# Ensemble Theory: Arguing Across and Within Scenarios

Peter McBurney
Department of Computer Science
University of Liverpool
Liverpool L69 7ZF UK
p.j.mcburney@csc.liv.ac.uk
Tel: + 44 151 794 6760

Simon Parsons
Sloan School of Management
Massachusetts Institute of Technology
Cambridge MA USA
sparsons@mit.edu

**A paper presented at the**

**AN INTERNATIONAL CONFERENCE AT THE UNIVERSITY OF STRATHCLYDE GRADUATE SCHOOL OF BUSINESS IN GLASGOW, UK**

***Probing the Future: Developing Organizational Foresight in the Knowledge Economy***

11-13th July 2002

Key words: Argumentation, Decision-Support Systems, Scenarios.

## Abstract

Scenario planning provides a widely-used means of summarizing the multitude of possible ways in which the future may arrive. Given certain assumptions about the world as it is now in some domain, and assumptions about the factors which may influence it and their inter-relationships, we can articulate an argument for how the present may evolve into the future. With different starting assumptions and different influencing factors, we may articulate alternative arguments for this possible evolution, each corresponding to one or more evolutionary paths.

A major challenge for organizational or public policy use of scenario planning is that different people have different starting assumptions and believe different influencing factors are important; thus they articulate or opt for different future evolutionary paths. How might these different paths be compared and synthesized? Drawing on the philosophy of argumentation, we have developed a theory to support comparison and synthesis of alternative future scenarios. Our approach enables argument both within one scenario and across alternative scenarios.

Our theory, which we call ensemble theory, is defined in a computational manner, making it suitable for implementation itself in intelligent computer systems. We describe an example of a computer application in a clinical medical domain, for an intelligent system to advise patients on the consequences of their lifestyle options, medical tests and treatments for breast cancer.

## 1. Introduction

In many domains, the absence of hard data or the presence of conflicting perceived interests makes reaching agreement on the quantification of uncertainty difficult. Argumentation formalisms have been proposed for the qualitative representation of uncertainty in these circumstances (Krause *et al.* 1995) and have found application in intelligent systems, for example in medical and safety analysis domains (Carbogim *et al.* 2000). In McBurney & Parsons (2001b), we proposed a formalism using dialectical argumentation for representing and resolving the arguments for and against uncertain propositions. Our formalism was particularly appropriate for application domains involving missing or ambiguous data, or where multiple stakeholders have divergent interests. We grounded our representation in specific theories of rational human discourse and centered it on an electronic space for presentation of arguments, termed an *Agora*. In addition, we demonstrated that this formalism had several desirable properties when used for inference and decision-making. In this paper, we extend this framework to enable dialectical argumentation under and between multiple sets of circumstances, or scenarios.

The notion of scenario has found widespread application in business forecasting, in public policy determination, and in scientific domains (Schwartz 1991, McBurney & Parsons 2002c). An early use of the methods of scenario analysis may be seen in nineteenth-century statistical mechanics, where research sought to determine if the properties of a physical system, such as its entropy at a given time, depended on the system's initial state. Ludwig Boltzmann (1872) tackled this problem by comparing the given system to a collection of alternative, imaginary systems, each having different initial conditions — i.e., what we would now call scenarios. By doing so, he could potentially assess the extent to which the system property of interest was independent of the initial system state. Josiah W. Gibbs (1902) formalized the concept of a collection of alternative systems with his notion of *ensemble*, a term we also use in this paper.

Perhaps the most important and complex recent application of scenario analysis has been in the work of the Intergovernmental Panel on Climate Change (IPCC) (McCarthy *et al.* 2001), the UN agency tasked with assessing the current and possible future states of the world's ecosystem, and with considering and recommending appropriate environmental regulatory policies. In this domain, scenario analysis has been used for scientific modeling and prediction, for the modeling of socio-economic variables and conditions, and for the assessment of proposed regulatory policies and targets (Carter *et al.* 2001)

Despite their widespread use, however, there appears to be no formal theory of scenarios or scenario analysis. Without a formal theory, many questions remain without rigorous answers, e.g., How should scenarios be constituted? How many scenarios should be considered? How should individual scenarios be analysed? How should any differences in the likelihood of occurrence of different scenarios be represented? How should their relative importance be represented? How should reasoning be undertaken *across* a collection of scenarios, or multiple collections of scenarios? In the absence of a formal theory of scenarios it is difficult to assess the validity or reliability of any particular application of scenario analysis, for example, the many analyses generated by the work of the IPCC. Moreover, because no computational theory of scenarios yet exists, application of scenario analysis in intelligent systems is limited.

The long-term aim of the research reported here is a rigorous, formal, computational theory of scenarios. In this paper, we take an initial step towards this aim, by considering one type of scenario, those based on dialectical argumentation systems. In Section 2, we review our model for qualitative inference in uncertain domains, which uses dialectical argumentation to represent conflicting, ambiguous or contested information. Section 3 defines our notions of scenario and ensemble, while Section 4 discusses how we may reason across multiple scenarios. An example of the application of these ideas to medical decisions support systems is presented in Section 5, and Section 6 concludes the paper.

## 2. Dialectical Argumentation

In this section we briefly summarize the Agora framework for the qualitative representation of uncertainty which we presented in earlier work (McBurney & Parsons 2001b). In this framework, arguments for and against claims are articulated by participants in an electronic space, an *Agora*, with claims expressed as formulae in a propositional language. By means of defined locutions, participants in the Agora can variously posit, assert, contest, justify, rebut, undercut, qualify and retract claims, just as happens in real discourse. For example, a debate participant *P* could demonstrate her argument supporting a claim *c*, an argument to which she was committed with strength *D*, by means of the locution:

**show_arg**( P : Arg(*c*))

In this formalism, we use the symbols "Arg(*c*)" to denote the argument presented by P in support of the claim *c*. To manipulate arguments such as these computationally requires a model of argument, and so our Agora framework draws on the well-known model of Stephen Toulmin (1958). In this model, a claim is supported by specific premises (called "data"), which are connected to the claim by means of a "warrant". Thus, to support a claim that a particular medical patient has a high chance of developing breast cancer, we may provide data in the form of the patient's family history of the disease, along with a warrant which connects any person's chances of developing cancer to that person's family history. Such a warrant may be supported by epidemiological data or genetic theory. Because data and warrants may only provide limited support for claims, Toulmin's model also permits claims to be qualified by "modalities", such as "probably" or "almost certainly". The Agora framework also permits the articulation of arguments expressing the consequences of some claim, such as the likelihood of the patient's children also developing breast cancer if the patient does.

With this model of an argument we can consider the manner in which arguments may relate to each other, and here we used terminology which has become standard in computational argumentation. An argument is said to be **consistent** if its various premises do not contradict one another. Given any argument in support of some claim *c*, a **rebuttal** is an argument in support of the negation of the claim, i.e., an argument that $c$ is not true.[1] Given an argument for a claim $c$ which relies on some premise $q$, an undercutting argument, or an **undercut**, is an argument for the negation of the premise, i.e., an argument that $q$ is not true. Rebuttals and undercuts of an argument are said to **attack** that argument, and are called **attackers**.

In the Agora framework, the rules governing the use of permitted locutions are expressed in terms of a formal dialogue-game between the participants. Dialogue games involve formalized interactions between two or more participants, in which each player "moves" by making utterances in a conversation, according to various pre-defined rules. Only utterances from some set of defined locutions are permitted. These games were first studied by Aristotle (1928) and then by medieval philosophers, before being taken up my modern philosophers for the study of fallacious arguments (Hamblin 1970) and as a game-theoretic semantics for formal logic (Lorenzen and Lorenz 1978). Because they are rule-governed, they are particularly attractive for computer science, and have been used, for example, to represent legal reasoning (Bench-Capon *et al.* 2000) and to design protocols for interaction between different computational entities, for example, for automated consumer purchase negotiations (McBurney, van Eijk *et al.* 2002).

We assume that the Agora participants begin a debate with a set of agreed facts, or assumptions, and an agreed set of inference rules (or warrants). Because we want to model many forms of reasoning, these warrents need not be deductive and may themselves, in our Agora formulation, be the subject of argument. Similarly, participants may question and contest the premises, the modalities and the consequences of claims presented in arguments by other participants.

We demonstrated the use of this framework for the representation of uncertainty by defining a set of qualitative uncertainty labels assigned to claims on the basis of the arguments presented for and against them in the Agora. Essentially, one could say that claims have more credibility (and hence less uncertainty) the fewer and the weaker are the arguments against them. While any set of labels could be so defined, we drew on earlier work in computational argumentation (Krause *et al.* 1995) and defined the collection of labels:[2]

*Accepted, Probable, Plausible, Supported, Open.*

Our definitions of these labels were as follows. A claim $c$ is said to be:

- **Open** at time $t$ if there are no arguments presented for $c$ in the Agora up to this time.

- **Supported** at time $t$ if an argument in support of $c$ has been presented in the Agora by this time.

- **Plausible** at time $t$ if a consistent argument in support of $c$ has been presented in the Agora by this time.

[1] For simplicity, we are here assuming classical, two-valued, logic, where every statement is either true or false and not both and not neither. Other logics, such as intuitionistic logic or multiple-valued logics, are also possible.
[2] It is important to stress that nothing in our approach precludes the use of labels other than these.

- **Probable** at time *t* if a consistent argument in support of *c* has been presented in the Agora by this time, and no undercuts or rebuttals have been presented against the argument by this time.

- **Accepted** at time *t* if a consistent argument in support of *c* has been presented in the Agora by this time, and every argument attacking this argument is itself attacked.

These labels express increasing certainty about a claim being true, at any given time. However, as arguments for and against a proposition are presented to the Agora, the status of a proposition may rise or fall: a claim considered *Probable* at one time may be only *Plausible* later, and then be *Accepted* later again. We therefore defined the **truth-valuation** of a claim *c* at time *t*, denoted $v\_t(c)$, to be 1 if *c* had the label *Accepted* at this time, otherwise it was 0. Such a valuation summarizes the knowledge of the community of debate participants at the particular time, since it incorporates, via the definitions of the labels, all the arguments for and against ☙ articulated to that time. Consequently, assessing the truth-status of a claim at a particular time can be viewed as taking a *snapshot* of an Agora debate. Of course, because these definitions are time-dependent, and arguments may be articulated in the Agora at any time, such an assignment of uncertainty labels and truth valuation must be defeasible. Claims accepted at one time may be overturned at another, in the light of new information learnt or arguments presented subsequently.

In using the Agora framework to represent uncertainty, attention will focus on the truth valuation function over the long-run.[3] The sequence of truth-values $(v\_t(c) \mid t = 1, 2, \ldots)$ may or may not converge as time *t* heads to infinity. Suppose that it does converge, and denote its limit value by $v(c)$. What will the value of a snapshot taken at time *t*, namely $v\_t(c)$, tell us about $v(c)$? Of course, since any finite snapshot risks being overtaken by subsequent information or arguments, we cannot infer with complete accuracy from the finite snapshot to the infinite value. However, we have shown (McBurney & Parsons 2001b) that, under certain conditions, we can place a bound on the likelihood that such an inference is in error. The conditions essentially require that: firstly, the snapshot is taken at a time after commencement sufficient for all the arguments using the initial information to have been presented, and secondly, there is a bound on the probability that new information arises following the snapshot.[4]

This result is analogous to the standard, Neyman-Pearson, procedures for statistical hypothesis testing (Cox and Hinckley 1974). When testing a scientific hypothesis we can never be certain that inferences based on a sample, drawn randomly from a given population, are valid when applied to that population. But the Neyman-Pearson procedures enable us to estimate an upper bound on the probability that such inferences are invalid. In other words, we can bound the probability of inference error. Likewise, in the case of Agora debates, this proposition provides us with some confidence in our use of finite snapshots to make inferences about the long-run truth-

[3] Strictly, we are assuming throughout that time in the Agora is discrete, and can be represented by a countably-infinite set.
[4] This result is proved as Proposition 7 of McBurney & Parsons (2001b).

valuation function for a debate. While such inference is not deductively valid, at least its likelihood of error may also be bounded.

## 3. Ensembles

The framework we have just outlined provides a means to represent the diverse arguments that may be derived from a given set of assumptions, by means of a given set of inference rules (deductive or otherwise). If we were to start with a different set of assumptions, and/or permit the use of different inference rules, the arguments presented in the Agora could well be different. As a result, the uncertainty labels and truth values assigned to formulae could also be different, both when taken at finite snapshots and in the limit. Each collection of alternative sets of assumptions and inference rules we call a scenario, which we define as follows:

> A **Scenario *S*** for a given domain consists of a set of assumptions and a set of inference rules, with which participants are equipped at the commencement of an Agora debate over propositions in that domain. For any scenario ***S***, an Agora debate undertaken with the assumptions and inference rules of that scenario, is said to be the Agora **associated** with ***S***. We assume only one Agora debate is conducted in association with any scenario.

Because we wish to reason across multiple scenarios, we also define:

> An **Ensemble *E*** is a finite collection of distinct Scenarios relating to a common domain. We assume that, associated with each scenario is a real-number between zero and one, called its **scenario weight**. We call the vector of these weights for all the scenarios in a given ensemble ***E*** the **ensemble weights vector** of ***E***.

What interpretation we give to the weights depends upon the meanings we give to the logical language, to the scenarios and to arguments for claims in the corresponding Agora debates. Three possible interpretations are as follows:

- **Resources:** The assumptions and claims may represent objects in the physical world, and the inference rules physical manipulations of these objects, such as actual construction of new objects from existing ones. Scenarios can thus be interpreted as different sets of resourcing assumptions, with claims being well-defended in an Agora debate when the objects they represent are able to be constructed with the assumed resources. In this interpretation, the weights attached to scenarios may be the relative costs or benefits of different resources, or their likelihoods of occurrence.

- **Interaction protocols:** A second interpretation could arise where the scenarios represent alternative sets of rules of procedure for interaction between a group of participants, for example, in a legal domain or in automated negotiation. Here the rules of inference may represent different allowable modes of reasoning, such as reasoning by analogy or reasoning from authority. The weights may represent the extent of compliance of each scenario with some set of principles of rational discourse, such as those articulated by Alexy (1990) or Hitchcock (1991), or with some normative economic or political theory.

- **Uncertain assumptions:** Finally, a third interpretation would have the scenarios as different descriptions of some uncertain domain, for example different scientific theories, with propositions being statements about the domain, and the inference rules representing different causal mechanisms. The scenario weights could be relative likelihoods of occurrence, or valuations of relative importance or utility. This third interpretation is the one we will consider in this paper.

We do not assume the weights sum to unity across the scenarios in an ensemble, although they may do so. For example, the first interpretation may be appropriate for applications in robotics, where robots are engaged in identifying the tasks necessary to achieve some pre-defined goal or goals, an activity known as *Planning* in Artificial Intelligence (Allen *et al*. 1990). In the case where the robots are uncertain of each other's resources and capabilities, the scenario weights could represent relative feasibility of different planning assumptions. If the scenarios in an ensemble were mutually exclusive and exhaustive of possible resourcing assumptions, then it would be reasonable for the weights to sum to unity across the scenarios in the ensemble.

The weights may vary with time, but, if so, we assume that their assignment to scenarios is independent of the dialectical status of claims in the corresponding debates. This assumption of independence is made because the assignment of weights to scenarios should be on the basis of characteristics of the scenarios themselves, such as their assumptions and inference rules, not on the basis of arguments which ensue or don't ensue in the associated Agora debates. In other words, we may assign weights to a scenario on the basis of arguments *about* the scenario, not on the arguments conducted *under* the scenario.

Finally, there are two technical issues which we mention only briefly here although they are of major importance in any computational application. Firstly, we have not discussed procedures for the creation of scenarios. If they represent possible assumptions about some uncertain domain, then it may be sensible to define a scenario for each possible combination of the truth values of the uncertain assumptions. Doing this, however, quickly leads to a combinatorial increase in the numbers of scenarios, a state-space explosion problem. Articulating procedures robust against this problem is not something we (or anyone) have yet achieved. Secondly, our definition of an Ensemble requires that the included scenarios must be distinct. This is necessary so that when aggregating across scenarios we do not engage in "double-counting" of separate scenarios which are really the same. How to determine whether two scenarios are indeed distinct is a difficult philosophical and practical question. We have proposed a decision rule for determining an answer to this question for scenarios in an argumentation context in McBurney & Parsons (2002b). More work will be needed to extend this decision rule to other types of scenarios.

## 4. Arguing across Scenarios

Given we have an ensemble of distinct scenarios, under each of which an Agora debate has been conducted, how may we reason across the ensemble? In other words, how may we coherently aggregate the information contained in the entire ensemble? The answer to this question depends upon what we are seeking to achieve in such reasoning. Since we are working in argumentation context, we may wish to determine the dialectical status of some claim *c* across the range of the different scenarios in the ensemble. As with the situation *within* one individual Agora debate, where we assign the uncertainty labels: *{Accepted, Probable, Plausible, Supported, Open}*, we may similarly assign qualitative uncertainty labels *across* Agora debates on the basis of dialectical argumentation status. We present some possibilities here.

Recall that, within any one Agora debate, arguments for and against a claim may be articulated. This problem is amplified when we consider Agora debates under multiple scenarios: a claim may have a very different dialectical status under different scenarios. This thought leads one to consider labels of the following sort. Given an Ensemble *E* of distinct scenarios, a claim *c* is said to be:

- **Open** at time *t* if there are no arguments presented for *c* up to this time in any Agora associated with a scenario in *E*.

- **Possible** at time *t* if *c* is assigned the status *Accepted* at this time in at least one Agora associated with a scenario in *E*.

- **Inevitable** at time *t* if *c* is assigned the status *Accepted* at this time in every Agora associated with a scenario in *E*.

- **Impossible** at time *t* if *not-c*, the claim expressing the negation of *c*, is assigned the status *Accepted* at this time in every Agora associated with a scenario in *E*.

If we further assume that the scenarios in ensemble E are assigned a vector of ensemble weights, then we can take the weights into account in determining the likelihood of a claim *c* being true across the ensemble. Thus, our labels could include, for example:

- A claim *c* is said to be **more-likely-than-not** at time *t* if *c* is assigned the status *Accepted* at this time in Agora debates associated with scenario in *E* having total weight greater than the total of those in which *c* is not assigned the status *Accepted*.

Extending this idea, we could define a whole family of qualitative labels, for every percentage *a* between zero and 100, as follows

- A claim *c* is said to be ***a*%-likely** at time *t* if *c* is assigned the status *Accepted* at this time in Agora debates associated with scenario in *E* having total weight at least *a* percent of the total weight of all scenarios.

So, for instance, if the weights of the scenarios where a particular claim *c* is accepted comprise 35% of the total scenario weights across the entire ensemble, we would say that *c* is 35%-likely under *E*.

These definitions assign labels to claims on the basis of the weighted proportion of Agora debates in which the claims are *Accepted*. In Section 2, however, the qualitative uncertainty labels we defined represented various dialectical statuses. Using these, therefore, it would be possible to define a multitude of labels for the status of a claim across an ensemble, depending on the particular mix of statuses achieved by the claim in the debates under the various scenarios. The primary challenge of such an approach is comprehending the resulting multi-dimensional description of the ensemble status of a claim.[5]

One possible response to these proposals is that scenario analysis is unnecessary in an argumentation context, because these frameworks have been developed precisely to represent conflicting or uncertain information, and to resolve any inconsistency in the resulting conclusions. Such a view is mistaken. In a typical application, we are not merely trying to decide whether the possible arguments for some given claim are, on balance, stronger or weaker than the arguments against it; we are also trying to identify the *circumstances* (the assumptions and allowed warrants) under which arguments exist for or against the claim, and the *circumstances* under which those arguments for it are stronger than those against it. To do this rigorously, we need to clearly demarcate the sets of possible circumstances — i.e., the scenarios — from one another and to compare them.

## 5. An Application

In this section we present an example of the application of these ideas in the domain of clinical medicine. As is well-known, many illnesses and medical conditions are caused or influenced by genetic factors, and being able to test for the presence of these factors greatly facilitates risk assessment and treatment. For example, mutations in the BRCA1 and BRCA2 genes are responsible for 5—10% of all breast and ovarian cancers in women. Women with these genes have a lifetime risk of contracting breast cancer of up to 80%, compared with about 10% for all women in the population.[6] One of the results of the revolution in molecular genetics is the availability of tests for the presence of these and other genes, tests which are increasingly becoming available. Receiving a positive result before the cancer has started increases the patient's monitoring and treatment options, and thus her chances of survival.

However, a positive test result is not necessarily better than no test at all (Brody 1999, MacDonald 2000). Employers, medical insurers and Governments may all discriminate against people with a known higher risk of some medical condition. Even other family members – whose own risks may increase with the knowledge that one member of the family has had a positive test result – may not welcome the news, since it may change their own circumstances. Thus, the advent of genetic testing can lead to increased decision-making requirements upon patients, an area in which most

[5] We presented one approach to this problem in McBurney & Parsons (2001c).
[6] Center for Cancer Risk Analysis, Massachusetts General Hospital, Boston, MA, USA.

patients most of the time do not follow decision-theoretic best practice (Schneider 1998). Accordingly, medical professionals and others have seen a need for forms of genetic counselling, to provide individual patients with estimates of their personal risks of developing a disease under different circumstances, and to assist them to choose between alternative courses of monitoring, testing and treatment, and between alternative lifestyle options. This requirement has led to the development of computer-assisted decision-support systems, able to provide information and guidance to patients and their medical care-givers in these decisions, e.g., Emery *et al.* (1999), Coulson *et al.* (2001).

In providing dispassionate advice to patients, such genetic counselling systems need to incorporate the major factors influencing the assessment of risk of contracting the disease in question. As mentioned, these factors typically include the presence or absence of certain genes which predispose the patient to the disease. However, for many diseases they also include lifestyle and demographic factors, such as whether or not the patient smokes, the number of children a woman may have, and the age at which they were born. These factors are, to a greater or lesser extent, within the control of the patient, at least in so far as they concern future decisions. It is therefore reasonable to consider each combination of patient decision-options – whether or not to have a genetic test; whether or not to have further children; when to have children; etc. – as an alternative scenario. Under each scenario, there will be arguments supporting the claim that the patient will contract the disease and arguments against this claim; these arguments may be resolved through the calculation (from relevant epidemiological data) of a quantitative estimate of lifetime risk of contracting the disease. Likewise, under each scenario, there may be arguments for and against particular treatment options or lifestyle choices. Having a breast removed, for example, may reduce the patient's risk of cancer but at the possible expense of a significant loss of quality of life.

What we have just described is a high-level specification for a computer decision-support system which assists patients to undertake scenario-based decision-making in an argumentation context. In order to develop a computer system with this capability, the system needs to be able to articulate arguments for and against various testing, monitoring, treatment and lifestyle decision-options within each scenario in an ensemble. The model we presented in Sections 2, 3 and 4 provides the formal structure for such a computer system, and it is one we are currently developing in collaboration with the Advanced Computation Laboratory of Cancer Research UK, in London.

## 6. Conclusion

In this paper, we have reported on our current research effort attempting to formalize scenario analysis. We seek to do this primarily so that scenarios and scenario analysis may be incorporated in a coherent manner into intelligent computer systems, such as those supporting human decision-making in complex domains. An application of these ideas is to the development of computer-based decision-support systems for genetic counselling, to advise patients on their testing, prophylactic, treatment and lifestyle options under a range of possible circumstances and outcomes (Emery *et al.* 1999, Coulson *et al.* 2001). Other examples include systems to support public deliberation over environmental issues (McBurney & Parsons 2001a) and the design of tests of scientific hypotheses (McBurney & Parsons 2002a). A secondary benefit of a formal theory of scenarios would be the development of analytical tools to assess the validity of particular applications of scenario analysis and planning. The Kyoto Agreements on climate change, for instance, made extensive use of scenario analysis (Carter *et al.* 2001), and with possibly catastrophic consequences for us all; it is impossible to assess the validity or appropriateness of this application of scenario analysis without a formal theory of scenarios.

## References

R. Alexy [1990]: “A theory of practical discourse”, pp.151—190 in S. Benhabib and F. Dallmayr (Editors): *The Communicative Ethics Controversy*. Studies in Contemporary German Social Thought. Cambridge, MA, USA: MIT Press. (Translation by D. Frisby of: “Eine Theorie des praktischen Diskurses”, In: W. Oelmüler (Editor), *Normenbegründung-Normendurchsetzung*. Paderborn, Germany: Schöningh, 1978.)

J. Allen, J. Hendler and A. Tate [1990]: *Readings in Planning*. San Mateo, CA: Morgan Kaufmann.

Aristotle [1928]: *Topics.* Edited by W. D. Ross. Oxford, UK: Clarendon Press.

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

L. Boltzmann [1872]: “Weitere Studien uber das warmegleichgewicht unter Gasmolekulen”, *Wissenschaftliche Abhandlungen*, **1**: 316—402. English translation in: S. G. Brush [1966]: *Kinetic Theory, Volume 2*. Oxford, UK: Pergamon Press.

J. E. Brody [1999]: “Cancer gene tests turn out to be far from simple”, *New York Times*, 17 August 1999.

D. V. Carbogim, D. S. Robertson and J. R. Lee [2000]: “Argument-based applications to knowledge engineering”, *Knowledge Engineering Review*, **15 (2)**: 119—149.

T. R. Carter, E. L. La Rovere, R. N. Jones, R. Leemans, L. O. Mearns, N. Nakicenovic, A. B. Pittock, S. M. Semenov and J. Skea [2001]: “Chapter 3: Developing and Applying Scenarios”, pp. 145—190 in: J. McCarthy, O. F. Canziani, N. A. Leary, D. J. Dokken and K. S. White (Editors): *Climate Change 2001: Impacts, Adaptation and Vulnerability*. *Intergovernmental Panel on Climate Change, Working Group II, Third Assessment Report.* Cambridge, UK: Cambridge University Press.

A. S. Coulson, D. W. Glasspool, J. Fox and J. Emery [2001]: “RAGs: A novel approach to computerized genetic risk assessment and decision support from pedigrees”, *Methods of Information in Medicine*, **40**: 315—322.

D. R. Cox and D. V. Hinkley [1974]: *Theoretical Statistics*. London, UK: Chapman and Hall.

J. Emery, R. Walton, A. Coulson, D. W. 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.

J. W. Gibbs [1902]: *Elementary Principles in Statistical Mechanics, Developed with Especial Reference to the Rational Foundation of Thermodynamics*. New Haven, CT, USA: Yale University Press. Reprinted by Dover, New York, 1962.

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

D. Hitchcock [1991]: "Some principles of rational mutual inquiry", pp. 236—243 in: F. van Eemeren, R. Grootendorst, J. A. Blair and C. A. Willard (Editors): *Proceedings of the Second International Conference on Argumentation (ISSA-1991)*. Amsterdam, The Netherlands: SICSAT: International Society for the Study of Argumentation.

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

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

A. S. MacDonald [2000]: "Human genetics and insurance issues", in: I. Torrance (Editor): *Bio-Ethics for the New Millenium*. Board of Social Responsibility, Church of Scotland. Scotland, UK: St. Andrew Press.

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

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

P. McBurney and S. Parsons [2001b]: "Representing epistemic uncertainty by means of dialectical argumentation", *Annals of Mathematics and Artificial Intelligence* (Special Issue on Representations of Uncertainty), **32 (1—4)**: 125—169.

P. McBurney and S. Parsons [2001c]: "Reasoning across scenarios in planning under uncertainty", pp. 85—92 in: C. Gomes and T. Walsh (Editors): *Using Uncertainty within Computation: Papers from the 2001 AAAI Fall Symposium*. Technical Report FS-01-04. Menlo Park, CA, USA: American Association for Artificial Intelligence Press.

P. McBurney and S. Parsons [2002a]: "On hypothesis testing procedures in risk assessment: a research agenda", *Law, Probability and Risk*, **1 (1)**. In press.

P. McBurney and S. Parsons [2002b]: "Formalizing scenario analysis", in: A. Darwiche and N. Friedman (Editors): *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence (UAI-2002)*. San Francisco, CA, USA: Morgan Kaufmann. In press.

P. McBurney and S. Parsons [2002c]: "Using Belief Functions to forecast demand for Mobile Satellite Services", pp. 281—315 in: R. P. Srivastava and T. J. Mock (Editors): *Belief Functions in Business Decisions.* Volume 88 in Studies in Fuzziness and Soft Computing Series. New York, NY, USA: Physica Verlag.

J. McCarthy, O. F. Canziani, N. A. Leary, D. J. Dokken and K. S. White (Editors) [2001]: *Climate Change 2001: Impacts, Adaptation and Vulnerability. Intergovernmental Panel on Climate Change, Working Group II, Third Assessment Report*. Cambridge, UK: Cambridge University Press.

C. E. Schneider [1998]: *The Practice of Autonomy: Patients, Doctors and Medical Decisions.* Oxford, UK: Oxford University Press.

P. Schwartz [1991]: *The Art of the Long View*. New York, NY, USA: Doubleday.

S. E. Toulmin [1958]: *The Uses of Argument*. Cambridge, UK: Cambridge University Press.