Imperfect information :
Imprecision - Uncertainty

Philippe Smets
IRIDIA
Université Libre de Bruxelles

We can do everything. What we cannot do is Useless. So... why to worry about Uncertainty? *The Information Systems Radical.*	T R U T H	What IS can do toady is highly limited. Useless confused with Cannot Be Done. No future without using Uncertainty. *The Uncertainty Zealot.*

Summary: This report surveys various forms of imperfect data, be it imprecision or uncertainty. To that end, a structured thesaurus is proposed. The models that have been proposed to represent imprecision and uncertainty are briefly presented, and their meanings are discussed.

1. Imperfection in Data.

Imperfection, be it imprecision or uncertainty, pervades real world scenarios and must be incorporated into every information system that attempts to provide a complete and accurate model of the real world. But yet, this is hardly achieved by today's information system products. A major reason might be found in the difficulty to understand the various aspects of imprecision and uncertainty.

Is there imprecision and uncertainty in the real world? This is an open question. Whatever the answer, it must be recognized that our picture of the world, that corresponds to the only information we can cope with, never reaches perfection. Data as available for an information system are always somehow imperfect.

Until recently, almost all aspects of imperfect data were modeled by probability theory but in the last 20 years, many new models have been developed to represent imperfect data. The large number of models reflects the recent acknowledgment that there exist many aspects of imperfection and that probability theory, as good as it is, is not the unique normative model that can cope with all of them.

Why should we worry about all these new models when it comes to incorporating them in an information system? The use of inappropriate, unjustified, or purely ad hoc models can lead to outputs that might be misunderstood by the end user. It was the case with the Certainty Factor, where the real meaning of the numbers was not defined, and therefore it lead to the abandon of the model.

Newcomers in the domain are overwhelmed by the multitude of models. Their reaction may be to accept one of them as THE model and use it in every context. Another reaction is to accept all of them and to apply them more or less at random. Both attitudes are wrong and could be seriously misleading.

This paper will present some aspects of imperfection. A full inventory is not practically possible. We propose a classification where imprecision, inconsistency and uncertainty are the major groups. We then present the various approaches that have been proposed to model imprecision and uncertainty. These models are grouped into two large categories : the symbolic and the quantitative models. Details about these models and their use for specific information system application are presented in the subsequent papers in this book. This general introduction focuses only on the ideas underlying the various models, not on their technical details. A detailed analysis of imperfect data is presented in Smithson (1989). Clark (1990) presents a survey of the probability oriented linguistic terms. Some elements of this report were previously published in a shorter version in Smets (1991a). A recent review on the representation of uncertainty in artificial intelligence can be found in Krause and Clark (1993).

2. Variety of imperfect information.

Table 1 presents a structured thesaurus of the aspects of imperfection one can find in a piece of information. Table 2 presents the definition of each aspect as can be found in a classical dictionary.

We use imperfection as the most general label. Information is perfect when it is precise and certain. Imperfection can be due to imprecision, inconsistency and uncertainty, the major aspects of imperfect data.

Imprecision and inconsistency are properties related to the content of the statement: either more than one world or no world is compatible with the available information, respectively. Uncertainty is a property that results from a lack of information about the world for deciding if the statement is true or false. Imprecision and inconsistency are essentially properties of the information itself whereas uncertainty is a property of the relation between the information and our knowledge about the world.

To illustrate the difference between imprecision and uncertainty, consider the following two situations :

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 case 1, the number of children is imprecise but certain. In case 2, the number of children is precise but uncertain. Both aspects can coexist but are distinct. Often the more imprecise you are, the most certain you are, and the more precise, the less certain. There seems to be some Information Maximality Principle that requires that the ‘product’ of precision and certainty cannot be beyond a certain critical level. Any increase in one is balanced by a decrease in the other.

We consider successively and separately the various aspects of imprecision, inconsistency and uncertainty.

2.1. Imprecision.

Imprecision can be characterized by the presence of absence of an error component.

2.1.1. Imprecision without error.

The information ‘the food is hot’ is **ambiguous** (synonymous to **amphibologic**) as the food might be spicy or warm.

‘Age = in the 30’s’ is **approximate** if the age is 36. ‘Age = close to 30’ is a **fuzzy** information. In the first case, it is always **decidable** if the information is correct or not. ‘Age = In the 30’s’ is correct for someone who is 36, whereas it is incorrect for someone who is 28. Decidability is lost with fuzzy information. ‘Age = close to 30’ is more or less correct for both individuals but more correct for the 28 year-old individual than for the 36 year-old one. Correctness admits some kind of degrees once fuzziness is involved.

Data can be strictly **missing** but also **incomplete** as if ‘spouse name = Joan or Jill’ for John. According to the context, incompleteness can be associated with **deficiency** or not. If you want to list only bachelors, the information available for the spouse’s name field, even though incomplete, is not deficient as far as you know that John is not a bachelor. If you want to address John’s wife by her name, the information is not only incomplete but also deficient as you do not know if her name is Joan or Jill.

2.1.2. Imprecision combined with error.

So far, the information is not erroneous. The true value is compatible with the available information. When errors can also be present, many aspects of imprecision can be described. Data will be **erroneous** or **incorrect** when it is just wrong as in ‘age = 37’ when the age is 25. **Inaccurate** data are somehow wrong but the error is small,, for example, ‘Age = 37’ when the actual age is 36, is of course erroneous but not too much, so it might be qualified as inaccurate.

Invalid data are not only erroneous data but data that might potentially be linked to unacceptable conclusions. ‘John’s marital status = widower’ when John is a bachelor is erroneous, but also invalid as a pension alimony will be paid abusively.

Distortion in data is analogous to inaccuracy combined with invalidity. Data is **biased** if all data were subjected to a systematic error. It would be the case if all ages were given as the true age - 2 years. It would be obtained if the age were computed as a difference between the birth date and today’s date, and today’s date is wrong by two years.

Nonsensical and **meaningless** data are extreme aspects of erroneous data. However, the value is so extreme that the user can discover the error instantaneously: ‘age = 245 years’, ‘marital status = apple’.... Meaningless is a less pernicious type of error as it has the flavor of irrelevancy.

2.2. Inconsistency.

When several statements are combined, new aspects of imperfection can appear, in which cases some kind of error is always involved. An information can be **conflicting** : ‘marital status = bachelor’, ‘spouse name = Joan’. The conflict in the data leads to an **incoherence** in the conclusions. Indeed the conclusion drawn from the data is that John is a ‘married bachelor’ as he has a spouse.

Inconsistency is better used in a context when time is involved: at 3 p.m. the eggs were boiled and at 3.15 p.m. the same eggs were fresh.

Logicians used **inconsistency** to define the incoherence that results from a conflicting information, like when you learn that at 3 p.m. the eggs were boiled and at 3.15 p.m. the same eggs were fresh.

Conclusions will be **confused** when incoherent and when the involved incoherence can be recovered by some small modifications of the data : you announce you arrive in Brussels by train at 3.05 p.m. but the train is scheduled to arrive at 3.15 p.m. The incoherence is much smaller than would be the case if there was no train arriving in the afternoon. In the first case, I accept that you will arrive just after 3 p.m. whereas in the second case I hardly know what to accept.

2.3. Uncertainty.

The third aspect of informational imperfection, uncertainty, concerns the state of knowledge of an agent (denoted You, but the agent could even be a computer) about the relation between the world and the statement about the world. The statement is either true or false, but Your knowledge about the world does not allow You to decide if the statement is true or false. Certainty is full knowledge of the true value of the data. Uncertainty is partial knowledge of the true value of the data. Uncertainty results in **ignorance** (etymologically not knowing). It is essentially, if not always, an epistemic property induced by a lack of information. A major cause of uncertainty is imprecision in the data. Whether uncertainty is an objective or a subjective property is a still debated philosophical question left aside here.

2.3.1. Objective Uncertainty.

Some specialists have argued that uncertainty related to **randomness** is an objective property and the term **likely** qualifies an event that will probably occur. They defend that the fact that '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 **propensity** of an event is covered by such objective randomness.

Before discussing the propensity of some event, its **dispositionality** might be considered. Only possible events can be probable. **Possibility** concerns the ability of the event to occur, its 'happen ability' (whereas probability concerns its tendency to occur). Identically, it concerns the ability of a proposition to be true. **Necessity** is the dual of possibility: necessity is the impossibility of the contrary.

2.3.2. 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 true value of the data as derived from the available information.

Data are **believable** or **probable** if You accept them, maybe temporarily. Data are **doubtful** if not believable or at worst if You would be willing to accept them, but then with a very strong reluctance.

The relation between **probability** (equated to belief in the subjective context) and possibility as encountered in objective uncertainty can also be described in the subjective context. **Possibility** and **necessity** are the epistemic properties that reflect Your opinion about the truth statement. In particular, only possible statements can be believed.

Unreliability reflects Your opinion about the source of the data, opinion that is transferred secondarily to the data.

Irrelevance concerns Your opinion about the data and **decidability** concerns Your ability to decide if the information is true or false.

The major source of uncertainty is imprecision. Consider the simplest case, the ambiguity case. When You learn that the food is hot, You are in a state of uncertainty as the food can be either spicy or warm. Nevertheless, that state of uncertainty can be refined. Usually you know more than just the fact that the food is spicy or warm. If you are in a Thaï restaurant, hot will probably mean spicy whereas in an English restaurant, hot will probably mean warm.

Uncertainty usually admits some kind of ordering, therefore it is the privileged domain of application for quantitative modeling, the probability model being the most famous, but hardly the only one today. Imprecision induces uncertainty but the nature of this uncertainty and its quantification will depend on the type of imprecision.

3. Modeling.

Models for imperfect data can be separated into symbolic-qualitative and numeric-quantitative models. Most quantitative models concern uncertainty. An exception is fuzzy sets theory that addresses imprecision. Symbolic models rather concern deduction based on soft knowledge than data representation but of course one cannot create new deduction models without appropriately adapting the data representation by introducing new operators.

3.1. The Symbolic Approach: Non Monotonic Logics.

In the past, logicians have focused on developing deduction schemata that permit the deduction of true conclusions from true premises. However, when it comes to apply these methods to commonsensical problems, the whole procedure collapses.

In classical logic, if you know that Tweety is a bird, you cannot deduce that Tweety flies as the rule ‘every bird flies’ is wrong (because there are exceptions).

Nevertheless, it is a pragmatic commonsense attitude to conclude that Tweety flies until there are reasons to believe otherwise and to go on deducing other facts that are deduced from the fact that Tweety flies. Of course, the new deductions could be defeated by a new piece of evidence. That Tweety flies is a temporary defeasible deduction then can nevertheless be assumed as a working hypothesis. Therefore new deduction operators had to be defined that could model this defeasible reasoning, deduction operators that are non-monotonic in that the set of deductions they allow does not increase monotonically by the adjunction of new pieces of information but could even decrease (when a new piece of information justifies the retraction of previously deduced facts). As an example, the information “Tweety is a penguin” coupled with the information “penguins do not fly” justifies the retraction of the previously non-monotonically deduced fact that Tweety flies and the deduction that Tweety does not fly.

A whole class on non-monotonic logic has been then introduced since the late 70’s. (Bobrow (1980), Ginsberg (1988), Lukaszewicz (1990), Reiter (1987)) The properties of the non-monotonic deduction operators have been defined, special techniques like default logic, hypothetical reasoning, defeasible reasoning have been proposed. Their aim is to reason from rules with non explicit exceptions within classical logic. This logic tries to deduce as much as possible from these rules in that they apply them except when it can be proved that the data are indeed members of the set of exceptions or when inconsistency would be deduced.

The introduction of non-monotonic reasoning is conceptually satisfactory but its cost might become prohibitively expensive. Indeed, in classical logic, any fact, once deduced, will stay true whatever new facts are introduced. With non-monotonic reasoning, this is not a valid strategy. All assumptions used to deduce a fact must be recorded. So whenever an assumption previously accepted turns out to be inapplicable, the deduced facts must be reconsidered. A fact will be retracted, except if another chain of reasoning could still be applied that leads to its deduction. Of course, the storage of all underlying assumptions used to deduce the fact can cause heavy overhead.

3.2. The Quantification of Imprecision: Fuzzy sets.

Imprecision is essentially represented by a disjunctive information that characterizes a set of possible values in which the actual value is known to belong. Recently the classical concept of set has been extended into fuzzy sets that have been used to characterize ‘ordered’ disjunctive information.

Classically, sets are crisp in the sense that one element either belongs to a set or is excluded from it. Zadeh (1965) introduces the idea of non-crisp sets., called fuzzy sets. Fuzziness is a property related to the use of vague predicates like in 'John is tall'.

The predicates are vague, fuzzy because the words used to define them are themselves ill defined, vague, fuzzy (Black 1937). The idea is that belonging to a set admits a degree that is not necessarily just 0 or 1 as it is the case in classical set theory. For some elements of the universe of discourse, one cannot say that it belongs or not to the set. At most one can assess some degree of membership $\mu_A(x)$ of the element x to the fuzzy set A. This function generalizes the classical indicator function $I_A(x)$ of a set:

$$I_A(x) = 1 \text{ if } x \in A$$
$$I_A(x) = 0 \text{ if } x \notin A$$

Zadeh replaces the range {0, 1} by the interval [0, 1].

New concepts like fuzzy numbers (e.g. several, few), fuzzy probability (likely), fuzzy quantifiers (most), fuzzy predicates (tall), and the impact of linguistic hedges (very) can be formalized (Dubois and Prade, 1980). Classical set operators like union, intersection and negation have been generalized. The most classical solution is based on the min-max operators:

$$\mu_{\overline{A}}(x) = 1 - \mu_A(x)$$
$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$$
$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$$

Other operators have been proposed that belong to the family of triangular norms and co-norms (Dubois and Prade, 1985b, Yager, 1991). The generalization of the implication operator turns out to be less obvious, especially when it is considered in the context of the modus ponens as encountered in approximate reasoning (Smets 1991b).

The law of excluded middle does not apply to fuzzy sets. Indeed $\mu_{A \cap \overline{A}}(x) = \min(\mu_A(x), \mu_{\overline{A}}(x))$ can be larger than 0. This must look odd at first sight. This translates nothing but the fact that one can be somehow tall and not tall simultaneously, a perfectly valid property.

Mathematically fuzzy sets theory generalizes the concept of set. The model can be used wherever sets can be used, and therefore is not restricted to any particular form of imperfect data. Its simplest domain of application is the modeling of imprecision and vagueness. Fuzziness creates an order among the possible values in which the actual value is known to belong.

Several authors have tried to disregard fuzzy sets theory by claiming that it is subsumed by probability measure. This Procustean[1] attitude completely misfires.

---

[1] Procuste was a greek bandit. He wanted that travellers passing through his place sleep in his bed. But he wants them to fit the bed perfectly, so when they were smaller

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. Introducing random sets does not change the conceptual picture. Of course there are mathematical relations between the two theories but the problem is not with the mathematical comparison but with a comparison of the problems they try to model. Fuzziness deals with imprecision, probability with uncertainty. Of course, fuzziness induces uncertainty. One could defend that when I know that John is tall, I can build a probability measure on John's height. This does not mean that the grade of membership is a probability (Smets 1985)

3.3. The Quantification of Uncertainty: Sugeno's Fuzzy measures.

Another concept developed by Sugeno (1977) has received the label fuzzy. Sugeno studied functions that express uncertainty associated with a statement 'x belongs to S' where S is a crisp set (generalization to fuzzy sets S is possible but not important here) and x is a particular arbitrary element of X which is not *a priori* located in any of the subset of X. The Sugeno measure g satisfies the following

G1: $g(\emptyset) = 0 \qquad g(X) = 1$

G2: for all $A,B \subseteq X$, if $A \subseteq B$, then $g(A) \leq g(B)$

G3: for all $A_i \subseteq X$, $i \in N$, if $A_1 \subseteq A_2 \subseteq \ldots$ or $A_1 \supseteq A_2 \supseteq \ldots$,
then $\lim_{i \to \infty} g(A_i) = g(\lim_{i \to \infty} A_i)$

The Sugeno measure for finite X is just a normalized measure, monotonous for inclusion. It fits with probability measures, possibility measures, necessity measures, belief functions, possibility measures..... It has been called 'fuzzy' measure but should not be confounded with fuzzy sets.

3.4. Possibility and Necessity Measures.

3.4.1. Possibility measure.

Incomplete information such as "John's height is above 170" implies that any height h above 170 is possible and any height equal 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). Ignorance results from the lack of precision, of specificity of the information "above 170".

---

than the bed, he stretched them, and when they were longer, he chopped off the excess.

When the predicate is vague like in "John is tall', possibility can admit degrees, the largest the degree, the largest the possibility. But even though possibility is often associated with fuzziness, the fact that non fuzzy (crisp) events can admit different degrees of possibility is shown in the following example. Suppose there is a box in which you try to squeeze soft balls. You can say: it is possible to put 20 balls in it, impossible to put 30 balls, quite possible to put 24 balls, but not so possible to put 26 balls...These degrees of possibility are degrees of realizability and totally unrelated to any supposedly underlying random process.

Identically ask a salesman about his forecast about next year sales. He could answer: it is possible to sell about 50K, impossible to sell more than 100K, quite possible to sell 70K, hardly possible to sell more than 90K... His statements express what are the possible values for next year sales. What the values express are essentially the sale capacity. Beside, he could also express his belief about what he will actually sell next year, but this concerns another problem for which the theories of probability and belief functions are more adequate.

Let $\Pi:2^{\Omega}\rightarrow[0, 1]$ be the **possibility measure** defined on a space Ω with $\Pi(A)$ for $A\subseteq\Omega$ being the degree of possibility that A (is true, occurs...). The fundamental axiom is that the possibility $\Pi(A\vee B)$ of the disjunction of two propositions A and B is the maximum of the possibility of the individual propositions $\Pi(A)$ and $\Pi(B)$. (Zadeh 1978, Dubois and Prade, 1985a):

$$\Pi(A\vee B) = \max (\Pi(A) , \Pi(B)). \tag{3.1}$$

Usually one requires also $\Pi(\Omega) = 1$.

As in modal logic, where the **necessity** of a proposition is the negation of the possibility of its negation, one defines the necessity measure N(A) given to a proposition A by:

$$N(A) = 1 - \Pi(\neg A)$$

In that case, one has the following:

$$N(A\wedge B) = \min (N(A) , N(B))$$

Beware that one has only:

$$\Pi(A\wedge B) \leq \min (\Pi(A) , \Pi(B))$$

$$N(A\vee B) \geq \max (N(A) , N(B)).$$

Let Ω be the universe of discourse on which a possibility measure Π is defined. Related to the possibility measure $\Pi:2^{\Omega}\rightarrow[0, 1]$, one can define a **possibility distribution** $\pi:\Omega\rightarrow[0, 1]$,

$$\pi(x) = \Pi(\{x\}) \qquad \text{for all } x\in\Omega.$$

Thanks to (3.1), one has

$$\Pi(A) = \max_{x\in A} \pi(x) \qquad \text{for all A in } \Omega.$$

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 intrinsically 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.

As an example of the use of possibility measure versus probability measure, consider the number of eggs X that Hans is going to order tomorrow morning (Zadeh 1978). Let $\pi(u)$ be the degree of ease with which Hans can eat u eggs. Let p(u) be the probability that Hans will eat u eggs at breakfast tomorrow. Given our knowledge, assume the values of $\pi(u)$ and p(u) are those of table 4.

Table 4: The possibility and probability distributions associated with X.

u	1	2	3	4	5	6	7	8
$\pi(u)$	1	1	1	1	.8	.6	.4	.2
p(u)	.1	.8	.1	0	0	0	0	0

We observe that, whereas the possibility that Hans may eat 3 eggs for breakfast is 1, the probability that he may do so might be quite small, e.g., 0.1. Thus, a high degree of possibility does not imply a high degree of probability, nor does a low degree of probability imply a low degree of possibility. However, if an event is impossible, it is bound to be improbable. This heuristic connection between possibilities and probabilities may be stated in the form of what might be called the **possibility/probability consistency principle** (Zadeh 1978).

3.4.2. Physical and Epistemic Possibility.

Two forms of (continuous valued) possibility have been described: the physical and the epistemic. These 2 forms of possibility can be recognized by their different linguistic uses: it is **possible that** and it is **possible for** (Hacking 1975). When I say it is possible that Paul's height is 170, it means that for all I know, Paul's height may be 170. When I say it is possible for Paul's height to be 170, it means that physically, Paul's height may be 170. The first form, 'possible that', is related to our state of knowledge and is called epistemic. The second form, 'possible for', deals with actual abilities independently of our knowledge about them. It is a degree of realizability. The distinction is not unrelated to the one between the epistemic concept of probability (called here the credibility) and the aleatory one (called here chance). These forms of possibilities are evidently not independent concepts, but the exact structure of their interrelations is not yet clearly established.

3.5. Relation between fuzziness and possibility .

Zadeh has introduced both the concept of fuzzy set (1965) and the concept of possibility measure (1978). The first allows one to describe the grade of membership of a well-known individual to an ill-defined set. The second allows one to describe what are the individuals that satisfy some ill-defined constraints or that belong to some ill-defined sets.

For instance $\mu_{\mathrm{Tall}}(h)$ quantifies the membership of a person with height h to the set of *Tall* men and $\pi_{\mathrm{Tall}}(h)$ quantifies 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 :

$$\pi_{Tall}(h) = \mu_{Tall}(h) \qquad \text{for all } h \in H$$

where H is the set of height = $[0, \infty)$

The writing is often confusing and would have been better written as

$$\pi(h|Tall) = \mu(Tall|h) \quad \text{for all } h \in H$$

or still better

$$\text{If } \mu(Tall|h) = x \text{ then } \pi(h|Tall) = x \qquad \text{for all } h \in H$$

The last expression avoids the confusion between the two concepts. It shows that they share the same scale without implying that a possibility is a membership and *vice versa*. The previous expression clearly indicates the domain of the measure (sets for the grade of membership μ and height for the possibility distribution π) and the background knowledge (the height for μ and the set for π). The difference is analogous to the difference between a probability distribution $p(x|\theta)$ (the probability of the observation x given the hypothesis θ) and a likelihood function $l(\theta|x)$ (the likelihood of the hypothesis θ given the observation x) in which case Zadeh's possibilistic principle becomes the likelihood principle:

$$l(\theta|x) = p(x|\theta)$$

The likelihood of an hypothesis θ given an observation x is equal to the probability of the observation x given the hypothesis θ.

The initial writing of Zadeh's possibilistic principle is the one most usually encountered but its meaning should be interpreted with care. It states that the possibility that a tall man has a height h is equal numerically to the grade of membership of a man with height h to the set of tall men.

The possibility measure Π is related to the possibility distribution π by

$$\Pi_A(X) = \max_{x \in X} \pi_A(x),$$

where X is a crisp set. It generalizes into :

$$\Pi_A(X) = \max_{x \in \Omega} \min(\pi_A(x), \mu_X(x)),$$

where Ω is the domain of x and X is a fuzzy subset of Ω. One can thus express the possibility that the height of a person is about 180 cm given the person is tall.

In general, one has a relation

$$\Pi_A(X \cup Y) = \max (\Pi_A(X) , \Pi_A(Y))$$
$$\Pi_{A \cup B}(X) = \max (\Pi_A(X) , \Pi_B(X))$$

One does not have a similar relation with intersection operators. It is true that

$$\mu_{A \cap B}(X) = \min (\mu_A(X) , \mu_B(X))$$

but one should not deduce that

$$\Pi_{A \cap B}(X) = \min (\Pi_A(X) , \Pi_B(X))$$

as erroneously assumed by careless writers. Nor do we have

$$\Pi_A(X \cap Y) = \min (\Pi_A(X) , \Pi_A(Y)).$$

By duality we have also the following relation between the necessity measure

$$N_A(X \cap Y) = \min (N_A(X) , N_A(Y))$$
$$N_{A \cap B}(X) = \min (N_A(X) , N_B(X))$$

but there is no similar relation for the union operators.

The link between fuzzy set and possibility measure is established through Zadeh's possibilistic principle. An identical principle could also be used to link both fuzzy sets and possibility measure with partial truths. Let us assume.

$$\nu(\text{John is } \mathit{Tall} \mid \text{height(John)} = h) = \mu_{Tall}(h) = \pi_{Tall}(h)$$

So the degree of truth, if such a thing exists, of a proposition 'John is tall' knowing that 'John's height is h' is equated numerically to the grade of membership of a person with height h to the set of *Tall* men and therefore to the possibility that the height of a *Tall* person is h.

3.6. Probability Theory.

Since its beginning as a model for uncertainty in the 17th century, probability has been given at least four different meanings.

Probability measure quantifies the degree of probability P(A) (whatever probability means) that an arbitrary element $X \in \Omega$ belongs to a well-defined subset $A \subseteq \Omega$. It satisfies the following property :

P1: $P(\emptyset) = 0 \qquad P(\Omega) = 1$
P2: For all $A, B \subseteq \Omega$, if $A \cap B = \emptyset$, $P(A \cup B) = P(A) + P(B)$
P3: For all $A, B \subseteq \Omega$, if $P(B) > 0$, then $P(A|B) = P(A \cap B) / P(B)$

where $P(A|B)$ is the probability of that $X \in A$ given it is known that $X \in B$. Such definition can be extended to fuzzy events (Zadeh, 1968, Smets 1982) which further enhances, if still needed, the difference between probability and fuzziness. As an example consider the probability that the next man who enters the room is a tall man. Could we say that such a probability is .7 or is that probability itself a fuzzy probability? This is still unresolved, this might explain today's lack of interest in that concept.

Related to the probability measure $P: 2^{\Omega} \rightarrow [0, 1]$, one defines a probability distribution $p: \Omega \rightarrow [0, 1]$ such that :

$$p(x) = P(\{x\}) \qquad \text{for all } x \in \Omega.$$

By property P2,

$$P(A) = \sum_{x \in A} p(x) \qquad \text{for all } A \subseteq \Omega.$$

Notice that the relation between P and p is the same as the relation between Π and π, (but not as the one between bel and m (as bel and m are both defined on the same frame 2^{Ω}) see the chapter on belief functions).

3.6.1. 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...). When symmetry cannot be applied, the Principle of Insufficient Reason is evoked (also called Principle of Indifference (Keynes 1962)). It essentially states that alternatives are considered as equiprobable if there is no reason to expect or prefer any one over the other. As nice as it might seem, the Principle of Insufficient Reason is a very dangerous tool whose application has led to most errors described in probability theory. It is hardly defended today.

Just to show the danger of its application, suppose all you know about John's wife is that her name is either Joan or Jill or Joey. In that case, it is acceptable that none of the three names is considered as more likely than any other. But is there any reason for the fact 'the name is Joan' to be less likely than the fact 'the name is either Jill or Joey'? Probability theory requires a "yes" answer as the Principle of Insufficient Reason and axiom P2 allocates a probability 1/3 to the first event and 2/3 to the second. Common sense does not require such a clear "yes".

3.6.2. Relative frequency theory.

Probability is essentially the convergence limit of relative frequencies under repeated independent trials (Reichenbach 1949, von Mises 1957). It is not concerned with capturing commonsensical notions, it is a purely prescriptive definition. The definition tries to comply with the operationalist version of scientific positivism: theoretical concepts must be reducible to concrete operational terms. It is strongly related to the concept of proportion, and its direct generalization, measurability (limited to its objective form).

It is by far the most widely accepted definition even though it has been shown not to resist criticisms. Convergence limits cannot be observed, it postulates that past observed propensities for events to occur will continue on into 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 must be a long run that will converge to its limit... Nevertheless, it 'works' and this pragmatic argument explains its popularity.

3.6.3. Subjective (Bayesian, personal) probability.

For the Bayesian school of probability, the probability measure quantifies Your (You is the agent) **credibility** that an event will occur, that a proposition is true. It is a subjective, personal measure.

The additivity of the credibility measure (axiom P2) is essentially based on **betting behavior** arguments. Bayesians define P(A) as the fair price p You propose that a player should pay to play 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 **Dutch book** (i.e., a set of simultaneous bets that would lead to a sure loss), You must assess the probability of the subset of Ω according to P1 and P2. The justification of P3 by diachronic[2] Dutch books (Jeffrey 1988, Teller 1973, 1976) is

---

[2] i.e., when time is involved, where there might be some bets before and after some event.

less convincing as it is based on a Temporal (Diachronic) Coherence postulate (Jeffrey 1988, Earman 1992) that can be objected to. It requires that the way You commit yourself now to organize Your bets after A has occurred if it occurs should be the same as the bets You would accept once A has occurred. The Temporal Coherence claims that hypothetical bets (bets on the hypothesis that A occurs) should be equated to factual bets (bets after A has occurred) (Savage 1954, De Finetti 1974).

Another algebraic justification for the use of probability measure to quantify credibility is based on **Cox's axiom** (Cox 1946). It states essentially that the credibility of $\overline{A}$ (not A) should be a function of the credibility of A, and the credibility of 'A and B' should be a function of the credibility of A given B and the credibility of B. Adding a strict monotonic requirement leads to the conclusion that the probability measure is the only measure that satisfies both requirements (Dubois et al. 1991).

As compelling as Cox's justification seems to be, it can nevertheless be criticized. Strict monotony kills possibility measures and of course possibility theory or belief functions theory advocates reject the first requirement as being not so obvious (Clarke et al. 1991).

3.6.4. Logical probabilities.

Some attempts have been proposed to avoid the subjective component of the Bayesian probability. It fits with the objectivity one likes to defend for scientific rationalism.

Keynes (1962) defined probability as a **logical relation** between a proposition and a corpus of evidence. While propositions are ultimately either true or false (no fuzzy propositions are involved here), we express them as being probable in relation to our current knowledge. A proposition is probable with respect to a given body of evidence regardless of whether anyone thinks so.

Bayesians consider the same kind of relation between knowledge and a proposition but admit it is subjective and therefore that the probability of a proposition is not an objective property that exists 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 with the overall schema of defining a logical measure of probability.

This program, as intellectually attracting as it seems, unfortunately fails to explain **how to define the probability weight** to be given for these relations. On that point the strongest are the Bayesians who 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 and possibility theory, in upper and lower probabilities theory, and in belief functions theory has been the source of serious criticisms (see nevertheless Smets and Magrez (1988) for fuzzy sets theory and Smets and Kennes (1990) for the transferable belief model).

3.7. 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 interval 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 $\mathcal{P}$ 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 $P \in \mathcal{P}$. Should Y_2 learn then that an event $A \subseteq \Omega$ has occurred, $\mathcal{P}$ should be updated to $\mathcal{P}_A$ where $\mathcal{P}_A$ is this set of conditional probability measures obtained by conditioning the probability measures $P \in \mathcal{P}$ on A. (Smets 1987, Fagin and Halpern 1991, Jaffray 1992).

One obtains a similar model by assuming that one’s belief is not described by a single probability measure as do the Bayesians but by a family of probability measures (usually the family is assumed to be convex). Conditioning on some event $A \subseteq \Omega$ is obtained as in the previous case.

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 to many mapping M from X to Y. Then the lower probability of A in 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 in Y 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. In the Artificial Intelligence community, this theory is often called the Dempster-Shafer theory.

A generalization of a upper and lower probability model to **second-order probability** model is quite straightforward. Instead of just acknowledging that $P \in \mathcal{P}$, one can accept a probability measure P* on $\mathbb{P}_\Omega$, the set of probability measures on Ω. So for all $\mathcal{A} \subseteq \mathbb{P}_\Omega$, one can define the probability $P^*(\mathcal{A})$ that the actual probability P on Ω belongs to the subset $\mathcal{A}$ of probability measures on Ω. In that case, the information $P \in \mathcal{P}$ induces a conditioning of P* into $P^*(\mathcal{A}|\mathcal{P}) = P^*(\mathcal{A} \cap \mathcal{P})/P^*(\mathcal{P})$.

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 introducing second-order probabilities directly leads 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.

3.8. Credibility: the Transferable Belief Model.

Information can induce some subjective, personal **credibility** (hereafter called belief) that a proposition is true. Sometimes, its origin can be found either in the **random nature** of the underlying event or in the **partial reliability** that we give to the source of information.

In the first case, one ends up with a probability measure if one accepts the **frequency principle** (Hacking 1965) that, given the chance that a random event X might occur is p, our degree of belief that it will occur is p.

IF chance(X)=p THEN belief(X) = p

This is one of the fundamental requirements for the **classical Bayesian model**, as it relates chance and belief.

When randomness is not involved, there is no necessity for beliefs at the **credal states** (the psychological level where beliefs are entertained) to be quantified by probability measures (Levi 1984). The coherence principle advanced by the Bayesians to justify probability measures is adequate in a context of decision (Degroot 1970), but it cannot be used when all one wants to describe is a cognitive process. Beliefs can be entertained outside any decision context. In the **transferable belief model** (Smets 1988) we assume that beliefs at the credal level are quantified by belief functions (Shafer 1976). When decisions must be made, our belief held at the credal level induces a probability measure held at the so-called **'pignistic' level** (the level at which decisions are made). This probability measure will be used in order to make decisions using expected utilities theory. Relations between belief functions held at the credal level and probabilities held at the pignistic level are given in Smets (1990).

4. Combining models of ignorance.

The various forms of ignorance can be encountered simultaneously and it is necessary to be able to integrate them. In common-sense reasoning, two forms of ignorance, sometime three, are often encountered in the same statement. Just to give an idea of the problem, consider the following example of generalized modus ponens.

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

∴ What can I say about the salary of Company X employees?

This example is probably too complex to be encountered in practice, but it includes most forms of ignorance.

Another combination that is more realistic enhances the links between possibility, beliefs and surprise. Let the two statements:
S1 = ‘It is *quite* possible that John will come *but* I don’t believe he will.’
S2= ‘It is *hardly* possible that John will come *and* I don’t believe he will.’
Suppose You tell me ‘John comes’. My reaction after learning that fact could be translated by: in case 1, ‘well, so he came!’, in case 2, ‘really, are you sure!’. In case 2 I am strongly surprise, whereas in case 1, I am not so surprise. Therefore a modelisation of the concept of ‘surprise’ might have to take in consideration both the degrees of possibility and of beliefs.

To deal with problems like this, beliefs, possibilities, fuzziness need to be combined, and **a set of metalanguages** must be constructed. Care must be given however to what the domains are of each operator. For instance, probability deals with two domains, the set of propositions (as are usually mentioned) and the truth domain (that is usually disregarded as it contains only two elements, but must be considered once fuzzy propositions are accepted).

The first problem is to investigate the **connections** between the probability theory in its frequency approach and the physical possibility theory. The next problem is to investigate the connections between subjective probability measures, belief functions and epistemic possibility measures. Finally, one must establish the connections between the physical properties and the epistemic properties. There is further the problem of extending all these theories when the propositions involved are fuzzy.

Almost no work has been done 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 into consideration the redundancies, the correlations and the contradictions? All these problems must be studied and the implementation of potential solutions tested.

Understanding the meaning of statements and their translation into appropriate models is delicate, if not hazardous. For example, how do we translate "usually bald men are old". Which of P(bald|old) or P(old|bald) is somehow large? "When x shaves, usually x does not die". Which conditioning is appropriate: Pl(dead|shaving) or Pl(shaving|dead)? Is it a problem of plausibility or possibility? These examples are just illustrative of the kind of problems that must be addressed.

5. Conclusions.

The conclusions are more in the form of a plea for ecumenism applicable to both the Information Science and the Uncertainty communities. A plain rejection of the problem of Uncertainty by the first group is highly optimistic. Null values problems are quite old and messy, and many of the proposed solutions are often too shallow to express reality. Beside the use of sophisticate solutions might be unbearable when implemented. As indicated in the introductory remark, truth lays halfway between the Information Sciences Radical and the Uncertainty Zealot attitudes.

We have shown that imprecision and uncertainty are really multiform, and that none of the models available today can fit with all forms of imperfect data. The real problem is to have an open mind attitude and to avoid the dogmatic attitude that leads to claims like: 'I can do *everything* with *my* theory' or 'the forms not covered by *their* models are useless'. We will not insist on the arrogance that underlies such claims. To force fuzzy sets into the probabilitisc mold, or to claim that fuzzy sets theory embeds all other theories are usually unfounded. Each model aims at describing some forms of imperfection. They are complementary, not concurrent.

When confronted with imperfect data, the user should first try to realize the form of imperfection he/she is facing, then see which model is the most appropriate. The real challenge is in recognizing the nature of the imprecision and uncertainty encountered in a given problem. This paper tried to give some hints in that direction. The following papers on uncertainty will present more in depth studies of the major models available today: those based on logic, on probability functions, on possibility functions and on belief functions. Some synthesis is presented in the last papers.

Acknowledgments:
The author is indebted to Hector Garcia-Molina, Joerg Gebhardt, Frank Klawonn, Rudolf Kruse, Henri Prade and the UMIS participants for many suggestions about this presentation.

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.
CLARK D.A. (1990) Verbal Uncertainty Expressions: A review of two decades of research. Current Psychology: Researc and Reviews, 9:203-235.
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. (1985a) *Theorie des possibilités.* Masson, Paris. (English Version: Theory of Possibility, Plenum, London (1988))
DUBOIS D. and PRADE H. (1985b) A review of fuzzy sets aggregation connectives. Information Sciences 36:85-121.
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.
HACKING I. (1965) *Logic of statistical inference.* Cambridge UNiversity Press, Cambridge.
HACKING I.(1975) *The emergence of probability.* Cambridge UNiversity Press, Cambridge.
HECKERMAN D. (1986) Probabilistic interpretation for MYCIN's certainty factors, in KANAL L.N. and LEMMER J.F. (eds) Uncertainty in Artificial Intelligence, North Holland, Amsterdam, pp.167-196.
JAFFRAY J.Y. (1992) Bayesian updating and belief functions. IEEE Trans. SMC, 22:1144-1152.

JEFFREY R. (1988) Conditioning, kinematics, and exchangeability. in Causation, Chance, and Credence, Skyrms B. and Harper W.L. eds., Reidel, Dordrecht,,vol.1, 221-255.
KEYNES J.M. (1962) *A Treatise on Probability.* New York: Harper and Row.
KLIR G.J. (1987) Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like. Int. J. Fuzzy Sets and Systems 24: 141-160.
KRAUSE P. and CLARK D. (1993) Representing Uncertain Knowledge. An AI approach. Intellect, Oxford.
LEVI I. (1984) *Decisions and revisions: philosophical essays on knowledge and value.* Cambridge University Press, Cambridge.
LOPEZ DE MANTARAS R. (1990) *Approximate reasoning models.* Ellis Horwood, Chichester.
LUKASZEWICZ W. (1990) *Nonmonotonic Reasoning*, Ellis Horwood.
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.
SAVAGE L.J. (1954) *Foundations of Statistics.* Wiley, New York.
SHAFER G. (1976) *A mathematical theory of evidence*. Princeton Univ. Press, Princeton, NJ.
SHAFER G. and TVERSKY A. (1985) Languages and designs for probability. Cognitive Sc. 9:309-339.
SHORTLIFFE E.H. and BUCHANAN B.G. (1975) A model for inexact reasoning in medicine. Mathematical Biosciences 23:351-379.
SMETS P. (1982) Probability of a Fuzzy Event : an Axiomatic Approach. Fuzzy Sets and systems, 7 : 153-164.
SMETS P. (1985) Probability of a Fuzzy Event. In : Systems and Control Encyclopedia. M. G. Singh ed. Pergamon, Oxford, 1802-1805.
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) Constructing the pignistic probability function in a context of uncertainty. *Uncertainty in Artificial Intelligence 5*, Henrion M., Shachter R.D., Kanal L.N. and Lemmer J.F. eds, North Holland, Amsterdam, , 29-40.
SMETS P. (1991a) Varieties of ignorance. Information Sciences. 57-58:135-144.
SMETS P. (1991b) Implication and modus ponens in fuzzy logic. in GOODMAN I.R., GUPTA M.M., NGUYEN H.T. and ROGERS G.S. Conditional logic in expert systems. Elsevier, Amsterdam, pg. 235-268.
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.
SUGENO M. (1977) Fuzzy measures and fuzzy integrals: a survey. in M.M. Gupta, G.N. Saridis and B. R. Gaines, eds., Fuzzy Automata and Decision Processes. North Holland, Amsterdam.89-102.
TELLER P. (1973) Conditionalization and Observation. Synthesis 26:218-258.

TELLER P. (1976) Conditionalization, Observation and Change of Prefetrence. in W Harper and C.A. Hooker, eds., Foundations of Probability Theory, Statistical Inference, and Statistical Theory of Science. Reidel, Doordrecht. 205-259.

VON MISES R. (1957) *Probablity, Statistics and Truth.* (Second Ed.) London, Allen and Unwin.

WALLEY P. (1991) Statistical reasoning with imprecise probabilities. Chapman and Hall, London.

YAGER R. (1991) Connectives and quantfiers in fuzzy sets. Int. J. Fuzzy Sets and Systems 40:39-76.

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.

Table 1 : STRUCTURED THESAURUS OF IMPERFECTION.

Imperfect Information.

I. Imprecision: Related to the content of the statement.
Informational property, External world, Negligence.
Several worlds satisfy the statement.

I.1. Data without Error.

Vagueness :

Ambiguous	Has several meanings.
Amphibologic	Has several meanings.
Approximate	Close to reality and well defined.
Vague	Not well defined.

Missing :

Incomplete	Something missing.
Deficient	Missing, applies when needed.

I.2. Data tainted with Error.

Erroneous	Just wrong.
Incorrect	Just wrong.
Inaccurate	Essentially imprecise, but not completely erroneous.
Invalid	Would lead to unacceptable conclusions.
Distorted	Wrong but not far from correct.
Biased	Tainted by a systematic error.
Nonsensical	Cannot be fitted to reality.
Meaningless	Cannot be fitted to reality.

II. Inconsistency.
No world satisfies the statement.

Conflicting	Disagreement among the data.
Incoherent	A property of the conclusions drawn from the data
Inconsistent	Incoherent with a temporal connotation.
Confused	A milder form of incoherence.

III. Uncertainty: Induced by a lack of information, by some imprecision.
Ordering on the several worlds that satisfy the statement.

Property of the information. Objective. External uncertainty.

Propensity.

Random	Subject to change.
Likely	Will probably occur.

Disposition.

Possible	Ability to occur, to be true
Necessary	Negation is impossible, not possible.

Property of the observer. Subjective. Internal uncertainty.

Ignorance, Epistemic property, Internal state of knowledge.

Believable	Observer accepts the data, but is ready to reconsider it.
Probable	Observer accepts the data, but is ready to reconsider it.
Doubtful	Observer can hardly accept the data.
Possible	Observer's considers that the data could be true.
Unreliable	Observer opinion about the source of the data.
Irrelevant	Observer doesn't care about the data.
Undecidable	Inability to decide if true or false.

Table 2 : THESAURUS ON UNCERTAINTY AND INCOMPLETENESS.

Imperfect :
Something is imperfect if it is **incomplete**, **faulty**.

Negligent :
Someone who is negligent fails to deal with something or someone with the right amount of care or concern, or fails to do something which they ought to do.
A **lack of proper care** or attention.

Imprecise :
Something that is imprecise is **not clear**, **accurate**, or precise, not accurately expressed, not scrupulous in being **inexact**.

Vague :
Vague is used to describe things that people say or write that are **not clearly explained or expressed**, so that they can be **understood in different ways**.
It results in **uncertain** or **ill-defined meaning**.

Ambiguous :
Something that is ambiguous is **unclear** or **confusing** because it can have **more than one possible meaning**.
It can be due to vagueness.

Amphibologic :
Synonymous to ambiguous.

Approximate :
An approximate number, amount, time, position, etc. is **close (or similar) to the correct** number, amount, etc., but is probably slightly different from it because it has been calculated quickly rather than exactly.
An idea or description of something that is approximate provides some indication of what it is like but is **not intended to be absolutely precise or accurate**.

Fuzzy :
If your thoughts are fuzzy or what you are thinking about is fuzzy, you are **confused** and cannot see an idea clearly or make a decision. You also describe something as fuzzy when it is **not clearly defined** and is **indistinct** or **vague**.

Missing :
If something is missing, it is not in its place, it is lost, **not present**.
You say that something is missing from a statement, report, etc. when it has not been included in it and you think that it should have been.

Incomplete :
Something that is incomplete does **not have all the parts** that it should have,
Not entered, not filled in.

Deficient :
If someone or something is deficient in a particular thing, they do not have the full amount of it that they need in order to function normally or work properly.
Someone or something that is deficient is not good enough for a particular purpose or standard.
Incomplete or insufficient in **some essential respect**.

Erroneous :

Beliefs, opinions, methods etc. that are erroneous are **incorrect** or **only partly correct.**

Incorrect :
Something that is incorrect is **wrong**, **untrue** , **inaccurate** .

Inaccurate :
Something that is inaccurate is not correct, not precise and **not conforming exactly to a standard** or to truth.

Invalid :
If an argument, conclusion, result, is invalid, it is **not acceptable**, because it is based on a **mistake** . Not sound logically.

Distorted :
If an argument or a statement is distorted, its meaning becomes different and misrepresenting of what it should be.

Biased :
Subject to a **constant error**.

Nonsensical :
That do not make sense, absurd, foolish, stupid, ridiculous, untrue.

Meaningless :
Without any meaning, but also: without importance or relevance.

Conflicting :
If two or more things are in conflict, they are **very different and not compatible.** It seems impossible for each of them to be true, impossible for them to exist together, or for each of them to be believed by the same person.

Incoherent :
If something is incoherent, it is **unclear** and difficult to understand, **rambling** in speech or reasoning.

Inconsistent :
Someone who is inconsistent is **unpredictable** and **behaves differently** is a particular situation each time it happens, rather than doing or saying the same thing each time.
Not compatible or not in harmony
Not constant to the same principles of thought or action.

Confused :
Something is confused if it does not have any order or pattern and is difficult to understand because of this.

Random :
Something that is random happens or is chosen without a definite plan, pattern, or purpose. Made or done without method or conscious choice.

Likely :
Indicates that something is probably the case or will probably happen in a particular situation.

Believable:
Something you think is likely.

Doubtful :
That seems unlikely or uncertain.

Unreliable :
If people, machines, or methods are unreliable, you cannot trust them or rely on them.
That may be not relied on.

Irrelevant :
An irrelevant fact, remark, is not connected with what you are focusing or dealing with, and is therefore not important.
Not related to the matter in hand.

Ignorance :
Lack of knowledge.

Undecidable :
Which validity or truth cannot be decided, is questionable, or on which you cannot make your mind.

Table 3 : THEORIES OF PROBABILITY.

A Taxonomy of the Normative or Prescriptive Interpretations of the Concept of Probability.

Classical Theory.
Definition: Probability of A is the ratio of the number of 'favorable' cases for A to the number of all 'equipossible' cases.
Assumes the existence of a fundamental set of 'equipossible' events.
Equipossibility can be justified by some arguments of symmetry.
Extended to cover any situation characterized by a lack of prior knowledge about the propensity of fundamental events.

Relative Frequency Theory.
Definition: The convergence limit of relative frequency under repeated independent trials.
Related to proportion and (objective) measures.

Subjective (Bayesian) Theory.
Definition: The probability is the degree of belief, of credibility that a proposition is true, that an event will occur, given the agent's corpus of evidence.
The relation between the proposition (event) and the corpus of evidence is subjective.

Operational Definition: P(A) is the prize a player is ready to pay for buying a game where he receives $1 if A occurs, and $0 otherwise.
Probability Measures necessary to satisfy rationality in the context of Decision Making.

Logical (or A Priori) Theory
Definition: The probability of A is a weighted logical relation between the proposition A and a corpus of evidence.
The relation between the proposition and the corpus of evidence is objective.

Table 4 : MODELS FOR UNCERTAINTY ON FINITE FRAMES.
The major axioms of each model (not complete)

Sugeno Measures : $g:2^{\Omega}\rightarrow[0, 1]$

$g(\emptyset) = 0 \qquad g(\Omega) = 1$

$g(A)\leq g(B) \qquad \forall A,B\subseteq\Omega,\ A\subseteq B.$

Special cases:

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

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

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

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

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

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

Belief function,

in DB A present with proba .6,,, A then B with .7 so B present with .42, cf Clarke.

To be considered as to be added???

? RMS (JTMS, ATMS...) Kruse?
Certainty Factors?

3. Application Example
Relational DB — Field Uncertainty
Tuple Uncertainty.
Key Uncertainty?
Query.

? Learning. ≠ Updating KM
Revision AGM
Conditioning. Bayes rule

CLARK D.A. (1990) Verbal Uncertainty Expressions: A critical review of two decades of research.

Appendix.

An Unified Representation of Quantified Imprecision and Uncertainty.

In order to describe the possible interrelations between the various forms of quantified representations for uncertainty and imprecision, we present a data representation that covers many cases, if not all. It is evidently a very heavy representation. Its generality is obtained at the cost of its impracticality. Nevertheless its interest is in the fact it enhances the exact meaning of the various quantities.

The overall idea is to represent families of labelled subsets of some space X. The set X is called the **object-domain**. The labels themselves come from the **label-domains** and qualified the elements and subsets of the object-domain X.

Unfortunately none of the various measures of uncertainty and imprecision has been endowed with any **unit**. Should they have been, many errors and confusion would have been avoided. Nobody is going to equate $ and £ or inches and cm, or worse inches and $. So why not to create those units? May we suggest to call them respectively but adding the postfix 'it' for 'information unit' like in the bit. (We use probit even though the term is used in a totally different meaning in statistics.)

fuzzit : the unit for the grade of membership of an element to a fuzzy set.
probit : the unit for a probability measure
possit: : the unit of a possibility measure
necit : the unit of a necessity measure
belit : the unit of belief measure

plausit : the unit of a plausibility measure

For crisp sets, if $x \in A$, its grade of membership to A is 1 fuzzit, if $x \notin A$, 0 fuzzit.
Due to the dual nature between possibility and necessity measures, one has:

" $\Pi(A) = \alpha$ possit " is equivalent to " $N(\neg A) = (1 - \alpha)$ necit "

Due to the dual nature between belief and plausibility measures, one has:

" $bel(A) = \alpha$ belit " is equivalent to " $pl(\neg A) = (1 - \alpha)$ plausit "

One should not write $bel(\Omega) = pl(\Omega)$, but "$bel(\Omega) = 1$ belit" is equivalent to "$pl(\Omega) = 1$ plausit".

The **general form for a family of labelled subsets** of some space X is:

$[\![\, [\, \{(x_1, \mu_1), (x_2, \mu_2), ...(x_n, \mu_n)\}, \{(w_1, \nu_1), (w_2, \nu_2), ...(w_m, \nu_m)\} \,], [....] \, ... \, [....] \,]\!]$

We admit the convention that any pair (x, 0) contained in a list { } has been omitted.

The symbols mean :

x an element of the object-domain X.
μ an object-weight in fuzzit,
the grade of membership of the element x to a subset of X.
(x, μ) (object-element, object-weight)
$\{(x, \mu) ...( )\}$ a subset of X (crisp if all μ are 1, otherwise fuzzy)

w an element of the label-domain L
(in probit, possit or belit depending on L)
ν a label-weight in fuzzit,
the grade of membership of the element w to a subset of L.
(w, ν) (label-element, label-weight)
$\{(w, \nu), ...( )\}$ a label in L that applies to the subset of X.(a label is a subset of L)

[{ }, { }] = [subset S of X, label of S] a labelled subset of X

$[\![$ [subset S_1 of X, label of S_1], [subset S_2 of X, label of S_2] ... [ ] $]\!]$
a family of labelled subsets of X.

Example 1: A subset.
Let X = (1, 2, 3, 4, 5). Let C be the set {1, 2, 3}.
The representation is:
⟦ [{(1, 1), (2, 1), (3, 1)}, {(1, 1)}] ⟧
? problem what unit for the (1, 1)?is it just a set, or a set with a proba = 1 of plaus or bel or nec or poss = 1.

Example 2: A fuzzy subset.
Let X = (1, 2, 3, 4, 5). Let A be the fuzzy set with $\mu_A(1) = 1.$, $\mu_A(2) = .6$, $\mu_A(3) = .3$.
The representation is:
⟦ [{(1, 1), (2, .6), (3,.3)}, {(1, 1)}] ⟧

Example 3: A probability distribution over X.
Let X = (1, 2, 3, 4, 5).
Let p be a probability distribution with p(1) = .2, p(2) = .5, p(3) = .3.
The representation is:
⟦ [{(1, 1)} , {(.2, 1)}] , [{(2, 1)} , {(.5, 1)}] , [{(3, 1)} , {(.3, 1)}] ⟧

Example 4: A fuzzy probability distribution over X.
Let X = (1, 2, 3, 4, 5).
Let p be a probability distribution with
p(1) = .2,
p(2) is the fuzzy set: $\mu_{p(2)}(.5) = 1.$, $\mu_{p(2)}(.6) = .3$, $\mu_{p(2)}(.7) = .1$, $\mu_{p(2)}(.4) = .8$.
p(3) is the fuzzy set: $\mu_{p(3)}(.3) = 1.$, $\mu_{p(3)}(.2) = .3$, $\mu_{p(3)}(.1) = .1$, $\mu_{p(3)}(.4) = .8$.
The representation is:
⟦ [{(1, 1)} , {(.2, 1)}] ,
[{(2, 1)} , {(.5, 1) , (.6, .3) , (.7, .1) , (.4, .8)}] ,
[{(3, 1)} , {(.3, 1) , (.2, .3) , (.1, .1) , (.4, .8)}] ⟧

Example 5: A probability distribution over fuzzy subsets of X (fuzzy random sets)
Let X = (1, 2, 3, 4, 5).
Let the fuzzy subsets
A with $\mu_A(1) = 1.$, $\mu_A(2) = .6$, $\mu_A(3) = .3$
B with $\mu_B(2) = .7$, $\mu_B(3) = 1.$, $\mu_B(4) = .5$
Let p be a probability distribution with
p(A) = .2,
p(B) = .8
The representation is:
⟦ [{(1, 1), (2, .6), (3,.3)}, {(.2, 1)}] ,
[{(2, .7), (3, 1.), (4,.5)}, {(.8, 1)}] ⟧

A dubious example as $p(A) + p(B) = 1$ but $A \cup B \neq X$.

Example 6: A possibility distribution over X.
Let X = (1, 2, 3, 4, 5).
Let π be the possibility distribution with $\pi(1) = .4$, $\pi(2) = 1.$, $\pi(3) = .3$.
The representation is:
$[\![$ [{(1, 1)} , {(.4, 1)}] , [{(2, 1)} , {(1., 1)}] , [{(3, 1)} , {(.3, 1)}] $]\!]$

Example 7: A possibility measure over crisp subsets of X.
Let X = (1, 2, 3, 4, 5).
Let $\Pi(\{1, 2\}) = .7$, $\Pi(\{2,3,4\}) = .5$.
The representation is:
$[\![$ [{(1, 1) , (2, 1)} , {(.7, 1)}] , [{(2, 1) , (3, 1), (4, 1)} , {(.5, 1)}] $]\!]$

Example 8: A possibility measure over fuzzy subsets of X.
Let X = (1, 2, 3, 4, 5).
Let the fuzzy subsets

A with $\mu_A(1) = 1.$, $\mu_A(2) = .6$, $\mu_A(3) = .3$
B with $\mu_B(2) = .7$, $\mu_B(3) = 1.$, $\mu_B(4) = .5$

Let Π be a possibility measure with
$\Pi(A) = .4$,
$\Pi(B) = 1.$
The representation is:
$[\![$ [{(1, 1), (2, .6), (3,.3)}, {(.4, 1)}] ,
[{(2, .7), (3, 1.), (4,.5)}, {(1., 1)}] $]\!]$

Example 9: A basic belief assignment over crisp subsets of X.
Let X = (1, 2, 3, 4, 5).
Let a bba with $m(\{1\}) = .2$, $m(\{1, 2\}) = .1$, $m(\{1, 4, 5\}) = .7$
The representation is:
$[\![$ [{(1, 1)}, {(.2, 1)}] , [{(1, 1) , (2, 1)}, {(.1, 1)}]
[{(1, 1), (4, 1), (5, 1)}, {(.7, 1)}] $]\!]$

Example 10: A basic belief assignment over fuzzy subsets of X.
Let X = (1, 2, 3, 4, 5).
Let the fuzzy subsets

A with $\mu_A(1) = 1.$, $\mu_A(2) = .6$, $\mu_A(3) = .3$

B with $\mu_B(2) = .7$, $\mu_B(3) = 1.$, $\mu_B(4) = .5$

Let Π be a possibility measure with

m(A) = .4,

m(B) = .6

The representation is:

$\llbracket$ [{(1, 1), (2, .6), (3,.3)}, {(.4, 1)}] ,

[{(2, .7), (3, 1.), (4,.5)}, {(.6, 1)}] $\rrbracket$