Computing with Words and its Applications

Lotfi A. Zadeh

Computer Science Division
Department of EECS
UC Berkeley

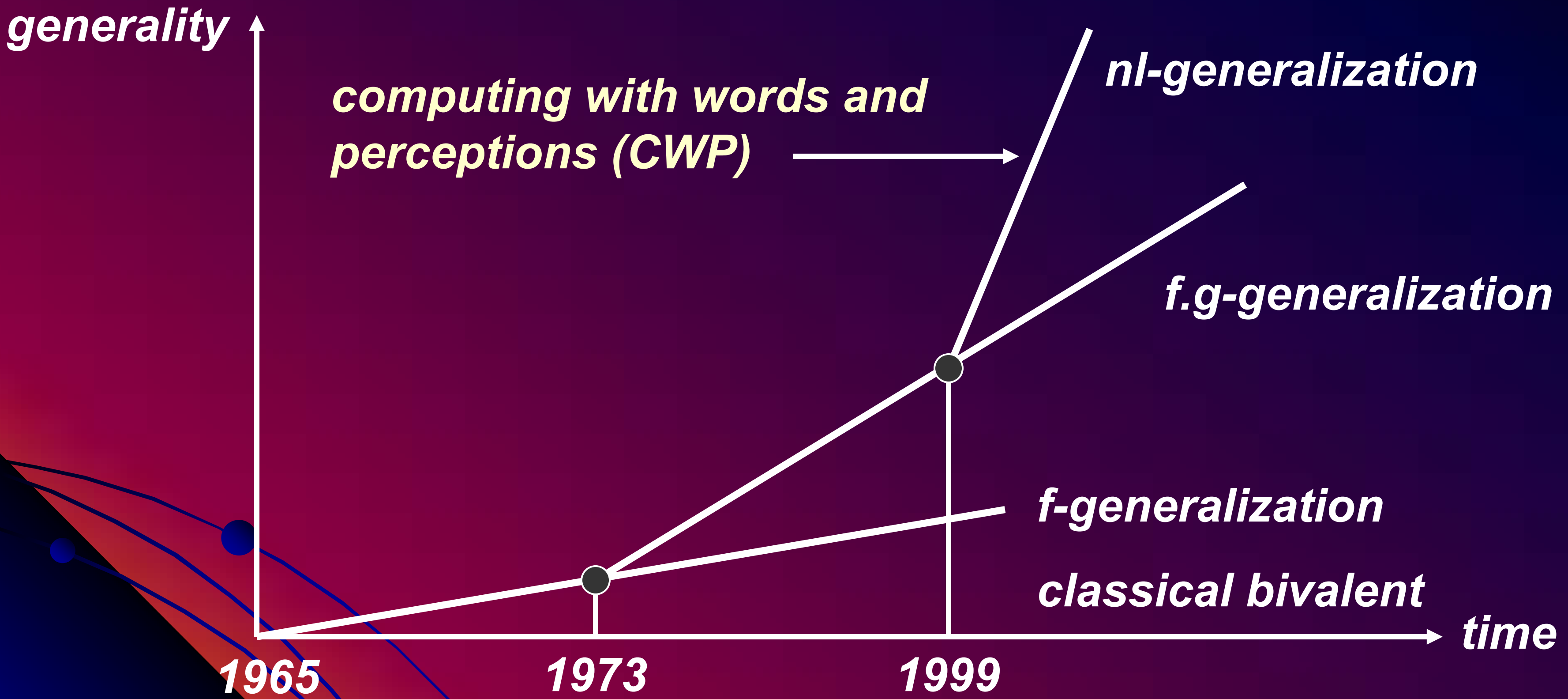
April 21, 2004

4th WSEAS International Conference on
Soft Computing, Optimization and Manufacturing Systems
Miami, Florida

www.wseas.org

BACKDROP

EVOLUTION OF FUZZY LOGIC—A PERSONAL PERSPECTIVE



1965: crisp sets → fuzzy sets
1973: fuzzy sets → granulated fuzzy sets (linguistic variable)
1999: measurements → perceptions

COMPUTING WITH WORDS (CW)
COMPUTING WITH WORDS AND PERCEPTIONS (CWP)

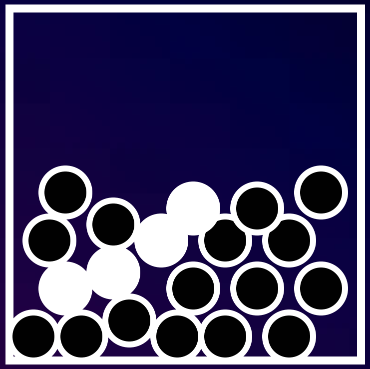
1973 — **CW** — **1999** — **CWP**

- 1973:
  - the concept of a linguistic variable
  - calculus of fuzzy if-then rules
- 1999:
  - perception-based information
  - computing with propositions

- CW: objects of computation are words
- CWP: objects of computation are words and perceptions
- example: usually Robert returns from work at about 6 pm
- What is the probability that Robert is home at 6:15pm?

WHAT IS CWP?

THE BALLS-IN-BOX PROBLEM



Version 1. Measurement-based

- *a box contains 20 black and white balls*
- *over 70% 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 drawn at random is white?*

CONTINUED

Version 2. Perception-based

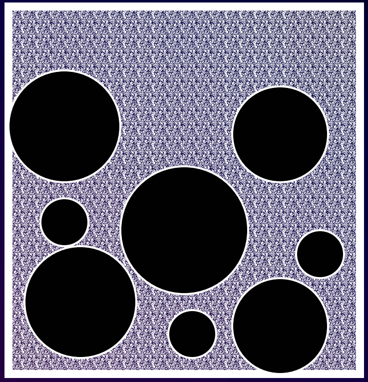
- *a box contains about 20 black and white balls*
- *most are black*
- *there are several times as many black balls as white balls*

- *what is the number of white balls?*
- *what is the probability that a ball drawn at random is white?*

CONTINUED

Version 3. Perception-based

box



- *a box contains about 20 black balls of various sizes*
- *most are large*
- *there are several times as many large balls as small balls*

- *what is the number of small balls?*
- *what is the probability that a ball drawn at random is small?*

MEASUREMENT-BASED

- 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 1)

- 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 1)

- *measurement-based*

  X = number of black balls

  Y_2 number of white balls

  $X \geq 0.7 \cdot 20 = 14$

  $X + Y = 20$

  $X = 3Y$

  $X = 15$; $Y = 5$

  $p = 5/20 = .25$

- *perception-based*

  X = number of black balls

  Y = number of white balls

  *X = most × 20**

  *X = several *Y*

  *X + Y = 20**

  P = Y/N

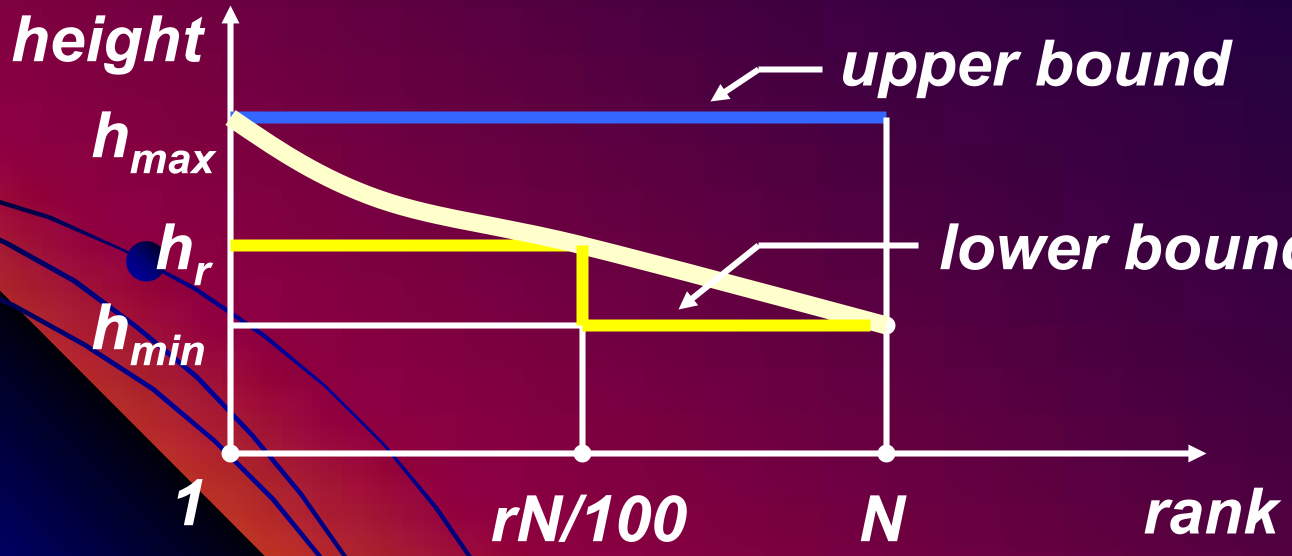
THE TALL SWEDES PROBLEM

MEASUREMENT-BASED	PERCEPTION-BASED
IDS p_1: Height of Swedes ranges from h_{min} to h_{max}	p_1: Height of Swedes ranges from approximately h_{min} to approximately h_{max}
p_2: Over 70% are taller than h_{tall}	p_2: Most are tall (taller than approximately h_{tall})
TDS q_1: What fraction are less than h_{tall}	q_1: What fraction are not tall (shorter than approximately h_{tall})
q_2: What is the average height of Swedes	q_2: What is the average height of Swedes

$X \longrightarrow$ approximately X

THE TALL SWEDES PROBLEM

- *measurement-based version*
  - *height of Swedes ranges from h_{min} to h_{max}*
  - *over r% of Swedes are taller than h_r*
  - *what is the average height, h_{ave}, of Swedes?*



$$rh_r+(i-r)h_{min} \le h_{ave} \le h_{max}$$

CONTINUED

- *most Swedes are tall* → $\int_{h_{min}}^{h_{max}} g(u)\mu_{tall}(u)du$ *is most*

  ↑ *fraction of tall Swedes*

- *average height* → $\int_{h_{min}}^{h_{max}} ug(u)du$

- *constraint propagation*

$$\frac{\int_{h_{min}}^{h_{max}} g(u)\mu_{tall}(u)du \text{ is most}}{\int_{h_{min}}^{h_{max}} ug(u)du \text{ is ? } h_{ave}}$$

CONTINUED

- *Solution: application of extension principle*

$$\mu_{h_{ave}}(v) = \sup_g \left(\mu_{most}\left(\int_{h_{min}}^{h_{max}} g(u)\mu_{tall}(u)du \right) \right)$$

subject to

$$\int_{h_{min}}^{h_{max}} g(u)du = 1$$

$$v = \int_{h_{min}}^{h_{max}} ug(u)du$$

BASIC PERCEPTIONS

attributes of physical objects

- *distance*
- *time*
- *speed*
- *direction*
- *length*
- *width*
- *area*
- *volume*
- *weight*
- *height*
- *size*
- *temperature*

sensations and emotions

- *color*
- *smell*
- *pain*
- *hunger*
- *thirst*
- *cold*
- *joy*
- *anger*
- *fear*

concepts

- *count*
- *similarity*
- *cluster*
- *causality*
- *relevance*
- *risk*
- *truth*
- *likelihood*
- *possibility*

DEEP STRUCTURE OF PERCEPTIONS

- *perception of likelihood*
- *perception of truth (compatibility)*
- *perception of possibility (ease of attainment or realization)*
- *perception of similarity*
- *perception of count (absolute or relative)*
- *perception of causality*

subjective probability = quantification of perception of likelihood

MEASUREMENT-BASED VS. PERCEPTION-BASED INFORMATIO

INFORMATION

measurement-based	perception-based
numerical	linguistic

• it is 35 C°	• It is very warm
• Eva is 28	• Eva is young
• probability is 0.8	• probability is high
•	• it is cloudy
•	• traffic is heavy
	• it is hard to find parking near the campus

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

MEASUREMENT-BASED VS. PERCEPTION-BASED CONCEPTS

measurement-based	***perception-based***
expected value	*usual value*
stationarity	*regularity*
continuous	*smooth*

Example of a regular process

$T = (t_0 , t_1 , t_2 \ldots)$

t_i ***= travel time from home to office on day i.***

PSEUDONUMBERS

- A pseudonumber is a symbol which has the appearance of a number and is used as a label of an interval, fuzzy set, or more generally, an arbitrary object

examples

- Room number 326
- Checkout time is 1 pm
- Speed limit is 100km/hour
- Intensity of earthquake was 6.3
- Probability is 0.7

CONTINUED

- *in many instances, a pseudonumber functions as a trigger of expectations*
  - *probability is 0.8: label of an interval or a distribution*
  - *check-out time is 1 pm: trigger of expectations*
  - *speed limit is 100km/hour: trigger of expectations*
  - *strength of earthquake was 6.5*

many of the numbers used in decision analysis and economics are pseudonumbers

EXAMPLE

speed limit is 100km/hour

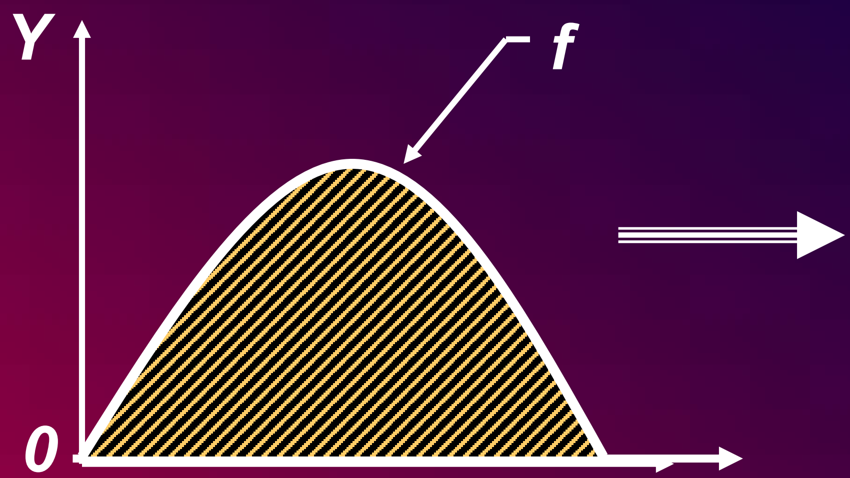
Expectation graph: S0 ×E0 +S1 ×E1 + ··· +Si ×Ei+ ···

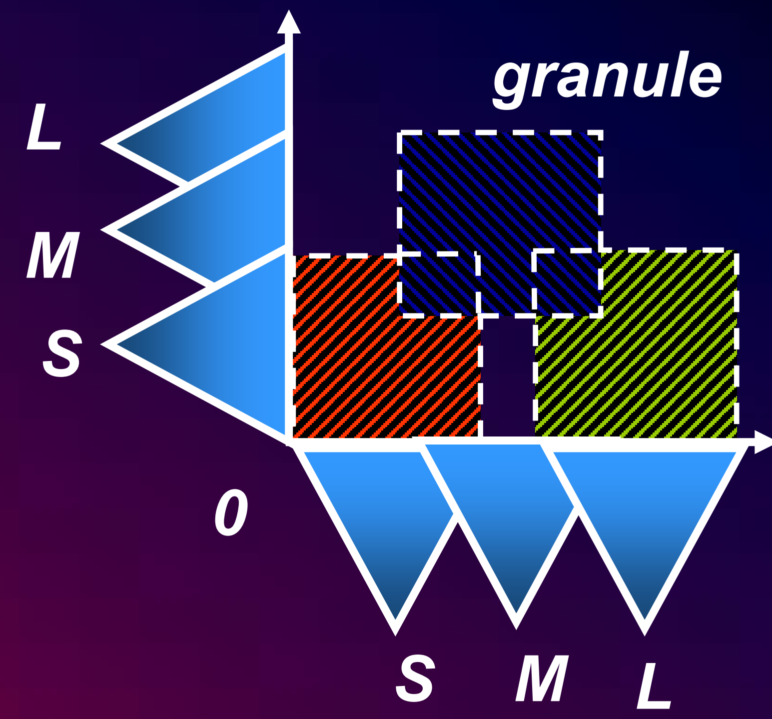
Ei: linguistic description of expectation

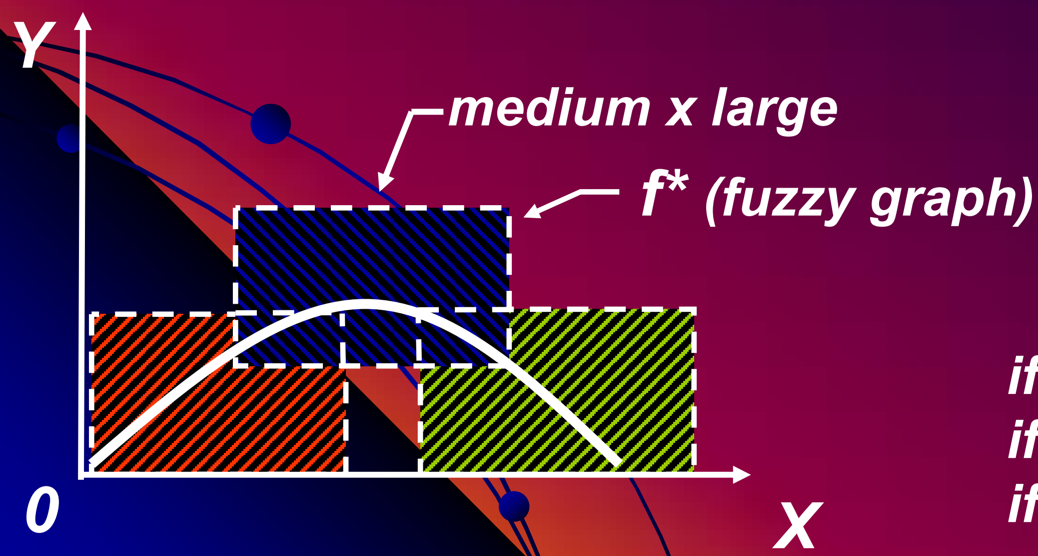
LI(Ei): loss index of Ei

LossIndex graph: S0 ×LI(E0) +S1 ×LI(E1) + ··· +Si ×LI(Ei) + ···

PERCEPTION OF MATHEMATICAL CONCEPTS: PERCEPTION OF FUNCTION



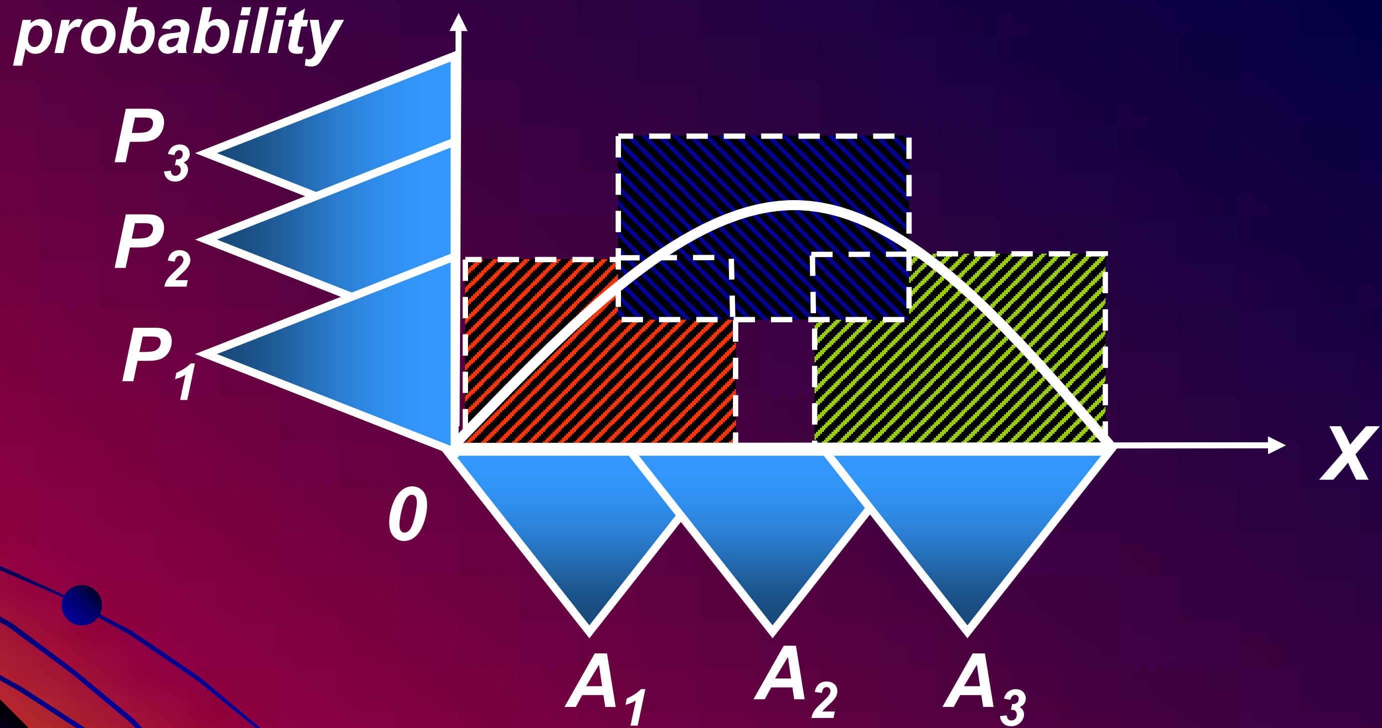




$f \xrightarrow{\text{perception}} f^*$:

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

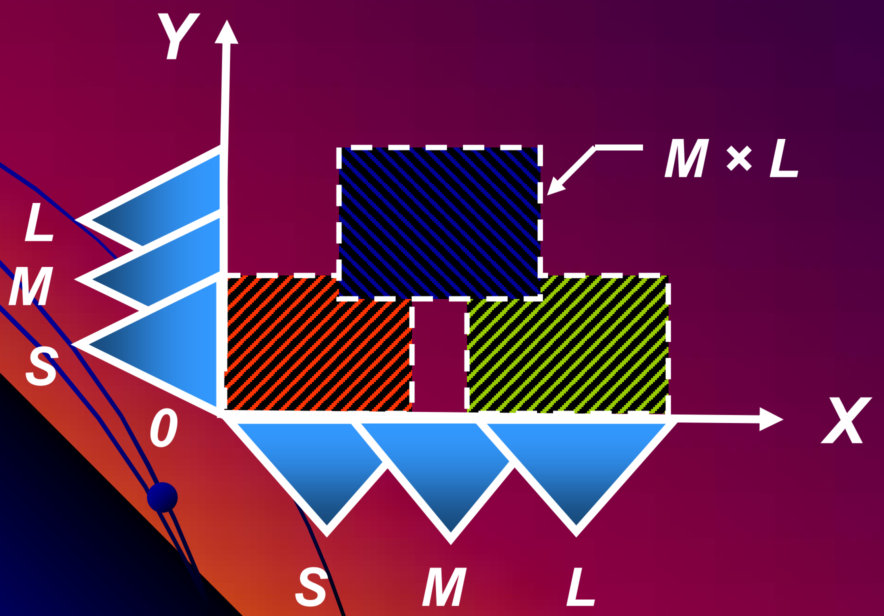
TEST PROBLEM

- *A function, Y=f(X), is defined by its fuzzy graph expressed as*

  f_1 if X is small then Y is small
  if X is medium then Y is large
  if X is large then Y is small

  (a) what is the value of Y if X is not large?

  (b) what is the maximum value of Y



BASIC POINTS

- *Computing with words and perceptions, or CWP for short, is a mode of computing in which the objects of computation are words, propositions and perceptions described in a natural language.*

CONTINUED

- *Perceptions play a key role in human cognition. Humans—but not machines—have a remarkable capability to perform a wide variety of physical and mental tasks without any measurements and any computations. Everyday examples of such tasks are driving a car in city traffic, playing tennis and summarizing a book.*

CONTINUED

- In computing with words and perceptions, the objects of computation are words, propositions, and perceptions described in a natural language
- A natural language is a system for describing perceptions
- In CWP, a perception is equated to its description in a natural language

CONTINUED

- *in science, it is a deep-seated tradition to strive for the ultimate in rigor and precision*
- *words are less precise than numbers*
- *why and where, then, would words be used in preference to numbers?*

CONTINUED

- *when the available information is not precise enough to justify the use of numbers*

- *when precision carries a cost and there is a tolerance for imprecision which can be exploited to achieve tractability, robustness and reduced cost*

- *when the expressive power of words is greater than the expressive power of numbers*

CONTINUED

- ***One of the major aims of CWP is to serve as a basis for equipping machines with a capability to operate on perception-based information. A key idea in CWP is that of dealing with perceptions through their descriptions in a natural language. In this way, computing and reasoning with perceptions is reduced to operating on propositions drawn from a natural language.***

CONTINUED

- ***In CWP, what is employed for this purpose is PNL (Precisiated Natural Language.) In PNL, a proposition, p, drawn from a natural language, NL, is represented as a generalized constraint, with the language of generalized constraints, GCL, serving as a precisiation language for computation and reasoning, PNL is equipped with two dictionaries and a modular multiagent deduction database. The rules of deduction are expressed in what is referred to as the Protoform Language (PFL).***

KEY POINTS

- *decisions are based on information*
- *in most realistic settings, decision-relevant information is a mixture of measurements and perceptions*
- *examples: buying a house; buying a stock*
- *existing methods of decision analysis are measurement-based and do not provide effective tools for dealing with perception-based information*

- *a decision is strongly influenced by the perception of likelihoods of outcomes of a choice of action*

KEY POINTS

- *in most realistic settings:*

  a) *the outcomes of a decision cannot be predicted with certainty*
  b) *decision-relevant probability distributions are f-granular*
  c) *decision-relevant events, functions and relations are f-granular*

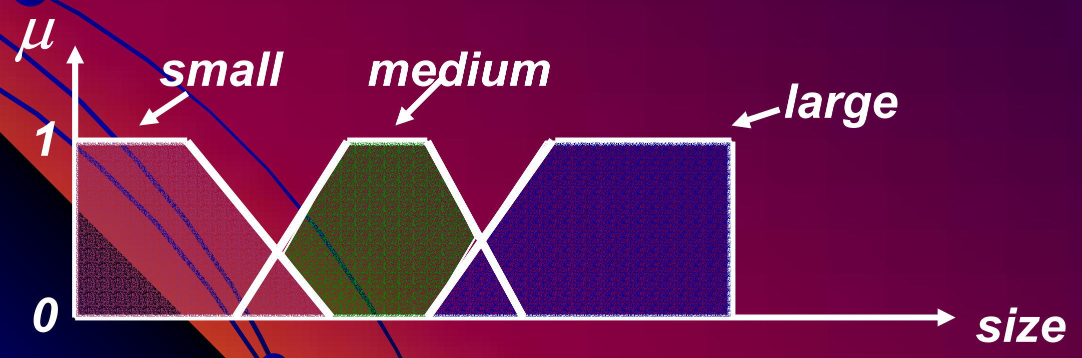
- *perception-based probability theory, PTp, is basically a calculus of f-granular probability distributions, f-granular events, f-granular functions, f-granular relations and f-granular counts*

OBSERVATION

- *machines are driven by measurements*
- *humans are driven by perceptions*
- *to enable a machine to mimic the remarkable human capability to perform a wide variety of physical and mental tasks using perception-based information, it is necessary to have a means of converting measurements into perceptions*

BASIC PERCEPTIONS / F-GRANULARITY

- *temperature: warm+cold+very warm+much warmer+…*
- *time: soon + about one hour + not much later +…*
- *distance: near + far + much farther +…*
- *speed: fast + slow +much faster +…*
- *length: long + short + very long +…*



CONTINUED

- *similarity: low + medium + high +…*
- *possibility: low + medium + high + almost impossible +…*
- *likelihood: likely + unlikely + very likely +…*
- *truth (compatibility): true + quite true + very untrue +…*
- *count: many + few + most + about 5 (5*) +…*

subjective probability = perception of likelihood

CONTINUED

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

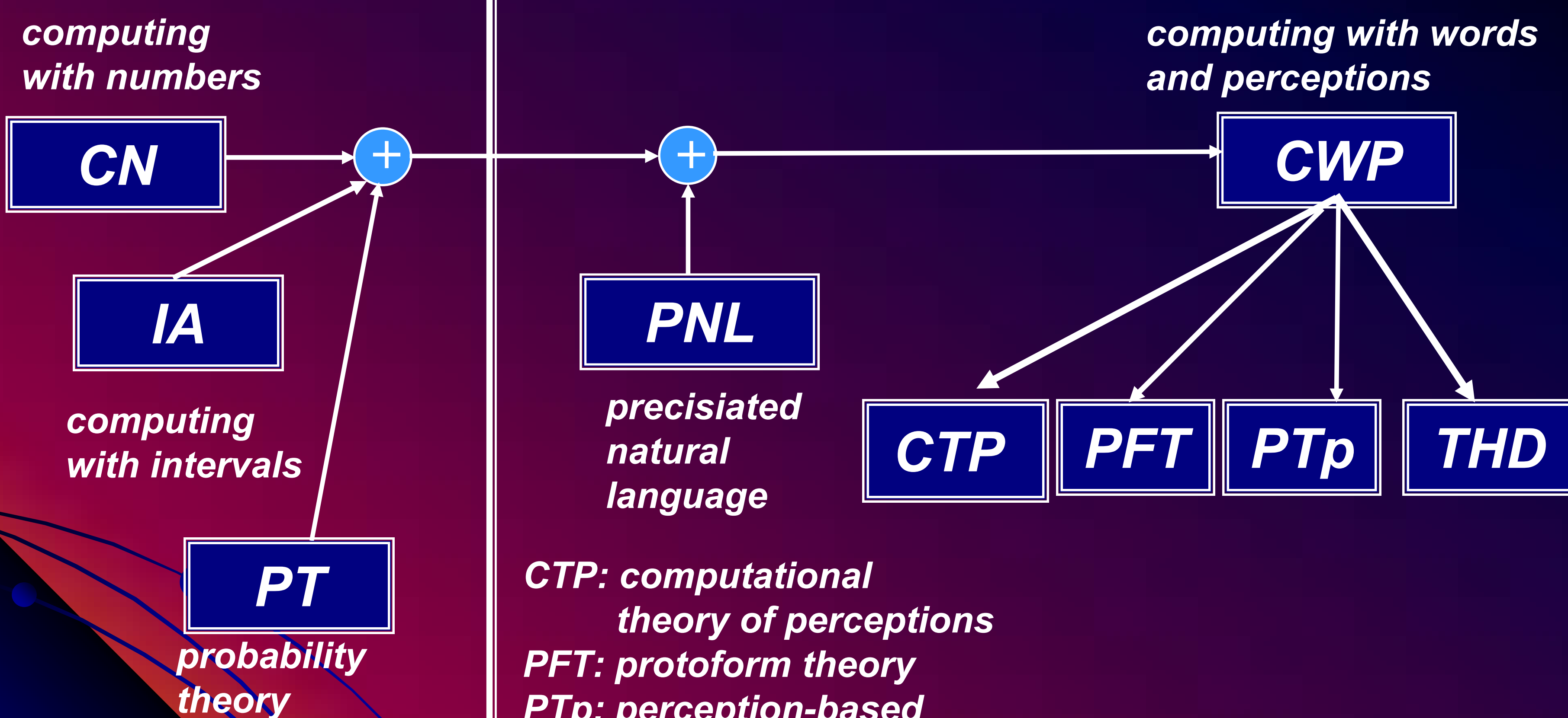
  (X is small, Y is large)
- *probability distribution: low \ small + low \ medium + high \ large +…*
- *Count \ attribute value distribution: 5* \ small + 8* \ large +…*

PRINCIPAL RATIONALES FOR F-GRANULATION

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

New Tools

NEW TOOLS



CTP: computational theory of perceptions
PFT: protoform theory
PTp: perception-based probability theory
THD: theory of hierarchical definability

 LAZ 2/15/2004

GRANULAR COMPUTING GENERALIZED VALUATION

valuation = assignment of a value to a variable

X = 5	0 ≤ X ≤ 5	X is small	X isr R
point	interval	fuzzy interval	generalized

singular value

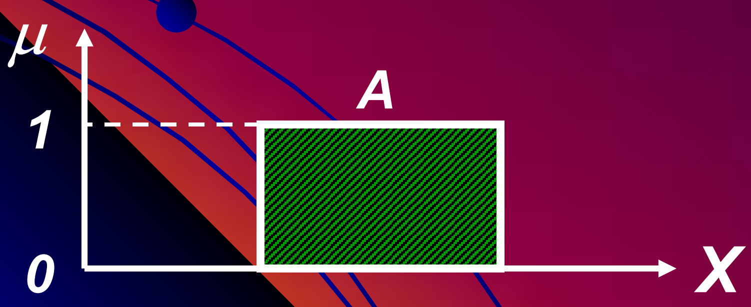
measurement-based

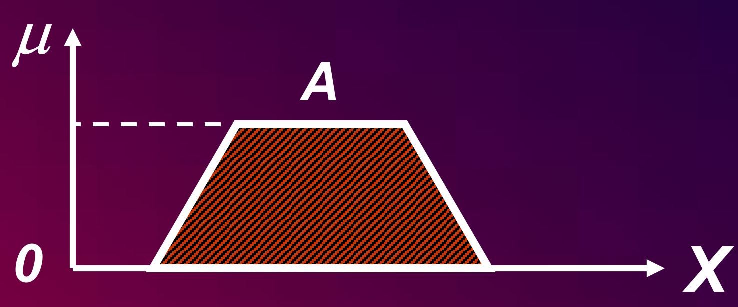
granular values

perception-based

F-GENERALIZATION

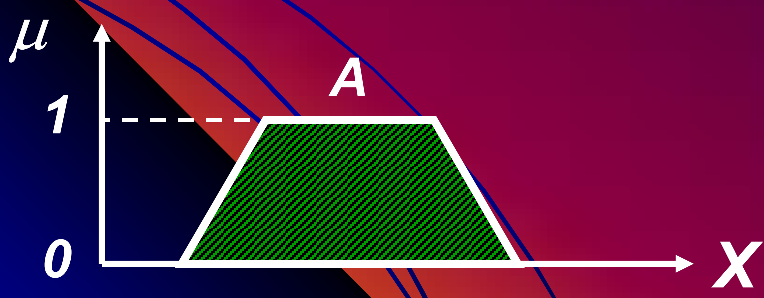
- *f-generalization of a theory, T, involves an introduction into T of the concept of a fuzzy set*
- *f-generalization of PT, PT^+, adds to PT the capability to deal with fuzzy probabilities, fuzzy probability distributions, fuzzy events, fuzzy functions and fuzzy relations*

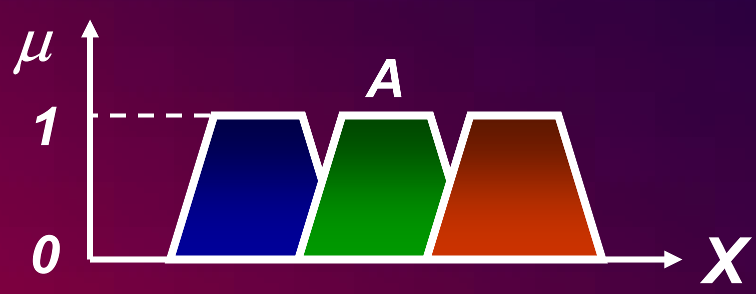




F.G-GENERALIZATION

- *f.g-generalization of T, T^{++}, involves an introduction into T of the concept of a granulated fuzzy set*
- *f.g-generalization of PT, PT^{++} , adds to PT^{+} the capability to deal with f-granular probabilities, f-granular probability distributions, f-granular events, f-granular functions and f-granular relations*





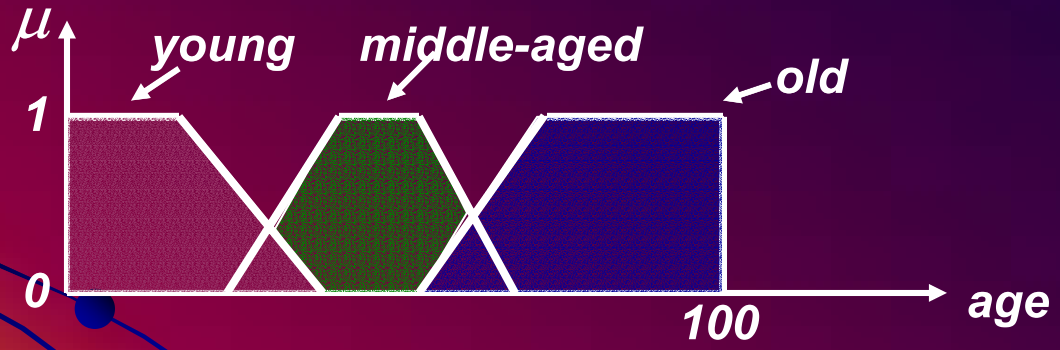
EXAMPLES OF F-GRANULATION (LINGUISTIC VARIABLES)

color: red, blue, green, yellow, …

age: young, middle-aged, old, very old

size: small, big, very big, …

distance: near, far, very, not very far, …



- *humans have a remarkable capability to perform a wide variety of physical and mental tasks, e.g., driving a car in city traffic, without any measurements and any computations*
- *one of the principal aims of CTP is to develop a better understanding of how this capability can be added to machines*

NL-GENERALIZATION

- *nl-generalization of T. T_{nl}, involves an addition to T^{++} of a capability to operate on propositions expressed in a natural language*
- *nl-generalization of T adds to T^{++} a capability to operate on perceptions described in a natural language*
- *nl-generalization of PT, PT_{nl}, adds to PT^{++} a capability to operate on perceptions described in a natural language*
- *nl-generalization of PT is perception-based probability theory, PTp*
- *a key concept in PTp is PNL (Precisiated Natural Language)*

PRECISIATED NATURAL LANGUAGE

PNL

WHAT IS PRECISIATED NATURAL LANGUAGE (PNL)? PRELIMINARIES

- ***a proposition, p, in a natural language, NL, is precisiable if it translatable into a precisiation language***
- ***in the case of PNL, the precisiation language is the Generalized Constraint Language, GCL***
- ***precisiation of p, p*, is an element of GCL (GC-form)***

WHAT IS PNL?

- ***PNL is a sublanguage of precisiable propositions in NL which is equipped with two dictionaries: (1) NL to GCL; (2) GCL to PFL (Protoform Language); and (3) a modular multiagent database of rules of deduction (rules of generalized constrained propagation) expressed in PFL.***

GENERALIZED CONSTRAINT

- *standard constraint:* $X \in C$
- *generalized constraint: X isr R*

X isr R

copula

GC-form (generalized constraint form of type r)

type identifier

constraining relation

constrained variable

- $X = (X_1, \ldots, X_n)$
- *X may have a structure: X=Location (Residence(Carol))*
- *X may be a function of another variable: X=f(Y)*
- *X may be conditioned: (X/Y)*
- r := / ¡Ü ... / □¼□½blank / v / p / u / rs / fg / ps / ...

GC-FORM (GENERALIZED CONSTRAINT FORM OF TYPE 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

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)*

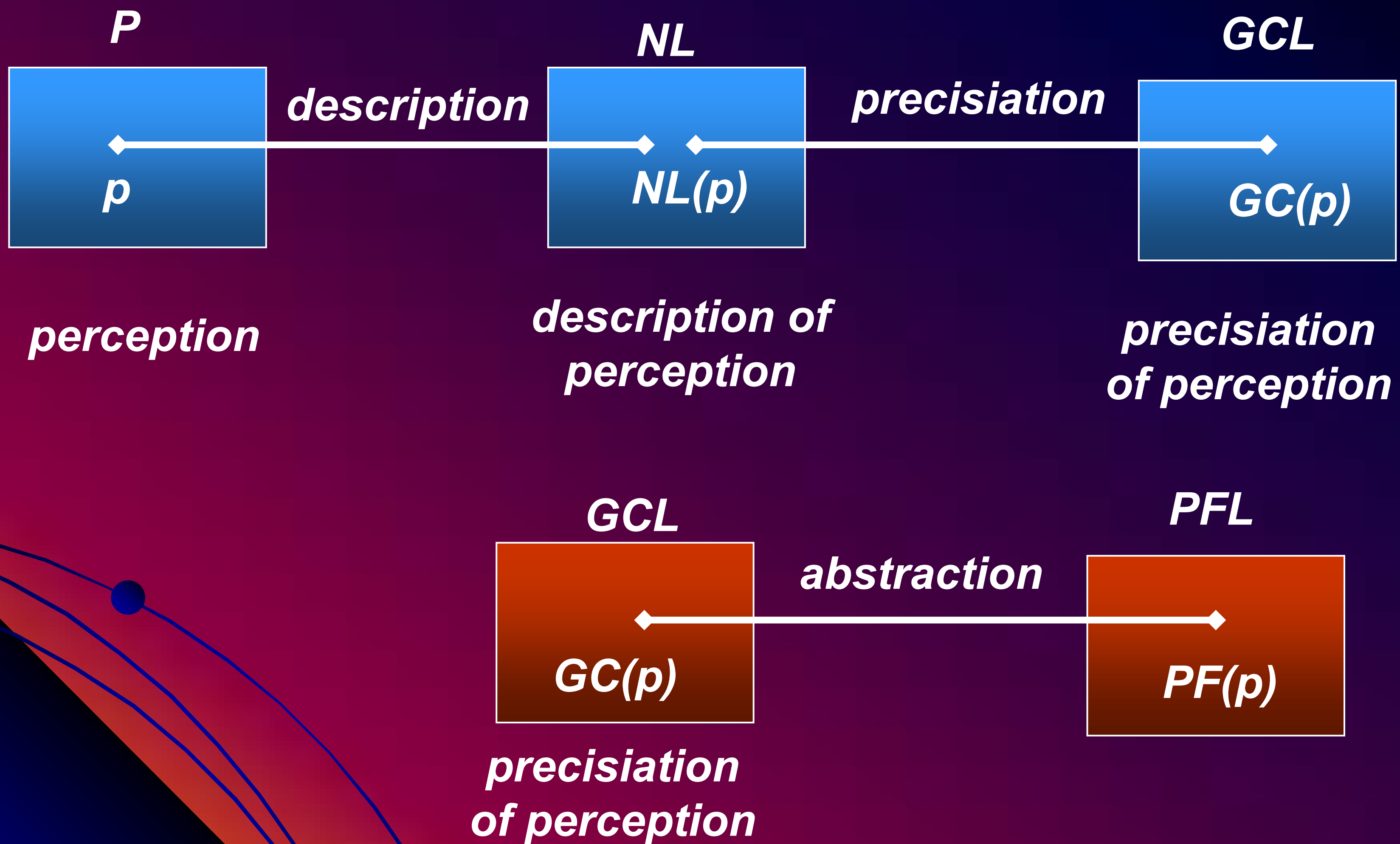
r: ps *Pawlak set constraint: X isps ($\underline{X}$, $\overline{X}$) means that X is a set and $\underline{X}$ and $\overline{X}$ are the lower and upper approximations to X*

GENERALIZED CONSTRAINT LANGUAGE (GCL)

- *GCL is generated by combination, qualification and propagation of generalized constraints*
- *in GCL, rules of deduction are the rules governing generalized constraint propagation*
- *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 small then Y is large*

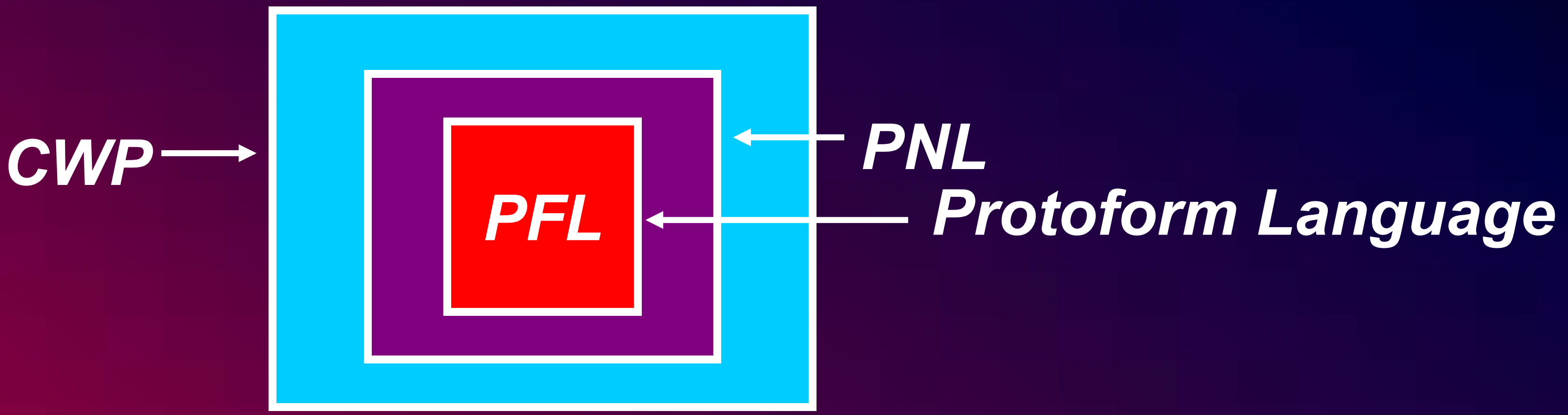
- *the language of fuzzy if-then rules is a sublanguage of PNL*

THE BASIC IDEA



GCL (Generalized Constrain Language) is maximally expressive

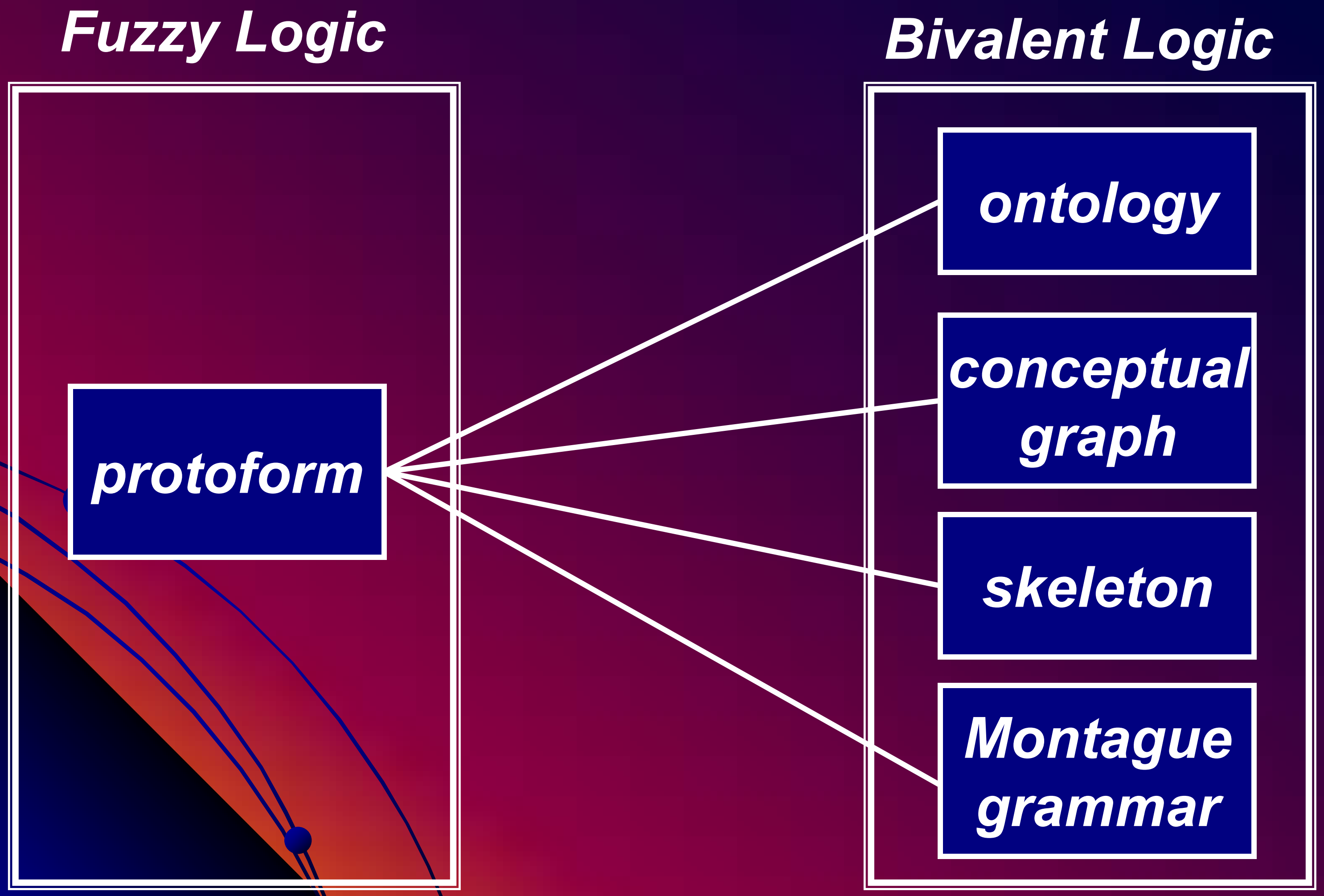
THE CONCEPT OF A PROTOFORM



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*

THE CONCEPT OF PROTOFORM AND RELATED CONCEPTS



TRANSLATION FROM NL TO PFL

examples

Most Swedes are tall ⟶ *Count (A/B) is Q*

Eva is much younger than Pat ⟶ *(A (B), A (C)) is R*

Age Eva Age Pat much younger

usually Robert returns from work at about 6pm ⟶

Prob {A is B} is C

usually

about 6 pm

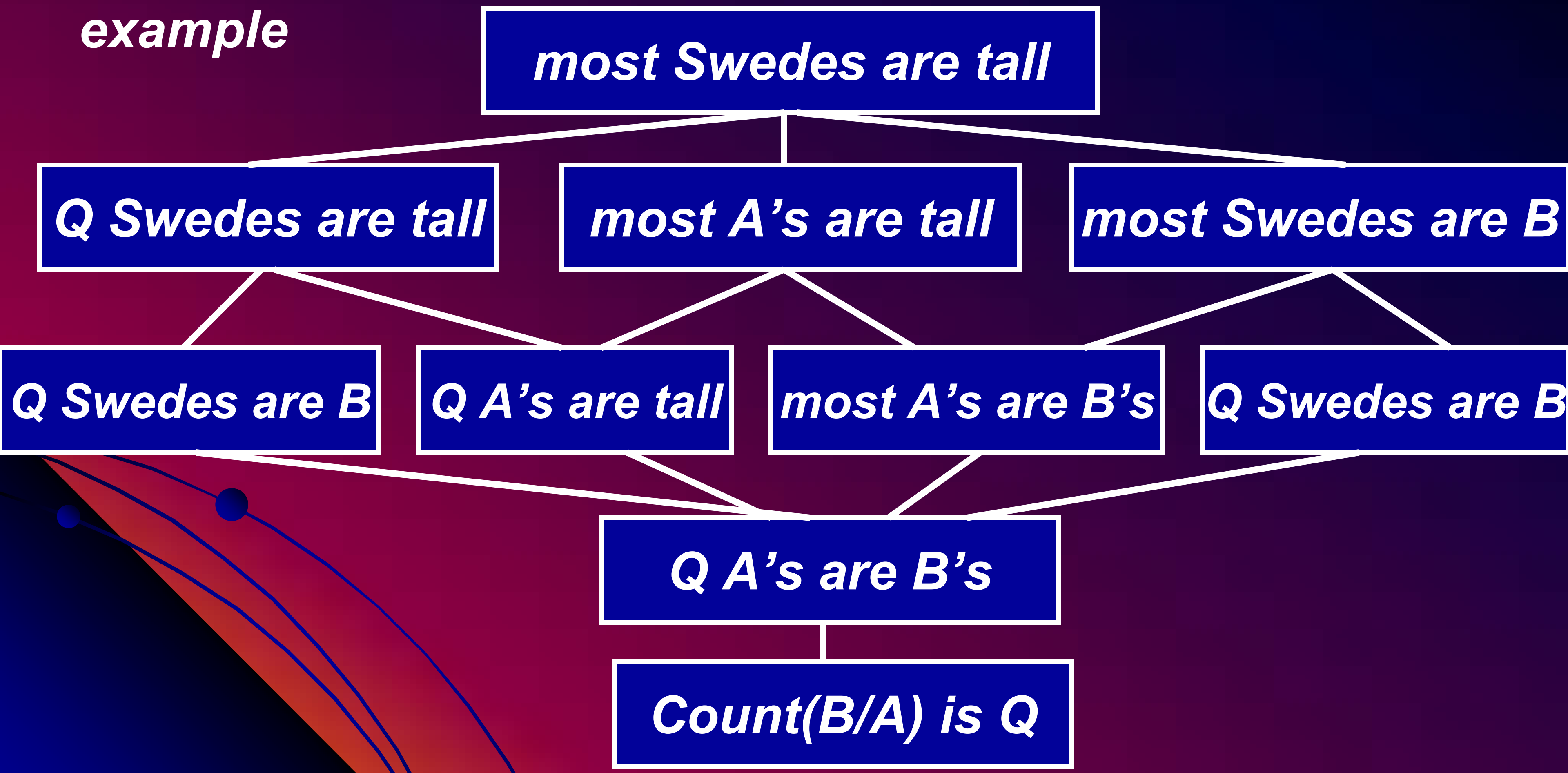
Time (Robert returns from work)

MULTILEVEL STRUCTURES

- An object has a multiplicity of protoforms
- Protoforms have a multilevel structure
- There are three principal multilevel structures
- Level of abstraction (α)
- Level of summarization (σ)
- Level of detail (δ)
- For simplicity, levels are implicit
- A terminal protoform has maximum level of abstraction
- A multilevel structure may be represented as a lattice

ABSTRACTION LATTICE

example



LEVELS OF SUMMARIZATION

example

p: it is very unlikely that there will be a significant increase in the price of oil in the near future

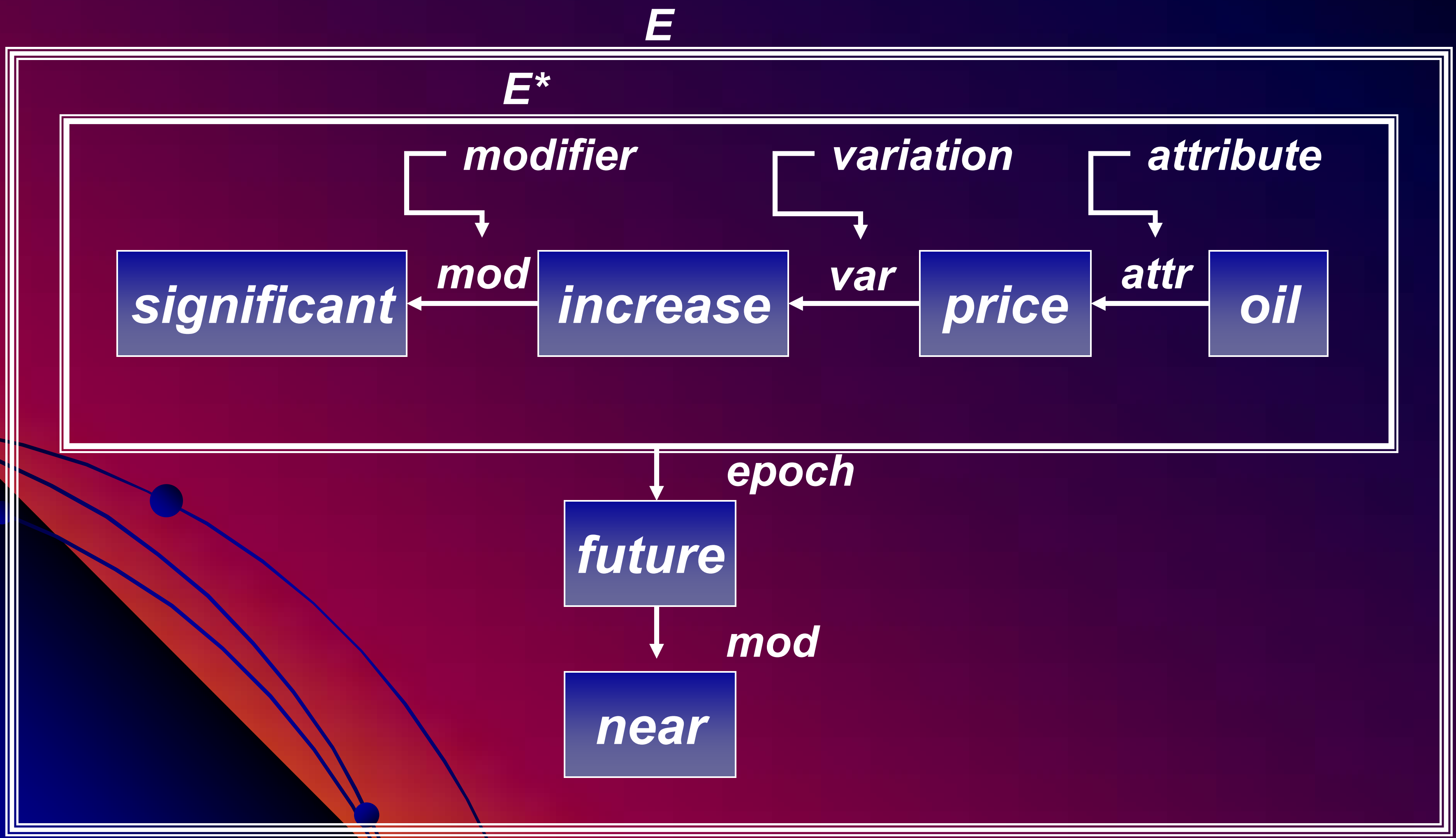
PF(p): Prob(E) is A

very.unlikely

significant increase in the price of oil in the near future

CONTINUED

semantic network representation of E



CONTINUED

PF(E): B(C) is D

near.future

significant.increase.price.oil

epoch

PF(C): H(G(D))

oil

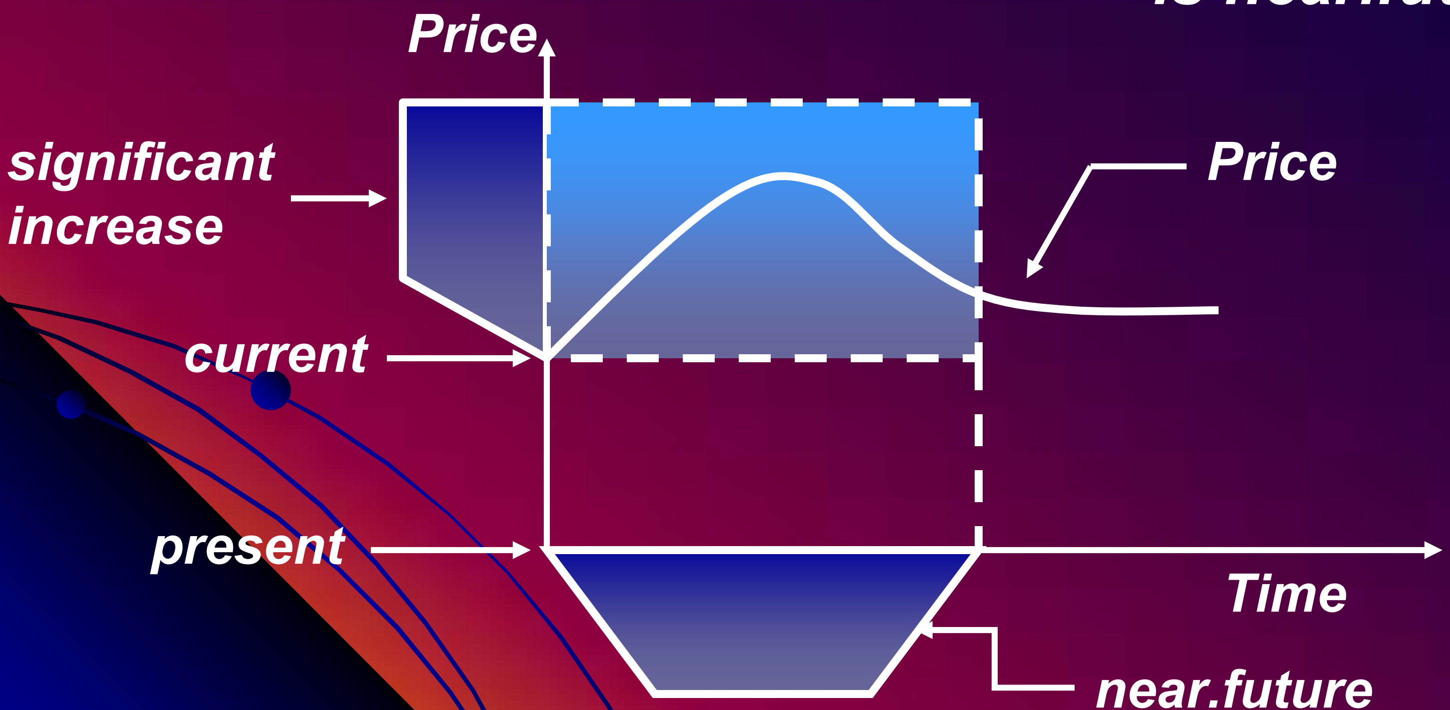
price

significant.increase

CONTINUED

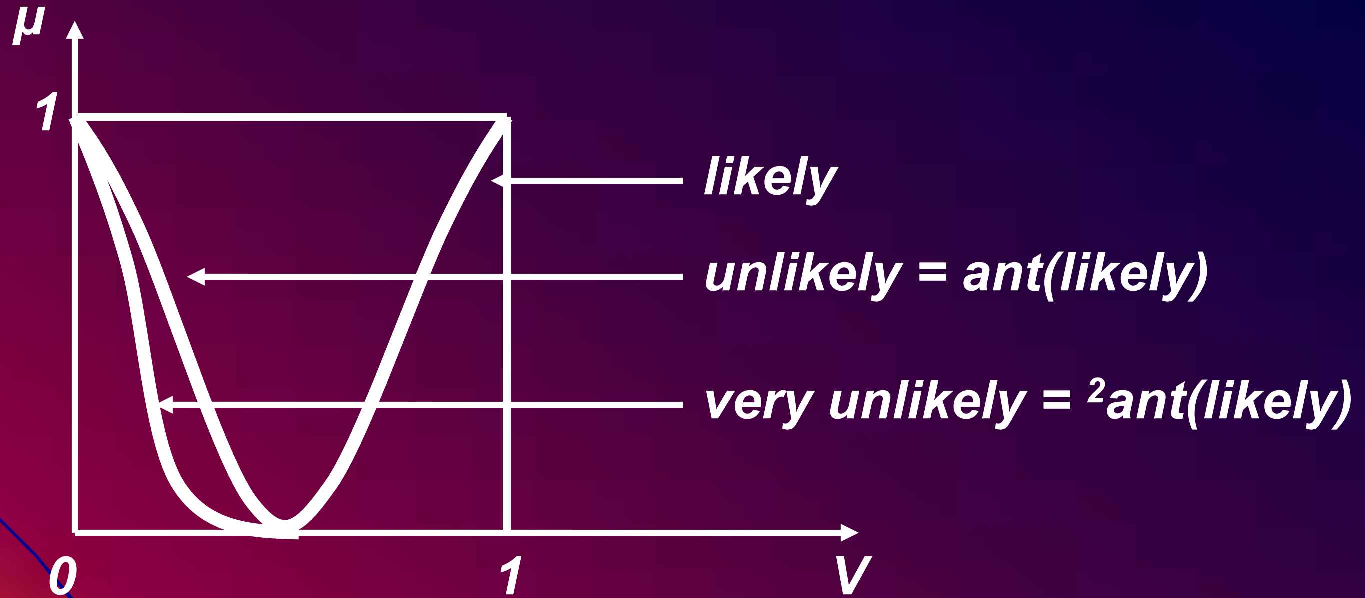
Precisiation (f.b.-concept)

E: Epoch (Variation (Price (oil)) is significant.increase) is near.future*



CONTINUED

precisiation of very unlikely



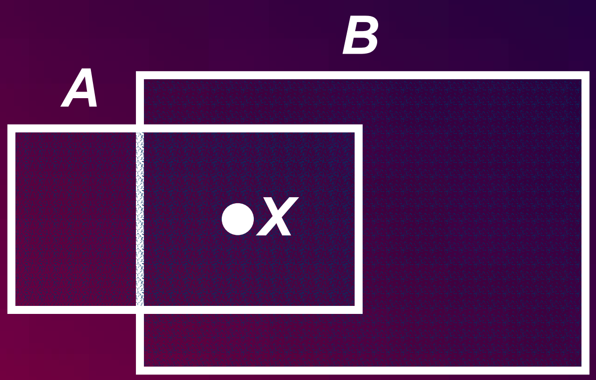
$$\mu_{very.unlikely}(v) = (\mu_{likely}(1-v))^2$$

PROTOFORM OF A DECISION PROBLEM

- *buying a home*
- *decision attributes*
  - *measurement-based: price, taxes, area, no. of rooms, …*
  - *perception-based: appearance, quality of construction, security*
- *normalization of attributes*
- *ranking of importance of attributes*
- *importance function: w(attribute)*
- *importance function is granulated: L(low), M(medium), H(high)*

PROTOFORM OF A QUERY

- *largest port in Canada?*
- *second tallest building in San Francisco*



?X is selector (attribute (A/B))

2^{nd} tallest — *height* — *buildings* — *San Francisco*

TEST QUERY (GOOGLE)

- *population of largest city in Spain: failure*
- *largest city in Spain: Madrid, success*
- *population of Madrid: success*

PROTOFORM OF A DECISION PROBLEM

- buying a house
- decision attributes

  measurement-based: price, taxes, area, no. of rooms, …

  perception-based: appearance, quality of construction, security
- normalization of attributes
- ranking of importance of attributes
- importance function: w(attribute)

importance function is granulated: L(low), M (medium), H (high)

DICTIONARIES

1:

proposition in NL	precisiation
p	p^* (GC-form)
most Swedes are tall	Σ Count (tall.Swedes/Swedes) is most

2:

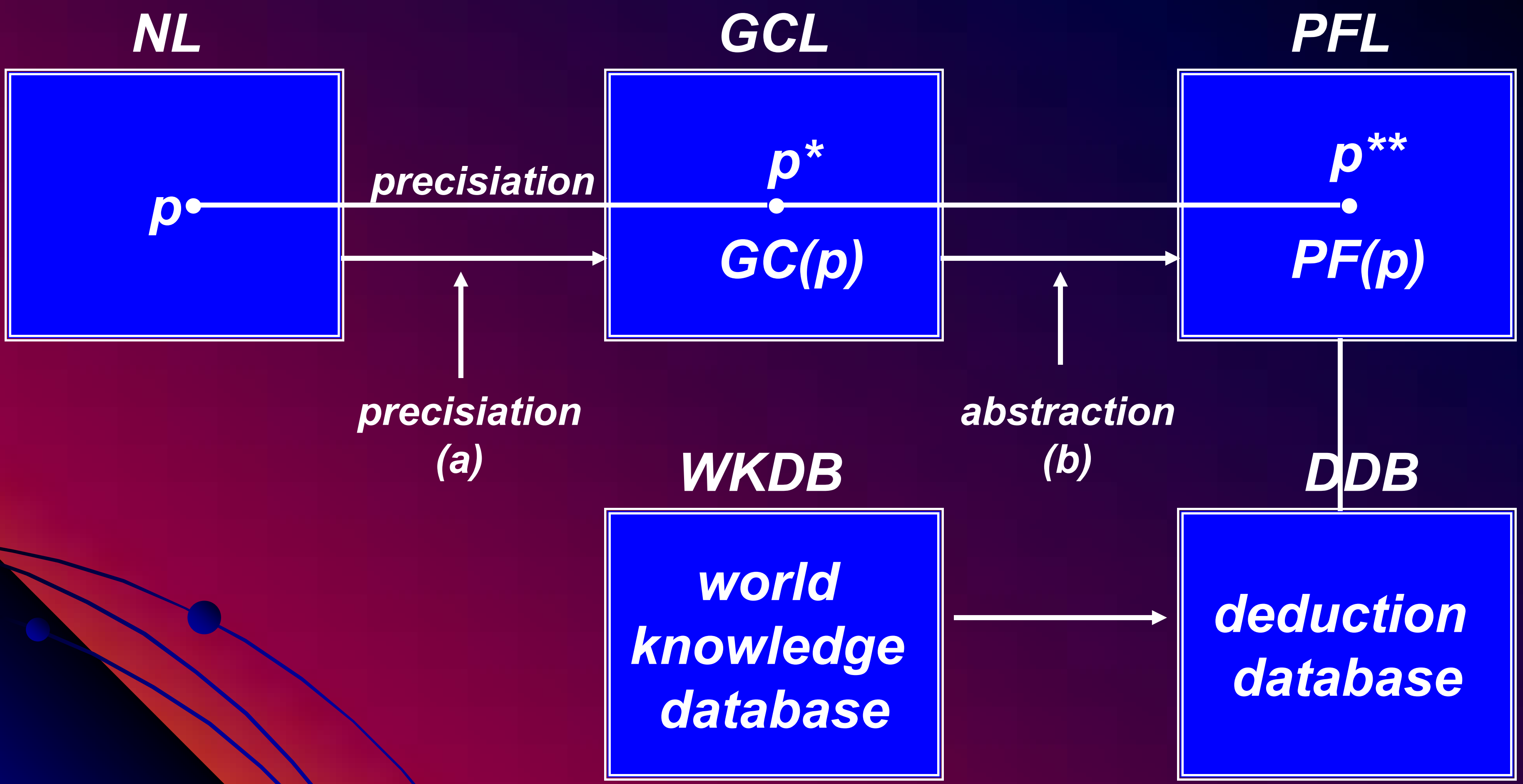
precisiation	protoform
p^* (GC-form)	PF(p^*)
Σ Count (tall.Swedes/Swedes) is most	Count(A/B) is Q

EXAMPLE OF TRANSLATION

- P: usually Robert returns from work at about 6 pm
- P*: Prob {(Time(Return(Robert)) is 6 pm} is usually
- PF(p): Prob {X is A} is B
- X: Time (Return (Robert))
- A: 6 pm
- B: usually

$p \in NL$
$p^* \in GCL$
$PF(p) \in PFL$

BASIC STRUCTURE OF PNL



- *In PNL, deduction=generalized constraint propagation*

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

WKDB: PNL-based

WORLD KNOWLEDGE

examples

- *icy roads are slippery*
- *big cars are safer than small cars*
- *usually it is hard to find parking near the campus on weekdays between 9 and 5*
- *most Swedes are tall*
- *overeating causes obesity*
- *Ph.D. is the highest academic degree*
- *an academic degree is associated with a field of study*
- *Princeton employees are well paid*

WORLD KNOWLEDGE

KEY POINTS

- *world knowledge—and especially knowledge about the underlying probabilities—plays an essential role in disambiguation, planning, search and decision processes*

- *what is not recognized to the extent that it should, is that world knowledge is for the most part perception-based*

WORLD KNOWLEDGE: EXAMPLES

specific:

- *if Robert works in Berkeley then it is likely that Robert lives in or near Berkeley*
- *if Robert lives in Berkeley then it is likely that Robert works in or near Berkeley*

generalized:

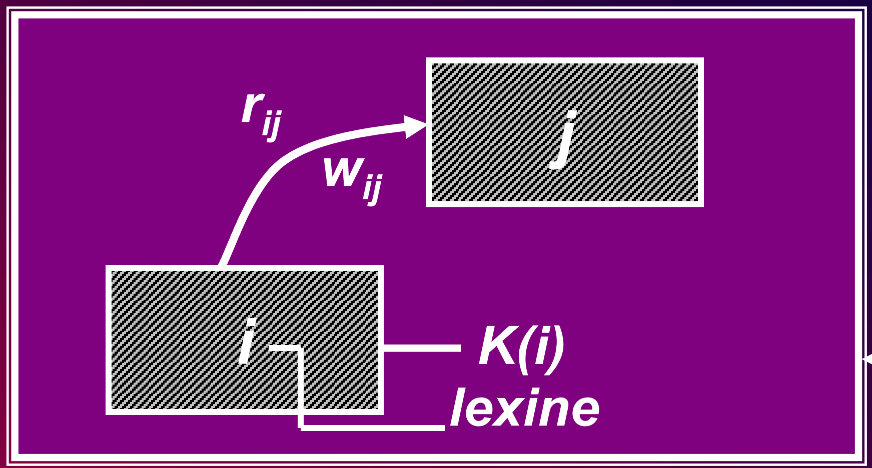
if A/Person works in B/City then it is likely that A lives in or near B

precisiated:

Distance (Location (Residence (A/Person), Location (Work (A/Person) isu near

protoform: F (A (B (C)), A (D (C))) isu R

ORGANIZATION OF WORLD KNOWLEDGE EPISTEMIC (KNOWLEDGE-DIRECTED) LEXICON (EL) (ONTOLOGY-RELATED)

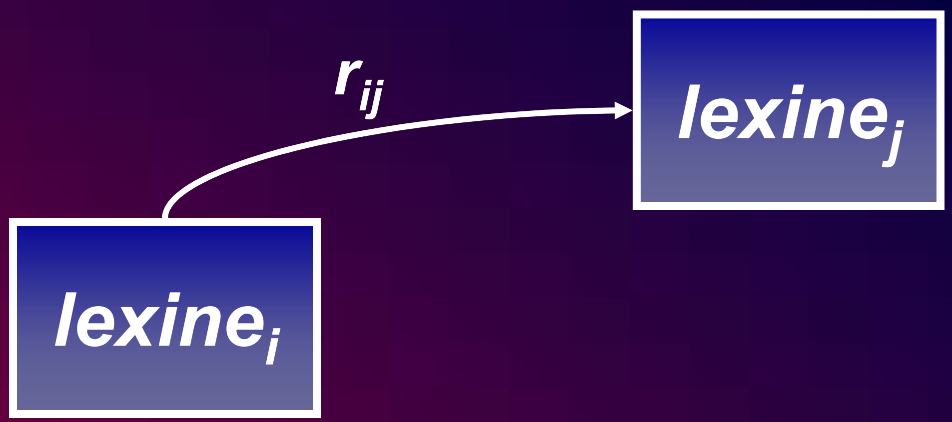


w_{ij}= granular strength of association between i and j

network of nodes and links

- *i (lexine): object, construct, concept (e.g., car, Ph.D. degree)*
- *K(i): world knowledge about i (mostly perception-based)*
- *K(i) is organized into n(i) relations $R_{ii}, \ldots, R_{in}$*
- *entries in R_{ij} are bimodal-distribution-valued attributes of i*
- *values of attributes are, in general, granular and context-dependent*

EPISTEMIC LEXICON



r_{ij}: *i is an instance of j* *(is or isu)*
i is a subset of j *(is or isu)*
i is a superset of j *(is or isu)*
j is an attribute of i
i causes j *(or usually)*
i and j are related

EPISTEMIC LEXICON

FORMAT OF RELATIONS

perception-based relation

lexine	A_1	…	A_m
	G_1		G_m

attributes

granular values

example

car	*Make*	*Price*	
	ford	*G*	
	chevy		

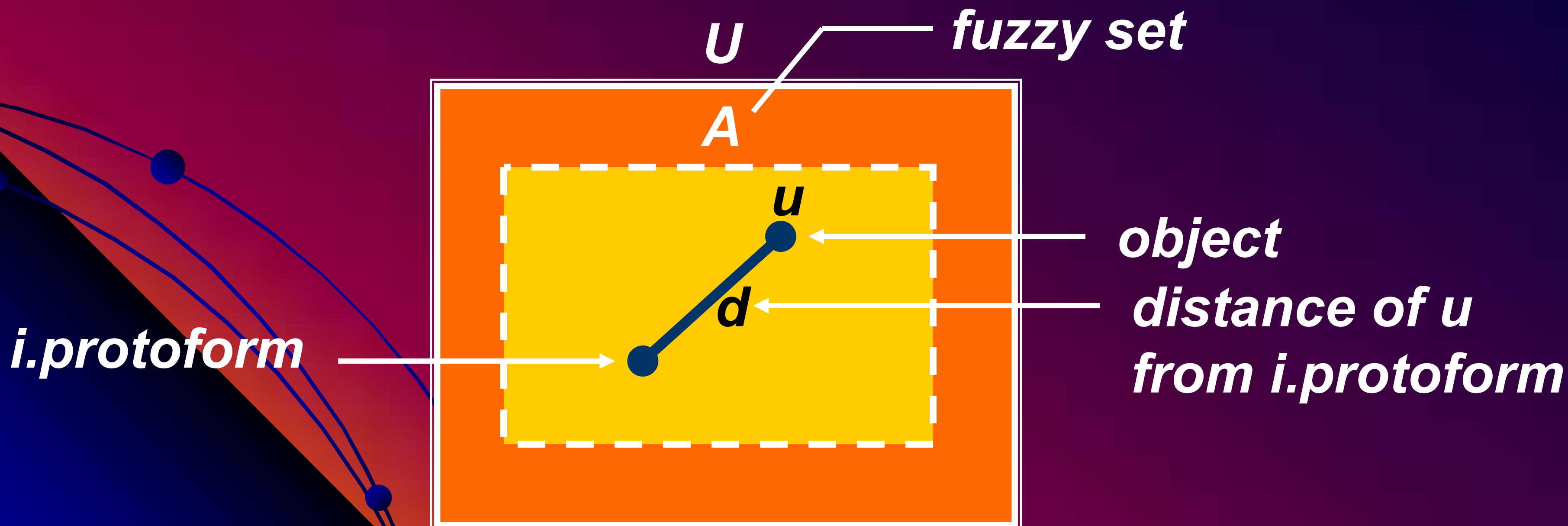
$$G: 20^*\% \setminus \angle 15k^* + 40^*\% \setminus [15k^*, 25k^*] + \bullet\bullet\bullet$$

granular count

PROTOFORM-BASED DEDUCTION

THE CONCEPT OF i.PROTOFORM

- *i.protoform: idealized protoform*
- *the key idea is to equate the grade of membership, $\mu_A(u)$, of an object, u, in a fuzzy set, A, to the distance of u from an i.protoform*
- *this idea is inspired by E. Rosch's work (ca 1972) on the theory of prototypes*



- *d is defined via PNL*

PROTOFORM-CENTERED CONCEPTS EXAMPLE: EXPECTED VALUE (f.f-concept)

- *X: real-valued random variable with probability density g*

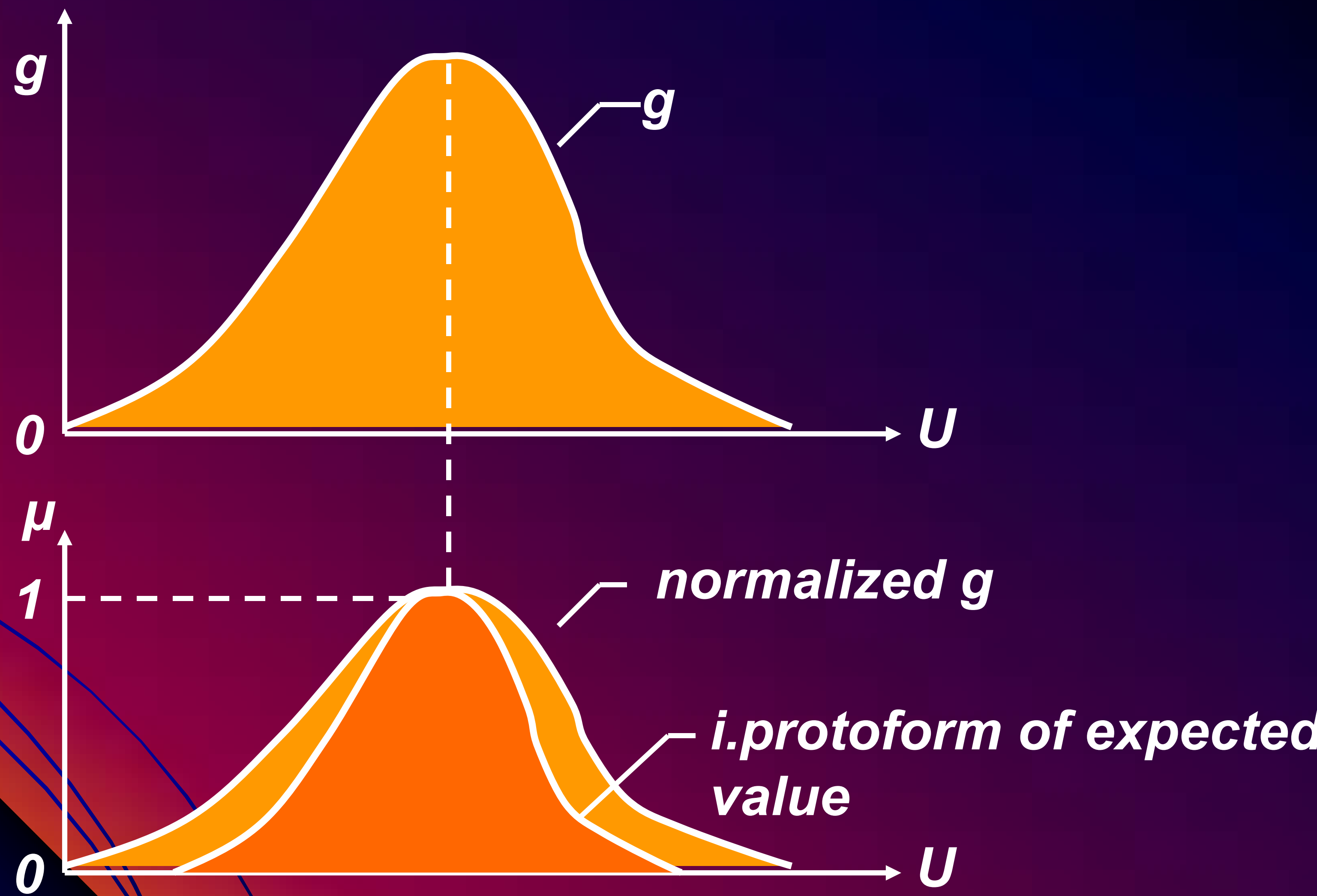
  standard definition of expected value of X

$$E(X) = \int_U ug(u)du$$

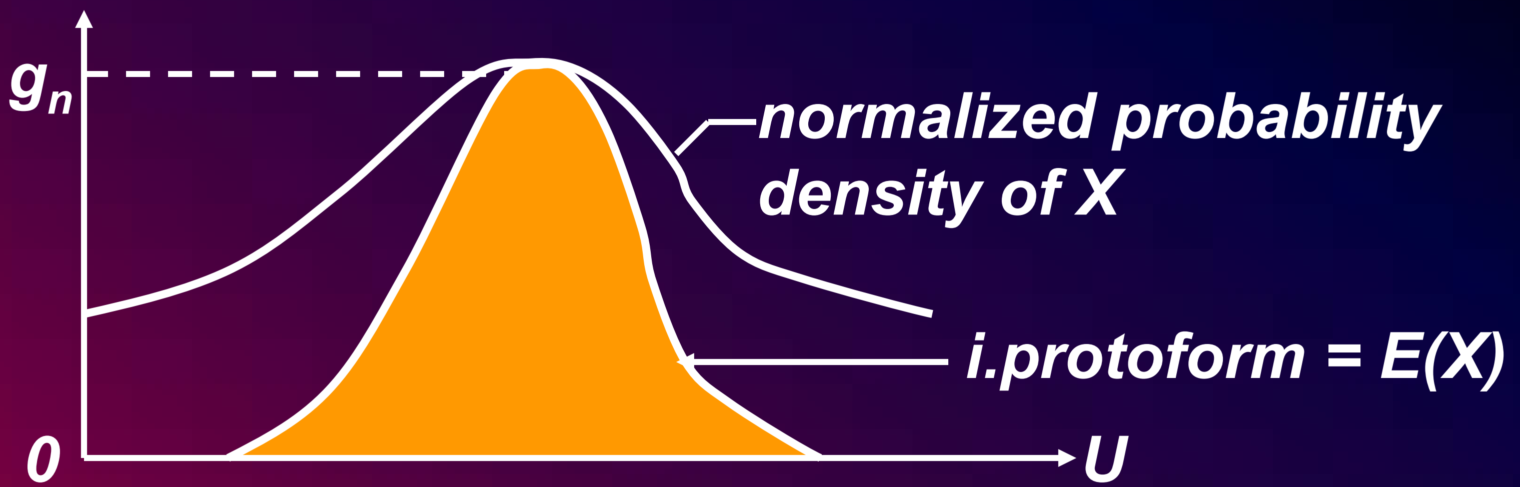
$E(X)$ = *average value of X*

- *the label "expected value" is misleading*

i.PROTOFORM-BASED DEFINITION OF EXPECTED VALUE



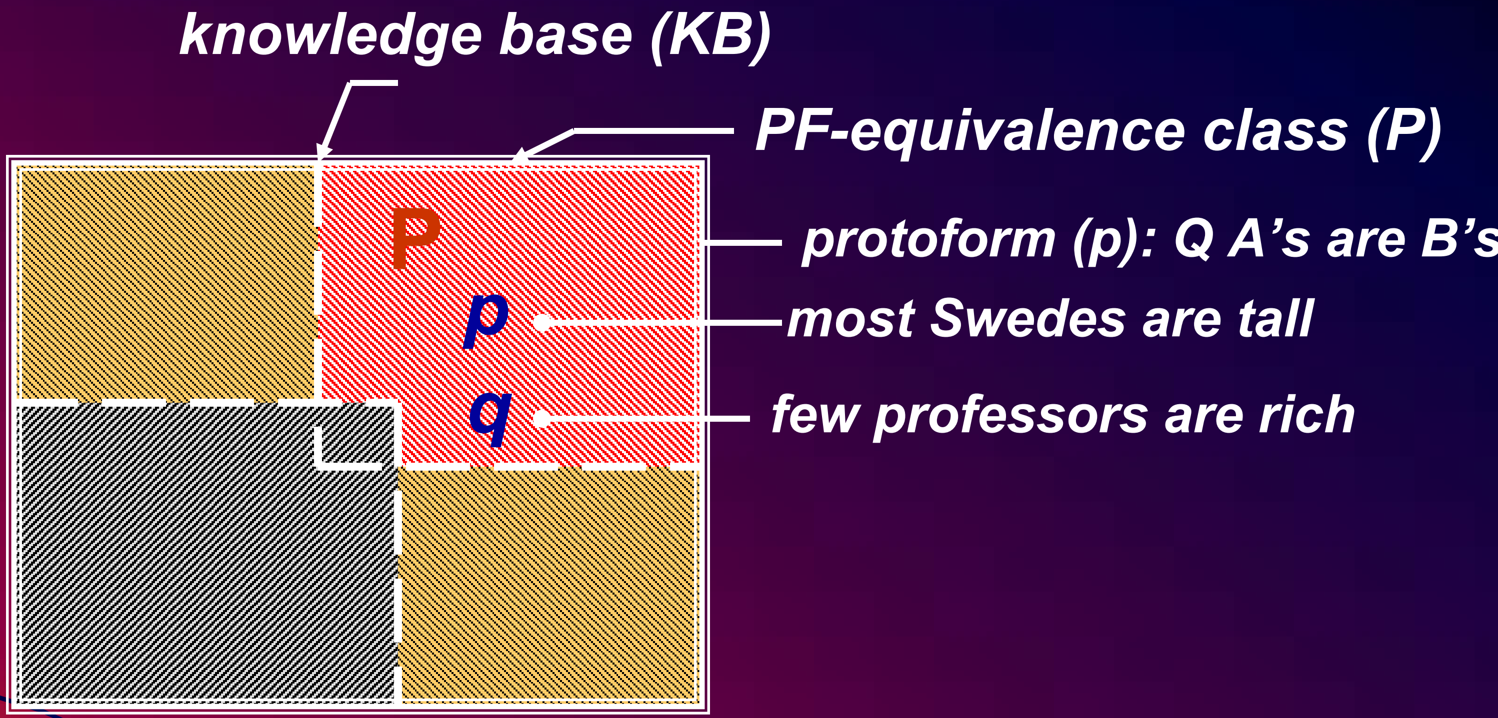
CONTINUED



- E(X) is a fuzzy set

- grade of membership of a particular function, E*(X), in the fuzzy set of expected value of X is the distance of E*(X) form best-fitting i.protoform

PROTOFORM AND PF-EQUIVALENCE



- P is the class of PF-equivalent propositions
- P does not have a prototype
- P has an abstracted prototype: Q A's are B's
- P is the set of all propositions whose protoform is: Q A's are B's

PF-EQUIVALENCE

Scenario A:

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?

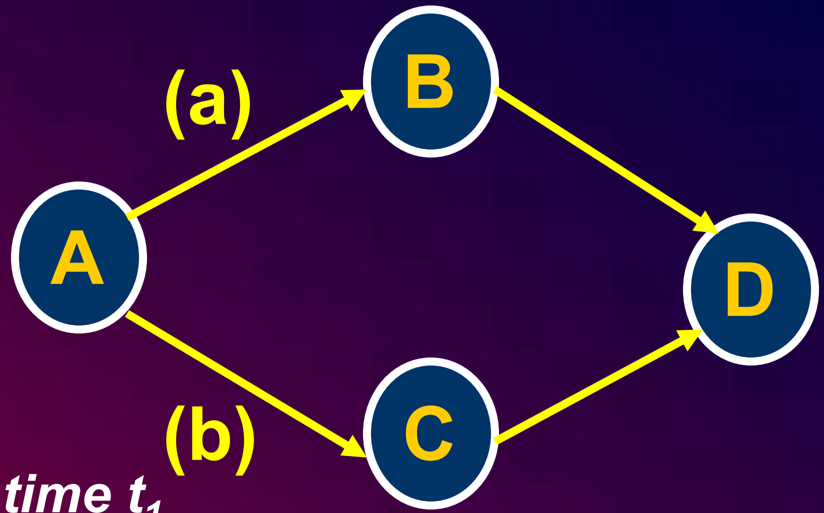
PF-EQUIVALENCE

Scenario B:

Alan needs to fly from San Francisco to St. Louis and has to get there as soon as possible. One option is fly to St. Louis via Chicago and the other through Denver. The flight via Denver is scheduled to arrive in St. Louis at time a. The flight via Chicago is scheduled to arrive in St. Louis at time b, with a<b. However, the connection time in Denver is short. If the flight is missed, then the time of arrival in St. Louis will be c, with c>b. Question: Which option is best?

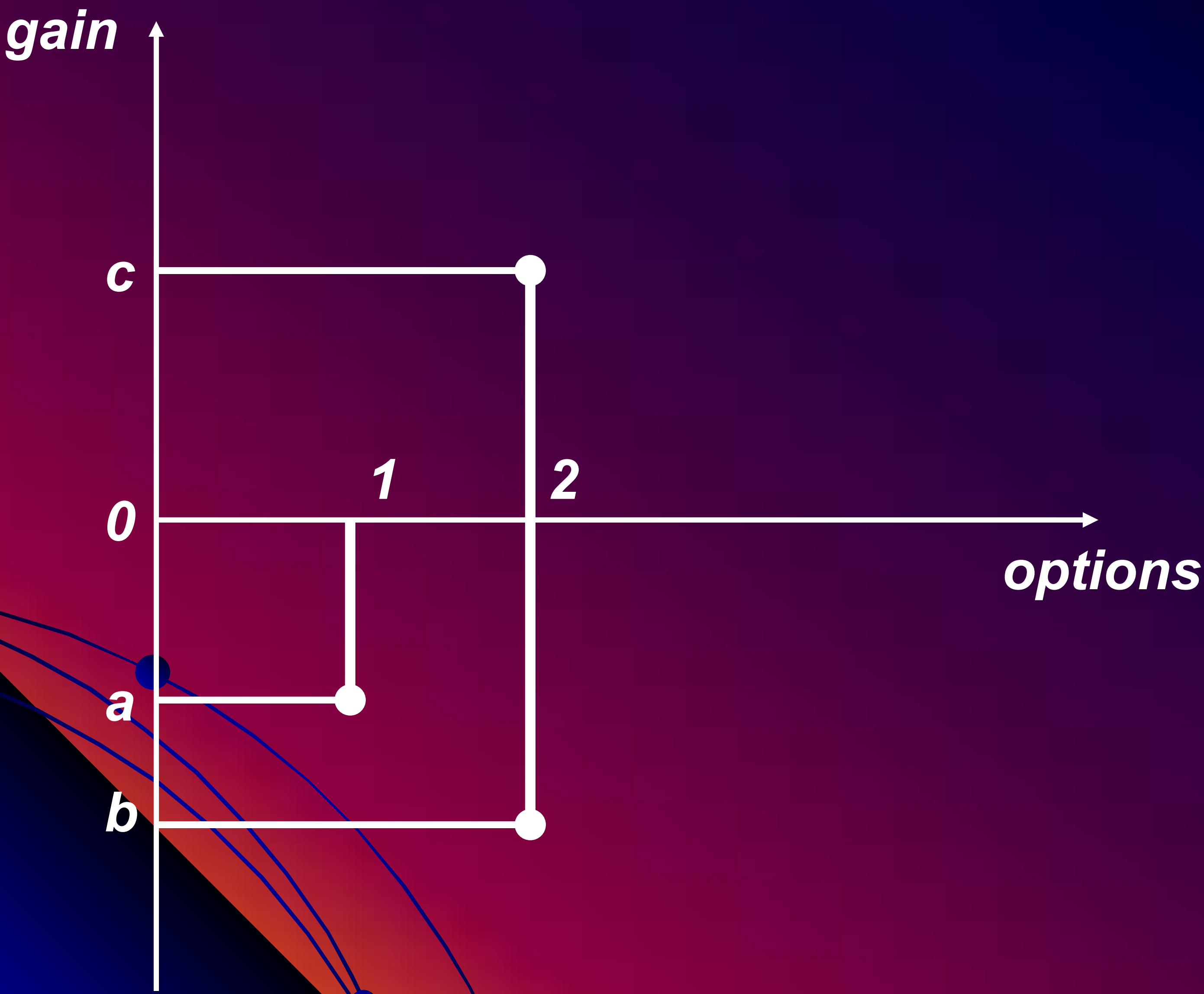
THE TRIP-PLANNING PROBLEM

- *I have to fly from A to D, and would like to get there as soon as possible*
- *I have two choices: (a) fly to D with a connection in B; or*
  (b) fly to D with a connection in C

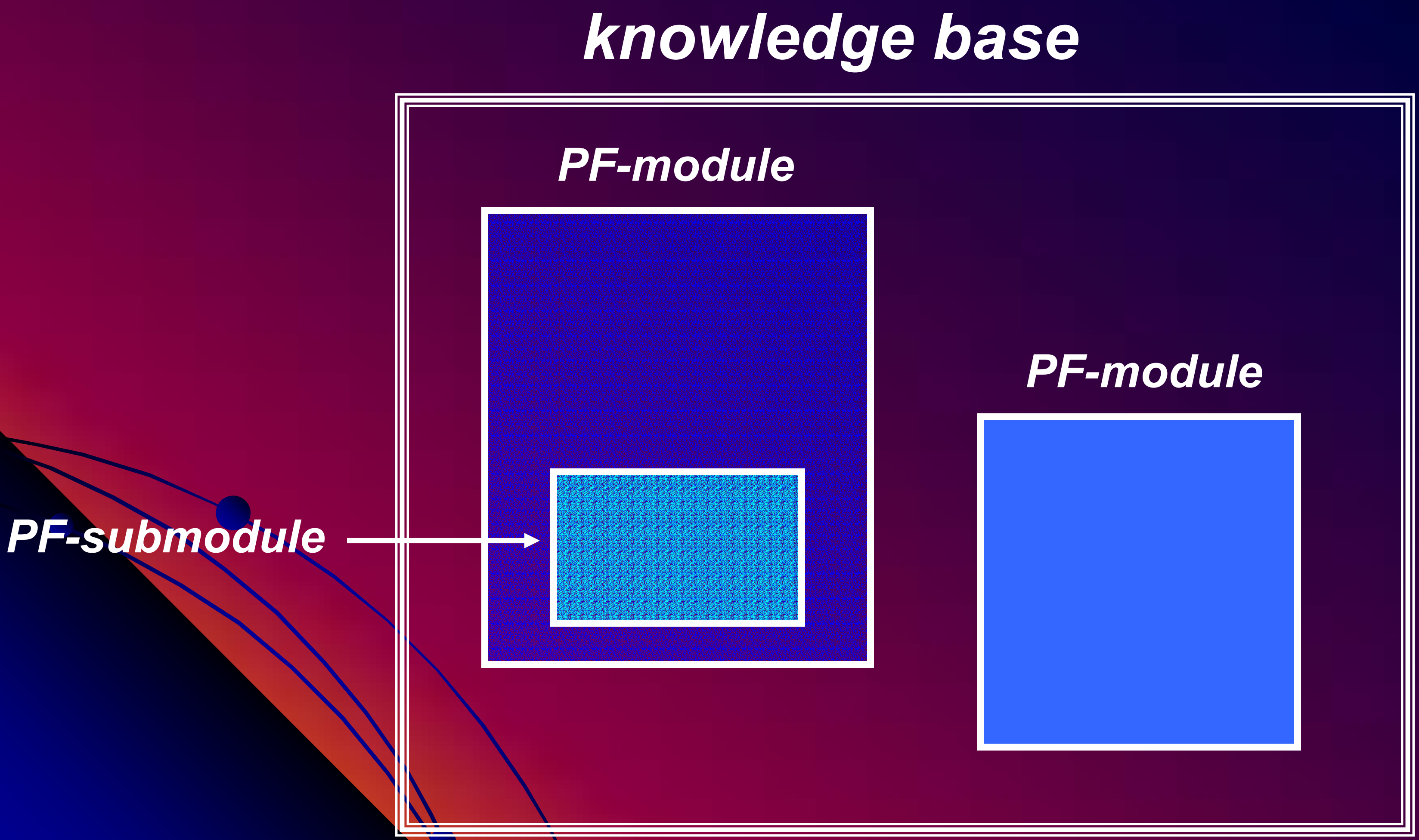


- *if I choose (a), I will arrive in D at time t_1*
- *if I choose (b), I will arrive in D at time t_2*
- *t_1 is earlier than t_2*
- *therefore, I should choose (a) ?*

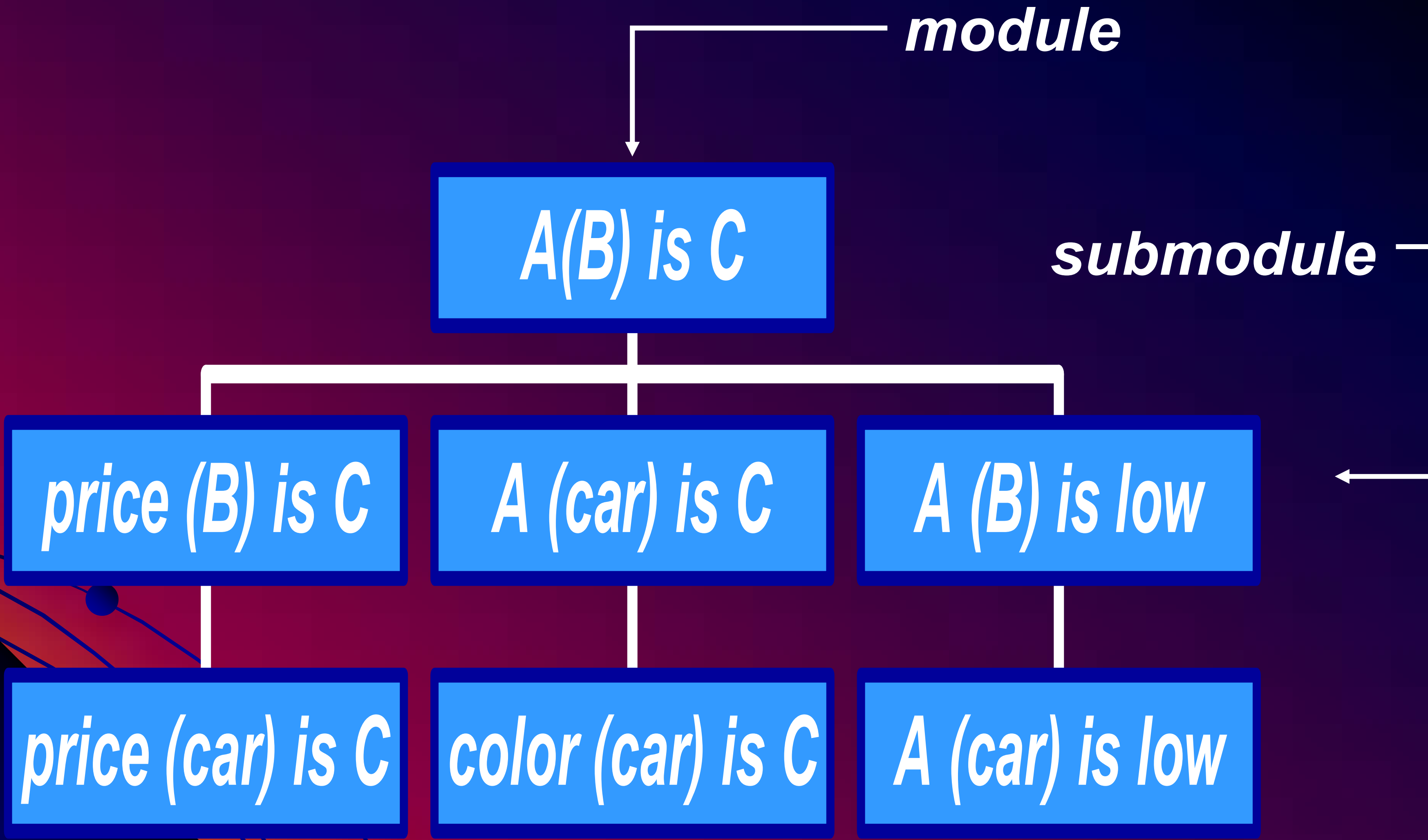
PROTOFORM EQUIVALENCE



PROTOFORM-CENTERED KNOWLEDGE ORGANIZATION



EXAMPLE



set of cars and their prices

BASIC STRUCTURE OF PNL

DICTIONARY 1

NL	GCL
p	GC(p)

DICTIONARY 2

GCL	PFL
GC(p)	PF(p)

MODULAR DEDUCTION DATABASE

- POSSIBILITY MODULE
- PROBABILITY MODULE (agent)
- FUZZY ARITHMETIC MODULE
- SEARCH MODULE
- FUZZY LOGIC MODULE
- EXTENSION PRINCIPLE MODULE

TEST QUERY (GOOGLE)

- *distance between largest city in Spain and largest city in Portugal: failure*
- *largest city in Spain: Madrid (success)*
- *largest city in Portugal: Lisbon (success)*
- *distance between Madrid and Lisbon*

  (success)

PROTOFORMAL SEARCH RULES

example

query: What is the distance between the largest city in Spain and the largest city in Portugal?

protoform of query: ?Attr (Desc(A), Desc(B))

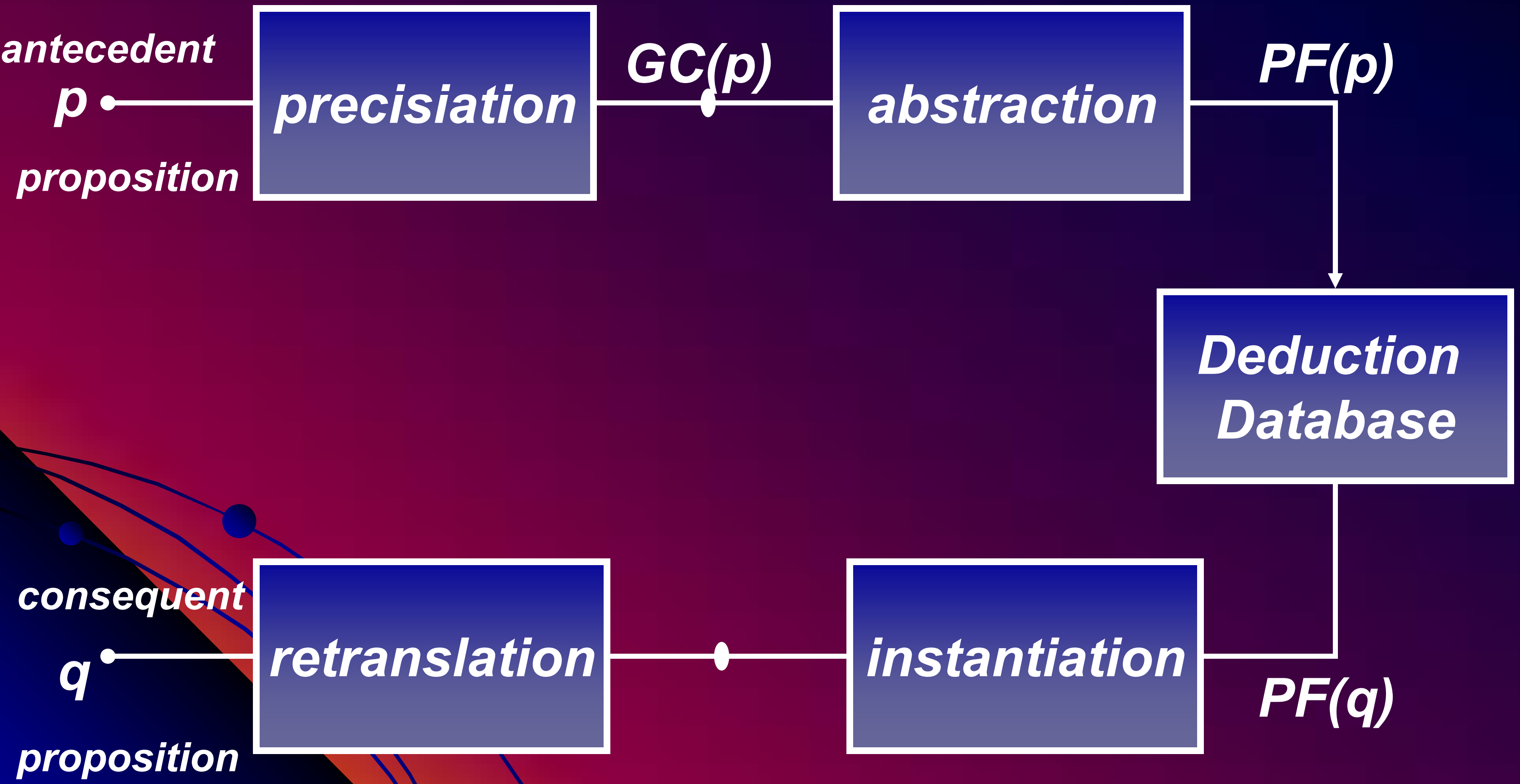
procedure

query: ?Name (A)|Desc (A)

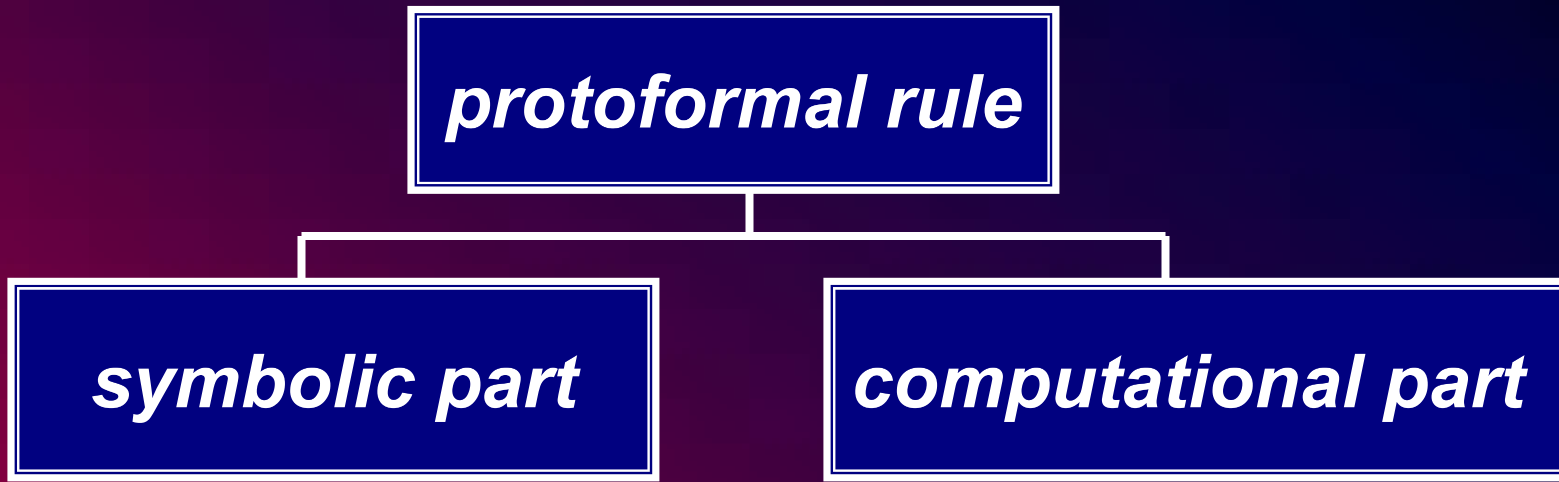
query: Name (B)|Desc (B)

query: ?Attr (Name (A), Name (B))

PROTOFORMAL (PROTOFORM-BASED) DEDUCTION



FORMAT OF PROTOFORMAL DEDUCTION RULES



PROTOFORM DEDUCTION RULE: GENERALIZED MODUS PONENS

fuzzy logic

classical

$$\frac{\begin{array}{l} A \\ A \rightarrow B \end{array}}{B}$$

symbolic

$$\frac{\begin{array}{l} X \text{ is } A \\ \text{If } X \text{ is } B \text{ then } Y \text{ is } C \end{array}}{Y \text{ is } D}$$

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

symbolic part

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

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

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

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

$$\text{subject to: } v = \int_U \mu_c(u)g(u)du$$

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

computational part

PROTOFORM-BASED (PROTOFORMAL) DEDUCTION

- *Rules of deduction in the Deduction Database (DDB) are protoformal*

examples: (a) compositional rule of inference

symbolic →

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

$$\mu_B(v) = sup(\mu_A(u) \wedge \mu_B(u,v))$$

computational

(b) extension principle

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

$$\mu_y(v) = sup_u(\mu_A(u))$$

Subject to: $v = f(u)$

symbolic

computational

THE TALL SWEDES PROBLEM

p: most Swedes are tall

Q: What is the average height of Swedes?

Try

p ⟶ p**: Count (B/A) is Q*

q ⟶ q**: F(C/A) is ?R*

Swedes

height attribute

functional of height attribute

*answer to q** cannot be inferred from p***

level of summarization of p has to be reduced

CONTINUED

$p \xrightarrow{\text{precisiation}} p^* =$ *Prop(tall.Swedes/Swedes) is most*

$q \xrightarrow{\text{precisiation}} q^* =$ *Ave.height is ?R*

$p^* \xrightarrow{\text{abs}} p^{**}$: *Prob F(B/A) is ?Q*

$q^* \xrightarrow{\text{abs}} q^{**}$: *Ave F(B/A) is ?R*

protoformal deduction rule

symbolic:

$$\frac{\text{Prop (F(B/A)) is Q}}{\text{Ave F(B/A) is R}}$$

computational:

$$\mu_{ave}(v) = \sup_g \mu_Q\left(\int \mu_B(u)g(u)du\right)$$

subject to

$$v = \frac{l}{M}\int g(u)du$$

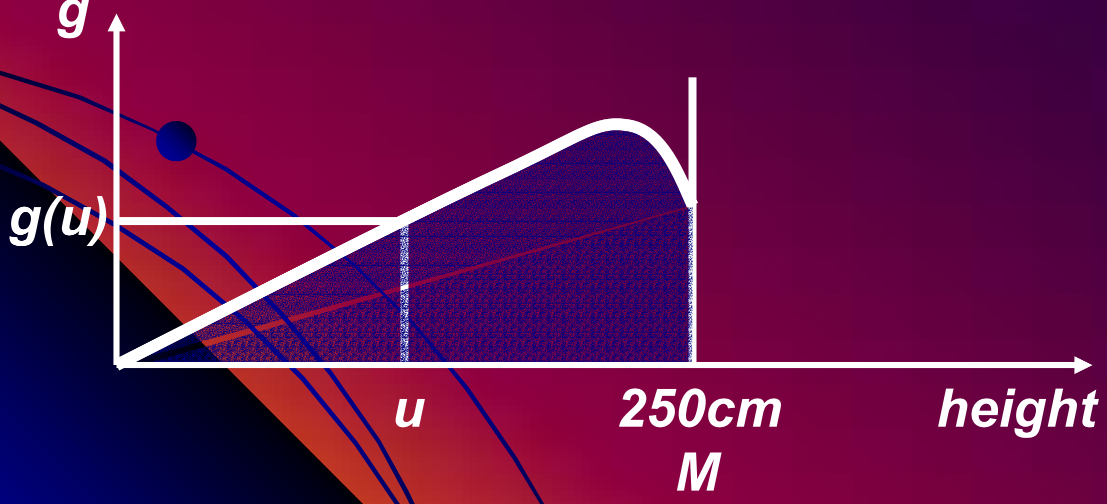
CONTINUED

- *example*

  IDS p: Most Swedes are tall

  TDS q: What is the average height of Swedes?

  g(u)= count density *g(u)du= number of Swedes whose height is between u and u+du*



PARTICULARIZATION (LAZ 1975)

P: population of objects

R: relation describing P

example

R: population of Swedes

R [Height; weight; age; …]

R^+: particularized R

R^+: [Height is tall]: population of tall Swedes

CONTINUED

$p \longrightarrow p^* = Count(Swedes[Height\ is\ tall]/Swedes)\ is\ most$

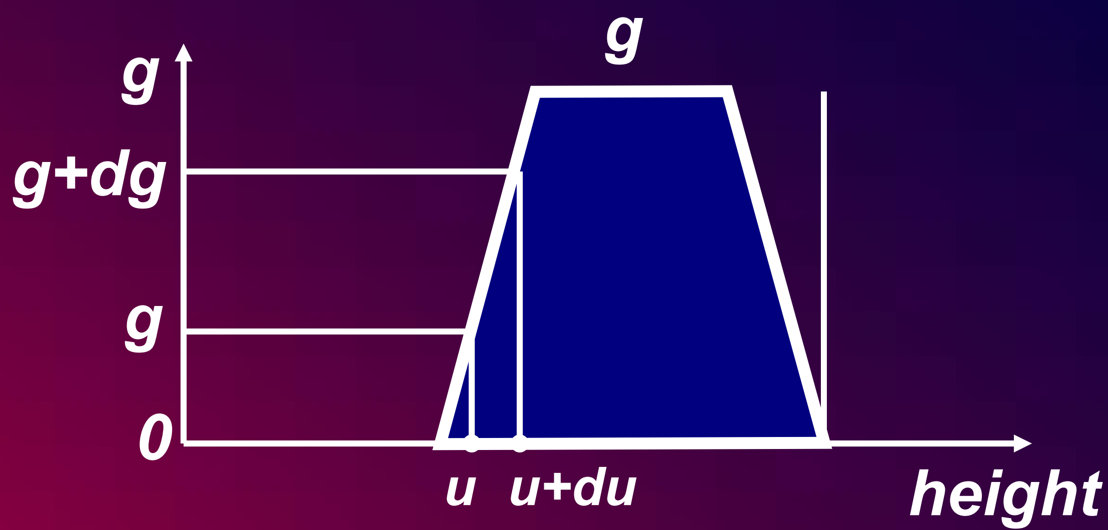
$p^{**}: Count(R[A\ is\ B]/R)\ is\ Q$

$q^* \longrightarrow q^{**}: ?\ Ave\ (R[A\ is\ B];\ A)$

rule:

$$\boxed{\begin{array}{c} Count(R[A\ is\ B]/R)\ is\ Q \\ Ave(R[A\ is\ B]\ is\ ?C \end{array}}$$

CONTINUED



g(u)= height distribution

$\frac{l}{M}\int^{M} g(u)\mu_{tall}(u)du$ *is most*

$\frac{l}{M}\int^{M} g(u)du$ *is ?C*

CONTINUED

$$\mu_C(v) = sup_g\, \mu_Q(\int_0^M g(u)\mu_Q(u)du$$

subject to

$$v = \frac{1}{M}\int_0^M g(u)du$$

RULES OF DEDUCTION

- *Rules of deduction are basically rules governing generalized constraint propagation*
- *The principal rule of deduction is the extension principle*

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

symbolic

$$\mu_B(v) = sup_u(\mu_A(u))$$

Subject to: $v = f(u)$

computational

GENERALIZATIONS OF THE EXTENSION PRINCIPLE

information = constraint on a variable

$f(X)$ is A	← given information about X
$g(X)$ is B	← inferred information about X

$$\mu_B(v) = sup_u(\mu_A(f(u)))$$

Subject to: $v = g(u)$

CONTINUED

$$\frac{f(X_1, \ldots, X_n) \text{ is } A}{g(X_1, \ldots, X_n) \text{ is } B}$$

$$\mu_B(v) = sup_u(\mu_A(f(u)))$$

Subject to: $v = g(u)$

$(X_1, \ldots, X_n)$ is A

$g_j(X_1, \ldots, X_n)$ is Y_j , $j=1, \ldots, n$

$(Y_1, \ldots, Y_n)$ is B

$$\mu_B(v) = sup_u(\mu_A(f(u)))$$

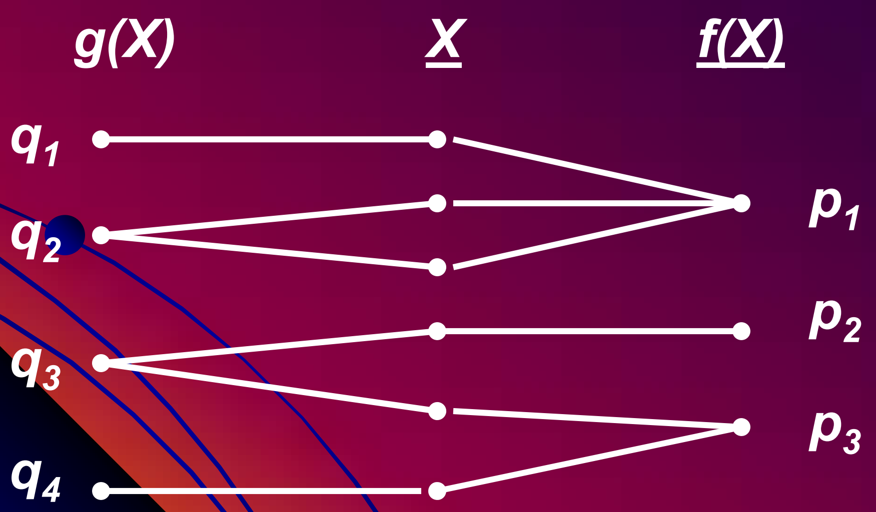
Subject to: $v = g(u)$

$j = 1,\ldots,n$

PROBLEM

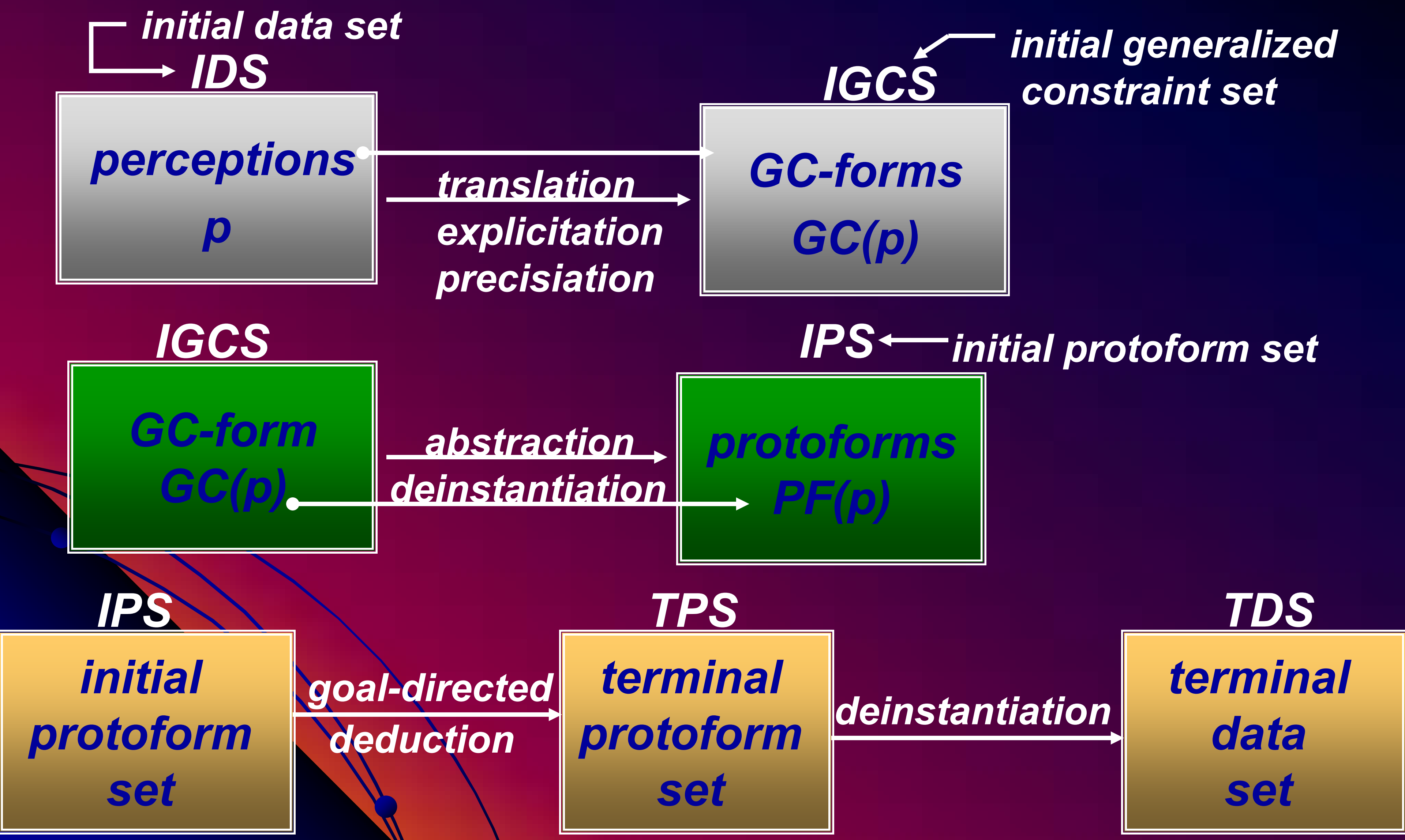
X: real-valued random variable

$$\frac{f(X) \text{ isp } P}{g(X) \text{ isr } ?Q}$$



$q_1 \leq p_1$; $q_2 \leq p_1$; $q_1 + q_2 = p_1 \cdots$

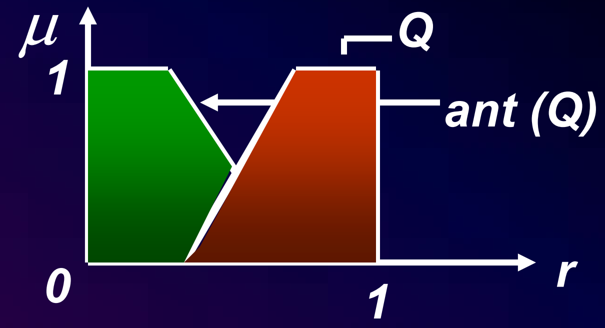
REASONING WITH PERCEPTIONS: DEDUCTION MODULE



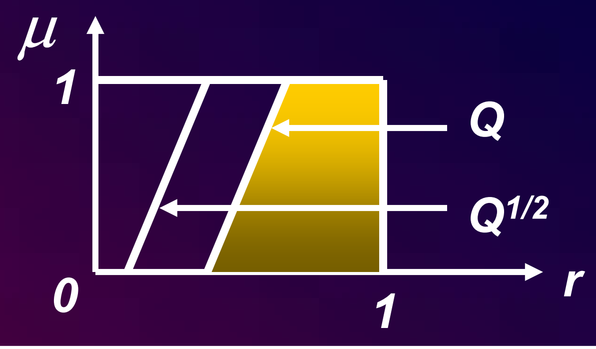
COUNT-AND MEASURE-RELATED RULES

crisp

$$\frac{Q\ A\text{'s are } B\text{'s}}{ant\ (Q)\ A\text{'s are not } B\text{'s}}$$



$$\frac{Q\ A\text{'s are } B\text{'s}}{Q^{1/2}\ A\text{'s are }{}^{2}B\text{'s}}$$



$$\frac{\text{most Swedes are tall}}{\text{ave (height) Swedes is ?h}} \longrightarrow \frac{Q\ A\text{'s are } B\text{'s}}{\text{ave } (B|A) \text{ is ?C}}$$

$$\mu_{ave}(v) = sup_a\, \mu_Q\left(\frac{1}{N} \Sigma_i\, \mu_B(a_i)\right) \quad , a = (a_1, \dots, a_N)$$

$$v = \frac{1}{N}\left(\Sigma_i\, a_i\right)$$

CONTINUED

not(QA's are B's) $\longleftrightarrow$ *(not Q) A's are B's*

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

$$\frac{\begin{array}{l} Q_1 \ \text{A's are B's} \\ Q_2 \ \text{A's are C's} \end{array}}{(Q_1 + Q_2 - 1) \ \text{A's are (B\&C)'s}}$$

PROBABILITY MODULE

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

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

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)

$$\frac{X \text{ isp } P \qquad Y = f(X)}{Y \text{ isp } f(P)}$$

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

$$\frac{X \text{ isp } P \qquad (X,Y) \text{ is } R}{Y \text{ isrs } S}$$

$$\frac{X \text{ isu } A \qquad Y = f(X)}{Y \text{ isu } f(A)}$$

PROBABILISTIC CONSTRAINT PROPAGATION RULE
(a special version of the generalized extension principle)

$$\int_U g(u)\mu_A(u)du \quad \text{is } R$$

$$\int_U g(u)\mu_B(u)du \quad \text{is } ?S$$

$$\mu_S(v) = \sup_g(\mu_R(\int_U g(u)\mu_A(u)du))$$

subject to

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

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

PROTOFORMAL DEDUCTION RULES

possibilistic extension principle

$$\frac{\begin{array}{l} X \text{ is } (\Sigma_i\, \mu_i / u_i) \\ Y = f(X) \end{array}}{Y \text{ is } (\Sigma_i\, \mu_i / f(u_i))}$$

$$\mu_i / u_i + \mu_j / u_i = (\mu_i \vee \mu_j) / u_i$$

probabilistic extension principle

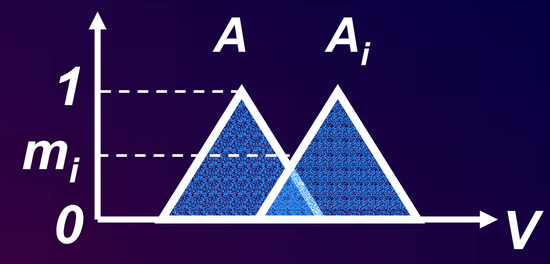
$$\frac{\begin{array}{l} X \text{ isp } (\Sigma_i\, p_i \backslash u_i) \\ Y = f(X) \end{array}}{Y \text{ isp } (\Sigma_i\, p_i \backslash f(u_i))}$$

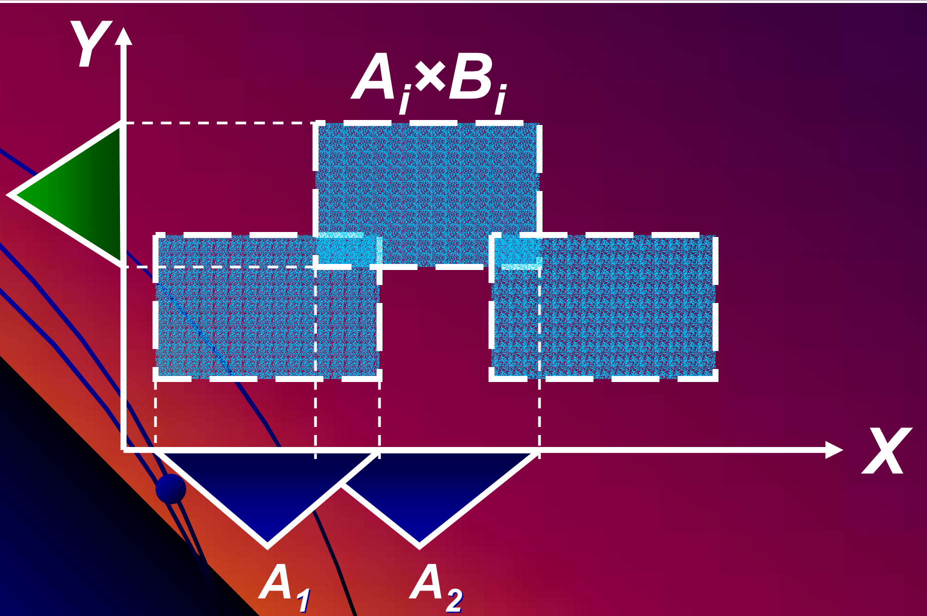
$$p_i \backslash u_i + p_j \backslash u_i = (p_i + p_j) \backslash u_i$$

COMPUTATION WITH PERCEPTIONS PROTOFORMAL RULE OF DEDUCTION

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

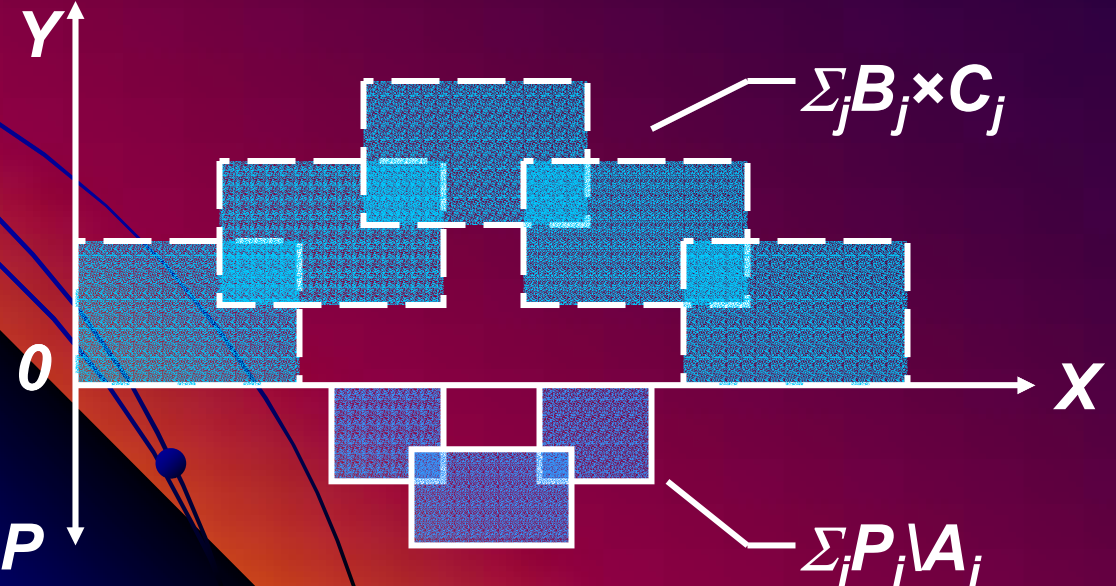
$$m_i = \sup (A \cap A_i)$$





PROTOFORMAL DEDUCTION RULE

$$\frac{\begin{array}{c} X \text{ ispa } (\Sigma_i P_i \backslash A_i) \\ Y \text{ isfq } (\Sigma_j B_j \times C_j) \end{array}}{Y \text{ isr } ?D}$$



PROTOFORMAL DEDUCTION RULE

$$\frac{\begin{array}{l} X \ \textit{ispa} \ (\Sigma_i P_i \backslash A_i) \\ Y = f(X) \end{array}}{Y \ \textit{isr} \ \ ?B}$$

$$\frac{\begin{array}{l} X \ \textit{ispb} \ (\Sigma_i P_i \backslash\backslash A_i) \\ Y = f(X) \end{array}}{Y \ \textit{isr} \ \ ?C}$$

PROTOFORMAL CONSTRAINT PROPAGATION

p	GC(p)	PF(p)
Dana is young	Age (Dana) is young	X is A
Tandy is a few years older than Dana	Age (Tandy) is (Age (Dana)) +few	Y is (X+B)

$$\frac{\begin{array}{c}X \text{ is } A \\ Y \text{ is } (X+B)\end{array}}{Y \text{ is } A+B} \longrightarrow \text{Age (Tandy) is (young+few)}$$

$$\mu_{A+B}(v) = \sup_u(\mu_A(u) \wedge \mu_B(v-u))$$

THE ROBERT EXAMPLE

PROTOFORMAL DEDUCTION THE ROBERT EXAMPLE

- ***The Robert example is intended to serve as an illustration of protoformal deduction. In addition, it is intended to serve as a test of ability of standard probability theory, PT, to operate on perception-based information***

IDS: Usually Robert returns from work at about 6 pm

TDS: What is the probability that Robert is home at about t pm?

SOLUTION

1. **Precisiation**

   p: usually Robert returns from work at about 6 pm

   p→p*: Prob(Return.Robert.from.work is about.6 pm is usually)

   X: Return.Robert.from.work; A: about.6 pm; B: usually

2. ***What is the probability that Robert is home at about t pm?***

   q→q*: Prob(Robert.home.at.about.t pm) is ? D

   Y: Robert.home; C: about.t pm; D: ? D

3. **Abstraction**

   p*→p**: Prob(X is A) is B

   q*→q**: Prob(Y is C) is ?D

CONTINUED

4. **Search in Deduction Database**

- ***desired rule:*** $\dfrac{\textit{Prob(X is A) is B}}{\textit{Prob(Y is C) is ?D}}$

- ***top-level agent reports that desired rule is not in DDB, but that a variant rule,***

  $\dfrac{\textit{Prob(X is A) is B}}{\textit{Prob(X is C) is ?D}}$,

  is in DDB

- ***Can the desired rule be linked to the variant rule?***

CONTINUED

5. **Computation**

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

computational part *(g: probability density of X)*

$$\mu_D(u) = \sup, (\mu_A(\int \mu_B(u)g(u)du))$$

subject to

$$v = \int \mu_B(u)q(u)du$$

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

CONTINUED

6. Search for linkage
   - If Robert does not leave his home after returning from work, then

     Robert is at home at about.t pm =

     Robert returns from work at.or.before t pm

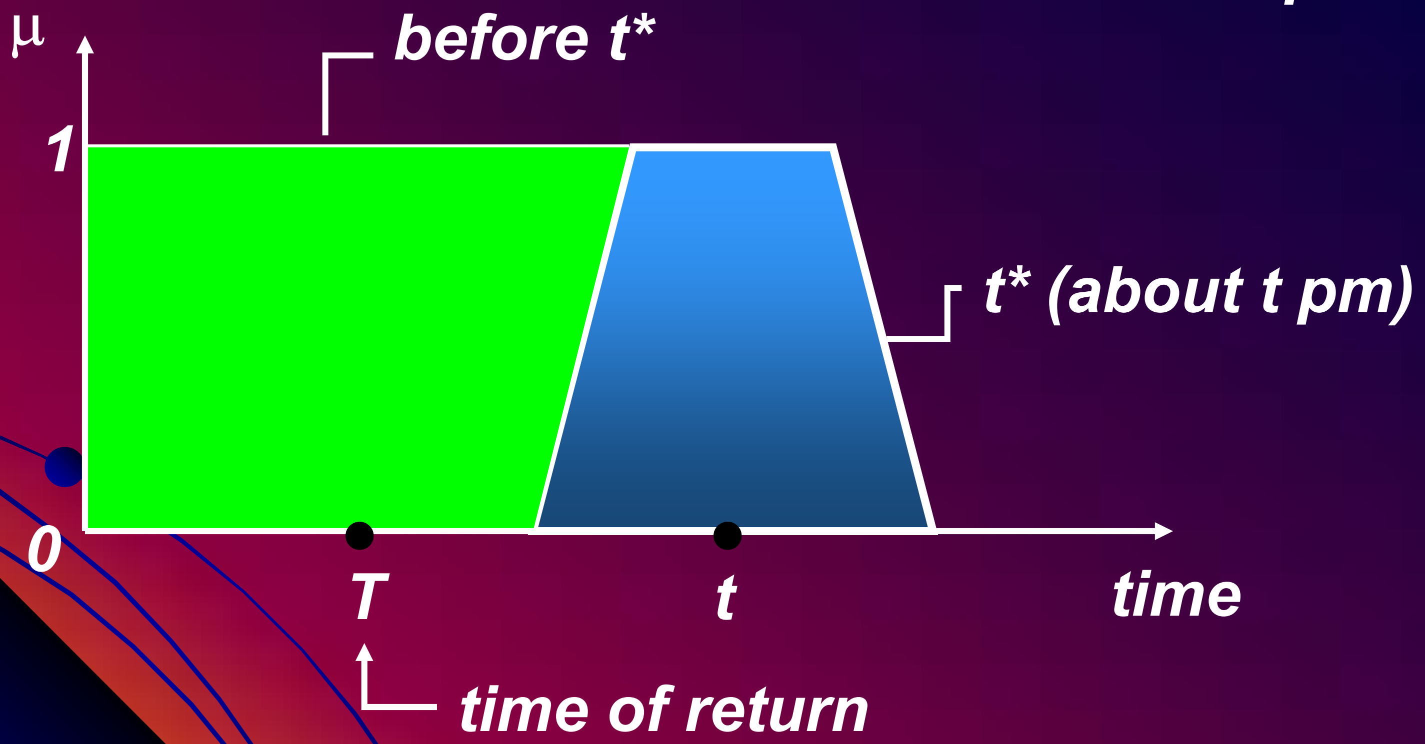
consequently

Y is about t pm= X is $\leq$ about.t pm

THE ROBERT EXAMPLE

event equivalence

Robert is home at about t pm= Robert returns from work before about t pm



Before about t pm= ≤ ο about t pm

CONTINUED

7. **Answer**

$$\mu_D(v) = \sup_g\left(\mu_{usually}\left(\int_{12\,am}^{12\,pm} \mu_{about\ .tpm}(u)g(u)du\right)\right)$$

subject to

$$v = \int_{12\,pm}^{12\,am} \mu_{about\,.tpm}(u)g(u)du$$

$$\int_{12\,pm}^{12\,am} g(u)du = 1$$

8. *Instantiation: D= Prob {Robert is home at about t}*

X= Time (Robert returns from work)

*A= 6**

B= usually

C= $\leq t^*$

SUMMATION

KEY POINTS

- *humans have a remarkable capability—a capability which machines do not have—to perform a wide variety of physical and mental tasks using only perceptions, with no measurements and no computations*

- *perceptions are intrinsically imprecise, reflecting the bounded ability of sensory organs, and ultimately the brain, to resolve detail and store information*

CONTINUED

- *imprecision of perceptions stands in the way of constructing a computational theory of perceptions within the conceptual structure of bivalent logic and bivalent-logic-based probability theory*

- *this is why existing scientific theories—based as they are on bivalent logic and bivalent-logic-based probability theory—provide no tools for dealing with perception-based information*

CONTINUED

- *in computing with words and perceptions (CWP), the objects of computation are propositions drawn from a natural language and, in particular, propositions which are descriptors of perceptions*
- *computing with words and perceptions is a methodology which may be viewed as (a) a new direction for dealing with imprecision, uncertainty and partial truth; and (b) as a basis for the analysis and design of systems which are capable of operating on perception-based information*

STATISTICS

Count of papers containing the word "fuzzy" in title, as cited in INSPEC and MATH.SCI.NET databases. (data for 2003 are not complete)

Compiled by Camille Wanat, Head, Engineering Library, UC Berkeley, November 20, 2003

	INSPEC/fuzzy	Math.Sci.Net/fuzzy
1970-1979	569	443
1980-1989	2,404	2,465
1990-1999	23,207	5,479
2000-present	9,945	2,865
1970-present	36,125	11,252

STATISTICS

- *Count of books containing the words "soft computing" in title, or published in series on soft computing. (source: Melvyl catalog)*

1994= 4
1995= 2
1996= 7
1997= 12
1998= 15
1999= 23
2000= 36
2001= 43
2002= 42
Total= 184

Compiled by Camille Wanat, Head, Engineering Library, UC Berkeley, October 12, 2003

- *Count of papers containing "soft computing" in title or published in proceedings of conferences on soft computing*

2494 (1994-2002)