Computer Decision-Support Systems for Public Argumentation: Assessing Deliberative Legitimacy

William Rehg
Department of Philosophy
St. Louis University
St. Louis, MO 63156-0907 USA.
Email: rehgsp@slu.edu

Peter McBurney (Communicating author)
Department of Computer Science
University of Liverpool
Liverpool L69 3BX, United Kingdom.
Fax: + 44 151 794 3715
Email: p.j.mcburney@csc.liv.ac.uk

and

Simon Parsons
Department of Computer and Information Science
Brooklyn College,
City University of New York
Brooklyn NY 11210 USA.
Email: s.d.parsons@csc.liv.ac.uk

Abstract

Recent proposals for computer-assisted argumentation have drawn on dialectical models of argumentation. When used to assist public policy planning, such systems also raise questions of political legitimacy. Drawing on deliberative democratic theory, we elaborate normative criteria for deliberative legitimacy and illustrate their use for assessing two argumentation systems. Full assessment of such systems requires experiments in which system designers draw on expertise from the social sciences and enter into the policy deliberation itself at the level of participants.

Keywords: Argumentation, Computer-Mediated Decisions, Decision Support Systems.

1. Introduction

Human reliance on computer technology is by now a well-established cultural fact. Indeed, we now rely on computers even for certain forms of decision-making: so-called "knowledge-based expert systems," or "decision support systems," have been developed over the last three decades to support, or sometimes even replace human decision-making (Alty and Coombs 1984; Buchanan and Shortliffe 1984). As the name suggests, expert systems attempt to automate (by means of a knowledge base and inference mechanisms) the knowledge and reasoning skills of experts in a given domain, such as medical diagnosis, marketing decisions, and so on.[1] In this paper we are concerned with recent attempts to develop decision-support systems for processes of public policy argumentation. Like earlier expert systems, these argumentation-support systems incorporate both knowledge bases and inference mechanisms. Unlike the earlier systems, however, they place greater emphasis on the processes used for reasoning and inference than on the database of knowledge from which conclusions are drawn; thus we might designate them as "argumentation systems" in contrast to the earlier "knowledge systems."[2] Dialectical approaches in particular have drawn attention in this developing area of Artificial Intelligence research.[3] Indeed, some commentators have recommended the employment of such systems, via the Internet, to enable democratic participation in public policy decision-making processes (e.g. Ess 1996). The kind of argumentation systems we examine here are noteworthy for their attempt to draw explicitly upon dialectical theories of argumentation as a framework for public deliberation involving multiple parties and interests.

The use of argumentation systems to assist public deliberation raises a number of conceptual and social-ethical questions—including the question of their political legitimacy.[4] On

the one hand, argumentation systems provide a deliberative forum whose results can advise and, one hopes, improve the qulaity of decisions. If appropriately designed, such systems should be able to assist debate by tracking the various claims and arguments, by searching databases for relevant information, and by continually updating and assessing the overall state of the debate. They thereby help participants argue in a dialectically responsible manner, while offering them ample scope for modifying aspects of the system (e.g., inference rules, proof standards). On the other hand, the value-laden, controversial character of public policy argumentation—and as we shall see, the ideal nature of dialectical rules—poses interesting challenges for assessing "appropriate design." In this paper we draw on deliberative democratic theories of legitimacy to clarify the criteria for such assessment. To focus and concretize the discussion, we start by describing two current proposals for computer-mediated argumentation and decision-making, the Zeno system of Gordon, Karacapilidis, and their colleagues at the GMD-German National Research Centre for Information Technology and the Risk Agora of McBurney and Parsons (sec. 2). We then sketch a deliberative model of legitimacy whose basic evaluative dimensions are appropriate for the policymaking contexts in which such systems are to be used (sec. 3). This model implies four deliberative criteria, and to illustrate their application we use them to assess Zeno and the Risk Agora (secs. 4-7). Analysis of the gaps between a system's formal design and actual use in context shows that the assessment of argumentation systems requires AI researchers to engage in interdisciplinary experimentation not only with argumentation theorists, but also with scholars working in areas such as poltical science, democratic theory and policy studies, social psychology and sociology. Not only that, the researcher must engage in dialogue with the citizen-participants themselves over the dialectical substance of policy deliberation (secs. 8, 9).

2. Two Argumentation-Based Decision Support Systems

(2.1) We begin by describing the two argumentation systems. The first is the Zeno system of the GMD National Research Centre for Information Technology (Gordon and Karacapilidis 1997, Gordon *et al.* 1997, Karacapilidis *et al.* 1997). Zeno was developed to support decision-making in urban planning, as part of a larger European Community-funded project to develop innovative information systems infrastructure for public collaborative environmental planning. In these domains, there are multiple interested parties, with diverse professional or private backgrounds, interests, preferences and viewpoints, and they are often geographically dispersed. Because the application domain of Zeno involves urban planning decisions, the system has to integrate information which is spatially indexed with information which is not. And because the users may be diverse and geographically distributed, the system requires intuitive and easy-to-use interfaces, provided, preferably, across an Internet platform. Neither of these elements of software design was technically straightforward, but these issues will not concern us here. Our focus is on the argumentation and decision support elements of Zeno.

The developers of Zeno define their system as "a mediation system":

> a kind of computer-based discussion forum with particular support for argumentation. In addition to the generic functions for viewing, browsing and responding to messages, a mediation system uses a formal model of argumentation to facilitate retrieval, to show and manage dependencies between arguments, to provide heuristic information focusing the discussion on solutions which appear most promising, and to assist human mediators in providing advice about the rights and obligations of the participants in formally

regulated decision making procedures. (Gordon and Karacapilidis 1997, 10).

The argumentation model used by Zeno is a formal version of the informal Issue-Based Information System (IBIS) model of Rittel and Webber (1973), modified for the urban-planning domain. The IBIS model identifies several atomic elements of a discourse: *Issues*, the topic about which a discussion is conducted; *Positions*, which express some statement relevant to an issue; and *Arguments*, which present statements in favour or against particular *Positions*. Thus (using an example from Karacapilidis *et al.* 1997), an Issue may be: "Which site should the airport be located?"; Positions may then be statements designating different alternative sites, or groups of possible sites; and Arguments may be positive and/or negative attributes about these site alternatives and statements, such as: "Has easy access" or "Is private land." Each element of this model can be attached to each other element at any time, so that, for example, a new issue can be raised at any point in a discussion, thus creating a subsidiary discussion on the new issue.

The IBIS model is well-suited for the display of a discussion as a hierarchical graph; the implementation of hypertext links enables a user to move easily around this graph from one thread of a discussion to another, or to access background data, supporting documents or contextual information, etc, associated with any element of the discussion. Among the objectives of the Zeno project is to enable snapshots of a debate: "One important goal is to provide easy access to the current state of the planning process, at any time." (Gordon et al. 1997, Section 2). In order to provide for these, the designers of Zeno modified the IBIS model to permit the expression of preferences. Positions have the form of logical propositions but they do not have a context-independent truth status. Their meaning is defined by their role in a particular thread of discussion. Preferences are defined as particular types of positions with an

internal structure of comparison of two (non-preference) positions. For example, two planning options for siting an airport, such as "public site" and "easy access" might enter into a preference that considers easy access as "more important than" a public site. This preference then constitutes a qualitative constraint, which may or may not be supported by further arguments (i.e., positions), and which may or may not be consistent with other constraints. These preferences and constraints, being positions, may themselves be the subject of discussion, via the articulation of arguments and the raising of further issues.

Zeno provides users with an overview of the argumentative status of positions, preferences, and constraints. By considering the extent to which each position satisfies the articulated constraints, Zeno permits positions to be labeled as acceptable or not at any time in a discussion. These position-labels can then be aggregated in various ways to assign labels to Issues, and Zeno does this to indicate the extent to which their current argument support meets defined standards of proof. The Zeno developers argue that no set of proof-standards is applicable across all application domains, and so they adopted a set of five labels from the field of jurisprudence, namely: (1) *Scintilla of Evidence*; (2) *Preponderance of Evidence*; (3) *No Better Alternative*; (4) *Best Choice*; and (5) *Beyond a Reasonable Doubt* (Gordon and Karacapilidis 1997). Each label is provided with a formal definition in terms of the presence or absence of positions and arguments (although these definitions are not claimed to be instantiations of the legal definitions of these terms). With this argumentation structure, it is then straightforward for the system to present to a user the current status of a discussion and to show this as it changes.

The designers of Zeno identify the generic speech acts involved in contributing to a discussion, for example, "Raise an issue", "Assert a position", "State a preference", etc (Gordon

and Karacapilidis 1997). However, they do not, in the work published to date, articulate a definitive list of such speech acts or the rules that govern their use. By contrast, our second example of an argumentation system, the Risk Agora of McBurney and Parsons (2000, 2001a, 2001b) is fully specified in this manner, and we now discuss this system.

(2.2) The Risk Agora has been proposed as a system to support deliberations over the potential health and environmental risks of new chemicals and substances, and the appropriate regulation of these substances. Determination of these issues typically first involves debate within the relevant scientific community over whether or not a significant correlation exists between the putative chemical cause and any observed health effects, and then, if a significant relationship is identified, a subsequent debate in the wider community over the consequences of alternative regulatory options then occurs. Initial development of the system (McBurney and Parsons 2000, 2001b) has focused on the first of these debates, the scientific debate, where the designers adopted explicit philosophies of science and of rational discourse. The philosophy of science they draw on is Pera's model of science as a three-party interaction, where progress is made through the work of, firstly, scientists undertaking experiments, whose experiments provoke reactions from, secondly, Nature, whose responses are in turn mediated through, thirdly, the scientific community (Pera 1994). Because the members of the scientific community are (presumed to be) rational and willing participants in the process, and because the assertions in these scientific discourses are all subject to contestation and defence, the authors adopt a philosophy of rational discourse for such debates. For this they draw on Habermas's theory of discourse ethics (Habermas 1983/1990), whose rules were first fully articulated by Alexy (1978/1990), and the principles of rational mutual inquiry of Hitchcock (1991).

Within this framework, the authors then articulate the locutions and rules of a dialogue-game, in the style of Hamblin (1971) and MacKenzie (1979), specifying the pre-conditions necessary for the execution of each locution and the changes each locution effects. The locutions of this game permit assertion, contestation and defence of propositions, modes of inference, assumptions and consequences, in what is essentially a persuasion dialogue (Walton and Krabbe 1995). Certain locutions incur requirements on the speaker to defend a statement if subsequently contested, and the formalism permits this to be in the form of an argument for the statement. Statements, assumptions, consequences and modes of inference may all be asserted with an attached modality, expressing the speaker's degree of confidence in the assertion. As in the Hamblin games, "Commitment stores" track assertions made in the course of the debate. McBurney and Parsons have also begun the task of the detailed specification of a system to support discussion over regulatory options (McBurney and Parsons 2001a). In this work, they have drawn on Habermas's theory of communicative action (Habermas 1981/1984) to define types of locutions (i.e., speech acts) appropriate for such discussions. Because these discussions are about actions, their formalization as dialogue games requires models of deliberation dialogues (Walton and Krabbe 1995, Hitchcock, McBurney and Parsons 2001) and means by which dialogues may be combined (McBurney and Parsons 2002b). Thus, the work of specifying the Agora to support regulatory debates is on-going.

Unlike the Zeno system, the Risk Agora is not intended to support debates in real-time, and the designers have also not yet developed the intuitive, graphical interfaces present in Zeno. Rather, the Agora is intended to formally model and represent debates in the risk domain, so as:

1. To understand the logical implications of the scientific knowledge relating to the particular issue, and the arguments concerning the consequences and value-assignments

of alternative regulatory options.

2. To consider the various arguments for and against a particular claim (including regulatory options), how these arguments relate to each other, their respective degrees of certainty, and their relative strengths and weaknesses.

3. To develop an overall case for a claim, combining all the arguments for it and against it.

4. To enable interested members of the public to gain an overview of the debate on an issue.

5. To support group deliberation on the issue, for example in Citizens' Panels.

6. To support risk assessment and regulatory determination by government regulatory agencies. (McBurney and Parsons 2001a).

Although the Risk Agora has different aims from Zeno, its designers also desire to enable snap-shots of a debate to be taken at any time. This requirement has added importance in the risk domain, where regulatory decisions must be made even though a final determination of the state of scientific knowledge on a particular issue is not possible. For this reason, and like Zeno, the Agora defines a set of labels which are attached to each claim on the basis of the arguments presented for and against the claim up to that time. *Probable* claims, for example, are those for which an argument has been presented in the Agora, but for which no rebutting arguments (arguments for the negation of the claim) or undercutting arguments (arguments for negations of an assumption or an intermediate premise) have been presented. In this way, the dialectical status of a claim can be assessed at any time, thus providing a snapshot of the debate at that time. The designers of the Agora then examine the likelihood that a snapshot, taken at some finite time

after commencement of a debate, is indicative of the longer-term state of the debate, assuming such a stable state is achieved. They show that the Agora has desirable properties when used for inference from finite snapshots to longer-term states in this way (McBurney and Parsons 2001b).

(2.3) Before moving on, notice that each system supports argumentation in a number of ways, each corresponding to a different role in deliberation. We single out three main roles. First, then, argumentation systems may *support participants.* In fact, the Zeno system provides support in a number of ways (see Gordon and Karacapilidis 1997, 10). As a mediation system, Zeno supports a human moderator or mediator (where this person is not one of the decision-makers), for example by identifying common assumptions across different arguments. Such mediators play an important role in public policy deliberations. Urban planners, for instance, often assist community groups to reach a consensus in this manner (Forester 1999), and Märker *et al.* (2002) emphasize the necessity of a human "moderator" for GeoMed, an urban planning mediation system based on Zeno. However, Zeno also supports all the other human participants, presumably in a neutral fashion, by tracking the deliberation as it unfolds—the changing commitments, open questions, and so on. In addition, it can provide a kind of partisan support insofar as it assists participants in their construction of arguments and counterarguments. Systems that provide this type of support have been called argument-assistance systems in artificial intelligence (Verheij 1999). Although the Risk Agora has a somewhat different aim, by reconstructing arguments it not only helps all participants track a debate but can also support individual users as they construct their own positions and arguments. And when the Agora system is fully specified, and thus able to support deliberation over the consequences of regulatory options, it may also support those tasked with facilitating such decisions.

Second, the system may play a record-keeping role, or what we call an *orrery role,* on the analogy of mechanical models of the solar system. Record keeping allows users to understand the reasoning used in the decision-process. Each of the two systems records the reasoning used to reach a decision, Zeno in real-time and with complete accuracy, the Agora in a formal reconstruction of the process. Indeed, Agora's main role is the orrery one.[5]

Third, an argumentation system may support the entire process of decision-making by providing *a forum in which to undertake dialogue,* with defined protocols for this discourse. In this role the system provides something like a structured space in which participants interact. The forum-creating role is central to Zeno's design. In contrast, although the Risk Agora is designed as a forum for discussion, this is only in an ideal sense, as it provides a forum for the reconstruction of arguments rather than for real-time support.

Note that neither system plays the role of participant or decision-maker. There are, to be sure, argumentation systems envisioned for this more active role, for example the StAR system designed for the automated prediction of chemical properties such as toxicity.[6] Such systems may replace a human decision-maker, a prospect that has raised concern on the part of philosophers and computer scientists.[7] When the system plays the role of participant, it needs to be able to generate, evaluate, contest and defend arguments itself. This is true even for the mediation role, as a mediator may need to find common ground between different positions; for example, he or she may need to argue that two opposing positions share common assumptions, or that one implies another. How well a given system executes these tasks will be important in any evaluation. Neither Zeno nor the Agora appear to have the capability to generate, evaluate, contest or defend arguments. For the Agora, acting as an orrery, these capabilities are not required. For Zeno, because the designers do not seek to automate the role of mediator, the

system can also operate without these capabilities (Gordon and Karacapilidis 1997, 11). Indeed, from a computer science perspective, the generation of appropriate arguments in a dialogue, even using only the limited sets of locutions of Zeno or the Agora, is a challenging research problem, and one that is not yet solved.

Rather, both Zeno and the Agora support users in meeting their responsibility to explain decisions reached through use of the system only insofar as the participants justify their statements to each forum. The proof standards in Zeno, as well as the game-like structure of the system itself, should lead users to meet this responsibility. The dialogue rules in the Agora require those participants asserting claims to provide arguments justifying these claims when questioned or contested.

3. Deliberative Legitimacy and Argumentation Systems: Dimensions for Evaluation

Along what dimensions might one analyze and evaluate such systems? For software systems designed for well-defined, decomposable and measurable tasks (e.g., for the production of bills for use of an electricity network), assessment of system competence and quality is straightforward; standardized methods have been developed, are in widespread use, and influence good software design (e.g. Kirwan and Ainsworth 1993). According to Parker (2000), these methods are inapplicable for most decision support systems, since most decisions and decision-processes are not amenable to such reductionist task analysis. Whether one accepts that verdict or not, methods for evaluating decision support systems remain in the early stages of development.[8] Groothuis and Svensson, in research presented as recently as December 2000 and investigating the quality of computer-supported welfare assistance decisions in the Netherlands, say: “To our knowledge this is the first investigation into the extensive use of expert systems in

the daily practice of handling a very complex administrative task" (Groothuis and Svensson 2000, 9).

It is true that designers of what we have termed knowledge systems — expert systems that encode some body of expertise — will typically compare performance of the system against a group of human experts through a number of test cases. However, such comparisons are fraught with difficulties. In some domains, what counts as "expertise" may be open to dispute, or subject to cultural and contextual variation.[9] In other domains, normative decision theory (e.g., Lindley 1985) does not reflect the decision methods people actually use. Although one may simply take this as an illustration of deficiency of human decision-makers, in some cases one may not have any means of assessing the normative methods as superior.[10] Moreover, normative decision theories have tended to focus only on those elements that are quantifiable, and so may ignore much that is salient to good decision-making (cf. Rehg 1997a). Assessing the quality of advice-giving systems poses further specific difficulties. If advice is not taken, is it necessarily of low quality or unhelpful to the decision-maker? And how does one assess the quality of advice if the world changes in a salient way between the giving of the advice and its execution?[11] Still further difficulties are raised by the fact that advice is often linked with certain contextual limitations or caveats, which users sometimes ignore.[12] How does one assess the quality of advice in such cases? Finally, how does one assess advice for extreme situations or rare events?[13]

Beneath the foregoing methods of assessment and their difficulties lurks a shared assumption, namely that there is a correct or right or "true" decision to be reached by following a set of context-independent (albeit domain-specific) inference rules. However, if assessments based on this straightforward model run into difficulties for knowledge systems, we should

expect even greater difficulties when we attempt to evaluate argumentation systems for public deliberative domains that involve multiple interested parties, controversial value-laden issues, and complex legal-political contexts. In fact, experiments and pilot projects with Zeno-based systems in real-world settings have had to confront a number of technical and legal difficulties (Gordon 2003; Märker *et al.* 2002; Schmidt-Belz *et al.* 1998; Schneider, n.d.). The precise problem that interests us here, however, is the lack of an inherent, or independently accessible standared of truth or correctness for urban planning decisions. Any viewpoint from which we may judge such decisions will invariably be partial and never completely disinterested; only the process used to reach the decision can tell us whether or not it deserves to stand (Forester 1999; Bohman 1996).

Although the intended applications are different, both Zeno and the Risk Agora seek to support deliberative decision-making by multiple participants in a public policy domain. Given the difficulties we have just sketched, any evaluation of argumentation systems such as Zeno or the Agora must begin with a more sophisticated understanding of the relevant kind of decision-making at issue. Specifically, assessment criteria must draw not only upon argumentation theory but also on an understanding of the process and social-organizational context of public policy formation—as the designers of GeoMed, a Zeno-based system for urban planning, have noticed. Schmidt-Belz, Rinner, and Gordon (1998, 86), for example, conclude that such mediation systems require the development of "model procedures and new standards." Among the process considerations bearing on the quality of outcomes they list efficiency, transparency, noncoerciveness, and equality of participation.

In deliberative democratic theory, process considerations such as those above provide the basic dimensions for a theory of legitimate decision-making. By drawing on this strand of

political theory, we hope to clarify the "new standards" that the GeoMed experimenters seek; in clarifying the standards, we also articulate the theory of "deliberative legitimacy" implicit in the use of argumentation systems for policy planning. Before sketching the standards, some background remarks on legitimacy are in order.

Legitimacy refers to the "rightfulness" or validity of an authority, system of governance, or legal-political outcome (law, judicial decision, policy, etc.). Beyond this broad notion, however, legitimacy admits of diverse theoretical interpretations as well as sociohistorical variation (Beetham 1998; Bohman and Rehg 1997, introduction; Weber 1978, 1:31-38). As an empirical sociological concept, the legitimacy of an outcome might be measured by the observable degree of its acceptability to a public. Robust social-psychological findings show that public acceptability involves normative attitudes on the part of citizens, who judge legitimacy partly on the basis of their perceptions of the interpersonal quality and fairness of the procedure that generated the outcome; these judgments then affect level of compliance (Lind and Tyler 1988; Tyler 1990). Political philosophers, on the other hand, propose normative *theories* of legitimacy, ideal models that set standards citizens ought to adopt when they evaluate outcomes as legitimate or not. Theorists like Rawls (1996) and Habermas (1992/1996) attempt to maintain a link between citizen attitudes and theoretical standards by argung that the latter simply articulate what citizens already accept, at least in their better moments or in an inchoate manner. Their approach allows one to challenge citizen attitudes in particular cases by appealing to ideals that citizens supposedly do not regard as completely alien.

To explain the deliberative model of legitimacy, it helps to distinguish three levels or degrees of legitimacy (Rehg 1997b). First, many political theorists (and most citizens, we assume) accept a minimal condition that legitimate outcomes must issue from authoritative

enactments or properly executed institutional procedures, as defined by the relevant legal framework (e,g., a constitution). Deliberative legitimacy insists on an additional requirement: that procedurally valid outcomes also be substantively reasonable. Unlike elitists (e.g., Schumpeter 1942/1976), however, deliberative theorists hold that technical expertise does not always suffice to meet the reasonableness condition: to be reasonable, an outcome must issue from a dialogue or argumentation process that includes all the affected persons or stakeholders. Even when an issue can be defined in primarily technical terms, scientific knowledge may require lay knowledge for its proper application in context; evaluative aspects only heighten the need for lay participation (cf. Wynne 1989; Mayo and Hollander 1991; Morgan 1993; Collins and Evans 2002). Finally, some deliberative models go still further and link legitimacy with the supposition that consensus could be achieved under ideal conditions of discourse (Habermas 1992/1996, 110; Cohen 1989). The model we employ does not insist on this final level, however: although deliberation aims at consensus, this goal often cannot be expected (Warnke 1999; McCarthy 1999). In such cases, deliberative legitimacy is achieved insofar as the participants learn from one another and can accept the outcome as one they can live with for the sake of ongoing social cooperation (cf. Bohman 1996).

To specify our model of deliberative legitimacy further, we explain how deliberative processes are both transformatory and reasonable. In the remainder of this section we sketch the four major criterial considerations or dimensions of reasonable deliberative transformation. In subsequent sections (4-7) we further illustrate these dimensions by applying them to Zeno and the Agora. Our analysis takes as its point of departure Forester's (1999, chap. 6) application of the deliberative model to mediated public policy formation and dispute resolution.

(3.1) *Self-transformation:* A deliberative model of policy formation contrasts with the conventional pluralist model that conceived public policy questions as matters for negotiation (understood as bargaining).[14] Deliberative approaches deny that all public issues and choices reduce to bargaining. Although decisions over the division of scarce resources, or involving the conflict of particular (non-generalizable) interests, are typically resolved by resort to bargaining, decisions over what actions to take in some circumstance require deliberations (cf. Walton and Krabbe 1995). According to Forester (1999), mediators should view public planning and policy formation as deliberations. This approach has important implications for what one expects of both the process and its possible outcomes.

Although the distinction between these two modes of conflict resolution is sharper in theory than in actual practice, each individual in a bargaining process acts as a utility maximer who seeks the personally most favorable outcome that others will allow, given their aims to maximize their respective outcomes; the ensuing struggle of wills issues in a compromise position that balances competing preferences. Deliberative political processes, by contrast, require participants to adopt a civic standpoint oriented toward the transformation of individual preferences and interests in the direction of reaching agreement on a common good or general interest (e.g., Elster 1986; Michelman 1988). Forester (1999, 184) calls this "the self-transformative condition." The transformation involves the appeal to, and recognition of, shared reasons and values that go beyond the narrow self-regarding aims of individuals who bargain. As Michelman puts it:

> Deliberation . . . refers to a certain attitude toward social cooperation, 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. (Michelman 1989, 293)

The reference to pooling of judgments implies that participants in deliberations are open to learning from each other and even from the very fact of interacting. Because of this openness to change, participants to a deliberation – assuming they are rational – should be willing to share information, a strategy that may not be in the self-interests of participants to a negotiation. To the extent that participants are unwilling to share their knowledge and preferences, we may consider the decision process to be a negotiation rather than a deliberation.[15]

However, genuine learning requires self-transformation to be reasonable: if participants change their views in response to others' input, then these changes should lead toward a substantively better outcome. This observation leads to the second dimension of deliberation, namely its substantive dialectical quality (3.2). However, as we shall see, substantive quality cannot be assessed independently of considerations bearing on inclusion (3.3), which in turn point to difficult questions concerning the process itself and its "vulnerability to insinuations of power" (Forester 1999, 177) (3.4).

(3.2) *Substantive dialectical quality:* For any particular dispute resolution or policy question, one can identify specific considerations tied to the particular issue or dispute at stake, and an adequate deliberative outcome must, presumably, take all these relevant considerations—or at the least, those which are most pressing or salient—into account. An environmental dispute, for example, will typically turn on particular scientific facts about the ecosystem in question, fallible prognostications about the impact of different alternative actions on that ecosystem, and economic assessments about the costs and benefits of different options for

different affected parties. In addition, one can also expect that the parties will disagree over quality-of-life issues, fundamental goods and values (e.g., aesthetic beauty vs. economic growth), and even, at a deeper level, over basic ontologies and worldviews (Kriesberg *et al.* 1989; Rehg 1999). How well a deliberation addresses the range of relevant considerations—the pertinent questions and objections—determines the adequacy of its outcome.

The idea of relevance obviously plays a crucial role in this context—but what does it involve? Although the conception of argumentative relevance remains an open research question (Johnson 2000, 199-204; cf. Hitchcock 1992; Tindale 1999), the factual dimensions of policy argument suggest, on the one hand, a process-independent conception: a consideration is relevant if failure to take it into account in a deliberation can lead to an unsuccessful policy. On the other hand, relevance is also partly internal to the deliberative process: at least some kinds of considerations, such as interests and values, must be relevant *for the participants* if they are to be relevant at all.

This twofold conception of relevance points toward a more precise characterization of the substantive quality of an outcome or decision. Notice, to begin with, that substantive quality should probably not be reduced to "truth." Rather, given the complexity of the various relevant considerations, one does better to characterize substantive adequacy in terms of the quality and scope of dialectical testing of arguments and subsequent success of a policy choice. The standard of success stems from the external aspect of relevance. Because policy decisions partly rest on factual assumptions, policies can, in some sense, fail to be "correct," given the way the world actually is. An environmental dispute resolution may rest on false scientific assumptions, or it may depend on an economic forecast that proves mistaken. To this extent, the renewed interest of argumentation theorists in truth is not entirely misplaced (Goldman 1999, chapter 5;

Johnson 2000). But in policy-making contexts, factual claims intertwine with other sorts of evaluative claims, which are more readily assessed as "justified" (cogent) versus "arbitrary," rather than "true" versus "false." Consequently, what matters is that, in taking all the relevant considerations into account, participants construct and reconstruct their viewpoints in a manner that is publicly justified: not only logically consistent but, more importantly, *dialectically responsible, or "responsive"* (Goldman 1994). To put this idea in a nutshell, participants hold a dialectically responsible position insofar as they have addressed all the relevant questions and objections and thereby reached the most justifiable, and thus reasonable, outcome relative to the alternatives.

One may, to be sure, link dialectically responsible positions with a broadened notion of truth, for example along the lines of Rescher's dialectical account of plausible reasoning or Hintikka's game-theoretic semantics (Rescher 1976; 1977; Hintikka, 1968). According to such approaches, claims that hold up in a process of dialectical reasoning as more plausible, or that have the support of a winning strategy for the given argumentation system, enjoy a (defeasible) presumption of truth, relative to the appropriate burden of proof. Arguably, this broader concept of "truth" is applicable to many, if not most, decisions in the public policy domain.

Notice that the internal aspect of relevance (the relevant interests and values) is closely linked with the public acceptability of outcomes. For unless the perspectives and concerns of the affected parties are adequately addressed, those parties are not likely to find policy outcomes acceptable. Inasmuch as deficits in legitimacy tend to have a negative effect on compliance, the dialectical quality of a deliberative outcome is also crucial for its successful implementation. Thus, the substantive quality of the deliberation—its adequacy in taking all the relevant considerations into account—naturally leads into the question of participant inclusiveness, how

well a process has given the various affected parties the chance to voice their concerns and affect the outcome. This takes us into the third and fourth dimensions of evaluation.

(3.3) *Inclusion*: Who is admitted to the deliberative process? According to many deliberative models, all those affected by the issue—the stakeholders—should have access to, and voice in, the process. The size of this group will affect the choice of argumentation system, decision-making procedures, and perhaps division of roles. As Forester's case studies (1999) makes clear, public policy deliberations may involve numerous parties with quite different perspectives and interests. Including all the affected parties and giving them voice in the deliberation is crucial to the legitimacy and consequent success of such deliberations. But evaluating argumentation systems for overt or "formal" exclusions—exclusions that are explicit in the distribution of roles and entitlements—only goes part way. More important are those subtle forms of exclusion and coercion that might be built into the design itself as a system of rules. This issue brings us to the role of the system in deliberation, in particular the power of the system in fostering or impeding the expression and reasonable processing of information and viewpoints—an issue we take up under the fourth evaluative dimension.[16]

(3.4) *Noncoercivenss*: On a deliberative model, legitimacy-enhancing processes must not only grant all the affected parties entrance but also give them effective opportunities to voice their opinions and, still further, foster the participants' capacity for learning from one another. This evaluative dimension focuses on the *freedom* participants have to express and discuss their views so that the trend of opinion is both reasonable and based on mutual understanding. One thus asks whether a deliberative design, *precisely as an interaction,* harbors subtle forms of

coercion that could obstruct participants' effective self-expression and reciprocal learning—whether the coercion issues from the background social pressures that participants bring with them to the process, user-interface difficulties, or personal psychological factors. Such obstacles represent a kind of exclusion or inequality that occurs *within* a process whose formal design appears overtly inclusive.

To uncover sources of coercion one must attend to the kinds of power argumenation systems exert on the public deliberation they assist. If such systems in some sense "mediate" or at least facilitate deliberation, then they wield considerable power over the process and its outcome. As Forester (1999, 180) points out: "If parties to public dispute-resolution processes not only construct agreements but reconstruct themselves—in part as a result of being exposed to new information, in part as a result of the constellation of participants—then the *political significance and power* of the mediator-facilitator's role is more important to understand than ever before." This statement also applies to the argumentation system itself.

4. Deliberative Self-Transformation

Consider now some brief illustrations of how one might employ these dimensions to assess the two argumentation systems described earlier (sec. 2). Our intention in these remaining sections is not to provide an exhaustive assessment, but simply to sketch some of the more obvious moves, as an illustration of the basic idea. In fairness, we should also point out the experimental nature of these systems; our aim is not so much to criticize the system designs as to further clarify the use of the evaluative dimensions for deliberative quality and legitimacy.

The first dimension of evaluation concerned the self-transformative character of deliberative processes. Here we note briefly that, at least on the surface, or in terms of their *formal design*

features, both argumentation systems are designed to allow for deliberative self-transformation and not for bargaining or negotiation.[17] Because the Risk Agora is not intended to support real-time debates, its "participants" may in fact be representatives of positions, rather than real persons. However, the system does allow retraction of claims previously asserted or accepted, and thus permits self-transformation. For Zeno, in the paper that describes the model of decision-making used in the domain, revision of constraints is explicitly permitted so as to eliminate any inconsistencies in these (Gordon *et al.* 1997). Although the Zeno Argumentation System does not explicitly support retraction, the capacity to enter new positions in response to arguments models a kind of self-transformation (Gordon and Karacapilidis 1997).

5. Substantive Dialectical Quality

The analysis of substantive quality developed in Section 3.2 suggests that, to begin with, one assess argumentation systems for public deliberation by their capacity to include all the relevant considerations, that is, the relevant factual considerations, empirical prognostications, and values. On the one hand, the external aspect of relevance—empirical, process-independent considerations—suggests that the system should be able to incorporate all the factual information that would be relevant to a successful policy, insofar as success depends on the truth or accuracy of certain assumptions about the world and the persons affected by the policy. Concretely, this means that the system should be able to draw upon the relevant expertise and make this available to participants. On the other hand, as we have seen such expertise is not the only source of relevant information, nor is it usually decisive for a *legitimate* outcome. Because relevance also has an internal (i.e., participant-relative) sense, substantive quality requires that the value-orientations, interests, and other considerations that the participants themselves consider relevant

should be taken into account by the argumentation system, and thus opened up to public discussion. Again, the formal design features of both systems, Zeno and the Risk Agora, seem to allow for such substantive inclusiveness. Zeno does not seem to restrict the kind of information that can count as a "position," which can include both facts and preferences, and thus it can incorporate, at least indirectly, value-orientations. To be sure, Zeno may prove to be limited in its ability to distinguish between mere preferences and deeper values—a limitation that may be a problem in some contexts. Here the Risk Agora goes further, supporting the expression of various types of speech acts: factual claims, claims about what is right, or valuable, and so on (McBurney and Parsons 2001a).

Secondly, substantive quality also requires that the relevant information and arguments be processed in a dialectically responsible manner. For systems engaged in supporting a human participant or participants to construct, evaluate, contest and defend arguments, we may propose the following list of evaluative questions for an argumentation system (building on Verheij 1999, 43): Does the system track the issues raised? Does it track the assumptions made and the reasons adduced? Does it track the conclusions drawn and the counterarguments adduced? Does it track the justification status of the statements made and the commitments incurred? Does it verify that users obey the pertaining rules of argument? Does it identify omissions and weaknesses in arguments? Does it identify counter-arguments? Both Zeno and the Agora meet most of these requirements, though one can identify some areas for improvement. For example, the Risk Agora does not identify weaknesses or omissions in arguments. And the flexibility Zeno provides for rule violations may sometimes be an advantage (see sec. 7 below; Gordon and Karacapilidis 1997, 16-17), but without some kind of checks it could undermine reasonableness.

6. Inclusive Design

The third dimension of evaluation is concerned with the extent to which all the stakeholders are represented in the world of the computer-aided decision-process. The formal design features of neither Zeno nor the Agora appear to limit participation in any overt way, so that all those people interested in the discussion topic and willing to accept the rules of the dialogue are permitted to participate (in the case of Zeno) or able to have their views represented (in the case of the Agora). In other words, each system is formally inclusive: it gives anyone an equal right to enter the process. This criterion of formal inclusivity, however, is quite limited. In particular, it leaves the problem of actual use in context unresolved.

How well participants of diverse backgrounds and capacities make *actual* use of this formalism in concrete contexts of use remains an open research question to which AI researchers are sensitive (e.g., Schmidt-Belz *et al.* 1998). A formally inclusive design does not mean that all affected parties will in fact avail themselves of the system. On the one hand, exclusion may arise from inequalities among participants in their familiarity with rule systems and their computer skills, which tend to favor more educated participants. On the other hand, we should also ask whether all the relevant conversations take place through the medium of the system, as opposed to off-line discussion. The answers to these questions for any system will depend on the specifics of their implementation. We could imagine, for example, a situation where use of Zeno was mandated for an urban planning decision, with all stakeholders forced to conduct conversations occurring in and through it. The system has been designed to support this level of use, and the technical design ensures incorporation and integration of information of heterogeneous types (maps, blueprints, reports, email messages, etc). However, without such mandating of use, there is no guarantee that many, and especially many important, interactions

between stakeholders would not occur away from the system. In fact, the designers recognize off-line debate as necessary so long as some stakeholders lack access to the Internet or computer skills (Schmidt-Belz *et al.* 1998, 85). This situation becomes problematic, however, if all the relevant contributions are not represented in the system, for then one cannot expect it to provide an accurate record of the debate or of the status of various positions.[18]

7. Noncoercive Process

So far we have examined three normative dimensions that are incorporated into formal system design. Self-transformation can occur insofar as the system allows participants to retract commitments or adopt new positions, substantive quality can emerge insofar as the system allows the requisite dialectical and logical moves, and at least formal inclusion is acknowledged insofar as the system allows any user to participate. Similarly for noncoerciveness: at the level of formal design, both systems seek to be noncoercive. For example, the possibility of violating rules in Zeno could enhance the freedom of participants. The Risk Agora contains a number of rules that stipulate the freedom of the participants to make different kinds of speech acts. Speakers are free, for example, to challenge any assertion that has been introduced; they may also introduce any assertion without supporting argument, and may introduce or challenge rules of inference (McBerney and Parsons 2001b).

Of course, none of these features, taken as formal design, can ensure the deliberative quality of an actual process, given the human capacity for obtuseness and obstruction. As we saw above with inclusion, violations of noncoerciveness can arise from the social-psychological and institutional context in which systems are used. There are a complex set of issues here, bearing on the local legal-political setting and how the system is actually used (or ignored) in

context (Märker *et al.* 2002; Schmidt-Belz *et al.* 1998). More specifically, one might ask: What is the legal-institutional status of computational-deliberative results? On whose behalf is the system developed and deployed? Who should have rights to its findings?—and so on.[19] These questions illustrate the issues that may arise from an examination of the social and institutional relationships surrounding the use of any decision-support technology. None of these questions has a straightforward answer, and just resolution may only be possible on a case-by-case basis. Exploring them, one is confronted with those whom in sociology have been called "the locally powerful" (Bell 1978) —people who influence or control the decisions of others, even if only in a particular context. In the present context, one would expect the locally powerful to include users who are especially adept at employing the argumentation system to support their viewpoint. The danger in such inequality is that the more powerful players—perhaps even despite their good intentions—would fail to give some viewpoints the consideration they deserve on the merits. Non-standard approaches would thus be subtly forced out of the deliberative forum.

Behind these pitfalls, however, lies a deeper challenge for assessment, to which we now turn.

8. A Challenge forAssessment: Formal versus Pragmatic Rules

The deeper problem to which we refer also lies in a gap between procedural designs and actual performance, but the gap is less obvious than those mentioned above, which result from carelessness or an inhospitable social context. Failure to appreciate this subtler gap has had disasterous results, as in the Challenger space shuttle disaster. In her detailed case study, Vaughan (1996) shows how a well-intentioned, precisely followed formal design (for risk

assessment) actually undermined the aims of the design. Debunking the standard view that laid the blame on managerial violation of safety rules, she demonstrates that post-disaster investigators and journalists did not understand NASA's risk assessment categories and procedures; in fact, *no procedural violations* preceded the launch. On the contrary: in the years leading to the launch, formal risk assessment procedures were followed *more exactly* than in earlier days. This tightening of formal procedure had the perverse effect of inhibiting the flow of information—in particular difficult-to-justify tacit knowledge—from engineers to management. Although not the only factor, this aspect of organizational culture set up the mistakes that were made on the eve of the launch. The engineers argued against a launch, but under time pressure they failed to marshall persuasive arguments for their position.

Applied to our context, Vaughan's analysis suggests that subtle modes of coercion can emerge in the use of an argumentation system precisely *because* users strictly adhere to formal procedures. We focus now on the Risk Agora because it attempts to build rules of equal participation and freedom into the formalism itself.[20]

Specifically, the design features of the Agora attempt to formalize the kind of process standards articulated by Habermas (1983/1990, 88—89) and Alexy (1978/1990). McBurney and Parsons (2001b) are quite explicit about this: they show that the Risk Agora implements an applicable subset of Alexy's rules of discourse ethics and 15 of Hitchcock's 18 principles of rational mutual enquiry (Hitchcock 1991).[21] In the Alexy/Habermas model, noncoerciveness is expressed by the rule that prohibits any internal or external coercion that would prevent participants from exercising the twin freedoms to introduce or challenge any claim (Alexy 1978/1990, 167; Habermas 1983/1990, 89). This condition is satisfied in the Agora by specific allowable moves in the formalism. Nothing in the Agora syntax prevents speakers from making

these moves: that is, the moves that allow introduction of claims do not impose any precondition on the speaker, and the challenging moves only presuppose the introduction of the claim being challenged. Thus the Risk Agora avoids internal coercions resulting from system design.

However, by translating noncoercivenss into a formal argumentation system, designers attempt to bridge a challenging divide. Alexy and Habermas understand their rules as "pragmatic rules," that is, as attempts to articulate the presuppositions that anyone who seriously engages in argumentative discourse tacitly accepts (Alexy 1978/1990, 154; Habermas 1991/1993, 49-57). More precisely, pragmatic rules state "idealizations," ideal conditions that we can at best approximate in real discourse, but which we can never guarantee we have approximated through some verifiable test. Nonetheless, we must presume our discourse has sufficiently approximated such conditions, unless we have evidence to the contrary. By contrast, when these pragmatic rules are formalized as a set of moves in the Risk Agora, their observance is guaranteed insofar as they define what participants can say at all.[22] At the formal level, assessment of a noncoercive support system can look to a verifiable test: does the formal syntax correspond to pragmatic rules for unhindered participation?

Assessing a system for its noncoercive quality, however, requires the evaluator to ask two further types of question. One type we have noticed above (secs. 6, 7) when we noted inequalities and exclusions stemming from differences in user skills and the like: do the participants actually avail themselves of the possibilities built into the argumentation system? One might notice, for example, that in the course of a particular deliberation, users never employ certain types of allowable moves. This observation would raise the suspicion that users do not fully understand the formalism and its possibilities. Most cases of this type of coercion may not be so evident: to notice and address them AI researchers can benefit from the social sciences.

The deeper question for assessment provoked by Vaughan's study takes us still further, beyond both the testing of syntax and the examination of skill and power differences. One asks, that is, whether *technically competent* users, adept at employing all the resources of the system, have nonetheless been coerced by system design. This third sort of question requires not only sociological analysis but also a close study of the *content* of the deliberation.

That is, one must ask whether the particular formalism itself can present a form of exclusion in virtue of deep assumptions about reasonable argumentation. One can frame a range of questions at this level. Recalling the three maim roles of the system (sec. 2.3), one might ask questions such as the following: Do the participant-support mechanisms favor some parties over others? Here the partisan support role merits particular scrutiny. Is the tracking or record keeping genuinely neutral—that is, can each stakeholder perceive that the system has represented its position, interests, values, and arguments accurately? And does the forum structure (e.g., the sequencing of links at the user-interface) give some participants greater opportunity to influence the deliberation?

In the remainder of this section we address one such question, namely whether the system is flexible enough to capture all the styles of argument that might plausibly merit consideration. Consider in particular the argument forms the system permits. If a particular viewpoint rests on styles of argument that are excluded—or simply not representable—in the system, then those viewpoints are excluded from the start.[23] This issue becomes important for types of inferences or arguments whose validity is highly context-dependent (e.g., appeals to authority, to emotions, etc.). Insofar as the reasonableness of an appeal to authority, or to emotion, depends on the particular dialectical context (Walton 1995), they present a challenge to formal rule systems, albeit not an impossible challenge (cf. Reed 1998; Grasso 2002).

In this context, the fact that the Agora allows participants to introduce (without precondition), challenge, and revise inference rules takes on considerable significance (McBurney and Parsons 2001b). The possibility of such revision not only introduces greater flexibility into the system, it also opens up the possibility of responding to the subtle modes of coercion built into the rule system itself. Revising inference rules, to be sure, will not address all the possible sources of coercion in a system, but it does lead to a general point about the nature of such revisions.

When participants argue about a rule of inference, in effect they engage in a kind of assessment of noncoerciveness of the system they are using for a particular policy deliberation. Insofar as such engagement is crucial for a noncoercive use of the Agora in context, the capacity of participants for reasonably assessing the quality of their assisted deliberation is a necessary condition for noncoercive use in context. Consequently, if AI researchers want to assess whether a system such as the Agora satisfies the noncoerciveness condition at this deeper level, then they must assess the quality of participants' revisionary assessments. What we envision, in other words, are experiments (both laboratory and real-world) in which the AI researcher must evaluate how citizens actually use the Agora, and whether that use is noncoercive. But if such experiments are to address the deeper question raised by Vaughan, then the researchers must assess how well the participants themselves assess the noncoerciveness of the system in context.

Such revisionary moves are highly context- and issue-dependent. To assess participants' revisionary assessments, then, AI researchers must understand the policy setting as well as its substance. Moreover, they would have to understand the reasons participants have for revising a rule of inference in this context. Thus researchers presumably would have to engage the participants on their own turf, in a dialogue centered on the rule revision and its relation to the

particpants' assessment of noncoercive process. When the researcher engages citizens in dialogue at this level, however, computational expertise no longer provides a privileged vantage point on the reasonableness of the rule at issue. In a word, then, the researcher cannot remain an outside observer of citizens who use the system, but must become a particpant in the deliberation, at least to some extent. As a participant, however, the researcher is subject to the same pragmatic rules of persuasion dialogue as the citizens under study. The researcher would thus have to strive to reach a mutual agreement with citizens about the reasonableness of a particular rule revision.

One cannot rule out the possibility that researchers and citizen-users might not always be able to reach agreement about the reasonableness of a particular rule revision, and thus about the coerciveness or noncoerciveness of a design. In that case, in order to carry out an assessment of the noncoerciveness of their system, researchers would face some difficult questions: Are the citizens in this case simply unreasonable? Or have our own biases led us to dismiss their mode of argumentation? Or is this simply an intractable issue? Such questions are inherent in democratic debate, and it should not surprise us if attempts to design argumentation-support systems for such debate occasionally take the designers into such treacherous waters.

Argumentation-system designers cannot entirely eliminate such questions so long as they intend to build systems that are supposed to *improve* deliberation. A given design thus contains a *normative* theory of argumentation for deliberative contexts—hence, a normative model of deliberative legitimacy. This normative orientation puts the researcher in the position of a political theorist calling citizens to their better selves. But given that citizens typically have normative attitudes of their own about legitimacy, theoretical models are in turn open to a kind of reverse critique: citizen-users can complain that the theorist has not understood or properly

articulated deliberative ideals. Indeed, citizens might disagree among themselves about what makes for a good deliberative process. In their case study of different attitudes about good process for public participation in environmental policy, Webler, Tuler, and Krueger (2001) conclude that process designers must enter into discussion with citizens in order to determine their views about good process. Our conclusion is similar: the AI researcher cannot avoid being drawn into critical discussion as a participant on a par with citizens-users.

9. Conclusion

This paper has explored a number of evaluative issues associated with the use of computer decision support systems for public deliberation, in particular the Zeno system of Gordon and Karacapilidis and their colleagues at GMD and the Risk Agora of McBurney and Parsons. We have argued that assessments of the quality of such systems present challenges to system designers and users, and we have suggested a conceptual structure within which to undertake such assessments. This structure draws on deliberative democratic theory for assessment of deliberative planning processes. Our analysis indicates that the formal design features of the two argumentation systems capture many of the features of reasonable argumentation—mechanisms for ensuring substantive inclusiveness of relevant considerations, dialectically responsible argumentation, and an overtly open forum of participation. But to assess the deliberative quality of such systems in use, AI researchers must not only scrutinize formal design; they must also examine the social-institutional context for disparities among different user groups and possible misuses of the system, and they must assess the capacity of users to adapt the system to their local context by reasonable revisions of system features. To examine social-institutional context requires interdisciplinary work with political scientists,

sociologists, legal scholars, and policy planners. To assess the revisionary capacity of participants, however, interdisciplinary researchers must engage citizens at the particpant level, and give them a voice in system design and assessment.

Acknowledgments

This work was partly funded by the British Engineering and Physical Sciences Research Council (EPSRC) through a PhD studentship. We gratefully acknowledge this support. We are also grateful for discussions on these topics with Trevor Bench-Capon, Rod Girle, Muki Haklay, David Hitchcock and Bart Verheij. We also thank Thomas Gordon and R. Prescott Loui for their feedback on an earlier draft of this pape.

References

Alexy R (1978/1990): A theory of practical discourse. In: S. Benhabib and F. Dallmayr (Editors): *The Communicative Ethics Controversy,* pp. 151—190. Cambridge, MA, USA: MIT Press. (Translation by D. Frisby of: "Eine Theorie des praktischen Diskurses." In: W. Oelmuller, Editor, *Normenbegrundung-Normendurchsetzung.* Paderborn, Germany: Schoningh, 1978).

Alty JL and MJ Coombs (1984): *Expert Systems: Concepts and Examples.* Manchester, UK: National Computer Centre Publications.

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

Beetham D (1998): Legitimacy. In E Craig (ed.), *Routledge Encyclopedia of Philosophy*. London: Routledge. Retrieved July 21, 2004, from <www.rep.routledge.com/article/S034SECT5>

Bell C (1978): Studying the locally powerful: Personal reflections on a research career. In: C. Bell and S. Encel (eds.): *Inside the Whale: Ten Personal Accounts of Social Research.* Sydney, Australia: Pergamon Press.

Bench-Capon TJM, J Freeman, H Hohmann and H Prakken (2003): Computational models, argumentation theories and legal practice. In: C. Reed and T. Norman (eds.): *Argumentation Machines: New Frontiers in Argument and Computation.* Kluwer, Dordrecht.

Benfer RA, EE Brent, and L Furbee (1990). *Expert Systems.* Newbury Park: Sage.

Bledsoe CH and KM Robey (1986): Arabic literacy and secrecy among the Mende of Sierra Leone. *Man*, 21 (2) (n.s.): 202—226.

Bohman J (1996): *Public Deliberation: Pluralism, Complexity and Democracy.* Cambridge, MA,USA: MIT Press.

Bohman J and W Rehg (eds.) (1997): *Deliberative Democracy: Essays on Reason and Politics*. Cambridge, Mass.: MIT Press.

Buchanan BG and EH Shortliffe (1984): *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Reading, MA: Addison-Wesley.

Carbogim DV, DS Robertson and JR Lee (2000): Argument-based applications to knowledge engineering. *The Knowledge Engineering Review*, 15 (2): 119—149.

Cohen J (1989): Deliberation and democratic legitimacy. In A Hamlin and P Pettit (eds.), *The Good Polity,* 17-34. Oxford: Blackwell. Reprinted in Bohman and Rehg (1997).

Collins, HM, and R Evans (2002): The third wave of science studies: Studies of expertise and experience. *Social Studies of Science* 32:235-296.

Dahl RA (1956): *A Preface to Democratic Theory.* Chicago: University of Chicago Press.

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

Dung PM (1995): On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-persons games. *Artificial Intelligence*, 77: 321—357.

Elster J (1986). The market and the forum. In: J Elster and A Aanund (eds.): *The Foundations of Social Choice Theory.* Cambridge: Cambridge University Press, pp. 103—132. (Reprinted in Bohman and Rehg 1997, pp. 3—33.)

Emery J, R Walton, A Coulson, DW Glasspool, S Ziebland and J Fox (1999): Computer support for recording and interpreting family histories of breast and ovarian cancer in primary care (RAGs): qualitative evaluation with simulated patients. *British Medical Journal*, 319: 32—36.

Ess C (1996): The political computer: Democracy, CMC, and Habermas. In: C Ess (ed.): *Philosophical Perspectives on Computer-Mediated Communication,* pp. 197—230. Albany, NY, USA: SUNY Press.

Fairhead J, and M Leach (1998): *Reframing Deforestation: Global Analysis and Local Realities: Studies in West Africa.* London, UK: Routledge.

Feenberg A (1999): *Questioning Technology.* New York, NY, USA: Routledge.

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

Fox J and S Das (2000): *Safe and Sound: Artificial Intelligence in Hazardous Applications.* Cambridge, USA: MIT Press.

Gardenfors PG (1994): The role of expectations in reasoning. In: M Masuch and L Polos (eds.): *Knowledge Representation and Reasoning under Uncertainty: Logic at Work*, pp. 1—16. Berlin, Germany: Springer. (Lecture Notes in Artificial Intelligence 808).

Girle R, D Hitchcock, P McBurney and B Verheij (2003): Decision support for Practical Reasoning. In: C Reed and T Norman (eds.): *Argumentation Machines: New Frontiers in Argument and Computation.* Kluwer, Dordrecht.

Goldman AI (1994): Argumentation and social epistemology. *Journal of Philosophy*, 91: 27—49.

Goldman AI (1999): *Knowledge in a Social World.* Oxford, UK: Oxford University Press.

Gordon TF (1994): The Pleadings Game: An exercise in computational dialectics. *Artificial Intelligence and Law*, 2: 239—292.

Gordon TF (2003): An open, scalable and distributed platform for public discourse. Retrieved July 16, 2004 from <www.tfgordon.de/publications/Gordon2003a.pdf.>

Gordon TF, and N Karacapilidis (1997). The Zeno argumentation framework. *Proceedings of the Sixth International Conference on AI and Law,* pp. 10—18.. New York, NY, USA: ACM Press.

Gordon TF, N Karacapilidis, H Voss and A Zauke (1997): Computer-mediated cooperative spatial planning. In: H Timmermans (ed.): *Decision Support Systems in Urban Planning*, pp. 299—309. E and FN Spon.

Grasso (2002): Towards a computational rhetoric. *Informal Logic* 22: 195-229.

Greenberg HJ (1987): Validation of Decision Support Systems. In: G. Mitra (ed.): *Mathematical Models for Decision Support,* pp. 641-657. Berlin: Springer (NATO ASI Series).

Groothuis MM and JS Svensson (2000): Expert system support and juridical quality. In: J Breuker (ed.): *Proceedings of the Thirteenth International Conference on Legal Knowledge-Based Systems (JURIX-2000).* Enschede, The Netherlands.

Habermas J (1981/1984): *The Theory of Communicative Action: Volume 1: Reason and the Rationalization of Society.* London, UK: Heinemann. (Translation by T McCarthy of: *Theorie des Kommunikativen Handelns, Band I, Handlungsrationalitat und gesellschaftliche Rationalisierung.* Frankfurt-am-Main: Suhrkamp, 1981.)

Habermas J (1983/1990): *Moral Consciousness and Communicative Action.* Cambridge, MA, USA: MIT Press. (Translation by C Lenhardt and SW Nicholsen of: *Moralbewusstsein un kommunikatives Handeln.* Frankfurt-am-Main: Suhrkamp, 1983.)

Habermas J (1991/1993): *Justification and Application: Remarks on Discourse Ethics.* Cambridge, MA, USA: MIT Press. (Translation by CP Cronin of *Erläuterung zur Diskursethik*. Frankfurt-am-Main: Suhrkamp, 1991.)

Habermas J (1992/1996): *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy.* Cambridge, MA, USA: MIT Press. (Translation by W Rehg of *Faktizität und Geltung,* Frankfurt-am-Main: Suhrkamp, 1992.)

Haklay M. (2001): *Public Environmental Information Systems: Challenges and Perspectives.* PhD Thesis. Department of Geography, University College, London, UK.

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

Hintikka KJJ (1968): Language-games for Quantifiers. *Americal Philosophical Quarterly Monograph Series 2: Studies in Logical Theory*, pp. 46—72. Oxford, UK: Blackwell.

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

Hitchcock D (1992): Relevance. *Argumentation* 6: 251-270.

Hitchcock D, P McBurney and S. Parsons (2001): A framework for deliberation dialogues. In: HV Hansen, CW Tindale, JA Blair and RH Johnson (eds.): *Proceedings of the Fourth Biennial Conference of the Ontario Society for the Study of Argumentation (OSSA 2001).* Windsor, Ontario, Canada.

Jackson P (1986): *Introduction to Expert Systems.* Wokingham, UK: Addison-Wesley.

Jennings NR, P Faratin, AR Lomuscio, S Parsons, C Sierra and M Wooldridge (2001): Automated negotiation: prospects, methods and challenges. *Group Decision and Negotiation*, 10(2): 199—215.

Johnson R (2000): *Manifest Rationality: A Pragmatic Theory of Argument.* Mahwah, NJ, USA: Erlbaum.

Kahneman D, P Slovic and A Tversky (1982) (Editors): *Judgment under Uncertainty: Heuristics and Biases.* Cambridge, UK: Cambridge University Press.

Karacapilidis N, D Papadias, T Gordon and H Voss (1997): Collaborative environmental planning with GeoMed. *European Journal of Operational Research*, 102(2): 335—346.

Kirwan B and LK Ainsworth (1993): *A Guide to Task Analysis.* London, UK: Taylor and Francis.

Krause P, S Ambler, M Elvang-Gorannson and J Fox (1995): A logic of argumentation for

reasoning under uncertainty. *Computational Intelligence*, 11(1): 113—131.

Kriesberg L, TA Northrup, and SJ Thorson (Editors) (1989): *Intractable Conflicts and Their Transformation.* Syracuse, USA: Syracuse University Press.

Kuflik A (1999): Computers in control: rational transfer of authority or irresponsible abdication of autonomy. *Ethics and Information Technology,* 1(3): 173—184.

Lind EA and TR Tyler (1988): *The Social Psychology of Procedural Justice.* New York, USA: Plenum.

Lindley DV (1985): *Making Decisions.* London, UK: Prentice-Hall. Second Edition.

Mackenzie JD (1979): Question-begging in non-cumulative systems. *Journal of Philosophical Logic,* 8: 117—133.

Märker O, H Hagedorn, M Trénel, and T F Gordon (2002): Internet-based citizen participation in the city of Esslingen: Relevance – moderation – software. In: M Schrenk (ed.) *CORP 2000 – who Plans Europe's Future?* Vienna: Institute für EDV-gestützte Metbhoden in Architektur und Raumplanung der Technischen Universität Wien.

Mayo DG, and RD Hollander (eds) (1991): *Acceptable Evidence:Science and Values in Risk Management.* Oxford: Oxford University Press.

McBurney P and S Parsons (2000): Risk Agoras: Dialectical argumentation for scientific reasoning. In: C Boutilier and M Goldszmidt (eds.): *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence (UAI-2000)*, pp. 371—379. Stanford, CA, USA: Morgan Kaufmann.

McBurney P and S Parsons (2001a): Intelligent systems to support deliberative democracy in environmental regulation." *Journal of Information and Communications Technology Law*, 10(1): 33—43.

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

McBurney P and S Parsons (2002a): Using Belief Functions to forecast demand for Mobile Satellite Services. In: RP Srivastava and T Mock (eds.): *Belief Functions for Business Decisions,* pp. 281—315. New York, NY, USA: Physica.

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

McCarthy T (1999): Legitimacy and diversity: Dialectical reflections on analytic distinctions. In M Rosenfeld and A Arato (eds), *Habermas on Law and Democracy*, 115-153. Berkeley: University of California Press.

McGlade J (1999). Indigenous knowledge and the coastal environment: the SimCoast system. *Presentation to ESRC/GEC Seminar on Environmental Knowledge: Uncertainty, Authority and Responsibility*. Centre for the Study of Environmental Change, University of Lancaster, UK. 22 November 1999.

Michelman FI (1988): Law's Republic. *Yale Law Review*, 97(8): 1493—1537.

Michelman FI (1989): Conceptions of democracy in American Constitutional argument: the case of pornography regulation. *Tennessee Law Review*, 56: 291—319.

Moor J (1979): Are there decisions computers should never make? *Nature and System*, 1: 217—229.

Morgon MG (1993): Risk analysis and risk management. *Scientific American* (July): 32-41.

NASA (1999). Computer program assumes spacecraft command. Press release, 17 May 1999. Pasadena, CA, USA: National Aeronautics and Space Administration. Available from:

www.jpl.nasa.gov/releases/99/remote.html

Nisbett R and L Ross (1980): *Human Inference: Strategies and Shortcomings of Social Judgment.* Englewood Cliffs, NJ, USA: Prentice-Hall.

Parker C (2000): Requirements analysis and evaluation of decision support systems: the question approach. In: J. Reid and M. Lalmas (Editors): *Proceedings of the British Computer Society/Institution of Electrical Engineering Workshop on the Evaluation of Information Management Systems.* Department of Computer Science, Queen Mary College, University of London, London, September 2000.

Parsons S, C Sierra and NR Jennings (1998): Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3): 261—292.

Pera M (1994): *The Discourses of Science.* Chicago, IL, USA: University of Chicago Press.

Prakken H and G Sartor (1998): Modelling reasoning with precedents in a formal dialogue game. *Artificial Intelligence and Law*, 6: 231—287.

Prakken H and G Vreeswijk (2001): Logics for defeasible argumentation. In: D. Gabbay (Editor): *Handbook of Philosophical Logic*, Second Edition. Dordrecht, The Netherlands: Kluwer Academic (in press).

Rawls J (1996): *Political Liberalism.* New York: Columbia University Press.

Reed C (1998): *Generating Arguments in Natural Language.* PhD thesis, University College, University of London.

Reed C, and T Norman (eds.): *Argumentation Machines: New Frontiers in Argument and Computation.* Kluwer, Dordrecht.

Rehg W (1997a). Reason and rhetoric in Habermas's Theory of Argumentation. In: W Jost and MJ Hyde (eds.): *Rhetoric and Hermeneutics in Our Time: A Reader,* pp. 358—377. New

Haven, CN, USA: Yale University Press.

Rehg W (1997b): Legitimacy and deliberation in epistemic conceptions of democracy: Between Habermas and Estlund. *The Modern Schoolman*, 74: 355—374.

Rehg W (1999): Intractable conflicts and moral objectivity: A dialogical, problem-based approach. *Inquiry*, 42: 229—258.

Rittel HWJ and MM Webber (1973): Dilemmas in a general theory of planning. *Policy Sciences*, 4: 155—169.

Sandholm TW (1999): Distributed rational decision making. In: G Weiss (ed.): *Multiagent Systems: A Modern Introduction to Distributed Artificial Intelligence,* pp. 201—258. Cambridge, MA, USA: MIT Press.

Schmidt-Belz B, C Rinner, andT F Gordon (1998): GeoMed for urban planning—first user experiences. In: *Proceedings of the Sixth ACM International Symposium on Advances in Geographic Information Systems.* Washington, D.C.

Schneider CE (1998). *The Practice of Autonomy: Patients, Doctors and Medical Decisions.* Oxford, UK: Oxford University Press.

Schneider F (no date): Democracy goes on-line: Inernet-based citizen participation in the zoning plan. Retrieved July 16, 2004, from <www.mediakomm.net/documents/kongress/esslingen/schneider_en.pdf.>

Schumpeter, JA (1942/1976): *Capitalism, Socialism and Democracy.* New York: Harper.

Seffers GI (1998): Crisis system for President: DARPA's Project Genoa would speed U.S. decision-making. *Defense News*, 30 March, 13(13): 1.

Sierra C, NR 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 No. 1365. Berlin, Germany: Springer.

Stathis K (2000): A game-based architecture for developing interactive components in computational logic. *Electronic Journal of Functional and Logic Programming*, 2000(5).

Taylor A (1991): Organizational and social implications. In: T Bench-Capon (ed.): *Knowledge-Based Systems and Legal Applications,* pp. 309-327. London, UK: Academic.

Tindale, CW (1999): *Acts of Arguing.* Albany: SUNY Press.

Tonnelier CAG, J Fox, PN Judson, PJ Krause, N Pappas and M Patel (1997): Representation of chemical structures in knowledge-based systems: the StAR system. *Journal of Chemical Information and Computer Sciences*, 37: 117-123.

Toulmin SE (1958): *The Uses of Argument.* Cambridge, UK: Cambridge University Press.

Tyler, TR (1990): *Why People Obey the Law.* New Haven: Yale University Press.

Vaughan D (1996): *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago, IL, USA: University of Chicago Press.

Verheij B (1998): ArguMed: a template-based argument mediation system for lawyers. In: JC Hage, TJM Bench-Capon, AW Koers, CNJ. de Vey Mestdagh and CAFM Grutters (eds.): *Proceedings of the Eleventh Conference of Legal Knowledge-Based Systems (JURIX-98)*, pp. 113—130. Nijmegen, The Netherlands: Gerard Noodt Instituut.

Verheij B (1999): Automated argument assistance for lawyers. In: *Proceedings of the Seventh International Conference on Artificial Intelligence and Law,* pp. 43—52. New York City, NY, USA: ACM Press.

Walton DN (1995): *A Pragmatic Theory of Fallacy.* Tuscaloosa: University of Alabama Press.

Walton DN (2000): The place of dialogue theory in logic, computer science and communication studies, *Synthese*, 123: 327-346.

Walton DN and ECW Krabbe (1995): *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning.* SUNY Series in Logic and Language. Albany, NY, USA: State University of New York Press.

Warnke G (1999): *Legitimate Differences*. Berkeley: University of California Press.

Weber, M (1978): *Economy and Society*. 2 vols. G. Roth and C. Wittich (eds). Berkeley: University of California Press.

Webler T, S Tuler, and R Krueger (2001): What is a good public particpation process? Five perspectives from the public. *Environmental Management* 27: 435-450.

Weizenbaum J (1976): *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: Freeman.

Wynne, B (1989): Sheepfarming after Chernobyl: A case study in communicatingt scientific information. *Environment* 31/2: 10-39.

---

[1] For an introduction, see Benfer, Brendt, and Furbee 1990.

[2] One can discern a natural evolution from knowledge systems to argumentation systems in the increasing concern with providing users with the underlying explanation for the system's conclusion (or advice). Although developers of knowledge systems sought from the start to have their systems explain or justify their conclusions, early expert systems offered little in the way of explanation beyond high-level traces of the inference rules invoked in the chain of reasoning (Jackson 1986). Applications using some model of argumentation have included systems for: medical domains (Krause *et al.* 1995, Fox and Das 2000); legal domains (Gordon 1994, Verheij 1998, Bench-Capon *et al.* 2003); public policy decision support (Gordon and Karacapilidis 1997, Gordon *et al.* 1997, Karacapilidis *et al.* 1997); geopolitical risk prediction (Seffers 1998); scientific discourse (McBurney and Parsons 2000, 2001b); software design (Stathis 2000); and autonomous agent dialogues (Parsons *et al.* 1998, Sierra *et al.* 1998, Dignum *et al.* 2000). Carbogim *et al.* (2000) present a review of such applications

[3] For example, Dung (1995) proposed dialectical argumentation within AI as a proof theory for non-monotonic reasoning; for a review of this area, see Prakken and Vreeswijk 2001. For broad overviews of developments in computational dialectics, see Walton 2000, Reed and Norman 2003.

[4] As Haklay notes in a review of information technology for environmental decision-making, the philosophical and ethical aspects of information systems design are rarely made explicit or even explored (Haklay 2001).
[5] Much of the research in the field of AI and Law appears devoted to developing systems which take an orrery role, formalizing the reasoning processes used by decision-makers so as to better understand them (Bench-Capon *et al.* 2003).
[6] Tonnelier *et al.* 1997. Likewise, systems of autonomous software agents assume decision-making is being undertaken by computer decision-makers, and much of the research effort of second-generation electronic commerce, for instance, is directed at such automated decision-making (Jennings *et al.* 2001).
[7] The question to what extent humans should delegate decision-making authority to computer systems has been discussed in ethical philosophy, though a well-known early objection was raised by AI researcher Joseph Weizenbaum 1976; see also Moor 1979, Kuflik 1999.
[8] E.g., Buchanan and Shortliffe 1984; in Chapter 30 of this work, the authors consider quality evaluation methods for decision support systems. Greenberg (1987) also considers evaluation issues for such systems, focusing his discussion on the validation of the inference rules used by the system. Taylor (1991) briefly discusses evaluation questions as part of an exploration of the wider organizational issues associated with deployment of these systems. Parker (2000) proposes a design methodology based on the types of questions likely to be asked by a user of the system. While promising, this approach seems more appropriate for what we have termed knowledge systems (those encoding expert knowledge) than for argumentation systems.
[9] For example, the SimCoast expert system of the UK's Centre for Coastal and Marine Sciences, was developed to provide assistance to marine coastal environment decision-makers in developing nations (McGlade 1999). However, the expertise encoded in the system embodies a specific *Weltanschauung*, that of a standardized western scientific ontology, which may not accord with the worldviews or ontologies of non-scientist users in the developing world. Moreover, Western expertise may actually distort the perception of the local reality, as shown by the example of deforestation in West Africa (Fairhead and Leach 1998; cf. Harding 1998). Using developing-country residents as the "experts" for the design phase may well have led to a different system.
[10] There has been considerable research on this variance (see, e.g., Nisbett and Ross 1980; Kahneman *et al.* 1982). In a recent review of the limited research into how people make important medical decisions, Schneider concluded that patients make decisions quite differently from experts or normative decision theory; however, we have few means to measure the quality of medical decisions. Selecting one procedure or course of treatment usually precludes the selection of alternatives, and so strict comparison of results of alternative decision options at an individual level is impossible. The diversity and complexity of individual circumstances and medical aetiologies make comparisons at an aggregated level also problematic (Schneider 1998, 92, 97).
[11] Developers of public mobile satellite communications networks such as Iridium and ICO, for example, had lead times of a decade for designing, manufacturing and deploying the innovative satellite technology they required (McBurney and Parsons 2002a). To guide this work, the intending investors sought the advice of market researchers on the size of the potential market for mobile satellite services. In the particular decade concerned (1989-1999) demand for terrestrial cellular services grew much faster than anyone had forecast at the outset of the period, to the detriment of demand for mobile satellite services at the end of the decade. In major part

this growth in terrestrial cellular demand was spurred by technological changes and the spread of cellular network coverage unanticipated in 1989. Was the advice given by the market researchers wrong because it did not predict unanticipated events, or, if anticipated, did not sufficiently emphasize unlikely events? The advice may not have been wrong if the world had been otherwise in the subsequent decade. Most management consultancy is not assessable or assessed for these two reasons (an observation based on the second author's decade of experience as a management consultant).

[12] In the one quality assessment of decision support systems known to us, Groothuis and Svensson (2000) assessed the quality of decisions made with the help of computer systems in the provision of welfare assistance in the Netherlands, comparing the actual determination reached in a sample of welfare assistance cases with the decision which should have been made under the relevant laws. The results of this assessment showed that decision quality varied according to the extent to which the system provided support for the complex administrative tasks involved. For those cases where the system provided full support, few human errors were made; in other cases, where only limited support was provided, decision errors were more frequent. The authors concluded that this was due to the human decision-makers trusting the advice of the system even when such trust was not warranted, thus revealing weaknesses in the human decision-makers rather than in the decision support systems.

[13] For example, a system designed to support water-flow management through a dam will only be required to recommend responsive actions for 200-year floods on average once every 200 years. There may be insufficient data to design the system or to predict its performance in these circumstances, and possibly only one case every 200 years on which to base a live assessment of that performance. Moreover, if a system is designed for an entirely new activity how does one assess the adequacy of its advice? The various intelligent agent systems currently being deployed by NASA for control of autonomous spacecraft are examples of systems which undertake completely new activities. How is it possible to rate their performance in any other but crude terms, such as overall mission success versus non-success? (NASA 1999).

[14] For a well-known statement of the pluralist model, which draws on James Madison's Federalist Paper no. 10, see Dahl 1956.

[15] In Hitchcock *et al.* (2001), the characteristic of sharing information is taken to be one that distinguishes deliberation from negotiation dialogues. In contrast, because much of the research focus in the area of intelligent multi-agent software systems has been on automated negotiation (Jennings *et al.* 2001), this self-transformative condition is not satisfied by all agent systems.

[16] See Bohman 1996, esp. chap. 3. Even a technology as apparently benign as literacy can be used by those possessing it to establish and maintain political power, as Bledsoe and Robey (1986) show in their study of Arabic literacy among the Mende of Sierra Leone.

[17] For a review of automated negotiation systems, see Jennings et al. 2001.

[18] By contrast, in systems of autonomous software agents using argumentation, such as the team-formation systems of Dignum *et al.* (2000), the entirety of the dialogue occurs inside the system.

[19] Such questions arise, for example, in the use of intelligent systems in medical domains. See, for example, Emery *et al.* 1999.

[20] We focus on the Risk Agora because the dialogue rules are explicit in the system formalism. This is not to deny that certain features of the Zeno Argumentation Framework could be translated into such rules.

[21] Although similar procedural criteria have been proposed for assessment of automatic electronic auction systems (Sandholm 1999), to our knowledge this is the first time such

proposals have been made for argumentation-based computer systems.

[22] Cf. Hitchcock's "externalization" property, which requires that rules in the system "be formulated in terms of verifiable linguistic behaviour" (Hitchcock 1991,).

[23] For example, Gardenfors 1994 discusses the reasoning processes used by illiterate Uzbeki peasants, and argues that apparent violations of deductive inference rules are, in fact, differential assessments of argument premises according to the perceived experience of the proponent of the premise. A formalism which did not represent this aspect of argument premises would not be able to represent the styles of arguments used by the Uzbekis.