Reliable Computing (2006) 12: 153–165
DOI: 10.1007/s11155-006-4874-2

Dealing with Imprecise Probabilities: Interval-Related Talks at ISIPTA'05

What was the conference? From July 20–24, 2005, the 4th International Symposium on Imprecise Probabilities and Their Applications (ISIPTA'05) was held on the campus of Carnegie Mellon University in Pittsburgh, Pennsylvania, USA. This biennial conference is organized by the Society for Imprecise Probability Theory and Applications (SIPTA) [13]. The main organizers of ISIPTA'05 were Fabio G. Cozman (University of Sao Paulo, Brazil), Robert Nau (Duke University, USA), and Teddy Seidenfeld (Carnegie Mellon University). Papers presented at ISIPTA'05 were authored by researchers from a very diverse list of countries: Belgium, Brazil, the Czech Republic, France, Germany, Israel, Italy, New Zealand, Russia, Slovenia, Spain, Switzerland, the UK, and the US.

This ISIPTA conference had an interesting format that had been found valuable in previous conferences. The first day was devoted to tutorials, starting with an introduction to imprecise probabilities given by Gert de Cooman, President of SIPTA. Starting the second day, each regular paper was presented in a plenary session, which was followed by a poster session where each presenter of that session also presented a poster and actively communicated with interested participants. A special session was devoted to open problems, and the conference was immediately followed by a half-day workshop on a key application area, financial risk assessment.

Next we review subjective probabilities as traditionally handled. With that as background, we review the imprecise probability approach. In this review, we mainly follow the tutorial given by G. de Cooman. This review is followed by problems, challenges, and steps toward their resolution presented by researchers at the conference.

Subjective probabilities through preferences and lotteries. To better understand the ideas and techniques behind imprecise probabilities, let us start by describing a traditional approach to subjective probabilities, with particular reference to an important application of probabilistic analysis, decision making; for details see, e.g., [11].

A person's rational decisions are based on the relative values to the person of different outcomes. In financial applications, the value is usually measured in monetary units such as dollars. However, the same monetary amount may have different values for different people: e.g., a single dollar is likely to have more value to a poor person than to a rich one. In view of this difference, in decision theory, to describe the relative values of different outcomes, researchers use a special *utility* scale instead of the more traditional monetary scales.

There are many different ways to elicit utility from decision makers. A common approach is based on preferences of a decision maker among *lotteries*. A simple way to define a lottery is as follows. Take a very undesirable outcome A^- and a very desirable outcome A^+, and then consider the lottery $A(p)$ in which we get A^+ with probability p and A^- with probability $1 - p$ (p is given and is usually understood as an "objective" probability). Clearly, the larger p, the more preferable $A(p)$: $p < p'$ implies $A(p) < A(p')$. Traditional decision theory is based on assumptions concerning preferences over lotteries. For example, the following two assumptions are usually adopted as axioms:

- the comparison amongst lotteries is a linear order—i.e., a person can always select one of the two alternatives, and
- the comparison is Archimedean—i.e. if for all $\varepsilon > 0$, an outcome B is better than $A(p - \varepsilon)$ and worse than $A(p + \varepsilon)$, then it is of the same quality as $A(p)$: $B \sim A(p)$ (where $A \sim B$ means that A and B are of the same quality).

Because of our selection of A^- and A^+, most reasonable outcomes are better than $A^- = A(0)$ and worse than $A^+ = A(1)$. Due to linearity, for every p, either $A(p) < B$, or $B \sim A(p)$, or $B < A(p)$. If we define the *utility* of outcome B as $u(B) \stackrel{\text{def}}{=} \sup\{p \mid A(p) < B\}$, we have $A(u(B) - \varepsilon) < B$ and $A(u(B) + \varepsilon) > B$; thus, due to the Archimedean property, we have $A(u(B)) \sim B$. This value $u(B)$ is called the *utility* of the outcome B.

Comment. As defined above utility always takes values within the interval $[0, 1]$. It is also possible to define utility to take values within other intervals. Indeed, note that the numerical value $u(B)$ of the utility depends on the choice of reference outcomes A^- and A^+. If we select a different pair of reference outcomes, then the resulting numerical utility value $u'(B)$ is different. The usual axioms of utility theory guarantee that two utility functions $u(B)$ and $u'(B)$ corresponding to different choices of the reference pair are related by a linear transformation: $u'(B) = a \cdot u(B) + b$ for some real numbers $a > 0$ and b. By using appropriate values a and b, we can then re-scale utilities to make the scale more convenient (e.g. in financial applications, closer to the monetary scale).

From our definition of the utility function, if an event E has an objective (frequency-based) probability p, the utility of the lottery "A^+ if E and A^- otherwise" is exactly p. Therefore, to gauge the subjective probability of an arbitrary event E, we can form a lottery "A^+ if E and A^- otherwise" and define the subjective probability of E as the utility of this lottery. In other words, the subjective probability $P(E)$ of E is the value p for which the lottery "A^+ if E, otherwise A^-" is equivalent to "A^+ with probability p, otherwise A^-."

Often, we have a "branching" situation involving n incompatible events $E_1, \dots, E_n$ such that exactly one of them will occur. E.g. coins can fall heads or tails, dice can show 1 to 6, etc. In such situations, for every n outcomes $B_1, \dots, B_n$, we can form a lottery by assigning outcome B_i if event E_i occurs. If we know the utility $u_i = u(B_i)$ of each outcome B_i, and we know the (subjective) probability $p_i = P(E_i)$ of each event E_i, then the utility of the corresponding lottery may be determined as follows.

We know the subjective probability p_i of each event E_i. Thus, the lottery "B_i if E_i" if equivalent to the lottery in which we get B_i with probability p_i. The fact that $u(B_i) = u_i$ means that each B_i is equivalent to getting A^+ with probability u_i and A^- with probability $1 - u_i$. By replacing each B_i with this new "lottery," we conclude that the lottery "if E_i then B_i" is equivalent to a two-step lottery in which we:

- first select E_i with probability p_i, and
- then, for each i, select A^+ with probability u_i and A^- with the probability $1 - u_i$.

In this two-step lottery, the probability of getting A^+ is equal to $p_1 \cdot u_1 + \cdots + p_n \cdot u_n$ (often this is obtained by adding suitable axioms on combination of lotteries, but the meaning should be intuitive here). Thus, by our definition of utility, the utility of the lottery "if E_i then B_i" is equal to $u = \sum_{i=1}^{n} p_i \cdot u_i = \sum_{i=1}^{n} p(E_i) \cdot u(B_i)$. In mathematical terms, u is the expected value of the utility, so this approach is often called the *expected utility approach*.

In the traditional approach, between several alternatives we select the one with the largest utility u, hence the one with the largest value of the expected utility.

The "traditional" approach in mathematical terms. Events can be naturally described as subsets E_i of the set Ω of possible states of the world; this set is called the *sample space*.

From this viewpoint, a lottery can be described as a function that assigns, to each state $\omega \in \Omega$, a value $f(\omega)$: E.g., a lottery "if E_i then B_i" means that $f(\omega) = u_i = u(B_i)$ when $\omega \in E_i$. Different researchers use different terms to describe such mappings, random variables being typical but random quantities also common. In the imprecise probability community, often such mappings $f : \Omega \to R$ are called *gambles*—a term made popular by P. Walley [15]. This term emphasizes the subjective component of most theories that deal with imprecision in probability values, and their reliance on preferences and choices. However, it should be pointed out that several researchers in the community of imprecise probabilities seek a more "objective" viewpoint in which "gambles" and "preferences" have no place. An important example is K. Weichselberger's theory of interval probabilities [16]. The term "random variable" is thus employed both by those with a more "objectivist" bent and by those who want to stick to a well-known term.

Now let's consider the expected utilities $u(f)$ of different gambles further. The terms *expected value* and *expectation* are widely used in this context; a less well-known term that is often used in the imprecise community is the term *prevision*. This term was originally proposed by B. de Finetti, one of the founders in the foundations of probability and decision making; the term has been made popular in the imprecise probability community again through the work of P. Walley. A prevision maps a gamble into a real number; a prevision $u(f) = \sum_{\omega} p(\omega) \cdot f(\omega)$ (introduced above) is a linear functional on the linear space of all gambles.

According to the above description, a natural way to describe this functional is to describe the probability $P(E)$ of each event $E \subseteq \Omega$—i.e., in mathematical terms, to describe a *probability measure* P.

It is well known that we do not have to describe all the values $P(E)$ for all subsets E: since $P(E) = \sum_{\omega \in E} p(\omega)$, it is sufficient to describe the *probability density*, i.e., the probabilities $p(\omega)$ of individual states ω.

Computational aspects of the "traditional" approach. From a purely mathematical viewpoint, all three approaches—prevision (linear functional), probability measure, and probability density—are equivalent (at least when the set Ω is finite).

However, from a computational viewpoint, there is a major difference between the three representations.

- When we have n possible states $\omega \in \Omega$, then, to describe the probability density $p(\omega)$, we need to store n values $p(\omega)$ corresponding to different states ω.
- To explicitly describe the probability measure $P(E)$, we need to store the values $P(E)$ corresponding to all 2^n subsets $E \subseteq \Omega$—i.e., we need to store 2^n values. For large enough n, e.g. $n \geq 500$ to 1000, this value exceeds the number of particles in the Universe and is thus not realistic.
- Finally, to explicitly describe the prevision, we must store infinitely many different values $u(f)$ for infinitely many gambles f. When we approximate each value $f(\omega)$ by a value from a k-element set S_k, then each approximate gamble $f(\omega)$ is a function from the n-element set Ω to a k-element set S_k. There are k^n such functions $f : \Omega \to S_k$; so, to directly describe a mapping $u(f)$, we need to store k^n values of $u(f)$. For $k > 2$, we have $k^n \gg 2^n$, so for large n, representing all these values is even less realistic than representing the probability measure.

As a result, in the traditional approach, the probability density is most commonly used in practical applications.

From the "traditional" approach to imprecise probabilities. Utility theory is typically based on the assumption that the preference relation is a *linear* order, so that we can always select between two given alternatives. In practice, if the consequences of two alternatives are similar, it is very difficult to select between them. Imprecise probabilities are obtained from such *partially* ordered preferences.

Each gamble $f(\omega)$ can still be compared with different lotteries "A^+ with probability u and A^- otherwise," but we can no longer guarantee that there is only one u for which the gamble is equivalent to this lottery. Instead there are, in general, two values: the largest value $\underline{u}(f)$ for which the lottery is worse than f or is equivalent to f, and the smallest value $\overline{u}(f)$ for which the lottery is better than f or equivalent to f. When $u \in (\underline{u}(f), \overline{u}(f))$, there is no guidance on whether the lottery is better or the gamble is better.

In other words, we have an *interval-valued* prevision $\mathbf{u}(f) = [\underline{u}(f), \overline{u}(f)]$ which maps each gamble f into an interval. In economic terms,

- $\underline{u}(f)$ is the largest price (in "utiles"—units of utility) that we are willing to pay to participate in the gamble f, and
- $\overline{u}(f)$ is the smallest price for which we are willing to sell our right to participate in the gamble.

Interval-valued previsions form a general description framework for imprecise probabilities (although, as we will remark later, an even more general description is sometimes needed). However, as mentioned earlier, explicitly describing an interval-valued prevision requires that we store an unrealistic number (k^n) of intervals.

In principle, similarly to representing a prevision in terms of a probability measure, we can represent an interval-valued prevision as a set of the corresponding probability measures—namely, probability measures that correspond to all the previsions $u(f) \in \mathbf{u}(f)$. However, this time, this representation does not seriously decrease the computational complexity, because we need arbitrary *convex* sets $\mathcal{P}$ of probability measures. Convex sets of probability measures (and sometimes non-convex sets as well) are often called *credal sets*, after the *credal states* introduced by I. Levi [9].

To decrease the computational complexity, it is often reasonable to restrict ourselves to a specific case of *interval-valued* probability measures $\mathbf{P}(E) = [\underline{P}(E), \overline{P}(E)]$, where $\mathcal{P}$ is the box $\{P \mid P(E) \in \mathbf{P}(E) \text{ for all } E\}$. Some natural conditions cannot be represented in this form however, because an interval-valued probability cannot express any arbitrary set of probabilities—and two different sets of probability measures represent two different interval-valued previsions. For example, the condition that $P(A) \geq 2P(B)$ for some events A and B is not a condition that interval-valued probability can exactly represent, because the constraint is over a linear combination of events, not simply over events.

Even representing an interval-valued probability measure still requires storing a large number (2^n) of intervals. To further decrease the computational complexity, we may want to consider an even narrower class of imprecise probabilities which correspond to interval-valued probabilities $\mathbf{p}(\omega) = [\underline{p}(\omega), \overline{p}(\omega)]$. In this case, to represent the uncertainty, we only need to store n intervals.

Other computationally interesting classes include *p-boxes*, where interval-valued probabilities are given only for sets $(-\infty, x)$—i.e., where we only know the bounds $\mathbf{F}(x) = [\underline{F}(x), \overline{F}(x)]$ on the cumulative distribution function (cdf) $F(x) \stackrel{\text{def}}{=} \text{Prob}(\xi \leq x)$.

Yet another class that has received great attention (due to conceptual and computational properties) is the class of Choquet capacities [2]. Sub-classes of special interest are 2-monotone capacities (which are set-functions with the properties of interval-valued measures plus the condition that $\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$ for disjoint events A and B), and the class of infinitely monotone capacities, also known as *belief functions* in Dempster-Shafer theory. In the Dempster-Shafer (DS) approach we have n sets $S_1, \ldots, S_n$ with "masses" m_i ($\sum m_i = 1$). Then, the class $\mathcal{P}$ of possible probability measures P is described as follows: We select n probability measures P_i such that P_i is located on the set S_i with probability 1 (i.e., $P_i(S_i) = 1$). Then, for every set $E \subseteq \Omega$, we define $P(E) \stackrel{\text{def}}{=} \sum\limits_{i=1}^{n} m_i \cdot P_i(E)$. The corresponding lower probability $\underline{P}(E) \stackrel{\text{def}}{=} \inf\limits_{P \in \mathcal{P}} P(E)$ defines an infinitely monotone capacity.

Interval-related comment on the DS approach. In DS, for each event E, we can compute the corresponding bounds for the probability $P(E)$ as $\overline{P}(E) = \sum\limits_{S_i \cap E \neq \emptyset} m_i$ and $\underline{P}(E) = \sum\limits_{S_i \subseteq E} m_i$. The lower bound is called the *belief* in E, and the upper bound is called the *plausibility* of E. Vice versa, if we know the bounds $\overline{P}(E)$ for each event E, we can uniquely reconstruct the sets S_i and the masses m_i.

From the interval viewpoint, it seems reasonable to consider an interval version of the DS approach, where for each set S_i, we only know the interval $\mathbf{m}_i = [\underline{m}_i, \overline{m}_i]$ of possible values of the corresponding mass. Here, the possible probability measures are $p(E) = \sum m_i \cdot p_i(E)$, where $m_i \in \mathbf{m}_i$ and each probability measure $p_i(E)$ is located on the set S_i. For this interval-valued DS approach, we can find the bounds on $P(E)$ as $\overline{P}(E) = \min\left(\sum\limits_{S_i \cap E \neq \emptyset} \overline{m}_i, 1\right)$ and $\underline{P}(E) = \sum\limits_{S_i \subseteq E} \underline{m}_i$. However, it is no longer possible to uniquely reconstruct the mass intervals $\mathbf{m}_i$ from these bounds. For example, for $\Omega = \{a, b, c\}$ and $\mathbf{m}(\{a, b\}) = \mathbf{m}(\{b, c\}) = \mathbf{m}(\{a, c\}) = [0, 0.6]$, we have $\overline{P}(E) = 1$ for all $E \neq \emptyset$ and $\underline{P}(E) = 0$ for all $E \neq \Omega$; however, we have the exact same bounds for $P(E)$ if we take, e.g., $\mathbf{m} = [0, 0.7]$ instead of $[0, 0.6]$.

Research problems. To use imprecise probabilities, we must elicit them, update them, and use them to make conclusions and produce decisions. Papers presented at IPIPTA'05 covered all stages of this process:

- how to *elicit* imprecise probabilities; in particular, how to set up a *learning process* that would eventually enable us to get a good description of imprecise probabilities;
- how to *update* them when new information appears; and
- how to make *inferences* from the known information and how to make *decisions* based on this information.

It is important to note that independence relations usually significantly decreases the computational complexity of a problem. For example, if we have two variables each of which takes n values, then:

- to describe a general probability distribution, we must describe n^2 values $p(x, y)$, but

- on the other hand, if we know that x and y are independent, we only need to describe $2n \ll n^2$ values $p(x)$ and $p(y)$, and we will then be able to reconstruct $p(x, y)$ as $p(x) \cdot p(y)$.

In view of this importance, the conference included papers describing how we can define and use independence for imprecise probabilities. An important aspect of this is that the concept of independence itself becomes ambiguous in the case of imprecise probabilities, and different kinds of independence can be specified and their differences understood.

A type of model that has received significant attention, and where independence relations are extensively explored, is the credal network model. A credal network can be understood as a generalization of the popular Bayesian network model, where a graph is used to specify distributions and independence relations in a visually pleasant and computationally efficient manner. Credal networks were the object of several papers presented at ISIPTA'05.

A lot of interesting mathematical and algorithmic results presented at ISIPTA'05 were interval-related; it is difficult to adequately describe all these results without making this report too long. We will therefore concentrate on the results that have an applied character.

Practical applications. Most current work on imprecise probability focuses on the representation of subjective uncertainty. Thus, the corresponding techniques are most useful in situations when we do not have objective (frequency-based) probabilities, and we have to rely on subjective human expertise. This often happens, e.g., when we try to predict consequences of behavior of difficult-to-analyze complex systems.

Applications described at the conference included the following.

- Applications to *climate change* models were described by A. P. Dempster (Harvard University), J. Lawry and G. Fu (University of Bristol, U.K.), J. W. Hall (University of Newcastle-upon-Tyne, U.K.), and E. Kriegler (Potsdam Institute of Climate Impact Research, Germany); in particular, A. P. Dempster noted the need for caution when deciding which part of climate change is due to natural variation and which is human-induced.
- Applications to *ecology*, in particular, to soil contamination, were described by C. Baudrit and D. Dubois (Institut de Recherche en Informatique de Toulouse, France).
- Applications to elicitation of *medical* knowledge were described by A. A. Silva and F. M. Campello de Souza (Universidade Federal de Pernambuco, Recife, Brazil).
- Application to *biometrics* (in particular, to pose estimation) were given by P. Cuzzolin (University of California at Los Angeles) and R. Frezza (Università di Padova, Italy).

- Applications to *financial* problems were described in the tutorials given by P. Cheridito (princeton University), S. D'Silva, J. B. Kadane, M. Schervish, T. Seidenfeld (Carnegie Mellon University), R. T. Rockafeller (University of Washington), S. Uryasev (University of Florida), P. Vicig (Universita di Trieste, Italy), and M. Zabarankin (Stevens Institue of Technology). Practical applications to the optimization of electric company portfolios were proposed by D. Berleant and G. Sheblé (Iowa State University) and M. Dancre and J. P. Argaud (Electricité de France).

In financial applications, there is an additional important issue. An interval-valued prevision describes the price at which the user is willing to sell his or her participation in a gamble and the price at which the user is willing to buy this participation. When these prices are reasonably small, it is reasonable to expect that there will be someone else willing to, correspondingly, buy or sell this participation. However, when we go to large-scale transactions, we may not find a seller or a buyer—or a transaction may even be banned by, e.g., anti-trust law. For example, one might expect to be able to sell a small piece of gold, but it is not so clear that there would be a ready buyer if the US decided to sell all its gold reserves.

To describe actual financial transactions, it is therefore desirable to describe not the price for which the user is willing to sell, but rather the price that the user can expect to get. This idea was discussed by P. Vicig. In his description, the main difference between a generalized prevision $u(f)$ describing this price and the standard prevision is that the homogeneity $u(\lambda \cdot f) = \lambda \cdot u(f)$ of a standard prevision is replaced by a weaker condition of convexity. Other relevant issues were discussed in detail at the workshop on financial risk assessment.

Beyond the standard approach to imprecise probabilities. In some practical situations, the interval-valued prevision model for describing subjective probabilities is too crude, and more complex models are needed to describe human preference and decision making.

Several such situations were described at the conference, both situations in which we need to go beyond interval-valued previsions, and situations where we need to take objective probabilities into account. Let us describe these situations in more detail.

Beyond interval-valued previsions. The traditional approach to decision making is based on the assumptions that the preference ordering of outcomes and lotteries is linear ($a \leq b$ or $b \leq a$) and that it satisfies the Archimedean property. Since the interval-valued prevision model is based on standard utility theory, these assumptions are, to some extent, preserved in the interval-valued prevision model—and both these assumptions are only approximately true.

Beyond the Archimedean assumption. The *Archimedean* assumption does not take into account the empirical phenomenon that people prefer choices with more information; this phenomenon is known as the *Ellsberg paradox*. An example follows.

Suppose that we have an event E with frequency $1/2$, and an event E' whose frequency is completely unknown. Empirically, most people prefer to get "A^+ if E, otherwise A^-" (i.e., "A^+ with probability 1/2, otherwise A^-") rather than "A^+ if E', otherwise A^-." We have defined subjective probability of an event E as the value u for which "A^+ if E, otherwise A^-" is equivalent to "A^+ with probability u, otherwise A^-." Thus, the subjective probability $P(E')$ of the event E' does not exceed $1/2$.

On the other hand, one can easily see that the same relation holds for the negations $\neg E$ and $\neg E'$: $\neg E$ holds with the frequency $1/2$, and the frequency of the event $\neg E'$ is completely unknown. Thus, most people will prefer "A^+ is not $\neg E$, otherwise A^-" to "A^+ if $\neg E'$, otherwise A^-." Hence, we have $P(\neg E') \leq 1/2$. Since $P(\neg E') = 1 - P(E')$, we thus have $P(E') \leq 1/2$ and $P(E') \geq 1/2$—i.e., $P(E') = 1/2$.

So, if we compare two lotteries: "1 utile if E and 0 otherwise" and "1 utile if E' and 0 otherwise," we see that both lotteries have exactly the same expected utility $1/2$, but the first is preferable to most persons. So, here we have a fine distinction between lotteries with the same value of expected utility. Such situations can be described by multi-dimensional utilities and/or multi-dimensional (e.g., infinitesimal) probabilities; see, e.g., [11].

This fine structure is especially important for events of probability 0—because there is a clear difference between an impossible event $E = \emptyset$ for which the probability is exactly 0 and a rare event (e.g., $E = \{0\}$ for a uniform distribution on the interval $[0, 1]$) which is theoretically possible. This distinction was described by B. Vantaggi (Università di Roma "La Sapienza," Italy).

In geometry, this distinction is captured by the notions of Hausdorff dimension and the corresponding Hausdorff measure. For example, within a 3-D space, a surface has Hausdorff dimension 2, and its Hausdorff measure is the area of this surface; a curve has Hausdorff dimension 1, and its length is the Hausdorff measure of this curve. It is therefore reasonable, for an event of probability 0, to consider its Hausdorff dimension and Hausdorff measure as a degree of its subjective probability. This was described by S. Doria from Università G. D'Annunzio, Chieti, Italy. In effect, this is similar to using multi-dimensional probabilities, because we have a continuum of new values where previously we had a single value $P(E) = 0$.

Another important situation is estimating the subjective probability of a statement S as the proportion of the experts who consider this statement to be true. When $t(n)$ out of n experts believe in S, then we take $p \approx t(n) / n$. We start with the best available experts. If there are too few of them to make a reasonable estimate for p, we would want to ask more experts. Usually $t(n)$ is proportional to n, so when n increases, we indeed get a better estimate for p. However for certain types of statements such as those describing surprising discoveries which have not yet become common knowledge, we may have $t(n) / n > 0$ when we start with the best experts, but, as n grows, $t(n) / n \to 0$. Such cases can be described, e.g., by the dependence $t(n) \approx a \cdot n^{\alpha}$, where $\alpha < 1$; in this case, $\lim t(n) / n = 0$, but clearly the subjective probability of this statement is higher than of a statement that everyone considers false ($t(n) = 0$ for every n); see, e.g., [12].

Beyond linearity. While we earlier loosened the assumption of linear ordering for events, we still assumed that in the gamble "B_i if E_i," each outcome B_i has a well-defined utility u_i, i.e., that utilities come from a linearly ordered set of real numbers. In practice, the utility of outcomes is also only partially ordered. To adequately describe the corresponding interval-valued (and more general) gambles, we can use multi-valued probabilities; see, e.g., [4], [5], [14].

Beyond crisp partial orders. In the interval-valued prevision approach, we assume that for every two alternatives a and b, the user will always select one of the three options:

- definitely prefer a,
- definitely prefer b, or
- definitely indicate that neither a not b can be preferred.

In real life, however, the transition is not crisp but *fuzzy*: When we decrease the advantages of a over b, the user's preference will not change abruptly but rather gradually change from "definitely better" to "probably better" to "somewhat better" to "I don't know." In other words, a user describes his or her preferences by using words from natural language. To describe such preferences, it is therefore reasonable to use fuzzy logic, a technique specifically designed to describe such words in computer understandable form. Such fuzzy-valued previsions have indeed been successfully used in several applications, including applications to climate change presented at the conference by J. Lawry, J. Hall, and G. Fu.

Objective probabilities. Issues in dealing with objective probabilities include combining them with subjective probabilities, and applying to them the imprecise probability concept.

Combining subjective probabilities with objective (frequency-based) probabilities. Often, we need to combine subjective probabilities with frequency-based ones. For example, the statement "80% of birds fly" describes a frequency, while the statement "I am 90% sure that this particular bird flies" describes a subjective probability, and the statement "I am 90% sure that at least 80% of birds fly" combines subjective and frequency-based probabilities. Such combinations are difficult to describe adequately, and when appropriately described, lead to difficult computational problems; these problems were enumerated in a talk by T. Lukasiewicz (Università di Roma "La Sapienza," Italy).

Can objective probabilities be imprecise? In standard physics, we assume that for enough random events, the frequencies tend to a limit—the (objective) probability. In situations like chaotic dynamics, however, it is reasonable to consider events for which the frequencies do not tend to any definite limit. Such situations were described and analyzed by L. C. Rego and T. L. Fine (Cornell University) in terms of Kolmogorov complexity (for an introduction to Kolmogorov complexity see, e.g., [10]).

Session on computational challenges. A special open problems session was devoted to computational challenges. The session tried to focus on a few challenges, rather than cover the (too many!) possible questions one might try to formulate concerning imprecise probabilities.

A quick introduction was given by the session organizer F. G. Cozman (University of Sao Paulo, Brazil), who noted that in many problems related to imprecise probabilities, we deal with characteristics $c(p)$ such as moments, the probabilities within a certain interval, values $F(t)$ of the cdf, etc., which are linear in terms of the unknown probabilities $p(\omega)$. A typical practical problem is that we know the ranges of some of these characteristics $\underline{a}_i \leq c_i(p) \leq \overline{a}_i$, and we want to find the range $[\underline{a}, \overline{a}]$ of the possible values of some other characteristic $c(p)$. For example, we might have the bounds on the marginal cdfs $F_x(t)$ and $F_y(s)$ of x and y, and we want to find the range of the cdf $F_{x+y}(z)$ for $x + y$ at a particular point z.

In such problems, to find $\underline{a}$ (correspondingly, $\overline{a}$), we must minimize (correspondingly, maximize) the linear objective function $c(p)$ under linear constraints—i.e., solve a linear programming (LP) problem. There are known efficient algorithms and software for solving LP problems, and they are actively used in imprecise probabilities.

However, there are important practical problems which lie outside LP, e.g., problems involving independence, when constraints are linear in $p(x, y) = p(x) \cdot p(y)$ and thus, bilinear in $p(x)$ and $p(y)$. Some of these problems were emphasized by B. Vantaggi. She also mentioned that even without independence, constraints on conditional probabilities $P(A \mid B)$ sometimes present the following computational challenges.

- If $P(B) > 0$, then, because of the definition of the conditional probability $P(A \mid B) = P(A \,\&\, B) \,/\, P(B)$, e.g., the constraint $P(A \mid B) \leq p_0$ is equivalent to a linear constraint $P(A \,\&\, B) \leq p_0 \cdot P(B)$.
- However when $P(B) = 0$, there is no such easy reduction. One possible way to solve this problem may be to consider probabilities whose values are not real numbers but rather elements of a lexicographically ordered vector space.

Another important problem she noted is to design clarifying graphical representations of the corresponding concepts and results, such as credal networks and similar models. These would help the decision maker to take the corresponding information into account.

T. Lukasiewicz emphasized the computational problems related to representations that allow combination of probabilities and first-order sentences (that is, sentences in logic that allow quantifiers "for all" and "there exists," plus relations and functions). Among other things, such representations can be used to combine statements about subjective and objective probabilities in expert systems. However, the computational complexity of the resulting problems is very high, and currently no viable algorithm exists that can process such general problems.

V. Kreinovich (University of Texas at El Paso) described the following computational challenges related to *data processing* under imprecise probabilities.

- We can extend formulas of interval arithmetic to the cases when in addition to the interval, we also know bounds on the cdf or on the first moment. However, it is still difficult to extend them to the case when we know *both* bounds on cdf and on the first moment, or when we also know bounds on the second moment.
- There are many open problems related to extending statistical formulas like population average, population variance, correlation, etc., to the case when we only know the intervals $\mathbf{x}_i$ of possible values of the sample values x_i.

D. Berleant (Iowa State University) described related challenges regarding arithmetic on p-bounds [1]. These include:

- how to deal with p-boxes which are only known with some limited certainty;
- how to make decisions under such uncertainty; and
- how to do *back-calculation*—i.e., find bounds on an input x_i given the desired uncertainty for the result $y = F(x_1, \ldots, x_n)$ and for other inputs.

Beyond the talks. The banquet was held in the one of the city's main attractions: the world-famous Andy Warhol Museum (Warhol was born and grew up in Pittsburgh). I. Levi gave the banquet talk, focusing on the arguments in favor and against convexity of credal sets.

Pittsburgh is the largest inland port in the US. The conference organized a dinner cruise on the city's picturesque "three rivers": the Allegheny and Monongahela Rivers which meet to form the Ohio River (an important tributary of the Mississippi River).

The next conference will be held in Summer 2007, most likely in Prague. See you there!

References

1. Berleant, D., Cheong, M.-P., Chu, C., Guan, Y., Kamal, A., Sheblé, G., Ferson, S., and Peters, J. F.: Dependable Handling of Uncertainty, *Reliable Computing* **9** (6) (2003), pp. 407–418.
2. Choquet, G.: Theory of Capacities, *Annales de l'Institute Fourier* **5** (1953), pp. 131–195.
3. Cozman, F. G., Nau, R., and Seidenfeld, T.: *Proceedings of the 4th International Symposium on Imprecise Probabilities and Their Applications ISIPTA'05*, Pittsburgh, July 20–24, 2005.
4. de la Mora, C., Wojciechowski, P., Kreinovich, V., Starks, S. A., Tanenbaum, P., and Kuzminykh, A.: Robust Methodology for Characterizing System Response to Damage: A Subjective (Fuzzy) Partial Ordered Modification of the Traditional Utility-Probability Scheme, in: *Proceedings of the 22nd International Conference of the North American Fuzzy Information Processing Society NAFIPS'2003, Chicago, Illinois, July 24–26, 2003*, pp. 413–419.
5. Kosheleva, O., Kreinovich, V., Nguyen, H. T., and Bouchon-Meunier, B.: How to Describe Partially Ordered Preferences: Mathematical Foundations, in: Nguyen, H. Ph. and Ohsato, A. (eds.),

Proceedings of the Vietnam-Japan Bilateral Symposium on Fuzzy Systems and Applications VJFUZZY'98, HaLong Bay, Vietnam, 30th September–2nd October, 1998, pp. 269–278.
6. Kuznetsov, V. P.: Auxiliary Problems of Statistical Data Processing: Interval Approach, in: *Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, TX, February 23–25, 1995*, *Reliable Computing* (1995), Supplement, pp. 123–129.
7. Kuznetsov, V. P.: Interval Methods for Processing Statistical Characteristics, in: *Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, TX, February 23–25, 1995*, *Reliable Computing* (1995), Supplement, pp. 116–122.
8. Kuznetsov, V. P.: *Interval Statistical Models*, Radio i Svyaz, Moscow, 1991 (in Russian).
9. Levi, I.: *The Enterprise of Knowledge*, MIT Press, Cambridge, 1980.
10. Li, M. and Vitanyi, P. M. B.: *An Introduction to Kolmogorov Complexity*, Springer, N.Y., 1997.
11. Luce, R. D. and Raifa, H.: *Games and Decisions*, Dover, N.Y., 1989.
12. Nguyen, H. T., Kreinovich, V., and Wu, B.: Fuzzy/Probability ~ Fractal/Smooth, *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems (IJUFKS)* **7** (4) (1999), pp. 363–370.
13. SIPTA: `http://www.sipta.org`.
14. Tanenbaum, P. J., de la Mora, C., Wojciechowski, P., Kosheleva, O., Kreinovich, V., Starks, S. A., and Kuzminykh, A.: Robust Methodology for Characterizing System Response to Damage: Approach Based on Partial Order, in: Lirkov, I., Margenov, S., Wasniewski, J., and Yalamov, P. (eds), *Large-Scale Scientific Computing, Proceedings of the 4th International Conference LSSC'2003, Sozopol, Bulgaria, June 4–8, 2003*, *Lecture Notes in Computer Science* **2907**, Springer, 2004, pp. 276–283.
15. Walley, P.: *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall, N.Y., 1991.
16. Weichselberger, K.: *Elementare Grundbegriffe einer allgemeinerem Wahrscheinlichkeitsrechnung I—Intervallwahrscheinlichkeit als umfassendes*, Physica, Heidelberg, 2001.

Daniel Berleant, Fabio G. Cozman, Olga Kosheleva, and Vladik Kreinovich