ELSEVIER

Fuzzy Sets and Systems 90 (1997) 111–127

FUZZY sets and systems

Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic

Lotfi A. Zadeh*[1,2]

Computer Science Division/Electronics Research Laboratory, Department of EECS, University of California, Berkeley, CA 94720-1776, USA

Received February 1997

Abstract

There are three basic concepts that underlie human cognition: granulation, organization and causation. Informally, granulation involves decomposition of whole into parts; organization involves integration of parts into whole; and causation involves association of causes with effects.

Granulation of an object A leads to a collection of granules of A, with a granule being a clump of points (objects) drawn together by indistinguishability, similarity, proximity or functionality. For example, the granules of a human head are the forehead, nose, cheeks, ears, eyes, etc. In general, granulation is hierarchical in nature. A familiar example is the granulation of time into years, months, days, hours, minutes, etc.

Modes of information granulation (IG) in which the granules are crisp (c-granular) play important roles in a wide variety of methods, approaches and techniques. Crisp IG, however, does not reflect the fact that in almost all of human reasoning and concept formation the granules are fuzzy (f-granular). The granules of a human head, for example, are fuzzy in the sense that the boundaries between cheeks, nose, forehead, ears, etc. are not sharply defined. Furthermore, the attributes of fuzzy granules, e.g., length of nose, are fuzzy, as are their values: long, short, very long, etc. The fuzziness of granules, their attributes and their values is characteristic of ways in which humans granulate and manipulate information.

The theory of fuzzy information granulation (TFIG) is inspired by the ways in which humans granulate information and reason with it. However, the foundations of TFIG and its methodology are mathematical in nature.

The point of departure in TFIG is the concept of a generalized constraint. A granule is characterized by a generalized constraint which defines it. The principal types of granules are: possibilistic, veristic and probabilistic.

The principal modes of generalization in TFIG are fuzzification (f-generalization); granulation (g-generalization); and fuzzy granulation (f.g-generalization), which is a combination of fuzzification and granulation. F.g-generalization underlies the basic concepts of linguistic variable, fuzzy if–then rule and fuzzy graph. These concepts have long played a major role in the applications of fuzzy logic and differentiate fuzzy logic from other methodologies for dealing with imprecision and uncertainty. What is important to recognize is that no methodology other than fuzzy logic provides a machinery for fuzzy information granulation.

*Tel: + 1-510-642-4959; fax: + 1-510-642-1712; e-mail: zadeh@cs.berkeley.edu.

[1]Research supported in part by NASA Grant NCC 2-275, ONR Grant N00014-96-1-0556, LLNL Grant 442427-26449, ARO Grant DAAH 04-961-0341, and the BISC Program of UC Berkeley.

[2]To Didier Dubois and Henri Prade, who have contributed in so many major ways to the development of fuzzy logic and its applications.

PII S0165-0114(97)00077-8

TFIG builds on the existing machinery for fuzzy information granulation in fuzzy logic but takes it to a significantly higher level of generality, consolidates its foundations and suggests new directions. In coming years, TFIG is likely to play an important role in the evolution of fuzzy logic and, in conjunction with computing with words (CW), may well have a wide-ranging impact on its applications. The impact of TFIG is likely to be felt most strongly in those fields in which there is a wide gap between theory and reality.

Keywords: Fuzzy logic; Fuzzy information granulation; Approximate reasoning

1. Preamble

As the papers in this issue make amply clear, during the past decade fuzzy logic has evolved into a well-structured system of concepts and techniques with a solid mathematical foundation and a widening array of applications ranging from basic sciences to engineering systems, social systems, biomedical systems and consumer products.

And yet there is a basic issue in fuzzy logic that has not been highlighted to the extent that it should. The issue is the centrality of the role of fuzzy information granulation – a mode of granulation which underlies the concepts of linguistic variable, fuzzy if–then rule and fuzzy graph. Clearly, the machinery of fuzzy information granulation has played and is continuing to play a pivotal role in the applications of fuzzy logic. But what is beginning to crystallize is a basic theory of fuzzy information granulation (TFIG) which casts fuzzy logic in a new light and, in time, may come to be recognized as its quintessence. This is the perception that I should like to articulate in this paper.

My perception may be viewed as an evolution of ideas rooted in my 1965 paper on fuzzy sets [24]; 1971 paper on fuzzy systems [26]; 1973–1976 papers on linguistic variables, fuzzy if–then rules and fuzzy graphs [27–30]; 1979 paper on fuzzy sets and information granularity [31]; 1986 paper on generalized constraints [32] and 1996 paper on computing with words [37]. Furthermore, it reflects many important contributions by others both to the foundations of fuzzy logic and its applications. Among my papers, the 1973 paper in which the basic concepts of linguistic variable and fuzzy if–then were introduced may be viewed as a turning point at which the foundation of TFIG was laid.

In what follows, what I will have to say should be viewed as a summary rather than a full exposition. A more detailed account of the theory of fuzzy information granulation is in the process of gestation.

2. Introduction

Among the basic concepts which underlie human cognition there are three that stand out in importance. The three are: granulation, organization and causation. In a broad sense, granulation involves decomposition of whole into parts; organization involves integration of parts into whole; and causation relates to association of causes with effects (Fig. 1).

Informally, granulation of an object A results in a collection of granules of A, with a granule being a clump of objects (or points) which are drawn together by indistinguishability, similarity, proximity or functionality (Fig. 2). In this sense, the granules of a human body are the head, neck, arms,

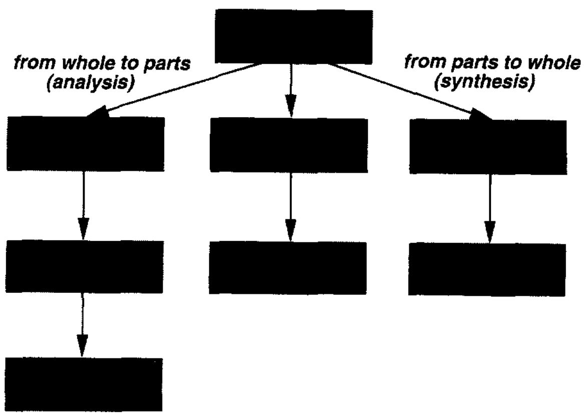


Fig. 1. Fundamental concepts in human cognition: granulation, organization and causation.

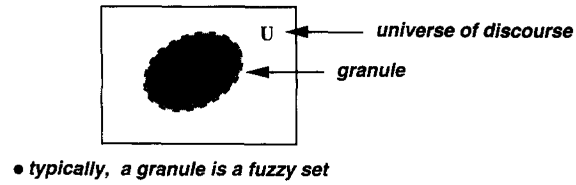


Fig. 2. A granule is a clump of objects (or points) which are drawn together by indistinguishability, similarity, proximity or functionality.

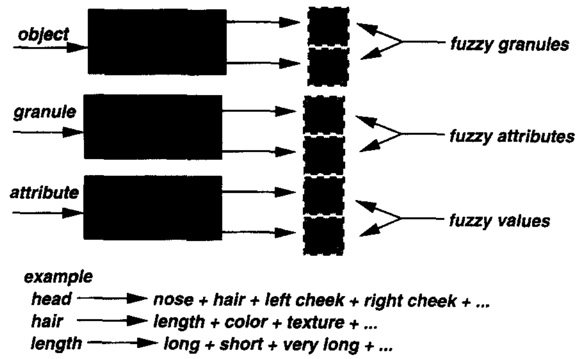


Fig. 3. Basic structure of fuzzy information granulation: granulation, attribution and valuation.

chest, etc. In turn, the granules of a head are the forehead, cheeks, nose, ears, eyes, hair, etc. In general, granulation is hierarchical in nature. A familiar example is granulation of time into years, years in months, months into days and so on.

Modes of information granulation (IG) in which the granules are crisp (c-granular) play important roles in a wide variety of methods, approaches and techniques. Among them are: interval analysis, quantization, rough set theory, diakoptics, divide and conquer, Dempster–Shafer theory, machine learning from examples, chunking, qualitative process theory, decision trees, semantic networks, analog-to-digital conversion, constraint programming, Prolog, cluster analysis and many others.

Important though it is, crisp information granulation (crisp IG) has a major blind spot. More specifically, it fails to reflect the fact that in much, perhaps most, of human reasoning and concept formation the granules are fuzzy (f-granular) rather than crisp. In the case of a human body, for example, the granules are fuzzy in the sense that the boundaries of the head, neck, arms, legs, etc. are not sharply defined. Furthermore, the granules are associated with fuzzy attributes, e.g., length, color and texture in the case of hair. In turn, granule attributes have fuzzy values, e.g., in the case of the fuzzy attribute length (hair), the fuzzy values might be long, short, very long, etc. The fuzziness of granules, their attributes and their values is characteristic of the ways in which human concepts are formed, organized and manipulated (Fig. 3).

A point that is worthy of note is that attributes may be associated with two or more granules, in which case they might be referred to as *intergranular* attributes. An example of an intergranular attribute is the distance between ears, with the understanding that ears are f-granules of head.

In human cognition, fuzziness of granules is a direct consequence of fuzziness of the concepts of indistinguishability, similarity, proximity and functionality. Furthermore, it is entailed by the finite capacity of the human mind and sensory organs to resolve detail and store information. In this perspective, fuzzy information granulation (fuzzy IG) may be viewed as a form of lossy data compression.

Fuzzy information granulation underlies the remarkable human ability to make rational decisions in an environment of imprecision, partial knowledge, partial certainty and partial truth. And yet, despite its intrinsic importance, fuzzy information granulation has received scant attention except in the domain of fuzzy logic, in which, as was pointed already, fuzzy IG underlies the basic concepts of linguistic variable, fuzzy if–then rule and fuzzy graph. In fact, the effectiveness and successes of fuzzy logic in dealing with real-world problems rest in large measure on the use of the machinery of fuzzy information granulation. This machinery is unique to fuzzy logic and differentiates it from all other methodologies. In this connection, what should be underscored is that when we talk about fuzzy information granulation we are not talking about a single fuzzy granule; we are talking about a collection of fuzzy granules which result from granulating a crisp or fuzzy object.

The theory of fuzzy information granulation (TFIG) outlined in this paper builds on the existing machinery of fuzzy IG in fuzzy logic but goes far beyond it. Basically, TFIG draws its inspiration from the informal ways in which humans employ fuzzy information granulation but its foundation and methodology are mathematical in nature.

In this perspective, fuzzy information granulation may be viewed as a mode of generalization which may be applied to any concept, method or theory. Related to fuzzy IG are the following principal modes of generalization.

(a) Fuzzification (f-generalization). In this mode of generalization, a crisp set is replaced by a fuzzy set (Fig. 4).
(b) Granulation (g-generalization). In this case, a set is partitioned into granules (Fig. 5).
(c) Randomization (r-generalization). In this case, a variable is replaced by a random variable.
(d) Usualization (u-generalization). In this case, a proposition expressed as X *is* A is replaced with usually (X *is* A).

These and other modes of generalization may be employed in combination. A combination that is of particular importance is the conjunction of fuzzification and granulation. This combination plays a pivotal role in the theory of fuzzy information granulation and fuzzy logic, and will be referred to as f.g-generalization (or f-granulation or fuzzy granulation).

As a mode of generalization, f.g-generalization may be applied to any concept, method or theory. In particular, in application to the basic concepts of variable, function and relation, f.g-generalization leads, in fuzzy logic, to the basic concepts of linguistic variable, fuzzy rule set and fuzzy graph (Fig. 6). These concepts are unique to fuzzy logic and play a central role in its applications.

The distinctive concepts of f-generalization, g-generalization, r-generalization and f.g-generalization make a significant contribution to a better understanding of fuzzy logic and its relation to other methodologies for dealing with uncertainty and imprecision. In particular, crisp g-generalization of set theory and relational models of data lead to rough set theory [18]. F-generalization of classical logic and set theory leads to multiple-valued logic, fuzzy logic in its narrow sense and parts of fuzzy set theory (Fig. 7). But it is f.g-generalization that leads to fuzzy logic (FL) in its wide sense and underlies most of its applications. This is a key point that is frequently overlooked in

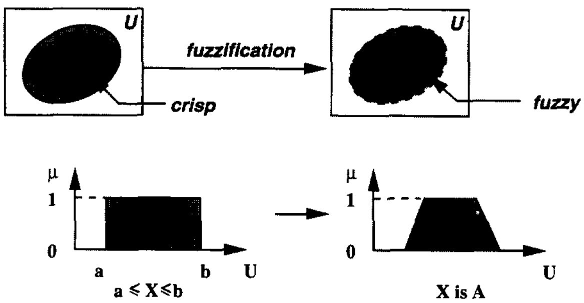


Fig. 4. Fuzzification: crisp set → fuzzy set.

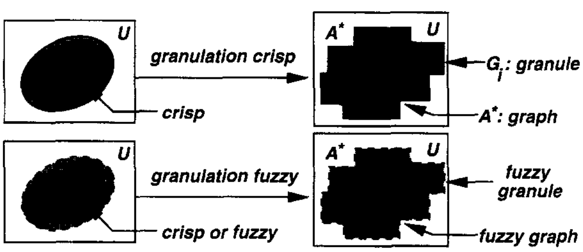


Fig. 5. Granulation. Crisp granulation: crisp set is partitional into crisp granules. Fuzzy granulation: crisp or fuzzy set is partitioned into fuzzy granules.

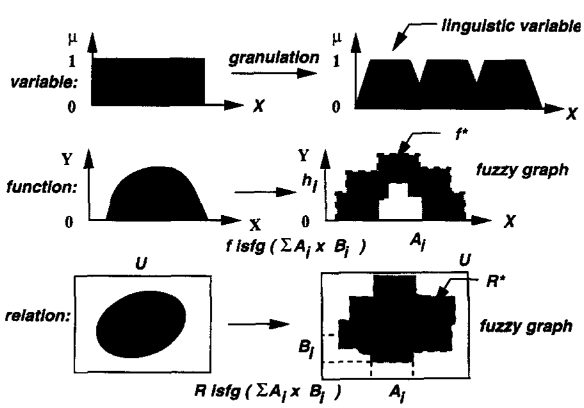


Fig. 6. Granulation of the basic mathematical concepts of variable function and relation. Linguistic variable = f-granular variable. A fuzzy graph may be represented as a fuzzy rule set and vice versa. R *isfg* T means that R is constrained by the fuzzy graph T.

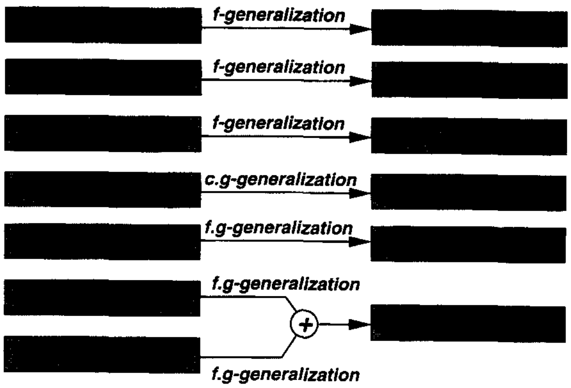


Fig. 7. Theories resulting from applying various modes of generalization.

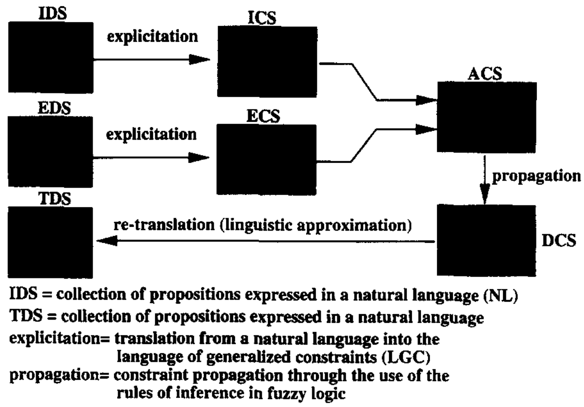


Fig. 8. Basic structure of computing with words (CW).

discussions about fuzzy logic and its relation to other methodologies.

The theory of fuzzy information granulation serves to highlight the centrality of the concept of fuzzy information granulation in fuzzy logic. More importantly, the theory provides a basis for computing with words (CW) [37]. In effect, CW is an integral part of TFIG. However, since it is discussed elsewhere [37], it will suffice in this paper to summarize its essential features.

The point of departure in CW is the observation that in a natural language words play the role of labels of fuzzy granules. In computing with words, a proposition is viewed as an implicit fuzzy constraint on an implicit variable. The meaning of a proposition is the constraint which it represents.

In CW, the initial data set (IDS) is assumed to consist of a collection of propositions expressed in a natural language. The result of computation, referred to as the terminal data set (TDS), is likewise a collection of propositions expressed in a natural language. To infer TDS from IDS the rules of inference in fuzzy logic are used for constraint propagation from premises to conclusions (Fig. 8).

There are two main rationales for computing with words. First, computing with words is a necessity when the available information is not precise enough to justify the use of numbers. And second, computing with words is advantageous when there is a tolerance for imprecision, uncertainty and partial truth that can be exploited to achieve tractability, robustness, low solution cost and better rapport with reality. In coming years, computing with words is likely to evolve into an important methodology in its own right with wide-ranging applications on both basic and applied levels.

Inspired by the ways in which humans granulate human concepts, we can proceed to granulate conceptual structures in various fields of science. In a sense, this is what motivates computing with words. An intriguing possibility is to granulate the conceptual structure of mathematics. This would lead to what may be called granular mathematics. Eventually, granular mathematics may evolve into a distinct branch of mathematics having close links to the real world. A subset of granular mathematics and a superset of computing with words is granular computing.

In the final analysis, fuzzy information granulation is central to fuzzy logic because it is central to human reasoning and concept formation. It is this aspect of fuzzy IG that underlies its essential role in the conception and design of intelligent systems. In this regard, what is conclusive is that there are many, many tasks which humans can perform with ease and that no machine could perform without the use of fuzzy information granulation.

A typical example is the problem of estimation of age from voice. More specifically, consider a common situation where A gets a telephone call from B, whom A does not know. After hearing B talk for 5–10 seconds, A would be able to form a rough estimate of B's age and express it as, say, "*B is old*" or "*It is very likely that B is old*", in which both age and probability play the role of linguistic, that is, f-granulated variables. Neither A nor any machine could come up with crisp estimates of B's age, e.g., "*B is* 63" or "*The probability that B is* 63 *is* 0.002". In this and similar cases, a machine would have to have a capability to process and reason with f-granulated information in order to come up with a machine solution to a problem that has a human solution expressed in terms of f-granulated variables.

A related point is that, in everyday decision-making, humans use that and only that information which is decision-relevant. For example, in playing golf, parking a car, picking up an object, etc., humans use fuzzy estimates of distance, velocity, angles, sizes, etc. In a pervasive way, decision-relevant information is f-granular. To perform such everyday tasks as effortlessly as humans can, a machine must have a capability to process f-granular information. A conclusion which emerges from these examples is that fuzzy information granulation is an integral part of human cognition. This conclusion has a thought-provoking implication for AI: Without the methodology of fuzzy IG in its armamentarium, AI cannot achieve its goals.

In what follows, we shall elaborate on the points made above and describe in greater detail the basic ideas underlying fuzzy information granulation and its role in fuzzy logic.

3. The concept of a generalized constraint

The point of departure in the theory of fuzzy information granulation is the concept of a generalized constraint [32]. For simplicity, we shall restrict our discussion to constraints which are unconditional.

Let X be a variable which takes values in a universe of discourse U. A generalized constraint on the values of X is expressed as X *isr* R, where R is the constraining relation, *isr* is a variable copula and r is a discrete variable whose value defines the way in which R constrains X.

The principal types of constraints and the values of r which define them are the following:

1. *Equality constraint*, $r = e$. In this case, X *ise* a means that $X = a$.

2. *Possibilistic constraint*, $r = blank$. In this case, if R is a fuzzy set with membership function $\mu_R : U \to [0, 1]$, and X is a disjunctive (possibilistic) variable, that is, a variable which cannot be assigned two or more values in U simultaneously, then

$$X \text{ is } R$$

means that R is the possibility distribution of X. More specifically,

$$X \text{ is } R \to Poss\{X = u\} = \mu_R(u), \quad u \in U.$$

A simple example of a possibilistic constraint is X *is small*. In this case, $Poss\{X = u\} = \mu_{small}(u)$. Constraints induced by propositions expressed in a natural language are for the most part possibilistic in nature. This is the reason why the simplest value, $r = blank$, is chosen to define possibilistic constraints.

3. *Veristic constraint*, $r = v$. In this case, if R is a fuzzy set with membership function μ_R and X is a conjunctive (veristic) variable, that is, a variable which can be assigned two or more values in U simultaneously, then $X \text{ isv } R \to Ver\{X = u\} = \mu_R(u)$, $u \in U$, where $Ver\{X = u\}$ is the verity (truth value) of $X = u$.

An example of a veristic constraint is the following. Let U be the universe of natural languages and let X denote the fluency of an individual in English, French and German. Then, X *isv* (1.0 English + 0.8 French + 0.6 Italian) means that the degrees of fluency of X in English, French and Italian are 1.0, 0.8 and 0.6, respectively.

It is important to observe that, in the case of a possibilistic constraint, the fuzzy set R plays the role of a possibility distribution, whereas in the case of a veristic constraint R plays the role of a verity distribution. What this implies is that, in general, any fuzzy, and ipso facto any crisp, set R admits of two different interpretations.[3] More specifically, in the possibilistic interpretation the grades of membership are possibilities, while in the veristic

[3] An insightful discussion of various possible interpretations of grades of membership in a fuzzy set is contained in the paper by D. Dubois and H. Prade, "The Semantics of Fuzzy Sets," in this issue.

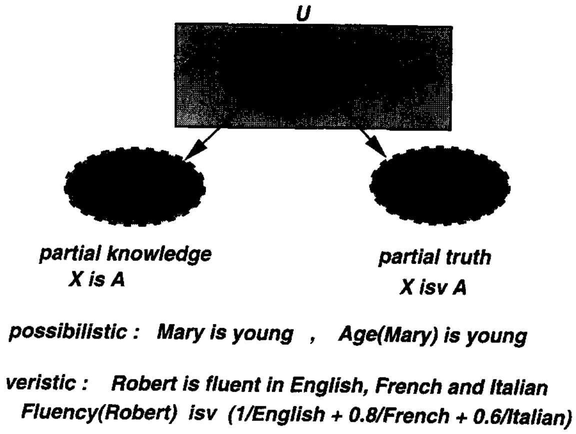


Fig. 9. Possibilistic and veristic interpretations of a fuzzy set.

interpretation the grades of membership are verities (truth values) (Fig. 9). Since in most cases constraints are possibilistic, the default assumption is that a fuzzy set plays the role of a possibility distribution.

4. *Probabilistic constraint*, $r = p$. In this case, X *isp* R means that X is a random variable and R is the probability distribution (or density) of X. For example, X *isp* $N(m, \sigma^2)$ means that X is a normally distributed random variable with mean m and variance σ^2. Similarly, X *isp* $(0.2\backslash a + 0.4\backslash b + 0.4\backslash c)$ means that X takes the values a, b, c with respective probabilities 0.2, 0.4 and 0.4.

5. *Probability value constraint*, $r = \lambda$. In this case, X *is*λ R signifies that what is constrained is the probability of a specified event, X *is* A. More specifically, X *is*λ $R \rightarrow$ *Prob*$\{X$ *is* $A\}$ *is* R. For example, if A = *small* and R = *likely*, then X *is*λ *likely* means that *Prob*$\{X$ *is small*$\}$ *is likely*.

6. *Random set constraint*, $r = rs$. In this case, X *isrs* R is a composite constraint which is a combination of probabilistic and possibilistic (or veristic) constraints. In a schematic form, a random set constraint may be represented as

$$Y \text{ isp } P$$

$$\frac{(X,Y) \text{ is } Q}{X \text{ isrs } R}$$

or

$$Y \text{ isp } P$$

$$\frac{(X,Y) \text{ isv } Q}{X \text{ isrs } R},$$

where Q is a joint possibilistic (or veristic) constraint on X and Y, and R is a random set, that is, a set-valued random variable. It is of interest to note that the Dempster–Shafer theory of evidence is in essence a theory of random set constraints.

7. *Fuzzy graph constraint*, $r = fg$. In this case, in X *isfg* R, X is a function and R is a fuzzy graph approximation to X (See Section 5). More specifically, if X is a function, $X : U \rightarrow V$, defined by a fuzzy rule set

if u *is* A_1, *then* v *is* B_1

if u *is* A_2, *then* v *is* B_2

...

if u *is* A_n, *then* v *is* B_n

where A_1 and B_1 are linguistic values of u and v, then R is the fuzzy graph [26, 28–30, 36],

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

where $A_i \times B_i$, $i = 1, \ldots, n$, is the cartesian product of A_i and B_i and $+$ represents disjunction or, more generally, an s-norm (Fig. 10).

A fuzzy graph constraint may be represented as a possibilistic constraint on the function which is approximated (Fig. 11). Thus, X *isfg* $R \rightarrow X$ *is* $(\sum_i A_i \times B_i)$.

In addition to the types of constraints defined above there are many others that are more specialized and less common. A question that arises is: What purpose is served by having a large variety of constraints to choose from.

A basic reason is that, in a general setting, information may be viewed as a constraint on a variable. For example, the proposition "Mary is young", conveys information about Mary's age by constraining the values that the variable Age (Mary) can take. Similarly, the proposition "Most Swedes are tall" may be interpreted as a

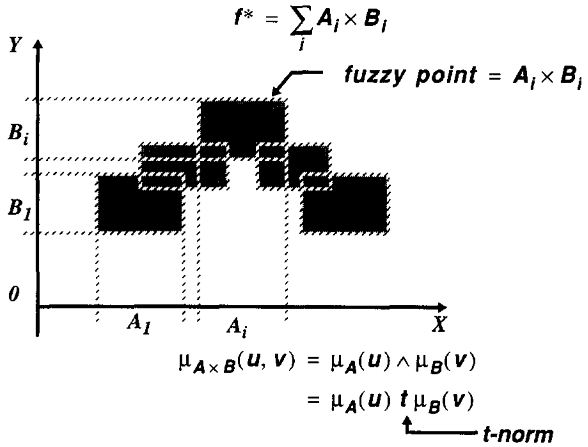


Fig. 10. A fuzzy graph f^* is a disjunction of cartesian products.

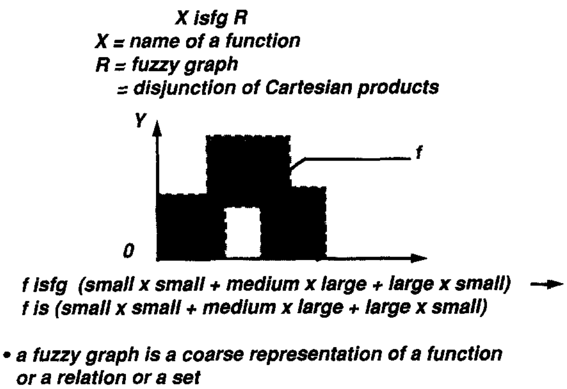


Fig. 11. Representation of a fuzzy graph constraint as a possibilistic constraint.

possibilistic constraint on the proportion of tall Swedes, that is,

most Swedes are tall

$\rightarrow$ *Proportion (tall Swedes/Swedes) is most*

in which the fuzzy quantifier *most* plays the role of a fuzzy number.

More generally, in the context of computing with words, a basic assumption is that a proposition, p, expressed in a natural language may be interpreted as a generalized constraint $p \rightarrow X\ isr\ R$. In this interpretation, $X\ isr\ R$ is the *canonical form* of p. The function of the canonical form is to place in evidence, i.e., explicitate, the implicit constraint which p represents.

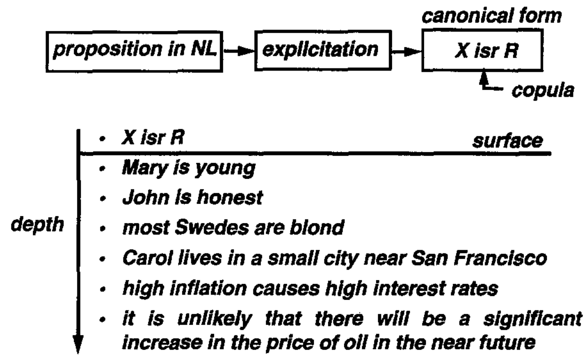


Fig. 12. Depth of explication of propositions in a natural language.

In CW [37], the depth of explicitation of a proposition is a measure of the effort involved in explicitating p, that is, translating, p into its canonical form. In this sense, the proposition $X\ isr\ R$ is a surface constraint (depth = zero). As shown in Fig. 12, the depth of explication increases in the downward direction. Thus, a proposition such as "Mary is young" is shallow, whereas "it is not very likely that there will be a significant increase in the price of oil in the near future" is not.

What we see, then, is that the information conveyed by a proposition expressed in a natural language is, in general, too complex to admit of representation as a simple, crisp constraint. This is the main reason why in representing the meaning of a proposition expressed in a natural language we need a wide variety of constraints which are subsumed under the rubric of generalized constraints.

4. Taxonomy of fuzzy granulation

The concept of generalized constraint provides a basis for a classification of fuzzy granules. More specifically, in the theory of fuzzy IG a granule, G, is viewed as a clump of points characterized by a generalized constraint. Thus,

$$G = \{X \mid X\ isr\ R\}.$$

In this context, the type of a granule is determined by the type of constraint which defines it (Fig. 13). In particular, possibilistic, veristic and

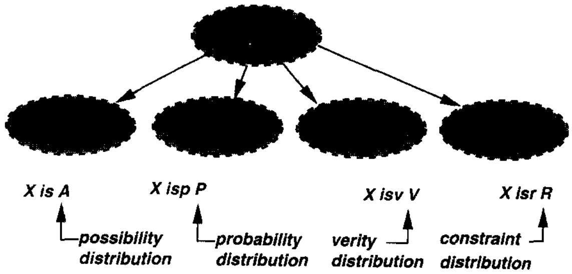


Fig. 13. Taxonomy of granulation.

probabilistic granules are defined, respectively, by possibilistic, veristic and probabilistic constraints. To illustrate, the granule

$G = \{X \mid X \text{ is } small\}$

is a possibilistic granule. The granule

$G = \{X \mid X \text{ isv } small\}$

is a veristic granule. And the granule

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

is a probabilistic (Gaussian) granule.

As a more concrete illustration consider the fuzzy granule *Nose* of a human head. If we associate with each point on the nose its grade of membership in *Nose*, the fuzzy granule *Nose* should be interpreted as a veristic granule. Now suppose that we associate with the attribute length (Nose) a fuzzy value *long*. The question is: What is the meaning of the proposition "Nose is long?"

Assume that the profile of *Nose*, N, has the form shown in Fig. 14. With each point p on the profile are associated two numbers: α, representing the grade of membership of p in *Nose*; and β, the degree of relevance of p to the value of the attribute length (*Nose*). In general, $\beta \leqslant \alpha$.

Now let $\bar{N}$ be a veristic fuzzy set which results from a rectification of the profile of *Nose* (Fig. 14). At this point, the original question reduces to "What is the length of $\bar{N}$?" This question is a familiar one in fuzzy logic. Assume for simplicity that the set is trapezoidal, as shown in Fig. 15. Then, by using the α-cuts of $\bar{N}$, its length may be represented as a veristic triangular fuzzy set $L(\bar{N})$ (Fig. 15). Thus, $L(\bar{N})$ is the answer to the original question.

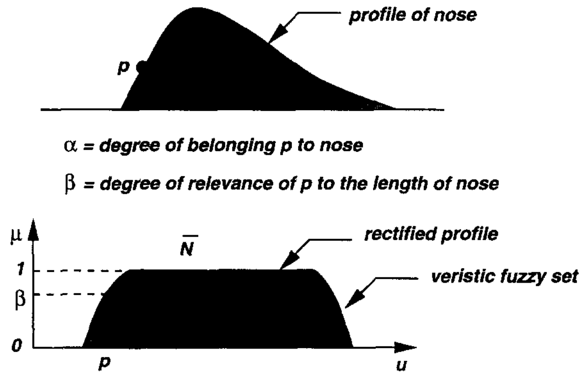


Fig. 14. Profile of *Nose* and its rectification.

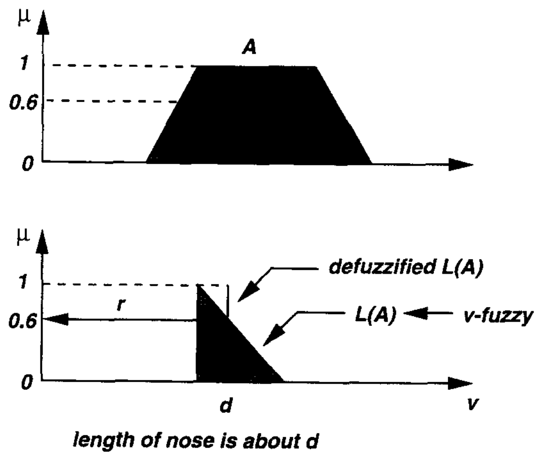


Fig. 15. Length of a trapezoidal fuzzy set and length of *Nose*.

However, if a single real value of the length of nose it required, $L(\bar{N})$ may be defuzzified using, say, the *COG* definition of defuzzification.

The purpose of this simple example is to show how a fuzzy value may be associated with a fuzzy attribute of a fuzzy granule. A more complex example would be an association of a fuzzy value *long* with the fuzzy attribute (*Hair*). In this case, the problem is very similar to that of associating a fuzzy value with the fuzzy attribute *unemployment* for a fuzzy segment of a population in a city, region or country.

In the foregoing discussion, classification of granules is based on the types of constraints which define them. A different mode of classification involves representation of complex granules as cartesian products or other combinations of simpler granules.

More specifically, let $G_1, \ldots, G_n$ be granules in $U_1, \ldots, U_n$, respectively. Then the granule $G = G_1 \times \cdots \times G_n$ is a cartesian granule. For simplicity, we shall assume that $n = 2$ (Fig. 16).

An important elementary property of cartesian granules relates to their α-cuts. Thus, if $G = G_1 \times G_2$ and G_α, $G_{1\alpha}$ and $G_{2\alpha}$ are α-cuts of G, G_1 and G_2, respectively, then

$$G_\alpha = G_{1\alpha} \times G_{2\alpha}.$$

A cartesian granule, G, may be rotated (Fig. 17). More generally, a cartesian granule, G, may be subjected to a coordinate transformation defined by

$$X \to f(X, Y),$$

$$Y \to g(X, Y).$$

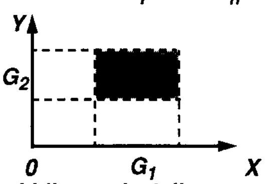
In this case, if G_1 and G_2 are defined by possibly different generalized constraints:

G_1: X *isr* A

G_2: X *iss* B

- ***a cartesian granule is a non-interactive conjunction (cartesian product) of granules***

$G = G_1 \times \ldots \times G_n$



example : G= middle-aged x tall

Fig. 16. Cartesian granule.

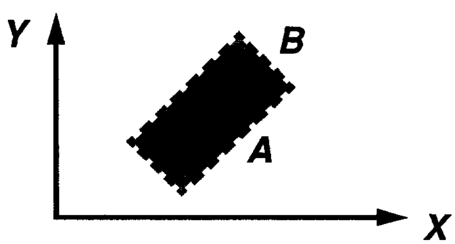


Fig. 17. Rotated cartesian granule.

then the transformed granule G^* is defined by

$$G^*: (f(X, Y)\ isr\ A) \times (g(X, Y)\ iss\ B).$$

A generalized constraint in which what is constrained is a function or a functional of a variable will be referred to as a *generalized functional constraint* (Fig. 18). Such constraints play an important role in computing with words.

The importance of the concept of a cartesian granule derives in large measure from its role in what might be called *encapsulation*.

More specifically, consider a granule, G, defined by a possibilistic constraint $G = \{(X, Y) \mid (X, Y)$ *is* $R\}$.

Let G_X and G_Y denote the projections of G on U and V, the domains of X and Y, respectively. Thus,

$$\mu_{G_X}(u) = \sup_v \mu_G(u, v), \quad u \in U, \ v \in V$$

$$\mu_{G_Y}(v) = \sup_u \mu_G(u, v).$$

Then, the cartesian granule G^+,

$$G^+ = G_X \times G_Y$$

encapsulates G in the sense that it is the least upper bound of cartesian granules which contain G. (Fig. 19). Invoking the entailment principle in fuzzy logic allows us to assert that

$$(X, Y)\ is\ G \ \Rightarrow\ (X, Y)\ is\ G^+.$$

Thus, G^+ can be used as an upper approximation to G [25]. It should be noted that in the case of veristic constraints the entailment principle asserts

if $F_1(X_1, \ldots, X_m)$ is C_{11} and ... $F_n(X_1, \ldots, X_m)$ is C_{1n} then Y_1 is D_{11} and ... Y_k is D_{1k}

...

if $F_{l1}(X_1, \ldots, X_m)$ is C_{l1} and ... $F_{ln}(X_1, \ldots, X_m)$ is C_{ln} then Y_1 is D_{l1} and ... Y_k is D_{lk}

- ***rule explosion***
- ***number of rules depends on the choice of features***

Fig. 18. Format of a fuzzy rule set representing a collection of possibilistic functional constraints.

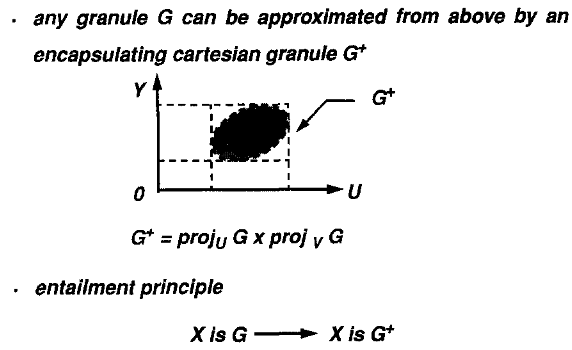


Fig. 19. A granule, G; its projection and its encapsulating granule, G^+.

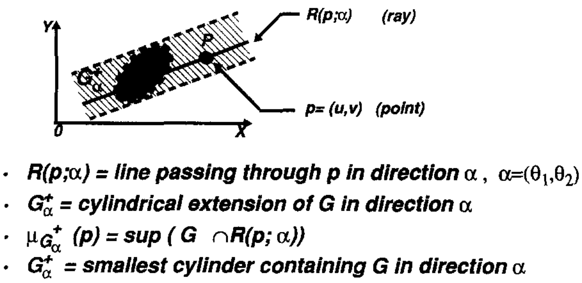


Fig. 20. G_α^+ is a cylindrical extension of G in direction α.

that

$$(X, Y)\, isv\, A \;\Rightarrow\; (X, Y)\, isv\, B$$

if $B \subset A$.

In a more general setting, we can construct a cylindrical extension of G in the manner shown in Fig. 20 [25]. More concretely, the cylindrical extension, G_α^+, of G in direction α is a cylindrical fuzzy set such that

$$\mu_{G_\alpha^+}(p) = \sup(G \cap R(p;\alpha))$$

where $R(p;\alpha)$ is a ray (line) passing through p in direction α, $\alpha = (\theta_1, \theta_2)$, where θ_1 and θ_2 are the angles that define α. By its construction, G_α^+ encapsulates G.

Let $G_{\alpha_1}^+, \ldots, G_{\alpha_u}^+$ be cylindrical extensions of G in directions $\alpha_1, \ldots, \alpha_u$, respectively. Then, the intersection of the $G_{\alpha_i}^+$ is a granule, G^+, that encapsulates G (Fig. 21). This concept of an encapsulating granule subsumes that of a cartesian encapsulating granule as a special case.

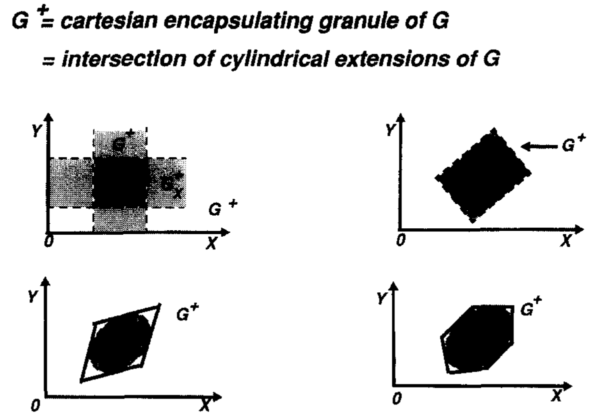


Fig. 21. Encapsulating granules generated by intersections of cylindrical extensions.

As shown in [25], an encapsulating granule G^+ may be viewed as an upper approximation to G. Dually, as shown in [25], one can define a lower approximation to G. However, these concepts of upper and lower approximation of fuzzy granules are different from those defined in the theory of rough sets [18].

5. Fuzzy graphs

One of the most basic facets of human cognition relates to the perception of dependencies and relations. In the theory of fuzzy information granulation, this facet of human cognition underlies the very basic concept of a *fuzzy graph*.

The concept of a fuzzy graph was introduced in [26] and was developed more fully in [28–30]. What might be called the *calculus of fuzzy graphs* [36] lies at the center of fuzzy logic and is employed in most of its applications.

In the context of fuzzy information granulation, a fuzzy graph may be viewed as the result of f.g-generalization of the concepts of function and relation (Fig. 6).

As the point of departure consider a function (or a relation) f which is defined by a table of the form

f	X	Y
	a_1	b_1
	a_2	b_2
	•	•
	a_n	b_n

F.g-generalization of f results in a function f^* whose defining table is of the form

f^*	X	Y
	A_1	B_1
	A_2	B_2
	•	•
	A_n	B_n

(1)

where X and Y play the role of linguistic (granular) variables, with the A_i and B_i, $i = 1, \dots, n$, representing their linguistic values. The defining table of f^* may be expressed as the fuzzy rule set

$$\begin{aligned} f^*\colon & \textit{if } X \textit{ is } A_1 \textit{ then } Y \textit{ is } B_1 \\ & \textit{if } X \textit{ is } A_2 \textit{ then } Y \textit{ is } B_2 \\ & \dots \\ & \textit{if } X \textit{ is } A_n \textit{ then } Y \textit{ is } B_n. \end{aligned} \tag{2}$$

It is important to note that in this context a fuzzy if–then rule of the form "*if X is A then Y is B*" is not a logical implication but a reading of the ordered pair (A, B). This point is discussed more fully in [28, 29].

As postulated in [28–30], the meaning of the defining table (1) and, equivalently, the fuzzy rule set (2), is the fuzzy graph (Fig. 22)

$$\begin{aligned} f^* &= A_1 \times B_1 + \cdots + A_n \times B_n \\ &= \sum_1 A_i \times B_i, \quad i = 1, \dots, n \end{aligned}$$

where $+$ represents disjunction. A point of key importance is that the fuzzy graph f^* may be

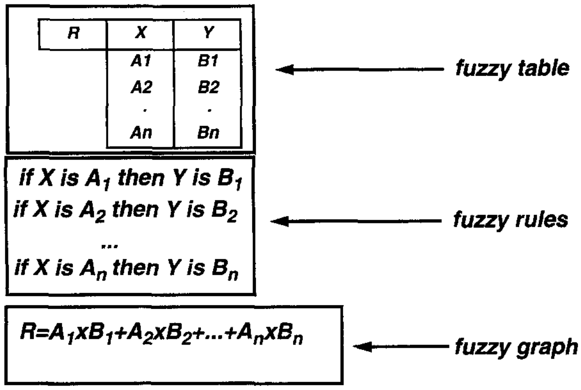


Fig. 22. Representation of a fuzzy function (or relation) as a fuzzy table, fuzzy rule set and a fuzzy graph.

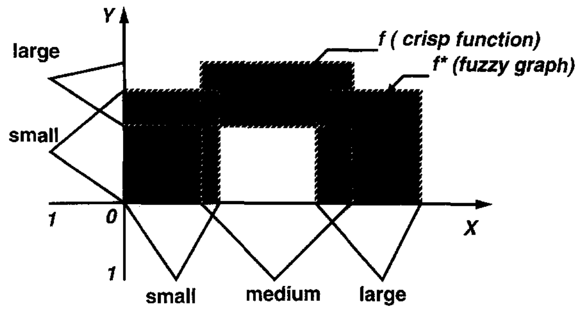


Fig. 23. A fuzzy graph approximation to a function.

viewed as a f-granular approximation of f. For example, in the case of the function shown in Fig. 23, the fuzzy-graph approximation may be expressed as

$$\begin{aligned} f^* &= \textit{small} \times \textit{small} + \textit{medium} \\ &\quad \times \textit{large} + \textit{large} \times \textit{small}. \end{aligned}$$

In this and other cases, the coarseness of granulation is determined by the desired degree of approximation.

There are four basic rationales for f.g-granulation of functions and relations.

1. *Crisp, fine-grained information is not available.* Examples: economic systems, everyday decision-making.
2. *Precise information is costly.* Examples: diagnostic systems, quality control, decision analysis.

3. *Fine-grained information is not necessary.* Examples: Parking a car, cooking, balancing.
4. *Coarse-grained information reduces cost.* Examples: Throw-away cameras, consumer products.

Underlying these rationales is the basic guiding principle of fuzzy logic:

Exploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness, low solution cost and better rapport with reality.

In the context of this principle, the importance of f-granulation derives principally from the fact that it paves the way for a far more extensive use of the machinery of fuzzy information granulation than is the norm at this juncture in both theory and applications.

A case in point relates to the use of crisply defined probability distributions in decision analysis. More specifically, although probability theory is precise and rigorous, its rapport with the real world is far from perfect, largely because most real-world probabilities are poorly defined or hard to estimate, For example, I may need to know the probability that my car may be stolen to decide on whether or not to insure it and for what amount. But probability theory provides no ways for estimating the probability in question. What it does offer is a way of elicitation of subjective probabilities but begs the question of how an estimate of subjective probability can be formed.

In this and similar cases what may work is f-granulation of probability distributions. More specifically, assume for simplicity that X is a discrete random variable taking values $a_1, \ldots, a_n$ with respective probabilities $p_1, \ldots, p_n$. Such distributions will be referred to as *singular* and the probabilistic constraint on X may be expressed as $X\ isp\,(p_1 \backslash a_1 + \cdots + p_n \backslash a_n)$. A probability distribution is *semi-granular* (singular\granular) if it is of the form $X\ isp\,(p_1 \backslash A_1 + \cdots + p_n \backslash A_n)$ where $A_1, \ldots, A_n$ are fuzzy granules. Semi-granular probability distributions of this type define a random set. Furthermore, they play an important role in the Dempster–Shafer theory of evidence.

A probability distribution is *semi-granular* (granular\singular) if it is of the form $X\ isp\,(P_1 \backslash a_1 + \cdots + P_n \backslash a_n)$ where $P_1, \ldots, P_n$ are granular probabilities.

A probability distribution is *granular* if it is of the form

$$X\ isp\,(P_1 \backslash A_1 + \cdots + P_n \backslash A_n) \tag{3}$$

signifying that X is a granular random variable, taking granular (linguistic) values $A_1, \ldots, A_n$ with granular (linguistic) probabilities $P_1, \ldots, P_n$. The granules $A_1, \ldots, A_n$ may be possibilistic or veristic. Granular probability distributions of the form (3) were discussed in [31] in the context of the Dempster–Shafer theory of evidence.

A simple example of a granular probability distribution is shown in Fig. 24. In this example,

$$X\ isp\,(P_1 \backslash A_1 + P_2 \backslash A_2 + P_3 \backslash A_3), \tag{4}$$

or, more specifically,

$$X\ isp\ (small \backslash small + large \backslash medium + small \backslash large).$$

An important concept in the context of granular probability distributions is that of *p-dominance*. More specifically, if in (4) there is a value, A_j, whose probability dominates that of all other values of X then A_j is said to be *p-dominant* or, equivalently, the *usual value* of X (Fig. 24). The importance of p-dominance derives from the fact that in everyday reasoning and discourse it is common practice to

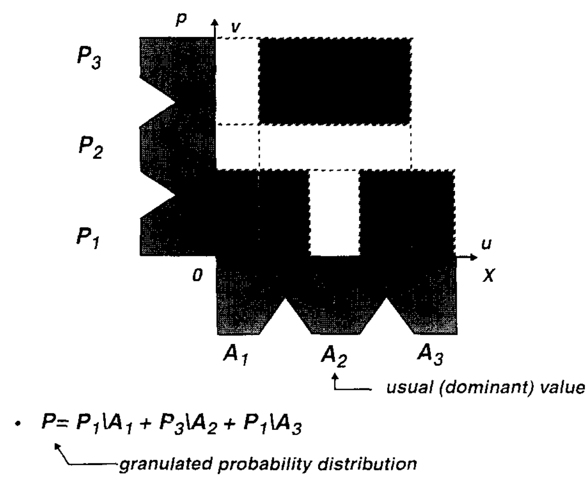


Fig. 24. A granulated (granular) probability distribution.

approximate to

$X\ isp\ (P_1\backslash A_1 + \cdots + P_n\backslash A_n)$

by

$X\ is\ A_j$

if A_j is a p-dominant value of X. For example, in the case of (4), one may assert that

$X\ is\ medium$ (5)

with the understanding that (5) is not a categorical statement but an approximation to

usually (X *is medium*)

where the fuzzy quantifier *usually* may be interpreted as a fuzzy number which represents the probability of the fuzzy event $\{X\ is\ medium\}$.

6. Fuzzy granulation in a general setting

As was alluded to already, the methodology of f-granulation of variables, functions and relations has played and is continuing to play a major role in the applications of fuzzy logic. Within the theory of fuzzy information granulation, the methodology of f-granulation is developed in a much more general setting, enhancing the applicability of f-granulation and widening its impact. This is especially true of f-granulation of functions, since the concept of a function is ubiquitous in all fields of science and engineering.

As a simple illustration of this point consider the standard problem of maximization of an objective function in decision analysis. Let us assume, as is frequently the case in real-world problems, that the objective function, f, is not well-defined and that what we know about f can be expressed as a fuzzy rule set

f^*: *if* X *is* A_1 *then* Y *is* B_1

if X *is* A_2 *then* Y *is* B_2

...

if X *is* A_n *then* Y *is* B_n

or, equivalently, as a fuzzy graph

$$f = \sum_i A_i \times B_i.$$

The question is: What is the point or, more generally, the maximizing set at which f is maximized, and what is the maximum value of f? (Fig. 25)

The problem can be solved by employing the technique of α-cuts. With reference to Fig. 26, if A_{i_α} and B_{i_α} are *α-cuts* of A_i and B_i, respectively, then the corresponding α-cut of f is given by $f_\alpha = \sum_i A_{i_\alpha} \times B_{i_\alpha}$. From this expression, the maximizing fuzzy set, the maximum fuzzy set and maximum value fuzzy set can readily be derived, as shown in Fig. 27.

In a similar vein, one can ask "What is the integral of f; What are the roots of f; etc.?" Problems of this type fall within the province of computing with words [37].

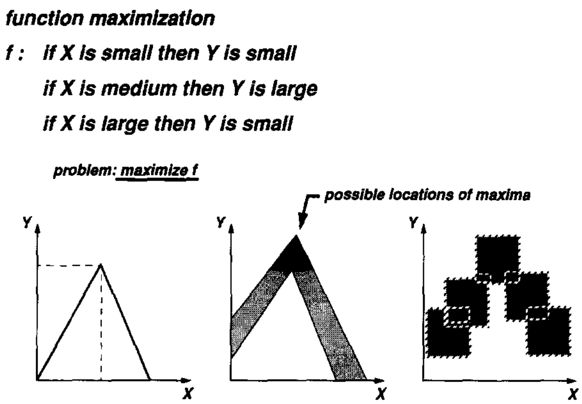


Fig. 25. Maximization of a function, f, defined by a fuzzy rule set or a fuzzy graph.

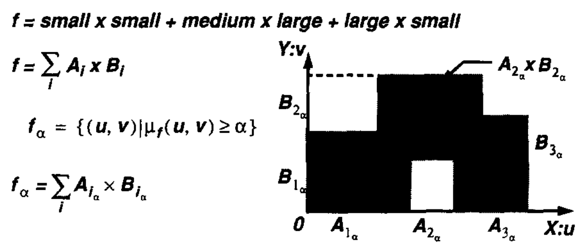


Fig. 26. α-cuts of the fuzzy graph of f.

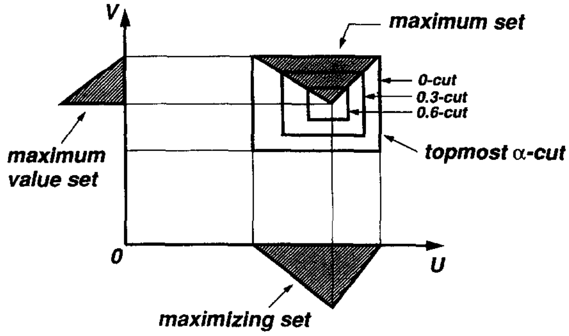


Fig. 27. The maximizing set, the maximum value set and the maximum set of a fuzzy graph.

Another illustration is provided by the extension principle [24, 36, 37], which is a basic rule of inference in fuzzy logic and is expressible as the inference schema

$$\frac{X \text{ is } A}{f(x) \text{ is } f(A)}$$

where $f: U \to V$ and

$$\mu_{f(A)}(v) = \sup_{u \mid v = f(u)} \mu_A(u).$$

Let us apply f-granulation to f, yielding the rule set

f: *if* X *is* A_i *then* Y *is* B_i, $\quad i = 1, \ldots, n$.

In this case, the problem reduces to the familiar interpolation schema in the calculus of fuzzy rules:

X *is* A

if X *is* A_1 *then* Y *is* B_1, $\quad i = 1, \ldots, n$

Y *is* $\sum_i m_i \Lambda B_i$

where the matching coefficient m_i is given by

$m_i = \sup(A \cap A_i)$.

The examples discussed above suggest an important direction in the development of TFIG. Specifically, the examples in question may be viewed as f.g-generalizations of standard problems and techniques. Thus, in the first example the standard problem is that of maximization, while in the second problem f.g-generalization is applied to the extension principle.

6.1. The airport shuttle problem

Another example in this spirit is what might be called the Airport Shuttle problem, a problem which may be viewed as an f.g-generalization of the standard Traveling Salesman problem. In this case, an airport shuttle picks up passengers at an airport and takes them to specified addresses. The objective of the driver is to return to the airport as soon as possible (Fig. 28).

The difference between this problem and the Traveling Salesman problem is that in the case of the Traveling Salesman problem the cost of going from node i to node j is known for all i, j, whereas in the Airport Shuttle problem the transit time from address i to address j has to be estimated by the driver. The driver does so by interpolating the data stored in the driver's memory, performing interpolation in both time and space (Fig. 29). In an intuitive way, the driver approximates to the transit

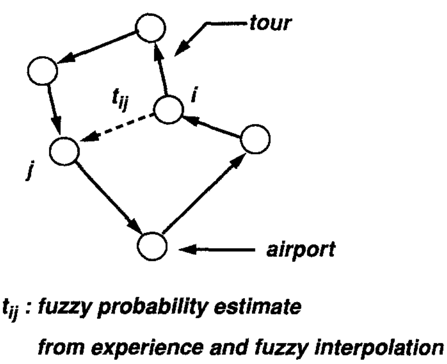


Fig. 28. The airport shuttle problem.

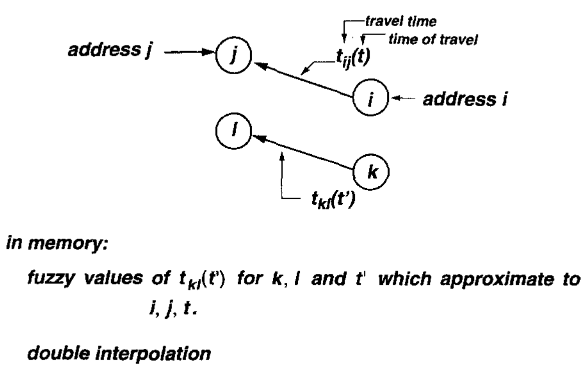


Fig. 29. Interpolation in time and space in the airport shuttle problem.

time by a coarse granular probability distribution. In arriving at a decision on the order in which the passengers should be taken to their destinations, the driver uses an intuitive form of p-dominance. This, of course, is merely a coarse perception of what goes on in the driver's mind.

In the problem under consideration, fuzzy information granulation in an intuitive form underlies the human solution. What this suggests is that no machine could solve the problem without using, as human do, the machinery of fuzzy information granulation. How this could be done in detail is a challenge that has not as yet been met.

6.2. The commute time problem

Another problem of this type, a problem which makes the same point, is what might be called the Commute Time problem.

The problem may be formulated in two versions: (a) unannotated; and (b) annotated.

In the unannotated version we are given a time series such as

$$T_a\colon \{15, 18, 21, 14, 20, \theta, \theta, 13, 0, 3, 18, 17, \theta, 19, \dots\}$$

with no knowledge of what the numbers represent or how they were obtained. The questions posed are the following

1. Does T_a represent the result of a random experiment?
2. If it does, what is the sample space? What are the random variables? Is T_a stationary?
3. Given the elements of T_a up to and including $t = i$, what would be an estimate of T_a at time $i + 1$?

The unannotated version has neither a human nor a machine solution. In particular, standard probability theory provides no answers to the posed questions. Nevertheless, there are programs which, given an unannotated time series, will come up with a prediction. It can be argued that such predictions have no justification.

In the annotated version, the time-series reads:

$$T_b\colon \{(Mon, 15), (Tue, 18), (Wed, 21), (Thu, 14), (Fri, 20), (Sat, \theta), (Sun, \theta), (Mon, 13), \dots, \}$$

and has the following meaning.

T_b represents a record of the time it took me to commute from my home to the compus, starting with Monday, 1 January, 1996; θ means that I did not go to the campus that day; it took longer on Wednesday, 3 January, because of rain; usually it takes longer on Fridays, etc.

Suppose that in the morning of Wednesday, 20 March, I had to estimate the commute time that day, knowing that it would be slightly shorter than 18 min on Wednesday, 13 March, because of the Spring recess which started on 18 March. Everything considered, my estimate might be: around 18 min.

The point of this example is that the problem has a human solution arrived at through human reasoning based on f-granulated information. Neither standard probability theory nor any methodology which does not employ the machinery of fuzzy information granulation can come up with a machine solution. The challenge, then, is to develop a theory of fuzzy information granulation which can model the ways in which human granulate information and reason with it. In a preliminary way, this is what we have attempted to do in this paper.

7. Concluding remark

The machinery of fuzzy information granulation, especially in the form of linguistic variables, fuzzy if–then rules and fuzzy graphs, has long played a major role in the applications of fuzzy logic. What has not been fully recognized, however, is the centrality of fuzzy information granulation in human reasoning and, ipso facto, its centrality in fuzzy logic. A related point is that no methodology other than fuzzy logic provides a conceptual framework and associated techniques for dealing with problems in which fuzzy information granulation plays, or could play, a major role. In the context of such problems, the way in which humans employ fuzzy information granulation to make rational decisions in an environment of partial knowledge, partial certainty and partial truth should be viewed as a role model for machine intelligence.

The theory of fuzzy information granulation outlined in this paper takes the existing machinery of fuzzy information granulation in fuzzy logic to a higher level of generality, consolidates its foundations and suggests new directions. In coming years, TFIG is likely to play an important role in the evolution of fuzzy logic and, in conjunction with computing with words, may eventually have a far-reaching impact on its applications.

References

[1] D. Driankov, H. Hellendoorn, M. Reinfrank, An Introduction to Fuzzy Control, Springer, Berlin, 1993.

[2] D. Dubois, H. Fargier, H. Prade, Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty, Applied Intelligence 6 (1996) 287–309.

[3] D. Dubois, J. Lang, H. Prade, Possibilistic Logic, in: Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 3, D.M. Gabbay, C.J. Hogger, J.A. Robinson, D. Nute (Eds.), Oxford Univ. Press, Oxford, 1994, 439–513.

[4] D. Dubois, H. Prade, Fuzzy sets in approximate reasoning, Part I: Interference with possibility distributions; Part II: Logical approaches (with J. Lang), Fuzzy Sets and Systems Vol. 40, Part I: pp. 143–202; Part II: pp. 203–244 (1991).

[5] D. Dubois, H. Prade, Putting rough sets and fuzzy sets together, in: Intelligent Decision Support – Handbook of Applications and Advances of the Rough Sets Theory, R. Slowinski, Ed., Kluwer, Dordrecht, 1992, 203–232.

[6] D. Dubois, H. Prade, R. Yager (Eds.), Readings in Fuzzy Sets for Intelligent Systems, Morgan Kaufmann, San Mateo, 1993.

[7] D. Dubois, H. Prade, R. Yager (Eds.), Fuzzy Information Engineering, Wiley, New York, 1997.

[8] J.A. Goguen, The logic of inexact concepts, Synthese 19 (1969) 325–373.

[9] M. Jamshidi, N. Vadiee, T. Ross (Eds.), Fuzzy Logic and Control, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[10] G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic, Prentice-Hall, Englewood Cliffs, 1995.

[11] B. Kosko, Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence, Prentice-Hall, Englewood Cliffs, NJ, 1991.

[12] R. Kruse, J. Gebhardt, F. Klawonn, Foundations of Fuzzy Systems, Wiley, New York, 1994.

[13] C.C. Lee, Fuzzy logic in control systems: Fuzzy logic controller, parts I and II, IEEE Trans. Systems, Man Cybernet. 20 (1990), 404–418.

[14] E. H. Mamdani, B.R. Gaines (Eds.), Fuzzy Reasoning and its Applications, Academic Press, London, 1981.

[15] M. Mares, Computation Over Fuzzy Quantities, CRC Press, Boca Raton, 1994.

[16] V. Novak, Fuzzy logic, fuzzy sets, and natural languages, Internat. J. General Systems 20 (1) (1991) 83–97.

[17] V. Novak, M. Ramik, M. Cerny, J. Nekola (Eds.), Fuzzy Approach to Reasoning and Decision-Making, Kluwer, Boston, 1992.

[18] Z. Pawlak, Rough sets, Internat. J. Comput. Inform. Sci. 11 (1982) 341–356.

[19] W. Pedrycz, Fuzzy Control and Fuzzy Systems, Wiley, New York, 1989.

[20] T. Terano, K. Asai, M. Sugeno, Fuzzy Systems Theory and its Applications, Academic Press, New York, 1992.

[21] C. von Altrock, Fuzzy Logic & Neurofuzzy Applications Explained, PTR Prentice-Hall, Englewood, NJ, 1995.

[22] L.-X. Wang, Adaptive Fuzzy Systems and Control: Design Stability Analysis, PTR Prentice-Hall, Englewood Cliffs, NJ, 1994.

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

[24] L.A. Zadeh, Fuzzy sets, Inform. and Control 8 (1965) 338–353.

[25] L.A. Zadeh, Shadows of fuzzy sets, Prob. Trans. Inform. 2 (1966) 37–44.

[26] L.A. Zadeh, Toward a theory of fuzzy systems, in: R.E. Kalman, N. DeClaris (Eds.), Aspects of Network and System Theory, Rinehart & Winston, New York, 1971, 469–490.

[27] L.A. Zadeh, Outline of a new approach to the analysis of complex system and decision processes, IEEE Trans. Systems. Man, Cybernet. 3 (1973) 28–44.

[28] L.A. Zadeh, On the Analysis of Large Scale Systems, in: H. Gottinger (Ed.), Systems Approaches and Environment Problems, Vandenhoeck & Ruprecht, Gottingen: 1974, 23–37.

[29] L.A. Zadeh, The concept of a linguistic variable and its applications to approximate reasoning Part I: Inform. Sci. 8 (1975) 199–249; Part II: Inform. Sci. 8 (1975) 301–357; Part III: Inform. Sci. 9 (1975) 43–80.

[30] L.A. Zadeh, A fuzzy-algorithmic approach to the definition of complex or imprecise concepts, Internat. J. Man–Machine Stud. 8 (1976) 249–291.

[31] L.A. Zadeh, Fuzzy Sets and Information Granularity, in: M. Gupta, R. Ragade, R. Yager, (Eds.), Advances in Fuzzy Set Theory and Applications, North-Holland, Amsterdam, 1979, 3–18.

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

[33] L.A. Zadeh, Outline of a theory of usuality based on fuzzy logic, Reidel, Dordrecht, Fuzzy Sets Theory and Applications, in: A. Jones, A. Kaufmann, H.J. Zimmerman (Eds.), 1986, 79–97.

[34] L.A. Zadeh, Fuzzy logic, neural networks and soft computing, Commun. ACM 37 (3) (1994) 77–84.

[35] L.A. Zadeh, Why the success of fuzzy logic is not paradoxical, IEEE Expert 9 (4) (1994) 43–45.

[36] L.A. Zadeh, Fuzzy logic and the calculi of fuzzy rules and fuzzy graphs, Multiple Valued Logic 1 (1996) 1–38.

[37] L.A. Zadeh, Fuzzy logic = computing with words, IEEE Trans. on Fuzzy Systems 4 (1996) 103–111.

[38] H.J. Zimmerman, Fuzzy Set Theory and Its Applications, 3rd ed., Kluwer-Nijhoff, Amsterdam, 1996.