Computation with imprecise probabilities*

Lotfi A. Zadeh

Department of EECS, University of California,
Berkeley, CA 94720-1776
E-Mail: zadeh@eecs.berkeley.edu

Extended abstract

An imprecise probability distribution is an instance of second-order uncertainty, that is, uncertainty about uncertainty, or uncertainty for short. Another instance is an imprecise possibility distribution. Computation with imprecise probabilities is not an academic exercise—it is a bridge to reality. In the real world, imprecise probabilities are the norm rather than exception. In large measure, real-world probabilities are perceptions of likelihood. Perceptions are intrinsically imprecise, reflecting the bounded ability of human sensory organs, and ultimately the brain, to resolve detail and store information. Imprecision of perceptions is passed on to perceived probabilities. This is why real-world probabilities are, for the most part, imprecise.

Peter Walley's seminal work "Statistical Reasoning with Imprecise Probabilities," published in 1991, sparked a rapid growth of interest in imprecise probabilities. Today, we see a substantive literature, conferences, workshops and summer schools. An exposition of mainstream approaches to imprecise probabilities may be found in the 2002 special issue of the Journal of Statistical Planning and Inference (JSPI), edited by Jean-Marc Bernard. My paper [3] is contained in this issue but is not a part of the mainstream. A mathematically rigorous treatment of elicitation of imprecise probabilities may be found in [1].

The approach which is outlined in the following is rooted in my 1975 paper [5] , and is in the spirit of [3]. The approach is a radical departure from the mainstream. Its principal distinguishing features are: (a) imprecise

*Research supported in part by ONR N00014-02-1-0294, BT Grant CT1080028046, Omron Grant, Tekes Grant, Chevron Texaco Grant and the BISC Program of UC Berkeley.
Dedicated to Peter Walley.

probabilities are dealt with not in isolation, as in the mainstream approaches, but in an environment of imprecision of events, relations and constraints; (b) imprecise probabilities are assumed to be described in a natural language. This assumption is consistent with the fact that a natural language is basically a system for describing perceptions. The capability to compute with information described in natural language opens the door to consideration of problems which are not well-posed mathematically. Following are very simple examples of such problems.

1. X is a real-valued random variable. What is known about X is: (a) usually X is much larger than approximately a; and (b) usually X is much smaller than approximately b, with $a < b$. What is the expected value of X?

2. X is a real-valued random variable. What is known is that $Prob(X$ is small) is low; $Prob(X$ is medium) is high; $Prob(X$ is large) is low. What is the expected value of X?

3. A box contains approximately twenty balls of various sizes. Most are small. There are many more small balls than large balls. What is the probability that a ball drawn at random is neither large nor small?

4. I am checking in for my flight. I ask the ticket agent: What is the probability that my flight will be delayed. He tells me: Usually most flights leave on time. Rarely most flights are delayed. What is the probability that my flight will be delayed?

To compute with information described in natural language we employ the formalism of Computing with Words (CW) [2] or, more generally, NL-Computation [6]. The formalism of Computing with Words in application to computation with information described in natural language involves two basic steps: (a) precisiation of meaning of propositions expressed in natural language; and (b) computation with precisiated propositions. Precisiation of meaning is achieved through the use of generalized-constraint-based semantics, or GCS for short. The concept of a generalized constraint is the centerpiece of GCS. Informally, generalized constraints, in contrast to standard constraints, have elasticity. What this implies is that in GCS everything is or is allowed to be graduated, that is, be a matter of degree. Furthermore, in GCS everything is or is allowed to be granulated. Granulation involves partitioning of an object into granules, with a granule being a clump of elements drawn together by indistinguishability, equivalence, similarity, proximity or functionality.

A generalized constraint is an expression of the form $XisrR$, where X is the constrained variable, R is the constraining relation and r is an indexical variable which defines the modality of the constraint, that is, its semantics. The principal modalities are: possibilistic ($r =$ blank), probabilistic ($r = p$), veristic ($r = v$), usuality ($r = u$) and group ($r = g$). The primary constraints are possibilistic, probabilistic and veristic. The standard constraints are bivalent possibilistic, probabilistic and bivalent veristic. In large measure, scientific theories are based on standard constraints.

Generalized constraints may be combined, projected, qualified, propagated and counterpropagated. The set of all generalized constraints, together with the rules which govern generation of generalized constraints from other generalized constraints, constitute the Generalized Constraint Language (GCL). Actually, GCL is more than a language—it is a language system. A language has descriptive capability. A language system has descriptive capability as well as deductive capability. GCL has both capabilities.

The concept of a generalized constraint plays a key role in GCS. Specifically, it serves two major functions. First, as a means of representing the meaning of a proposition p, as a generalized constraint; and second, through representation of p as a generalized constraint it serves as a means of dealing with p as an object of computation. Representing the meaning of p as a generalized constraint is equivalent to precisiation of p through translation into GCL. In this sense, GCL plays the role of a meaning precisiation language. More importantly, GCL provides a basis for computation with information described in a natural language, CW or more generally, NL-Computation.

A concept which plays an important role in computation with information described in a natural language is that of a granular value. Specifically, let X be a variable taking values in a space U. A granular value of X, $^*u$, is defined by a proposition, p, or more generally by a system of propositions drawn from a natural language. Assume that the meaning of p is precisiated by representing it as a generalized constraint, $GC(p)$. $GC(p)$ may be viewed as a definition of the granular value, $^*u$. For example, granular values of probability may be approximately 0.1, ... approximately 0.9, approximately 1. A granular variable is a variable which takes granular values. For example, young, middle-aged and old are granular values of the granular variable Age. The probability distribution in Example 2 is an instance of a granular probability distribution. In effect, computation with imprecise probability distributions may be viewed as an instance of computation with granular probability distributions.

In the CW-based approach to computation with imprecise probabilities, computation with imprecise probabilities reduces to computation with generalized constraints. What is used for this purpose is the machinery of GCL. More specifically, computation is carried out through the use of rules which govern propagation and counterpropagation of generalized constraints. The principal rule is the extension principle [4, 5]. In its general form, the extension principle is a computational schema which relates to the following question. Assume that Y is a given function of X, $Y = g(X)$. Let $^*g$ and $^*X$ be granular values of g and X, respectively. Compute $^*g(^*X)$.

In most computations involving imprecise probabilities what is needed is a special case of the extension principle which applies to possibilistic constraints. More specifically, assume that f is a given function and $f(X)$ is constrained by a possibility distribution, A. Assume that g is a given function, $g(X)$. The problem is to compute the possibility distribution of $g(X)$ given the possibility distribution of $f(X)$. The extension principle reduces the solution of this problem to solution of a variational problem [6].

In summary, the CW-based approach to computation with imprecise probabilities opens the door to computation with probabilities, events, relations and constraints which are described in a natural language. Progression from computation with precise probabilities, precise events, precise relations and precise constraints to computation with imprecise probabilities, imprecise events, imprecise relations and imprecise constraints is an important step forward—a step which has the potential for a significant enhancement of the role of natural languages in human-centric fields such as economics, decision analysis and operations research, law and medicine.

References

[1] G. de Cooman. A behavioural model for vague probability assessments *Fuzzy Sets and Systems*, 154(3):305-358, 2005

[2] L.A. Zadeh. From Computing with Numbers to Computing with Words—From Manipulation of Measurements to Manipulation of Perceptions *IEEE Transactions on Circuits and Systems —I: Fundamental Theory and Applications*, 45(1), 105–119, 1999

[3] L.A. Zadeh. Toward a perception-based theory of probabilistic reasoning with imprecise probabilities. *Journal of Statistical Planning and Inference* 105 (2002) 233–264.

[4] L.A. Zadeh. Fuzzy Sets. *Information and Control* 8, 338–353, 1965

[5] L.A. Zadeh. The Concept of a Linguistic Variable and its Applications to Approximate Reasoning—I. *Information Sciences* 8, 199–249, 1975

[6] L.A. Zadeh. Generalized theory of uncertainty (GTU)—principal concepts and ideas *Computational Statistics & Data Analysis* 51(1):15–46, 2006