Granular Computing and Rough Set Theory

Lotfi A. Zadeh

Computer Science Division Department of EECS UC Berkeley

RSEISP'07 Warsaw, Poland June 28, 2007 Dedicated to the memory of Prof. Z. Pawlak

URL: http://www-bisc.cs.berkeley.edu URL: http://www.cs.berkeley.edu/~zadeh/ Email: Zadeh@eecs.berkeley.edu

PREAMBLE





GRANULATION

 granulation: partitioning (crisp or fuzzy) of an object into a collection of granules, with a granule being a clump of elements drawn together by indistinguishability, equivalence, similarity, proximity or functionality.

RST GRC example: f-granulation bead+neck+chest+ans+…+feet.

c-granulation

LAZ 6/25/2007

Set

GRANULATION OF A VARIABLE (Granular Variable)

continuous → quantized → granulated

Example: Age



GRANULATION OF A FUNCTION GRANULATION=SUMMARIZATION



0

6 /73



*f (fuzzy graph)

f perception summarization *f :

if X is small then Y is small if X is medium then Y is large if X is large then Y is small

GRANULATION OF A PROBABILITY DISTRIBUTION



P(X) = low\small + high\medium + low\large LAZ 6/25/2007



probability distributions 8 /73

PRINCIPAL TYPES OF GRANULES

• Possibilistic

• X is a number in the interval [a, b]

• Probabilistic

- X is a normally distributed random variable with mean a and variance b
- Veristic
 - X is all numbers in the interval [a, b]
- Hybrid
 - X is a random set

SINGULAR AND GRANULAR VALUES

- X is a variable taking values in U
- a, aεU, is a singular value of X if a is a singleton
- A is a granular value of X if A is a granule, that is, A is a clump of values of X drawn together by indistinguishability, equivalence, similarity, proximity or functionality.
- A may be interpreted as a representation of information about a singular value of X.
- A granular variable is a variable which takes granular values
- A linguistic variable is a granular variable with linguistic labels of granular values.

SINGULAR AND GRANULAR VALUES



granular value of X

singular value of X

- universe of discourse

	singular	granular
unemploymen t	→ 7.3%	high
temperature [*]	102.5	very high
blood pressure	160/80	high

LAZ 6/25/2007

11 /73

ATTRIBUTES OF A GRANULE

- Probability measure
- Possibility measure
- Verity measure





12/73





if there is a tolerance for imprecision, exploited through granulation of X

CLARIFICATION—THE MEANING OF PRECISION



p: X is a Gaussian random variable with mean m and variance σ². m and σ² are precisely defined real numbers
 p is v-imprecise and m-precise

p: X is in the interval [a, b]. a and b are precisely defined real numbers
p is v-imprecise and m precise

14 /73

granulation = v-imprecisiation

MODALITIES OF m-PRECISIATION



mm-precise: mathematically well-defined



CLARIFICATION

• Rationale 2: if there is a tolerance for imprecision, exploited through granulation of X

 Rationale 2: if there is a tolerance for v-imprecision, exploited through granulation of X followed by mmprecisiation of granular values of X

Example: Lily is $25 \rightarrow$ Lily is young





RATIONALES FOR FUZZY LOGIC

RATIONALE 2—Fuzzy Logic Gambit



Fuzzy Logic Gambit: if there is a tolerance for imprecisiation, exploited by v-imprecisiation followed by mm-precisiation

Rationale 2 plays a key role in fuzzy control

CHARACTERIZATION OF A GRANULE

- granular value of X = information, I(X), about the singular value of X
- I(X) is represented through the use of an information description language, IDL. IDL



- BL: SCL (standard constraint language)
- FL: GCL (generalized constraint language)
- NL: PNL (precisiated natural language)

19/73

information = generalized constraint

EXAMPLE—PROBABILISTIC GRANULE

- Implicit characterization of a probabilistic granule via natural language
- X is a real-valued random variable
- Probability distribution of X is not known precisely. What is known about the probability distribution of X is: (a) usually X is much larger than approximately a; usually X is much smaller than approximately b.

In this case, information about X is mm-precise and implicit.

THE CONCEPT

OF A GENERALIZED

CONSTRAINT



PREAMBLE

 In scientific theories, representation of constraints is generally oversimplified. **Oversimplification of constraints is a** necessity because existing constrained definition languages have a very limited expressive power. The concept of a generalized constraint is intended to provide a basis for construction of a maximally expressive constraint definition language which can also serve as a meaning representation/precisiation language for naturaNanguages.

GENERALIZED CONSTRAINT (Zadeh 1986)

- Bivalent constraint (hard, inelastic, categorical:)
 - X ε C constraining bivalent relation
- Generalized constraint on X: GC(X)



CONTINUED

- constrained variable
 - X is an n-ary variable, X= (X₁, ..., X_n)
 - X is a proposition, e.g., Leslie is tall
 - 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 generalized constraint, X: Y isr R
 X is a group variable. In this case, there is a group, G: (Name₁, ..., Name_n), with each member of the group, Name_i, i =1, ..., n, associated with an attribute-value, h_i, of attribute H. h_i may be vector-valued. Symbolically

CONTINUED

G = (Name₁, ..., Name_n)

 $G[H] = (Name_1/h_1, ..., Name_n/h_n)$

 $G[H is A] = (\mu_A(h_i)/Name_1, \dots, \mu_A(h_n)/Name_n)$

Basically, G[H] is a relation and G[H is A] is a fuzzy restriction of G[H]

Example:

25 / 73

tall Swedes

Swedes[Height is tall]

GENERALIZED CONSTRAINT—MODALITY r



r: = *r*: ≤ $r:\subset$ r: blank

r: v

r: p

equality constraint: X=R is abbreviation of X is=R inequality constraint: $X \leq R$ subsethood constraint: X C R possibilistic constraint; X is R; R is the possibility distribution of X veristic constraint; X isv R; R is the verity distribution of X probabilistic constraint; X isp R; R is the probability distribution of X Standard constraints: bivalent possibilistic, bivalent veristic and probabilistic LAZ 6/25/2007

CONTINUED

r: bm bimodal constraint; X is a random variable; R is a bimodal distribution

r: rs random set constraint; X isrs R; R is the setvalued probability distribution of X

r: fg fuzzy graph constraint; X isfg R; X is a function and R is its fuzzy graph

r: u usuality constraint; X isu R means usually (X is R)

r: g group constraint; X isg R means that R constrains the attribute-values of the group

PRIMARY GENERALIZED CONSTRAINTS

- Possibilistic: X is R
- Probabilistic: X isp R
- Veristic: X isv R
- Primary constraints are formalizations of three basic perceptions: (a) perception of possibility; (b) perception of likelihood; and
 (c) perception of truth

 In this perspective, probability may be viewed as an attribute of perception of likelihood

STANDARD CONSTRAINTS

- Bivalent possibilistic: X ε C (crisp set)
- Bivalent veristic: Ver(p) is true or false
- Probabilistic: X isp R

Standard constraints are instances of generalized constraints which underlie methods based on bivalent logic and probability theory



EXAMPLES: POSSIBILISTIC

Monika is much younger than Maria —
(Age (Monika), Age (Maria)) is much younger *x x*

30 /73

EXAMPLES: VERISTIC

 Robert is half German, quarter French and quarter Italian
 Ethnicity (Robert) isv (0.5|German + 0.25|French + 0.25|Italian)

 Robert resided in London from 1985 to 1990
 Reside (Robert, London) isv [1985, 1990]

GENERALIZED CONSTRAINT LANGUAGE (GCL)

- GCL is an abstract language
- GCL is generated by combination, qualification, propagation and counterpropagation of generalized constraints
- examples of elements of GCL
 - X/Age(Monika) is R/young (annotated element)
 - (X isp R) and (X,Y) is S)
 - (X isr R) is unlikely) and (X iss S) is likely
 - If X is A then Y is B

the language of fuzzy if-then rules is a sublanguage of GCL

 deduction= generalized constraint propagation and counterpropagation

EXTENSION PRINCIPLE

• The principal rule of deduction in NL-Computation is the Extension Principle (Zadeh 1965, 1975).



 $\mu_{B}(v) = sup_{\mu}\mu_{A}(f(u))$

 $\frac{subject}{v} = g(u)$

33 /73



p: most Swedes are tall p*: ΣCount(tall.Swedes/Swedes) is most

further precisiation

X(h): height density function (not known) X(h)du: fraction of Swedes whose height is in [h, h+du], $a \le h \le b$

$$\int_{a}^{b} X(h) du = 1$$

34 /73

PRECISIATION AND CALIBRATION

- µ_{tall}(h): membership function of tall (known)
- µ_{most}(u): membership function of most (known)



CONTINUED

36 /73

$$\pi(h) = \mu_{most}(\int_a^b X(h)\mu_{tall}(h)dh)$$
DEDUCTION

q: What is the average height of Swedes?

 q^* : $\int_a^b X(h)hdh$ is ? Q

deduction:

∫<mark>a X(h)</mark>μ_{tall}(h)dh is most

 $\int_{a}^{b} X(h) h dh$ is ? Q



THE CONCEPT OF PROTOFORM

Protoform= abbreviation of prototypical form

$$\rightarrow$$
 summarization \rightarrow abstraction \rightarrow generalization \rightarrow Pro(p)

p: object (proposition(s), predicate(s), question
(s), command, scenario, decision problem, ...)

Pro(p): protoform of p

Basically, Pro(p) is a representation of the deep structure of p



• p: most Swedes are tall

$$p \rightarrow abstraction \rightarrow QA's are B's$$

QA's are B's \rightarrow generalization \rightarrow Count(G[H is R]/G) is Q



EXAMPLES

Monika is much younger than Robert \longrightarrow (Age(Monika), Age(Robert) is much.younger \longrightarrow D(A(B), A(C)) is E

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?



PROTOFORM EQUIVALENCE



at a given level of abstraction and summarization, objects p and q are PF-equivalent if PF(p)=PF(q)

example p: Most Swedes are tall q: Few professors are rich

Count (A/B) is Q Count (A/B) is Q

PROTOFORM EQUIVALENCE DECISION PROBLEM

 Pro(backpain)= Pro(surge in Iraq) = Pro (divorce) = Pro(new job)= Pro(new location)

Status quo may be optimal



DEDUCTION

• In NL-computation, deduction rules are protoformal Example: $1/n\Sigma$ Count(G[H is R]) is Q

1/nΣCount(G[H is S]) is T

 $\Sigma_i \mu_R(h_i)$ is Q $\Sigma_i \mu_S(h_i)$ is T

 $\mu_{T}(v) = \sup_{h_{1}, \dots, h_{n}}(\mu_{Q}(\Sigma_{i}\mu_{R}(h_{i})))$ subject to $v = \Sigma_{i}\mu_{S}(h_{i})$

values of H: h₁, ..., h_n

43 /73

PROBABILISTIC DEDUCTION RULE

$$\mu_{Q}(\mathbf{v}) = \mathbf{sup}_{g}(\mu_{P_{1}}(\int_{U}\mu_{A_{1}}(\mathbf{u})\mathbf{g}(\mathbf{u})\mathbf{du}) \wedge \cdots \wedge$$

$$\mu_{P_n} \int_U \mu_{P_n} \left(\int_U \mu_{A_n} (u) g(u) du \right)$$

subject to

$$U = \int_{U} \mu_{A}(u)g(u)du$$

PROTOFORMAL DEDUCTION RULE

• Syllogism

 $Q_1 A's$ are B's $Q_2 (A\&B)'s$ are C's $Q_1 Q_2 A's$ are (B&C)'s

Example

- Overeating causes obesity *precisiation most of those who overeat become obese*
- Overeating and obesity cause high blood pressure precisiation, most of those who overeat and are obese have high blood pressure
- I overeat and am obese. The probability that I will develop high blood pressure is most²

GRANULAR COMPUTING VS. NL-COMPUTATION

- In conventional modes of computation, the objects of computation are values of variables.
- In granular computing, the objects of computation are granular values of variables.
- In NL-Computation, the objects of computation are explicit or implicit descriptions of values of variables, with descriptions expressed in a natural language.
- NL-Computation is closely related to Computing with Words and the concept of Precisiated Natural Language (PNL).

PRECISIATED NATURAL LANGUAGE (PNL)

• PNL may be viewed as an algorithmic dictionary with three columns and rules of deduction

p	Pre(p)	Pro(p)
Lily is young	Age (Lily is young)	A(B) is C
	•••	

NL-Computation = PNL

47 /73

Principal Concepts And Ideas

LAZ 6/25/2007

48 /73

BASIC IDEA

?Z=f(X, Y)

Conventional computation
given: value of X
given: value of Y
given: f
compute: value of Z

*Z= *f(*X, *Y)

• NL-Computation

50 /73

given: NL(X) (information about the value of X described in natural language) *X
given: NL(Y) (information about the values of Y described in natural language) *Y
given: NL(X, Y) (information about the values of X and Y described in natural language) *(X, Y)
given: NL (f) (information about f described in natural language) *f

 computation: NL(Z) (information about the value of Z described in natural language) *Z LAZ 6/25/2007

EXAMPLE (AGE DIFFERENCE)

- Z= Age(Vera) Age(Pat)
- Age(Vera): Vera has a son in late twenties and a daughter in late thirties
- Age(Pat): Pat has a daughter who is close to thirty. Pat is a dermatologist. In practice for close to 20 years
- NL(W1): (relevant information drawn from world knowledge) child bearing age ranges from about 16 to about 42
- NL(W2): age at start of practice ranges from about 20 to about 40
- Closed (circumscribed) vs. open (uncircumscribed)
- Open: augmentation of information by drawing on world knowledge is allowed

51 /73

EXAMPLE (NL(f))

Y=f(X)

NL(f): if X is small then Y is small if X is medium then Y is large if X is large then Y is small NL(X): usually X is medium ?NL(Y)



EXAMPLE (balls-in-box)

 a box contains about 20 black and white balls. Most are black. There are several times as many black balls as white balls. What is the number of white balls?

EXAMPLE (chaining)

Overeating causes obesity

- Overeating and obesity cause high blood pressure
- I overeat. What is the probability that I will develop high blood pressure?

KEY OBSERVATIONS--PERCEPTIONS

- A natural language is basically 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 perceptions is passed on to the natural languages which is used to describe them
- Natural languages are intrinsically imprecise



•it is 35 C°

It is very warm

 Over 70% of Swedes are taller than 175
 most Swedes are tall **c***m* probability is high

•probability is 0.8

•it is cloudy

traffic is heavy

•it is hard to find parking near the campus

 measurement-based information may be viewed as a special case of perception-based information

 perception-based information is intrinsically imprecise 55 /73 LAZ 6/25/2007

NL-capability

 In the computational theory of perceptions (Zadeh 1999) the objects of computation are not perceptions per se but their descriptions in a natural language

 Computational theory of perceptions (CTP) is based on NL-Computation

 Capability to compute with perception-based information= capability to compute with information described in a natural language= NL-capability.

56 /73

KEY OBSERVATION—NL-incapability

- Existing scientific theories are based for the most part on bivalent logic and bivalentlogic-based probability theory
- Bivalent logic and bivalent-logic-based
 probability theory do not have NL-capability

 For the most part, existing scientific theories do not have NL-capability



DIGRESSION—HISTORICAL NOTE

• The point of departure in NL-Computation is my 1973 paper, "Outline of a new approach to the analysis of complex systems and decision processes," published in the IEEE Transactions on Systems, Man and Cybernetics. In retrospect, the ideas introduced in this paper may be viewed as a first step toward the development of NL-**Computation**.

 In the 1973 paper, two key ideas were introduced: (a) the concept of a linguistic variable; and (b) the concept of a fuzzy-if-then rule. These concepts play pivotal roles in dealing with complexity.

In brief



LINGUISTIC VARIABLE

 A linguistic variable is a variable whose values are fuzzy sets carrying linguistic labels
 example:

Age: young + middle-aged + old



hedging

- Age: young + very young + not very young + quite young + ...
- Honesty: honest + very honest + quite honest + ...

FUZZY IF-THEN RULES

Example: if X is small and Y is medium then Z is large

Rule set: if X is A1 and Y is B1 then Z is C1

if X is An and Y is Bn then Z is Cn

A rule set is a granular description of a function





Control Rules:

If (speed is low) and (shift is high) then (-3)
 If (speed is high) and (shift is low) then (+3)
 If (throt is low) and (speed is high) then (+3)
 If (throt is low) and (speed is low) then (+1)
 If (throt is high) and (speed is high) then (-1)
 If (throt is high) and (speed is low) then (-3)

FUZZY LOGIC TODAY

• Today linguistic variables and fuzzy if-then rules are employed in almost all applications of fuzzy logic, ranging from digital photography, consumer electronics, industrial control to biomedical instrumentation, decision analysis and patent classification. A metric over the use of fuzzy logic is the number of papers with fuzzy in title.

<u>INSPEC</u> 1970-1979: 569 1980-1989: 2,403 1990-1999: 23,210 2000-present: 21,919 Total: 51,096 <u>MathSciNet</u> 1970-1979: 443 1980-1989: 2,465 1990-1999: 5,487 2000-present: 5,714 Total: 14,612

INITIAL REACTIONS

• When the idea of a linguistic variable occurred to me in 1972, I recognized at once that it was the beginning of a new direction in systems analysis. But the initial reaction to my ideas was, for the most part, hostile. Here are a few examples. There are many more.

R.E. Kalman (1972)

I would like to comment briefly on **Professor Zadeh's presentation. His** proposals could be severely, ferociously, even brutally critisized from a technical point of view. This would be out of place here. But a blunt question remains: Is Professor Zadeh presenting important ideas or is he indulging in wishful thinking?

No doubt Professor Zadeh's enthusiasm for fuzziness has been reinforced by the prevailing climate in the U.S.—one of unprecedented permissiveness. "Fuzzification" is a kind of scientific pervasiveness; it tends to result in socially appealing slogans unaccompanied by the discipline of hard scientific work and patient observation.

Professor William Kahan (1975)

"Fuzzy theory is wrong, wrong, and pernicious." says William Kahan, a professor of computer sciences and mathematics at Cal whose Evans Hall office is a few doors from Zadeh's. "I can not think of any problem that could not be solved better by ordinary logic."

"What Zadeh is saying is the same sort of things 'Technology got us into this mess and now it can't get us out.'" Kahan says. "Well, technology did not get us into this mess. Greed and weakness and ambivalence got us into this mess. What we need is more logical thinking, not less. The danger of fuzzy theory is that it will encourage the sort of imprecise thinking that has brought us so much trouble."

 What my critics did not understand was that the concept of a linguistic variable was a gambit—the fuzzy logic gambit. Use of linguistic variables entails a sacrifice of precision. But what is gained is reduction in cost since precision is costly. The same rationale underlies the effectiveness of granular computing, rough-set-based techniques and NL-Computation.

SUMMATION

 In real world settings, the values of variables are rarely known with perfect certainty and precision. A realistic assumption is that the value is granular, with a granule representing the state of knowledge about the value of the variable. A key idea in Granular **Computing is that of defining a granule** as a generalized constraint. In this way, computation with granular values reduces to propagation and counterpropagation of generalized constraints.

RELATED PAPERS BY L.A. ZADEH (IN REVERSE CHRONOLOGICAL ORDER)

- Generalized theory of uncertainty (GTU)—principal concepts and ideas, to appear in Computational Statistics and Data Analysis.
- Precisiated natural language (PNL), Al Magazine, Vol. 25, No. 3, 74-91, 2004.
- Toward a perception-based theory of probabilistic reasoning with imprecise probabilities, Journal of Statistical Planning and Inference, Elsevier Science, Vol. 105, 233-264, 2002.
- A new direction in Al—toward a computational theory of perceptions, Al Magazine, Vol. 22, No. 1, 73-84, 2001.

- From computing with numbers to computing with words --from manipulation of measurements to manipulation of perceptions, IEEE Transactions on Circuits and Systems 45, 105-119, 1999.
- Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems, Soft Computing 2, 23-25, 1998.

 Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90, 111-127, 1997.
CONTINUED

- Outline of a computational approach to meaning and knowledge representation based on the concept of a generalized assignment statement, Proceedings of the International Seminar on Artificial Intelligence and Man-Machine Systems, M. Thoma and A. Wyner (eds.), 198-211. Heidelberg: Springer-Verlag, 1986.
- Precisiation of meaning via translation into PRUF, Cognitive Constraints on Communication, L. Vaina and J. Hintikka, (eds.), 373-402. Dordrecht: Reidel, 1984.
- Fuzzy sets and information granularity, Advances in Fuzzy Set Theory and Applications, M. Gupta, R. Ragade and R. Yager (eds.), 3-18. Amsterdam: North-Holland Publishing Co., 1979.