



# Dottorato di Ricerca in Ingegneria dell'Informazione

## Data Mining and Soft Computing

Francisco Herrera

Research Group on Soft Computing and  
Information Intelligent Systems (SCI<sup>2</sup>S)

Dept. of Computer Science and A.I.

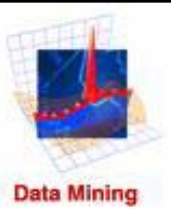
University of Granada, Spain

Email: [herrera@decsai.ugr.es](mailto:herrera@decsai.ugr.es)

<http://sci2s.ugr.es>

<http://decsai.ugr.es/~herrera>





# Data Mining and Soft Computing

## Summary

1. Introduction to Data Mining and Knowledge Discovery
2. Data Preparation
3. Introduction to Prediction, Classification, Clustering and Association
4. Data Mining - From the Top 10 Algorithms to the New Challenges
5. Introduction to Soft Computing. Focusing our attention in Fuzzy Logic and Evolutionary Computation
6. Soft Computing Techniques in Data Mining: Fuzzy Data Mining and Knowledge Extraction based on Evolutionary Learning
7. **Genetic Fuzzy Systems: State of the Art and New Trends**
8. **Some Advanced Topics I: Classification with Imbalanced Data Sets**
9. **Some Advanced Topics II: Subgroup Discovery**
10. **Some advanced Topics III: Data Complexity**
11. **Final talk: How must I Do my Experimental Study? Design of Experiments in Data Mining/Computational Intelligence. Using Non-parametric Tests. Some Cases of Study.**

# Genetic Fuzzy Systems: State of the Art and New Trends



## Outline

- ✓ Brief Introduction to Genetic Fuzzy Systems
- ✓ Tuning Methods: Basic and Advanced Approaches
- ✓ Genetic Fuzzy Systems Application to HVAC Problems
- ✓ GFSs: Current Trends and Prospects
- ✓ Concluding Remarks



## Outline

- ✓ Introduction to Genetic Fuzzy Systems
- ✓ Genetic Tuning Methods: Basic and Advanced Approaches
- ✓ Genetic Fuzzy Systems Application to HVAC Problems
- ✓ GFSs: Current Trends and Prospects
- ✓ Concluding Remarks

**F. Herrera, Genetic Fuzzy Systems: Taxonomy, Current Research Trends and Prospects. *Evolutionary Intelligence* 1 (2008) 27-46 [doi: 10.1007/s12065-007-0001-5](https://doi.org/10.1007/s12065-007-0001-5).**

# 1. Introduction to genetic fuzzy systems

---

- **Brief Introduction**
- **Taxonomy of Genetic Fuzzy Systems**
- **¿Why do we use GAs? GFSs versus Neural Fuzzy Systems**
- **The birth, GFSs roadmap, current state and most cited papers**

# 1. Introduction to genetic fuzzy systems

---

- **Brief Introduction**
- **Taxonomy of Genetic Fuzzy Systems**
- **¿Why do we use GAs? GFSs versus Neural Fuzzy Systems**
- **The birth, GFSs roadmap, current state and most cited papers**

# 1. Introduction to genetic fuzzy systems

## Brief Introduction

- The use of genetic/evolutionary algorithms (GAs) to design fuzzy systems constitutes one of the branches of the **Soft Computing** paradigm: **genetic fuzzy systems** (GFSs)
- The most known approach is that of **genetic fuzzy rule-based systems**, where some components of a fuzzy rule-based system (FRBS) are derived (**adapted or learnt**) using a GA
- Some other approaches include genetic fuzzy neural networks and genetic fuzzy clustering, among others

# 1. Brief introduction to genetic fuzzy systems

## Brief Introduction

### Evolutionary algorithms and machine learning:

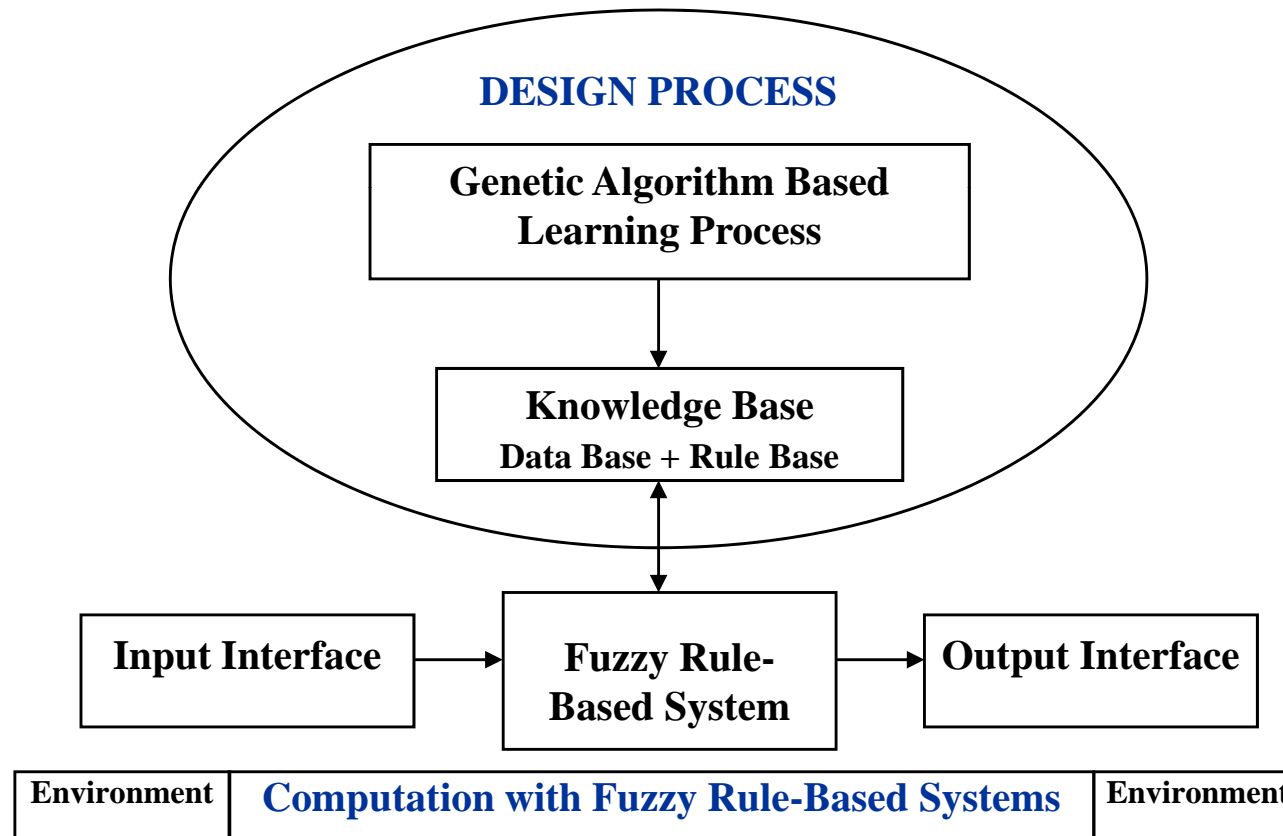
- Evolutionary algorithms were not specifically designed as machine learning techniques, like other approaches like neural networks
- However, it is well known that a learning task can be modelled as an optimization problem, and thus solved through evolution
- Their powerful search in complex, ill-defined problem spaces has permitted applying evolutionary algorithms successfully to a huge variety of machine learning and knowledge discovery tasks
- Their flexibility and capability to incorporate existing knowledge are also very interesting characteristics for the problem solving.



# 1. Brief introduction to genetic fuzzy systems

## Brief Introduction

### Genetic Fuzzy Rule-Based Systems:



# 1. Brief introduction to genetic fuzzy systems

## Brief Introduction

### Design of fuzzy rule-based systems:

- An FRBS (regardless it is a fuzzy model, a fuzzy logic controller or a fuzzy classifier), is comprised by two main components:
  - The **Knowledge Base (KB)**, storing the available problem knowledge in the form of fuzzy rules
  - The **Inference System**, applying a fuzzy reasoning method on the inputs and the KB rules to give a system output
- Both must be designed to build an FRBS for a specific application:
  - The KB is obtained from expert knowledge or by machine learning methods
  - The Inference System is set up by choosing the fuzzy operator for each component (conjunction, implication, defuzzifier, etc.)

**Sometimes, the latter operators are also parametric and can be tuned using automatic methods**

# 1. Brief introduction to genetic fuzzy systems

## Brief Introduction

The KB design involves two subproblems, related to its two subcomponents:

- Definition of the **Data Base** (DB):
  - Variable universes of discourse
  - Scaling factors or functions
  - Granularity (number of linguistic terms/labels) per variable
  - Membership functions associated to the labels
- Derivation of the **Rule Base** (RB): fuzzy rule composition

# 1. Brief introduction to genetic fuzzy systems

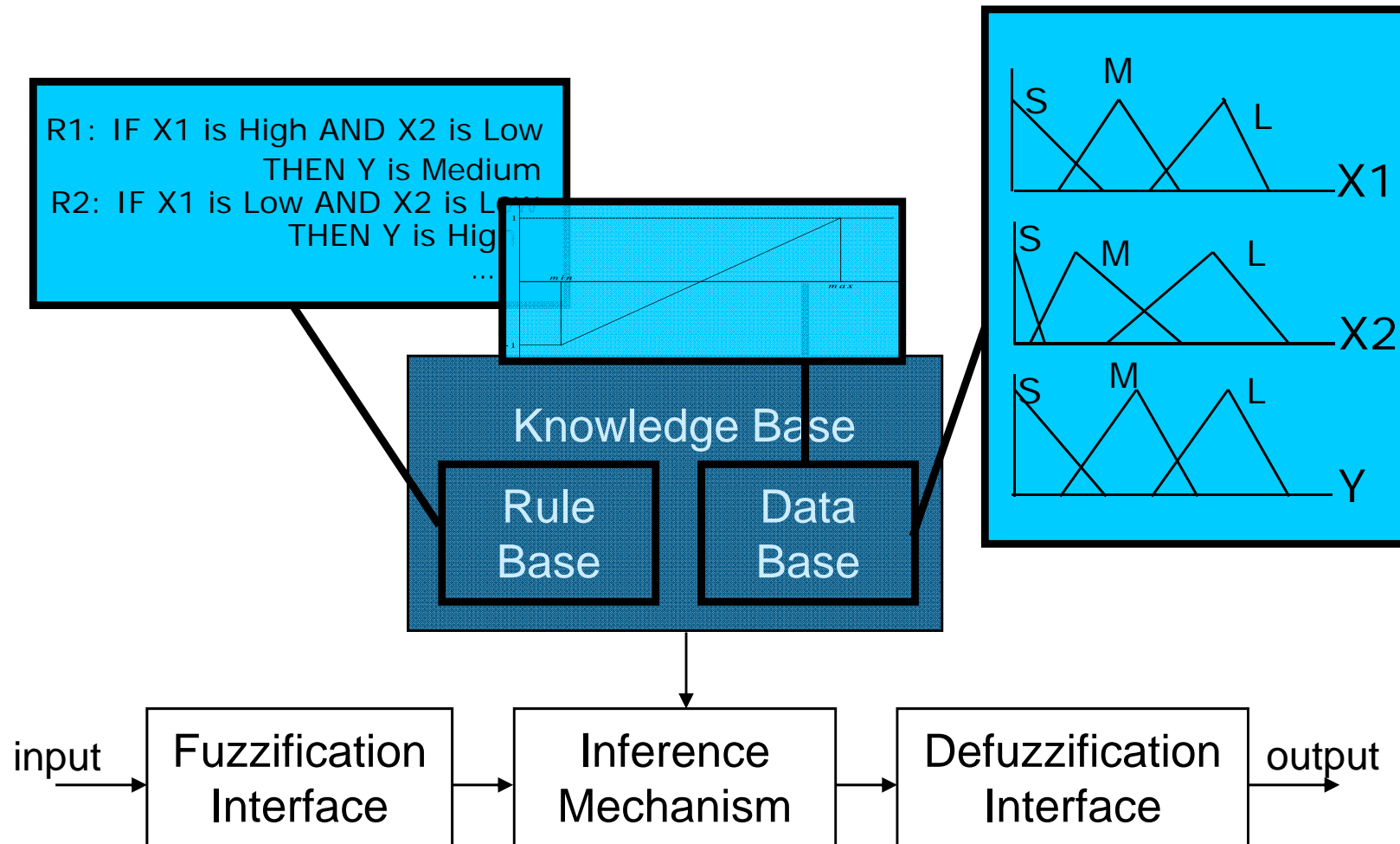
## **Brief Introduction**

As said, there are two different ways to design the KB:

- From **human expert** information
- By means of **machine learning methods** guided by the existing **numerical information** (fuzzy modeling and classification) or by a model of the system being controlled

# 1. Brief introduction to genetic fuzzy systems

## Brief Introduction



## Fuzzy rule-based system

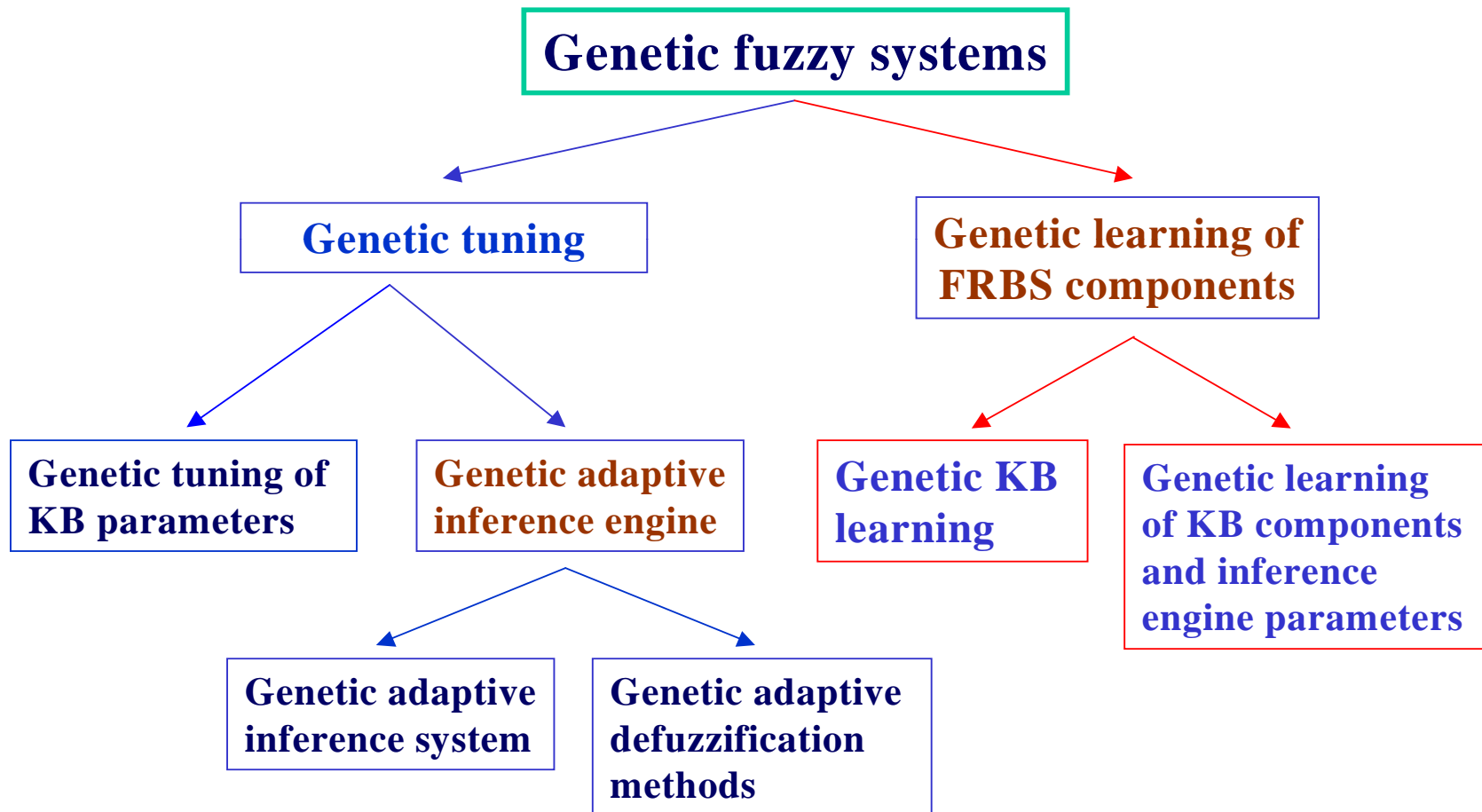
# 1. Introduction to genetic fuzzy systems

---

- **Brief Introduction**
- **Taxonomy of Genetic Fuzzy Systems**
- ¿Why do we use GAs? GFSs versus Neural Fuzzy Systems
- The birth, GFSs roadmap, current state and most cited papers

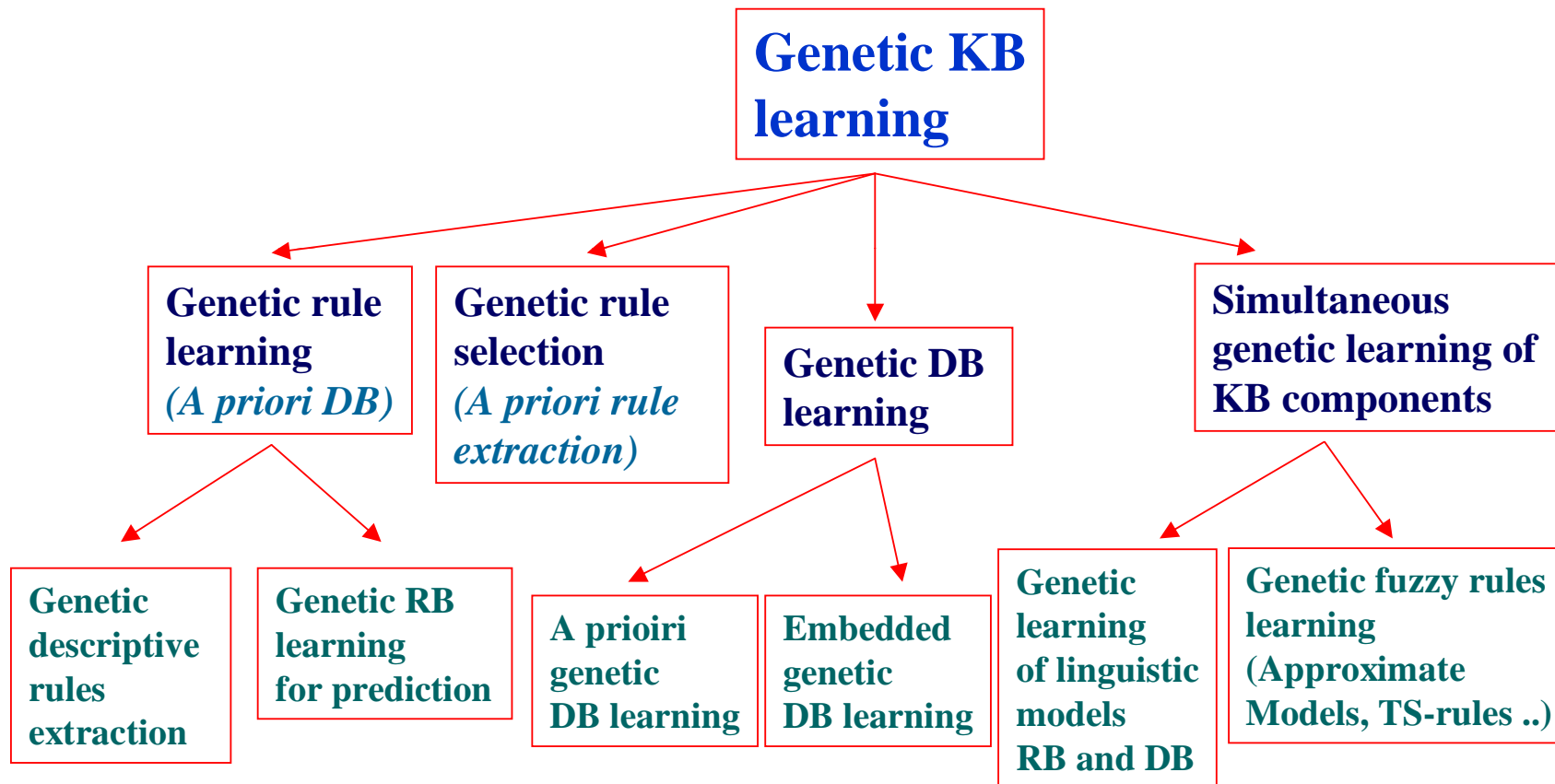
# 1. Brief introduction to genetic fuzzy systems

## Taxonomy of Genetic Fuzzy Systems



# 1. Brief introduction to genetic fuzzy systems

## Taxonomy of Genetic Fuzzy Systems



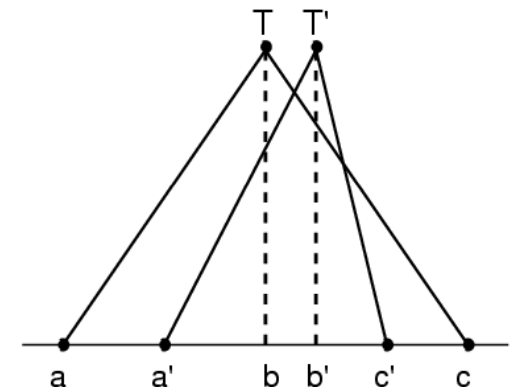
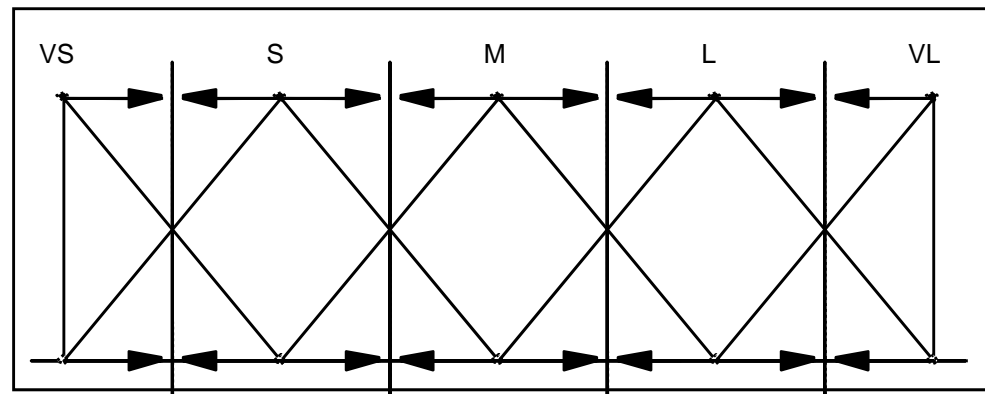


# 1. Brief introduction to genetic fuzzy systems

## 1. Genetic Tuning

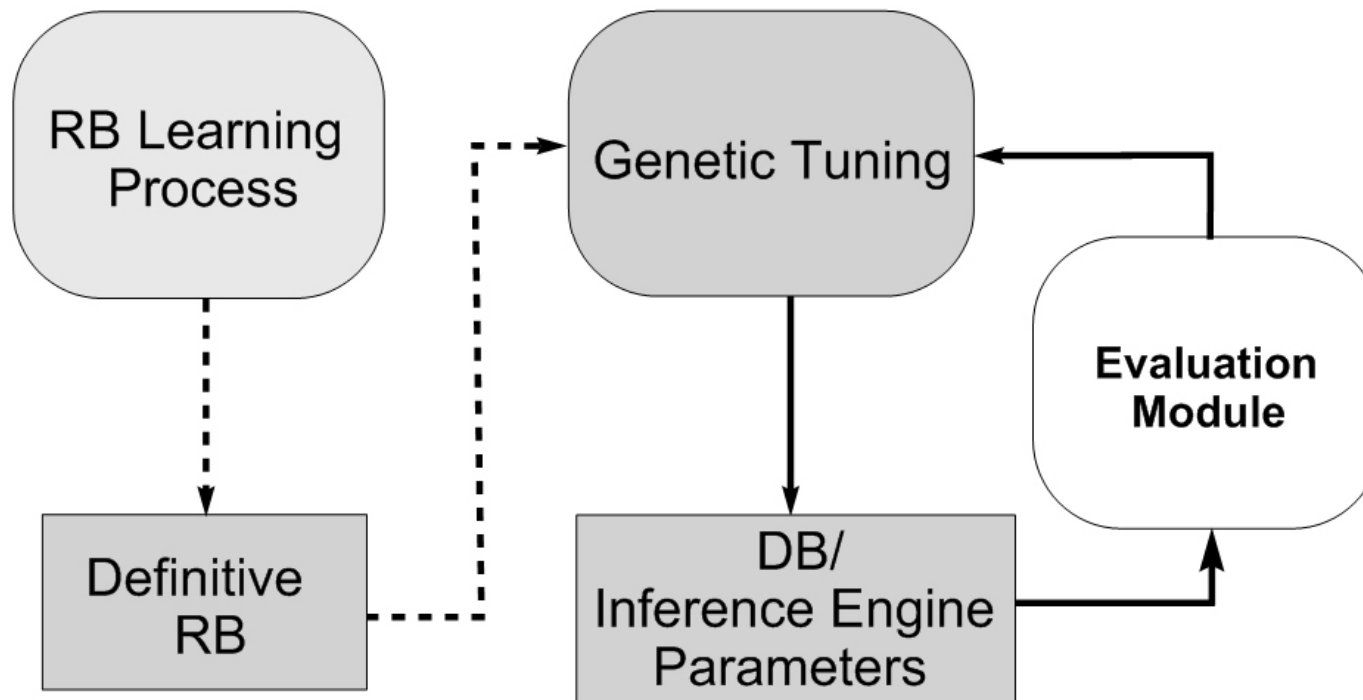
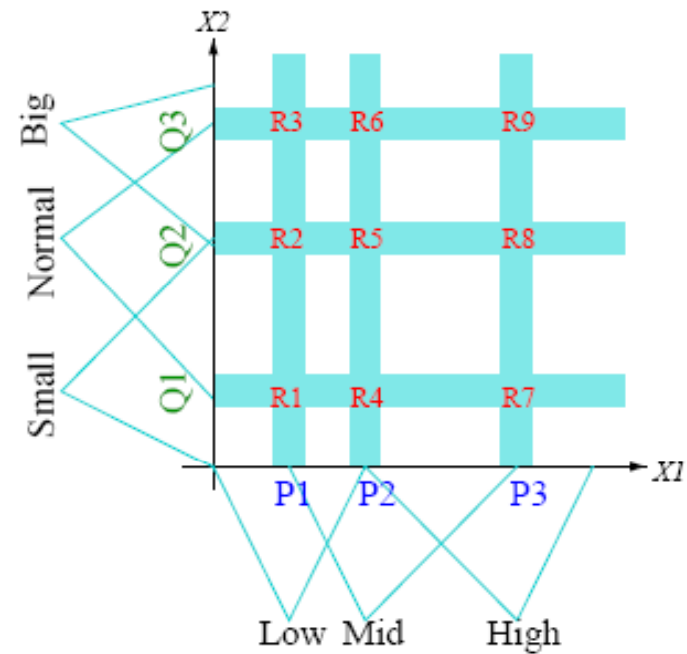
Classically:

- performed on a predefined DB definition
- **tuning** of the membership function shapes by a GA



- **tuning** of the inference parameters

# 1. Brief introduction to genetic fuzzy systems



# 1. Brief introduction to genetic fuzzy systems

## 2. Genetic Rule Learning

- A predefined Data Base definition is assumed
- The fuzzy rules (usually Mamdani-type) are derived by a GA

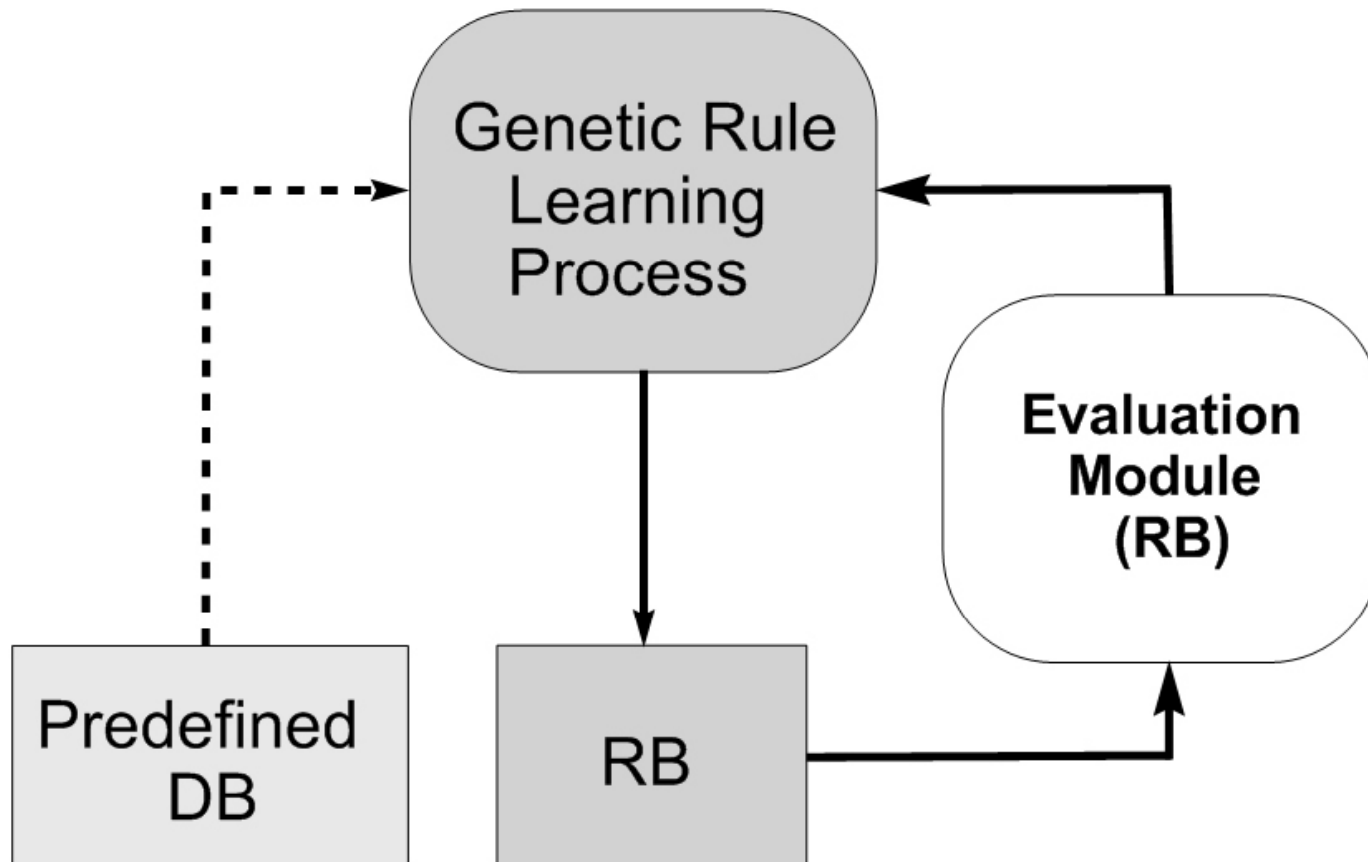
$X_2 \backslash X_1$	P	M	G
P		$S_1$ $B_1$	
M	$S_2$ $B_2$	$S_3$ $B_2$	
G			$S_4$ $B_3$



Rule Base

$R_1 = \text{IF } X_1 \text{ is M and } X_2 \text{ is P THEN } Y \text{ is } B_1$   
 $R_2 = \text{IF } X_1 \text{ is P and } X_2 \text{ is M THEN } Y \text{ is } B_2$   
 $R_3 = \text{IF } X_1 \text{ is M and } X_2 \text{ is M THEN } Y \text{ is } B_2$   
 $R_4 = \text{IF } X_1 \text{ is G and } X_2 \text{ is G THEN } Y \text{ is } B_3$

# 1. Brief introduction to genetic fuzzy systems

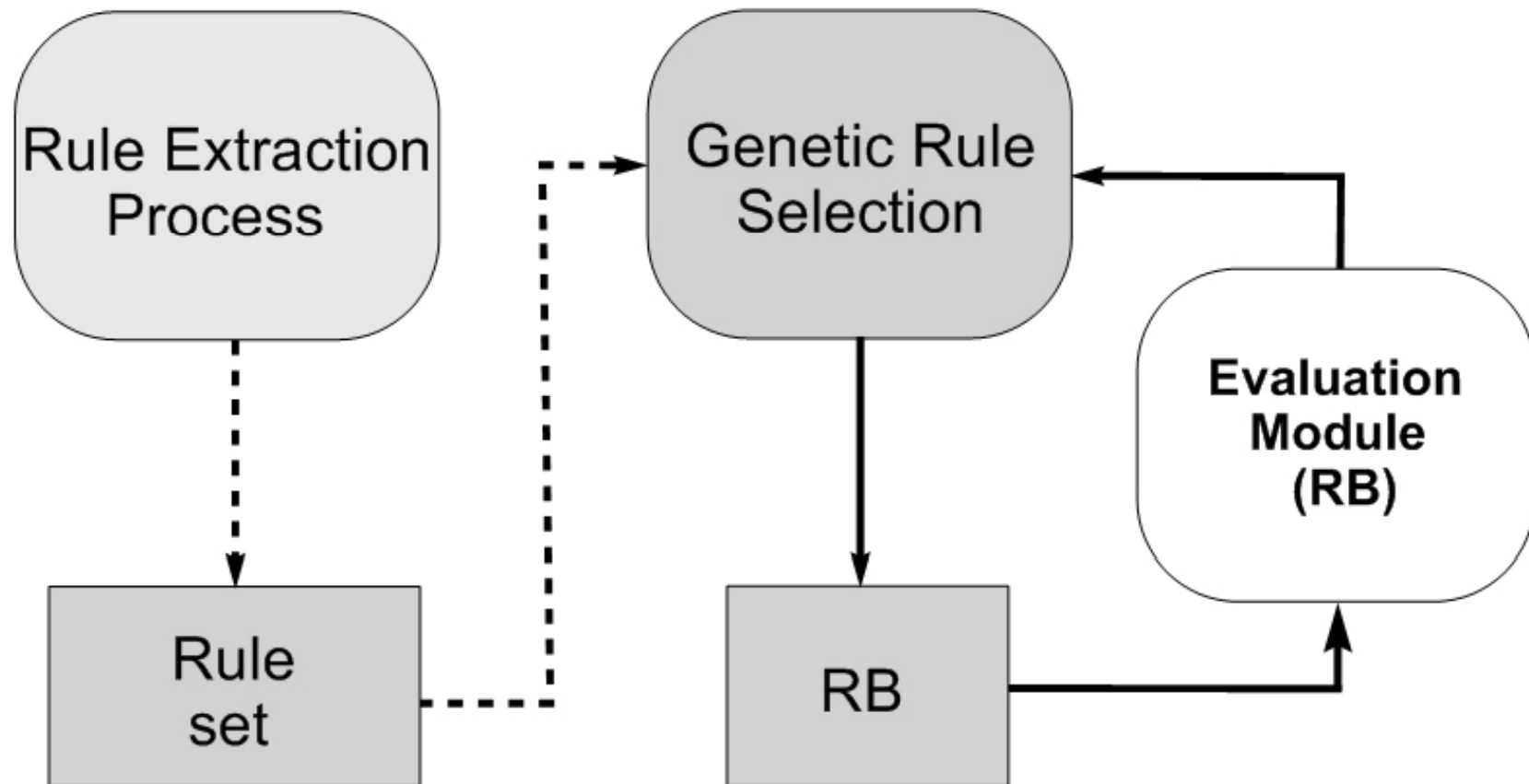


# 1. Brief introduction to genetic fuzzy systems

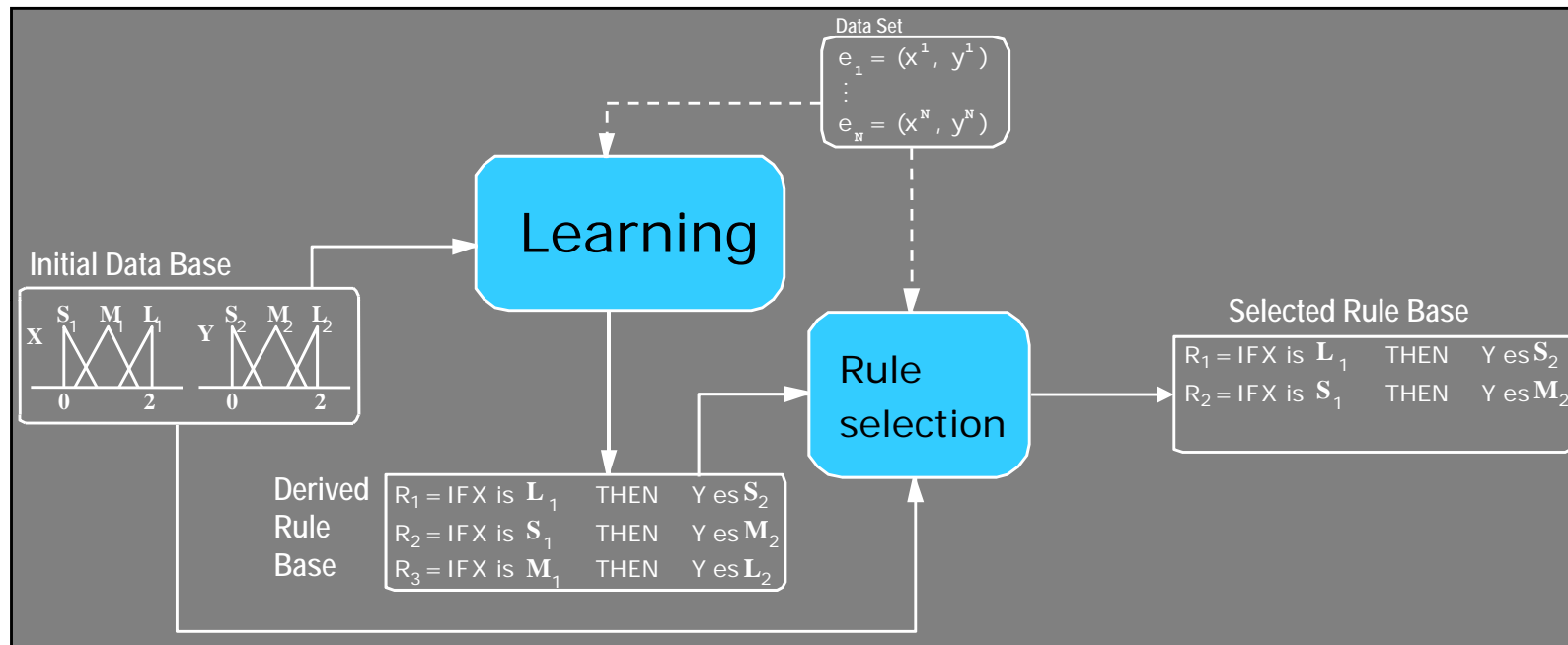
## 3. Genetic Rule Selection

- A predefined Rule Bases definition is assumed
- The fuzzy rules **are selection** by a GA for getting a compact rule base (more interpretable, more precise)

# 1. Brief introduction to genetic fuzzy systems



# 1. Brief introduction to genetic fuzzy systems

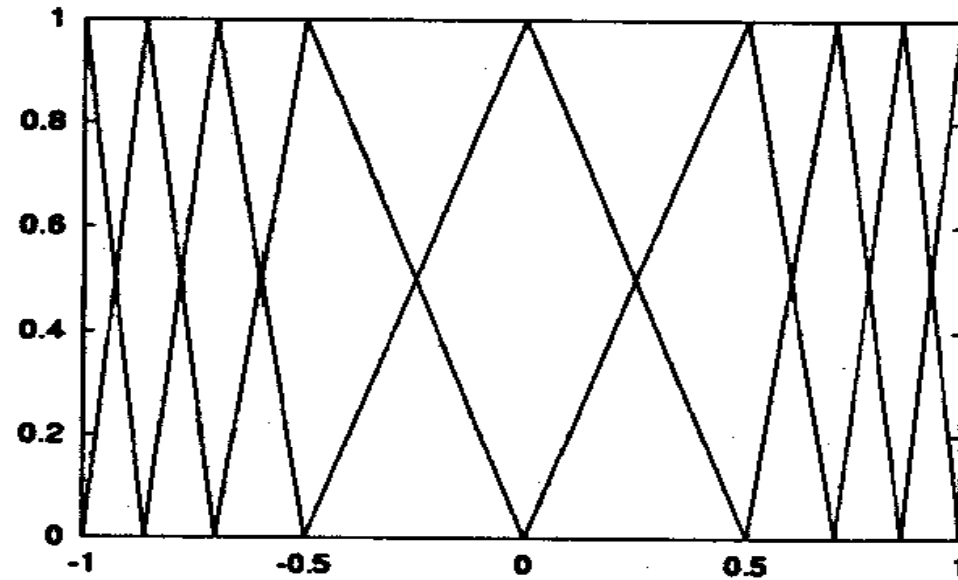
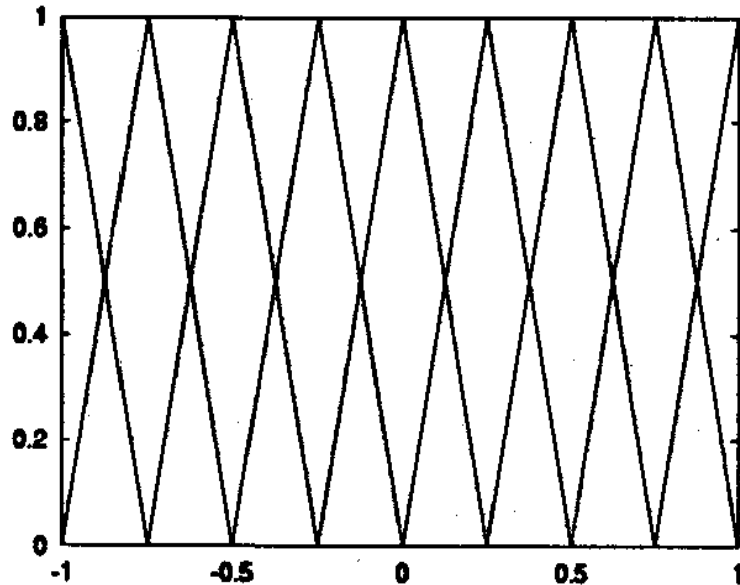


## Example of genetic rule selection

# 1. Brief introduction to genetic fuzzy systems

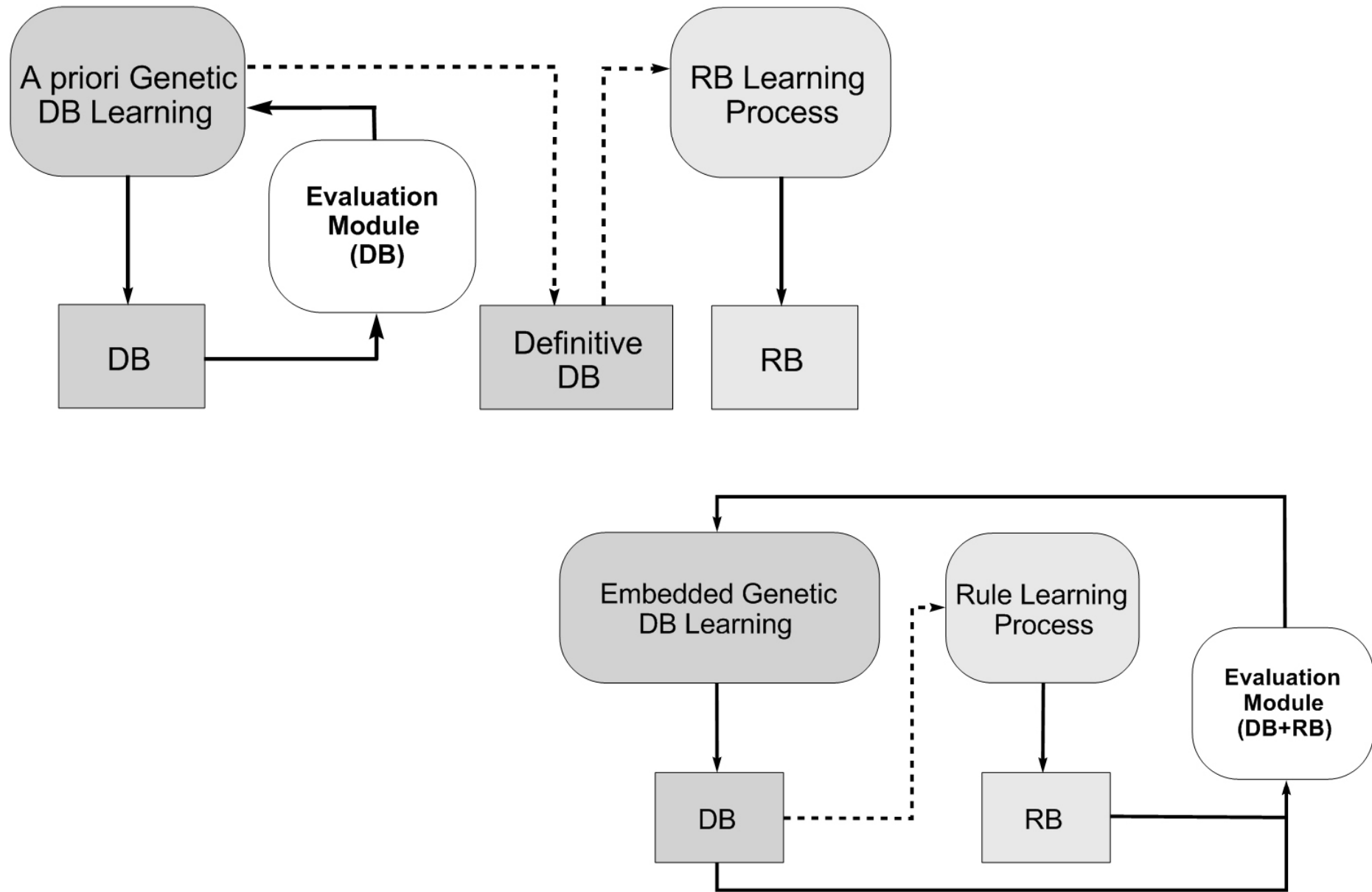
## 4. Genetic DB Learning

- Learning of the membership function shapes by a GA





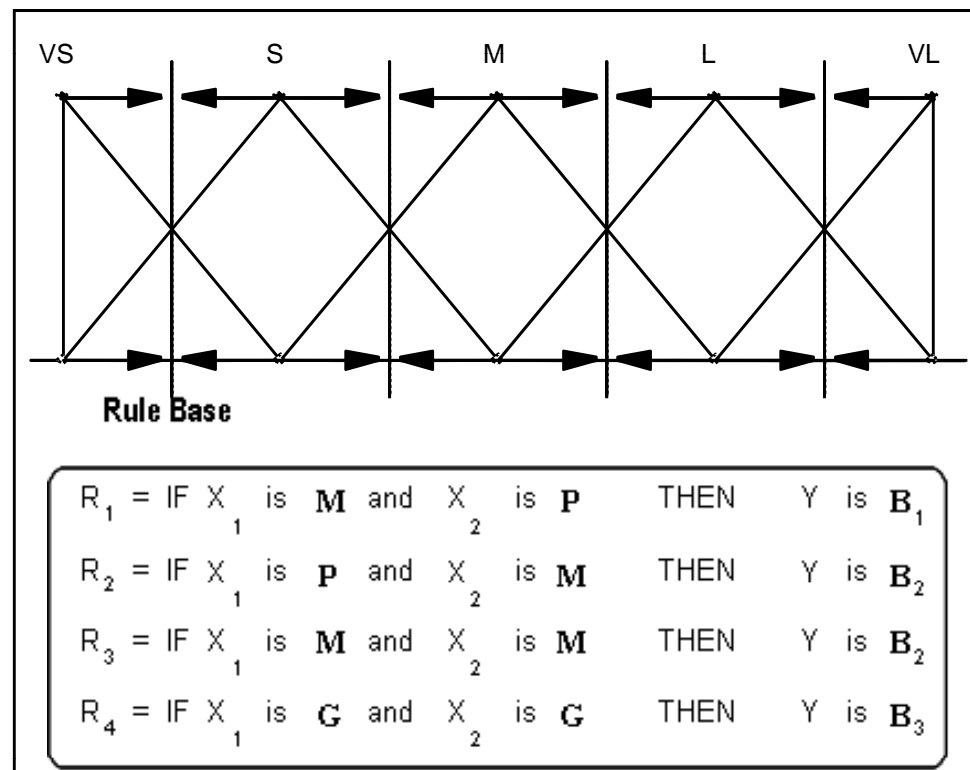
# 1. Brief introduction to genetic fuzzy systems



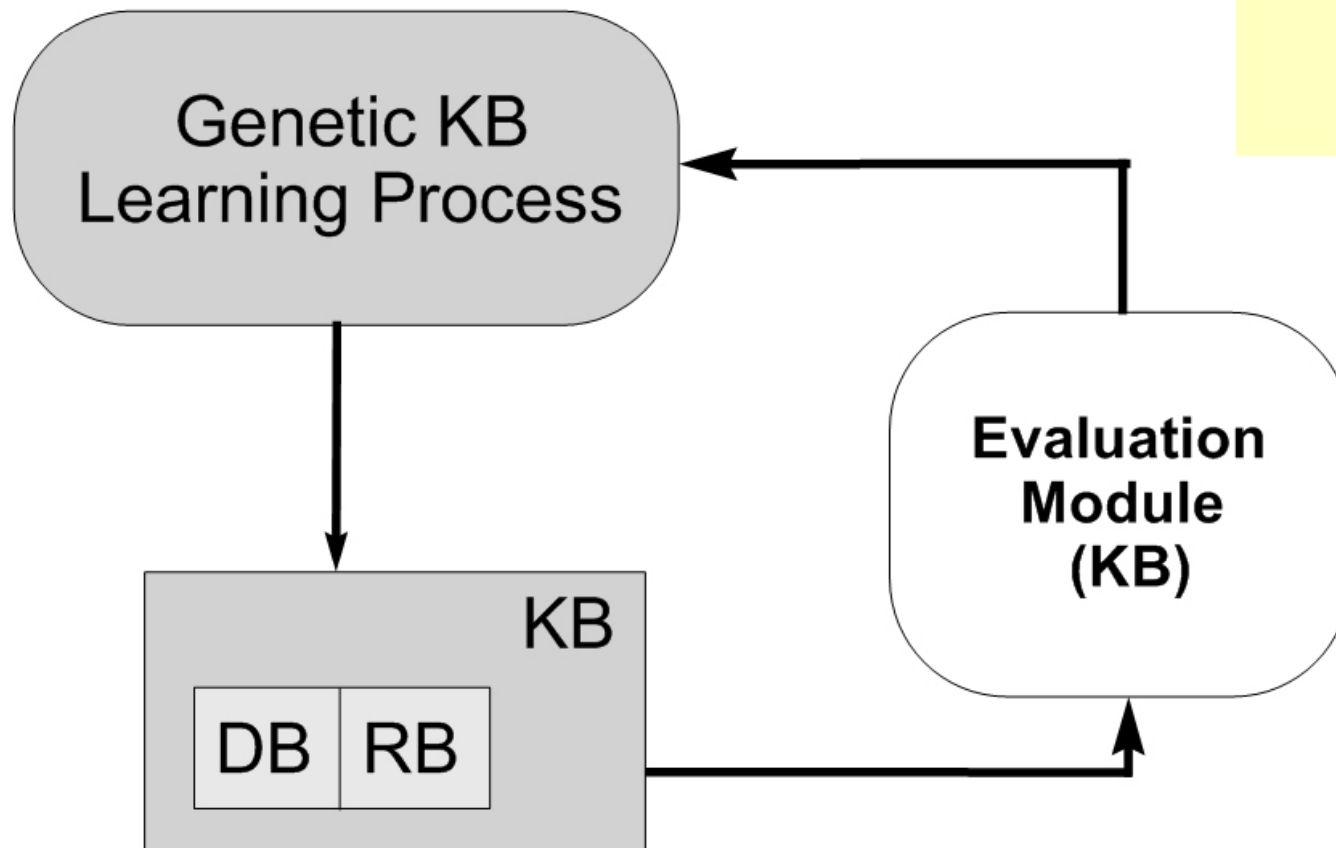
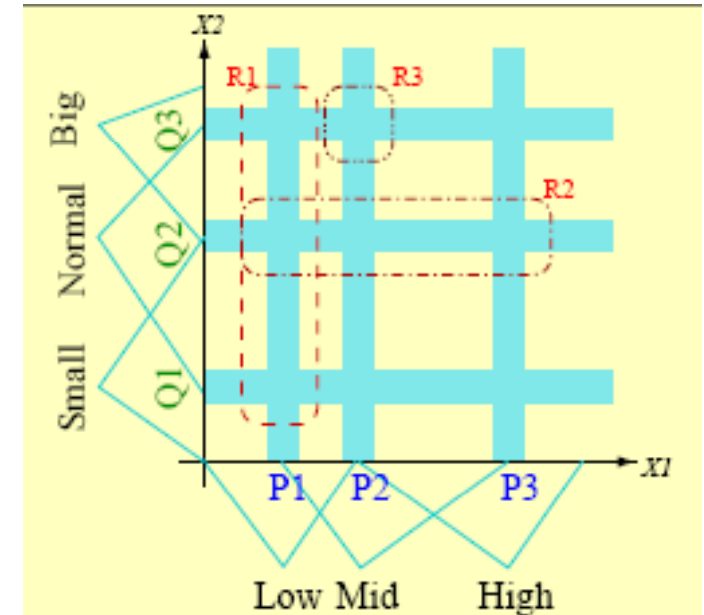
# 1. Brief introduction to genetic fuzzy systems

## 5. Simultaneous Genetic Learning of KB Components

- The simultaneous derivation properly addresses the **strong dependency** existing between the RB and the DB

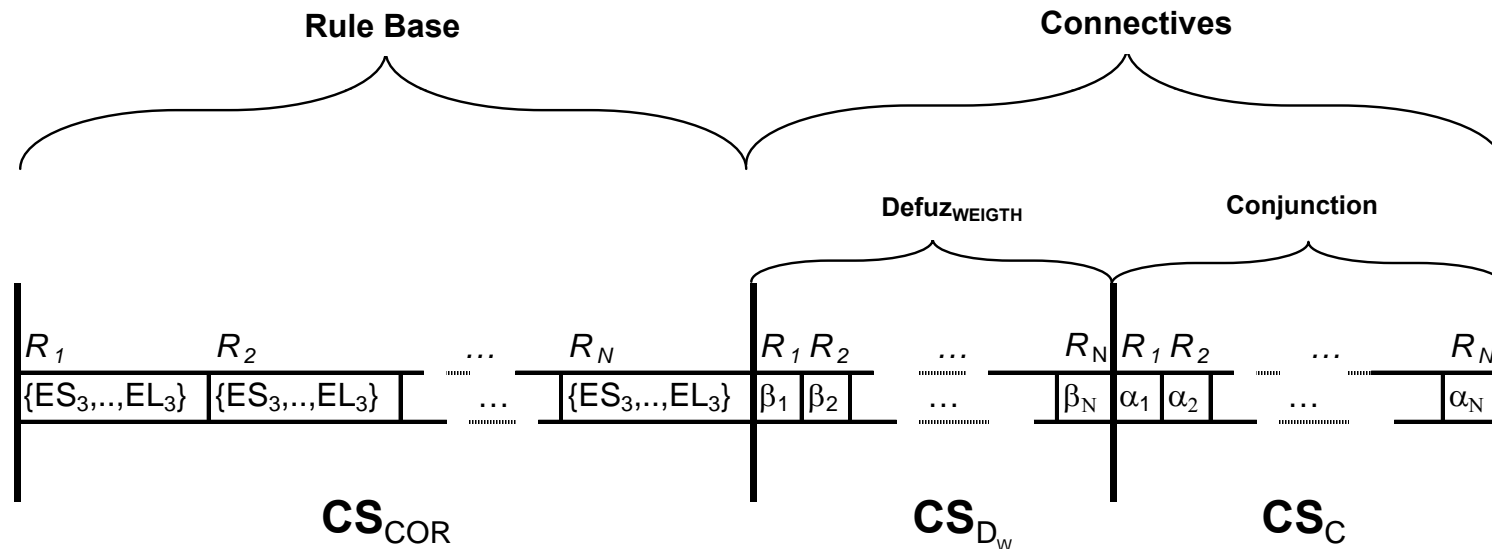


# 1. Brief introduction to genetic fuzzy systems



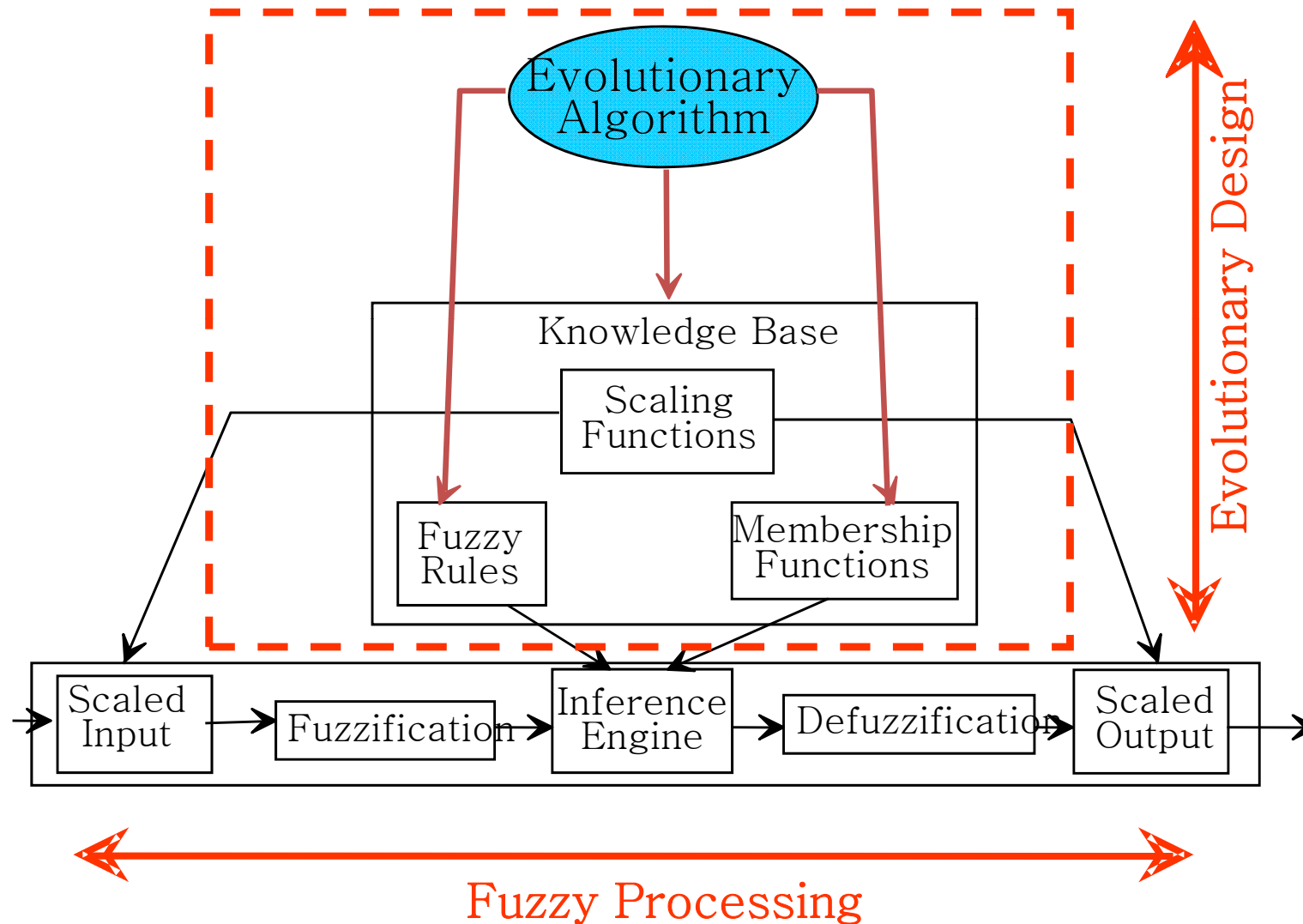
# 1. Brief introduction to genetic fuzzy systems

## 6. Genetic Learning of KB Components and Inference Engine Parameters



Example of the coding scheme for learning an RB and the inference  
connective parameters

# 1. Brief introduction to genetic fuzzy systems

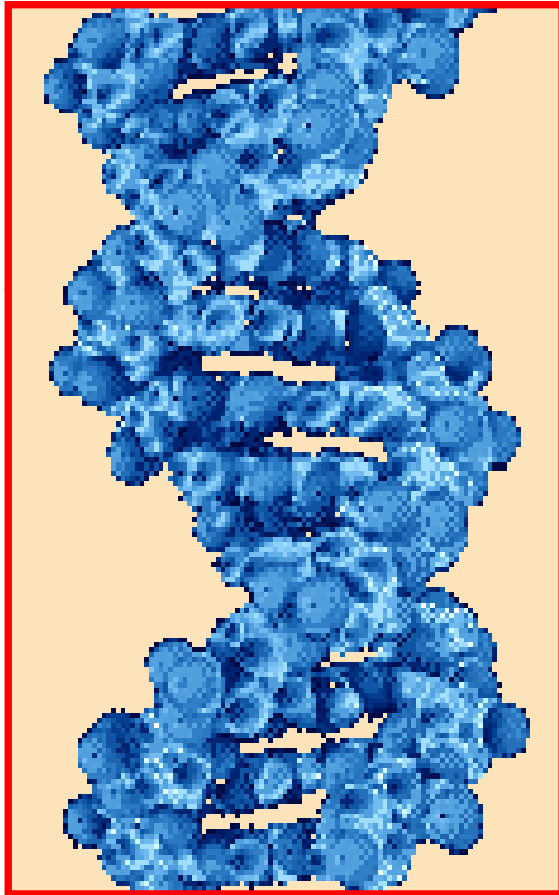


# 1. Introduction to genetic fuzzy systems

---

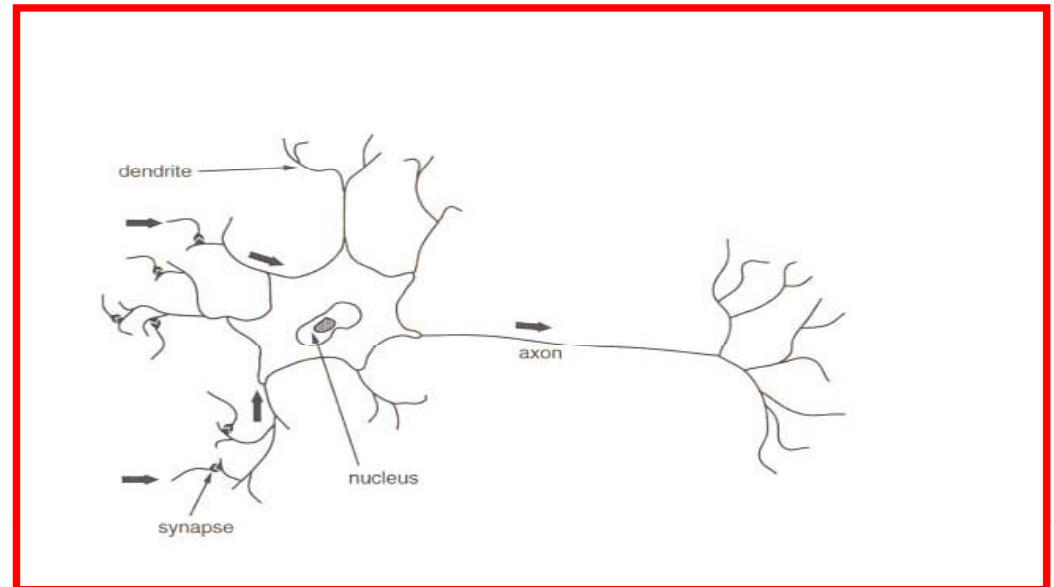
- **Brief Introduction**
- **Taxonomy of Genetic Fuzzy Systems**
- **¿Why do we use GAs? GFSs versus Neural Fuzzy Systems**
- **The birth, GFSs roadmap, current state and most cited papers**

# 1. Brief introduction to genetic fuzzy systems



¿Why do we use GAs?

**GFSs versus Neural Fuzzy Systems**



# 1. Brief introduction to genetic fuzzy systems

## ¿Why do we use GAs? GFSs versus Neural Fuzzy Systems

### Neuro Fuzzy Systems

- The most usual architecture:

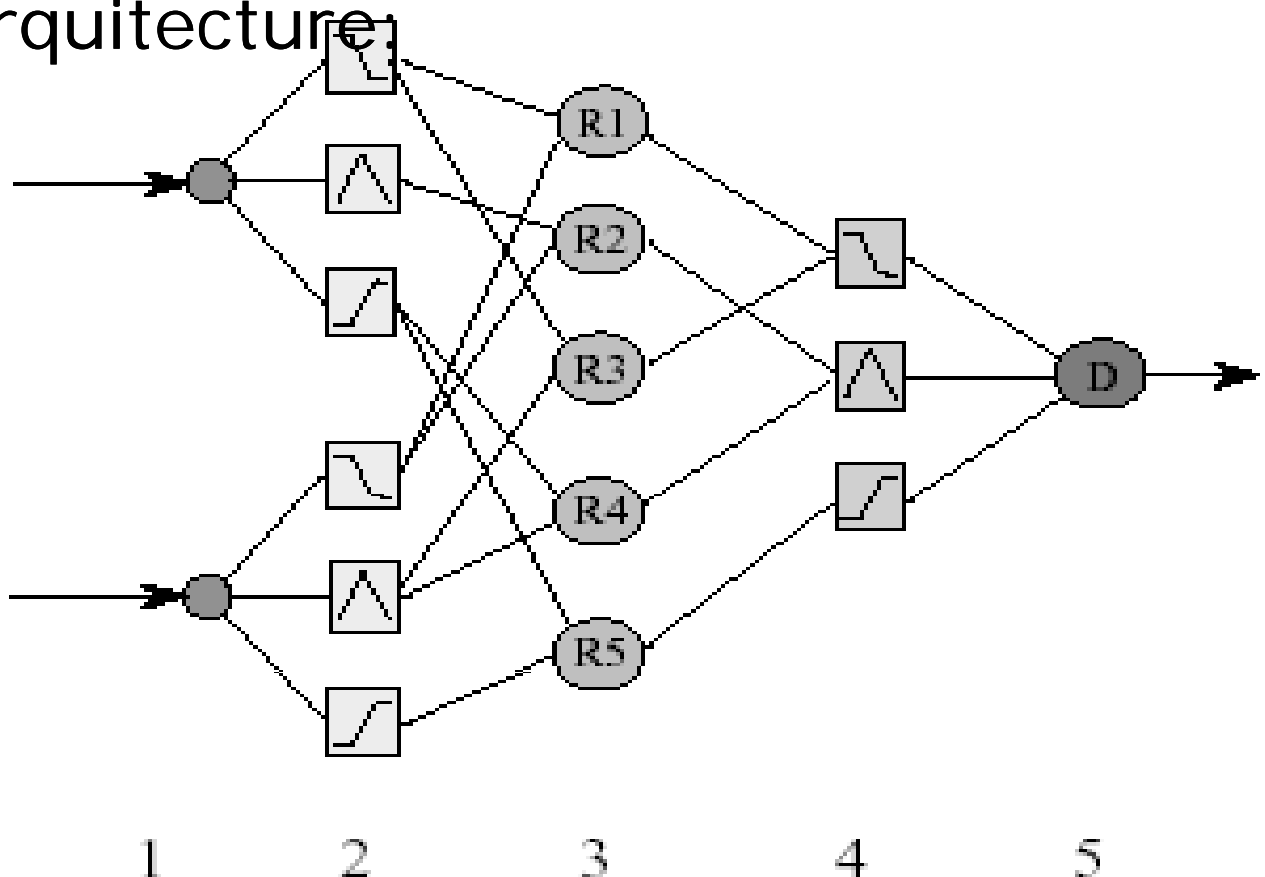
1. Variables

2. Fuzzification  
(Fuzzy Partition,  
Data Base)

3. Rules

4. Consequents

5. Defuzzification



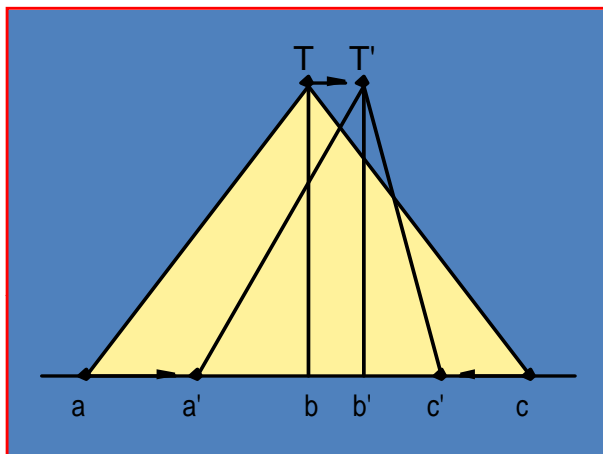


# 1. Brief introduction to genetic fuzzy systems

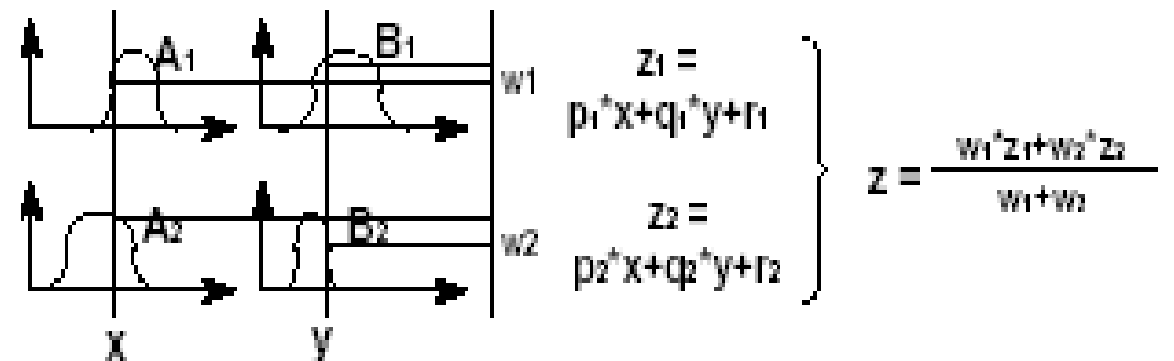
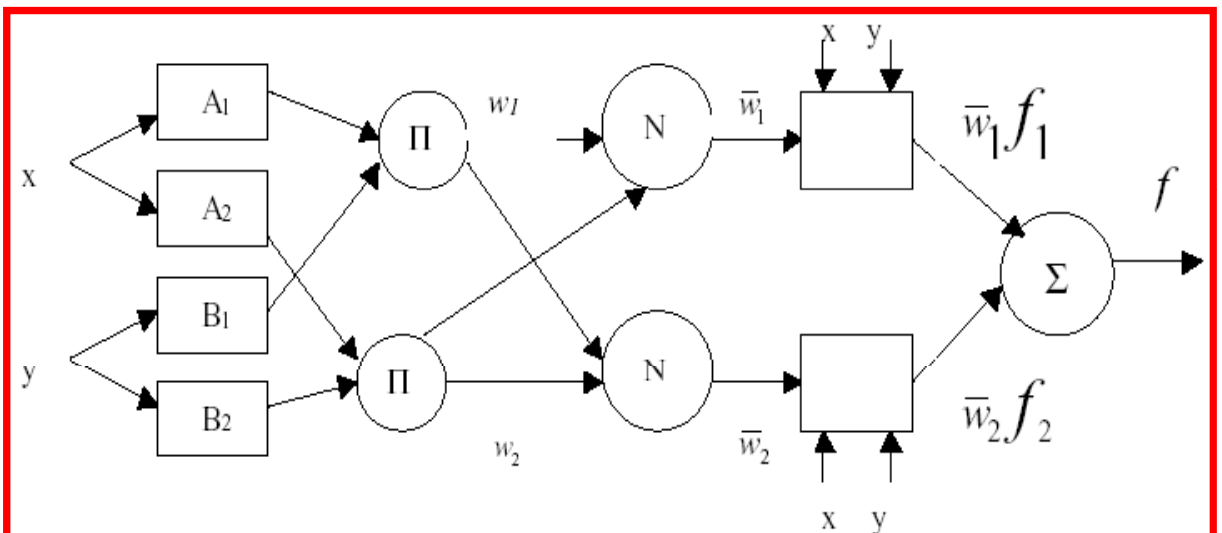
## ¿Why do we use GAs? GFSs versus Neural Fuzzy Systems

■ **ANFIS: Adaptive Network based Fuzzy Inference System**  
(Jyh-Shing Roger Jang, 1993)

- It uses a fixed number of linguistic labels per variable
- It only tunes the membership functions



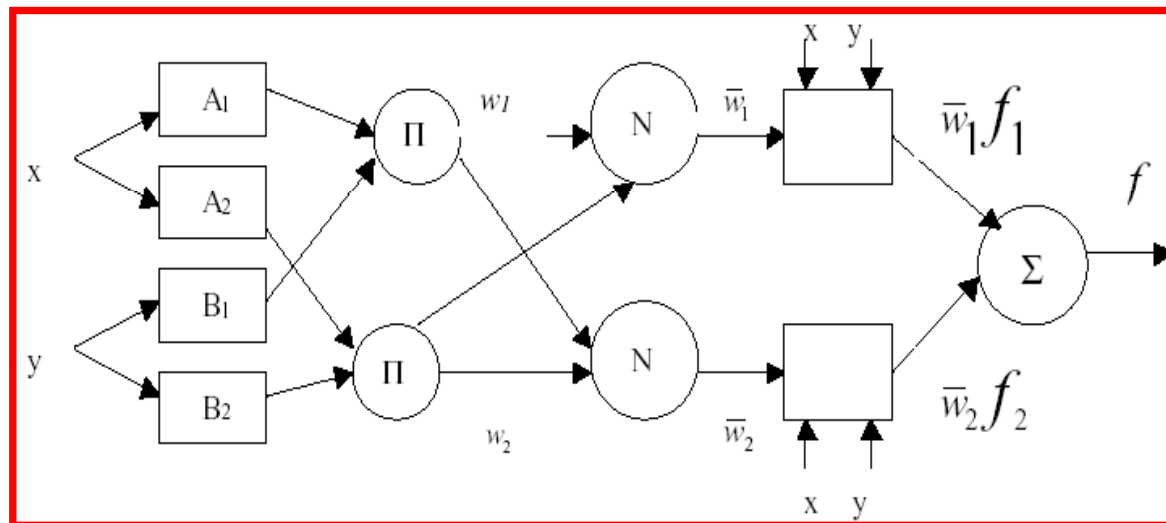
### Neuro Fuzzy Systems (ANFIS)



# 1. Brief introduction to genetic fuzzy systems

## ¿Why do we use GAs? GFSs versus Neural Fuzzy Systems Limitations of the Neuro Fuzzy Systems

- **Dimensionality problem:** They can manage a small number of variables (the complexity increases geometrically with the number of variables)
- **A necessity:** To know previously the number of labels per variable.
- **Difficulty for learning the rule structure:** Usually, NFS only learn the membership functions and rule consequent coefficients.

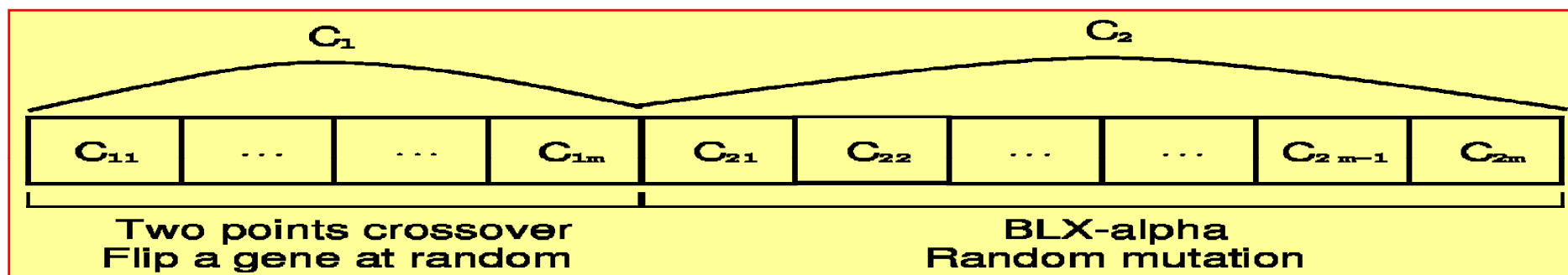


# 1. Brief introduction to genetic fuzzy systems

## ¿Why do we use GAs? GFSs versus Neural Fuzzy Systems Advantages of the Genetic Fuzzy Systems

- We can code different FS components in a chromosome:
  - Identify relevant inputs
  - Scaling factors
  - Membership functions, shape functions, optimal shape of membership funct., granularity (number of labels per variable)
  - Fuzzy rules, Any inference parameter, ....

We can define different mechanism for managing them  
(combining genetic operators, coevolution,...)

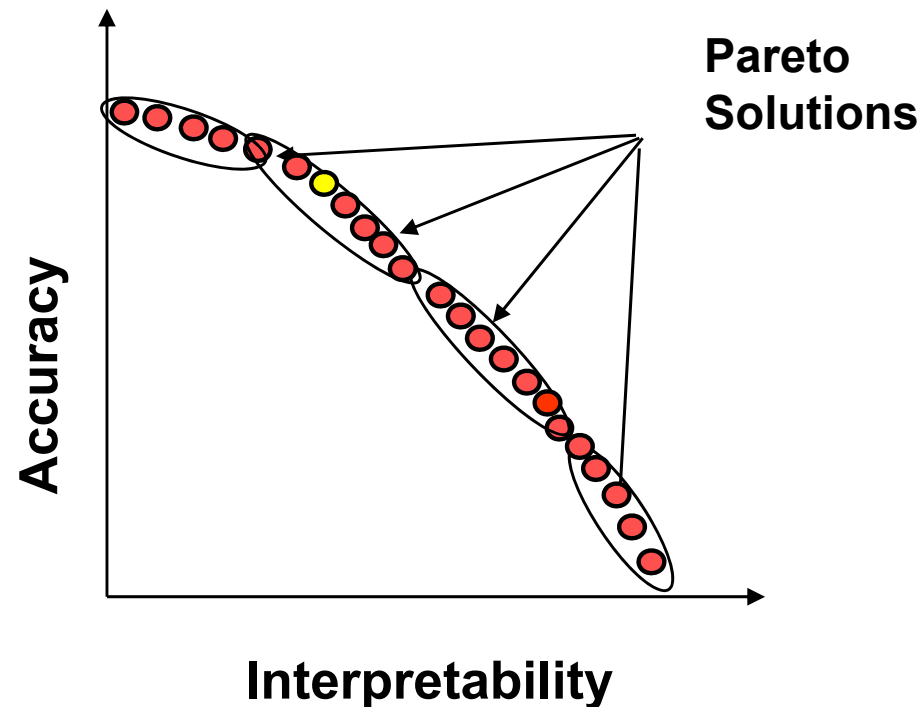


# 1. Brief introduction to genetic fuzzy systems

## ¿Why do we use GAs? GFSs versus Neural Fuzzy Systems

### Advantages of the Genetic Fuzzy Systems

- We can consider multiple objectives in the learning model (interpretability, precision, ....)



# 1. Introduction to genetic fuzzy systems

---

- **Brief Introduction**
- **Taxonomy of Genetic Fuzzy Systems**
- **¿Why do we use GAs? GFSs versus Neural Fuzzy Systems**
- **The birth, GFSs roadmap, current state and most cited papers**

# 1. Introduction to genetic fuzzy systems

The birth, GFSs roadmap, current status and most cited papers

## The birth of GFSs: 1991

- Thrift's ICGA91 paper (Mamdani-type Rule Base Learning. Pittsburgh approach)
- Thrift P (1991) **Fuzzy logic synthesis with genetic algorithms**. *In: Proc. of 4th International Conference on Genetic Algorithms (ICGA'91), pp 509-513*
- Valenzuela-Rendón's PPSN-I paper (Scatter Mamdani-type KB Learning. Michigan approach)
- Valenzuela-Rendon M (1991) **The fuzzy classifier system: A classifier system for continuously varying variables**. *In: Proc. of 4th International Conference on Genetic Algorithms (ICGA'91), pp 346-353*
- Pham and Karaboga's Journal of Systems Engineering paper (Relational matrix-based FRBS learning. Pittsburgh approach)
- Pham DT, Karaboga D (1991) **Optimum design of fuzzy logic controllers using genetic algorithms**. *Journal of Systems Engineering 1: 114-118*).
- Karr's AI Expert paper (Mamdani-type Data Base Tuning)
- Karr C (1991) **Genetic algorithms for fuzzy controllers**. *AI Expert 6(2):26-33*.

Almost the whole basis of the area were established in the first year!

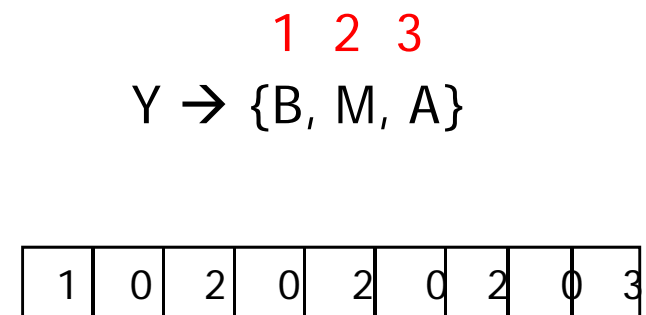
# 1. Introduction to genetic fuzzy systems

## Thrift's GFS:

P. Thrift, Fuzzy logic synthesis with genetic algorithms, Proc. Fourth Intl. Conf. on Genetic Algorithms (ICGA'91), San Diego, USA, 1991, pp. 509–513

- Classical approach: Pittsburgh – the decision table is encoded in a rule consequent array
- The output variable linguistic terms are numbered from 1 to n and comprise the array values. The value 0 represents the rule absence, thus making the GA able to learn the optimal number of rules
- The ordered structure allows the GA to use simple genetic operators

$X_2 \backslash X_1$	S	M	L
S	R <sub>1</sub> B	R <sub>2</sub> —	R <sub>3</sub> M
M	R <sub>4</sub> —	R <sub>5</sub> M	R <sub>6</sub> —
L	R <sub>7</sub> M	R <sub>8</sub> —	R <sub>9</sub> A



# 1. Introduction to genetic fuzzy systems

The birth, GFSs roadmap, current status and most cited papers

## GFSs roadmap

### **1991-1996/7: INITIAL GFS SETTING: KB LEARNING:**

- Establishment of the three classical learning approaches in the GFS field: Michigan, Pittsburgh, and IRL
- Different FRBS types: Mamdani, Mamdani DNF, Scatter Mamdani, TSK
- Generic applications: Classification, Modeling, and Control

### **1995-...: FUZZY SYSTEM TUNING:**

- First: Membership function parameter tuning
- Later: other DB components adaptation: scaling factors, context adaptation (scaling functions), linguistic hedges, ...
- Recently: **interpretability consideration**



# 1. Introduction to genetic fuzzy systems

The birth, GFSs roadmap, current status and most cited papers

## GFSs roadmap

### **1998-....: APPROACHING TO MATURITY? NEW GFS LEARNING APPROACHES:**

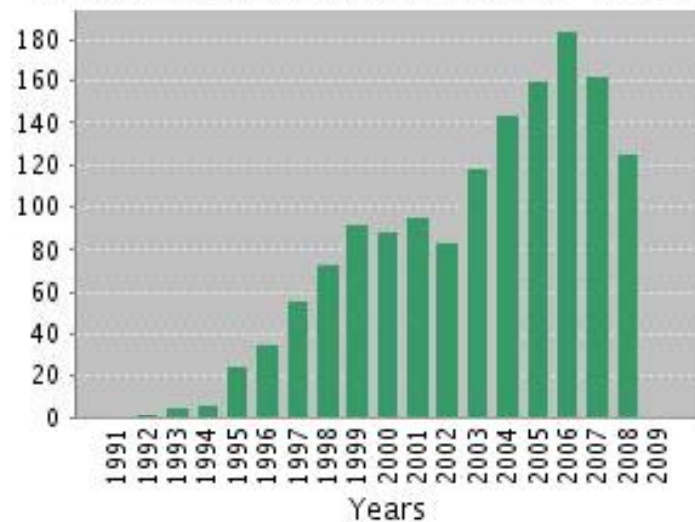
- New EAs: Bacterial genetics, DNA coding, Virus-EA, genetic local search (memetic algorithms), ...
- Hybrid learning approaches: a priori DB learning, GFNNs, Michigan-Pitt hybrids, ...
- Multiobjective evolutionary algorithms
- **Interpretability-accuracy trade-off** consideration
- Course of dimensionality (handling large data sets and complex problems):
  - Rule selection (1995-...)
  - Feature selection at global level and fuzzy rule level
  - Hierarchical fuzzy modeling
- “Incremental” learning

# 1. Brief introduction to genetic fuzzy systems

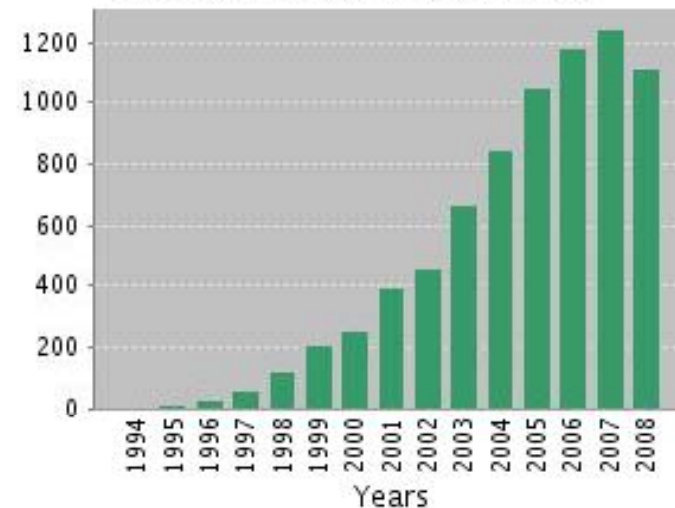
Current state of the GFS area

## Number of papers on GFSs published in JCR journals

Published Items in Each Year



Citations in Each Year



**Source:** The Thomson Corporation ISI Web of Knowledge

**Query:** (TS = (("GA-" OR "GA based" OR evolutionary OR "genetic algorithm\*" OR "genetic programming" OR "evolution strate\*" OR "genetic learning" OR "particle swarm" OR "differential evolutio\*" OR "ant system\*" OR "ant colony" OR "genetic optimi\*" OR "estimation of distribution algorithm\*") AND ("fuzzy rule\*" OR "fuzzy system\*" OR "fuzzy neural" OR "neuro-fuzzy" OR "fuzzy control\*" OR "fuzzy logic cont\*" OR "fuzzy class\*" OR "fuzzy if" OR "fuzzy model\*" OR "fuzzy association rule\*" OR "fuzzy regression")))

**Date:** October 15, 2008

**Number of papers:** 1459

41

**Number of citations:** 5,237,630 **Average citations per paper:** 5.23

# 1. Brief introduction to genetic fuzzy systems

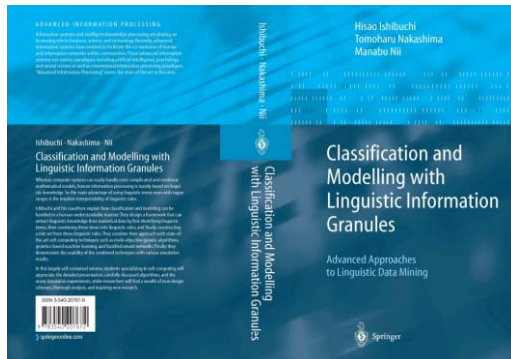
## Current state of the GFS area

### Most cited papers on GFSs

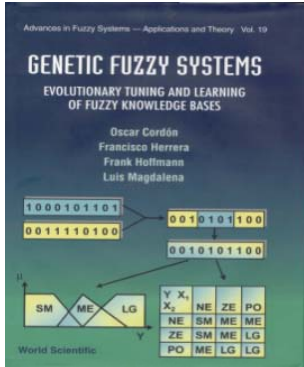
1. Homaifar, A., McCormick, E., Simultaneous Design of Membership Functions and rule sets for fuzzy controllers using genetic algorithms, IEEE TFS 3 (2) (1995) 129-139. **Citations: 184**
2. Ishibuchi, H., Nozaki, K., Yamamoto, N., Tanaka, H., Selecting fuzzy if-then rules for classification problems using genetic algorithms, IEEE TFS 3 (3) (1995) 260-270. **Citations: 164**
3. Setnes, M., Roubos, H., GA-fuzzy modeling and classification: complexity and performance, IEEE TFS 8 (5) (2000) 509-522 . **Citations: 101**
4. Ishibuchi, H., Nakashima, T., Murata, T., Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems, IEEE TSMC B 29 (5) (1999) 601-618. **Citations: 93**
5. Park, D., Kandel, A., Langholz, G., Genetic-based new fuzzy reasoning models with application to fuzzy control, IEEE TSMC B 24 (1) (1994) 39-47. **Citations: 86**
6. Herrera, F., Lozano, M., Verdegay, J.L., Tuning fuzzy-logic controllers by genetic algorithms, IJAR 12 (3-4) (1995) 299-315. **Citations: 71**
7. Shi, Y.H., Eberhart, R., Chen, Y.B., Implementation of evolutionary fuzzy systems, IEEE TFS 7 (2) (1999) 109-119. **Citations: 63**
8. Carse B., Fogarty, TC., Munro, A., Evolving fuzzy rule based controllers using genetic algorithms, FSS 80 (3) (1996) 273-293. **Citations: 63**
9. Juang, C.F., Lin, J.Y., Lin, C.T., Genetic reinforcement learning through symbiotic evolution for fuzzy controller design, IEEE TSMC B 30 (2) (2000) 290-302. **Citations: 59**
10. Cordon, O., Herrera, F., A three-stage evolutionary process for learning descriptive and approximate fuzzy-logic-controller knowledge bases from examples, IJAR 17 (4) (1997) 369-407. **Citations: 58**

# 1. Brief introduction to genetic fuzzy systems

## Some references



**GENETIC FUZZY SYSTEMS**  
**Evolutionary Tuning and Learning of Fuzzy Knowledge Bases.**  
O. Cordón, F. Herrera, F. Hoffmann, L. Magdalena  
World Scientific, July 2001



**H. Ishibuchi, T. Nakashima, M. Nii, Classification and Modeling with Linguistic Information Granules. Advanced Approaches to Linguistic Data Mining. Springer (2005)**

- F. Herrera, Genetic Fuzzy Systems: Taxonomy, Current Research Trends and Prospects. *Evolutionary Intelligence* 1 (2008) 27-46 doi: 10.1007/s12065-007-0001-5,
- F. Herrera, Genetic Fuzzy Systems: Status, Critical Considerations and Future Directions, *International Journal of Computational Intelligence Research* 1 (1) (2005) 59-67
- O. Cordón, F. Gomide, F. Herrera, F. Hoffmann, L. Magdalena, Ten Years of Genetic Fuzzy Systems: Current Framework and New Trends, *FSS* 141 (1) (2004) 5-31
- F. Hoffmann, Evolutionary Algorithms for Fuzzy Control System Design, *Proceedings of the IEEE* 89 (9) (2001) 1318-1333



# Genetic Fuzzy Systems: State of the Art and New Trends

## Outline

- ✓ Brief Introduction to Genetic Fuzzy Systems
- ✓ Genetic Tuning Methods: Basic and Advanced Approaches
- ✓ Genetic Fuzzy Systems Application to HVAC Problems
- ✓ GFSs: Current Trends and Prospects
- ✓ Concluding Remarks

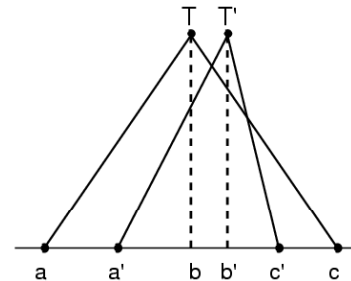
## 2. Evolutionary Tuning of FRBSs

### **Tuning of membership functions**

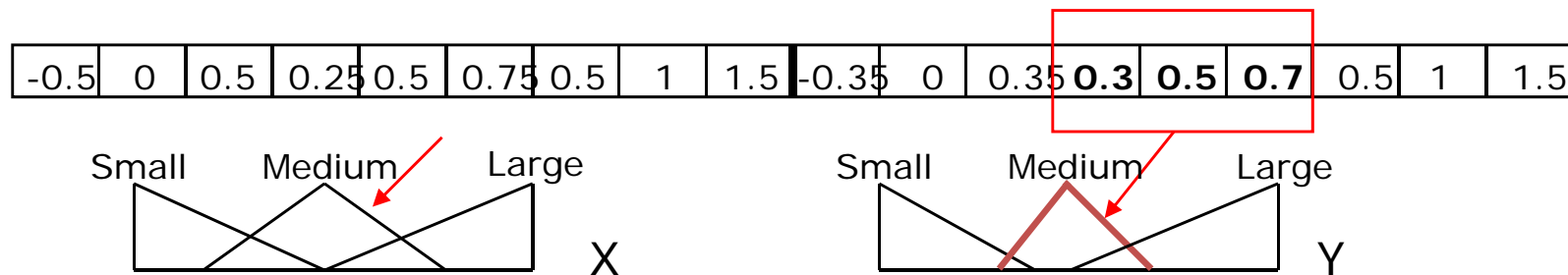
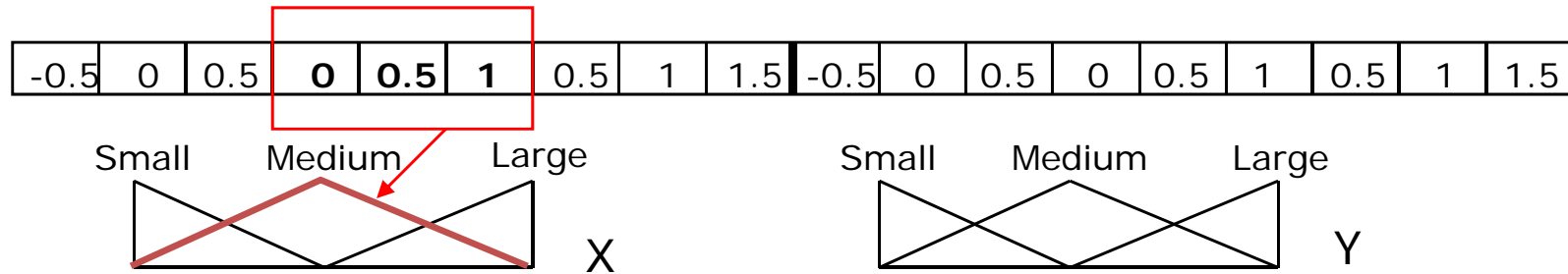
- A genetic tuning process that slightly adjusts the shapes of the membership functions of a preliminary DB definition
- Each chromosome encodes a whole DB definition by joining the partial coding of the different membership functions involved
- The coding scheme depends on:
  - The kind of membership function considered (triangular, trapezoidal, bell-shaped, ...) → different real-coded definition parameters
  - The kind of FRBS:
    - Grid-based: Each linguistic term in the fuzzy partition has a single fuzzy set definition associated
    - Non grid-based (free semantics, scatter partitions, fuzzy graphs): each variable in each rule has a different membership function definition

## 2. Evolutionary Tuning of FRBSs

- **Example:** Tuning of the triangular membership functions of a grid-based SISO Mamdani-type FRBS, with three linguistic terms for each variable fuzzy partition
- Each chromosome encodes a different DB definition:
  - $2$  (variables)  $\cdot$   $3$  (linguistic labels) =  $6$  membership functions
  - Each triangular membership function is encoded by  $3$  real values (the three definition points):
  - So, the chromosome length is  $6 \cdot 3 = 18$  real-coded genes (binary coding can be used but is not desirable)
- Either **definition intervals** have to be defined for each gene and/or appropriate genetic operators in order to obtain meaningful membership functions



## 2. Evolutionary Tuning of FRBSs



The RB remains unchanged!

R1: IF X1 is Small THEN Y is Large  
 R2: IF X1 is Medium THEN Y is Med  
 ...



## 2. Evolutionary Tuning of FRBSs

### References:

- C. Karr, Genetic algorithms for fuzzy controllers, *AI Expert* 6 (2) (1991) 26–33
- C. Karr, E.J. Gentry, Fuzzy control of pH using genetic algorithms, *IEEE TFSs* 1 (1) (1993) 46–53
- J. Kinzel, F. Klawonn, R. Kruse, Modifications of genetic algorithms for designing and optimizing fuzzy controllers, *Proc. First IEEE Conf. on Evolutionary Computation (ICEC'94)*, Orlando, FL, USA, 1994, pp. 28–33
- D. Park, A. Kandel, G. Langholz, Genetic-based new fuzzy reasoning models with application to fuzzy control, *IEEE TSMC* 24 (1) (1994) 39–47
- F. Herrera, M. Lozano, J.L. Verdegay, Tuning fuzzy controllers by genetic algorithms, *IJAR* 12 (1995) 299–315
- P.P. Bonissone, P.S. Khedkar, Y. Chen, Genetic algorithms for automated tuning of fuzzy controllers: a transportation application, in *Proc. Fifth IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE'96)*, New Orleans, USA, 1996, pp. 674–680
- O. Cordón, F. Herrera, A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples, *IJAR* 17 (4) (1997) 369–407
- H.B. Gurocak, A genetic-algorithm-based method for tuning fuzzy logic controllers, *FSS* 108 (1) (1999) 39–47

## 2. Evolutionary Tuning of FRBSs

### Genetic tuning of DB and RB using linguistic hedges

J. Casillas, O. Cordón, M.J. del Jesus, F. Herrera, Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction, *IEEE TFS* 13 (1) (2005) 13-29

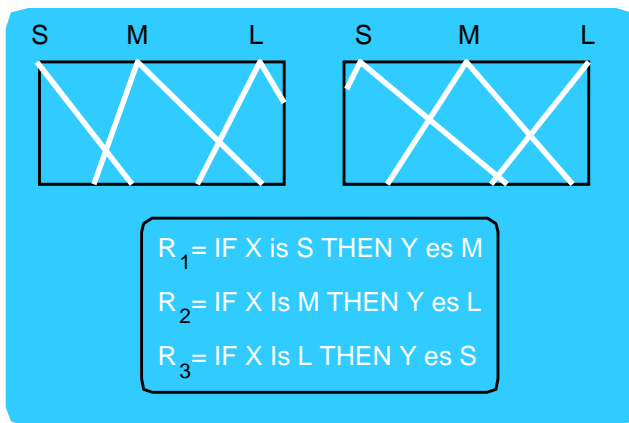
Genetic tuning process that refines a preliminary KB working at two different levels:

- **DB level:** Linearly or non-linearly adjusting the membership function shapes
- **RB level:** Extending the fuzzy rule structure using automatically learnt linguistic hedges

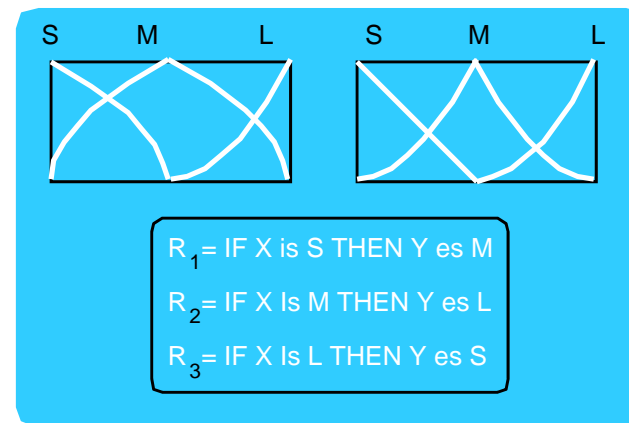
## 2. Evolutionary Tuning of FRBSs

- Tuning of the DB:

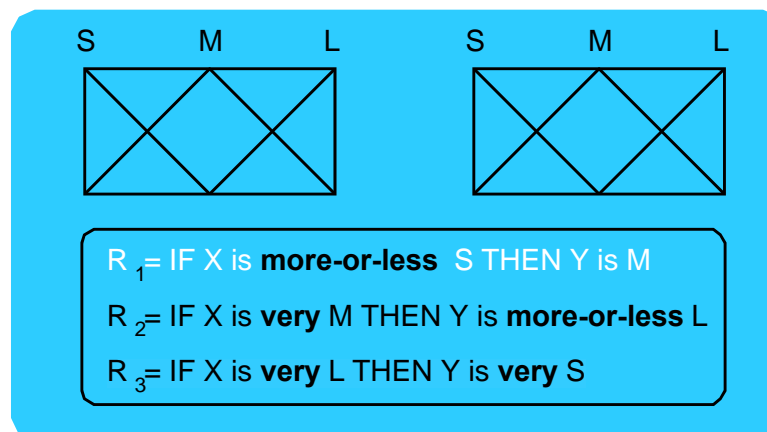
Linear tuning



Non-linear tuning



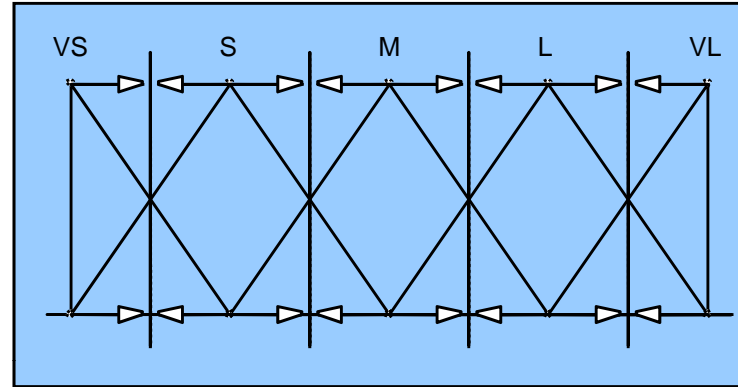
- Tuning of the RB: **linguistic hedges 'very' and 'more-or-less'**



## 2. Evolutionary Tuning of FRBSs

Triple coding scheme:

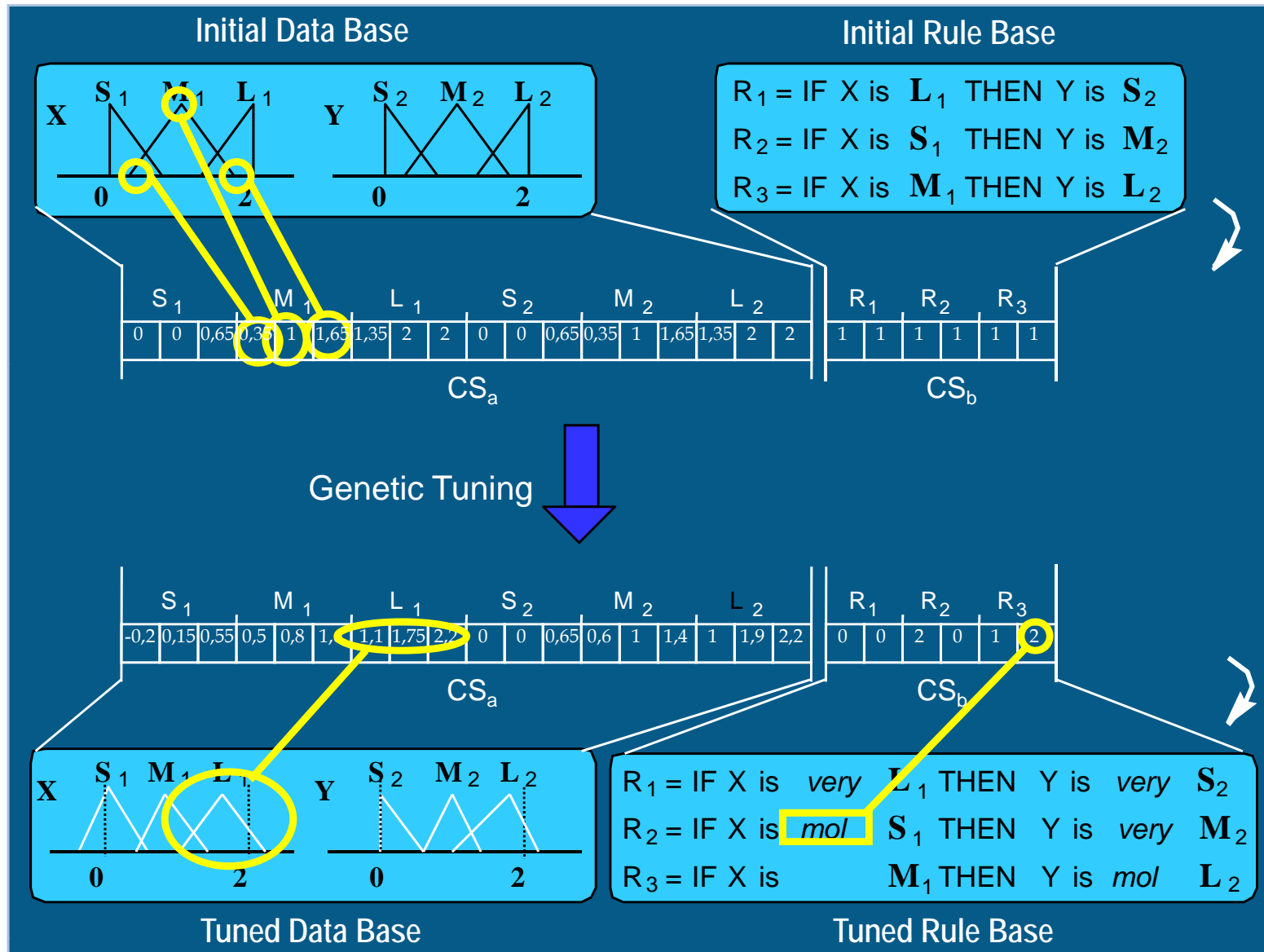
- Membership function parameters (**P**) (DB linear tuning): **real coding**
- Alpha values (**A**) (DB non linear tuning): **real coding**
- Linguistic hedges (**L**) (RB tuning): **integer coding**



$$\alpha = \begin{cases} 1 + c_{ij}^A, & \text{si } c_{ij}^A \in [-1,0] \\ 1 + 4 \cdot c_{ij}^A, & \text{si } c_{ij}^A \in ]0,1] \end{cases}$$

$c_{ij} = 0$	$\leftrightarrow$	'very'
$c_{ij} = 1$	$\leftrightarrow$	no hedge
$c_{ij} = 2$	$\leftrightarrow$	'more-or-less'

## 2. Evolutionary Tuning of FRBSs



## 2. Evolutionary Tuning of FRBSs

### Experimental study for the medium voltage line problem:

- Learning method considered: Wang-Mendel
- Tuning method variants:

TUNING PROCESSES CONSIDERED IN THIS EXPERIMENTAL STUDY

Method	Basic m.f. parameters	$\alpha$ m.f. parameter	Surface structure with linguistic hedges
P-tun	✓		
A-tun		✓	
L-tun			✓
PA-tun	✓	✓	
PL-tun	✓		✓
AL-tun		✓	✓
PAL-tun	✓	✓	✓

- Evaluation methodology: 5 random training-test partitions 80-20% (5-fold cross validation)  $\times$  6 runs = 30 runs per algorithm

## 2. Evolutionary Tuning of FRBSs

### Maintenance cost estimation for low and medium voltage lines in Spain:

O. Cerdón, F. Herrera, L. Sánchez, Solving electrical distribution problems using hybrid evolutionary data analysis techniques, *Appl. Intell.* 10 (1999) 5-24

- Spain's electrical market (before 1998): Electrical companies shared a business, Red Eléctrica Española, receiving all the client fees and distributing them among the partners
- The payment distribution was done according to some complex criteria that the government decided to change
- One of them was related to the maintenance costs of the power line belonging to each company
- The different producers were in trouble to compute them since:
  - As low voltage lines are installed in small villages, there were no actual measurement of their length
  - The government wanted the maintenance costs of the optimal medium voltage lines installation and not of the real one, built incrementally

## 2. Evolutionary Tuning of FRBSs

### **Low voltage line maintenance cost estimation:**

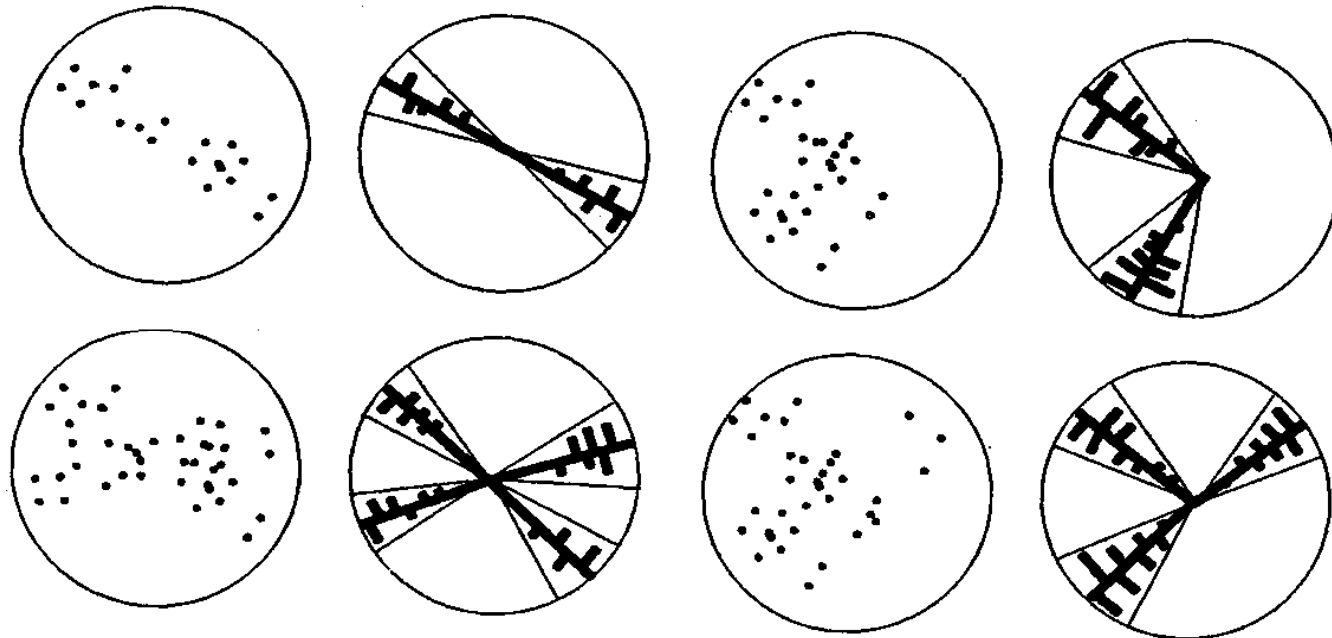
- **Goal:** estimation of the low voltage electrical line length installed in 1000 rural towns in Asturias
- **Two input variables:** number of inhabitants and radius of village
- **Output variable:** length of low voltage line
- Data set composed of **495** rural nuclei, **manually measured and affected by noise**
- **396** (80%) examples for **training** and **99** (20%) examples for **test** randomly selected
- **Seven** linguistic terms for each linguistic variable



## 2. Evolutionary Tuning of FRBSs

### Low voltage line maintenance cost estimation:

- **Classical solution:** numerical regression on different models of the line installation in the villages



## 2. Evolutionary Tuning of FRBSs

### Medium voltage line maintenance cost estimation:

- **Goal:** estimation of the maintenance cost of the **optimal** medium voltage electrical line installed in the Asturias' towns
- **Four input variables:** street length, total area, total area occupied by buildings, and supplied energy
- **Output variable:** medium voltage line maintenance costs
- Data set composed of **1059 simulated** cities
- **847** (80%) examples for **training** and **212** (20%) examples for test randomly selected
- **Five** linguistic terms for each linguistic variable

## 2. Evolutionary Tuning of FRBSs

Obtained results for the medium voltage line problem:

Tuning methods:

Method	Electrical Problem									
	$\bar{x}$				$\sigma_{\bar{x}_i}$			$\sigma_{x_i}$		
	#R	MSE <sub>tra</sub>	MSE <sub>tst</sub>	h:m:s	#R	MSE <sub>tra</sub>	MSE <sub>tst</sub>	#R	MSE <sub>tra</sub>	MSE <sub>tst</sub>
WM	<b>65</b>	56,135	56,359	0:00:00	<b>0.0</b>	1,498	4,685	—	—	—
WM+P-tun	65	18,395	22,136	<b>0:22:41</b>	0.0	778	3,200	—	1,110	1,988
WM+A-tun	65	37,243	38,837	0:33:58	0.0	455	1,816	—	<b>125</b>	<b>572</b>
WM+L-tun	65	20,967	23,420	<b>0:25:16</b>	0.0	632	3,207	—	336	1,439
WM+PA-tun	65	17,967	21,377	0:38:02	0.0	1,078	1,625	—	2,133	2,628
WM+PL-tun	65	<b>9,617</b>	<b>13,519</b>	0:25:33	0.0	<b>263</b>	3,153	—	694	1,509
WM+AL-tun	65	20,544	23,207	0:34:55	0.0	834	2,701	—	797	1,430
WM+PAL-tun	65	<b>11,222</b>	<b>14,741</b>	0:38:12	0.0	380	<b>1,315</b>	—	801	2,136

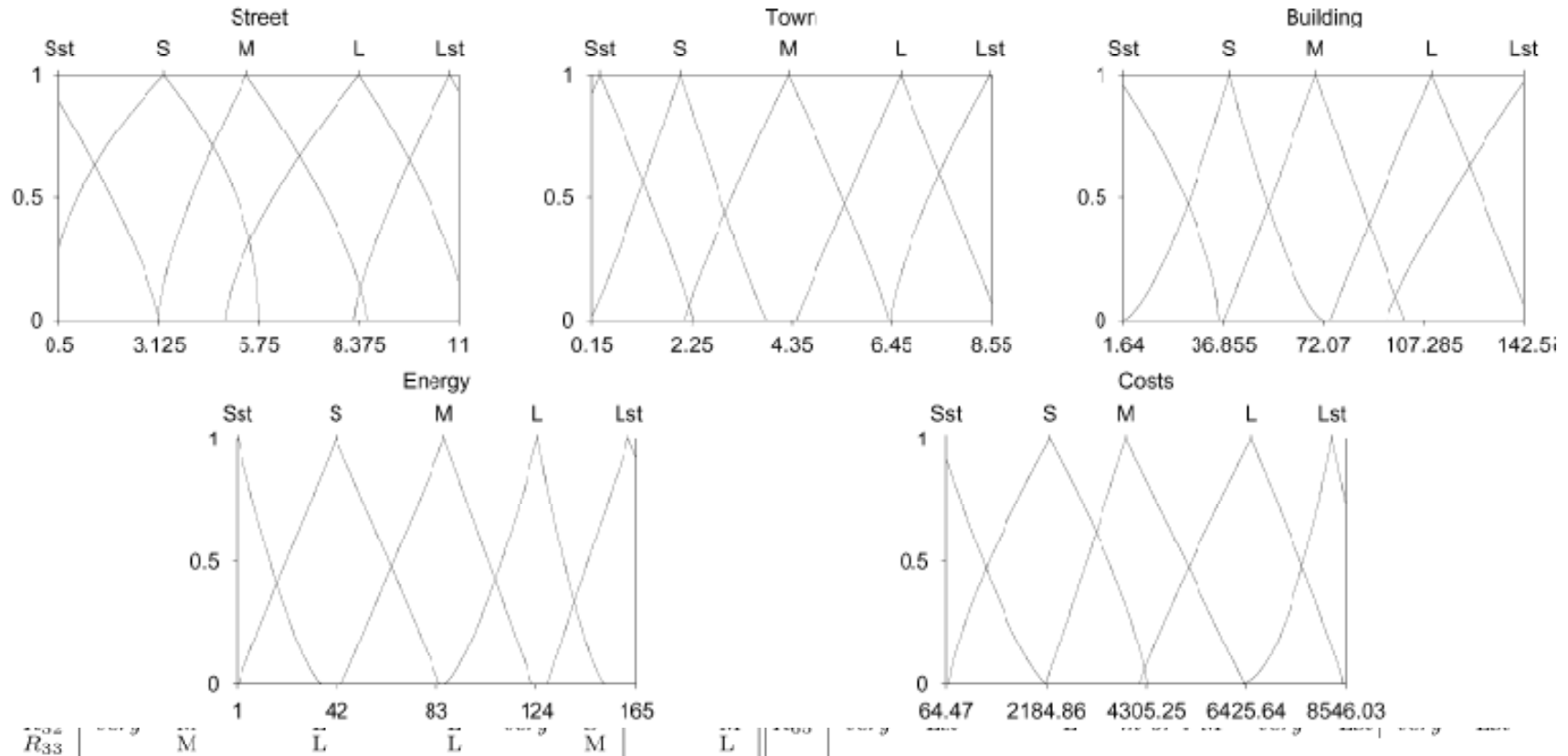
Other fuzzy modeling techniques and GFS:

Method	Electrical Problem									
	$\bar{x}$				$\sigma_{\bar{x}_i}$			$\sigma_{x_i}$		
	#R	MSE <sub>tra</sub>	MSE <sub>tst</sub>	h:m:s	#R	MSE <sub>tra</sub>	MSE <sub>tst</sub>	#R	MSE <sub>tra</sub>	MSE <sub>tst</sub>
Nozaki [5]	<b>532</b>	<b>26,705</b>	<b>27,710</b>	<b>0:00:00</b>	<b>0.0</b>	<b>764</b>	<b>2,906</b>	—	—	—
Thrift [38]	565.3	31,228	37,579	3:13:25	2.6	1,018	7,279	6.1	<b>2,110</b>	<b>3,609</b>
Liska [45]	<b>624.9</b>	49,263	56,089	<b>7:13:34</b>	0.1	2,356	4,628	<b>0.1</b>	7,522	11,191

## 2. Evolutionary Tuning of FRBSs

Obtained results for the medium voltage line problem:

Example of one KB derived from the WM+PAL-tun method:



Before tuning:  $MSE_{tra/test} = 58032 / 55150$

After tuning:  $MSE_{tra/test} = 11395 / 14465$

## 2. Evolutionary Tuning of FRBSs

### **New coding schemes: 2- and 3-tuples:**

**IDEA: New fuzzy rule representation model allowing a more flexible definition of the fuzzy sets of the linguistic labels**

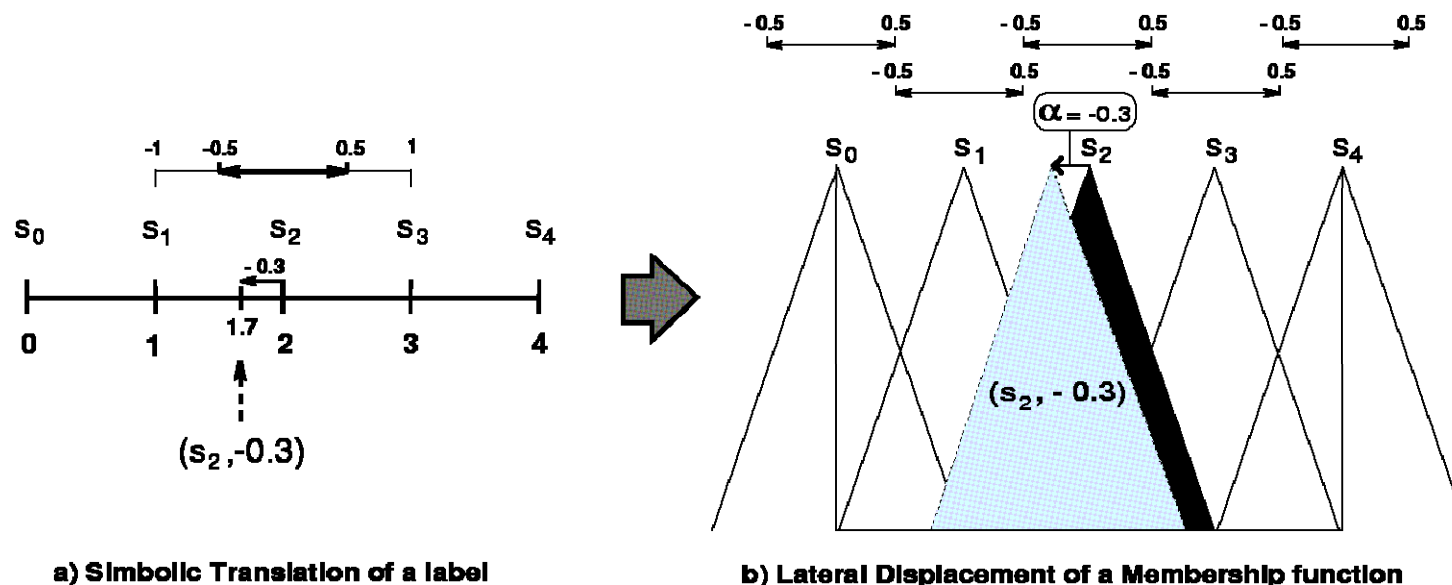
- **R. Alcalá, J. Alcalá-Fdez, M.J. Gacto, F. Herrera, Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation, *Soft Computing* 11 (5) (2007) 401-419**
- **R. Alcalá, J. Alcalá-Fdez, F. Herrera, A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection, *IEEE Transactions on Fuzzy Systems* 15:4 (2007) 616-635**

## 2. Evolutionary Tuning of FRBSs

### New coding schemes: 2- and 3-tuples

IDEA: New fuzzy rule representation model allowing a more flexible definition of the fuzzy sets of the linguistic labels

- **2-tuples**: label id.  $i$  and a displacement parameter  $\alpha_i \in [-0.5, 0.5]$

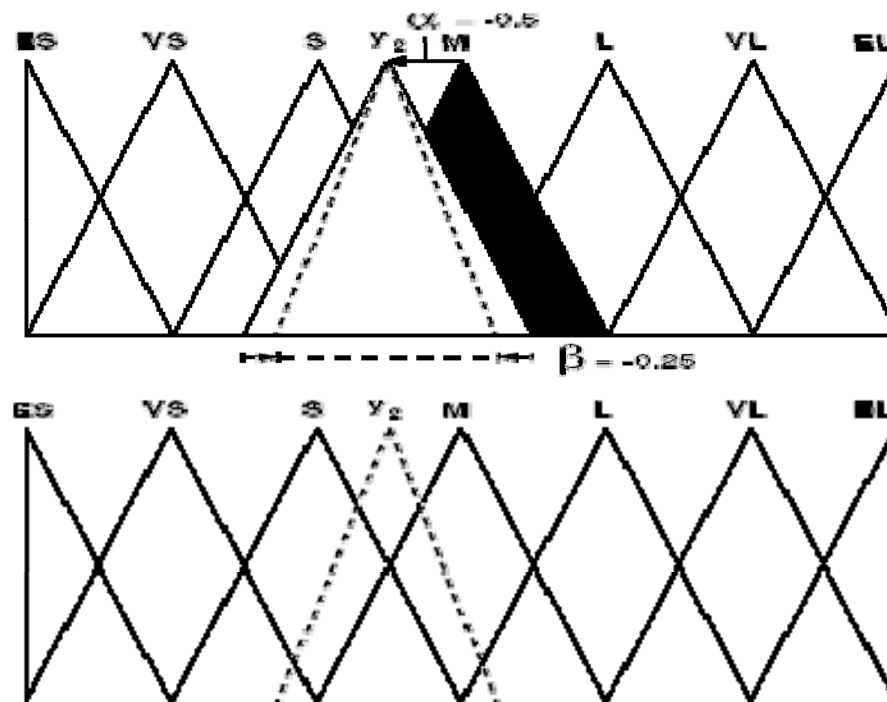


- New rule structure:

IF  $X_1$  IS  $(S^1_i, \alpha_1)$  AND ... AND  $X_n$  IS  $(S^n_i, \alpha_n)$  THEN  $Y$  IS  $(S^y_i, \alpha_y)$

## 2. Evolutionary Tuning of FRBSs

- **3-tuples:** label id.  $i$ , a displacement parameter  $\alpha_i \in [-0.5, 0.5]$ , and a width parameter  $\beta_i \in [-0.5, 0.5]$



- New rule structure:  
IF  $X_1$  IS  $(S^1_i, \alpha_1, \beta_1)$  AND ... AND  $X_n$  IS  $(S^n_i, \alpha_n, \beta_n)$  THEN  $Y$  IS  $(S^y_i, \alpha_y, \beta_y)$

## 2. Evolutionary Tuning of FRBSs

### **New coding schemes: 2- and 3-tuples**

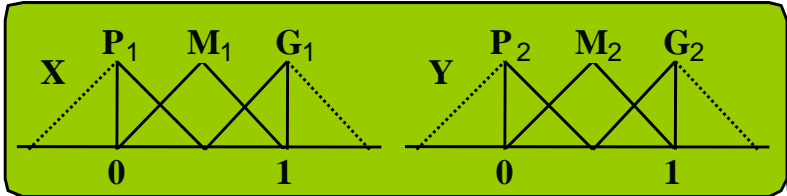
**COLATERAL PROBLEM:** Both structures decreases the KB learning/tuning large scale problem, since the fuzzy sets are encoded using a lower number of parameters

### **Existing proposals:**

- Genetic 2-tuple/3-tuple DB global tuning: adjustment of the global fuzzy sets → **full interpretability** (usual fuzzy partitions)
- Genetic 2-tuple/3-tuple DB tuning at rule level → **lower interpretability, higher flexibility** (like scatter Mamdani FRBSs)
- Genetic 2-tuple/3-tuple DB tuning + rule selection

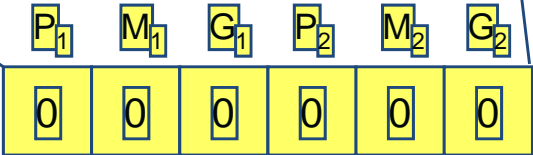


# 2. Evolutionary Tuning of FRBSs



Initial Data Base

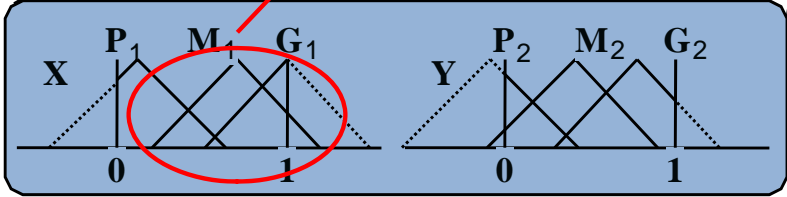
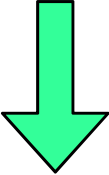
**Lateral  
Tuning  
Scheme**



Initial Rule Base

- $R_1 = \text{If } X \text{ is } G_1 \text{ THEN } Y \text{ is } P_2$
- $R_2 = \text{If } X \text{ is } P_1 \text{ THEN } Y \text{ is } M_2$
- $R_3 = \text{If } X \text{ is } M_1 \text{ THEN } Y \text{ is } G_2$

**Genetic  
Tuning**



**Tuned Data Base**

## 2. Evolutionary Tuning of FRBSs

Medium voltage electrical network in towns

Genetic 2-tuple tuning + rule selection method:

WM	Wang and Mendel Learning Method
S	Rule Selection Method
GL	Global Lateral Tuning
LL	Local Lateral Tuning
T	Classical Genetic Tuning
P A L	Tuning of: Parameters, Domains, and Linguistic Modifiers

Five labels per linguistic variable  
50000 Evaluations per run

Method	#R	MSE <sub>tra</sub>	$\sigma_{tra}$	t-test	MSE <sub>tst</sub>	$\sigma_{tst}$	t-test
WM	65	57605	2841	+	57934	4733	+
S	40.8	41086	1322	+	59942	4931	+
T	65	18602	1211	+	22666	3386	+
PAL	65	10545	279	+	13973	1688	+
T+S	41.9	14987	391	+	18973	3772	+
PAL+S	57.4	12851	362	+	16854	1463	+
GL	65	23064	1479	+	25654	2611	+
LL	65	3664	390	*	5858	1798	*
GL+S	49.1	18801	2669	+	22586	3550	+
LL+S	58.0	3821	385	=	6339	2164	=

5 data partitions 80% - 20%

6 runs per data partition

Averaged results from 30 runs

t-student Test with  $\alpha = 0.05$

## 2. Evolutionary Tuning of FRBSs

Obtained results for the low voltage line problem:

Genetic 2-tuple tuning + rule selection method:

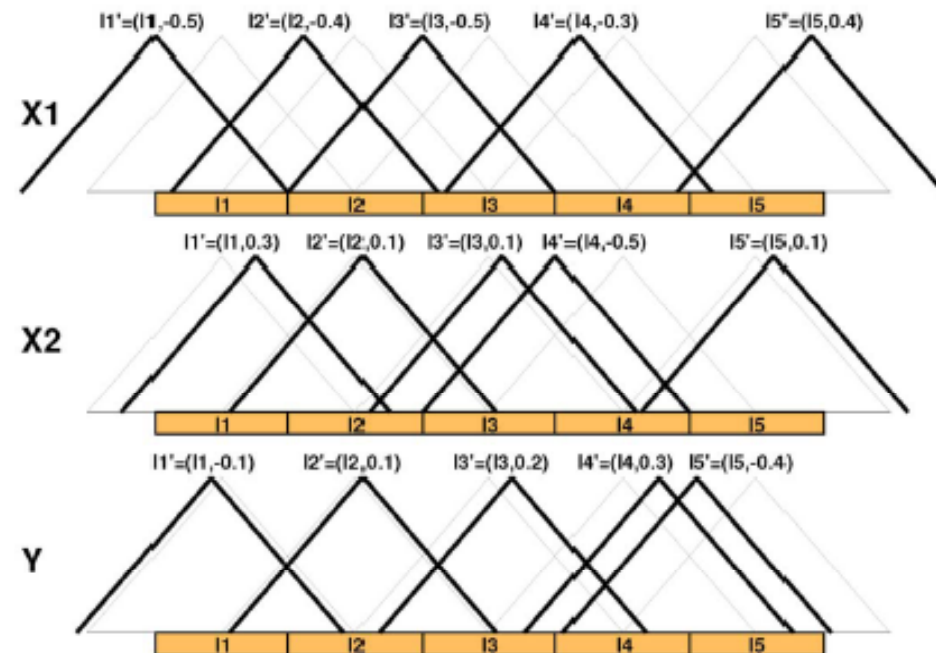
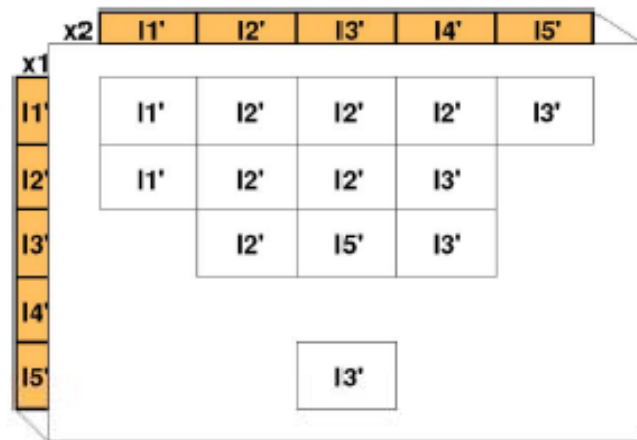
Method	#R	MSE <sub>tra</sub>	$\sigma_{tra}$	t-test	MSE <sub>tst</sub>	$\sigma_{tst}$	t-test
Approaches without tuning							
WM	12.4	234712	32073	+	242147	24473	+
S	10.0	226135	19875	+	241883	19410	+
Approaches with global semantics							
T	12.4	158662	6495	+	221613	29986	+
T+S	8.9	156313	2967	+	193477	49912	=
GL <sub>ddl</sub>	12.4	166674	11480	!	189216	14743	-
GL <sub>ddl</sub> +S	9.0	160081	7316	+	189844	22448	=
Approaches with local semantics							
PAL	12.4	141638	4340	+	189279	19523	-
PAL+S	10.6	145712	5444	+	191922	16987	-
LL <sub>ddl</sub>	12.4	139189	3155	*	191604	18243	-
LL <sub>ddl</sub> +S	10.5	141445	3444	-	186746	15762	*

- 5-fold cross validation  $\times$  6 runs = 30 runs per algorithm
- T-student test with 95% confidence

## 2. Evolutionary Tuning of FRBSs

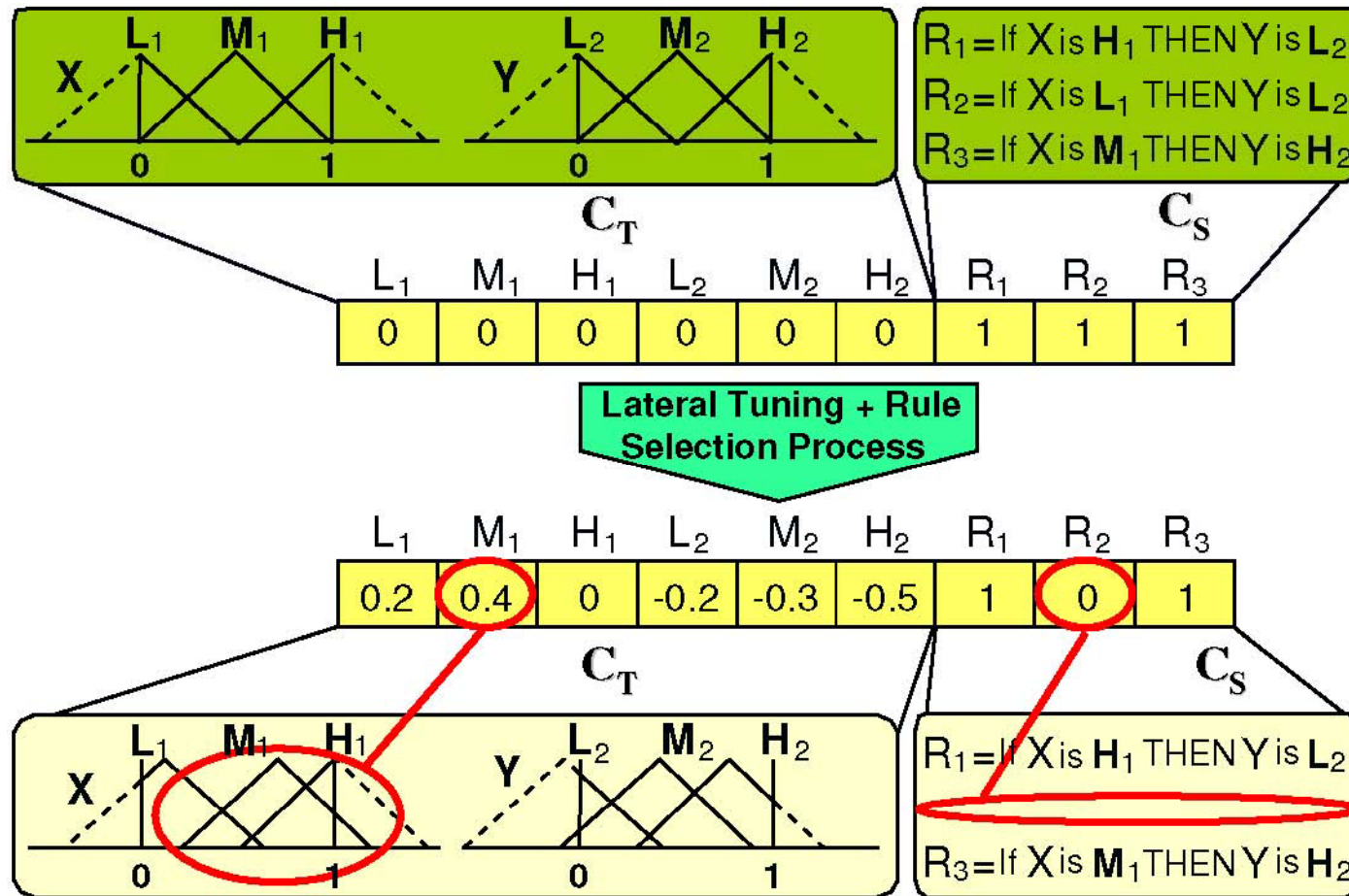
Obtained results for the low voltage line problem:

Example of one KB derived from the global tuning method:



After tuning+rule selection: #R=13;  $MSE_{tra/test} = 187494 / 176581$

## 2. Evolutionary Tuning of FRBSs



Example of genetic lateral tuning and rule selection

## 2. Evolutionary Tuning of FRBSs

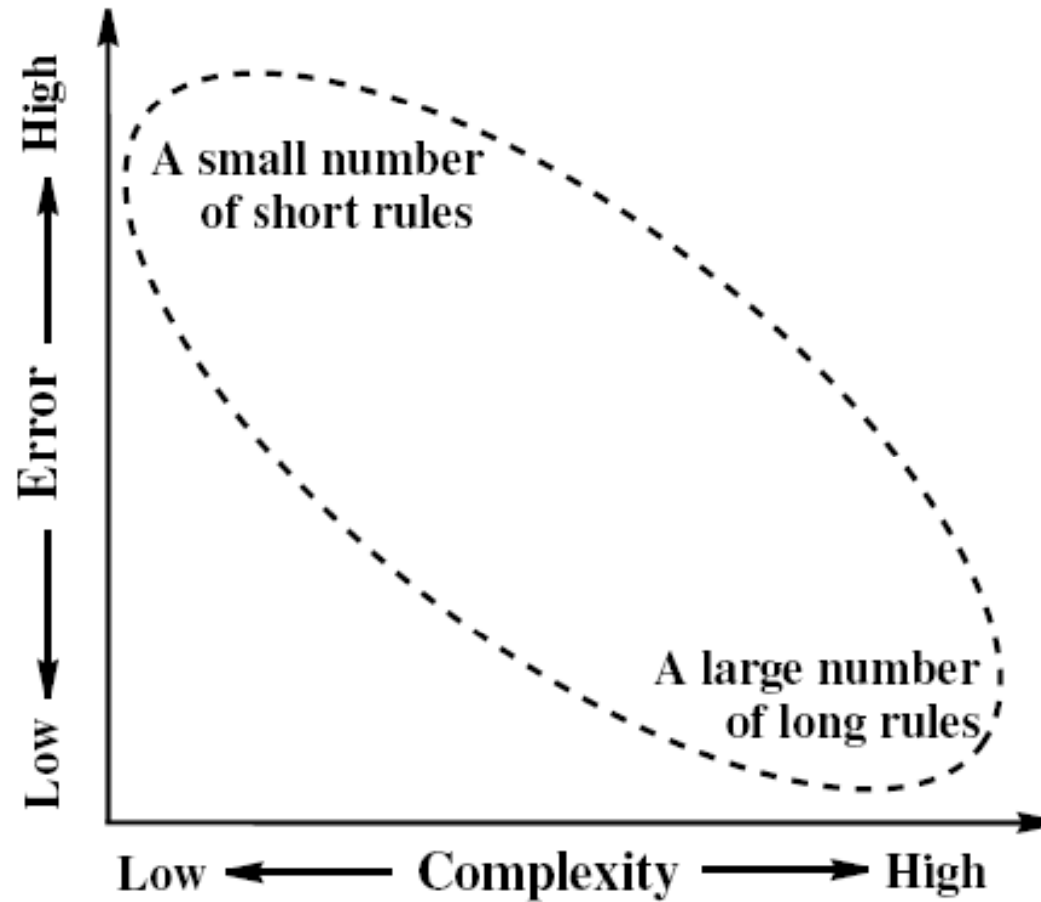
### **New Tuning Model: Multi-objective GFS for the interpretability-accuracy trade-off:**

R. Alcalá, J. Alcalá-Fdez, M.J. Gacto, F. Herrera, A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 15:5 (2007) 539–557,

Multi-objective EAs are powerful tools to generate GFSs but they are based on getting a large, well distributed and spread off, Pareto set of solutions

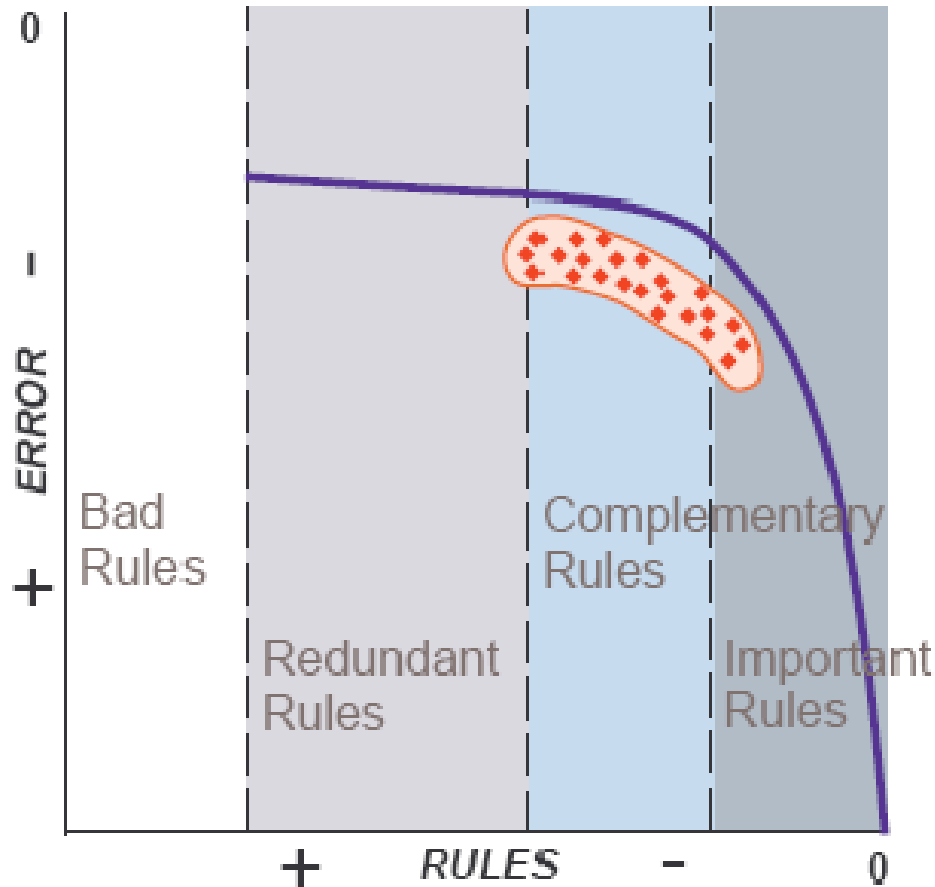
- The two criteria to optimize in GFSs are accuracy and interpretability. The former is more important than the latter, so many solutions in the Pareto set are not useful
- Solution: Inject knowledge through the MOEA run to bias the algorithm to generate the desired Pareto front part

## 2. Evolutionary Tuning of FRBSs



## 2. Evolutionary Tuning of FRBSs

### Pareto front classification in an interpretability-accuracy GFSs:



— Desired pareto zone  
— Optimal pareto frontier

- **Bad rules zone:** solutions with bad performance rules. Removing them improves the accuracy, so no Pareto solutions are located here
- **Redundant rules zone:** solutions with irrelevant rules. Removing them does not affect the accuracy and improves the interpretability
- **Complementary rules zone:** solutions with neither bad nor irrelevant rules. Removing them slightly decreases the accuracy
- **Important rules zone:** solutions with essential rules. Removing them significantly decreases the accuracy



## 2. Evolutionary Tuning of FRBSs

### Accuracy-oriented modifications performed:

- Restart the genetic population at the middle of the run time, keeping the individual with the highest accuracy as the only one in the external population and generating all the new individuals with the same number of rules it has
- In each MOGA step, the number of chromosomes in the external population considered for the binary tournament is decreased, focusing the selection on the higher accuracy individuals

## 2. Evolutionary Tuning of FRBSs

Obtained results for the medium voltage line problem:

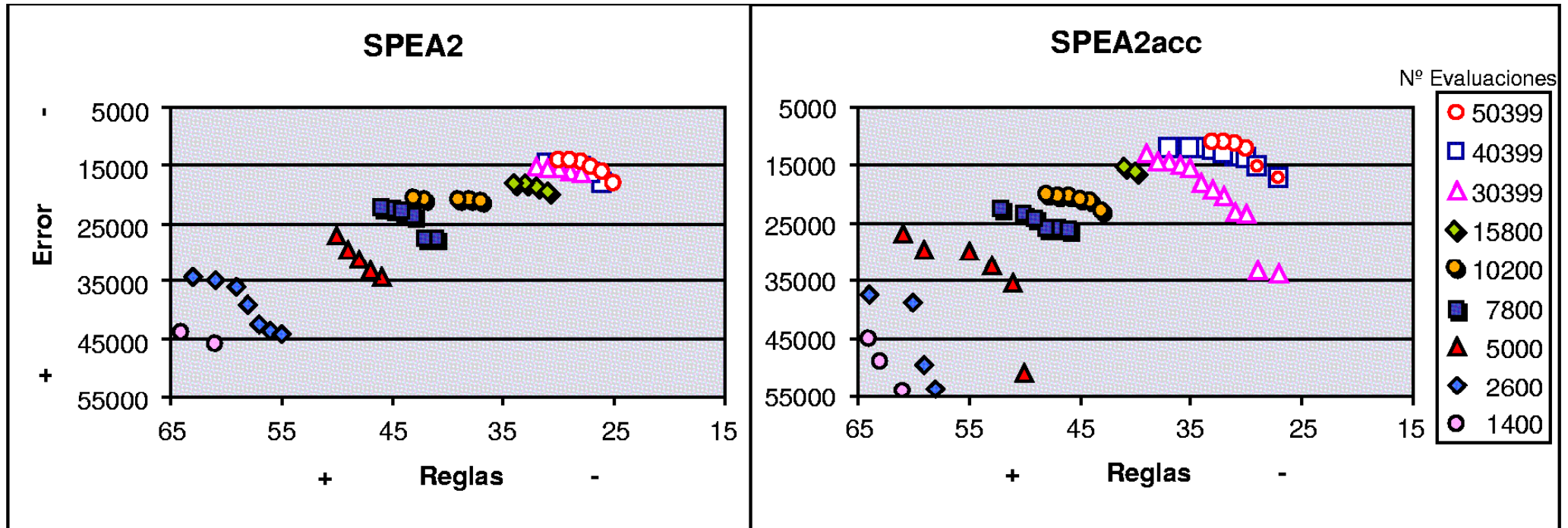
Multi-objective genetic tuning + rule selection method:

Method	#R	MSE <sub>tra</sub>	$\sigma_{tra}$	t-test	MSE <sub>tst</sub>	$\sigma_{tst}$	t-test
WM	65	57605	2841	+	57934	4733	+
WM+T	65	18602	1211	+	22666	3386	+
WM+S	40.8	41086	1322	+	59942	4931	+
WM+TS	41.9	14987	391	+	18973	3772	+
NSGAI	41.0	14488	965	+	18419	3054	+
NSGAI <sub>ACC</sub>	48.1	16321	1636	+	20423	3138	+
SPEA2	<b>33</b>	13272	1265	+	17533	3226	+
SPEA2 <sub>ACC</sub>	34.5	<b>11081</b>	1186	*	<b>14161</b>	2191	*

- 5-fold cross validation  $\times$  6 runs = 30 runs per algorithm
- T-student test with 95% confidence

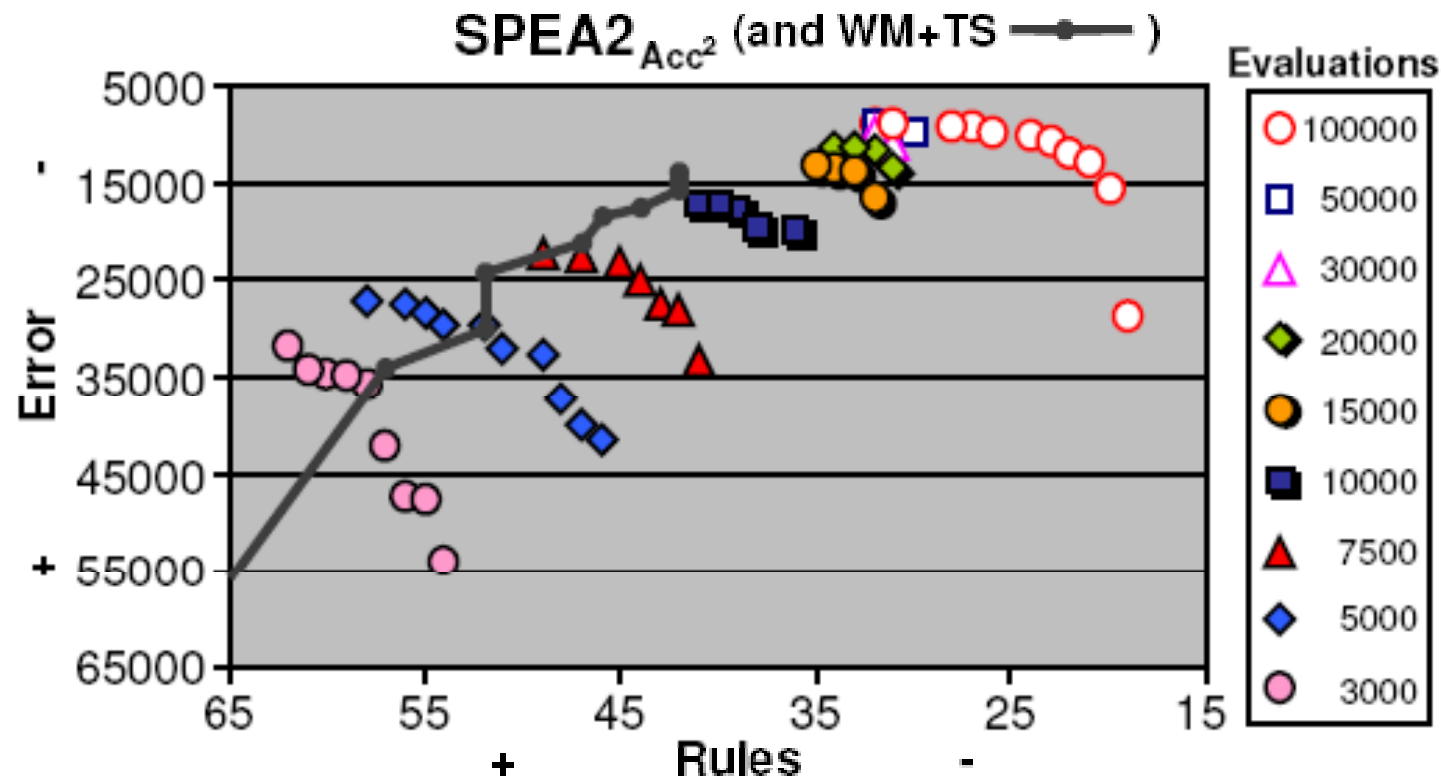
## 2. Evolutionary Tuning of FRBSs

Comparison of the SPEA2 – SPEA2acc convergence:



## 2. Evolutionary Tuning of FRBSs

Comparison of the SPEA2acc and classical GA for for the medium voltage line problem:



M.J. Gacto, R. Alcalá, F. Herrera, Adaptation and Application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems, *Soft Computing*, 2008, to appear.

## 2. Evolutionary Tuning of FRBSs

### **Future Studies:**

- ❑ To develop appropriate MOEAs for getting a pareto with a better trade-off between precision and interpretability, improving the precision.
- ❑ To design interpretability measures for including them into the MOEAs objectives.



## Outline

- ✓ Brief Introduction to Genetic Fuzzy Systems
- ✓ Genetic Tuning Methods: Basic and Advanced Approaches
- ✓ Genetic Fuzzy Systems Application to HVAC Problems
- ✓ GFSs: Current Trends and Prospects
- ✓ Concluding Remarks