B.1.2. Theories of Uncertainty.

Philippe Smets
IRIDIA. Université Libre de Bruxelles

Abstract: Data are imperfect whenever imprecision or uncertainty pervade them. We define the two concepts, imprecision is partial knowledge about the true state of the world, while uncertainty is an added meta-knowledge that expresses the agent's opinions about which state of affairs prevails. Fuzzy set theory is just an extension of the standard set theory, but it provides a fertile generalization of the conventional set theory. We present several theories for the representation of uncertainty and the mathematical models that have been proposed for it. We discuss their meaning and applicability.

B.1.2.1. Introduction

This section defines the concept of tha agent and of the actual world and compares the concepts of imprecision and uncertainty about the exact naure of the actual world. The presentation is based on propositional logic and possible worlds, even though it could also be focusing on events, subsets, etc.....

The agent and the actual world.

Imprecision and uncertainty deal with the determination by an agent of which world among a set of possible worlds is the actual world given the data available to the agent. The agent, denoted by 'You', is the user, the agent, the program, the robot... that must reason or act with imperfect data. (You is capitalized following Good's tradition and represents the individual, real or fictious, in whose srate of uncertainty we are ineresred (De Finetti, 1974, vol, 1, page 27)). The actual world, denoted by ω_0, is the world that corresponds to the actual state of affairs. It is unique, conceptually well defined but not knowable by You. The term 'world' is used in a general sense. It covers concepts like 'state of affairs', 'state of nature', 'situations', 'context', 'values of a variable'...

Formally, we suppose a finite propositional language $\mathtt{L}$, supplemented by the tautology and the contradiction. Extensions to infinite language are possible but useless here. Let Ω denote the set of worlds that correspond to the interpretations of $\mathtt{L}$. Ω is called the frame of discernment. It is built so that no two worlds in Ω denote logically equivalent propositions, i.e., for every pair of worlds in Ω, there exists a proposition in the language $\mathtt{L}$ that is true in one world and false in the other (this avoids useless repetition of worlds denoting logically equivalent propositions). Among those worlds in Ω, a particular one corresponds to the actual world ω_0. Because the data available to You are imperfect, You do not know exactly which world in a set of possible worlds is the actual world ω_0. All You can express is Your `opinion' about the fact that the world ω_0 belongs to the various subsets of Ω. This opinion can be representing some belief, probability, possibility, etc....

All that counts is that we assume there is an actual world, that this actual world is unique and perfectly well defined, but Your knowledge about its exact value is not perfect. That imprecision and/or uncertainty could be properties of the reality is an open philosophical question. Whatever the answer, it must be recognized that Your picture of the world, that corresponds to the only information You can cope with, never reaches perfection. Data as available to You are always somehow imperfect where 'data' is used as a generic term that covers 'information', 'statements', ...

Imprecision and uncertainty.

Data are perfect when they are precise and certain. Ideally, data should be such that they indicate, in a unique way and with certainty, which of the possible worlds corresponds to the actual world. Usually such perfection is not achieved and we must face imperfect data, i.e., imprecise and/or uncertain data.

To highlight the difference between imprecision and uncertainty, consider the following statements.

1. John has at least two children and I am sure about it.
2. John has three children but I am not sure about it.

In statement 1, the number of children is imprecise but certain. In statement 2, the number of children is precise but uncertain. Both aspects can coexist but are distinct. It is often the case that the more imprecise, the more certain, and the more precise, the less certain.

One must be careful to distinguish precise and certain from true. You can be precise and certain about the nature of ω_0 even though You are wrong. This would be the case should You state “I am sure John has three children” a precise and certain statement but false if John has only two children. When You are certain that a proposition p is true in ω_0, You can say ‘I believe p’, but not ‘I know p’. To believe p does not mean that p is indeed true in ω_0. You can be wrong. The concept of belief and knowledge are distinguished by the fact that when You say ‘I know p’, it means not only that You believe it, but that p is indeed true in ω_0. Uncertainty deals with beliefs, not with knowledge.

General framework.

The problem is to evaluate which worlds among all possible worlds in Ω is the actual world ω_0 , i.e., which world satisfies the data. Ideally, one and only one world in Ω should satisfy the data, in which case the data are precise and certain and ω_0 is known. When no world satisfies the data, data are inconsistent. This problem is not further elaborated here. When several worlds satisfy the data, data are imprecise. If ontop of this imprecision, You can produce weights and/or nuances that express Your opinion about which world among those that satisfy the data might be the actual world, You face uncertainty. This opinion can represent a belief, a probability, a possibility, etc... Imprecision is essenially a property of the data whereas uncertainty is a property of You, the agent.

B.1.2.2. Imprecision.

Idealy, data should be so that for every proposition p in L, one could conclude if p is true or false in the actual world. Whenever such an ideal situation does not hold, in which case more than one world are compatible with the data, data are said to be imprecise.

The concept of imprecision is context dependent as it depends on the level of granularity required for the problem under consideration. If I want to assess the age of Sue, the statement “Sue is born in 1914” is precise if I am satisfied with her age in years, but imprecise if I want to know when her next birthday is. If I want to invite married people for the evening, knowing 'John’s wife is Joan or Ann' is precise, but if I want John’s wife to sit on my right for dinner, it would be imprecise. This problem is the same as defining the appropriate level of specificity of the language in which a problem is to be handled. So precision is relative to the needs of the agent, to the context in which the information is used.

The most obvious way of representing imprecision is through classical set theory. So is it the case when data only indicate that ω_0 belongs to a subset of Ω containing several possible alternatives. Of course, ‘crisp’ set theory is not very expressive in representing imprecision. A major step was achieved when fuzzy sets were introduced, even though the forms of imprecision described by crisp sets or by fuzzy sets are conceptually identical. This similarity is well-illustrated when imprecise pieces of information are combined. If You know that Alex is aged between 41 and 49 from one source of information and between 45 and 52 from another source of information, the result of the combination of these two pieces of information is obtained by the intersection of the two intervals: You will conclude that John is aged between 45 and 49. When fuzzy sets replace the crisp sets like with 'John is not too old' and 'John is about 50', the result of their combination is also obtained by the intersection of the two fuzzy sets. Conceptually, fuzzy sets are just a generalization of conventional sets theory even though the generalization is extraordinarily powerful as it induces an order among those worlds that might be the actual world.

The fact You can only claim that ω_0 belongs to the crisp set A or the fuzzy set *A* reflects imprecision, not uncertainty. When the set is fuzzy, the ordering among the possible worlds could of course be the origin of some induced degrees of uncertainty about the exact nature of ω_0, but this unceratinty is only a secondary phenomena, related to imprecision but not equal to it.

The distinction between imprecision and uncertainty becomes clear in the simple case when the fuzzy set *A* is degenerated into the crisp set A, in which case the availble information '$\omega_0 \in A$' reflects only imprecision. The possible assertions '$\omega_0 \in B$ is probable', '$\omega_0 \in C$ is not very probable', etc... reflect an added uncertainty. Imprecision is not uncertainty. It often precedes it as it provides a set on which uncertainty can be built.

Vagueness and fuzziness.
Vagueness and fuzziness are two forms of imprecisions usually confused (Black, 1937). Nevertheless, Zadeh (see introduction in Kruse et al., 1994) defends that vague is related to ambiguity whereas fuzziness is related to the absence of sharp boundaries. The statement "I will see you at 4:00 p.m. or 5:00 p.m." is vague (or ambiguous) but not fuzzy. The statement "I will see you at about 5:00 p.m." is fuzzy but not vague. The statement "I will see you some time" is both vague and fuzzy.

B.1.2.3. Uncertainty.

Uncertainty results in ignorance (etymologically not knowing). It is essentially an epistemic property induced by a lack of information (Bonissone and Tong, 1985, Clarke et al., 1991, Krause and Clark, 1993, Lopez de Mantaras, 1990, Smithson, 1989).

Objective Uncertainty.
Some specialists have argued that uncertainty related to chance or randomness is an objective property and the term 'likely' qualifies an event that will probably occur. They make the point that the fact 'an event is likely' is independent of Your opinion about the occurrence of the event, and that likelihood (as well as randomness) is an objective property of the experimental set up that generates the event. The concept of the propensity of an event is covered by such objective randomness.

It seems obvious that only possible events can be probable. Therefore, before considering the propensity of some event, its dispositionality might be considered and the term 'possible' qualifies the disposition of an event to occur. This possibility concerns the ability, the capacity of the event to occur, its 'happen ability', whereas probability concerns its tendency to occur. This dispositional property can be seen either as a binary property - an event is or is not possible - or as a graded property. In the last case, some events are more possible than others and possibility can be ordered or even valuated.

Degrees og propensity and disposition are usually represented by probability and possibility measures, respectively.

Subjective Uncertainty.

Objective properties of uncertainty are supposedly linked to the world and to the information. Subjective properties of uncertainty are linked to Your opinion about the exact nature of ω_0 as derived from the information available to You. Some scientists even claim that objective probabilities do not exist, and that only subjective probabilities do (De Finietti, 1974). This subjectivist idea could even be extended to dispositionalities. The most classical model for representing subjective uncertainty is based on probability functions, but contenders have been developed recently, based on possibility functions and on belief functions.

Imprecision usually underpins uncertainty, so could it be that imprecision is a special case of uncertainty? Suppose the pure state of imprecision, when 'all You know is that $\omega_0 \in A$' for some A subset of Ω. One could defend that You are facing uncertainty in a degenerated form. But this level of uncertainty can hardly be described within probability theory. Without extra information, only the Principle of Insufficient Reason could be evoked and each world in A would receive the same probability. This assignment does not stand the test of a change in the granularity of the frame of discernment Ω, and leads to the many incoherences encountered once the Principle of Insufficient Reason is applied routinely. A contrario, the state of knowledge 'all You know is that $\omega_0 \in A$' is easily represented within possibility theory and belief function theory. Hence the concept of imprecision could be seen as just a special case of uncertainty if it were not for the fact this assimilation does not hold for probability theory. The latter explains why we keep the distinction between imprecision and uncertainty.

B.1.2.4. Modeling of Uncertainty.

Uncertainty is accepted here as a property that depends on both the data and the agent, You. You are unceratin about the truth status of a proposition p whenvever You do not know if p is true or not in the actual world ω_0. This unceratiny can be categorical or graded and the mathematical models proposed for representing uncertainty can thus be divided into symbolic-qualitative and numeric-quantitative models.

The symbolic approach is studied within the field of Non Monotonic Logics. Logicians used to focus on developing deduction schemata allowing for the deduction of true conclusions from true premises. However, when it comes to applying these methods to common sense problems, the whole procedure breaks down because it is inadequate in coping with implicit exceptions and statements that are 'usually', 'normally', typically' valid.

A whole class on non-monotonic logics was then introduced in the late 70s (see Bobrow (1980), Ginsberg (1988), Lukaszewicz (1990), Reiter (1987)).

Numerical approaches for representing uncertainty can be split into two families: the non-standard probability and the non-probability models (Kohlas and Monney, 1994). The models based on non-standard probabilities are extensions of models based on probability functions. These include the upper and lower probabilities models (Good, 1950, Smith, 1961, Kyburg, 1987, Walley, 1991, Voorbraak, 1993), Dempster-Shafer's models (Dempster, 1967, Shafer, 1976, Smets and Kennes, 1994), the Hints models (Kohlas and Monney, 1995), the probability of provability models (Ruspini, 1986, Pearl, 1988, Smets, 1991). The non-probabilistic models are not based on probability functions, but on alternative functions like the possibility functions and the belief functions. They include the transferable belief model (Smets, 1988, 1990, Smets and Kennes, 1994), and the possibility theory model (Zadeh, 1978, Dubois and Prade, 1985).

Each model is based on a measure that maps the subsets of the frame of discernment Ω to the [0,1] interval, and is monotone for inclusion. Their major properties are:

Probability Measures: $P:2^{\Omega}\rightarrow[0, 1]$

$$P(A\cup B) = P(A) + P(B) - P(A\cap B) \qquad \forall A,B\subseteq\Omega$$

Possibility Measures: $\Pi:2^{\Omega}\rightarrow[0, 1]$

$$\Pi(A\vee B) = \max (\Pi(A) , \Pi(B)) \qquad \forall A,B\subseteq\Omega$$

Belief Measures: $bel:2^{\Omega}\rightarrow[0, 1]$

$$bel(A\cup B) \geq bel(A) + bel(B) - bel(A\cap B) \qquad \forall A,B\subseteq\Omega$$

A dual of a measure F is defined as the measure DualF whose values are equal to the complement of the value given by F to the complement. In probability theory, the dual is: $DualP(A) = 1 - P(\overline{A})$. Hence DualP = P, and probability functions are auto dual. This property does not hold with the two other measures. The dual measures of the possibility and belief functions are the necessity and plausibility functions, respectively.

Necessity Measures: $N:2^{\Omega}\rightarrow[0, 1]$

$$N(A) = 1 - \Pi(\overline{A}) \qquad \forall A\subseteq\Omega$$

Plausibility Measures: $pl:2^{\Omega}\rightarrow[0, 1]$

$$pl(A) = bel(\Omega) - bel(\overline{A}) \qquad \forall A\subseteq\Omega$$

In fact, they do not provide any new information, but are sometimes more convenient to work with.

B.1.2.5. Possibility and Necessity Measures.

Possibility measure.

Imprecise information such as "John's height is above 170" implies that any height h above 170 is possible and any height equal to or below 170 is impossible. This can be represented by a ‘possibility’ measure defined on the height domain whose value is 0 if h < 170 and 1 if h is ≥ 170 (with 0 = impossible and 1 = possible). A very simple form of modal logic can be used for modeling such a state of knowledge.

When the predicate is vague like in 'John is tall', possibility can accommodate degrees, the largest the degree, the largest the possibility.

Possibility is normally associated with some fuzziness, either in the background knowledge on which the possibility is based, or in the set for which the possibility is assessed.

Given a fuzzy set A defined on Ω and a background knowledge, denoted BK, about which world might be the actual world, the possibility that $\omega_0 \in A$ is usually defined as:

$$\Pi(\omega_0 \in A \mid BK) = \max_{\omega \in \Omega} \min(A(\omega), \pi(\omega \mid BK))$$

where $\pi(. \mid BK)$ is the possibility distribution function derived from BK, and $\pi(x) = \Pi(\{x\})$. For instance, the possibility that ‘John is tall ‘ knowing ‘the height of John is about 180 cm.’ denoted *180*, is:

$$\Pi(\text{h(John)} \in \mathit{Tall} \mid \mathit{180}) = \max_{x \in [0, 300]} \min(\mathit{Tall}(x), \pi(x \mid \mathit{180}))$$

where h(John) denotes the height of John, *Tall* is the set of heights that characterize tall men, [0, 300] is the domain of heights, $\mathit{Tall}(x)$ is the grade of membership of a man of height x to the set of tall men, and $\pi(x \mid \mathit{180})$ is the possibility that h = x given h belongs to the set *180*.

When both BK and A are crisp, the domain of Π reduces itself to $\{0, 1\}$.

A very important point in possibility theory (and in fuzzy set theory) when only the max and min operators are used is the fact that the values given to the possibility measure or to the grade of membership are not essential. The only important element of the measure is the order they create among the elements of the domain. Indeed the orders are invariant under any strictly monotonous transformation. Therefore, a change of scale will not affect conclusions. This property explains why authors insist on the fact that possibility theory is essentially an ordinal theory, a nice property in general. This robustness property does not apply once addition and multiplication are introduced as is the case with probability and belief functions.

Relation between fuzziness and possibility .

Zadeh has introduced both the concept of fuzzy set (Zadeh, 1965, Dubois and Prade, 1980) and the concept of possibility measure (Zadeh, 1978, Dubois and Prade, 1985). The grade of membership express the strength with which a well-known world belongs to an ill-defined set. The degree of possibility express the strength of Your opinion about the exact nature of the actual world given You know that it belings to an ill-defined set.

For instance *Tall*(h) is the grade of membership of a person with height h to the fuzzy set of tall men and $\pi(h|Tall)$ is the possibility that the height of a person is h given the person belongs to the set of *Tall* men. Zadeh's possibilistic principle postulates the following equality:

If $Tall|(h) = x$ then $\pi(h|Tall) = x$ for all $h \in [0,300]$.

It is often written as

$Tall|(h) = \pi(h|Tall)$ for all $h \in H$.

The first expression avoids the confusion between the two concepts. It shows they share the same scale without implying that degrees of possibility and grades of membership are identical concepts.

B.1.2.6. Probability Theory.

Since it was first used as a model for uncertainty in the 17th century, probability has been given at least four different meanings (Fine, 1973, Smithson, 1989).

The classical theory.

The initial definition of probability, as defended by Laplace, assumes the existence of a fundamental set of equipossible events. The probability of an event is then the ratio of the number of favorable cases to the number of all equipossible cases. Of course, the concept of equipossible cases is hardly defined in general. It works with applications where symmetry can be evoked, as it is the case for most games of chance (dice, cards...). Where symmetry cannot be applied, the Principle of Insufficient Reason is evoked (also called Principle of Indifference (Keynes 1962)). Essentially, it states that alternatives are considered as equiprobable if there is no reason to expect the occurence of or prefer any one rother than any other. As attractive as it may seem, the Principle of Insufficient Reason is a very dangerous tool whose application has led to most errors described in probability theory. Very few are prepared to defend it today.

Relative frequency theory.

Probability is essentially the convergence limit of relative frequencies under repeated independent trials (Reichenbach 1949, von Mises 1957). It is strongly related to the concept of proportion. It is by far the most widely accepted definition even though it has been shown not to stand up to criticism. Convergence limits cannot be observed, it postulates propensities observed in the past for events to occur will continue in the future, it does not apply to single events, it suffers from the difficulty of specifying the appropriate reference class, it never explains how long a long run must be that will converge to its limit... Nevertheless, it ‘works’ and this pragmatic argument explains its popularity.

Subjective (Bayesian, personal) probability.
For the Bayesian school of probability, the probability measure quantifies Your belief that an event will occur, that a proposition is true. It is a subjective, personal measure. The justification of the additivity of the probability measures is essentially based on betting behavior arguments. Most Bayesians define P(A), the probability of A, as the fair price p You propose that a player should pay to enter a game against a banker, where the player receives $1 if A occurs and $0 if A does not occur. The concept of fairness is related to the fact that after deciding p You are ready to be either the player or the banker. In order to avoid a Dutch Book (i.e., a set of simultaneous bets that would lead to a sure loss), P must be a probability measure (DeFinetti, 1974, Savage, 1954, Earman, 1992).

Cox (1946) justifies probability measures by requiring essentially that the probability of $\overline{A}$ should be a function of the probability of A, and the probability of ‘A & B’ should be a function of the probability of A given B and the probability of B. Adding a strict monotony requirement leads to the conclusion that the probability measure is the only measure that satisfies both requirements. Dubois et al. (1991) show the limitations of such an argument: strict monotony is not as innocuous at it seems, and the first requirement is exactly what both possibility theory and belief functions theory reject.

Logical probabilities.
Some attempts have been proposed to avoid the subjective component of the Bayesian probability. Keynes (1962) defines the probability as a logical relation between a proposition and a corpus of knowledge. A proposition is probable with respect to a given body of evidence regardless of whether anyone thinks so. The concept of Corroboration introduced by Popper (1959) and the concept of Confirmation introduced by Carnap (1950) both fit in with the overall schema of defining a logical measure of 'probability'.

This program, as intellectually attractive as it may seem, unfortunately fails to explain how the probability weight for these relations can be defined. On that point, the Bayesians have the strongest case since they can use their betting behavior as a guideline on how to assess probabilities. The existence of such operational method to assess a measure of probability is important as it provides a meaning to the '.7' encountered in the proposition “the probability of A is .7”. The lack of such well established and widely accepted operational meaning in fuzzy set theory has been the source of serious criticism (nevertheless see Smets and Magrez, 1988).

Whatever its interpretation, the probability measure P(A) quantifies the degree of 'probability' (whatever 'probability' means) that the actual world ω_0 belongs to A. To enhance the difference between probability and fuzzy set theories, fuzzy set theory concerns the belonging of a well-defined individual to an ill-defined set, whereas probability concerns the belonging of a not yet defined individual to a well-defined set. Probabilities can be extended to fuzzy events (Zadeh, 1968, Smets 1982). As an example let us consider the probability that the next man too enter the room will be a ‘tall man’. Could we say that such a probability is .7 or is that probability itself a fuzzy probability? This is still unresolved, which might explain current lack of interest in that concept.

B.1.2.7. Upper and lower probability models.

Smith (1961, 1965), Good (1950, 1983) and Whaley (1991) suggested that personal degrees of belief cannot be expressed by a single number but that one can only assess intervals that bound them. The interval is described by its boundaries called the upper and lower probabilities. Such intervals can easily be obtained in a two-person situation when one person, Y_1, communicates the probability of some events in Ω to a second person, Y_2, by only saying that the probabilities P(A) belong to an interval, for all $A\subseteq\Omega$. Suppose Y_2 has no other information about the probability on Ω. In that case, Y_2 can only build a set $\mathbf{P}$ of probability measures on Ω compatible with the boundaries provided by Y_1. All that is known to Y_2 is that there exists a probability measure P and that $P\in\mathbf{P}$. Should Y_2 then learn that an event $A\subseteq\Omega$ has occurred, $\mathbf{P}$ should be updated to $\mathbf{P}_A$ where $\mathbf{P}_A$ is this set of conditional probability measures obtained by conditioning the probability measures $P\in\mathbf{P}$ on A. (Smets, 1987, Fagin and Halpern, 1991, Jaffray, 1992).

A similar model can be obtained by assuming that one’s belief is not described by a single probability measure as is the case with the Bayesians but by a family of probability measures.

A special case of upper and lower probabilities has been described by Dempster (1967, 1968). He assumes the existence of a probability measure on a space X and a one-o-many mapping M from X to Y. Then the lower probability of the subset A of Y is equal to the probability of the largest subset of X such that its image under M is included in A. The upper probability of A is the probability of the largest subset of X such that the images under M of all its elements have a non-empty intersection with A.

A generalization of an upper and lower probability model to a second-order probability model is quite straightforward. Instead of just acknowledging that $P \in \mathbf{P}$, one can assume the existence of a probability measure P* on P_Ω, the set of probability measures on Ω.

Second-order probabilities, i.e., probabilities over probabilities, do not enjoy the same support as subjective probabilities. Indeed, there seems to be no compelling reason to conceive a second-order probability in terms of betting and avoiding Dutch books. So the major justification for the subjective probability modeling is lost. Further introduction of second-order probabilities leads directly to a proposal for third-order probabilities that quantifies our uncertainty about the value of the second-order probabilities.... Such iteration leads to an infinite regress of meta-probabilities that cannot be easily avoided (and is identically encountered with higher order fuzzy sets).

B.1.2.8. The Transferable Belief Model.

Information can induce some subjective, personal opinion (henceforth called belief) that a proposition is true. The transferable belief model distinguishes between two mental levels, a ‘credal’ level where beliefs are entertained, and a ‘pignistic’ level where decisions are made (from credo, I believe, and pignus, a bet, both in Latin). The coherence principle advanced by the Bayesians to justify probability measures is adequate at the pignistic level, not at the credal level. Beliefs can be entertained outside any decision context. In the transferable belief model (Smets, 1988, Smets and Kennes, 1994), a non probabilistic model, we assume that beliefs at the credal level are quantified by belief functions (Shafer ,1976). When decisions need to be made, that belief induces a probability measure at the pignistic level and decisions are based on maximizing expected utilities (DeGroot, 1970). The model is very rich and very general. It covers imprecision and probability theory as special cases. Mathematically, it also covers possibility theory but that relation seems more mathematical than conceptual.

B.1.2.9. Combining models of ignorance.

The various forms of ignorance can be encountered simultaneously and it will be necessary to be able to integrate them. Just to give an idea of the problem, consider the following example of generalized modus ponens. What conclusions should be reached?

I strongly believe that:
‘If the company sales is large, then salaries are good'.
It is quite possible that:
‘Company X sales are very large'.

∴ Salaries of Company X employees are ...

To deal with problems like those, beliefs, possibilities, fuzziness need to be combined, and a set of metalanguages need to be constructed. Almost no work has been carried out in this area. However, its importance for data fusion is obvious: when several sensors provide information, how do we recognize the nature of the ignorance involved and select the appropriate model, how do we collapse them into more compact forms, how do we combine them, how do we take the redundancies into consideration, the correlations and the contradictions?

Word Count: 4905 words

Bibliography.

BLACK M. (1937) Vagueness: an exercice in logical analysis. Philosophy of Science, 4:427-455.

BOBROW D. G. (1980) Special Issue on Non-Monotonic Logics, Artificial Intelligence vol. 13.

BONISSONE P.P. and TONG R.M. (1985) Editorial: reasoning with uncertainty in expert systems. I.J.Man Machine Studies 22:241-250.

CARNAP R. (1950) *Logical Foundations of Probability.* Chicago Univ. Press.

CLARKE M.R.B., FROIDEVAUX C. GREGOIRE E. and SMETS Ph. (1991) (eds) *Uncertainty, Conditional and Non Monotonicity: Positions and Debates in Non-Standard Logics.* Special Issue of J. Applied Non-Classical Logics. Vol.1, N° 2.pg. 103-310.

COX R.T. (1946) Probability, frequency and reasonable expectation. Amer.J.Phys. 14:1-13.

DE FINETTI B. (1974) *Theory of Probability.* Wiley, London. Vol.1 and Vol. 2,

DEGROOT M. (1970) *Optimal statistical decisions.* McGraw Hill, New York.

DEMPSTER A.P. (1967) Upper and lower probabilities induced by a multiplevalued mapping. Ann. Math. Statistics 38:325-339.

DEMPSTER A.P. (1968) A generalization of Bayesian inference. J. Roy. Statist. Soc. B.30:205-247.

DUBOIS D. and PRADE H. (1980) *Fuzzy sets and systems: theory and applications.* Academic Press, New York.

DUBOIS D. and PRADE H. (1985) *Theorie des possibilités.* Masson, Paris. (English Version: Theory of Possibility, Plenum, London (1988))

DUBOIS D., GARBOLINO P., KYBURG H.E., PRADE H. and SMETS Ph. (1991) Quantified Uncertainty. J. Applied Non-Classical Logics 1:105-197.

EARMAN J. (1992) Bayes or Bust?. A Critical Examination of Bayesian Confirmation Theory. MIT Press, Cambridge, Mass.

FAGIN R. and HALPERN J. (1991) A new approach to updating beliefs. in Uncertainty in Artificial Intelligence 6 (Bonissone P.P., Henrion M., Kanal L.N. and Lemmer J.F. eds.) North Holland, Amsteram, 347-374.

FINE T. (1973) *Theories of probability*. Academic Press, New York.

GINSBERG M. L. (ed.) (1988) *Readings in Nonmonotonic Reasoning*, Morgan Kaufmann, San Mateo, Ca.

GOOD I.J. (1950) *Probability and the weighting of evidence*. Hafner.

GOOD I.J. (1983) *Good thinking: the foundations of probability and its applications.* Univ. Minnesota Press, Minneapolis.

JAFFRAY J.Y. (1992) Bayesian updating and belief functions. IEEE Trans. SMC, 22:1144-1152.
KEYNES J.M. (1962) *A Treatise on Probability.* New York: Harper and Row.
KOHLAS J. and MONNEY P.A. (1994) Theory of Evidence. A Survey of its Mathematical Foundations, Applications and Computations. ZOR-Mathematical Methods of Operatioanl Research 39:35-68.
KOHLAS J. and MONNEY P.A. (1995) A Mathematical Theory of Hints. An Approach to Dempster-Shafer Theory of Evidence. Lecture Notes in Economics and Mathematical Systems No. 425. Springer-Verlag.
KRAUSE P. and CLARK D. (1993) Representing Uncertain Knowledge. An AI approach. Intellect, Oxford.
KRUSE R., GEBHARDT J. and KLAWONN F. (1994) Foundations of Fuzzy Systems. Wiley, Chichester.
KYBURG H. (1987) Objective probabilities. IJCAI-87, 902-904.
LOPEZ DE MANTARAS R. (1990) *Approximate reasoning models.* Ellis Horwood, Chichester.
LUKASZEWICZ W. (1990) *Nonmonotonic Reasoning*, Ellis Horwood.
PEARL J. (1988) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Pub. San Mateo, Ca, USA.
POPPER K.R. (1959) *The Logic of Scientific Discovery.* New York, Basic Books.
REICHENBACH H. (1949) *The Theory of Probability.* Berkeley: University of California Press.
REITER R. (1987) Nonmonotonic Reasoning, Annual Review of Computer Science, pp. 147-186.
RUSPINI E.H. (1986) The logical foundations of evidential reasoning. Technical note 408, SRI International, Menlo Park, Ca.
SAVAGE L.J. (1954) *Foundations of Statistics.* Wiley, New York.
SHAFER G. (1976) *A mathematical theory of evidence*. Princeton Univ. Press, Princeton, NJ.
SMETS P. (1982) Probability of a Fuzzy Event : an Axiomatic Approach. Fuzzy Sets and systems, 7 : 153-164.
SMETS P. (1987) Upper and lower probability functions versus belief functions. Proc. International Symposium on Fuzzy Systems and Knowledge Engineering, Guangzhou, China, July 10-16, pg 17-21.
SMETS P. (1988) Belief functions. in *Non standard logics for automated reasoning,* edited by SMETS P., MAMDANI A., DUBOIS D. AND PRADE H., Academic Press, London, pp. 253-286.
SMETS P. (1990) The combination of evidence in the transferable belief model. IEEE-Pattern analysis and Machine Intelligence, 12:447-458.
SMETS P. (1991) Probability of provability and belief functions. Logique et Analyse, 133-134:177-195.
SMETS Ph. and KENNES R. (1994) The transferable belief model. Artificial Intelligence 66:191-234.
SMETS P. and MAGREZ P. (1988) The Measure of the Degree of Truth and of the Grade of Membership. Int. J. Fuzzy Sets and Systems. 25:67-72.
SMITH C.A.B. (1961) Consistency in statistical inference and decision. J. Roy. Statist. Soc. B23:1-37.
SMITH C.A.B. (1965) Personal probability and statistical analysis. J. Roy. Statist. Soc. A128, 469-499.
SMITHSON M. (1989) *Ignorance and Uncertainty: Emerging Paradigms.* Springer-Verlag, New York.
VON MISES R. (1957) *Probablity, Statistics and Truth.* (Second Ed.) London, Allen and Unwin.

VOORBRAAK F. (1993) As Far as I Know: Epistemic Logic and Uncertainty. Dissertation, Utrecht University.
WALLEY P. (1991) Statistical reasoning with imprecise probabilities. Chapman and Hall, London.
ZADEH L.A. (1965) Fuzzy sets. Inform.Control. 8:338-353.
ZADEH L.A. (1968) Probability measures of fuzzy events. J. Math. Anal. Appl. 23: 421-427.
ZADEH L.A. (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems. 1:3-28.