Fuzzy Rule Based Systems

Natural Language

- Natural language is often used for human communication,
- It comprises fundamental atomic terms that are vague, imprecise and ambiguous.
- These fundamental terms are referred to as "*atoms*".
- A collection of *atoms* will form "*molecules*" or phrases or composite terms of natural language.
- Example *atoms*
  - *slow*, *medium*, *young*, *beautiful*
- *Example of composite terms*
  - *very slow horse*, *at least about 5*, *fairly beautiful painting*, *very dangerous act of terrorism.*

- Atomic terms can be defined to exist as elements on a universe X and another universe say Z represent its interpretations, of natural language terms.
- The interpretation of these terms is vague and is best represented as fuzzy sets.
- The strength of these interpretation is defined by a membership value.
- Thus, Natural language can be expressed as a mapping from a set of atomic terms to a corresponding set of interpretations.
  - For an atomic term "α" in the universe of natural language X, there exists a fuzzy set $\underline{A}$ in the universe of interpretations, Y.
  - The relationship between "α" to $\underline{A}$ is a mapping $\underline{M}$, expressed as $\mu_{\underline{M}}(\alpha, y) = \mu_{\underline{A}}(y)$

Linguistic Hedges

- Adjectives/adverbs used to modify the atomic terms and their membership values are referred to as "linguistic hedges".
- Examples
  - *very*, *low*, *slight*, *almost*, *approximately*
- For an atomic term, "α" represented as the following are some of the hedges.

$$\alpha = \int_Y \frac{\mu_\alpha(y)}{y}$$

$$\text{"}Very\text{"} \quad \alpha = \alpha^2 = \int_Y \frac{|\mu_\alpha(y)|^2}{y} \qquad \text{"}Very, very\text{"} \quad \alpha = \alpha^4 = \int_Y \frac{|\mu_\alpha(y)|^4}{y}$$

$$\text{"}Plus\text{"} \quad \alpha = \alpha^{1.25} \qquad \text{"}Slightly\text{"} \quad \alpha = \sqrt{\alpha} \qquad \text{"}Minus\text{"} \quad \alpha = \alpha^{0.75}$$

- For any modifier h(a),
  - *If h(a) < a, then the modifier is* **strong**.
  - *If h(a) > a, then the modifier is* **weak**.
  - *If h(a) = a, then it is identity* **modifier**.
- A strong modifier strengthens a fuzzy predicate to which it applied but reduces the truth value of the associated proposition.
- Similarly a weak modifier, weakens a predicate and hence increase the truth value of the associated proposition.
  - ✓ E.g Let old(80)= 0.90,
    - ✓Very old = $0.9^2 = 0.81$
    - ✓Fairly old = $\sqrt{a}$ = 0.94

Fuzzy Operators

- Two fundamental fuzzy operators are in existence.
  - **Concentration** - which tends to concentrate the elements of a fuzzy set by reducing the degree of membership of all the elements that are "partly" in the set.
  - The less the membership, the more it is concentrated.
  - e,g. $\alpha^{2}[0.9 \rightarrow 0.81; 0.1 \rightarrow 0.01]$

- **Dilation** - which stretches a fuzzy set by increasing the degree of membership of all the elements that are "partly" in the set.
  - e,g.

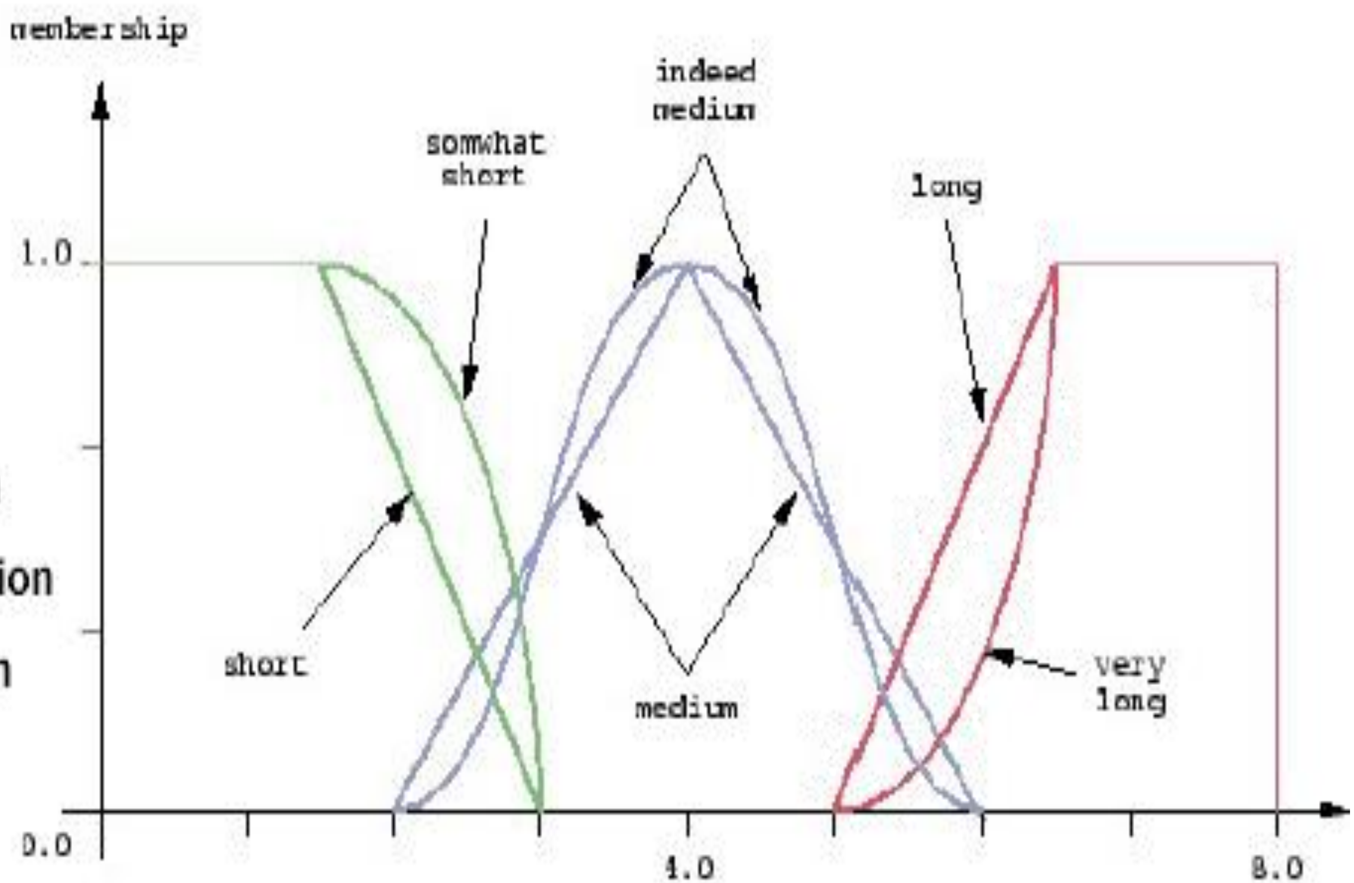
$$\sqrt{\alpha}[0.9 \rightarrow 0.95; 0.1 \rightarrow 0.32]$$

- **Intensification** - It is a combination of concentration and dilation. It increases the degree of membership of those elements in a fuzzy set with original membership values greater than 0.5, and decreases the degree of membership of those elements in the set with original membership less than 0.5.
- It increases the contrast between the elements of the set that have more/less than half membership.

- Mathematically,

$$\text{"Intensity"} \quad \alpha = \begin{cases} 2\mu_\alpha{}^2(y) & for \quad 0 \le \mu_\alpha(y) \le 0.5 \\ 1-2[1-\mu_\alpha(y)]^2 & for \quad 0.5 \le \mu_\alpha(y) \le 1 \end{cases}$$

- Somewhat: Dilation
- Indeed: Intensification
- Very: Concentration



- **Fuzzification** - It is opposite of **intensification**.

  Mathematically,

$$\mu_{FUZ(\underline{A})}(x)=\begin{cases}[\mu_{\underline{A}}(x)/2]^{1/2} & for \quad 0\le \mu_{\alpha}(y)\le 0.5\\ 1-2[(1-\mu_{\underline{A}}(x))/2]^{1/2} & for \quad 0.5\le \mu_{\alpha}(y)\le 1\end{cases}$$

- These operators are extensively used in several image processing, pattern recognition and artificial intelligent problems.

- ***A linguistic expression*** can be formed by using logical connectives between linguistic hedges.

- Example:
  - “*not costly*” *And* “*very high building*”

- Order of precedence:

$$Not \rightarrow And \rightarrow Or$$

Example

- Let the universe of integers, *Y*={1, 2, 3, 4, 5} and the following linguistic terms

$$"Small" = \{\frac{1}{1} + \frac{0.8}{2} + \frac{0.6}{3} + \frac{0.4}{4} + \frac{0.2}{5}\} \quad "Large" = \{\frac{0.2}{1} + \frac{0.4}{2} + \frac{0.6}{3} + \frac{0.8}{4} + \frac{1}{5}\}$$

Construct the *composite* term "*not very small and not very*, *very large*"

Solution

$$not \quad very \quad small = 1 - very \quad small = 1 - small^2$$

$$= \{\frac{0}{1} + \frac{0.36}{2} + \frac{0.64}{3} + \frac{0.84}{4} + \frac{0.96}{5}\}$$

$$not \quad very \quad very \quad large = 1 - very \quad very \quad large = 1 - large^4$$

$$= \{\frac{1}{1} + \frac{1}{2} + \frac{0.9}{3} + \frac{0.6}{4}\}$$

$\therefore$ The linguistic expression "*not very small And not very very large*"

$$= \{\frac{0}{1} + \frac{0.36}{2} + \frac{0.64}{3} + \frac{0.84}{4} + \frac{0.96}{5}\} \cap \{\frac{1}{1} + \frac{1}{2} + \frac{0.9}{3} + \frac{0.6}{4}\}$$

$$= \{\frac{0.36}{2} + \frac{0.64}{3} + \frac{0.6}{4}\}$$

Rule Based Systems

- The most common way to represent human knowledge is to form it into natural language expressions of the type

  IF premise (antecedent) THEN conclusion (consequent)

- It expresses an inference such that
  - ➢Given a fact (premise, hypothesis, antecedent) one can derive another fact called a conclusion (consequent).
- This form of knowledge representation is referred to as *shallow knowledge*.
  - ➢It uses linguistic variables as its antecedent and consequent.

Canonical Rule Forms

- Three canonical forms of rule based statements
  - ➢Assignment statements: Restricts the value of a variable.
    - ✓*x=large*; *color=yellow*
    - ✓*x=x*
  - ➢Conditional statements: Restricts the value of a variable based on a condition.
    - ✓*IF x is large THEN color=yellow*
  - ➢Unconditional statements: Restricts the value of a variable based on any condition.
    - ✓*x is large*
    - ✓*Stop*

Compound Rules

- Several linguistic connectives can be used to combine rules.
- These combined rules are referred to a ***compound rules***.
- Example:

  IF price is high THEN buying capacity (*bc*) *is very low.*

  IF price is high And quality is good THEN bc is high.

  IF price is low and quality is good THEN bc is medium.

  ELSE

  IF price is medium and quality is bad THEN bc is low.

Decomposition of Compound Rules

- The basic properties and operations defined for fuzzy sets can be used to decompose any compound rule into a number of simple canonical rules.

❖ **Multiple Conjunctive Antecedents**

IF x is $\underline{A}^1$ *And* $\underline{A}^2$ *And …. And* $\underline{A}^L$ *THEN y is* $\underline{B}^S$

Assuming a new fuzzy subset $\underline{A}^S$ as

$$\underline{A}^S = \underline{A}^1 \cap \underline{A}^2 \cap \ldots\ldots \cap \underline{A}^L$$

Thus t he compound rule can be rewritten as

IF $\underline{A}^S$ *THEN* $\underline{B}^S$

❖ Multiple Disjunctive Antecedents

IF x is $\underline{A}^1$ *OR* $\underline{A}^2$ *OR …. OR* $\underline{A}^L$ *THEN y is* $\underline{B}^S$

Assuming a new fuzzy set $\underline{A}^S$ as

$$\underline{A}^S = \underline{A}^1 \cup \underline{A}^2 \cup \ldots . \cup \underline{A}^L$$

Thus the compound rule can be rewritten as

IF $\underline{A}^S$ *THEN* $\underline{B}^S$

❖ Conditional ELSE statements

IF $\underline{A}^1$ *THEN* ($\underline{B}^1$ *ELSE* $\underline{B}^2$)

This can be decomposed into

IF $\underline{A}^1$ *THEN* $\underline{B}^1$ *OR IF NOT* $\underline{A}^1$ *THEN* $\underline{B}^2$

❖ Conditional UNLESS statements

$$IF\ \underline{A}^1\ (THEN\ \underline{B}^1)\mathrm{UNLESS}\ \underline{A}^2)$$

This can be decomposed into

$$IF\ \underline{A}^1\ THEN\ \underline{B}^1\ OR\ IF\ \underline{A}^2\ THEN\ NOT\ \underline{B}^1$$

$$IF\ \underline{A}^1\ THEN\ \ (\underline{B}^1\ ELSE\ IF\ \underline{A}^2\ \ THEN\ (\underline{B}^2))$$

This may be further simplified as

$$\mathrm{IF}\ \underline{A}^1\ THEN\ \underline{B}^1\ OR\ IF\ NOT\ \underline{A}^1\ AND\ \underline{A}^2\ THEN\ \underline{B}^2$$

❖ Nested IF-THEN Rules

$$IF\ \underline{A}^1\ THEN\ (IF\ \underline{A}^2\ THEN\ (\underline{B}^1))$$

It can be put in the form

$$IF\ \underline{A}^1\ AND\ \underline{A}^2\ THEN\ \underline{B}^1$$

Likelihood and Truth Qualification

- *Atomic* and *composite* terms may also be modified by **likelihood linguistic variables**, such as
  - “*likely*”, “*very likely*”, “*highly likely*” or “*unlikely*”.

- They may also be modified by **truth qualification** statements such as
  - “*true*”, “*fairly true*”, “*very true*”, “*false*”, “*fairly false*” or “*very false*”.

- These likelihood labels are based on the notions of probability.

Example

- Let the universe of discourse is given by

  U={0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1.0}

  where the elements represent probabilities.

  Compute the meaning of the linguistic variable x, where x="*highly unlikely*"

  with

$$"likely" = \{\frac{1}{1} + \frac{1}{0.9} + \frac{1}{0.8} + \frac{0.8}{0.7} + \frac{0.6}{0.6} + \frac{0.5}{0.5} + \frac{0.3}{0.4} + \frac{0.2}{0.3}\}$$

  and "*highly*" and "*unlikely*" defined as

  – "*highly*" = "*minus very very*" = (*very very*)$^{0.75}$

  – "*unlikely*" = "*not likely*"

Solution

$$"unlikely" = "1 - likely" = \{\frac{1}{0} + \frac{1}{0.1} + \frac{1}{0.2} + \frac{0.8}{0.3} + \frac{0.6}{0.4} + \frac{0.5}{0.5} + \frac{0.7}{0.6} + \frac{0.8}{0.7}\}$$

$$"very \quad very \quad unlikely" = "(unlikely)"^4$$

$$= \{\frac{1}{0} + \frac{1}{0.1} + \frac{1}{0.2} + \frac{0.4}{0.3} + \frac{0.2}{0.4}\}$$

$$"highly \quad unlikely" = \quad "\min us \quad very \quad very \quad unlikely"$$

$$= \left(\frac{1}{0} + \frac{1}{0.1} + \frac{1}{0.2} + \frac{0.4}{0.3} + \frac{0.2}{0.4}\right)^{0.75} = \left(\frac{1}{0} + \frac{1}{0.1} + \frac{1}{0.2} + \frac{0.5}{0.3} + \frac{0.3}{0.4}\right)$$

Aggregation of Fuzzy Rules

- Mostly a rule based system involves more than one rule.

- The process of obtaining overall consequent from the individual consequents is known as aggregation of rules.

❖ Conjunctive system of rules

- These rules must be jointly satisfied and are connected by "*and*" connectives.
- The aggregated output (consequent), y, is found by the fuzzy intersection of all the individual rule consequents, y^i, where i = 1, 2, …., r, as

$$y = y^1 \quad and \quad y^2 \quad and \quad .... \quad and \quad y^r$$

or

$$y = y^1 \cap y^2 \cap .... \cap y^r$$

In terms of membership

$$\mu_y(y) = \min(\mu_{y^1}(y), \mu_{y^2}(y), ..., \mu_{yr}(y)) \quad for \quad y \in Y$$

❖ Disjunctive system of rules

- At least one rule is required to be satisfied and these rules are connected by "*or*" connectives.
- The aggregated output (consequent), y, is found by the fuzzy union of all the individual rule consequents, y^i, where i = 1, 2, …., r, as

$$y = y^1 \quad or \quad y^2 \quad or \quad .... \quad or \quad y^r$$

or

$$y = y^1 \cup y^2 \cup .... \cup y^r$$

In terms of membership

$$\mu_y(y) = \max(\mu_{y^1}(y), \mu_{y^2}(y), ..., \mu_{yr}(y)) \quad for \quad y \in Y$$

Graphical Methods of Inference

- Inferences from a family of rules can also be achieved by graphical techniques due to Mamdani.
- The individual antecedents are represented as membership functions in the graphical view.
- So are the consequents.
- The inference is obtained using Mamdani’s implication method of inference.
- The inputs applied to the system, in conjunction with the membership functions of the antecedents and consequents determine the inference.

Example

- Consider a system guided by a disjunctive set of rules manifested by *r* linguistics of the form

$$If \quad x_1 \quad is \quad \underline{A}_1^k \quad and \quad x_2 \quad is \quad \underline{A}_2^k \quad THEN \quad y^k \quad is \quad \underline{B}^k \quad for \quad k = 1, 2, \ldots, r$$

The inputs to the system are crisp delta functions, x_1 and x_2, whose memberships are given by

$$\mu(x_1) = \delta(x_1 - input(i)) = \begin{cases} 1, & x_1 = input(i) \\ 0, & otherwise \end{cases}$$

$$\mu(x_2) = \delta(x_2 - input(i)) = \begin{cases} 1, & x_2 = input(j) \\ 0, & otherwise \end{cases}$$

Find the crisp output using Mamdani approach.

Solution

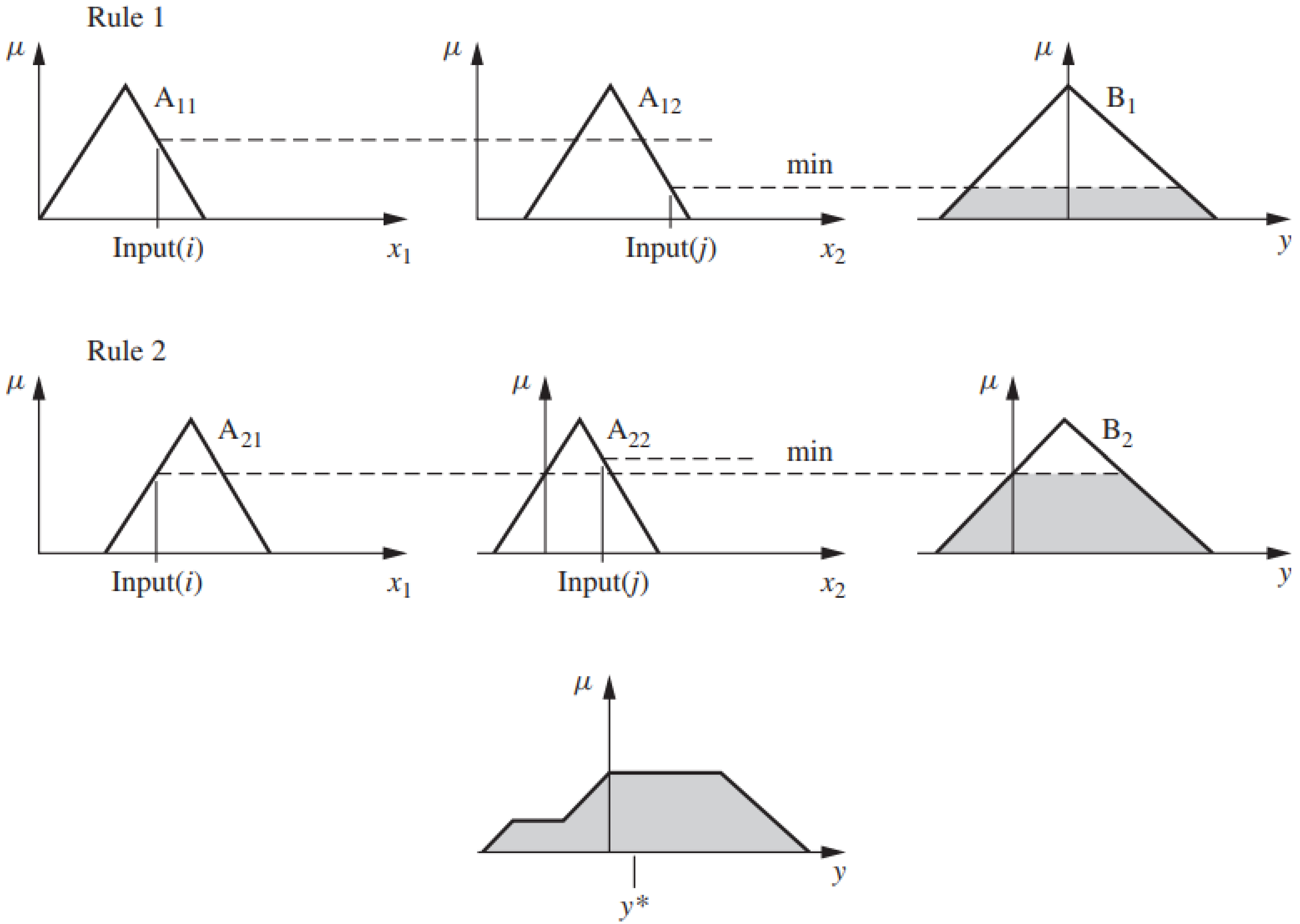
Based on Mamdani's implication method of inference for a set of disjunctive rules, the aggregated output for the *r* rules will be given by

$$\mu_{\underset{\sim}{B}^k}(y) = \max_k [\min[\mu_{\underset{\sim}{A}_1^k}(input(i)), \mu_{\underset{\sim}{A}_2^k}(input(j))]] \quad k = 1,2,\ldots.,r$$

Basically, it implies that the following steps are to be taken.

- Find out the individual minimum of intersection of all the individual antecedents with the inputs.
- Find out the maximum of all these *k* minima.
- Defuzzify the fuzzy output to a crisp value.

Solution


Rule 1
μ
A11
Input(i)
x1
μ
A12
Input(j)
x2
min
μ
B1
y
Rule 2
μ
A21
Input(i)
x1
μ
A22
Input(j)
x2
min
μ
B2
y
μ
y
y*