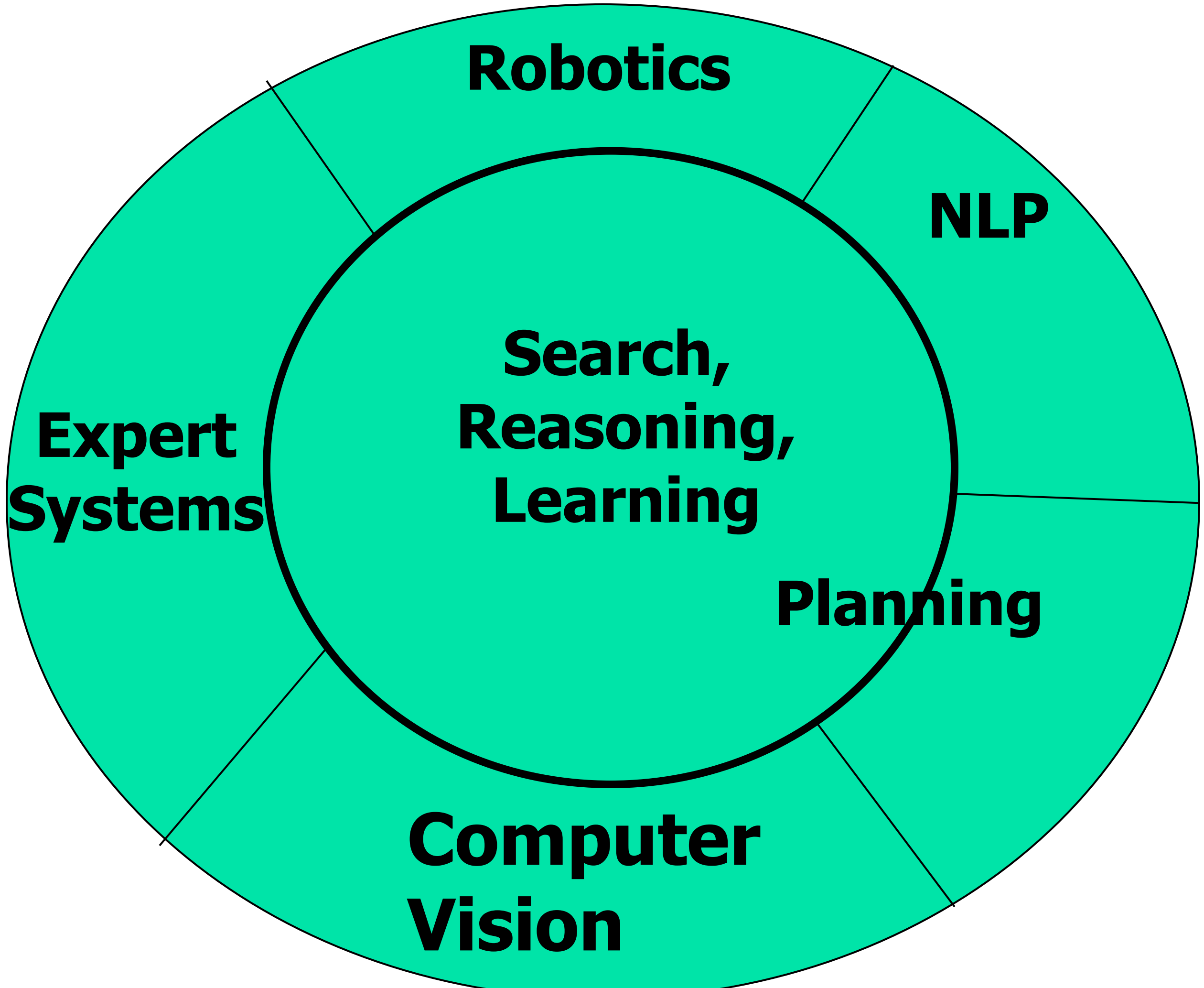
CS344: Introduction to Artificial Intelligence

(associated lab: CS386)

Pushpak Bhattacharyya
CSE Dept.,
IIT Bombay

Lecture–2: Fuzzy Logic and Inferencing

Disciplines which form the core of AI- inner circle
Fields which draw from these disciplines- outer circle.



Allied Disciplines

Philosophy	Knowledge Rep., Logic, Foundation of AI (is AI possible?)
Maths	Search, Analysis of search algos, logic
Economics	Expert Systems, Decision Theory, Principles of Rational Behavior
Psychology	Behavioristic insights into AI programs
Brain Science	Learning, Neural Nets
Physics	Learning, Information Theory & AI, Entropy, Robotics
Computer Sc. & Engg.	Systems for AI

Fuzzy Logic tries to capture the human ability of reasoning with imprecise information

- Models Human Reasoning
- Works with imprecise statements such as:

  In a process control situation, "*If* the temperature is <u>moderate</u> and the pressure is <u>high</u>, *then* turn the knob <u>slightly right</u>"
- The rules have "Linguistic Variables", typically adjectives qualified by adverbs (adverbs are <u>hedges</u>).

Underlying Theory: Theory of Fuzzy Sets

- Intimate connection between logic and set theory.
- Given any set 'S' and an element 'e', there is a very natural predicate, $\mu_s(e)$ called as the *belongingness predicate*.
- The predicate is such that,

$$\mu_s(e) = 1, \quad \textit{iff } e \in S$$
$$= 0, \quad \textit{otherwise}$$

- For example, $S = \{1, 2, 3, 4\}$, $\mu_s(1) = 1$ and $\mu_s(5) = 0$
- A predicate $P(x)$ also defines a set naturally.

  $S = \{x \mid P(x) \text{ is } \textit{true}\}$

  For example, *even(x)* defines $S = \{x \mid x \text{ is even}\}$

Fuzzy Set Theory (contd.)

- Fuzzy set theory starts by questioning the fundamental assumptions of set theory *viz.*, the belongingness predicate, μ, value is 0 or 1.
- Instead in Fuzzy theory it is assumed that,

$$\mu_s(e) = [0, 1]$$

- Fuzzy set theory is a generalization of classical set theory also called Crisp Set Theory.
- In real life *belongingness* is a fuzzy concept.

Example: Let, T = set of "tall" people

$$\mu_T(\text{Ram}) = 1.0$$

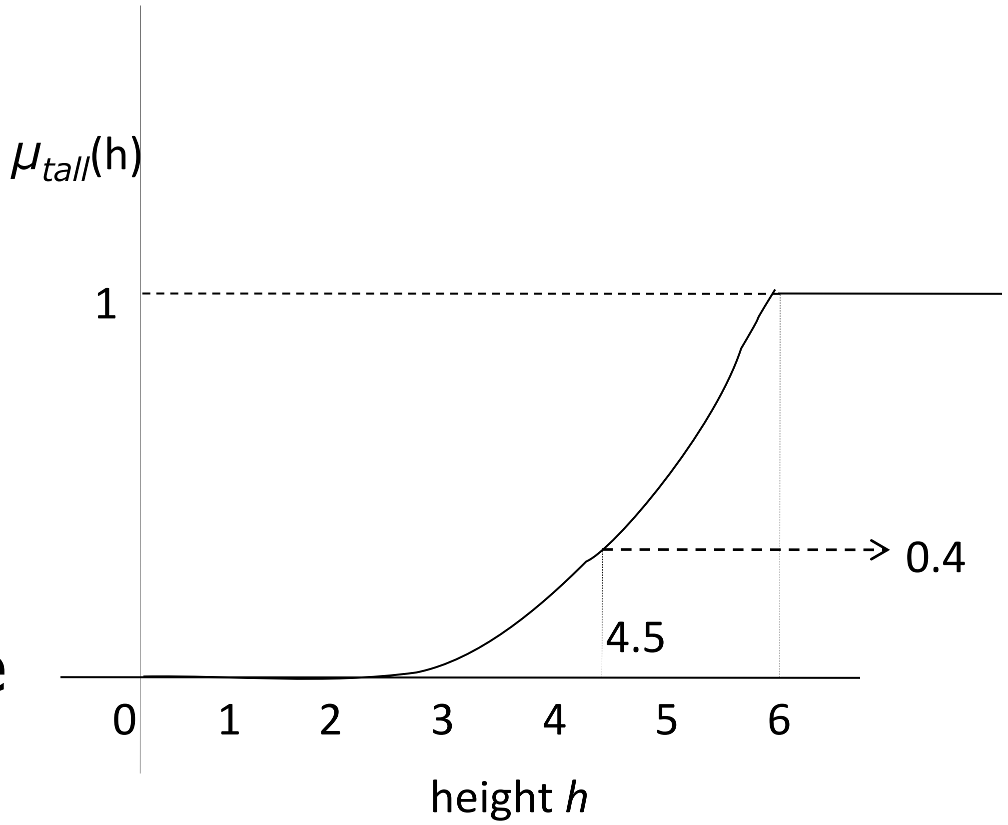
$$\mu_T(\text{Shyam}) = 0.2$$

Shyam belongs to T with degree *0.2*.

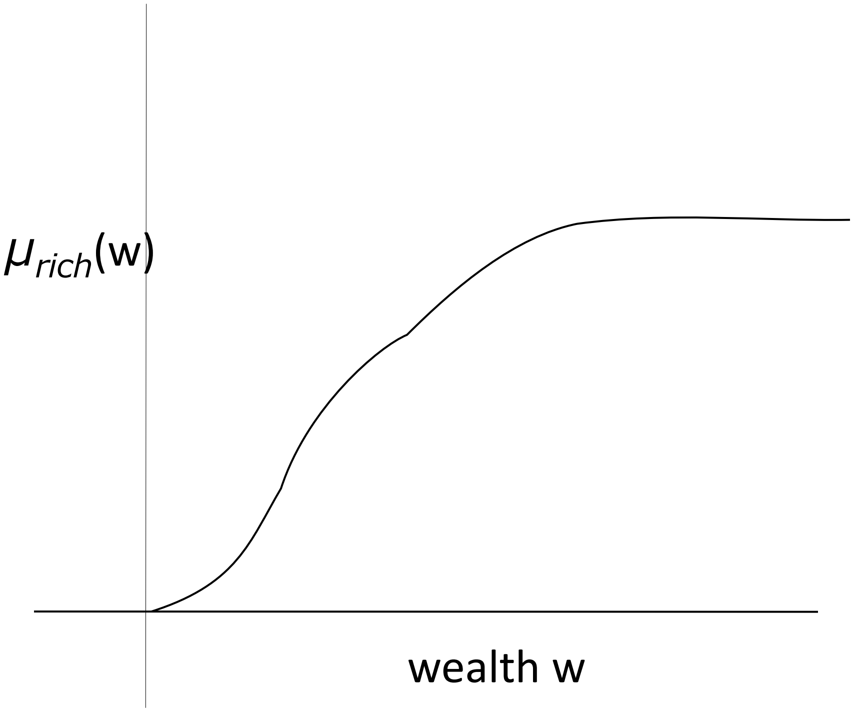
Linguistic Variables

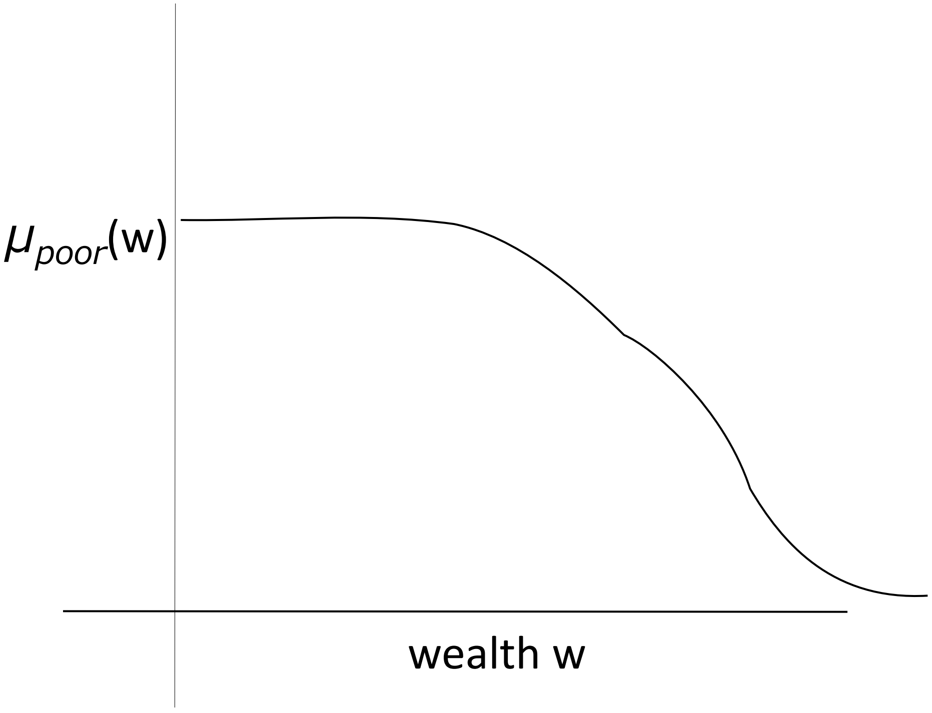
- Fuzzy sets are named by Linguistic Variables (typically adjectives).
- Underlying the LV is a numerical quantity

  E.g. For 'tall' (LV), 'height' is numerical quantity.
- Profile of a LV is the plot shown in the figure shown alongside.

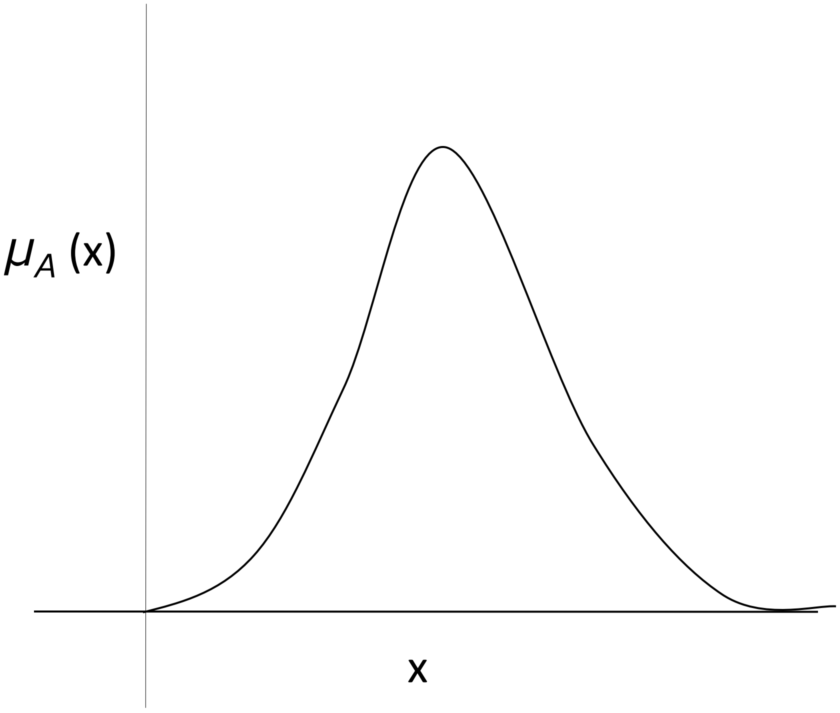


Example Profiles

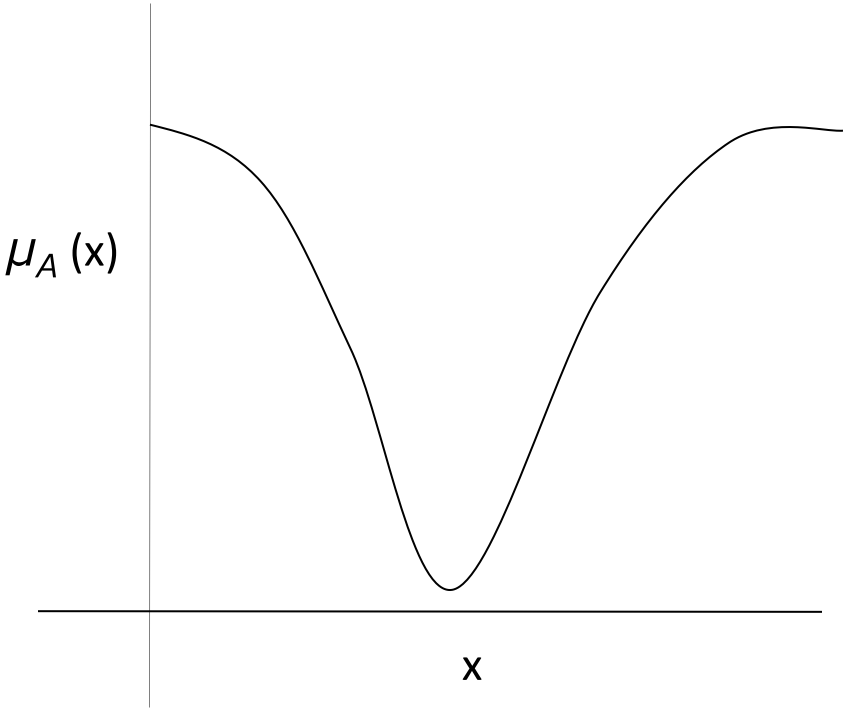




Example Profiles



Profile representing
moderate (*e.g.* moderately rich)



Profile representing
extreme

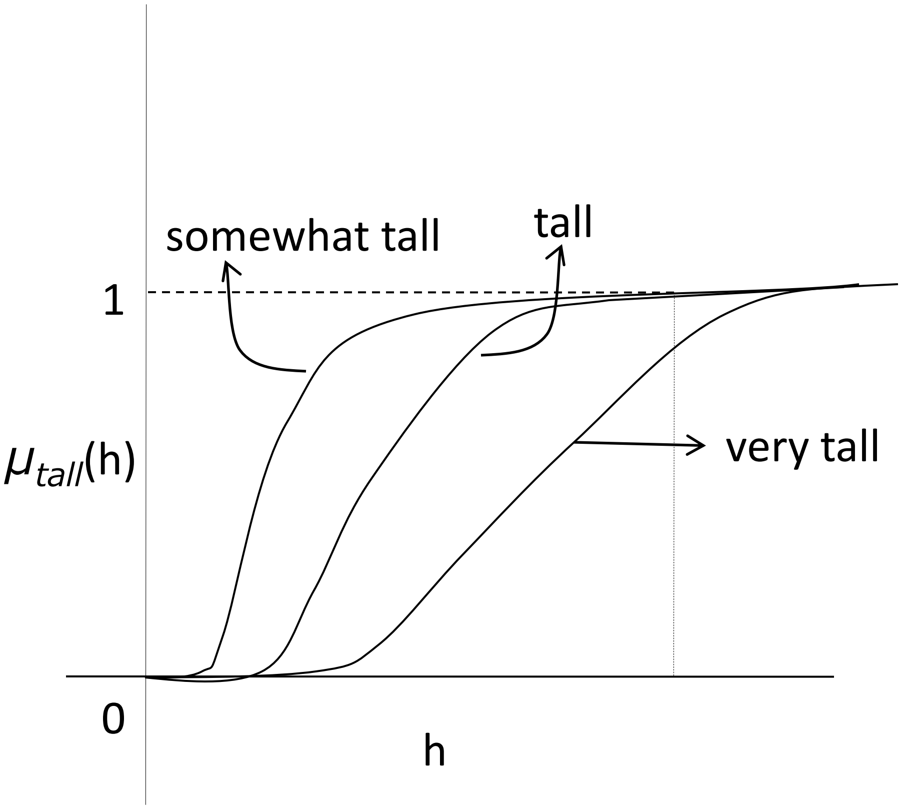
Concept of Hedge

- Hedge is an intensifier
- Example:

  LV = tall, LV_1 = very tall, LV_2 = somewhat tall
- 'very' operation:

  $\mu_{very\ tall}(x) = \mu^2_{tall}(x)$
- 'somewhat' operation:

$\mu_{somewhat\ tall}(x) = \sqrt{(\mu_{tall}(x))}$



Representing sets

- 2 ways of representing sets
- By extension – actual listing of elements
  - A = {2, 4, 6, 8,....}
- By intension – assertion of properties of elements belonging to the set
  - $A = \{x | x \textit{ mod } 2 = 0 \}$

Belongingness Predicate

- Let U = {1,2,3,4,5,6}
- Let A = {2,4,6}
- A = {0.0/1, 1.0/2, 0.0/3, 1.0/4, 0.0/5, 1.0/6}

- Every subset of U is a point in a 6 dimensional space

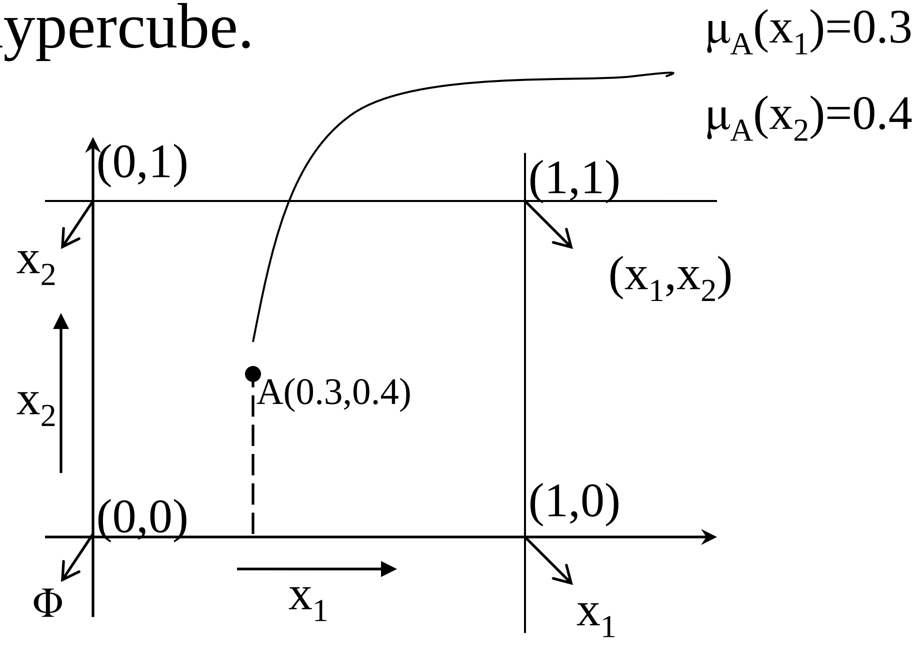
Representation of Fuzzy sets

Let $U = \{x_1, x_2, \ldots\ldots, x_n\}$

$|U| = n$

The various sets composed of elements from U are presented as points on and inside the n-dimensional hypercube. The crisp sets are the corners of the hypercube.

$U=\{x_1,x_2\}$



A fuzzy set A is represented by a point in the n-dimensional space as the point $\{\mu_A(x_1), \mu_A(x_2), \ldots\ldots \mu_A(x_n)\}$

Degree of fuzziness

The centre of the hypercube is the "most fuzzy" set. Fuzziness decreases as one nears the corners

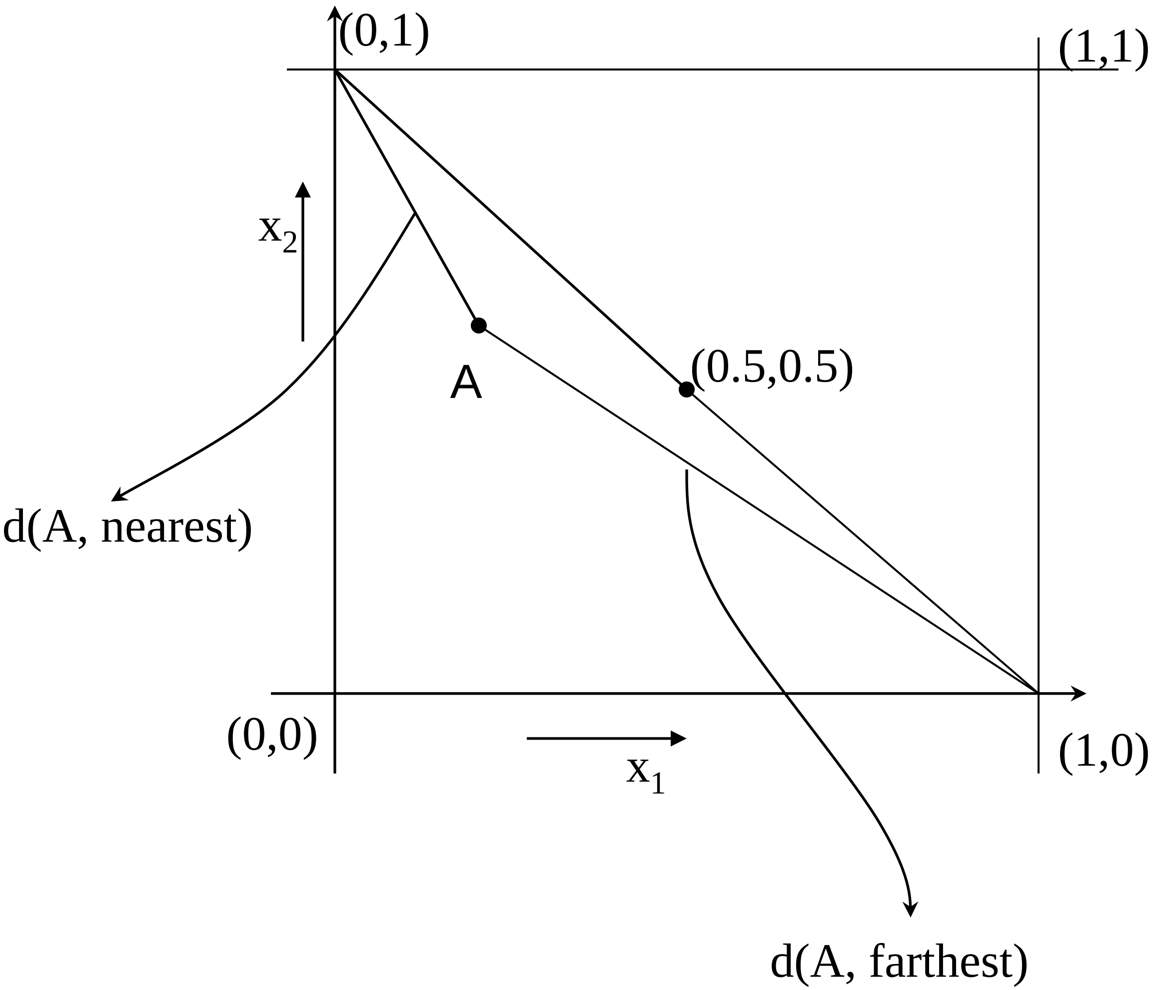
Measure of fuzziness

Called the entropy of a fuzzy set

Fuzzy set — Farthest corner

$$E(S) = d(S, nearest) / d(S, farthest)$$

Entropy — Nearest corner


(0,1)
(1,1)
x_2
A
(0.5,0.5)
d(A, nearest)
(0,0)
x_1
(1,0)
d(A, farthest)

Definition

Distance between two fuzzy sets

$$d(S_1, S_2) = \sum_{i=1}^{n} | \mu_{s_1}(x_i) - \mu_{s_2}(x_i) |$$

L_1 - norm

Let C = fuzzy set represented by the centre point

d(c,nearest) = |0.5-1.0| + |0.5 – 0.0|

= 1

= d(C,farthest)

=> E(C) = 1

Definition

Cardinality of a fuzzy set

$$m(s) = \sum_{i=1}^{n} \mu_s(x_i)$$ [generalization of cardinality of classical sets]

Union, Intersection, complementation, subset hood

$$\mu_{s_1 \cup s_2}(x) = \max[\mu_{s_1}(x), \mu_{s_2}(x)] \forall x \in U$$

$$\mu_{s_1 \cap s_2}(x) = \min[\mu_{s_1}(x), \mu_{s_2}(x)] \forall x \in U$$

$$\mu_{s^c}(x) = 1 - \mu_s(x)$$

Note on definition by extension and intension

$S_1 = \{x_i | x_i \bmod 2 = 0\}$ – Intension

$S_2 = \{0,2,4,6,8,10,\ldots\ldots\ldots\ldots\}$ – extension

How to define subset hood?

Conceptual problem

$\mu_B(x) <= \mu_A(x)$ means

$B \varepsilon P(A)$, *i.e.*, $\mu_{P(A)}(B)=1;$

Goes against the grain of fuzzy logic

History of Fuzzy Logic

- Fuzzy logic was first developed by Lofti Zadeh in 1967
- μ took values in [0,1]
- Subsethood was given as

  $$\mu B(x) <= \mu A(x) \text{ for all } x$$

- This was questioned in 1970s leading to Lukasiewitz formula

Lukasiewitz formula for Fuzzy Implication

- t(P) = truth value of a proposition/predicate. In fuzzy logic t(P) = [0,1]
- $t(P \rightarrow Q) = \min[1, 1 - t(P) + t(Q)]$

Lukasiewitz definition of implication

Fuzzy Inferencing

- Two methods of inferencing in classical logic
  - Modus Ponens
    - Given *p* and $p \rightarrow q$, infer *q*
  - Modus Tolens
    - Given $\sim q$ and $p \rightarrow q$, infer $\sim p$
- How is fuzzy inferencing done?

Classical Modus Ponens in tems of truth values

- Given $t(p)=1$ and $t(p \rightarrow q)=1$, infer $t(q)=1$
- In fuzzy logic,
  - given $t(p)>=a$, $0<=a<=1$
  - *and* $t(p \rightarrow >q)=c$, $0<=c<=1$
  - What is $t(q)$
- How much of truth is transferred over the channel

$$p \Rightarrow q$$

Use Lukasiewitz definition

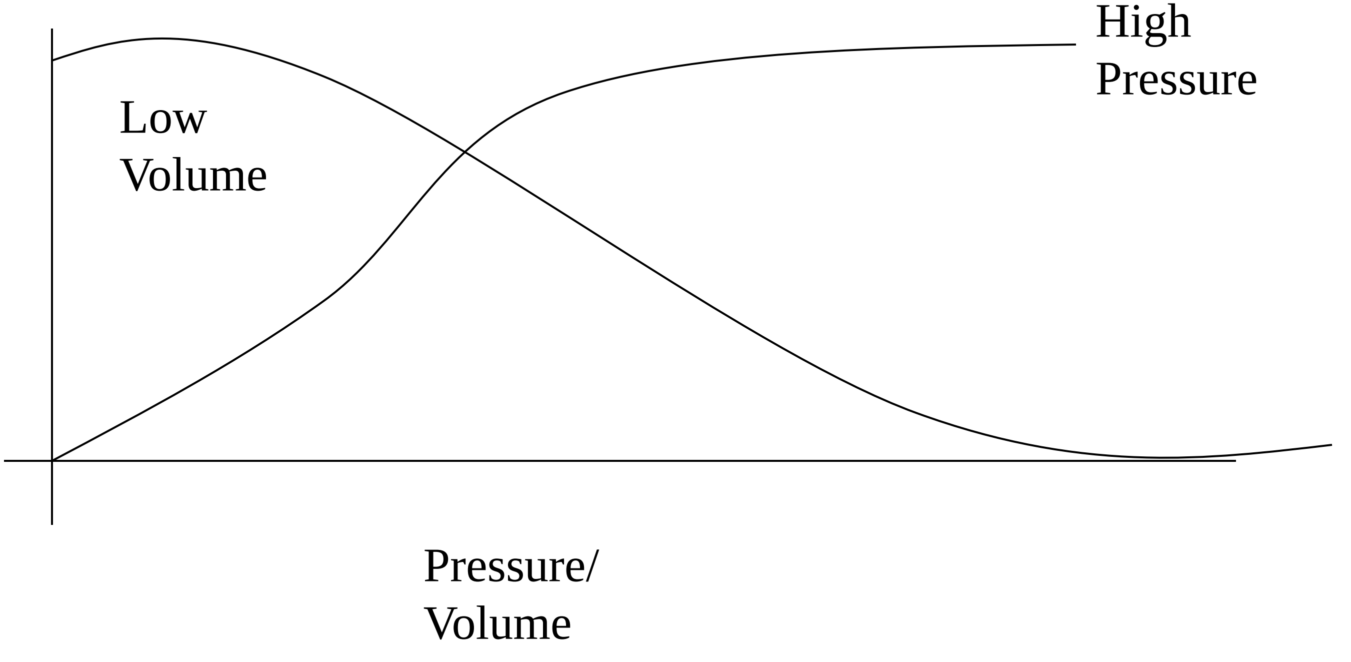
- $t(p \rightarrow q) = min[1, 1 - t(p) + t(q)]$
- We have $t(p\text{->}q)=c$, *i.e.*, $min[1,1-t(p)+t(q)]=c$
- Case 1:
- $c=1$ gives $1-t(p)+t(q)>=1$, *i.e.*, $t(q)>=a$
- Otherwise, $1-t(p)+t(q)=c$, *i.e.*, $t(q)>=c+a-1$
- Combining, $t(q)=max(0,a+c-1)$
- This is the amount of truth transferred over the channel $p \rightarrow q$

ANDING of Clauses on the LHS of implication

$$t(P \wedge Q) = \min(t(P), t(Q))$$

Eg: If pressure is high then Volume is low

$$t(high(pressure) \rightarrow low(volume))$$



Fuzzy Inferencing

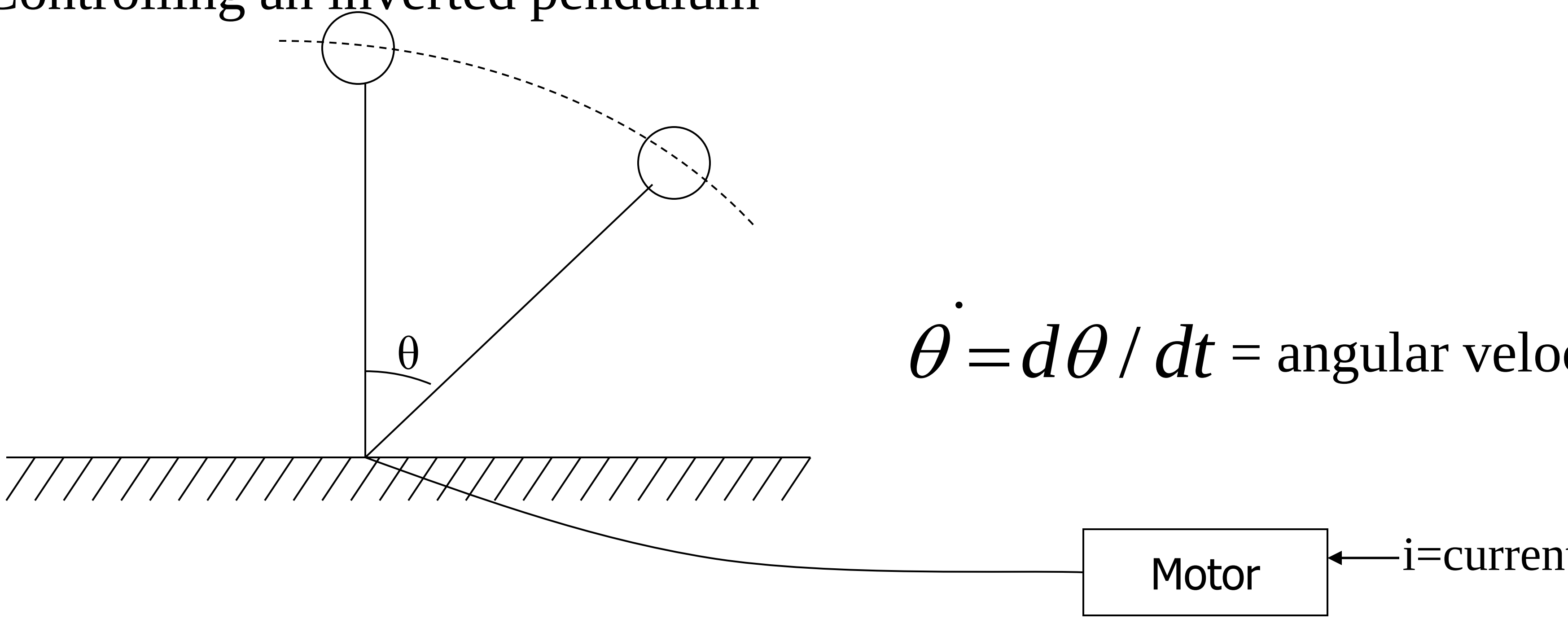
Core

The Lukasiewitz rule

$$t(P \rightarrow Q) = \min[1, 1 + t(P) - t(Q)]$$

An example

Controlling an inverted pendulum



$$\dot{\theta} = d\theta / dt = \text{angular velocity}$$

The goal: To keep the pendulum in vertical position (θ=0) in dynamic equilibrium. Whenever the pendulum departs from vertical, a torque is produced by sending a current 'i'

Controlling factors for appropriate current

Angle θ, Angular velocity $\dot{\theta}$

Some intuitive rules

If θ is +ve small and $\dot{\theta}$ is –ve small

then current is zero

If θ is +ve small and $\dot{\theta}$ is +ve small

then current is –ve medium

Control Matrix

$\dot{\theta}$ \ θ	-ve med	-ve small	Zero	+ve small	+ve med	
-ve med						
-ve small		+ve med	+ve small	Zero		
Zero		+ve small	Zero	-ve small		
+ve small		Zero	-ve small	-ve med		
+ve med						

Region of interest

Each cell is a rule of the form

If θ is <> and $\dot{\theta}$ is <>

then i is <>

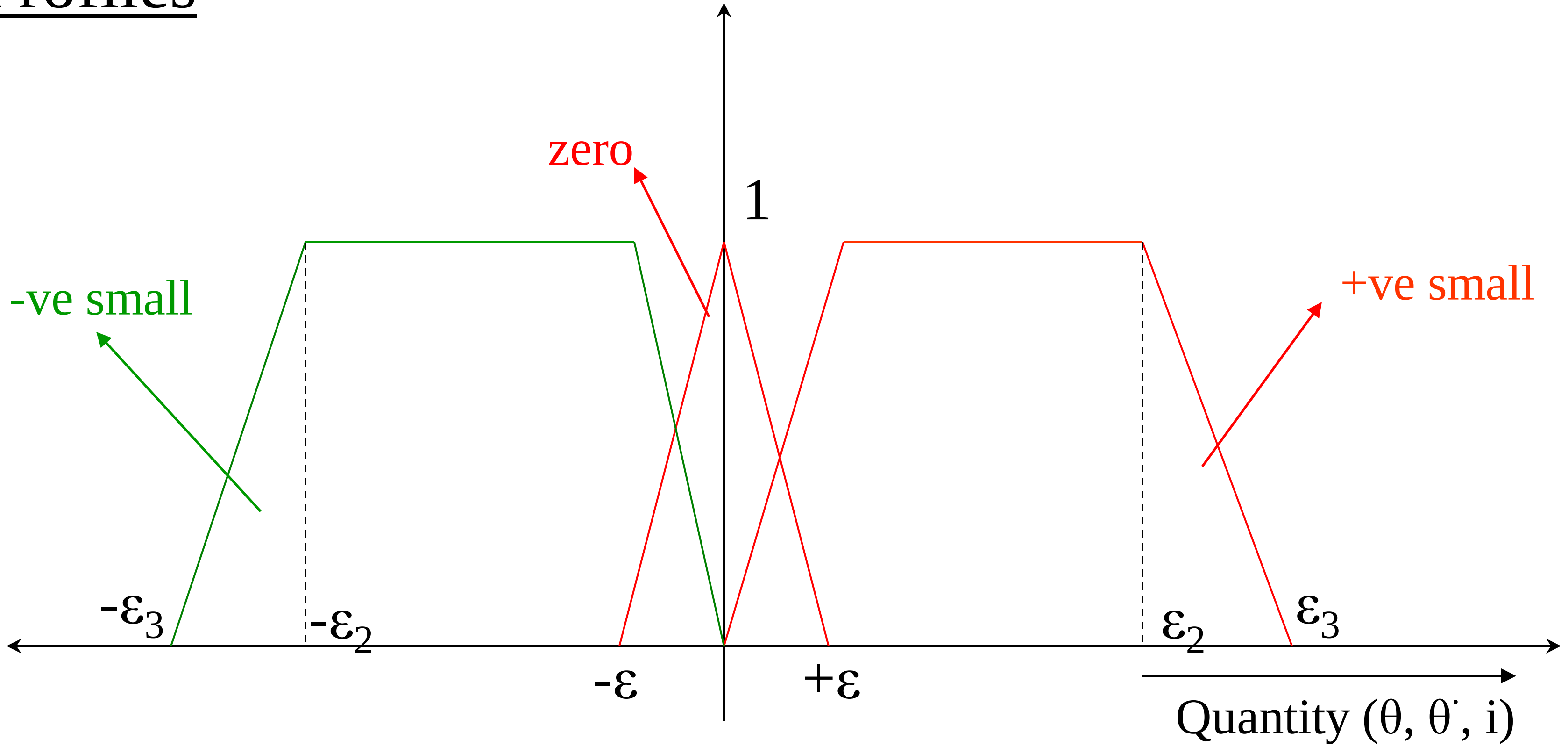
4 “Centre rules”

1. if θ = = Zero and $\dot{\theta}$ = = Zero then i = Zero
2. if θ is +ve small and $\dot{\theta}$ = = Zero then i is –ve small
3. if θ is –ve small and $\dot{\theta}$ = = Zero then i is +ve small
4. if θ = = Zero and $\dot{\theta}$ is +ve small then i is –ve small
5. if θ = = Zero and $\dot{\theta}$ is –ve small then i is +ve small

Linguistic variables

1. Zero
2. +ve small
3. -ve small

<u>Profiles</u>

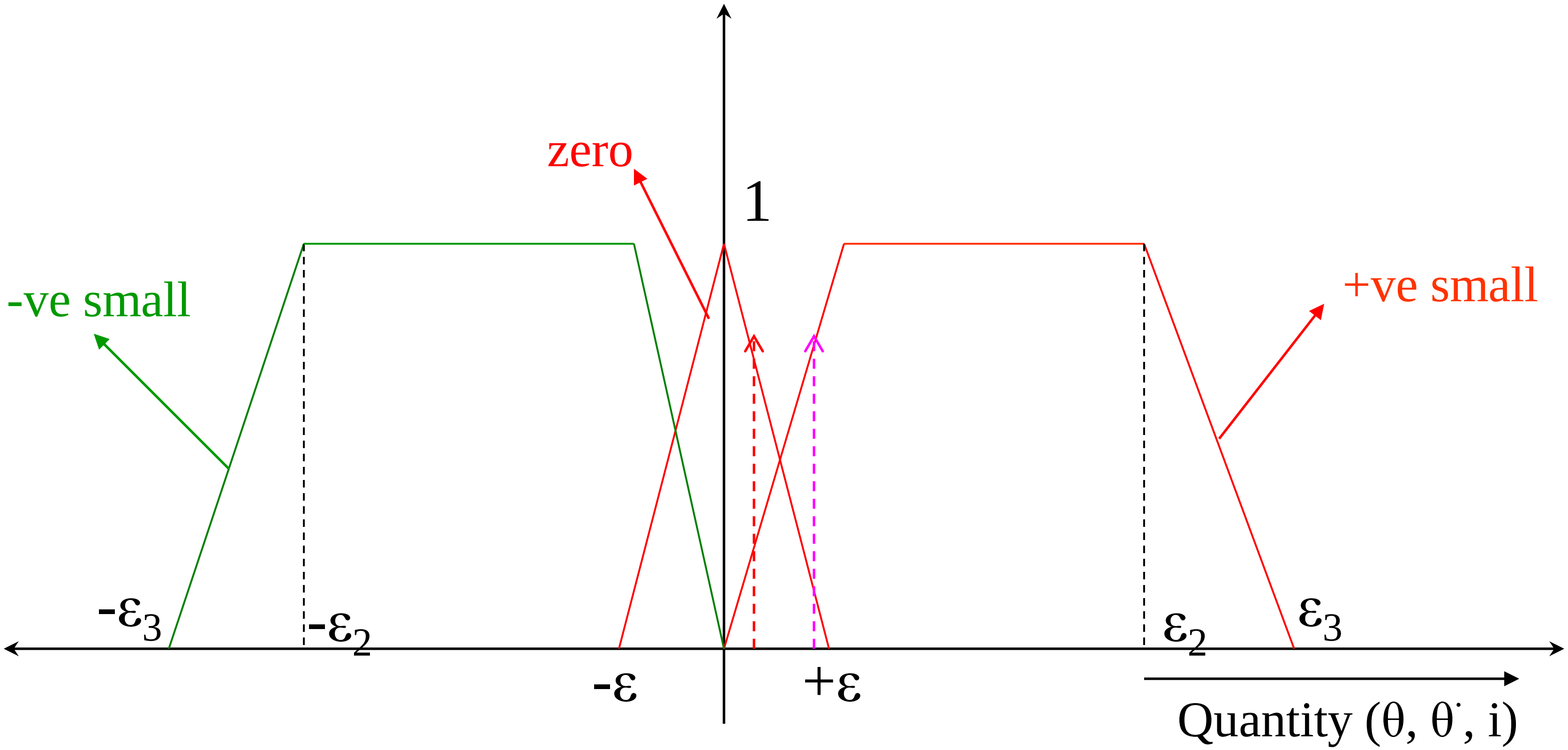


Inference procedure

1. Read actual numerical values of θ and $\dot{\theta}$
2. Get the corresponding μ values μ_{Zero}, $\mu_{(+ve\ small)}$, $\mu_{(-ve\ small)}$. This is called FUZZIFICATION
3. For different rules, get the fuzzy I-values from the R.H.S of the rules.
4. "Collate" by some method and get ONE current value. This is called DEFUZZIFICATION
5. Result is one numerical value of 'i'.

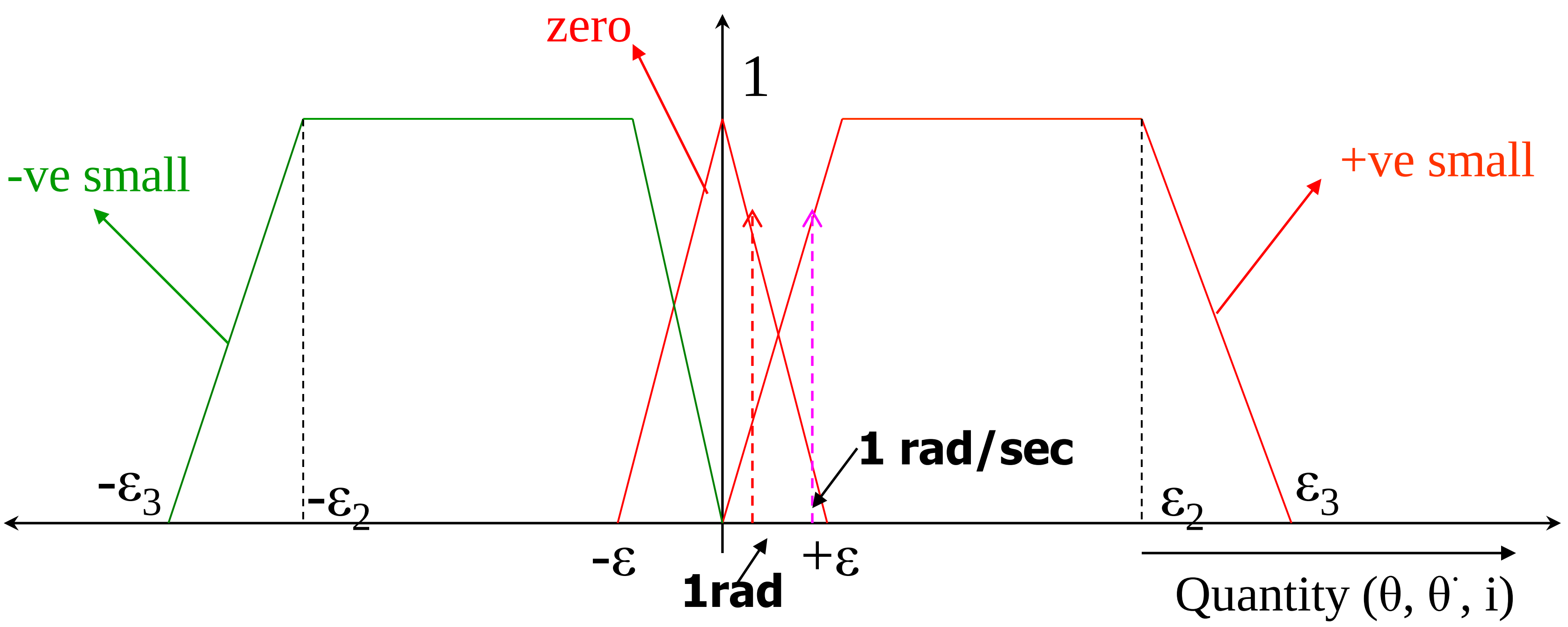
Rules Involved

if θ is Zero and dθ/dt is Zero then i is Zero
if θ is Zero and dθ/dt is +ve small then i is –ve small
if θ is +ve small and dθ/dt is Zero then i is –ve small
if θ +ve small and dθ/dt is +ve small then i is -ve medium



Fuzzification

Suppose θ is 1 radian and dθ/dt is 1 rad/sec
$\mu_{zero}(\theta = 1) = 0.8$ (say)
$M_{+ve\text{-}small}(\theta = 1) = 0.4$ (say)
$\mu_{zero}(d\theta/dt = 1) = 0.3$ (say)
$\mu_{+ve\text{-}small}(d\theta/dt = 1) = 0.7$ (say)

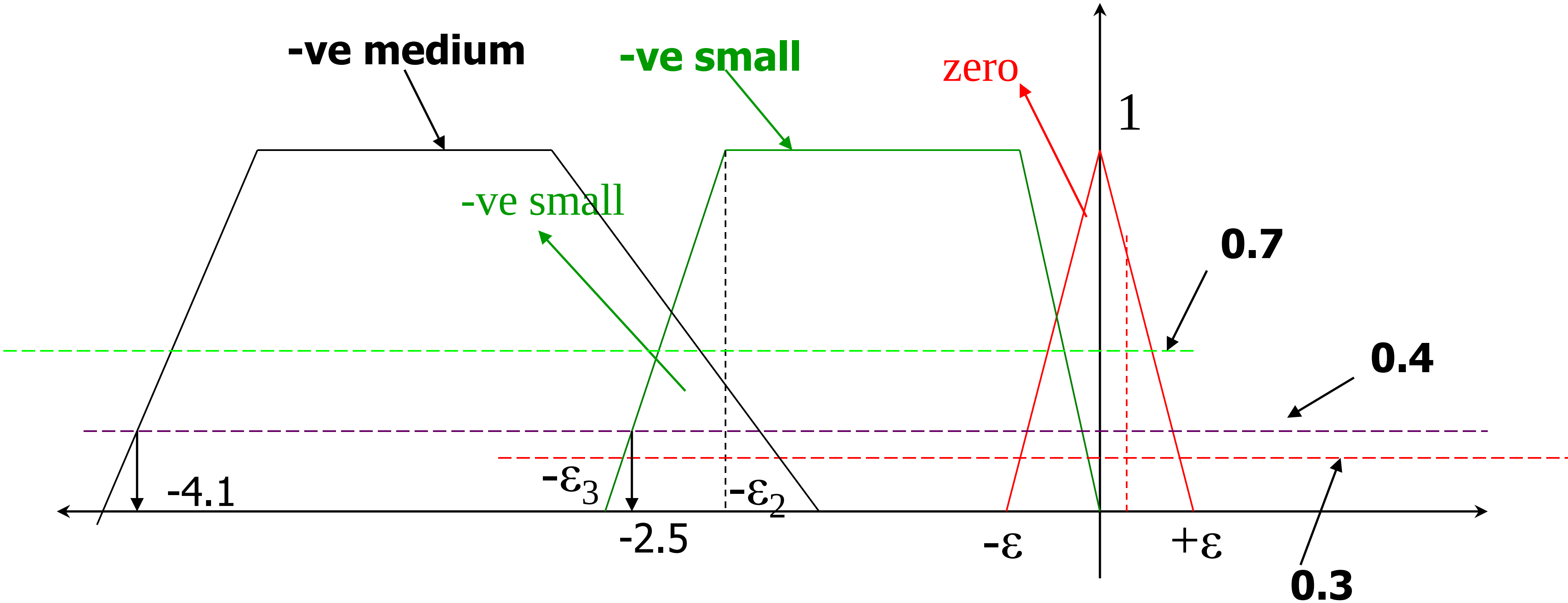


Fuzzification

Suppose θ is 1 radian and dθ/dt is 1 rad/sec
$\mu_{zero}(\theta = 1) = 0.8$ (say)
$\mu_{+ve\text{-}small}(\theta = 1) = 0.4$ (say)
$\mu_{zero}(d\theta/dt = 1) = 0.3$ (say)
$\mu_{+ve\text{-}small}(d\theta/dt = 1) = 0.7$ (say)

if θ is Zero and dθ/dt is Zero then i is Zero
min(0.8, 0.3)=0.3
hence ***$\mu_{zero}(i) = 0.3$***
if θ is Zero and dθ/dt is +ve small then i is –ve small
min(0.8, 0.7)=0.7
hence ***$\mu_{-ve\text{-}small}(i) = 0.7$***
if θ is +ve small and dθ/dt is Zero then i is –ve small
min(0.4, 0.3)=0.3
hence ***μ-ve-small(i)=0.3***
if θ +ve small and dθ/dt is +ve small then i is -ve medium
min(0.4, 0.7)=0.4
hence ***$\mu_{-ve\text{-}medium}(i) = 0.4$***

Finding *i*



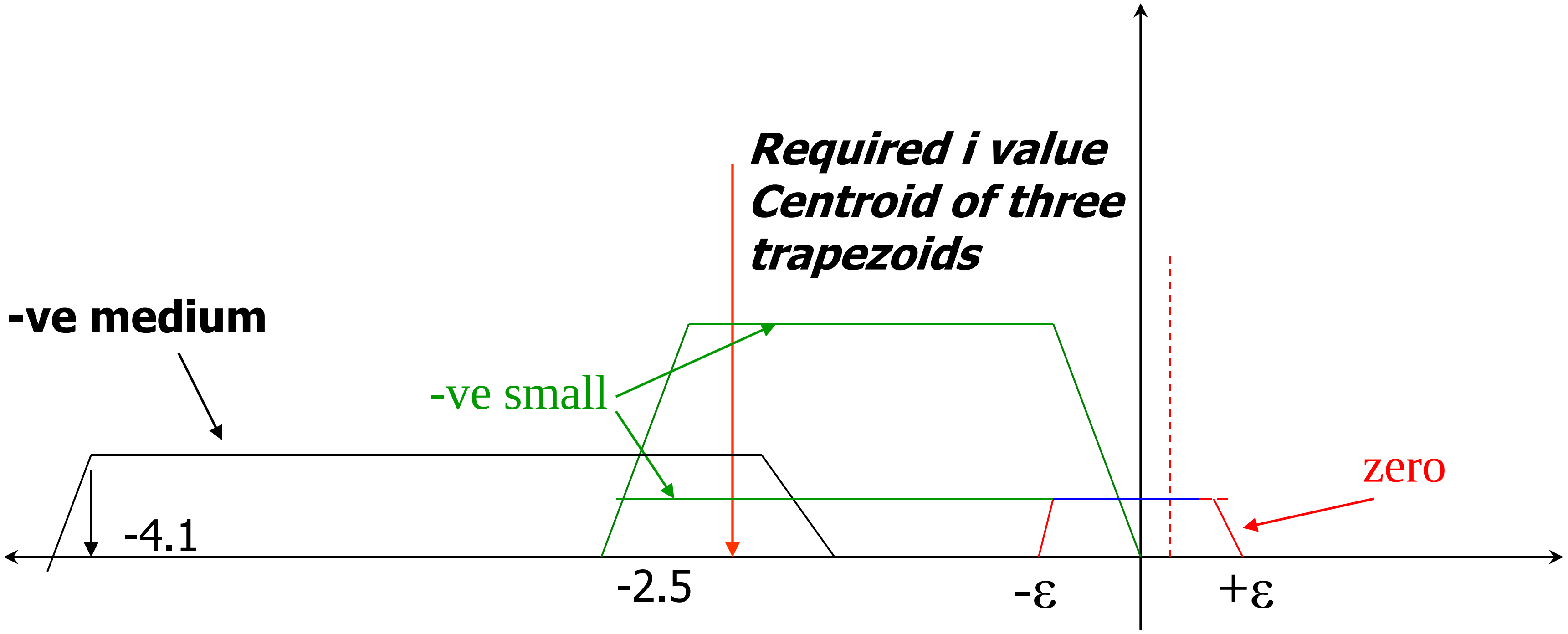
Possible candidates:

i=0.5 and -0.5 from the "zero" profile and μ=0.3

i=-0.1 and -2.5 from the "-ve-small" profile and μ=0.3

i=-1.7 and -4.1 from the "-ve-small" profile and μ=0.3

Defuzzification: Finding *i* by the *centroid* method



Possible candidates:

i is the x-coord of the centroid of the areas given by the blue trapezium, the green trapeziums and the black trapezium