Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2006) 15 – 46
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Generalized theory of uncertainty (GTU)—principal concepts and ideas

Lotfi A. Zadeh*

Department of EECS, University of California, Berkeley, CA 94720-1776, USA

Available online 15 May 2006

To Dan Ralescu, Anca Ralescu and the memory of Professor Grigore Moisil

Abstract

Uncertainty is an attribute of information. The path-breaking work of Shannon has led to a universal acceptance of the thesis that information is statistical in nature. Concomitantly, existing theories of uncertainty are based on probability theory. The generalized theory of uncertainty (GTU) departs from existing theories in essential ways. First, the thesis that information is statistical in nature is replaced by a much more general thesis that information is a generalized constraint, with statistical uncertainty being a special, albeit important case. Equating information to a generalized constraint is the fundamental thesis of GTU. Second, bivalence is abandoned throughout GTU, and the foundation of GTU is shifted from bivalent logic to fuzzy logic. As a consequence, in GTU everything is or is allowed to be a matter of degree or, equivalently, fuzzy. Concomitantly, all variables are, or are allowed to be granular, with a granule being a clump of values drawn together by a generalized constraint. And third, one of the principal objectives of GTU is achievement of NL-capability, that is, the capability to operate on information described in natural language. NL-capability has high importance because much of human knowledge, including knowledge about probabilities, is described in natural language. NL-capability is the focus of attention in the present paper. The centerpiece of GTU is the concept of a generalized constraint. The concept of a generalized constraint is motivated by the fact that most real-world constraints are elastic rather than rigid, and have a complex structure even when simple in appearance. The paper concludes with examples of computation with uncertain information described in natural language.

Keywords: Fuzzy logic; Generalized constraint-based computation; Natural language computation; Precisiation; Statistics

1. Introduction

There is a deep-seated tradition in science of dealing with uncertainty—whatever its form and nature—through the use of probability theory. Successes of this tradition are undeniable. But as we move further into the age of machine intelligence and automated decision-making, a basic limitation of probability theory becomes a serious problem. More specifically, in large measure, standard probability theory, call it PT, cannot deal with information described in natural language; that is, to put it simply, PT does not have NL-capability. Here are a few relatively simple examples:

Trip planning: I am planning to drive from Berkeley to Santa Barbara, with stopover for lunch in Monterey. Usually, it takes about two hours to get to Monterey. Usually, it takes about one hour to have lunch. It is likely that it will take

* Tel.: +1 510 642 4959; fax: +1 510 642 1712.
E-mail address: zadeh@eecs.berkeley.edu.

doi:10.1016/j.csda.2006.04.029

about five hours to get from Monterey to Santa Barbara. At what time should I leave Berkeley to get to Santa Barbara, with high probability, before about 6 p.m.?

Balls-in-box: A box contains about 20 balls of various sizes. Most are large. What is the number of small balls? What is the probability that a ball drawn at random is neither small nor large?

Temperature: Usually, the temperature is not very low and not very high. What is the average temperature?

Tall Swedes: Most Swedes are tall. How many are short? What is the average height of Swedes?

Flight delay: Usually, most United Airlines flights from San Francisco leave on time. What is the probability that my flight will be delayed?

Maximization: f is a function from reals to reals described as: If X is small then Y is small; if X is medium then Y is large; if X is large then Y is small. What is the maximum of f?

Expected value: X is a real-valued random variable. Usually, X is much larger than approximately a and much smaller than approximately b, where a and b are real numbers, with $a < b$. What is the expected value of X?

Vera's age: Vera has a son who is in mid-twenties, and a daughter, who is in mid-thirties. What is Vera's age? This example differs from other examples in that to answer the question what is needed is information drawn from world knowledge. More specifically: (a) child-bearing age ranges from about 16 to about 42; and (b) age of mother is the sum of the age of child and the age of mother when the child was born.

In recent years, important contributions have been made to enhancement of capabilities of PT (Bouchon-Meunier et al., 2000; Colubi et al., 2001; Dubois and Prade, 1992, 1994; Nguyen, 1993; Nguyen et al., 2003; Puri et al., 1993; Smets, 1996; Singpurwalla and Booker, 2004; Yager, 2002). Particularly worthy of note are random-set-based theories (Orlov, 1980; Wang and Sanchez, 1982; Goodman and Nguyen, 1985), among them the Dempster–Shafer theory (Dempster, 1967; Shafer, 1976); Klir's generalized information theory (Klir, 2004, 2006); and theories of imprecise probabilities (Walley, 1991; de Cooman, 2005). The generalized theory of uncertainty (GTU) differs from other theories in three important respects. First, the thesis that information is statistical in nature is replaced by a much more general thesis that information is a generalized constraint (Zadeh, 1986), with statistical uncertainty being a special, albeit important case. Equating information to a generalized constraint is the fundamental thesis of GTU. In symbolic form, the thesis may be expressed as

$$I(X) = \mathrm{GC}(X)$$

where X is a variable taking values in U; $I(X)$ is information about X; and $\mathrm{GC}(X)$ is a generalized constraint on X.

Second, bivalence is abandoned throughout GTU, and the foundation of GTU is shifted from bivalent logic to fuzzy logic (Zadeh, 1975a,b; Novak et al., 1999). As a consequence, in GTU everything is or is allowed to be a matter of degree or, equivalently, fuzzy. Concomitantly, all variables are, or are allowed to be granular, with a granule being a clump of values defined by a generalized constraint (Zadeh, 1979a,b, 1997, 1999).

And third, one of the principal objectives of GTU is achievement of NL-capability. Why is NL- capability an important capability? Principally because much of human knowledge and real-world information is expressed in natural language. Basically, a natural language is a system for describing perceptions. Perceptions are intrinsically imprecise, reflecting the bounded ability of human sensory organs, and ultimately the brain, to resolve detail and store information. Imprecision of perception is passed on to natural languages. It is this imprecision that severely limits the ability of PT to deal with information described in natural language. NL-capability of GTU is the focus of attention in the present paper.

A concomitant of GTU's NL-capability is its ability to deal with perception-based information (Fig. 1). Much of information about subjective probabilities is perception-based. In an earlier paper, a generalization of PT, which leads to a perception-based theory, PTp, of probabilistic reasoning with imprecise probabilities, is described (Zadeh, 2002), PTp is subsumed by GTU.

What follows is a precis of GTU. An exposition of an earlier but more detailed version of GTU may be found in Bouchon-Meunier et al. (2000). This paper forms the basis for the present paper.

The centerpiece of GTU is the concept of a generalized constraint—a concept drawn from fuzzy logic. The principal distinguishing features of fuzzy logic are (a) graduation and (b) granulation. More specifically, in fuzzy logic everything is, or is allowed to be, graduated, that is, be a matter of degree or, more or less equivalently, fuzzy. Furthermore, in fuzzy logic all variables are allowed to be granulated, with a granule being a clump of values drawn together by indistinguishability, similarity, proximity or functionality (Fig. 2). Graduation and granulation underline the concept of a linguistic variable (Zadeh, 1973)—a concept which plays a key role in almost all applications of fuzzy logic

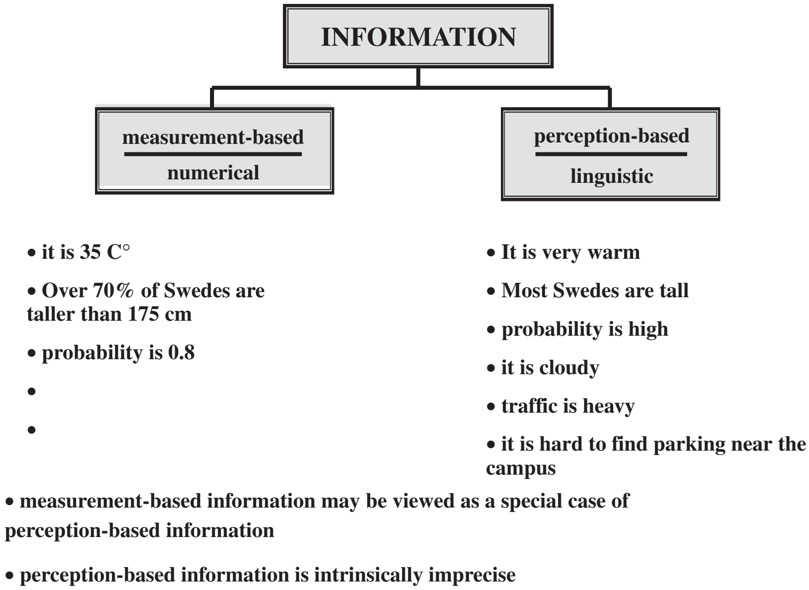


Fig. 1. Measurement-based vs. perception-based information.

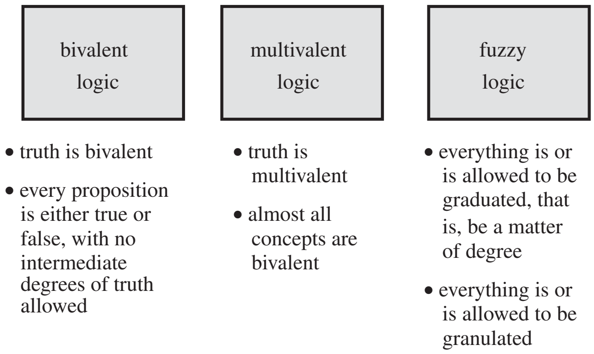


Fig. 2. Logical systems.

(Yen et al., 1995). More fundamentally, graduation and granulation have a position of centrality in human cognition. This is one of the basic reasons why fuzzy logic may be viewed in a model of human reasoning.

NL-Computation is the core of PNL (precisiated natural language) (Zadeh, 2004a,b). Basically, PNL is a fuzzy-logic-based system for computation and deduction with information described in natural language. A forerunner of PNL is PRUF (Zadeh, 1984). We begin with a brief exposition of the basics of NL-Computation in the context of GTU.

2. The concept of NL-Computation

NL-Computation has a position of centrality in GTU. The basic structure of NL-Computation, viewed as the core of PNL, is shown in Fig. 3. The point of departure is a given proposition or, more generally, a system of propositions, p, which constitutes the initial information set described in a natural language (INL). In addition, what is given is a query, q, expressed in a natural language (QNL). The problem is to compute an answer to q given p, $\text{ans}(q|p)$. In GTU, deduction of $\text{ans}(q|p)$ involves these modules: (a) precisiation module, P; (b) protoform module, Pr; and (c) computation/deduction module, C/D. Informally, precisiation is an operation which precisiates its operand. The operand and the result of precisiation are referred to as precisiand and precisiand, respectively. The precisiation module

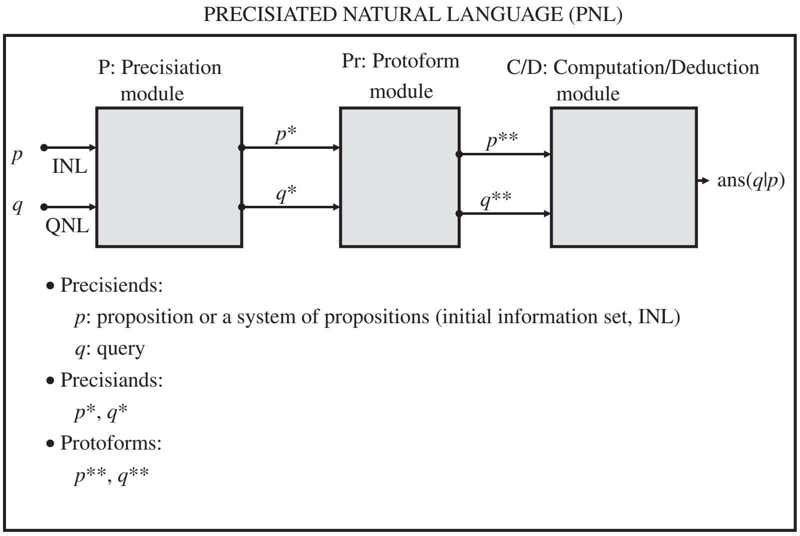


Fig. 3. NL computation—basic structure (PNL).

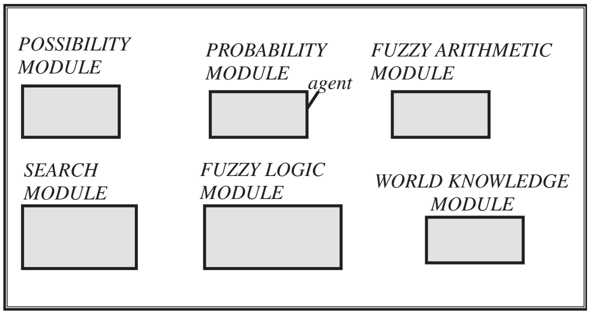


Fig. 4. Computational/deduction module.

operates on the initial information set, p, expressed as INL, and results in a precisiand, p^*. The protoform module serves as an interface between the precisiation module and the computation/deduction module. The input to Pr is a precisiand, p^*, and its output is a protoform of p^*, that is, its abstracted summary, p^{**}. The computation/deduction module is basically a database (catalog) of deduction rules which are, for the most part, rules which govern generalized constraint propagation and counterpropagation. The principal deduction rule is the Extension Principle (Zadeh, 1965, 1975b). The rules are protoformal, with each rule having a symbolic part and a computational part. Protoformal rules are grouped into modules, with each module comprising rules which are associated with a particular class of generalized constraints, e.g., possibilistic constraints, probabilistic constraints, veristic constraints, usuality constraints, etc. (Fig. 4). The inputs to the C/D module are p^{**} and q^{**}. A module which plays an important role in C/D is the world knowledge module (WK). World knowledge is the knowledge which humans acquire through experience, education and communication (Zadeh, 2004a,b). Much of the information in WK is perception-based. Organization of knowledge in WK is a complex issue which is not addressed in the present paper.

3. The concept of precisiation

The concept of precisiation has few precursors in the literature of logic, probability theory and philosophy of languages (Carnap, 1950; Partee, 1976). The reason is that the conceptual structure of bivalent logic—on which the literature is based—is much too limited to allow a full development of the concept of precisiation. In GTU what is used for this purpose is the conceptual structure of fuzzy logic.

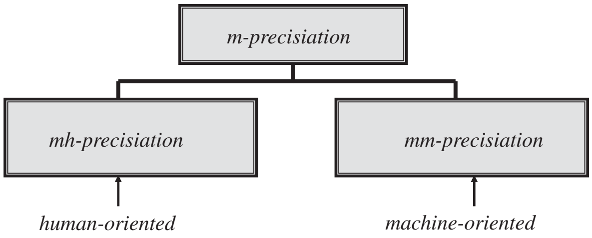


Fig. 5. *mh*- and *mm*-percisiation.

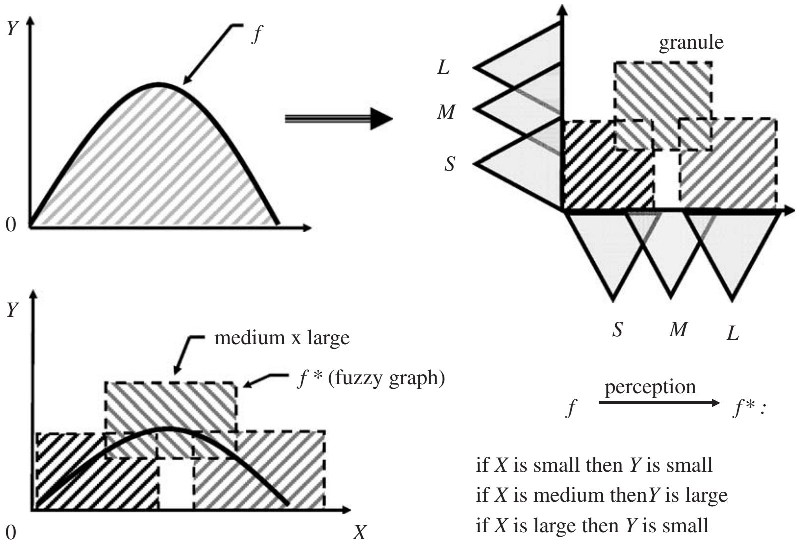


Fig. 6. Granular definition of a function.

Precisiation and precision have many facets. More specifically, it is expedient to consider what may be labeled λ-precisiation, with λ being an indexical variable whose values identify various modalities of precisiation. In particular, it is important to differentiate between precision in value (v-precision) and precision in meaning (m-precision). For example, proposition $X = 5$ is v-precise and m-precise, but proposition $2 \leqslant X \leqslant 6$, is v-imprecise and m-precise. Similarly, proposition "*X* is a normally distributed random variable with mean 5 and variance 2," is v-imprecise and m-precise.

A further differentiation applies to *m*-precisiation. Thus, *mh*-precisiation is human-oriented meaning precisiation, while *mm*-precisiation is machine-oriented or, equivalently, mathematically based meaning precisiation (Fig. 5). A dictionary definition may be viewed as a form of *mh*-precisiation, while a mathematical definition of a concept, e.g., stability, is *mm*-precisiation whose result is *mm*-precisiand of stability.

A more general illustration relates to representation of a function as a collection of fuzzy if–then rules—a mode of representation which is widely used in practical applications of fuzzy logic (Dubois and Parde, 1996; Yen and Langari, 1998). More specifically, let *f* be a function from reals to reals which is represented as (Fig. 6).

f: if *X* is small then *Y* is small,
if *X* is medium than *Y* is large,
if *X* is large than *Y* is small,

where small, medium and large are labels of fuzzy sets. In this representation, the collection in question may be viewed as *mh*-precisiand of *f*. When the collection is interpreted as a fuzzy graph (Zadeh, 1974, 1996) representation of *f*

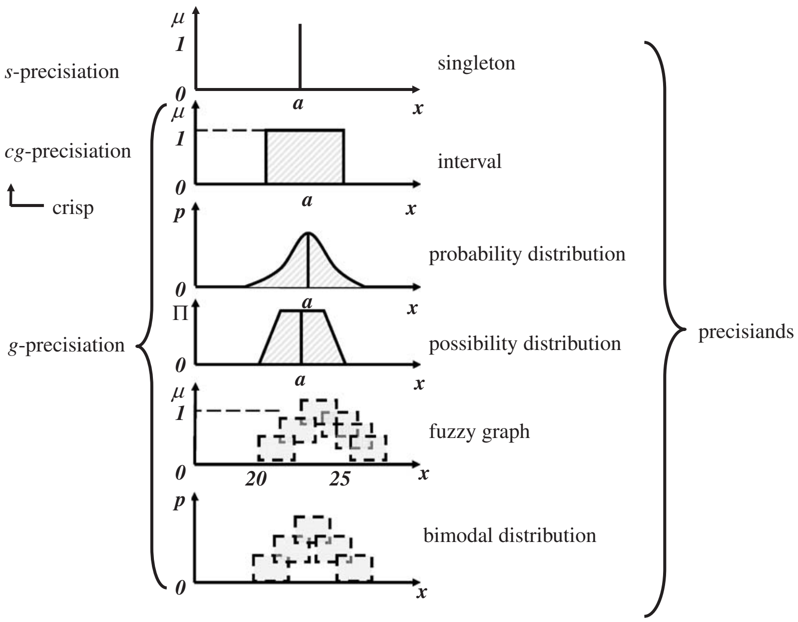


Fig. 7. Granular precisiation of "approximately *a*." $^*a$.

assumes the form.

$$f : \text{small} \times \text{small} + \text{medium} \times \text{large} + \text{large} \times \text{small}$$

which is a disjunction of Cartesian products of small, medium and large. This representation is *mm*-precisiand of *f*.

In general, a precisiend has many precisiands. As an illustration consider the proposition "*X* is approximately *a*," or "*X* is $^*a$" for short, where *a* is a real number. How can "*X* is $^*a$" be precisiated?

The simplest precisiand of "*X* is $^*a$" is "$X = a$," (Fig. 7). This mode of precisiation is referred to as *s*-precisiation, with *s* standing for singular. This is a mode of precisiation which is widely used in science and especially in probability theory. In the latter case, most real-world probabilities are not known exactly but in practice are frequently computed with as if they are exact numbers. For example, if the probability of an event is stated to be 0.7, then it should be understood that 0.7 is actually *0.7, that is, approximately 0.7. The standard practice is to treat *0.7 as 0.7000 . . ., that is, as an exact number.

Next in simplicity is representation of $^*a$ is an interval centering on *a*. This mode of precisiation is referred to *cg*-precisiation, with *cg* standing for crisp-granular. Next is *fg*-precisiation of $^*a$, with the precisiand being a fuzzy interval centering on *a*. Next is *p*-precisiation of $^*a$, with the precisiand being a probability distribution centering on *a*, and so on.

An analogy is helpful in understanding the relationship between a precisiend and its precisiands. More specifically, a *mm*-precisiand, p^*, may be viewed as a model of precisiend, *p*, in the same sense as a differential equation may be viewed as a model of a physical system.

In the context of modeling, an important characteristic of a model is its "goodness of fit." In the context of NL-computation, an analogous concept is that of cointension. The concept is discussed in the following.

4. The concept of cointensive precisiation

Precisiation is a prerequisite to computation with information described in natural language. To be useful, precisiation of a precisiend, *p*, should result in a precisiand, p^*, whose meaning, in some specified sense, should be close to that of *p*. Basically, cointension of p^* and *p* is the degree to which the meaning of p^* fits the meaning of *p*.

In dealing with meaning, it is necessary to differentiate between the intension or, equivalently, the intensional meaning, *i*-meaning, of *p*, and the extension, or, equivalently, the extensional, *e*-meaning of *p*. The concepts of extension and intension are drawn from logic and, more particularly, from modal logic and possible world semantics (Cresswell, 1973; Lambert and Van Fraassen, 1970; Belohlavek and Vychodil, 2006). Basically, *e*-meaning is attribute-free and *i*-meaning is attribute-based. As a simple illustration, if A is a finite set in a universe of discourse, U, then the *e*-meaning of A, that is, its extension is the list of elements of A, $\{u_1, \ldots, u_n\}$, u_i being the name of ith element of A, with no attributes associated with u_i. Let $a\,(u_i)$ be an attribute-vector associated with each u_i. Then the intension of A is a recognition algorithm which, given $a\,(u_i)$, recognizes whether u_i is or is not an element of A. If A is a fuzzy set with membership function μ_A then the *e*-meaning and *i*-meaning of A may be expressed compartly as

$$e\text{-meaning of } A : A = \left\{\mu_A(u_i)/u_i\right\},$$

where $\mu_A(u)/u$ means that $\mu_A(u)$ is the grade of membership of u_i in A; and

$$i\text{-meaning of } A : A = \left\{\mu_A(a(u_i))/u_i\right\},$$

with the understanding that in the *i*-meaning of A the membership function, μ_A is defined on the attribute space. It should be noted that when A is defined through exemplification, it is said to be defined ostensively. Thus, *o*-meaning of A consists of exemplars of A. An ostensive definition may be viewed as a special case of extensional definition. A neural network may be viewed as a system which derives *i*-meaning from *o*-meaning.

Clearly, *i*-meaning is more informative than *e*-meaning. For this reason, cointension is defined in terms of intensions rather than extensions of precisiend and precisiand. Thus, meaning will be understood to be *i*-meaning, unless stated to the contrary. However, when the precisiend is a concept, which plays the role of definiendum and we know its extension but not its intension, cointension has to involve the extension of the definiendum (precisiend) and the intension of the definiens (precisiand).

As an illustration, let p be the concept of bear market. A dictionary definition of p—which may be viewed as a *mh*-precisiand of p—reads “A prolonged period in which investment prices fall, accompanied by widespread pessimism.” A widely accepted quantitative definition of bear market is: We classify a bear market as a 30% decline after 50 days, or a 13 % decline after 145 days. (Shuster) This definition may be viewed as a *mm*-precisiand of bear market. Clearly, the quantitative definition, p^*, is not a good fit to the perception of the meaning of bear market which is the basis for the dictionary definition. In this sense, the quantitative definition of bear market is not cointensive.

Intensions are more informative than extensions in the sense that more can be inferred from propositions whose meaning is expressed intensionally rather than extensionally. The assertion will be precisiated at a later point. For the present, a simple example will suffice.

Consider the proposition p: Most Swedes are tall. Let U be a population of n Swedes, $U = (u_1, \ldots, u_n)$, $u_1 =$ name of ith Swede.

A precisiand of p may be represented as

$$\frac{1}{n}\text{Count(tall.Swedes) is most},$$

where most is a fuzzy quantifier which is defined as a fuzzy subset of the unit interval (Zadeh, 1983a,b). Let $\mu_{\text{tall}}\,(u_i)$, $i = (1, \ldots, n)$ be the grade of membership of u_i in the fuzzy set of tall Swedes. Then the *e*-meaning of tall Swedes may be expressed in symbolic form as

$$\text{tall.Swedes} = \mu_{\text{tall}}\,(u_1)\,/u_1 + \cdots + \mu_{\text{tall}}\,(u_n)\,/u_n.$$

Accordingly, the *i*-precisiand of p may be expressed as

$$\frac{1}{n}\left(\mu_{\text{tall}}\,(u_1) + \cdots + \mu_{\text{tall}}\,(u_n)\right) \text{ is most.}$$

Similarly, the *i*-precisiand of p may be represented as

$$\frac{1}{n}\left(\mu_{\text{tall}}\,(h_1) + \cdots + \mu_{\text{tall}}\,(h_n)\right) \text{ is most,}$$

where h_i is the height of u_i.

As will be seen later, given the *e*-precisiend of p we can compute the answer to the query: How many Swedes are not tall? The answer is 1-most (Fig. 8). However, we cannot compute the answer to the query: How many Swedes are

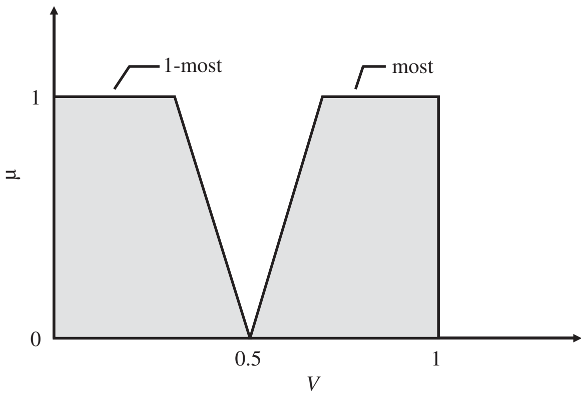


Fig. 8. “most” and antonym of “most”.

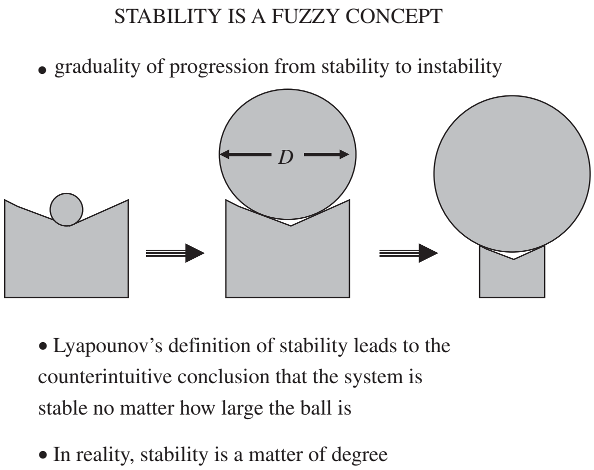


Fig. 9. Stability in a fuzzy concept.

short? The same applies to the query: What is the average height of Swedes? As will be shown later, the answers to these queries can be computed given the *i*-precisiand of *p*.

The concept of cointensive precisiation has important implications for the way in which scientific concepts are defined. The standard practice is to define a concept within the conceptual structure of bivalent logic, leading to a bivalent definition under which the universe of discourse is partitioned into two classes: objects which fit the concept and those which do not, with no shades of gray allowed. Such definition is valid when the concept that is defined, the definiendum, is crisp, that is, bivalent. The problem is that in reality most scientific concepts are fuzzy, that is, are a matter of degree. Familiar examples are the concepts of causality, relevance, stability, independence and bear market. In general, when the definiendum (precisiend) is a fuzzy concept, the definiens (precisiand) is not cointensive, which is the case with the bivalent definition of bear market. More generally, bivalent definitions of fuzzy concepts are vulnerable to the Sorites (heap) paradox (Sainsbury, 1995).

As an illustration, consider a bottle whose mouth is of diameter *d*, with a ball of diameter *D* placed on the bottle (Fig. 9). When *D* is slightly larger than *d*, based on commonsense the system is stable. As *D* increases, the system becomes less and less stable. But Lyapounov's definition of stability leads to the conclusion that the system is stable for all values of *D* so long as *D* is greater than *d*. Clearly, this conclusion is counterintuitive. The problem is that under Lyapounov's bivalent definition of stability, a system is either stable or unstable, with no degrees of stability allowed.

What this example points to is the need for redefinition of many basic concepts in scientific theories. To achieve cointension, bivalence must be abandoned.

5. A key idea—the meaning postulate

In GTU, a proposition, p, is viewed as an answer to a question, q, of the form "What is the value of X?" Thus, p is a carrier of information about X. In this perspective, the meaning of p, $M(p)$, is the information which p carries about X. An important consequence of the fundamental thesis of GTU is what is referred to as the meaning postulate. In symbolic form, the postulate is expressed as $M(p) = \text{GC}(X(p))$, where $\text{GC}(X(p))$ is a generalized constraint on the variable which is constrained by p. In plain words, the meaning postulate assents that the meaning of a proposition may be represented as a generalized constraint. It is this postulate that makes the concept of a generalized constraint the centerpiece of GTU.

A point which should be noted is that the question to which p is an answer is not uniquely determined by p; hence, $X(p)$ is not uniquely defined by p. Generally, however, among the possible questions there is one which is most likely. For example, if p is "Monika is young," then the most likely question is "How old is Monika?" In this example, X is Age(Monika).

6. The concept of a generalized constraint

Constraints are ubiquitous. A typical constraint is an expression of the form $X \in C$, where X is the constrained variable and C is the set of values which X is allowed to take. A typical constraint is hard (inelastic) in the sense that if u is a value of X then u satisfies the constraint if and only if $u \in C$.

The problem with hard constraints is that most real-world constraints are not hard, meaning that most real-world constraints have some degree of elasticity. For example, the constraints "check-out time is 1 p.m.," and "speed limit is 100 km/hr," are, in reality, not hard. How can such constraints be defined? The concept of a generalized constraint is motivated by questions of this kind.

Real-world constraints may assume a variety of forms. They may be simple in appearance and yet have a complex structure. Reflecting this reality, a generalized constraint, $\text{GC}(X)$, is defined as an expression of the form.

$$\text{GC}(X) : X \text{ isr } R,$$

where X is the constrained variable; R is a constraining relation which, in general, is non-bivalent; and r is an indexing variable which identifies the modality of the constraint, that is, its semantics.

The constrained variable, X, may assume a variety of forms. In particular,

- X is an n-ary variable, $X = (X_1, \ldots, X_n)$,
- X is a proposition, e.g., $X =$ Leslie is tall,
- X is a function,
- X is a function of another variable, $X = f(Y)$,
- X is conditioned on another variable, X/Y,
- X has a structure, e.g., $X =$ Location(Residence(Carol)),
- X is a group variable. In this case, there is a group, $G[A]$; with each member of the group, Name_i, $i = 1, \ldots, n$, associated with an attribute-value, A_i. A_i may be vector valued. Symbolically,

  $$G[A] : \text{Name}_1/A_1 + \cdots + \text{Name}_n/A_n.$$

  Basically, $G[A]$ is a relation.
- X is a generalized constraint, $X = Y \text{ isr } R$.

A generalized constraint is associated with a test-score function, $\text{ts}(u)$, (Zadeh, 1981a,b) which associates with each object, u, to which the constraint is applicable, the degree to which u satisfies the constraint. Usually, $\text{ts}(u)$ is a point in the unit interval. However, if necessary, the test-score may be a vector, an element of a semiring (Rossi and Codognet, 2003), an element of a lattice (Goguen, 1969) or, more generally, an element of a partially ordered set, or a bimodal distribution—a constraint which will be described later. The test-score function defines the semantics of the constraint with which it is associated.

The constraining relation, R, is, or is allowed to be, non-bivalent (fuzzy). The principal modalities of generalized constraints are summarized in the following.

7. Principal modalities of generalized constraints

(a) *Possibilistic* ($r =$ blank)

$$X \text{ is } R$$

with R playing the role of the possibility distribution of X. For example,

$$X \text{ is } [a, b]$$

means that $[a, b]$ is the set of possible values of X. Another example is

$$X \text{ is small.}$$

In this case, the fuzzy set labeled small is the possibility distribution of X (Zadeh, 1978; Dubois and Prade, 1988). If μ_{small} is the membership function of small, then the semantics of "X is small" is defined by

$$\text{Poss}\{X = u\} = \mu_{\text{small}}(u),$$

where u is a generic value of X.

(b) *Probabilistic* ($r = p$)

$$X \text{ isp } R,$$

with R playing the role of the probability distribution of X. For example,

$$X \text{ isp } N(m, \sigma^2)$$

means that X is a normally distributed random variable with mean m and variance σ^2.

If X is a random variable which takes values in a finite set $\{u_1, \ldots, u_n\}$ with respective probabilities $p_1, \ldots, p_n$, then X may be expressed symbolically as

$$X \text{ isp } (p_1 \backslash u_1 + \cdots + p_n \backslash u_n),$$

with the semantics

$$\text{Prob}\,(X = u_i) = p_i, \quad i = 1, \ldots, n.$$

What is important to note is that in GTU a probabilistic constraint is viewed as an instance of a generalized constraint. When X is a generalized constraint, the expression

$$X \text{ isp } R$$

is interpreted as a probability qualification of X, with R being the probability of X, (Zadeh, 1979a, 1981a,b). For example,

$$(X \text{ is small) isp likely},$$

where small is a fuzzy subset of the real line, means that probability of the fuzzy event $\{X \text{ is small}\}$ is likely. More specifically, if X takes values in the interval $[a, b]$ and g is the probability density function of X, then the probability of the fuzzy even "X is small" may be expressed as (Zadeh, 1968)

$$\text{Prob}(X \text{ is small}) = \int_a^b \mu_{\text{small}}(u) g(u)\,\mathrm{d}u.$$

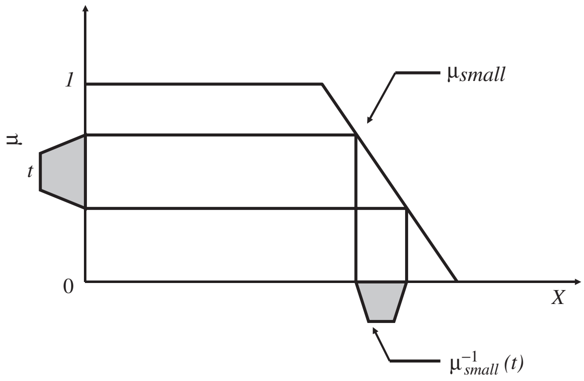


Fig. 10. Truth-qualification: (*X* is small) is *t*.

Hence,

$$\mathrm{ts}(g) = \mu_{\text{likely}}\left(\int_b^a g(u)\mu_{\text{small}}(u)\,\mathrm{d}u\right).$$

This expression for test-score function defines the semantics of probability qualification of a possibilistic constraint.

(c) *Veristic* $(r = v)$

$$X \text{ isv } R,$$

where R plays the role of a verity (truth) distribution of X. In particular, if X takes values in a finite set $\{u_1, \ldots, u_n\}$ with respective verity (truth) values $t_1, \ldots, t_n$, then X may be expressed as

$$X \text{ isv } (t_1 | u_1 + \cdots + t_n | u_n),$$

meaning that Ver $(X = u_i) = t_i$, $i = 1, \ldots, n$.
For example, if Robert is half-German, quarter-French and quarter-Italian, then

$$\text{Ethnicity(Robert) isv } 0.5|\text{German} + 0.25|\text{French} + 0.25|\text{Italian}.$$

When X is a generalized constraint, the expression

$$X \text{ isv } R$$

is interpreted as verity (truth) qualification of X. For example,

$$(X \text{ is small) isv very.true},$$

should be interpreted as "It is very true that X is small." The semantics of truth qualification is defined by (Zadeh, 1979b)

$$\mathrm{Ver}(X \text{ is } R) \longrightarrow X \text{ is } \mu_R^{-1}(t),$$

where μ_R^{-1} is inverse of the membership function of R, and t is a fuzzy truth value which is a subset of [0, 1], Fig. 10.
Note: There are two classes of fuzzy sets: (a) possibilistic, and (b) veristic. In the case of a possibilistic fuzzy set, the grade of membership is the degree of possibility. In the case of a veristic fuzzy set, the grade of membership is the degree of verity (truth). Unless stated to the contrary, a fuzzy set is assumed to be possibilistic.

(d) *Usuality* $(r = u)$

$$X \text{ isu } R.$$

The usuality constraint presupposes that X is a random variable, and that probability of the event $\{X \text{ isu } R\}$ is usually, where usually plays the role of a fuzzy probability which is a fuzzy number (Kaufmann and Gupta, 1985). For example.

$$X \text{ isu small}$$

means that "usually X is small" or, equivalently,

$$\text{Prob}\{X \text{ is small}\} \text{ is usually.}$$

In this expression, small may be interpreted as the usual value of X. The concept of a usual value has the potential of playing a significant role in decision analysis, since it is more informative than the concept of an expected value.

(e) *Random set* $(r = vs)$
In

$$X \text{ isrs } R$$

X is a fuzzy-set-valued random variable and R is a fuzzy random set.

(f) *Fuzzy graph* $(r = fg)$
In

$$X \text{ isfg } R$$

X is a function, f, and R is a fuzzy graph (Zadeh, 1974, 1996) which constrains f (Fig. 11). A fuzzy graph is a disjunction of Cartesian granules expressed as

$$R = A_1 \times B_1 + \cdots + A_n \times B_n,$$

where the A_i and B_i, $i = 1, \ldots, n$, are fuzzy subsets of the real line, and $\times$ is the Cartesian product. A fuzzy graph is frequently described as a collection of fuzzy if–then rules (Zadeh, 1973, 1996; Pedrycz and Gomide, 1998; Bardossy and Duckstein, 1995).

$$R : \text{if } X \text{ is } A_i \text{ then } Y \text{ is } B_i, \quad i = 1, \ldots, n.$$

The concept of a fuzzy-graph constraint plays an important role in applications of fuzzy logic (Bardossy and Duckstein, 1995; Di Nola et al., 1989; Filev and Yager, 1994; Jamshidi et al., 1997; Ross, 2004; Yen et al., 1995).

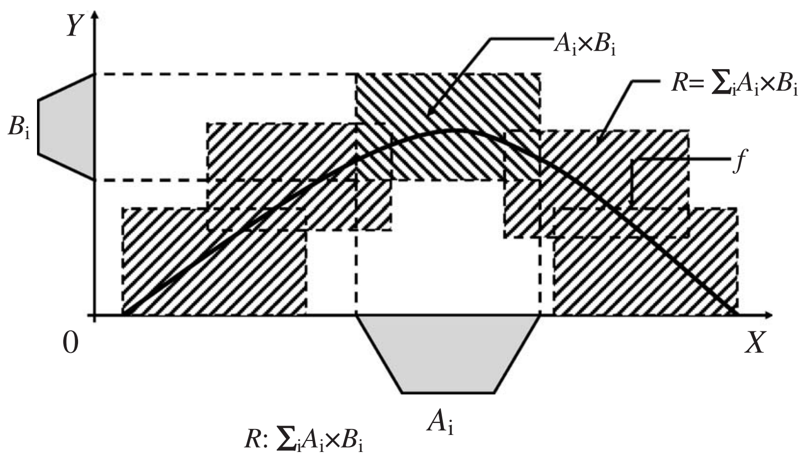


Fig. 11. Fuzzy graph.

8. The concept of bimodal constraint/distribution

In the bimodal constraint,

$$X \text{ isbm } R,$$

R is a bimodal distribution of the form

$$R : \Sigma_i P_i \backslash A_i, \quad i = 1, \ldots, n.$$

with the understanding that Prob $(X \text{ is } A_i)$ is P_i. (Zadeh, 2002), that is, P_i is a granular value of Prob $(X \text{ is } A_i)$, $i = 1, \ldots, n$. (See next section for definition of granular value.)

To clarify the meaning of a bimodal distribution it is expedient to start with an example. I am considering buying Ford stock. I ask my stockbroker, "What is your perception of the near-term prospects for Ford stock?" He tells me, "A moderate decline is very likely; a steep decline is unlikely; and a moderate gain is not likely." My question is: What is the probability of a large gain?

Information provided by my stock broker may be represented as a collection of ordered pairs:

- Price: ((unlikely, steep.decline), (very.likely, moderate.decline), (not.likely, moderate.gain)).

In this collection, the second element of an ordered pair is a fuzzy event or, generally, a possibility distribution, and the first element is a fuzzy probability. The expression for Price is an example of a bimodal distribution.

The importance of the concept of a bimodal distribution derives from the fact that in the context of human-centric systems, most probability distributions are bimodal. Bimodal distributions can assume a variety of forms. The principal types are Type 1, Type 2 and Type 3 (Zadeh, 1979a,b, 1981a). Type 1, 2 and 3 bimodal distributions have a common framework but differ in important detail (Fig. 12). A bimodal distribution may be viewed as an important generalization of standard probability distribution. For this reason, bimodal distributions of Type 1, 2, 3 are discussed in greater detail in the following.

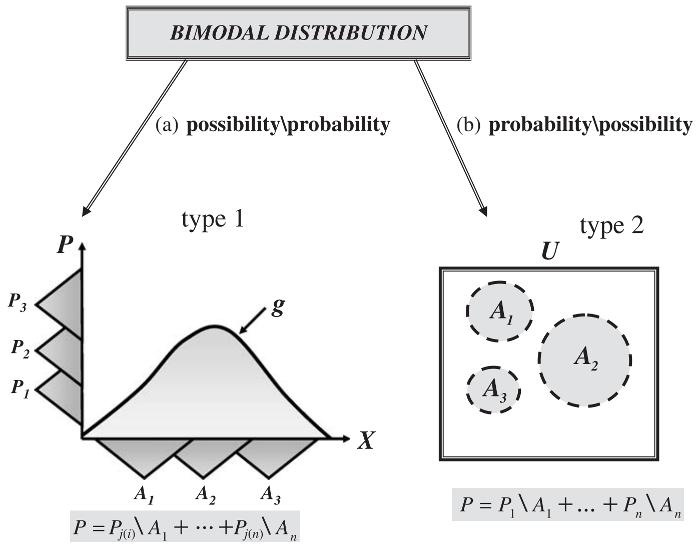


Fig. 12. Type 1 and Type 2 bimodal distributions.

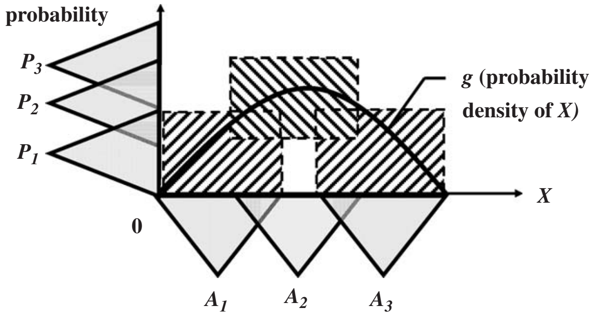


Fig. 13. Basic bimodal distribution.

- Type 1 (default): X is a random variable taking values in U

$A_1, \ldots, A_n, A$ are events (fuzzy sets),
$p_i = \text{Prob}(X \text{ is } A_i), \quad i = 1, \ldots, n,$
$\Sigma_i p_i$ is unconstrained,
$P_i =$ granular value of P_i.

BD: bimodal distribution: $((P_1, A_1), \ldots, (P_n, A_n))$
or, equivalently,

$$X \text{ isbm } (P_1\backslash A_1 + \cdots + P_n\backslash A_n).$$

Problem: What is the granular probability, P, of A? In general, this probability is fuzzy-set-valued.

A special case of bimodal distribution of Type 1 is the basic bimodal distribution (BBD). In BBD, X is a real-valued random variable, and X and P are granular (Fig. 13).

- Type 2 (fuzzy random set): X is a fuzzy-set-valued random variable with values

$A_1, \ldots, A_n$ (fuzzy sets),
$p_i = \text{Prob}(X = A_i), \quad i = 1, \ldots, n,$
$P_i :$ granular value of p_i.

BD: X isrs $(P_1\backslash A_1 + \cdots + P_n\backslash A_n)$,
$\Sigma_i p_i = 1$.
Problem: What is the granular probability, P, of A? P is not definable. What are definable
are (a) the expected value of the conditional possibility of A given BD,
and (b) the expected value of the conditional necessity of A given BD.

- Type 3 (Dempster–Shafer) (Dempster, 1967; Shafer, 1976; Schum, 1994): X is a random variable taking values $X_1, \ldots, X_n$ with probabilities $p_1, \ldots, p_n$

X_i is a random variable taking values in $A_i, \quad i = 1, \ldots, n.$
Probability distribution of X_i in $A_i, \quad i = 1, \ldots, n,$ is not specified,
X isp $(p_1\backslash X_1 + \cdots + p_n\backslash X_n)$.

Problem: What is the probability, p, that X is in A? Because probability distributions of the X_i in the A_i are not specified, p is interval valued. What is important to note is that the concepts of upper and lower probabilities break down when the A_i are fuzzy sets (Zadeh, 1979a).

Note: In applying Dempster–Shafer theory, it is important to check on whether the data fit Type 3 model. In many cases, the correct model is Type 1 rather than Type 3.

The importance of bimodal distributions derives from the fact that in many realistic settings a bimodal distribution is the best approximation to our state of knowledge. An example is assessment of degree of relevance, since relevance is generally not well defined. If I am asked to assess the degree of relevance of a book on knowledge representation to summarization, my state of knowledge about the book may not be sufficient to justify an answer such as 0.7. A better approximation to my state of knowledge may be "likely to be high." Such an answer is an instance of a bimodal distribution.

What is the expected value of a bimodal distribution? This question is considered in the section on protoformal deduction rules.

9. The concept of a group constraint

In

$$X \text{ isg } R,$$

X is a group variable, $G[A]$, and R is a group constraint on $G[A]$. More specifically, if X is a group variable of the form

$$G[A] : \text{Name}_1/A_i + \cdots + \text{Name}_n/A_n$$

or

$$G[A] : \Sigma_i \text{Name}_i/A_i, \ \text{for short}, \quad i = 1, \ldots, n,$$

then R is a constraint on the A_i, written in $G[A \text{ is } R]$. To illustrate, if we have a group of n Swedes, with Name_i being the name of ith Swede, and A_i being the height of Name_i, then the proposition "most Swedes are tall," is a constraint on the A_i which may be expressed as (Zadeh, 1983a, 2004a)

$$\frac{1}{n} \sum \text{Count (tall.Swedes) is most}$$

or, more explicitly,

$$\frac{1}{n} \left(\mu_{\text{tall}} (A_1) + \cdots + \mu_{\text{tall}} (A_n) \right) \text{ is most,}$$

where $A_i = \text{Height} \ (\text{Name}_i)$, $i = 1, \ldots, n$, and most is a fuzzy quantifier which is interpreted as a fuzzy number (Zadeh, 1983a,b).

10. Primary constraints, composite constraints and standard constraints

Among the principal generalized constraints there are three that play the role of primary generalized constraints. They are:

$$\text{Possibilistic constraint} : X \text{ is } R,$$
$$\text{Probabilistic constraint} : X \text{ isp } R$$

and

$$\text{Veristic constraint} : X \text{ isv } R.$$

A special case of primary constraints is what may be called standard constraints: bivalent possibilistic, probabilistic and bivalent veristic. Standard constraints form the basis for the conceptual framework of bivalent logic and probability theory.

A generalized constraint is composite if it can be generated from other generalized constraints through conjunction, and/or projection and/or constraint propagation and/or qualification and/or possibly other operations. For example, a random-set constraint may be viewed as a conjunction of a probabilistic constraint and either a possibilistic or veristic constraint. The Dempster–Shafer theory of evidence is, in effect, a theory of possibilistic random-set constraints. The derivation graph of a composite constraint defines how it can be derived from primary constraints.

The three primary constraints—possibilistic, probabilistic and veristic—are closely related to a concept which has a position of centrality in human cognition—the concept of partiality. In the sense used here, partial means: a matter of degree or, more or less equivalently, fuzzy. In this sense, almost all human concepts are partial (fuzzy). Familiar examples of fuzzy concepts are: knowledge, understanding, friendship, love, beauty, intelligence, belief, causality, relevance, honesty, mountain and, most important, truth, likelihood and possibility. Is a specified concept, C, fuzzy? A simple test is: If C can be hedged, then it is fuzzy. For example, in the case of relevance, we can say: very relevant, quite relevant, slightly relevant, etc. Consequently, relevance is a fuzzy concept.

The three primary constraints may be likened to the three primary colors: red, blue and green. In terms of this analogy, existing theories of uncertainty may be viewed as theories of different mixtures of primary constraints. For example, the Dempster–Shafer theory of evidence is a theory of a mixture of probabilistic and possibilistic constraints. GTU embraces all possible mixtures. In this sense, the conceptual structure of GTU accommodates most, and perhaps all, of the existing theories of uncertainty.

11. The generalized constraint language and standard constraint language

A concept which has a position of centrality in GTU is that of generalized constraint language (GCL). Informally, GCL is the set of all generalized constraints together with the rules governing syntax, semantics and generation. Simple examples of elements of GCL are

$(X \text{ is small})$ is likely
$((X, Y) \text{ isp } A) \wedge (X \text{ is } B)$
$(X \text{ isp } A) \wedge ((X, Y) \text{ isv } B)$
$\text{Proj}_Y((X \text{ is } A) \wedge (X, Y) \text{ isp } B)$

where $\wedge$ is conjunction.

A very simple example of a semantic rule is:

$$(X \text{ is } A) \wedge (Y \text{ is } B) \longrightarrow \text{Poss}(X \text{ is } A) \wedge \text{Poss}(Y \text{ is } B) = \mu_A(u) \wedge \mu_B(v),$$

where u and v are generic values of X, Y; and μ_A and μ_B are the membership functions of A and B, respectively.

In principle, GCL is an infinite set. However, in most applications only a small subset of GCL is likely to be needed.

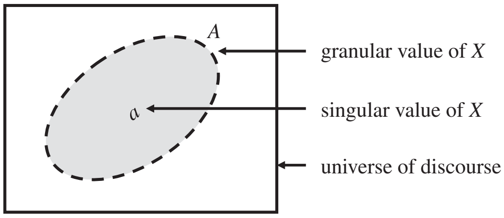
A key idea which underlies NL-Computation is embodied in the meaning postulate—a postulate which asserts that the meaning of a proposition, p, drawn from a natural language is representable as a generalized constraint. In the context of GCL, the meaning postulate asserts that p may be precisiated through translation into GCL. Transparency of translation may be enhanced through annotation. Simple example of annotation,

$$\text{Monika is young} \longrightarrow X/\text{Age (Monika) is } R/\text{young}$$

In GTU, the set of all standard constraints together with the rules governing syntax, semantics and generation constitute the Standard Constraint Language (SCL). SCL is a subset of GCL.

12. The concept of granular value

The concept of a generalized constraint provides a basis for an important concept—the concept of a granular value. Let X be a variable taking values in a universe of discourse U, $U = \{u\}$. If a is an element of U, and it is known that the value of X is a, then a is referred to as a singular value of X. If there is some uncertainty about the value of X, the available information induces a restriction on the possible values of X which may be represented as a generalized constraint $\text{GC}(X)$, X isr R. Thus a generalized constraint defines a granule which is referred to as a granular value of X, $\text{Gr}(X)$ (Fig. 14). For example, if X is known to lie in the interval $[a, h]$, then $[a, h]$ is a granular value of X. Similarly, if X isp $N(m, \sigma^2)$, then $N(m, \sigma^2)$ is a granular value of X. What is important to note is that defining a granular value in terms of a generalized constraint makes a granular value *mm*-precise. It is this characteristic of granular values that



- singular: X is a ⟶ singleton
- granular: X isr A ⟵ granule
- a granule is defined by a generalized constraint

example:

X: unemployment

a: 7.3%

A: high

Fig. 14. A granule defined as a generalized constraint.

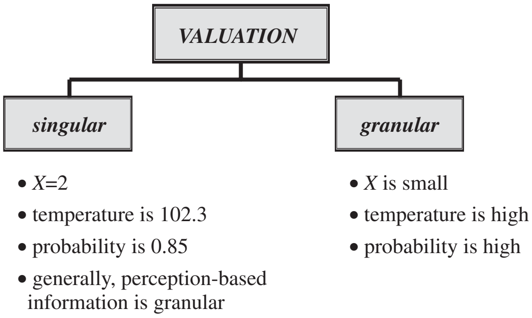


Fig. 15. Singular and granular values.

underlies the concept of a linguistic variable (Zadeh, 1973). Symbolically, representing a granular value as a generalized constraint may be expressed as $\mathrm{Gr}(X) = \mathrm{GC}(X)$. It should be noted that, in general, perception-based information is granular (Fig. 15).

The importance of the concept of a granular value derives from the fact that it plays a central role in computation with information described in natural language. More specifically, when a proposition expressed in a natural language is represented as a system of generalized constraints, it is, in effect, a system of granular values. Thus, computation with information described in natural language ultimately reduces to computation with granular values. Such computation is the province of Granular Computing. (Zadeh, 1979a,b, 1997, 1998; Lin, 1998; Bargiela and Pedrycz, 2002; Lawry, 2001; Lawry et al., 2003; Mares, 1994; Yager, 2006).

13. The concept of protoform

The term "protoform" is an abbreviation of "prototypical form." Informally, a protoform, A, of an object, B, written as $A = \mathrm{PF}(B)$, is an abstracted summary of B. (Fig. 6). Usually, B is a proposition, a system of propositions, question, command, scenario, decision problem, etc. More generally, B may be a relation, system, case, geometrical form or an object of arbitrary complexity. Usually, A is a symbolic expression, but, like B, it may be a complex object. The primary function of $\mathrm{PF}(B)$ is to place in evidence the deep semantic structure of B (Fig. 16).

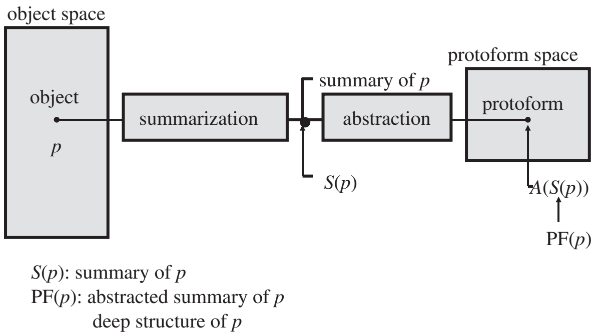
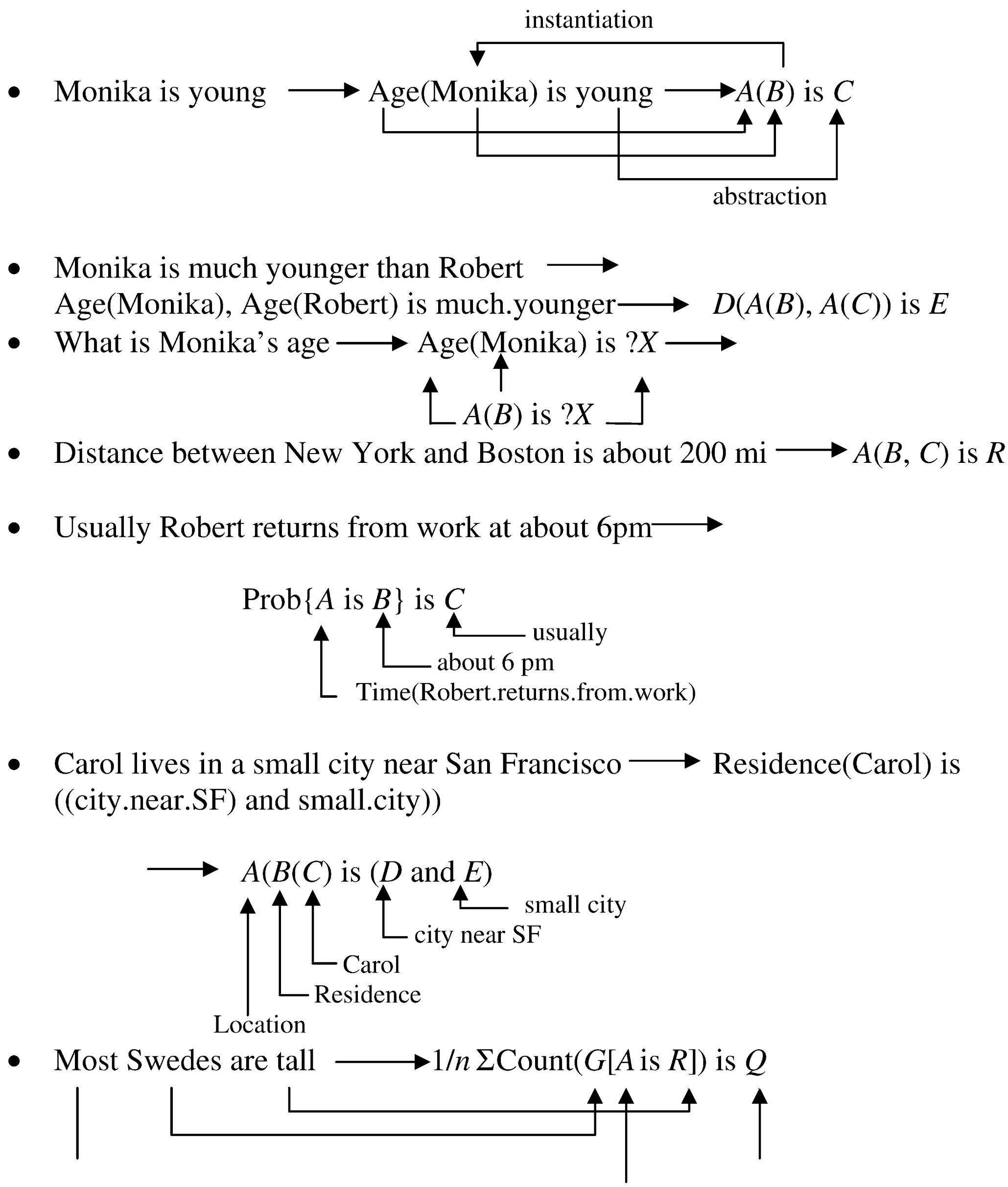
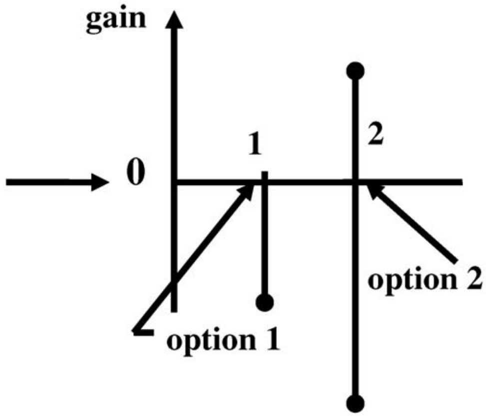
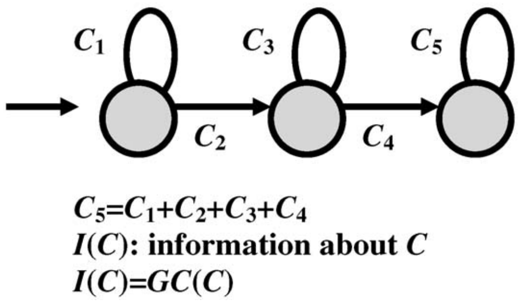


Fig. 16. Definition of protoform of p.



- Alan has severe back pain. He goes to see a doctor. The doctor tells him that there are two options: (1) do nothing; and (2) do surgery. In the case of surgery, there are two possibilities: (a) surgery is successful, in which case Alan will be pain free; and (b) surgery is not successful, in which case Alan will be paralyzed from the neck down. Question: Should Alan elect surgery?





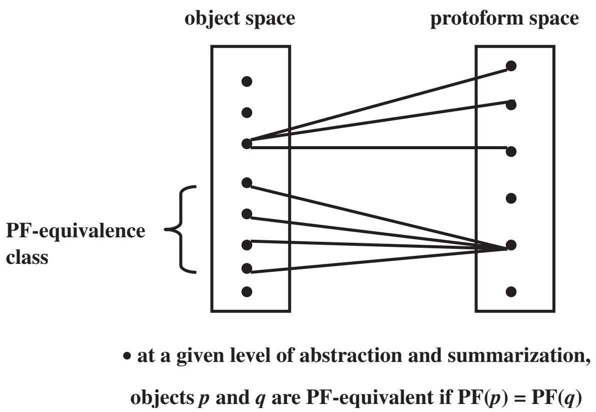


Fig. 17. Protoforms and PF-equivalence.

- I am planning to drive from Berkeley to Santa Barbara, with stopover for lunch in Monterey. Usually, it takes about two hours to get to Monterey. Usually, it takes about one hour to have lunch. It is likely that it will take about five hours to get from Monterey to Santa Barbara. At what time should I leave Berkeley to get to Santa Barbara, with high probability, before about 6 p.m.?

Abstraction has levels, just as summarization does, including no summarization and/or no abstraction. For this reason, an object may have a multiplicity of protoforms (Fig. 17). Conversely, many objects may have the same protoform.

Such objects are said to be protoform-equivalent, or PF-equivalent, for short. For example, p: Most Swedes are tall, and q: Few professors are rich, are PF-equivalent.

A protoform may be extensional (e-protoform) or intensional (i-protoform). For example,

- **Most Swedes are tall** → **Q A's are B's** **(attribute-free; extensional)**
  → **$\frac{1}{n}$ Count(G[H is A]) is Q** **(attribute-based; intensional)**

As in the case of meaning, an e-protoform is less informative than an i-protoform.

The concept of a protoform serves two important functions. First, it provides a basis for organization of knowledge, with PF-equivalent propositions placed in the same class. And second, in NL-Computation the concept of a protoform plays a pivotal role in computation/deduction.

14. The concept of generalized-constraint-based computation

In GTU, computation/deduction is treated as an instance of question-answering. With reference to Fig. 3, the point of departure is a proposition or a system of propositions, p, described in a natural language p is referred to as the initial information set, INL. The query, q, is likewise expressed in a natural language. As was stated earlier, the first step in NL-Computation involves precisiation of p and q, resulting in precisiands p^* and q^*, respectively. The second step involves construction of protoforms of p^* and q^*, p^{**} and q^{**}, respectively. In the third and last step, p^{**} and q^{**} are applied to the computation/deduction module, C/D. An additional internal source of information is world knowledge, wk. The output of C/D is an answer, ans$(q|p)$.

Examples

p: Monika is young	p^*: Age(Monika) is young,
	p^{**}: A(B) is C,
p: Most Swedes are tall	p^*: Count(tall.Swedes/Swedes) is most,
	p^{**}: Count ($G[X$ is $A]/G[X])$ is Q.

The key idea in NL-Computation—the meaning postulate—plays a pivotal role in computation/deduction in GTU. More specifically, p^* may be viewed as a system of generalized constraints which induces a generalized constraint on ans$(q|p)$. In this sense, computation/deduction in GTU may be equated to generalized constraint propagation. More concretely, generalized constraint propagation is governed by what is referred to as the deduction principle. Informally, the basic idea behind this principle is the following.

Deduction principle

Assume that the answer to q can be completed if we know the values of variables $u_i, \ldots, u_n$. Thus,

$$\text{ans}(q|p) = f\,(u_i, \ldots, u_n)\,.$$

Generally, what we know are not the values of the u_i but a system of generalized constraints which represent the precisiand of p, p^*. Express the precisiand, p^*, as a generalized constraint on the u_i.

$$p^*\!: \ \text{GC}(u_i, \ldots, u_n).$$

At this point, what we have is GC$(u_i, \ldots, u_n)$ but what we need is the generalized constraint on ans$(q|p)$, ans$(q|p) = f(u_i, \ldots, u_n)$. To solve this basic problem—a problem which involves constraint propagation—what is needed is the extension principle of fuzzy logic (Zadeh, 1965, 1975b). This principle will be discussed at a later point. At this juncture, a simple example will suffice.

Assume that

p: Most Swedes are tall

and

q: What is the average height of Swedes?

Assume that we have a population of Swedes, $G=(u_i, \ldots, u_n)$, with h_i, $i=1, \ldots, n$, being the height of ith Swede. Precisiends of p and q may be expressed as

$$p^*: \frac{1}{n}\left(\mu_{\text{tall}}(h_1)+\cdots+\mu_{\text{tall}}(h_n)\right) \text{ is most,}$$

$$q^*: \ \text{ans}(q|p)=\frac{1}{n}(h_1+\cdots+h_n).$$

In this instance, what we are dealing with is propagation of the constraint on p^* to a constraint on ans$(q|p)$. Symbolically, the problem may be expressed as

$$\frac{\frac{1}{n}\left(\mu_{\text{tall}}(h_1)+\cdots+\mu_{\text{tall}}(h_n)\right) \text{ is most}}{\frac{1}{n}(h_1+\cdots+h_n)}$$

with the understanding that the premise and the consequent are fuzzy constraints. Let $\mu_{\text{ave}}(v)$ be the membership function of the average height. Application of this extension principle reduces computation of the membership function of ans$(q|p)$ to the solution of the variational problem

$$\mu_{\text{ave}}(v)=\sup_h\left(\mu_{\text{most}}\left(\frac{1}{n}\left(\mu_{\text{tall}}(h_1)+\cdots+\mu_{\text{tall}}(h_n)\right)\right)\right)$$

subject to

$$v=\frac{1}{n}(h_1+\cdots+h_n), \quad h=(h_1, \ldots, h_n).$$

In this simple example, computation of the answer to query requires the use of just one rule of deduction—the extension principle. More generally, computation of the answer to a query involves application of a sequence of deduction rules drawn from the Computation/Deduction module. The Computation/Deduction module comprises a collection of agent-controlled modules and submodules, each of which contains protoformal deduction rules drawn from various fields and various modalities of generalized constraints. (Fig. 4) A protoformal deduction rule has a symbolic part, which is expressed in terms of protoforms; and a computational part which defines the computation that has to be carried out to arrive at a conclusion. The principal protoformal deduction rules are described in the following.

15. Protoformal deduction rules

There are many ways in which generalized constraints may be combined and propagated. The principal protoformal deduction rules are the following:

(a) *Conjunction* (*possibilistic*)

Symbolic	Computational
X is R	$T=R\times S$,
Y is S	
(X, Y) is T	

where × is the Cartesian product.

(b) *Projection* (*possibilistic*)

Symbolic	Computational
$\frac{(X,Y) \text{ is } R}{X \text{ is } S}$,	$\mu_S(u)=\mu_{\text{Proj}_X R}(u)=\max_v \mu_R(u, v)$,

where μ_R and μ_S are the membership functions of R and S, respectively.

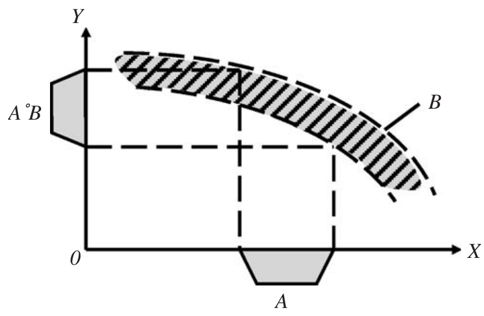


Fig. 18. Compositional rule of inference.

(c) *Projection* (*probabilistic*)

Symbolic Computational

$$\frac{(X, Y) \text{ isp } R}{X \text{ isp } S}, \quad p_S(u) = \int p_R(u, v)\, \mathrm{d}v,$$

where X and Y are real-valued random variables, and R and S are probability densities of (X, Y) and X, respectively.

(a) *Computational rule of inference* (Zadeh, 1965)

Symbolic Computational

$$\frac{\begin{array}{l} X \text{ is } A \\ (X, Y) \text{ is } B \end{array}}{Y \text{ is } C} \quad \mu_C(v) = \max_u \left(\mu_A(u) \wedge \mu_B(u, v)\right)$$

A, B and C are fuzzy sets with respective membership functions μ_A, μ_B, μ_C; $\wedge$ is min or t-norm (Fig. 18).

(b) *Intersection/product syllogism* (Zadeh, 1983a,b)

Symbolic Computational

$$\frac{\begin{array}{l} Q_1 A\text{' s are } B\text{' s} \\ Q_2 (A\&B)\text{' s are } C\text{' s} \end{array}}{Q_3 A\text{' s are } (B\&C)\text{' s}} \quad Q_3 = Q_1 * Q_2$$

Q_1 and Q_2 are fuzzy quantifiers; A, B, C are fuzzy sets; $*$ is product in fuzzy arithmetic (Kaufmann and Gupta, 1985).

(c) *Basic extension principle* (Zadeh, 1965)

Symbolic Computational

$$\frac{X \text{ is } A}{g(X) \text{ is } B} \quad \mu_B(v) = \sup_u \left(\mu_A(u)\right)$$

subject to

$v = g(u)$

g is a given function or functional; A and B are fuzzy sets (Fig. 19).

(d) *Extension principle* (Zadeh, 1975b)

This is the principal deduction rule governing possibilistic constraint propagation (Fig. 20)

Symbolic Computational

$$\frac{f(X) \text{ is } A}{g(X) \text{ is } B} \quad \mu_B(v) = \sup_u \left(\mu_B(f(u))\right)$$

subject to

$v = g(u)$

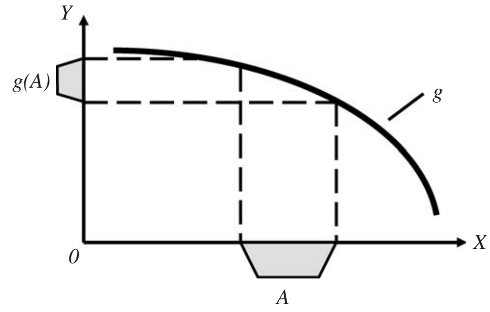


Fig. 19. Basic extension of principle.

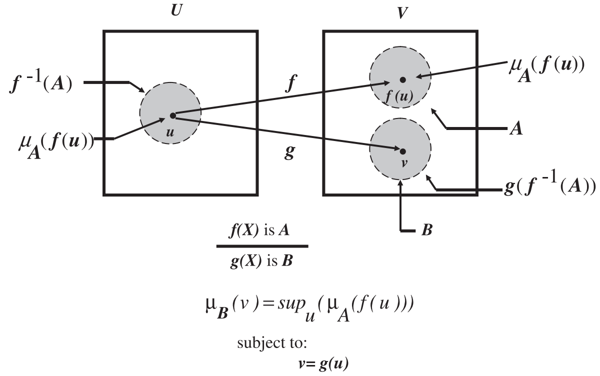


Fig. 20. Extension principle.

The extension principle is an instance of the generalized extension principle

$Y = f(X)$

$\text{Gr}(Y)$ isr $\text{Gr}(X)$

The generalized extension principle may be viewed as an answer to the following question: If f is a function from $U = \{X\}$ to $V = \{Y\}$ and I can compute the singular value of Y given a singular value of X, what is the granular value of Y given a granular value of X?

Note: The extension principle is a primary deduction rule in the sense that many other deduction rules are derivable from the extension principle. An example is the following rule:

(e) *Probability rule*

Symbolic	Computational
$\dfrac{\text{Prob}(X \text{ is } A) \text{ is } B}{\text{Prob}(X \text{ is } C) \text{ is } D}$	$\mu_D(v) = \sup_r \left(\mu_B\left(\int_U \mu_A(u)r(u)\,\mathrm{d}u\right)\right)$ subject to $v = \int_U \mu_C(u)r(u)\,\mathrm{d}u$ $\int_U r(u)\,\mathrm{d}u = 1.$

where X is a real-valued random variable; A, B, C and D are fuzzy sets: r is the probability density of X; and $U = \{u\}$. To derive this rule, we note that

$$\text{Prob}(X \text{ is } A) \text{ is } B \longrightarrow \int_U r(u)\mu_A(u)\,\mathrm{d}u \text{ is } B$$

$$\text{Prob}(X \text{ is } C) \text{ is } D \longrightarrow \int_U r(u)\mu_C(u)\,\mathrm{d}u \text{ is } D$$

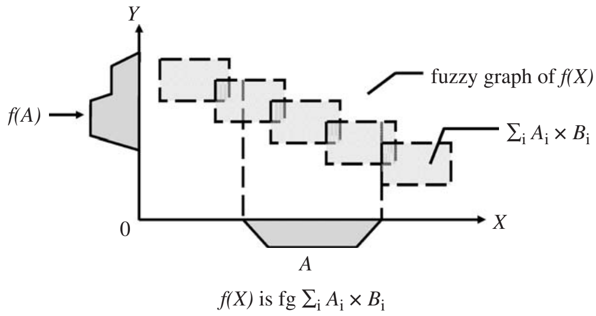


Fig. 21. Fuzzy-graph interpolation.

which are generalized constraints of the form

$$f(r) \text{ is } B$$

$$g(r) \text{ is } D.$$

Applying the extension principle to these expressions, we obtain the expression for D which appears in the basic probability rule.

(f) *Fuzzy-graph interpolation rule*
This rule is the most widely used rule in applications of fuzzy logic (Zadeh, 1975a, 1976). We have a function, $Y = f(X)$, which is represented as a fuzzy graph (Fig. 21). The question is: What is the value of Y when X is A? The A_i, B_i and A are fuzzy sets.
Symbolic part

$$X \text{ is } A,$$

$$Y = f(X),$$

$$\frac{f(X) \text{ isfg } \sum_i A_i \times B_i}{Y \text{ is } C}.$$

Computational part

$$C = \sum_i m_i \wedge B_i,$$

where m_i is the degree to which A matches A_i,

$$m_i = \sup_u \left(\mu_A(u) \wedge \mu_{A_i}(u)\right), \quad i = 1, \ldots, n.$$

When A is a singleton, this rule reduces to

$$X = a,$$

$$Y = f(X),$$

$$f(X) \text{ isfg } \sum_i A_i \times B_i, \quad i = 1, \ldots, n.$$

$$Y = \sum_i \mu_{A_i}(a) \wedge B;$$

In this form, the fuzzy-graph interpolation rule coincides with the Mamdani rule—a rule which is widely used in control and related applications (Mamdani and Assilian, 1975b) (Fig. 22).

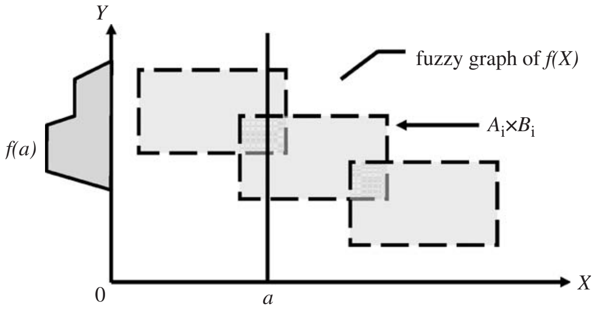


Fig. 22. Mamdani interpolation.

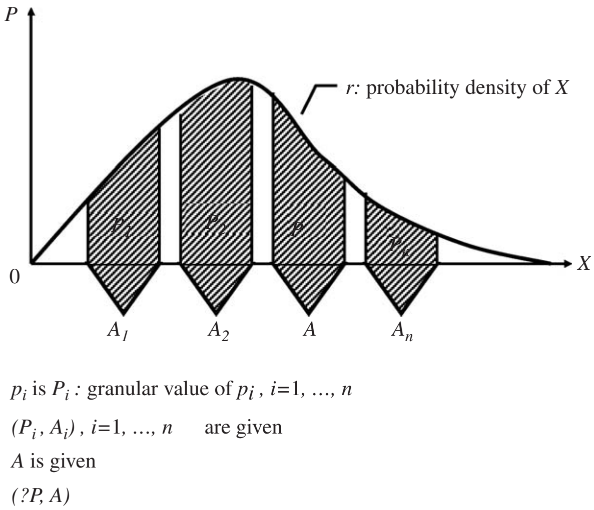


Fig. 23. Interpolation of a bimodal distribution.

In addition to basic rules, the Computation/Deduction module contains a number of specialized modules and submodules. Of particular relevance to GTU are Probability module and Usuality submodule. A basic rule in Probability module is the bimodal distribution interpolation rule which is stated in the following.

(g) *Bimodal distribution interpolation rule*
With reference to Fig. 23, the symbolic and computational parts of this rule are:
Symbolic

$$\frac{\text{Prob}(X \text{ is } A_i) \text{ is } P_i}{\text{Prob}(X \text{ is } A) \text{ is } Q}, \quad i = 1, \ldots, n$$

Computational

$$\mu_Q(v) = \sup_r \left(\mu_{P_1} \left(\int_U \mu_{A_1}(u) r(u)\, \mathrm{d}u \right) \right) \wedge \cdots \wedge \mu_{P_n} \left(\int_U \mu_{A_n}(u) r(u)\, \mathrm{d}u \right)$$

subject to

$$v = \int_U \mu_A(u) r(u)\, \mathrm{d}u$$

$$\int_U r(u)\, \mathrm{d}u = 1$$

In this rule, X is a real-valued random variable; r is the probability density of X; and U is the domain of X.

What is the expected value, $E(X)$, of a bimodal distribution? The answer follows through application of the extension principle:

$$\mu_{E(X)}(v) = \sup_r \left(\mu_{P_1}\left(\int_U \mu_{A_1}(u) r(u)\,\mathrm{d}u \right)\right) \wedge \cdots \wedge \mu_{P_n}\left(\int_U \mu_{A_n}(u) r(u)\,\mathrm{d}u \right)$$

subject to

$$v = \int_U u r(u)\,\mathrm{d}u$$

$$\int_U r(u)\,\mathrm{d}u = 1$$

Note: $E(X)$ is a fuzzy subset of U.

16. Examples of computation/deduction

The following relatively simple examples are intended to illustrate application of deduction rules.

16.1. The Robert example

p: Usually, Robert returns from work at about 6:00 pm. What is the probability that Robert is home at about 6:15 pm.?

First, we find the protoforms of the information set and the query.

Usually, Robert returns from work at about 6:00 pm. ⟶

⟶ Prob(Time(Return(Robert)) is *6:00 pm.) is usually *B*/usually,

which in annotated form reads

⟶ Prob(*X*/Time(Return(Robert)) is $A/{}^*6{:}\,00$ pm.) is *B*/usually, where $^*a$ is an abbreviation of about a.

Likewise, for the query, we have

Prob(Time(Return(Robert)) is $\leqslant \circ {}^*6{:}\,15$ pm.) is ? D

which in annotated form reads

⟶ Prob(*X*/Time(Return(Robert)) is $C/\leqslant \circ\, {}^*6{:}\,15$ pm.) is *D*/usually,

where $\circ$ is the operation of composition (Pedrycz and Gomide, 1998) Searching the computation/deduction module, we find that the basic probability rule matches the protoforms of the data and the query

$$\frac{\text{Prob } (X \text{ is } A) \text{ is } B}{\text{Prob } (X \text{ is } C) \text{ is } D},$$

where

$$\mu_D(v) = \sup_r \left(\mu_B \left(\int_U \mu_A(u) r(u)\,\mathrm{d}u \right)\right)$$

subject to

$$v = \int_U \mu_C(u) r(u)\,\mathrm{d}u$$

$$\int_U r(u)\,\mathrm{d}u = 1$$

and r is the probability density of X.

Instantiating A, B, C and D, we obtain the answer to the query:
Probability that Robert is home at about 6:15 pm. is D, where

$$\mu_D(v) = \sup_r \left(\mu_{\text{usually}} \left(\int_U \mu_{*6:00\ \text{pm.}}(u) r(u)\, \mathrm{d}u \right) \right)$$

subject to

$$v = \int_U \mu_{\leqslant \circ *6:15\ \text{pm.}}(u) r(u)\, \mathrm{d}u$$

and

$$\int_U r(u)\, \mathrm{d}u = 1.$$

16.2. The tall Swedes problem

We start with the information set

p: Most Swedes are tall.

Assume that the queries are:

q_1: How many Swedes are not tall

q_2: How many are short

q_3: What is the average height of Swedes

In our earlier discussion of this example, we found that p translates into a generalized constraint on the count density function, h.

Thus,

$$p \longrightarrow \int_a^b h(u) \mu_{\text{all}}(u)\, \mathrm{d}u \text{ is most,}$$

where a and b are the lower and upper bounds on the height of Swedes.

Precisiands of q_1, q_2 and q_3 may be expressed as

$$q_1 : \longrightarrow \int_a^b h(u) \mu_{\text{not.tall}}(u)\, \mathrm{d}u,$$

$$q_2 : \longrightarrow \int_a^b h(u) \mu_{\text{short}}(u)\, \mathrm{d}u,$$

$$q_3 : \longrightarrow \int_a^b u h(u)\, \mathrm{d}u,$$

Considering q_1, we note that

$$\mu_{\text{not.tall}}(u) = 1 - \mu_{\text{tall}}(u)$$

Consequently,

$$q_1 \longrightarrow 1 - \int_a^b h(u) \mu_{\text{tall}}(u)\, \mathrm{d}u$$

which may be rewritten as

$$\text{ans}(q_1) \longrightarrow \text{1-most},$$

where 1-most plays the role of the antonym of most (Fig. 8).

Considering q_2, we have to compute

$$A\colon \int_a^b h(u)\mu_{\text{short}}(u)\,\mathrm{d}u$$

given that $\int_a^b h(u)\mu_{\text{tall}}(u)\,\mathrm{d}u$ is most

Applying the extension principle, we arrive at the desired answer to the query:

$$\mu_A(v) = \sup_h \left(\mu_{\text{most}}\left(\int_a^b h(u)\mu_{\text{tall}}(u)\,\mathrm{d}u\right)\right)$$

subject to

$$v = \int_a^b h(u)\mu_{\text{short}}(u)\,\mathrm{d}u$$

and

$$\int_a^b h(u)\,\mathrm{d}u = 1$$

Likewise, for q_3 we have as the answer

$$\mu_A(v) = \sup_h \left(\mu_{\text{most}}\left(\int_a^b h(u)\mu_{\text{tall}}(u)\,\mathrm{d}u\right)\right)$$

subject to

$$v = \int_a^b uh(u)\,\mathrm{d}u$$

and

$$\int_a^b h(u)\,\mathrm{d}u = 1.$$

As an illustration of application of protoformal deduction to an instance of this example, consider

p: Most Swedes are tall

q: How many Swedes are short?

We start with the protoforms of p and q (see earlier example):

Most Swedes are tall $\longrightarrow \frac{1}{n}\Sigma\text{Count}(G[A \text{ is } R]) \text{ is } Q$,

T Swedes are short $\longrightarrow \frac{1}{n}\Sigma\text{Count}(G[A \text{ is } S]) \text{ is } T$, where

$$G[A] = \Sigma_i \text{ Name}_i/A_i, \quad i = 1, \ldots, n.$$

An applicable deduction rule in symbolic form is

$$\frac{\frac{1}{n}\Sigma\text{Count}(G[A \text{ is } R]) \text{ is } Q}{\frac{1}{n}\Sigma\text{Count}(G[A \text{ is } S]) \text{ is } T}.$$

The computational part of this rule is expressed as

$$\frac{\frac{1}{n}\Sigma_i \mu_R(A_i) \text{ is } Q}{\frac{1}{n}\Sigma_i \mu_S(A_i) \text{ is } T},$$

where

$$\mu_T(v) = \sup_{A_i,\ldots,A_n} \mu_Q(\Sigma_i \mu_R(A_i))$$

subject to

$$v = \Sigma_i \mu_S(A_i).$$

What we see is that computation of the answer to the query, q, reduces to the solution of a variational problem, as it does in the earlier discussion of this example—a discussion in which protoformal deduction was not employed.

16.3. Tall Swedes and tall Italians

p: Most Swedes are much taller than most Italians
q: What is the difference in the average height of Swedes and the average height of Italians?
Step 1: Precisiation: translation of p into GCL

$$S = \{S_1, \ldots, S_n\}: \text{ population of Swedes,}$$

$$I = \{I_1, \ldots, I_n\}: \text{ population of Italians,}$$

$$g_i = \text{ height of } S_i, \quad g = (g_1, \ldots, g_m),$$

$$h_j = \text{ height of } I_j, \quad h = (h_1, \ldots, h_n),$$

$$\mu_{ij} = \mu_{\text{much.taller}}(g_i, h_j) = \text{ degree to which } S_i \text{ is much taller than } I_j,$$

$$r_i = \frac{I}{n}\Sigma_j \mu_{ij} = \text{Relative } \Sigma\text{Count of Italians in relation to whom } S_i \text{ is much taller,}$$

$$t_i = \mu_{\text{most}}(r_i) = \text{ degree to which } S_i \text{ is much taller than most Italians,}$$

$$v = \frac{I}{m}\Sigma t_i = \text{ Relative } \Sigma\text{Count of Swedes who are much taller than most Italians,}$$

$$\text{ts}(g, h) = \mu_{\text{most}}(v),$$

$$p \longrightarrow \text{ generalized constraint on } S \text{ and } I,$$

$$q : d = \frac{I}{m}\Sigma_i g_i - \frac{1}{n}\Sigma_j h_j.$$

Step 2: Deduction via extension principle

$$\mu_q(d) = \sup_{g,h} \text{ts}(g, h)$$

subject to

$$d = \frac{1}{m}\Sigma_i g_i - \frac{1}{n}\Sigma_j h_j.$$

16.4. Simplified trip planning

Probably it will take about two hours to get from San Francisco to Monterey, and it will probably take about five hours to get from Monterey to Los Angeles. What is the probability of getting to Los Angeles in less than about seven hours?

BD : (probably,*2) + (probably,*5)

↑ ↑

X Y

$$\underset{w}{\overset{\uparrow}{Z}} = \underset{u}{\overset{\uparrow}{X}} + \underset{v}{\overset{\uparrow}{Y}} \qquad p_Z(w) = \int p_X(u)\,p_Y(w-u)\,\mathrm{d}u$$

query: $\int p_Z(w)\mu_{\leqslant \circ * 7}(w)\,\mathrm{d}w$ is $?A$

$$\text{query relevant information: } \begin{cases} \pi_{p_X} = \mu_{\text{probably}}\left(\int \mu_{*2}(u)\,p_X(u)\,\mathrm{d}u\right) \\ \pi_{p_Y} = \mu_{\text{probably}}\left(\int \mu_{*5}(v)\,p_Y(v)\,\mathrm{d}v\right) \end{cases}$$

$$\mu_A(t) = \sup_{p_X, p_Y}(\pi_X \wedge \pi_Y)$$

subject to

$$t = \int p_Z(w)\mu_{\leqslant \circ * 7}(w)\,\mathrm{d}w.$$

17. Concluding remark

The theory of uncertainty which is outlined in this paper may be viewed as a radical step toward abandonment of bivalence and shifting the foundation of the theory from bivalent logic to fuzzy logic. Though only a step, it is a step which has wide-ranging ramifications. Standing out in importance is achievement of NL-capability. This capability opens the door to extensive enlargement of the role of natural languages in themes of uncertainty, decision analysis, economics and other fields in which human perceptions play an important role.

Acknowledgements

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

References

Bardossy, A., Duckstein, L., 1995. Fuzzy Rule-Based Modelling with Application to Geophysical Biological and Engineering Systems. CRC Press, Boca Raton, FL.

Bargiela, A., Pedrycz, W., 2002. Granular Computing. Kluwer Academic Publishers, Dordrecht.

Belohlavek, R., Vychodil, V., 2006. Attribute implications in a fuzzy setting. In: Ganter, B., Kwuida, L. (Eds.), ICFCA 2006, Lecture Notes in Artificial Intelligence, vol. 3874. Springer, Heidelberg, pp. 45–60.

Bouchon-Meunier, B., Yager, R.R., Zadeh, L.A. (Eds.), 2000. Uncertainty in Intelligent and Information Systems. Advances in Fuzzy Systems—Applications and Theory, vol. 20, World Scientific, Singapore.

Carnap, R., 1950. The Logical Foundations of Probability. University of Chicago Press, Chicago.

Colubi, A., Santos Domínguez-Menchero, J., López-Díaz, M., Ralescu, D.A., 2001. On the formalization of fuzzy random variables. Inf. Sci. 133 (1–2), 3–6.

Cresswell, M.J., 1973. Logic and Languages. Methuen, London, UK.

de Cooman, G., 2005. A behavioural model for vague probability assessments. Fuzzy Sets and Systems 154 (3), 305–358.

Dempster, A.P., 1967. Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Statist. 38, 325–329.

Di Nola, A., Sessa, S., Pedrycz, W., Sanchez, E., 1989. Fuzzy Relation Equations and Their Applications to Knowledge Engineering. Kluwer, Dordrecht.

Dubois, D., Prade, H., 1988. Representation and combination of uncertainty with belief functions and possibility measures. Comput. Intell. 4, 244–264.

Dubois, D., Prade, H., 1992. Gradual inference rules in approximate reasoning. Inf. Sci. 61 (1–2), 103–122.

Dubois, D., Prade, H., 1994. Non-Standard Theories of Uncertainty in Knowledge Representation and Reasoning. KR, 634–645.

Dubois, D., Parde, H. (Eds.), 1996. Fuzzy Information Engineering: A Guided Tour of Applications. Wiley, New York.

Filev, D., Yager, R.R., 1994. Essentials of Fuzzy Modeling and Control. Wiley-Interscience, New York.

Goguen, J.A., 1969. The logic of inexact concepts. Synthese 19, 325–373.

Goodman, I.R., Nguyen, H.T., 1985. Uncertainty Models for Knowledge-Based Systems. North-Holland, Amsterdam.

Jamshidi, M., Titli, A., Zadeh, L.A., Boverie, S. (Eds.), 1997. Applications of Fuzzy Logic—Towards High Machine Intelligence Quotient Systems. Environmental and Intelligent Manufacturing Systems Series, vol. 9. Prentice-Hall, Upper Saddle River, NJ.

Kaufmann, A., Gupta, M.M., 1985. Introduction to Fuzzy Arithmetic: Theory and Applications. Von Nostrand, New York.

Klir, G.J., 2004. Generalized information theory: aims, results, and open problems. Reliab. Eng. System Saf. 85 (1–3), 21–38.

Klir, G.J., 2006. Uncertainty and Information: Foundations of Generalized Information Theory. Wiley-Interscience, Hoboken, NJ.

Lambert, K., Van Fraassen, B.C., 1970. Meaning relations, Possible Objects and Possible Worlds. Philos. Probl. Logic 1–19.

Lawry, J., 2001. A methodology for computing with words. Internat. J. Approx. Reason. 28, 51–89.

Lawry, J., Shanahan, J.G., Ralescu, A.L. (Eds.), 2003. Modelling with Words—Learning, Fusion, and Reasoning within a Formal Linguistic Representation Framework, Springer, Berlin.

Mamdani, E.H., Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. Internat. J. Man–Machine Studies 7, 1–13.

Mares, M., 1994. Computation Over Fuzzy Quantities. CRC, Boca Raton, FL.

Nguyen, H.T., 1993. On Modeling of Linguistic Information Using Random Sets, Fuzzy Sets for Intelligent Systems. Morgan Kaufmann Publishers, San Mateo, CA, pp. 242–246.

Nguyen, H.T., Kreinovich, V., Di Nola, A., 2003. Which truth values in fuzzy logics are definable? Internat. J. Intell. Systems 18 (10), 1057–1064.

Novak, V., Perfilieva, I., Mockor, J., 1999. Mathematical Principles of Fuzzy Logic. Kluwer, Boston/Dordrecht.

Orlov, A.I., 1980. Problems of Optimization and Fuzzy Variables. Znaniye, Moscow.

Partee, B., 1976. Montague Grammar. Academic, New York.

Pedrycz, W., Gomide, F., 1998. Introduction to Fuzzy Sets. MIT Press, Cambridge, MA.

Puri, M.L., Ralescu, D.A., 1993. Fuzzy Random Variables, Fuzzy Sets for Intelligent Systems. Morgan Kaufmann Publishers, San Mateo, CA, pp. 265–271.

Ross, T.J., 2004. Fuzzy Logic with Engineering Applications. second ed. Wiley, New York.

Rossi, F., Codognet, P., 2003. Soft Constraints, Special issue on Constraints, vol. 8. no. 1. Kluwer, Dordrecht.

Sainsbury, R.M., 1995. Paradoxes. Cambridge University Press, Cambridge.

Schum, D., 1994. Evidential Foundations of Probabilistic Reasoning. Wiley, New York.

Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press, Princeton, NJ.

Singpurwalla, N.D., Booker, J.M., 2004. Membership functions and probability measures of fuzzy sets. J. Amer. Statist. Assoc. 99 (467), 867–889.

Smets, P., 1996. Imperfect information: imprecision and uncertainty, Uncertain. Manag. Inf. Systems 225–254.

Walley, P., 1991. Statistical Reasoning with Imprecise Probabilities. Chapman & Hall, London.

Wang, P.Z., Sanchez, E., 1982. In: Gupta, M.M., Sanchez, E. (Eds.), Treating a Fuzzy Subset as a Projectable Random Set, Fuzzy Information and Decision Processes. North-Holland, Amsterdam, pp. 213–220.

Yager, R.R., 2002. Uncertainty representation using fuzzy measures. IEEE Trans. Systems, Man Cybernet. Part B 32, 13–20.

Yager, R.R., 2006. Perception based granular probabilities in risk modeling and decision making. IEEE Trans. Fuzzy Systems 14, 129–139.

Yen, J., Langari, R., 1998. Fuzzy Logic: Intelligence, Control and Information. first ed. Prentice-Hall, Englewood Cliffs, NJ.

Yen, J., Langari, R., Zadeh, L.A. (Eds.), 1995. Industrial Applications of Fuzzy Logic and Intelligent Systems, IEEE, New York.

Zadeh, L.A., 1965. Fuzzy sets. Inform. and Control 8, 338–353.

Zadeh, L.A., 1968. Probability measures of fuzzy events. J. Math. Anal. Appl. 23, 421–427.

Zadeh, L.A., 1973. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. Systems, Man Cybernet. SMC-3, 28–44.

Zadeh, L.A., 1974. On the analysis of large scale systems. In: Gottinger, H. (Ed.), Systems Approaches and Environment Problems. Vandenhoeck and Ruprecht, Gottingen, pp. 23–37.

Zadeh, L.A., 1975a. Fuzzy logic and approximate reasoning. Synthese 30, 407–428.

Zadeh, L.A., 1975b. The concept of a linguistic variable and its application to approximate reasoning. Part I: Inf. Sci. 8, 199–249 Part II: Inf. Sci. 8, 301–357; Part III: Inf. Sci. 9, 43–80.

Zadeh, L.A., 1976. A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. Internat. J. Man–Machine Studies 8, 249–291.

Zadeh, L.A., 1978. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3–28.

Zadeh, L.A., 1979a. Fuzzy sets and information granularity. In: Gupta, M., Ragade, R., Yager, R. (Eds.), Advances in Fuzzy Set Theory and Applications. North-Holland Publishing Co., Amsterdam, pp. 3–18.

Zadeh, L.A., 1979b. A theory of approximate reasoning. In: Hayes, J., Michie, D., Mikulich, L.I. (Eds.), Machine Intelligence, vol. 9. Halstead Press, New York, pp. 149–194.

Zadeh, L.A., 1981a. Possibility theory and soft data analysis. In: Cobb, L., Thrall, R.M. (Eds.), Mathematical Frontiers of the Social and Policy Sciences. Westview Press, Boulder, CO, pp. 69–129.

Zadeh, L.A., 1981b. Test-score semantics for natural languages and meaning representation via PRUF. In: Rieger, B. (Ed.), Empirical Semantics, 1982. Brockmeyer, Bochum, W. Germany, pp. 281–349 Also Technical Memorandum 246, AI Center, SRI International, Menlo Park, CA.

Zadeh, L.A., 1983a. A computational approach to fuzzy quantifiers in natural languages. Comput. Math. 9, 149–184.

Zadeh, L.A., 1983b. A fuzzy-set-theoretic approach to the compositionality of meaning: propositions, dispositions and canonical forms. J. Semantics 3, 253–272.

Zadeh, L.A., 1984. Precisiation of meaning via translation into PRUF. In: Vaina, L., Hintikka, J. (Eds.), Cognitive Constraints on Communication. Reidel, Dordrecht, pp. 373–402.

Zadeh, L.A., 1986. Outline of a computational approach to meaning and knowledge representation based on the concept of a generalized assignment statement. In: Thoma, M., Wyner, A. (Eds.), Proceedings of the International Seminar on Artificial Intelligence and Man–Machine Systems. Springer, Heidelberg, pp. 198–211.

Zadeh, L.A., 1996. Fuzzy logic and the calculi of fuzzy rules and fuzzy graphs. Multiple-Valued Logic 1, 1–38.

Zadeh, L.A., 1997. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets and Systems 90, 111–127.

Zadeh, L.A., 1998. Some reflections on soft computing granular computing and their roles in the conception, design and utilization of information/intelligent systems. Soft Comput. 2, 23–25.

Zadeh, L.A., 1999. From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions. IEEE Trans. Circuits Systems 45, 105–119.
Zadeh, L.A., 2002. Toward a perception-based theory of probabilistic reasoning with imprecise probabilities. J. Statist. Plann. Inference 105, 233–264.
Zadeh, L.A., 2004a. A note on web intelligence, world knowledge and fuzzy logic. Data Knowledge Eng. 50, 291–304.
Zadeh, L.A., 2004b. Precisiated natural language (PNL). AI Mag. 25 (3), 74–91.
Zadeh, L.A., 2005. Toward a generalized theory of uncertainty - an outline. Inf. Sciences 172, 1–40.