TOWARD A GENERALIZED THEORY OF UNCERTAINTY (GTU)

Lotfi A. Zadeh

Computer Science Division
Department of EECS
UC Berkeley

April 15, 2005
Polish Academy of Sciences
Warsaw, Poland

URL: http://www-bisc.cs.berkeley.edu

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

Email: Zadeh@eecs.berkeley.edu

BACKDROP

LAZ 4/7/2005

PREAMBLE

- Almost all decisions are made in an environment of partial uncertainty, partial knowledge and partial truth. As we move further into the age of machine intelligence and mechanized decision-making, the need for a better understanding of how to deal with uncertainty is growing in urgency and importance.

KEY POINTS

- *Uncertainty is an attribute of information*
- *Traditionally, it is assumed that information is statistical in nature*
- *A logical consequence of this assumption is that uncertainty should be dealt with through the use of probability theory*
- *A key question which arises is: Is probability theory (PT) sufficient for dealing with uncertainty?*

CONTINUATION

- *There are some who believe that it is. In the words of Professor Dennis Lindley, an eminent Bayesian:*

  The only satisfactory description of uncertainty is probability. By this I mean that every uncertainty statement must be in the form of a probability; that several uncertainties must be combined using the rules of probability; and that the calculus of probabilities is adequate to handle all situations involving uncertainty…probability is the only sensible description of uncertainty and is adequate for all problems involving uncertainty. All other methods are inadequate…anything that can be done with fuzzy logic, belief functions, upper and lower probabilities, or any other alternative to probability can better be done with probability.

CONTINUATION

- There are many, including myself, who in one way or another, do not share this view. Indeed, in recent years a number of theories have been proposed which may be viewed as generalizations of probability theory or additions to it. Can these theories be unified? This is what the generalized theory of uncertainty, attempts to do. But first, a counterpoint and a challenge to those who share Professor Lindley's view.

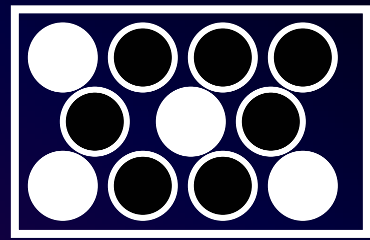
A QUICK COUNTERPOINT

- *To begin with, uncertainty need not be associated with randomness*

  Example: suppose that I am uncertain about the value of X, but know that it is between 10 and 12. No randomness is involved
- *Interval analysis deals with uncertainty but is not probabilistic. Interval analysis is possibilistic*
- *The real test of a theory of uncertainty is its ability to solve test problems. A few simple test problems are discussed in the following*

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***

CONTINUED

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

p_1: *there are about 20 balls*

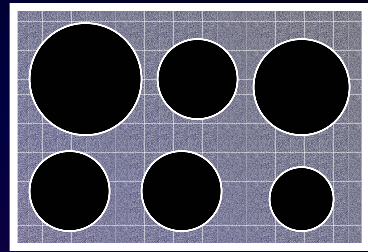
p_2: *most are black*

p_3: *there are several times as many black balls as white balls*

PT's solution?

CONTINUED

Version 3. Measurement-based



The balls have the same color but different sizes

You are allowed as much time as you need to count the balls

q_1: How many balls are large?

q_2: What is the probability that a ball drawn at random is large

PT's solution?

CONTINUED

Version 4. Perception-based

You are allowed n seconds to look at the box. n seconds is not enough to allow you to count the balls

Your perceptions are:

p_1: there are about 20 balls
p_2: most are small
p_3: there are several times as many small balls as large balls
q_1: how many are large?
q_2: what is the probability that a ball drawn at random is large?

A SERIOUS LIMITATION OF PT

- *Version 4 points to a serious short coming of PT*
- *In PT there is no concept of cardinality of a fuzzy set*
- *How many large balls are in the box?*

- *There is no underlying randomness*

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?

COMPUTATION (version 2)

- *measurement-based*

  X = number of black balls

  Y_2 number of white balls

  $X \geq 0.7 \bullet 20 = 14$

  $X + Y = 20$

  $X = 3Y$

  $X = 15 \quad ; \quad Y = 5$

  $p = 5/20 = .25$

- *perception-based*

  X = number of black balls

  Y = number of white balls

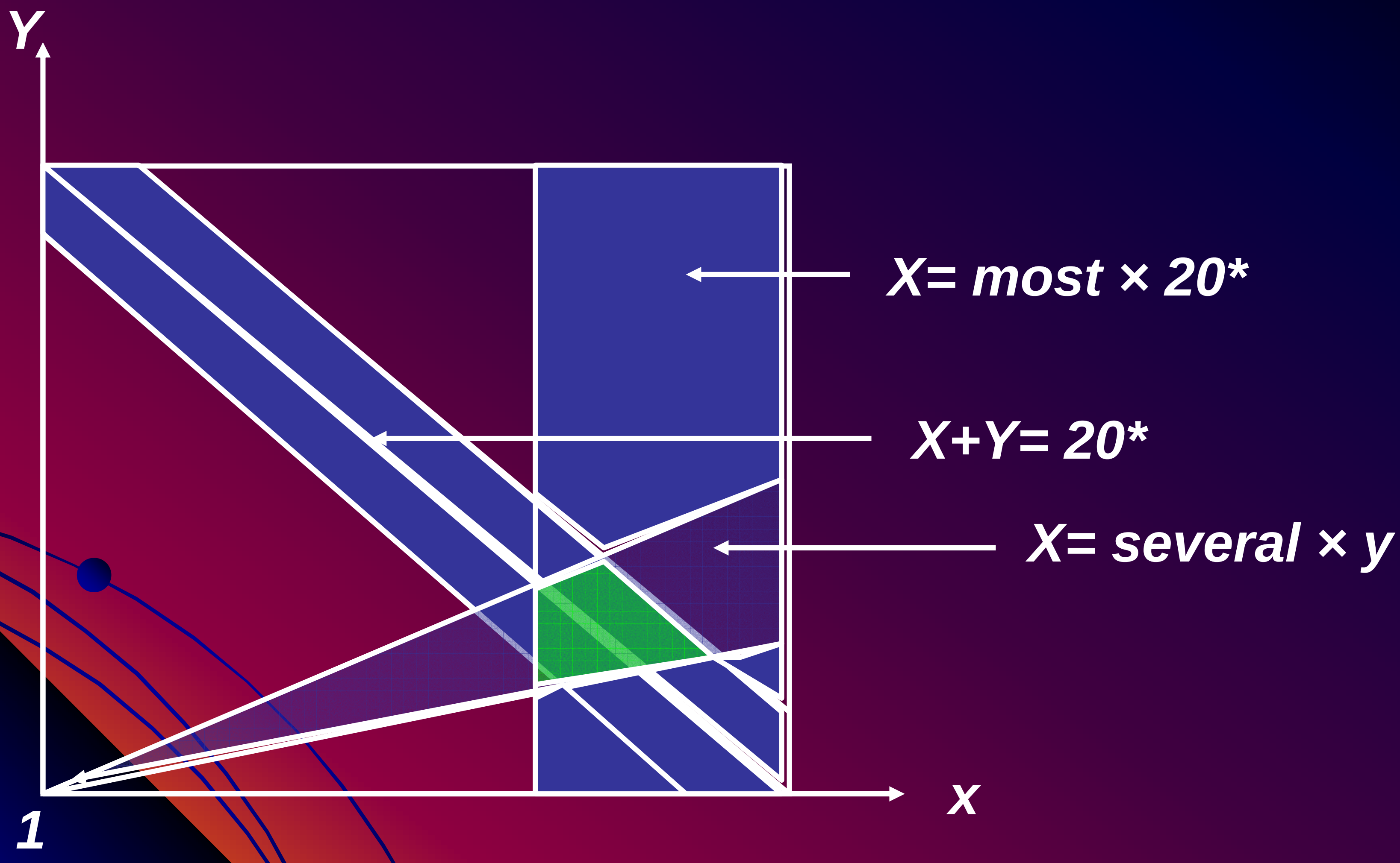
  *X = most × 20**

  *X = several *Y*

  *X + Y = 20**

  P = Y/N

FUZZY INTEGER PROGRAMMING



PROBLEM

- *X, a and b are real numbers, with a < b*
- *Find an X or X's such that X >> *a and X << *b*
- **a: approximately a; *b: approximately b*
- *Use probability theory?*
- *Fuzzy logic solution*

$$\mu_X(u) = \mu_{>>*a}(u) \wedge \mu_{<<*b}(u)$$

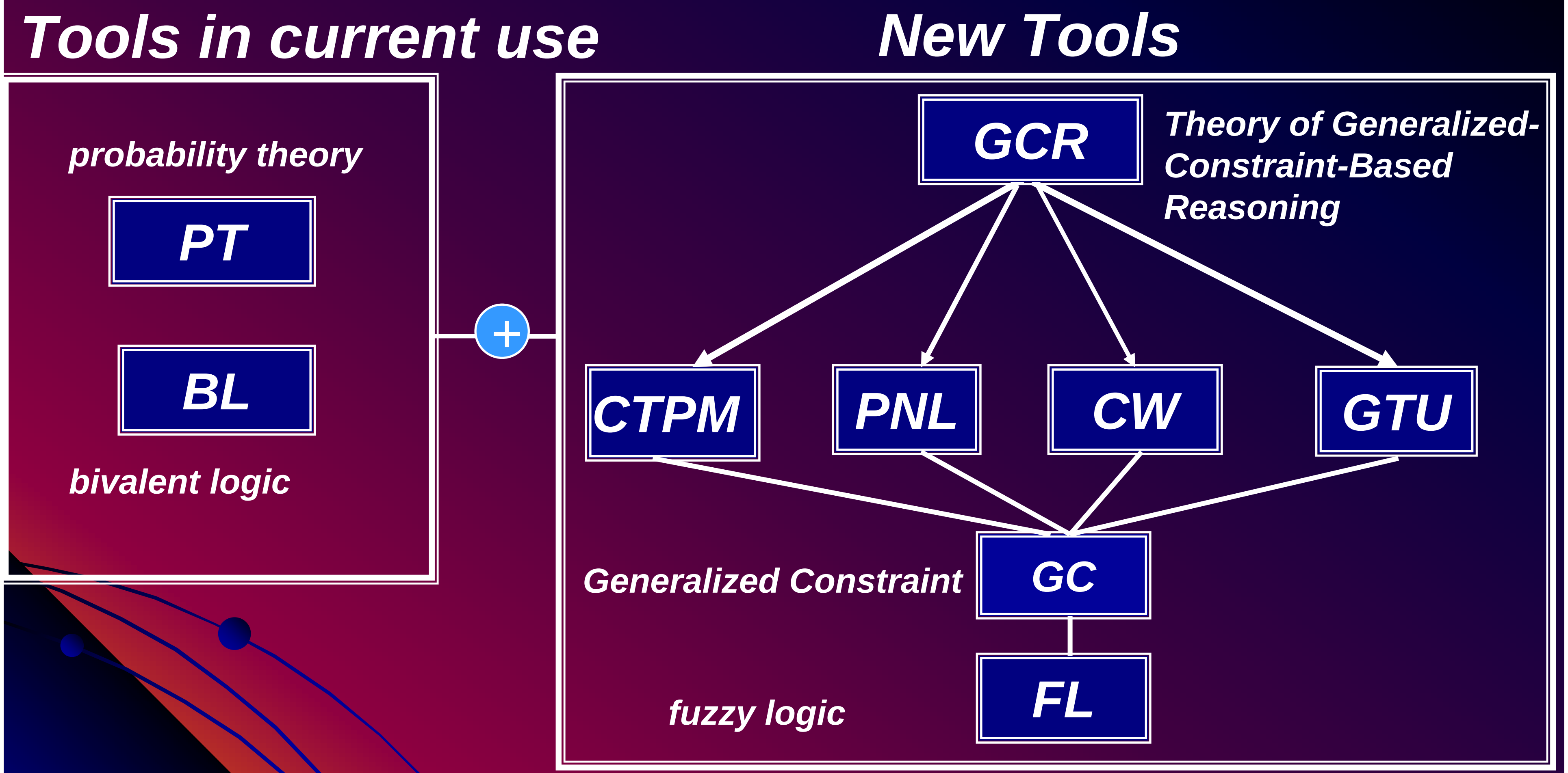
CONTINUED

- In PT, there are no tools for dealing with semantics of natural languages
- What is the meaning of: There are several times as many small balls as large balls?
- PT is not equipped to operate on perception-based information expressed in a natural language
- PT has a limited capability to deal with world knowledge, since much of world knowledge is perception-based

KEY POINT

- A basic difference between GTU and bivalent-logic-based theories of uncertainty relates to the role of natural languages. In GTU, the semantics of natural languages plays a pivotal role. The underlying reason is that GTU's capability to operate on perception-based information is directly dependent on GTU's ability to understand natural language, since a natural language is basically a system for describing perceptions.
- To deal with information expressed in a natural language, GTU employs new tools drawn from fuzzy logic. The centerpiece of these tools is the concept of a generalized constraint. A concept which underlies the concept of a generalized constraint is the concept of precisiation

STRUCTURE OF GTU



PT: standard bivalent-logic-based probability theory
CTPM : Computational 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

THE CONCEPT OF PRECISIATION

- ***The concepts of precision and imprecision have a position of centrality in science and, more generally, in human cognition. But what is not in existence is the concept of precisiation—a concept whose fundamental importance becomes apparent when we move from bivalent logic to fuzzy logic.***

WHAT IS PRECISE?

PRECISE

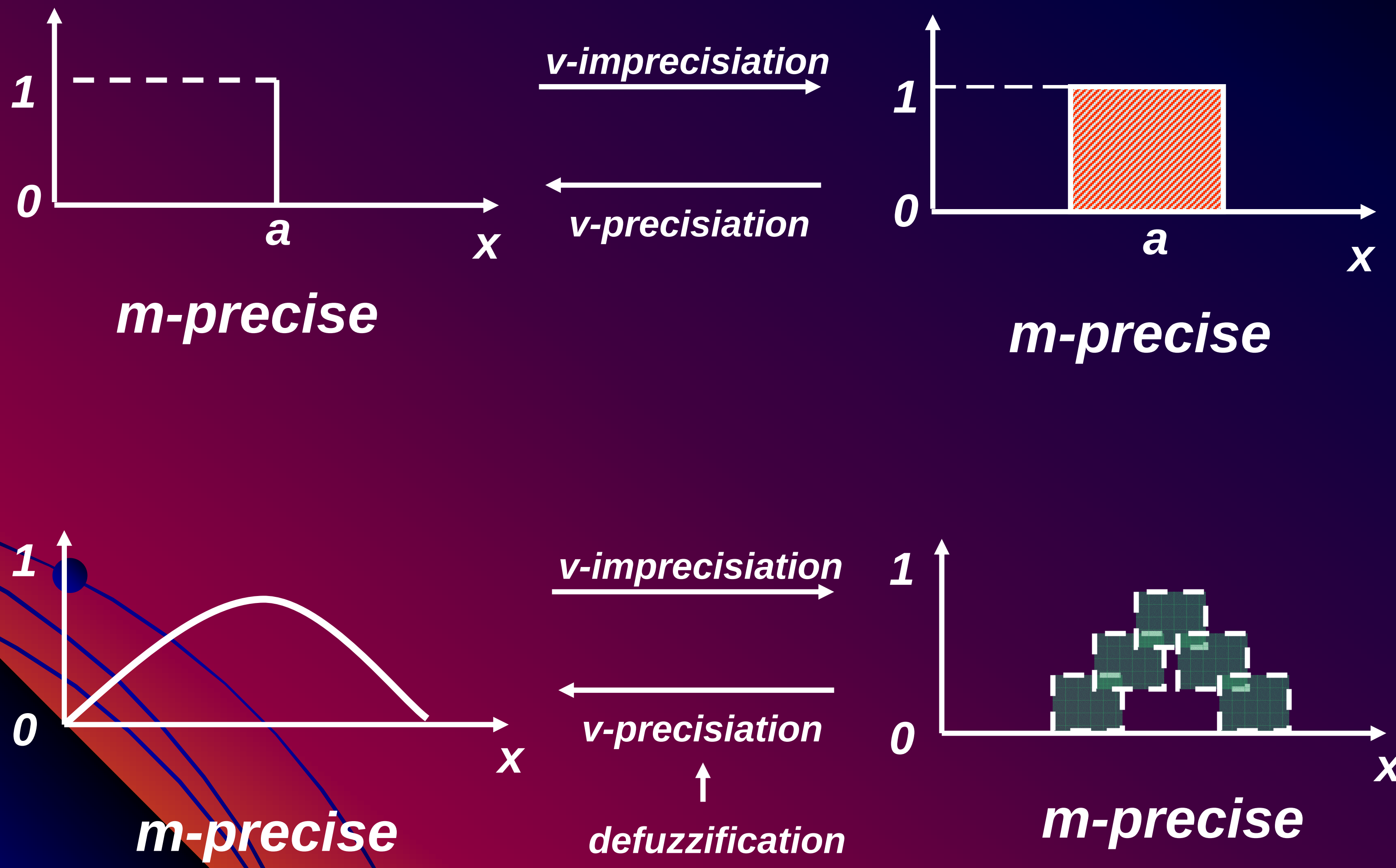
v-precise — **m-precise**

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

PRECISIATION AND IMPRECISIATION



KEY POINTS

In PNL

precisiation = m-precisiation

- *a proposition, p, is p precisiated by representing its meaning as a generalized constraint*
- *precisiation of meaning does not imply precisiation of value*
- *"Andrea is tall" is precisiated by defining "tall" as a fuzzy set*
- *A desideratum of precisiation is cointension*
- *Informally, p and q are cointensive if the intension (meaning) of p is approximately the same as the intension (meaning) of q*

NEED FOR PRECISIATION

- fuzzy commands, instructions
  - take a few steps
  - slow down
  - proceed with caution
  - raise your glass
  - use with adequate ventilation
- fuzzy commands and instructions cannot be understood by a machine
- to be understood by a machine, fuzzy commands and instructions must be precisiated

VALIDITY OF DEFINITION

- *If C is a concept and Def(C) is its definition, then Def(C) is a valid definition if it is cointensive with C*

IMPORTANT CONCLUSION

- *In general, cointensive, i.e., valid, definitions of fuzzy concepts cannot be formulated within the conceptual structure of bivalent logic and bivalen-logic-based probability theory*
- *This conclusion applies to such basic concepts as*
  - *Randomness*
  - *Causality*
  - *Relevance*
  - *Summary*
  - *Independence*
  - *Mountain*

PRECISIATION OF MEANING VS. UNDERSTANDING OF MEANING

- Precisiation of meaning $\neq$ Understanding of meaning
  - I understand what you said, but can you be more precise
- Use with adequate ventilation
- Unemployment is high
- Most Swedes are tall
- Most Swedes are much taller than most Italians
- Overeating causes obesity
- Causality
- Relevance
- Bear market
- Mountain
- Edge
- Approximately 5

fuzzy concepts

PRECISIATION OF MEANING

BASIC POINT

- ***The meaning of a proposition, p, may be precisiated in many different ways***

$$p \xrightarrow{\textit{precisiation}} \left.\begin{array}{l} Pre_1(p) \\ Pre_2(p) \\ \ldots \\ Pre_n(p) \end{array}\right\} \textit{precisiands of } p$$

- ***Conventional, bivalent-logic-based precisiation has a limited expressive power***

CHOICE OF PRECISIANDS

BASIC POINT

- *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 in which achieves a compromise between complexity and cointension*

EXAMPLE OF CONVENTIONAL DEFINITION OF FUZZY CONCEPTS

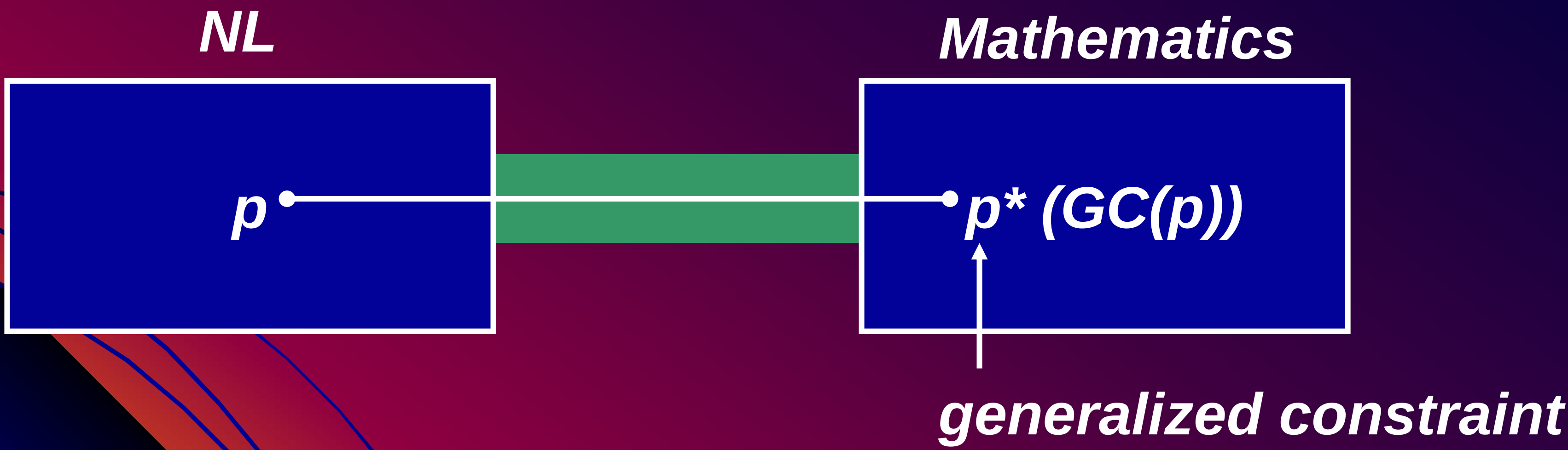
Robert Shuster (Ned Davis Research)

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*

THE KEY IDEA

- *In PNL, 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 PNL*

PRECISIATION AND GRANULAR COMPUTING

KEY IDEA

example

most Swedes are tall $\xrightarrow{\textit{precisiation}}$

Count(tall.Swedes/Swedes) is most $\rightarrow$ $\boxed{\int_a^b h(u)\mu_{tall}(u)du}$ *is most*

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

h: height density function

h(u)du = fraction of Swedes whose height lies in the interval [u, u+du]

- *In granular computing, the objects of computation are not values of variables but constraints on values of variables*

THE CONCEPT

OF A GENERALIZED

CONSTRAINT

GENERALIZED CONSTRAINT (Zadeh 1986)

- **Bivalent constraint (hard, inelastic, categorical:)**

X ε *C* — **constraining bivalent relation**

- **Generalized constraint:**

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

- *R*: **constraining non-bivalent (fuzzy) relation**
- *r*: **index of modality (defines semantics)**
- *X*: **constrained variable**

r: $\underbrace{\varepsilon \mid = \mid \leq \mid \geq \mid \subset \mid \ldots}_{\text{bivalent}}$ | **blank | p | v | u | rs | fg | ps |…**

primary: blank, p, v

CONTINUED

- ***constrained variable***

  - ***X is an n-ary variable, $X= (X_1, \ldots, X_n)$***
  - ***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, $G[A]: (Name_1, \ldots, Name_n)$, with each member of the group, $Name_i$, $i =1, \ldots, n$, associated with an attribute-value, A_i. A_i may be vector-valued. Symbolically***

$$G[A]: (Name_1/A_1+\ldots+Name_n/A_n)$$

Basically, X is a relation

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*

GENERALIZED CONSTRAINT—MODALITY r

X 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

CONTINUED

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

PRIMARY GENERALIZED CONSTRAINTS

Possibilistic examples:

- ***Monika is young*** $\longrightarrow$ ***Age (Monika) is young***

  X R

- ***Monika is much younger than Maria*** $\longrightarrow$

  (Age (Monika), Age (Maria)) is much younger

  X R

- ***most Swedes are tall***

  $\longrightarrow$ ***ΣCount (tall.Swedes/Swedes) is most***

  R

  X

EXAMPLES: PROBABILISITIC

- *X is a normally distributed random variable with mean m and variance* σ^2 $\longrightarrow$

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

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

  $$X \text{ isp } (p_1\backslash u_1+p_2\backslash u_2+p_3\backslash u_3)$$

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—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

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

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

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

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

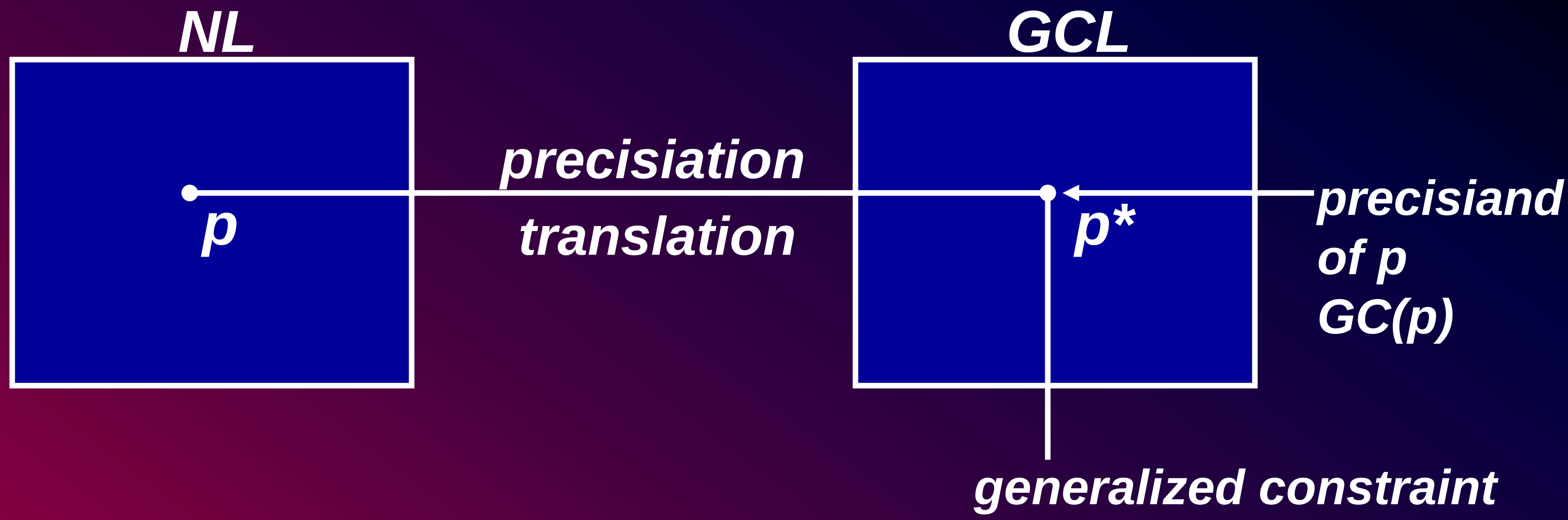
GENERALIZED CONSTRAINT LANGUAGE (GCL)

- 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

PRECISIATION = TRANSLATION INTO GCL

BASIC STRUCTURE



annotation

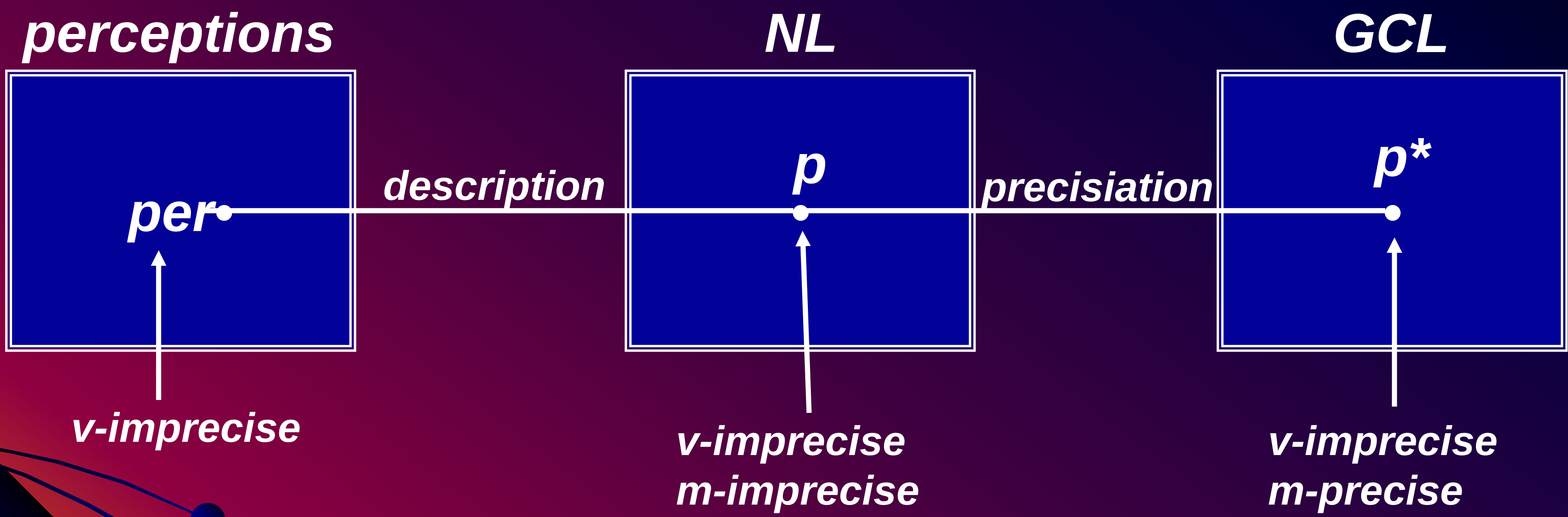
p ⟶ *X/A isr R/B* ⟵ *GC-form of p*

example

p: Carol lives in a small city near San Francisco

X/Location(Residence(Carol)) is R/NEAR[City] $\wedge$ *SMALL[City]*

STAGES OF PRECISIATION



COINTENSIVE PRECISIATION

- *In general, precisiand of p is not unique. If $GC_1(p)$, …, $GC_n(p)$ are possible precisiands of p, then a basic question which arises is: which of the possible precisiands should be chosen to represent the meaning of p? There are two principal criteria which govern the choice: (a) Simplicity and (b) Cointension. Informally, the cointension of $GC_i(p)$, I=1, …, n, is the degree to which the meaning of $GC_i(p)$ approximates to the intended meaning of p. More specifically, $GC_i(p)$ is coextensive with p, or simply coextensive, if the degree to which the intension of $GC_i(p)$, with intension interpreted in its usual logical sense, approximates to the intended intension of p.*

DIGRESSION: COINTENSION

CONCEPT

C

human perception of C
p(C)

definition of C
d(C)

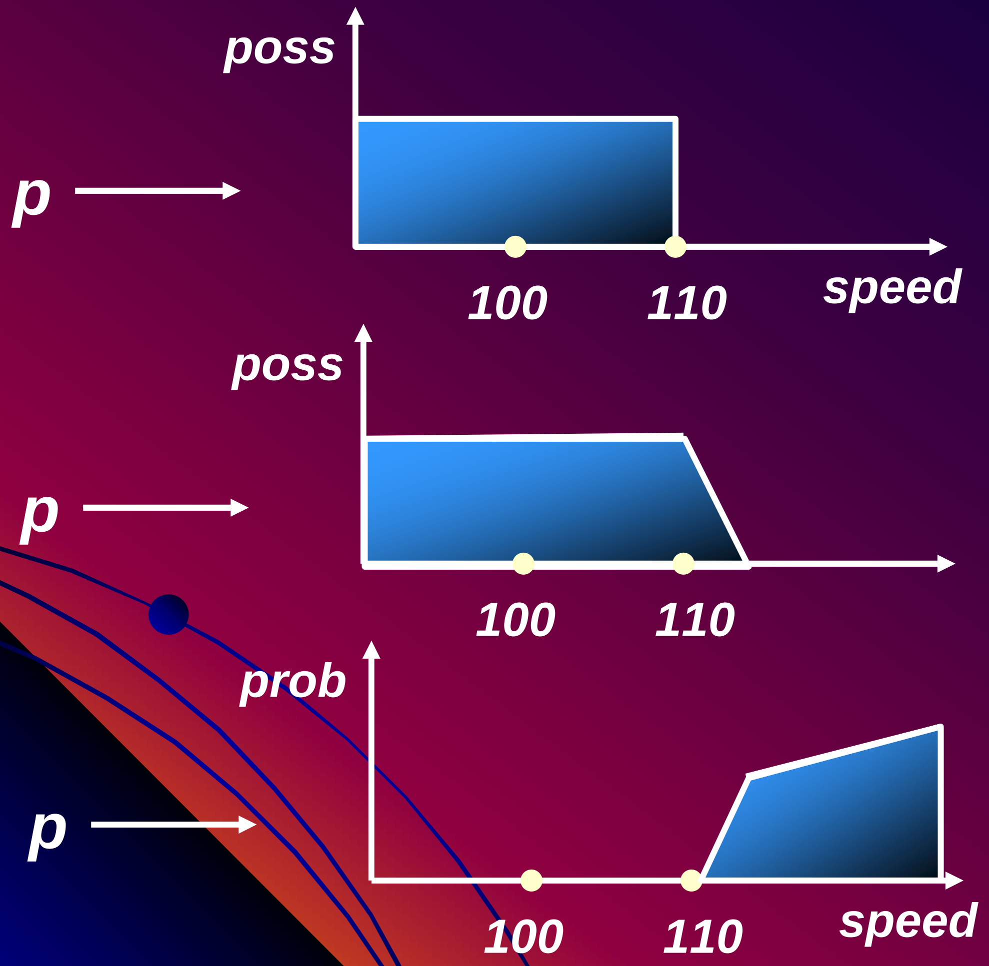
intension of p(C)

intension of d(C)

cointension: coincidence of the intended intensions of the definiendum, c, and the intension of the definiens d(C)

EXAMPLE

- *p: Speed limit is 100 km/h*

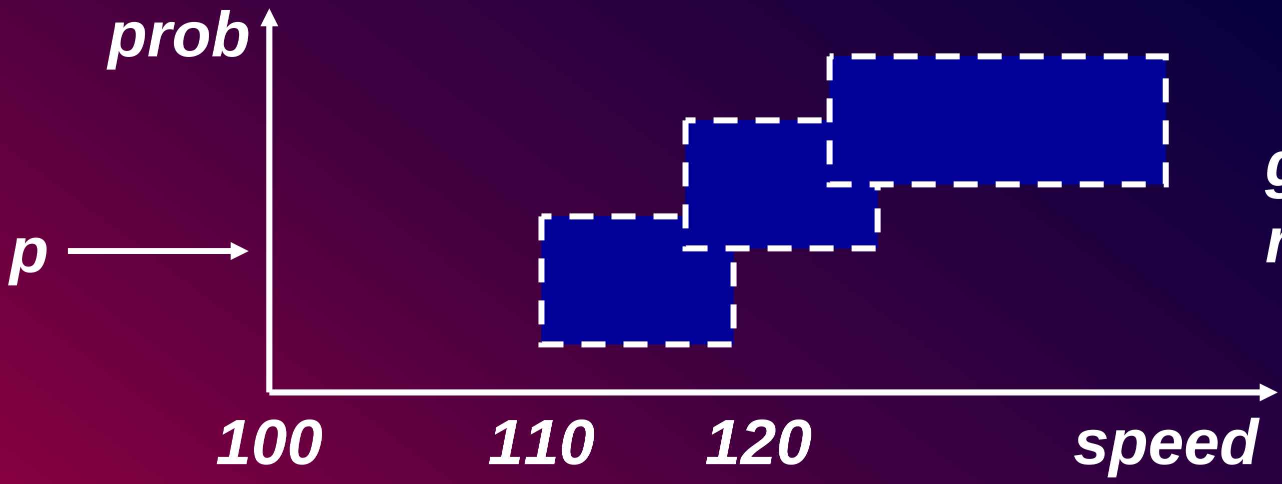


cg-precisiation
r = blank (possibilistic)

g-precisiation
r = blank (possibilistic)

g-precisiation
r = p (probabilistic)

CONTINUED



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

PRECISIATION

s-precisiation

conventional (degranulation)

$\underbrace{*\,a}_{\text{approximately } a} \xrightarrow{\text{precisiation}} a$

approximately a

g-precisiation

GCL-based (granulation)

$*a,\ p \xrightarrow{\text{precisiation}} \underbrace{X \text{ isr } R}_{\text{GC-form}}$

p — *proposition*

GC-form

common practice in probability theory

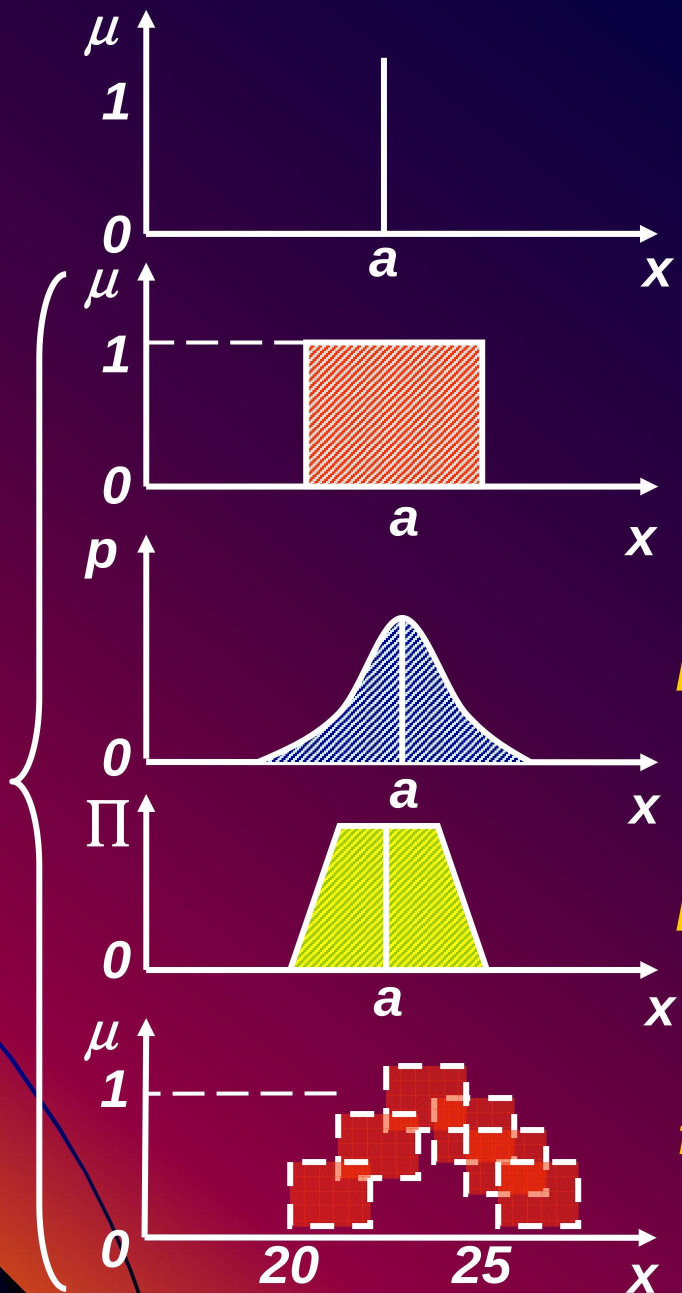
- *cg-precisiation: crisp granular precisiation*

PRECISIATION OF "approximately a," *a

s-precisiation

cg-precisiation

g-precisiation



singleton

interval

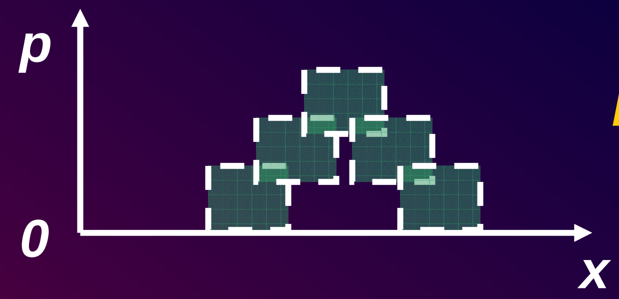
probability distribution

possibility distribution

fuzzy graph

CONTINUED

g-precisiation



bimodal distribution

GCL-based (maximal generality)

$$*a \xrightarrow{\text{g-precisiation}} \underbrace{X \text{ isr } R}_{\text{GC-form}}$$

THEORIES OF UNCERTAINTY

- *Standard, bivalent-logic-based probability theory, PT*

  Modality

  X isp P

  PT is unimodal

- *D-S theory*

  X isp P

  (X,Y) is Q

  D-S theory is unimodal if Q is crisp (bivalent)

  D-S theory is bimodal if Q is fuzzy

CONTINUED

- *Perception-based probability theory, PTp*

X isp P

X isv Q

X is R

PTp is trimodal

Unimodality of PT has the effect of limiting its problem-solving capability

PNL

THE BASICS OF PNL

- The point of departure in PNL is the key idea:
  - A proposition, p, drawn from a natural language, NL, is precisiated by expressing its meaning as a generalized constraint

$$p \longrightarrow X \text{ isr } R$$

- X: constrained (focal) variable
- isr — r: Identifier of modality (type of constraint)
- R: constraining relation

  - In general, X, R, r are implicit in p
  - precisiation of p ⟷ explicitation of X, R, r

SIMPLE EXAMPLE

- Monika is young ⟶ Age(Monika) is young

X   r (blank)   R

- Annotated representation

  X/Age(Monika) is R/young

KEY POINTS

- *A proposition is an answer to a question*

example:

p: Monika is young

is an answer to the question

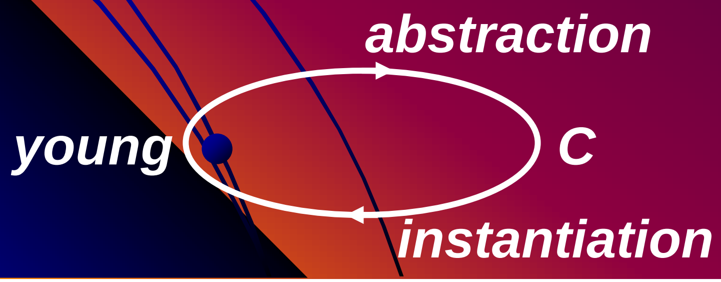
q: How old is Monika?

- *The concept of a generalized constraint serves as a basis for generalized-constraint-based semantics of natural languages*

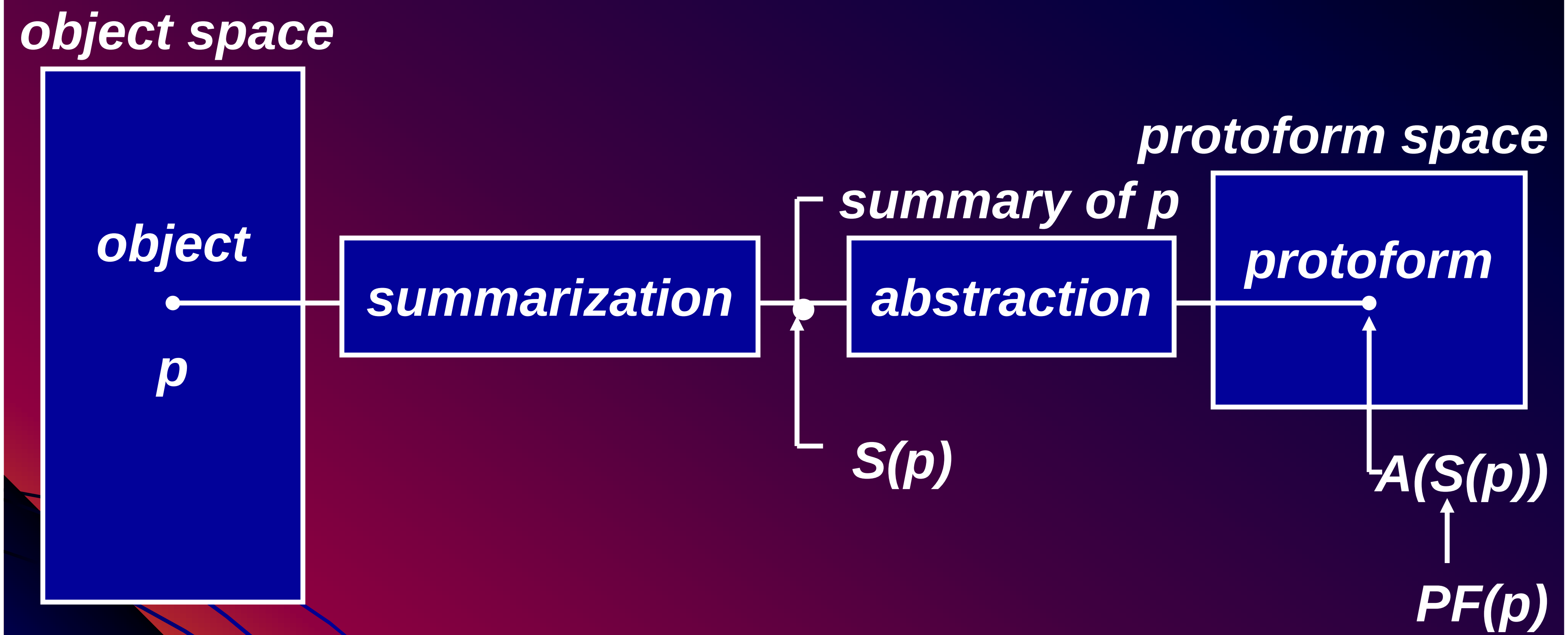
THE CONCEPT OF A PROTOFORM AND ITS BASIC ROLE IN KNOWLEDGE REPRESENTATION, DEDUCTION AND SEARCH

- *Informally, a protoform—abbreviation of prototypical form—is an abstracted summary. More specifically, a protoform is a symbolic expression which defines the deep semantic structure of a construct such as a concept, proposition, command, question, scenario, case or a system of such constructs*
- *Example:*

Monika is young $\longrightarrow$ *A(B) is C*



CONTINUED



PF(p): abstracted summary of p
deep structure of p

- *protoform equivalence*
- *protoform similarity*

EXAMPLES

- Monika is young → Age(Monika) is young A(B) is C

instantiation

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

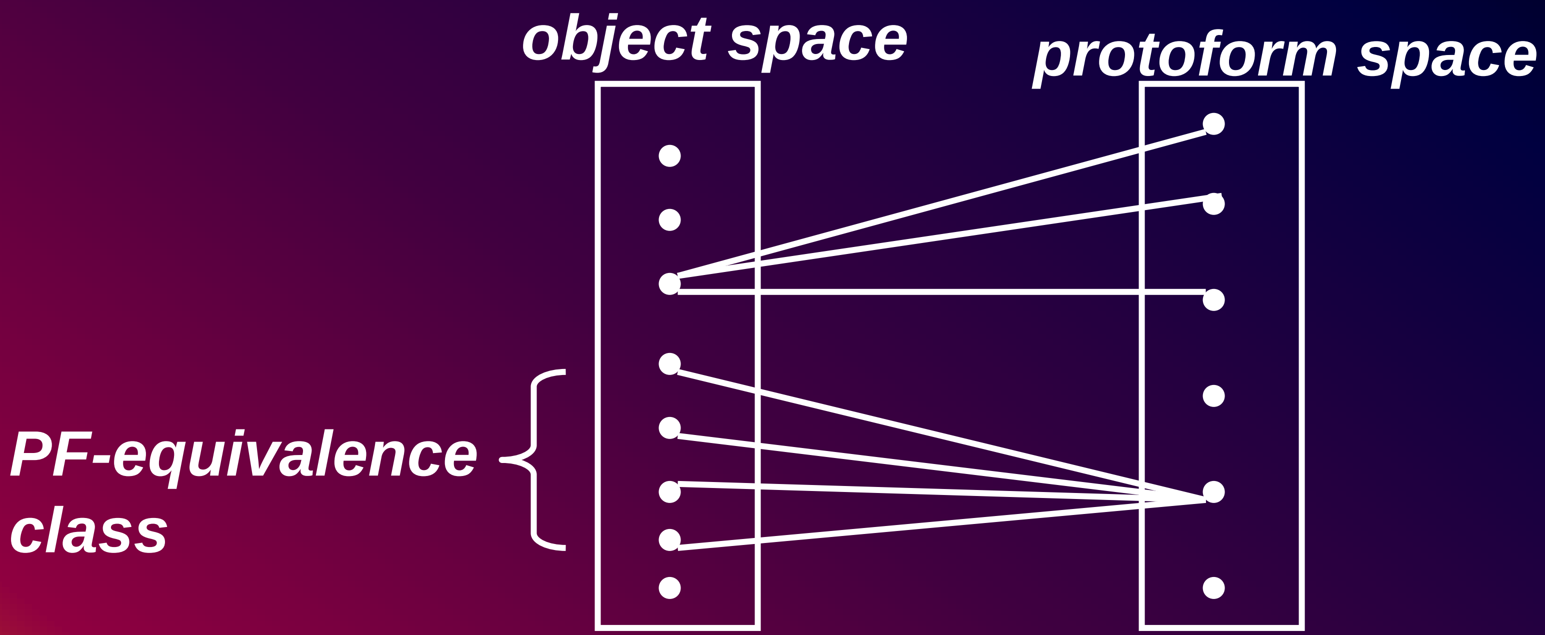
usually

6:15*

Return(Robert)

Time

PROTOFORMS



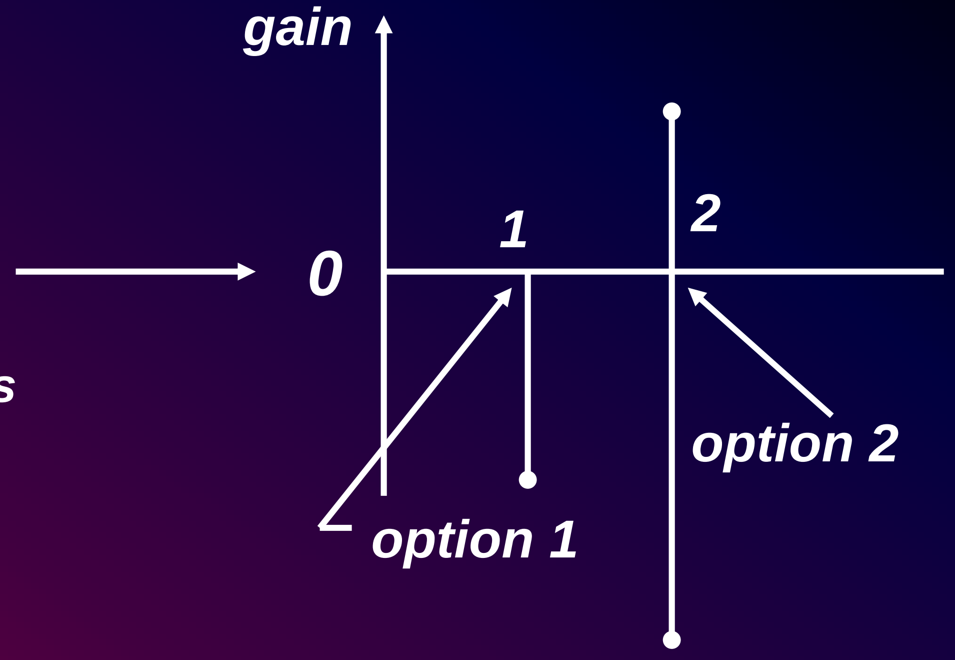
- *at a given level of abstraction and summarization, objects p and q are PF-equivalent if PF(p)=PF(q)*

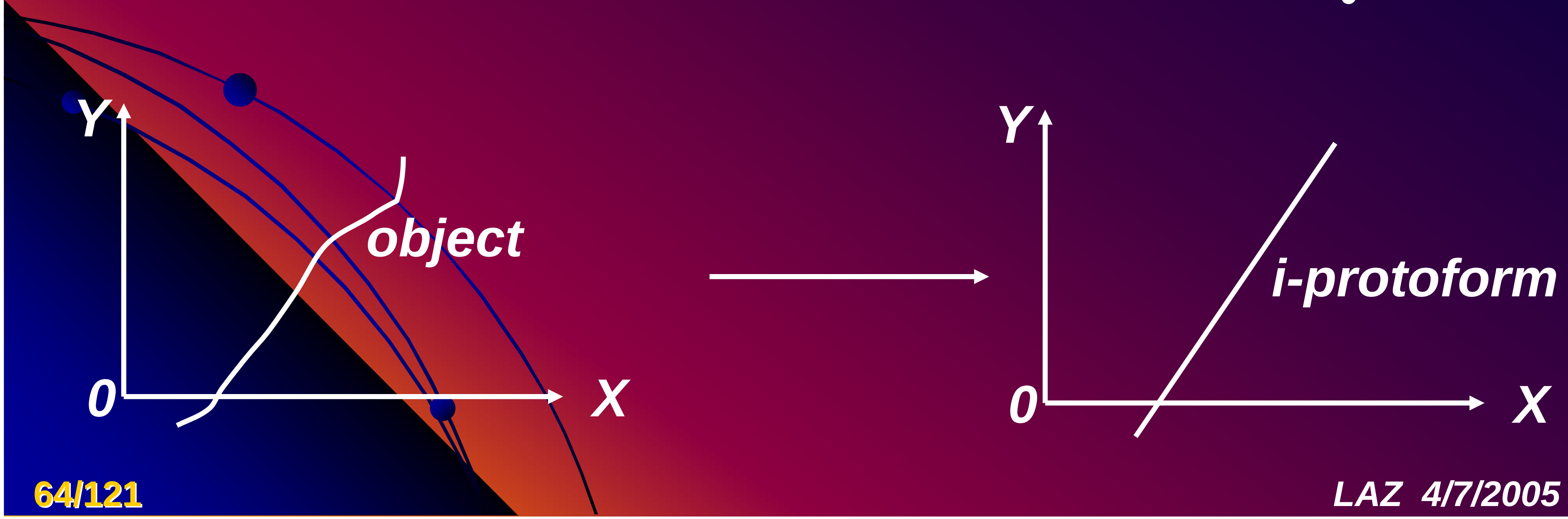
example

p: Most Swedes are tall	*Count (A/B) is Q*
q: Few professors are rich	*Count (A/B) is Q*

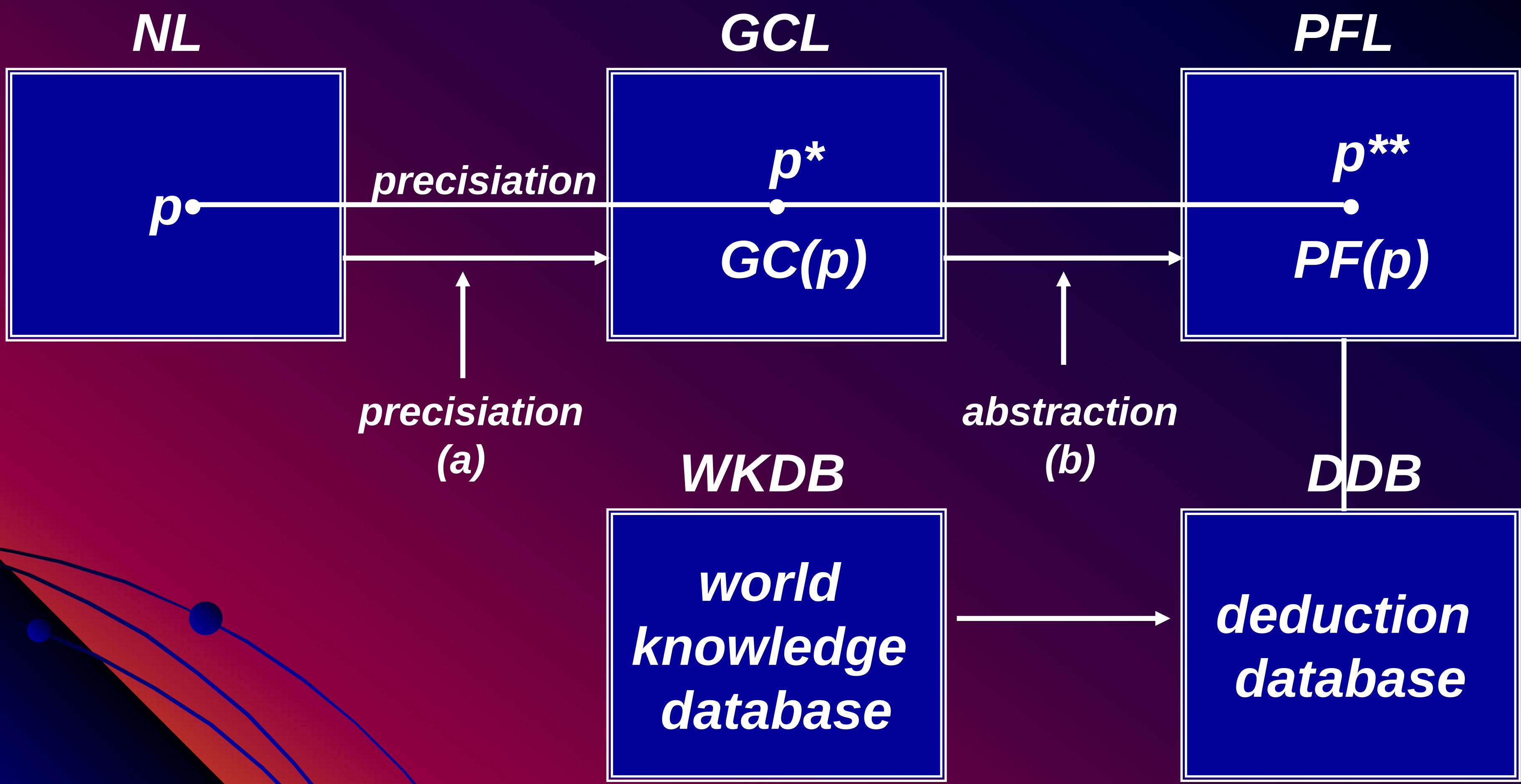
EXAMPLES

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?





BASIC STRUCTURE OF PNL

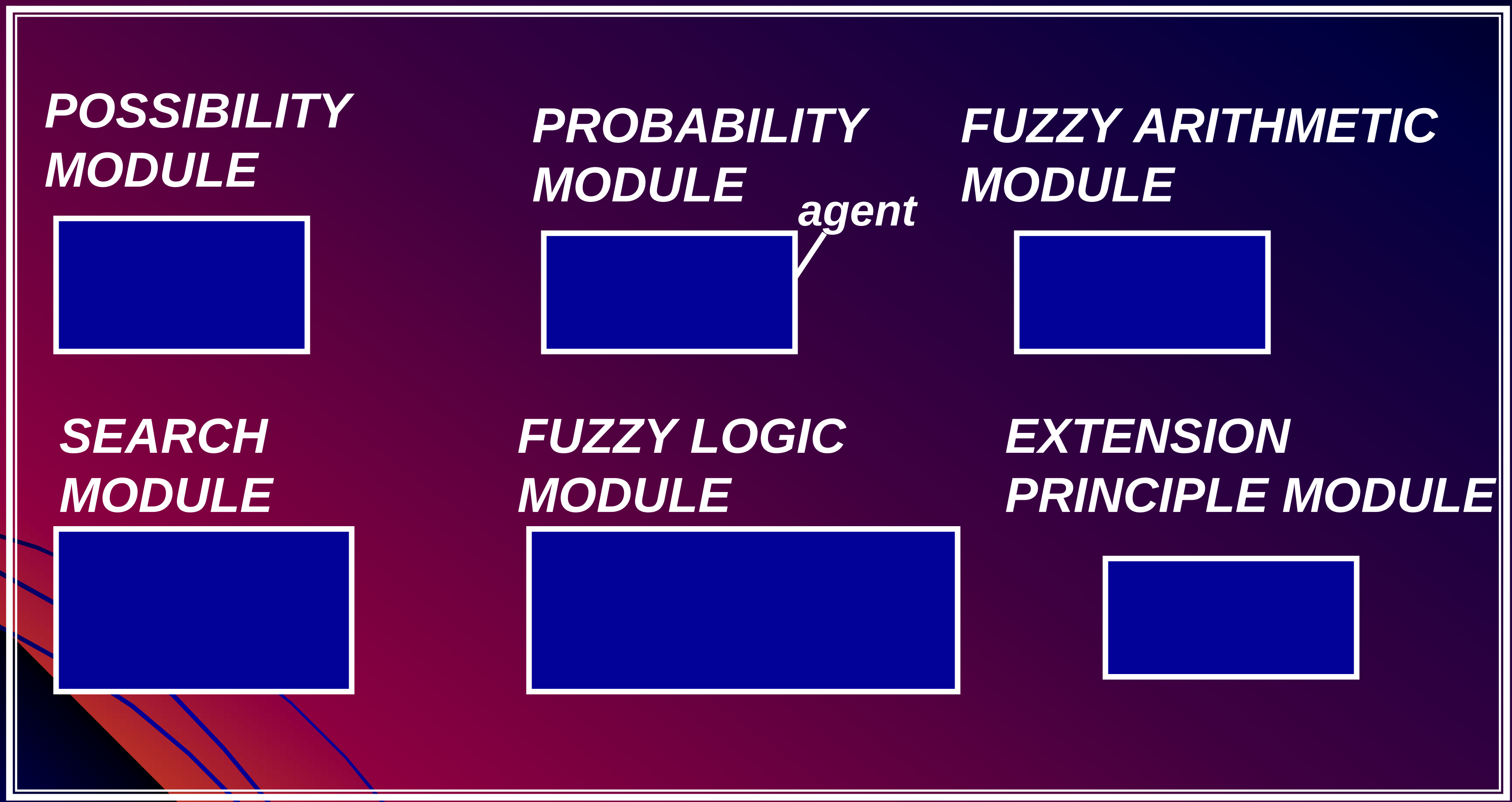


- *In PNL, deduction=generalized constraint propagation*

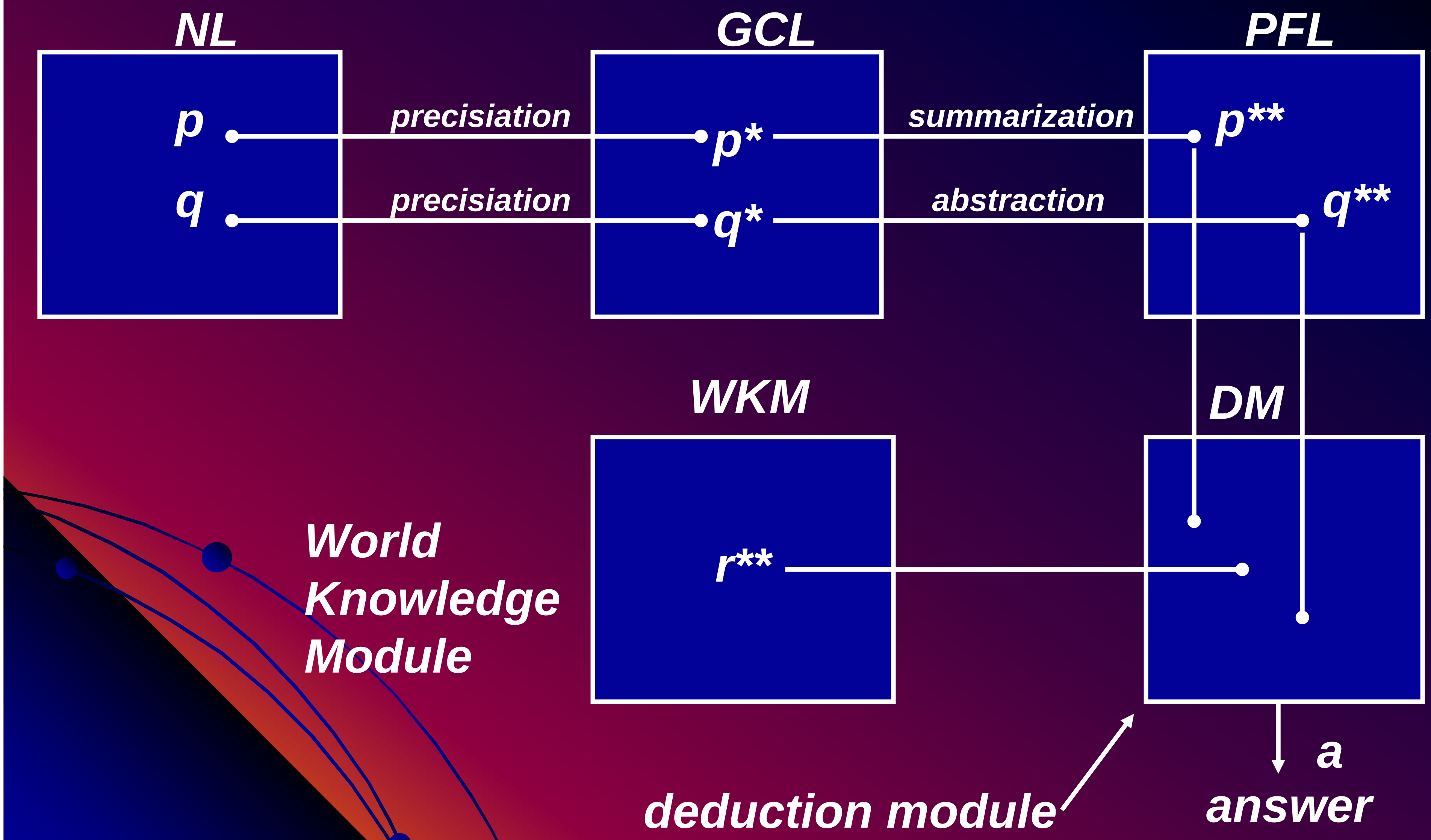
DDB: deduction database=collection of protoformal rules governing generalized constraint propagation

WKDB: PNL-based

MODULAR DEDUCTION DATABASE



PROTOFORM-BASED DEDUCTION



FORMAT OF PROTOFORMAL DEDUCTION RULES

- *protoformal rule*
  - *symbolic part*
  - *computational part*

PROTOFORM DEDUCTION RULE: GENERALIZED MODUS PONENS

fuzzy logic

classical

$$\frac{A \qquad A \rightarrow B}{B}$$

$$\frac{X \text{ is } A \qquad \text{If } X \text{ is } B \text{ then } Y \text{ is } C}{Y \text{ is } D}$$

← *symbolic*

computational 1 → $D = A \circ (B \times C)$ *(fuzzy graph; Mamdani)*

computational 2 → $D = A \circ (B \Rightarrow C)$ *(implication; conditional relation)*

PROTOFORMAL RULES OF DEDUCTION

examples

$$\frac{\begin{array}{c}X \text{ is } A\\ (X, Y) \text{ is } B\end{array}}{Y \text{ is } A \circ B}$$

$$\mu_{A \circ B}(v) = \max_u(\mu_A(u) \wedge \mu_B(u,v))$$

symbolic part

computational part

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

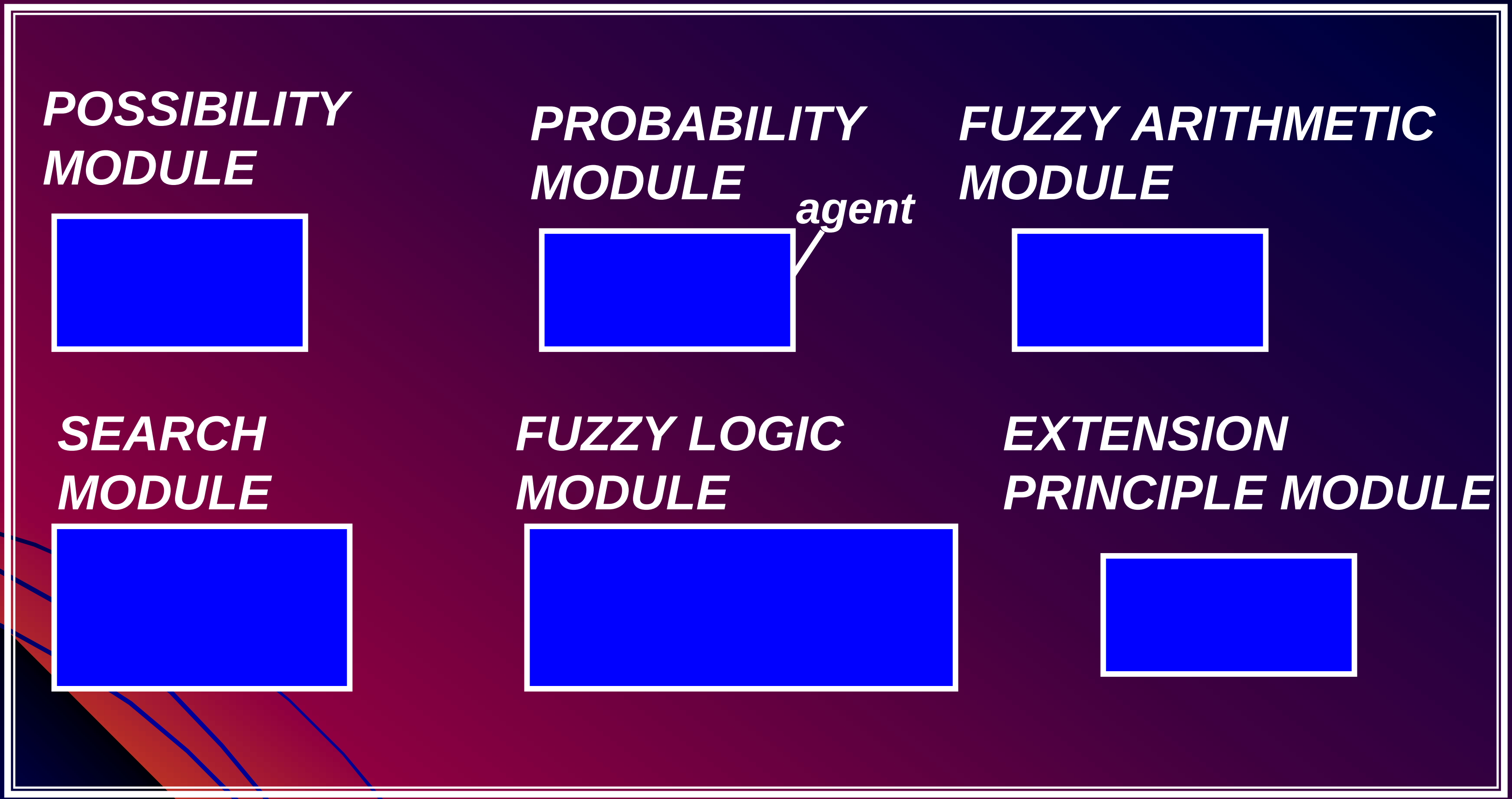
$$\mu_D(u) = \max_q(\mu_B(\int_U \mu_A(u)g(u)du))$$

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

$$\int_U g(u)du = 1$$

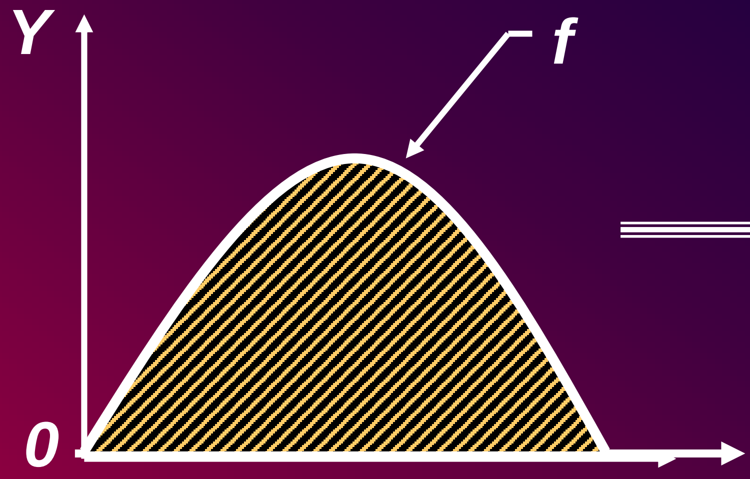
PROTOFORM-BASED DEDUCTION

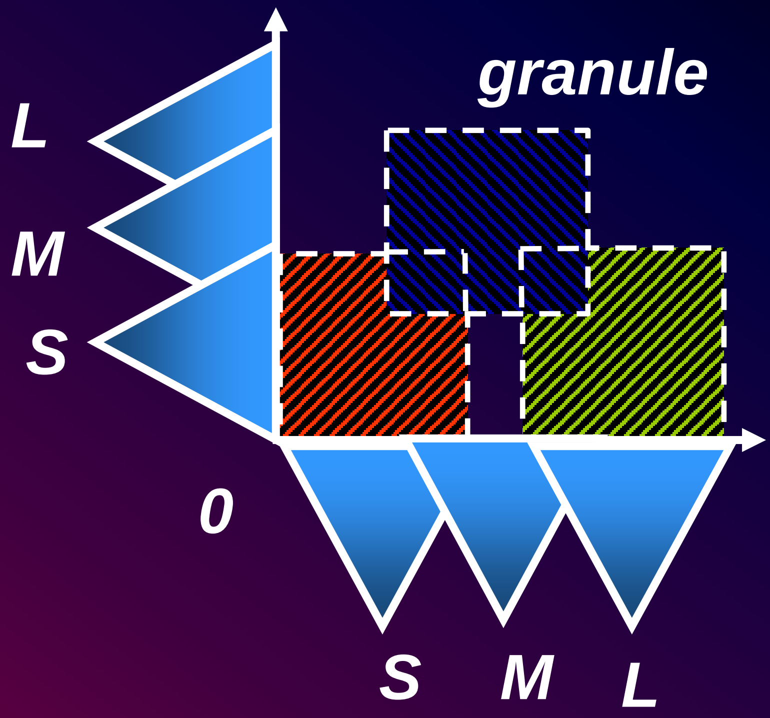
MODULAR DEDUCTION DATABASE

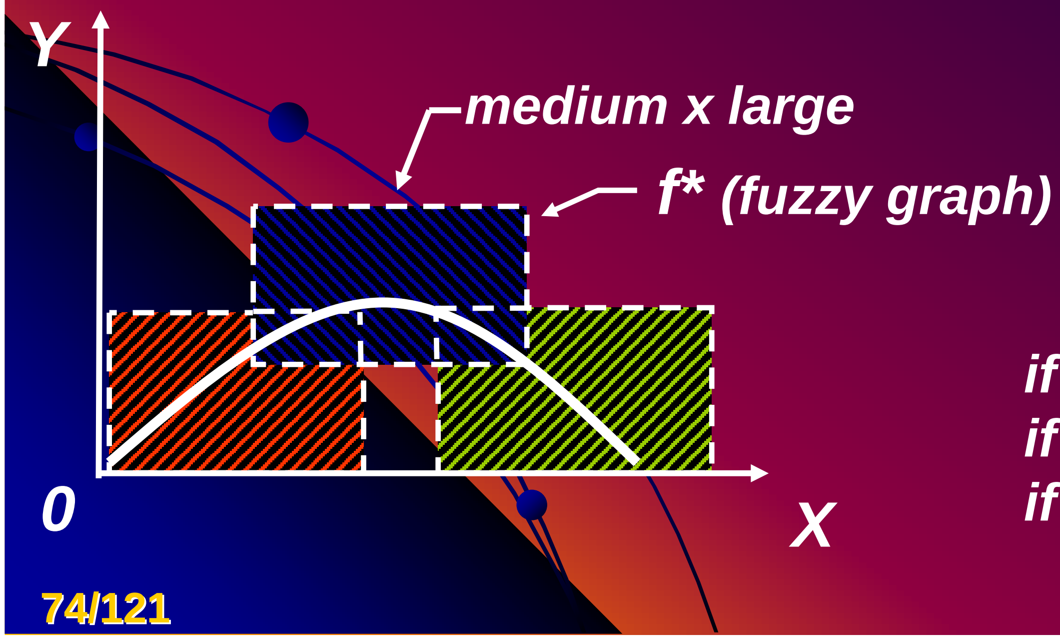


PROBABILITY MODULE

PERCEPTION OF A FUNCTION



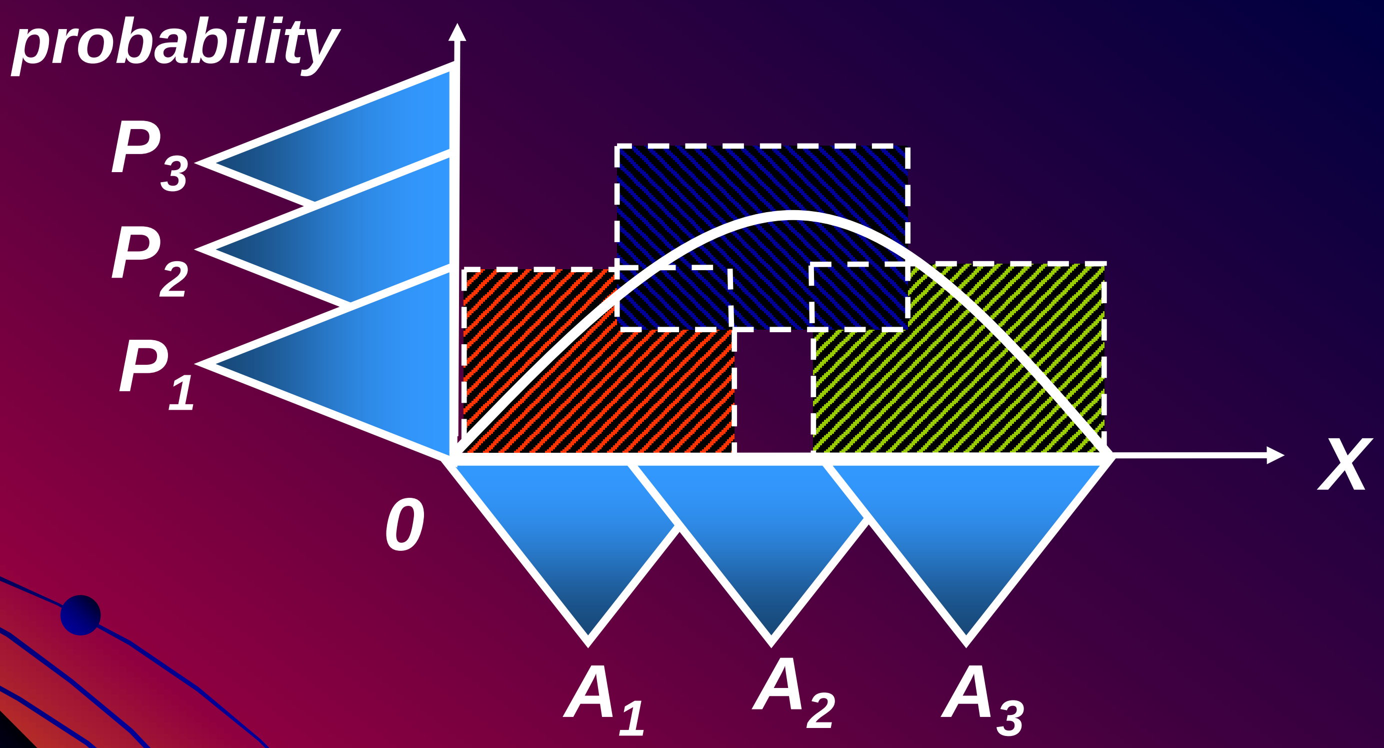




$f \xrightarrow{\text{perception}} f^*$:

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

BIMODAL DISTRIBUTION (PERCEPTION-BASED PROBABILITY DISTRIBUTION)



$P(X) = P_{i(1)} \backslash A_1 + P_{i(2)} \backslash A_2 + P_{i(3)} \backslash A_3$

Prob $\{X \text{ is } A_i\}$ is $P_{j(i)}$

P(X)= low\small + high\medium + low\large

CONTINUED

- function: if X is small then Y is large +…

  (X is small, Y is large)
- probability distribution: low \ small + low \ medium + high \ large +…
- Count \ attribute value distribution: 5* \ small + 8* \ large +…

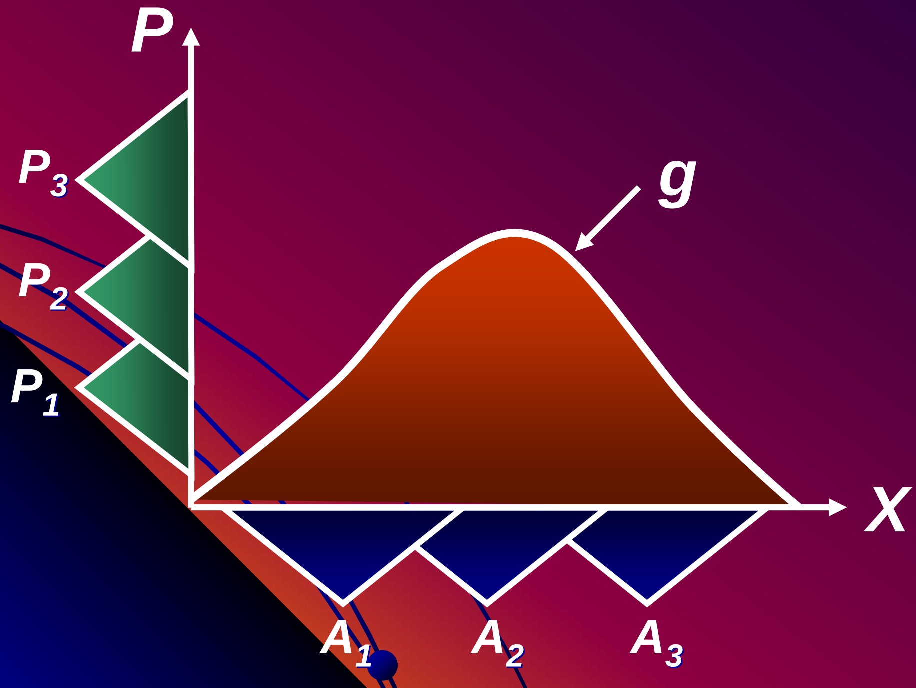
PRINCIPAL RATIONALES FOR F-GRANULATION

- detail not known
- detail not needed
- detail not wanted

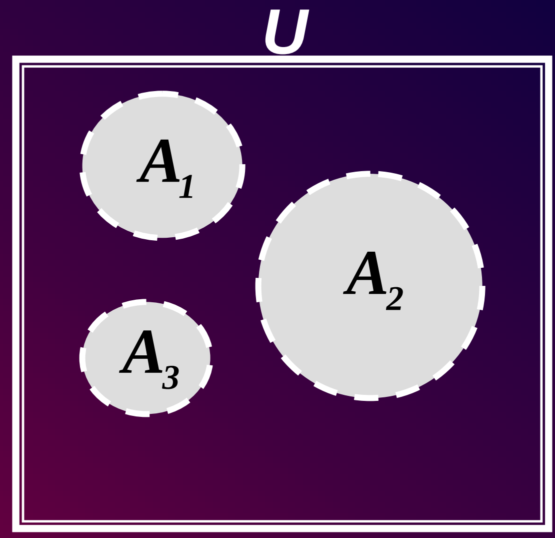
BIMODAL PROBABILITY DISTRIBUTIONS (LAZ 1981)

(a) possibility\probability

(b) probability\\possibility



$$P^* = P_1 \backslash A_1 + \cdots + P_n \backslash A_n$$

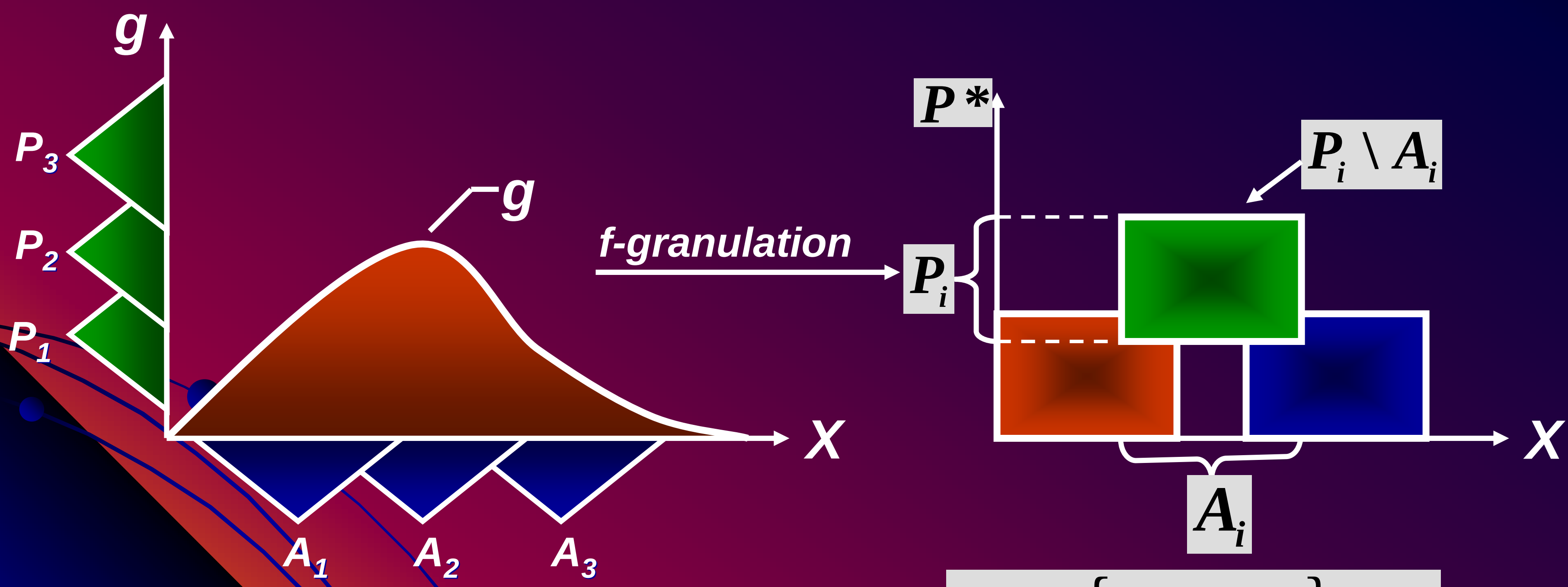


$$P = P_1 \backslash\backslash A_1 + \cdots + P_n \backslash\backslash A_n$$

BIMODAL PROBABILITY DISTRIBUTION

X*: a random variable taking values in *U

g*: probability density function of *X



$$P^* = \Sigma_i P_i \setminus A_i$$

$$\mathrm{Prob}\{X \text{ is } \mathrm{A_i}\} \text{ is } \mathrm{P_i}$$

$$\mathrm{Pr}\,ob\{\mathrm{X} \text{ is } A_i\} = \int_U \mu_{A_i}(u) g(u) du$$

CONTINUED

P defines a possibility distribution of g*

$$\pi(g) = \mu_{P_i}\left(\int_U \mu_{A_i}(u)g(u)du\right) \wedge \cdots \wedge \mu_{P_u}\left(\int_U \mu_{A_n}(u)g(u)du\right)$$

problems

a) what is the probability of a perception-based event A in U

b) what is the perception-based expected value of X

PROBABILITY OF A PERCEPTION-BASED EVENT

problem: **Prob {X is A} is ?B** ***knowing $\pi(g)$***

$$\text{Prob}\{\text{X is A}\} = \int_U \mu_A(u)g(u)du = f(g)$$

Extension Principle

$$\frac{\pi_1(g)}{\pi_2(f(g))}$$

$$\pi_2(v) = sup_g\, \pi_1(g)$$

subject to: $v = f(g)$

CONTINUED

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

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

subject to

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

EXPECTED VALUE OF A BIMODAL PD

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

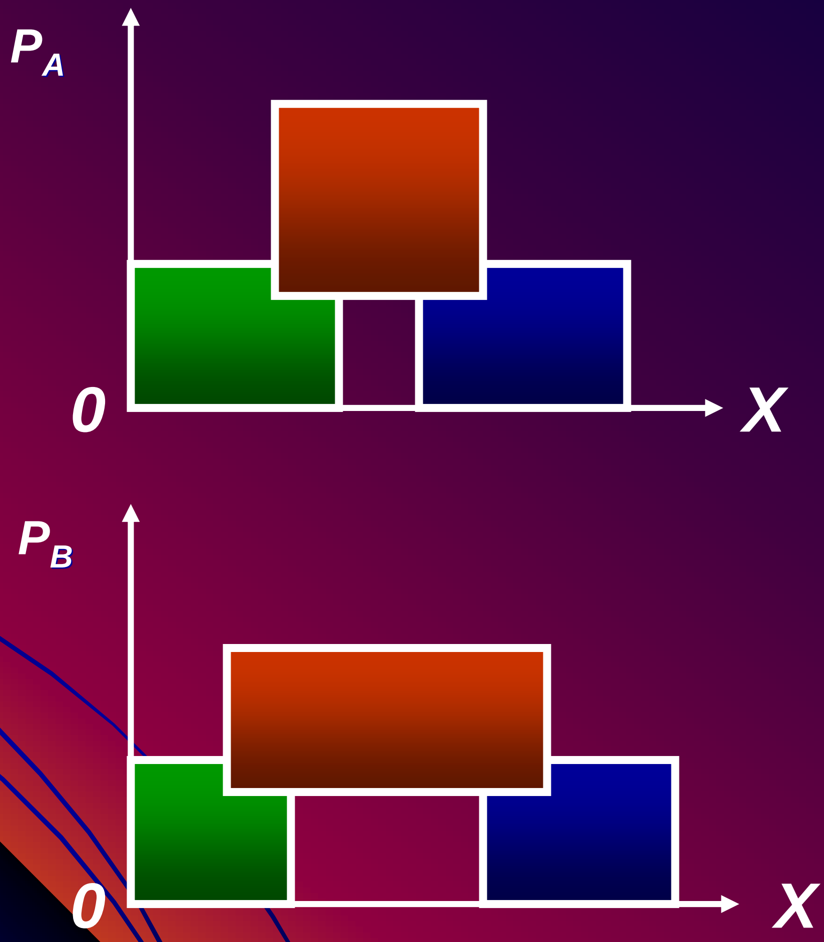
Extension Principle

$$\mu_{E(P^*)}(v) = \sup_g \left(\mu_{p_1}\left(\int_U \mu_{A_1}(u)g(u)du\right) \wedge \cdots \wedge \mu_{P_n}\left(\int_U \mu_{A_n}(u)g(u)du\right)\right)$$

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

PERCEPTION-BASED DECISION ANALYSIS

ranking of f-granular probability distributions



maximization of expected utility ⟶ *ranking of fuzzy numbers*

USUALITY CONSTRAINT PROPAGATION RULE

X: random variable taking values in U
g: probability density of X

$$\frac{X \text{ isu } A}{\text{Prob }\{X \text{ is } B\} \text{ is } C}$$

X isu A ⟶ ***Prob {X is A} is usually*** ⟶

$$\pi(g) = \mu_{usually}\left(\int_U \mu_A(u)g(u)du\right)$$

$$\mu_C(v) = sup_g\left(\mu_{usually}\left(\int_U \mu_A(u)g(u)du\right)\right)$$

subject to:

$$v = \int_U \mu_B(u)g(u)du$$

PROBABILITY MODULE

X: real-valued random variable

g: probability density function of X

A_1, …, A_n, A: perception-based events in U

P_1, …, P_n, P: perception-based probabilities in U

$$\frac{\begin{array}{c}\text{Prob } \{X \text{ is } A_1\} \text{ is } P_{j(1)} \\ \cdots \\ \text{Prob } \{X \text{ is } A_n\} \text{ is } P_{j(n)}\end{array}}{\text{Prob } \{X \text{ is } A\} \text{ is } P}$$

CONTINUED

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

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

subject to:

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

PROBABILITY MODULE (CONTINUED)

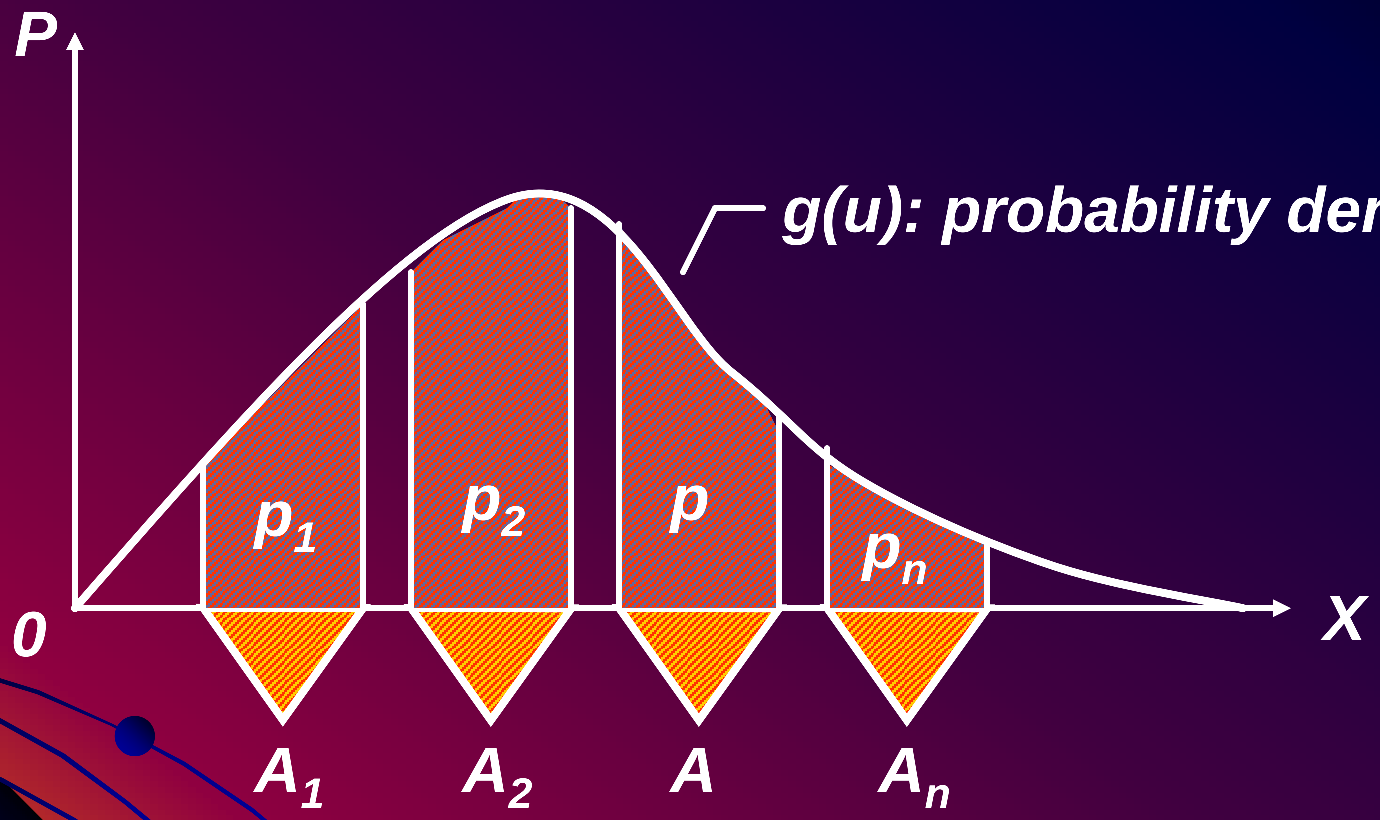
$$\frac{\begin{array}{c} X \text{ isp } P \\ Y = f(X) \end{array}}{Y \text{ isp } f(P)}$$

$$\frac{\text{Prob } \{X \text{ is } A\} \text{ is } P}{\text{Prob } \{f(X) \text{ is } B\} \text{ is } Q}$$

$$\frac{\begin{array}{c} X \text{ isp } P \\ (X,Y) \text{ is } R \end{array}}{Y \text{ isrs } S}$$

$$\frac{\begin{array}{c} X \text{ isu } A \\ Y = f(X) \end{array}}{Y \text{ isu } f(A)}$$

INTERPOLATION OF BIMODAL DISTRIBUTIONS



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)

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

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

USUALITY SUBMODULE

CONJUNCTION

$$\frac{\begin{array}{c} X \text{ is } A \\ X \text{ is } B \end{array}}{X \text{ is } A \cap B}$$

$$\frac{\begin{array}{c} X \text{ isu } A \\ X \text{ isu } B \end{array}}{X \text{ isr } A \cap B}$$

- *determination of r involves interpolation of a bimodal distribution*

USUALITY CONSTRAINT

X is A
X is B

X is A ⋂ B

X isu A
X isu B

X isp P

(A ⋂ B) ispv Q

g: probability density function of X

π(g): possibility distribution function of g

- $\pi(g) = \sup_g (\mu_{usually} (\int_U g(u)\mu_A(u)du) \wedge \mu_{usually} (\int_U g(u)\mu_B(u)du))$

subject to: $\int_U g(u)du = 1$

$\mu_Q(v) = \sup_g (\pi(g))$

subject to: $v = \int_U g(u)(\mu_A(u) \wedge \mu_B(u))du$

USUALITY — QUALIFIED RULES

$$\frac{X \text{ isu } A}{X \text{ isun } (\text{not } A)}$$

$$\frac{\begin{array}{c} X \text{ isu } A \\ Y = f(X) \end{array}}{Y \text{ isu } f(A)}$$

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

USUALITY — QUALIFIED RULES

X isu A

Y isu B

Z = f(X,Y)

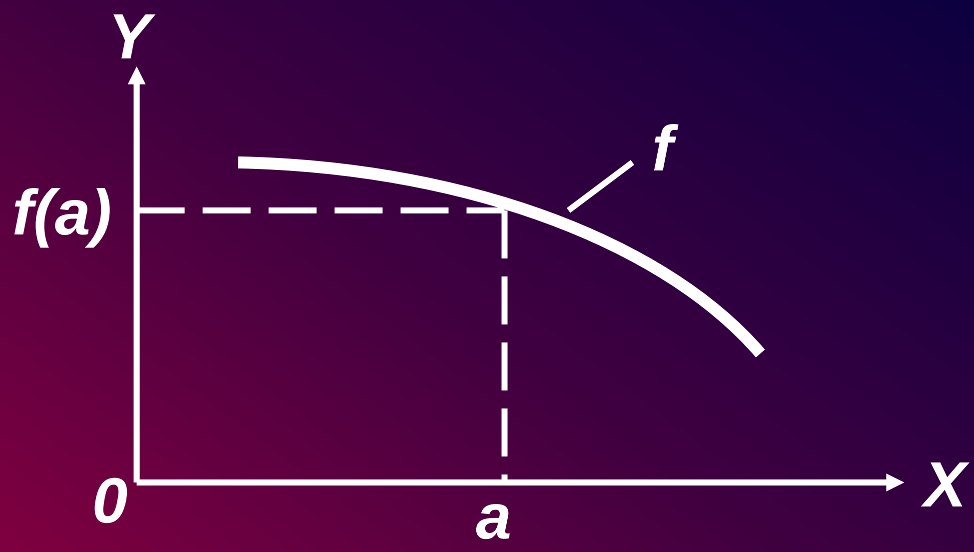
Z isu f(A, B)

$$\mu_Z(w) = \sup_{u,v|w=f(u,v)}(\mu_A(u) \wedge \mu_B(v))$$

EXTENSION PRINCIPLE MODULE

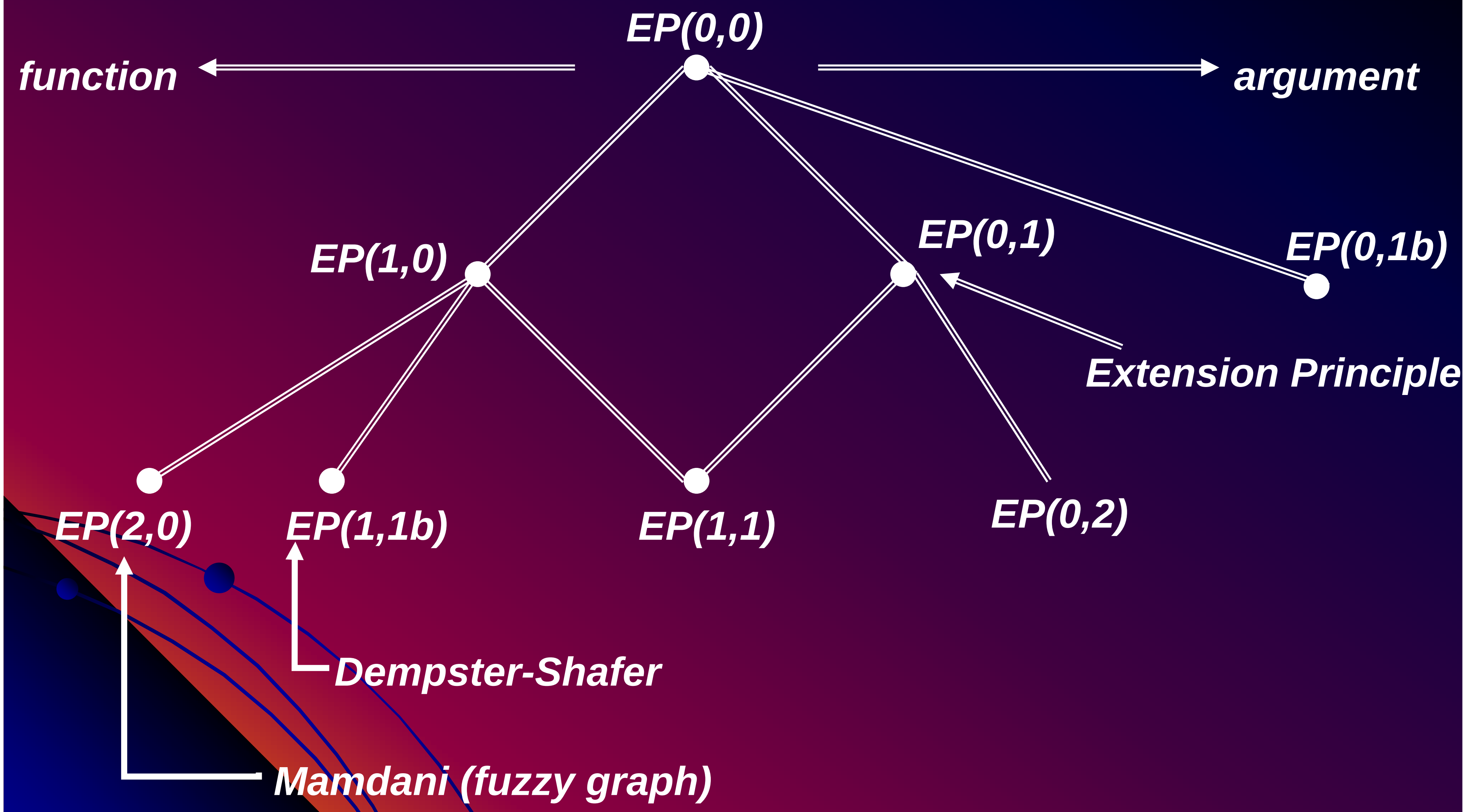
PRINCIPAL COMPUTATIONAL RULE IS THE EXTENSION PRINCIPLE (EP)

point of departure: function evaluation

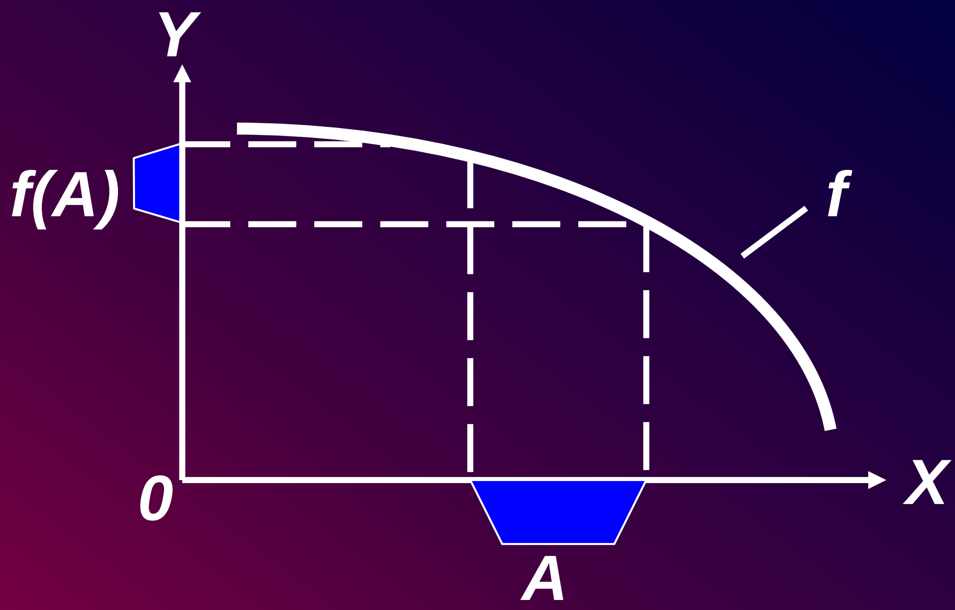


$$\frac{\begin{array}{c}X=a\\ Y=f(X)\end{array}}{Y=f(a)}$$

EXTENSION PRINCIPLE HIERARCHY



VERSION EP(0,1) (1965; 1975)



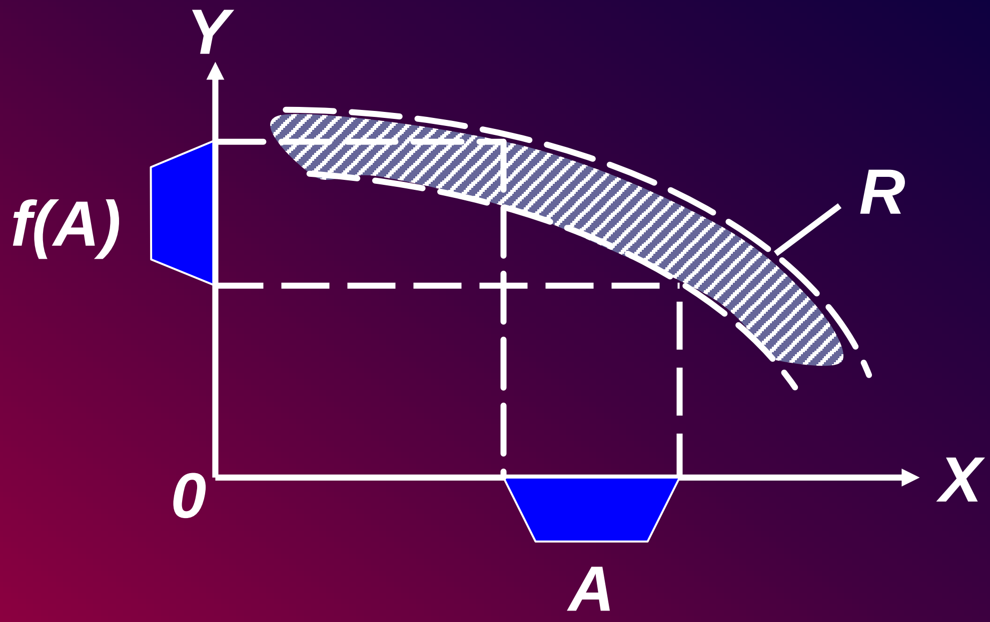
$$\frac{\begin{array}{c}X \text{ is } A \\ Y=f(X)\end{array}}{Y=f(A)}$$

$$\mu_{f(A)}(v) = \sup_u(\mu_A(u))$$

subject to

$$v = f(u)$$

VERSION EP(1,1) (COMPOSITIONAL RULE OF INFERENCE) (1965)



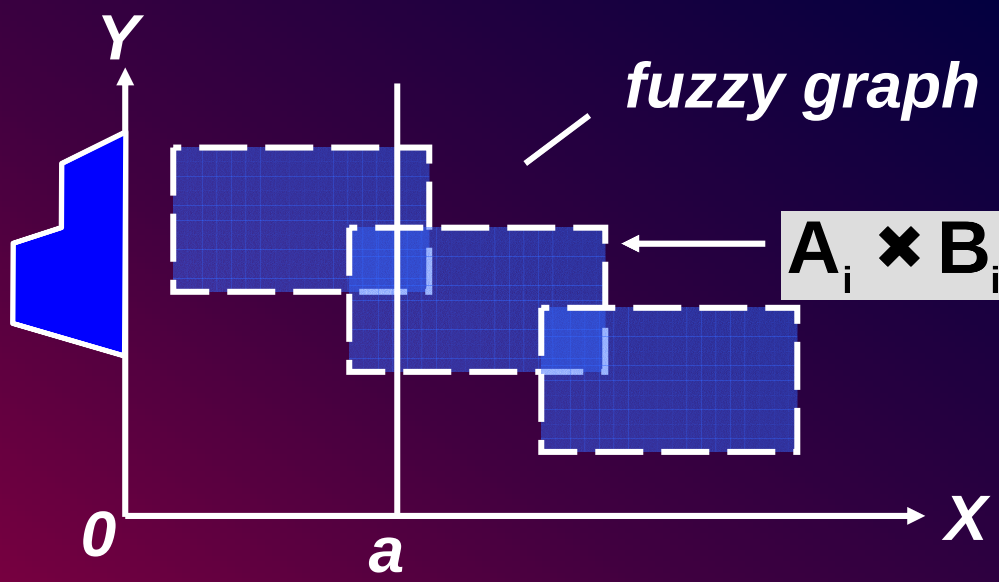
X is A

(X,Y) is R

Y is A ∘ R

$$\mu_Y(v) = \sup_u(\mu_A(u) \wedge \mu_R(u,v))$$

EXTENSION PRINCIPLE EP(2,0) (Mamdani)



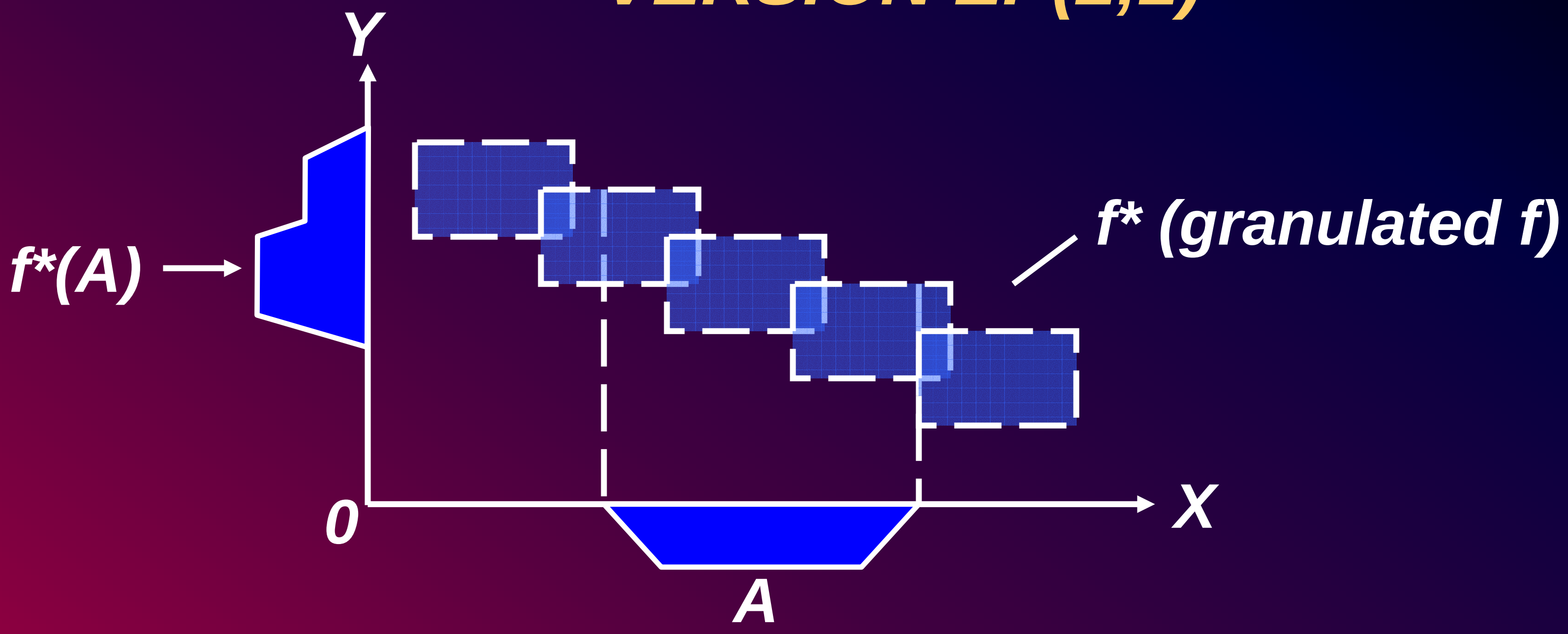
$f^* = \sum_i A_i \times B_i$

(if X is A_i then Y is B_i)

$X = a$

$Y = \sum_i {}^{\mu}A_i(a) \wedge B_i$

VERSION EP(2,1)



$$\frac{\begin{array}{c}X \text{ is } A \\ (X, Y) \text{ is } R\end{array}}{Y \text{ is } \Sigma_i m_i \wedge B_i}$$

$R = \Sigma_i A_i \times B_i$

$m_i = \sup_u (\mu_A(u) \wedge \mu_{Ai}(u))$: *matching coefficient*

VERSION EP(1,1b) (DEMPSTER-SHAFER)

$$X \text{ isp } (p_1\backslash u_1 + \ldots + p_u\backslash u_n)$$

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

$$Y \text{ isp } (p_1\backslash R(u_1) + \ldots + p_n\backslash R(u_n))$$

Y is a fuzzy-set-valued random variable

$$\mu_{R(u_i)}(v) = \mu_R(u_i, v)$$

VERSION GEP(0,0)

$$\frac{f(X) \text{ is } A}{g(X) \text{ is } g(f^{-1}(A))}$$

$$\mu_{g(f^{-1}(A))}(v) = \sup_u (\mu_A(f(u)))$$

subject to

$$v = g(u)$$

GENERALIZED EXTENSION PRINCIPLE

$f(X)$ is A
$g(Y)$ is B
$Z=h(X,Y)$

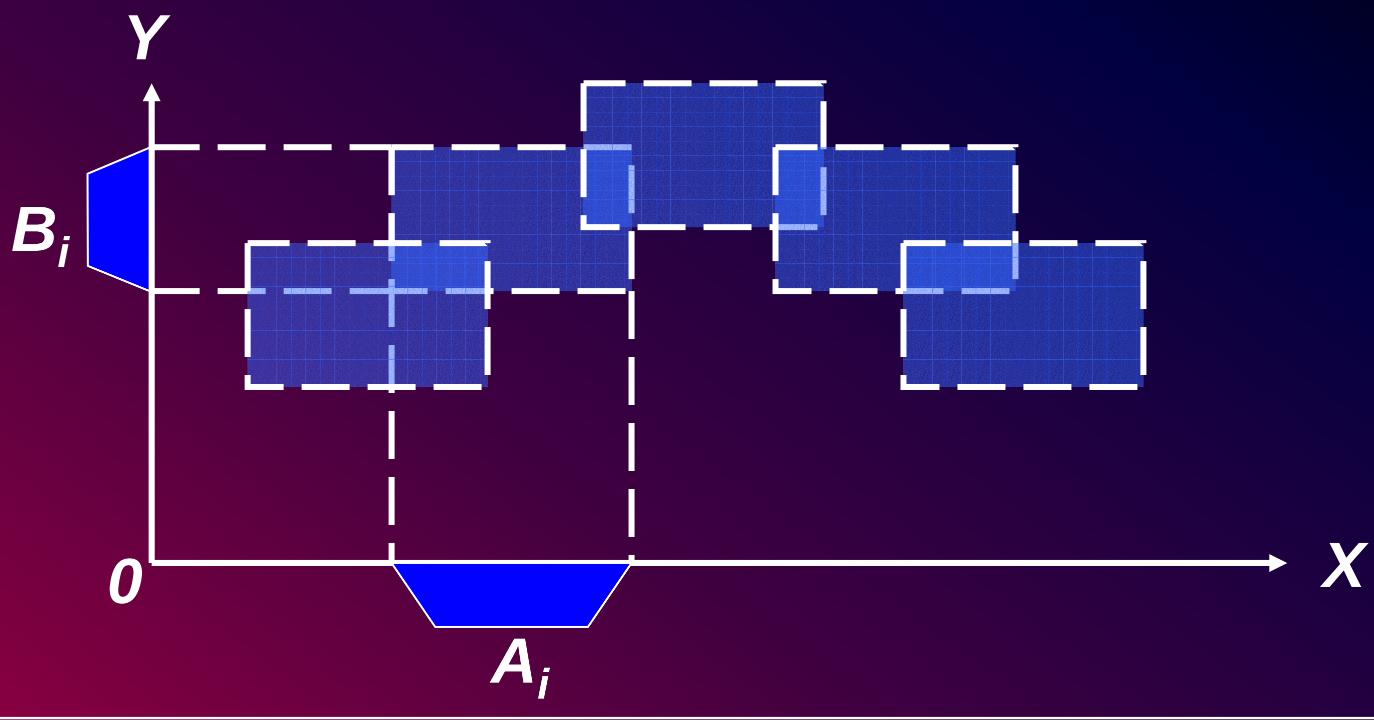
Z is $h\,(f^{-1}(A),\, g^{-1}\,(B))$

$$\mu_Z(w) = \sup_{u,v}\,(\mu_A(f(u)) \wedge \mu_B(g(u)))$$

subject to

$$w = h(u,v)$$

U-QUALIFIED EXTENSION PRINCIPLE



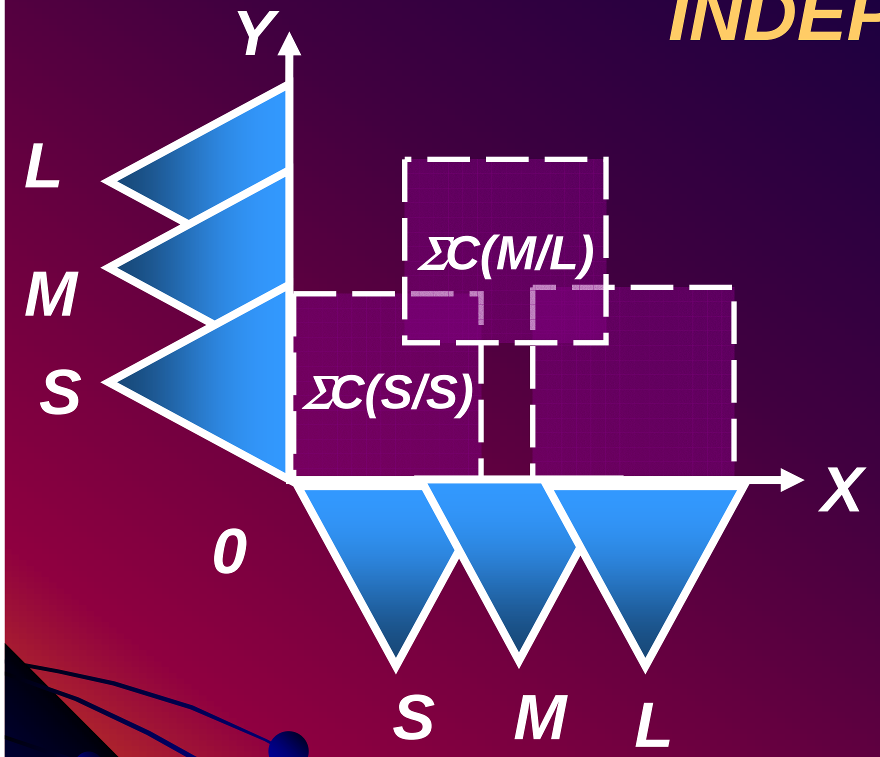
If X is A_i then Y isu B_i, $i=1,\ldots, n$

X isu A

Y isu $\Sigma_i\, m_i \wedge B_i$

$m = \sup_u (\mu_A(u) \wedge \mu_{A_i}(u))$: matching coefficient

PNL-BASED DEFINITION OF STATISTICAL INDEPENDENCE



contingency table

3	L/S	L/M	L/L
2	M/S	M/M	M/L
1	S/S	S/M	S/L
	1	2	3

$$\Sigma(M/L) = \frac{\Sigma C\ (M \times L)}{\Sigma C\ (L)}$$

- *degree of independence of Y from X= degree to which columns 1, 2, 3 are identical*

→ *PNL-based definition*

WHAT IS A RANDOM SAMPLE?

- In most cases, a sample is drawn from a population which is a fuzzy set, e.g., middle class, young women, adults
- In the case of polls, fuzziness of the population which is polled may reflect the degree applicability of the question to the person who is polled
- example (Atlanta Constitution 5-29-95)
  Is O.J. Simpson guilty?

Random sample of 1004 adults polled by phone.
61% said “yes.” Margin of error is 3%

- to what degree is this question applicable to a person who is n years old?

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?

PNL-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

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

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

CONTINUED

Step 2. Deduction via extension principle

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

subject to

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

DEDUCTION PRINCIPLE

- Point of departure: question, q
- Data: D = (X_1/u_1, …, 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, …, u_n, we could express Ans(q) as a function of the u_i

$$Ans(q)=g(u_1, \ldots, u_n) \qquad u=(u_1, \ldots, u_n)$$

- 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

$$f(u)=f(u_1, \ldots, u_n)$$

CONTINUED

- *Use the extension principle*

$$\mu_{Ans(q)}(v) = sup_u(ts(u))$$

subject to

$$v = g(u)$$

SUMMATION

- ***A basic difference between GTU and bivalent-logic-based theories of uncertainty relates to the role of natural languages. In GTU, the semantics of natural languages plays a pivotal role. The underlying reason is that GTU's capability to operate on perception-based information is directly dependent on GTU's ability to understand natural language, since a natural language is basically a system for describing perceptions.***
- ***Another basic difference relates to the conceptual framework of GTU. In GTU, the basic concepts, e.g., the concepts of independence are defined, for the most part, through the use of PNL. As a consequence, most of the basic concepts in GTU are context-dependent. All existing theories of uncertainty may be viewed as specializations of GTU.***

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

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

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

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

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 .

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.

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.

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.