Soft Computing
in Fault Detection and Isolation

PART IV

Fuzzy and neuro-fuzzy systems in fault diagnosis

Marcin Witczak

8 października 2012

OUTLINE

- ➥ Concepts and history of fuzzy and neuro-fuzzy systems
- ➥ Fuzzy set theory in fault diagnosis
- ➥ Fault detection with fuzzy systems
  - Fuzzy and neuro-fuzzy models
  - Fuzzy and neuro-fuzzy observers
- ➥ Fault isolation with fuzzy systems
  - Fuzzy residual evaluation
  - Neuro-fuzzy classifiers
  - Fuzzy clustering

☛ CONCEPTS AND HISTORY OF FUZZY AND NEURO-FUZZY SYSTEMS

➜ Outline of fuzzy logic history

1965: Fuzzy logic: Zadeh

1977: Fuzzy control: Mamdani

1981: Fuzzy clustering algorithms: Bezdek

1985: Development of Takagi-Sugeno fuzzy models: Takagi and Sugeno

1990: Development of neuro-fuzzy techniques: Czogała, Rutkowski, Nauck, Babuška

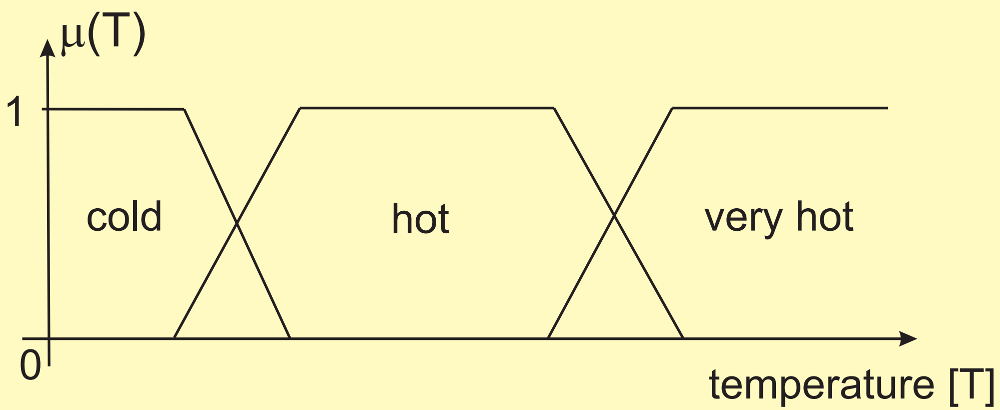
1991: Fuzzy control system for the Sendai railway, Japan

1996: Computing with words: Zadeh

2001: Granular computing: Pedrycz

➜ Fuzzy sets

- Fuzzy set theory was formalised by Professor Lofti Zadeh at the University of California in 1965.
- Fuzzy sets are an extension of the classical set theory used in fuzzy logic. A fuzzy set is characterized by a membership function, which maps the members of the universe into the unit interval $[0, 1]$. For the universe X and given the membership-degree function $\mu : X \rightarrow [0, 1]$, the fuzzy set A is defined as $A = \{(x, \mu(x))|x \in X\}$.
- Membership functions and linguistic labels:



➜ Fuzzy knowledge representation

- Fuzzy rule:

$$\text{IF (fuzzy antecedent) THEN (fuzzy consequent)}$$

- Sample fuzzy rule:

$$\text{IF } x \text{ is } A \text{ THEN } y \text{ is } B,$$

  where A, B - fuzzy sets

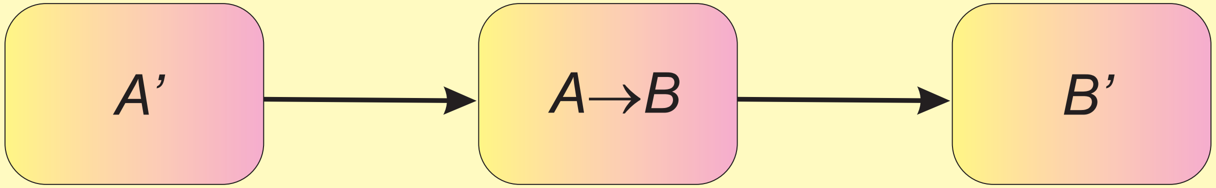
- Fuzzy rule is a fuzzy relation $R : (X \times Y) \to [0, 1]$:

$$\mu_R(x, y) = \mu_A(x)\mu_B(y) \text{ – Larsen implication}$$

$$\mu_R(x, y) = \min(\mu_A(x), \mu_B(y)) \text{ – Mamdani implication}$$

$$\mu_R(x, y) = \min(1, 1 - \mu_A(x) + \mu_B(y)) \text{ – Łukasiewicz implication}$$

➜ Fuzzy inference mechanism



where A' and B' stand for fuzzy sets, and $A \to B$ is fuzzy implication.

- inference mechanism is based on the *modus ponens* rule:

  given the rule

  $$\text{R: IF } x \text{ is } A \text{ THEN } y \text{ is } B$$

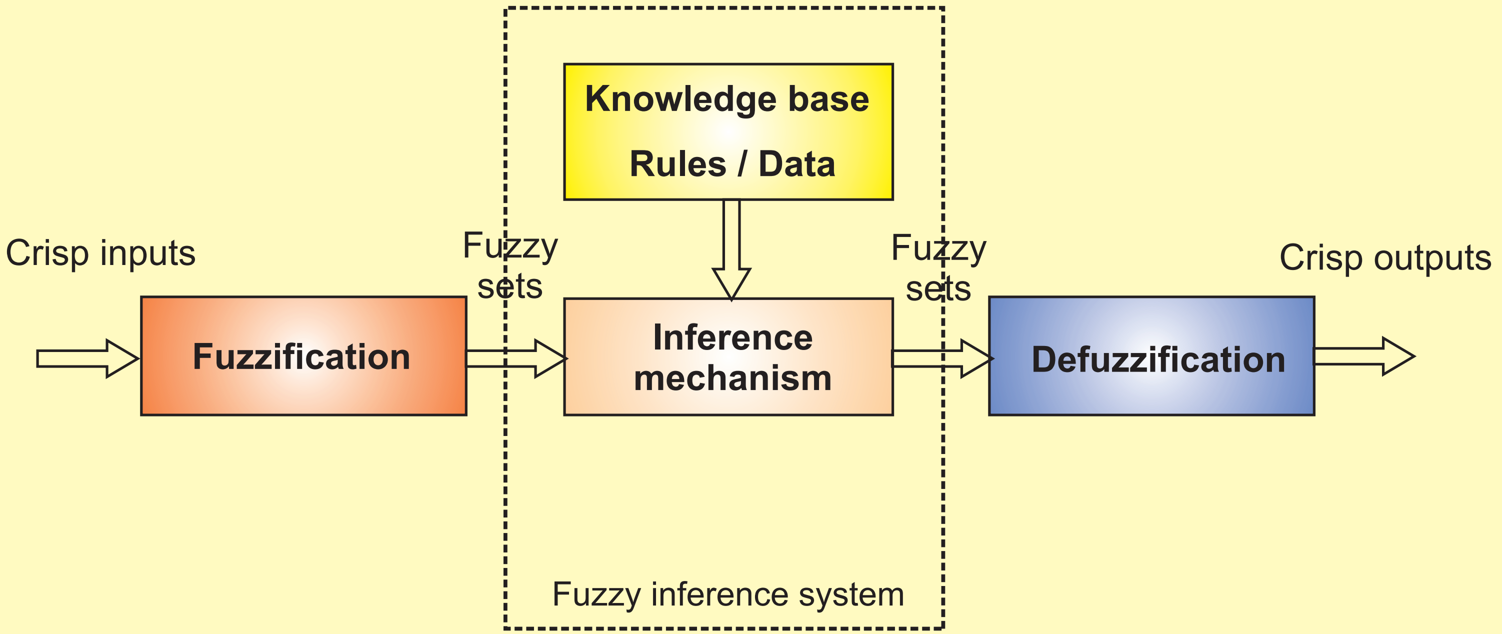
  and the fact

  $$x \text{ is } A',$$

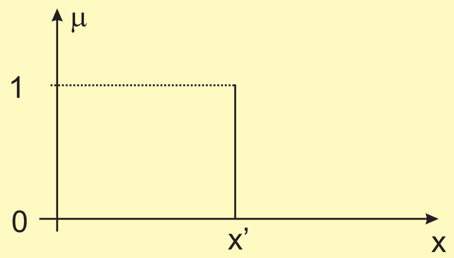
  the output fuzzy set B' is derived using relational composition:

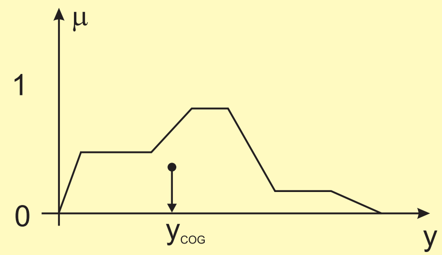
  $$B' = A' \circ R.$$

➜ Fuzzy inference with crisp inputs and outputs



- fuzzification: the singleton method
- defuzzification: the Center of Gravity





☛ FUZZY SET THEORY IN FAULT DIAGNOSIS

➜ Why use fuzzy logic for fault diagnosis?

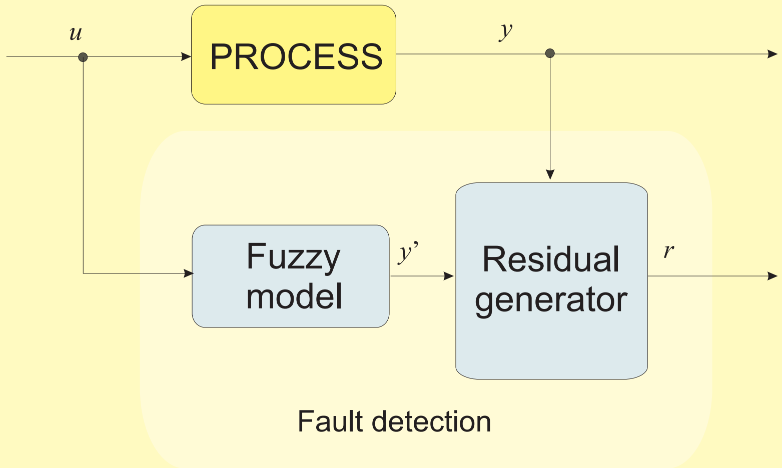
- Transparent representation of the system under study
- Linguistic interpretation in the form of fuzzy sets
- Ability to combine quantitative and qualitative knowledge
- Rules extracted from data can be validated by an expert
- Non-linear mappings
- Ability to represent some kind of uncertainty present in real processes

➜ Fuzzy and neuro-fuzzy systems in fault diagnosis

- Fault detection:
  - input-output fuzzy and neuro-fuzzy models
  - fuzzy observers
- Fault isolation:
  - fuzzy and neuro-fuzzy classifiers
  - fuzzy residual evaluation
  - fuzzy decision-making (fuzzy expert systems)
  - fuzzy pattern recognition

☛ FAULT DETECTION WITH FUZZY SYSTEMS

➜ Fuzzy and neuro-fuzzy models



- Linguistic models: Zadeh (1973), Frank, (1996), Isermann (1998)
- Relational models: Pedrycz (1984), Amann *et al.* (2001)
- Takagi-Sugeno models: Takagi and Sugeno (1985), Babuška (1998)

➜ Linguistic fuzzy model

The linguistic model gives a qualitative description of the process and is usually used to describe the knowledge obtained from process operators.

- Knowledge representation:

$$R_k : \text{ IF } x_1 \text{ is } A_1^k \text{ AND } \ldots \text{ AND } x_n \text{ is } A_n^k \text{ THEN } y \text{ is } B^k$$

- Fuzzy inference:

$$\mu_{B'^k}(y) = \max_{\boldsymbol{x}\in X} \min_{\boldsymbol{x}\in X, y\in Y} (\mu_{A^k}(\boldsymbol{x}), \mu_{B^k}(y))$$

$$B' = \bigcup_{k=1}^{N} B'^k$$

- Center of Gravity defuzzification:

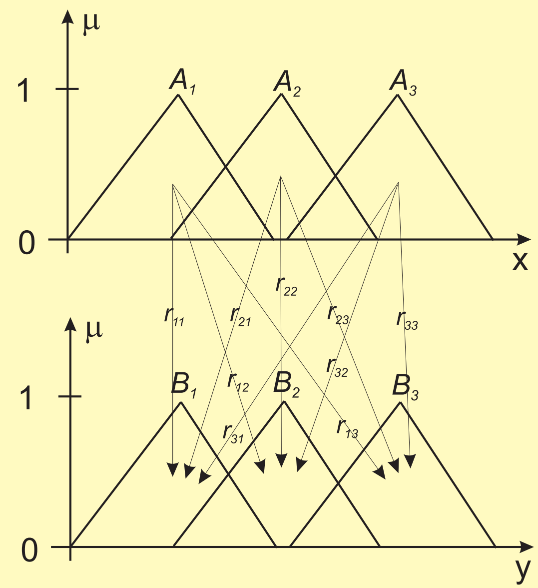
$$\hat{y} = \frac{\int_Y y\mu_{B'}(y)\mathrm{d}y}{\int_Y \mu_{B'}(y)\mathrm{d}y}$$

➜ Fuzzy relational model

Fuzzy relational models describe the relationship between input and output variables by using the relational matrix

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ \vdots & \vdots & \dots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{bmatrix},$$

where the elements of the matrix $R \in [0,1]^{m \times n}$ represent the strength of association between fuzzy sets.



➜ Fuzzy relational model

- Knowledge representation:

$$\text{IF } x \text{ is } A_i \text{ THEN } y \text{ is } B_1(r_{i1}) \text{ AND } B_2(r_{i2}) \text{ AND } \ldots \text{ AND } B_m(r_{im})$$

- Inference mechanism:
  The output fuzzy set B' is derived using relational composition (e.g. the max-min composition)

$$B' = A' \circ R$$

  where $A' = [\mu_{A_1}(x), \ldots, \mu_{A_n}(x)]$ and $B' = [\mu_{B_1}(y), \ldots, \mu_{B_m}(y)]$

- Defuzzification

$$\hat{y} = \frac{\sum_{i=1}^{m} \mu_i(y) y_{B_i}}{\sum_{i=1}^{m} \mu_i(y)},$$

  where y_{B_i} is the center of gravity of the fuzzy set B_i

➜ Takagi-Sugeno fuzzy model

The Takagi-Sugeno model can be interpreted in terms of local linear models, thus is well suited for mathematical analysis

- Knowledge representation:

$$R_k: \text{ IF } x_1 \text{ is } A_1^k \text{ AND } \ldots \text{ AND } x_n \text{ is } A_n^k \text{ THEN } y_k = f(\boldsymbol{x})$$

- Inference mechanism and defuzzification operation:

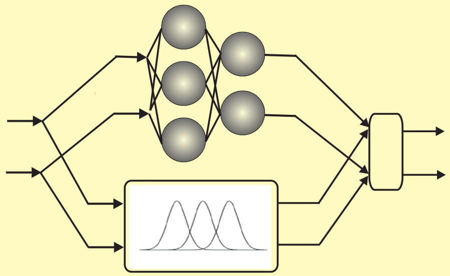
$$\hat{y} = \frac{\sum_{i=1}^{N} \mu_i(\boldsymbol{x}) y_i}{\sum_{i=1}^{N} \mu_i(\boldsymbol{x})},$$

where

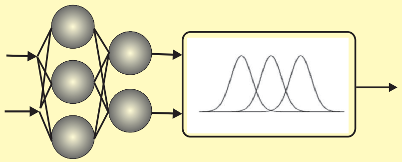
$$\mu_i(\boldsymbol{x}) = \mu_{A_1^i}(x_1) \wedge \mu_{A_2^i}(x_2) \wedge \ldots \wedge \mu_{A_n^i}(x_n)$$

➜ Neuro-fuzzy models

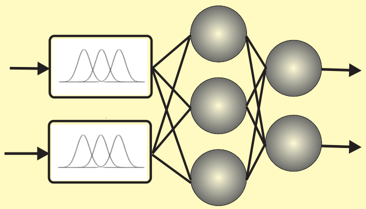
- Paralel connection of a fuzzy system and a neural network:



- Cascade connection of a fuzzy system and a neural network:

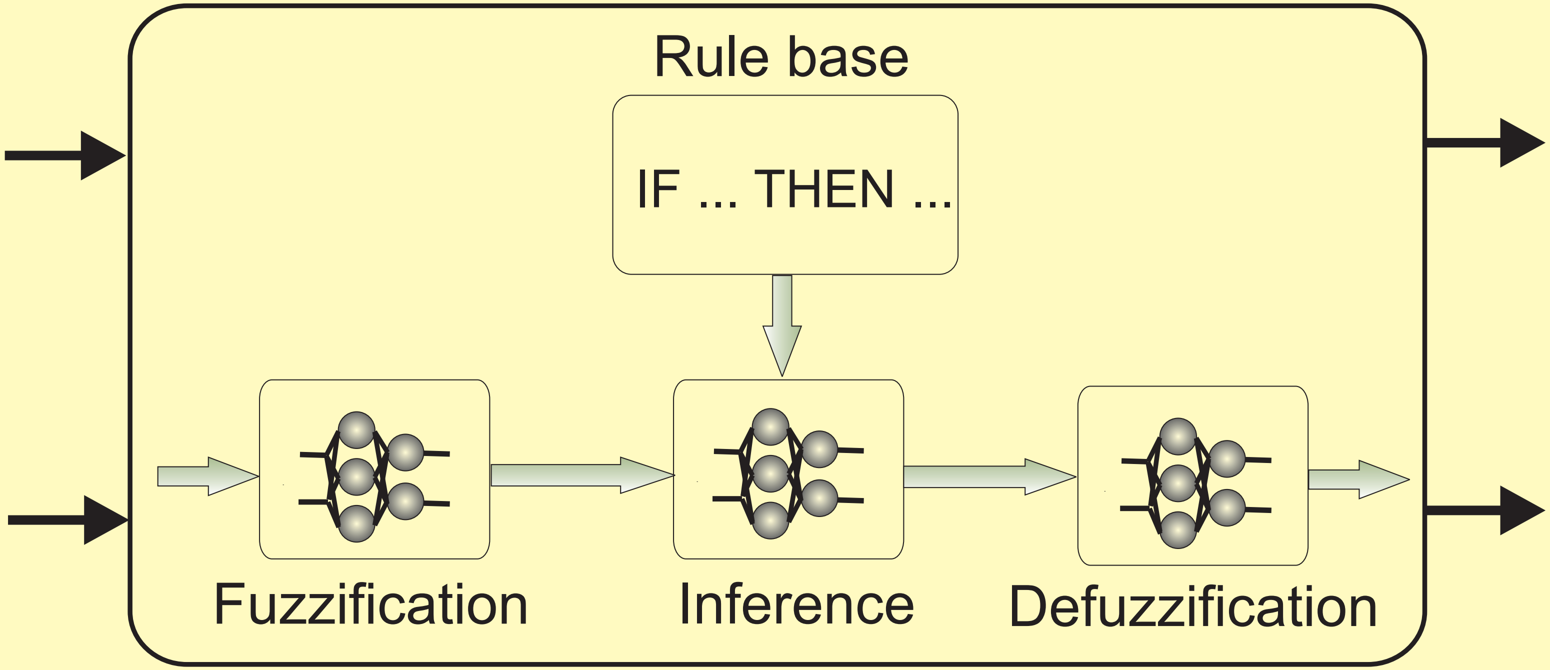


- Neural network with fuzzy inputs:



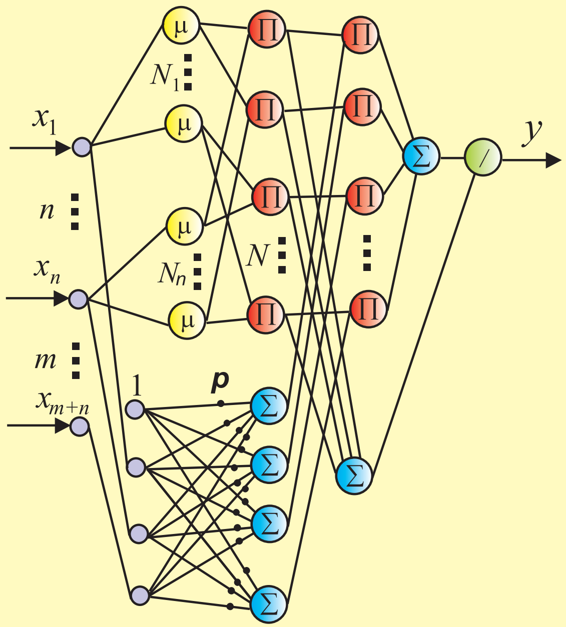
➜ Neuro-fuzzy models

- Fuzzy inference using neural networks:



➜ Takagi-Sugeno neuro-fuzzy model

- The Takagi-Sugeno neuro-fuzzy network is the most commonly used neuro-fuzzy model for fault detection.
- Sample Takagi-Sugeno neuro-fuzzy network with linear consequents:



where

x_i – input variable, y – output variable, N – number of fuzzy rules, N_j – number of fuzzy partitions, μ – membership function, $\boldsymbol{p}_a$ – parameters of membership functions, $\boldsymbol{p}_c$ – parameters of linear consequents

➜ Building fuzzy and neuro-fuzzy models

- Design procedure:
  - choosing the type of the fuzzy or neuro-fuzzy model
  - developing the knowledge base
- Choice of the fuzzy model type:
  - linguistic and relational models give a more qualitative description and are preferred to describe the knowledge obtained from process experts
  - Takagi-Sugeno models are usually used when only measurements are available and mathematical analysis of the model is required
- Developing the knowledge base:
  - knowledge-based approach
  - data-driven approach
    - gradient descent algorithms
    - clustering algorithms
    - Wang-Mendel method
    - evolutionary algorithms

➜ Dynamic fuzzy models – knowledge representation

- Dynamic linguistic model:

$$\begin{aligned}\text{IF } u(k) \text{ is } A_1 \text{ AND } \dots \text{ AND } u(k-n_u) \text{ is } A_{n_u} \text{ AND} \\ y(k-1) \text{ is } A_{n_u+1} \text{ AND } \dots \text{ AND } y(k-n_y-1) \text{ is } A_{n_u+n_y} \\ \text{THEN } y(k) \text{ is } B\end{aligned}$$

- Takagi-Sugeno NARX (*Nonlinear AutoRegressive with eXogenous input*) model:

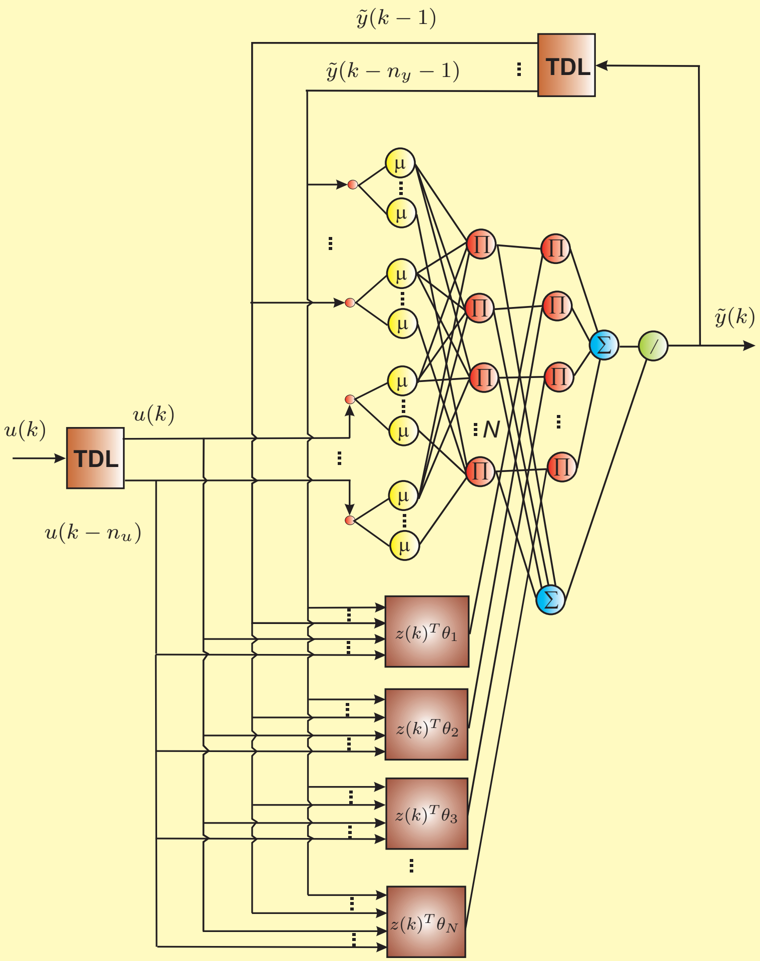
$$\begin{aligned}\text{IF } u(k) \text{ is } A_1 \text{ AND } \dots \text{ AND } u(k-n_u) \text{ is } A_{n_u} \text{ AND} \\ y(k-1) \text{ is } A_{n_u+1} \text{ AND } \dots \text{ AND } y(k-n_y-1) \text{ is } A_{n_u+n_y} \\ \text{THEN } y(k) = \boldsymbol{z}^T(k)\boldsymbol{\theta},\end{aligned}$$

  where

  $\boldsymbol{z}(k) = [u(k), \dots, u(k-n_u), y(k-1), \dots y(k-n_y-1)]^T$, $\boldsymbol{\theta}$ is the vector of parameters, n_u and n_y define the order of the system

➜ Dynamic Takagi-Sugeno neuro-fuzzy network

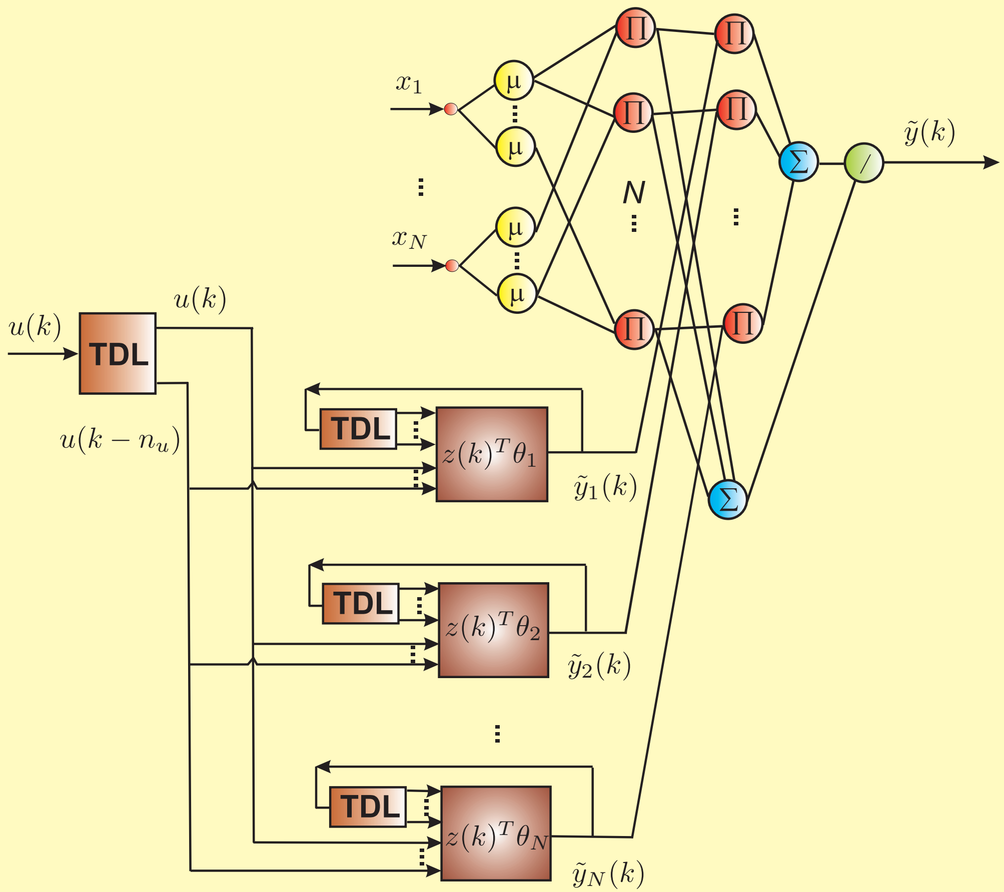
- Global dynamic model:



TDL - Tapped Delay Line

➜ Dynamic Takagi-Sugeno neuro-fuzzy network

- Set of local linear dynamic models:



➜ Problems with the model-based approach for fault detection

- Residual signal:

$$r(k) = y(k) - \tilde{y}(k)$$

- Objectives:

$$r(k) = 0 \text{ for the fault-free case}$$

$$r(k) \neq 0 \text{ when a fault occurs}$$

- Objectives cannot be fulfilled due to problems with:
  - disturbances that corrupt the measurements
  - modelling errors (regardless of the identification method used there is always model-reality mismatch)

➜ Robust fault detection using the Takagi-Sugeno model and adaptive threshold

- Problem: Takagi-Sugeno model uncertainty
- Solution: The adaptive threshold technique
- Let us consider the following Takagi-Sugeno model:

$$\tilde{y}(k) = \sum_{i=1}^{N} \phi_i(k)\tilde{y}_i(k)$$

  where $\tilde{y}_i(k)$ is the output of the i-th rule
- The model can be viewed as a model linear in parameters if the parameters of membership functions are constant:

$$\tilde{y}(k) = \boldsymbol{x}^T(k)\boldsymbol{\theta}.$$

- Output error:

$$\varepsilon(k) = y(k) - \boldsymbol{x}^T(k)\boldsymbol{\theta}$$

➜ Bounded-error approach

- Feasible set of parameters:

$$\mathbb{P} = \{\boldsymbol{\theta} \in \mathbb{R}^n \,|\, y(k) - \varepsilon \leqslant \boldsymbol{x}^T(k)\boldsymbol{\theta} \leqslant y(k) + \varepsilon, k = 1, \ldots, N\}$$

- Confidence interval for the system output:

$$\boldsymbol{x}^T(k)\boldsymbol{\theta}^{\min}(k) \leqslant \boldsymbol{x}^T(k)\boldsymbol{\theta} \leqslant \boldsymbol{x}^T(k)\boldsymbol{\theta}^{\max}(k),$$

where

$$\boldsymbol{\theta}^{\min}(k) = \arg\min_{\boldsymbol{\theta}\in\mathbb{W}} \boldsymbol{x}^T(k)\boldsymbol{\theta},$$

$$\boldsymbol{\theta}^{\max}(k) = \arg\max_{\boldsymbol{\theta}\in\mathbb{W}} \boldsymbol{x}^T(k)\boldsymbol{\theta}$$

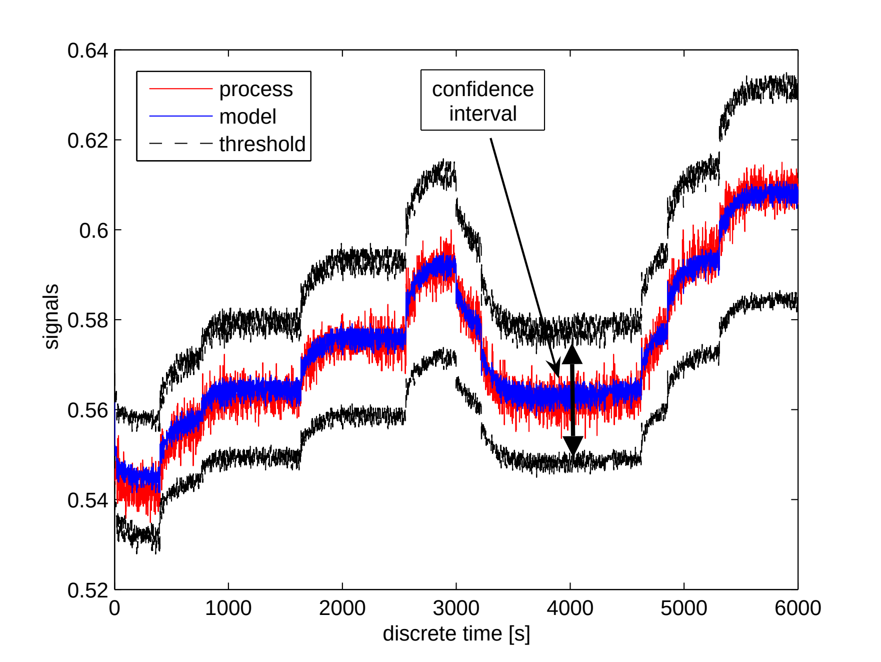
- Residuals:

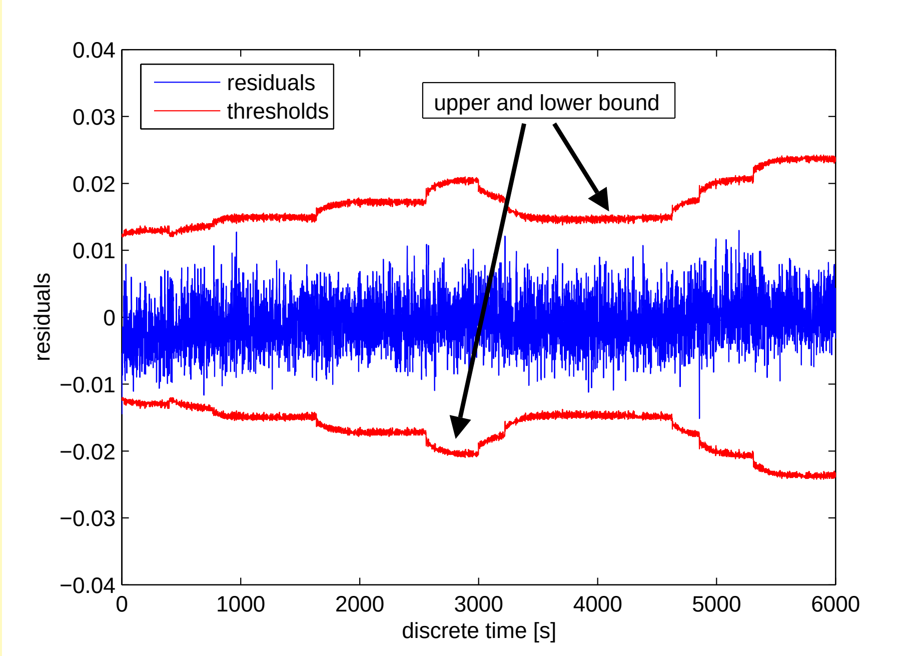
$$r(k) = y(k) - \tilde{y}(k)$$

- Adaptive threshold:

$$\boldsymbol{x}^T(k)\boldsymbol{\theta}^{\min}(k) - \varepsilon - \tilde{y}(k) \leqslant r(k) \leqslant \boldsymbol{x}^T(k)\boldsymbol{\theta}^{\max}(k) + \varepsilon - \tilde{y}(k)$$

➜ Illustration of the adaptive threshold method



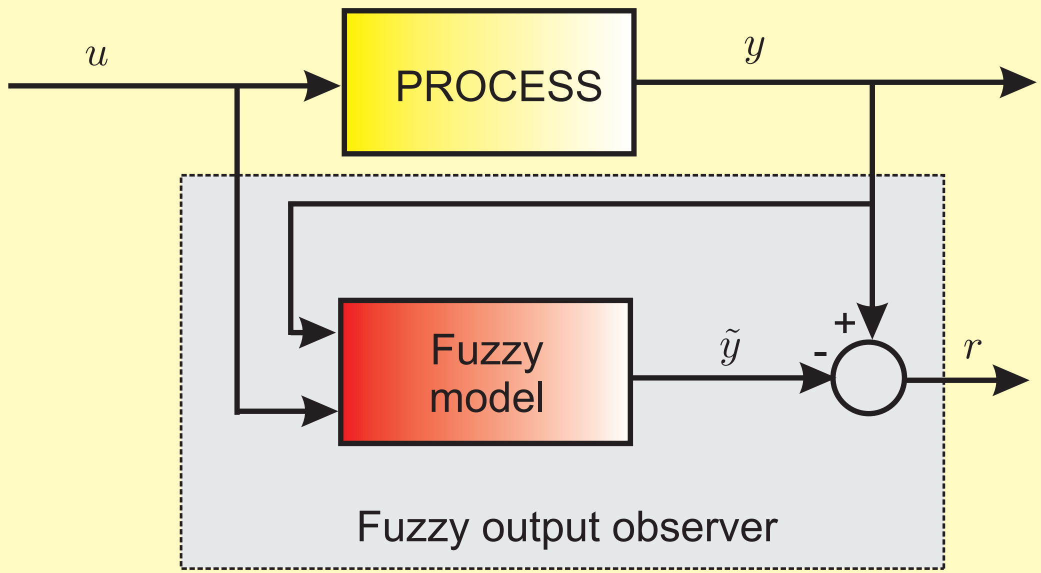


confidence interval for the system output

residuals

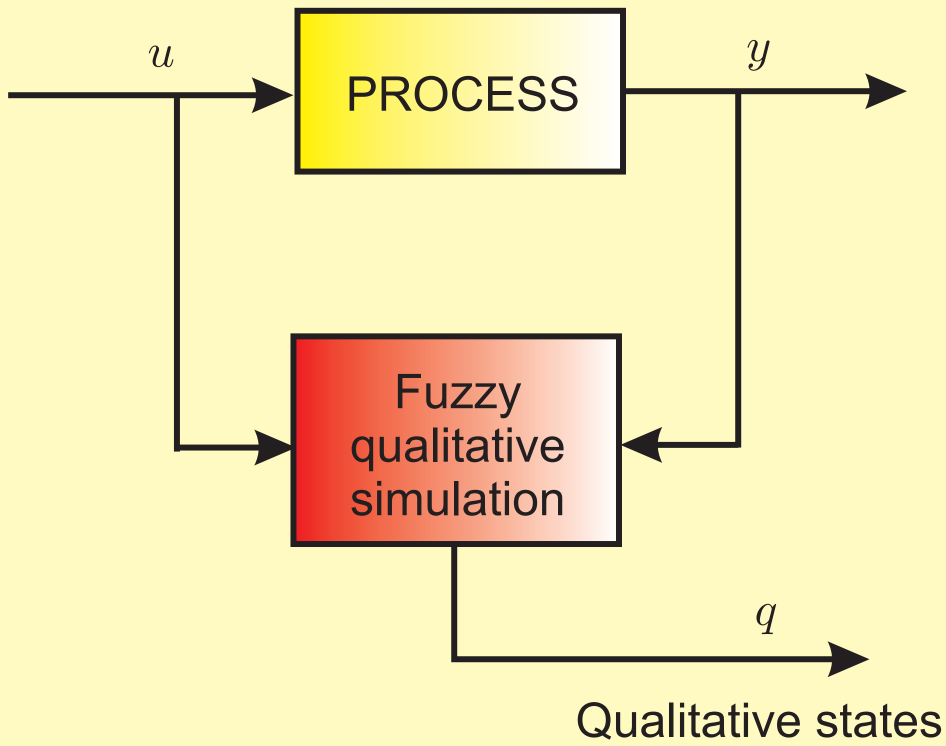
➜ Fuzzy-observer-based residual generation

- Qualitative observer: Frank and Köppen-Seliger (1997), Zhuang *et al* (1997)
- Takagi-Sugeno model-based fuzzy observer: Chen and Patton (1999)
- Relational observer: Amann *et al* (2000)

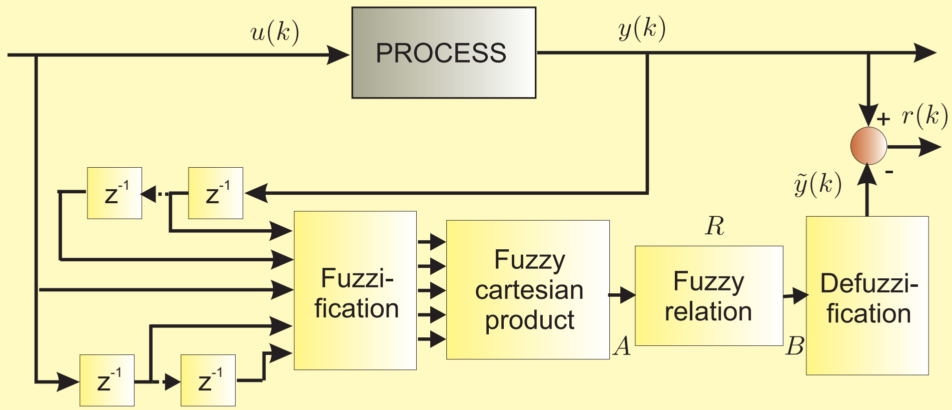


➜ Qualitative observer

- The dynamic behavior of the process is described by a small number of symbols or qualitative values specified by fuzzy sets:



➜ Fuzzy relational observer



- The core problem in fuzzy relational model design is the search for the best-matching relational matrix R:

$$\frac{\partial J(R)}{\partial R} = 0 \Rightarrow R = R_{opt}$$

  - iterative search Dubois (1992)
  - recursive least squares Jang *et al.* (1997)
  - first-order gradient descent method Isermann (1988)

➜ Takagi-Sugeno fuzzy state-space model

- Another approach to describe dynamic processes
- Knowledge representation:

$$R_i : (i = 1, \ldots, N) \text{ IF } w(k) \text{ is } M_i \text{ THEN } \begin{cases} x(k+1) = A_i x(k) + B_i u(k) \\ y(k) = C_i x(k) + D_i u(k) \end{cases},$$

where $x(k) \in \mathbb{R}^n, u(k) \in \mathbb{R}^r, y(k) \in \mathbb{R}^m$, and A_i, B_i, C_i and D_i are time invariant matrices of appropriate dimensions, $w(k)$ is a premise variable, M_i is a fuzzy set

- The global state and output of the system are inferred as follows:

$$\begin{cases} x(k+1) = \sum_{i=1}^{N} \mu_i(w(k))[A_i x(k) + B_i u(k)] \\ y(k) = \sum_{i=1}^{N} \mu_i(w(k))[C_i x(k) + D_i u(k)] \end{cases}.$$

- The membership grade functions $\mu_i(w(k))$ satisfy the constraints

$$\begin{cases} \sum_{i=1}^{N} \mu_i(w(k)) = 1 \\ 0 \leqslant \mu_i(w(k)) \leqslant 1, \forall i = 1, 2, \ldots, N \end{cases}.$$

➜ Fuzzy Takagi-Sugeno observer

- Each local observer is associated with a fuzzy rule:

$$R_i : \text{IF } w(k) \text{ is } M_i \text{ THEN } \begin{cases} \hat{x}(k+1) = A_i\hat{x}(k) + B_i u(k) + K_i[y(k) - \hat{y}(k)] \\ \hat{y}(k) = C_i\hat{x}(k) + D_i u(k) \end{cases}$$

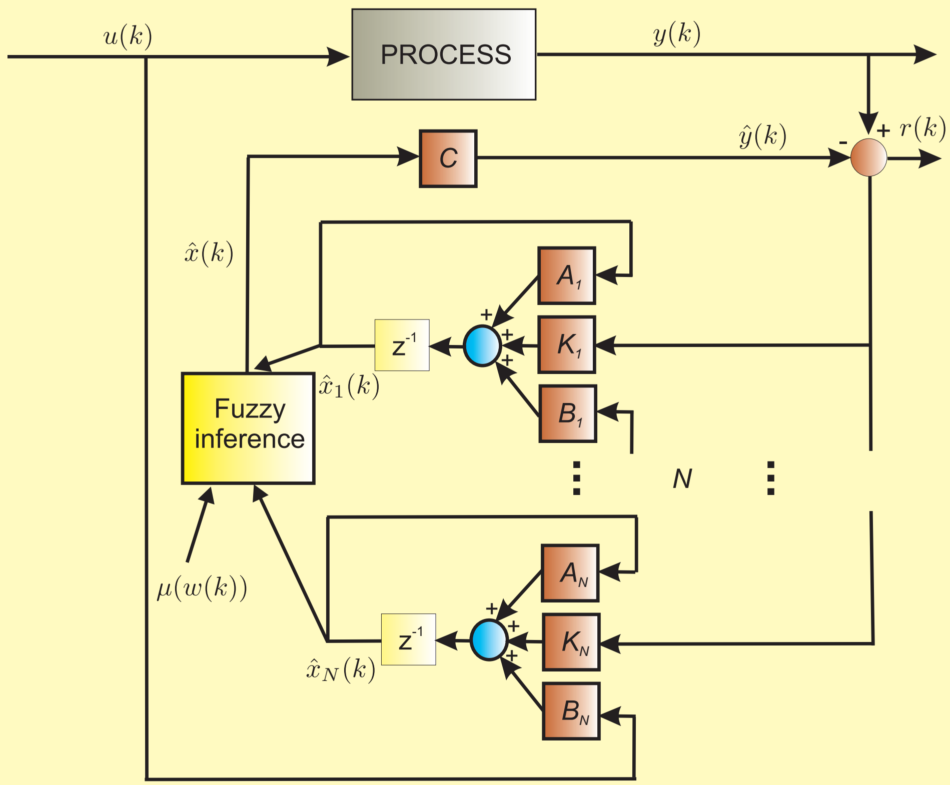
- Global fuzzy Takagi-Sugeno observer Chen and Patton (1999):

$$\begin{cases} \hat{x}(k+1) = \sum_{i=1}^{N} \mu_i(w(k))[A_i\hat{x}(k) + B_i u(k) + K_i(y(k) - \hat{y}(k))] \\ \hat{y}(k) = \sum_{i=1}^{N} \mu_i(w(k))[C_i\hat{x}(k) + D_i u(k)] \end{cases}$$

- The global observer is simplified if there is no uncertainty nor non-linearity included in the output equation:

$$\begin{cases} \hat{x}(k+1) = \sum_{i=1}^{N} \mu_i(w(k))[A_i\hat{x}(k) + B_i u(k) + K_i(y(k) - \hat{y}(k))] \\ \hat{y}(k) = \sum_{i=1}^{N} \mu_i(w(k))C_i\hat{x}(k) \end{cases}$$

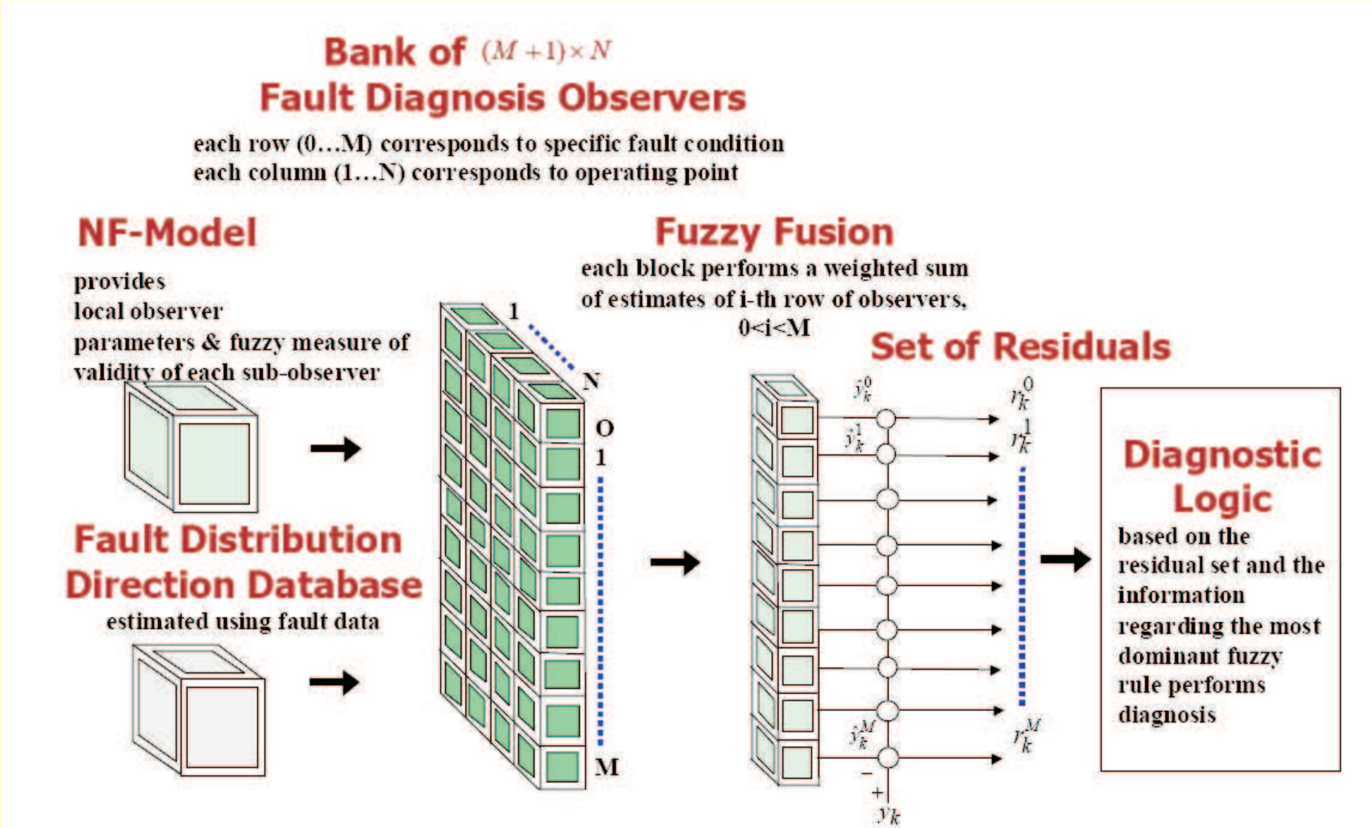
➜ Residual generation



- The fuzzy inference mechanism is used for the fusion of N local linear observers

➜ Neuro-fuzzy multiple-model observer approach

- A diagnostic system consists of $M + 1$ fault diagnosis observers (M – no. of faults)
  Uppal and Patton (2005): J. Adaptive Control and Signal Processing, Vol. 19
- Each fault detection observer: a non-linear system comprising a number of linear sub-observers, each one corresponding to a different operating point of the process



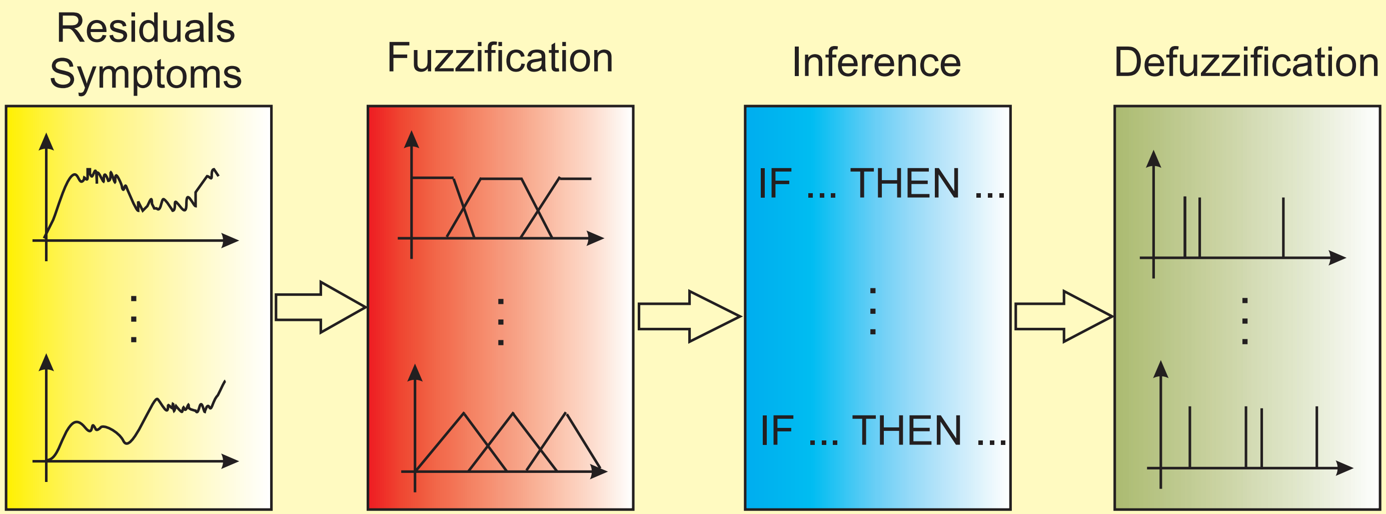
☛ FAULT ISOLATION WITH FUZZY SYSTEMS

➜ Fuzzy and neuro-fuzzy approaches used in fault isolation

- fuzzy residual evaluation: Schneider and Frank (1996), Isermann (1998)
- fuzzy and neuro-fuzzy classifiers: Nauck and Kruse (1998)
- fuzzy clustering algorithms: Bezdek (1981)

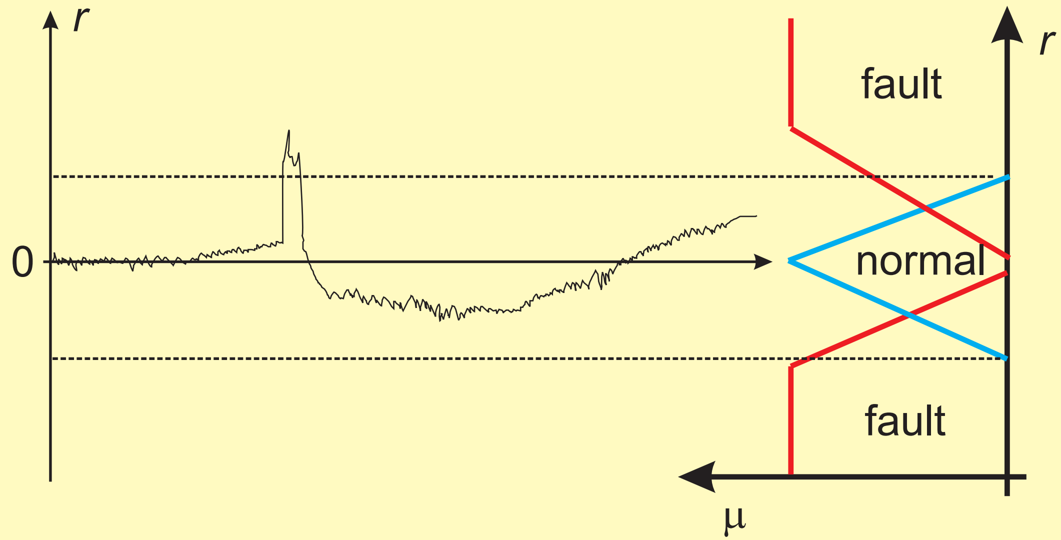
➜ Residual evaluation using fuzzy logic

- Residual evaluation is a logic decision-making process that transforms quantitative knowledge into qualitative statements
- The principle of residual evaluation using fuzzy logic consists of a three-step procedure: Frank and Koppen-Seliger (1997)



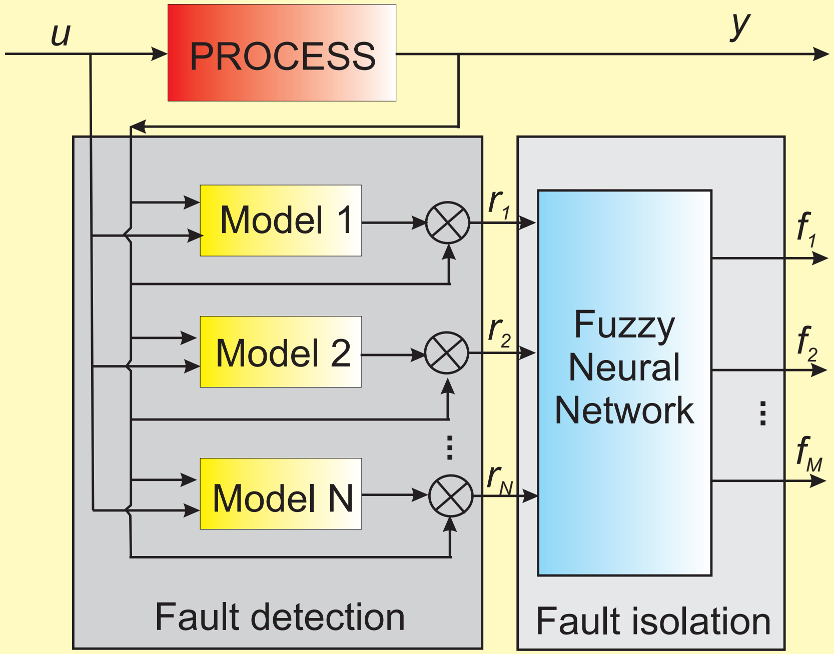
➜ Fuzzy threshold for residuals

- The fuzzy threshold tackles the problem of residual signal uncertainties
- Residual signals are described by linguistic variables Isermann (1998):



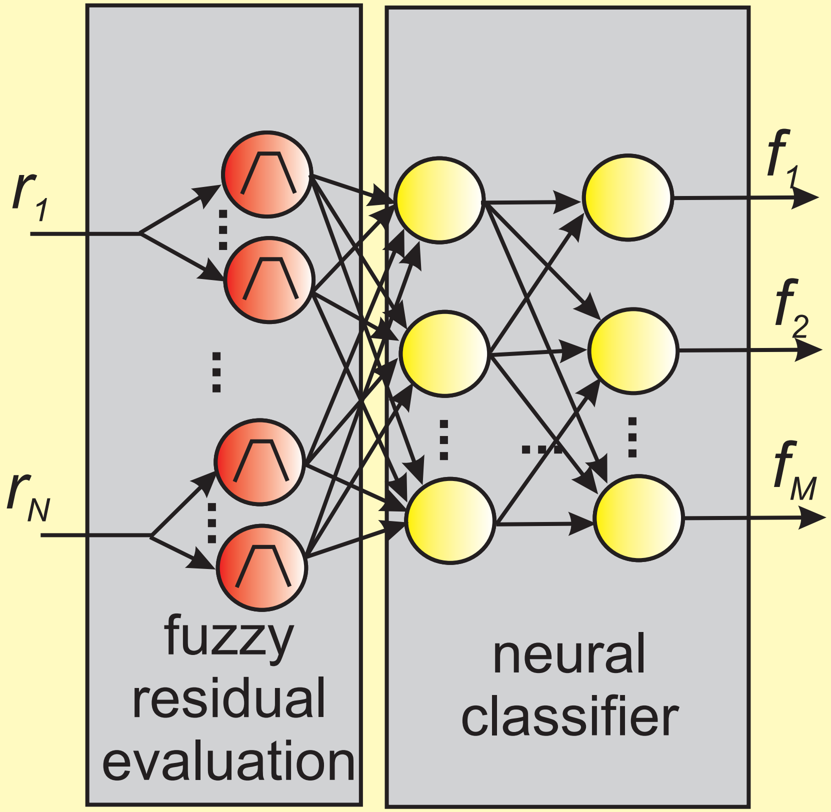
➜ Fault isolation with the use of fuzzy neural networks

- The expert's knowledge can be directly coded in the form of fuzzy rules
- Weights can be tuned using algorithms known for neural networks



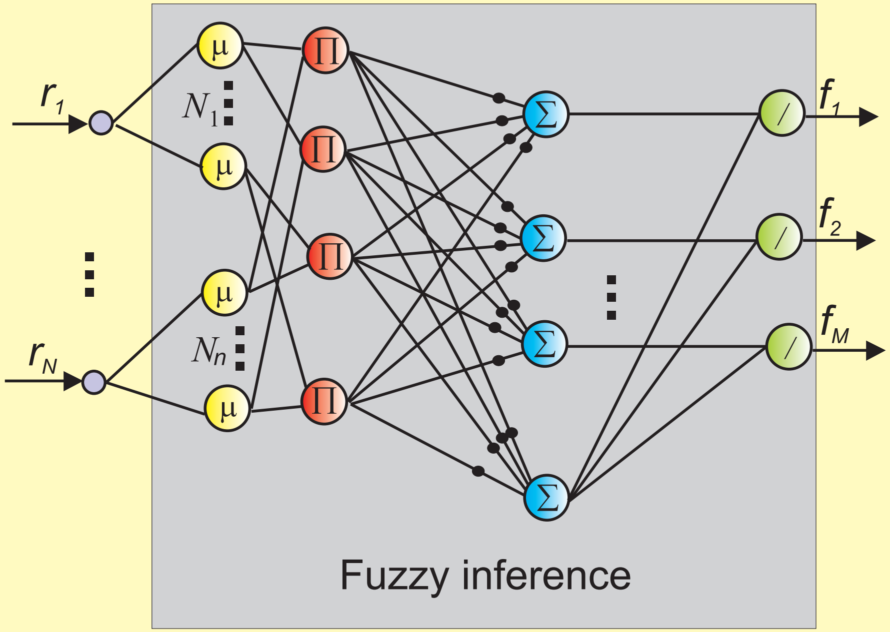
➜ Fuzzy neural network

- Residual evaluation and fault classification in the hybrid structure:



➜ Neuro-fuzzy classifier

- Qualitative and quantitative knowledge can be used to build classifiers as well
- The parameters of the classifier can be tuned using algorithms known for neural networks, so the required accuracy can be guaranteed
- Rules can be extracted from experimental data using a clustering algorithm:



➜ Fuzzy clustering

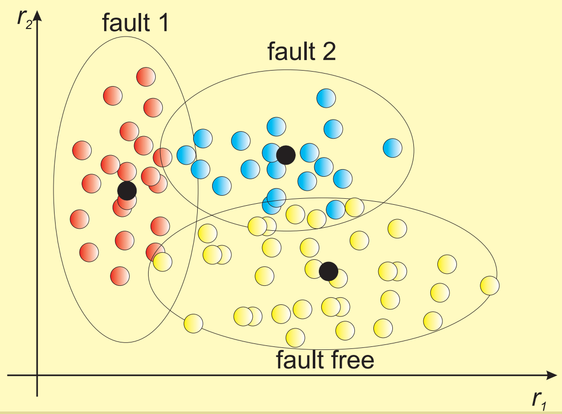
- Fuzzy clustering algorithms base on the minimization of the fuzzy *c-means* functional formulated by Bezdek (1981):

$$J(\boldsymbol{X};\boldsymbol{U},\boldsymbol{V}) = \sum_{i=1}^{c}\sum_{k=1}^{N}\mu_{ik}^{m}D_{ik}^{2},$$

where $\boldsymbol{U}$ is a matrix, which contains the membership degrees of data points from the matrix $\boldsymbol{X}$, $\boldsymbol{V}$ is a matrix which defines the cluster centers, and

$$D_{ik}^{2} = \|\boldsymbol{x_k} - \boldsymbol{v_i}\|^2 = (\boldsymbol{x_k} - \boldsymbol{v_i})^T\boldsymbol{A}(\boldsymbol{x_k} - \boldsymbol{v_i}).$$

- Fuzzy clusters can be converted to fuzzy rules, thus a fuzzy classifier can be built using clustering algorithms Babuška (1998):



☛ CONCLUDING REMARKS

- ❐ Fuzzy logic is an attractive tool for designing fault detection and isolation systems
  - Transparent knowledge in the form of fuzzy rules
  - Experts can formulate formal knowledge using linguistic values
  - Ability to simulate uncertainty
- ❐ Neuro-fuzzy approaches combine the advantages of fuzzy and neural techniques and thus are often used to build models, observers or classifiers for FDI tasks
- ❐ The uncertainty of the fuzzy system must be considered in order to ensure reliable fault diagnosis

Thank you