5th IEEE International Workshop on Genetic and Evolutionary Fuzzy Systems
April 15 2011 - Paris

Multi-objective Evolutionary Learning of Fuzzy Rule-based Systems for Regression Problems

Francesco Marcelloni

Computational Intelligence Group

Department of Information Engineering

University of Pisa

Italy

E-mail: f.marcelloni@ing.unipi.it

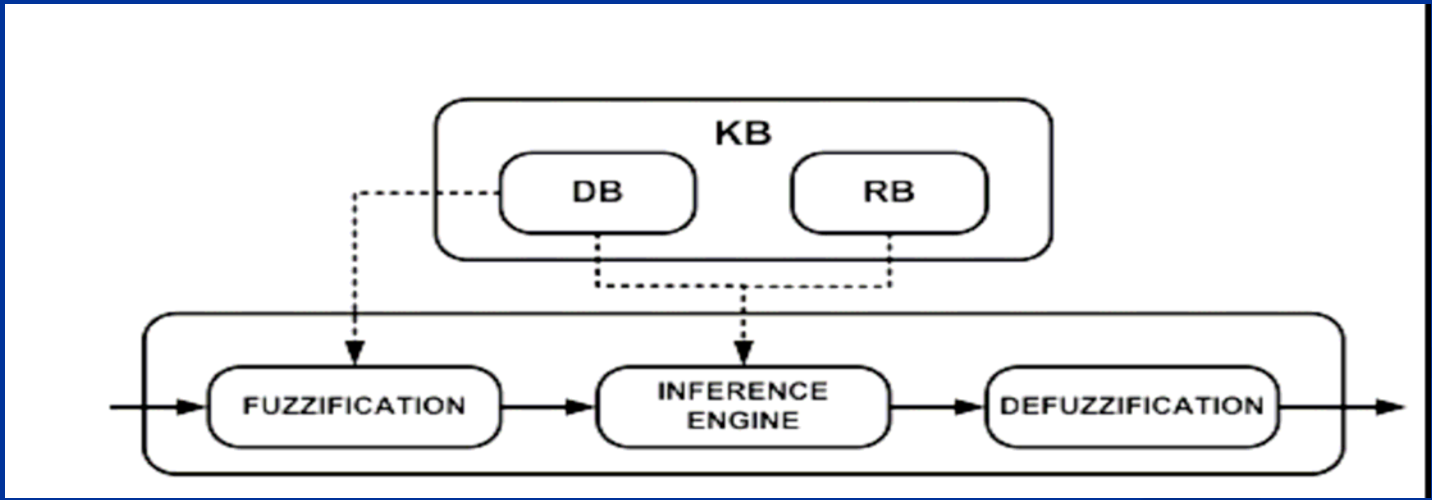
Outline

- Background
  - Fuzzy Rule-Based Systems (FRBSs) and Genetic Fuzzy Systems
  - Interpretability Issues in FRBS design
- Multi-Objective Evolutionary Fuzzy Systems
  - Multi-objective Evolutionary Data Base Tuning
  - Multi-objective Evolutionary Data Base Learning
  - Multi-objective Evolutionary Rule Selection
  - Multi-objective Evolutionary Rule Learning
  - Multi-objective Evolutionary Rule Selection and Data Base Tuning
  - Multi-objective Evolutionary Knowledge Base Learning
- How can we compare different approaches?
- Hot Topics and New Challenges
  - Large Datasets
  - High-dimensional Datasets

Fuzzy Rule-Based Systems

Fuzzy Rule-Based Systems (FRBSs) consist of:

- a rule base (RB) containing the fuzzy rules
- a data base (DB) containing the fuzzy sets associated with the linguistic terms used in the RB
- a fuzzy logic inference engine



Formally, an FRBS is a mathematical model that, given an input vector, computes an output value, exploiting the knowledge coded in the RB and in the DB, and an inference process based on fuzzy logic.

Francesco Marcelloni

Fuzzy Rule-

Mamdani Rules

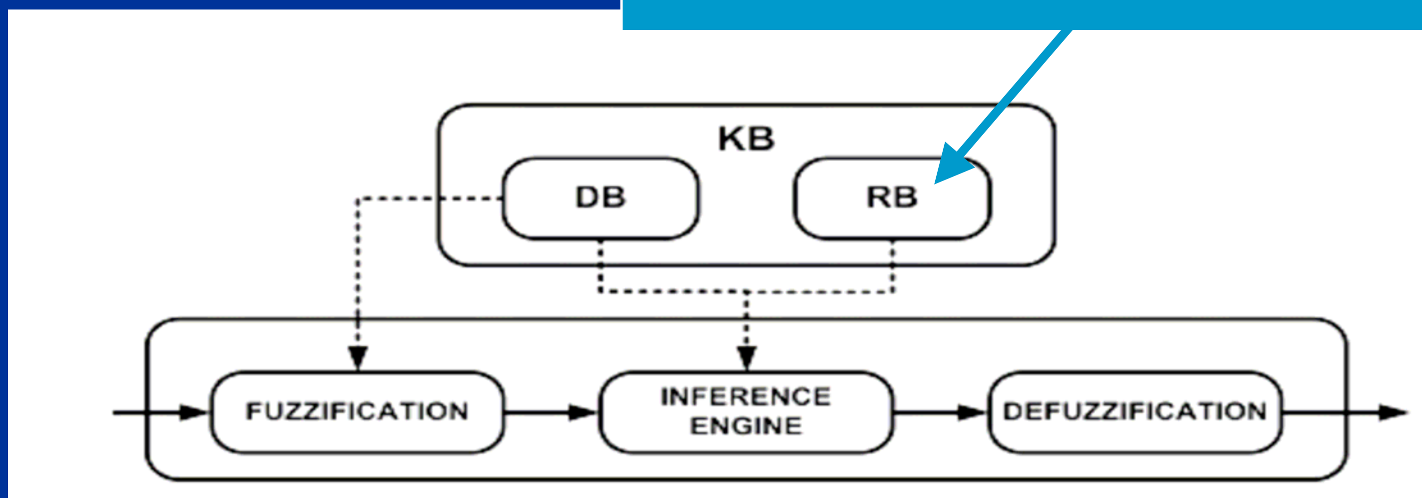
$$R_1: \text{IF } X_1 \text{ is } A_{1,2} \text{ and } X_2 \text{ is } A_{2,1} \text{ THEN } X_3 \text{ is } A_{3,1}$$

$$R_2: \text{IF } X_1 \text{ is } A_{1,2} \text{ and } X_2 \text{ is } A_{2,2} \text{ THEN } X_3 \text{ is } A_{3,1}$$

$$R_3: \text{IF } X_1 \text{ is } A_{1,1} \text{ and } X_2 \text{ is } A_{2,1} \text{ THEN } X_3 \text{ is } A_{3,2}$$

Fuzzy Rule-Based Systems (FR

- a rule base (RB) containing the f
- a data base (DB) containing the linguistic terms used in the RB
- a fuzzy logic inference engine



Formally, an FRBS is a mathematical model that, given an input vector, computes an output value, exploiting the knowledge coded in the RB and in the DB, and an inference process based on fuzzy logic.

FRBS Design Process

- The objective of the FRBS design process is to identify both the Rule Base (RB) and the Data Base (DB) of FRBSs from numerical data, when tackling problems where the knowledge provided by human experts is low or missing
- The RB design process consists of identifying the optimal set of rules that manage to reproduce the behaviour of the approached problem framework
- The DB design process consists of finding the correct number of fuzzy sets for each linguistic variable and their parameters

Evolutionary Learning

- The automatic definition of the structure and parameters of FRBSs can be considered as an optimization process

- Evolutionary algorithms (EAs) have proved to be very effective to search for optimal solutions in complex search space

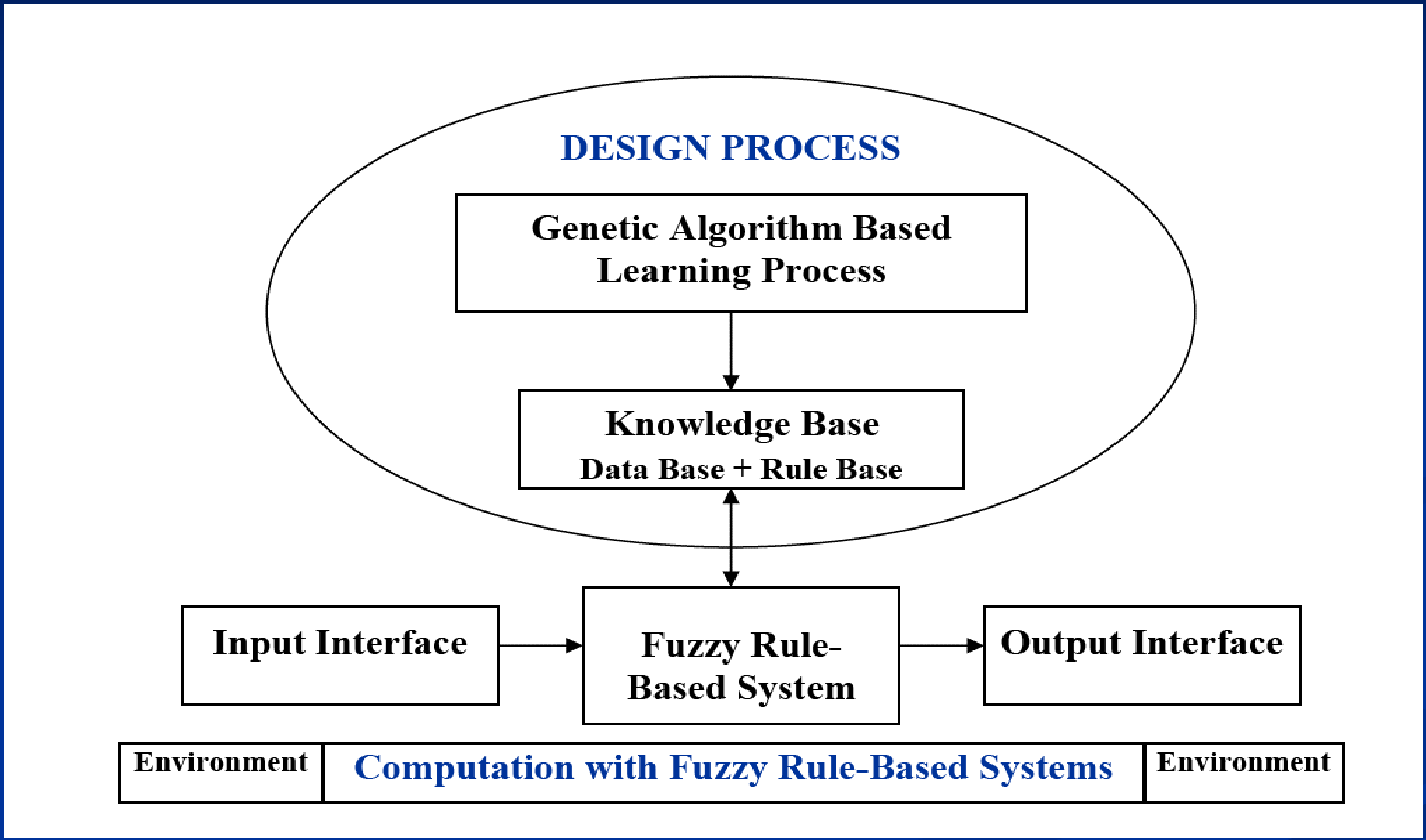
- Genetic/evolutionary algorithms have been so extensively used to design FRBSs that a specific term has been introduced in Computational Intelligence

Genetic algorithms + Fuzzy Systems

$\Downarrow$

Genetic Fuzzy Systems

Genetic Fuzzy Rule-Based Systems



F. Herrera, "Genetic fuzzy systems: taxonomy, current research trends and prospects," Evolutionary Intelligence, vol. 1, pp. 27–46, 2008.

GFS Taxonomy

GFS

Genetic Tuning

Given a KB, a genetic tuning process is applied to improve the performance but without changing the existing RB (adjusting the DB definition or adapting the parameters of the inference engine)

Genetic Learning

Gen. DB Learning

Embedded learning: the DB generation process wraps an RB learning process.

Gen. RB Selection

Selects relevant rules from a predefined RB

Gen. RB Learning

Uses a predefined DB

Gen. KB Learning

Simultaneous learning of RB and DB components

F. Herrera, "Genetic fuzzy systems: taxonomy, current research trends and prospects," Evolutionary Intelligence, vol. 1, pp. 27–46, 2008.

What should you optimise?

- Accuracy?
  - Yes, of course.
  - But, we would also like to understand how the FRBS works. Thus, we would like to optimize interpretability of the FRBS
- How?
  - We need a universally accepted measure of interpretability of the FRBS
- Does this measure exist?
  - Uhhhmmm... Interpretability is subjective
  - Thus, researchers have focused their attention on discussing some factors which characterize interpretability and on proposing some constraints which have to be satisfied for these factors.

Interpretability Factors

The interpretability of an FRBS is related to the following factors:

- **Comprehensibility/Integrity** of fuzzy partitions (e.g., linguistic interpretability of each fuzzy set, separation of neighbouring fuzzy sets, number of fuzzy sets per each variable)
- **Simplicity/Complexity** of the system (e.g., number of input variables, number of fuzzy if-then rules)
- **Simplicity of fuzzy if-then rules** (e.g., type of fuzzy if-then rules, number of antecedent conditions in each fuzzy if-then rule)
- **Simplicity of fuzzy reasoning** (e.g., selection of a single winner rule, voting by multiple rules)

H. Ishibuchi, and T. Yamamoto, "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining," Fuzzy Sets and Systems, vol. 141, pp. 59–88, 2004.

Reviews on Interpretability

J.V. de Oliveira, "Semantic constraints for membership function optimization," IEEE Trans. Syst. Man. Cybern. Part A, vol. 29, n.1, pp 128-138, 1999.

S.M. Zhou and J.Q. Gan, "Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modeling," Fuzzy Sets and Systems, vol. 159, pp. 3091-3131, 2008.

C. Mencar and A.M. Fanelli, "Interpretability constraints for fuzzy information granulation," Information Sciences, vol. 178, pp. 4585-4618, 2008.

J.M. Alonso, L. Magdalena and G. Gonzalez-Rodriguez, "Looking for a good fuzzy system interpretability index: An experimental approach," Int. J. Approx. Reason, vol. 51, pp. 115-134,2009

M.J. Gacto, R. Alcalá, and F. Herrera, "Interpretability of Linguistic Fuzzy Rule-Based Systems: An Overview of Interpretability Measures," Information Sciences, in press (2011).

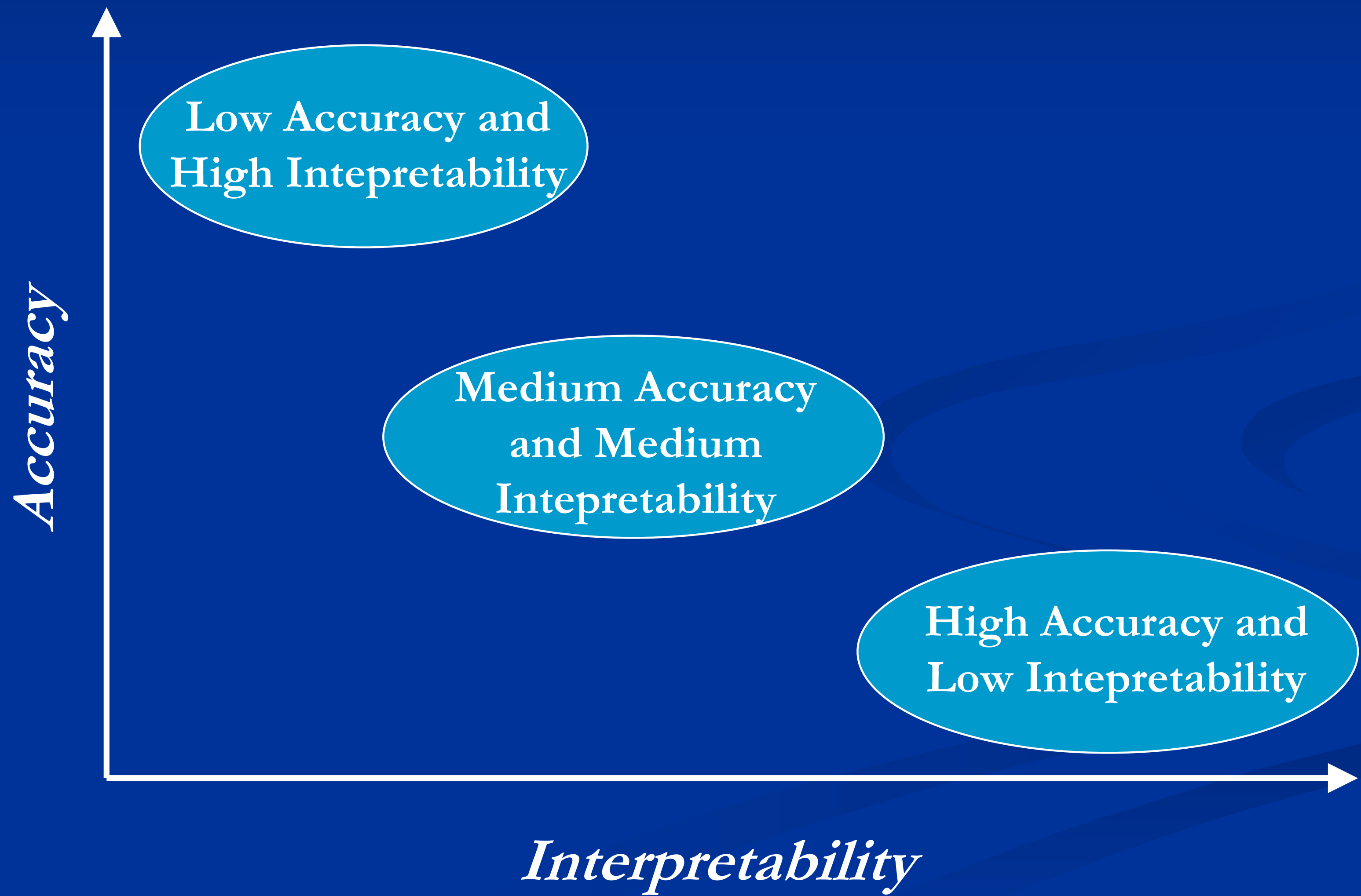
Interpretability: A taxonomy

	Rule Base Level	Data Base Level
Complexity	Number of Rules Number of Conditions Average Rule Length	Number of Features Number of Membership Functions
Semantic	Consistency of Rules Number of Rules Fired at the Same Time Transparency of the Structure Cointension	Coverage Normalization Distinguishability Order Relative Measures

M.J. Gacto, R. Alcalá, and F. Herrera, "Interpretability of Linguistic Fuzzy Rule-Based Systems: An Overview of Interpretability Measures," Information Sciences, in press (2011)

Multi-objective FRBS Design

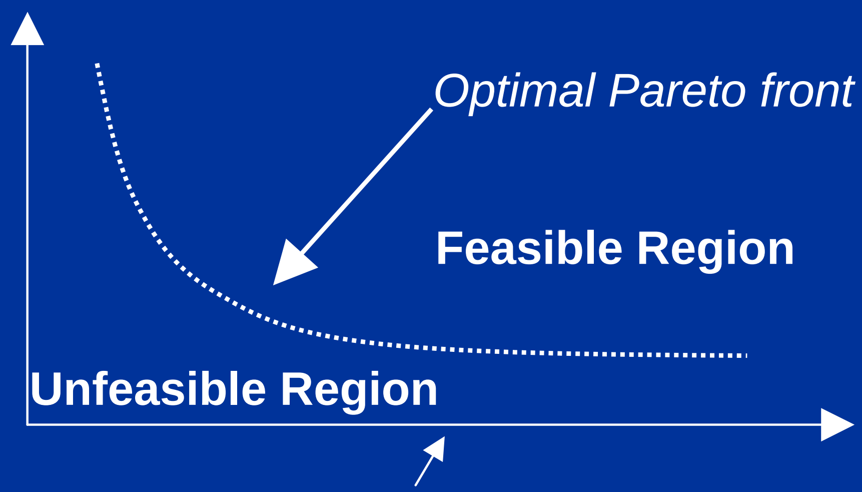
- Unfortunately, increasing interpretability and improving accuracy are often conflicting objectives



Pareto Optimality

- There exists a set of trade-off solutions, i.e. with different compromises between the objectives
- Two solutions are compared by using the notion of Pareto dominance:

"s_1 dominates s_2 if s_1 is not worse than s_2 on all the objectives and is better than s_2 in at least one objective"



Here the goal is the minimization of both the objectives

- A solution is said to be Pareto-optimal if it is not dominated by any other possible solution
- The set of Pareto-optimal solutions is denoted as *Pareto optimal set* and the corresponding objective vectors form the *Optimal Pareto front*

Multi-objective Evolutionary Fuzzy Systems (MOEFSs)

- **MOEFSs** extend the GFS paradigm exploiting MOEAs to design sets of FRBSs with different **trade-offs** between **accuracy and intepretability**

MOEFS

MOE Tuning

Given a KB, optimization is performed without changing the existing RB (typically, adjusting the DB definition)

MOE Learning

MOE DB Learning

Embedded learning: the DB generation process wraps an RB learning process.

MOE RB Selection

Selects relevant rules from a predefined RB

MOE RB Learning

Uses a predefined DB

MOE KB Learning

Simultaneous learning of RB and DB components

Multi-objective Evolutionary Fuzzy Systems (MOEFSs)

MOEFS

MOE Tuning

Given a KB, optimization is performed without changing the existing RB (typically, adjusting the DB definition)

MOE Learning

MOE DB Learning

Embedded learning: the DB generation process wraps an RB learning process.

MOE RB Selection

Selects relevant rules from a predefined RB

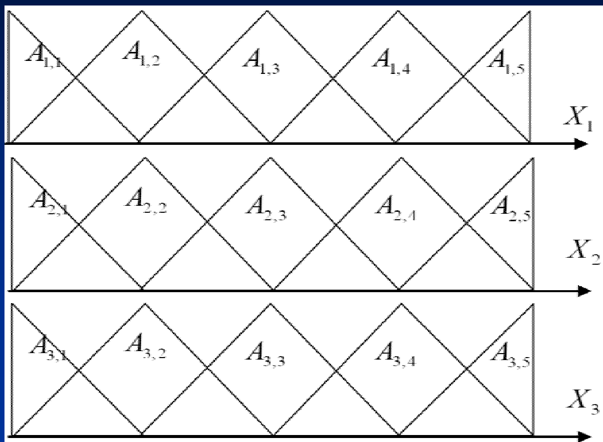
MOE RB Learning

Uses a predefined DB

MOE KB Learning

Simultaneous learning of RB and DB components

MOE DB Tuning - Description



Initial Data Base

R_1 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

R_2 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3 : IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$

Predefined Rule Base

Training set

MOE DB Tuning

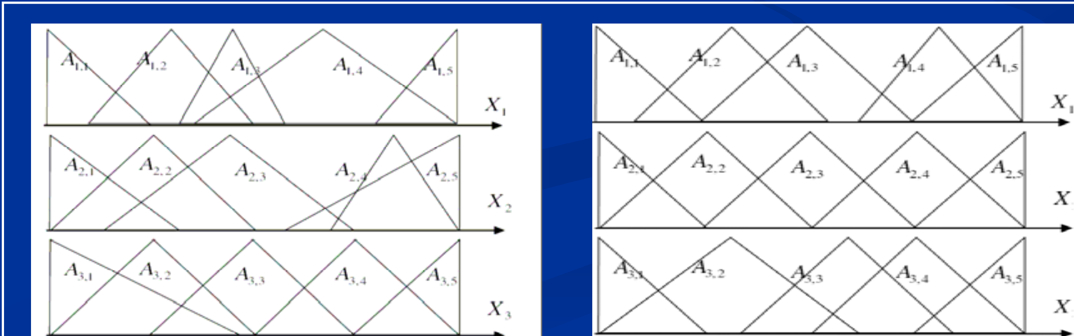
Real chromosome which codifies:

- Linear or non linear tranformation
- Core position and/or support width and position
- Linguistic modifiers
- Lateral displacement

DB Fitness Evaluation

Accuracy: Mean Square Error

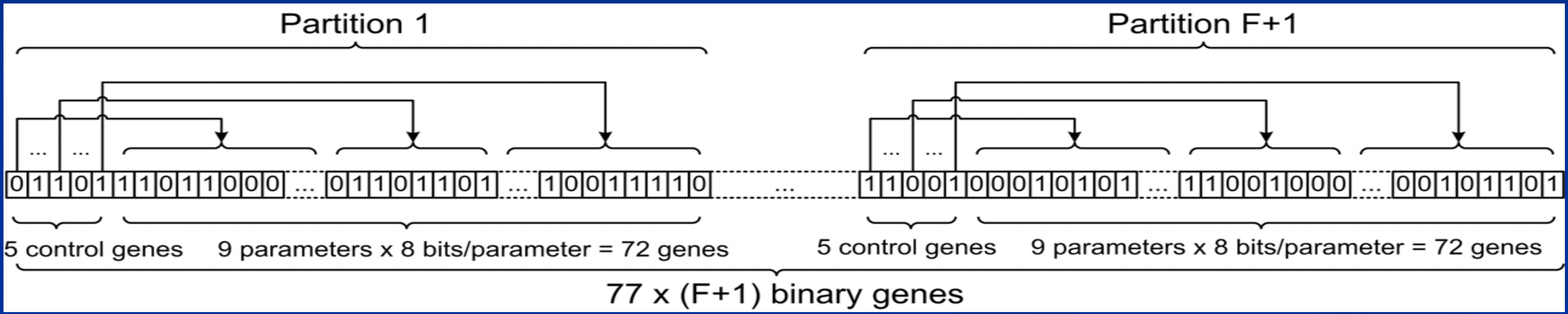
Interpretability: Integrity of the Partitions, Relative Measures



Non-Dominated Final Data Bases

MOE DB Tuning – Example 1(a)

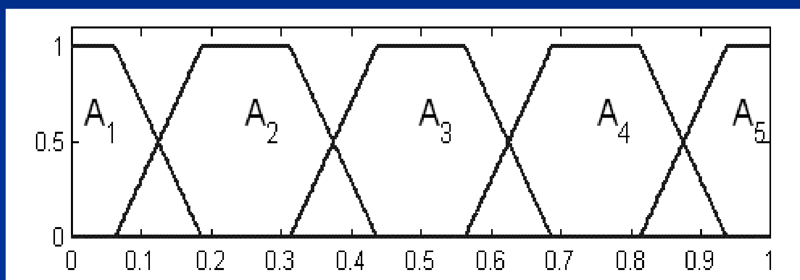
A. Botta, B. Lazzerini, F. Marcelloni, and D. Stefanescu, "Context adaptation of fuzzy systems through a multiobjective evolutionary approach based on a novel interpretability index," Soft Comput., vol. 13, no. 5, pp. 437–449, 2009



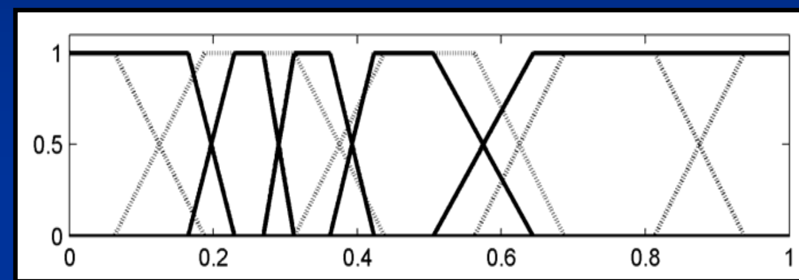
- The chromosome codifies the parameters of 5 operator (a non-linear scaling function and 4 **linguistic modifiers)** used to adapt the DB
- The first five bits, one for each operator, control whether the corresponding operator is applied or not on each fuzzy partition
- The other 72 bits are organized in sub-strings of 8 bits: each sub-string determines the value of a different parameter, via **Gray decoding** and quantization

MOE DB Tuning – Example 1(b)

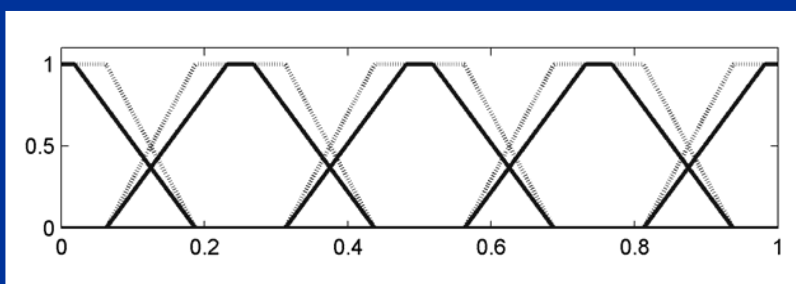
A. Botta, B. Lazzerini, F. Marcelloni, and D. Stefanescu, "Context adaptation of fuzzy systems through a multiobjective evolutionary approach based on a novel interpretability index," Soft Comput., vol. 13, no. 5, pp. 437–449, 2009



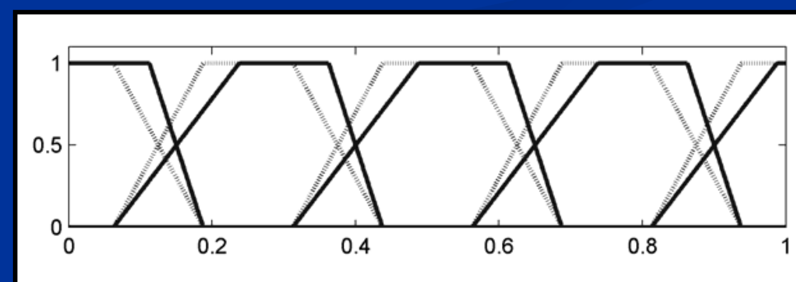
Normalized partition



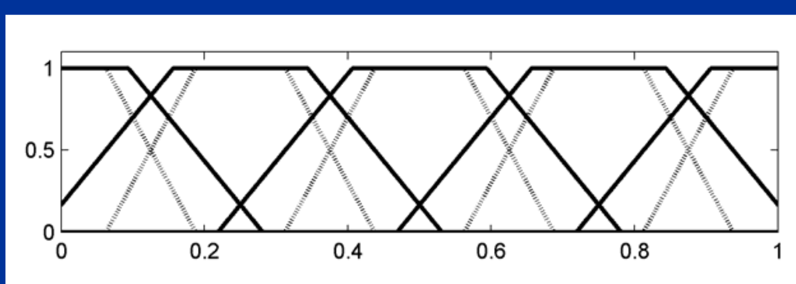
Non-linear Scaling function



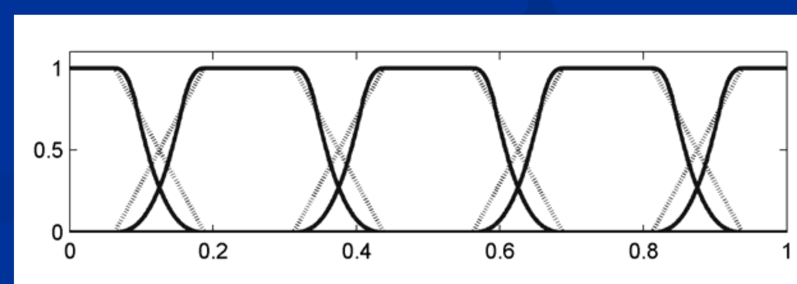
Core-width modifier



Core-position modifier



Support-width modifier



Generalized positively modifier

MOE DB Tuning – Example 1(c)

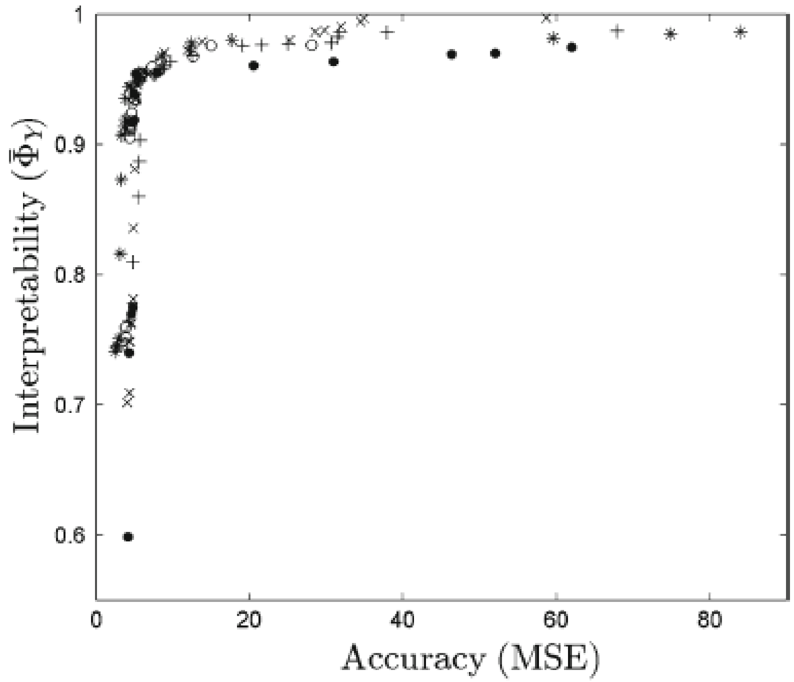
A. Botta, B. Lazzerini, F. Marcelloni, and D. Stefanescu, "Context adaptation of fuzzy systems through a multiobjective evolutionary approach based on a novel interpretability index," Soft Comput., vol. 13, no. 5, pp. 437–449, 2009

- **NSGA-II** is applied to find a set of non dominated and context-adapted Mamdani FRBSs
- **Accuracy** is evaluated in terms of mean square error
- **Interpretability** is evaluated using a **novel index** which exploits an **ordering measure** and an empirical binding between **crossing points** and ordering to assess the **integrity** level of the adapted partitions

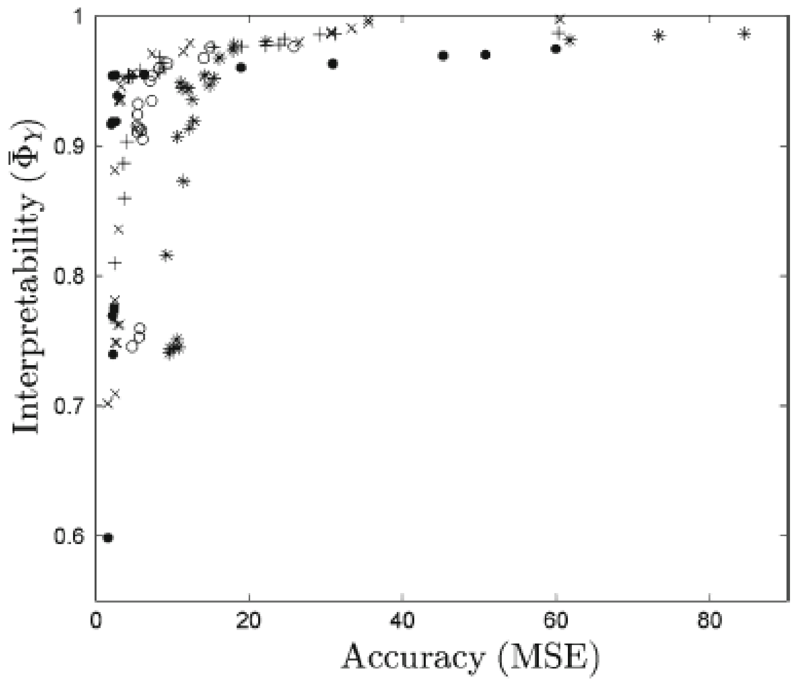
MOE DB Tuning – Example 1(d)

A. Botta, B. Lazzerini, F. Marcelloni, and D. Stefanescu, "Context adaptation of fuzzy systems through a multiobjective evolutionary approach based on a novel interpretability index," Soft Comput., vol. 13, no. 5, pp. 437–449, 2009

Results on the **Fuel Efficiency dataset** which contains 387 input-output patterns (4 input variables)



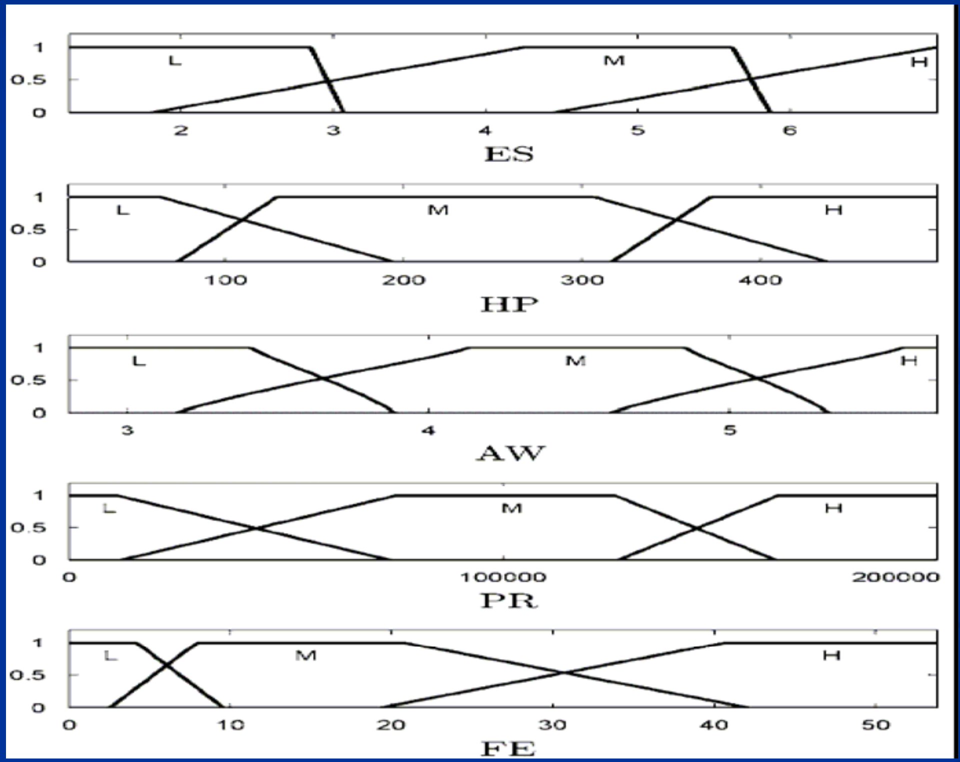
(a) Training set



(b) Test set

MOE DB Tuning – Example 1(e)

A. Botta, B. Lazzerini, F. Marcelloni, and D. Stefanescu, "Context adaptation of fuzzy systems through a multiobjective evolutionary approach based on a novel interpretability index," Soft Comput., vol. 13, no. 5, pp. 437–449, 2009



An example of context adapted DB

Multi-objective Evolutionary Fuzzy Systems (MOEFSs)

MOEFS

MOE Tuning

Given a KB, optimization is performed without changing the existing RB (typically, adjusting the DB definition)

MOE Learning

MOE DB Learning

Embedded learning: the DB generation process wraps an RB learning process.

MOE RB Selection

Selects relevant rules from a predefined RB

MOE RB Learning

Uses a predefined DB

MOE KB Learning

Simultaneous learning of RB and DB components

MOE DB Learning – Description

MOE Granularity and MF Parameters Learning

The chromosomes can be formed by different parts

Granularity part: Integer

MF parameter part: Real

Set of DBs

Rule Learning Method

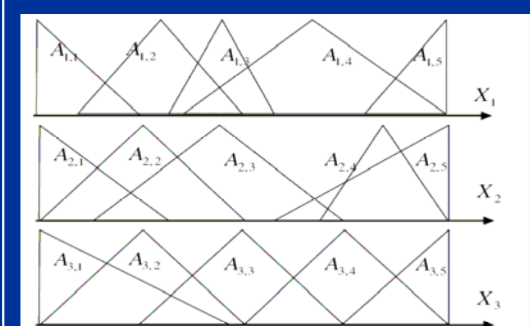
Set of KBs

Training set

KB Fitness Evaluation

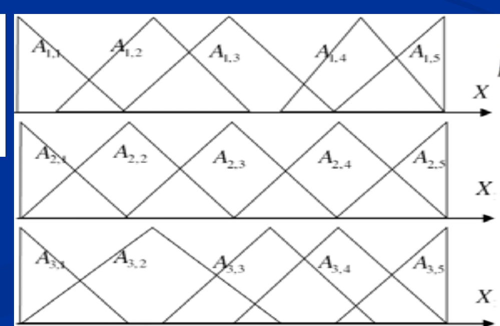
Accuracy: Mean Square Error

RB Interpretability: Number of Rules



R_2: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3: IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$



R_1: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

Non-Dominated Final Knowledge Bases

MOE DB Learning – Example 1

R. *Alcala, M. J. Gacto, and F. Herrera, "A Fast and Scalable Multi-Objective Genetic Fuzzy System for Linguistic Fuzzy Modeling in High-Dimensional Regression Problems," IEEE Transactions on Fuzzy System,* doi: 10.1109/TFUZZ.2011.2131657, *in press (2011).*

Double coding scheme:

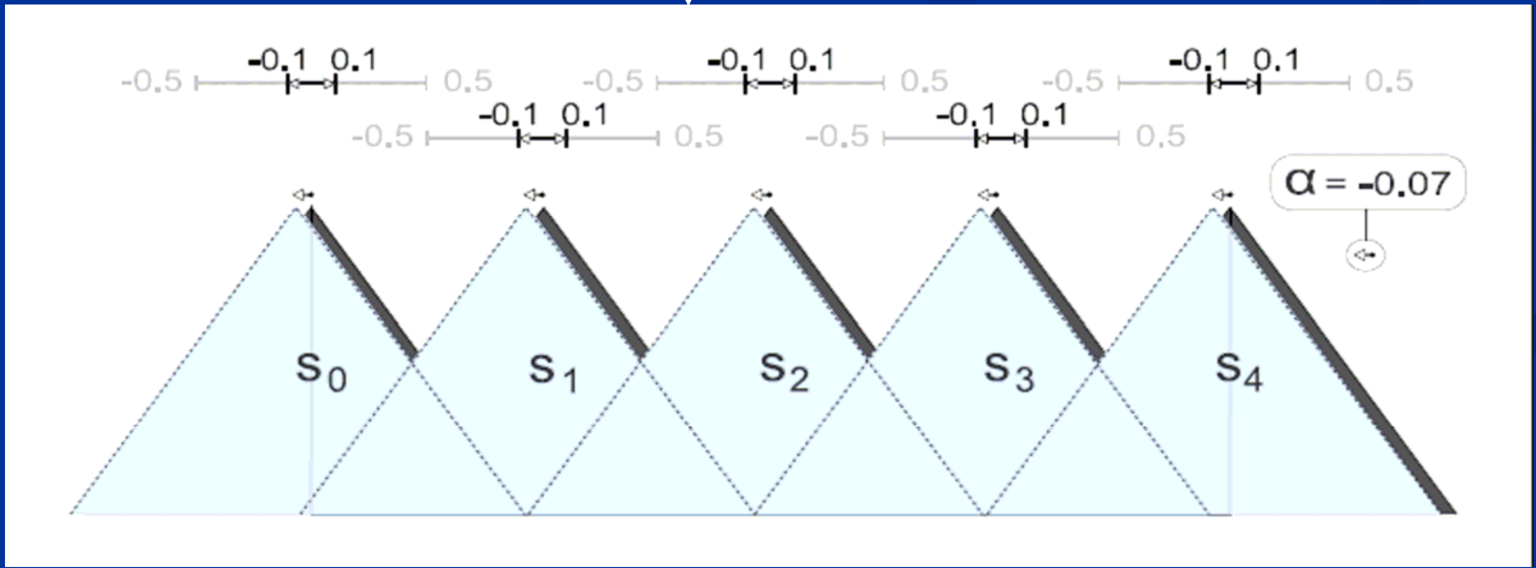
$C_1 = (L^1, \ldots, L^N)$

Integer Coding for Granularity Learning and

Input Variable Selection: $L^i \in \{1,\ldots,7\}$ for $i = 1 \ldots N-1$ and $L^N \in \{2,\ldots,7\}$

$C_2 = (\alpha^1, \ldots, \alpha^N)$

Real Coding for Lateral Displacement, where $\alpha^i \in [-0.1, 0.1]$



Multi-objective Evolutionary Fuzzy Systems (MOEFSs)

MOEFS

MOE Tuning

Given a KB, optimization is performed without changing the existing RB (typically, adjusting the DB definition)

MOE Learning

MOE DB Learning

Embedded learning: the DB generation process wraps an RB learning process.

MOE RB Selection

Selects relevant rules from a predefined RB

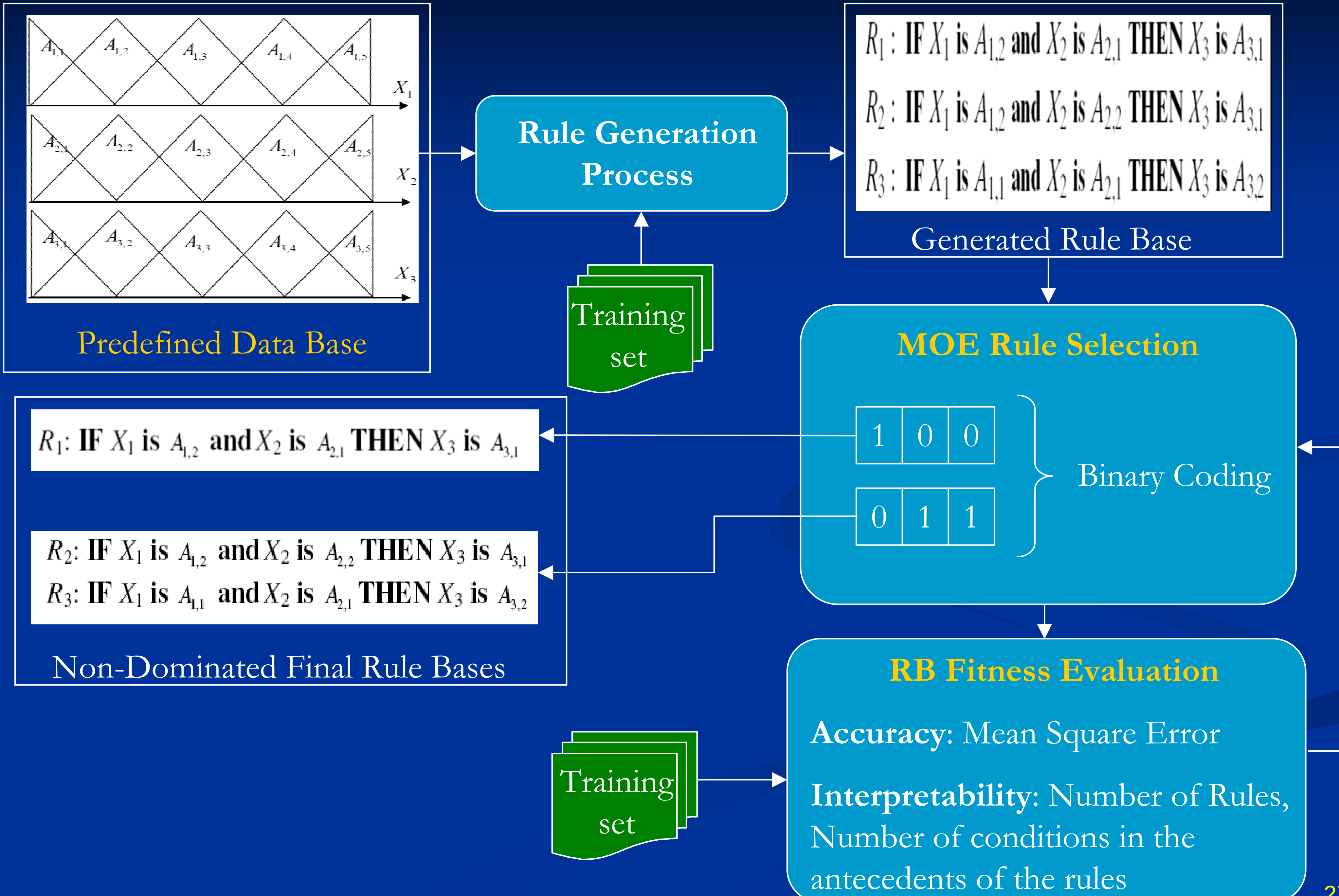
MOE RB Learning

Uses a predefined DB

MOE KB Learning

Simultaneous learning of RB and DB components

MOE Rule Selection - Description



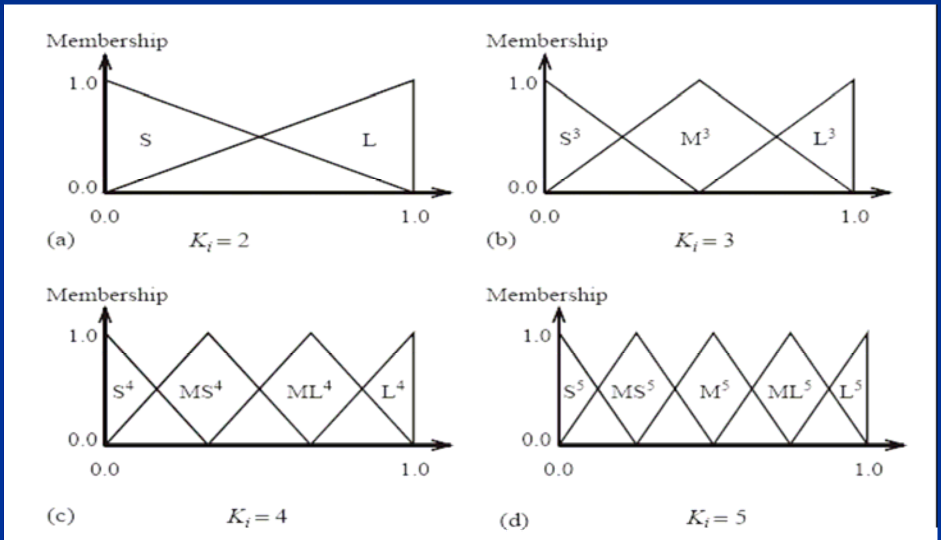
MOE Rule Selection – Example 1

H. Ishibuchi, T. Murata, and I. B. Turksen, "Single-objective and two objective genetic algorithms for selecting linguistic rules for pattern classification problems," Fuzzy Sets Syst., vol. 89, no. 2, pp. 135–150, 1997.

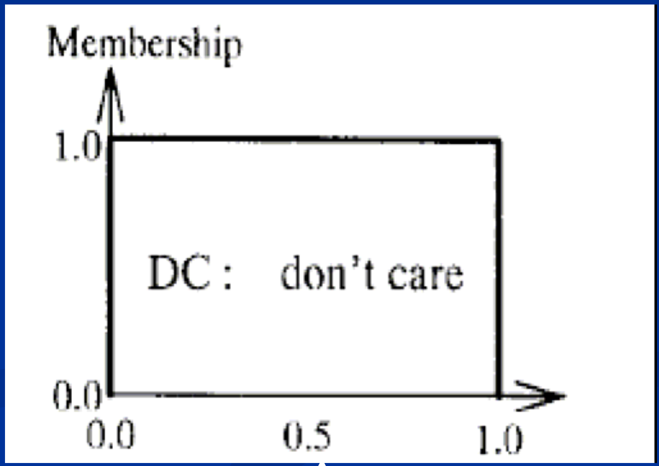
- **Pioneer** work in multi-objective evolutionary fuzzy systems
- Application to **Classification** problems
- **Selection based on a weighted fitness function (**Number of correctly classified training patterns and number of rules**)**
- Tentative set of **non-dominated solutions** preserved externally
- **Elitist strategy**: N_{elite} individuals of the population are randomly replaced with N_{elite} individuals randomly extracted from the tentative set of non-dominated solutions

MOE Rule Selection – Example 2

H. Ishibuchi, and T. Yamamoto, "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining," Fuzzy Sets and Systems, vol. 141, pp. 59–88, 2004.



Multiple Granularities for each Selected Rule



Don't Care Condition

Objectives:

Accuracy: *Percentage of Corrected Classified Patterns*

Interpretability: *Number of Rules, Total Rule Length*

Algorithm: Multi-objective Genetic Local Search (MOGLS)

Multi-objective Evolutionary Fuzzy Systems (MOEFSs)

MOEFS

MOE Tuning

Given a KB, optimization is performed without changing the existing RB (typically, adjusting the DB definition)

MOE Learning

MOE DB Learning

Embedded learning: the DB generation process wraps an RB learning process.

MOE RB Selection

Selects relevant rules from a predefined RB

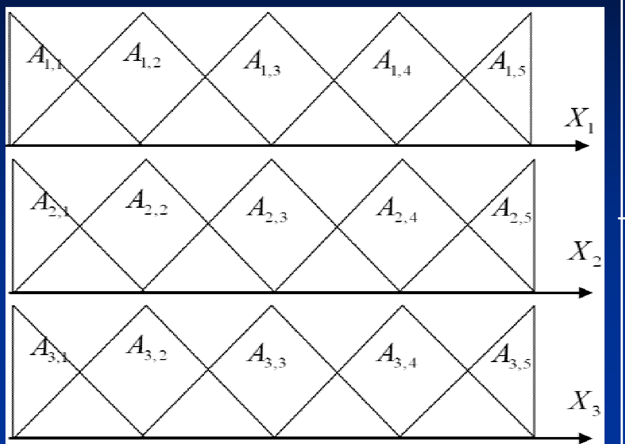
MOE RB Learning

Uses a predefined DB

MOE KB Learning

Simultaneous learning of RB and DB components

MOE Rule Selection and DB Tuning



Initial Data Base

MO-Evolutionary Rule Selection and DB Tuning

Double coding chromosome
Selected Rules: Binary coding
Membership Function Parameters: real coding

R_1 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

R_2 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3 : IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$

Initial Rule Base

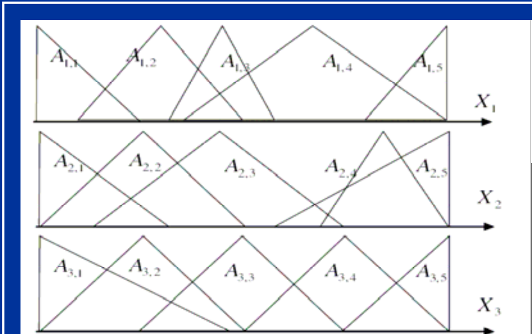
KB Fitness Evaluation

Accuracy: Mean Square Error

RB Interpretability: Number of Rules, Number of Conditions

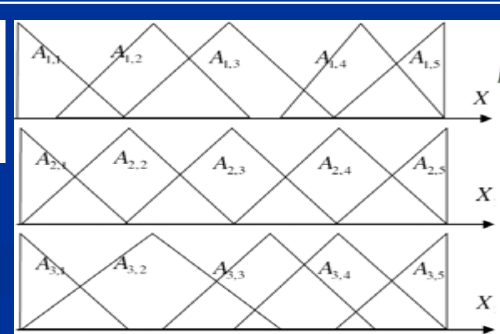
DB Interpretability: Integrity of the Partitions, Relative Measures

Training set



R_2: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3: IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$



R_1: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

Non-Dominated Final Knowledge Bases

MOE Rule Selection and DB Tuning – Example 1(a)

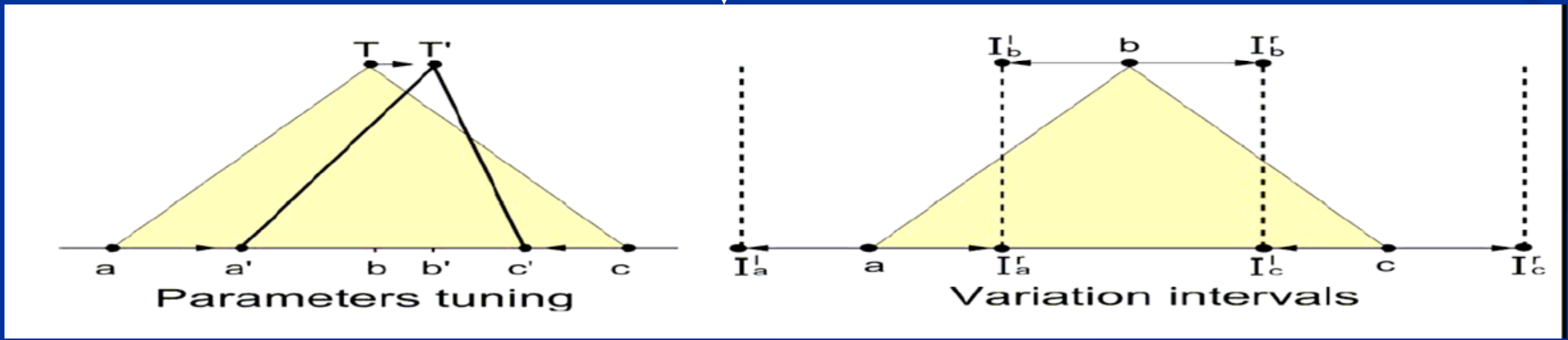
- *R. Alcala, M. J. Gacto, F. Herrera, and J. Alcala-Fdez, "A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems," Int. J. Uncertainty, Fuzziness Knowl.-Based Syst., vol. 15, no. 5, pp. 539–557, 2007.*
- *M. J. Gacto, R. Alcala, and F. Herrera, "Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems," Soft Comput., vol. 13, no. 5, pp. 419–436, 2009.*

Double coding scheme:

$C_T^p = C_1 C_2 \ldots C_n$ **Binary Coding** for Rule Selection

$$C_i = (a_1^i, b_1^i, c_1^i, \ldots, a_{m^i}^i, b_{m^i}^i, c_{m^i}^i), \qquad i = 1, \ldots, n.$$

Real Coding for Membership Function Parameters Tuning

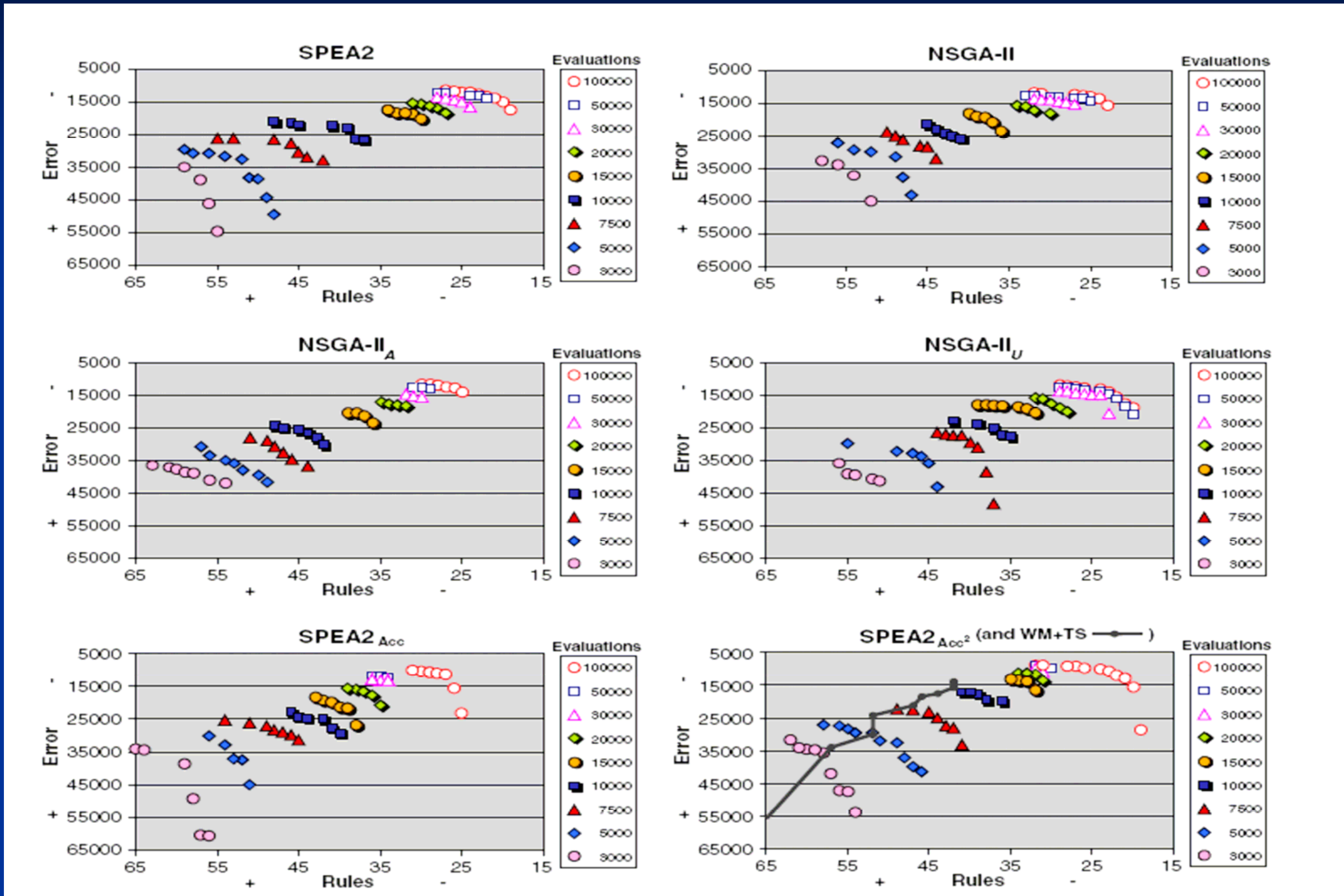


MOE Rule Selection and DB Tuning – Example 1(b)

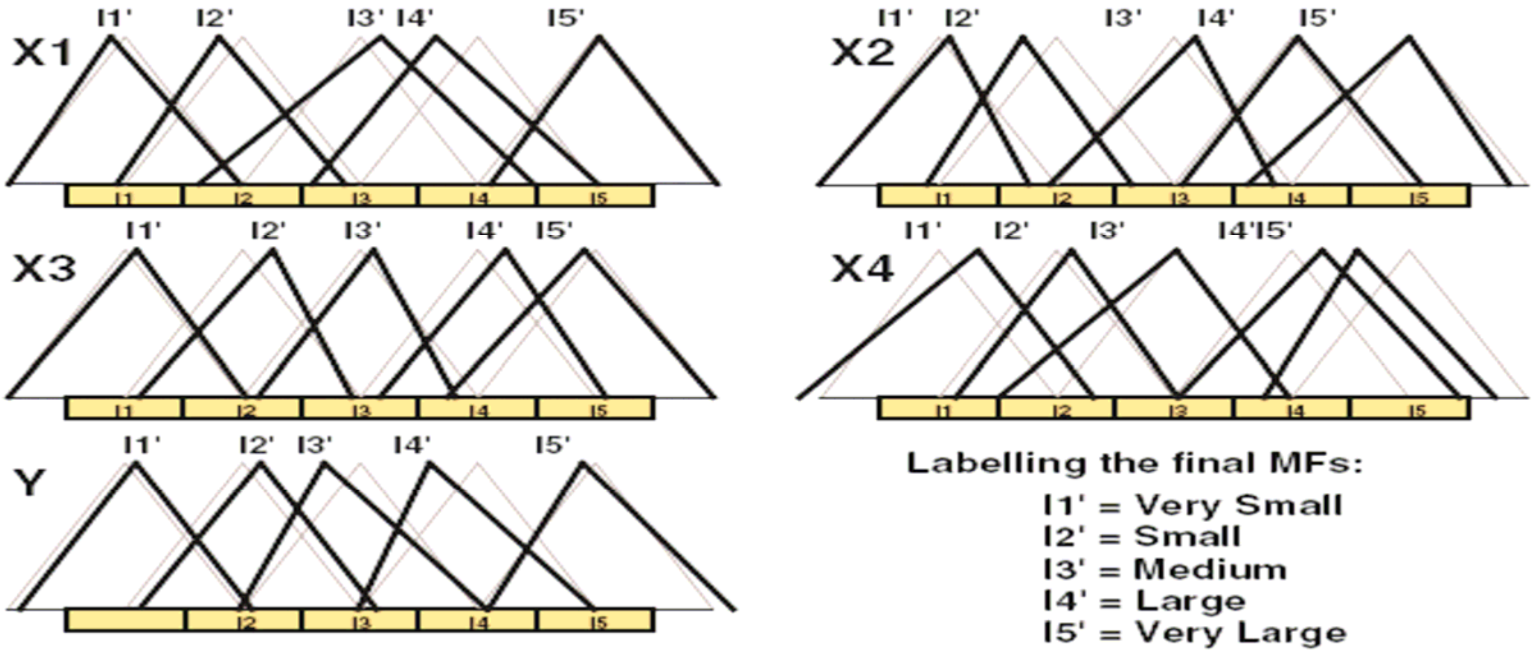
- R. *Alcala, M. J. Gacto, F. Herrera, and J. Alcala-Fdez, "A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems," Int. J. Uncertainty, Fuzziness Knowl.-Based Syst., vol. 15, no. 5, pp. 539–557, 2007.*
- *M. J. Gacto, R. Alcala, and F. Herrera, "Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems," Soft Comput., vol. 13, no. 5, pp. 419–436, 2009.*

- Ad-hoc modified versions of **NSGA-II** and **SPEA2** are applied to find sets of Mamdani FRBSs with different trade-offs between accuracy and interpretability
- **Accuracy** is evaluated in terms of mean square error
- **Rule Base Interpretability** is evaluated in terms of number of rules
- **Partition Intepretability** is ensured by using constraints for the variation intervals of the genes in the DB part of the chromosome

MOE Rule Selection and DB Tuning – Example 1(c)



MOE Rule Selection and DB Tuning – Example 1(d)



Labelling the final MFs:

l1' = Very Small
l2' = Small
l3' = Medium
l4' = Large
l5' = Very Large

#R: 43   *MSE-tra: 11383*   *MSE-tst: 13416*

X1	X2	X3	X4	Y
l1'	l1'	l1'	l1'	l1'
l1'	l1'	l1'	l2'	l2'
l1'	l2'	l1'	l1'	l1'
l1'	l2'	l2'	l2'	l2'
l2'	l1'	l1'	l1'	l1'
l2'	l1'	l1'	l2'	l2'
l2'	l1'	l2'	l1'	l2'
l2'	l1'	l2'	l2'	l2'
l2'	l2'	l1'	l1'	l1'
l2'	l2'	l1'	l2'	l2'
l2'	l2'	l2'	l1'	l2'
l2'	l2'	l2'	l2'	l2'
l2'	l3'	l3'	l2'	l3'
l3'	l2'	l1'	l1'	l1'

X1	X2	X3	X4	Y
l3'	l2'	l1'	l2'	l2'
l3'	l2'	l1'	l3'	l2'
l3'	l2'	l2'	l3'	l3'
l3'	l3'	l2'	l1'	l2'
l3'	l3'	l2'	l2'	l2'
l3'	l3'	l2'	l3'	l3'
l3'	l3'	l3'	l2'	l3'
l3'	l4'	l3'	l2'	l3'
l3'	l4'	l4'	l3'	l4'
l4'	l2'	l2'	l1'	l2'
l4'	l2'	l2'	l2'	l2'
l4'	l3'	l2'	l1'	l2'
l4'	l3'	l2'	l3'	l3'
l4'	l3'	l2'	l4'	l3'

X1	X2	X3	X4	Y
l4'	l3'	l3'	l2'	l3'
l4'	l3'	l3'	l3'	l4'
l4'	l4'	l3'	l1'	l3'
l4'	l4'	l3'	l4'	l4'
l4'	l4'	l4'	l2'	l4'
l4'	l4'	l4'	l4'	l5'
l4'	l5'	l4'	l2'	l3'
l4'	l5'	l4'	l3'	l4'
l4'	l5'	l5'	l2'	l5'
l4'	l5'	l5'	l3'	l5'
l5'	l2'	l2'	l5'	l4'
l5'	l2'	l3'	l2'	l3'
l5'	l2'	l3'	l5'	l4'
l5'	l4'	l3'	l4'	l4'
l5'	l4'	l3'	l5'	l5'

An Example of KB for the **ELE2 dataset**

MOE Rule Selection and DB Tuning – Example 2(a)

M. J. Gacto, R. Alcala, and F. Herrera, "Integration of an index to preserve the semantic interpretability in the multi-objective evolutionary rule selection and tuning of linguistic fuzzy systems," IEEE Trans. Fuzzy. Syst. vol. 18, n.3, pp. 515-531, 2010.

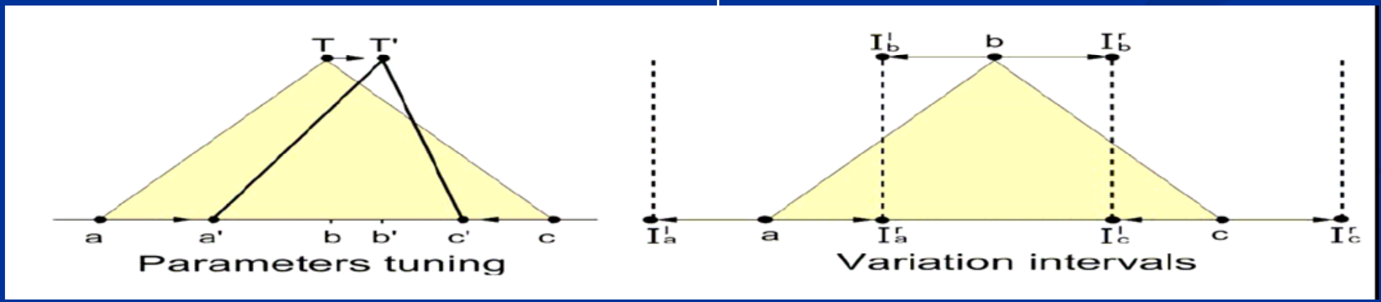
Double coding scheme:

$$C_T^p = C_1 C_2 \ldots C_n$$

Binary Coding for Rule Selection

$$C_i = (a_1^i, b_1^i, c_1^i, \ldots, a_{m^i}^i, b_{m^i}^i, c_{m^i}^i), \qquad i = 1, \ldots, n.$$

Real Coding for Membership Function Parameters Tuning



A new **semantic interpretability index** is exploited in the MOE rule selection and DB tuning process together with the **number of rules** and the **MSE**

MOE Rule Selection and DB Tuning – Example 2(b)

M. J. Gacto, R. Alcala, and F. Herrera, "Integration of an index to preserve the semantic interpretability in the multi-objective evolutionary rule selection and tuning of linguistic fuzzy systems," IEEE Trans. Fuzzy. Syst. vol. 18, n.3, pp. 515-531, 2010.

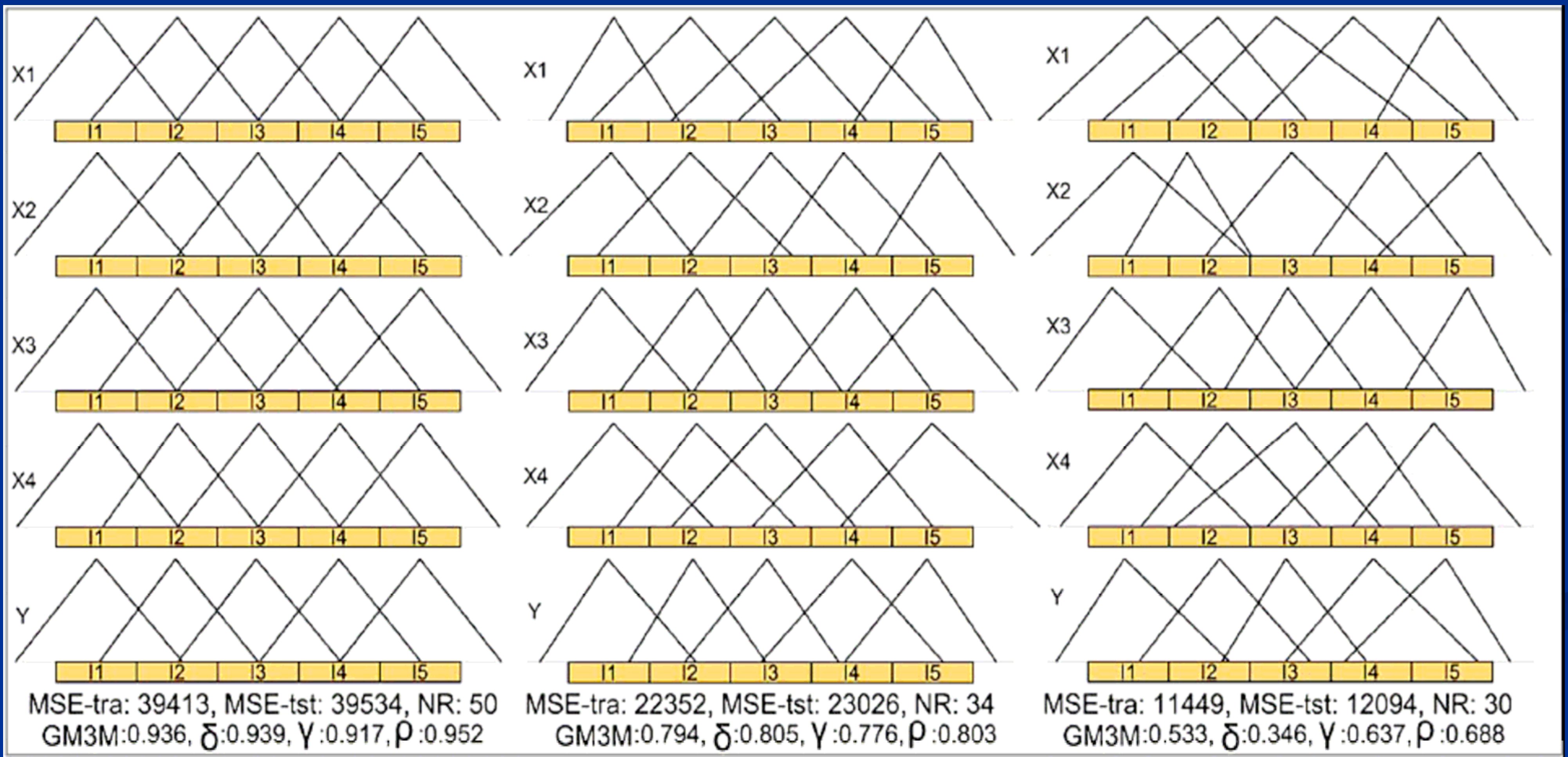
The semantic interpretability index **GM3M, defined in [0,1]** aggregates the following metrics:

- MF **centroids displacement** with respect to the original MFs
- MFs **lateral amplitude rate** (ratio between left and right parts of the MF supports equal to the original one)
- MFs **area similarity** between the new and original MF areas

A new version of **SPEA2,** which includes **incest prevention** and **restarting** strategies, is used to generate sets of Mamdani FRBSs with different trade-offs between accuracy, RB complexity and partition semantic interpretability

MOE Rule Selection and DB Tuning – Example 2(c)

M. J. Gacto, R. Alcala, and F. Herrera, "Integration of an index to preserve the semantic interpretability in the multi-objective evolutionary rule selection and tuning of linguistic fuzzy systems," IEEE Trans. Fuzzy. Syst. vol. 18, n.3, pp. 515-531, 2010.



Examples of DBs for the **ELE2 dataset**

Multi-objective Evolutionary Fuzzy Systems (MOEFSs)

MOEFS

MOE Tuning

Given a KB, optimization is performed without changing the existing RB (typically, adjusting the DB definition)

MOE Learning

MOE DB Learning

Embedded learning: the DB generation process wraps an RB learning process.

MOE RB Selection

Selects relevant rules from a predefined RB

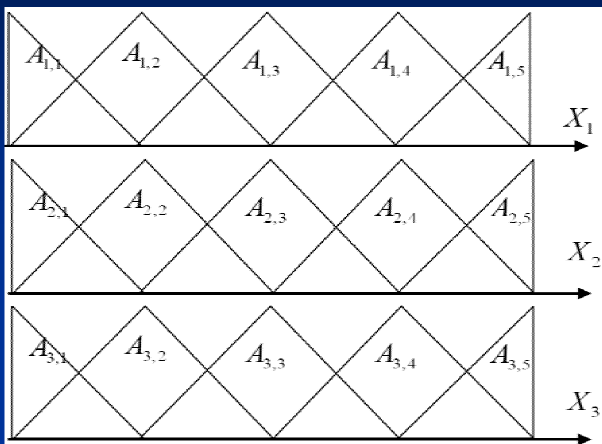
MOE RB Learning

Uses a predefined DB

MOE KB Learning

Simultaneous learning of RB and DB components

MOE Rule Learning - Description



Predefined Data Base

MOE Rule Learning

Integer chromosome coding

Pittsburgh Approach: Each chromosome codifies a rule base

Michigan Approach: Each chromosome codifies a rule

Training set

RB Fitness Evaluation

Accuracy: Mean Square Error

Interpretability: Number of Rules, Number of Conditions

R_1 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$
R_2 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$
R_3 : IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$

R_2: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$
R_3: IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$

R_1: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

Non-Dominated Final Rule Bases

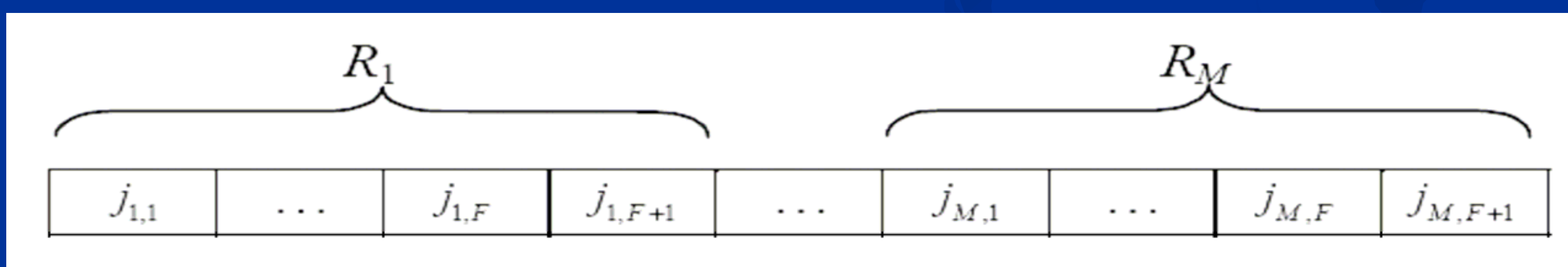
MOE Rule Learning – Example 1(a)

M. Cococcioni, P. Ducange, B. Lazzerini, and F. Marcelloni, "A Pareto-based multi-objective evolutionary approach to the identification of Mamdani fuzzy systems," Soft Comp.,vol. 11, no. 11, pp. 1013–1031, 2007.

Given a **Mamdani** fuzzy rule:

- R_m: ***IF*** X_1 ***is*** $A_{1,j_{m,1}}$ ***and … and*** X_F ***is*** $A_{F,j_{m,F}}$ ***THEN*** X_{F+1} ***is*** $A_{F+1,j_{m,F+1}}$

Each RB can be codified with an integer chromosome C :



$j_{m,f}$ identifies the index of the fuzzy set which has been selected for variable X_f in rule R_m

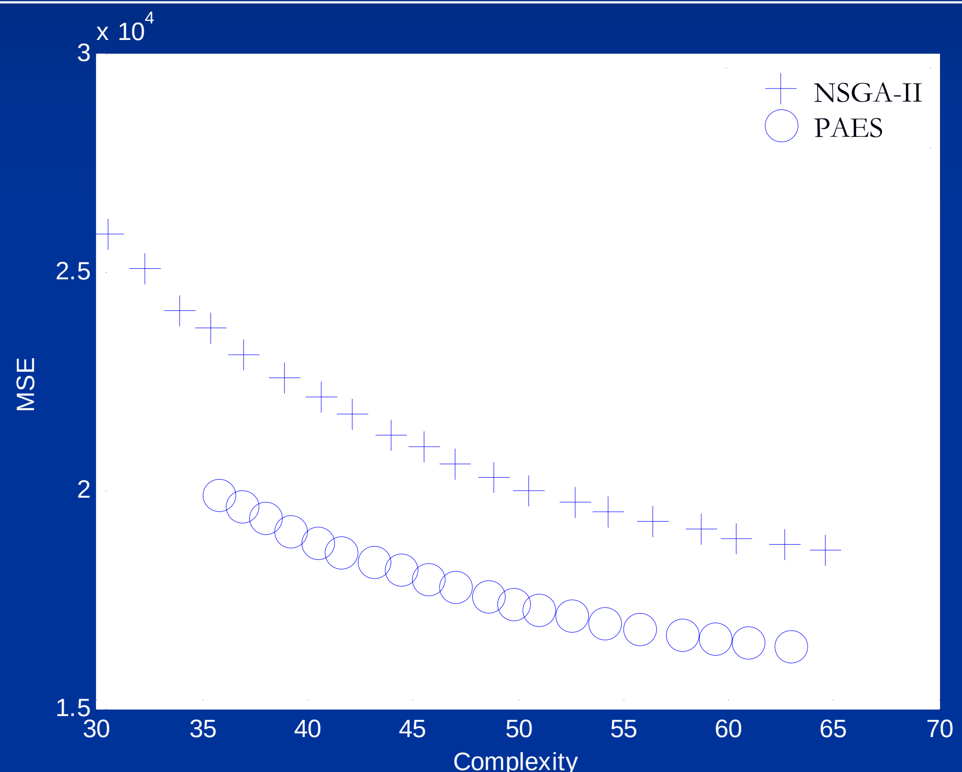
MOE Rule Learning – Example 1(b)

M. Cococcioni, P. Ducange, B. Lazzerini, and F. Marcelloni, "A Pareto-based multi-objective evolutionary approach to the identification of Mamdani fuzzy systems," Soft Comp.,vol. 11, no. 11, pp. 1013–1031, 2007.

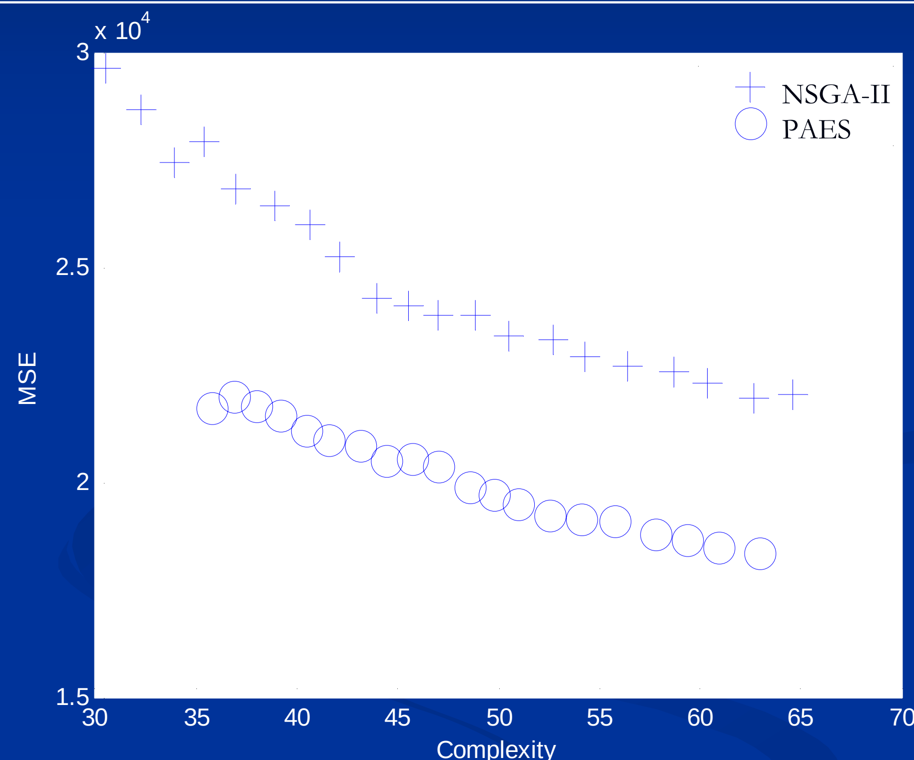
- **The interpretability** is calculated in terms of complexity of the RB, i.e., as the sum of the conditions in the overall RB
- **Accuracy** is evaluated in terms of mean square error
- **Ad-hoc genetic operators** are implemented for the specific integer chromosome coding
- A modified version of the well-known **2+2 PAES** is proposed (**(2+2)M-PAES**)

MOE Rule Learning – Example 1(c)

M. Cococcioni, P. Ducange, B. Lazzerini, and F. Marcelloni, "A Pareto-Based Multi-Objective Evolutionary Approach to the Identification of Mamdani Fuzzy Systems," Soft Comp.,vol. 11, no. 11, pp. 1013–1031, 2007.



Average Pareto fronts on Training Sets



Average Pareto fronts on Test Sets

- ✓ Results on a real word regression problem that consists of estimating the maintenance costs of medium voltage lines in some Spanish towns (**ELE2 dataset**)
- ✓ The data set contains 1059 input-output patterns described with 4 features

MOE Rule Learning – Example 2(a)

J. Casillas, P. Martínez, and A.D. Benítez, "Learning consistent, complete and compact sets of fuzzy rules in conjunctive normal form for regression problems," Soft Computing, vol. 13, n. 5, pp 451-465, 2009

This work deals with DNF-type fuzzy rules

$$\text{IF } X_1 \text{ is } \widetilde{A_1} \text{ and } \ldots \text{ and } X_n \text{ is } \widetilde{A_n} \text{ THEN } Y \text{ is } B$$

where X_i takes as values a set of linguistic terms:

$$\widetilde{A}_i = \{A_{i1} \ or \ldots or \ A_{il_i}\}$$

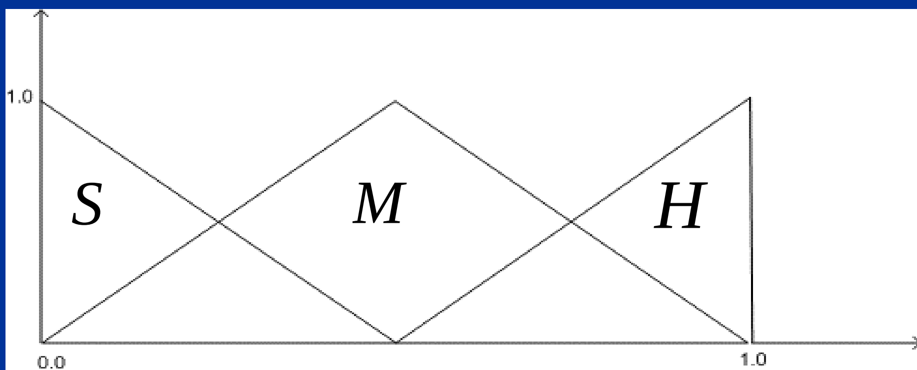
***T*-conorm**

- Each rule is encoded by a **binary string** for the antecendent and an **integer coding** scheme for the consequent
- The chromosome is formed by a concatenation of encoded rules

MOE Rule Learning – Example 2(b)

J. Casillas, P. Martínez, and A.D. Benítez, "Learning consistent, complete and compact sets of fuzzy rules in conjunctive normal form for regression problems," Soft Computing, vol. 13, n. 5, pp 451-465, 2009

An example of rule coding



Fuzzy partition for each linguistic variable

[IF X_1 is S and X_2 is {M or L} THEN Y is M]

Rule

[100|011||2]

Encoded rule

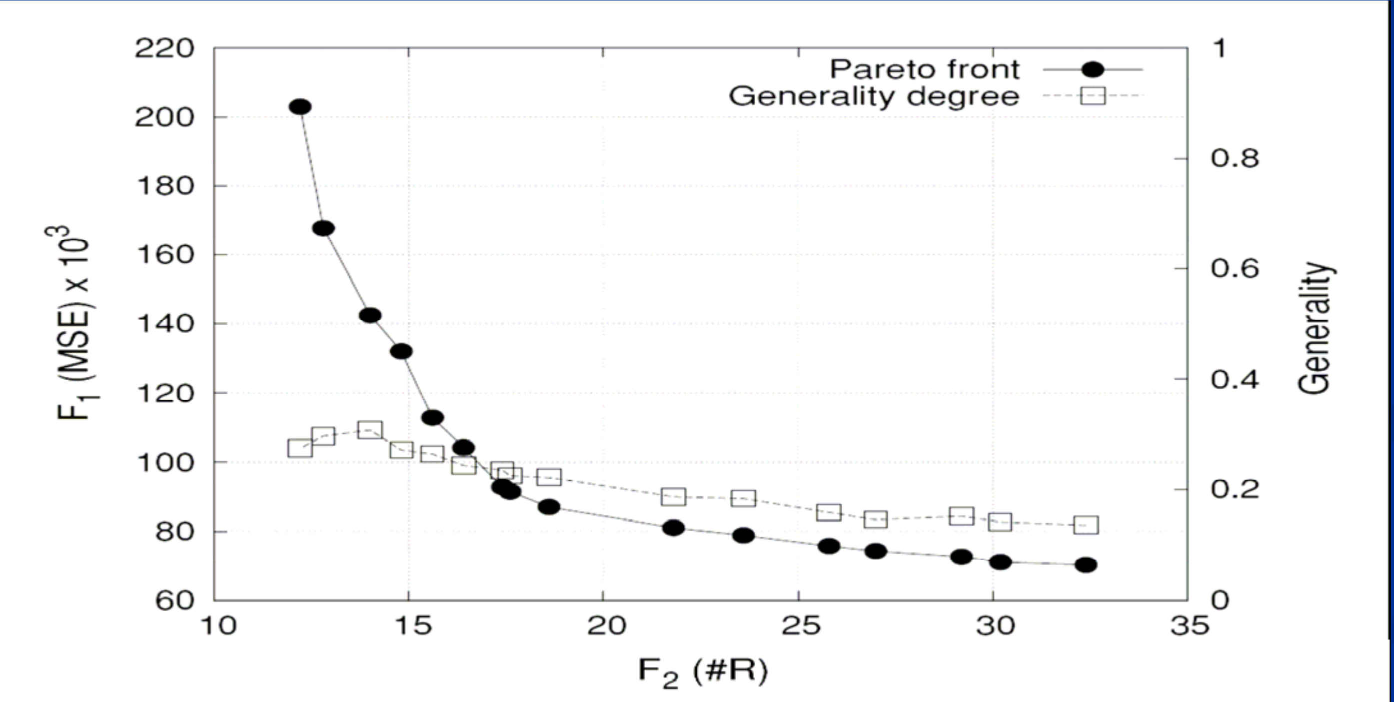
MOE Rule Learning – Example 2(c)

J. Casillas, P. Martínez, and A.D. Benítez, "Learning consistent, complete and compact sets of fuzzy rules in conjunctive normal form for regression problems," Soft Computing, vol. 13, n. 5, pp 451-465, 2009

- A set of non-dominated RBs are generated by using NSGA-II
- **Accuracy** and **interpretability** are evaluated in terms of mean square error and number of rules, respectively
- **Ad-hoc genetic operators** are implemented for the specific mixed coding chromosome
- The **Wang and Mendel** algorithm is exploited for generating the initial population
- The proposed method allows controlling the **consistency**, the **completeness**, the **compactness** and the **over-generality** of the generated RBs

MOE Rule Learning – Example 2(d)

J. Casillas, P. Martínez, and A.D. Benítez, "Learning consistent, complete and compact sets of fuzzy rules in conjunctive normal form for regression problems," Soft Computing, vol. 13, n. 5, pp 451-465, 2009



Average Pareto front (*solid circles*) and generality degrees (*empty squares*) obtained in the Ele2 problem

Generality degree: mean number of linguistic terms used per variable in each rule
0 maximum specificity 1 maximum generality

Multi-objective Evolutionary Fuzzy Systems (MOEFSs)

MOEFS

MOE Tuning

Given a KB, optimization is performed without changing the existing RB (typically, adjusting the DB definition)

MOE Learning

MOE DB Learning

Embedded learning: the DB generation process wraps an RB learning process.

MOE RB Selection

Selects relevant rules from a predefined RB

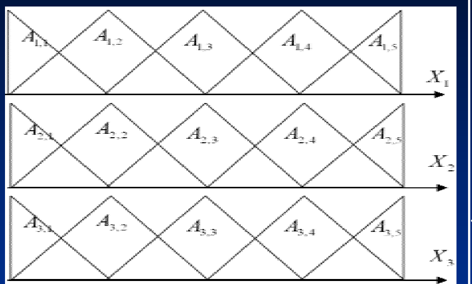
MOE RB Learning

Uses a predefined DB

MOE KB Learning

Simultaneous learning of RB and DB components

MOE Knowledge Base Learning



Initial Data Base

Optional

R_1 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

R_2 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3 : IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$

Initial Rule Base

Optional

MOE RB, Granularity and MF Parameters Learning

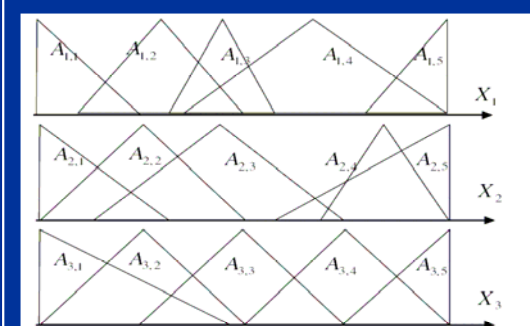
The chromosomes can be formed by different parts

- **RB part**: Integer
- **Granularity part**: Integer
- **MF parameter part**: Real

Training set

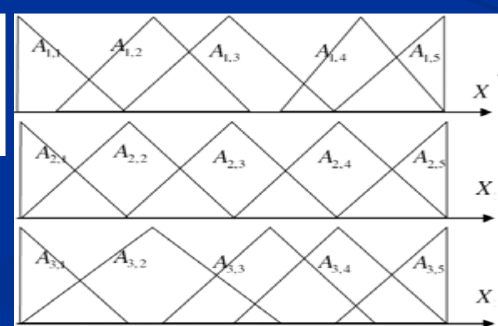
KB Fitness Evaluation

- **Accuracy**: Mean Square Error
- **RB Interpretability**: Number of Rules, Number of Conditions
- **DB Interpretability**: Integrity of the Partitions, Relative Measures



R_2: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3: IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$



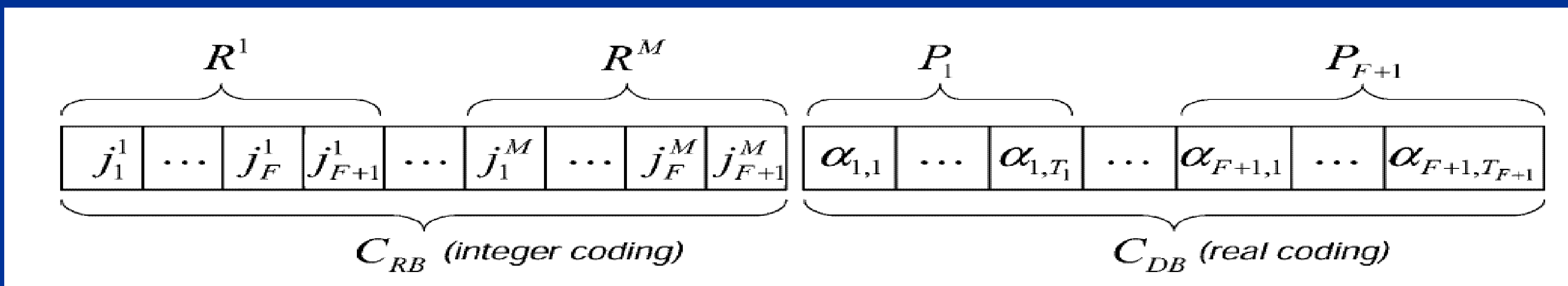
R_1: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

Non-Dominated Final Knowledge Bases

MOE Knowledge Base Learning – Example 1(a)

R. Alcalá, P. Ducange, F. Herrera, B. Lazzerini, and F. Marcelloni, "A Multi-objective evolutionary approach to concurrently learn rule and data bases of linguistic fuzzy rule-based systems," IEEE Trans. Fuzzy. Syst., vol. 17, n. 5, pp. 1106–1122, 2009.

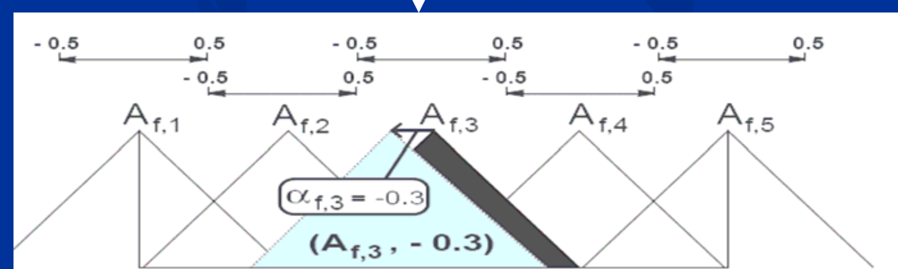
Coding scheme:



R_1 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

R_2 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3 : IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$

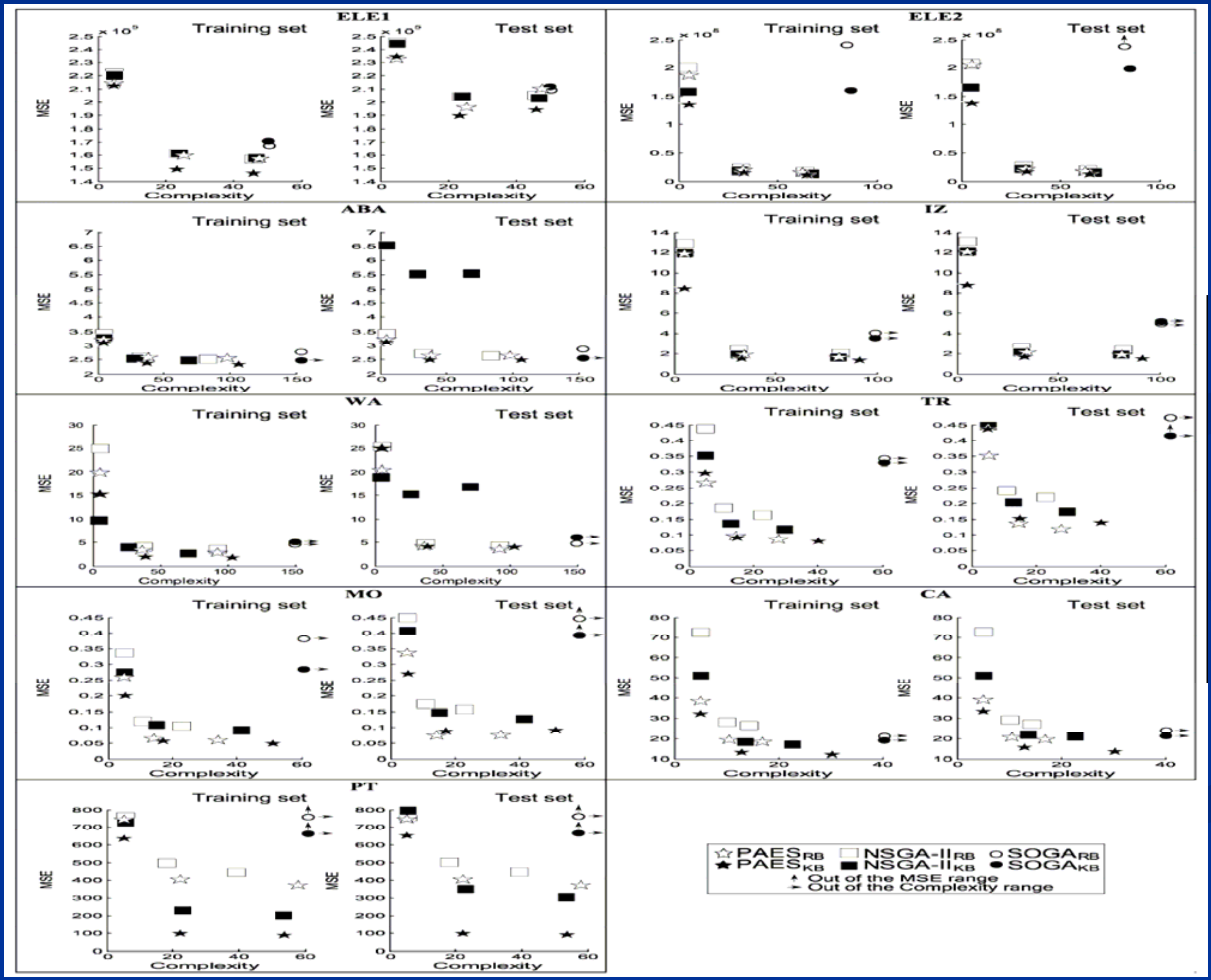


The symbolic translation of an MF is performed by means of the 2-tuple approach:

$$(A_{f,j};\alpha_{f,j}), A_{f,j} \in P_f, \alpha_{f,j} \in [-0.5, 0.5)$$

MOE Knowledge Base Learning – Example 1(b)

R. *Alcalá, P. Ducange, F. Herrera, B. Lazzerini, and F. Marcelloni, "A Multi-objective evolutionary approach to concurrently learn rule and data bases of linguistic fuzzy rule-based systems,"* IEEE Trans. Fuzzy. Syst., vol. 17, n. 5, pp. 1106–1122, 2009.



MOE Knowledge Base Learning – Example 1(c)

R. Alcalá, P. Ducange, F. Herrera, B. Lazzerini, and F. Marcelloni, "A Multi-objective evolutionary approach to concurrently learn rule and data bases of linguistic fuzzy rule-based systems," IEEE Trans. Fuzzy. Syst., vol. 17, n. 5, pp. 1106–1122, 2009.

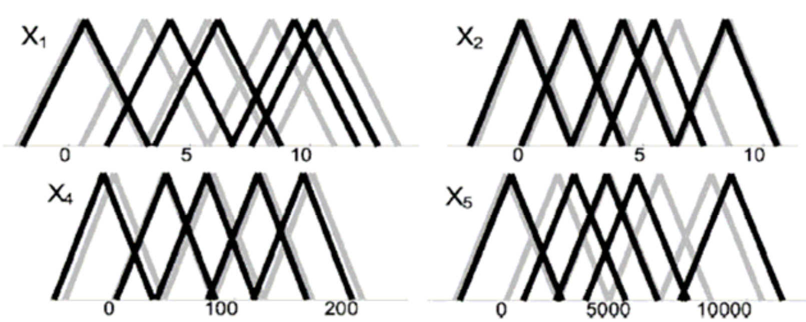
ELE2 dataset

X_1	X_2	X_3	X_4	X_5	X_1	X_2	X_3	X_4	X_5
VH	-	LO	VH	VH	-	-	**HI**	**LO**	**MM**
-	-	-	**VH**	**VH**	-	VH	MM	-	MM
-	-	**VH**	**HI**	**VH**	-	-	LO	LO	LO
HI	-	VH	-	VH	-	-	-	**LO**	**LO**
-	-	**VH**	-	**VH**	HI	-	-	VL	LO
-	HI	HI	-	VH	-	**HI**	**LO**	-	**LO**
HI	LO	MM	-	VH	-	VL	-	MM	VL
-	**VH**	-	-	**VH**	-	VL	LO	VL	VL
-	-	MM	HI	HI	-	-	LO	VL	VL
-	-	VL	HI	HI	LO	MM	-	VL	VL
-	-	-	HI	HI	LO	HI	VL	-	VL
-	-	-	MM	HI	-	HI	VL	-	VL
LO	LO	MM	-	HI	-	MM	VL	-	VL
-	-	LO	-	HI	-	LO	-	-	VL
-	-	**MM**	**MM**	**MM**	-	VL	-	-	VL

Rule Base with $PAES_{KB}$

X_1	X_2	X_3	X_4	X_5	X_1	X_2	X_3	X_4	X_5
HI	VH	MM	VH	VH	VH	MM	-	VL	MM
VH	-	-	VH	VH	-	-	LO	MM	LO
-	-	-	**VH**	**VH**	LO	VL	LO	LO	LO
VH	MM	-	MM	VH	MM	-	VL	LO	LO
-	VH	VH	-	VH	-	-	VL	LO	LO
-	-	**VH**	**HI**	**VH**	-	-	-	**LO**	**LO**
-	-	**VH**	-	**VH**	-	MM	-	VL	LO
-	-	HI	-	VH	-	**HI**	**LO**	-	**LO**
-	**VH**	-	-	**VH**	-	-	LO	-	LO
-	-	-	HI	HI	VH	LO	VL	-	LO
-	-	-	MM	HI	LO	HI	-	MM	VL
-	-	MM	-	HI	-	-	VL	VL	VL
-	-	**MM**	**MM**	**MM**	VH	HI	-	VL	VL
-	-	**HI**	**LO**	**MM**	-	VH	MM	-	VL
HI	-	LO	LO	MM	-	-	VL	-	VL

Rule Base with $PAES_{RB}$



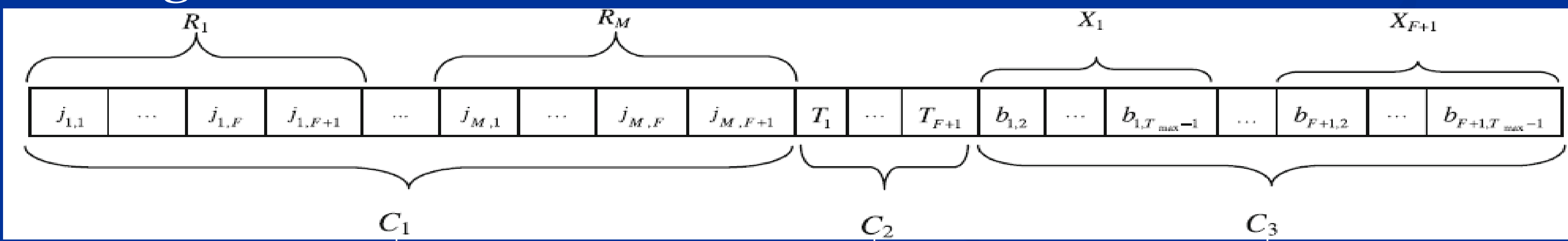
	MSE_{Tr}	MSE_{Ts}	C	R
$PAES_{KB}$	12130.31	13429.70	57	30
$PAES_{RB}$	16377.47	17099.98	60	30

Fig. 6. Example of KBs generated by $PAES_{KB}$ and $PAES_{RB}$ on the same data partition and seed (DBs with and without MF parameter learning are represented in black and gray, respectively). Rules that appear in both the KBs are in bold.

MOE Knowledge Base Learning – Example 2(a)

-M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, "Learning concurrently partition granularities and rule bases of Mamdani fuzzy systems in a multi-objective evolutionary framework," Int. J. Approx. Reason., vol. 50, n. 7, pp. 1066–1080, 2009.
-M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, "Multi-objective evolutionary learning of granularity, membership function parameters and rules of Mamdani fuzzy systems," Evolutionary Intelligence, vol. 2, n. 1-2, pp. 21–37, 2009.

Coding scheme:

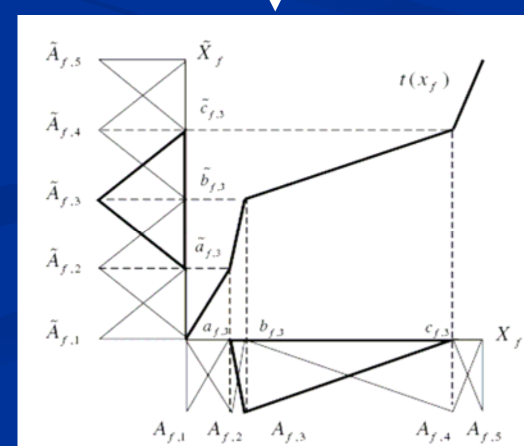


R_1 : IF X_1 is $\tilde{A}_{1,5}$ and X_2 is $\tilde{A}_{2,5}$ THEN X_3 is $\tilde{A}_{3,1}$

R_2 : IF X_1 is $\tilde{A}_{1,4}$ and X_2 is $\tilde{A}_{2,4}$ THEN X_3 is $\tilde{A}_{3,1}$

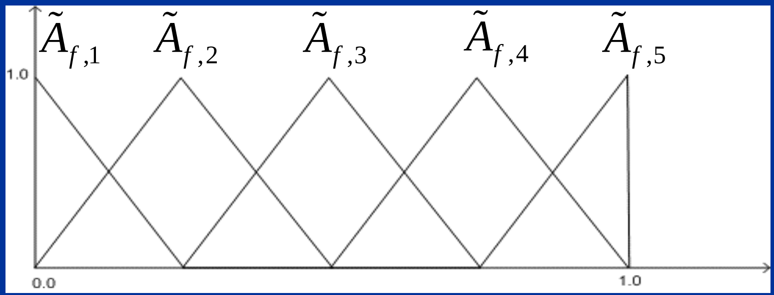
R_3 : IF X_1 is $\tilde{A}_{1,1}$ and X_2 is $\tilde{A}_{2,1}$ THEN X_3 is $\tilde{A}_{3,5}$

Integer genes for the granularity

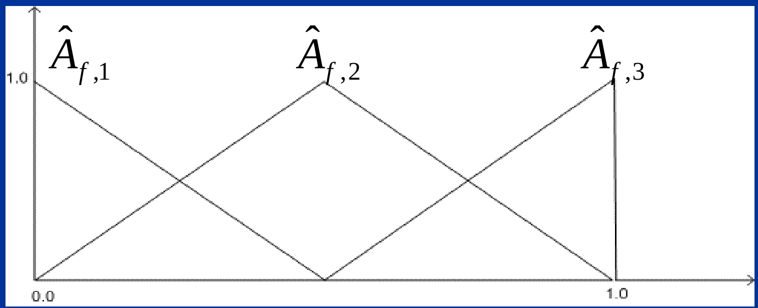


MOE Knowledge Base Learning – Example 2(d)

-M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, "Learning concurrently partition granularities and rule bases of Mamdani fuzzy systems in a multi-objective evolutionary framework," Int. J. Approx. Reason., vol. 50, n. 7, pp. 1066–1080, 2009.
-M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, "Multi-objective evolutionary learning of granularity, membership function parameters and rules of Mamdani fuzzy systems," Evolutionary Intelligence, vol. 2, n. 1-2, pp. 21–37, 2009.



Virtual Partition



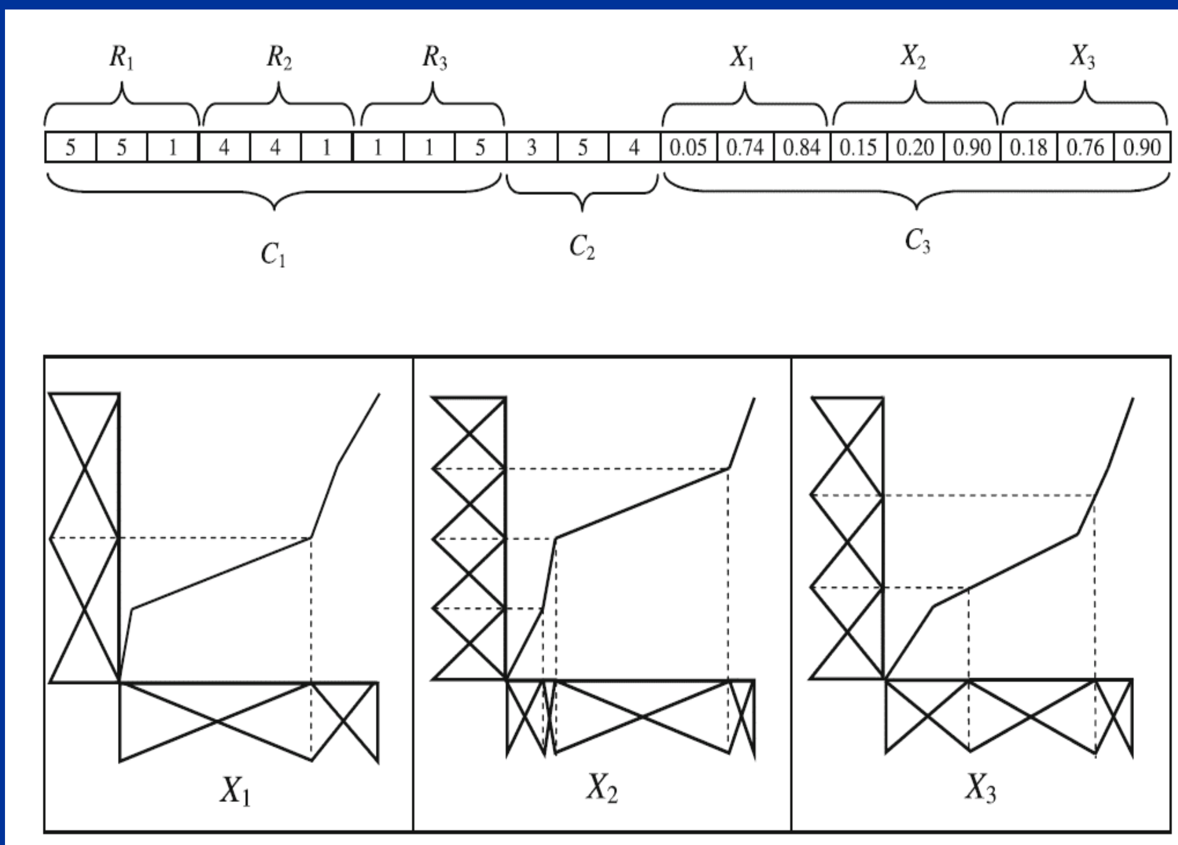
Actual Partition

- "**if** X_f **is** $\tilde{A}_{f,3}$" $\Rightarrow$ "**if** X_f **is** $\hat{A}_{f,2}$"
- "**if** X_f **is** $\tilde{A}_{f,4}$" $\Rightarrow$ "**if** X_f **is** $\hat{A}_{f,2}$ or $\hat{A}_{f,3}$"

MOE Knowledge Base Learning – Example 2(e)

-M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, "Learning concurrently partition granularities and rule bases of Mamdani fuzzy systems in a multi-objective evolutionary framework," Int. J. Approx. Reason., vol. 50, n. 7, pp. 1066–1080, 2009.

-M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, "Multi-objective evolutionary learning of granularity, membership function parameters and rules of Mamdani fuzzy systems," Evolutionary Intelligence, vol. 2, n. 1-2, pp. 21–37, 2009.



R_1 : IF X_1 is $\tilde{A}_{1,5}$ and X_2 is $\tilde{A}_{2,5}$ THEN X_3 is $\tilde{A}_{3,1}$

R_2 : IF X_1 is $\tilde{A}_{1,4}$ and X_2 is $\tilde{A}_{2,4}$ THEN X_3 is $\tilde{A}_{3,1}$

R_3 : IF X_1 is $\tilde{A}_{1,1}$ and X_2 is $\tilde{A}_{2,1}$ THEN X_3 is $\tilde{A}_{3,5}$

Virtual Rule Base

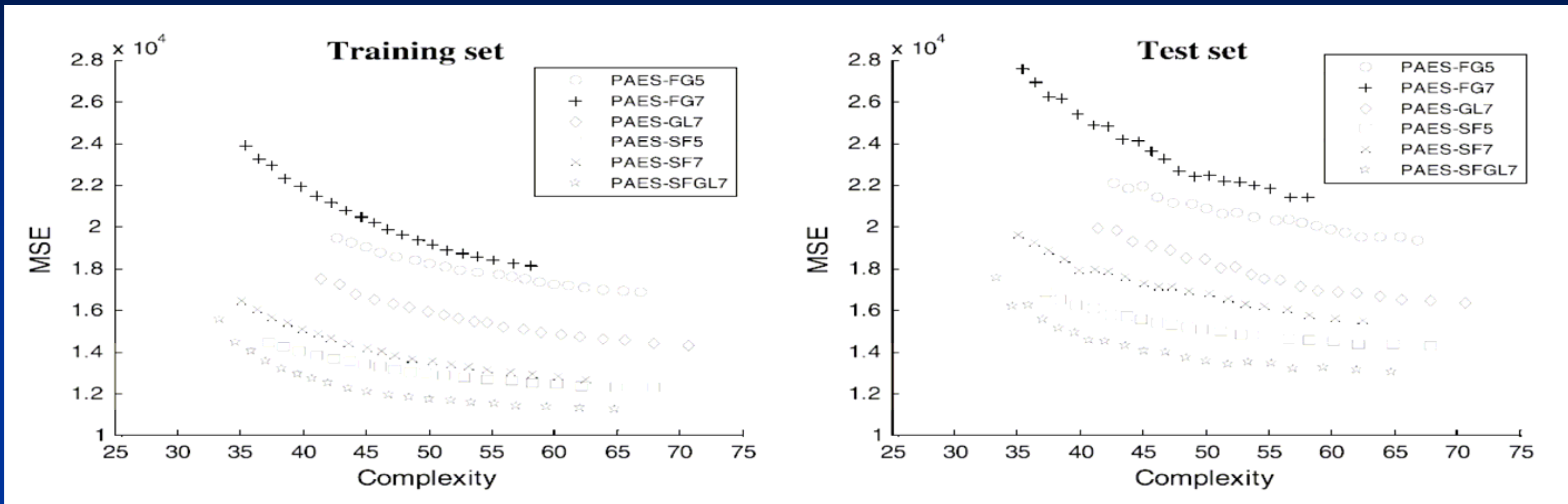
R_1 : IF X_1 is $A_{1,3}$ and X_2 is $A_{2,5}$ THEN X_3 is $A_{3,1}$

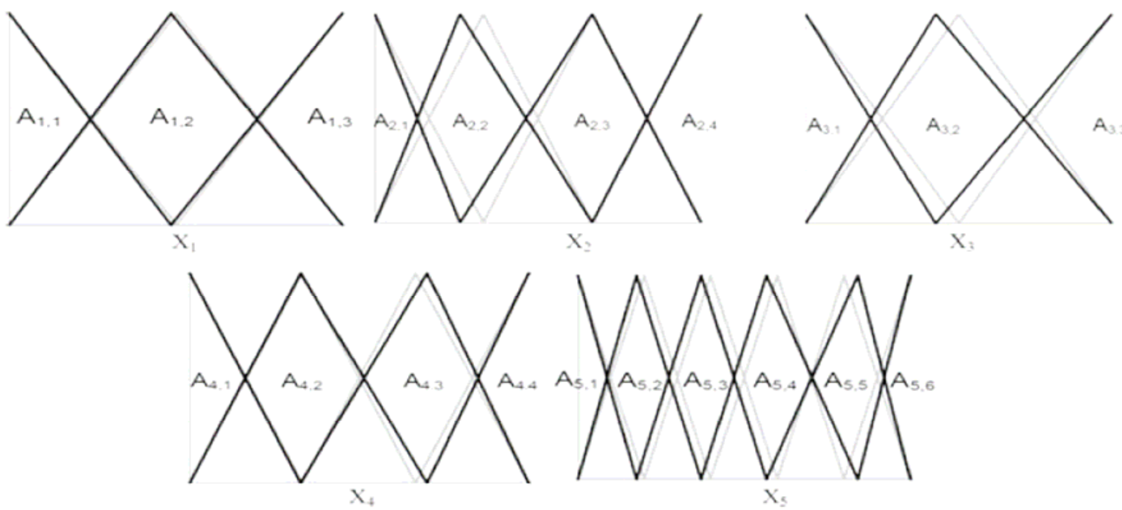
R_2 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,4}$ THEN X_3 is $A_{3,1}$

R_3 : IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,4}$

Actual Rule Base

MOE Knowledge Base Learning – Example 2(f)





(a)

X_1	X_2	X_3	X_4	X_5
–	$A_{2,3}$	$A_{3,1}$	$A_{4,4}$	$A_{5,6}$
–	–	–	$A_{4,4}$	$A_{5,6}$
$A_{1,2}$	$A_{2,3}$	$A_{3,3}$	$A_{4,2}$	$A_{5,6}$
–	–	$A_{3,3}$	–	$A_{5,6}$
$A_{1,3}$	$A_{2,3}$	$A_{3,3}$	–	$A_{5,6}$
–	$A_{2,2}$	$A_{3,2}$	$A_{4,3}$	$A_{5,5}$
$A_{1,3}$	$A_{2,1}$	$A_{3,2}$	$A_{4,2}$	$A_{5,4}$
$A_{1,3}$	–	$A_{3,2}$	$A_{4,1}$	$A_{5,4}$
–	–	$A_{3,2}$	–	$A_{5,4}$
$A_{1,1}$	–	$A_{3,1}$	$A_{4,2}$	$A_{5,3}$
–	$A_{2,3}$	–	$A_{4,2}$	$A_{5,3}$
$A_{1,2}$	$A_{2,2}$	–	$A_{4,2}$	$A_{5,3}$
$A_{1,2}$	$A_{2,3}$	$A_{3,1}$	$A_{4,3}$	$A_{5,2}$
–	$A_{2,4}$	$A_{3,2}$	$A_{4,1}$	$A_{5,2}$
–	–	–	$A_{4,1}$	$A_{5,1}$
–	–	$A_{3,1}$	–	$A_{5,1}$

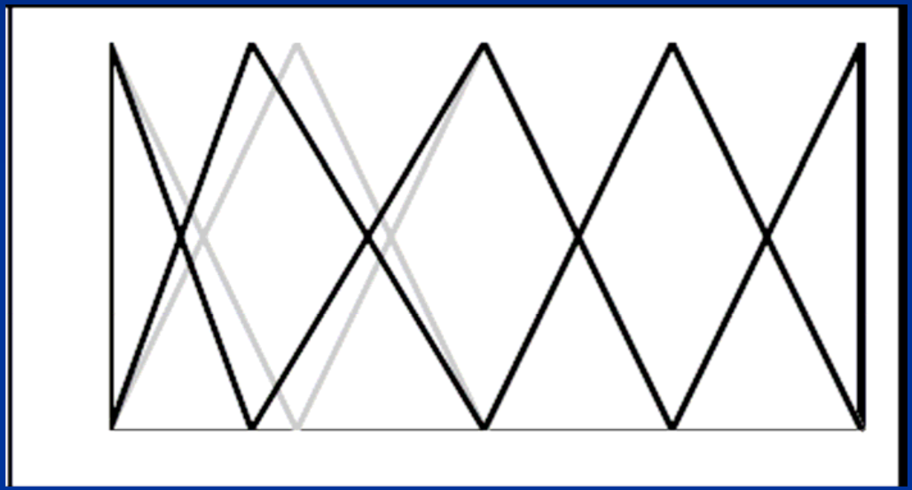
(b)

An example of DB (a) and RB (b) generated by PAES-SFGL7 on a representative trial

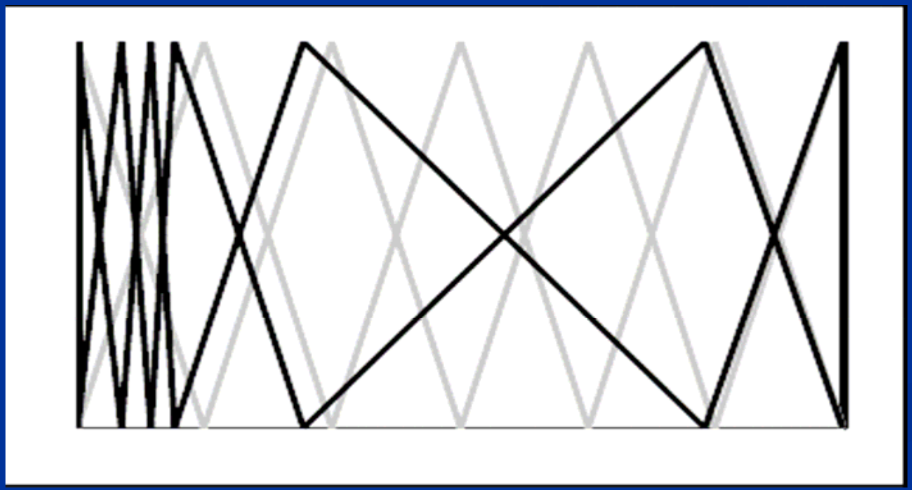
Results on the **ELE2 dataset**

MOE Knowledge Base Learning – Example 3(a)

M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, "Learning knowledge bases of multi-objective evolutionary fuzzy systems by simultaneously optimizing accuracy, complexity and partition integrity, Soft Computing, DOI: 10.1007/s00500-010-0665-0, in press.



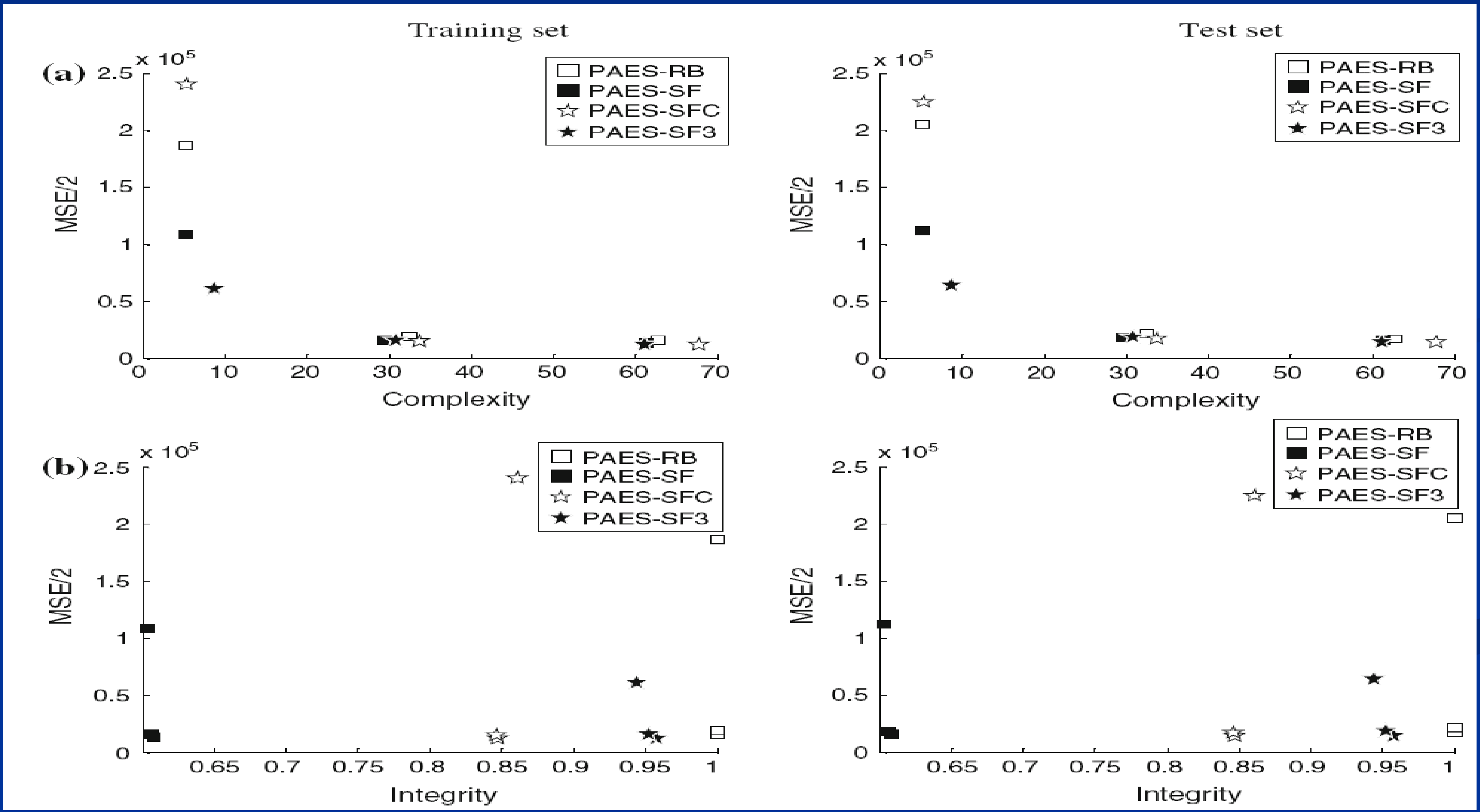
Fuzzy Strong Partitions with a high regularity level



Fuzzy Strong Partitions with a low regularity level

MOE Knowledge Base Learning – Example 3(b)

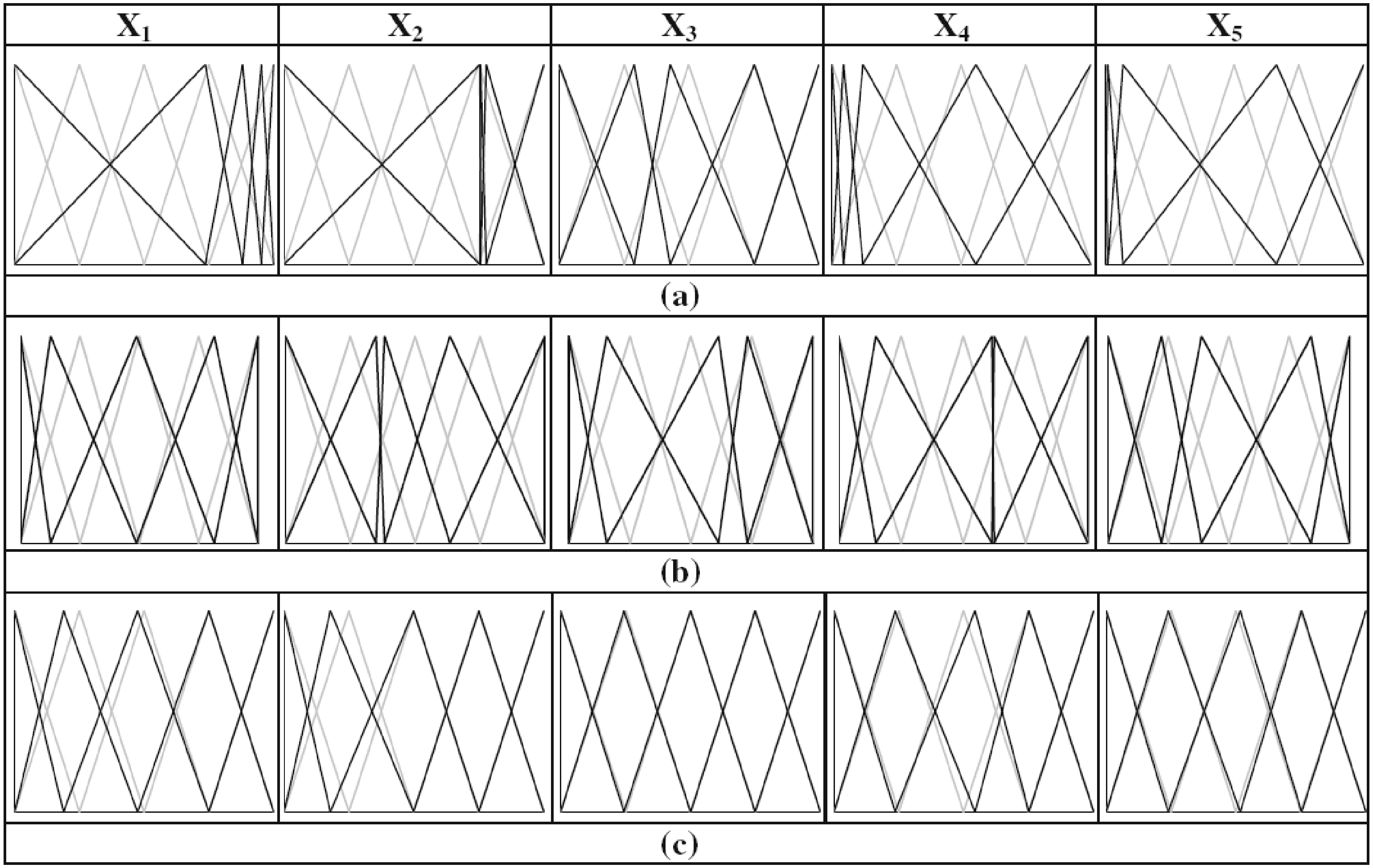
M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, 'Learning knowledge bases of multi-objective evolutionary fuzzy systems by simultaneously optimizing accuracy, complexity and partition integrity, Soft Computing, DOI: 10.1007/s00500-010-0665-0, in press.



Results on the **ELE2 dataset**

MOE Knowledge Base Learning – Example 3(c)

M. Antonelli, P. Ducange, B. Lazzerini, and F. Marcelloni, "Learning knowledge bases of multi-objective evolutionary fuzzy systems by simultaneously optimizing accuracy, complexity and partition integrity, Soft Computing, DOI: 10.1007/s00500-010-0665-0, in press.



Three examples of DB for one of the most accurate MFRBSs generated on a fold for the ELE2 dataset by, respectively, PAES-SF (**a**), PAES-SFC (**b**) and PAES-SF3 (**c**)

MOE Knowledge Base Learning – Example 4(a)

P. Pulkkinen, and H. Koivisto, "A dynamically constrained multiobjective genetic fuzzy system for regression problems," IEEE Trans. Fuzzy. Syst., vol. 18, n. 1, pp. 161–177, 2010.

Coding scheme:

Rule Base: Integer genes

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n_s} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n_s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{R,1} & a_{R,2} & \cdots & a_{R,n_s} \end{bmatrix}$$

Antecedent Conditions

$$\mathbf{s} = [s_1, s_2, \ldots, s_R]^T$$

Consequents (singleton fuzzy sets)

Data Base: Real genes

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,\delta} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,\delta} \\ \vdots & \vdots & \ddots & \vdots \\ p_{\rho,1} & p_{\rho,2} & \cdots & p_{\rho,\delta} \end{bmatrix}$$

MF parameters of the generalized bells for the input variables

$$\mathbf{o} = [o_1, o_2, \ldots, o_{M_{out}}]^T$$

Positions of output singleton MF

MOE Knowledge Base Learning – Example 4(c)

P. Pulkkinen, and H. Koivisto, "A dynamically constrained multiobjective genetic fuzzy system for regression problems," IEEE Trans. Fuzzy. Syst., vol. 18, n. 1, pp. 161–177, 2010.

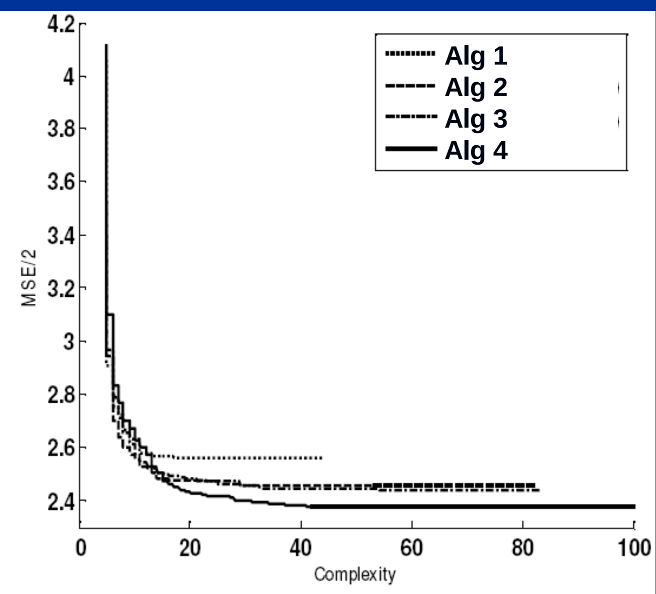
- **NSGA-II** is applied to find sets of FRBSs with different trade-offs between accuracy and interpretability
- **Accuracy** is evaluated in terms of mean square error
- **Rule Base Interpretability** is evaluated in terms total rule length
- **Semantic Partition Intepretability** is ensured by using dynamic constraints which enable three-parameter MF tuning to improve the accuracy while guaranteeing the transparency of fuzzy partitions
- **The initial population** is generated using an approach based on the **C4.5** and the **Wang and Mendel** algorithms. In this way, the initial fuzzy partitions are transparent, and the initial FRBSs contain less rules, rule conditions, and input variables than when only the WM algorithm is used

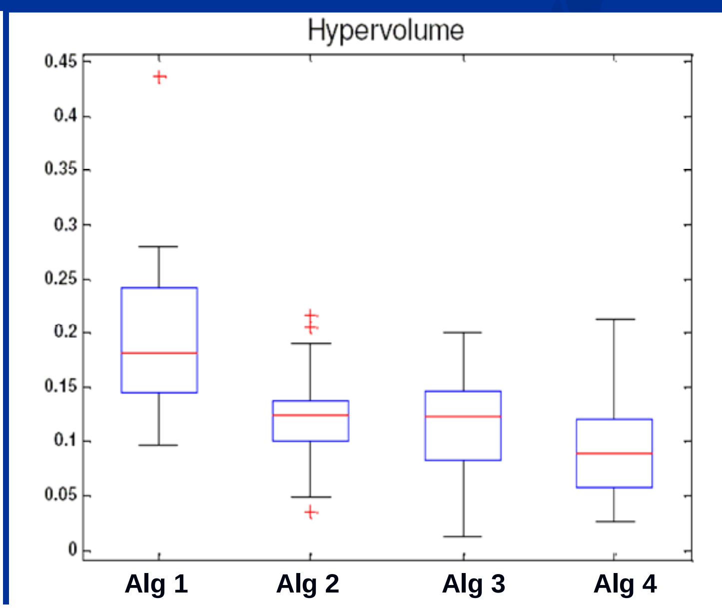
How can we compare different approaches?

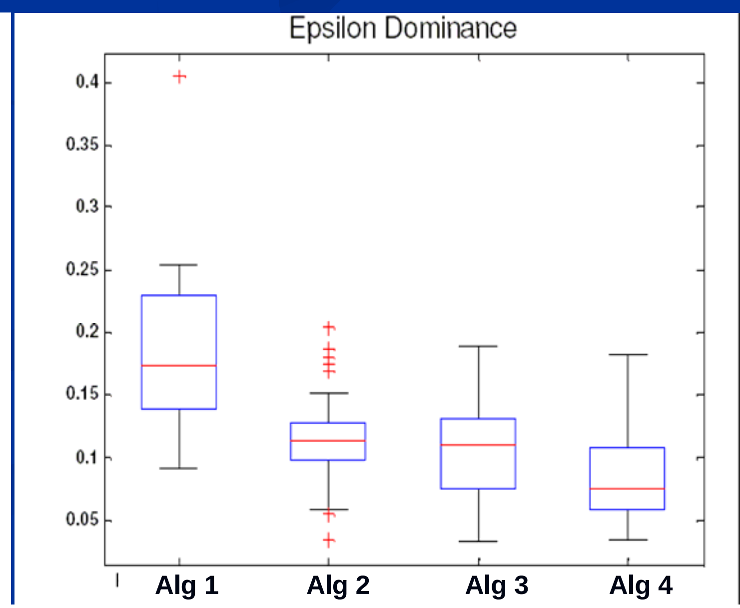
How can we compare different approaches?

Exploiting MOEA indexes

- Attainment surfaces - divides the objective space into two parts which contain, respectively, the solutions that are weakly dominated and the solutions that dominate with a frequency of at least 50%
- Hypervolume - measures the hypervolume of the portion of the objective space that is weakly dominated by a Pareto front approximation
- Epsilon dominance - is a measure of the smallest distance one would need to translate every solution in a front B so that B dominates a reference front A.



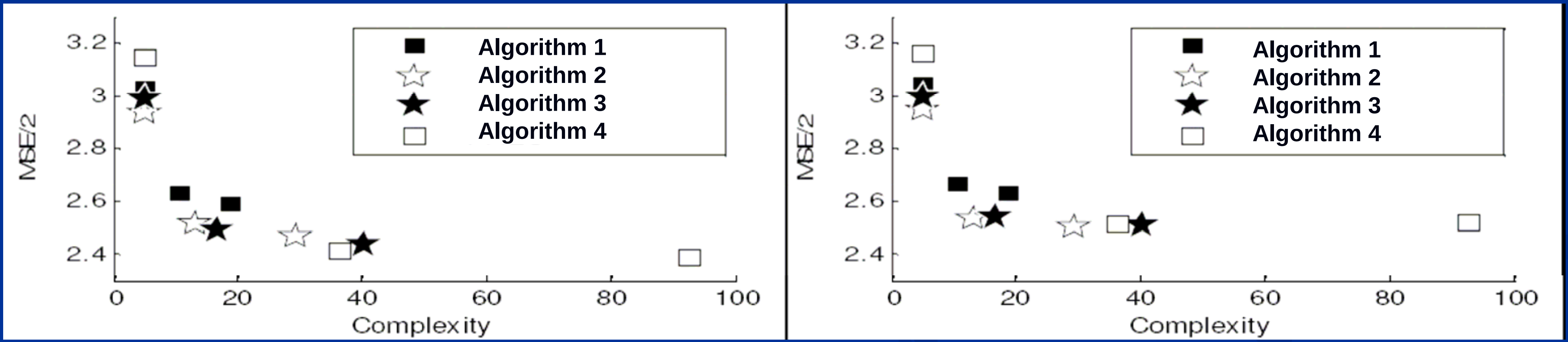




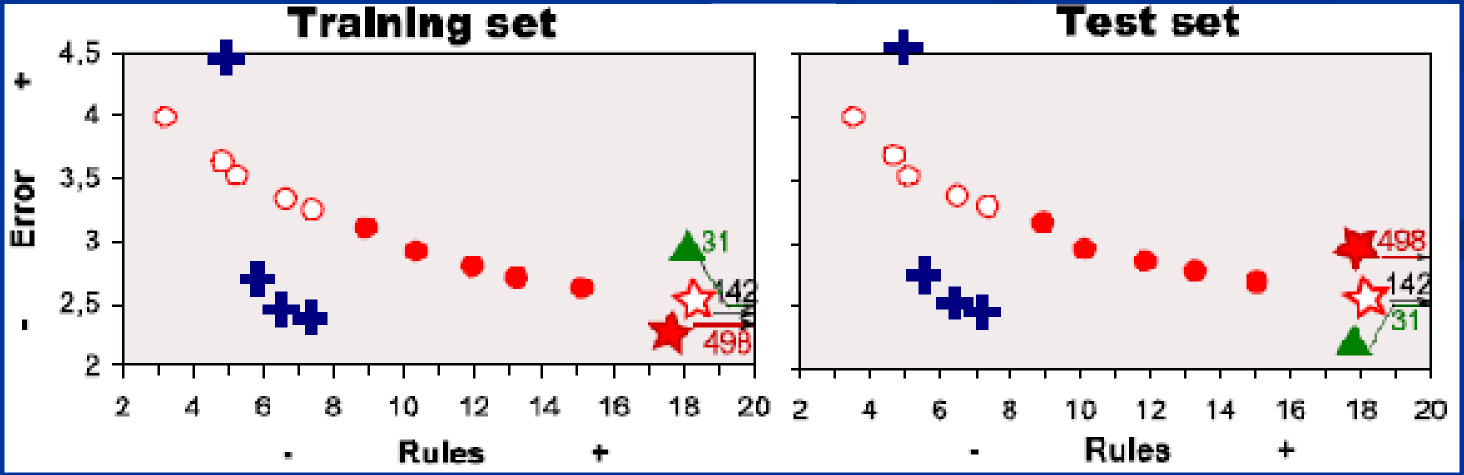
How can we compare different approaches?

The indexes evaluate the exploration and exploitation capabilities of the MOEA. But we are also interested in generalization capabilities of the FRBSs

- Choice of representative solutions (Most accurate, most interpretable, intermediate) and statistical analysis



- Average Pareto fronts



Hot Topics and New Challenges

Large Datasets

When the dataset is large, the computation of the fitness is very expensive in terms of **computational cost.**

Different approaches have been proposed to reduce this drawback

Parallel evolutionary algorithms

- I.Robles, R. Alcalá, J.M. Benítez, and F. Herrera, "Evolutionary parallel and gradually distributed lateral tuning of fuzzy rule-based systems," Evolutionary Intelligence, vol. 2, n. 1-2, pp. 5-19, 2009.

- M.A. de Vega, J.M Bardallo, F.A. Marquez, and A. Peregrin, "Parallel distributed two-level evolutionary multiobjective methodology for granularity learning and membership functions tuning in linguistic fuzzy systems," in Proc. in Proc. of ISDA 2009, Pisa (Italy), 30 Nov. –2 Dec., 2009, pp. 134–139.

- Y. Nojima, H. Ishibuchi, and I. Kuwajima, "Parallel distributed genetic fuzzy rule selection," Soft Computing, vol. 13, pp. 511-519, 2008.

Fitness approximation

- M. Cococcioni, B. Lazzerini, and F. Marcelloni, "On reducing computational overhead in multi-objective genetic Takagi–Sugeno fuzzy systems," Applied Soft Computing, vol. 11, n. 1, pp. 675-688, 2011.

Instance Selection (IS) techniques

-Y. Nojima, and H Ishibuchi, "Effects of data reduction on the generalization ability of parallel distributed genetic fuzzy rule selection," in Proc. of ISDA 2009, Pisa (Italy), 30 Nov. –2 Dec., 2009, pp. 96-101.

- M. Antonelli, P. Ducange, F. Marcelloni, "Exploiting a coevolutionary approach to concurrently select training instances and learn rule bases of Mamdani fuzzy dystems," in: Proc. of the IEEE World Congress on Computational Intelligence, Barcelona (Spain), 18 – 23 Jul., 2010, pp. 1366–1372.

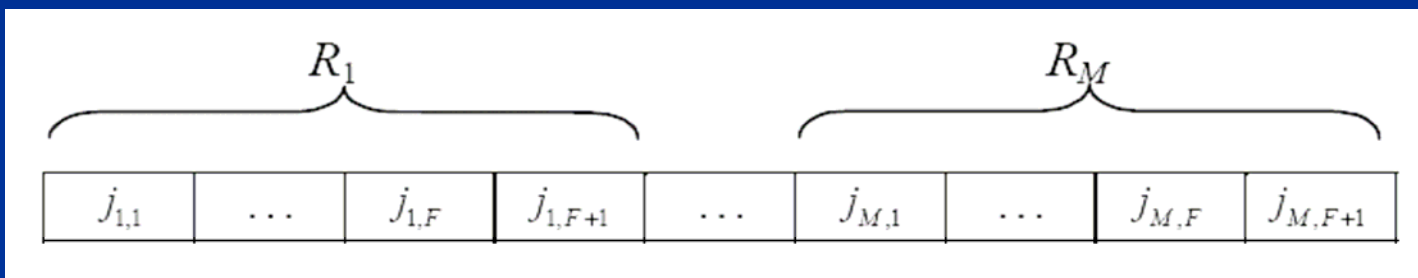
Large Datasets – Example 1(a)

M. Antonelli, P. Ducange, F. Marcelloni, "Exploiting a coevolutionary approach to concurrently select training instances and learn rule bases of Mamdani fuzzy dystems," in: Proc. of the IEEE World Congress on Computational Intelligence, Barcelona (Spain), 18 – 23 Jul., 2010, pp. 1366–1372.

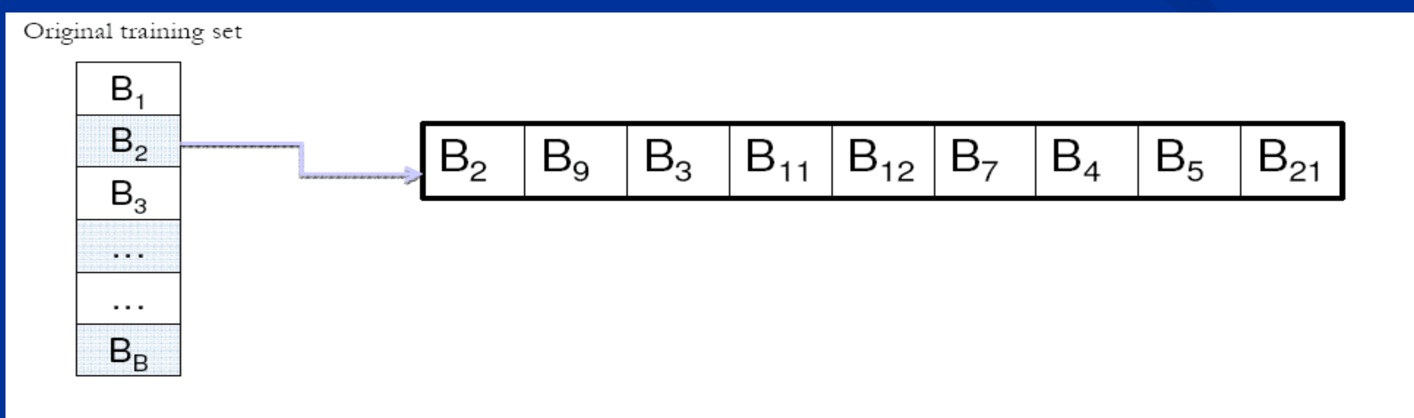
- The proposed coevolutionary approach is made of
  - A single objective genetic algorithm (SOGA), which perform the **training set selection**
  - A multi-objective evolutionary algorithm ((2+2)M-PAES) which **learn the RB** with different trade off between accuracy and complexity
- The SOGA and the ((2+2)M-PAES) are cyclically executed one after the other

Large Datasets – Example 1(b)

M. Antonelli, P. Ducange, F. Marcelloni, "Exploiting a coevolutionary approach to concurrently select training instances and learn rule bases of Mamdani fuzzy dystems," in: Proc. of the IEEE World Congress on Computational Intelligence, Barcelona (Spain), 18 – 23 Jul., 2010, pp. 1366–1372.

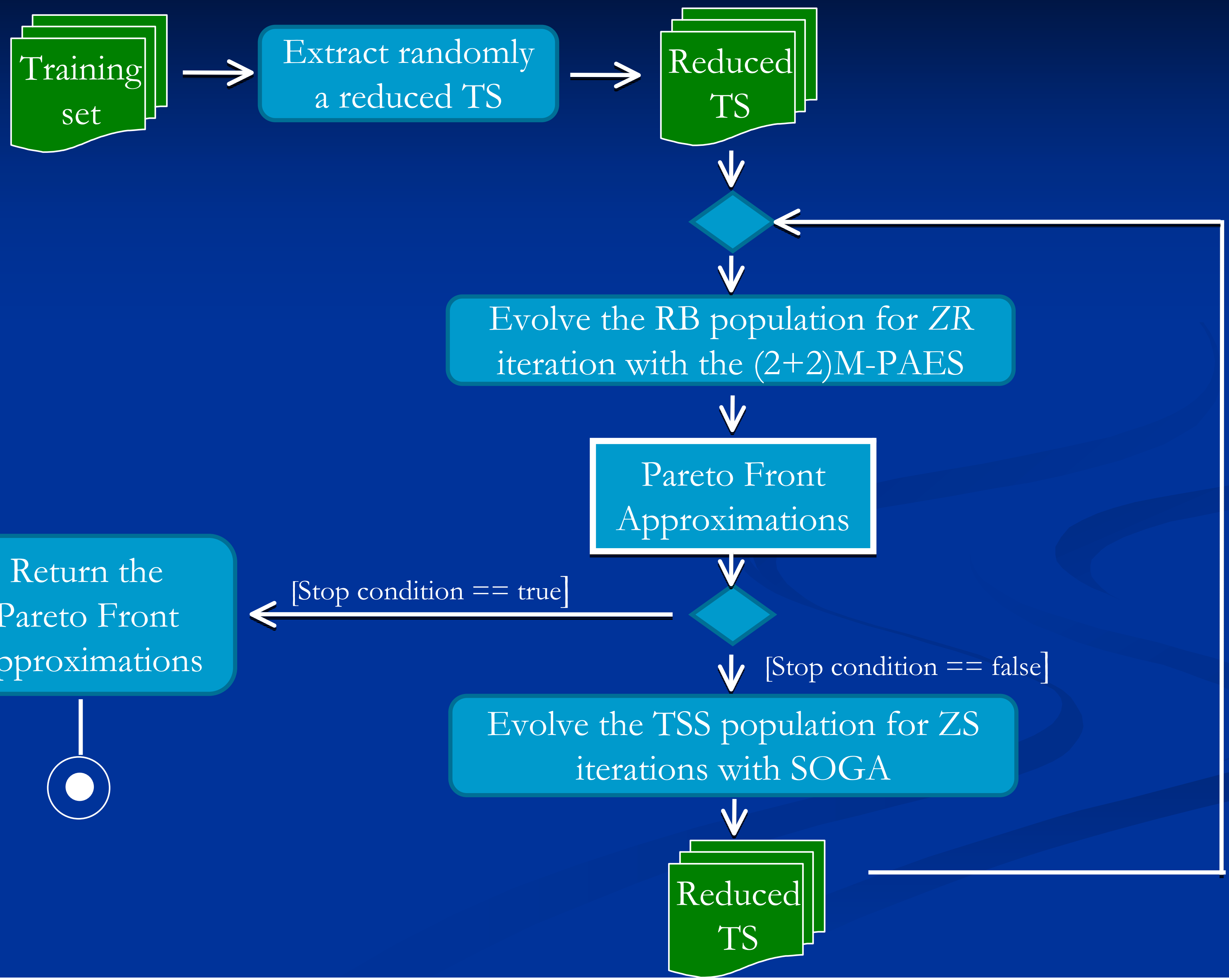


Chromosome coding for the (2+2)M-PAES



Chromosome coding for the SOGA

Large Datasets – Example 1(c)

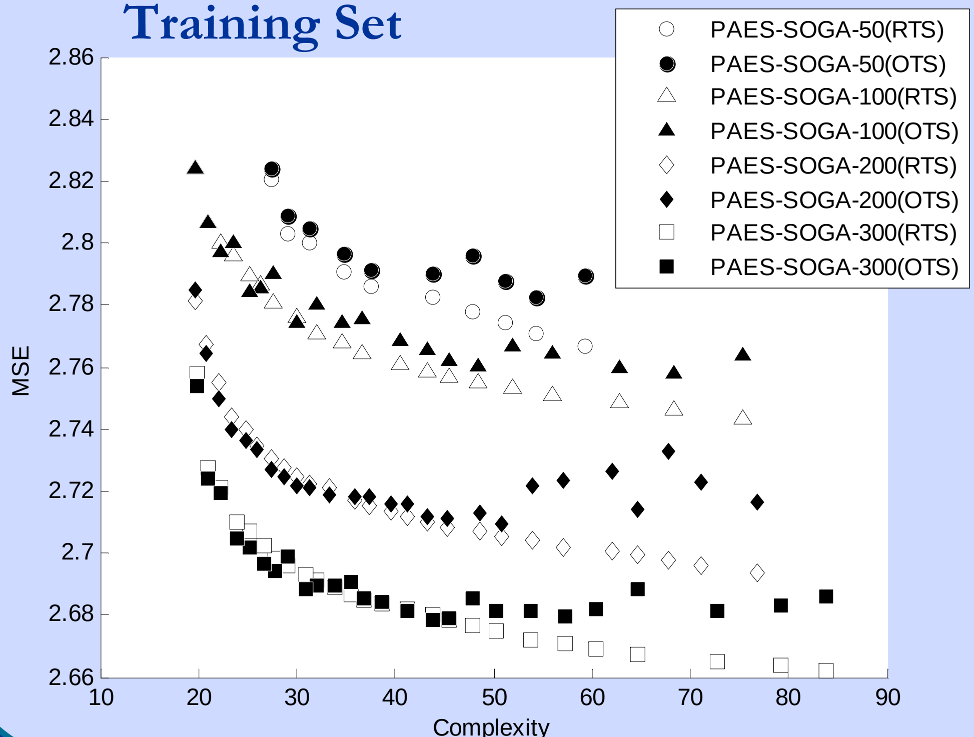


Large Datasets – Example 1(d)

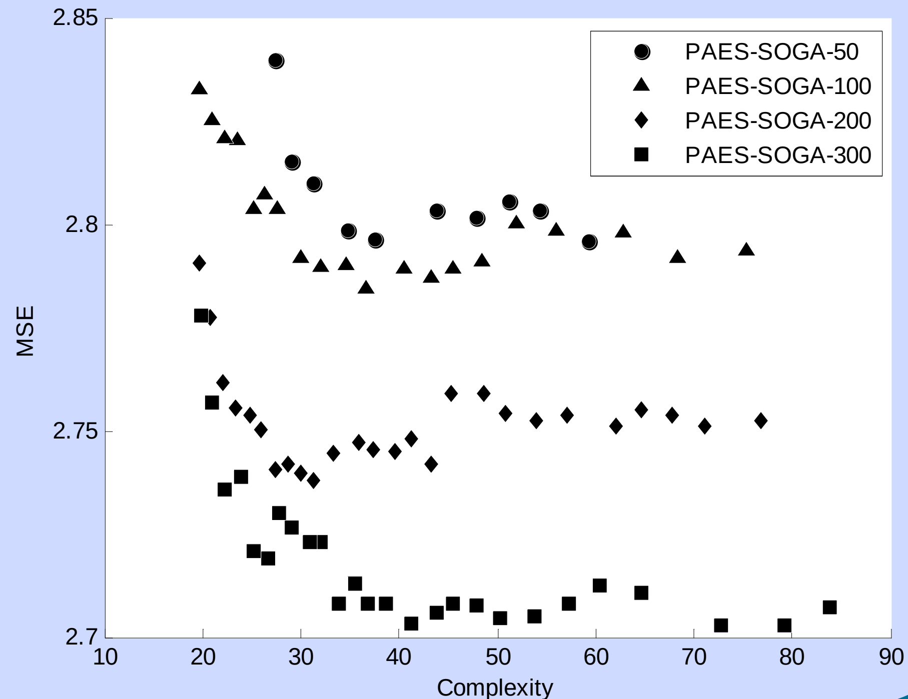
M. Antonelli, P. Ducange, F. Marcelloni, "Exploiting a coevolutionary approach to concurrently select training instances and learn rule bases of Mamdani fuzzy systems," in: Proc. of the IEEE World Congress on Computational Intelligence, Barcelona (Spain), 18 – 23 Jul., 2010, pp. 1366–1372.

Results on the **Abalone dataset** which contains 4177 input-output patterns described by 8 input variables

Training Set



Test Set

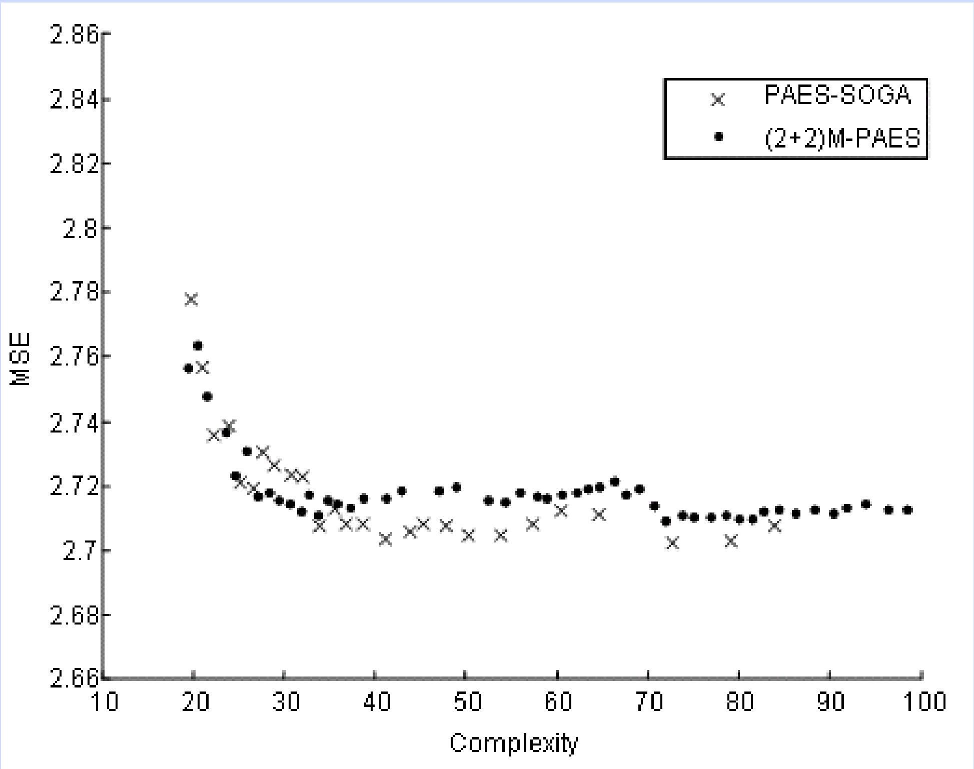


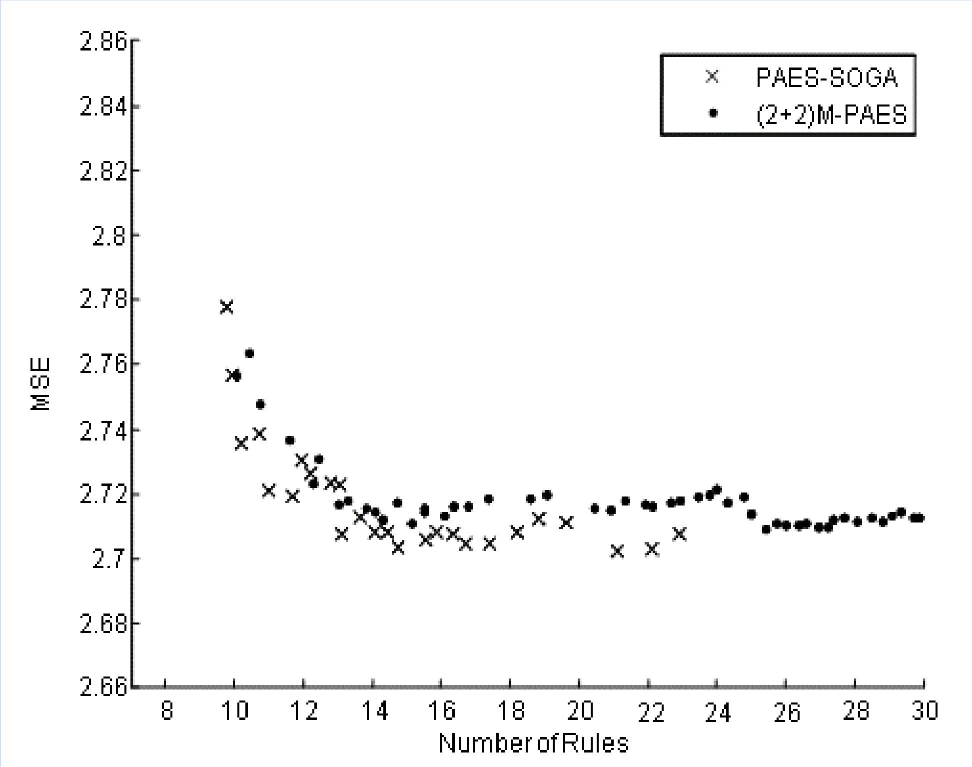
The SOGA selects a reduced training set (RTS) which contains only the 20% of the overall training set (OTS)

Large Datasets – Example 1(e)

M. Antonelli, P. Ducange, F. Marcelloni, "Exploiting a coevolutionary approach to concurrently select training instances and learn rule bases of Mamdani fuzzy dystems," in: Proc. of the IEEE World Congress on Computational Intelligence, Barcelona (Spain), 18 – 23 Jul., 2010, pp. 1366–1372.

Test Set





Comparison between PAES-SOGA and (2+2)M-PAES on the test set

High Dimensional Data Sets

- The search space grows as the number of features increases, leading to a slow and possibly difficult convergence of the algorithms
- Furthermore, the FRBSs could suffer from the exponential rule explosion due to the large numbers of features and data examples
- The following approaches have been proposed to reduce these drawbacks

Using feature selection

- J. Casillas, O. Cordón, M.J. del Jesus, and F. Herrera, "Genetic feature selection in a fuzzy rule-based classification system learning process, " Information Sciences, vol. 136, pp.135–157, 2001.

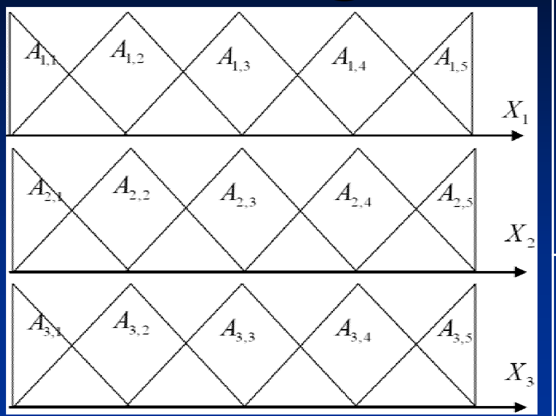
- O. Cordón, and A. Quirin, "Comparing Two Genetic Overproduce-and-choose Strategies for Fuzzy Rule-based Multiclassification Systems Generated by Bagging and Mutual Information-based Feature Selection", International Journal of Hybrid Intelligent Systems, vol. 7, pp.45–64, 2011.

Exploiting ad-hoc modified MOEAs

-M. Antonelli, P. Ducange, B. Lazzerini, F. Marcelloni, "Multi-objective Evolutionary Generation of Mamdani Fuzzy Rule-Based Systems based on Rule and Condition Selection," in: Proc. of 5th IEEE GEFS 2011, Paris (France), 11 – 15 Apr., 2011.

-R. Alcala, M. J. Gacto, and F. Herrera, "A Fast and Scalable Multi-Objective Genetic Fuzzy System for Linguistic Fuzzy Modeling in High-Dimensional Regression Problems," IEEE Transactions on Fuzzy System, doi: 10.1109/TFUZZ.2011.2131657, *in press (2011).*

High Dimensional Data Sets – Example 1(a)



Initial Data Base

MOE Rule and Condition Selection together with MF Learning

Chromosome coding :

- **Selected Rules**: Integer
- **Selected Conditions**: Binary

Membership Function Parameters: Real coding

R_1 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

R_2 : IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3 : IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$

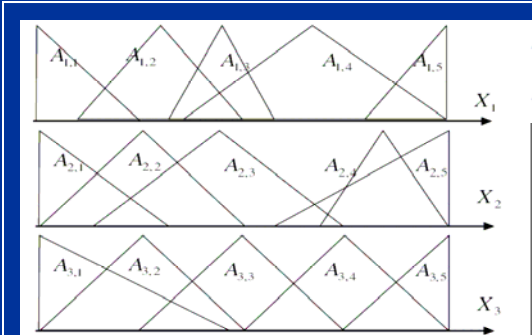
Initial Rule Base

Training set

KB Fitness Evaluation

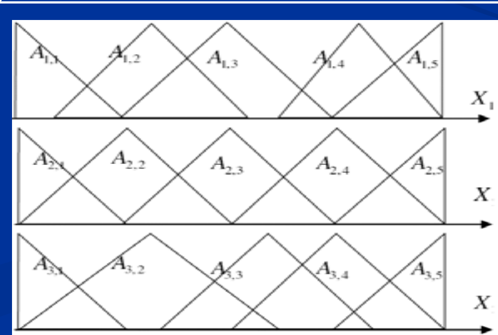
Accuracy: Mean Square Error

RB Interpretability: Number of Conditions



R_1: IF X_1 is $A_{1,2}$ THEN X_3 is $A_{3,1}$

R_3: IF X_1 is $A_{1,1}$ THEN X_3 is $A_{3,2}$



R_1: IF X_1 is $A_{1,2}$ THEN X_3 is $A_{3,1}$

R_2: IF X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3: IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$

Non-Dominated Final Knowledge Bases

High Dimensional Data Sets – Example 1(b)

-M. Antonelli, P. Ducange, B. Lazzerini, F. Marcelloni, "Multi-objective Evolutionary Generation of Mamdani Fuzzy Rule-Based Systems based on Rule and Condition Selection," in: Proc. of 5th IEEE GEFS 2011, Paris (France), 11 – 15 Apr., 2011.

Coding scheme:

The RB chromosome is a vector of pairs:

$$\mathbf{p}_m = (k_m, \mathbf{v}_m)$$

$$k_m \in [0, \ldots, M_{WM}]$$

Integer genes for rule selection

$$\mathbf{v}_m = [v_{m,1}, \ldots, v_{m,F}]$$

Binary genes for condition selection

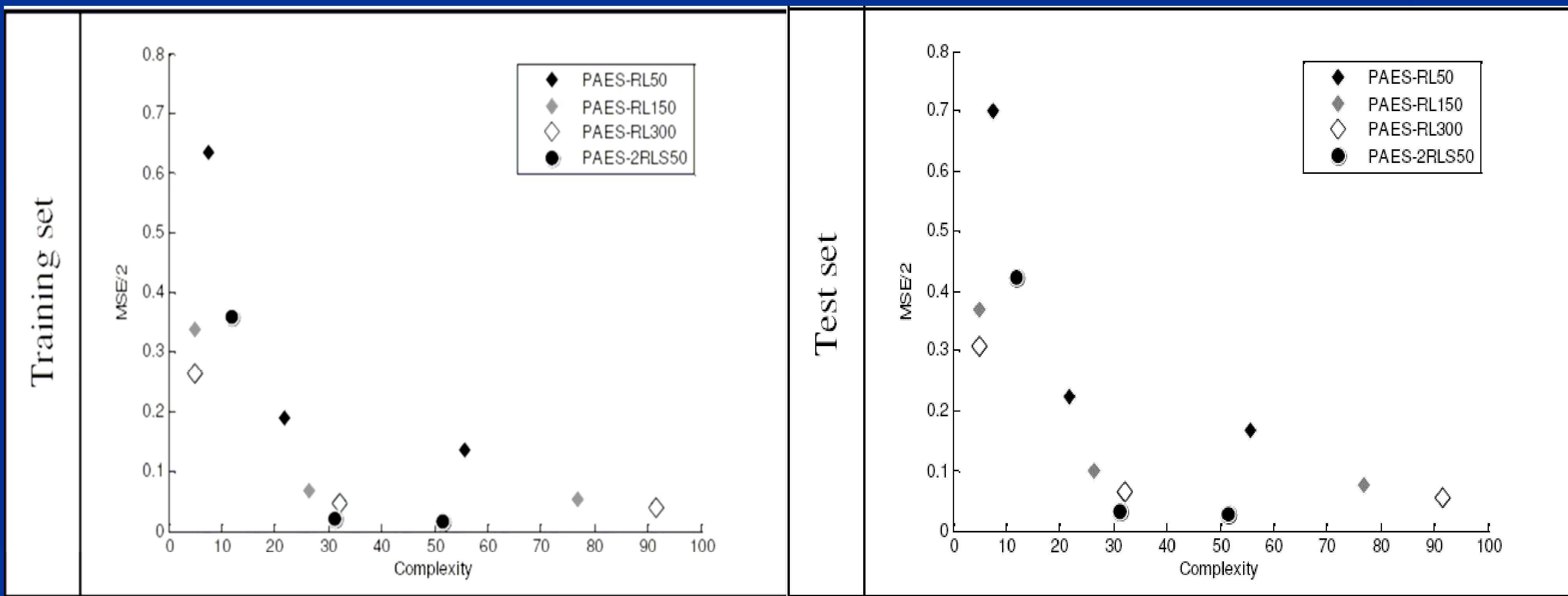
The DB chromosome is a vector of **real genes** which codify the position of the cores of strong fuzzy partitions

$$\overbrace{\boxed{b_{1,2}}\ \boxed{\cdots}\ \boxed{b_{1,T_1-1}}}^{X_1}\ \boxed{\cdots}\ \overbrace{\boxed{b_{F+1,2}}\ \boxed{\cdots}\ \boxed{b_{F+1,T_{F+1}-1}}}^{X_{F+1}}$$

High Dimensional Data Sets – Example 1(c)

-M. Antonelli, P. Ducange, B. Lazzerini, F. Marcelloni, "Multi-objective Evolutionary Generation of Mamdani Fuzzy Rule-Based Systems based on Rule and Condition Selection," in: Proc. of 5th IEEE GEFS 2011, Paris (France), 11 – 15 Apr., 2011.

Results on the **Mortgage dataset** which contains 1049 input-output patterns described by 15 input variables



High Dimensional Data Sets – Example 2(a)

MOE Granularity and MF Parameters Learning

The chromosomes can be formed by different parts

Granularity part: Integer

MF parameter part: Real

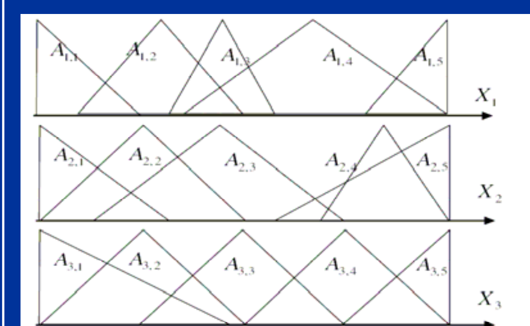
Set of DBs → **Rule Learning Method** → Set of KBs

Training set

KB Fitness Evaluation

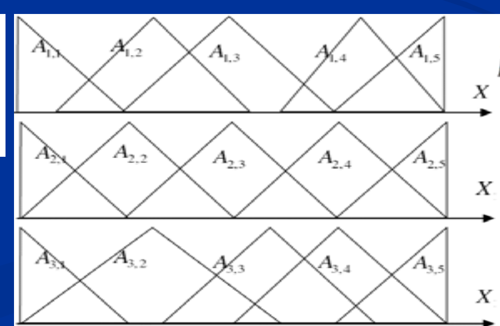
Accuracy: Mean Square Error

RB Interpretability: Number of Rules



R_2: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,2}$ THEN X_3 is $A_{3,1}$

R_3: IF X_1 is $A_{1,1}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,2}$



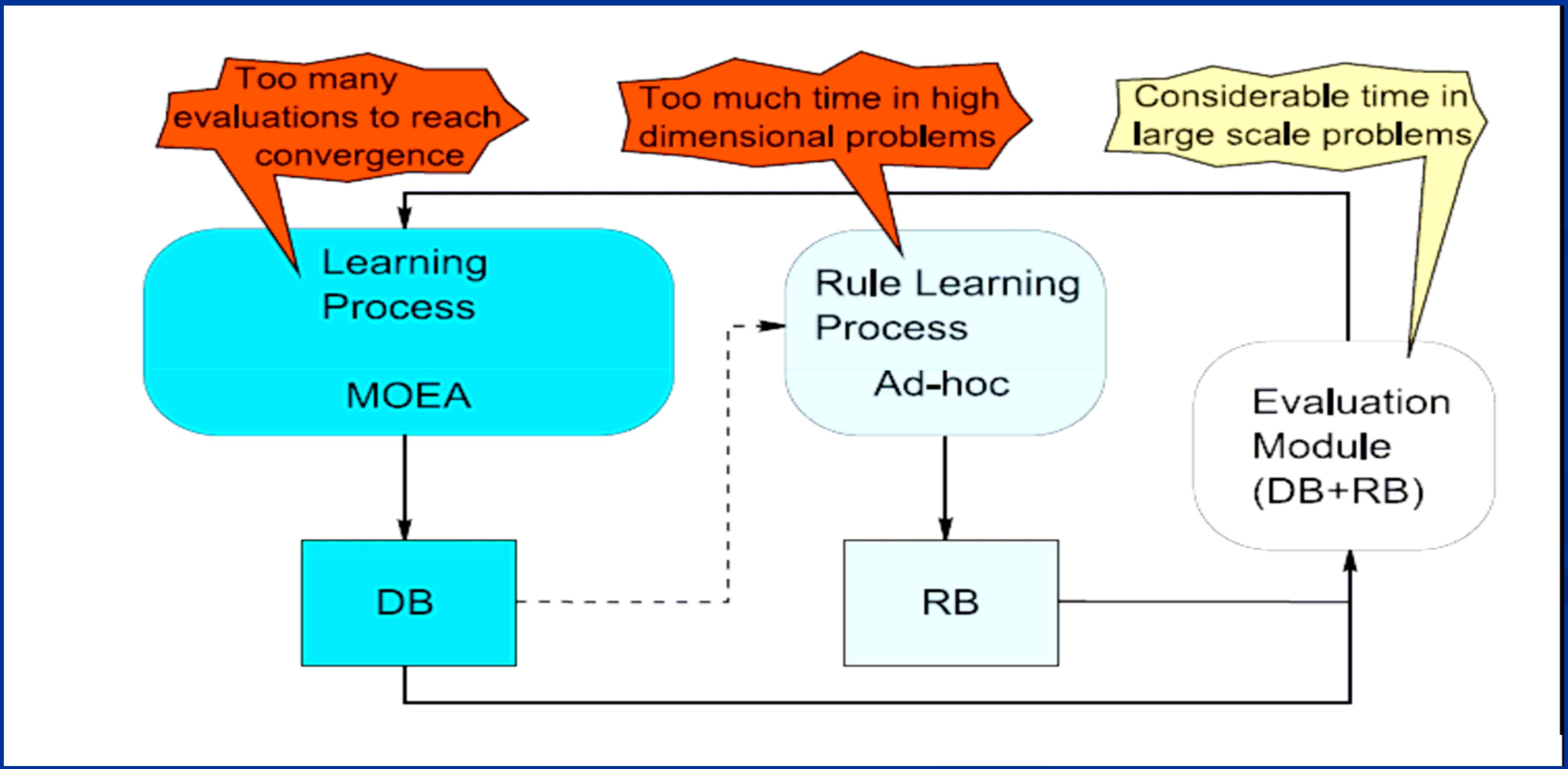
R_1: IF X_1 is $A_{1,2}$ and X_2 is $A_{2,1}$ THEN X_3 is $A_{3,1}$

Non-Dominated Final Knowledge Bases

High Dimensional Data Sets – Example 2(b)

R. *Alcala, M. J. Gacto, and F. Herrera, "A Fast and Scalable Multi-Objective Genetic Fuzzy System for Linguistic Fuzzy Modeling in High-Dimensional Regression Problems," IEEE Transactions on Fuzzy System,* doi: 10.1109/TFUZZ.2011.2131657, *in press (2011).*

Problems



High Dimensional Data Sets – Example 2(c)

R. *Alcala, M. J. Gacto, and F. Herrera, "A Fast and Scalable Multi-Objective Genetic Fuzzy System for Linguistic Fuzzy Modeling in High-Dimensional Regression Problems," IEEE Transactions on Fuzzy System,* doi: 10.1109/TFUZZ.2011.2131657, *in press (2011).*

Proposed Solutions:

- **Incest prevention**, **restarting strategy** and **stopping condition** have been integrated into the well-known SPEA2 algorithm
- The granularities and a partition displacement are codified into a **double coded chromosome** and learnt concurrently during the evolutionary process
- The RB is generated exploiting an ad-hoc modification of the **Wang and Mendel (WM**) algorithm which includes a **cropping criterion** in the RB generation process so as to avoid the generation of excessively large RBs
- A **new error estimation procedure** to reduce the computational effort for large datasets is also proposed

High Dimensional Data Sets – Example 2(d)

R. *Alcala, M. J. Gacto, and F. Herrera, "A Fast and Scalable Multi-Objective Genetic Fuzzy System for Linguistic Fuzzy Modeling in High-Dimensional Regression Problems," IEEE Transactions on Fuzzy System*, doi: 10.1109/TFUZZ.2011.2131657, *in press (2011).*

Double coding scheme:

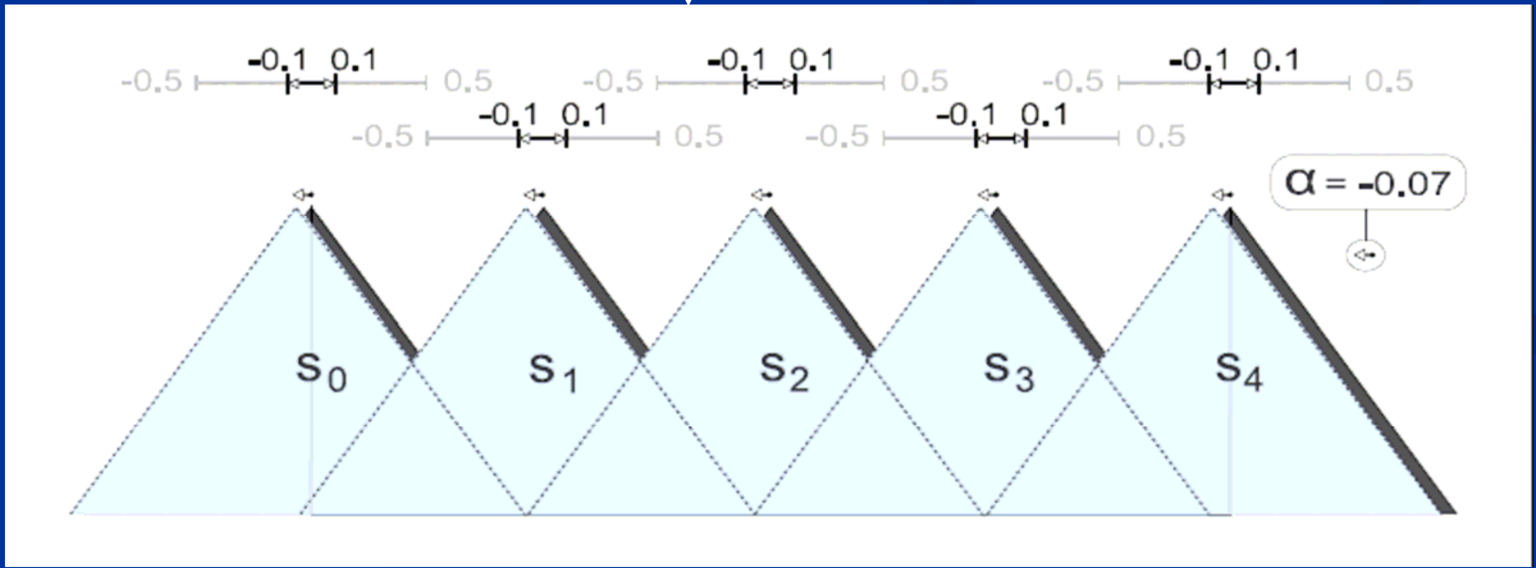
$$C_1 = (L^1, \ldots, L^N)$$

Integer Coding for Granularity Learning and

Input Variable Selection: $L^i \in \{1,\ldots,7\}$ for $i = 1 \ldots N-1$ and $L^N \in \{2,\ldots,7\}$

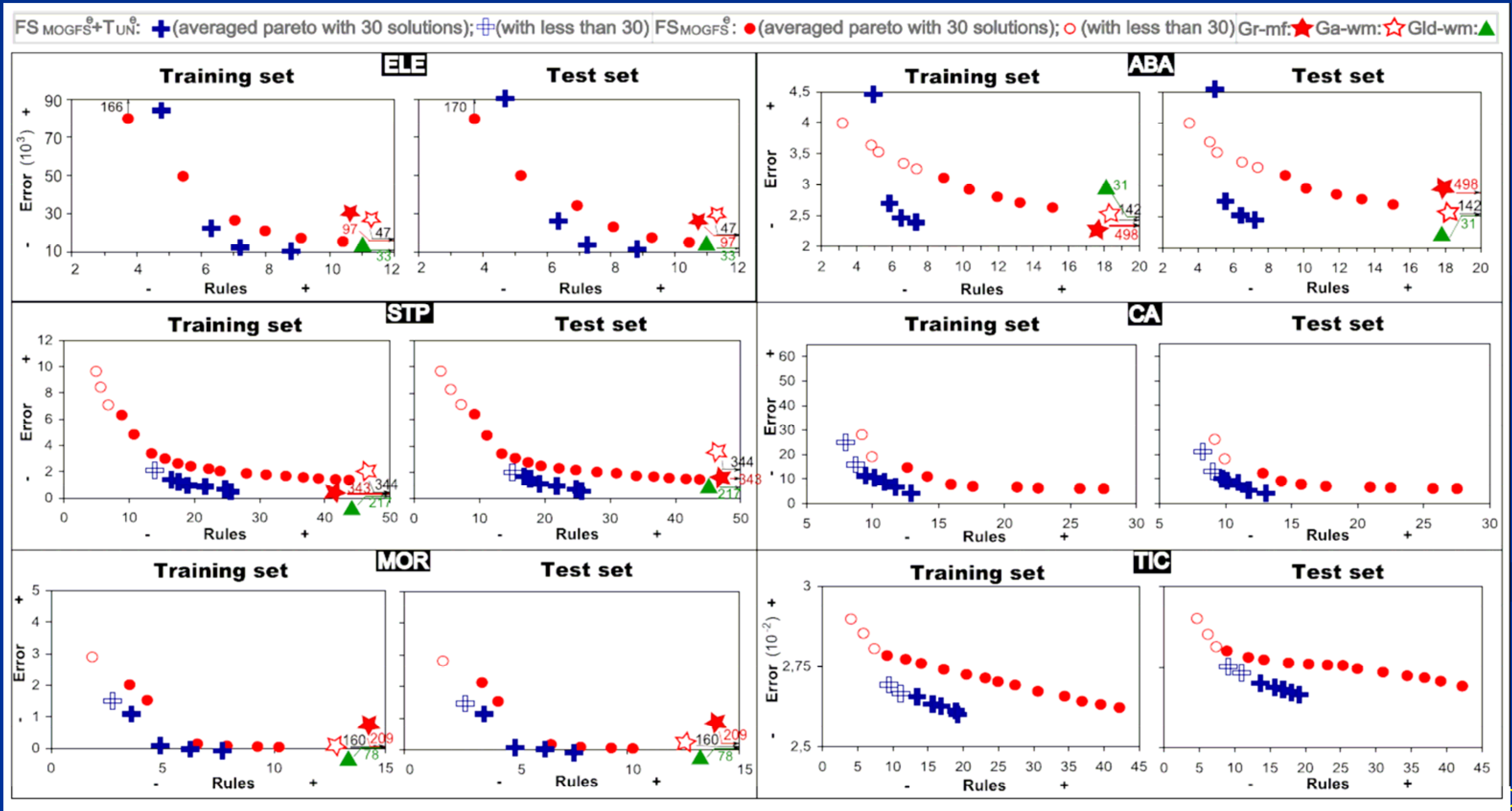
$$C_2 = (\alpha^1, \ldots, \alpha^N)$$

Real Coding for Lateral Displacement, where $\alpha^i \in [-0.1, 0.1]$



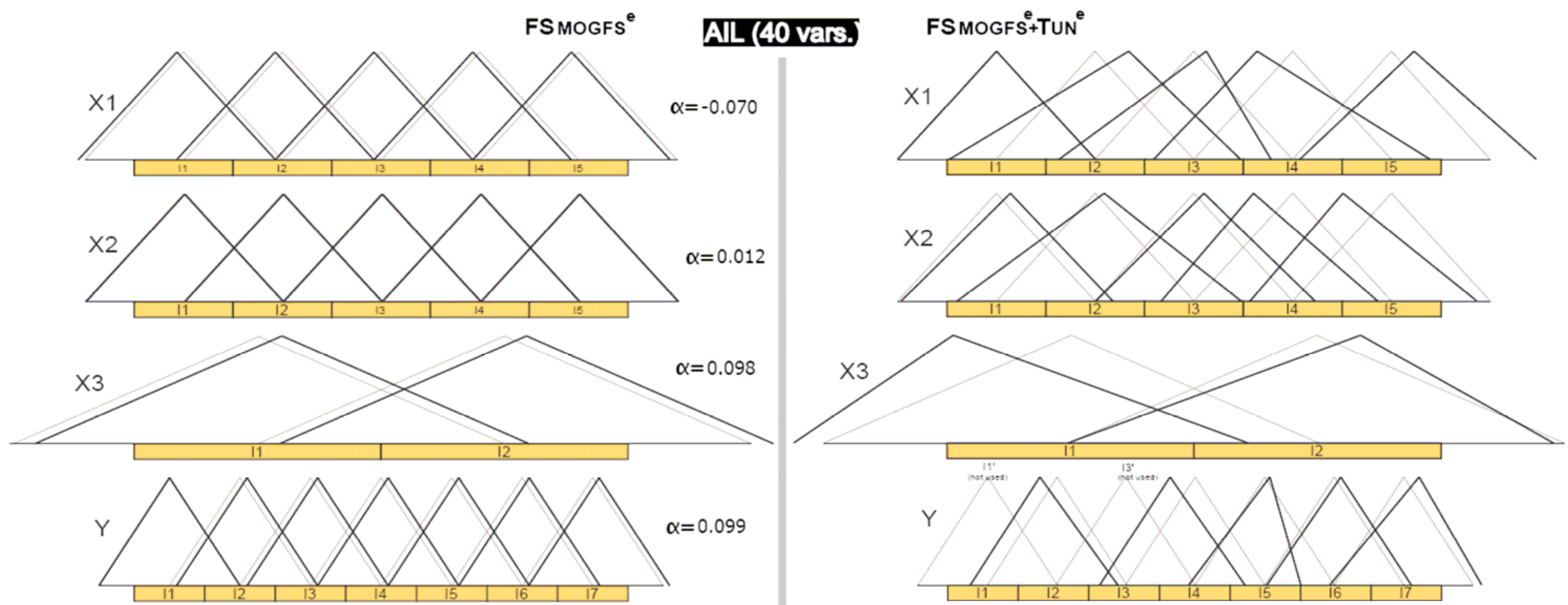
High Dimensional Data Sets – Example 2(e)

R. *Alcala, M. J. Gacto, and F. Herrera, "A Fast and Scalable Multi-Objective Genetic Fuzzy System for Linguistic Fuzzy Modeling in High-Dimensional Regression Problems," IEEE Transactions on Fuzzy System,* doi: 10.1109/TFUZZ.2011.2131657, *in press (2011).*



High Dimensional Data Sets – Example 2(f)

R. *Alcala, M. J. Gacto, and F. Herrera, "A Fast and Scalable Multi-Objective Genetic Fuzzy System for Linguistic Fuzzy Modeling in High-Dimensional Regression Problems," IEEE Transactions on Fuzzy System,* doi: 10.1109/TFUZZ.2011.2131657, *in press (2011).*



X1	X2	X3	Y	X1	X2	X3	Y	X1	X2	X3	Y
l1'	l3'	l1'	l7'	l3'	l3'	l1'	l6'	l4'	l3'	l2'	l3'
l2'	l1'	l1'	l5'	l3'	l3'	l2'	l3'	l4'	l4'	l1'	l6'
l2'	l2'	l1'	l6'	l3'	l4'	l1'	l6'	l4'	l4'	l2'	l4'
l2'	l3'	l1'	l6'	l3'	l4'	l2'	l5'	l4'	l5'	l1'	l6'
l2'	l3'	l2'	l4'	l3'	l5'	l1'	l6'	l4'	l5'	l2'	l5'
l2'	l4'	l1'	l6'	l4'	l1'	l1'	l4'	l5'	l1'	l1'	l3'
l2'	l5'	l1'	l6'	l4'	l1'	l2'	l2'	l5'	l2'	l1'	l5'
l3'	l1'	l1'	l5'	l4'	l2'	l1'	l5'	l5'	l2'	l2'	l3'
l3'	l2'	l1'	l5'	l4'	l2'	l2'	l3'	l5'	l3'	l1'	l4'
l3'	l2'	l2'	l3'	l4'	l3'	l1'	l5'	l5'	l4'	l1'	l5'

#R: 30
MSE-tra: 2.357
MSE-tst: 2.398

X1	X2	X3	Y	X1	X2	X3	Y
l2'	l2'	l1'	l7'	l4'	l2'	l1'	l6'
l2'	l4'	l1'	l7'	l4'	l2'	l2'	l4'
l2'	l5'	l1'	l7'	l4'	l3'	l1'	l6'
l3'	l2'	l1'	l6'	l4'	l3'	l2'	l4'
l3'	l2'	l2'	l4'	l4'	l4'	l1'	l7'
l3'	l3'	l1'	l7'	l4'	l5'	l2'	l6'
l3'	l5'	l1'	l7'	l5'	l3'	l1'	l5'
l4'	l1'	l2'	l2'				

#R: 15
MSE-tra: 1.944
MSE-tst: 1.992

New Challenges

- Exploit different or new interpretability measures in MOEFS
- Dealing with more than two or three objectives in the framework of MOEFS
- To exploit and/or to develop ad-hoc MOEAs for improving exploration of specific parts of the approximated Pareto fronts
- To provide tools to visualize and statistically compare the results of different MOEFSs

New Challenges

- Real applications (killer application?)
  - *M.J. Gacto, R. Alcalá, and F. Herrera, "A multi-objective evolutionary algorithm for an effective tuning of fuzzy logic controllers in heating, ventilating and air conditioning systems," Applied Intelligence, doi: 10.1007/s10489-010-0264-x, in press (2011)*
  - *J. Casillas, and F.J. Martínez-López, "Mining uncertain data with multiobjective genetic fuzzy systems to be applied in consumer behaviour modelling," Expert Systems with Applications, vol. 36, n. 2, pp. 1645-165, 2009.*
  - *P. Fazendeiro, J.V. de Oliveira, W. Pedrycz, "A multiobjective design of a patient and anaesthetist-friendly neuromuscular blockade controller," Transactions on Biomedical Engineering , vol. 54, pp. 1667–1678, 2007*

Thanks

- I would like to thank very much Dr. Pietro Ducange for the valuable help in preparing these slides.

Thank you very much for your attention.

Questions?