Lotfi A. ZADEH

Professor in the Graduate School
Director, Berkeley Initiative in
Soft Computing (BISC)
Address: 729 Soda Hall #1776
Computer Science Division
Department of Electrical
Engineering and Computer
Sciences
University of California
Berkeley, CA 94720-1776
E-mail: *zadeh@eecs.berkeley.edu*

FROM SEARCH ENGINES TO QUESTION-ANSWERING SYSTEMS – THE PROBLEMS OF WORLD KNOWLEDGE, RELEVANCE, DEDUCTION AND PRECISIATION

Note: *The paper was presented as Invited Conference to the IEEE International Workshop on Soft Computing Applications (IEEE SOFA 2005) 27-30 August, 2005, Szeged –Hungary and Arad – Romania*

CURRICULUM VITAE

Lotfi A. Zadeh

Lotfi A. Zadeh joined the Department of Electrical Engineering at the University of California, Berkeley, in 1959, and served as its chairman from 1963 to 1968. Earlier, he was a member of the electrical engineering faculty at Columbia University. In 1956, he was a visiting member of the Institute for Advanced Study in Princeton, New Jersey. In addition, he held a number of other visiting appointments, among them a visiting professorship in Electrical Engineering at MIT in 1962 and 1968; a visiting scientist appointment at IBM Research Laboratory, San Jose, CA, in 1968, 1973, and 1977; and visiting scholar appointments at the AI Center, SRI International, in 1981, and at the Center for the Study of Language and Information, Stanford University, in 1987-1988. Currently he is a Professor in the Graduate School, and is serving as the Director of BISC (Berkeley Initiative in Soft Computing).

Until 1965, Dr. Zadeh's work had been centered on system theory and decision analysis. Since then, his research interests have shifted to the theory of fuzzy sets and its applications to artificial intelligence, linguistics, logic, decision analysis, control theory, expert systems and neural networks. Currently, his research is focused on fuzzy logic, soft computing, computing with words, and the newly developed computational theory of perceptions and precisiated natural language.

An alumnus of the University of Tehran, MIT, and Columbia University, Dr. Zadeh is a fellow of the IEEE, AAAS, ACM, AAAI and IFSA, and a member of the National Academy of Engineering. He held NSF Senior Postdoctoral Fellowships in 1956-57 and 1962-63, and was a Guggenheim Foundation Fellow in 1968. Dr. Zadeh was the recipient of the IEEE Education Medal in 1973 and a recipient of the IEEE Centennial Medal in 1984. In 1989, Dr. Zadeh was awarded the Honda Prize by the Honda Foundation, and in 1991 received the Berkeley Citation, University of California.

In 1992, Dr. Zadeh was awarded the IEEE Richard W. Hamming Medal "For seminal contributions to information science and systems, including the conceptualization of fuzzy sets." He became a Foreign Member of the Russian Academy of Natural Sciences (Computer Sciences and Cybernetics Section) in 1992, and received the Certificate of Commendation for AI Special Contributions Award from the International Foundation for Artificial Intelligence. Also in 1992, he was awarded the Kampe de Feriet Prize and became an Honorary Member of the Austrian Society of Cybernetic Studies.

In 1993, Dr. Zadeh received the Rufus Oldenburger Medal from the American Society of Mechanical Engineers "For seminal contributions in system theory, decision analysis, and theory of fuzzy sets and its applications to AI, linguistics, logic, expert systems and neural networks." He was also awarded the Grigore Moisil Prize for Fundamental Researches, and the Premier Best Paper Award by the Second International Conference on Fuzzy Theory and Technology. In 1995, Dr. Zadeh was awarded the IEEE Medal of Honor "For pioneering development of fuzzy logic and its many diverse applications." In 1996, Dr. Zadeh was awarded the Okawa Prize "For outstanding contribution to information science through the development of fuzzy logic and its applications."

In 1997, Dr. Zadeh was awarded the B. Bolzano Medal by the Academy of Sciences of the Czech Republic "For outstanding achievements in fuzzy mathematics." He also received the J.P. Wohl Career Achievement Award of the IEEE Systems, Science and Cybernetics Society. He served as a Lee Kuan Yew Distinguished Visitor, lecturing at the National University of Singapore and the Nanyang Technological University in Singapore, and as the Gulbenkian Foundation Visiting Professor at the New University of Lisbon in Portugal. In 1998, Dr. Zadeh was awarded the Edward Feigenbaum Medal by the International Society for Intelligent Systems and the

Richard E. Bellman Control Heritage Award by the American Council on Automatic Control. In addition, he received the Information Science Award from the Association for Intelligent Machinery and the SOFT Scientific Contribution Memorial Award from the Society for Fuzzy Theory in Japan. In 1999, he was elected to membership in Berkeley Fellows and received the Certificate of Merit from IFSA (International Fuzzy Systems Association). In 2000, he received the IEEE Millennium Medal; the IEEE Pioneer Award in Fuzzy Systems; the ASPIH 2000 Lifetime Distinguished Achievement Award; and the ACIDCA 2000 Award for the paper, "From Computing with Numbers to Computing with Words – From Manipulation of Measurements to Manipulation of Perceptions." In addition, he received the Chaos Award from the Center of Hyperincursion and Anticipation in Ordered Systems for his outstanding scientific work on foundations of fuzzy logic, soft computing, computing with words and the computational theory of perceptions. In 2001, Dr. Zadeh received the ACM 2000 Allen Newell Award for seminal contributions to AI through his development of fuzzy logic. In addition, he received a Special Award from the Committee for Automation and Robotics of the Polish Academy of Sciences for his significant contributions to systems and information science, development of fuzzy sets theory, fuzzy logic control, possibility theory, soft computing, computing with words and computational theory of perceptions. In 2003, Dr. Zadeh was elected as a foreign member of the Finnish Academy of Sciences, and received the Norbert Wiener Award of the IEEE Society of Systems, Man and Cybernetics "For pioneering contributions to the development of system theory, fuzzy logic and soft computing." In 2004, Dr. Zadeh was awarded Civitate Honoris Causa by Budapest Tech (BT) Polytechnical Institution, Budapest, Hungary. Also in 2004, he was awarded the V. Kaufmann Prize, International Association for Fuzzy-Set Management and Economy (SIGEF).

Dr. Zadeh is a recipient of twenty-three honorary doctorates from: Paul-Sabatier University, Toulouse, France;

State University of New York, Binghamton, NY; University of Dortmund, Dortmund, Germany; University of Oviedo, Oviedo, Spain; University of Granada, Granada, Spain; Lakehead University, Canada; University of Louisville, KY; Baku State University, Azerbaijan; the Silesian Technical University, Gliwice, Poland; the University of Toronto, Toronto, Canada; the University of Ostrava, the Czech Republic; the University of Central Florida, Orlando, FL; the University of Hamburg, Hamburg, Germany; the University of Paris(6), Paris, France; Jahannes Kepler University, Linz, Austria; University of Waterloo, Canada; and the University of Aurel Vlaicu, Arad, Romania; Lappeenranta University of Technology, Lappeenranta, Finland; Muroran Institute of Technology, Muroran, Japan; Hong Kong Baptist University, Hong Kong, China.

Dr. Zadeh has single-authored over two hundred papers and serves on the editorial boards of over fifty journals. He is a member of the Advisory Committee, Center for Education and Research in Fuzzy Systems and Artificial Intelligence, Iasi, Romania; Senior Advisory Board, International Institute for General Systems Studies; the Board of Governors, International Neural Networks Society; and is the Honorary President of the Biomedical Fuzzy Systems Association of Japan and the Spanish Association for Fuzzy Logic and Technologies. In addition, he is a member of the Advisory Board of the National Institute of Informatics, Tokyo; a member of the Governing Board, Knowledge Systems Institute, Skokie, IL; and an honorary member of the Academic Council of NAISO-IAAC.

Professor in the Graduate School and Director, Berkeley Initiative in Soft Computing (BISC), Computer Science Division, Department of EECS, University of California, Berkeley, CA 94720-l776; Telephone: 5l0-642-4959; Fax: 5l0-642-l7l2;

E-mail: zadeh@eecs.berkeley.edu

http://www.cs.berkeley.edu/~zadeh/

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

Abstract

FROM SEARCH ENGINES TO QUESTION-ANSWERING SYSTEMS – THE PROBLEMS OF WORLD KNOWLEDGE, RELEVANCE, DEDUCTION AND PRECISIATION

Lotfi A. Zadeh*

Existing search engines, with Google at the top, have many truly remarkable capabilities. Furthermore, constant progress is being made in improving their performance. But what is not widely recognized is that there is a basic capability which existing search engines do not have: deduction capability – the capability to synthesize an answer to a query by drawing on bodies of information which reside in various parts of the knowledge base. By definition, a question-answering system, or a Q/A system for short, is a system which has deduction capability. Can a search engine be upgraded to a question-answering system through the use of existing tools – tools which are based on bivalent logic and probability theory? A view which is articulated in the following is that the answer is: No.

The first obstacle is world knowledge – the knowledge which humans acquire through experience, communication and education. Simple examples are: "Icy roads are slippery," "Princeton usually means Princeton University," "Paris is the capital of France," and "There are no honest politicians." World knowledge plays a central role in search, assessment of relevance

---

* Professor in the Graduate School and Director, Berkeley initiative in Soft Computing (BISC), Computer Science Division and the Electronics Research Laboratory, Department of EECS, University of California, Berkeley, CA 94720-1776; Telephone: 510-642-4959; Fax: 510-642-1712; E-Mail: zadeh@cs.berkeley.edu .

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

and deduction. The problem with world knowledge is that it is, for the most part, perception-based. Perceptions–and especially perceptions of probabilities–are intrinsically imprecise, reflecting the fact that human sensory organs, and ultimately the brain, have a bounded ability to resolve detail and store information. Imprecision of perceptions stands in the way of using conventional techniques – techniques which are based on bivalent logic and probability theory – to deal with perception-based information. A further complication is that much of world knowledge is negative knowledge in the sense that it relates to what is impossible and/or non-existent. For example, “A person cannot have two fathers,” and “Netherlands has no mountains.”

The second obstacle centers on the concept of relevance. There is an extensive literature on relevance, and every search engine deals with relevance in its own way, some at a high level of sophistication. But what is quite obvious is that the problem of assessment of relevance is quite complex and far from solution.

There are two kinds of relevance: (a) question relevance and (b) topic relevance. Both are matters of degree. For example, on a very basic level, if the question is *q*: “Number of cars in California?” and the available information is *p*: “Population of California is 37,000,000,” then what is the degree of relevance of *p* to *q*? Another example: To what degree is a paper entitled “A New Approach to Natural Language Understanding” of relevance to the topic of machine translation.

Basically, there are two ways of approaching assessment of relevance: (a) semantic; and (b) statistical. To illustrate, in the number of cars example, relevance of *p* to *q* is a matter of semantics and world knowledge. In existing search engines, relevance is largely a matter of statistics, involving counts of links and words, with little if any consideration of semantics. Assessment of semantic relevance presents difficult problems whose solutions lie beyond the reach of bivalent logic and probability theory. What should be noted is that assessment of topic relevance is more amendable to the use of statistical techniques, which explains why existing search engines are much better at assessment of topic relevance then question relevance.

The third obstacle is deduction from perception-based information. As a basic example, assume that the question is *q*: What is the average height of Swedes?, and the available information is *p*: Most adult Swedes are tall. Another example is: Usually Robert returns from work at about 6 pm. What is the probability that Robert is at home at 6:15 pm? Neither bivalent logic nor probability theory provide effective tools for dealing with problems of this type. The difficulty is centered on deduction from premises which are both uncertain and imprecise.

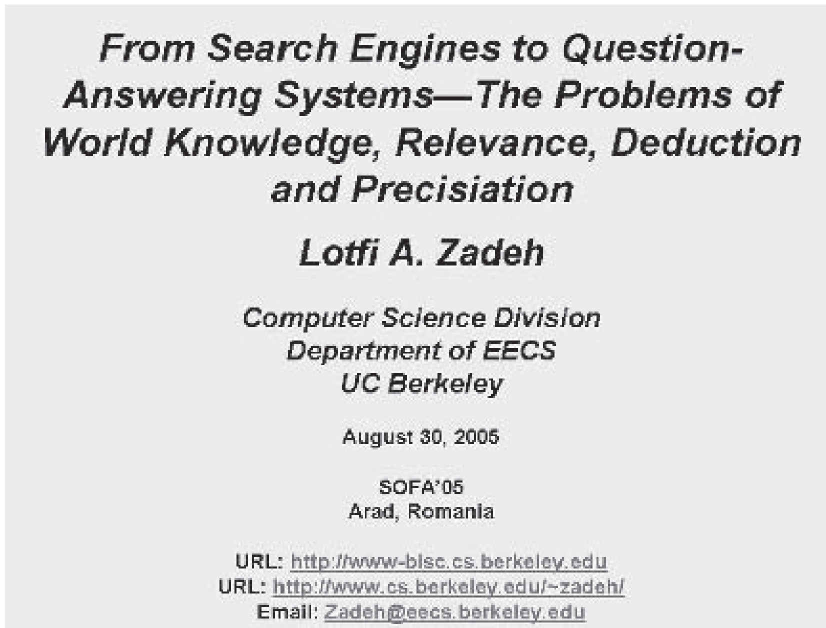
Underlying the problems of world knowledge, relevance and deduction is a very basic problem – the problem of natural language understanding. Much of world knowledge and web knowledge is expressed in a natural language. A natural language is basically a system for describing perceptions. Since perceptions are intrinsically imprecise, so are natural languages.

A prerequisite to mechanization of question-answering is mechanization of natural language understanding, and a prerequisite to mechanization of natural language understanding is precisiation of meaning of concepts and proposition drawn from a natural language. To deal effectively with world knowledge, relevance, deduction and precisiation, new tools are needed. The principal new tools are: Precisiated Natural Language (PNL); Protoform Theory (PFT); and the Generalized Theory of Uncertainty (GTU). These tools are drawn from fuzzy logic–a logic in which everything is, or is allowed to be, a matter of degree.

The centerpiece of the new tools is the concept of a generalized constraint. The importance of the concept of a generalized constraint derives from the fact that in PNL and GTU it serves as a basis for generalizing the universally accepted view that information is statistical in nature. More specifically, the point of departure in PNL and GTU is the fundamental premise that, in general, information is representable as a system of generalized constraints, with statistical information constituting a special case. This, much more general, view of information is needed to deal effectively with world knowledge, relevance, deduction, precisiation and related problems.

In summary, the principal objectives of this paper are: (a) to make a case for the view that a quantum jump in search engine IQ cannot be achieved through the use of methods based on bivalent logic and probability theory; and (b) to introduce and outline a collection of non-standard concepts, ideas and tools which are needed to achieve a quantum jump in search engine IQ.

Scientific and Technical Bulletin Series: Electrotechnics, Electronics, Automatic Control and Computer Science, Vol. 2, No. 2, 2005, ISSN 1584-9198



KEY ISSUE—DEDUCTION CAPABILITY

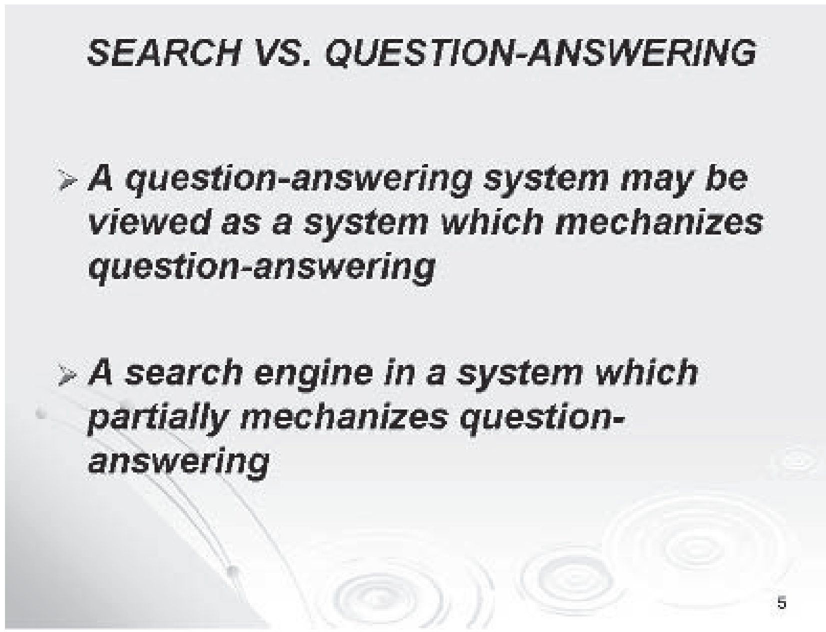
Existing search engines, with Google at the top, have many truly remarkable capabilities. Furthermore, constant progress is being made in improving their performance. But what should be realized is that existing search engines do not have an important capability—deduction capability—the capability to synthesize an answer to a query by drawing on bodies of information which reside in various parts of the knowledge base.

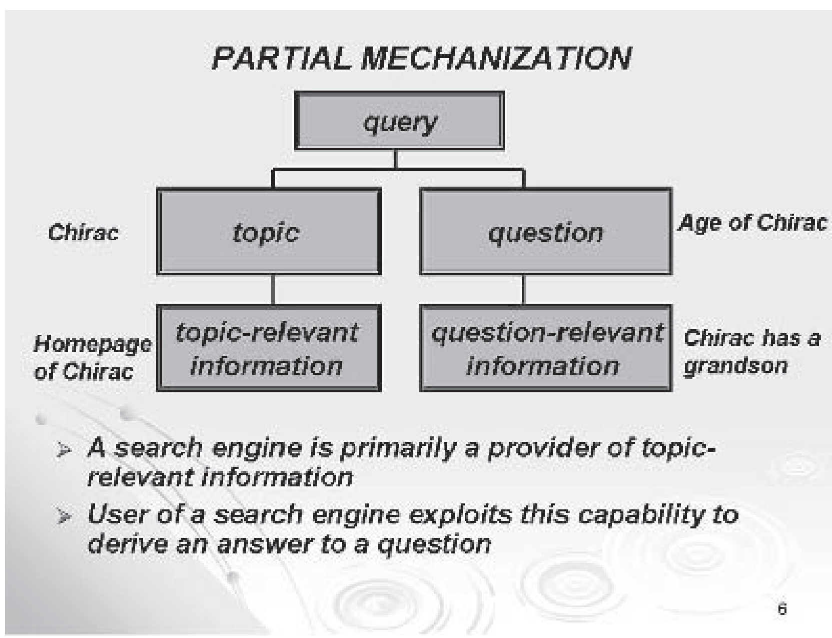
3

CONTINUED

- *What should be noted, however, is that there are many widely used special purpose question-answering systems which have limited deduction capability. Examples of such systems are driving direction systems, reservation systems, diagnostic systems and specialized expert systems, especially in the domain of medicine.*

4


SEARCH VS. QUESTION-ANSWERING
A question-answering system may be viewed as a system which mechanizes question-answering
A search engine in a system which partially mechanizes question-answering
5


PARTIAL MECHANIZATION
query
Chirac
topic
question
Age of Chirac
Homepage of Chirac
topic-relevant information
question-relevant information
Chirac has a grandson
A search engine is primarily a provider of topic-relevant information
User of a search engine exploits this capability to derive an answer to a question
6

COMPLEXITY OF UPGRADING

- *Addition of deduction capability to a search engine is a highly complex problem—a problem which is a major challenge to computer scientists and logicians*
- *A view which is articulated in the following is that the challenge cannot be met through the use of existing methods—methods which are based on bivalent logic and probability theory*
- *To add deduction capability to a search engine it is necessary to (a) generalize bivalent logic; (b) generalize probability theory*

7

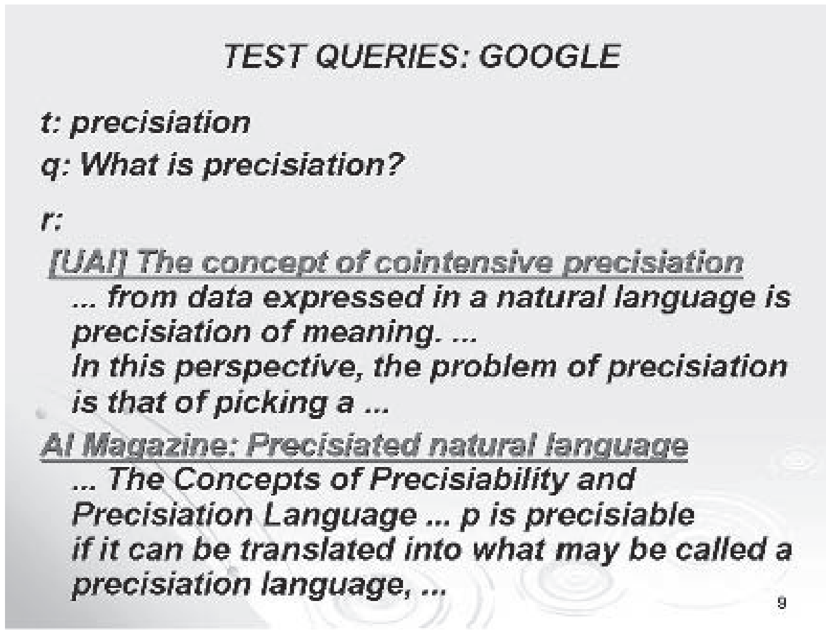
HISTORICAL NOTE

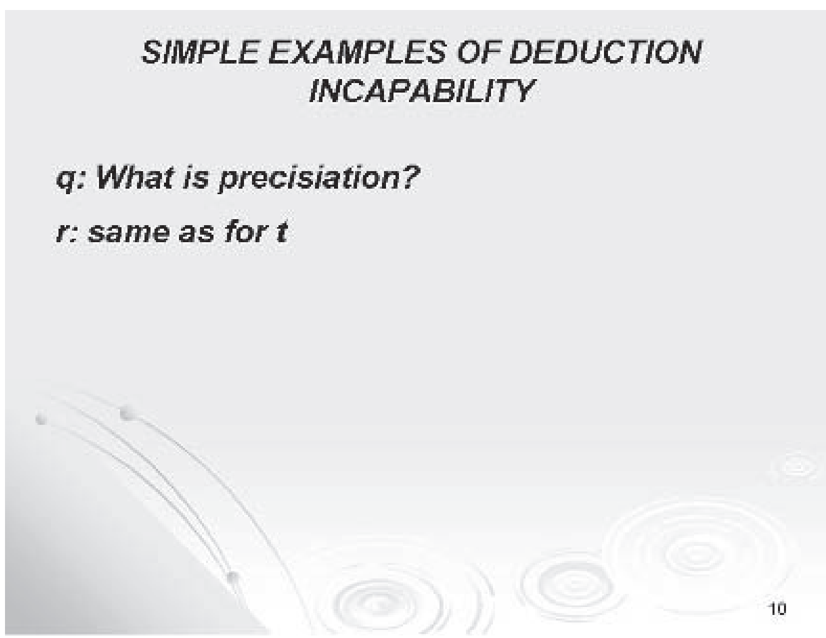
- *1970-1980 was a period of intense interest in question-answering and expert systems*
- *There was no discussion of search engines*

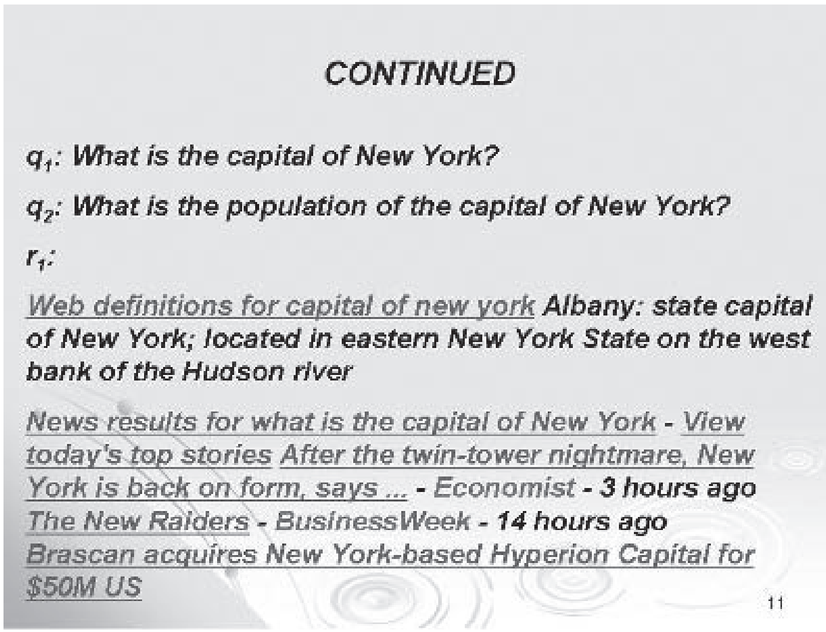
Example: L.S. Coles, "Techniques for Information Retrieval Using an Inferential Question-Answering System with Natural Language Input," SRI Report, 1972

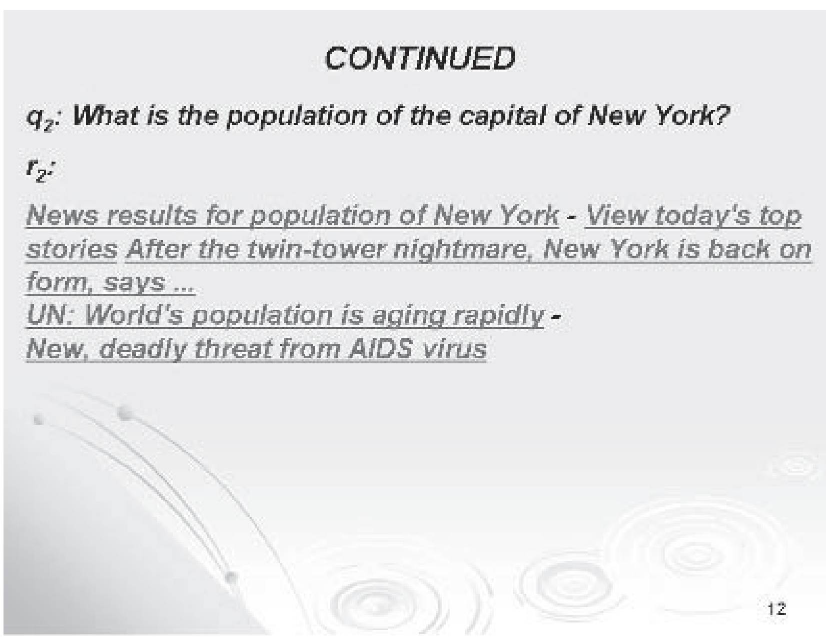
- *M. Nagao, J. Tsujii: Mechanism of Deduction in a Question-Answering System with Natural Language Inputd. IJCAI 1973: 285-290.*
- *J. R. McSkimin, J. Minker: The Use of a Semantic Network in a Deductive Question- Answering System. IJCAI 1977: 50-58.*
- *A. R. Aronson, B. E. Jacobs, J. Minker: A Note on Fuzzy Deduction. J. ACM 27(4): 599-603 (1980)*
- *W.J.H.J. Bronnenberg, H.C. Bunt, S.P.J. Lendsbergen, R.J.H. Scha, W.J. Schoenmakers and E.P.C. van Utteren. The Question Answering System PHLIQA1. In L. Bolc (editor), Natural Language Question Answering Systems. Macmillan, 1980.*

8

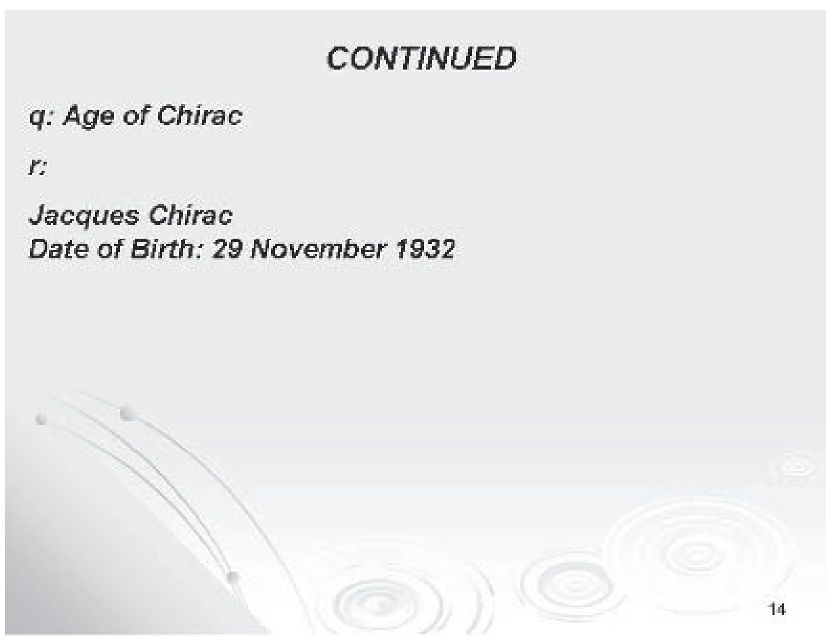

TEST QUERIES: GOOGLE
t: precisiation
q: What is precisiation?
r:
[UAI] The concept of cointensive precisiation
... from data expressed in a natural language is precisiation of meaning. ...
In this perspective, the problem of precisiation is that of picking a ...
AI Magazine: Precisiated natural language
... The Concepts of Precisiability and Precisiation Language ... p is precisiable if it can be translated into what may be called a precisiation language, ...
9

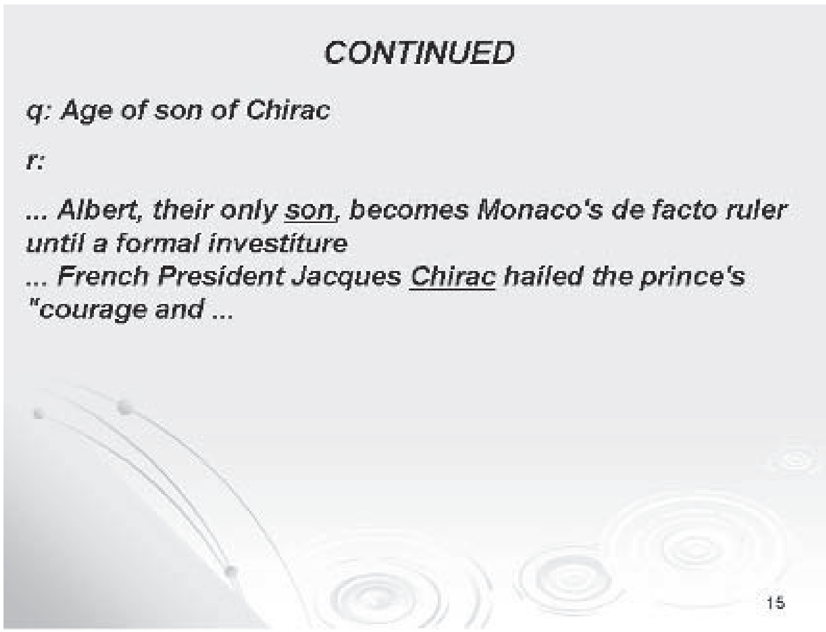

SIMPLE EXAMPLES OF DEDUCTION INCAPABILITY
q: What is precisiation?
r: same as for t
10

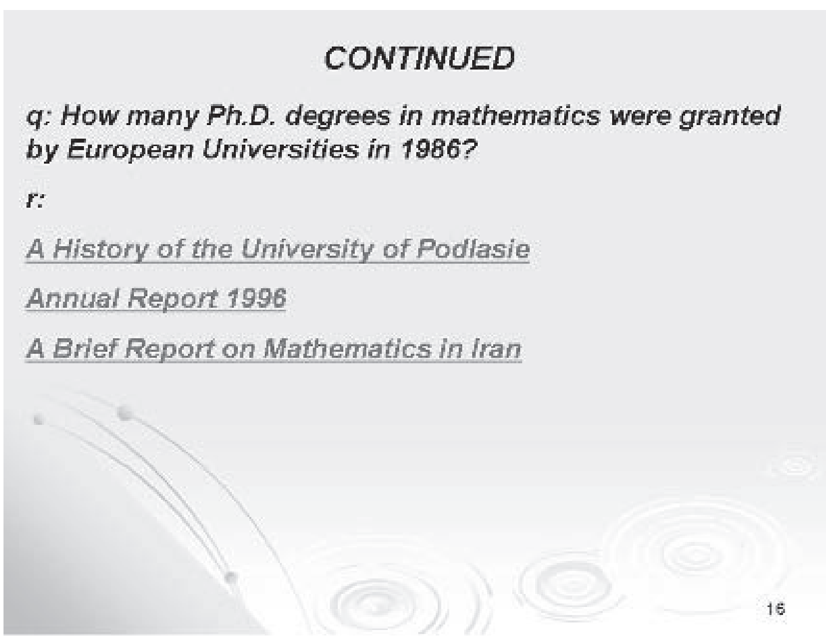

CONTINUED
q1: What is the capital of New York?
q2: What is the population of the capital of New York?
r1:
Web definitions for capital of new york Albany: state capital of New York; located in eastern New York State on the west bank of the Hudson river
News results for what is the capital of New York - View today's top stories After the twin-tower nightmare, New York is back on form, says ... - Economist - 3 hours ago
The New Raiders - BusinessWeek - 14 hours ago
Brascan acquires New York-based Hyperion Capital for $50M US
11

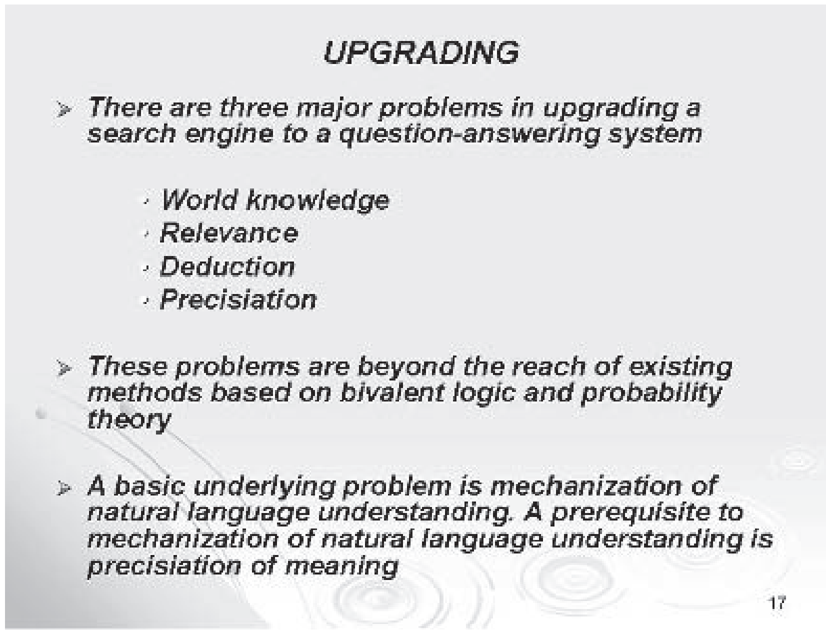

CONTINUED
q2: What is the population of the capital of New York?
r2:
News results for population of New York - View today's top stories After the twin-tower nightmare, New York is back on form, says ...
UN: World's population is aging rapidly -
New, deadly threat from AIDS virus
12

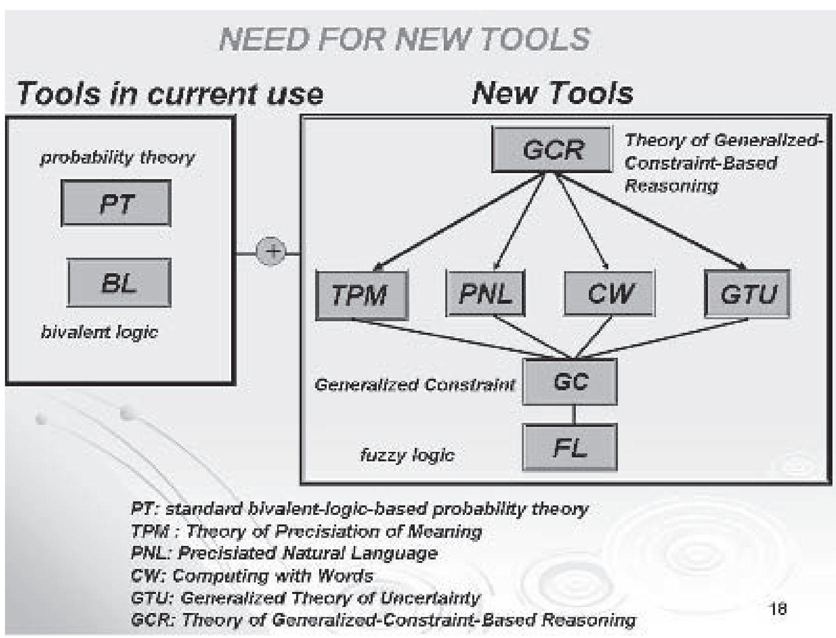
Scientific and Technical Bulletin Series: Electrotechnics, Electronics, Automatic Control and Computer Science, Vol. 2, No. 2, 2005**,** ISSN 1584-9198

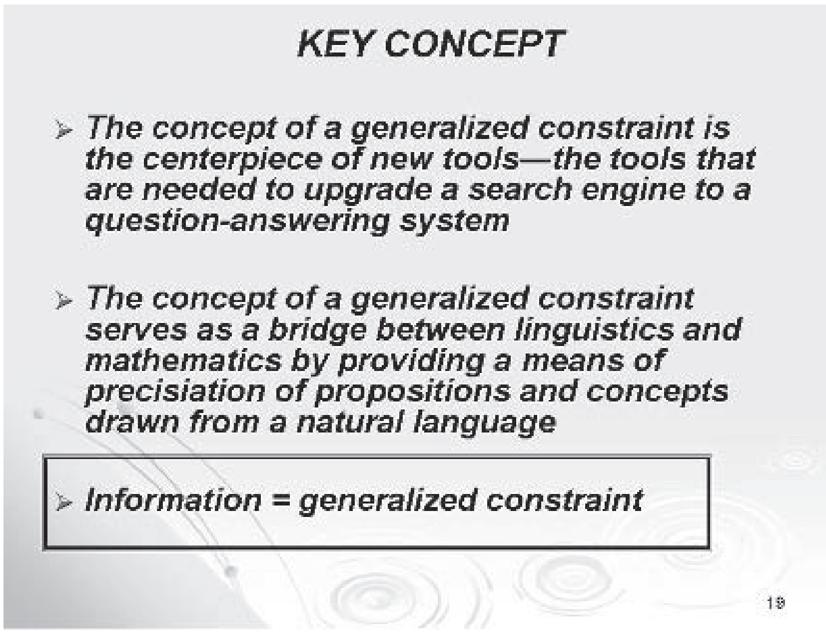



CONTINUED
q: Age of son of Chirac
r:
... Albert, their only son, becomes Monaco's de facto ruler until a formal investiture
... French President Jacques Chirac hailed the prince's "courage and ...
15

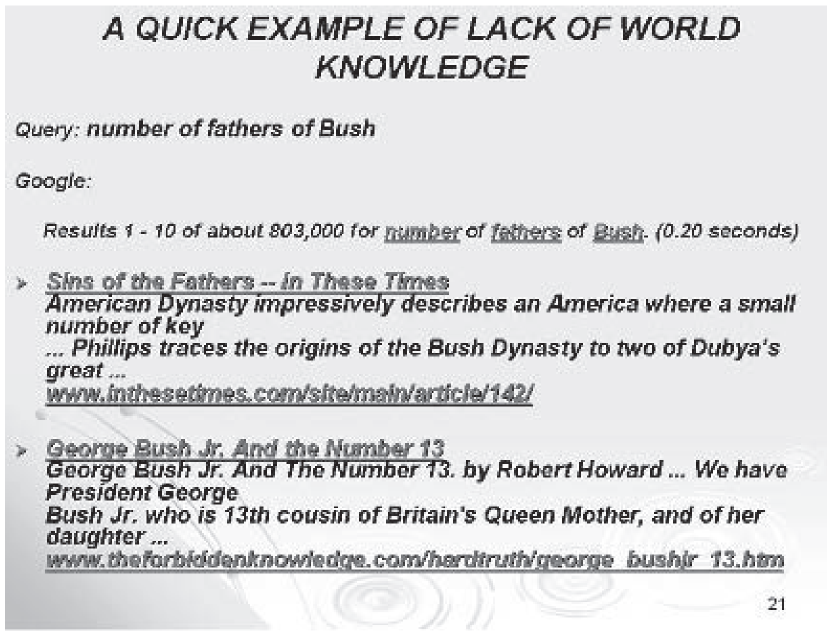

CONTINUED
q: How many Ph.D. degrees in mathematics were granted by European Universities in 1986?
r:
A History of the University of Podlasie
Annual Report 1996
A Brief Report on Mathematics in Iran
16


UPGRADING
There are three major problems in upgrading a search engine to a question-answering system
World knowledge
Relevance
Deduction
Precisiation
These problems are beyond the reach of existing methods based on bivalent logic and probability theory
A basic underlying problem is mechanization of natural language understanding. A prerequisite to mechanization of natural language understanding is precisiation of meaning
17


NEED FOR NEW TOOLS
Tools in current use
New Tools
probability theory
PT
BL
bivalent logic
+
GCR
Theory of Generalized-Constraint-Based Reasoning
TPM
PNL
CW
GTU
Generalized Constraint
GC
fuzzy logic
FL
PT: standard bivalent-logic-based probability theory
TPM : Theory of Precisiation of Meaning
PNL: Precisiated Natural Language
CW: Computing with Words
GTU: Generalized Theory of Uncertainty
GCR: Theory of Generalized-Constraint-Based Reasoning
18

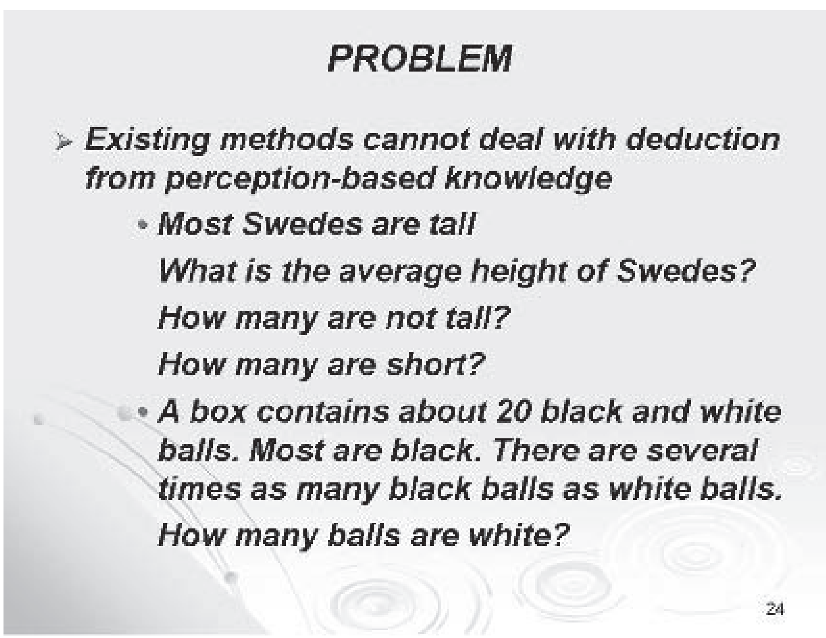

KEY CONCEPT
The concept of a generalized constraint is the centerpiece of new tools—the tools that are needed to upgrade a search engine to a question-answering system
The concept of a generalized constraint serves as a bridge between linguistics and mathematics by providing a means of precisiation of propositions and concepts drawn from a natural language
Information = generalized constraint
19

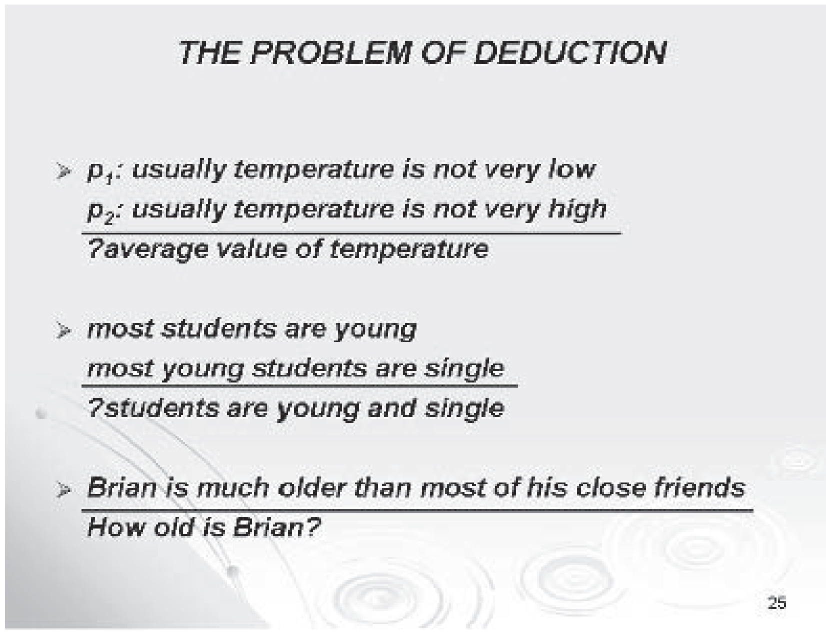

WORLD KNOWLEDGE
World knowledge is the knowledge acquired through experience, education and communication
Few professors are rich
There are no honest politicians
It is not likely to rain in San Francisco in midsummer
Most Swedes are tall
There are no mountains in Holland
Usually Princeton means Princeton University
A person can have only one father
20

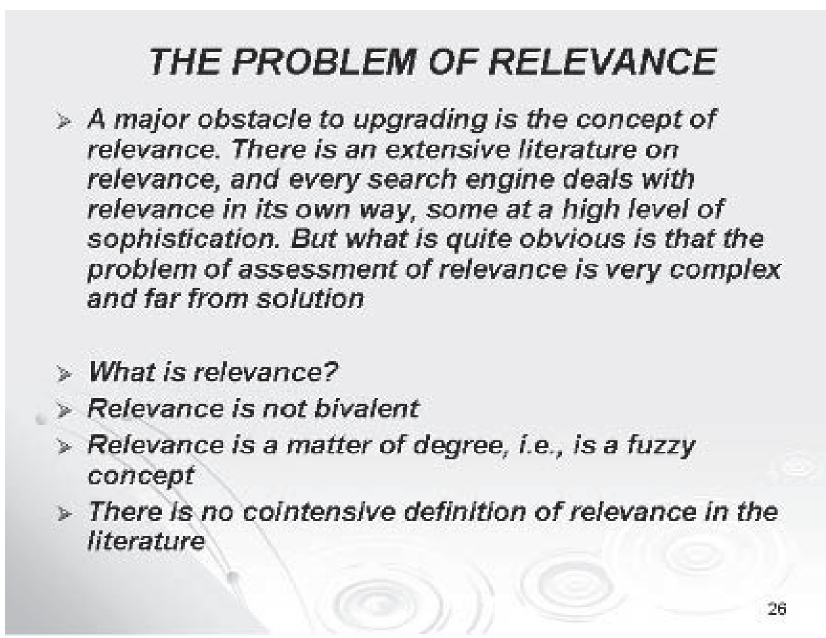

A QUICK EXAMPLE OF LACK OF WORLD KNOWLEDGE
Query: number of fathers of Bush
Google:
Results 1 - 10 of about 803,000 for number of fathers of Bush. (0.20 seconds)
Sins of the Fathers -- In These Times
American Dynasty impressively describes an America where a small number of key
... Phillips traces the origins of the Bush Dynasty to two of Dubya's great ...
www.inthesetimes.com/site/main/article/142/
George Bush Jr. And the Number 13
George Bush Jr. And The Number 13. by Robert Howard ... We have President George
Bush Jr. who is 13th cousin of Britain's Queen Mother, and of her daughter ...
www.theforbiddenknowledge.com/hardtruth/george_bushjr_13.htm
21

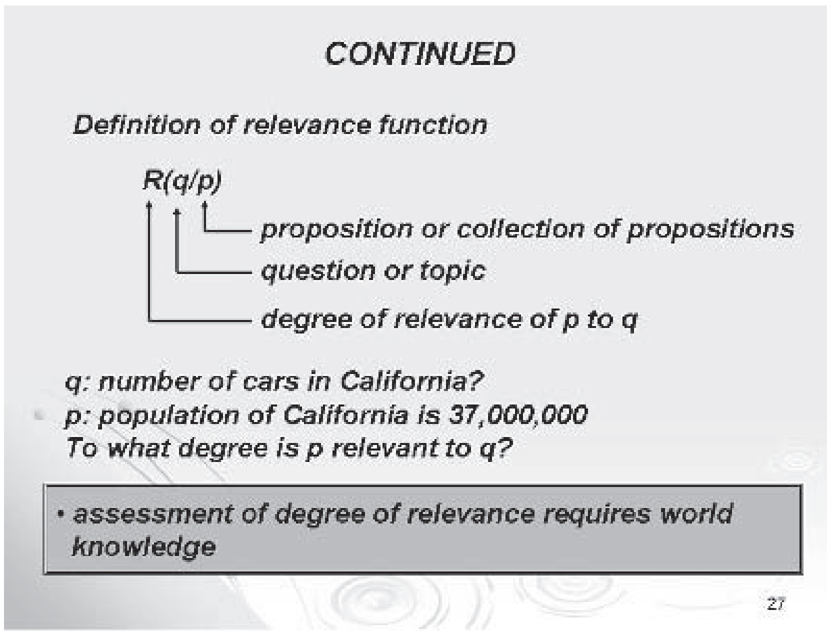

COMPONENTS OF WORLD KNOWLEDGE
Propositional
Paris is the capital of France
Conceptual
Climate
Ontological
Rainfall is related to climate
Existential
A person cannot have more than one father
Contextual
Tall
22

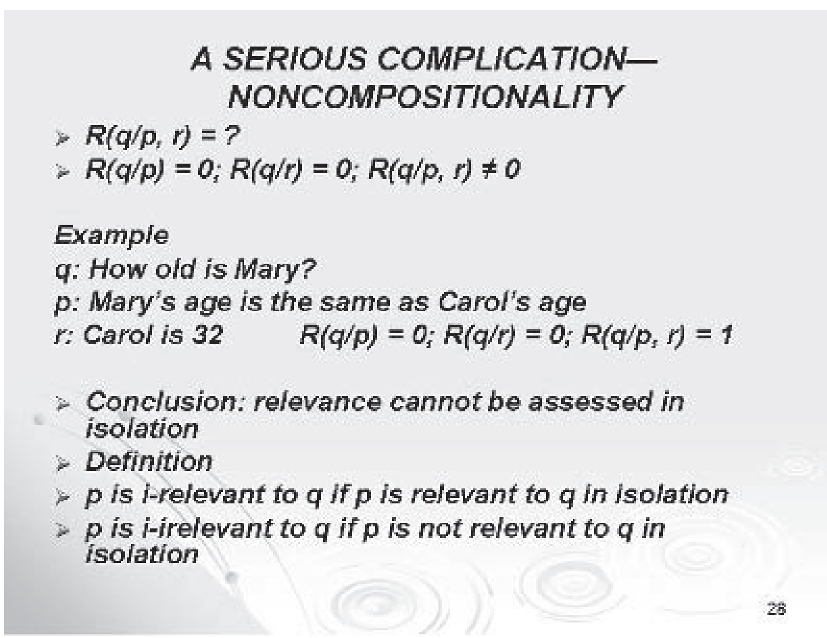

CONTINUED
Much of world knowledge is perception-based
Most Swedes are tall
Most Swedes are much taller than most Italians
Usually a large house costs more than a small house
Much of world knowledge is negative, i.e., relates to impossibility or nonexistence
A person cannot have more than one father
There are no honest politicians
Much of world knowledge is expressed in a natural language
23

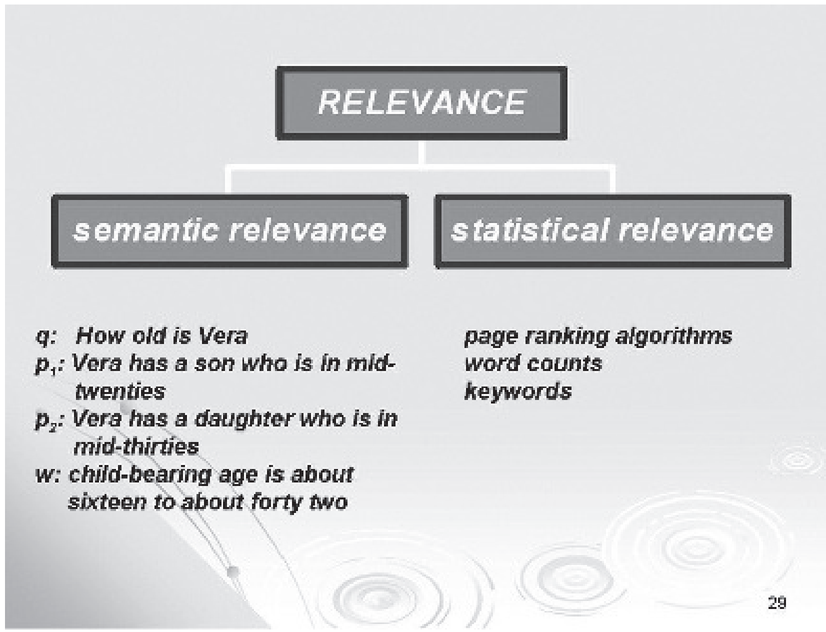

PROBLEM
Existing methods cannot deal with deduction from perception-based knowledge
Most Swedes are tall
What is the average height of Swedes?
How many are not tall?
How many are short?
A box contains about 20 black and white balls. Most are black. There are several times as many black balls as white balls.
How many balls are white?
24

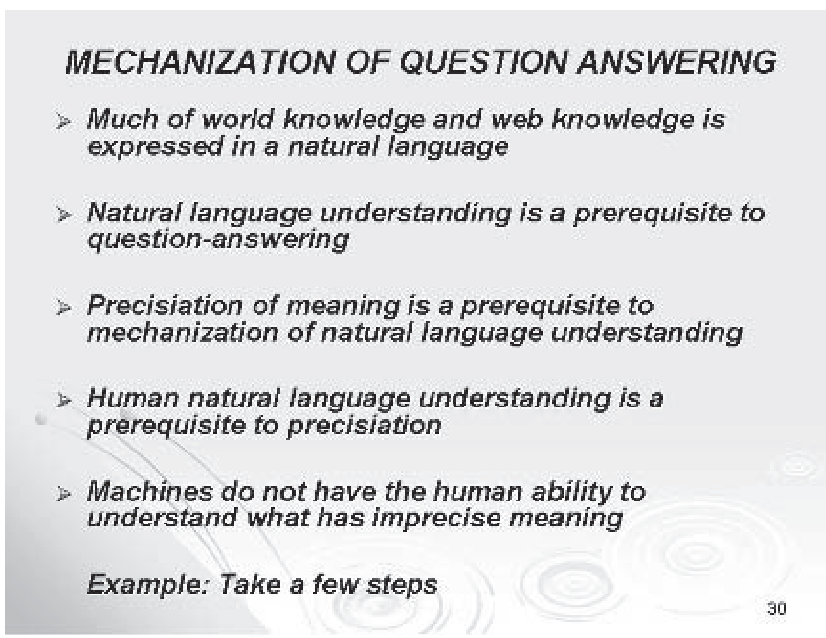

THE PROBLEM OF DEDUCTION
p1: usually temperature is not very low
p2: usually temperature is not very high
?average value of temperature
most students are young
most young students are single
?students are young and single
Brian is much older than most of his close friends
How old is Brian?
25

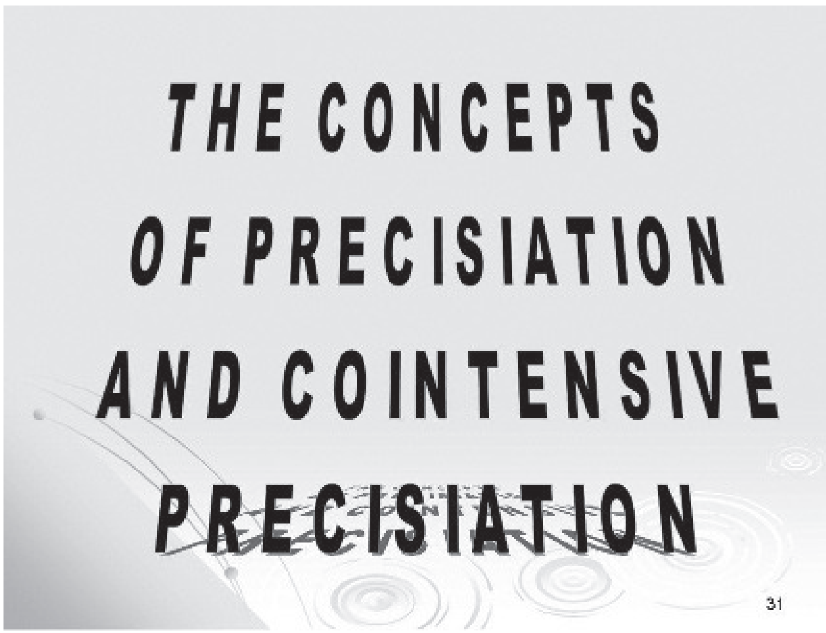

THE PROBLEM OF RELEVANCE
A major obstacle to upgrading is the concept of relevance. There is an extensive literature on relevance, and every search engine deals with relevance in its own way, some at a high level of sophistication. But what is quite obvious is that the problem of assessment of relevance is very complex and far from solution
What is relevance?
Relevance is not bivalent
Relevance is a matter of degree, i.e., is a fuzzy concept
There is no cointensive definition of relevance in the literature
26

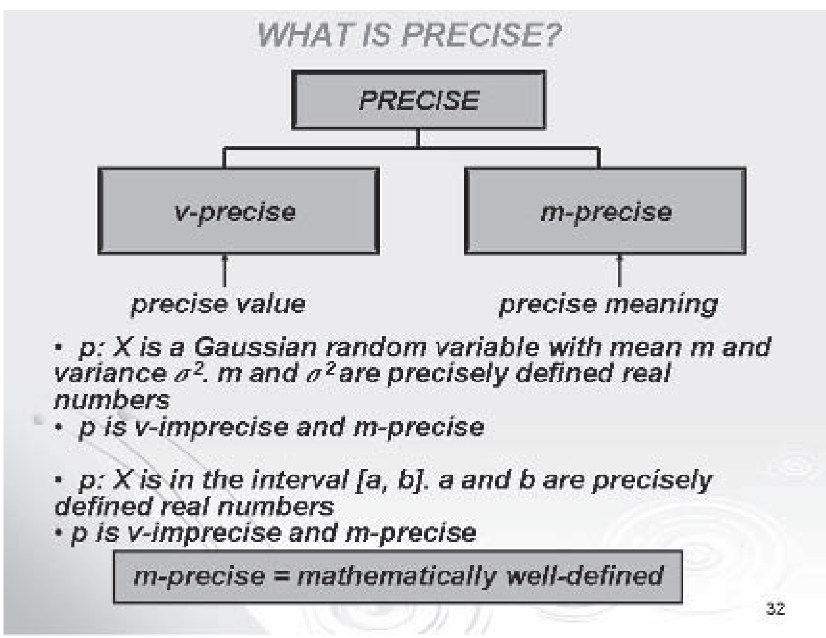

CONTINUED
Definition of relevance function
R(q/p)
proposition or collection of propositions
question or topic
degree of relevance of p to q
q: number of cars in California?
p: population of California is 37,000,000
To what degree is p relevant to q?
• assessment of degree of relevance requires world knowledge
27

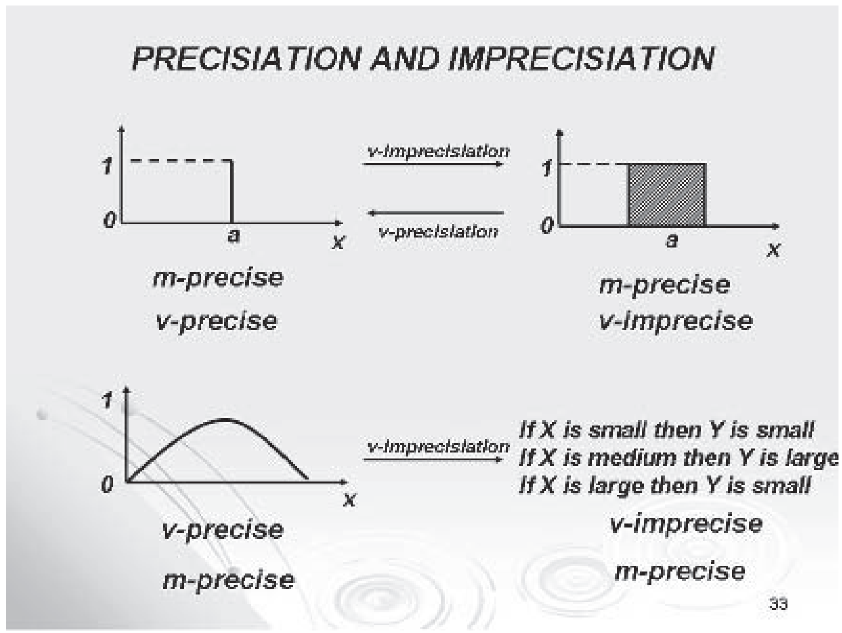

A SERIOUS COMPLICATION—NONCOMPOSITIONALITY
R(q/p, r) = ?
R(q/p) = 0; R(q/r) = 0; R(q/p, r) ≠ 0
Example
q: How old is Mary?
p: Mary's age is the same as Carol's age
r: Carol is 32
R(q/p) = 0; R(q/r) = 0; R(q/p, r) = 1
Conclusion: relevance cannot be assessed in isolation
Definition
p is i-relevant to q if p is relevant to q in isolation
p is i-irelevant to q if p is not relevant to q in isolation
28

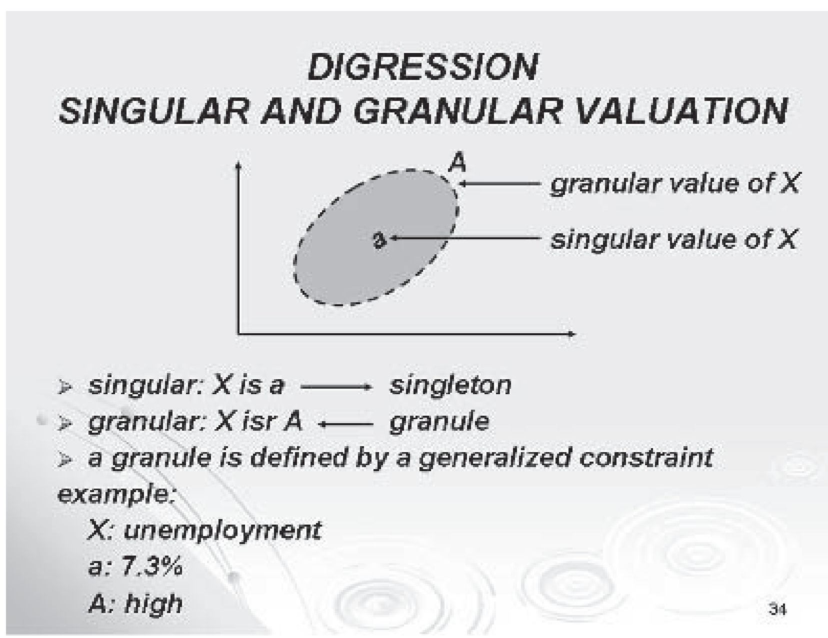

RELEVANCE
semantic relevance
statistical relevance
q: How old is Vera
p_1: Vera has a son who is in mid-twenties
p_2: Vera has a daughter who is in mid-thirties
w: child-bearing age is about sixteen to about forty two
page ranking algorithms
word counts
keywords
29

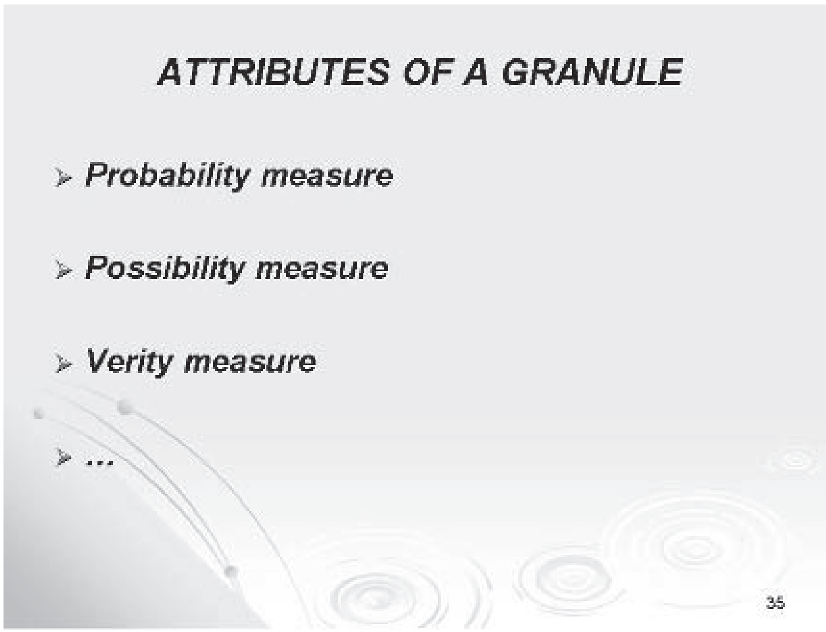

MECHANIZATION OF QUESTION ANSWERING
Much of world knowledge and web knowledge is expressed in a natural language
Natural language understanding is a prerequisite to question-answering
Precisiation of meaning is a prerequisite to mechanization of natural language understanding
Human natural language understanding is a prerequisite to precisiation
Machines do not have the human ability to understand what has imprecise meaning
Example: Take a few steps
30

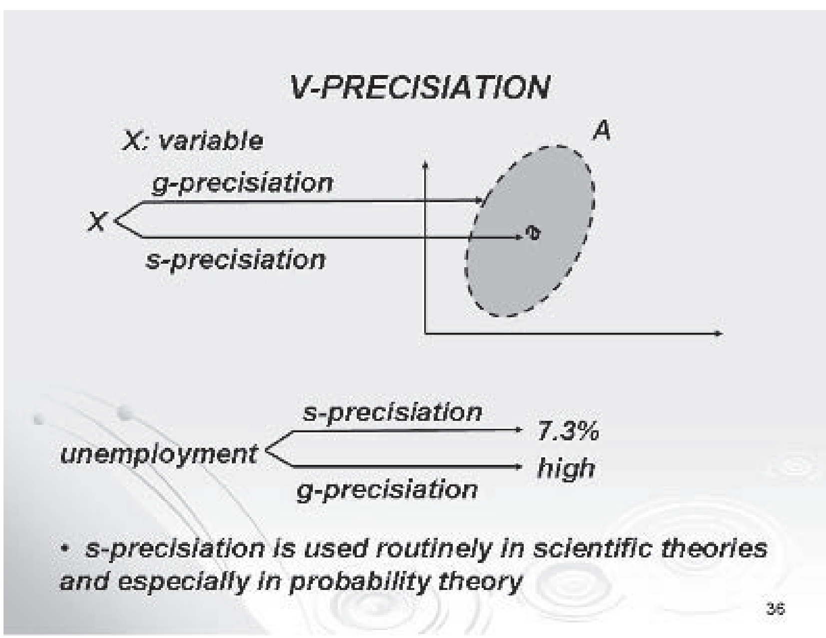

THE CONCEPTS
OF PRECISIATION
AND COINTENSIVE
PRECISIATION
31

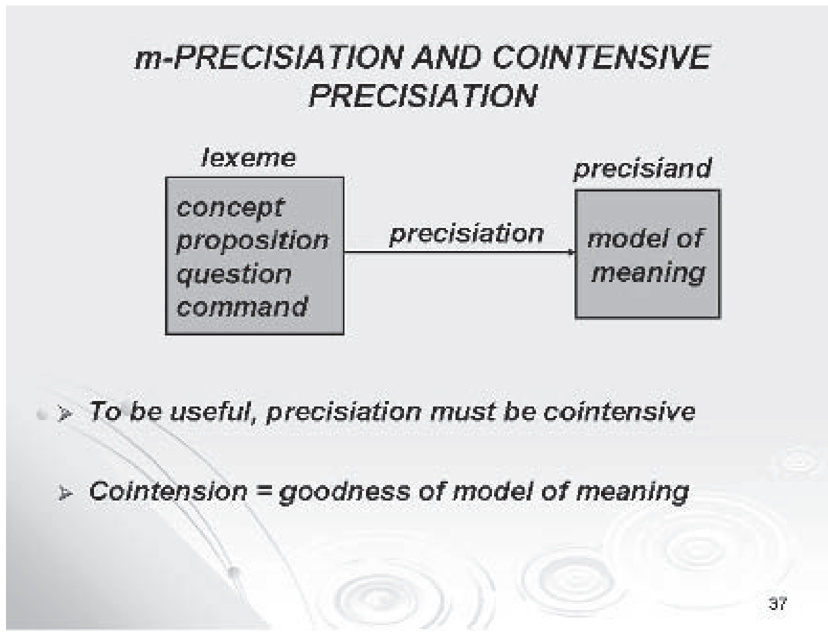

WHAT IS PRECISE?
PRECISE
v-precise
m-precise
precise value
precise meaning
• p: X is a Gaussian random variable with mean m and variance σ^2. m and σ^2 are precisely defined real numbers
• p is v-imprecise and m-precise
• p: X is in the interval [a, b]. a and b are precisely defined real numbers
• p is v-imprecise and m-precise
m-precise = mathematically well-defined
32

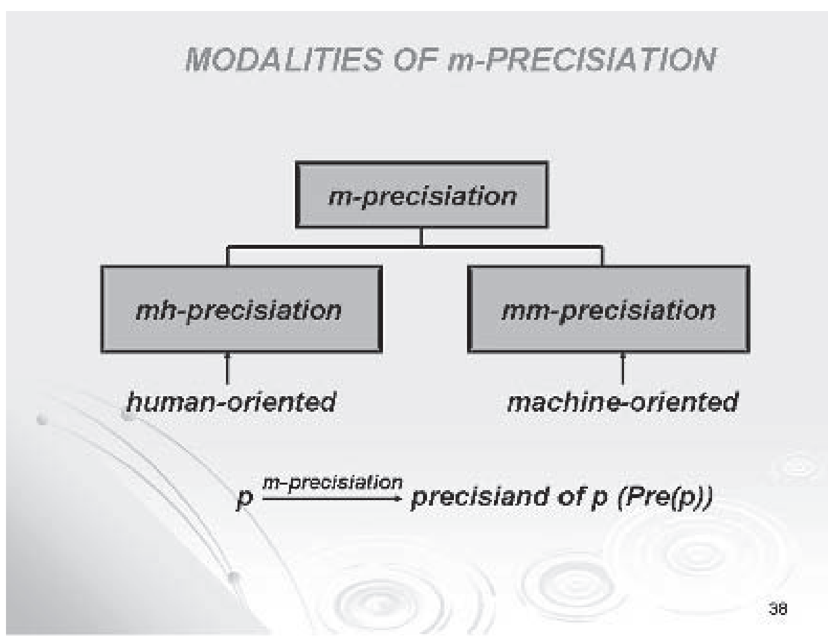

PRECISIATION AND IMPRECISIATION
1
0
a
X
v-imprecisiation
v-precisiation
1
0
a
X
m-precise
v-precise
m-precise
v-imprecise
1
0
X
v-imprecisiation
If X is small then Y is small
If X is medium then Y is large
If X is large then Y is small
v-precise
m-precise
v-imprecise
m-precise
33

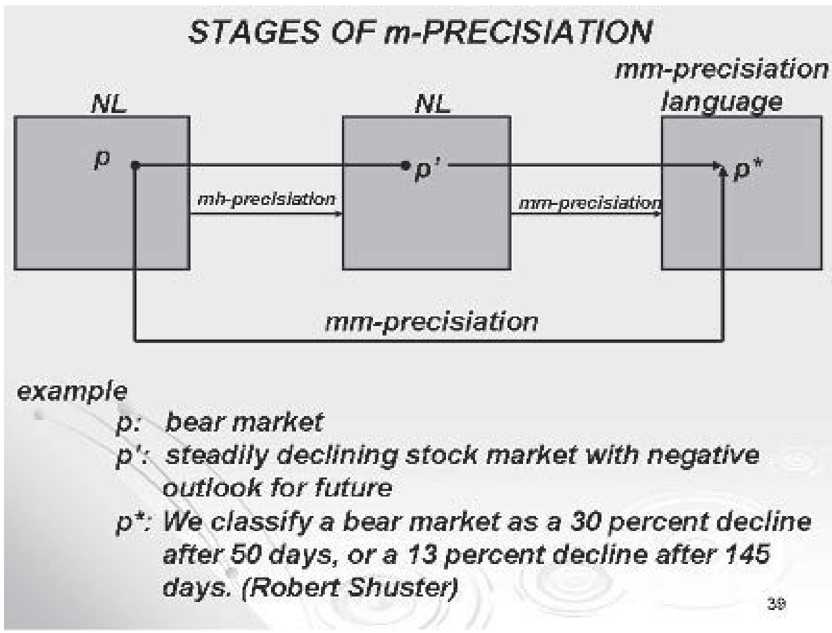

DIGRESSION
SINGULAR AND GRANULAR VALUATION
A
granular value of X
a
singular value of X
singular: X is a ⟶ singleton
granular: X isr A ⟵ granule
a granule is defined by a generalized constraint
example:
X: unemployment
a: 7.3%
A: high
34

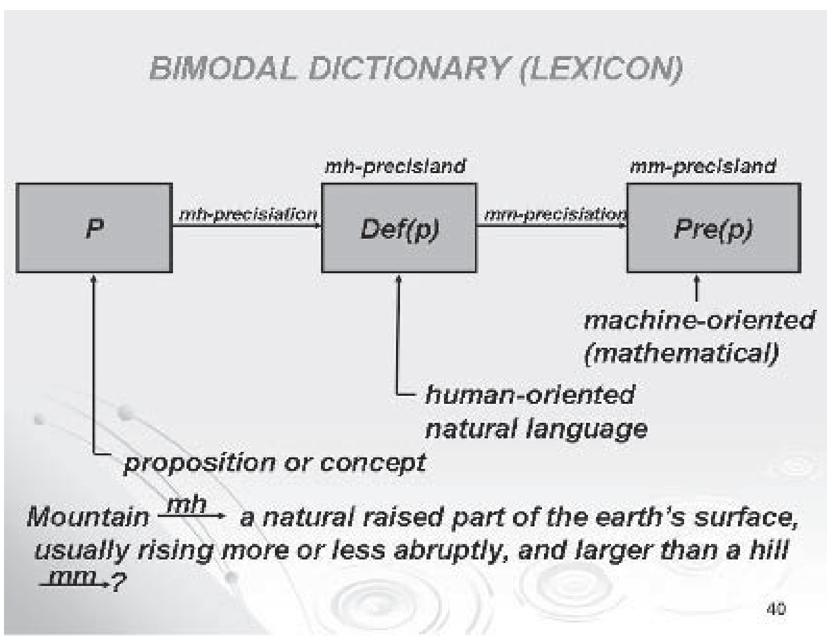

ATTRIBUTES OF A GRANULE
Probability measure
Possibility measure
Verity measure
...
35

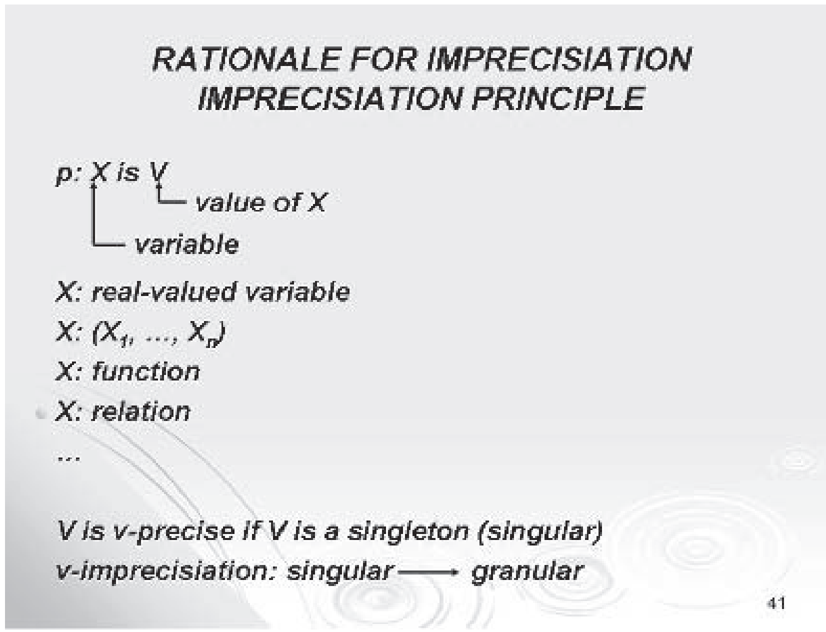

V-PRECISIATION
X: variable
A
g-precisiation
X
a
s-precisiation
s-precisiation
unemployment
7.3%
high
g-precisiation
• s-precisiation is used routinely in scientific theories and especially in probability theory
36

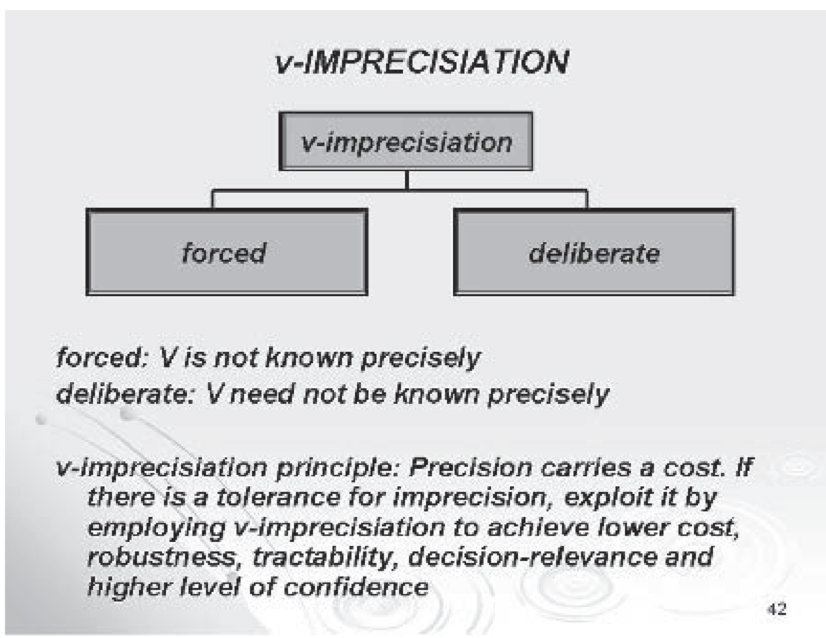

m-PRECISIATION AND COINTENSIVE PRECISIATION
lexeme
concept
proposition
question
command
precisiation
precisiand
model of meaning
To be useful, precisiation must be cointensive
Cointension = goodness of model of meaning
37

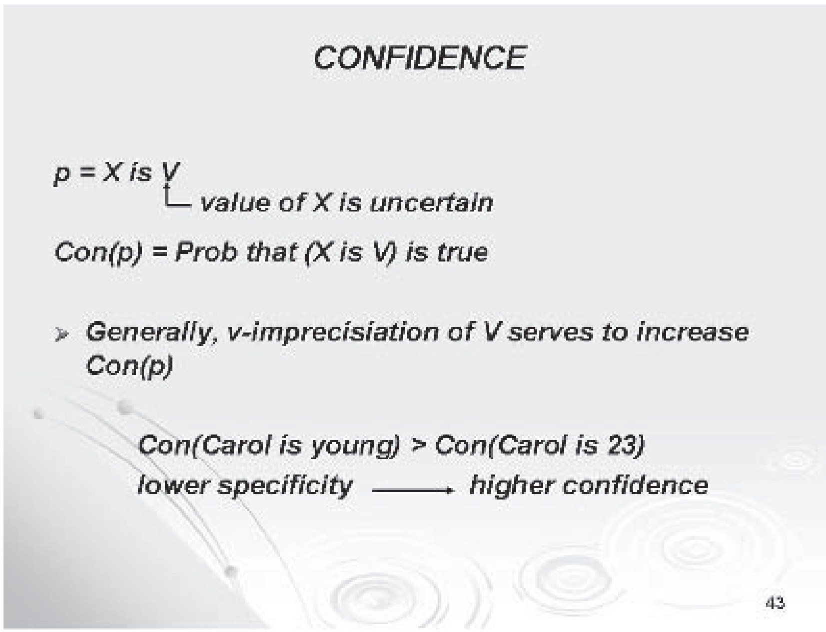

MODALITIES OF m-PRECISIATION
m-precisiation
mh-precisiation
mm-precisiation
human-oriented
machine-oriented
p $\xrightarrow{m\text{-precisiation}}$ precisiand of p (Pre(p))
38

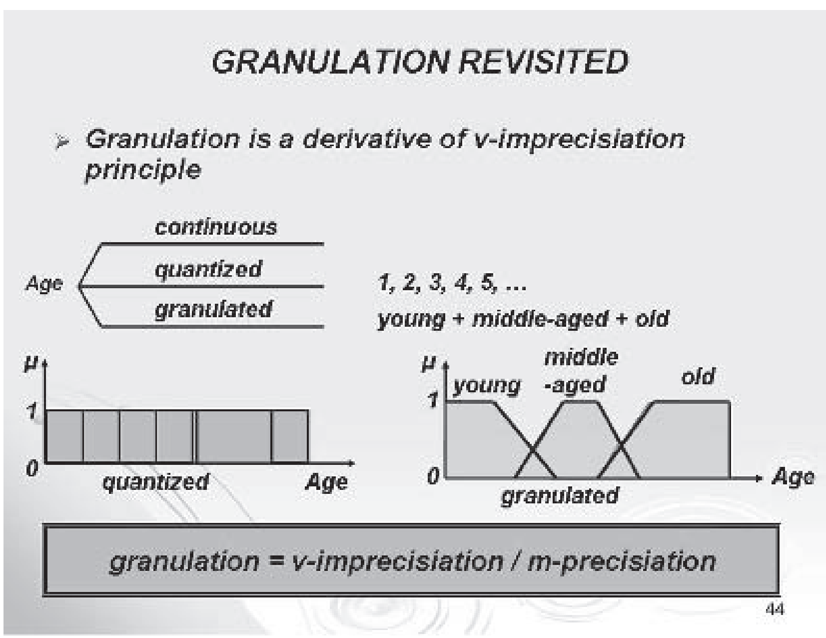

STAGES OF m-PRECISIATION
mm-precisiation language
NL
NL
p
p'
p*
mh-precisiation
mm-precisiation
mm-precisiation
example
p: bear market
p': steadily declining stock market with negative outlook for future
p*: We classify a bear market as a 30 percent decline after 50 days, or a 13 percent decline after 145 days. (Robert Shuster)
39

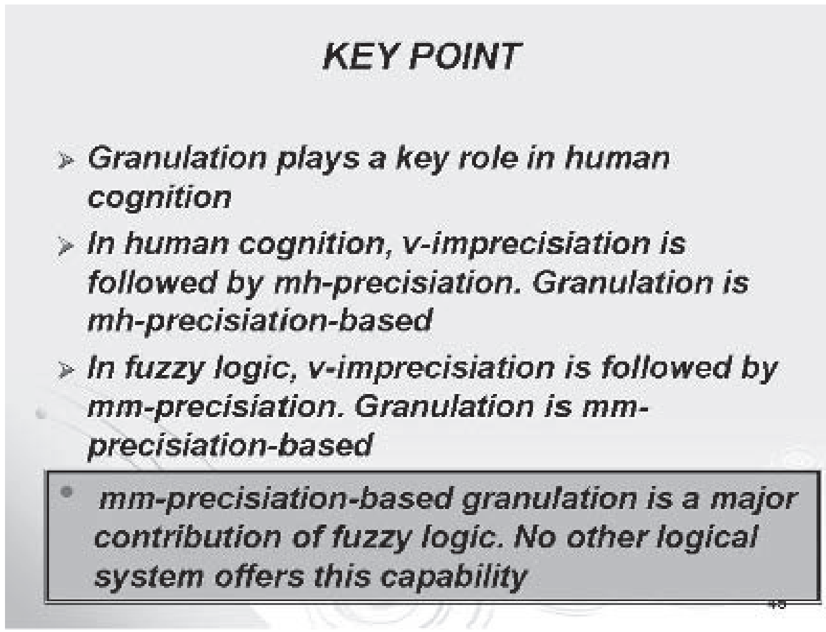

BIMODAL DICTIONARY (LEXICON)
mh-precisiand
mm-precisiand
P
mh-precisiation
Def(p)
mm-precisiation
Pre(p)
machine-oriented (mathematical)
human-oriented natural language
proposition or concept
Mountain mh a natural raised part of the earth's surface, usually rising more or less abruptly, and larger than a hill
mm ?
40

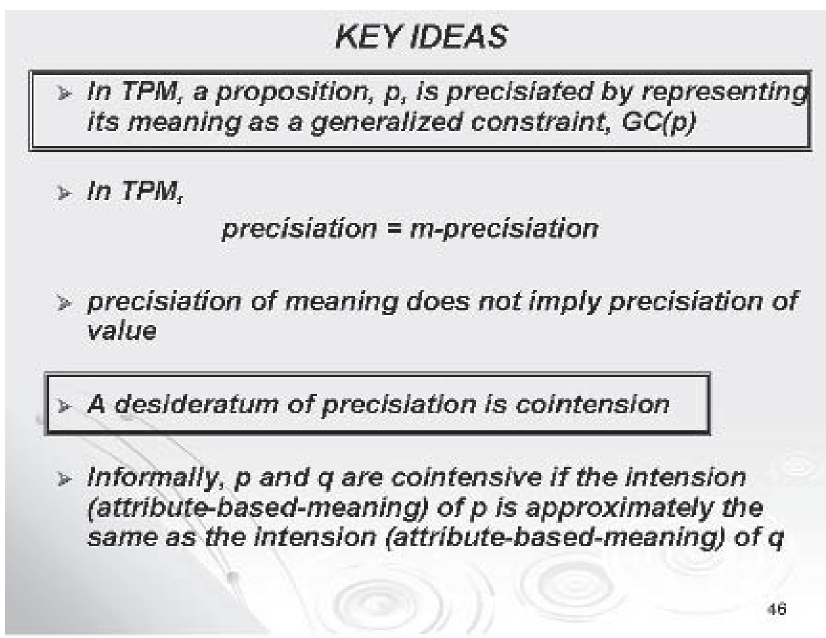

RATIONALE FOR IMPRECISIATION
IMPRECISIATION PRINCIPLE
p: X is V
value of X
variable
X: real-valued variable
X: $(X_1, \ldots, X_n)$
X: function
X: relation
...
V is v-precise if V is a singleton (singular)
v-imprecisiation: singular ⟶ granular
41

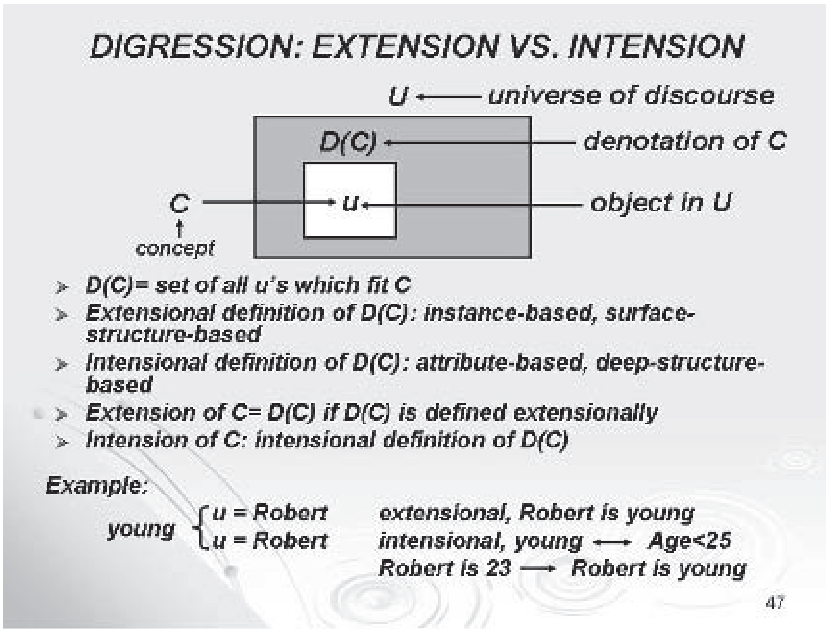

v-IMPRECISIATION
v-imprecisiation
forced
deliberate
forced: V is not known precisely
deliberate: V need not be known precisely
v-imprecisiation principle: Precision carries a cost. If there is a tolerance for imprecision, exploit it by employing v-imprecisiation to achieve lower cost, robustness, tractability, decision-relevance and higher level of confidence
42

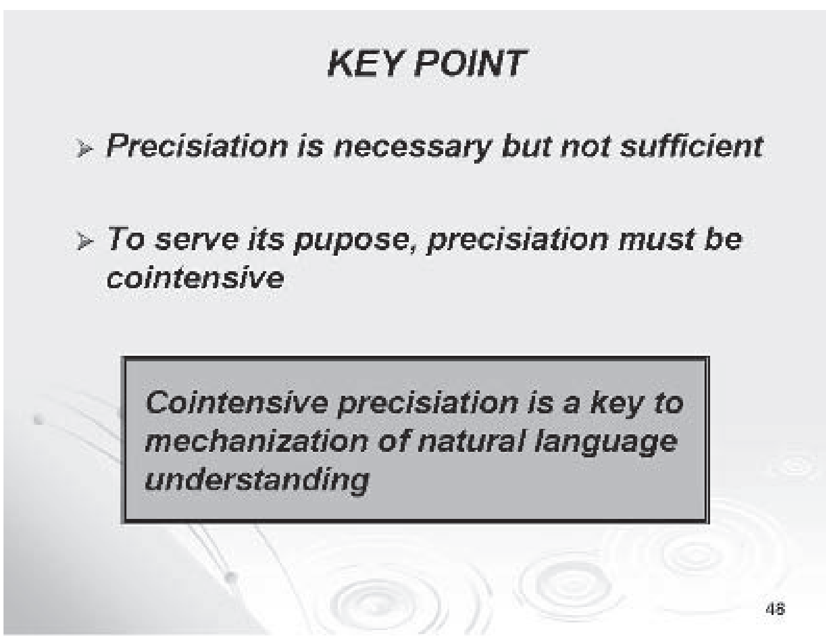

CONFIDENCE
p = X is V
value of X is uncertain
Con(p) = Prob that (X is V) is true
Generally, v-imprecisiation of V serves to increase Con(p)
Con(Carol is young) > Con(Carol is 23)
lower specificity → higher confidence
43

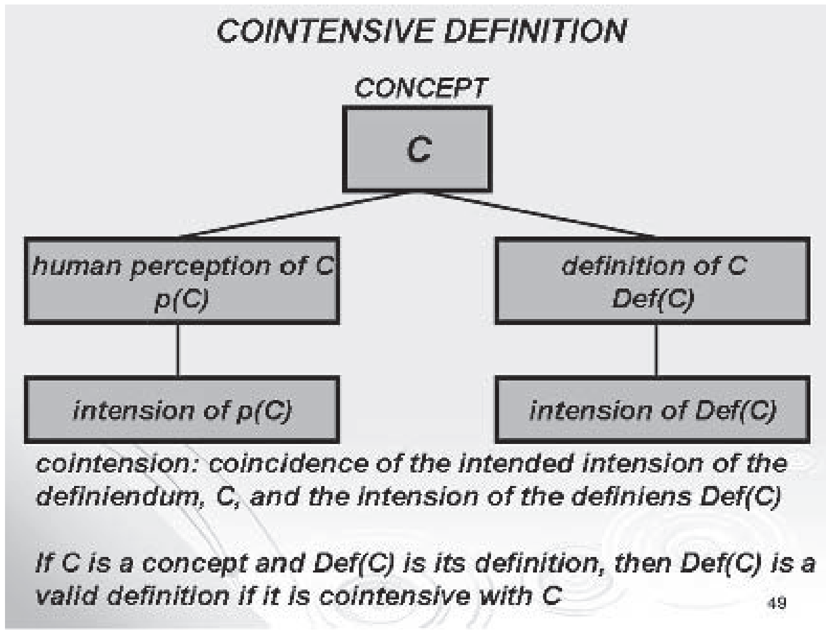

GRANULATION REVISITED
Granulation is a derivative of v-imprecisiation principle
Age
continuous
quantized
granulated
1, 2, 3, 4, 5, …
young + middle-aged + old
μ
1
0
quantized
Age
μ
1
0
young
middle -aged
old
Age
granulated
granulation = v-imprecisiation / m-precisiation
44

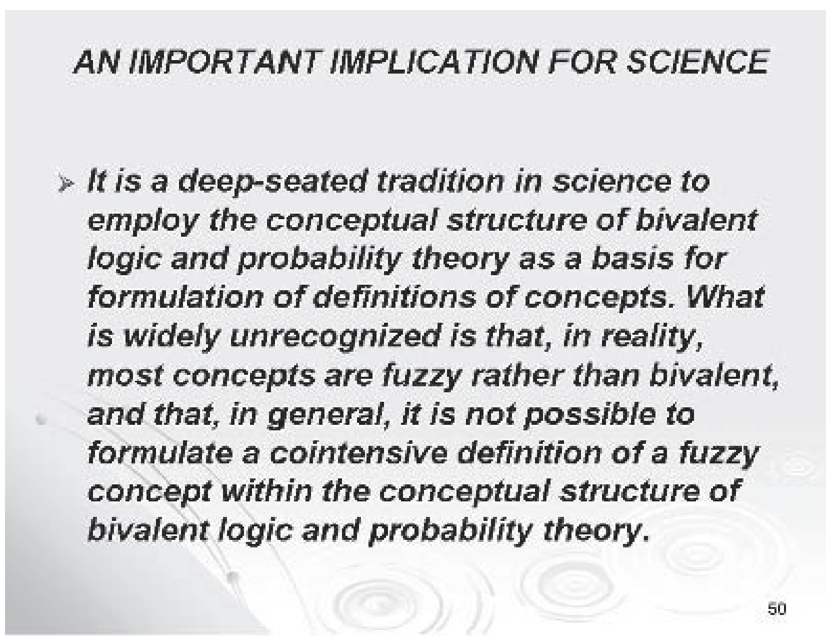

KEY POINT
Granulation plays a key role in human cognition
In human cognition, v-imprecisiation is followed by mh-precisiation. Granulation is mh-precisiation-based
In fuzzy logic, v-imprecisiation is followed by mm-precisiation. Granulation is mm-precisiation-based
mm-precisiation-based granulation is a major contribution of fuzzy logic. No other logical system offers this capability


KEY IDEAS
In TPM, a proposition, p, is precisiated by representing its meaning as a generalized constraint, GC(p)
In TPM,
precisiation = m-precisiation
precisiation of meaning does not imply precisiation of value
A desideratum of precisiation is cointension
Informally, p and q are cointensive if the intension (attribute-based-meaning) of p is approximately the same as the intension (attribute-based-meaning) of q
46


DIGRESSION: EXTENSION VS. INTENSION
U ← universe of discourse
D(C) ← denotation of C
C → u ← object in U
concept
D(C)= set of all u's which fit C
Extensional definition of D(C): instance-based, surface-structure-based
Intensional definition of D(C): attribute-based, deep-structure-based
Extension of C= D(C) if D(C) is defined extensionally
Intension of C: intensional definition of D(C)
Example:
young
u = Robert
u = Robert
extensional, Robert is young
intensional, young ⟷ Age<25
Robert is 23 ⟶ Robert is young
47

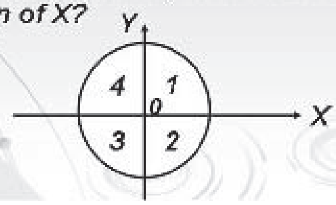

KEY POINT
Precisiation is necessary but not sufficient
To serve its pupose, precisiation must be cointensive
Cointensive precisiation is a key to mechanization of natural language understanding
48


COINTENSIVE DEFINITION
CONCEPT
C
human perception of C
p(C)
definition of C
Def(C)
intension of p(C)
intension of Def(C)
cointension: coincidence of the intended intension of the definiendum, C, and the intension of the definiens Def(C)
If C is a concept and Def(C) is its definition, then Def(C) is a valid definition if it is cointensive with C
49


AN IMPORTANT IMPLICATION FOR SCIENCE
It is a deep-seated tradition in science to employ the conceptual structure of bivalent logic and probability theory as a basis for formulation of definitions of concepts. What is widely unrecognized is that, in reality, most concepts are fuzzy rather than bivalent, and that, in general, it is not possible to formulate a cointensive definition of a fuzzy concept within the conceptual structure of bivalent logic and probability theory.
50

TEST PROBLEMS (PROBABILITY THEORY)

- *X is a real-valued random variable. What is known about X is: a(usually X is much larger than approximately a; b usually X is much smaller than approximately b, where a and b are real numbers with a < b. What is the expected value of X?*
- *X and Y are random variables. (X,Y) takes values in the unit circle. Prob(1) is approximately 0.1; Prob(2) is approximately 0.2; Prob(3) is approximately 0.3; Prob(4) is approximately 0.4. What is the marginal distribution of X?*

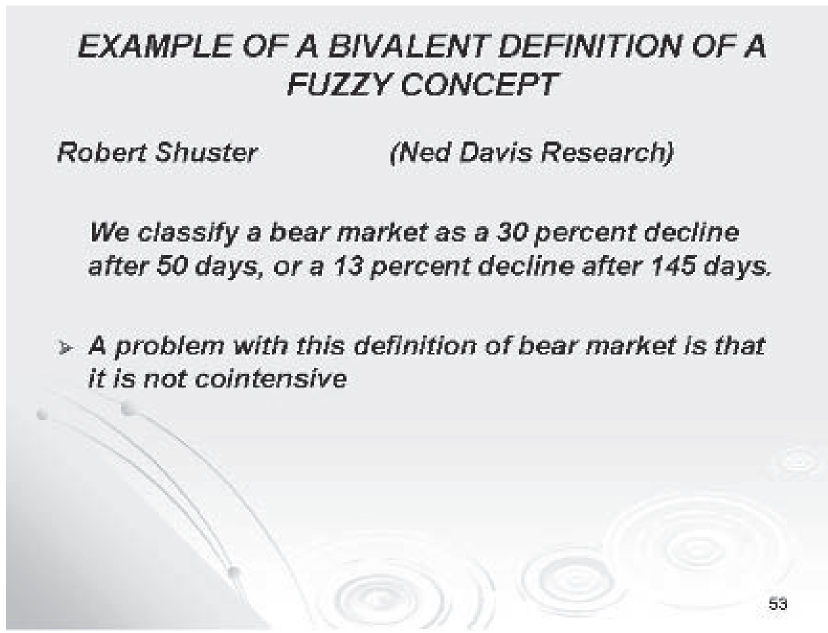


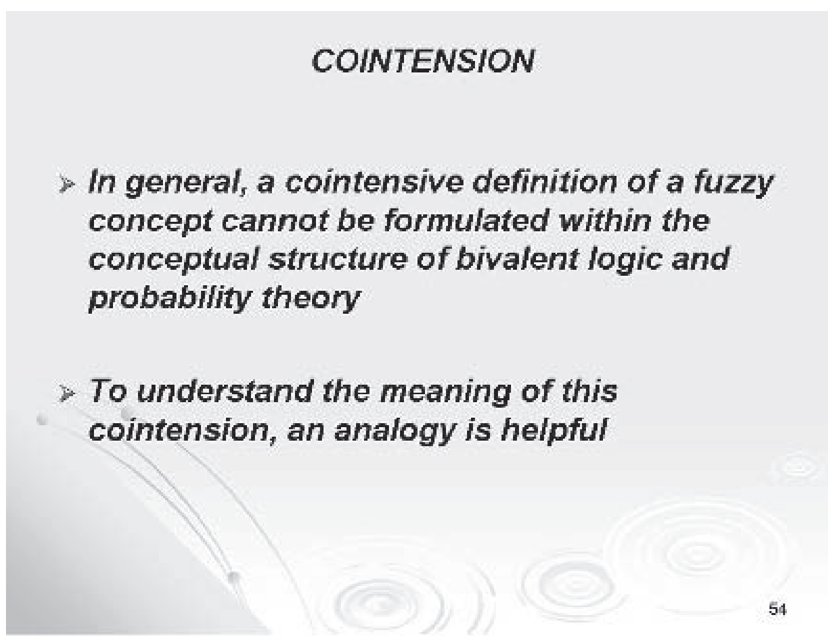
51

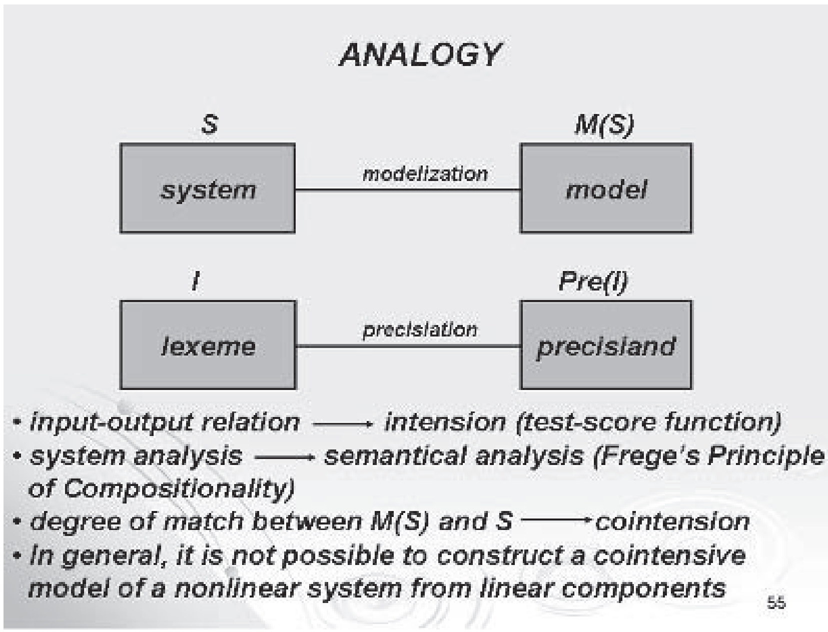
EXAMPLES OF FUZZY CONCEPTS WHOSE STANDARD, BIVALENT-LOGIC-BASED DEFINITIONS ARE NOT COINTENSIVE

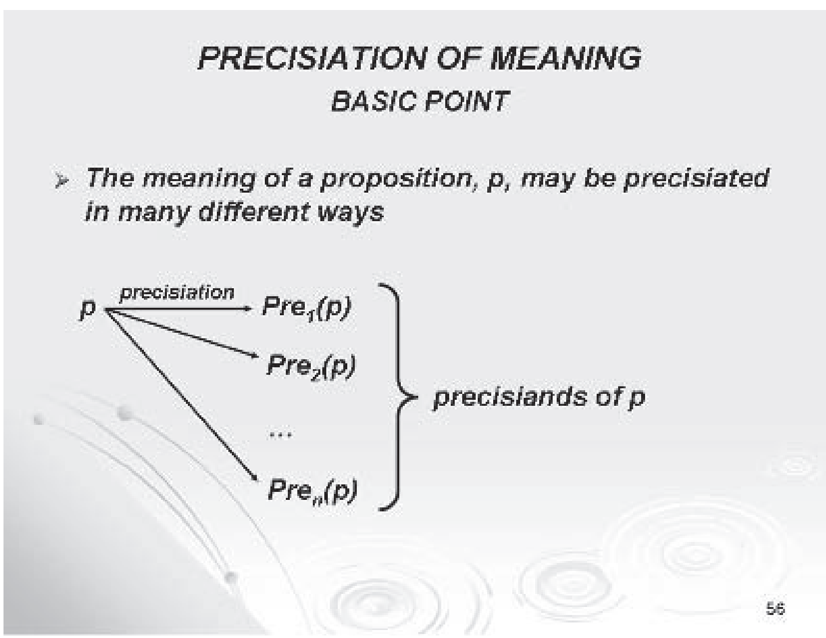
- *stability*
- *causality*
- *relevance*
- *bear market*
- *recession*
- *mountain*
- *independence*
- *stationarity*
- *cluster*

52


EXAMPLE OF A BIVALENT DEFINITION OF A FUZZY CONCEPT
Robert Shuster (Ned Davis Research)
We classify a bear market as a 30 percent decline after 50 days, or a 13 percent decline after 145 days.
A problem with this definition of bear market is that it is not cointensive
53


COINTENSION
In general, a cointensive definition of a fuzzy concept cannot be formulated within the conceptual structure of bivalent logic and probability theory
To understand the meaning of this cointension, an analogy is helpful
54


ANALOGY
S
M(S)
system
modelization
model
l
Pre(l)
lexeme
precisiation
precisiand
• input-output relation ⟶ intension (test-score function)
• system analysis ⟶ semantical analysis (Frege's Principle of Compositionality)
• degree of match between M(S) and S ⟶ cointension
• In general, it is not possible to construct a cointensive model of a nonlinear system from linear components
55


PRECISIATION OF MEANING
BASIC POINT
➢ The meaning of a proposition, p, may be precisiated in many different ways
p
precisiation
Pre1(p)
Pre2(p)
...
Pren(p)
precisiands of p
56

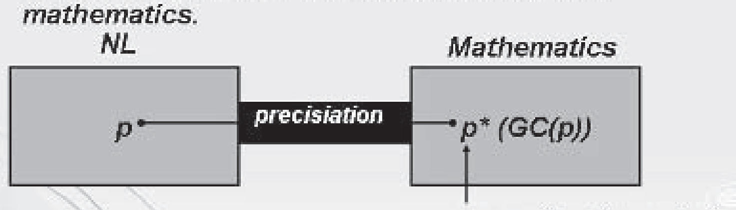
CHOICE OF PRECISIANDS

- ***In TPM, Pre(p) is equated to GC(p)***
- ***The concept of a generalized constraint opens the door to an unlimited enlargement of the number of ways in which a proposition may be precisiated***
- ***An optimal choice is one which achieves a compromise between complexity and cointension***

57

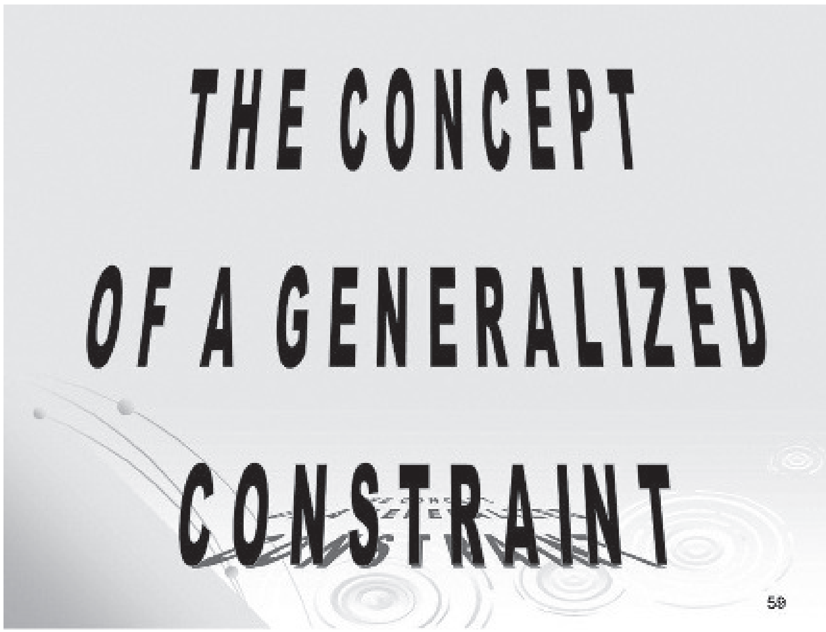
THE KEY IDEA

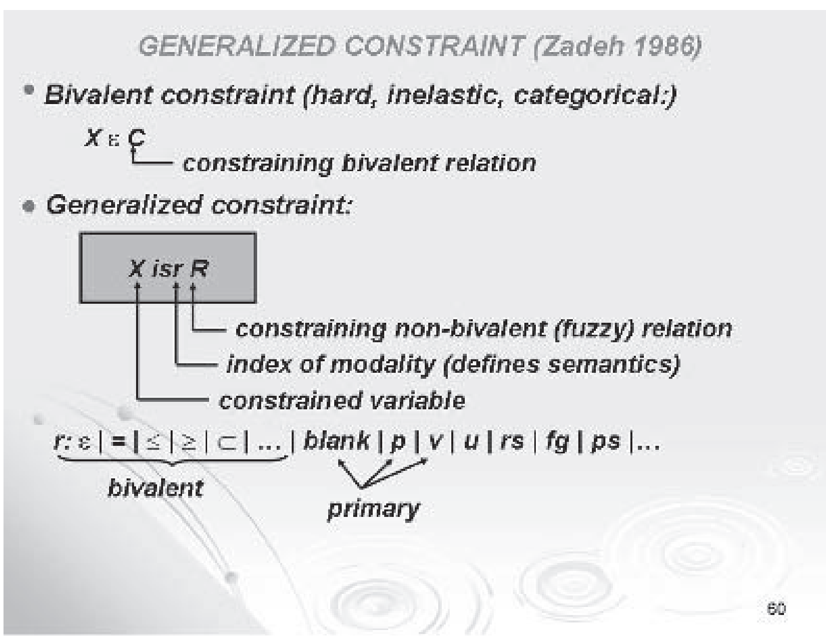
- *In TPM, a proposition, p, is precisiated by expressing its meaning as a generalized constraint. In this sense, the concept of a generalized constraint serves as a bridge from natural languages to mathematics.*

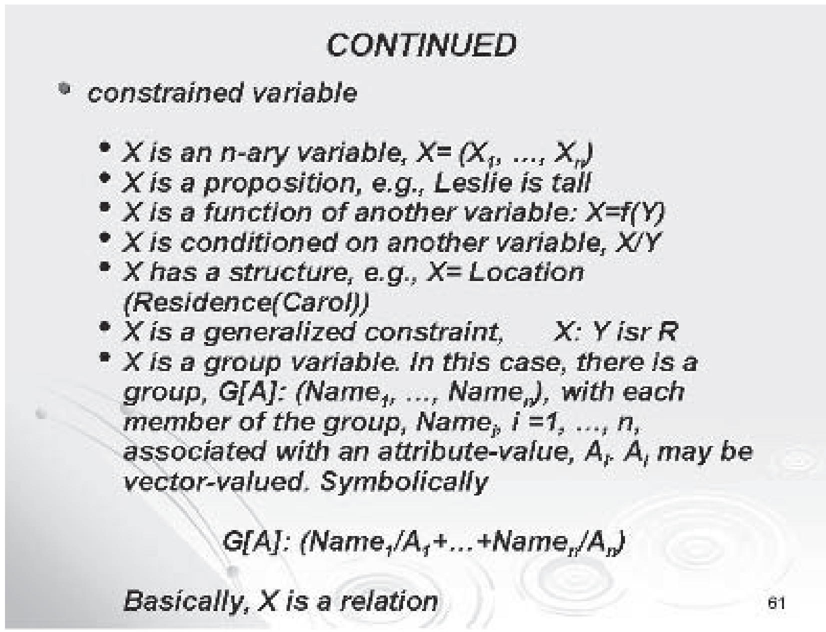


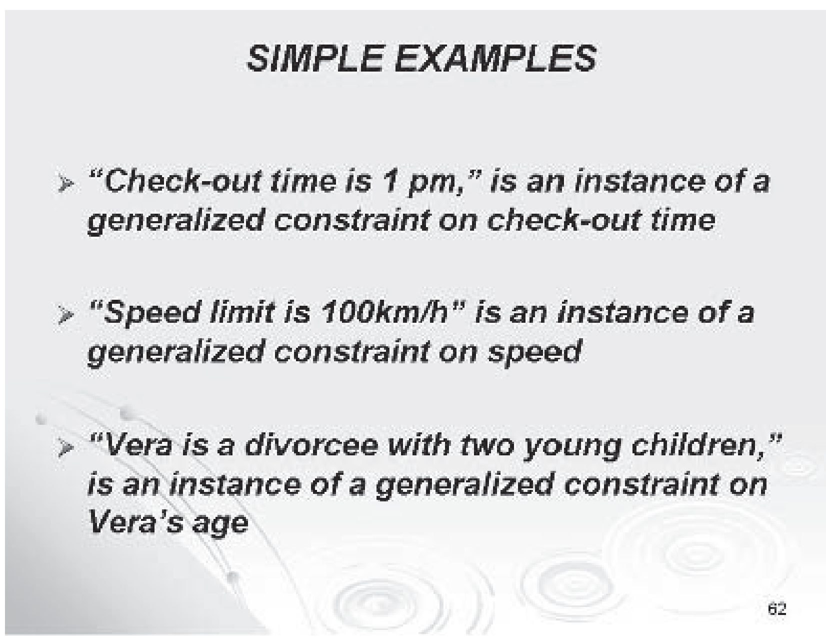
- *The concept of a generalized constraint is the centerpiece of TPM*

58


THE CONCEPT
OF A GENERALIZED
CONSTRAINT
59


GENERALIZED CONSTRAINT (Zadeh 1986)
• Bivalent constraint (hard, inelastic, categorical:)
X ε C
constraining bivalent relation
• Generalized constraint:
X isr R
constraining non-bivalent (fuzzy) relation
index of modality (defines semantics)
constrained variable
r: ε | = | ≤ | ≥ | ⊂ | ... | blank | p | v | u | rs | fg | ps |...
bivalent
primary
60


CONTINUED
constrained variable
X is an n-ary variable, X= (X1, ..., Xn)
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[A]: (Name1, ..., Namen), with each member of the group, Namei, i =1, ..., n, associated with an attribute-value, Ai. Ai may be vector-valued. Symbolically
G[A]: (Name1/A1+...+Namen/An)
Basically, X is a relation
61


SIMPLE EXAMPLES
"Check-out time is 1 pm," is an instance of a generalized constraint on check-out time
"Speed limit is 100km/h" is an instance of a generalized constraint on speed
"Vera is a divorcee with two young children," is an instance of a generalized constraint on Vera's age
62

GENERALIZED CONSTRAINT—MODALITY r

$$\boxed{X \text{ isr } R}$$

r: =	*equality constraint: X=R is abbreviation of X is=R*
r: ≤	*inequality constraint: X ≤ R*
r:⊂	*subsethood constraint: X ⊂ R*
r: blank	*possibilistic constraint; X is R; R is the possibility distribution of X*
r: v	*veristic constraint; X isv R; R is the verity distribution of X*
r: p	*probabilistic constraint; X isp R; R is the probability distribution of X*

Standard constraints: bivalent possibilistic, bivalent veristic and probabilistic

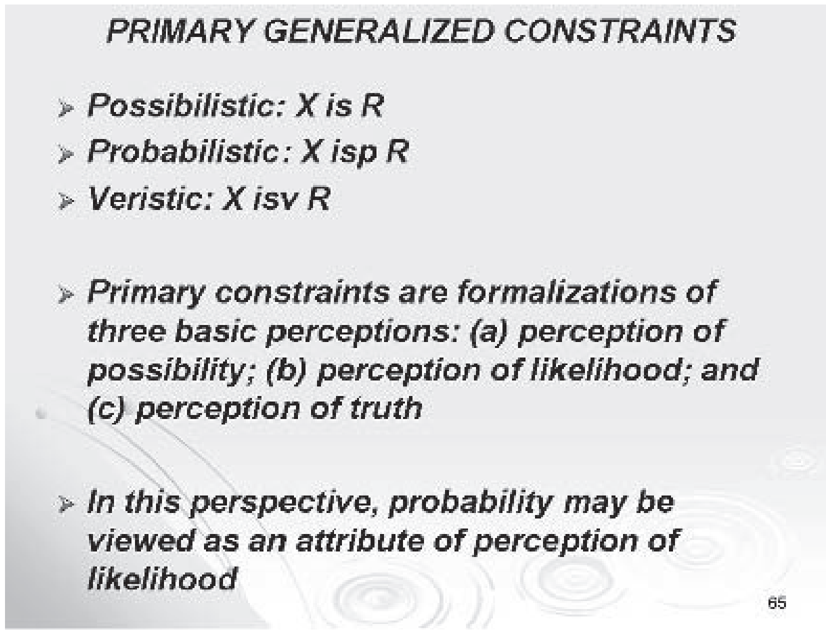
63

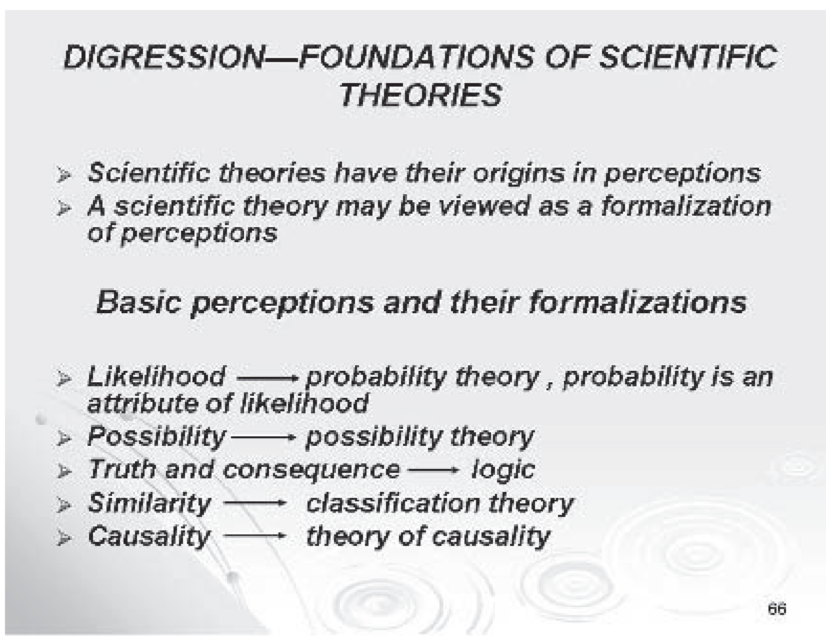
CONTINUED

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 set-valued 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*

64

Scientific and Technical Bulletin Series: Electrotechnics, Electronics, Automatic Control and Computer Science, Vol. 2, No. 2, 2005, ISSN 1584-9198





EXAMPLES: POSSIBILISTIC

- ***Monika is young*** $\longrightarrow$ ***$\underbrace{\text{Age (Monika)}}_{X}$ is $\underbrace{\text{young}}_{R}$***

- ***Monika is much younger than Maria*** $\longrightarrow$
  $\underbrace{\text{(Age (Monika), Age (Maria))}}_{X}$ is $\underbrace{\text{much younger}}_{R}$

- ***most Swedes are tall***
  $\longrightarrow$ ***$\underbrace{\Sigma\text{Count (tall.Swedes/Swedes)}}_{X}$ is $\underbrace{\text{most}}_{R}$***

67

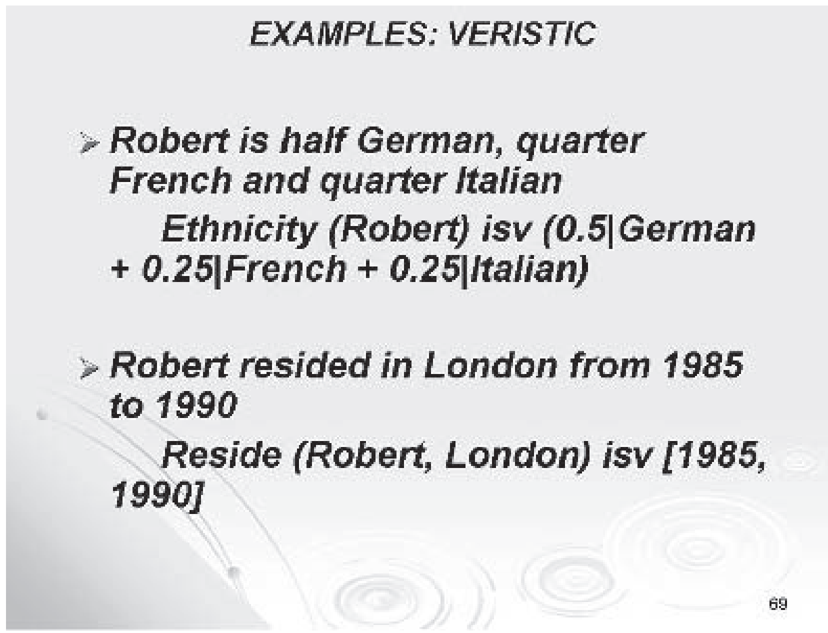
EXAMPLES: PROBABILISITIC

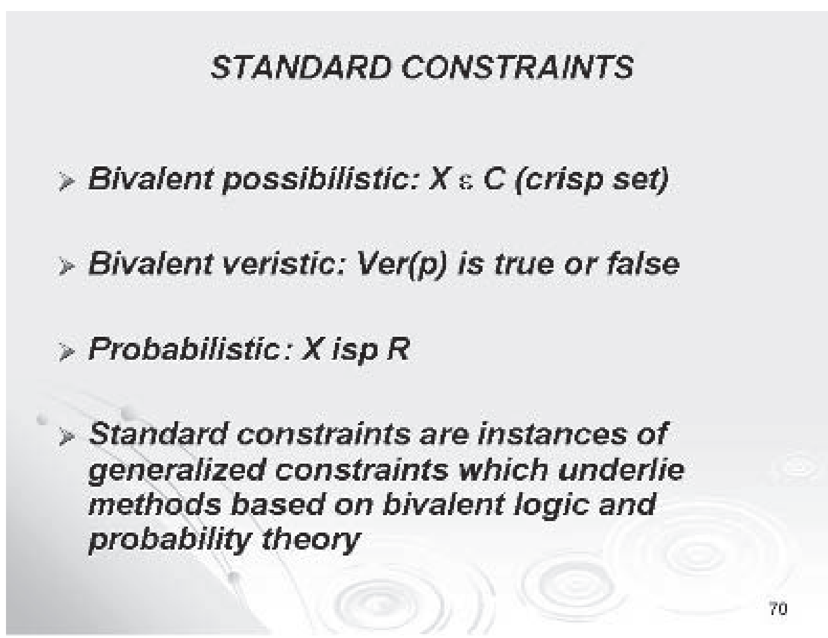
- ***X is a normally distributed random variable with mean m and variance*** σ^2 $\longrightarrow$
  X ***isp*** $N(m, \sigma^2)$

- ***X is a random variable taking the values*** u_1, u_2, u_3 ***with probabilities*** p_1, p_2 ***and*** p_3, ***respectively*** $\longrightarrow$

  X ***isp*** $(p_1\backslash u_1+p_2\backslash u_2+p_3\backslash u_3)$

68


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]
69


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
70

GENERALIZED CONSTRAINT—SEMANTICS

A generalized constraint, GC, is associated with a test-score function, ts(u), which associates with each object, u, to which the constraint is applicable, the degree to which u satisfies the constraint. Usually, ts(u) is a point in the unit interval. However, if necessary, it may be an element of a semi-ring, a lattice, or more generally, a partially ordered set, or a bimodal distribution.

example: possibilistic constraint, X is R

$$X \text{ is } R \longrightarrow Poss(X=u) = \mu_R(u)$$

$$ts(u) = \mu_R(u)$$

71

TEST-SCORE FUNCTION

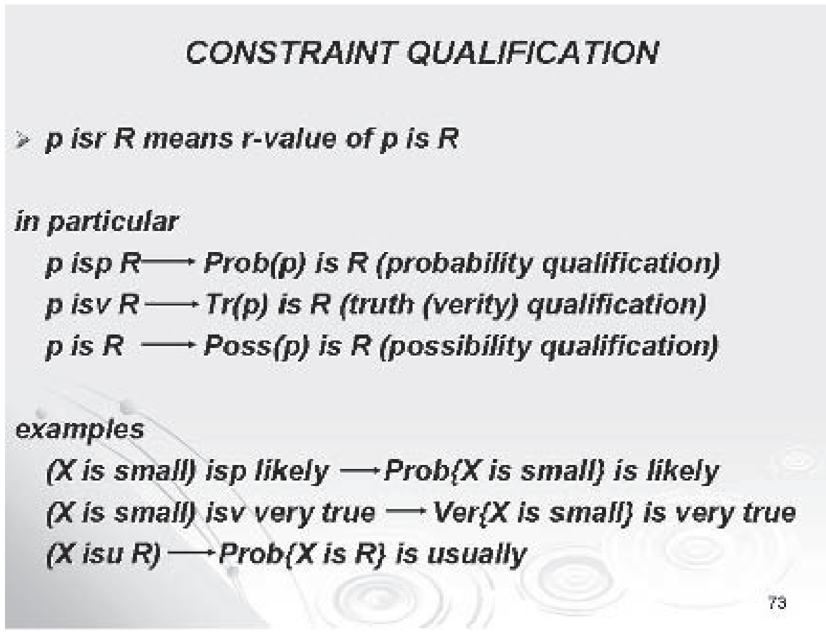
- *GC(X): generalized constraint on X*
- *X takes values in U= {u}*
- *test-score function ts(u): degree to which u satisfies GC*
- *ts(u) may be defined (a) directly (extensionally) as a function of u; or indirectly (intensionally) as a function of attributes of u*

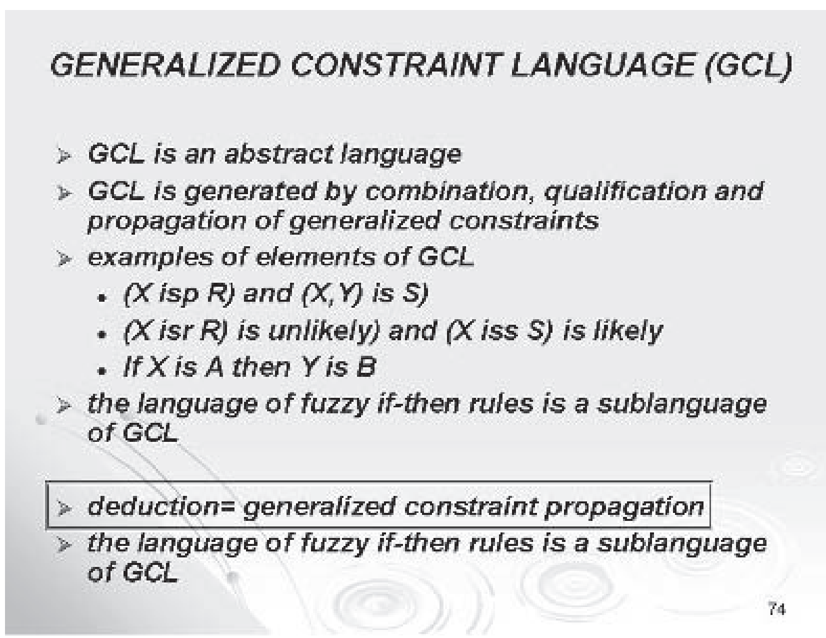
    intensional definition=attribute-based definition

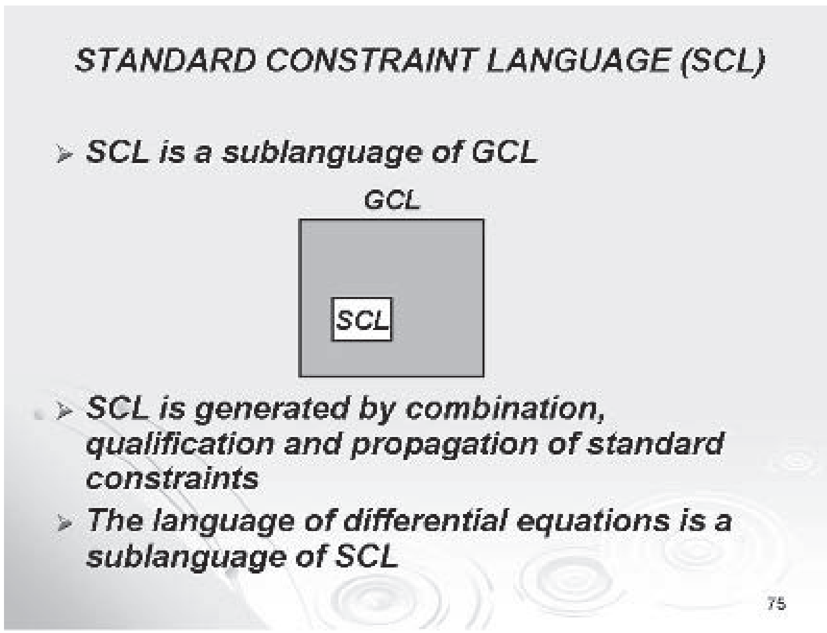
- *example (a) Andrea is tall 0.9*

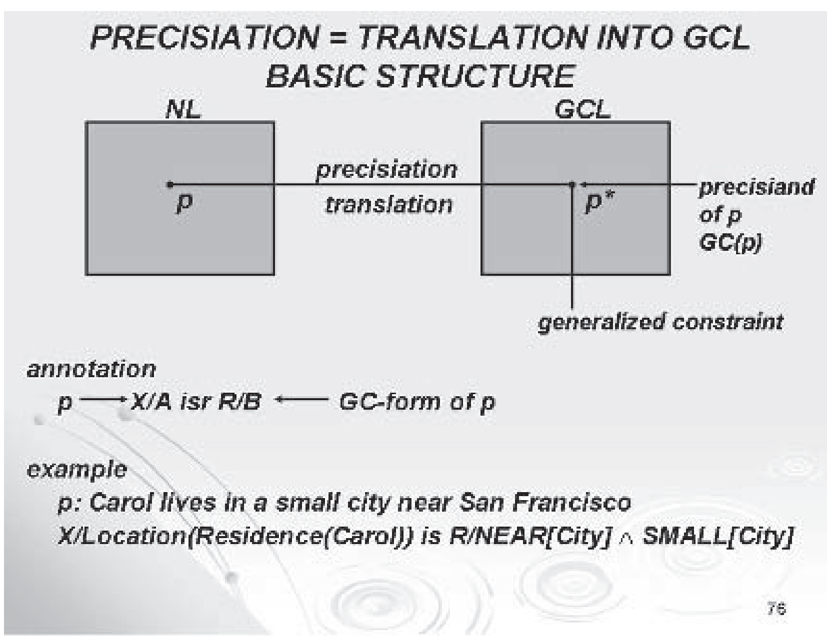
    (b) Andrea's height is 175cm; $\mu_{tall}(175)=0.9$; Andrea is 0.9 tall

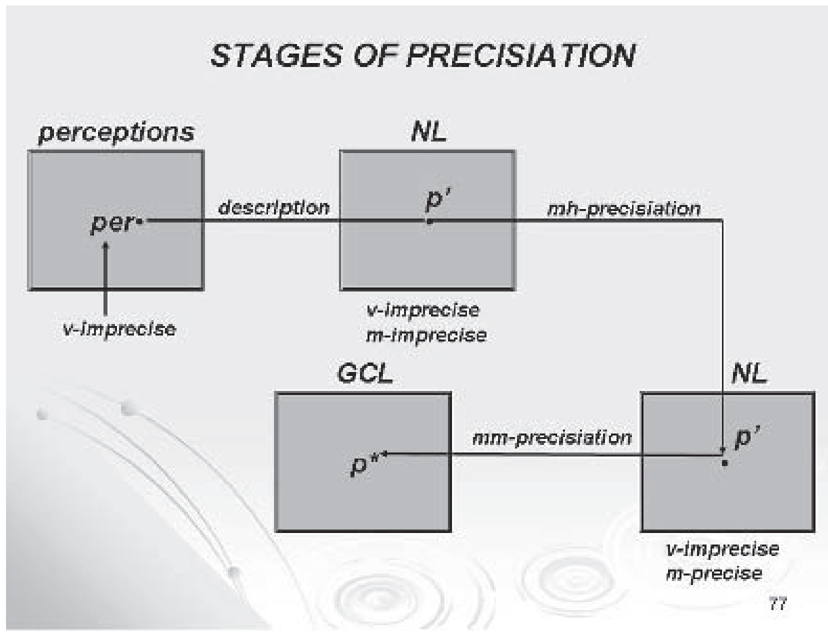
72

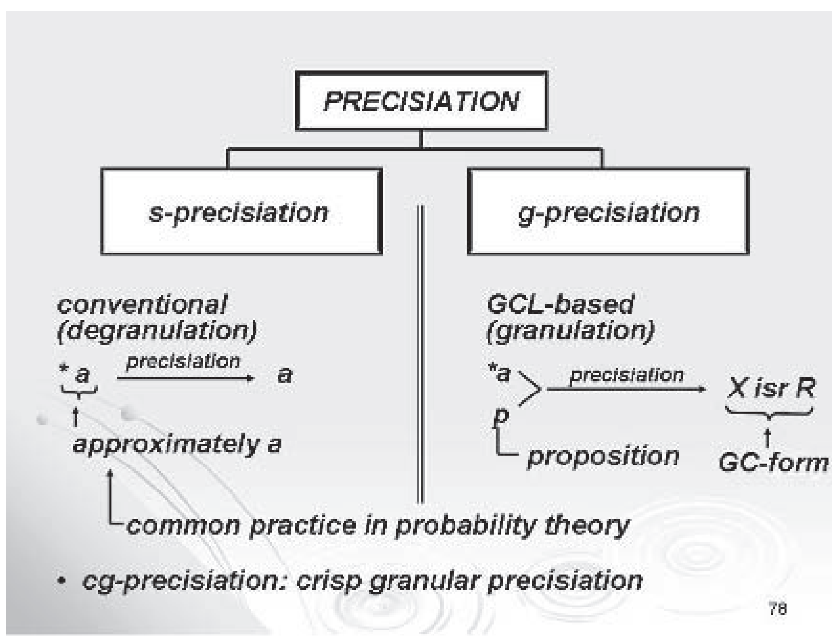

CONSTRAINT QUALIFICATION
p isr R means r-value of p is R
in particular
p isp R ⟶ Prob(p) is R (probability qualification)
p isv R ⟶ Tr(p) is R (truth (verity) qualification)
p is R ⟶ Poss(p) is R (possibility qualification)
examples
(X is small) isp likely ⟶ Prob{X is small} is likely
(X is small) isv very true ⟶ Ver{X is small} is very true
(X isu R) ⟶ Prob{X is R} is usually
73

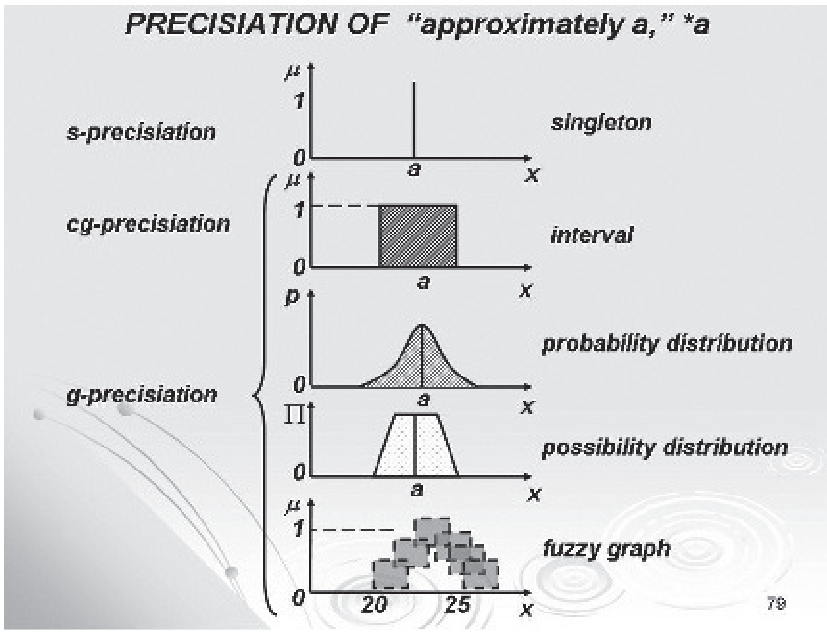

GENERALIZED CONSTRAINT LANGUAGE (GCL)
GCL is an abstract language
GCL is generated by combination, qualification and propagation of generalized constraints
examples of elements of GCL
(X isp R) and (X,Y) is S)
(X isr R) is unlikely) and (X iss S) is likely
If X is A then Y is B
the language of fuzzy if-then rules is a sublanguage of GCL
deduction= generalized constraint propagation
the language of fuzzy if-then rules is a sublanguage of GCL
74

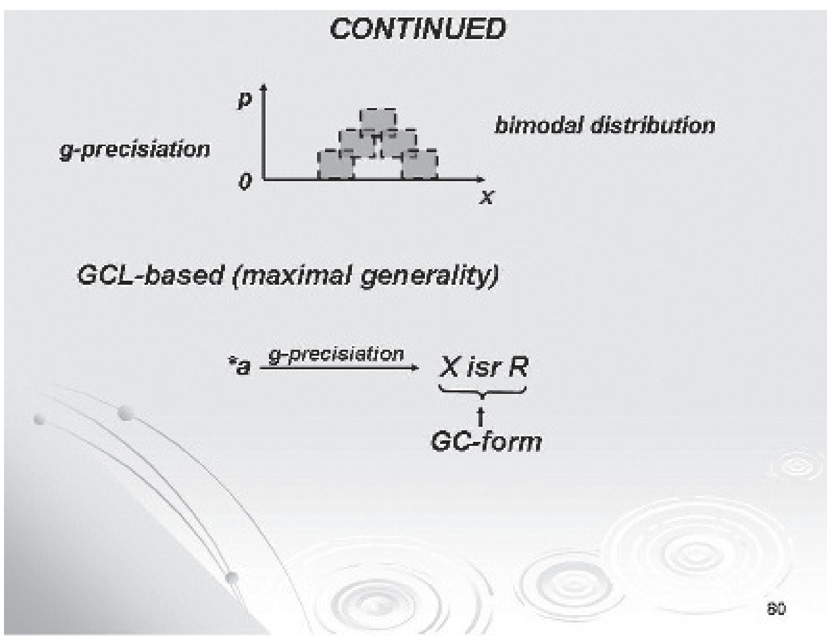

STANDARD CONSTRAINT LANGUAGE (SCL)
SCL is a sublanguage of GCL
GCL
SCL
SCL is generated by combination, qualification and propagation of standard constraints
The language of differential equations is a sublanguage of SCL
75

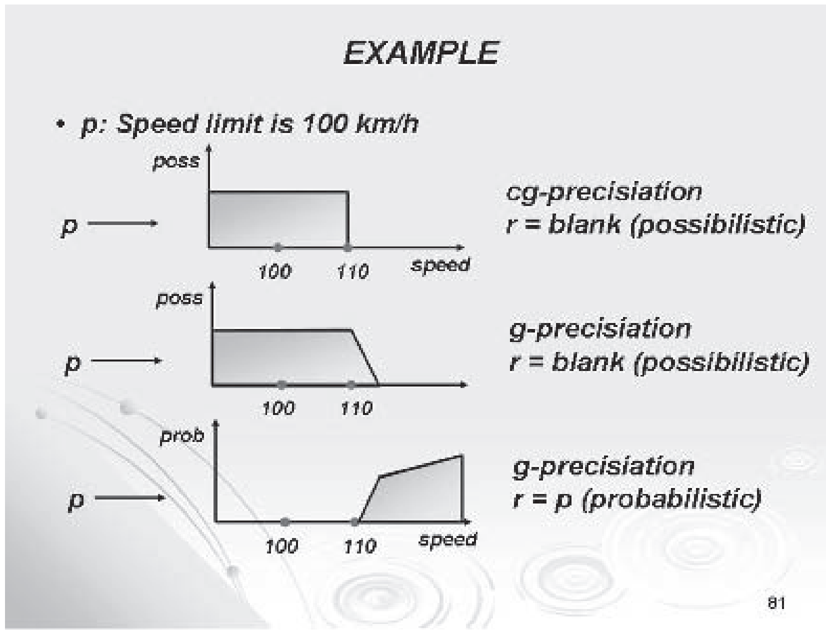

PRECISIATION = TRANSLATION INTO GCL
BASIC STRUCTURE
NL
GCL
precisiation
translation
p
p*
precisiand
of p
GC(p)
generalized constraint
annotation
p → X/A isr R/B ← GC-form of p
example
p: Carol lives in a small city near San Francisco
X/Location(Residence(Carol)) is R/NEAR[City] ∧ SMALL[City]
76

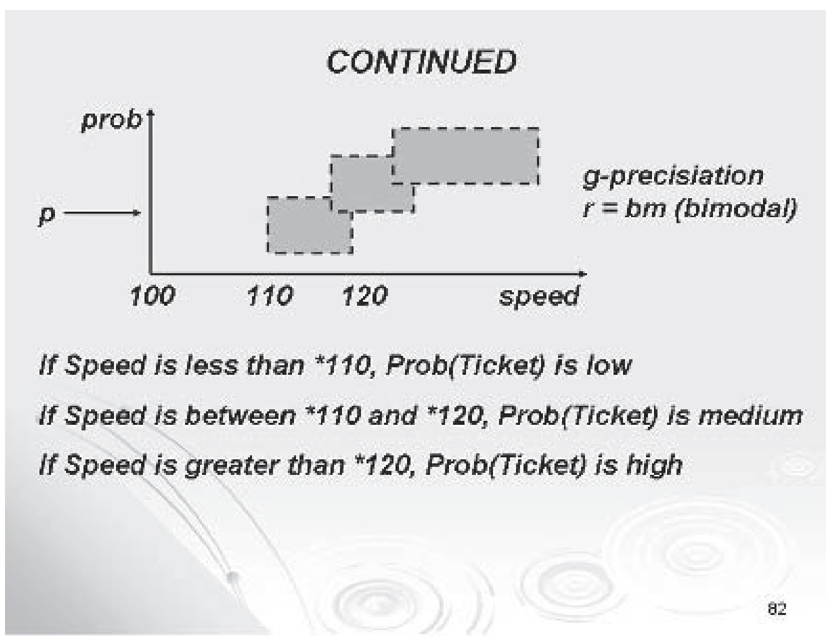

STAGES OF PRECISIATION
perceptions
NL
per
description
p'
mh-precisiation
v-imprecise
v-imprecise
m-imprecise
GCL
NL
p*
mm-precisiation
p'
v-imprecise
m-precise
77

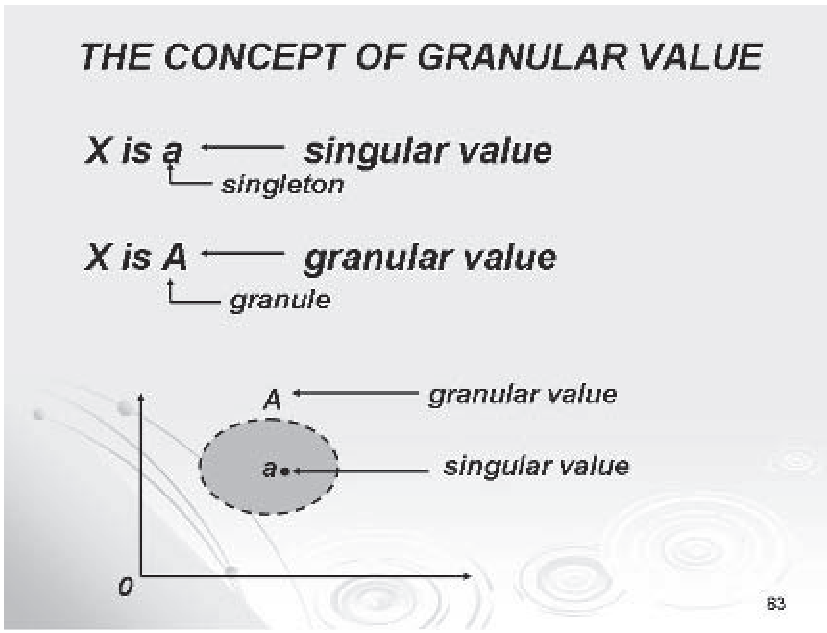

PRECISIATION
s-precisiation
g-precisiation
conventional
(degranulation)
*a
precisiation
a
approximately a
common practice in probability theory
GCL-based
(granulation)
*a
p
precisiation
X isr R
proposition
GC-form
• cg-precisiation: crisp granular precisiation
78

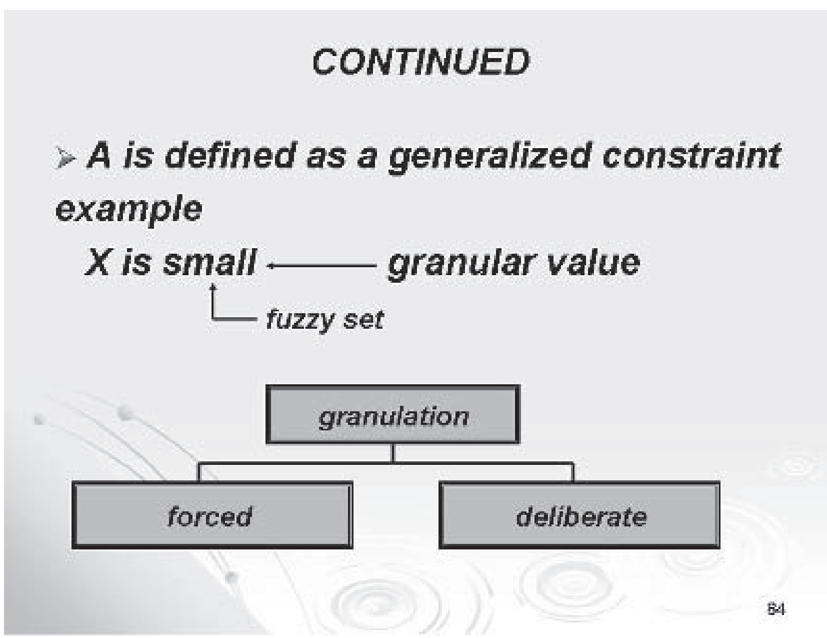

PRECISIATION OF "approximately a," *a
s-precisiation
singleton
cg-precisiation
interval
probability distribution
g-precisiation
possibility distribution
fuzzy graph
μ
1
0
a
x
p
Π
20
25
79


CONTINUED
g-precisiation
bimodal distribution
p
0
x
GCL-based (maximal generality)
*a
g-precisiation
X isr R
GC-form
80


EXAMPLE
• p: Speed limit is 100 km/h
poss
p
100
110
speed
cg-precisiation
r = blank (possibilistic)
poss
p
100
110
g-precisiation
r = blank (possibilistic)
prob
p
100
110
speed
g-precisiation
r = p (probabilistic)
81


CONTINUED
prob
p
100
110
120
speed
g-precisiation
r = bm (bimodal)
If Speed is less than *110, Prob(Ticket) is low
If Speed is between *110 and *120, Prob(Ticket) is medium
If Speed is greater than *120, Prob(Ticket) is high
82

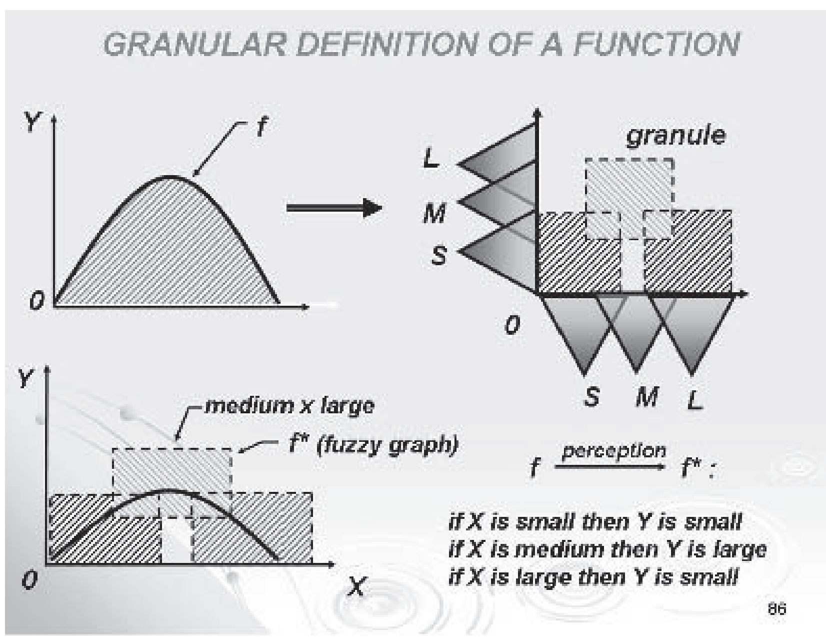

THE CONCEPT OF GRANULAR VALUE
X is a ← singular value
singleton
X is A ← granular value
granule
A ← granular value
a ← singular value
0
83


CONTINUED
A is defined as a generalized constraint
example
X is small ← granular value
fuzzy set
granulation
forced
deliberate
84

GRANULAR COMPUTING (GrC)

- *The objects of computation in granular computing are granular values of variables and parameters*
- *Granular computing has a position of centrality in fuzzy logic*
- *Granular computing plays a key role in precisiation and deduction*

85



PRECISIATION AND DEDUCTION

- ***p: most Swedes are tall***

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

further precisiation

h(u): height density function

h(u)du: fraction of Swedes whose height is in [u, u+du], $a \leq u \leq b$

$$\int_a^b h(u)du = 1$$

67

CONTINUED

- *ΣCount(tall.Swedes/Swedes) =* $\int_a^b h(u)\mu_{tall}(u)du$

- ***constraint on h***

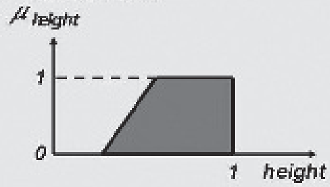
$\int_a^b h(u)\mu_{tall}(u)du$ ***is most***

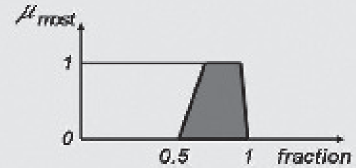
$$\pi(h) = \mu_{most}\left(\int_a^b h(u)\mu_{tall}(u)du\right)$$

68

CALIBRATION / PRECISIATION

• calibration





• precisiation

most Swedes are tall ⟶ $\pi(h) = \mu_{most}(\int_a^b h(u)\mu_{tall}(u)du)$

h: count density function

• Frege principle of compositionality—precisiated version
• precisiation of a proposition requires precisiations (calibrations) of its constituents

89

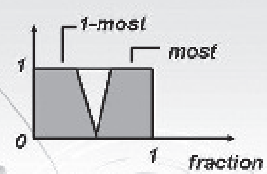
DEDUCTION

q: How many Swedes are not tall

q*: $\int_a^b h(u)\mu_{not.tall}(u)du$ is ? Q

solution: $\int_a^b h(u)(1-\mu_{tall}(u))du =$

$\int_a^b h(u)du - \int_a^b h(u)\mu_{tall}(u)du = 1 - most$



90

DEDUCTION

q: How many Swedes are short

q:* $\int_a^b h(u)\mu_{short}(u)du$ *is ? Q*

solution: $\int_a^b h(u)\mu_{tall}(u)$ *is most*

$\int_a^b h(u)\mu_{short}(u)$ *is ? Q*

extension principle

$$\mu_Q(v) = \sup_u(\mu_{most}(\int_a^b h(u)\mu_{tall}(u)du))$$

subject to

$$v = \int_a^b h(u)\mu_{short}(u)du$$

91

CONTINUED

q: What is the average height of Swedes?

q:* $\int_a^b h(u)udu$ *is ? Q*

solution: $\int_a^b h(u)\mu_{tall}(u)du$ *is most*

$\int_a^b h(u)udu$ *is ? Q*

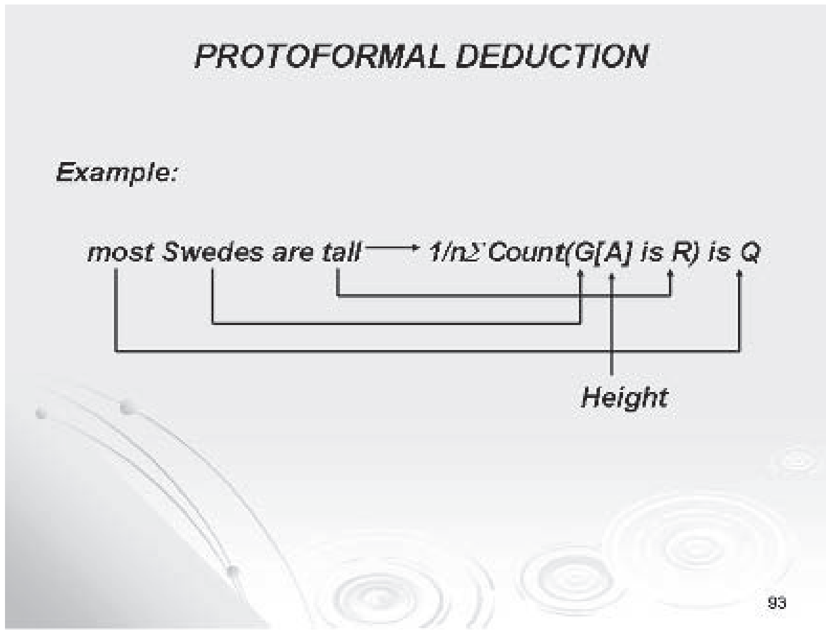
extension principle

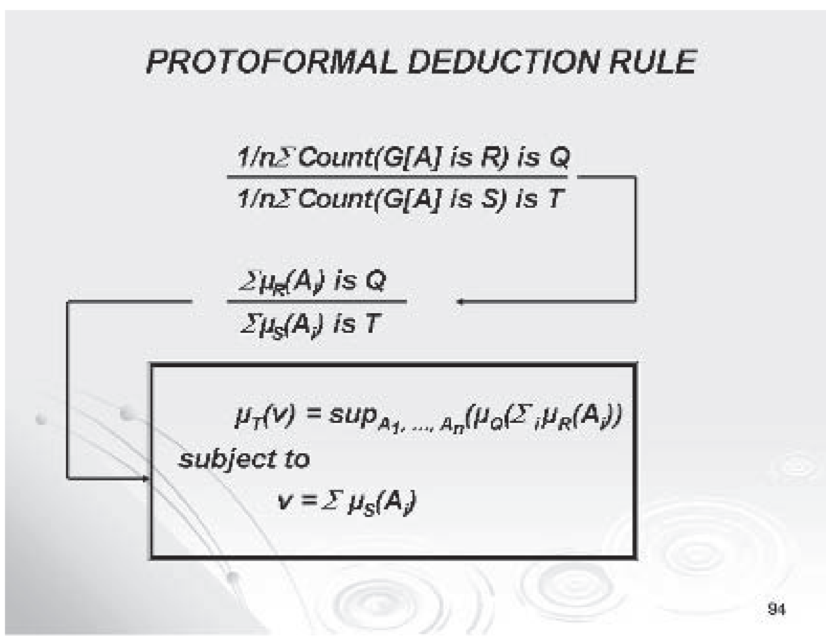
$$\mu_Q(v) = \sup_h(\mu_{most}(\int_a^b h(u)\mu_{tall}(u)du))$$

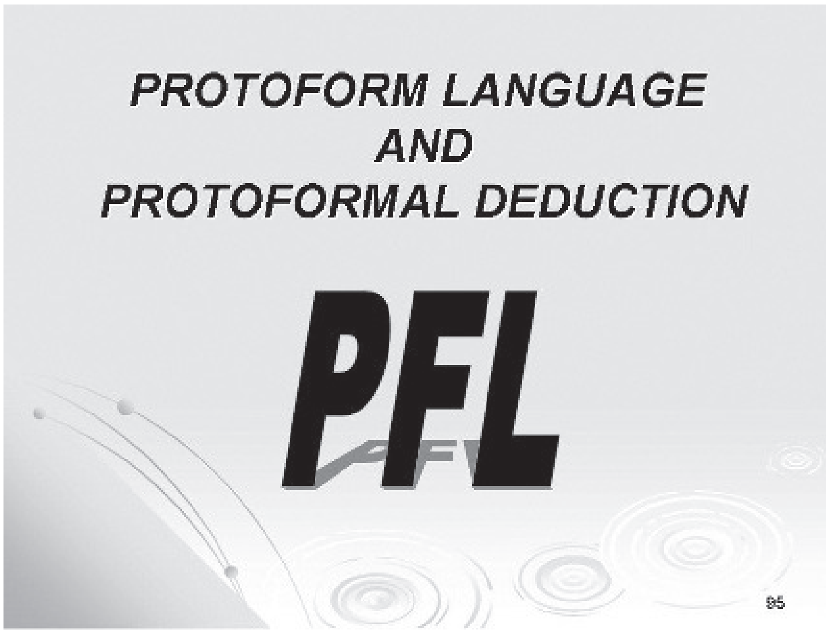
subject to

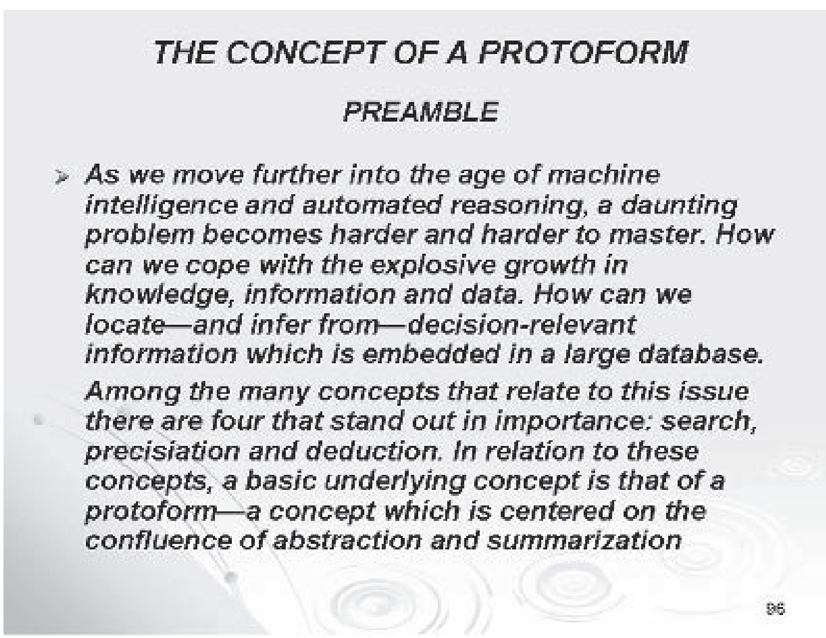
$$v = \int_a^b h(u)udu$$

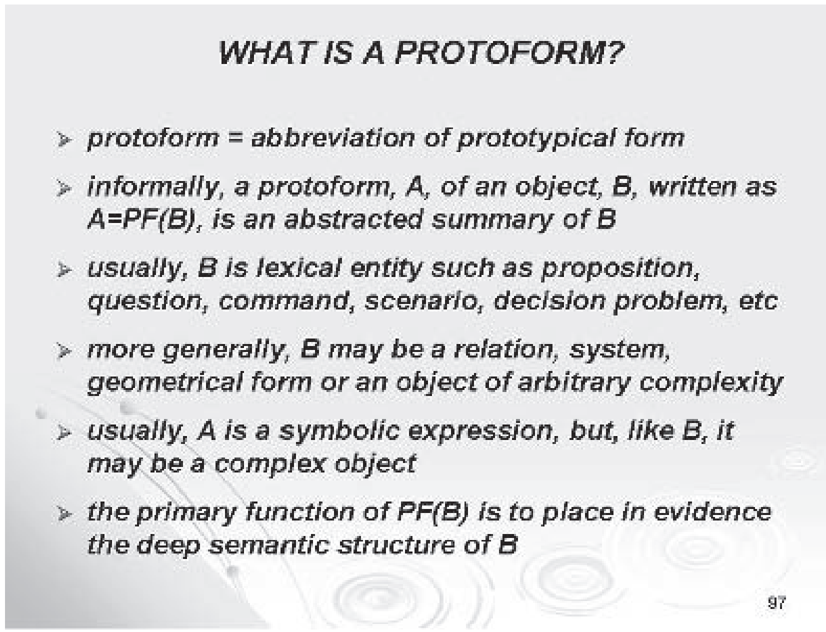
92

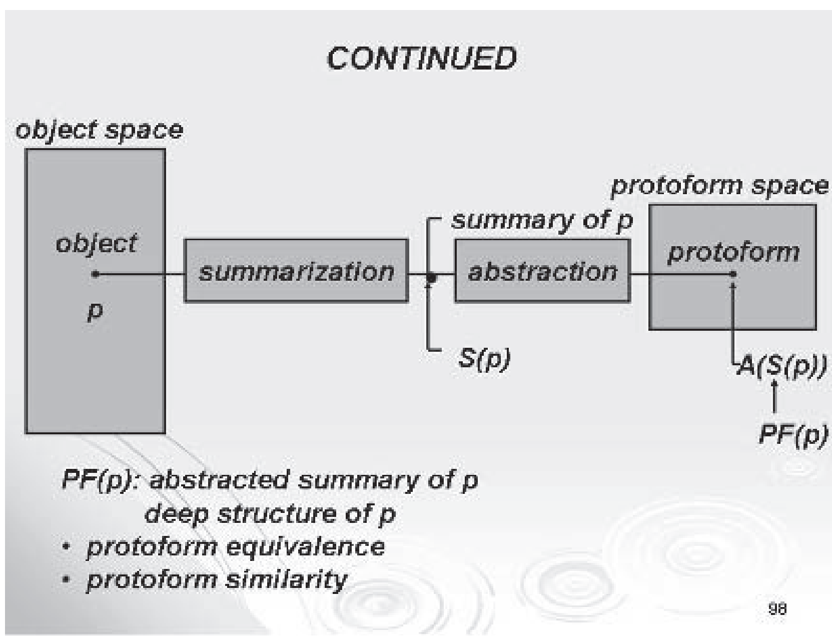

PROTOFORMAL DEDUCTION
Example:
most Swedes are tall
1/nΣCount(G[A] is R) is Q
Height
93

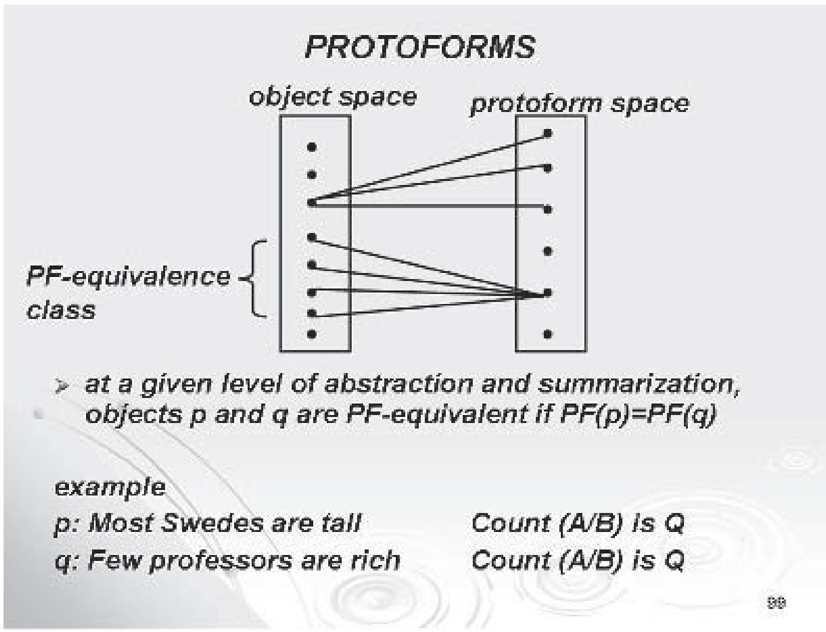

PROTOFORMAL DEDUCTION RULE
1/nΣCount(G[A] is R) is Q
1/nΣCount(G[A] is S) is T
$\Sigma\mu_R(A_i)$ is Q
$\Sigma\mu_S(A_i)$ is T
$\mu_T(v) = \sup_{A_1, \ldots, A_n}(\mu_Q(\Sigma_i \mu_R(A_i))$
subject to
$v = \Sigma\, \mu_S(A_i)$
94

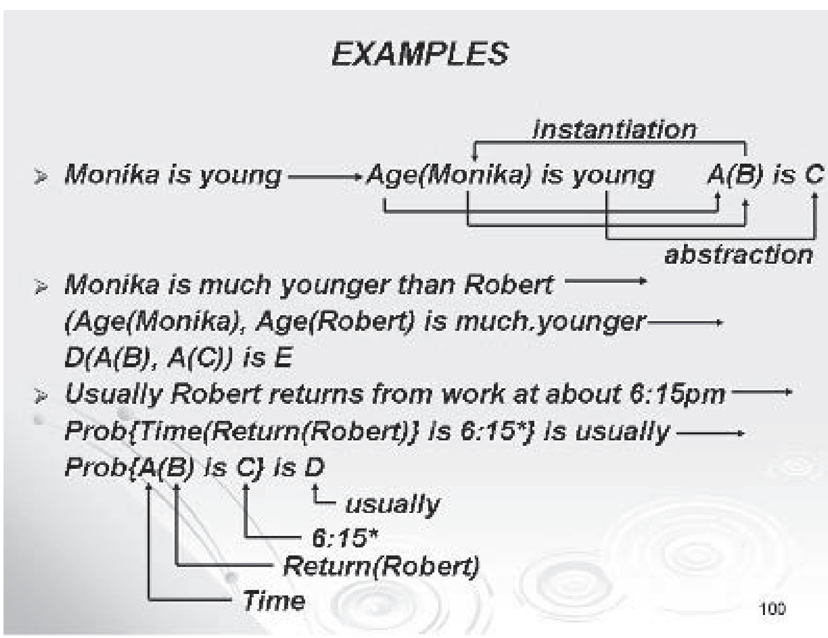

PROTOFORM LANGUAGE
AND
PROTOFORMAL DEDUCTION
PFL
95

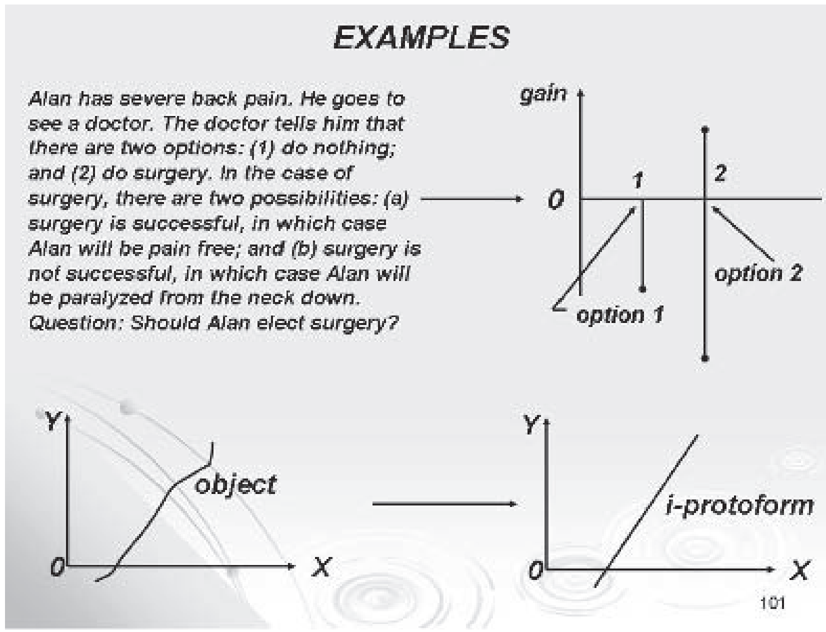

THE CONCEPT OF A PROTOFORM
PREAMBLE
➢ As we move further into the age of machine intelligence and automated reasoning, a daunting problem becomes harder and harder to master. How can we cope with the explosive growth in knowledge, information and data. How can we locate—and infer from—decision-relevant information which is embedded in a large database.
Among the many concepts that relate to this issue there are four that stand out in importance: search, precisiation and deduction. In relation to these concepts, a basic underlying concept is that of a protoform—a concept which is centered on the confluence of abstraction and summarization
96

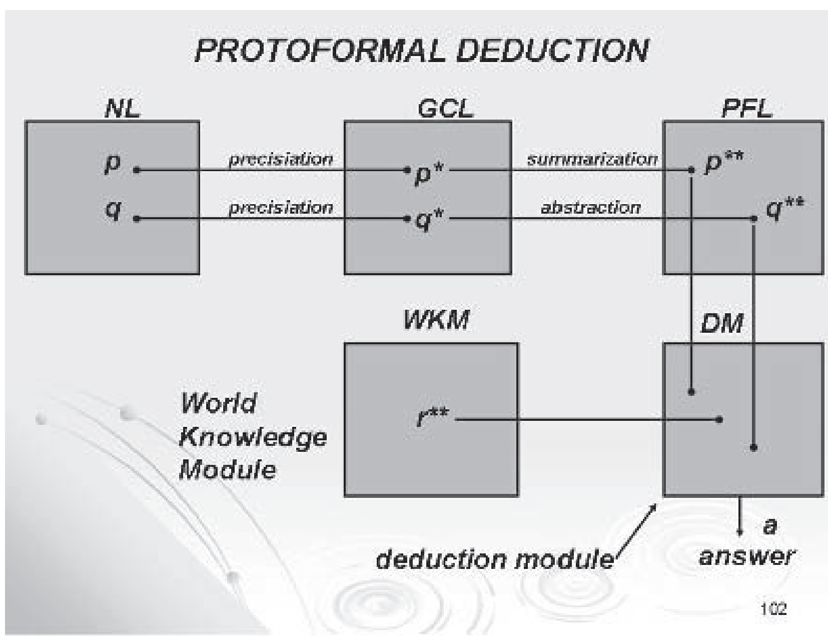

WHAT IS A PROTOFORM?
protoform = abbreviation of prototypical form
informally, a protoform, A, of an object, B, written as A=PF(B), is an abstracted summary of B
usually, B is lexical entity such as proposition, question, command, scenario, decision problem, etc
more generally, B may be a relation, system, 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 PF(B) is to place in evidence the deep semantic structure of B
97

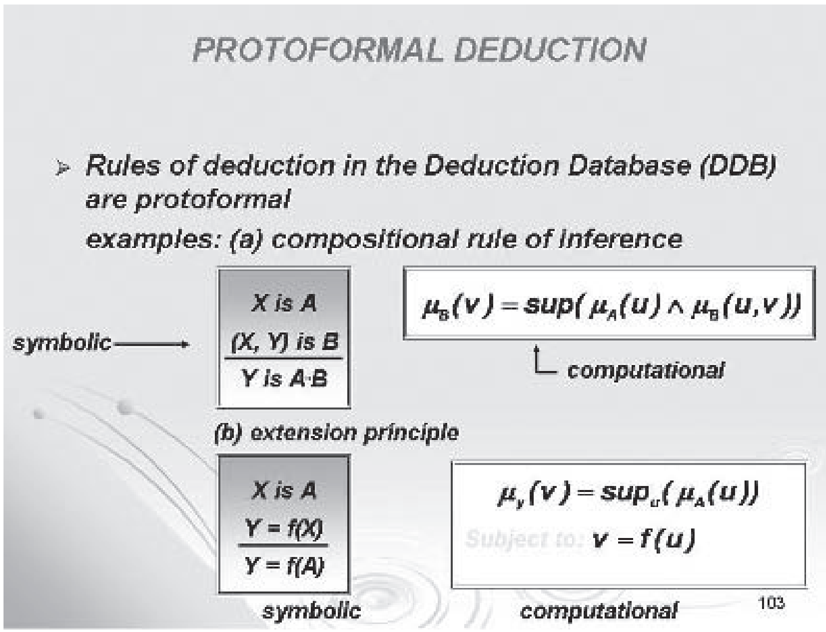

CONTINUED
object space
object
p
summarization
summary of p
S(p)
abstraction
protoform space
protoform
A(S(p))
PF(p)
PF(p): abstracted summary of p
deep structure of p
• protoform equivalence
• protoform similarity
98

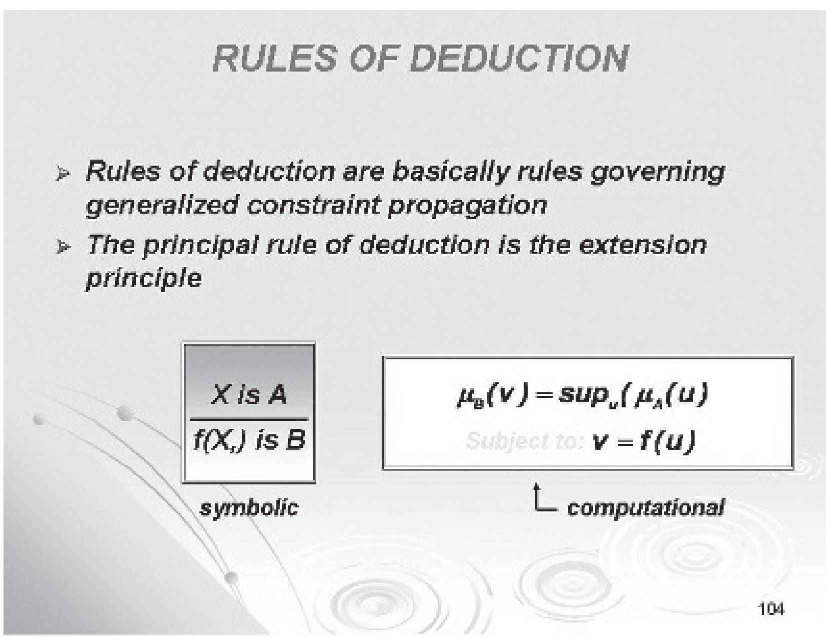

PROTOFORMS
object space
protoform space
PF-equivalence class
at a given level of abstraction and summarization, objects p and q are PF-equivalent if PF(p)=PF(q)
example
p: Most Swedes are tall
Count (A/B) is Q
q: Few professors are rich
Count (A/B) is Q
99

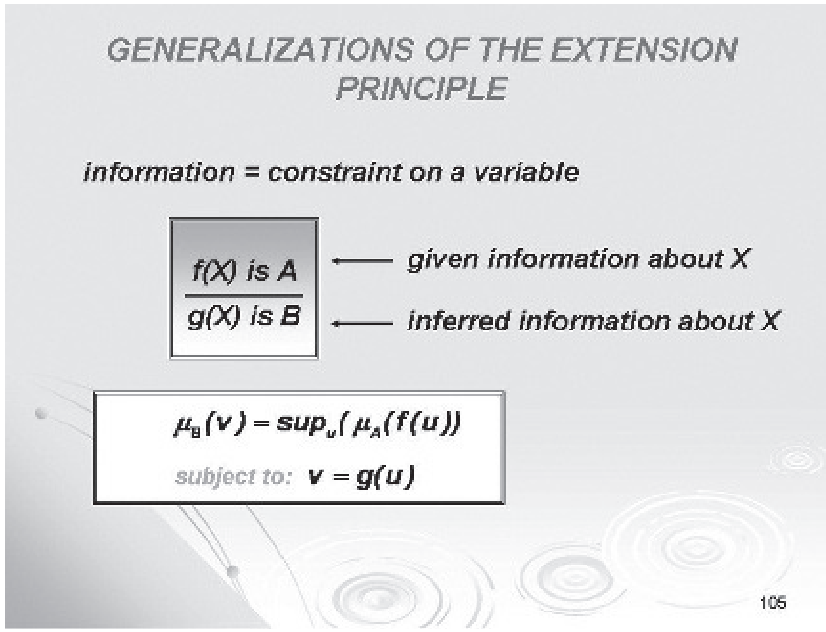

EXAMPLES
instantiation
Monika is young → Age(Monika) is young
A(B) is C
abstraction
Monika is much younger than Robert →
(Age(Monika), Age(Robert) is much.younger →
D(A(B), A(C)) is E
Usually Robert returns from work at about 6:15pm →
Prob{Time(Return(Robert)} is 6:15*} is usually →
Prob{A(B) is C} is D
usually
6:15*
Return(Robert)
Time
100

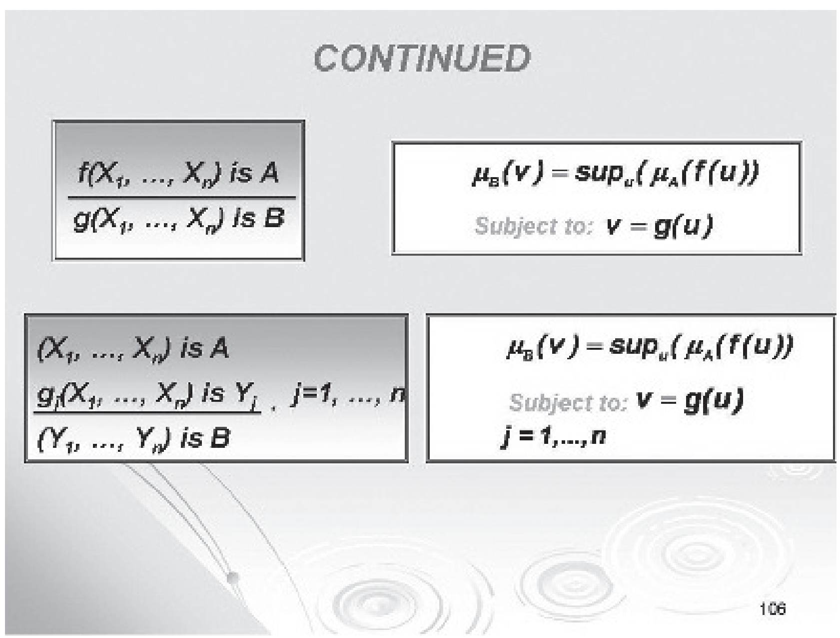

EXAMPLES
Alan has severe back pain. He goes to see a doctor. The doctor tells him that there are two options: (1) do nothing; and (2) do surgery. In the case of surgery, there are two possibilities: (a) surgery is successful, in which case Alan will be pain free; and (b) surgery is not successful, in which case Alan will be paralyzed from the neck down. Question: Should Alan elect surgery?
gain
0
1
2
option 1
option 2
Y
X
0
object
i-protoform
101

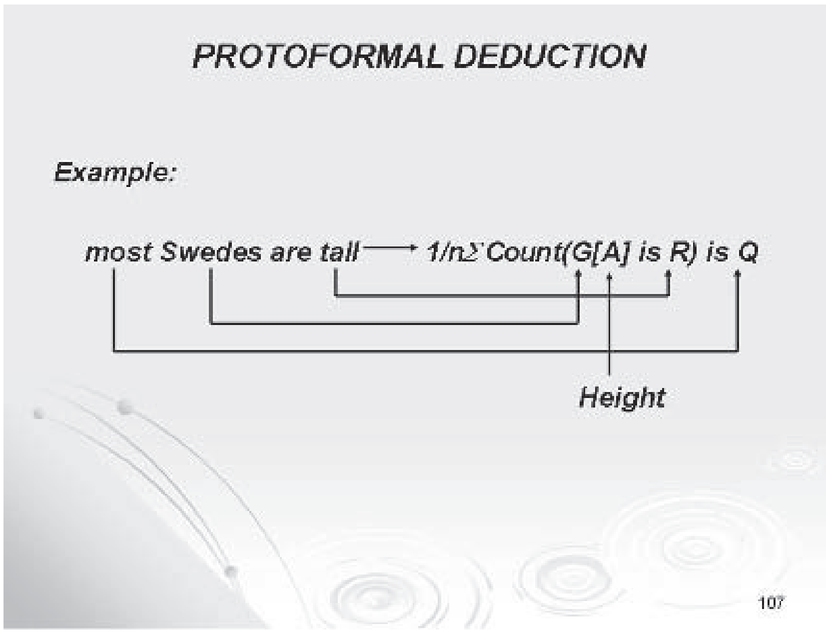

PROTOFORMAL DEDUCTION
NL
GCL
PFL
p
q
precisiation
precisiation
p*
q*
summarization
abstraction
p**
q**
WKM
DM
World Knowledge Module
r**
deduction module
a
answer
102

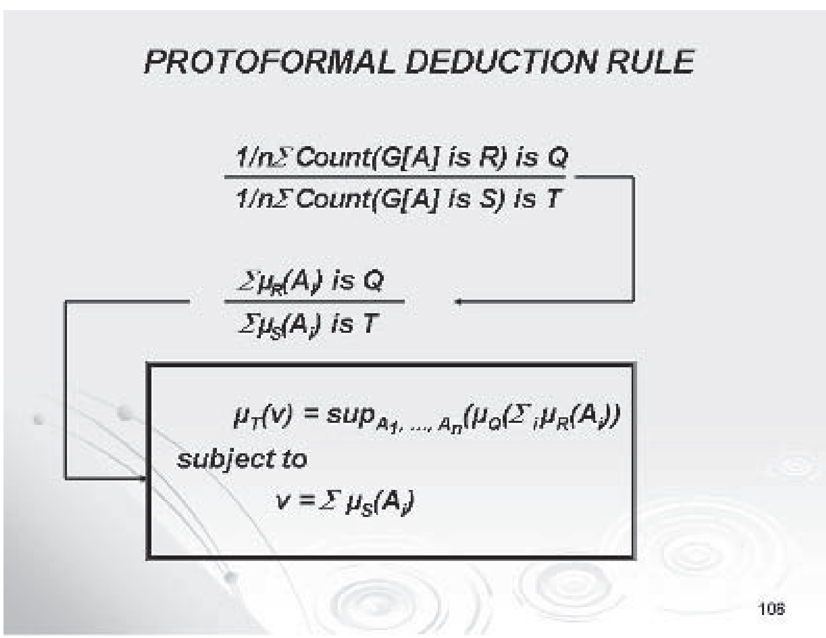

PROTOFORMAL DEDUCTION
Rules of deduction in the Deduction Database (DDB) are protoformal
examples: (a) compositional rule of inference
symbolic
X is A
(X, Y) is B
Y is A·B
$\mu_B(v) = sup(\mu_A(u) \wedge \mu_B(u,v))$
computational
(b) extension principle
X is A
Y = f(X)
Y = f(A)
$\mu_Y(v) = sup_u(\mu_A(u))$
Subject to: $v = f(u)$
symbolic
computational
103


RULES OF DEDUCTION
Rules of deduction are basically rules governing generalized constraint propagation
The principal rule of deduction is the extension principle
X is A
f(X,) is B
$\mu_B(v) = sup_u(\mu_A(u)$
Subject to: $v = f(u)$
symbolic
computational
104


GENERALIZATIONS OF THE EXTENSION PRINCIPLE
information = constraint on a variable
f(X) is A
g(X) is B
given information about X
inferred information about X
$\mu_B(v) = sup_u(\mu_A(f(u))$
subject to: $v = g(u)$
105


CONTINUED
$f(X_1, \ldots, X_n)$ is A
$g(X_1, \ldots, X_n)$ is B
$\mu_B(v) = sup_u(\mu_A(f(u))$
Subject to: $v = g(u)$
$(X_1, \ldots, X_n)$ is A
$g_j(X_1, \ldots, X_n)$ is Y_j , j=1, ..., n
$(Y_1, \ldots, Y_n)$ is B
$\mu_B(v) = sup_u(\mu_A(f(u))$
Subject to: $v = g(u)$
j = 1,...,n
106


PROTOFORMAL DEDUCTION
Example:
most Swedes are tall → 1/nΣCount(G[A] is R) is Q
Height
107


PROTOFORMAL DEDUCTION RULE
1/nΣCount(G[A] is R) is Q
1/nΣCount(G[A] is S) is T
$\Sigma\mu_R(A_i)$ is Q
$\Sigma\mu_S(A_i)$ is T
$\mu_T(v) = \sup_{A_1, \ldots, A_n}(\mu_Q(\Sigma_i \mu_R(A_i))$
subject to
$v = \Sigma\, \mu_S(A_i)$
108

EXAMPLE OF DEDUCTION

p: Most Swedes are much taller than most Italians

q: What is the difference in the average height of Swedes and Italians?

TPM-based solution

Step 1. precisiation: translation of p into GCL

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

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

g_i = height of S_i , $g = (g_1, \ldots, g_n)$

h_j = height of I_j , $h = (h_1, \ldots, h_n)$

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

109

CONTINUED

$r_i = \frac{1}{n}\Sigma_j\mu_{ij}$ = Relative ΣCount of Italians in relation to whom S_i is much taller

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

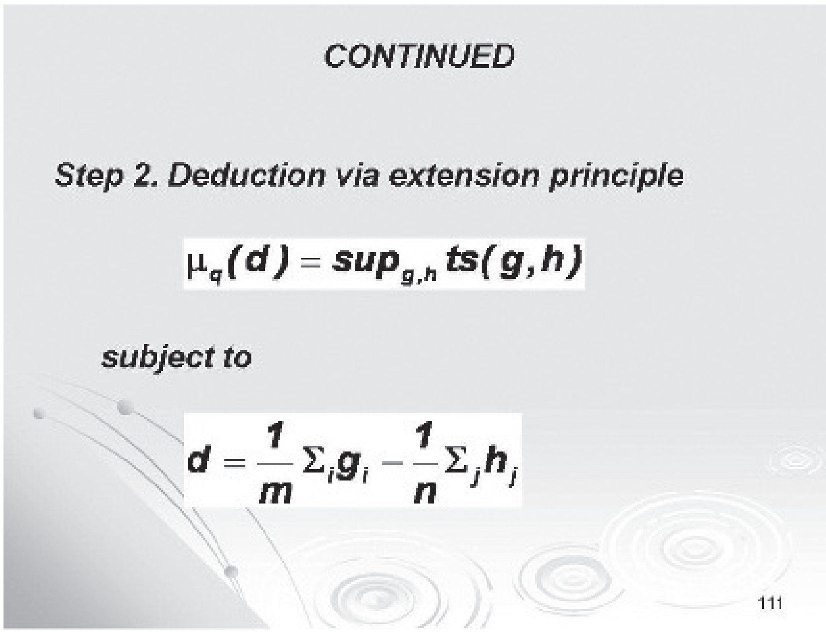
$v = \frac{1}{m}\Sigma t_i$ = Relative ΣCount of Swedes who are much taller than most Italians

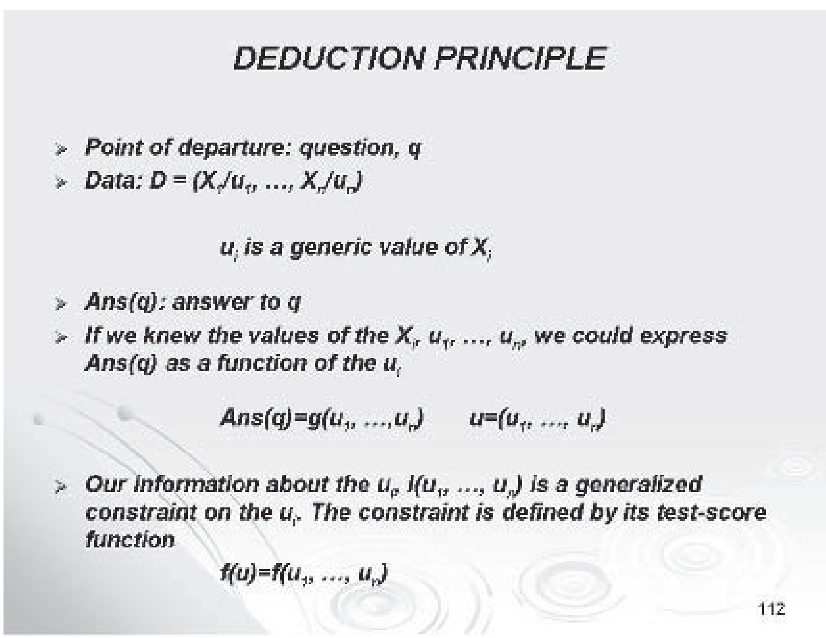
$ts(g, h) = \mu_{most}(v)$

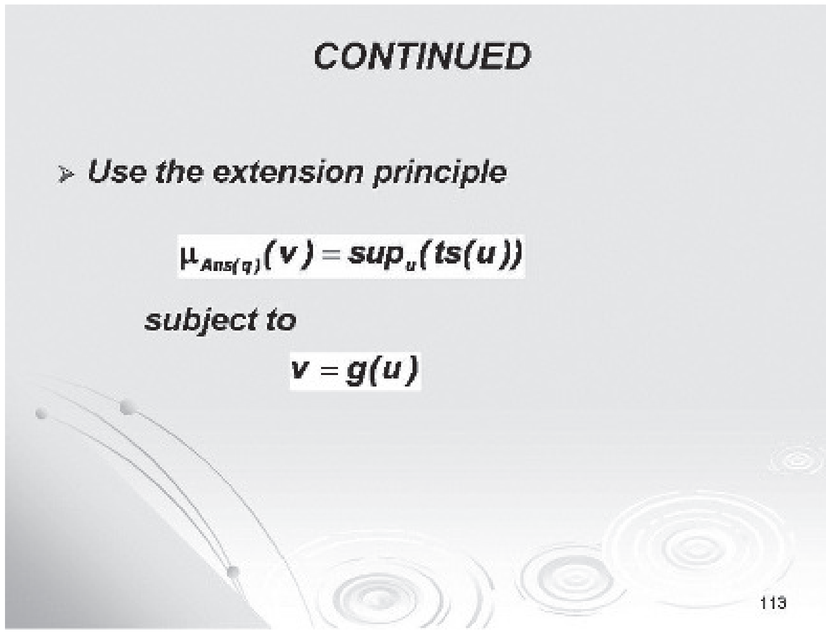
$p \longrightarrow$ generalized constraint on S and I

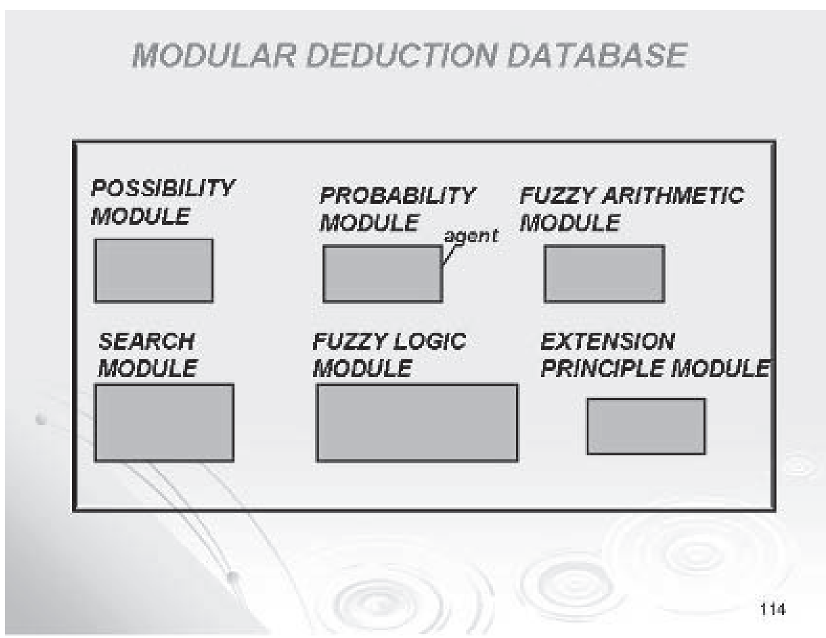
q: $d = \frac{1}{m}\Sigma_i g_i - \frac{1}{n}\Sigma_j h_j$

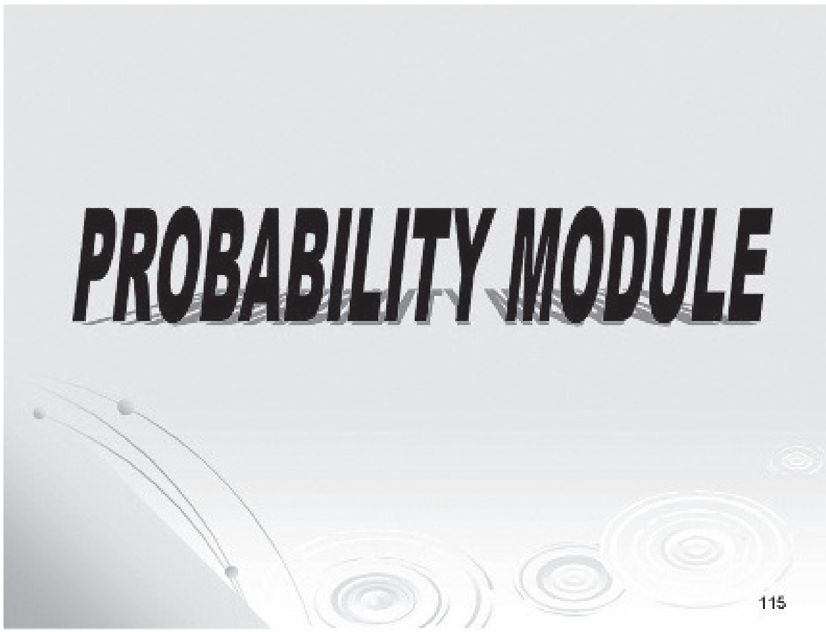
110

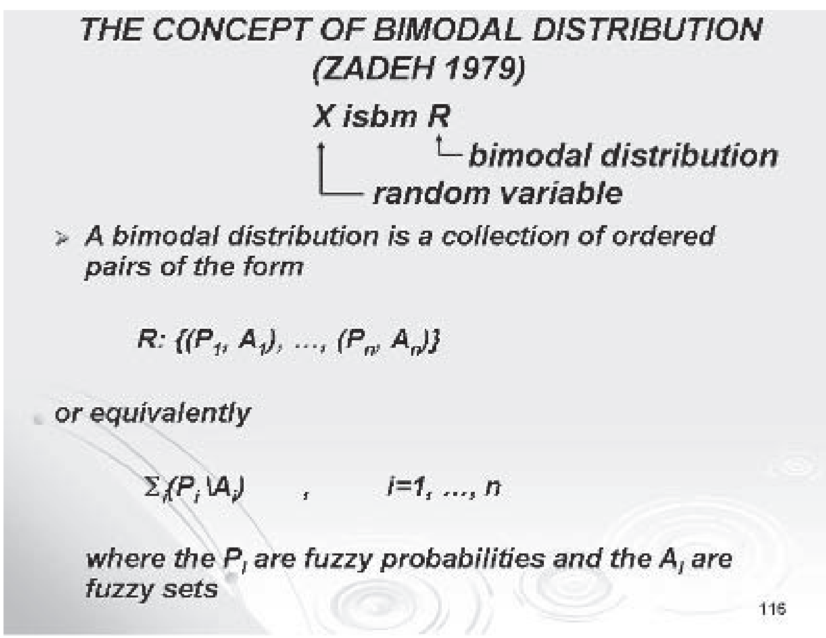

CONTINUED
Step 2. Deduction via extension principle
$\mu_q(d) = \sup_{g,h} ts(g,h)$
subject to
$d = \frac{1}{m}\Sigma_i g_i - \frac{1}{n}\Sigma_j h_j$
111


DEDUCTION PRINCIPLE
Point of departure: question, q
Data: $D = (X_1/u_1, \ldots, X_n/u_n)$
u_i is a generic value of X_i
Ans(q): answer to q
If we knew the values of the X_i, $u_1, \ldots, u_n$, we could express Ans(q) as a function of the u_i
$Ans(q)=g(u_1, \ldots, u_n)$ $u=(u_1, \ldots, u_n)$
Our information about the u_i, $I(u_1, \ldots, u_n)$ is a generalized constraint on the u_i. The constraint is defined by its test-score function
$f(u)=f(u_1, \ldots, u_n)$
112


CONTINUED
Use the extension principle
$\mu_{Ans(q)}(v) = sup_u(ts(u))$
subject to
$v = g(u)$
113


MODULAR DEDUCTION DATABASE
POSSIBILITY MODULE
PROBABILITY MODULE
agent
FUZZY ARITHMETIC MODULE
SEARCH MODULE
FUZZY LOGIC MODULE
EXTENSION PRINCIPLE MODULE
114


PROBABILITY MODULE
115


THE CONCEPT OF BIMODAL DISTRIBUTION (ZADEH 1979)
X isbm R
bimodal distribution
random variable
A bimodal distribution is a collection of ordered pairs of the form
$R: \{(P_1, A_1), \ldots, (P_n, A_n)\}$
or equivalently
$\Sigma_i(P_i \backslash A_i)$, $i=1, \ldots, n$
where the P_i are fuzzy probabilities and the A_i are fuzzy sets
116

CONTINUED

Special cases:

1. *The P_i are crisp; the A_i are fuzzy*
2. *The P_i are fuzzy; the A_i are crisp*
3. *The P_i are crisp; the A_i are crisp*

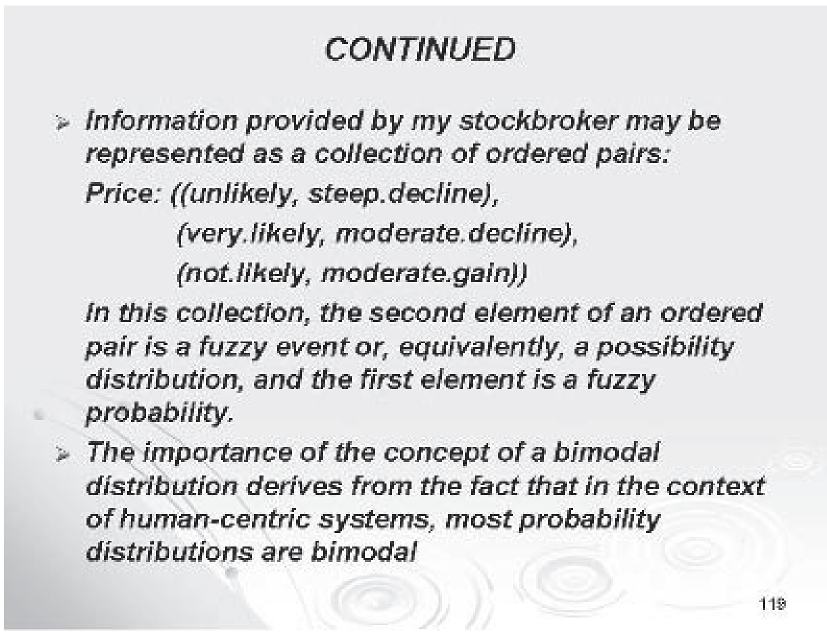
- *The Demspter-Shafer theory of evidence is basically a theory of crisp bimodal distributions*

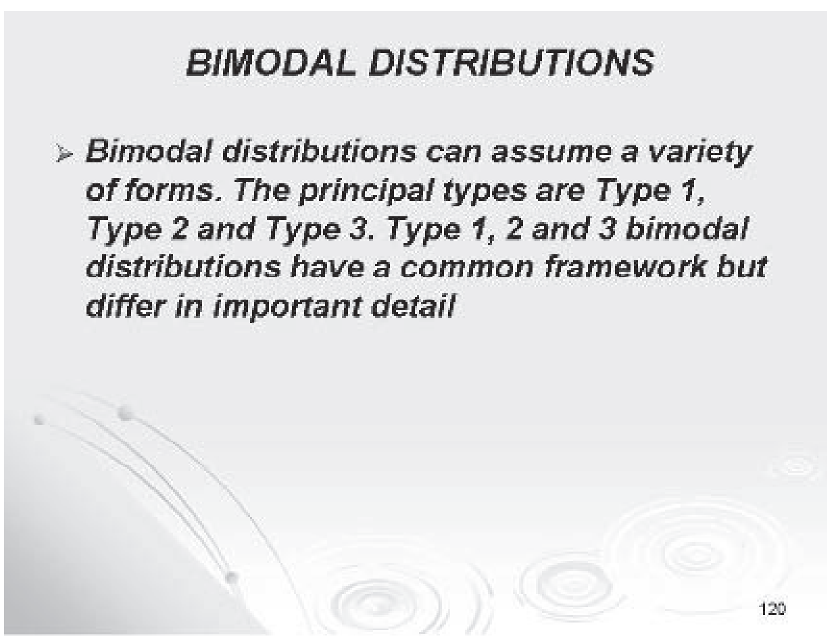
117

EXAMPLE: FORD STOCK

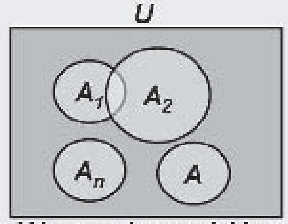
- *I am considering buying Ford stock. I ask my stockbroker, “What is your perception of the near-term prospects for Ford stock?” He tells me, “A moderate decline is very likely; a steep decline is unlikely; and a moderate gain is not likely.” My question is: What is the probability of a large gain?*

118


CONTINUED
Information provided by my stockbroker 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, equivalently, a possibility distribution, and the first element is a fuzzy probability.
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
119


BIMODAL DISTRIBUTIONS
Bimodal distributions can assume a variety of forms. The principal types are Type 1, Type 2 and Type 3. Type 1, 2 and 3 bimodal distributions have a common framework but differ in important detail
120

BIMODAL DISTRIBUTIONS (Type 1, 2, 3)



- *Type 1 (default): X is a random variable taking values in U*
  $A_1, \ldots, A_n, A$ are events (fuzzy sets)
  p_i = Prob(X is A_i), $i = 1, \ldots, n$
  $\Sigma_i p_i$ is unconstrained
  p_i is P_i (granular probability)

BMD: bimodal distribution: $((P_1, A_1), \ldots, (P_n, A_n))$
X isbm $(P_1 \backslash A_1 + \cdots + P_n \backslash A_n)$

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

121

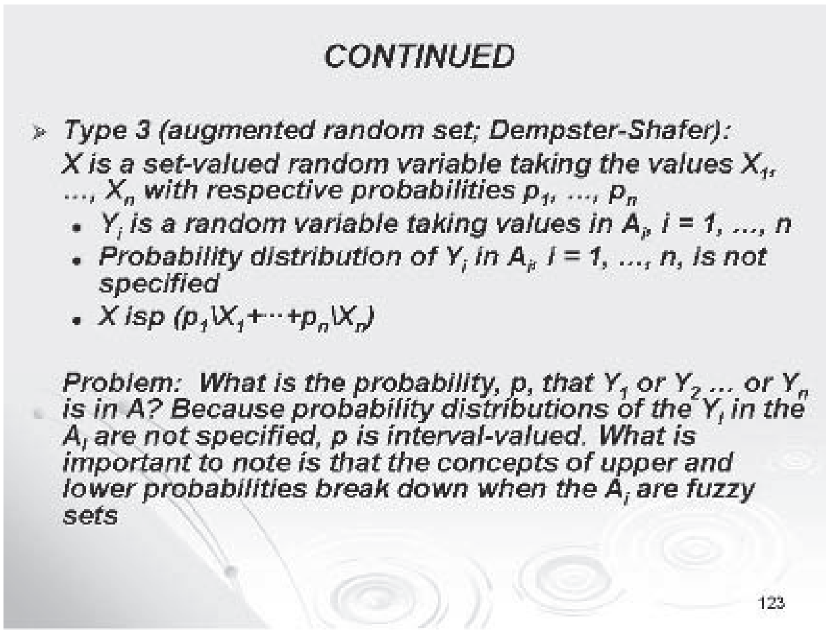
CONTINUED

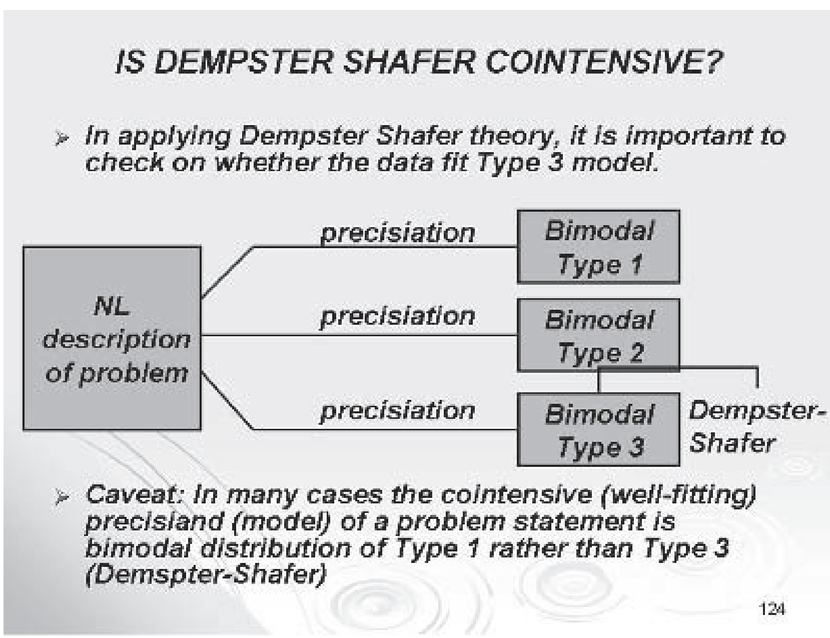
- *Type 2 (fuzzy random set): X is a fuzzy-set-valued random variable with values $A_1, \ldots, A_n$ (fuzzy sets)*
  p_i = Prob(X = A_i), $i = 1, \ldots, n$

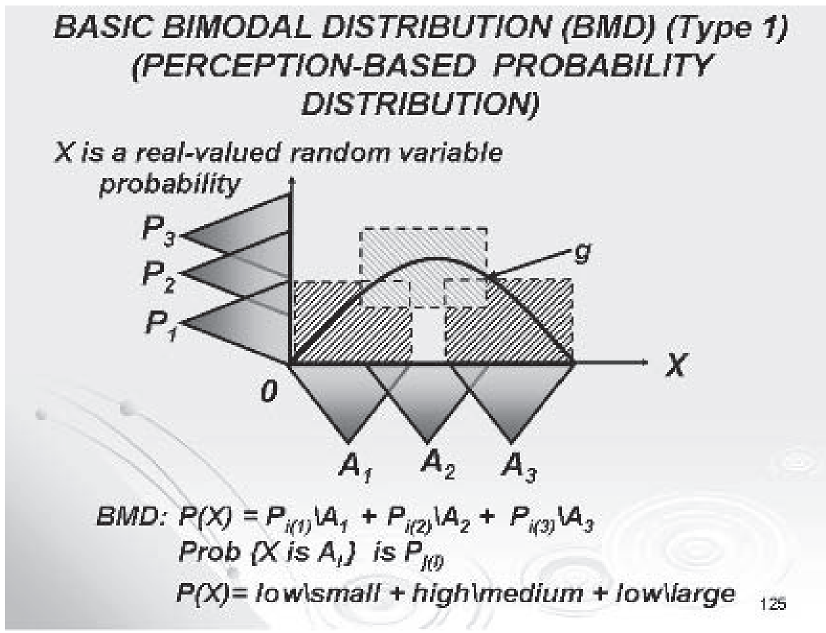
BMD: X isrs $(p_1 \backslash A_1 + \cdots + p_n \backslash A_n)$
$\Sigma_i p_i = 1$

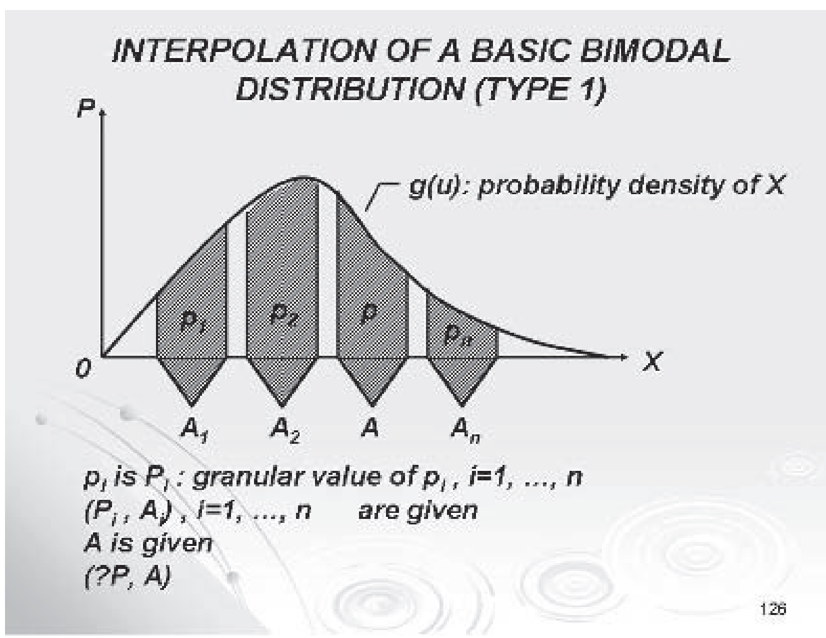
Problem: What is the probability, p, of A? p is not definable. What are definable are (a) the expected value of the conditional possibility of A given BMD, and (b) the expected value of the conditional necessity of A given BMD

122


CONTINUED
Type 3 (augmented random set; Dempster-Shafer):
X is a set-valued random variable taking the values X_1, ..., X_n with respective probabilities p_1, ..., p_n
Y_i is a random variable taking values in A_i, i = 1, ..., n
Probability distribution of Y_i in A_i, i = 1, ..., n, is not specified
X isp ($p_1\backslash X_1+\cdots+p_n\backslash X_n$)
Problem: What is the probability, p, that Y_1 or Y_2 ... or Y_n is in A? Because probability distributions of the Y_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
123


IS DEMPSTER SHAFER COINTENSIVE?
In applying Dempster Shafer theory, it is important to check on whether the data fit Type 3 model.
NL description of problem
precisiation
Bimodal Type 1
precisiation
Bimodal Type 2
precisiation
Bimodal Type 3
Dempster-Shafer
Caveat: In many cases the cointensive (well-fitting) precisiand (model) of a problem statement is bimodal distribution of Type 1 rather than Type 3 (Demspter-Shafer)
124


BASIC BIMODAL DISTRIBUTION (BMD) (Type 1)
(PERCEPTION-BASED PROBABILITY DISTRIBUTION)
X is a real-valued random variable
probability
P_3
P_2
P_1
g
0
X
A_1 A_2 A_3
BMD: $P(X) = P_{i(1)} \backslash A_1 + P_{i(2)} \backslash A_2 + P_{i(3)} \backslash A_3$
Prob {X is A_i} is $P_{j(i)}$
P(X)= low\small + high\medium + low\large
125


INTERPOLATION OF A BASIC BIMODAL DISTRIBUTION (TYPE 1)
P
g(u): probability density of X
p_1 p_2 p p_n
0
X
A_1 A_2 A A_n
p_i is P_i : granular value of p_i , i=1, ..., n
(P_i , A_i) , i=1, ..., n are given
A is given
(?P, A)
126

INTERPOLATION MODULE AND PROBABILITY MODULE

Prob {X is A_i} is P_i , i = 1, ..., n
Prob {X is A} is Q

$$\mu_Q(v) = sup_g(\mu_{P_1}(\int_U \mu_{A_1}(u)g(u)du) \wedge \cdots \wedge$$

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

subject to

$$U = \int_U \mu_A(u)g(u)du$$

127

OPERATIONS ON BIMODAL DISTRIBUTIONS

P(X) defines possibility distribution of g

$$\pi(g) = \mu_{P_1}(\int_U \mu_{A_1}(u)g(u)du) \wedge \cdots \wedge \mu_{P_n}(\int_U \mu_{A_n}(u)g(u)du)$$

problem

a) what is the expected value of X

128

EXPECTED VALUE OF A BIMODAL DISTRIBUTION

$$E(P^*) = \int_U ug(u)du = f(g)$$

Extension Principle

$$\mu_{E(P^*)}(v) = \sup_g \left(\mu_{P_1}\left(\int_U \mu_{A_1}(u)g(u)du\right) \wedge \cdots \wedge \mu_{P_n}\left(\int_U \mu_{A_n}(u)g(u)du\right)\right)$$

subject to: $v = \int_U ug(u)du$

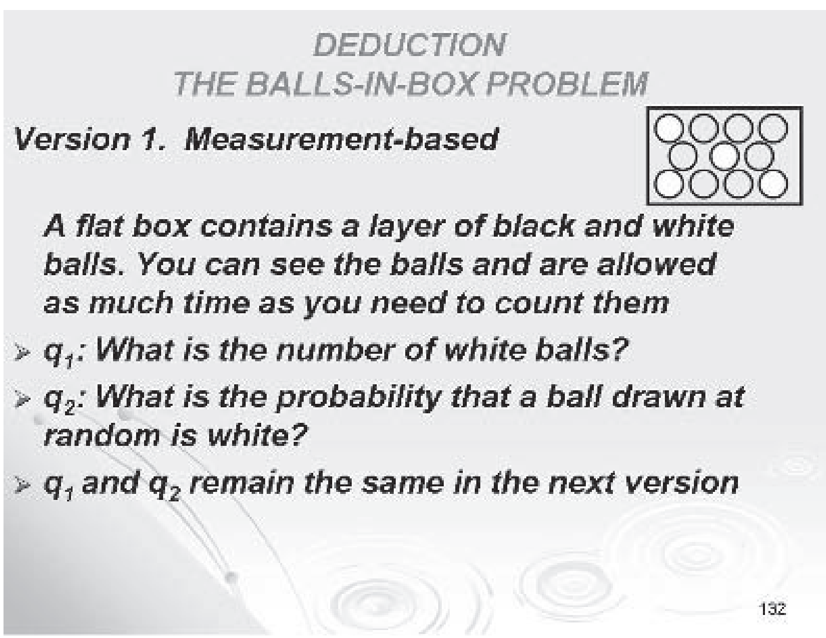
129

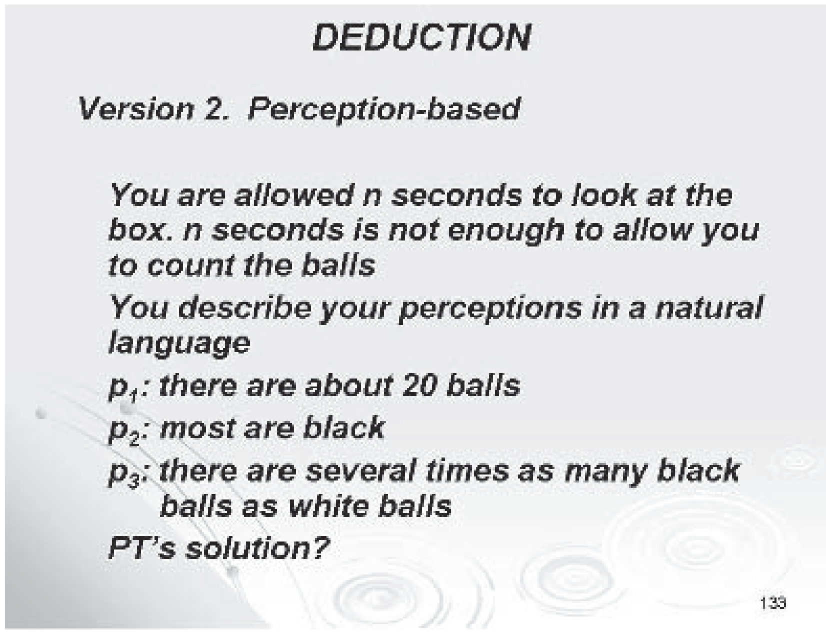
SUMMATION

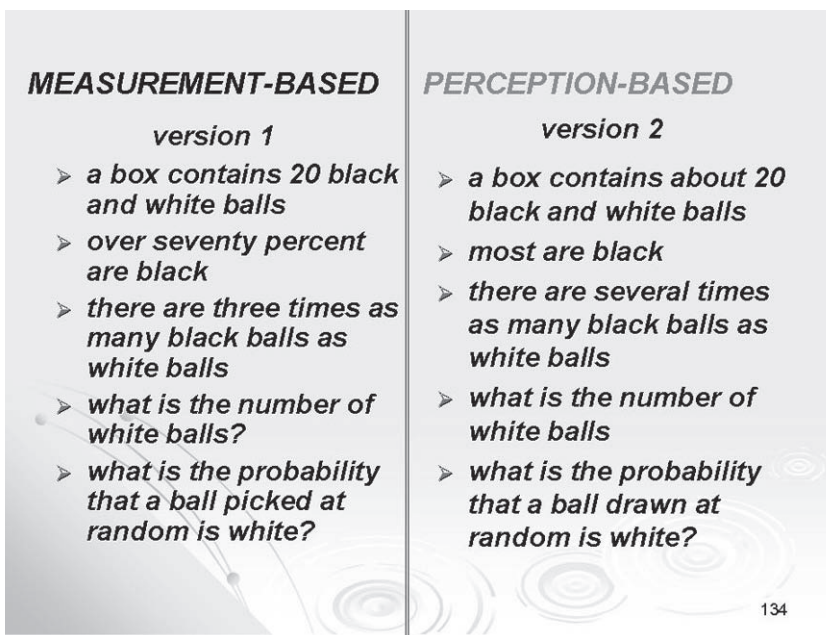
- *addition of significant question-answering capability to search engines is a complex, open-ended problem*
- *incremental progress, but not much more, is achievable through the use of bivalent-logic-based methods*
- *to achieve significant progress, it is imperative to develop and employ new methods based on computing with words, protoform theory, precisiated natural language and computational theory of precisiation of meaning*
- *The centerpiece of new methods is the concept of a generalized constraint*

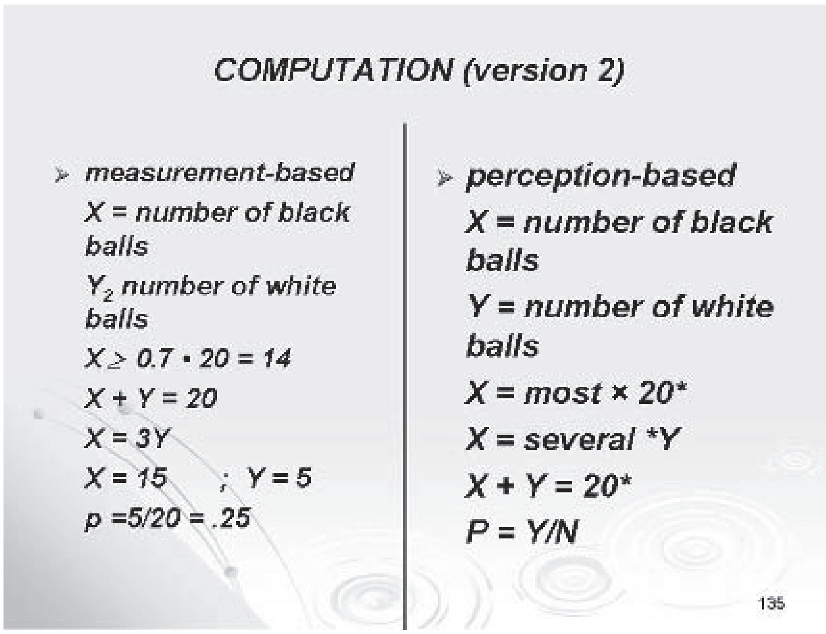
130

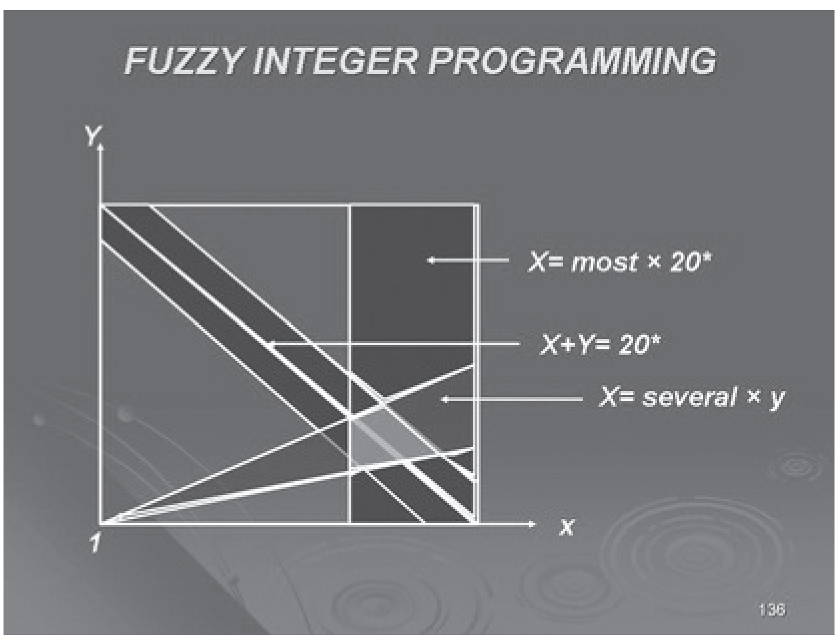

APPENDIX
131

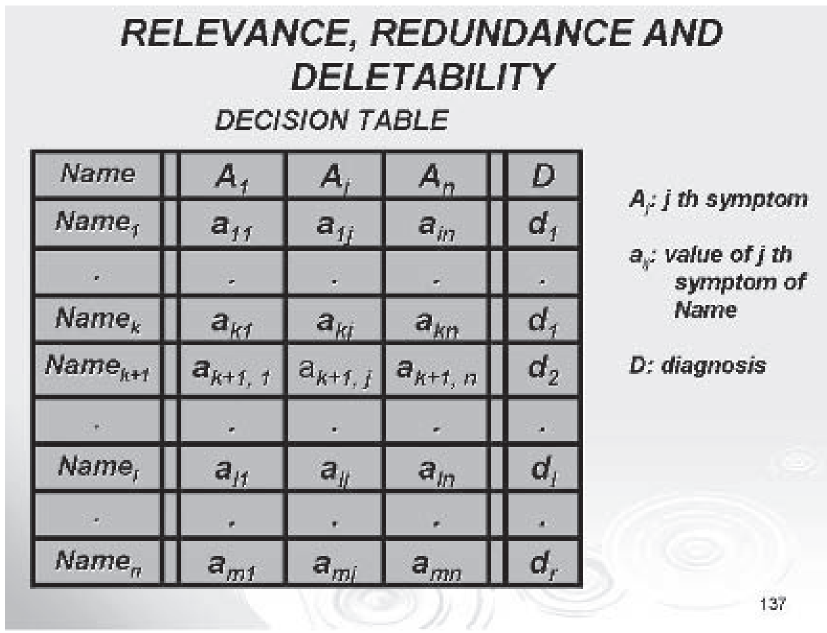

DEDUCTION
THE BALLS-IN-BOX PROBLEM
Version 1. Measurement-based
A flat box contains a layer of black and white balls. You can see the balls and are allowed as much time as you need to count them
➢ q_1: What is the number of white balls?
➢ q_2: What is the probability that a ball drawn at random is white?
➢ q_1 and q_2 remain the same in the next version
132


DEDUCTION
Version 2. Perception-based
You are allowed n seconds to look at the box. n seconds is not enough to allow you to count the balls
You describe your perceptions in a natural language
p1: there are about 20 balls
p2: most are black
p3: there are several times as many black balls as white balls
PT's solution?
133


MEASUREMENT-BASED
version 1
a box contains 20 black and white balls
over seventy percent are black
there are three times as many black balls as white balls
what is the number of white balls?
what is the probability that a ball picked at random is white?
PERCEPTION-BASED
version 2
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
what is the probability that a ball drawn at random is white?
134


COMPUTATION (version 2)
measurement-based
X = number of black balls
Y₂ number of white balls
X ≥ 0.7 • 20 = 14
X + Y = 20
X = 3Y
X = 15 ; Y = 5
p =5/20 = .25
perception-based
X = number of black balls
Y = number of white balls
X = most × 20*
X = several *Y
X + Y = 20*
P = Y/N
135


FUZZY INTEGER PROGRAMMING
Y
X= most × 20*
X+Y= 20*
X= several × y
x
1
136



RELEVANCE, REDUNDANCE AND DELETABILITY

DECISION TABLE

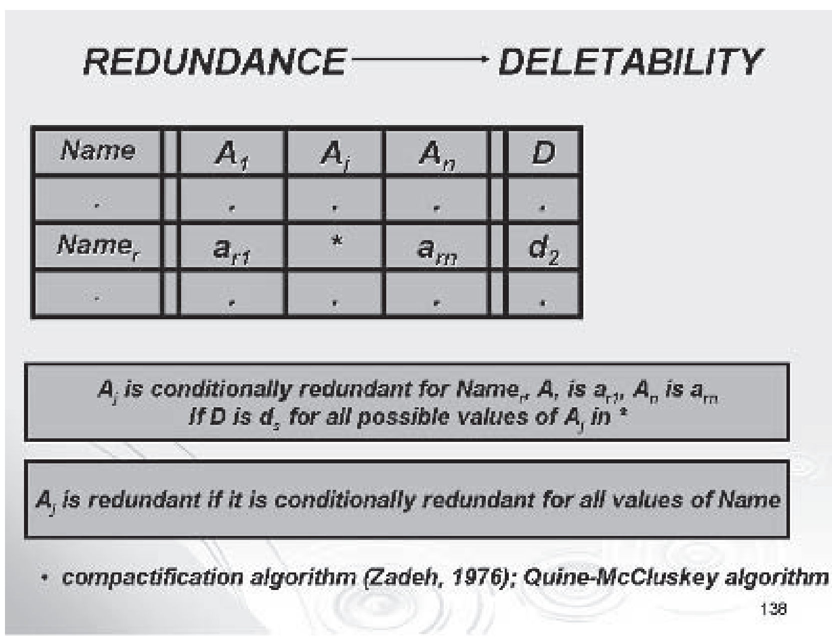
Name	A_1	A_j	A_n	D
$Name_1$	a_{11}	a_{1j}	a_{in}	d_1
.	.	.	.	.
$Name_k$	a_{k1}	a_{kj}	a_{kn}	d_1
$Name_{k+1}$	$a_{k+1,\,1}$	$a_{k+1,\,j}$	$a_{k+1,\,n}$	d_2
.	.	.	.	.
$Name_l$	a_{l1}	a_{lj}	a_{ln}	d_l
.	.	.	.	.
$Name_n$	a_{m1}	a_{mj}	a_{mn}	d_r

A_j: j th symptom

a_{ij}: value of j th symptom of Name

D: diagnosis

137



REDUNDANCE ⟶ DELETABILITY

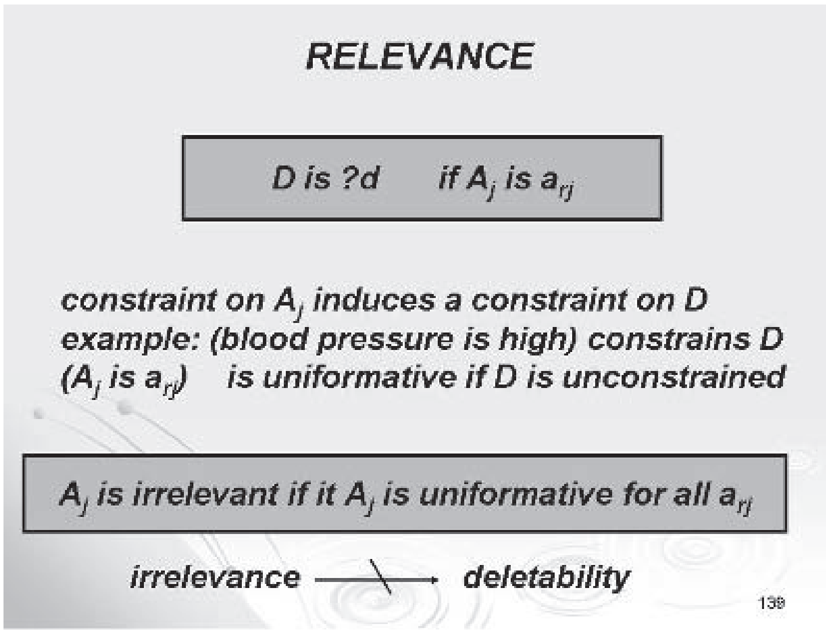
Name	A_1	A_j	A_n	D
.	.	.	.	.
$Name_r$	a_{r1}	*	a_{rn}	d_2
.	.	.	.	.

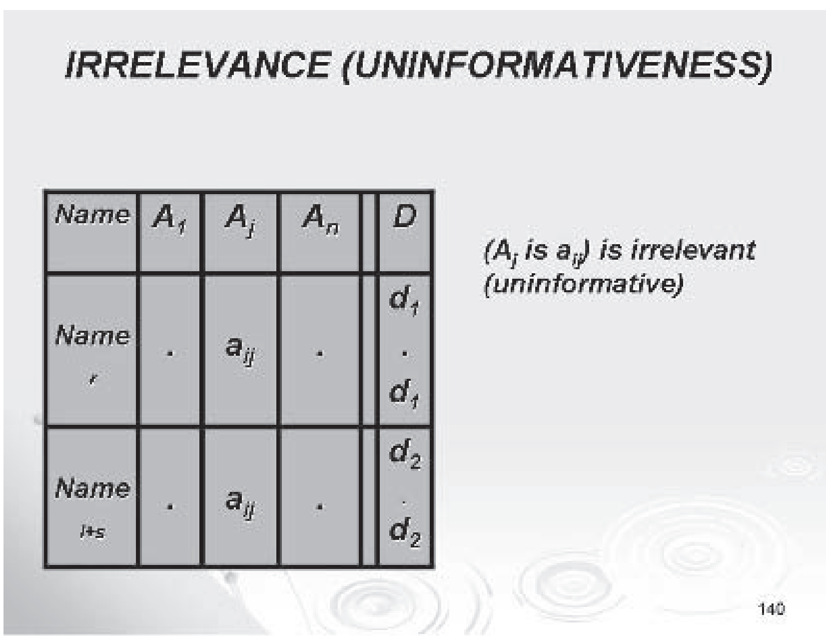
A_j is conditionally redundant for $Name_r$, A_1 is a_{r1}, A_n is a_{rn}
If D is d_s for all possible values of A_j in *

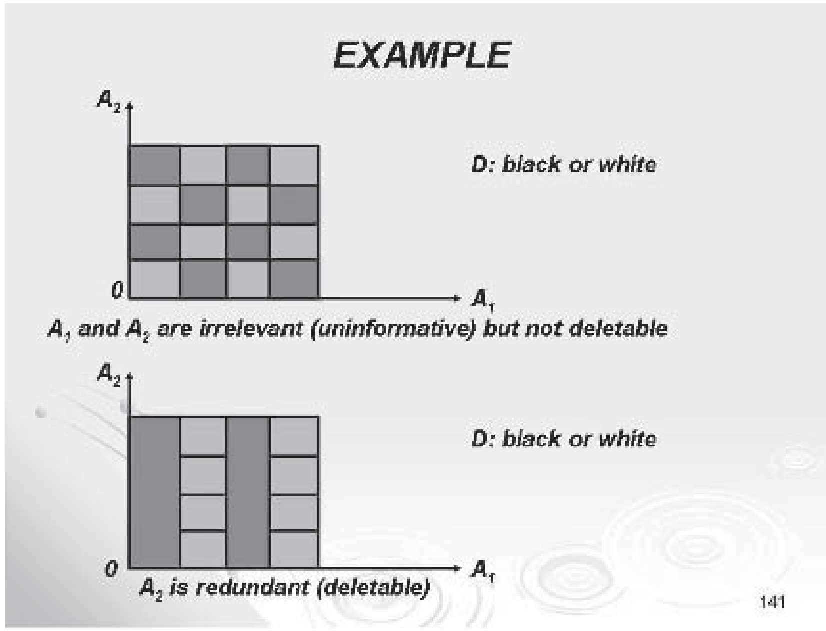
A_j is redundant if it is conditionally redundant for all values of Name

- compactification algorithm (Zadeh, 1976); Quine-McCluskey algorithm

138


RELEVANCE
D is ?d if A_j is a_{rj}
constraint on A_j induces a constraint on D
example: (blood pressure is high) constrains D
(A_j is a_{rj}) is uniformative if D is unconstrained
A_j is irrelevant if it A_j is uniformative for all a_{rj}
irrelevance $\not\rightarrow$ deletability
139


IRRELEVANCE (UNINFORMATIVENESS)
Name
A_1
A_j
A_n
D
Name r
.
a_{ij}
.
d_1
.
d_1
Name l+s
.
a_{ij}
.
d_2
.
d_2
(A_j is a_{ij}) is irrelevant (uninformative)
140



KEY POINT—THE ROLE OF FUZZY LOGIC

- *Existing approaches to the enhancement of web intelligence are based on classical, Aristotelian, bivalent logic and bivalent-logic-based probability theory. In our approach, bivalence is abandoned. What is employed instead is fuzzy logic—a logical system which subsumes bivalent logic as a special case.*
- *Fuzzy logic is not fuzzy*
- *Fuzzy logic is a precise logic of fuzziness and imprecision*
- *The centerpiece of fuzzy logic is the concept of a generalized constraint.*

142

- *In bivalent logic, BL, truth is bivalent, implying that every proposition, p, is either true or false, with no degrees of truth allowed*
- *In multivalent logic, ML, truth is a matter of degree*
- *In fuzzy logic, FL:*
  - *everything is, or is allowed to be, to be partial, i.e., a matter of degree*
  - *everything is, or is allowed to be, imprecise (approximate)*
  - *everything is, or is allowed to be, granular (linguistic)*
  - *everything is, or is allowed to be, perception based*

143

CONTINUED

- *The generality of fuzzy logic is needed to cope with the great complexity of problems related to search and question-answering in the context of world knowledge; to deal computationally with perception-based information and natural languages; and to provide a foundation for management of uncertainty and decision analysis in realistic settings*

144

January 26, 2005

Factual Information About the Impact of Fuzzy Logic

PATENTS

- *Number of fuzzy-logic-related patents applied for in Japan: 17,740*
- *Number of fuzzy-logic-related patents issued in Japan: 4,801*
- *Number of fuzzy-logic-related patents issued in the US: around 1,700*

145

PUBLICATIONS

Count of papers containing the word "fuzzy" in title, as cited in INSPEC and MATH.SCI.NET databases.

Compiled by Camille Wanat, Head, Engineering Library, UC Berkeley, December 22, 2004

Number of papers in INSPEC and MathSciNet which have "fuzzy" in their titles:

INSPEC - "fuzzy" in the title
1970-1979: 569
1980-1989: 2,404
1990-1999: 23,207
2000-present: 14,172
Total: 40,352

MathSciNet - "fuzzy" in the title
1970-1979: 443
1980-1989: 2,465
1990-1999: 5,483
2000-present: 3,960
Total: 12,351

146

JOURNALS ("fuzzy" or "soft computing" in title)

1. ***Fuzzy Sets and Systems***
2. ***IEEE Transactions on Fuzzy Systems***
3. ***Fuzzy Optimization and Decision Making***
4. ***Journal of Intelligent & Fuzzy Systems***
5. ***Fuzzy Economic Review***
6. ***International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems***
7. ***Journal of Japan Society for Fuzzy Theory and Systems***
8. ***International Journal of Fuzzy Systems***
9. ***Soft Computing***
10. ***International Journal of Approximate Reasoning--Soft Computing in Recognition and Search***
11. ***Intelligent Automation and Soft Computing***
12. ***Journal of Multiple-Valued Logic and Soft Computing***
13. ***Mathware and Soft Computing***
14. ***Biomedical Soft Computing and Human Sciences***
15. ***Applied Soft Computing***

147

APPLICATIONS

The range of application-areas of fuzzy logic is too wide for exhaustive listing. Following is a partial list of existing application-areas in which there is a record of substantial activity.

1. *Industrial control*
2. *Quality control*
3. *Elevator control and scheduling*
4. *Train control*
5. *Traffic control*
6. *Loading crane control*
7. *Reactor control*
8. *Automobile transmissions*
9. *Automobile climate control*
10. *Automobile body painting control*
11. *Automobile engine control*
12. *Paper manufacturing*
13. *Steel manufacturing*
14. *Power distribution control*
15. *Software engineerinf*
16. *Expert systems*
17. *Operation research*
18. *Decision analysis*
19. *Financial engineering*
20. *Assessment of credit-worthiness*
21. *Fraud detection*
22. *Mine detection*
23. *Pattern classification*
24. *Oil exploration*
25. *Geology*
26. *Civil Engineering*
27. *Chemistry*
28. *Mathematics*
29. *Medicine*
30. *Biomedical instrumentation*
31. *Health-care products*
32. *Economics*
33. *Social Sciences*
34. *Internet*
35. *Library and Information Science*

148

Product Information Addendum 1

This addendum relates to information about products which employ fuzzy logic singly or in combination. The information which is presented came from SIEMENS and OMRON. It is fragmentary and far from complete. Such addenda will be sent to the Group from time to time.

SIEMENS:

* *washing machines, 2 million units sold*
* *fuzzy guidance for navigation systems (Opel, Porsche)*
* *OCS: Occupant Classification System (to determine, if a place in a car is occupied by a person or something else; to control the airbag as well as the intensity of the airbag). Here FL is used in the product as well as in the design process (optimization of parameters).*
* *fuzzy automobile transmission (Porsche, Peugeot, Hyundai)*

OMRON:

* *fuzzy logic blood pressure meter, 7.4 million units sold, approximate retail value $740 million dollars*

Note: If you have any information about products and or manufacturing which may be of relevance please communicate it to Dr. Vesa Niskanen vesa.a.niskanen@helsinki.fi and Masoud Nikravesh Nikravesh@cs.berkeley.edu .

149

Product Information Addendum 2

This addendum relates to information about products which employ fuzzy logic singly or in combination. The information which is presented came from Professor Hideyuki Takagi, Kyushu University, Fukuoka, Japan. Professor Takagi is the co-inventor of neurofuzzy systems. Such addenda will be sent to the Group from time to time.

Facts on FL-based systems in Japan (as of 2/06/2004)

1. Sony's FL camcorders

Total amount of camcorder production of all companies in 1995-1998 times Sony's market share is the following. Fuzzy logic is used in all Sony's camcorders at least in these four years, i.e. total production of Sony's FL-based camcorders is 2.4 millions products in these four years.

1,228K units X 49% in 1995
1,315K units X 52% in 1996
1,381K units X 50% in 1997
1,416K units X 51% in 1998

2. FL control at Idemitsu oil factories

Fuzzy logic control is running at more than 10 places at 4 oil factories of Idemitsu Kosan Co. Ltd including not only pure FL control but also the combination of FL and conventional control.

They estimate that the effect of their FL control is more than 200 million YEN per year and it saves more than 4,000 hours per year.

150

3. Canon

Canon used (uses) FL in their cameras, camcorders, copy machine, and stepper alignment equipment for semiconductor production. But, they have a rule not to announce their production and sales data to public.

Canon holds 31 and 31 established FL patents in Japan and US, respectively.

4. Minolta cameras

Minolta has a rule not to announce their production and sales data to public, too.

whose name in US market was Maxxum 7xi. It used six FL systems in a camera and was put on the market in 1991 with 98,000 YEN (body price without lenses). It was produced 30,000 per month in 1991. Its sister cameras, alpha-9xi, alpha-5xi, and their successors used FL systems, too. But, total number of production is confidential.

151

5. FL plant controllers of Yamatake Corporation

Yamatake-Honeywell (Yamatake's former name) put FUZZICS, fuzzy software package for plant operation, on the market in 1992. It has been used at the plants of oil, oil chemical, chemical, pulp, and other industries where it is hard for conventional PID controllers to describe the plan process for these more than 10 years.

They planed to sell the FUZZICS 20 - 30 per year and total 200 million YEN.

As this software runs on Yamatake's own control systems, the software package itself is not expensive comparative to the hardware control systems.

6. Others

Names of 225 FL systems and products picked up from news articles in 1987 - 1996 are listed at http://www.adwin.com/elec/fuzzy/note_10.html in Japanese.)

Note: If you have any information about products and or manufacturing which may be of relevance please communicate it to Dr. Vesa Niskanen vesa.a.niskanen@helsinki.fi and Masoud Nikravesh Nikravesh@cs.berkeley.edu , with cc to me.

152