VARIETIES OF POSSIBILITY: HOW ITERATED MODALITIES SOLVE A METHODOLOGICAL DILEMMA OF SIMULATING UNDER UNCERTAINTY

GREGOR BETZ (UNIVERSITÄT STUTTGART)

Introduction

There are many things about the future we know for sure. It is, for instance, certain that several million children will die of hunger next year, or that humans will not be able to live on earth once the sun has become a red giant in seven billion years. Other bits of our foreknowledge are, however, not of deterministic, but of probabilistic nature. Thus, it is very improbable that an asteroid hits the earth and extinguishes all forms of advanced life within the next decade—a probability forecast established from astronomical observations as well as geological evidence for similar impacts (Napier 2008). Still, like deterministic prediction, probabilistic foreknowledge faces limitations: There are statements about the future to which we cannot reliably assign probabilities. Section 1 is devoted to defending this thesis in some detail. Where even probabilistic prediction fails, foreknowledge is (at most) possibilistic in kind; i.e. we know some future events to be possible, and some other events to be impossible. This tripartite classification of foreknowledge is nothing but Frank Knight's classic distinction of different epistemic modes (Knight 1921). And I shall adopt Knight's terminology, "certainty—risk—uncertainty", in the following. The classification raises, in particular, the methodological question how to obtain and justify possibilistic statements. What we need is what I suggest to call a "modal methodology". How does one rationally reason in the epistemic mode of uncertainty? I will argue in section 2 that this is anything but a trivial question—and that it actually gives rise to a methodological dilemma. Solving this dilemma is the main purpose of my paper.

This paper's central theme can be approached from different angles, the previous paragraph being just one among others. I shall try to highlight some of these perspectives in this introductory section, spelling out how a modal methodology relates to other, well-studied philosophical issues.

So, an alternative way to see the urgency of clarifying how to establish possibility statements is by considering the debate about the precautionary principle. One argument against what Stephen Gardiner (Gardiner 2006) identifies as the core precautionary principle, basically Rawls' maximin rule, criticises that this very principle gives "mere possibilities" too important a weight in policy consideration (Manson 2002). But responsible decision making, the so-called mere-possibility-argument says, surely shouldn't rely on pure fiction: Although we can *imagine* that hot chocolate causes brain cancer, this is by no means relevant for health and food policies, or is it? Gardiner, in defence of the precautionary principle, rightly notes that (i) the application of the precautionary principle requires a range of "realistic possibilities" to be established, and that (ii) this is required by any principle for decision making under uncertainty whatsoever. In other words, he stresses the need for a modal methodology.

The thoughts unfolded in this paper are closely related to another point: One might worry whether Knight's tripartite classification is really exhaustive. Two issues give rise to this general worry. *First*, Knight's classification seems to be very coarse. In particular, the gap between the epistemic modes of risk and uncertainty appears to be rather wide. This impression probably drives some scientists to make probabilistic forecasts by all means, hence, they might reason: "Well, I know for sure more about the system than mere possibilities. Maybe my understanding is not enough, strictly spoken, to establish reliable probability forecasts—but if that's the only alternative ... And surely, I'm supposed to express *all* I know about the system." The question is, of course: Is there really no intermediate mode between probabilistic and possibilistic prediction? Important conceptual work in epistemology and decision theory has been devoted to this question; and the theory of imprecise probabilities (e.g. Levi 1980; Walley 1991) can be considered as one intermediate epistemic mode, extending Knight's classification. *Second*, one might hold Knight's classification to be limited with regard to its extreme: Is there really no other mode of reasoning in which we know less than in the mode of uncertainty? It has been suggested, notably by scholars in environmental sciences (e.g. Wynne 1993; Faber, Manstetten et al. 1996; Healy 1999), to extend Knight's classification by a fourth mode, namely ignorance. Accordingly, "uncertainty" refers, more precisely, to a situation where all possibilities (but no probabilities) are known, whereas in the epistemic mode of ignorance one's knowledge of future possibilities is incomplete, i.e. some possibilities are ignored. This paper's proposal as to how to solve the methodological dilemma presented in section 2 takes into account these shortcomings of Knight's original classification. It amounts to modifying Knight's distinction and providing us a conceptual framework for expressing non-probabilistic, possibilistic knowledge in a more differentiated and nuanced way. Clearly, this more fine-grained conceptual framework raises the same methodological question as the Knightian one, i.e.: How do we come up with and justify the respective knowledge claims under uncertainty? This leads, as will be shown in section 4, to a plurality of modal methods.

Yet another perspective from which the epistemological reflections of this paper can be motivated starts from an observation about scientific reasoning. In some scientific disciplines, especially those which try to understand and to predict complex systems, scientists work simultaneously with different and incompatible models of one and the same system without being able to rank these models according to significant epistemic criteria. They face model-underdetermination. This triggers the question as to the status of their results. E.g., does model-underteremination compel one to conditionalise every prediction relative to the specific model which was used to generate it? Or is there a way of interpreting these results in a more general and unconditionalised way? Can simulations based on different, conflicting models show that some statement is possible, or even probable? The specific science I had in mind when developing these thoughts was climate science, where, clearly, model-underdetermination reigns (Betz 2009a). The question of how we adequately express and characterise our knowledge about the climate's future becomes ever more pressing given the policy issues involved. However, my hypothesis is that model-underdetermination represents, at least in so far complex systems are the object of scientific inquiry, such as in earth sciences (Oreskes, Shrader-Frechette et al. 1994), the rule rather than the exception. Thus I will illustrate my arguments with examples from two different disciplines: climate science, in particular the prediction of anthropogenic climate change (IPCC 1990, 1996, 2001, 2007), on the one hand, and geology, specifically the study of carbon storage in geological reservoirs by computer simulations (IPCC 2005), on the other hand. Like climate science, the latter represents a discipline, where conflicting models and simulations are used side by side (see Gaus, Audigane et al. 2008).

1. (There are) Boundaries of probabilistic foreknowledge

Modern science can hardly be imagined without probability. From quantum mechanics, through statistical treatment of large ensembles of individuals—be it in statistical mechanics or in econometrics, to statistical tests of deterministic hypotheses: Probabilistic methods seem to be irreplaceable in today's scientist's methodological toolbox. And arguably, these methods are highly successful, allowing us in particular to better understand and assess the uncertainties we are facing, and thus to take more effective public or private decisions. Whenever probabilistic knowledge is available, it would be straight forward irresponsible not to make use of it.

This undeniable and ubiquitous success of probabilistic methods might spur the hope that they are universally applicable. That is, whenever we are unable to establish reliable deterministic results, we will at least get probabilistic statements. Likewise, (O'Hagan and Oakley 2004) write:

> In principle, probability is uniquely appropriate for the representation and quantification of all forms of uncertainty; it is in this sense that we claim that 'probability is perfect'. (p. 239)

[Lee Clarke -> probabilism, Larry Laudan]

What explains, besides successful applications, that probabilistic methods are thought to be universally applicable? The rise of Bayesianism, which licenses the attribution of (subjective) probabilities to every statement whatsoever, might be one reason for the very belief that 'probability is universal'; a second one could lie in a growing awareness that knowledge is sometimes implicit, and that expert judgement, including probability estimates, can be reliable without the expert being able to give an explicit justification of his judgement—facts the philosophical and scientific community has become increasingly aware of in the 2nd half of the twentieth century.

This section's aim is simply to remind us that that is not so. There is no reason to believe that 'probability is universal', or "perfect", even not "in principle". We cannot justifiably or reliably assign probabilities to every sentence which is of importance to us. One case in point are climate predictions. Here, Bayesian methods have been applied to generate probabilistic climate forecasts (cf. Webster, Forest et al. 2003; Hegerl, Crowley et al. 2006). The ultimate reason why these fail in providing justified probabilistic climate predictions, which don't hold merely conditional to a certain model, is that the posterior probabilities still significantly depend on the (arbitrary) prior. And expert elicitation, which has also been applied in climatology to establish probability forecasts (e.g. Zickfeld, Levermann et al. 2007), leads us nowhere, either. Because, in spite of being experts for many things, climate scientists are definitely not well-trained and experienced experts for future, unprecedented anthropogenic climate change: 21st century climate change is a singular event. We have here a major dis-analogy to the well-trained medical doctor who has, during his career, internalised a tremendous amount of information *regarding similar cases* on which he can implicitly found his informed judgment. Where similar cases, and thence a track record of relevant experiences, are lacking, there is absolutely no reason—it would even be entirely naïve—to accept a merely implicit judgement of a so-called expert. Unfortunately, the fourth IPCC assessment report made use of probabilistic predictions (IPCC 2007). I think this was a mistake (Betz 2007). However, climate scientists are critically discussing the scope of probabilistic methods, assessing, with regard to their own domain, the limits of probabilistic knowledge. So do, for example, the organisers of the climateprediction.net–Project in two recent articles, whose positive proposal for how o interpret climate simulation results will be discussed in section 5 (Stainforth, Allen et al. 2007; Stainforth, Downing et al. 2007). The authors are making a succinct and clear argument against probabilistic projections of anthropogenic climate change. They see, rightly, that this is required to safeguard the credibility of climate science, and its policy advice. As the case of climate prediction makes clear: Accepting the limits of probabilistic methods and refusing to make probabilistic forecasts where those limits are exceeded, originates, ultimately, from the virtue of truthfulness, and from the requirements of scientific policy advice in a democratic society.

2. Justifying possibilistic statements: the dilemma

Because our probabilistic knowledge is limited, i.e. because many systems cannot be described and predicted probabilistically in a reliable way, Knight's third category, possibilistic foreknowledge, is not simply swept by the probabilistic mode. We cannot, at least not currently, establish reliable probabilistic climate projections. But we can reasonably ask which future evolutions of the climate system are possible, and whether certain other climate projections—or "scenarios", as I shall also say in the following—are impossible.

"Possibility", here, means neither logical nor metaphysical possibility, but simply consistency with our relevant background knowledge. So, if we ask: can global mean temperature possibly rise more than 7° if the CO2-concentration doubles, we inquire as to the consistency of this statement with what we know about the climate (which may comprise deterministic, probabilistic or other possibilistic statements).

But how are we supposed to answer such a question? And how could we justify a given answer? [As a possibilistic prediction is, in some situations, all we can reliably establish, the methodological question how to justify scientific possibility statements pops up the agenda.] Since I equate possibility with relative consistency, the modal methodology seems to be, at first glance, a simple application of deductive logic. That is, unfortunately, untrue. Logic does not fully determine a modal methodology for, say, climate change projections, because it is unclear whether the corresponding range of possibilities should contain (i) all future scenarios, which are positively shown to be relatively consistent, or (ii) those, which have not been shown to be inconsistent with the relevant background knowledge. This distinction gives rise to two alternative modal methodologies: modal inductivism and modal falsificationism.

According to modal inductivism, a (complex) statement is considered as scientifically possible if and only if it is positively shown to be possible. This means that those storylines which are not explicitly demonstrated to be possible won’t figure in the list of future possibilities.

Modal falsificationism, in contrast, holds that a statement has to be considered as scientifically possible if and only if it is not positively shown to be impossible. Accordingly, the construction of the scenario range proceeds in two steps. In a first step, one imagines as many different future storylines as possible; in a second step, these hypothetical scenarios are systematically tested for consistency with what we know. Only those storylines which survive these tests are included in the range of future possibilities.

Modal inductivism has been the preferred methodology of the IPCC and was coherently implemented in its first three Assessment Reports (IPCC 1990, 1996, 2001). A study of the future scenario ranges that are communicated in the IPCC reports suggests that the IPCC simply identifies the range of possible climate scenarios with the range of simulation results, thus assuming that, *first*, every climate simulation shows the respective result to be consistent with our background knowledge, and, *second*, a scenario is scientifically possible in a policy-relevant way if and only if its consistency is verified.

Modal inductivism and modal falsificationism give rise to a significant methodological dilemma. First of all, the two methods represent a real alternative and result, if implemented, in completely different possibility ranges. Ultimately, this will affect the policy decisions we take. Consider, for example, the melting of the West Antarctic ice-sheet. Such a scenario can neither be *shown* to be relatively consistent, because we lack the appropriate models of ice-dynamics, nor can we demonstrate that it is inconsistent (cf. IPCC 2007, Box 10.1, Section 10.7.4.4). So whether the melting of the West Antarctic ice-sheet is considered as possible merely depends on whether modal inductivism or modal falsificationism is implemented. Still, that makes a difference of 7m in sea-level rise projections, and will therefore crucially influence climate mitigation and adaptation policies.

So we have a real alternative. Now this represents a dilemma, because both methodologies seem to be problematic. Modal inductivism, on the one hand, systematically underestimates the range of possibilities. It’s a folly to consider a specific behaviour of some complex system as impossible just because we have not yet developed an appropriate (not necessarily perfect or unique) model that can reproduce the behaviour. Modal inductivism's systematic bias becomes also apparent in the IPCC reports, who typically had to admit that previous reports underestimated uncertainties; scenario ranges tended to be too narrow and were, sometimes, corrected in the following report (Betz 2009b).

Modal falsificationism, on the other hand, does not really do much better, it appears. This methodology licenses the inclusion of all sorts of statements and storylines in the range of future possibilities and leads to a proliferation of possibilities—which is, by the way, precisely the reason for which the "mere-possibility-argument" criticised the precautionary principle. Modal falsificationism therefore avoids the typical error of modal inductivism, namely to overlook a possibility, but it is prone to another type of error: that is to consider a statement as possible which actually isn't.

3. Dropping the dilemma's underlying assumption

Some philosophers of science might regard the situation as depicted so far symptomatic, because it seems to show, once more, that methodological choices are ultimately based on value judgements—in this case, value judgements regarding the avoidance of which type of error (false impossibility or false possibility) should be preferred.

The following sections, however, run counter to such a view. They aim at avoiding the dilemma exposed above and will, if they succeed, eliminate this specific value-component in modal methodology. By doing so, they will draw a richer and more nuanced picture of our modal knowledge than one typically encounters, and we have encountered so far.

There is an implicit assumption, hardly ever questioned, which underlies and generates the dilemma. Namely,

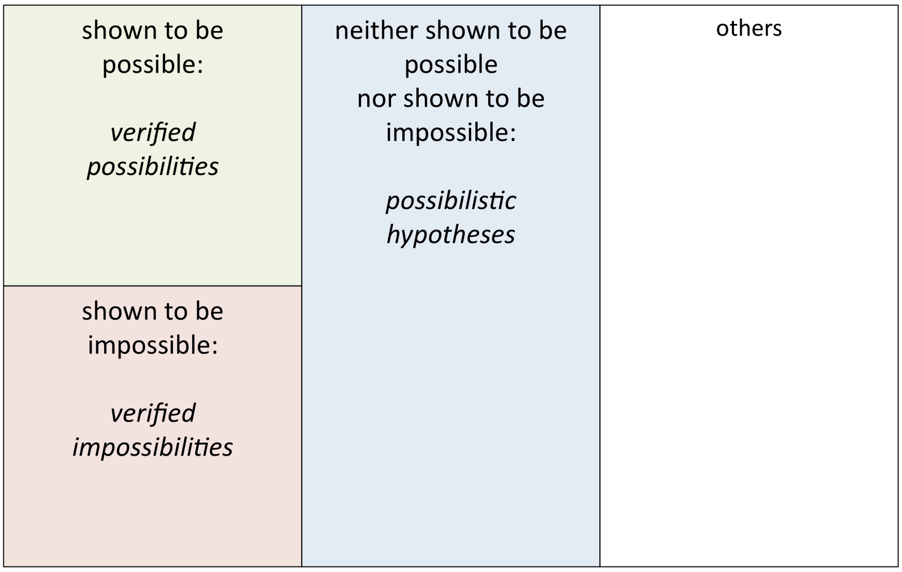
(1) Possibilistic statements about the future fall into two classes: a class of possible, and a class of impossible statements.

This basic assumption about the logical geography of statements about the future implies that a modal methodology consists in rules which prescribe how to sort statements by assigning them to one of these two classes.

My suggestion is that we should replace (1) by a more differentiated view.

(2) Possibilistic statements about the future fall into three classes: (i) verified possibilities, i.e. statements which are shown to be possible, (ii) verified impossibilities, i.e. statements which are shown to be impossible, and (iii) possibilistic hypotheses, i.e. statements which are articulated, but neither shown to be possible, nor shown to be impossible.

I understand that this classification is mutually exclusive, though not necessarily exhaustive. Moreover, it makes use of iterated modalities—but in a rather loose way, and I don't think that formalizing these issues gives significant new insights. When dealing with the methodology and dynamics of future possibilities in the following sections, it will be convenient to visualise the classification (2) as follows,



The fact that the classificatory scheme (2) does not pretend to be exhaustive is reflected by the inclusion of an area "others" in the diagram.

4. The variety of modal methods

Modal methodology contains, generally, prescriptions for classifying statements about the future according to the types of possibilities and impossibilities which are postulated. In line with the more fine-grained view expressed in (2), such a methodology consists in at least three different methods which state conditions for assigning a statement to one of the three postulated categories: the articulation of possibilistic hypotheses, the verification of possibilistic hypotheses, and the falsification of possibilistic hypotheses.

The *articulation* of a possibilistic hypothesis is the most fundamental method of a modal methodology. A possibility statement cannot be submitted to any further examination unless it is articulated. In other words, we have to *think of* a possibility in the very first place. This sounds trivial only to the inexperienced. Many future scenarios, which had initially not even been imagined, later turned out to be possible, if not even true. The paradigmatic case is, obviously, chlorofluorocarbons (CFCs) triggering the depletion of the atmospheric ozone layer—a possibility no one even dreamt of in the 1930s. The articulation of possibilistic hypotheses does not require formal mathematical reasoning, sound argumentation, precise measurement, or any other virtue typically associated with scientific reasoning; rather, it appeals to the virtues of fantasy, and creativity. No doubt, here is a place where laymen can contribute to scientific progress as much as trained experts can. It is worth noting, though, that the articulation of possibilistic hypotheses does not require *but* the virtues of fantasy, and creativity. And this relates to a fact Philip Kitcher recently stressed, and reminded us of (Kitcher 2001): We do not simply expect scientific results to be accurate or true, but we also expect them to be *significant*. Specifically, possibilistic hypotheses should be significant, too. Scenarios, it seems to me, can be insignificant in different ways. Consider, again, climate scenarios: they might simply not address the policy question at all (sales of Beatles-songs might skyrocket in ten years); or they might be articulated on an inappropriate level of detail (one extreme: it's, globally and on average, possibly going to be more windy; the other extreme: weather forecasts for all cities and every day in the year 2100). Whether there are other ways in which a scenario can be insignificant, and whether judging a possibilistic hypothesis significant involves certain types of arguments, or is simply based on *Urteilskraft*, seem to me open questions. They represent one issue to be addressed when further elaborating the method of articulation, which is part of a modal methodology.

Once a possibilistic hypothesis has been articulated, it can be verified, or falsified. Its *verification* consists in the demonstration that it is consistent with the relevant background knowledge; the articulated possibilistic hypothesis becomes a verified possibility. This might be achieved in different ways, e.g. by deducing the hypothesis from the background knowledge and other already verified possibilities, or by constructing a model (in the sense of formal semantics) which makes the background knowledge as well as the respective possibilistic hypothesis true. Verifying the possibility that more than 50% of the CO_2 stored in a geological reservoir escapes within the next century might, for instance, consist in deducing this scenario from (i) knowledge about the geological reservoir, (ii) knowledge about the way supercritical CO_2 behaves, (iii) the proven possibility that the caprock contains faults which expand and widen under increased pressure.

The demonstration that an articulated possibilistic hypothesis is inconsistent with our background knowledge amounts to its *falsification*. The possibilistic hypothesis becomes a definite impossibility. To prove an inconsistency, it suffices to deduce a contradiction. And such a deduction might merely start from a single fact. That is the reason why falsification, as a modal method, potentially allows to make use of heterogeneous and diverse evidence by integrating it into a methodological framework. It turns it into "evidence for use" (Cartwright 2006). Thus, to give a hypothetical example, a seemingly simple fact about a CO_2-storage site, such as whether a certain mineral occurs in the formation, can be sufficient, at least in principle, to rule out an entire future storage scenario in so far its implications contradict that single fact.

So by identifying an assumption which underlies the dilemma of possibilistic forecasting, and by replacing it with an alternative view, we found that a modal methodology contains a variety of different methods. But what does this imply for the initial dilemma? It is not difficult to see that it vanishes. Specifically, it presented us a false alternative. Instead of being incompatible methods, modal inductivism as well as modal falsificationism have both a role to play in modal methodology, namely when it comes to verifying or falsifying possibilistic hypotheses. Instead of choosing to base our decisions either on the results obtained by modal inductivism or those obtained by modal falsificationism, we should appreciate that both results contain valuable information about the future and should therefore be part of a more nuanced picture of what might going to happen.

5. The role of simulation in modal methodology

As opposed to the articulation of possibilistic hypotheses, both their verification and falsification rely on scientific observation and reasoning, including accurate measurements, statistical calculations, inductive and deductive arguments, mathematical deductions, and so on. This section addresses the question whether and how modal methods can make use of a particular tool for scientific reasoning: computer simulations. More specifically, I focus on cases where uncertainty stems from, or at least corresponds to, model-underdetermination, that is cases where there is no single, empirically and structurally adequate model of the simulated system, but a plurality of different, incompatible, and typically complex models giving rise to a variety of different simulations. What is the function these simulations might serve within a modal methodology?

As long as we suppose that each model is at least possible, if not true, simulations might serve to verify possibilistic hypotheses by deducing them from a set of collectively possible statements. Similarly, these simulations based on complex models might be considered as models (or possible worlds) in the sense of formal semantics—thence demonstrating the possibility, respectively relative consistency, of statements about the future. 'True in a model simulation' implies 'definitely possible'. But this seems to require that the models are at least possible—arguably a rather weak requirement. Or so it seems. Actually, some complex climate models (GCMs) are violating fundamental principles of mass and energy conservation. In which sense can these models be considered possible? And do corresponding simulations really show that some statement is consistent with our background knowledge? Things look even worse when we recall that every climate model is empirically inadequate and contradicts some empirical facts about the climate system (Stainforth, Allen et al. 2007). Every GCM is known to be inconsistent with our background knowledge. How could these models be used to demonstrate consistency with our background knowledge? I don't have a satisfying answer to this question. One might try to defend climate simulation's purposefulness by a pragmatic reasoning in line with Eric Winsberg (Winsberg 1999, 2006), that is by pinpointing that models of this kind, constructed in this or that way (e.g. by integrating unrealistic flux-adjustments), have a successful track record *regard-*

ing the verification of possibilistic hypotheses. Still, as a matter of fact, there is no such track record. (And it is not quite obvious how such a track record of successful verifications of modal hypotheses might even look like.)

The value of simulations vis-à-vis *falsification* of modal hypotheses is even more doubtful. If, in a classic deduction, one premiss is merely possible, the conclusion cannot hold necessarily. Likewise, building on premisses which are themselves only known to be consistent with some background knowledge, you cannot prove that some other statement is inconsistent with the respective background knowledge. Or, a consideration of one possible world won't allow you to infer anything about all possible worlds; although that is, besides, what Plantinga tried to do in his version of the ontological argument, namely via the concept of "maximal greatness" (Plantinga 1977, p. 108). These logical facts entail that simulations, when model-underdetermination reigns, cannot be used to falsify possibilistic hypotheses. The falsification of modal hypotheses, e.g. that climate sensitivity is larger than 10°C, cannot rely on complex climate models, the scenario range has to be constrained by other, "robust" methods.

These conclusions as to the role of simulations in modal methodology resonate with a recent assessment of climate model projections mentioned above (Stainforth, Allen et al. 2007). Firmly rejecting attempts to infer probabilistic predictions from GCM ensembles, the authors suggest that the range of simulated climate projections represents "possibilities for future real-world climate" (p. 2155), and gives us "a lower bound on the maximum range of uncertainty" (p. 2156). Thus, I understand, these authors take climate simulations to show that some future scenarios are possible (positive role in modal verification) without claiming that *only* these simulated scenarios are possible (rejecting modal inductivsm). Moreover, Stainforth et al. seem to conclude that future scenarios cannot be shown to be impossible by GCM simulations (no role in modal falsification): While GCM ensembles provide a lower bound to the range of possibilities, "objective and robust methods" are required "to constrain them" (p. 2159).

To sum up, simulations might have a role to play in modal verification; they definitely won't contribute to modal falsification. But what about the *articulation* of modal hypotheses? I said that this method primarily relies on the virtues of fantasy and creativity rather than on the classic epistemic virtues of scientific reasoning. But then, computer simulations seem to be irrelevant here. For they help us with complex computation, the integration and visualization of large amounts of very precise data, the design of huge experiments, etc.—in brief, they enhance scientific observation and reasoning (cf. Humphreys 2004). But they don't enhance fantasy. Or do they? Complex simulations have an interesting epistemic feature, which makes them apt for supporting our imagination, our creativity. They can give rise to so-called emergent phenomena, where I shall follow Mark Bedau (Bedau 1997) in defining a "weakly emergent phenomenon" P of some system S which is governed by a microdynamic D as follows:

> Macrostate P of S with microdynamic D is weakly emergent iff P can be derived from D and S 's external conditions but only by simulation. (378)

Emergent phenomena which arise in complex systems whose microdynamics are known typically overstrain our imagination and capacity of foresight—we cannot even imagine the kinds of self-organizing behaviour cellular automata exhibit from knowledge of their simplistic rules only. As a consequence, simulating complex systems based on their microdynamics might suggest possibilistic hypotheses we have not even thought about! A candidate for a case in point where simulations already have successfully enhanced our imagination is the possible shift of the intertropical convergence zone due to anthropogenic climate change and a resulting drying of the Amazonian rain forest: a scenario which appears to have been suggested by simulations of the Hadley Center climate model in the 1990s (Cox, Betts et al. 2000; Cox, Betts et al. 2004) and which was—but that is only a hypothesis and requires a more detailed case study—possibly not even articulated beforehand. Given the difficulties of applying simulations in order to verify modal hypotheses and the impossibility to falsify them based on model simulation, I suggest that 'creative simulation' which aims at articulating unseen possibilities—without pretending that they are really possible, or impossible—might be the ultimate and foremost function of computer simulations in the epistemic mode of uncertainty. It is clearly not the function they are supposed to fulfil today. And taking that task seriously would require rethinking the design strategies of simulations; epistemic constraints on simulation building would be loosened; and progress would consist not in convergence of simulation results but in a proliferation of the underlying models, and of the scenarios they generate.

6. Surprises

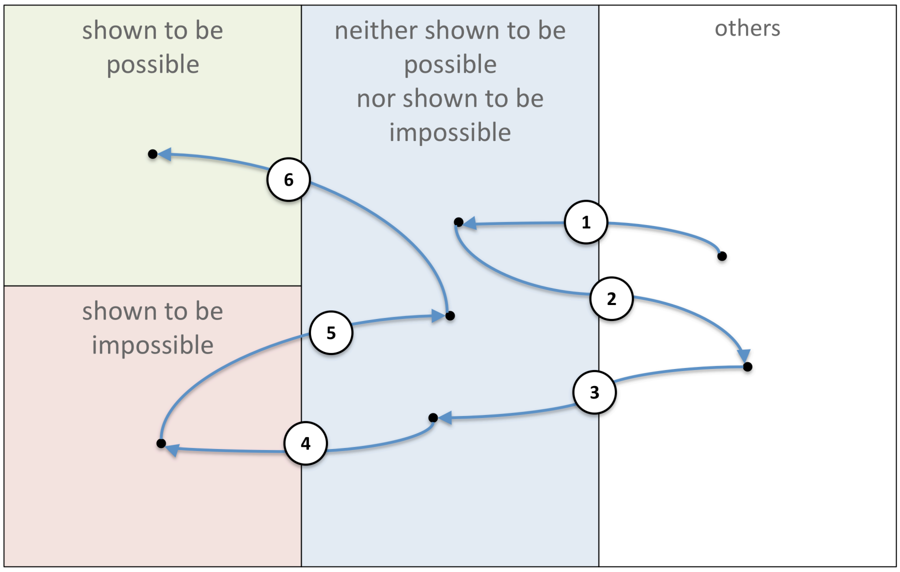
Surprises go hand in hand with a lack of knowledge, in particular foreknowledge. Were we gifted, or punished, with perfect foresight (Knight's epistemic mode of certainty), nobody would ever be surprised. A modal methodology, a methodology for imperfect and uncertain knowledge acquisition, has to deal, somehow, with surprises. So how do surprises fit into the picture as unfolded so far?

There are two types of surprises we can distinguish within the framework of a modal methodology and it is important to keep both in mind when assessing our foreknowledge. A surprise of the *first type* occurs if a possibility that had not even been articulated becomes true. I briefly touched this issue above when stressing the importance of hypothesis articulation as a modal method. Hypothesis articulation is, essentially, the business of avoiding surprises (of this first type).

There is, however, a *second type* of surprise that does not simply extend the picture we've drawn so far, but rather shakes it. By defining possibility as relative consistency we assessed possibilistic hypotheses with regard to some body of background knowledge that was assumed to be stable. Still, this assumption is, obviously, unrealistic. Our scientific knowledge is constantly changing, whereas that change is not cumulative: scientific progress also comprises refuting, correcting, and abandoning previous scientific beliefs. Now a readjustment of the background knowledge questions the entire former assessment of possibilistic hypotheses. Statements, which were compatible with the old body of background knowledge, might not be so any more relative to the corrected one. Similarly, falsified possibilistic hypotheses might not be falsifiable anymore—or even turn out to be definitely consistent with the new background knowledge. Scientific progress shakes the entire assessment of modal hypotheses and requires their re-evaluation. So, scientific change, being in itself unpredictable (to some extent for a priori reasons, as Popper (1982) argued), has the potential to generate surprises vis-à-vis our modal knowledge.

As fallibilists, we must accept the meta-possibility that what we consider as definitely impossible today, might turn out to be possible some when in the future.

Summing up the previous sections, the following diagram illustrates the potentially complex dynamics of our possibilistic knowledge with an abstract, hypothetical example.



The figure shows the changing epistemic status of some statement *P*. In step 1, *P* was articulated for the first time; what was before an unseen possibility or impossibility then became at least an articulated possibilistic hypothesis. But, for whatever reason, *P* was not passed on to next generations and was, soon after its initial articulation, completely forgotten (step 2). Centuries later, *P* was articulated again (step 3), and quickly dismissed, given the background knowledge of that time, as impossible (step 4). The possibilistic hypothesis *P* was falsified. Yet, scientific theories changed and in the course of the following decades, it became less and less clear that *P* really contradicted scientific knowledge (step 5). Due to scientific progress, the falsified possibility *P* became, again, merely an articulated possibilistic hypothesis. And it took, eventually, modern computer simulations to demonstrate that *P* is definitely consistent with state of the art scientific knowledge (step 6).

7. A new challenge for rational decision making

The previous sections developed a conceptual framework and sketched some methods which allowed us to express our uncertain, non-probabilistic foreknowledge in a nuanced and differentiated way. Since responsible decision making should surely be based on all available evidence, scientific policy advice should incorporate a full description of our modal knowledge along the lines of this paper, including the identification of verified, falsified, and merely articulated hypotheses. This gives rise to further problems, as I will try to show in this section.

These problems arise, mainly, because traditional principles for decision making under uncertainty, such as the maximin or the minimax regret principles (cf. Luce and Raiffa 1957; Savage 1972), assume that possibilistic predictions consist in simple possibility statements. They prescribe which alternative action to choose given their various possible outcomes, without distinguishing different kinds of possibilities. As a consequence, these traditional decision principles are not applicable to the decision situations under uncertainty when described along the lines of this paper. Adapting or replacing traditional principles so that they make use of all the information conveyed in the more detailed description of uncertainties generates non-trivial difficulties. I will try to highlight these with a hypothetical decision situation and an application of the precautionary/maximin principle.

Consider a decision situation where an agent faces to alternative options, A and B. Depending on the 'state of the world', these options' consequences are valued very differently, as depicted in the following table. Some states of the world are shown to be possible (S_1,S_2,S_3), others are neither shown to be possible, nor shown to be impossible (T_1,T_2,T_3), and some are, finally, definitely impossible (U_1,U_2,U_3). Let us try to apply the maximin rule to this situation.

	Utility of consequences according to states of the world								
	Verified possibilities			*Unverified & unfalsified possibilities*			*Falsified possibilities*		
	S_1	S_2	S_3	T_1	T_2	T_3	U_1	U_2	U_3
A	**-100**	10	5	**-10**	0	15	10	**-100**	50

B	20	0	**-10**	**-200**	-10	5	**-2000**	10	0

First, we consider but the verified possibilities. The worst case of option A is much worse than the worst case of action B, since we have A(S_1) = -100 < -10 = B(S_3). So B maximises the minimal outcome and should be adopted.

If we consider, *second*, the unverified & unfalsified possibilities, things turn upside down, though. The worst case of B with regard of these is T_1, which is also A's worst case. However, B(T_1) is much worse than A(T_1) and, moreover, definitely worse than A(S_1), i.e. A's worst case of all verified possibilities. So no matter whether you consider but the unverified & unfalsified possibilities or the verified as well as the unverified & unfalsified possibilities: the maximin rule prescribes to opt for A, not for B.

To illustrate another consideration that might become relevant, we consider, *third*, the falsified hypotheses. If the state of the world U_1—against all we know—became true, option B would trigger a catastrophe an order of magnitude worse than the worst cases considered before. Sure, U_1 has been shown to be impossible. Yet, the remarks of the last section reminded us that this judgement is fallible, and future scientific progress might trigger the surprising insight that U_1 is far from being definitely impossible. Thus, it seems, based on the full, detailed picture, a very cautious person could legitimately opt for A on these very grounds. Or couldn't she?

The decision deliberation starts to become messy and complicated. It is not clear to me, whether there are general principles which can guide rational decisions in such situations at all. This, however, must not serve as an excuse for simplifying the epistemic situation we face. If a policy decision requires a complex normative judgement, then democratically legitimised policy makers have arguably a hard job; it is, nevertheless, *their* job to balance and weigh the diverse risks of the alternative options. That is not the job of scientific policy advisers who might be tempted to simplify the situation, thereby predetermining the complex value judgements.

Yet, in terms of complexity of the decision process, it is even getting worse if we consider, finally, the first type of surprises I identified in section 6. Options might trigger consequences we had not even articulated when deliberating our decision. It appears to me that an informed decision would not only require the communication of all verified, falsified and articulated possibilistic hypotheses. It should, ideally, take into account an estimate of the potential to generate surprises for each of the respective options. On the one hand, this is obviously a weird idea, almost self-refuting and bordering on the paradoxical: We should assess, compare, and count statements we have not even articulated? Admittedly, I can't think of any detailed prescription for how to do that. On the other hand, this idea stresses an important point, as I'll try to make clear with the following example. Reconsider the options A and B from above. I made plausible that, in the light of U_1, a very cautious person might reasonably prefer A over B. But assume that A actually involves constructing and running a machine that has never been built before. Assume that machine will accelerate particles to speeds that have never been reached before in this part of the universe. Our best theories don't tell us anything about the kind situation that will be generated. Option B, we shall assume, just involves conventional means. Is it really irrational to refute option A on the grounds that we would be entering territory where we don't even know what might happen, where we have not the slightest experience of what sort of phenomena we might trigger, and where, therefore, we might eventually face consequences which exceed, today, our conceptual and imaginative capacity? Although such a conclusion is not compulsory, it does not seem wholly unreasonable, either. This example puts another, this paper's final, item on the agenda of a modal methodology: It is important to clarify when, and how, one should estimate an option's potential for surprise.

Conclusion

This paper sketched a conceptual framework for expressing uncertain, possibilistic knowledge. A framework which allows to express our foreknowledge in a more nuanced way than simply by labelling some statements about the future as possible. I suggested that we should adopt this framework for stating and communicating scientific results in the epistemic mode of uncertainty. This avoids the methodological dilemma between modal inductivism and modal falsificationism. The framework's conceptual variety of possibilities triggers a methodological variety, a plurality of modal methods. Some of these rely on traditional virtues of scientific reasoning, others don't. In some disciplines, computer simulations might, surprisingly, be most profitably applied in and contribute to the creative methods, rather than the strict and formal ones.

This paper also generates a couple of questions which deserve further investigation. They include:

- What counts as a significant hypothesis? If articulated possibilistic hypotheses are supposed to be relevant (e.g. to the decision problem), what kind of relevance or significance is referred to?
- Can simulations really be creative? Can they contribute to reducing our ignorance by articulating possibilities we had not even thought about? Are there paradigmatic examples? And how would taking this function serious affect the design of simulation studies?
- Can we estimate (the two different) potentials of surprise we face in a given epistemic situation?
- Can traditional decision principles be adapted so that they take into account the detailed possibilistic information conveyed in the conceptual framework exposed in this paper? Or is the deliberation of decisions under uncertainty becoming too complicated to be guided by rules of rational choice? And, if not by normative principles, how else could a theory of rational choice support decision making under uncertainty?

Bibliography

Bedau, M. A. (1997). "Weak Emergence." Philosophical Perspectives **11**: 375-399.

Betz, G. (2007). "Probabilities in Climate Policy Advice: A Critical Comment." Climatic Change **85**: 1-9.

Betz, G. (2009a). "Underdetermination, Model-ensemble, and Surprises – On the Epistemology of Scenario-analysis in Climatology." Journal for General Philosophy of Science **submitted**.

Betz, G. (2009b). "What range of future scenarios should climate policy be based on? Modal falsificationism and its limitations." Philosophy of Science **submitted**.

Cartwright, N. (2006). "Well-Ordered Science: Evidence for Use." Philosophy of Science **73**: 981-990.

Cox, P. M., R. A. Betts, et al. (2004). "Amazonian forest dieback under climate-carbon cycle projections for the 21st century." Theoretical and Applied Climatology **78**(1-3): 137-156.

Cox, P. M., R. A. Betts, et al. (2000). "Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model." Nature **408**(6809): 184-187.

Faber, M., R. Manstetten, et al. (1996). Ecological Economics: Concepts and Methods, Edward Elgar.

Gardiner, S. M. (2006). "A Core Precautionary Principle." The Journal of Political Philosophy **14**(1): 33-60.

Gaus, I., P. Audigane, et al. (2008). "Geochemical and solute transport modelling for CO2 storage, what to expect from it?" International Journal of Greenhouse Gas Control **2**(4): 605-625.

Healy, S. (1999). "Extended peer communities and the ascendance of post-normal politics." Futures **31**(7): 655-669.

Hegerl, G. C., T. J. Crowley, et al. (2006). "Climate sensitivity constrained by temperature reconstructions over the past seven centuries." Nature **440**: 1029-1032.

Humphreys, P. (2004). Extending Ourselves: Computational Science, Empiricism, and Scientific Method. New York, Oxford University Press.

IPCC (1990). Scientific Assessment of Climate change; Report of Working Group I Cambridge ; New York, Cambridge University Press.

IPCC (1996). Climate change 1995 : the science of climate change; Contribution of Working Group I to the Second Assessment of the Intergovernmental Panel on Climate Change. Cambridge ; New York, Cambridge University Press.

IPCC (2001). Climate Change 2001: The Scientific Basis; Contribution of Working Group I to the Third Assessment Report of the

Intergovernmental Panel on Climate Change. Cambridge, Cambridge University Press.

IPCC (2005). IPCC special report on carbon dioxide capture and storage. Cambridge, Cambridge University Press, for the Intergovernmental Panel on Climate Change.

IPCC (2007). Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge; New York, Cambridge University Press.

Kitcher, P. (2001). Science, Truth, and Democracy. Oxford, Oxford University Press.

Knight, F. (1921). Risk, uncertainty and profit, Houghton Mifflin.

Levi, I. (1980). The Enterprise of Knowledge. An Essay on Knowledge, Credal Probability and Chance, MIT Press.

Luce, R. D. and H. Raiffa (1957). Games and Decisions: Introduction and Critical Survey. New York, John Wiley.

Manson, N. A. (2002). "Formulating the Precautionary Principle." Environmental Ethics **24**: 263-274.

Napier, W. (2008). Hazards from comets and asteroids. Global catastrophic risks. N. Bostrom and M. M. Cirkovic. Oxford ; New York, Oxford University Press**:** 222-237.

O'Hagan, A. and J. E. Oakley (2004). "Probability is perfect, but we can't elicit it perfectly." Reliability Engineering & System Safety **85**(1-3): 239-248.

Oreskes, N., K. Shrader-Frechette, et al. (1994). "Verification, Validation, and Confirmation of Numerical Models in Earth Sciences." Science **263**: 641-646.

Plantinga, A. C. (1977). God, freedom, and evil. Grand Rapids, Mich., Eerdmans.

Popper, K. R. (1982). The Open Universe. Totowa, Rowan and Littlefield.

Savage, L. J. (1972). The foundations of statistics. New York, Dover Publications.

Stainforth, D. A., M. R. Allen, et al. (2007). "Confidence, uncertainty and decision-support relevance in climate predictions." Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences **365**(1857): 2145-2161.

Stainforth, D. A., T. E. Downing, et al. (2007). "Issues in the interpretation of climate model ensembles to inform decisions." Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences **365**(1857): 2163-2177.

Walley, P. (1991). Statistical Reasoning with Imprecise Probabilities, Chapman and Hall.

Webster, M. D., C. E. Forest, et al. (2003). "Uncertainty analysis of climatic change and policy responses." Climatic Change **61**(3): 295-320.

Winsberg, E. (1999). "Sanctioning Models: The Epsiemology of Simulation." Science in Context **12**(2): 275-292.

Winsberg, E. (2006). "Models of Success vs. the Success of Models: Reliability Without Truth." Synthese **152**: 1-19.

Wynne, B. (1993). Uncertainty and Environmental Learning: Reconceiving Science and Policy in the Preventive Paradigm. Clean Production Strategies: Developing Preventive Environmental Management in the Industrial Economy. T. Jackson, Lewis.

Zickfeld, K., A. Levermann, et al. (2007). "Present state and future fate of the Atlantic meridional overturning circulation as viewed by experts." Climatic Change **82**: 235-265.