Fuzzy Logic as a Basis for Theory of Precisiation of Meaning (TPM)

Lotfi A. Zadeh
Professor in the Graduate School, Computer Science Division
Department of Electrical Engineering and Computer Sciences
University of California
Director, Berkeley Initiative in Soft Computing (BISC)
zadeh@eecs.berkeley.edu

Abstract: The concept of precision is ubiquitous. It has a position of centrality not only in science but, more generally, in almost all domains of human reasoning and discourse. And yet, there are some fundamental issues which relate to the concept of precision that have received very little, if any, attention. One such issue is that of precisiation of natural languages. Another basic issue relates to definition of concepts. The Theory of Precisiation of Meaning (TPM) may be viewed as an attempt to construct a conceptual framework for dealing with these and related issues.

Much of human knowledge is expressed in a natural language. Close relationship between human knowledge and natural language is one of the principal reasons why mechanization of natural language understanding has long been one of the important objectives of AI.

Over the years, impressive progress has been made toward achievement of this objective. But what is widely unrecognized is that there is a fundamental limitation to what can be achieved through the use of commonly-employed methods of meaning representation. The aim of this paper is, first, to highlight this limitation and, second, to suggest ways of removing it.

To understand the nature of the limitation, two facts have to be considered. First, a natural language, NL, is basically a system for describing perceptions; and second, perceptions are intrinsically imprecise, reflecting the bounded ability of human sensory organs, and ultimately the brain, to resolve detail and store information. More specifically, perceptions are f-granular in the sense that (a) the boundaries of perceived classes are unsharp (fuzzy); and (b) the values of perceived attributes are granular, with a granule being a clump of values drawn together by indistinguishability, similarity, proximity or functionality.

Imprecision of perceptions is passed on to natural languages. What this implies is that imprecision of semantics of natural languages is rooted in imprecision of perceptions. Semantic imprecision of natural languages is not a problem for humans, but it is a major problem for machines.

To clarify the issue, let *p* be a proposition, concept, question or command. For *p* to be understood by a machine, it must be precisiated, that is, expressed in a mathematically well-defined language. A precisiated form of *p*, Pre(*p*), will be referred to as a precisiand of *p* and will be denoted as *p**.

To precisiate *p* we can employ a number of meaning-representation languages, e.g., Prolog, predicate logic, semantic networks, conceptual graphs, LISP, SQL, etc. The commonly-used meaning-representation languages are bivalent, i.e., are based on bivalent logic. Are we moving in the right direction when we employ such languages for mechanization of natural language understanding? The answer is: No. The reason relates to an important issue which we have not addressed: cointension of *p**, with intension used in its logical sense as attribute-based meaning. More specifically, cointension is a measure of the goodness of fit of the intension of a precisiand, *p**, to the intended intension of *p*. Thus, cointension is a desideratum of precisiation. What this implies is that mechanization of natural language understanding requires more than precisiation—it requires cointensive precisiation. Note that definition is a form of precisiation. In plain words, a definition is cointensive if its meaning is a good fit to the intended meaning of the definiendum.

Here is where the fundamental limitation which was alluded to earlier comes into view. In a natural language, NL, most *p*'s are fuzzy, that is, in one way or another, are a matter of degree. Simple examples: propositions "most Swedes are tall;" and "overeating causes obesity;" concepts "mountain;" and "honest;" question "is Albert honest?" and command "take a few steps."

Employment of commonly-used meaning-representation languages to precisiate a fuzzy *p* leads to a bivalent (crisp) precisiand *p**. The problem is that, in general, a bivalent *p** is not cointensive. As a simple illustration, consider the concept of recession. The standard definition of recession is: A period of general economic decline; specifically, a decline in GDP for two or more consecutive quarters. Similarly, a definition of bear market is: We classify a bear market as a 30 percent decline after 50 days, or a 13 percent decline after 145 days. (Robert Shuster, Ned Davis Research.) Clearly, neither definition is cointensive. Another example is the classical definition of stability. Consider a ball of diameter *D* which is placed on an open bottle whose mouth is of diameter *d*. If *D* is somewhat larger than *d*, the configuration is stable: Obviously, as *D* increases, the configuration becomes less and less stable. But, according to Lyapounov's bivalent definition of stability, the configuration is stable for all values of *D*

greater than *d*. This contradiction is characteristic of crisp definitions of fuzzy concepts—a well-known example of which is the Greek sorites (heap) paradox. The magnitude of the problem becomes apparent where we consider that many concepts in scientific theories are fuzzy, but are defined and treated as if they are crisp. This is particularly true in fields in which the concepts which are defined are descriptions of perceptions.

To remove the fundamental limitation, bivalence must be abandoned. Furthermore, new concepts, ideas and tools must be developed and deployed to deal with the issues of cointensive precisiation, definability and deduction. The principal tools are Precisiated Natural Language (PNL); Protoform Theory (PFT); and the Generalized Theory of Uncertainty (GTU). These tools form the core of what may be called the Computational Theory of Precisiation of Meaning (TPM). The centerpiece of TPM is the concept of a generalized constraint.

A generalized constraint, GC, is an expression of the form *X* isr *R*, where *X* is the constrained variable, *R* is a constraining relation and *r* is an indexing variable whose value defines the modality of the constraint, that is, its semantics. The principal constraints are possibilistic (*r*=blank); veristic (*r*=*v*); probabilistic (*r*=*p*); usuality (*r*=*u*); random set (*r*=*rs*); fuzzy graph (*r*=*fg*); bimodal (*r*=*bm*); and group (*r*=*g*). The primary constraints are probabilistic, possibilistic and veristic. The standard constraints, which are the constraints that underlie standard logic and probability theory, are bivalent possibilistic, bivalent veristic and probabilistic. The semantics of a generalized constraint, GC, is defined by its test-score function, *ts*(*u*), in which *u* is an object to which the constraint applies and *ts*(*u*) is the degree to which *u* satisfies GC. The set of all generalized constraints together with the rules governing combination, qualification and constraint propagation constitutes the Generalized Constraint Language (GCL). Similarly, the set of all standard constraints together with the rules governing combination, probability qualification and constraint propagation constitutes the Standard Constraint Language (SCL).

The concept of a generalized constraint plays a key role in TPM by providing a basis for precisiation of meaning. More specifically, if *p* is a proposition or a concept, its precisiand, Pre(*p*), is represented as a generalized constraint, GC. Thus, Pre(*p*)=GC. In this sense, the concept of a generalized constraint may be viewed on a bridge from natural languages to mathematics.

Representing precisiands of *p* as elements of GCL is the pivotal idea in TPM. Each precisiand is associated with the degree to which it is cointensive with *p*. Given *p*, the problem is that of finding those precisiands which are cointensive, that is, have a high degree of cointension. If *p* is a fuzzy proposition or concept then in general there are no cointensive precisiands in SCL.

In TPM, a refinement of the concept of precisiation is needed. First, a differentiation is made between v-precision (precision in value) and *m*-precision (precision in meaning). For example, proposition *p*: *X* is 5, is both *v*-precise and *m*-precise; *p*: *X* is between 5 and 7, is *v*-imprecise and *m*-precise; and *p*: *X* is small, is both *v*-imprecise and *m*-imprecise; however, *p* can be *m*-precisiated by defining small as a fuzzy set or a probability distribution. A perception is *v*-imprecise and its description is *m*-imprecise. PNL makes it possible to *m*-precisiate descriptions of perceptions.

Granulation of a variable, e.g., representing the values of age as young, middle-aged and old, may be viewed as a form of *v*-imprecisiation. Granulation plays an important role in human cognition by serving as a means of (a) exploiting tolerance for imprecision through omission of irrelevant information. (b) lowering precision and thereby lowering cost; and (c) facilitating understanding and articulation. In fuzzy logic, granulation is *m*-precisiated through the use of the concept of a linguistic variable.

Further refinement of the concept of precisiation relates to two modalities of *m*-precisiation: (a) human-oriented, denoted as *mh*-precisiation; and (b) machine-oriented, denoted as *mm*-precisiation. Unless stated to the contrary, in TPM, precisiation should be understood as *mm*-precisiation.

In a bimodal dictionary or lexicon, the first entry, *p*, is a concept or proposition; the second entry, *p**, is *mh*-precisiand of *p*; and the third entry is *mm*-precisiand of *p*. To illustrate, the entries for recession might read: *mh*-precisiand—a period of general economic decline; and *mm*-precisiand—a decline in GDP for two or more consecutive quarters.

Viewed in a broader perspective, what should be noted is that precisiation of meaning is not the ultimate goal—it is an intermediate goal. Once precisiation of meaning is achieved, the next goal is that of deduction of decision-relevant information. The ultimate goal is decision.

In TPM, a concept which plays a key role in deduction is that of a protoform—an abbreviation for prototypical form. Briefly, a protoform of *p*, PF(*p*), is an abstracted summary of *p*. For example, for *p*: Monika is young, PF(*p*) is *A*(*B*) is *C*, where *A* is abstraction of age, *B* is abstraction of Monika and *C* is abstraction of young. Basically, a protoform of *p* serves to place in evidence the deep semantic structure of *p*.

In TPM, a deduction rule has two components: symbolic and computational. For the most part, a deduction rule is a rule which governs generalized constraint propagation from premises to conclusion.

Protoformal deduction in TPM may be viewed as a computationally-extended version of symbolic reasoning.

There is a simple analogy which helps to understand the meaning of cointensive precisiation. Specifically, a proposition, *p*, is analogous to a system, *S*; precisiation is analogous to modelization; a precisiand, expressed as a generalized constraint, GC(*p*), is analogous to a model, *M*(*S*), of *S*; test-score function is analogous to input-output relation; cointensive precisiand is analogous to well-fitting model; GCL is analogous to the class of all fuzzy-logic-based systems; and SCL is analogous to the subclass of all bivalent-logic-based systems. To say that, in general, a cointensive definition of a fuzzy concept cannot be formulated within the conceptual structure of bivalent logic and probability theory, is similar to saying that, in general, a linear system cannot be a well-fitting model of a nonlinear system.

Ramifications of the concept of cointensive precisiation extend well beyond mechanization of natural language understanding. A broader basic issue is validity of definitions in scientific theories, especially in the realms of human-oriented fields such as law, economics, medicine, psychology and linguistics. More specifically, the concept of cointensive precisiation calls into question the validity of many of the existing definitions of basic concepts—among them the concepts of causality, relevance, independence, stability, complexity and optimality.

About the speaker: LOTFI A. ZADEH is a Professor in the Graduate School, Computer Science Division, Department of EECS, University of California, Berkeley. In addition, he is serving as the Director of BISC (Berkeley Initiative in Soft Computing).

Lotfi Zadeh is an alumnus of the University of Tehran, MIT and Columbia University. He held visiting appointments at the Institute for Advanced Study, Princeton, NJ; MIT; IBM Research Laboratory, San Jose, CA; SRI International, Menlo Park, CA; and the Center for the Study of Language and Information, Stanford University. His earlier work was concerned in the main with systems analysis, decision analysis and information systems. His current research is focused on fuzzy logic, computing with words and soft computing, which is a coalition of fuzzy logic, neurocomputing, evolutionary computing, probabilistic computing and parts of machine learning.

Lotfi Zadeh is a Fellow of the IEEE, AAAS, ACM, AAAI, and IFSA. He is a member of the National Academy of Engineering and a Foreign Member of the Russian Academy of Natural Sciences and the Finnish Academy of Sciences. He is a recipient of the IEEE Education Medal, the IEEE Richard W. Hamming Medal, the IEEE Medal of Honor, the ASME Rufus Oldenburger Medal, the B. Bolzano Medal of the Czech Academy of Sciences, the Kampe de Feriet Medal, the AACC Richard E. Bellman Control Heritage Award, the Grigore Moisil Prize, the Honda Prize, the Okawa Prize, the AIM Information Science Award, the IEEE-SMC J. P. Wohl Career Achievement Award, the SOFT Scientific Contribution Memorial Award of the Japan Society for Fuzzy Theory, the IEEE Millennium Medal, the ACM 2001 Allen Newell Award, the Norbert Wiener Award of the Systems, Man and Cybernetics Society, Civitate Honoris Causa by Budapest Tech (BT) Polytechnical Institution, Budapest, Hungary, the V. Kaufmann Prize, International Association for Fuzzy-Set Management and Economy (SIGEF), other awards and twenty-three honorary doctorates. He has published extensively on a wide variety of subjects relating to the conception, design and analysis of information/intelligent systems, and is serving on the editorial boards of over fifty journals.