Democracy in Open 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 & Information Science
Brooklyn College, City University of New York
Brooklyn NY 11210 USA
parsons@sci.brooklyn.cuny.edu

ABSTRACT

What sort of democracies are open agent societies? We present three normative models of democracy from political philosophy and consider their relevance for open multi-agent systems: democracy as wise rule by an elite; democracy as the exercise of rational consumer choices by voters; and democracy as deliberative decision-making by citizens. We consider the implications of these different models for the design of open systems, in terms of the communications language, the interaction protocol, and the conflict-resolution mechanism used by the agents involved. We also consider the issue of verifiability of the internal semantics of communications languages, and argue that a model of agent democracy based on deliberative democracy provides the basis for a form of verifiability which is stronger than a social semantics.

ACM CATEGORIES AND SUBJECT DESCRIPTORS:
I.2.11 **[Artificial Intelligence]** Distributed Artificial Intelligence: *Coherence and co-ordination; Languages and Structures; Multiagent systems*.

GENERAL TERMS: Languages, Theory.

KEYWORDS: Agent communications languages, deliberative democracy, open agent systems, political philosophy, philosophical foundations, rational choice, rationality, social semantics.

1. INTRODUCTION

Open agent systems are multi-agent systems which any software agent may join. Entry may require compliance with particular stated conventions, such as use of an agent communication language and interaction protocol, or the making of a financial deposit. However, subject only to such conventions, any software agent may join. Because such agents may represent different human principals and typically will have been constructed by different software design teams, they may, in general, have conflicting goals, interests, beliefs and values. In these circumstances, which agent's goals or beliefs prevail in the interaction will depend on the nature of the social and political relationships between the participants. In situations where the agents adhere to some hierarchical relationship inside the agent system, that agent or agents at the top of the hierarchy may have final decision-making authority. For example, in an auction interaction, the auctioneer may have the explicit power to determine the final allocation of the scarce resources being sought by the bidders. Power such as this may not reside in particular agents, but accrue to certain roles within the agent system, as in the electronic institutions of [29].

However, if human interaction is any guide, in many open environments there will either be no such hierarchy, or what hierarchies there are may be contested by some participants. Indeed, this is already true of existing agent societies on the Internet. What structures are appropriate for agent societies in these circumstances? The absence of hierarchy means that the relationship between the participants is closer to one of equality; this in turn suggests that some form of democracy is appropriate when we consider the structure of these agent systems. Within the discipline of political philosophy, human democracy is a notion much debated, and there are several alternative normative theories of democracy. A designer of an open agent system intending to permit democratic participation by the agents in the system therefore has a choice of theories of democracy to encode in the system. In this paper, we explore these alternatives from political theory, in order to identify what structures they provide for, and what constraints they place on, designers of multi-agent systems. In Section 2 we present the primary normative theories of democracy developed by political philosophers, and then discuss, in Section 3, their implications for the design of open agent systems. It happens that one theory, the Deliberative Model of democracy, stresses the joint and discursive nature of decision-making in a democracy, with participants exchanging arguments for and against various policy proposals, and forming preferences on the basis of these exchanges. The structure that this model provides to the agent system designer creates the means necessary to develop a strong form of social semantics, thereby increasing the extent to which a mentalistic semantics of an agent communications languages can be verified. This view is explained in Section 4. The paper concludes with a summary in Section 5.

2. MODELS OF DEMOCRACY

Political philosophers have articulated several normative models of democracy. In their most abstract form, these models all view a polity as comprising just two entities: *Society* and *the State*. Society is the collection of individuals, organizations and companies in a polity, together with the panoply of relationships between them, while the State is the apparatus of public-sector administration. The key question for political theory is then: *What should be the process of formation of political will?* or *How should Society program*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.
Copyright 2002 The Authors ..

the State? Supporters of democracy, beginning with the French Enlightenment philosopher, Jean-Jacques Rousseau [34], believe that these questions should be answered with the use of democratic procedures, such as elections based on universal adult suffrage. But if such procedures are used, what is the nature of the relationship between citizens and their elected representatives? Rousseau had assumed that the people have a "general will" which their elected representatives should seek to implement, but this is at best a high-level approximation to the multifarous cacophony which is modern democracy.

The first modern political theory of democracy which sought to answer this question was proposed in 1942 by Austrian-American economist Joseph Schumpeter [35]. Schumpeter's theory, possibly in reaction to the mass populism of Nazism and Communism and to his own failed political career, was disdainful of ordinary people and their views:[1] *"Thus the typical citizen drops down to a lower level of mental performance as soon as he enters the political field. He argues and analyzes in a way which he would readily recognize as infantile within the sphere of his real interests. He becomes a primitive again."* [35, p. 262]. Consequently, Schumpeter proposed that elected officials should act as a technocratic elite, making decisions on behalf of the general public and in accordance with what the elite believes are the public's best interests. Apart from voting, the people are entirely passive in Schumpeter's model of democracy, which has rightly acquired the label *elitist* [7, p. x]. We call this the *Wise Elite Model* of democracy.

In contrast to Schumpeter's inherently non-democratic view of democracy, Anthony Downs proposed an economic-theoretic model of political will-formation in a democracy in which citizens were more than simply passive objects [9]. This model has since been called a *rational-choice* or *liberal* model [16], and it views democracy as akin to the operation of an economic market. Downs proposes a theory of democracy where political parties and interest groups act as entrepreneurs, offering alternative "products" in the form of bundles of state-instructions (or equivalently, ideologies, which are philosophies of bundle-formation), to voters who then "purchase" their preferred bundle when they vote. That bundle with the greatest "market-share" — in the form of popular votes — becomes the set of instructions used to program the State.

Downs explicitly assumed that voter-consumers in a free and democratic society make their political choices on the basis of their perceived self-interest, and act according to the now-standard definition of rational economic behavior, e.g., [4, 23]. In other words, voters are assumed to always vote so as to maximize their perceived expected utility from the outcome of the election. In addition to consuming bundles of state-instructions, citizens also consume information about policies, ideologies, political parties and candidates to the extent necessary to make their voting decisions. And, as for any other good, such consumption may be subject to time-, resource-, or processing-constraints, and cost-benefit trade-offs.

The rational-choice model affords citizens a greater role than does the elitist model, namely that of consumers of relevant political information and of recipients of the effects of policies enacted by their representatives. But citizens, in the rational-choice model of democracy, are not regarded as *producers* of political information or public policies. This viewpoint ultimately stems, we believe, from Downs' adoption of Kenneth Arrow's operational definition of economic rationality [4], which assumes that a decision-maker's preferences and utilities are given and precede the task of selection of a decision-option. In many, if not all, public policy determinations, however, the preferences and utilities of voters may only be formed in the very process of decision-making, as participants learn about feasible decision-options and about the effects of various decision-options on one another and on others not involved in the decision process. Moreover, to the extent that a person's utility depends on the welfare of others, a rational voter would not finally determine his or her preferences until hearing from those others about their own utilities and preferences.[2]

In contrast, the *deliberative democracy* model of political will-formation emphasizes the manner in which beliefs and preferences of participants are formed or change through the very process of interacting together [6, 8]. In this model, citizens do not merely interact to exchange their preferences at election time, and to consume political information, as is the case with the rational-choice model. Rather, they are also producers of political information and policies, as they participate in political processes and debate, identify and publicize issues of personal or social concern, exchange arguments for and against various policy options, and generally seek to influence the outcomes of political decision processes. Seeking to influence and persuade other participants means that they must themselves be open to persuasion, and thus undergo what has been called *self-transformation* [14, p. 184]. As Frank Michelman defined it [27, p. 293]:

"Deliberation . . . refers to a certain attitude toward social co-operation, namely, that of openness to persuasion by reasons referring to the claims of others as well as one's own. The deliberative medium is a good faith exchange of views — including participants' reports of their own understanding of their respective vital interests — . . . in which a vote, if any vote is taken, represents a pooling of judgments."

The rational-choice and deliberative models of democracy embody different notions of rationality. As mentioned above, Downs' economic theory of democracy was based explicitly on Arrow's [4, Chapters 1 and 2] definition of rational behaviour as the maximization of expected perceived utility. This in turn was an operationalization of Lange's notion of rationality [23, p. 30]: *"A unit of economic decision is said to act rationally when its objective is the maximization of a magnitude."* This notion of rationality, although predominant in economics,[3] is not how the word is understood in the philosophy of argumentation [21, 15]. For example, Ralph Johnson [21, p. 14] gives this definition: *"Rationality is the ability to engage in the practice of giving and receiving reasons."* The deliberative model of democracy, because it construes democracy as the joint determination of public policy by the citizenry through debate and exchanges of views, embodies the philosophers' notion of rationality rather than the economists' understanding, although the latter is not precluded. For agents who know the reasons for their own actions, the maximum-expected-utility notion of rationality is a special case of argumentation-theoretic rationality, since the method used to select an action-options can be advanced as a reason for the option by a speaker in a debate. However, the converse is not true, because there may be many reasons for selecting an option which does not maximize expected utility, e.g., that it avoids catastrophic downside loss.[4]

These three models of political will-formation in a democracy can be seen as offering alternative roles to the citizens who comprise the Society. In the Wise Elite model, the people are seen as

[1] Given these views, it is hard for a contemporary believer in universal human equality to understand why Schumpeter's book achieved such prominence and respect.

[2] For example, one person's utility from a network good, such as a fax machine, depends on whether or not other people have them.

[3] and influential in AI. It has been subject to much criticism, however; see, for example, [20].

[4] Habermas [15, p. 10] calls these two notions of rationality *cognitive-instrumental* and *communicative*, respectively.

completely passive, except when choosing the Elite. In the Rational-Choice model, the people are viewed as consumers of policies, ideologies and information. In the Deliberative model, the people are viewed as both consumers and producers of policies, ideologies and political information.

3. DESIGN IMPLICATIONS FOR MAS

Which of these theories of democracy is appropriate for the design of open multi-agent systems? The answer, of course, depends on the intended objectives of the system, and the nature of the application. In systems with differentiated agent roles and an explicit hierarchy some version of the Wise Elite model of democracy may be appropriate. An example here are public auction sites, in so far as the auctioneer makes decisions on behalf of the agents comprising the auction. Bidders have the freedom to join or not any given auction site, and so may be viewed as "electing" the elite in the form of the auctioneer. Once the auction is underway, however, bidders express preferences through their bids, and the rules of the auction mechanism may resolve these into a collective decision, rather than leaving the resolution to the wisdom of the auctioneer. Another example of the Wise Elite model is the use of agents for security and access-policing functions, similar to the space-administration objects of [10]. One can readily imagine the participants in an open agent system agreeing to delegate a limited amount of their joint power to a group of policing agents, who exercise that power in pursuit of collective aims of security and confidentiality which all agree are essential. For example, the interaction protocol rules may permit any participant to speak at any time, but a policing agent could prevent participants monopolizing the microphone by limiting usage from verbose (or badly-coded) agents. Thus, the collective goal of fair distribution of microphone access takes precedence over an individual agent's goal of exclusive use of the microphone, and the policing agents act to ensure this on behalf of all agents.

However, in most open agent systems, such voluntary ceding of power will not occur on matters of concern to the participants. Participants are likely to disagree with one another over such issues, and will wish to express their own beliefs, preferences and intentions. Agent autonomy means that software agents cannot in general be *ordered* to fulfill requests; they may be requested and, at best, *persuaded* to do so. If their beliefs, preferences and intentions are predetermined and fixed, no amount of persuasion will change these, and so the Rational Choice model of democracy would be appropriate. Here the only sensible interaction mechanism between the participating agents would be some form of preference aggregation or voting, since the exchange of reasons for beliefs or preferences would not alter decision outcomes. Auction mechanisms are examples of open multi-agent societies where agents are usually assumed to have pre-determined (although not necessarily fixed) preferences, and where no party seeks to persuade another to change these. In this case, democracy as the expression of preferences, as in the Rational Choice model, is sufficient to represent the relationships between the agents concerned.

In other domains, however, agents may well seek to influence the beliefs, preferences or intentions of others. These could involve domains where the agents interact in order to achieve a joint belief, as in a scientific inquiry [25], or where they seek to form a joint intention, as in a employer-union negotiation [37], or where they seek to co-ordinate their separate intentions in order, for example, to make effective use of shared resources [30]. Whenever the relationship between agents in a multi-agent system is one of equality, and where agents seek to influence the beliefs, preferences or intentions of one another, then a Deliberative model of democracy will be the most appropriate model for the design of the system.

Adoption of a particular model of democracy for a multi-agent system has a number of design implications, which we explore in the next three subsections. These relate to the design of the communications language, any interaction protocol, and any disagreement-resolution mechanisms. We consider each of these topics in turn.

3.1 Communications Languages

The different models of democracy place different requirements on the communications language required for agent interaction. The Wise Elite Model requires that agents have some means to select the elite. But, other than this, no other expressions of beliefs, preference, intentions, etc, need be expressed by the participants, since all decisions are made by the elite. Under the Rational Choice model, participants express preferences for or between policy options but not necessarily arguments for these preferences. Thus, the communications language needs to be able to support the expression of preferences, either directly or by means of acceptance or rejection of particular proposals. Auction protocols, such as the FIPA ACL English auction protocol [13], typically permit the expression of preferences through utterances of acceptance of particular proposals.

Adoption of a deliberative model for democracy in a multi-agent system requires that the communications language allows each agent to express, not only its beliefs, intentions or preferences, but also its arguments for or against these. Participants require the ability to question or challenge the statements of others, and to defend and justify their own statements. Thus, the communications language needs to be able support the expression of arguments for statements, as well as expression of the statements themselves. There have been a number of proposals for agent communications languages providing this capability in recent years, e.g., [2, 24, 30].

In addition, if participants are to engage in debate with the potential to persuade one another to adopt new beliefs, intentions or preferences, then agents need to be able to withdraw prior statements and utter replacements in their stead. If one agent's goal in an interaction is to persuade a second agent of some belief which that other currently does not endorse, then the communications language should enable the second to express any changes of belief it makes as a result of the interaction; otherwise, why would the first agent seek to persuade the second? This self-transformation capability has not typically been a feature of agent communications languages or interaction protocols. In recent work [26], we proposed expression of self-transformation as a design criterion for multi-agent systems using dialogue game protocols, and assessed various interaction protocols and languages against this criterion (among others). As noted there, the FIPA Agent Communications Language FIPA ACL [12] provides only limited capability for participants to question one another and to express any changes in beliefs. Because FIPA ACL lacks retraction illocutions, changes in agent opinions can only be expressed by successive, and possibly contradictory, utterances of belief. An agent who believes the sky is blue utters an *inform* illocution to this effect; if it subsequently comes to believe the sky is red, it can only express this with a second, contradictory, *inform* illocution. How is a listener to such a sequence of contradictory utterances to interpret it? The sequence may be evidence of updated beliefs by the speaker, or it may be due to malice, whimsy, or simply faulty code. Explicit retraction locutions can ensure no such ambiguity of interpretation. Agent languages based on formal dialogue games have a better record in this regard [26], perhaps due to their origin as protocols for the conduct of debates in philosophy.

3.2 Interaction Protocols

We distinguish between the communications language used by the agents in an agent society to make utterances, and any rules which govern the combination of utterances, which, when combined with the language, we call the interaction protocol. The FIPA ACL [12], for example, has no such rules, with the result that any utterance by any agent may follow any other by any agent. Of course, such rules may defined to overlay the FIPA ACL, as with the various auction protocols defined by FIPA, e.g., [13]. As with the communications language, the requirements placed on any interaction protocol will differ according to which model of democracy is used. Because the Wise Elite model does not require any expression of opinions or of arguments, there are no requirements on the interaction protocol. The Rational Choice model only requires expression of opinions or preferences, and so any interaction protocol would need to be enable expression of these in an orderly fashion. Auction protocols, for example, typically proceed through a series of rounds, with constraints on what can be uttered at each round [13].

The Deliberative model, by contrast, leads to the most extensive requirements on any interaction protocol. If participants are able to question and challenge one another's statements, and to defend their own statements when challenged, then an orderly interaction will require rules relating one type of utterance to another, and specifying when particular utterances are required or prohibited. For instance, the rules may specify the circumstances under which a question seeking the reasons for some claim must be answered by the agent which made the claim; without such a rule, the questioner would have no guarantee that the question would receive an answer. Interaction rules such as these have received considerable attention from philosophers of argumentation, e.g., [1, 17], work which has, in turn, influenced the design of agent interaction protocols, e.g., [2, 25]

3.3 Resolution Mechanisms

Agents in a multi-agent system may have different beliefs and intentions; accordingly, agent researchers have designed mechanisms to enable agents to share their opinions and justifications, and to engage in persuasion and negotiation dialogues, e.g., [30, 37]. However, differences of opinion may persist even after exchanges of justifications and attempts at persuasion. In circumstances where a collective judgment must be made, such as where a group of agents need to agree a joint course of action, then the agents require some mechanism for resolving their differences. The mechanisms which are feasible and appropriate differ according to the model of democracy used.

Under the Wise Elite model, all decisions are taken by the elite so, as far as the other participants concerned, there is no need for a conflict-resolution mechanism.[5] In an auction with one seller and many potential buyers, for example, the auctioneer determines the winner, usually (but not necessarily) on the basis of previously-published rules; the auctioneer thus resolves the difference of opinion between the buyers unilaterally. Under the Rational Choice Model, agents choose between policies as if the agents were consumers and the policies were products. Although agents may receive information about policy options, the Rational Choice model does not assume they necessarily engage in debate or argument about these. The final resolution of any differences is made by each agent choosing whichever policy it most favors. If these choices differ, than the appropriate resolution mechanism is a vote by the agents, selecting that policy, for example, with the greatest numerical support. For instance, Hunsberger and Zancanaro propose a voting procedure for pooling agent judgements over alternative partial plans in undertaking joint planning activities [18].

[5]The Elite itself, if comprised of more than one agent, may require such a mechanism.

Under the Deliberative Model, however, it is assumed that agents may engage in debate over policy choices, and so there may be an exchange of arguments prior to determination of a collective judgment. As with the Rational Choice Model, a voting mechanism may be used to make this final determination. But the exchange of arguments means that other mechanisms are also feasible, relying on the argument-aggregation procedures from argumentation theory, e.g., [11, 22]. An example of such a procedure may make this clear. In [22], claims are classified into one of several mutually-exclusive classes on the basis of the arguments presented for or against them. In this framework, one argument B rebuts another A if B is an argument for the claim $\neg\theta$ and A is an argument for the claim θ. An argument C undercuts A if C is an argument for a claim $\neg\gamma$, where γ is a premise of argument A. We can then define the argument-status of a claim θ at time t as follows:

- If there have been no arguments uttered for or against θ up to time t, then the claim is *Open*.
- If there has been at least one argument uttered for θ up to time t, then the claim is *Supported*.
- If there has been at least one argument whose premises are consistent uttered for θ up to time t, then the claim is *Plausible*.
- If there has been at least one argument whose premises are consistent uttered for θ up to time t, and no undercutting or rebutting arguments have been uttered against θ by this time, then the claim is *Probable*.
- If there has been at least one argument whose premises are consistent uttered for θ up to time t, and any undercutting or rebutting arguments uttered against θ by this time have themselves been rebutted or undercut, then the claim is *Confirmed*.

The motivation for this approach is that the more and the stronger are the arguments for a claim, then the more support it has among the participants. The labels used, *Open, Supported*, etc, are entirely arbitrary and any set of qualitative labels could be defined in this way. Such a classification of arguments can be used as a conflict-resolution mechanism when the agents concerned are unable to agree on a claim. In [25], we explored the formal properties of such an argument-based resolution mechanism, particularly the circumstances under which the labels assigned to a claim would converge over time.

Some political theorists claim that being open to persuasion requires participants with conflicting views to see each other as adversaries rather than as enemies, engaged in argument in particular interactions in the joint knowledge that every interaction may be followed by others [19, 28]. Participants therefore need to achieve a feasible middle ground between striving for an impossible consensus and refusing to interact with one another. Accordingly (these theorists argue), democratic political institutions and processes need to be able to permit participants to express what may be very different preferences and goals, and to participate in joint political processes despite such differences. Argument-classification systems, such as the one above, facilitate this by incorporating *all* the views and arguments expressed, even those which are not in the majority.

In summary, the model of democracy adopted has implications for the types of mechanisms which are feasible for resolving conflicts of agent opinion. Such differences of opinion are ignored under a Wise Elite model. Voting is essentially the only mechanism possible under a Rational Choice model, while the exchange of arguments under a Deliberative Model enables the additional use in these systems of resolution mechanisms based on argument classification from argumentation theory.

4. SEMANTIC VERIFIABILITY

Verification of semantic requirements of agent communications languages and interaction protocols is problematic [32, 38]. This is essentially because a sufficiently-clever agent can always simulate insincerely any required mental state. In response to this problem, Munindar Singh [36] proposed a *social semantics* for agent communications languages, in which each participant to an interaction makes a public statement of its beliefs and intentions. Other agents can then use these public declarations to ensure that each agent is consistent in its utterances in an interaction. Of course, an agent may still make insincere declarations, but at least it can be called to account for inconsistencies between its declarations and its subsequent statements, as one of us has formalized in [3]. In the vernacular: *Liars need good memories*.

If a deliberative model of democracy forms the basis of an open agent system with a social semantics, then we are able to obtain a stronger form of semantic verifiability. Under a deliberative model, agents making claims may be questioned and challenged by other agents about the reasons for those claims; these reasons could be arguments for a belief of the agent, or reasons for an intention. Consequently, not only can the consistency of declarations and other utterances in the interaction be verified, but also the degree to which the declarations themselves — beliefs or intentions — are justified. We call this form of verification *contestability*, since social semantic declarations may be contested or challenged by other agents.[6] Of course, insincere declarations are still possible, but agents making false declarations may also need to fabricate a set of arguments for them. To be convincing to others, an insincere agent needs to create a set of inter-locking arguments for its statements, and other agents may only accept these arguments in defined circumstances, as is the case with the formal argumentation systems of [3, 31]. For example, if the listeners are skeptical regarding what arguments they accept [31], then they will only accept those arguments which defeat (in a precise sense) any attacking argument. Creating a set of interlocking arguments which convince a skeptical agent will usually not be easy for an insincere agent. In the Shakespearean vernacular: *"Oh what a tangled web we weave, when we first practice to deceive."*

One might view this approach as fine for beliefs and intentions which have a reasoned basis. But what of agent mental states such as preferences or values, which (some would argue) have no rational basis. The first point to make is that many more of these may have a rational basis than is commonly perceived; for example, a consumer may justify a preference for a white-colored motor car on the basis that he or she lives in a hot climate, and light-colored cars are generally cooler inside than dark-colored cars. Secondly, some philosophers argue that rational debate is possible even on matters of profound disagreement over fundamental values, e.g., [33]. Indeed, it is possible to show that some arguments may defeat all others[7] regardless of the values of the participants [5]. Thus, even when participants disagree over fundamental values, there can be non-trivial claims which are accepted by everyone. However, even if an agent's mental states resist justification, the possibility of facing contestation may still reduce the likelihood that the agent declares them falsely. Of course, there are situations where agents may wish to declare their mental states insincerely, as in a negotiation where an agent provides false misleading information about its preferences or intentions in order to gain an advantage.

Our argument in this Section can be summarized as follows: The Deliberative Model of democracy emphasizes the joint and discursive nature of decision-making in a society, with participants exchanging arguments for and against various policy proposals, and forming preferences on the basis of these exchanges. The structure provided by this model to the designer of an agent system creates the means necessary to develop a strong form of social semantics. This is because every assertion by an agent may be questioned or contested by others, thereby making insincere utterances harder to sustain. We call this feature of a dialogue system, contestability. One agent can never finally verify the mental states of another, and thus can never verify semantic compliance with an Agent Communications Language defined in terms of such states, such as the Semantic Language SL of the FIPA ACL [12]. However, the use of a social semantics increases the extent to which compliance with the semantics of the agent communications language can be verified; the use of a deliberative democracy model providing contestability of utterances increases the degree of verifiability again, above that provided by a social semantics. Auction mechanism designers in economics have traditionally dealt with this situation by aiming to design the interaction mechanism so that insincere declarations by an agent are not in that agent's best interest. The use of a social semantics and contestability can be viewed as analogous to this goal in the context of agent conversations, interactions which are generally far less structured than are auctions.

5. CONCLUSION

This paper has considered the problem of how to structure open agent societies. Open multi-agent systems are those where participation is open to any agent (possibly satisfying some conventions), and thus, in particular, to agents designed by a different design team to that responsible for the system itself. Because of this diversity, agents in an open agent system are likely to incorporate very different beliefs, goals, preferences, decision-processes and decision-constraints. In some such systems, social decision-making processes may be hierarchical and uncontested. In many systems, however, the relationship between the participating agents will be one of equality, and the question arises as to what organizing structures are appropriate in these domains. As a step towards answering this question, we have explored, for the first time in the agent literature, alternative normative models of democracy taken from political philosophy and considered their implications for open agent societies.

The three most influential abstract theories of democracy current in political theory are: the notion of democracy as wise rule by an elected elite, acting in the best interests of the populace, a theory due originally to Joseph Schumpeter [35]; the notion of democracy as the exercise of rational consumer choices by voters, choosing between alternative "products" in the form of policies or ideologies, a theory due to Anthony Downs [9]; and the notion of democracy as joint deliberation by the citizenry, deciding collectively, through argument and debate, on appropriate public policies for given situations. This deliberative theory has arisen through work by a number of authors in philosophy, political theory and judicial theory, particularly Jürgen Habermas [15, 16] and Frank Michelman [27].

We then considered these alternative normative models of democ-

[6] We also use this word because of the analogy with its meaning in economics. A contestable market is one to which new entrants may join at any time, a prospect which should lead existing suppliers to act as if competitors were already present. Thus a monopolist in a contestable market may behave as if in a competitive market.

[7] in the sense of argumentation theory [11].

racy for their relevance to the design of open agent systems. In particular, we considered their implications for the design of agent communications languages, for interaction protocols, and for any mechanisms for the resolution of conflicts. The different models place very differing requirements on the design of these aspects of an agent system, and thus allow for different ways of structuring open agent societies. Following this, we also discussed the notion of rationality in deliberative models of democracy, which we argued could provide a form of semantic verifiability of the communications language used by agents in an interaction. This form of verifiability, contestability, is stronger than Singh's social semantics, because an agent's utterances may not only be checked for consistency over the course of a dialogue, but also questioned and contested by other agents. As a consequence, insincere utterances become harder to sustain.

We believe the primary value of this paper is to raise awareness among designers of open agent systems of the availability of alternative conceptualizations of the notion of democracy, and the possibilities they provide for organizing open societies. Without such awareness, system designers are likely to encode one or other model implicitly, which may subsequently limit the functionality of the agent system. A second value of this paper is our notion of contestability, which provides a form of semantic verifiability for agent communications languages stronger than previous forms. In future work, we hope to formalize the argument we have made here regarding the relative verification effectiveness of contestable and social semantics.

Acknowledgments

We are grateful to William Rehg for discussions on some of these topics.

6. 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. Published in German 1978.

[2] L. Amgoud, N. Maudet, and S. Parsons. Modelling dialogues using argumentation. In E. Durfee, editor, *Proc. 4th Intern. Conf. on Multi-Agent Systems (ICMAS 2000)*, pages 31–38, Boston, MA, USA, 2000. IEEE Press.

[3] L. Amgoud, N. Maudet, and S. Parsons. An argumentation-based semantics for agent communications languages. In F. van Harmelen, editor, *Proc. 15th European Conf. on AI (ECAI 2002)*, Toulouse, France, 2002.

[4] K. J. Arrow. *Social Choice and Individual Values*. Wiley, New York, 1951.

[5] T. J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. Technical Report ULCS-02-017, Department of Computer Science, University of Liverpool, UK, 2002.

[6] J. Bessette. Deliberative Democracy: The majority principle in republican government. In R. A. Goldwin and W. A. Schambra, editors, *How Democratic is the Constitution?*, pages 102–116. AEI, Washington, DC, USA, 1980.

[7] J. Bohman and W. Rehg. Introduction. pages ix–xxx. In [8].

[8] J. Bohman and W. Rehg, editors. *Deliberative Democracy: Essays on Reason and Politics*. MIT Press, Cambridge, MA, USA, 1997.

[9] A. Downs. *An Economic Theory of Democracy*. Harper and Row, New York City, NY, USA, 1957.

[10] S. Ducasse, T. Hofmann, and O. Nierstrasz. OpenSpaces: an object-oriented framework for reconfigurable coordination spaces. In A. Porto and G-C. Roman, editors, *Coordination Languages and Models*, LNCS 1906, pages 1–19. Springer, Berlin, Germany, 2000.

[11] 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.

[12] FIPA. Communicative Act Library Specification. Technical Report XC00037H, Foundation for Intelligent Physical Agents, 10 August 2001.

[13] FIPA. English Auction Interaction Protocol Specification. Technical Report XC00031F, FIPA, 10 August 2001.

[14] J. Forester. *The Deliberative Practitioner: Encouraging Participatory Planning Processes*. MIT Press, Cambridge, MA, USA, 1999.

[15] J. Habermas. *The Theory of Communicative Action: Volume 1: Reason and the Rationalization of Society*. Heinemann, London, UK, 1984. Published in German 1981.

[16] J. Habermas. *The Inclusion of the Other: Studies in Political Theory*. MIT Press, Cambridge, MA, USA, 1998. Edited by C. Cronin and P. De Greiff.

[17] D. Hitchcock. Some principles of rational mutual inquiry. In F. van Eemeren et al., editors, *Proc. 2nd Intern. Conf. on Argumentation*, pages 236–243, Amsterdam, The Netherlands, 1991. SICSAT.

[18] L. Hunsberger and M. Zancanaro. A mechanism for group decision making in collaborative activity. In *Proc. 17th National Conf. AI (AAAI-2000)*, pages 30–35, Menlo Park, CA, USA, 2000. AAAI Press.

[19] R. L. Ivie. Rhetorical deliberation and democratic politics in the here and now. *Rhetoric and Public Affairs*, 5(2), Summer 2002.

[20] C. C. Jaeger, O. Renn, E. A. Rosa, and T. Webler. *Risk, Uncertainty, and Rational Action*. Earthscan Publications, London, UK, 2001.

[21] R. Johnson. *Manifest Rationality: A Pragmatic Theory of Argument*. LEA, Mahwah, NJ, USA, 2000.

[22] P. Krause, S. Ambler, M. Elvang-Gørannson, and J. Fox. A logic of argumentation for reasoning under uncertainty. *Computational Intelligence*, 11 (1):113–131, 1995.

[23] O. Lange. The scope and method of economics. *The Review of Economic Studies*, 13(1):19–32, 1945–1946.

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

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

[26] P. McBurney, S. Parsons, and M. Wooldridge. Desiderata for agent argumentation protocols. In C. Castelfranchi and W. L. Johnson, editors, *Proc. 1st Intern. Joint Conf. Autonomous Agents and MAS (AAMAS 2002), Bologna, Italy*, pages 402–409, New York, USA, 2002. ACM Press.

[27] F. I. Michelman. Conceptions of democracy in American Constitutional argument: the case of pornography regulation. *Tennessee Law Review*, 56:291–319, 1989.

[28] C. Mouffe. *The Return of the Political*. Verso, London, UK, 1993.

[29] P. Noriega and C. Sierra. Towards layered dialogical agents.

In J. P. Müller, M. J. Wooldridge, and N. R. Jennings, editors, *Intelligent Agents III*, LNAI 1193, pages 173–188, Berlin, Germany, 1997. Springer.

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

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

[32] J. Pitt and A. Mamdani. Some remarks on the semantics of FIPA's Agent Communications Language. *J. Autonomous Agents and Multi-Agent Systems*, 2(333–356), 1999.

[33] H. S. Richardson. *Practical Reasoning about Final Ends*. Cambridge University Press, Cambridge, UK, 1994.

[34] J. J. Rousseau. *Discourse on Political Economy and The Social Contract*. Oxford University Press, Oxford, UK, 1994. Published 1762. Translated by Christopher Betts.

[35] J. Schumpeter. *Capitalism, Socialism, and Democracy*. Harper, New York, USA, third edition, 1950. First edition 1942.

[36] M. P. Singh. A social semantics for agent communications languages. In F. Dignum, B. Chaib-draa, and H. Weigand, editors, *Proc. Workshop on Agent Communication Languages, IJCAI-99*, pages 75–88, 1999.

[37] K. Sycara. Persuasive argumentation in negotiation. *Theory and Decision*, 28:203–242, 1990.

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