Probability Theory and *Intrinsically* Non-Stochastic Phenomena in AI: How to Reject a Metaphysical Issue.

Carlotta Piscopo
IRIDIA, Université Libre de Bruxelles
Brussels, Belgium

Mauro Birattari
Intellektik/Informatik, TU Darmstadt
Darmstadt, Germany

Abstract The paper discusses some arguments that have been raised against the use of probability theory in artificial intelligence. We identify two main elements in these arguments: The first is related to practical issues on the use of probability theory, while the second is clearly of philosophical nature and concerns the contraposition stochastic/deterministic and the status of the *stochastic object*. We call these two elements *technical* and *methodological*, respectively. The methodological element has not been addressed by the AI community, and our paper is intended precisely to fill this gap. In the methodological element we spot the misleading assumption that since probability theory supposes that data are observations of a *stochastic object*, a probabilistic model is adequate only for dealing with *intrinsically* stochastic phenomena.

In the paper, on the basis of ideas borrowed from the empiricist and instrumentalist epistemology, we point out that the central hypothesis formulated when adopting a probabilistic model, namely the existence of the stochastic object, is simply a *working hypothesis*. It follows that a probabilistic model is to be evaluated only on the basis of its predicting power, leaving aside any further metaphysical consideration on the *truth* of the underlying working hypothesis.

1 Introduction

In artificial intelligence, the problem of uncertainty arises whenever an agent is required to acquire information on the environment, and to use such information to make decisions and act. In the general case, making decisions necessitates the ability of predicting the consequences of different possible actions, where the prediction is based, at least partially, on empirical knowledge extracted from the observation of the environment. At the age of birth of artificial intelligence, probability theory seemed the natural candidate for carrying out the task of inferring effective predictions from empirical data. At that time, probability theory was a well-established 300-years old theory with a respectable record of successes in treating practical problems emerging in many scientific domains ranging from statistical mechanics to information theory.

Nevertheless, following McCarthy and Hayes [10], some researchers considered inadequate for AI applications the theory of probability and, accordingly, they proposed alternative formalisms. Even if most of the arguments raised against probability theory were rejected

Technical Report AIDA-2002-04 of Fachgebiet Intellektik, Fachbereich Informatik, Technische Universität Darmstadt, 64289 Darmstadt, Germany. March 2002.

by the majority of the AI community [4], the discussion on this issue should be considered, in our view, very fruitful. Important contributions, as for instance Bayesian Networks [13], emerged also as a result of this debate.

Some of the arguments raised against the use of probability theory where clearly of philosophical nature. They can be summarized in the claim that in AI uncertainty is not related to stochastic phenomena but it has an intrinsically subjective and epistemic nature. The aim of this paper is to discuss such a claim that, up to now, has not found any clear and thorough analysis in the literature. Since the claim is undoubtedly philosophical, our answer will stay on the same level and will adopt therefore the concepts, the language, and the arguments of the epistemological analysis. In particular, in the paper we put forward and we argue in favor of the thesis that the above stated claim is metaphysical and that, as such, it should be excluded from the debate on the adequacy of probabilistic models in AI. Still, the reader should not intend our statements as an endorsement of the use of one method for handling uncertainty rather than another: This paper should be intended strictly and only as a discussion on the specific claim under analysis.

In order to proceed to our argumentation, some concepts need to be briefly introduced here. The adoption of a probabilistic model, in AI as in any other field of research, entails, in some sense to be clarified presently, the assumption that the available data are observations of a *stochastic object*. In the paper, with the expression *stochastic object* we mean some entity that we perceive through some observable taking on, in time, *random* values governed by some probability measure. As a mathematical—and therefore formal—object, the stochastic object is clearly defined by the precise mathematical concept of probability measure [3].

On the other hand, the actual existence in Nature of the stochastic object is a purely metaphysical issue: It is impossible to falsify on the basis of empirical observations either the statement "the stochastic object exists in Nature" or its contrary "the stochastic object does not exist in Nature." If we are presented with a finite sequence of observations, we have no clear–cut way for deciding if the sequence was generated by a stochastic rule or by a deterministic one.[1] On the basis of these considerations, as we argue in the body of the paper, the existence of the stochastic object to which probability theory refers the actual observations, is to be intended as a *working hypothesis* and does not mean to have any ontological implication.

The paper is structured as follows. In Section 2 we introduce the criticisms moved to the use of probabilistic models in artificial intelligence. In particular, we introduce the claim that we discuss in the rest of the paper: "Probability theory is not adequate in AI because we deal with non-stochastic phenomena." In Section 3 we show that recurring to working hypotheses is common to all scientific domains and that probabilistic models used in AI are no exception. In this section, we argument in favor of the idea that the existence of the stochastic object is a working hypothesis rather than an ontological truth. Section 4 concludes the paper by showing that rejecting the use of probabilistic methods in AI on the basis of the statement that "AI deals with non-stochastic phenomena" opens the way to a rejection of the scientific method in general.

[1] To be precise, statistical tests exist for rejecting the null hypothesis that a given string of observations is random [8, 9]. Nevertheless, as it is always the case in hypothesis testing, when such tests reject the null hypothesis, they do it within some *confidence level* that by definition is strictly smaller than the unity: The complete certitude that the string is not random cannot ever be obtained. On the other hand, it is easy to show that it is impossible, on the basis of a finite string, to infer underlying deterministic rules. The well known argument of the "inductivist turkey" proposed by Russell [16] should be sufficient for convincing our reader.

2 Criticisms Against the Use of Probability in AI

Starting from the sixties, the belief spread among part of the AI community that probability theory was not adequate for handling the problem of decision making under uncertainty [10]. Such belief was supported by a variety of composite arguments. In this paper, we find convenient to highlight two main elements that were usually present in these arguments. The first is related to practical issues on the use of probability theory, while the second is clearly of philosophical nature and concerns the contraposition stochastic/deterministic and the status of the *stochastic object*. In the following we call these two elements *technical* and *methodological*, respectively.

The technical element has been widely discussed in the literature and is considered as refuted by the large majority of the AI community: Section 2.1 proposes a quick overview of the main issues connected to it. On the other hand, to the best of our knowledge, what we call here the *methodological element* has not been so far the subject of any analysis. Apparently, the AI community felt that the refutation of the technical element was to be considered as a refutation of the criticisms as a whole and therefore never moved to discuss the methodological one. We maintain that, from a speculative point of view, the general discussion on the adequacy of probability in AI cannot be accomplished without a thorough analysis of the methodological element. This paper aims precisely at filling such a gap.

2.1 The Technical Element

The AI community discussed in depth various technical issues. We refer the reader to Cheeseman [4] for a systematic analysis of such issues and for the relative answers. In the following, we limit ourselves to a quick overview of some of the points made clear by Cheeseman.

The general idea underlying the counter-arguments proposed in [4] is that the criticisms raised against probability theory derive from misconceptions of probability: "*. . . these supposed difficulties are common misconceptions of probability, generally springing from the inadequate frequency interpretation. A major aim of this paper is to put forward the older view (Bayes, Laplace etc.), that probability is a measure of belief in a proposition given particular evidence. This definition avoids the difficulties associated with the frequency definition and answers the objections of those who felt compelled to invent new theories.*"—[4]. A Bayesian interpretation, as Cheeseman points out, is profitable when we have to cope with little data available to an agent, like it is usually the case in AI applications.

Another claim that Cheeseman recognizes as mistaken is related to the idea that more than one number is needed to represent uncertainty [17]. This claim is usually supported by the argument that one number does not represent the accuracy with which probabilities value are known, and, moreover, does not allow a distinction between uncertainty and ignorance. According to Cheeseman, the necessity of alternative representations of nuances of uncertainty have to be justified by the fact that they have an effective influence on the process of making predictions or decisions: "*. . . how many numbers are needed to represent uncertainty depends on the question you are trying to answer with the uncertainty representation. To always calculate two numbers, as done in Schafer-Dempster approach, is often overkill, and in some cases, under-kill.*"—[4].

Finally, Cheeseman discusses the claim that probability seems not being normative since psychological experiments show that experts, as well as common people, express beliefs that systematically violate the axioms of probability [18]. According to Cheeseman: "*The fact that* *these subjective estimates will be poorly known is no excuse for not using them. Fortunately, the final probability values calculated on the basis of extensive new information are not very sensitive to the exact value of the priors.*"—[4].

2.2 The Methodological Element

Some authors distinguished between data that are intrinsically stochastic and data that are intrinsically non-stochastic in order to justify the adequacy of probabilistic models for the former, and the necessity of alternative methods for the latter.

Zadeh, for example, maintained that probabilistic techniques are not effective in treating data that are not intrinsically stochastic: "*... it is a standard practice to rely almost entirely on the techniques provided by probability theory and statistic, especially in applications relating to parameter estimation, hypothesis testing and system identification. It can be argued, however, as we do in the present paper, that such techniques cannot cope effectively with those problems in which the softness of data is nonstatistical in nature—in the sense that it relates, in the main, to the presence of fuzzy sets rather than to random measurement errors or data variability.*"—[21].

As we argue in Section 3, being stochastic or non-stochastic is not a property of the data, but it is a property of the model in which we *decide* to cast them. Any statement that pretends to grasp the intrinsic properties of the data and of the process that generate them is metaphysical, and do not have a scientific status. We recognize as metaphysical, for instance, the distinction between *chances* and *belief* made by Shafer: "*Chances then, must be conceived of as features of the world. They are not necessarily features of our knowledge or belief. And it would be quite untenable to claim that a chance is merely a feature of our knowledge or belief.*"—[17]. On the basis of such distinction, Shafer claims that problems involving chances and those involving beliefs are intrinsically different and for the first ones probability theory is the natural framework, while for the seconds other methods are needed that should not respect the axioms of probability theory. In Section 3 we show that the metaphysical contraposition stochastic/deterministic becomes marginal when the assumption that data are produced by a stochastic objects is recognized as a *working hypothesis*.

3 Probabilistic Representations as Working Hypotheses

Formulating hypotheses and using them to give an interpretation of observed phenomena, is typical of every scientific discipline that has to deal with empirical data. In the case of probabilistic models, the very existence of the stochastic object is meant as a working hypothesis through which data can be framed and used to produce testable predictions. The actual existence in Nature of the stochastic object, is not of concern in the probabilistic framework: What matters is only the effectiveness of the prediction that a probabilistic model consents.

Once it is clear that referring the data to a stochastic object is simply a working hypothesis, and that the stochastic object does not necessarily correspond directly to any physical entity, the claim that probabilistic models are suited only for treating data that are intrinsically stochastic becomes questionable.

We develop this argumentation in the following sections. More precisely, in Section 3.1, through some ideas borrowed from the empiricist and instrumentalist epistemology, we show that the procedure of referring empirical data to working hypotheses is not a prerogative of the construction of probabilistic models, but is typical of the scientific method in general. In Section 3.2, we show that the procedure of referring empirical data to working hypotheses is not arbitrary. On the contrary, it is deeply anchored to reality and it is justified by the fact that it allows to solve real world problems. In Section 3.3, we show that the idea of seeing the existence of the stochastic object as a working hypotheses is not an *ad hoc* trick that we propose here for justifying the adoption of probabilistic models in AI. On the contrary, such idea was already clear in the early works on probability by Laplace and even before by Pascal.

3.1 Empiricism, Instrumentalism, and Working Hypotheses in Science

The procedure followed in AI for building a probabilistic model from data is strongly related to the scientific method in general. Collecting empirical data, organizing them in a modeling framework, and formulating predictions are some of the characterizing elements of the scientific method. In this section, we analyze in detail such characterizing elements through some epistemological ideas emerged in a time window ranging from the 17th to the 20th century. With the following, we do not mean to endorse *as a whole* the philosophy of the thinkers we consider: More simply, we freely extract, from each, the elements that we find useful for composing the mosaic of our argumentation.

Galileo, who is widely considered as one of the founders of the scientific method and is often associated with extreme empiricist ideas, was inspired in his work by the conviction that any empirical observation is neutral. Still, Galileo recognized that experiments are always rationally guided: For him, observations are "*sensate esperienze*," meaning with this that collecting data is an active phase biased by the theoretical framework accepted by the experimenter [5].

Berkeley was another convinced empiricist and, as such, he thought that all abstract concepts not referring to physical observable objects have to be eliminated from science [1]. He refused, for example, the Newtonian concepts of absolute space and absolute time because these are not observable objects. However, later [2] he justified the rejection of these concepts in another way: absolute space and absolute time have to be avoided because they do not bring any effective contribution to the general theory. In other terms they have to be eliminated because, according to Berkeley, they are useless. Clearly, Berkeley modified his pure empiricist view [2] and started to accept that some abstract concepts, as for instance those of "force" or "gravity" as used by Newton, can be admitted if they are successful and *useful* for making calculations: "*And just as geometers for the sake of their art make use of many devices which they themselves cannot describe nor find in the nature of things, even so the mechanician makes use of certain abstract and general terms, imagining in bodies force, action, attraction, solicitation, etc. which are of first utility for theories and formulations, as also for computation about motion, even if the truth of things, and in bodies actually existing, they would be looked for in vain, just like geometers' fictions made by mathematical abstraction.*"—[2].

In this composite empiricist–instrumentalist position, Berkeley anticipated Mach in maintaining that scientific hypotheses and theories might be justified simply because they are useful, without any need, and indeed any possibility, of saying anything about the ontological status of the postulated entities. According to Mach, scientific models should not be meant as immediately referred to *real entities* but only as convenient tools for making predictions and for giving simple and unifying descriptions of phenomena [7]. The empiricist and the instrumentalist view of science stressed by Mach had an enormous impact on the history of science and philosophy. By clearly stating that scientific models are useful tools concerned with and only with measurable quantities, Mach eventually freed scientific models from the task of grasping any ontological property of Nature and assigned them the pragmatic task of making effective predictions.

Deeply inspired by Mach, the members of the Vienna Circle contributed to spread this empiricist view of science that, in their neo-positivist program, was indeed intended as a mean for banishing any metaphysical speculation from science [19]. According to the Vienna Circle, metaphysical statements are lacking any meaning precisely because they are not empirically *verifiable*.

In the same sense, the reflections proposed by Wittgenstein in the *Tractatus* [20] are deeply anti-metaphysical: Wittgenstein states that science recurs to principles that, without giving an intrinsic description of Nature, are useful rules into which we are able to cast the observations. An example of these principles is causality: looking at the empirical world in terms of causal relations, Wittgenstein says, allows for useful interpretations, but this does not imply that the physical world is really causal.

Once freed from metaphysical tasks, scientific models can and should be regarded, following Popper, as inventions of the human mind rather than discoveries of the ontological properties of Nature. Still, scientific models are not completely arbitrary as they hold a precise connection with reality in the fact that they are constantly checked against experience [15]. In the context of the analysis we propose with this article, Popper's position is particularly relevant: Scientific models are not obtained by induction from empirical observation, but are challenging speculations which are only required to lend themselves to empirical evaluation. Popper characterizes precisely as metaphysical, and thus as non scientific, every statement that cannot be tested and that cannot be possibly disproved by empirical evidence.[2]

In general, several possible models can match the available experimental observations. The selection of one among such models is precisely the free process of invention pleaded by Popper. Different extra-evidential criteria might regulate this process on the basis of some principle of economy as for instance simplicity or conservatism. Occam's razor is probably the most widely known and cited extra-evidential criterion. Even if in a slightly different sense, another principle of economy leads to prefer models and theories whose analytical formulation is easy to manage.

For summarizing, models and their underlying hypotheses cannot be judged on the basis of their capability of corresponding to physical objects. The only criteria contemplated by science are on the one hand the ability of yield predictions in agreement with observations, and on the other hand some extra-evidential criteria that can be traced back to a principle of economy. A scientific model is based on working hypotheses that do not express any ontological statements on Nature, but, as Poincaré clearly stated, they are conventions that are useful for making effective predictions: *"Peu nous importe que l'éther existe réellement, c'est l'affaire des métaphysiciens ; l'essentiel pour nous c'est que tout se passe comme s'il existait et que cette hypothèse est commode pour l'explication des phénomènes.*[3]*"*—[14].

[2] The work of Popper [15] is widely recognized as a milestone in philosophy and science. In the context of our epistemological view, the concept of falsification and the related demarcation between science and metaphysics hold the most prominent position. Thanks to these concepts, Popper rescued the empiricist idea from the slippery grounds of idealism and solipsism to which it was lead by Berkeley and Mach. At the same time, Popper criticized the Vienna Circle on the use of verification as criterion of demarcation, showing the *metaphysical* nature of their anti-metaphysical program…

[3] Speaking of Fresnel theory, Poincaré says: "The real existence of the ether is of little importance for us, this

The use of probabilistic models for treating uncertainty in artificial intelligence, is directed by the same guidelines. The hypothesis formulated when adopting a probabilistic model, namely the existence of the stochastic object, is simply a working hypothesis and should be considered only as an economical, simple, and useful convention.

3.2 Working Hypotheses for Solving Real–World Problems

Working hypotheses are abstract but not arbitrary, and are always firmly fastened to reality by what we can indeed call a *principle of reality* that is enforced by the goal itself for which such hypotheses are formulated, that is, solving a practical problem.

This is the idea that, for example, inspires mathematicians, logicians and physicists when they recur to the concept of infinite: Such a concept is convenient in calculations, and the problem of its existential status, whatever the answer is, does not interfere with its practical use. If we look back into the history of thinking, we see that, at least since Zeno, the concept of infinite entails paradoxes.[4] This paradoxical nature raised questions on the status of infinity: It is actual or potential? And again, does it have some ontological status in the real world or it is simply an invention of the human mind? However, the fact that the concept of infinity was profitable for formulating and solving analytically problems arising in different domains, prevailed upon any metaphysical difficulty.

Scientific concepts have a precise connection with the real world defined by the ability of making predictions and solving practical problems. This idea is clearly explained by Papoulis [11] through an example extracted from the discipline of circuit theory. In this domain, the typical practical problem that one wishes to solve consists in combining electrical devices in a circuit where the output (the potential of one point of the circuit) varies in time as a given function of the input (the potential of another point). The simplest of the electrical devices, the resistor, *"is commonly viewed as a two–terminal device whose voltage is proportional to the current:* $R = v(t)/i(t)$*. This, however, is only a convenient abstraction. A real resistor is a complex device with distributed inductances and capacitance having no clear specified terminals.* [The above mentioned] *relation can, therefore, be claimed only within certain errors, in certain frequencies ranges, and with a variety of other qualifications. Nevertheless, in the development of circuit theory we ignore all these uncertainties. We assume that the resistance* R *is a precise number satisfying* [the relation $R = v(t)/i(t)$,] *and we develop a theory based on* [such relation] *and on Kirchhoff's laws. It would not be wise, we all agree, if at each stage of the development of the theory we were concerned with the* true *meaning of* R*"*—[11].

Also in our every day life, we use working hypotheses in a very similar way, for example when, buying a new table for our kitchen, we measure a candidate for checking if it will fit into the available space. Measuring a table requires, at least implicitly, to assume a model in which the concept of length is clearly defined. Typically, we leave immediately aside considerations from quantum mechanics or relativity theory and we assume we are furnishing a "classical" kitchen. Further, we assume that the table is a cuboid (rectangular parallelepiped) living in

is the business of metaphysics; The essential for us is that things happen as if it existed and that this hypothesis is convenient for explaining phenomena." Here, by explanation Poincaré means ability of performing correct predictions. Elsewhere in the same work Poincaré states clearly that the first task of the scientist is to make predictions: "*Et avant tout le savant doit prévoir.*"—[14].

[4]Zeno (~ 450 B.C.), philosopher of the Eleatic school, is credited with creating several paradoxes on the subject of motion. Among them, the famous one of the Tortoise and Achilles usually interpreted as a critique of the idea of continuous motion in infinitely divisible space and time.

a three-dimensional Euclidean space. Such assumption implicitly states that some features of the table are irrelevant for solving the task, namely the shape of the legs, the smoothness of the surface and so on. We know that the perfect table does not exist, that the board of a real table is never a perfect rectangle, that it is extremely unlikely that the four legs are exactly the same length, etc… Still, we accept to assume that the imperfections of our table are negligible *in the context of the decision we have to take*. In other words, we accept the working hypothesis that the table is a cuboid: When measuring the table, we proceed *as if* we were measuring the length of the edges of a cuboid in a Euclidean three-dimensional space.

We are convinced that the table exists in Nature independently of ourselves and of our thoughts—we wouldn't buy it otherwise—nevertheless the parameters we measure are not directly features of the table but, strictly speaking, pertain to our model. Between such parameters and the real table there is no direct connection that goes beyond the ability of formulating correct predictions in the context of the problem we are solving: If in our formal reasoning the cuboid fits into our representation of the kitchen, we expect that the real table will indeed fit into the real kitchen, too.

The process described above is the same followed when adopting a probabilistic model. More precisely, the adoption of the working hypothesis that data are observations of a stochastic object, is equivalent, in the kitchen example, to the adoption of the hypothesis that a table is a cuboid. These fictitious assumptions are accepted only because we expect that they will allow us to formulate and to solve some real problem of interest.

3.3 Working Hypotheses in Early Writings on Probability Theory

Seeing the existence of the stochastic object as a working hypothesis is not an *ad hoc* trick for justifying the use of probabilistic models in artificial intelligence. On the contrary, this idea traces back to the early works on probability theory.

This is clear for example in Laplace [6], who is unquestionably one of the fathers of probability theory. Indeed, the stochastic object could find a collocation into Laplace's thinking only under the form of a working hypothesis. In fact, intended as an ontological truth, the existence of the stochastic object would have unacceptably clashed with the rigid determinism that characterized the philosophy of Laplace. For him, reality is governed by uniform laws that, if known, would enable the scientist to reduce the evolution of the world to a totally predictable chain of events. However, despite his deterministic metaphysics, Laplace recurred to probabilistic calculus for studying empirical phenomena that, in his views, were not easily amenable to a deterministic analysis. In other words, as a determinist Laplace could not accept the idea that *chance* plays any role in Nature, but still he accepted the stochastic object as a useful working hypothesis for solving practical problems.

The same idea can be found also in Pascal [12]. According to Pascal, all the situations involving uncertainty are isomorph to the games of chance. In his famous *argument du pari* [12], that is, the analysis of the situation in which a human being bets on the existence of God, Pascal makes precisely the assumption that the very existence of God is the result of a stochastic experiment. Independently from his/her religious beliefs, our reader will easily agree with us that Pascal could not reasonably think that God exists or does not exist as a result of a stochastic experiment. This assumption was not for sure suggested to Pascal by any philosophical or religious belief: Clearly, for Pascal this was only a convenient working hypothesis.

Pascal's *argument du pari* is useful here for discussing another issue on the use of probabilistic models in artificial intelligence. It is often claimed that the use of a probabilistic model in AI for solving problems when little data is available is an improper forcing fit. In our view, this claim can be discussed only on the basis of the record of actual successes and failures of the use of probabilistic models in such contexts. Here, on the basis of Pascal's *pari*, we simply want to point out that the use of probability in cases of full ignorance was already present in the very early works on probability. Therefore, such a practice shouldn't be seen as anything peculiar to artificial intelligence. In the *pari*, no empirical data is available. Frequency data would consist in a record of cases in which God existed and cases in which God did not exist: Again, such record is quite hard even to conceive, independently of one's religious belief. Nevertheless, Pascal describes the bet under the working hypothesis that a stochastic object is responsible for the existence of God. By such hypothesis, he managed to reduce the case of complete ignorance to a framework in which it is treatable as the uncertainty faced in games of chance where frequency data are available.

4 Concluding Remarks

The adoption of a probabilistic model in artificial intelligence implies the working hypothesis that we are dealing with a stochastic object. We stress that the reference to the stochastic object is simply a working hypothesis and by no means a metaphysical statement concerning the *true* nature of the process that generated the data. The statement "in AI we are not dealing with stochastic phenomena," is a metaphysical statement that, as such, is an unacceptable argument in the debate on the adequacy of probability in artificial intelligence. If such a metaphysical argument were declared acceptable, we would open the way to the rejection, on the basis of metaphysical arguments, of any scientific theory. In fact, since science is indeed about formulating working hypothesis on Nature, as we argue in Section 3, rejecting the use of probabilistic models in AI because these are based on fictitious hypotheses implies rejecting the scientific method in general.

Nevertheless, by declaring metaphysical and therefore unacceptable the statement "in AI we are not dealing with stochastic phenomena" we do not mean to reject such a statement *tout court*. We simply refuse to consider it as a valid argument against the use of probability theory in AI, but we accept that it might have, for instance, a *regulative* role. In this sense, we do not exclude that such a metaphysical idea could suggest new techniques for handling uncertainty, but we insist that such new techniques will have to be evaluated only for their ability of producing correct prediction: Metaphysical elements should play no role in their evaluation.

Acknowledgments: Carlotta Piscopo thanks Prof. Marco Dorigo and all the staff of IRIDIA. Mauro Birattari thanks Prof. Wolfgang Bibel, Dr. Thomas Stützle, and all the staff of INTELLEKTIK. The authors wish to thank in particular Gianni Di Caro for the illuminating discussions.

The work of Mauro Birattari was supported by the "Metaheuristics Network," a Research Training Network funded by the Improving Human Potential programme of the CEC, grant HPRN-CT-1999-00106. The information provided is the sole responsibility of the authors and does not reflect the Community's opinion. The Community is not responsible for any use that might be made of data appearing in this publication.

References

[1] G. Berkeley. *A Treatise Concerning the Principles of Human Knowledge*. 1710. Available in: A. A. Luce and T. E. Jessop, editors, *The works of George Berkeley Bishop of Cloyne*. Thomas Nelson and Sons Ltd, London, United Kingdom, 1948-1957.

[2] G. Berkeley. *De Motu; Sive de Motus Principio & Natura, et de Causa Communicationis Motuum*. 1721. Available in: A. A. Luce and T. E. Jessop, editors, *The works of George Berkeley Bishop of Cloyne*. Thomas Nelson and Sons Ltd, London, United Kingdom, 1948-1957.

[3] P. Billingsley. *Probability and Measure*. John Wiley & Sons, New York, NY, USA, 2nd edition, 1986.

[4] P. Cheeseman. In defence of probability. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 1002–1009, San Mateo, CA, USA, 1985. Morgan Kaufmann Publisher.

[5] G. Galilei. *Il Saggiatore*. 1623. Available as: *The Assayer*. In: Stillman Drake and C. D. O'Malley, translators, *The Controversy on the Comets of 1618*. University of Pennsylvania Press, Philadelphia, PA, USA, 1960.

[6] P. S. Laplace. *Essai Philosophique sur les Probabilités*. 1814. Available as: G. J. Toomer, editor, and A. D. Dale, translator, *Philosophical Essay on Probabilities*, Springer Verlag, Berlin, Germany, 1994.

[7] E. Mach. *Die Mechanik in ihrer Entwicklung historisch-kritisch dargestellt*. 1883. Available as: *The Science of Mechanics: A Critical and Historical Account of Its Development*, Open Court Classics, Chicago, IL, USA, 6th edition, 1988.

[8] G. Marsaglia. Die Hard: A battery of tests for random number generators. http://stat.fsu.edu/~geo/diehard.html.

[9] P. Martin-Löef. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[10] J. McCarthy and P. J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Michie, editors, *Machine Intelligence 4*, pages 463–502, Edinburgh, Scotland, 1969. Edinburgh University Press.

[11] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill International Editions, New York, NY, USA, 3rd edition, 1991.

[12] B. Pascal. *Pensées*. 1670. Available in: H. Levi, editor, *Pensées and Other Writings*, Oxford University Press, Oxford, United Kingdom, 1999.

[13] J. Pearl. *Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA, USA, 1988.

[14] H. Poincaré. *La Science et L'Hypothèse*. 1903. Available as: *Science and Hypothesis*, Dover Publications, New York, NY, USA, 1967.

[15] K. Popper. *Logik der Forschung*. 1935. Available as: *The Logic of Scientific Discovery*, Routledge, London, United Kingdom, 1999.

[16] B. Russell. *The Problems of Philosophy*. Williams and Nogate, London, United Kingdom, 1957.

[17] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, USA, 1976.

[18] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185:1124–1131, 1974.

[19] The Scientific World View. The Vienna Circle, 1929. H. Hahn, O. Neurath, and R. Carnap, editors. Manifesto of the Vienna Circle.

[20] L. Wittgenstein. *Tractatus Logico-Philosophicus*. Routledge and Kegan Paul, London, United Kingdom, 1922. German text with an English translation *en regard* by C.K. Ogden.

[21] L. A. Zadeh. Possibility theory and soft data analysis. In L. Cobb and R. M. Thrall, editors, *Mathematical Frontiers of the Social and Policy Sciences*, pages 69–129. Westview Press, Boulder, CO, USA, 1981.