Fuzzy Systems and Learning

Ning Xiong
Mädardalen University

Agenda

- Motivation
- Basic concepts of fuzzy logic and fuzzy set theory
- Fuzzy rule-based system and reasoning

- Types of fuzzy systems
- Fuzzy system optimization and learning

Motivation

Motivation

- **Human centric intelligent systems** is a hot trend in current research, entailing more computer-human interaction.

- Humans are more comfortable to express opinions in vague terms in natural language; they also prefer to receive information in linguistic and comprehensible descriptions.

- Complex, ill-defined processes are difficult for description and analysis by exact mathematical techniques

- Precision is not always the goal to strive for. Tolerance of imprecision can lead to tractability, robustness, and short computation/reaction time.

Significance of Fuzzy Theory

- We need Fuzzy logic and fuzzy theory as powerful means of manimulation of vague and uncertaion information in order to create intelligent systems that are much closer in spirit to human thinking and reasoning

- Fuzzy Logic is the theory of fuzzy sets, which accomodates vagueness and uncertainty.

Basic Concepts

Fuzzy vs. Crisp Sets

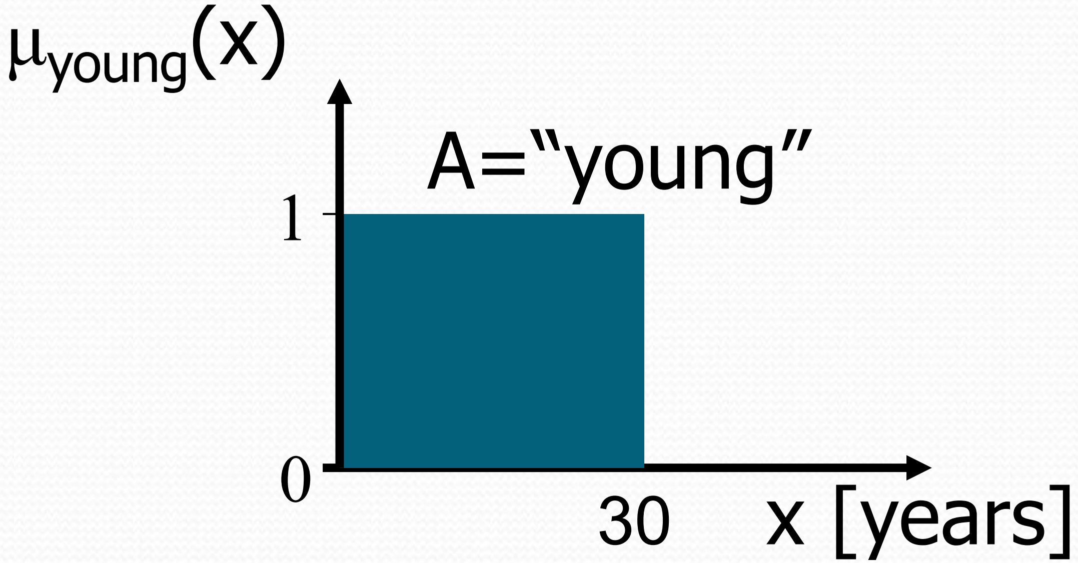
- Crisp set indicates whether an element belongs to it
- Fuzzy set indicates how much (to which extent) an element belongs to it.

Classical (Crisp) Set

young = { x | x $\leq$30 }

Membership function:

$$\mu_{young}(x) = \begin{cases} 1 : x \leq 30 \\ 0 : x > 30 \end{cases}$$



Fuzzy Set

Definition :

A fuzzy set F in a universe of discourse U is characterized by membership function μ_F, which takes values in the interval [0,1], i.e., μ_F: U$\rightarrow$ [0,1]

Example:

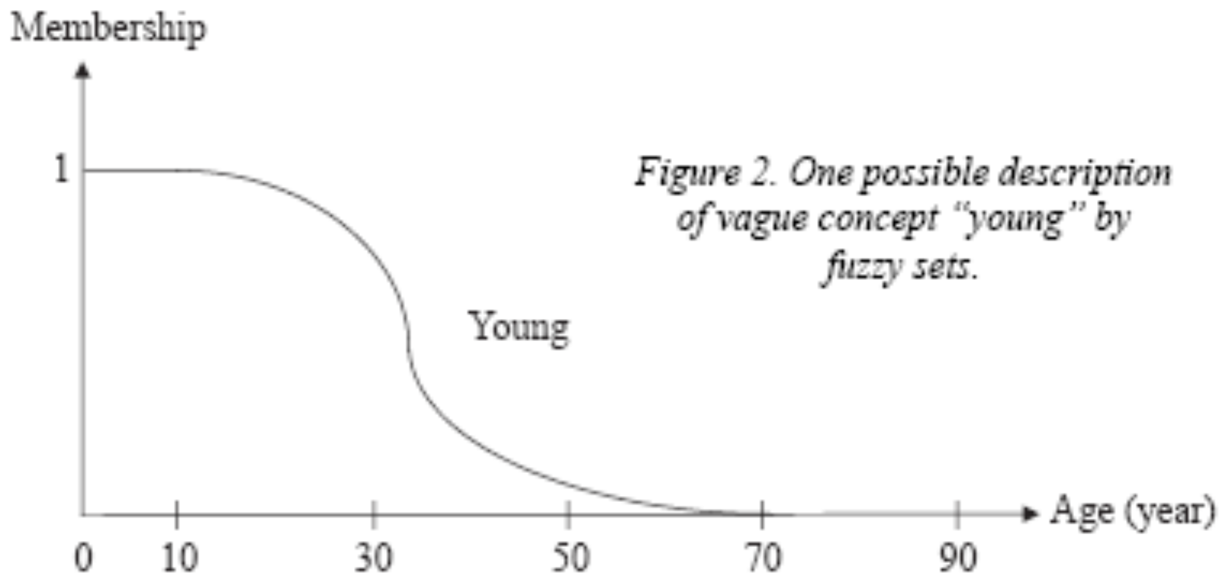


Figure 2. One possible description of vague concept "young" by fuzzy sets.

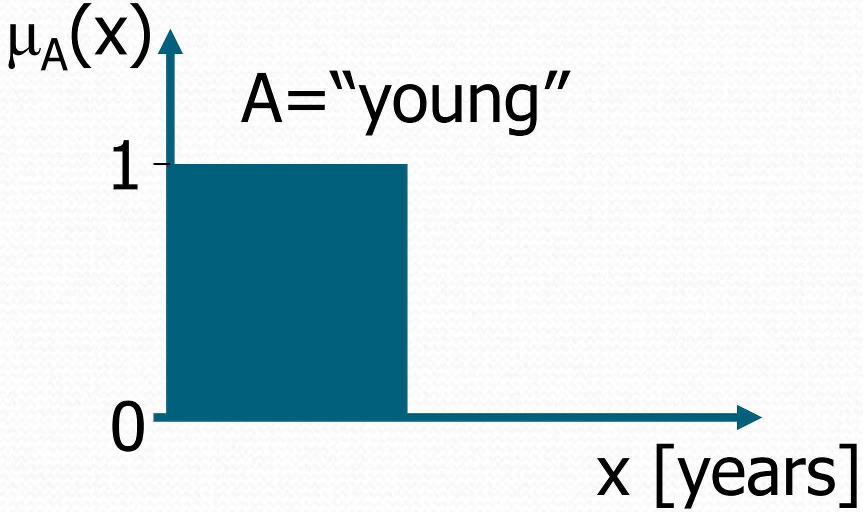
If U contains finite number of elements, fuzzy set F can be denoted by: $F=\{\mu_F(u_1)/u_1, \mu_F(u_2)/u_2, \ldots, \mu_F(u_n)/u_n\}$

Fuzzy versus Classical Logic

Classical Logic

Element x belongs to set A or it does not:

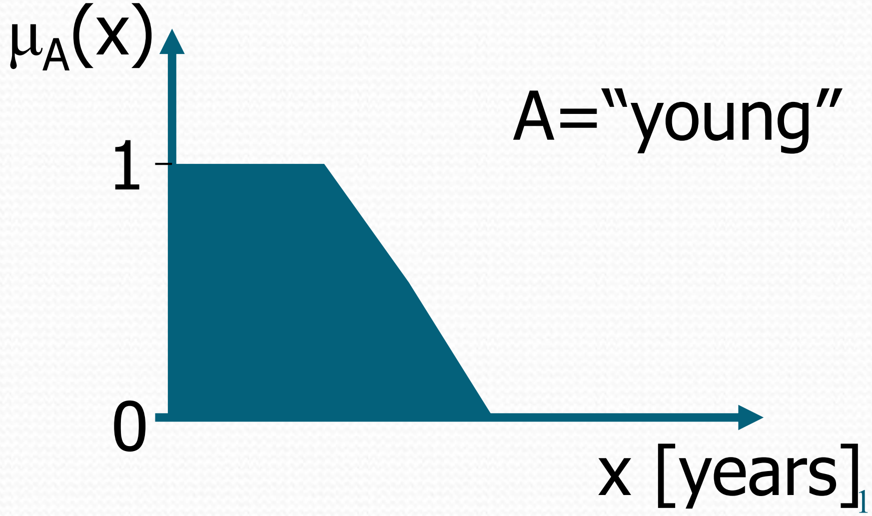
$$\mu_A(x) \in \{0,1\}$$



Fuzzy Logic

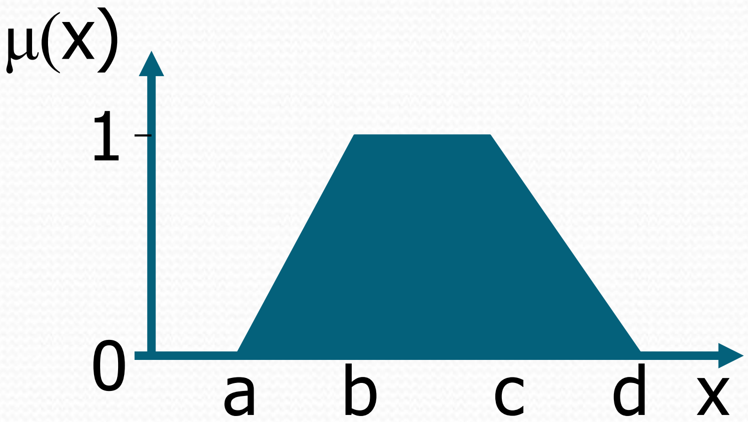
Element x belongs to set A with a certain *degree of membership*:

$$\mu_A(x \in 0,1]$$

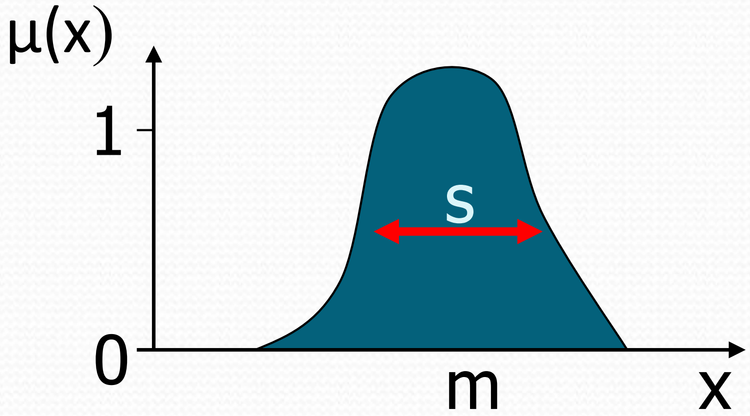


Types of Membership Functions

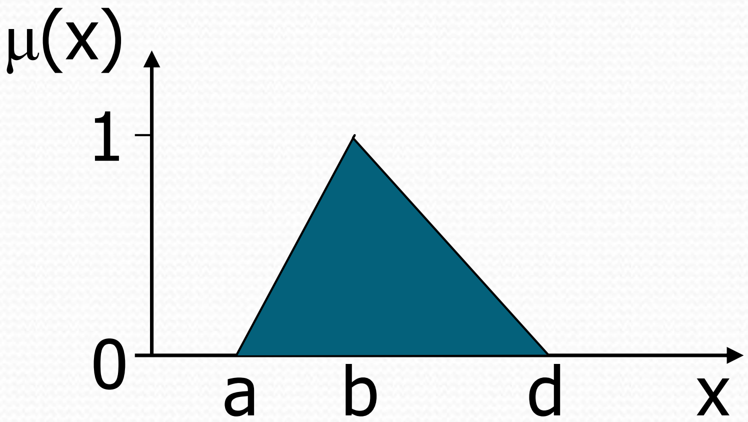
Trapezoid: <a,b,c,d>



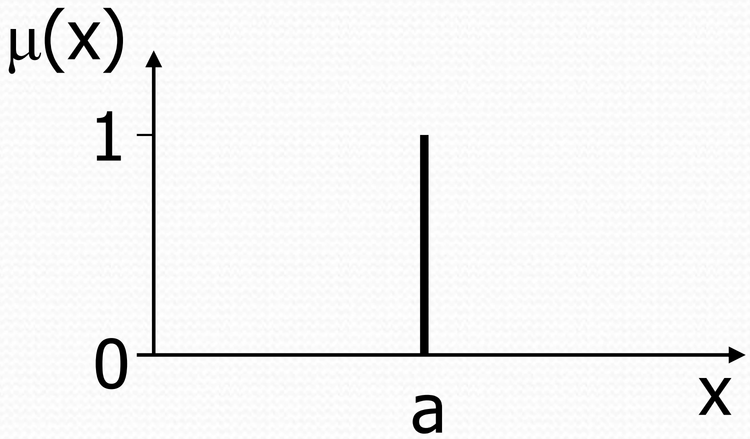
Gaussian: N(m,s)



Triangular: <a,b,b,d>



Singleton: (a,1)

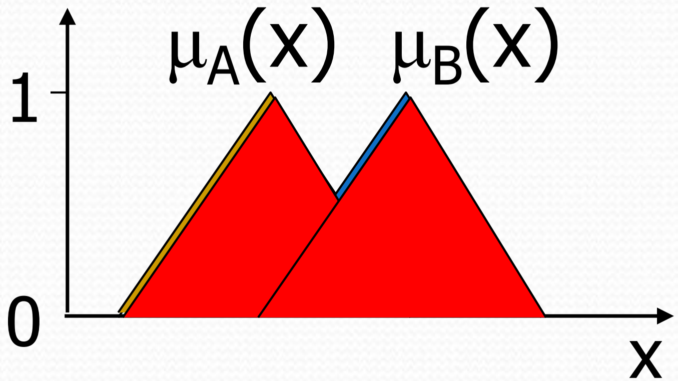


Operators on Fuzzy Sets

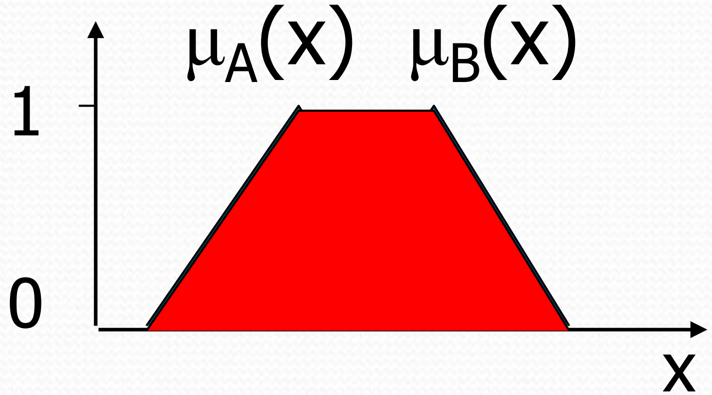
Union (of crisp sets): whether an element belongs to either set

Union (of fuzzy sets): How much an element belongs to either set

$\mu_{A \vee B}(x) = \max\{\mu_A(x), \mu_B(x)\}$



$\mu_{A \vee B}(x) = \min\{1, \mu_A(x) + \mu_B(x)\}$

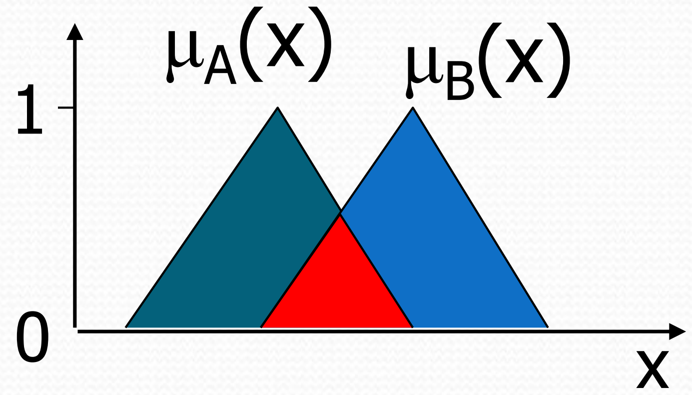


Operators on Fuzzy Sets

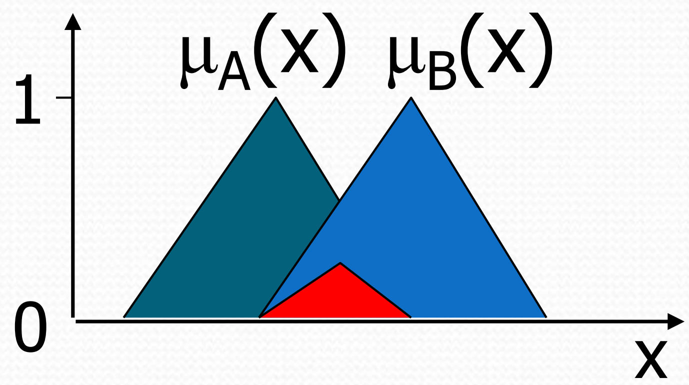
Intersection (of crisp sets): whether an element belongs to both sets

Intersection (of fuzzy sets): How much an element belongs to both sets

$$\mu_{A \wedge B}(x) = \min\{\mu_A(x), \mu_B(x)\}$$



$$\mu_{A \wedge B}(x) = \mu_A(x) \bullet \mu_B(x)$$

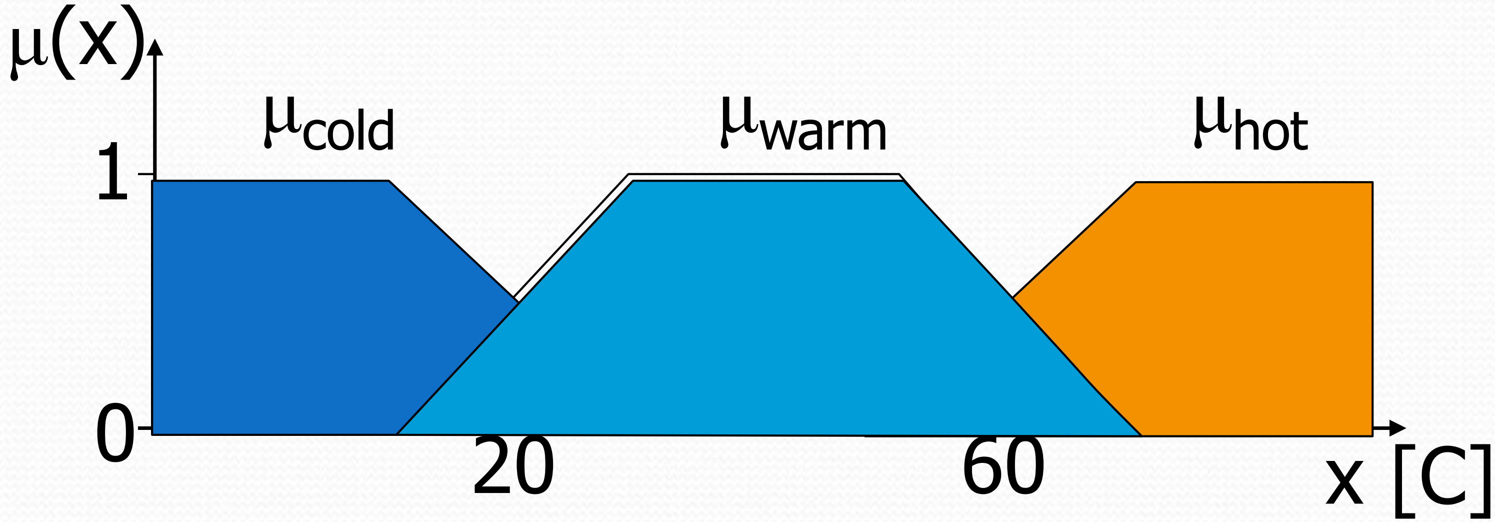


Linguistic Variables

- A linguistic variable has a set of linguistic terms/values.

- Each linguistic value corresponds to a fuzzy set and explained with its membership function

linguistic variable : temperature
linguistics terms (fuzzy sets) : { cold, warm, hot }



Fuzzy Rules

- A fuzzy rule is a linguistic expression of **causal dependencies between linguistic variables** in form of if-then statements
- General form:
  If <antecedent> then <consequence>
- Example:
  if temperature is cold and oil price is cheap
  then heating is high

linguistic variables

linguistic values/terms (fuzzy sets)

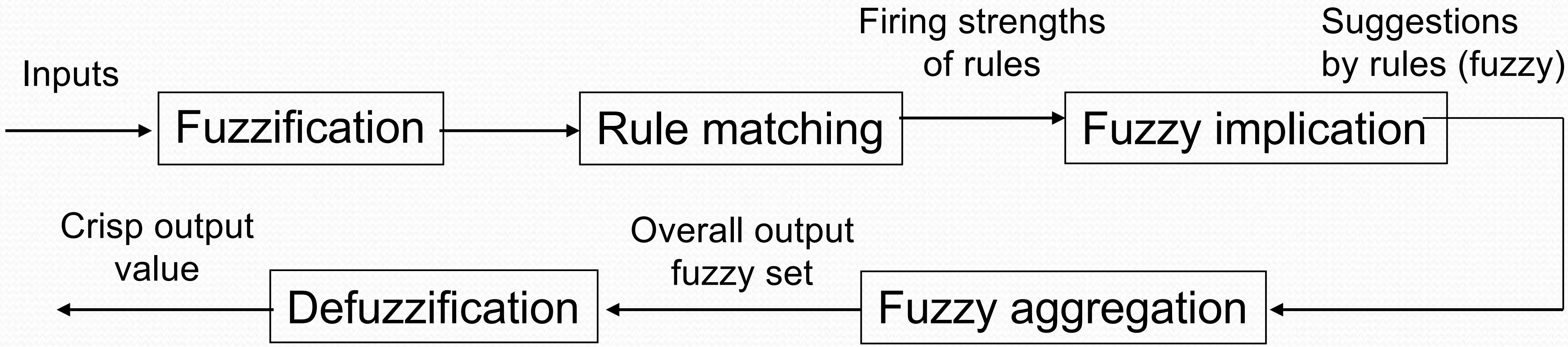
Fuzzy Rule Base

Heating Oil price:	Temperature : cold	warm	hot
cheap	high	high	medium
normal	high	medium	low
expensive	medium	low	low

A fuzzy rule base is a collection of fuzzy if-then rules

Fuzzy rule-based system and reasoning

Fuzzy Decision Making Procedure

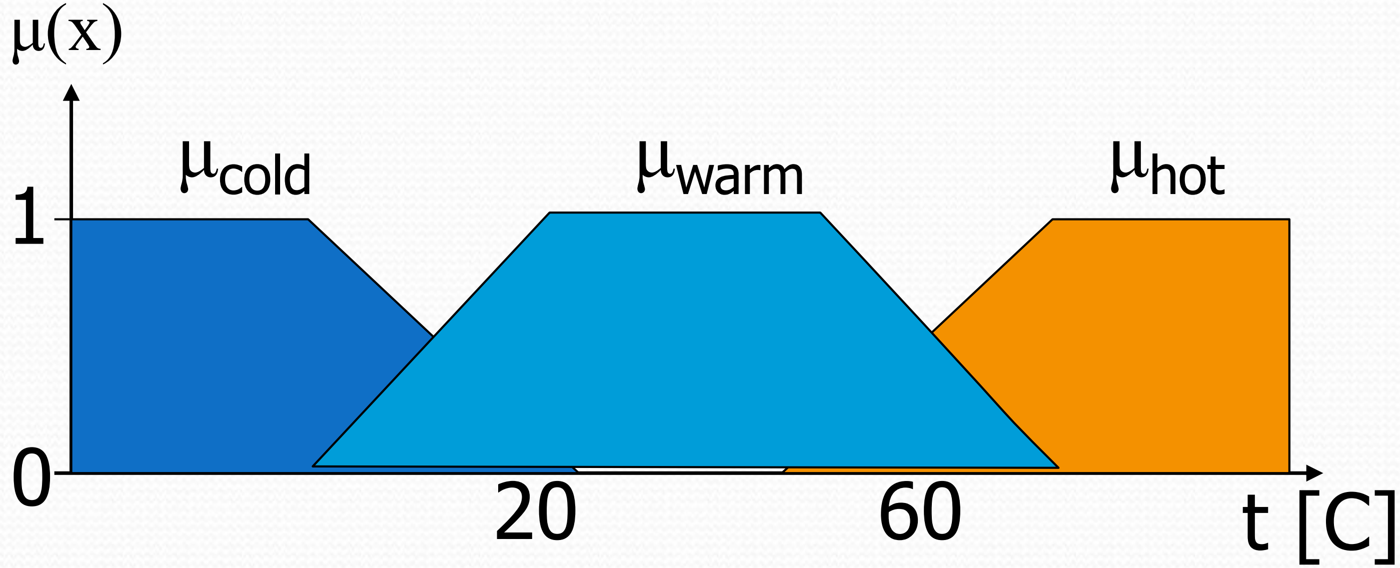


Fuzzification: compute the membership degrees for each input variable with respect to its linguistic terms.
Rule Matching: calculating firing strengths (degrees of satisfaction) of the individual rules
Fuzzy Implication: determine the suggestions of rules according to firing strengths and rule conclusions
Fuzzy Aggregation: combine suggestions from individual rules into an overall output fuzzy set
Defuzzification: determine a crisp value from the output membership function as the final result or solution.

Fuzzification

Determine the degree of membership for each fuzzy set of an input variable :

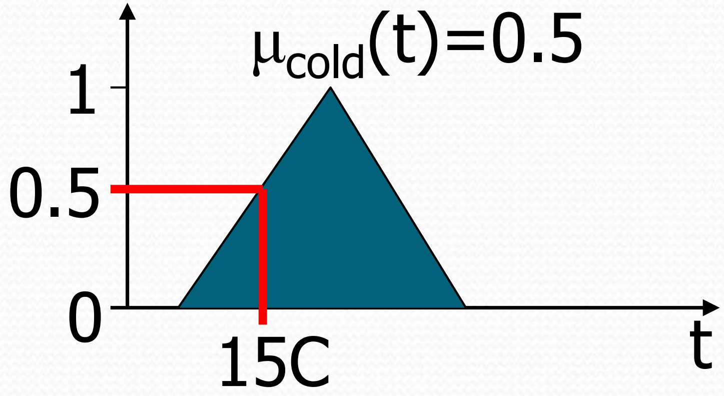
temperature : t=18 C

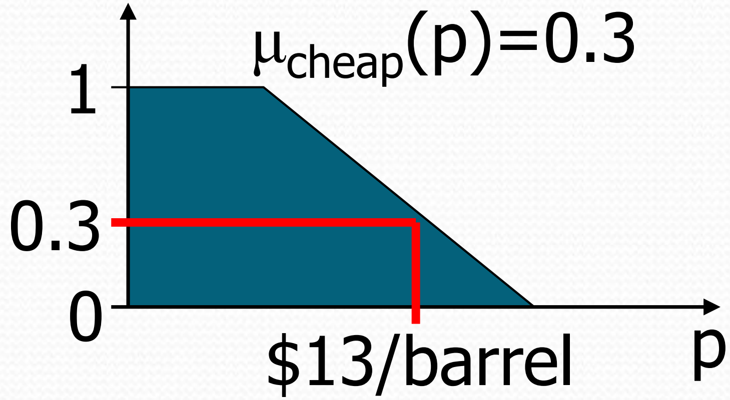


$\mu_{cold}(t)=0.2$, $\mu_{warm}(t)=0.6$, $\mu_{hot}(t)=0$

Rule Matching

Calculate the firing strength of every rule by combining the individual membership degrees for all terms involved in the condition part of the rule through fuzzy AND: min-operator



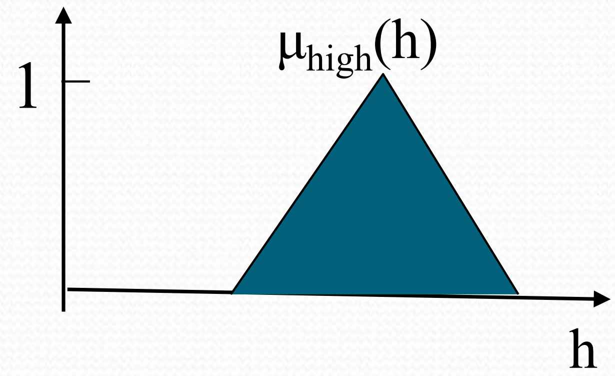


R_i if temperature is cold ... **and** oil is cheap ...

$t_i = \min\{\mu_{cold}(t), \mu_{cheap}(p)\} = \min\{0.5,0.3\} = 0.3$

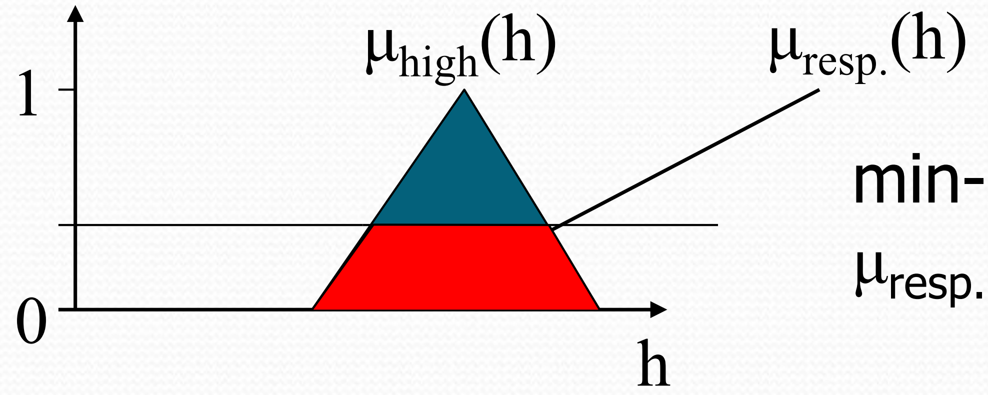
Fuzzy Implication

Implication step: For each rule apply the firing strength to modify its consequent fuzzy set, resulting in a new output fuzzy set as the suggestion of the rule.



Example: If ... then heating is high

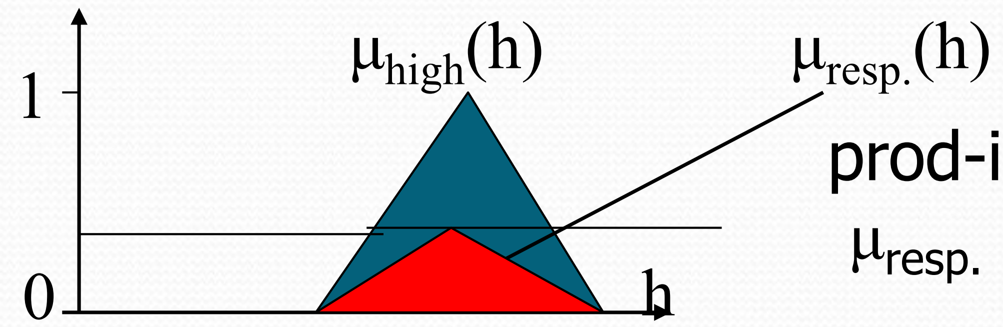
... t=0.3



min-implication:

$\mu_{resp.} = \min\{t , \mu_{high} \}$

... t=0.3



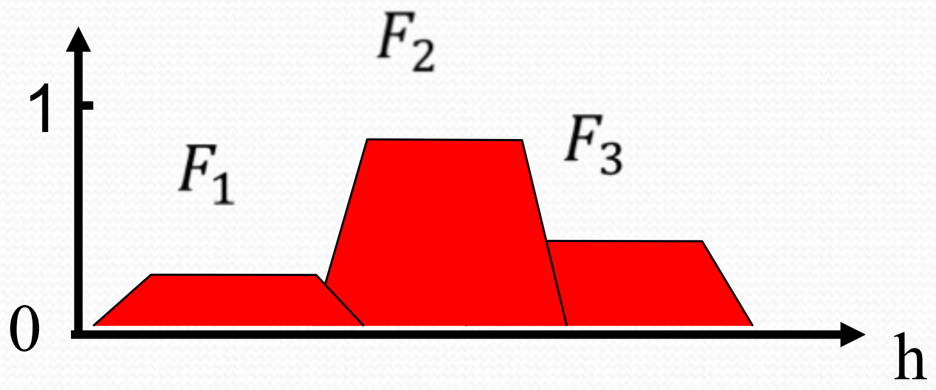
prod-implication:

$\mu_{resp.} = t \bullet \mu_{high}$

Fuzzy Aggregation

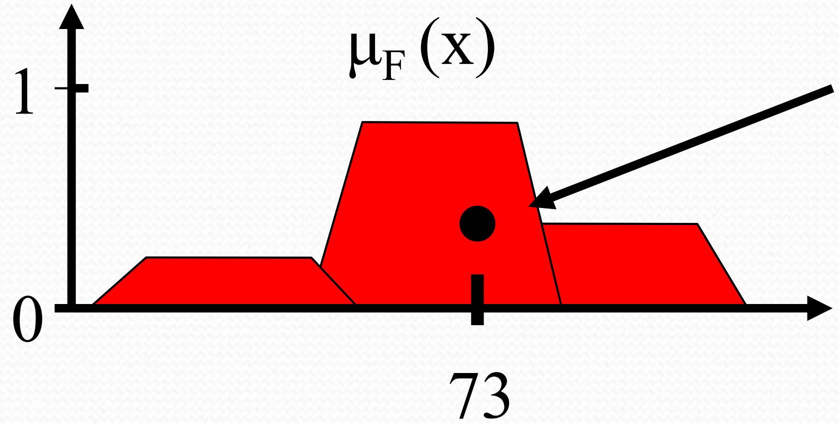
- Aggregation: combine suggestions of individual rules to yield an overall output fuzzy set using the max-operator for union. The overall fuzzy set is an aggregated effect from all rules.
- Suppose F_1, F_2, and F_n are suggestions from *n* fuzzy rules, the overall output fuzzy set is written as:

$$F = F_1 \cup F_2 \cup \cdots \cup F_n$$



Defuzzification

Determine crisp value from output membership function.
Commonly used is the "Center of Gravity" method:

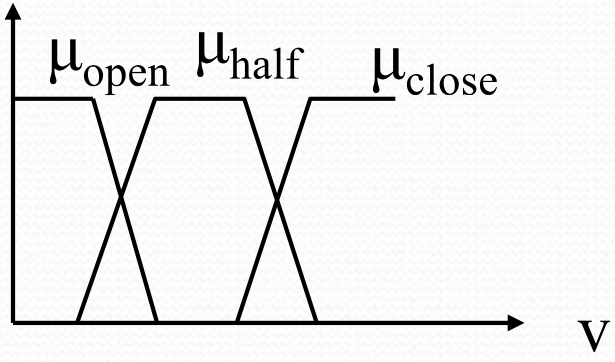


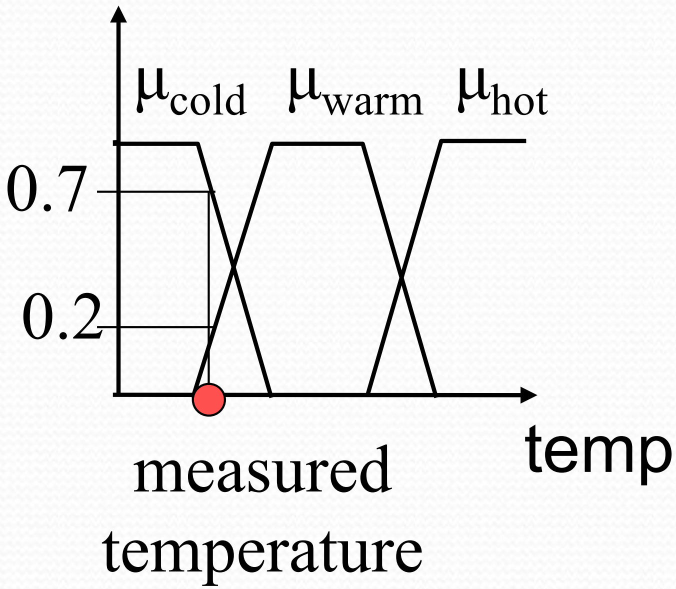
$$COG = \frac{\int_a^b \mu_F(x)\,x\,dx}{\int_a^b \mu_F(x)\,dx}$$

$$COG = \frac{\sum_{x=a}^{b} \mu_F(x)\,x}{\sum_{x=a}^{b} \mu_F(x)}$$

An Example of Fuzzy Reasoning

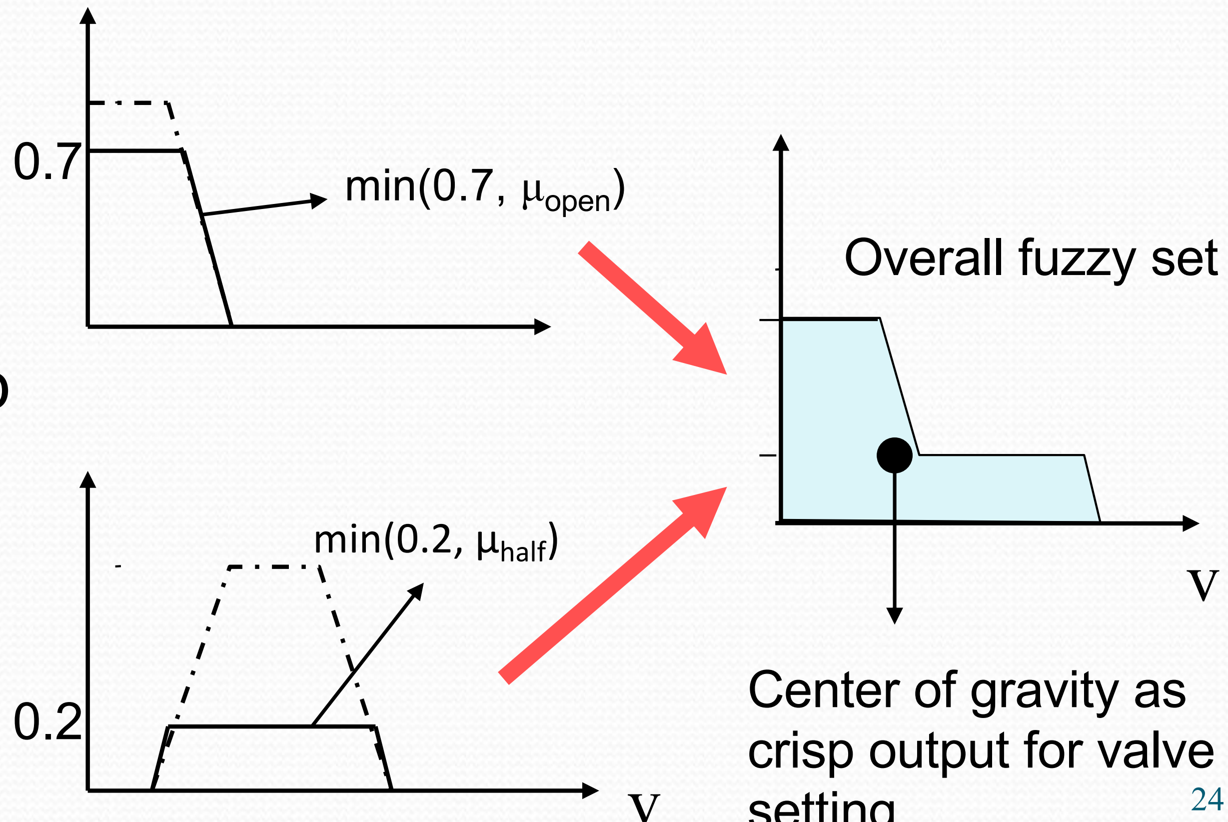
R1: if temp is cold then valve is open
R2: if temp is warm then valve is half
R3: if temp is hot then valve is close





t1=0.7 t2=0.2

t3 =0.0



Types of fuzzy systems

Types of Fuzzy Systems

1. Mamdani Fuzzy Systems: the conclusion of a fuzzy rule is a fuzzy subset.

2. Sugeno Fuzzy Systems: the conclusion of a fuzzy rule is a linear function of inputs.

3. Fuzzy classifier: the conclusion of a rule is a category or single class

Sugeno Fuzzy System

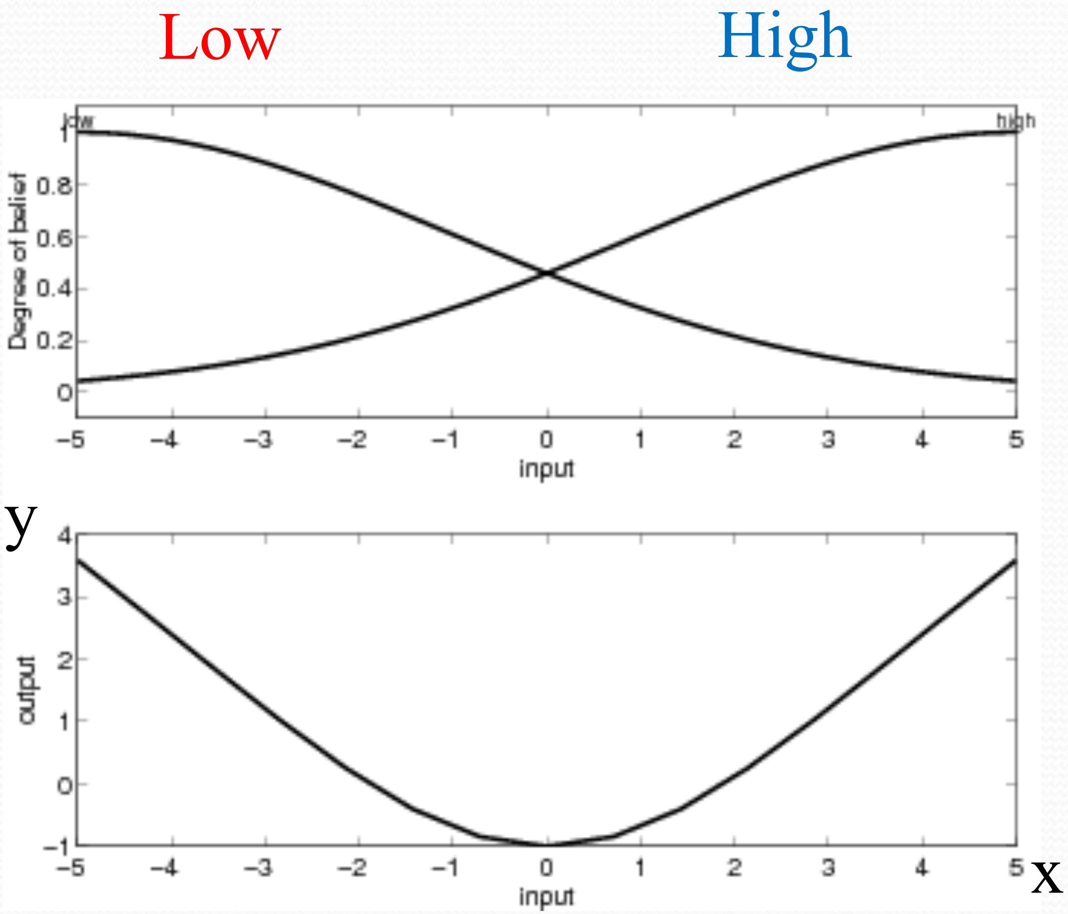
R_i: If x_1 is A_{i1} and x_2 is A_{i2} and x_n is A_{in}
Then $y = a_{i0} + a_{i1}x_1 + a_{i2}x_2 + \ldots.. + a_{in}x_n$

$$y = \frac{t_1 y(R_1) + t_2 y(R_2) + \cdots + t_m y(R_m)}{t_1 + t_2 + \cdots + t_m}$$

- Every fuzzy rule depicts a linear behavior in a fuzzy region as specified by its condition.

- The core idea is to approximate the nonlinear system behavior with a set of smoothly connected linear functions

A two-rule Sugeno model



If x is Low then y=-x-1

If x is High then y=x-1

Fuzzy Classifier

Fuzzy classification rule

R_i: If x_1 is A_{i1} and x_2 is A_{i2} and x_n is A_{in}
Then class B_k

- Key: consider a crisp class as a singleton fuzzy subset defined on the finite set of possible classes, i.e.

$$B_k = \left\{ \frac{0}{B_1}, \ldots\ldots, \frac{1}{B_k}, \ldots\ldots, \frac{0}{B_L} \right\}$$

Then we do fuzzy reasoning in the same manner as before to get an overall output fuzzy set.

- Finally we choose the class which has the highest membership degree for the overall output fuzzy set as the final decision.

Fuzzy system optimization and learning

Fuzzy System Development

Design of a fuzzy system entails:

1) Determining linguistic terrms for all variables
2) Defining all fuzzy set membership functions
3) Specifying conditions and conclusions for all rules;

Traditional design depends on expert knowledge

- Collecting explicit human experiences
- Translation in linguistic fuzzy rules

+ Comprehensibility guarantied

– Implicit feeling hard to express exactly

– Fine tuning of system performance is trial and error and time consuming

Recent Trends: Learning from Data

- Automatic generation (learning) from numerical data to ease knowledge acquisition bottleneck

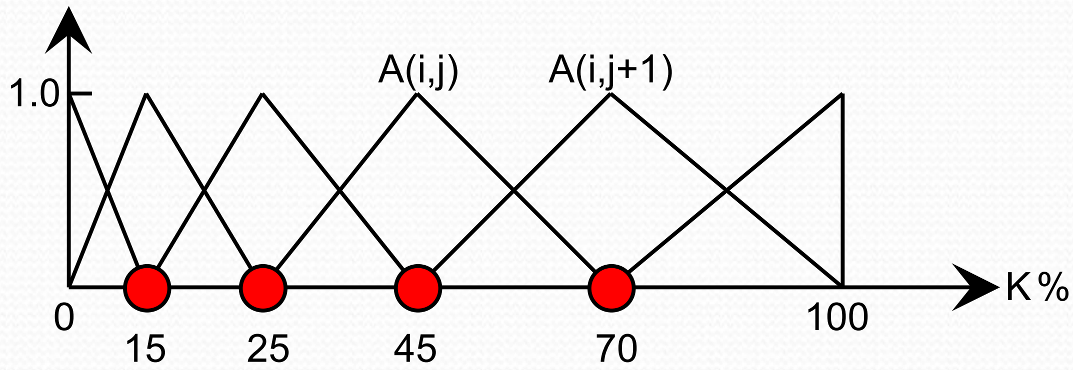
  - Learning of fuzzy set membership functions via GA

  - Wang-Mendel method for data-based learning

  - Extension of Wang-Mendel method in big data

Learning of Membership Functions

• Learning of fuzzy set membership functions under a fully defined rule set



• Usually the membership functions are depicted by critical end points as shown by red circles in the above figure.
• The learning of membership functions thus turns to finding a set of optimal parameter values
• Genetic code → real numbered parameters →complete fuzzy system → Total training error on the training data set
• GA aims to minimize the training error

Wang-Mendel Method

L.X. Wang and J.M. Mendel, ”Generating fuzzy rules by learning from examples”, IEEE Trans. Systems, Man, & Cybernetics, Vol. 22, No. 6, 1992, pp. 1414-1427.

This method requires pre-definition of fuzzy sets and their membership functions. Then it can learn fuzzy rules from data using these input/output fuzzy sets.

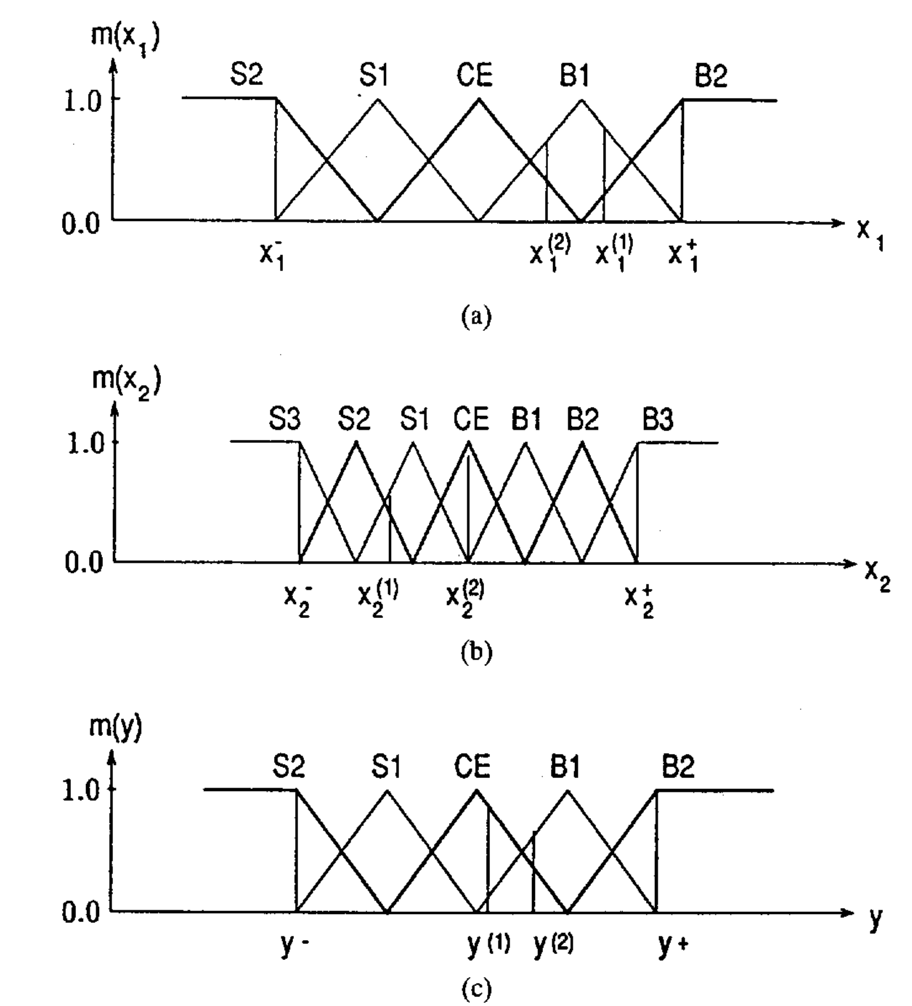
Step 1: Generate fuzzy rules from individual examples:

a) Determine the membership degrees of the given inputs and outputs with respect to their respective fuzzy subsets

b) For every given input/output find the fuzzy subset with the maximum membership degree

c) Using those identified fuzzy subsets to build a fuzzy rule from the given example.

Wang-Mendel Method



Example: $(x_1^{(1)}, x_2^{(1)}, y^{(1)})$

If x_1=B1 and x_2=S1
Then y=CE

Example: $(x_1^{(2)}, x_2^{(2)}, y^{(2)})$

If x_1=B1 and x_2=CE
Then y=B1

Wang-Mendel Method

Step 2: **Assign a truth degree to each rule**

If the rule "If x_1=A and x_2=B Then y=C" is generated from the example (x1, x2, y), then the degree of rule is defined as

$$D(\text{rule})=m_A(x_1)m_B(x_2)m_c(y)$$

Step 3: **Rule merging and conflict removing**

Merge the rules with the same conditions into one group. Take one rule with the highest degree from each group to enter the final knowledge base and discard the remaining rules in the group.

Extension of Wang-Mendel method for big data

V. Lopez, et al., Cost-sensitive fuzzy rule-based classification systems under the MapReduce framework for imbalanced big data, Fuzzy Sets and Systems, 258 (2015), pp. 5-38.

- Steps 1 and 3 are the same as in the Wang-Mendel method
- Step 2: new criteria for rule weight are proposed

$$Rw_j = \frac{\sum_{x \in Class C_j} \mu_{A_j}(x) - \sum_{x \overline{\in} Class C_j} \mu_{A_j}(x)\Big)}{\sum_{\forall x} \mu_{A_j}(x)\Big)}$$

$$Rw_j = \frac{\sum_{x \in Class C_j} \mu_{A_j}(x) C_x - \sum_{x \overline{\in} Class C_j} \mu_{A_j}(x) C_x}{\sum_{\forall x} \mu_{A_j}(x)\Big) C_x}$$

imbalanced data

where A_j is the condition of Rule j, C_j is the consequent class in Rule j, and C_x is the cost of misclassification for example *x*

Reading Guidance

Fuzzy part is not included in the Mitchell's book of Machine Learning. You can study the relevant knowledge from some other books. Examples are:

1.Chapter 4 "Fuzzy expert systems" in **Artificial Intelligence – A guide to intelligent systems,** by Michael Negnevitsky, Addison Wesley, 2002
2.Chapters 20-22**, in Computational Intelligence: An Introduction,** by Andries P. Engelbrecht, John Wiley &Sons, 2007.

Students should also read the paper (available in the blackboard):
L.X. Wang and J.M. Mendel, "Generating fuzzy rules by learning from examples", IEEE Trans. Systems, Man, & Cybernetics, Vol. 22, No. 6, 1992, pp. 1414-1427.

Other papers on genetic-fuzzy systems and fuzzy learning in big data are also available in blackboard for optional reading.