Strawmen and eidolons: using argumentation to reason across scenarios

Peter McBurney and **Simon Parsons**
Department of Computer Science
University of Liverpool
Liverpool L69 7ZF United Kingdom
{P.J.McBurney,S.D.Parsons}@csc.liv.ac.uk

Abstract

We propose a dialectical argumentation formalism for qualitative reasoning under uncertainty in a context of alternative scenarios. Our formalism extends prior work representing knowledge uncertainty using dialectical argumentation in participant interaction spaces called Agoras. We define the notion of a scenario in this framework and consider its formal properties. In particular, we ask when can two scenarios be considered the same, and how might we aggregate conclusions from debates conducted under different scenarios. Our formalism is a qualitative, argumentation-based generalization of notions of probability of provability.[1]

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 (see (Carbogim, Robertson, & Lee 2000) for a review). In (McBurney & Parsons 2000), we proposed a formalism using dialectical argumentation for representing and resolving the arguments for and against given propositions. This representation was grounded in specific theories of rational human discourse and was centered on an electronic space for presentation of arguments, which we termed an *Agora*. In subsequent work (McBurney & Parsons 2001b), we extended this formalism and showed that it had several desirable properties when used for inference and decision-making. In this paper, we further extend this framework to enable dialectical argumentation under and between multiple circumstances, or scenarios.

The notion of scenario has found widespread application in business forecasting (Schwartz 1991). In these applications scenarios are usually treated as alternative possible futures, and they usually differ according to the propositions assumed true in each. The differential implications of each scenario are then explored in order to guide the development and selection of business strategy or public policy. This can be valuable for sensitivity-assessment of strategies, even when the scenario is considered to be highly improbable. Used in this way, such scenarios are sometimes called *strawmen*, proposed only to be "burnt down" in discussion. Scenarios can also be used for worst-case contingency planning, in which case they may act to scare management; we call these scenarios *eidolons*, from the Greek for phantom or specter. The use of scenario-based reasoning is also found in scientific debates where agreed knowledge of presumed causal mechanisms may be absent. In these applications, one scenario differs from another not only in the propositions assumed true, but in the causal mechanisms assumed to operate. For example, in the recent debate in bio-medicine over the origins of Bovine Spongiform Encephalopathy (BSE) and its possible links with human Creutzfeldt-Jakob Disease (CJD), several alternative causal mechanisms were proposed and experiments undertaken to test these (Ironside 1999).

Our objective is to develop a framework for reasoning between and across different scenarios, building on our argumentation-based representation for uncertainty. Section 2 presents a brief summary of the previous work on Agoras which is being extended here. Section 3 defines scenarios in the Agora context. Section 4 explores the question of when two scenarios can be considered the same, and we propose a decision rule for determining this. Section 5 then explores reasoning across multiple distinct scenarios, while Section 6 presents an example of our approach. Section 7 concludes the paper with a discussion of related and future work.

[1]This work was partly funded by the British EPSRC under grant GR/L84117 and a PhD studentship.

2 AGORAS

In this section we briefly summarize the Agora framework for the qualitative representation of uncertainty which we presented in (McBurney & Parsons 2000; 2001b). In this framework, arguments for and against claims are articulated by participants in an electronic space, called an *Agora*. Claims are expressed as formulae in a propositional language, typically denoted by lower-case Greek letters. 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 $\mathcal{P}_i$ could demonstrate her argument $\mathcal{A}(\rightarrow \theta)$ supporting a claim θ, an argument to which she was committed with strength D, by means of the locution:

$$\textbf{show\_arg}(\mathcal{P}_i : \mathcal{A}(\rightarrow \theta, D)).$$

The rules governing the use of each permitted locution are expressed in terms of a dialogue-game between the participants (Hamblin 1971). We assume that the Agora participants begin a debate with a set of agreed facts, or assumptions, and a set of inference rules. Because we want to model many forms of reasoning, these rules need not be deductive and may themselves, in our Agora formulation, be the subject of argument.

We demonstrated the use of this framework for the representation of uncertainty by defining a set of uncertainty labels, which are 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 argumentation and used the set: {*Accepted, Probable, Plausible, Supported, Open*}, with the elements listed in decreasing order of certainty. For example, a claim was regarded as *Probable* at a particular time if at least one consistent argument had been presented for it in the Agora by that time, but no arguments for its negation (rebuttals) nor for the negation of any of its assumptions (undercuts) had been presented by then. We defined a claim as *well-defended* if there was an argument for it and any rebuttals or undercuts were themselves subject to counter-rebuttals or -undercuts. *Accepted* claims were defined as those which are well-defended.

We then defined the truth-valuation of a claim θ at time t, denoted $v_t(\theta)$, to be 1 if θ had the label *Accepted* at this time, otherwise it was 0. Such a valuation summarizes the knowledge of the community of participants at the particular time, since it incorporates, via the definitions of the labels, all the arguments for and against θ 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.[2] The sequence $(v_t(\theta) \mid t = 1, 2, \ldots)$ may or may not converge as $t \rightarrow \infty$. Suppose that it does converge, and denote its limit value by $v_\infty(\theta)$. What will the value of a snapshot taken at time t, namely $v_t(\theta)$, tell us about $v_\infty(\theta)$? Of course, any finite snapshot risks being overtaken by subsequent information or arguments, we 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: (a) the snapshot is taken at a time after commencement sufficient for all the arguments using the initial information to have been presented, and (b) there is a bound on the probability that new information arises following the snapshot. This result is proved as Proposition 7 of (McBurney & Parsons 2001b), which we reproduce here. For this, we first need some definitions.

Definition 1: *We write* LE_θ *for the statement:* "The function $v_t(\theta)$ converges to a finite limit as $t \rightarrow \infty$." *We also write* $\mathcal{X}_{t,\theta}$ *for the statement:* "New information relevant to θ becomes known to an Agora participant after time t."

In general, at any time s, we do not know whether new evidence will become available to Agora participants at a later time t or not. Consequently, the variables $\mathcal{X}_{t,\theta}$, for t not in the past, represent uncertain events. Also uncertain for the same reason are statements concerning the future values of $v_t(\theta)$ for any θ. Because these events are uncertain, we assume the existence of a probability function over them, i.e. a real-valued measure function mapping to $[0, 1]$ which satisfies the axioms of probability.

Definition 2: $Pr(.)$ *is a probability function defined over statements of the form* $\mathcal{X}_{t,\theta}$ *and statements concerning the values of* $v_t(\theta)$*, for any formula* θ.

Proposition 1: [Proposition 7 of (McBurney & Parsons 2001b)] *Let* θ *be a formula and suppose that all arguments pertaining to* θ *and using the information available at commencement are articulated by participants by some time* $s > 0$. *Suppose further that* $v_{t_m}(\theta) = 1$ *for some* $t_m \geq s$. *Also, assume that* $Pr(\mathcal{X}_{t_m,\theta}) \leq \epsilon$, *for some* $\epsilon \in [0, 1]$. *Then the following inequalities hold:*

$$Pr(LE_\theta \text{ and } v_\infty(\theta) = 1 \mid v_{t_m}(\theta) = 1) \geq 1 - \epsilon.$$

$$Pr(LE_\theta \text{ and } v_\infty(\theta) = 0 \mid v_{t_m}(\theta) = 1) \leq \epsilon. \qquad \square$$

[2] Strictly, we are assuming throughout that time in the Agora is discrete, and can be represented by a countably infinite set.

Like the standard procedures for statistical hypothesis testing, this proposition provides us with some confidence in our use of finite snapshots to make inferences about the long-run truth-valuation function for a debate. While such inference is not deductively valid, at least its likelihood of error may be bounded.[3] In the sections below, we will be comparing the results of debates in more than one Agora. We therefore assume that we have a single probability function Pr defined across all the relevant variables. We will also index symbols with superscripts (^{1},^{2}, etc) to denote the Agora to which they refer. The next Section defines the concept of Scenario.

3 SCENARIOS

The framework we have just outlined provides a means to represent the diverse arguments that may arise from a given set of assumptions, and using 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 a different set of inference rules, the arguments presented in the Agora may well be different. As a result, the uncertainty labels and truth values assigned to formulae may well also be different, both when taken at finite snapshots and in the limit. We begin with some definitions:

Definition 3: *A* ***Scenario*** *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 formulae in that domain. We denote scenarios for a given domain by $\mathcal{S}^1, \mathcal{S}^2, \ldots$, etc. For each scenario, $\mathcal{S}^i$, an Agora debate undertaken with the assumptions and inference rules of that scenario, is said to be the* ***associated Agora****, denoted $\mathcal{A}^i$. We assume only one Agora debate is conducted in association with any scenario.*

In this paper, we will be assuming that all scenarios, and all the resulting Agora debates, relate to the same application domain. For this domain, suppose we are interested in a particular proposition θ. We imagine we have a number of scenarios in parallel, each with a different set of starting assumptions and possibly also different inference mechanisms. We now allow the associated Agora debates to proceed up to a certain time t, when we take a finite snapshot of each debate. It would be expected that the truth status of θ would be different under different scenarios. Not only are the assumptions and inference mechanisms different, but not all arguments may have been presented to each Agora debate at the time of the snapshot. We are thus faced with the question: *Given these different truth assignments under different scenarios, what overall truth-label should be assigned to θ?* Our problem is thus one of aggregation across multiple scenarios, and in Section 5 we present a method of doing this. We permit scenarios to be weighted differentially, for example, according to their relative importance, their likelihood, or their feasibility, etc. However, to aggregate results arising from multiple scenarios, we first need to ensure that each distinct scenario is only counted once, i.e. that no "double-counting" of scenarios takes place. In other words, we need a means to decide whether two scenarios are the same or not. We consider this question next.

[3] One may object that we can never know the value of ϵ. While this is true, participants in a debate are often quite willing to provide subjective estimates for such probability bounds. Scientists, for example, will often estimate the chance that new information will arise which overturns an established theory.

4 COMPARING SCENARIOS

When are two scenarios the same? Obviously, we may consider them to be the same when they have identical sets of assumptions and inference rules. But two scenarios identical in this fashion may result in very different Agora debates, as different arguments may be presented in each, or the same arguments presented at different times. It is not therefore clear that identical scenarios will lead to identical assignments of truth-labels, even over the long-run; we show that, under certain conditions, they will do so. Throughout this section $\mathcal{S}^1$ and $\mathcal{S}^2$ will be two scenarios of interest, and $\mathcal{A}^1$ and $\mathcal{A}^2$ their associated Agora debates.

Proposition 2: *Let θ be a formula. Suppose that $\mathcal{S}^1$ and $\mathcal{S}^2$ are identical scenarios, i.e. they have identical sets of assumptions and identical sets of inference rules. Suppose that in the corresponding Agora debates, $\mathcal{A}^1$ and $\mathcal{A}^2$, all possible arguments based on the initial assumptions and using the inference rules are eventually articulated. Suppose further that no new information is presented to either debate following commencement. Then, the long-run truth status of θ in each debate is the same.*

Outline of Proof.[4] Given the premises, the only way the two debates will potentially differ will be in the order that arguments are articulated. It then follows, from the definitions, that the long-run truth status of a claim does not depend upon the order of presentation of the arguments for and against it. □

If we relax the assumption that no new information arrives in either debate our conclusion acquires a probabilistic qualification. While this does not guarantee that two identical scenarios always lead to identical truth assignments, it does bound the likelihood that this is not the case.

Proposition 3: *Let θ, $\mathcal{S}^1$ and $\mathcal{S}^2$ be as before. Suppose there exist upper bounds $\epsilon_i \in [0,1]$ for the probability that new information arrives after commencement in debate i, i.e. that $Pr(\mathcal{X}^i_{0,\theta}) \leq \epsilon_i$, for $i = 1, 2$. Then,*

$$Pr(v^1_\infty(\theta) = v^2_\infty(\theta)) \geq 1 - \epsilon_1 - \epsilon_2.$$

Outline of Proof. By the previous result, the two long-run

[4] Full proofs are contained in (McBurney & Parsons 2001c).

assignments of truth to θ are only different if one or other debate receives new information. The probability that this occurs is less than or equal to the sum of the probabilities that either debate receives new information less the probability that they both do. This latter event has probability greater than or equal to zero, and the inequality follows by algebraic manipulation. $\square$

Alternatively, we may have two different scenarios, whose differences are only small. For example, we might consider the world as it is today compared with what it might have been had the Julian calendar, adopted by the Roman empire in 46 BCE and then by the Catholic church, had been different by a few days. This is only a small change, and, one might think, unlikely to have any large consequences. But, one reason for the rapid success of the Spanish invaders of Mexico led by Cortes in 1519 was that they landed in Veracruz on Good Friday. Consequently Cortes was dressed in black, which confirmed a prediction of the Aztec mythology that they would be visited by a destructive God dressed in black. A few days difference in the calendar may have resulted in different clothes for the invaders, thereby invalidating a key prediction of the Aztec mythology; this may have led to the Aztecs being less willing to submit to Spanish domination. Therefore, to determine whether two nonidentical scenarios should be considered as the same or not, one needs to consider not only the magnitude of any differences in assumptions and inference rules, but also their importance, relative to the domain and to any question of interest. We combine both these elements in the following definition:

Definition 4: *Let $\mathcal{S}^1$ and $\mathcal{S}^2$ be non-identical scenarios, i.e. either their assumption sets and/or their sets of inference rules are not identical. We say that the difference between these two scenarios is* major *if the differences between assumption sets and/or inference rules is large (relative to the size of the sets concerned) or important (relative to the domain or to any formula of interest). Otherwise, the differences between $\mathcal{S}^1$ and $\mathcal{S}^2$ is said to be* minor.

Of course, deciding whether the differences between scenarios is major or minor will depend on context — the domain, the question of interest, and the assumptions and inference rules – and will be subjective. In this paper we do not propose a rule for making this determination, but assume simply that the differences between any two scenarios can be classified as one or the other. However, this is not sufficient to decide whether the two scenarios should be considered the same. Even small differences between assumption-sets or inference rules may lead to critical differences in the Agora debates conducted using those assumptions and rules. To decide whether two scenarios should be treated as the same or not, therefore, we need to compare the outcomes of the Agora debates conducted under them. We consider this problem next.

4.1 Comparing Agora debates

How may we compare two Agora debates? We can think about this in the light of what we may feasibly observe about a debate. Firstly, we could take a snapshot at a particular finite time after commencement. Or, secondly, we could take a number of such snapshots. Or, thirdly, we could examine the actual arguments used in a debate from commencement up to a particular time. For each of these three methods, we could observe both debates and compare the results. We consider the first two of these approaches in the sub-sections which follow. Comparing the arguments presented in two debates is, we believe, quite complex, and so we leave this approach for another occasion. As before, we denote the long-run truth status of a formula θ in debate i, if this limit exists, by $v^i_\infty(\theta) = \lim_{t\to\infty} v^i_t(\theta)$. The sub-sections which follow will discuss finite estimators of this long-run value, estimators we denote by $\hat{v}^i_\infty(\theta)$.

4.1.1 A single snapshot

We might compare two Agora debates, undertaken under different scenarios, by taking a finite snapshot of each at a particular finite time-point t after their commencement. Each snapshot, $v^i_t(\theta)$, will give us an assessment of the long-term truth-status of a claim θ in each debate $\mathcal{A}^i$, for $i = 1, 2$. That is, we set $\hat{v}^i_\infty(\theta) = v^i_t(\theta)$. Proposition 1 tells us that we can bound the probability of error in infering from the finite estimate $\hat{v}^i_\infty(\theta)$ to the true infinite limit value $v^i_\infty(\theta)$.

The rule we present in the next section will compare the estimates for the two debates, asking whether $\hat{v}^1_\infty(\theta) = \hat{v}^2_\infty(\theta)$. As with the individual estimates, an answer to this question based on a finite snapshot is not necessarily also true for the long-run. Can we also bound the probability of error when inferring from a comparison of the finite values? The next proposition provides such bounds.

Proposition 4: *Let θ be a formula, and let $\mathcal{A}^1, \mathcal{A}^2$ be two Agora debates associated with scenarios $\mathcal{S}^1, \mathcal{S}^2$ respectively. Suppose that, in debate $\mathcal{A}^i$ ($i = 1, 2$), all arguments pertaining to θ and using the information available at commencement are articulated by participants by some time $s_i > 0$. Suppose further that there is a time $m \geq max(s_1, s_2)$, such that $v^1_m(\theta) = v^2_m(\theta)$. Also, assume that there exist $\epsilon_1, \epsilon_2 \in [0, 1]$, possible dependent on m, such that each $Pr(\mathcal{X}^i_{m,\theta}) \leq \epsilon_i$. Then the following four inequalities hold:*

1. $Pr(LE^1_\theta \text{ and } LE^2_\theta \text{ and } v^1_\infty(\theta) = v^2_\infty(\theta) \mid v^1_m(\theta) = v^2_m(\theta)\ ) \geq 1 - \epsilon_1 - \epsilon_2$.
2. $Pr(LE^1_\theta \text{ and } LE^2_\theta \text{ and } v^1_\infty(\theta) \neq v^2_\infty(\theta) \mid v^1_m(\theta) = v^2_m(\theta)\ ) \leq \epsilon_1 + \epsilon_2$.

3. $Pr(LE^1_\theta$ *and* LE^2_θ *and* $v^1_\infty(\theta) = v^2_\infty(\theta) \mid v^1_m(\theta) \neq v^2_m(\theta)) \leq \epsilon_1 + \epsilon_2$.

4. $Pr(LE^1_\theta$ *and* LE^2_θ *and* $v^1_\infty(\theta) \neq v^2_\infty(\theta) \mid v^1_m(\theta) \neq v^2_m(\theta)) \geq 1 - \epsilon_1 - \epsilon_2$.

Proof. Arguments similar to Proposition 3. □

In words, using simultaneous finite snapshots of two debates to make an inference about the long-run truth-status of a formula is a process prone to error. This result says that, under certain circumstances, we can bound the probability of such errors. The "certain circumstances" relate to the timing of the snapshots — which must be long enough into the two debates for all the arguments based on the initial information to have been presented — and to the probabilities of new information being presented to each debate subsequent to the snapshots being taken. As one would expect, the error bounds are functions of these probabilities.

In proving this result, we have not assumed that the event of new information being presented to one Agora debate is independent of new information being presented to the other. If we were able to make such an assumption, our error bounds would be tighter, with $\epsilon_1\epsilon_2$ added to the right-hand side of the first and fourth inequalities and subtracted from the second and third. Independence of these two events is a function of how "distinct" are the two scenarios. Scenarios in the same domain which are very similar are likely to experience new information concurrently; participants in the corresponding debates are also likely to make similar assessments of the relevance of such new information.

4.1.2 Multiple snapshots

Suppose now that we have multiple snapshots of each debate, i.e. for each $\mathcal{A}^i$, $i = 1, 2$, we have a sequence of observations of the truth-status of θ: $v^i_1(\theta), v^i_2(\theta), v^i_3(\theta), \ldots, v^i_n(\theta)$. These values are all either zero or one, and each is an estimate for the long-run truth status $v^i_\infty(\theta)$ of θ, as in the subsection above. Given such multiple estimates, there are a number of ways we could combine them to produce a single estimate of $v^i_\infty(\theta)$, for example: (a) **The mean**, which is the sum of the observed snapshot values, divided by n; (b) **An α-trimmed mean**, which is the mean calculated after deleting α % of the observations (Huber 1981). For instance, we may delete those at the beginning of the sequence, on an assumption that early values of $v^i_t(\theta)$ will oscillate as all the relevant arguments are presented to the Agora; (c) **the mode**, the most common value, i.e. whichever of 0 or 1 appears most frequently in the n observations; etc.

However, each truth-valuation v^i_k is defined in terms of the arguments presented to the Agora up to time k, so, in some sense, each observation summarizes all the information relevant to θ up to and including the time the observation was made. We should therefore expect the final observation, v^i_k to contain the most information, and so to be the best estimator (in some sense) of the long-run value, v^i_∞.[5] This is indeed the case, as the following theorem shows. For simplicity, we omit θ and i from the notation.

Proposition 5: *Let $v_1, v_2, v_3, \ldots, v_n$ be a sequence of n snapshot values concerning θ taken from a debate $\mathcal{A}^i$. Suppose the limit $v_\infty = \lim_{n\to\infty} v_n$ exists. Let $\hat{v}_n$ be the estimator of v_∞ using only the final observation in such a sequence. Further, let $\hat{v}^*_n$ be any estimator of v_∞ based on these n observations which converges to a finite limit as $n \to \infty$. Then:*

$$\lim_{n\to\infty} Pr(|\hat{v}_n - v_\infty| \leq |\hat{v}^*_n - v_\infty|) = 1.$$

Proof. We prove this by contradiction. If the result does not hold, then there must exist $\epsilon > 0$ such that:

$$\lim_{n\to\infty} Pr(|\hat{v}_n - v_\infty| > |\hat{v}^*_n - v_\infty|) = \epsilon$$

Then, there must be infinitely many m such that $|\hat{v}_m - v_\infty| > |\hat{v}^*_m - v_\infty|$. Since the sequence of final values v_n is a sequence of zeros or ones, and it converges to v_∞, then for each of these m, we have one of two cases:

$v_\infty = 1$**:** By the strict inequality, we must have $\hat{v}_m = 0$.

$v_\infty = 0$**:** Likewise, we must have $\hat{v}_m = 1$

But this happens for infinitely many values m, which contradicts the assumption that v_n converges to v_∞. □

This result shows that using the final observation of a sequence of snapshots is at least as good as using any other convergent estimator of v_∞ based on this sequence. We have therefore reduced the multiple snapshots case to the single case of the subsection above.[6]

4.2 Are Two Scenarios the Same?

We now integrate the results of the previous subsections, to produce a decision rule for determining if two scenarios $\mathcal{S}^1$ and $\mathcal{S}^2$ are the same or not. This decision rule classifies scenarios into two classes, labeled *distinct* and *non-distinct*. In Section 5, we will present rules for combining the results of debates across Scenarios, and these will require that all

[5] Note that the final observation is itself a trimmed mean.

[6] One could also view each sequence as arising from a random process, and so view the decision problem as a test of an hypothesis that the two sequences are governed by the same probability distribution. Because the form of the distributions is not specified, the appropriate test would be nonparametric, e.g. the Kolmogorov-Smirnov two-sample test (Gibbons 1985). However, the asymptotic theory for even these tests requires that the underlying distributions be continuous and that the two samples be drawn independently. Neither assumption is appropriate here.

the scenarios are *distinct*. The rule proposed for determination of distinctness of scenarios uses four criteria (in order of application): (a) whether or not the two scenarios have identical assumptions and inference rules; (b) in the case where they do, whether or not either scenario is judged to have a high probability of receiving new information; (c) in the case where the two scenarios do not have identical assumptions and inference rules, whether their difference is judged to be *major* or *minor*; and (d) in the case where it is *minor*, whether the estimated long-run truth values of the two Agora debates are the same or not.

Case 1A: $\mathcal{S}^1 = \mathcal{S}^2$ **and** $Pr(\mathcal{X}^1_{m,\theta}), Pr(\mathcal{X}^1_{m,\theta})$ **both small.** In this case, the likelihood of new information arising in either scenario is small and Proposition 3 allows us to infer that $v^1_\infty(\theta) = v^2_\infty(\theta)$. Conclude that the two scenarios are *non-distinct*.

Case 1B: $\mathcal{S}^1 = \mathcal{S}^2$ **and one or both of** $Pr(\mathcal{X}^1_{m,\theta}), Pr(\mathcal{X}^1_{m,\theta})$ **large.** In this case, the likelihood of new information arising in at least one scenario is large, and thus is it is unlikely that $v^1_\infty(\theta) = v^2_\infty(\theta)$. Conclude that the two scenarios are *distinct*.

Case 2A: $\mathcal{S}^1 \neq \mathcal{S}^2$**, with the difference *major*.** Conclude that the two scenarios are *distinct*.

Case 2B-1: $\mathcal{S}^1 \neq \mathcal{S}^2$**, with the difference *minor*, and** $\hat{v}^1_\infty(\theta) = \hat{v}^2_\infty(\theta)$**.** Conclude that the two scenarios are *non-distinct*.

Case 2B-2: $\mathcal{S}^1 \neq \mathcal{S}^2$**, with the difference *minor*, and** $\hat{v}^1_\infty(\theta) \neq \hat{v}^2_\infty(\theta)$**.** Conclude that the two scenarios are *distinct*.

In the first case (Case 1A), where the underlying assumptions and inference rules are the same in the two scenarios, Proposition 2 says that the long run truth assignments for θ in the corresponding Agora debates, if they exist, will be identical, provided no new information is presented in either Agora debate following commencement. Even if new information is possible, if it is unlikely, then it is likely that the two truth assignments will still be equal, by Proposition 3. If such new information in either scenario is likely, however, we treat the two scenarios as *distinct* (Case 1B). In the other three cases, the underlying assumptions and inference rules are not the same. Our assessment of whether the scenarios are distinct or not depends on the size and importance of the differences between the two scenarios. Any difference which is *major* (Case 2A) leads to a determination that the two scenarios are *distinct*. If the differences are only *minor*, however, the determination depends on whether the estimated long-run truth values of the formula of interest are equal or not. If these two estimates are equal (Case 2B-1), the scenarios are determined to be *non-distinct*, whereas if they are not (Case 2B-2), the scenarios are determined to be *distinct*. These rules aim to capture the understanding that long-run truth assignments in an Agora debate may be sensitive to small changes in the assumptions or inferences rules under which the debate was conducted. Our only means of assessing this sensitivity, however, is to consider debates conducted under similar scenarios, and estimate the long-run truth assignments in each. If such debates produce the same estimated values, the truth assignment is not likely to be sensitive to the initial assumptions or inference rules.

5 REASONING ACROSS SCENARIOS

Now that we can distinguish scenarios from each other, we present a formalism for considering claims across multiple scenarios. We assume we have a finite set $\mathcal{S}$ of m of distinct scenarios for the one domain, $\{\mathcal{S}^1, \mathcal{S}^2, \ldots, \mathcal{S}^m\}$, each $\mathcal{S}^i$ with an associated Agora debate $\mathcal{A}^i$. As before, we are interested in the dialectical-argumentation status of a formula θ in the domain, but across all m debates, not simply in one. We assume that, associated with each scenario $\mathcal{S}^i$ is a real-number, $a^i \in [0, 1]$, called its scenario weight. We call $\tilde{a} = (a^1, a^2, \ldots, a^m)$ the scenario weights vector. W e do not assume the weights sum to unity across the m scenarios, although they may do so, in which case the weights could be probabilities. Note that the weights may vary with time, but we assume that their assignment to scenarios is independent of the dialectical status of claims in the corresponding debates. We discuss possible meanings for these weights below. We next define the *weighted global measure* of a claim at a particular time.

Definition 5: *The* ***weighted global measure*** *for θ at time t relative to a set $\mathcal{S}$ of distinct scenarios, is defined by:*

$$a^{\mathcal{S}}_t(\theta) = \frac{\sum_{i=1}^m a^i v^i_t(\theta)}{\sum_{i=1}^m a^i}.$$

This measure is an assessment of the dialectical status of a claim, adjusted according to the weights assigned to those scenarios in which the claim is well-defended in the associated debates. It is indexed by t and by $\mathcal{S}$ because its value may change with time and with a different set of scenarios. This measure has the following properties:

Proposition 6: *Given $\theta, \mathcal{S}$ and $\tilde{a}$ as above, the the weighted global measure, $a^{\mathcal{S}}_t(\theta)$ satisfies, for all $\theta, \mathcal{S}$ and t:*

1. $0 \leq a^{\mathcal{S}}_t(\theta) \leq 1$

2. $0 \leq a^{\mathcal{S}}_t(\theta) + a^{\mathcal{S}}_t(\neg\theta) \leq 2$*. In particular, it is not necessarily the case that:* $a^{\mathcal{S}}_t(\theta) = 1 - a^{\mathcal{S}}_t(\neg\theta)$.

Outline of Proof. Property 1 follows from the definition of $a^{\mathcal{S}}_t$. The other properties follow from the assignment of value 1 to $v^i_t(\theta)$ when θ is well-defended, and the possibility that both θ and $\neg\theta$ may be well-defended at the same time t in $\mathcal{A}^i$. □

What interpretation we give this measure depends upon the meanings applied to the logical language, to the scenarios and their weights, and to arguments for claims in the corresponding Agora debates. Three examples are: (a) 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 occurence; (b) The scenarios may 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 represent different allowable modes of reasoning, such as reasoning by analogy or from authority. The weights may represent the extent of compliance with some set of principles of rational discourse, or with some normative economic theory; (c) The scenarios could represent 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 consequence or utility.

In a typical scientific or business application, our interpretation is closest to the last example, with the scenario weights often being measures of likelihood of the scenario's occurrence. In this case, if the scenarios are mutually exclusive and comprehensive of the possible outcomes of the domain, then it would be reasonable to expect the weights to sum to unity across the m scenarios. Here, then, the weights could correspond to probabilities. In such a case, we would be interested in three questions: *Is θ possible? Is θ probable? Is θ inevitable?* We provide a means to answer these questions in terms of the arguments presented in the various Agora debates associated with the set of scenarios. At any one time-point t, we can take snapshots of each of the m debates and thus obtain an estimate $v_t^i(\theta)$ of the truth-status of θ in each Agora $\mathcal{A}^i$.

Definition 6: *Relative to a set $\mathcal{S}$ of distinct scenarios, we say that θ is **possible** at time t when $v_t^i(\theta) = 1$ for at least one Agora debate $\mathcal{A}^i$. We say that θ is **inevitable** at time t when $v_t^i(\theta) = 1$ for every Agora debate $\mathcal{A}^i$, $i = 1, \ldots, m$. Further, given a vector $\tilde{a}$ of scenario weights interpreted as likelihoods, we say that θ is **probable** (in the sense of being "more likely than not") at time t when θ is possible and the sum of the uncertainty weights assigned at time t to those scenarios $\mathcal{S}^j$ in which $v_t^j(\theta) = 1$ exceeds 0.5, i.e. whenever: $a_t^{\mathcal{S}}(\theta) > 0.5$.*

Proposition 6 will hold even when the scenario weights are probabilities, because of the argumentation-based definitions given to the truth-values $v_t^i(\theta)$ (McBurney & Parsons 2000; 2001b). Different properties for the weighted global measure $a_t^{\mathcal{S}}$ could arise with a different definition for v_t^i, e.g. one which precluded the possibility that both θ and $\neg\theta$ could be well-defended in the same Agora debate at the same time. However, a primary aim in our work is to model actual scientific or business strategy debates, a feature of which is that strong evidence may exist, at any one time, simultaneously for a claim and for its negation, or may not be present for either.

6 EXAMPLE

Given space limitations, our example is very simplified, and illustrates only the core aggregation idea. We consider the situation facing an intending operator of global mobile satellite-based telecommunications services (GMSS) in 1990 (McBurney & Parsons 2001d). Demand for these services was predicted to depend heavily on the extent to which terrestrial mobile communications services would expand, both in terms of customer numbers and the geographic area under coverage. One could imagine a number of scenarios for the future, under each of which there would be arguments for and against the claim that demand for GMSS would be large.

Scenario 1: Terrestrial mobile services expand rapidly and customers wish to use their phone everywhere. Argument: Large numbers of terrestrial customers leads to high demand for GMSS outside terrestrial coverage.

Scenario 2: Terrestrial mobile services expand rapidly and customers are happy with the terrestrial coverage. Argument: Large geographic coverage for these services leads to low demand for GMSS, as most areas have terrestrial coverage.

Scenario 3: Terrestrial mobile services do not expand. Argument: Small geographic coverage means high demand for GMSS.

Recall that scenario weights are assigned independently of the arguments under each, and assume these three scenarios are assigned the likelihood weights of $0.7, 0.7$ and 0.3, respectively. Let θ be the claim: *"GMSS experiences high demand."* Then, by Definition 6, we can say that θ is possible, but is not inevitable. Moreover, because $a_t^{\mathcal{S}}(\theta) = (0.7 + 0.3)/(0.7 + 0.7 + 0.3) = 0.59$, then we can say that θ is probable. This example has not illustrated the working of the argumentation apparatus within the Agora debates under each scenario, but has simply shown the basic mechanism for aggregation across scenarios.[7]

[7] The reality was that the explosive growth of terrestrial mobile services witnessed over the last decade did not lead to high

7 DISCUSSION

Building on prior work in using dialectical argumentation as a qualitative representation of uncertainty, we have presented a novel framework for comparing multiple scenarios and aggregating across debates conducted under these different scenarios. Our previous work had articulated a set of dialogue-game rules for the conduct of debates over scientific or other domains, drawing on theories of rational human discourse. In the framework presented here, scenarios are defined in terms of the assumptions and inference rules available to participants in such debate. We have demonstrated that these debates may be adequately compared on the basis of single snapshots taken at the most recent time, and that this comparison is at least as good (in a specific sense) as also using earlier snapshots. We have used this result to propose a rule for classifying two scenarios as distinct or not. With this rule, we are then able to aggregate the results of different debates while ensuring these do not arise from identical or closely-similar scenarios. Given a relative weighting of a set of distinct scenarios, e.g. their subjective probabilities of occurrence, we have defined a global measure $a_t^S(\theta)$ for a formula θ which is based on the argumentation-status of θ in each of the debates arising from the set $\mathcal{S}$ of scenarios at the time t of the status-assessment.

If the scenario weights are probabilities, then we can view $a_t^S(\theta)$ as the probability that a well-defended argument exists for θ under one or more scenarios at time t, i.e. that an argument exists for θ and any rebuttals or undercuts are themselves subject to counter-rebuttals or -undercuts. Thus, the weighted global measure $a_t^S(\theta)$ provides an argumentation equivalent to the notion of "probability of provability" (Pearl 1988). Here, "provability" of a claim corresponds to saying the claim *"has a well-defended argument, and hence an argumentation-based truth-value assignment of* 1 (defeasibly true), *in the Agora debate conducted under some scenario"*, and "probability" corresponds to *"the relative weighting of the scenarios in which this is the case"*, for a given weighting of scenarios. Similarly, our approach may be seen as an argumentation analog of the *Ents* model of belief of (Paris & Vencovská 1993), in which an agent's belief in a claim is determined by imagining scenarios in which the claim is decided, either true of false, and then setting the belief in the claim equal to the proportion of scenarios in which it is true.

There are several future paths we wish to tread. One relates to the third observational approach mentioned in Section 4.1, that of comparing the actual arguments advanced in two Agora debates. Here, it may be useful to measure the "distance" between the two sets of arguments, and the metrics used to measure distance between families of probability distributions in asymptotic statistical theory (Kass & Vos 1997) may be relevant to this. A second extension is to integrate the scenario framework with our work developing intelligent systems for public argumentation in environmental domains (McBurney & Parsons 2001a), so that participants to such debates can reason over alternative scientific and regulatory scenarios.

demand for GMSS, a fact which led to the business failures of the main intending providers of GMSS, Iridium, ICO and Globalstar.

References

Carbogim, D. V.; Robertson, D. S.; and Lee, J. R. 2000. Argument-based applications to knowledge engineering. *The Knowledge Engineering Review* 15(2):119–149.

Gibbons, J. D. 1985. *Nonparametric Statistical Inference*. New York: Marcel Dekker, second edition.

Hamblin, C. L. 1971. Mathematical models of dialogue. *Theoria* 37:130–155.

Huber, P. J. 1981. *Robust Statistics*. New York: Wiley.

Ironside, J. W. 1999. nvCJD: exploring the limits of our understanding. *The Biologist* 46(4):172–176.

Kass, R. E., and Vos, P. W. 1997. *Geometrical Foundations of Asymptotic Inference*. New York: Wiley.

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

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

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

McBurney, P., and Parsons, S. 2001b. Representing epistemic uncertainty by means of dialectical argumentation. *Annals of Mathematics and AI* (in press).

McBurney, P., and Parsons, S. 2001c. Using argumentation to reason across scenarios. Technical report, Department of Computer Science, University of Liverpool, U.K.

McBurney, P., and Parsons, S. 2001d. Using Belief Functions to forecast demand for Mobile Satellite Services. In Srivastava, R. P., and Mock, T., eds., *Belief Functions for Business Decisions*. New York: Physica (in press).

Paris, J., and Vencovská, A. 1993. A model of belief. *Artificial Intelligence* 64:197–241.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco: Morgan Kaufmann.

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