who wish to abstain from pronouncing on the issue, are free not to utter anything in reponse to the **move(.)** locution. Because in this model we have assumed unanimity of decision-making, the **Recommend** stage is only concluded successfully, and hence the dialogue only proceeds to the **Confirm** stage, in the case when all participants respond to the **move(.)** locution with the appropriate **assert(.)** locution.

We next show that our dialogue game framework implements the model for deliberation dialogues proposed in Section 4.

Proposition: *Each of the eight stages of the formal model of deliberation dialogues presented in Section 4 can be executed by judicious choice of these dialogue-game locutions.*

Proof: We consider each stage in turn:

- A dialogue opens with the locution **open_dialogue(Pi, *q?*)** and at least one utterance of **enter_dialogue(Pj,*q?*)**, for Pj and Pi distinct participants.

- The **Inform** stage consists of utterances of **propose(.)**, **assert(.)**, **retract(.)** and **ask_justify(.)** for some or all of the types *goal, constraint, perspective,* and *fact*.

- The **Propose** stage consists of one or more utterances of **propose(Pi, *action, t*)**.

- The **Consider** stage consists of utterances of **assert(Pi, *evaluation, e*)**, **prefer(Pj, a, b)** and **ask_justify(.)**.

- In the **Revise** stage, a revision *a2* to an action *a1* proposed earlier may be proposed by means of the locution **propose(Pi, *action, a2*)**.

- The **Recommend** stage consists of an execution of **move(Pi, *action, a*)**, possibly followed by utterances of **assert(Pj, *action, a*)**, for Pj and Pi distinct participants.

- The **Confirm** stage only occurs following a **Recommend** stage where all participants have indicated acceptance of the recommended action-option. It then consists of the utterance of **assert(Pj, *action, a*)** by every participant Pj, including the speaker of **move(Pi, *action, a*)**.

- The **Close** stage occurs whenever a participant Pi utters **withdraw_dialogue(Pi, *q?*).** A dialogue **closes** only when there remain two participants who have not uttered this locution, and one of them does so.

As mentioned at the beginning of this section, we have also defined for each locution the pre-conditions necessary for its legal utterance, and the post-conditions which occur upon its utterance, and these are presented in the Appendix. The locutions and the associated rules have been defined in accordance with the principles for rational mutual inquiry proposed by Hitchcock (1991).

6. Automated Dialogues

One objective of our work is the automation of deliberation dialogues between autonomous software agents. The formalism we have presented above provides the syntax for such dialogues, but is not sufficient for automation. Although there are some combination rules precluding or requiring locutions at various moves, agents still have a great deal of freedom in selecting utterances. We desire what we term a *generative* capability, so that dialogues can be generated automatically. We can achieve this by equipping the participating agents with mechanisms for deciding a preferred utterance at each move. In essence, these mechanisms are routines which are invoked by particular dialogue-game locutions, and, once invoked and executed, in turn invoke other locutions. We propose the following mechanism-types:

- **Recognize Need:** A mechanism which recognizes a need for a deliberation dialogue, enabling an agent to initiate or to enter such a dialogue.

- **Define Problem:** Mechanisms which identify and assess relevant questions, goals, constraints, facts and perspectives, enabling an agent to propose or consider these. As with purchase preferences provided in advance by a human principal to his or her delegated software agent in automated electronic commerce, an agent may enter the deliberation dialogue with these elements pre-determined.

- **Propose Option:** A mechanism to identify and assess possible action-options, enabling an agent to propose options, and to accept (or not) those proposed by others.

- **Consider Proposal:** Mechanisms to: (a) assert proposed questions, goals, constraints, facts and actions with respect to each other; (b) evaluate proposed actions with respect to goals, constraints or perspectives; and (c) seek appropriate justifications for the assertions of other participants.

- **Revise Proposal:** A mechanism to revise questions, goals, constraints, facts and perspectives in the light of **assert(Pi,*evaluation,e*)** and **prefer(Pi,a,b)** utterances in the dialogue-game.

- **Move Option:** A mechanism to select an action-option to move to the dialogue for joint acceptance.

- **Withdraw:** A mechanism to enable an agent to decide to withdraw from the dialogue at any time.

Proposition: *Autonomous software agents equipped with the mechanisms listed here can engage in deliberation dialogues of the form presented in Sections 4 and 5 automatically.*

Proof: On the basis of the mechanism definitions, the proof is straightforward; this can be seen by considering the impact of each permitted locution on the listed mechanisms, and vice versa.

This proposition says that the dialogue-game formalism we have presented for deliberation dialogues is generative (i.e. generates dialogues automatically) for agents with a certain high-level architecture. Similar approaches have been proposed in recent agent negotiation architectures which do not use argumentation. For example, Peyman Faratin (2000) equips agents engaged in automated negotiations with mechanisms for: (a) deciding their responses to multi-attribute offers; (b) proposing new offers involving different trade-offs of the same set of attributes as prior offers; and (c) proposing new offers having different attributes to prior offers. In Jennings *et al.* (2001), these mechanisms are called *heuristic* approaches to automated negotiation, and are distinguished from approaches using either economic game theory or argumentation. We believe our model is the first in which heuristic and argumentation approaches have been combined. In other work, two of us have proposed an evolutionary computational architecture as the basis for a generative mechanism (McBurney and Parsons 2001b).

7. Example

We consider a simplified example regarding what action to take regarding potential health hazards from the use of cellular phones. The dialogue moves are annotated following each move.

- **open_dialogue**(P1,*Do what about mobile phone health risk?*)

  This move is the first move in the Open stage of the dialogue.

- **enter_dialogue**(P2, *Do what about mobile phone health risk?*)

  With the entry of a second participant, the dialogue may be said to commence.

- **enter_dialogue**(P3, *Do what about mobile phone health risk?*)

  A third participant also enters the dialogue.

- **propose**(P2, *perspective, degree of risk*)

  Participant P2 proposes the degree of risk as a perspective from which to consider the question. With this move, the dialogue enters an Inform stage.

- **propose**(P3, *perspective, economic cost*)

  Participant P3 proposes economic cost as a perspective from which to consider the question.

- **propose**(P1, *action, prohibit sale*)

  Participant P1 proposes prohibition of sale of phones as an action-option. With this move, the dialogue enters a Propose stage.

- **propose**(P3, *action, do nothing*)

Participant P3 proposes doing nothing as an action-option.

- **assert**(P1, *evaluation, prohibit sale from a degree of risk perspective is lowest risk*)

  Participant P1 asserts that from the perspective of the degree of risk, prohibiting the sale of phones is the lowest risk action-option possible. With this move, the dialogue enters a Consider stage.

- **assert**(P3, *evaluation, prohibit sale from an economic cost perspective is high-cost*)

  Participant P3 asserts that from the perspective of economic cost, prohibiting sale is a high-cost option.

- **propose**(P1, *action, limit usage*)

  Participant P1 proposes limiting usage as an action-option, thus responding retroflexively to the previous two assert(Pi,*evaluation,e*) locutions. With this move, the dialogue enters a Revise stage.

- **propose**(P2, *perspective, feasibility*)

  Participant P2 proposes feasibility as a perspective from which to consider the question. With this move, the dialogue enters another Inform stage.

- **assert**(P2, *evaluation, limit usage from a feasibility perspective is impractical*)

  Participant P2 asserts that from the perspective of feasibility, limiting usage is not practical. With this move, the dialogue enters another Consider stage.

- **prefer**(P1, *prohibit sale, limit usage*)

  Participant P1 expresses a preference for the option of prohibiting the sale of phones over limiting their usage. The utterance is valid at this point, since each action-option has appeared as the first argument in a proposition *e* of type *evaluation* in an assert(Pi, *evaluation, e*) locution.

For reasons of space, we have not included any retraction locutions, nor included any goals, constraints or facts. For the same reason, we have not continued the example to the Recommend, Confirm and Close stages. Although very simplified, the example does show the usage of some types of locutions; it also demonstrates the way in which a dialogue may move between stages as it proceeds. Such cycling between stages is commonplace in human deliberations, where comments, arguments and preferences uttered by one participants are likely to provoke others to think of new goals, constraints, facts, perspectives and action-options.

8. Discussion

In this paper we have proposed the first formal model of a general deliberation dialogue, grounding it in the philosophy of argumentation and using a dialogue-game framework to ensure implementability. Our model creates a public space in which multiple participants may interact to jointly decide on a course of action, and our structure and rules seek to define the nature of these interactions. In enabling participants to contribute to a joint discussion which may proceed iteratively and to view each other's commitment stores, our model has some similarities with *"blackboard"* architectures in computer science (Nii 1986). We do not capture, in our model, all types of deliberation dialogues; nor is it sufficient for automated dialogues. To generate dialogues automatically we would need to equip the participating agents with mechanisms enabling them to choose between the permitted locutions at each move in the dialogue-game. We have outlined a portfolio of such mechanisms in this paper, analogous to recent work (not using argumentation) in devising frameworks for automated negotiation. Although at a high level, we believe our work is the first in which a generative formalism has been proposed for agents engaged in deliberative argument.

Much research effort in Artificial Intelligence (AI) over the last thirty years has concerned the task of designing robots so that, when given a specific goal, such as moving into the next room, they may determine a plan for achievement of this goal. Because this research area, known within AI as *"Planning"*, concerns consideration of possible actions, it would seem amenable to the application of deliberation dialogues. However, the only research program known to us which combines AI Planning with models of dialogues is the TRAINS project (Allen *et al.* 1995), which constructed an intelligent computer assistant for a human rail-freight scheduler. For this project, actual human-human conversations in the specific domain were first recorded and analyzed as a basis for the design of the machine-human interactions. Although the two participants in the TRAINS system, machine and human, discuss a course of action, and thus ostensibly engage in deliberation, the design of the system assumes that the machine and the human-user each begin the dialogue with a privately-developed proposal for action, which they then present to one another. Thus, in the terminology of Walton and Krabbe (1995), their conversation is closer to a two-way persuasion dialogue than to a true deliberation. In addition, the TRAINS system design assumes that the human user's goal is paramount, and that the machine participates in the dialogue to assist the human to find an effective plan for achievement of this goal. Thus, the model of dialogue assumes a specific relationship of inequality between the two participants. By contrast, our model of deliberation dialogue is not limited in this way.

Other work in Artificial Intelligence has also come close to developing a formal model of deliberation dialogues without yet doing so. Frank Dignum and colleagues have used dialogue models for the creation of a collective intention and for team-formation by autonomous software agents (Dignum *et al.* 2000a, 2000b) seeking to engage in some joint activity. This research also arises from within an AI Planning tradition, where the overall goal is assumed pre-determined. As with the TRAINS project, while such a context does not preclude use of deliberation dialogues, the focus of the research has been on other types of dialogue. The models of Dignum *et al.* are explicit combinations of persuasion and negotiation dialogues, the latter embedded within the former. The example dialogue of Parsons *et al.* (1998), in which agents collaborate to furnish a room, is a mixture of deliberation and negotiation, but this is not modeled formally. Later work by Parsons and colleagues (e.g. Amgoud *et al.* 2000a, 2000b) has presented dialogue-game models of negotiations. To date, negotiation has been the primary focus of AI researchers exploring automated interactions between autonomous software agents (Jennings *et al.* 2001), perhaps because of the potential applications to the design of electronic-commerce systems. This focus has led to work on distributed proof procedures in multi-agent negotiations (e.g. Fisher 2000) which may be seen as analogous to the model of distributed decision-making we have presented here. By contrast, AI researchers using dialogue-games in the legal domain have primarily focused on persuasion dialogues (e.g. Prakken 2000).

We are exploring a number of extensions of the work in this paper. Firstly, we seek to model and automate more general classes of deliberation dialogue. For example, many human deliberations exhibit strong disagreement between the participants over the relevance and importance of different perspectives. We will explore the question of how our dialogue-game may be extended to allow for arguments over these. Secondly, we plan to enable discussion over confirmation procedures, so that, for example, majority or plurality voting may be used instead of the unanimity now required in the Confirm stage. To do this will require the addition of a procedural discussion stage (or stages) to the dialogue model, along with locutions appropriate for such discussions. If a group of participants were to engage regularly in deliberation dialogues using the same decision-procedures, these procedural discussions would not need to be undertaken in each dialogue but could be assumed constant. Within the multi-agent systems community, systems for interaction between autonomous agents with such pre-determined rules of encounter have been called *Institutions* (Sierra *et al.* 1998).

A third extension could involve the formal modeling of trust and obligation among the participants. John Forester (1999) argues that, in the domain of land-use planning, a key task of planning professionals in facilitating public policy development and decision-making is the development of trust between the various stakeholders and the planning professional. Recent research in AI has looked at the formal modeling of trust and obligation (e.g. McNamara and Prakken 1998). Finally, our explicit typing of propositions (into facts, goals, constraints, etc) may facilitate the mathematical representation of dialogues under this model by means of the lambda-calculus (Church 1940), a representation we may also explore.

Acknowledgments

This work was partly funded by the European Union IST Programme, through the Sustainable Lifecycles in Information Ecosystems Project (IST-1999-10948), and by the British Engineering and Physical Sciences Research Council (EPSRC), through the *Symposium on Argument and Computation* (SAC) project (GR/N35441/01) and a Ph.D.

studentship. We gratefully acknowledge this support, and thank Chris Reed and Tim Norman for initiating and organizing the SAC. We also thank Joris Hulstijn and Harald Wohlrapp for their comments on an earlier draft of this paper.

Appendix:

In this section, we present the full list of pre-conditions and post-conditions for each locution defined in Section 5.

- **open_dialogue(.)**

  - **Locution:** **open_dialogue(Pi, *q?*)**

  - **Meaning:** Participant Pi proposes the opening of a deliberation dialogue to consider the governing question *q?*. A dialogue can only commence with this move.

  - **Preconditions:** No preconditions within the dialogue..

  - **Response:** None required. Other intending participants may respond with the **enter_dialogue(.)** locution.

  - **Commitment Store Update:** No effects.

- **enter_dialogue(.)**

  - **Locution:** **enter_dialogue(Pj, *q?*)**

  - **Meaning:** Intending participant Pj indicates a willingness to join a deliberation dialogue to consider the governing question *q?*. All intending participants other than the speaker of **open_dialogue(.)** must announce their participation with this move.

  - **Preconditions:** A participant Pi, where Pi and Pj are distinct, must previously have uttered the locution **open_dialogue(Pi, *q?*).**

  - **Response:** None required. This locution is a pre-condition for all locutions other than **open_dialogue(.)**, i.e. the intending speaker Pj of any other locution must have previously uttered **enter_dialogue(Pj, *q?*)**. As soon as one participant has uttered the **enter_dialogue(Pj, *q?*)** locution, the dialogue is said to be *Open.*

  - **Commitment Store Update:** No effects.

Since all the locutions listed below have a common precondition, namely that the speaker Pj have previously uttered either **open_dialogue(Pj, *q?*)** or **enter_dialogue(Pj, *q?*)**, we do not list this precondition under each locution, but only those specific to the locution.

- **propose(.)**
  - **Location:** **propose(Pi, *type, t*)**
  - **Meaning:** Participant Pi proposes proposition *t* as a valid instance of type *type*, where *type* is an element of the set {*question, goal, constraint, perspective, fact, action, evaluation*}.
  - **Preconditions:** No specific preconditions.
  - **Response:** None required.
  - **Commitment Store Update:** No effects.

- **assert(.)**
  - **Location:** **assert(Pi, *type, t*)**
  - **Meaning:** Participant Pi asserts proposition *t* as a valid instance of type *type*, where *type* is an element of the set {*question, goal, constraint, perspective, fact, action, evaluation*}.
  - **Preconditions:** No specific preconditions.
  - **Response:** None required.
  - **Commitment Store Update:** The 2-tuple *(type,t)* is inserted into *CS(Pi)*, the Commitment Store of Pi.

- **prefer(.)**
  - **Location:** **prefer(Pi, *a, b*)**
  - **Meaning:** Participant Pi indicates a preference for action-option *a* over action-option *b*.
  - **Preconditions:** Some participants Pj and Pk, possibly including Pi, must previously have uttered the locutions **assert(Pj, *evaluation, e*)** and **assert(Pk, *evaluation, f*)**, where *e* and *f* are evaluation propositions which refer respectively to action-options *a* and *b*.
  - **Response:** None required.

- **Commitment Store Update:** The 3-tuple (*prefer, a, b*) is inserted into *CS(Pi)*, the Commitment Store of Pi.

- **ask_justify(.)**

  - **Locution:** **ask_justify(Pj, Pi, *type, t*)**

  - **Meaning:** Participant Pj asks participant Pi to provide a justification of proposition *t* of type *type*, where *t* is in *CS(Pi)*.

  - **Preconditions:** Participant Pi has previously uttered the locution **assert(Pi, *type, t*).**

  - **Response:** Pi must either retract the proposition *t* or seek to persuade Pj in an embedded persuasion dialogue that proposition *t* is a valid instance of type *type.*

  - **Commitment Store Update:** No effect.

- **move(.)**

  - **Locution:** **move(Pi, *action, a*)**

  - **Meaning:** Participant Pi proposes that each participant pronounce on whether they assert proposition *a* as the action to be decided upon by the group.

  - **Preconditions:** Some participant Pj, possibly Pi, must previously have uttered either **propose(Pi, *action, a*)** or **assert(Pi, *action, a*).**

  - **Response:** None required. Other participants Pjk, distinct from Pi, who wish to support proposition *a* as the action to be decided upon by the group can respond with locution **assert(Pk, *action, a*).**

  - **Commitment Store Update:** The 2-tuple (*action,a*) is inserted into *CS(Pi).*

- **retract(.)**

  - **Locution:** **retract(Pi, *locution*)**

  - **Meaning:** Participant Pi expresses a retraction of a previous utterance *locution,* where *locution* is one of the following three locutions: **assert(Pi, *type, t*)**, **move(Pi, *action, a*)** or **prefer(Pi, *a, b*).**

  - **Preconditions:** Participant Pi must have previously uttered and not subsequently retracted the locution *locution.*

- **Response:** None required.

- **Commitment Store Update:** Depending on which of the three locutions mentioned in the preconditions was uttered previously, one of: (a) the 2-tuple (*type,t*); (b) the 2-tuple (*action,a*); or (c) the 3-tuple (*prefer,a,b*) is deleted from from *CS(Pi).*

▪ **withdraw_dialogue(.)**

  - **Locution:** **withdraw_dialogue(Pi, *q?*)**

  - **Meaning:** Participant Pi announces her withdrawal from the deliberation dialogue to consider the governing question *q?.*

  - **Preconditions:** No specific preconditions.

  - **Response:** None required. If only two participants remain in a dialogue and one of these utters this locution, the dialogue terminates.

  - **Commitment Store Update:** No effects.

References

Allen, J. F., L. K. Schubert, G. Ferguson, P. Heeman, C. H. Hwang, T. Kato, M. Light, N. G. Martin, B. W. Miller, M. Poesio and D. R. Traum (1995): "The TRAINS Project: a case study in building a conversational planning agent." *Journal of Experimental and Theoretical Artificial Intelligence*, 7: 7—48.

Amgoud, L., N. Maudet and S. Parsons (2000a): "Modelling dialogues using argumentation." In: E. Durfee (Editor): *Proceedings of the Fourth International Conference on Multi-Agent Systems* (ICMAS-2000), pp. 31—38. Boston: IEEE Press.

Amgoud, L., S. Parsons and N. Maudet (2000b): "Arguments, dialogue, and negotiation." In: W. Horn (Editor): *Proceedings of the Fourteenth European Conference on Artificial Intelligence (ECAI 2000)*, pp. 338—342. Berlin, Germany: IOS Press.

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.

Church, A. (1940): "A formulation of the simple theory of types." *Journal of Symbolic Logic*, 5: 56—68.

Dignum, F., B. Dunin-Keplicz and R. Verbrugge (2000a): "Agent theory for team formation by dialogue." In: C. Castelfranchi and Y. Lesperance (Editors): *Proceedings of the Seventh International Workshop on Agent Theories, Architectures, and Languages (ATAL-2000),* pp. 141—156. Boston, USA.

Dignum, F., B. Dunin-Keplicz and R. Verbrugge (2000b): "Creating collective intention through dialogue." In: J. Cunningham and D. Gabbay (Editors): *Proceedings of the International Conference on Formal and Applied Practical Reasoning (FAPR-2000)*, pp. 145—158. London, UK: Department of Computing, Imperial College, University of London.

Faratin, P. (2000): *Automated Service Negotiation Between Autonomous Computational Agents*. PhD Thesis. Department of Electronic Engineering, Queen Mary College, University of London, London UK.

Faratin, P., C. Sierra and N. Jennings (2000): "Using similarity criteria to make negotiation trade-offs." In: E. Durfee (Editor): *Proceedings of the Fourth International Conference on Multi-Agent Systems (ICMAS-2000)*, pp. 119—126. Boston, MA, USA: IEEE Press.

Fisher, M (2000): "Characterising simple negotiation as distributed agent-based theorem-proving -- a preliminary report." In: E. Durfee (Editor): *Proceedings of the Fourth International Conference on Multi-Agent Systems (ICMAS-2000).* Boston, MA, USA: IEEE Press.

Forester, J. (1999): *The Deliberative Practitioner: Encouraging Participatory Planning Processes.* Cambridge, MA, USA: MIT Press.

Hamblin, C. L. (1970): *Fallacies.* London, UK: Methuen.

Hamblin, C. L. (1971): "Mathematical models of dialogue." *Theoria*, 37: 130—155.

Hitchcock, D. (1991): "Some principles of rational mutual inquiry." In: F. van Eemeren et al. (Editors): *Proceedings of the Second International Conference On Argumentation*, pp. 236—243. Amsterdam: SICSAT.

Hulstijn, J. (2000): *Dialogue Models for Inquiry and Transaction.* PhD Thesis, Universiteit Twente, Enschede, The Netherlands.

Hulstijn, J., M. Dastani and L. van der Torre (2000): "Negotiation protocols and dialogue games." In: *Proceedings of the Belgian-Dutch Conference on Artificial Intelligence (BNAIC-2000)*. Kaatsheuvel, The Netherlands.
Jennings, N. R., P. Faratin, A. R. Lomuscio, S. Parsons, C. Sierra and M. Wooldridge (2001): "Automated negotiation: prospects, methods and challenges." *Journal of Group Decision and Negotiation*, 10 (2): 199—215.
Lakatos, I. (1976): *Proofs and Refutations: The Logic of Mathematical Discovery.* Cambridge, UK: Cambridge University Press.
MacKenzie, J. D. (1979): "Question-begging in non-cumulative systems." *Journal of Philosophical Logic*, 8: 117—133.
McBurney, P. and S. Parsons (2001a): "Agent Ludens: games for agent dialogues." In: P. Gmytrasiewicz and S. Parsons (Editors): *Proceedings of the Stanford Spring Symposium on Game-Theoretic and Decision-Theoretic Agents*, Menlo Park, CA, USA: AAAI Press.
McBurney, P. and S. Parsons (2001b): "Chance Discovery using dialectical argumentation." In: Y. Ohsawa (Editor): *Proceedings of the Workshop on Chance Discovery, Fifteenth Annual Conference of the Japanese Society for Artificial Intelligence (JSAI-2001)*, Japanese Society for Artificial Intelligence, Japan (to appear).
McNamara, P. and H. Prakken (Editors) (1998): *Norms, Logics and Information Systems: New Studies in Deontic Logic and Computer Science.* Amsterdam, The Netherlands: IOS Press.
Naess, A. (1947/1966): *Communication and Argument: Elements of Applied Semantics.* London, UK: Allen and Unwin. (Translation from the Norwegian of: *En del Elementaere Logiske Emner.* Universitetsforlaget, Oslo, Norway, 1947.)
Nii H.P. (1986): "Blackboard Systems (Part One): The blackboard model of problem solving and the evolution of blackboard architectures." *The AI Magazine*, Summer 1986, pp. 38—53.
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.
Prakken, H. (2000): "On dialogue systems with speech acts, arguments, and counterarguments." In: M. Ojeda-Aciego, M. I. P. de Guzman, G. Brewka and L. M. Pereira (Editors): *Proceedings of the seventh European Workshop on Logics in Artificial Intelligence (JELIA-2000).* Lecture Notes in Artificial Intelligence 1919. Berlin, Germany: Springer.
Reed, C. (1998): "Dialogue frames in agent communications." In: Y. Demazeau (Editor): *Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS-98)*, pp. 246—253. IEEE Press.
Rescher, N. (1977): *Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge.* Albany, NY, USA: State University of New York Press.
Sierra, C., N. R. Jennings, P. Noriega and S. Parsons (1998): "A framework for argumentation-based negotiation." In: M. P. Singh, A. Rao and M. J. Wooldridge (Editors): *Intelligent Agents IV: Agent Theories, Architectures, and Languages. Proceedings of the Fourth International ATAL Workshop*, pp. 177—192. Lecture Notes in Artificial Intelligence 1365. Berlin, Germany: Springer Verlag.
Stathis, K. (2000): "A game-based architecture for developing interactive components in computational logic." *Electronic Journal of Functional and Logic Programming*, 2000 (5).

Traum, D. R. and J. F. Allen (1992): "A speech acts approach to grounding in conversation. " In: *Proceedings of the Second International Conference on Spoken Language Processing*, pp. 137—140.
Walton, D. N. and E. C. W. Krabbe (1995): *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning.* Albany, NY: State University of New York Press.
Wohlrapp, H. (1998): "A new light on non-deductive argumentation schemes." *Argumentation*, 12: 341-350.