Data Mining and Soft Computing

Francisco Herrera

Research Group on Soft Computing and Information Intelligent Systems (SCI^{2}S)
Dept. of Computer Science and A.I.
University of Granada, Spain

Email: herrera@decsai.ugr.es
http://sci2s.ugr.es
http://decsai.ugr.es/~herrera

DECSAI
Universidad de Granada

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.

Outline

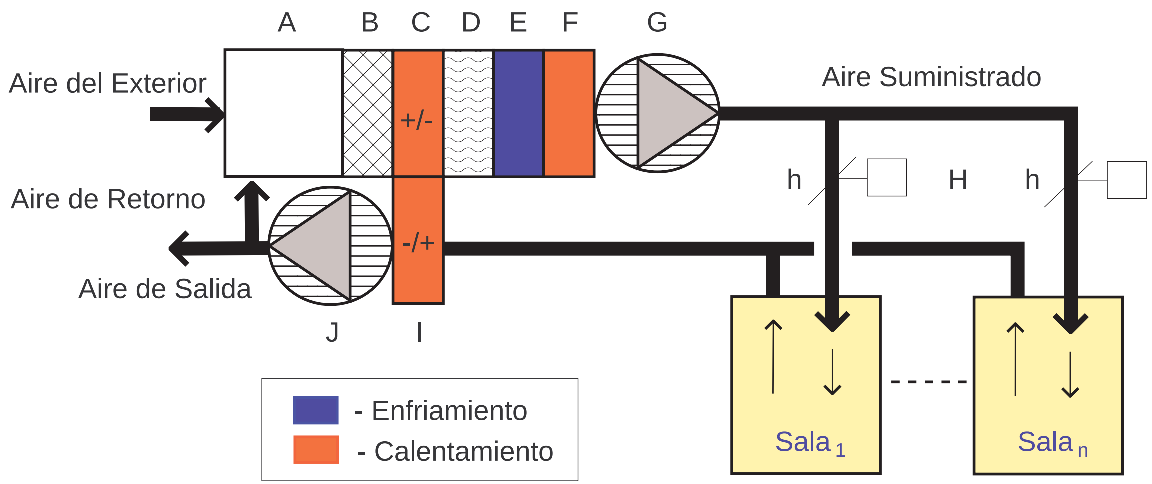
- ✓Brief Introduction to Genetic Fuzzy Systems
- ✓Tuning Methods: Basic and Advanced Approaches
- ✓Genetic Fuzzy Systems Application to HVAC Problems
- ✓Concluding Remarks

GENETIC FUZZY SYSTEMS: APPLICATION TO HVAC PROBLEM

Heating Ventilating and Air Conditioning Systems: Problem

GENESYS

JOULE-THERMIE JOE-CT98-0090



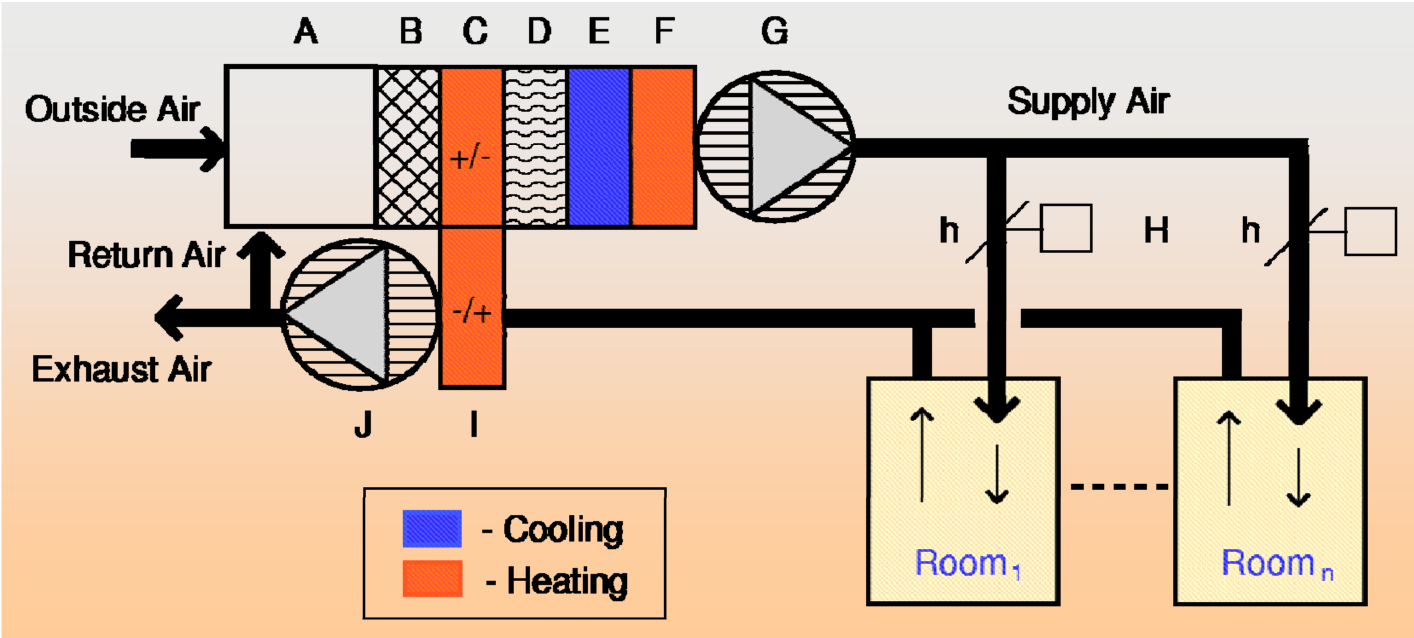
Fuzzy Logic Controllers for Energy Efficiency Consumption in Buildings

Heating Ventilating and Air Conditioning Systems: Problem

- Energy consumption in buildings is the 40% of the total and more than a half is for indoor climate conditions
- The use of specific technologies can save up to a 20% of the energy consumption
- The use of appropriate automatic control strategies could result in energy savings ranging 15-85 %
- Moreover, in current systems, several criteria are considered and optimized independently without a global strategy

Fuzzy Logic Controllers for Energy Efficiency Consumption in Buildings

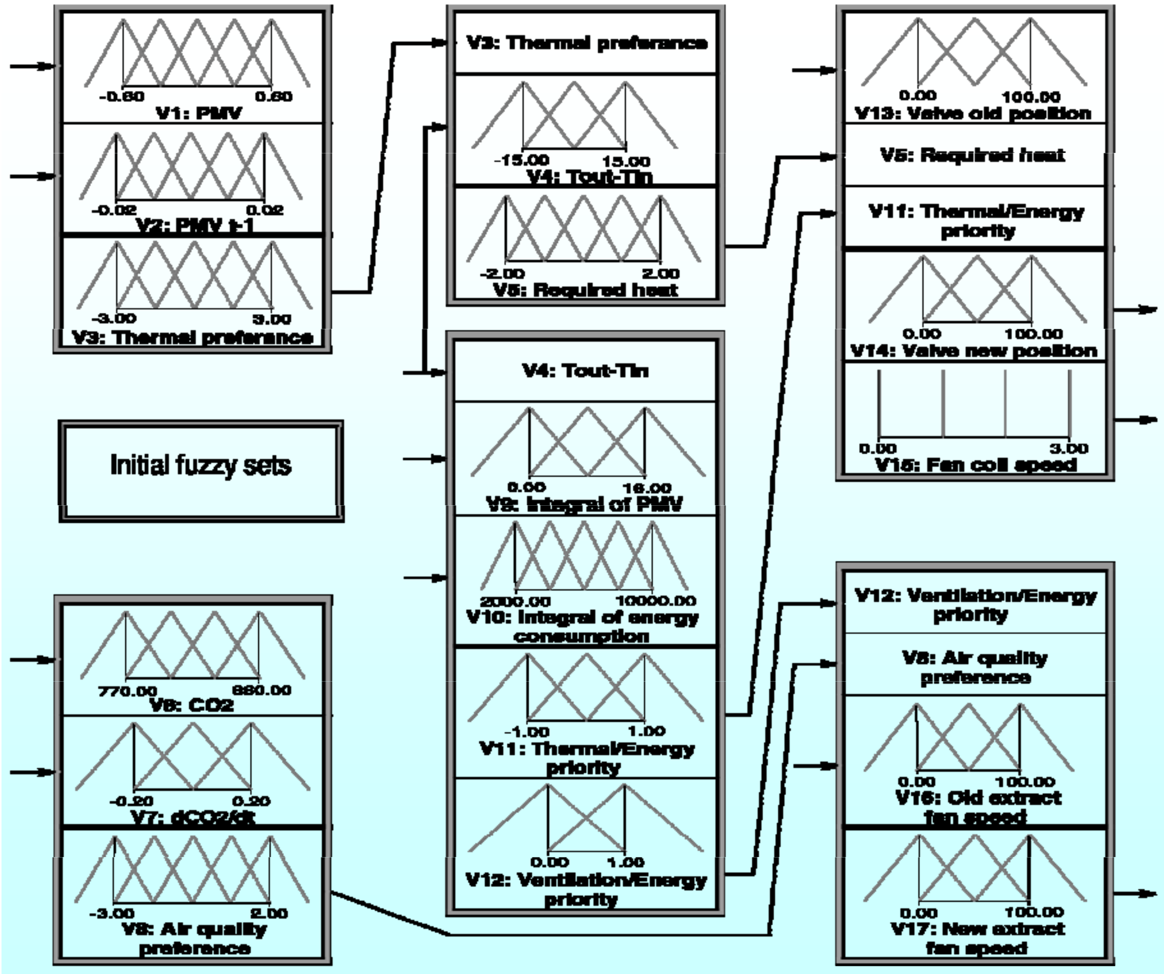
Generic Structure of an Office Building HVAC System



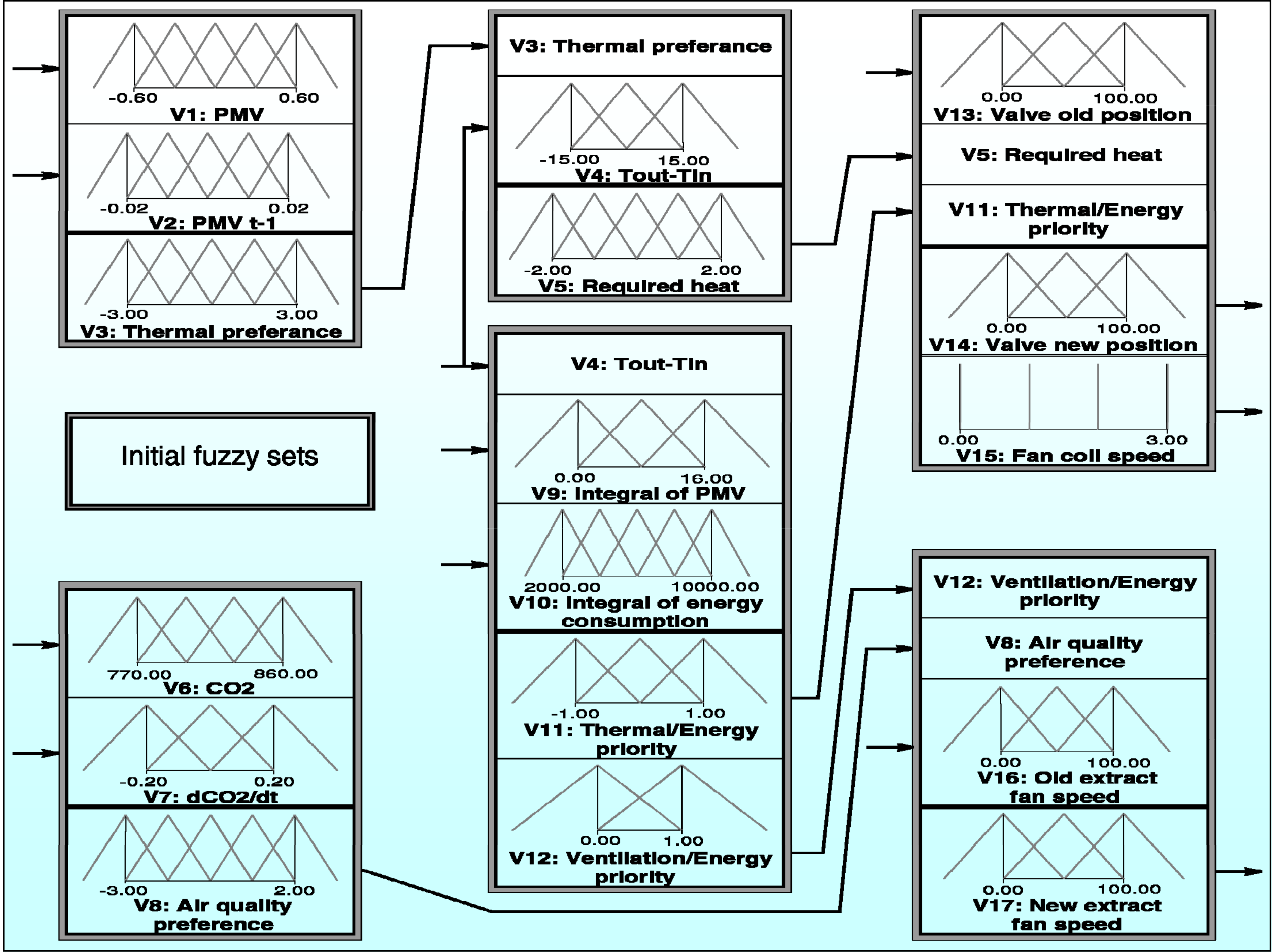
- It maintain a good thermal quality in summer and winter
- It dilutes and removes emissions from people, equipment and activities and supplies clean air

Fuzzy Logic Controllers for Energy Efficiency Consumption in Buildings

Initial Data Base



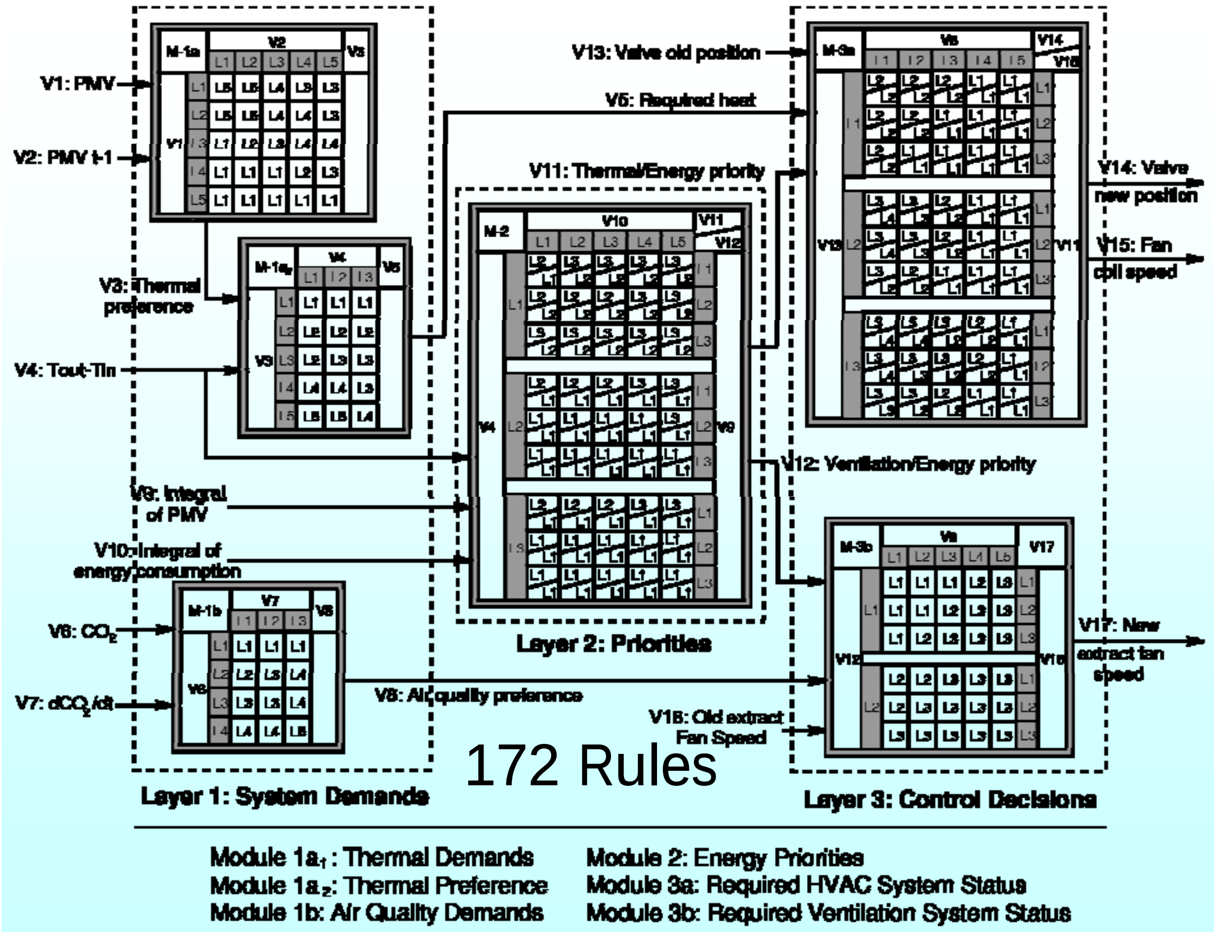
17 Variables



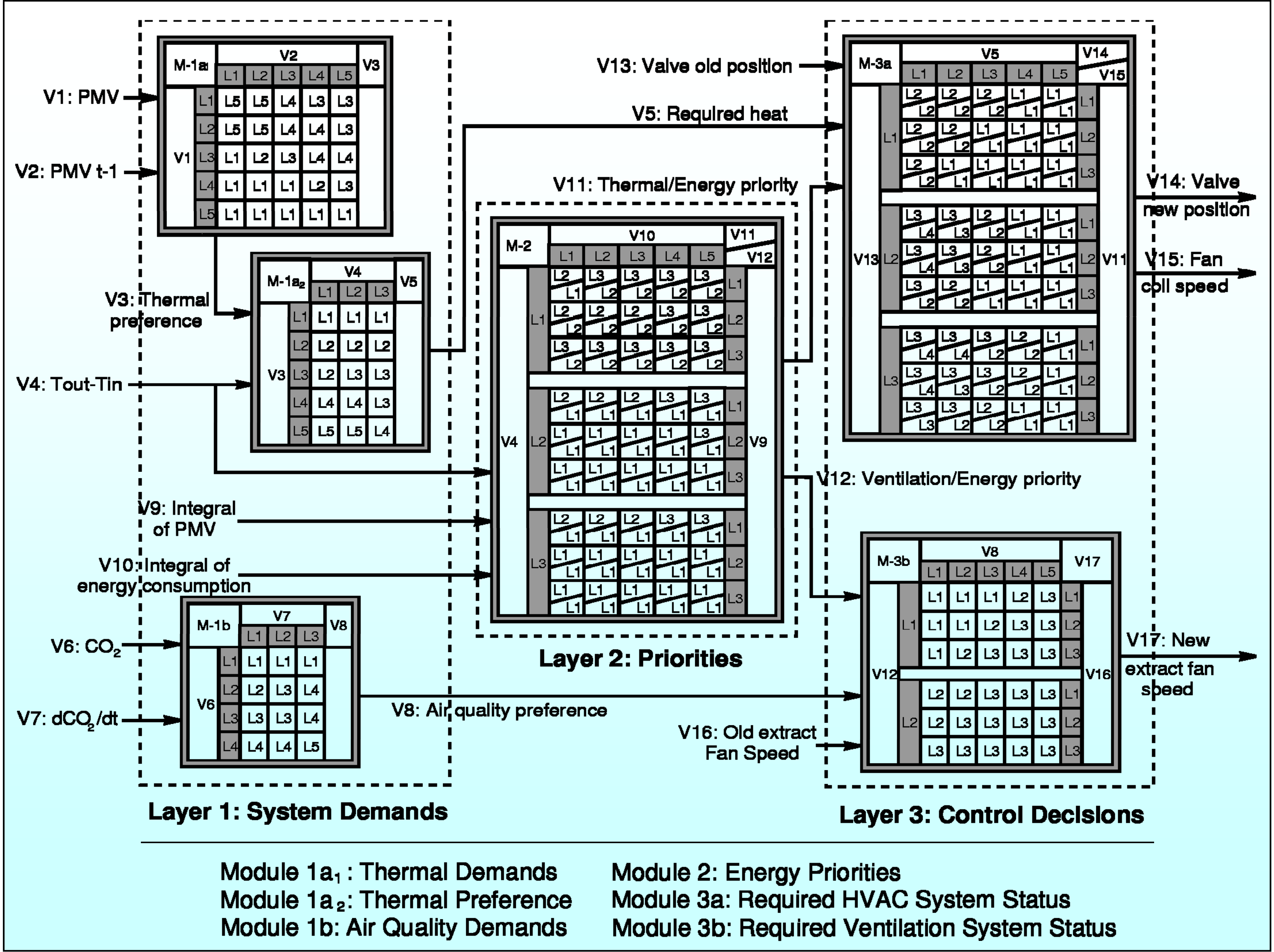
-0.60
0.60
V1: PMV
-0.02
0.02
V2: PMV t-1
-3.00
3.00
V3: Thermal preferance
Initial fuzzy sets
770.00
860.00
V6: CO2
-0.20
0.20
V7: dCO2/dt
-3.00
2.00
V8: Air quality preference
V3: Thermal preferance
-15.00
15.00
V4: Tout-Tin
-2.00
2.00
V5: Required heat
V4: Tout-Tin
0.00
16.00
V9: Integral of PMV
2000.00
10000.00
V10: Integral of energy consumption
-1.00
1.00
V11: Thermal/Energy priority
0.00
1.00
V12: Ventilation/Energy priority
0.00
100.00
V13: Valve old position
V5: Required heat
V11: Thermal/Energy priority
0.00
100.00
V14: Valve new position
0.00
3.00
V15: Fan coil speed
V12: Ventilation/Energy priority
V8: Air quality preference
0.00
100.00
V16: Old extract fan speed
0.00
100.00
V17: New extract fan speed

Fuzzy Logic Controllers for Energy Efficiency Consumption in Buildings

Initial Rule Base and FLC Structure



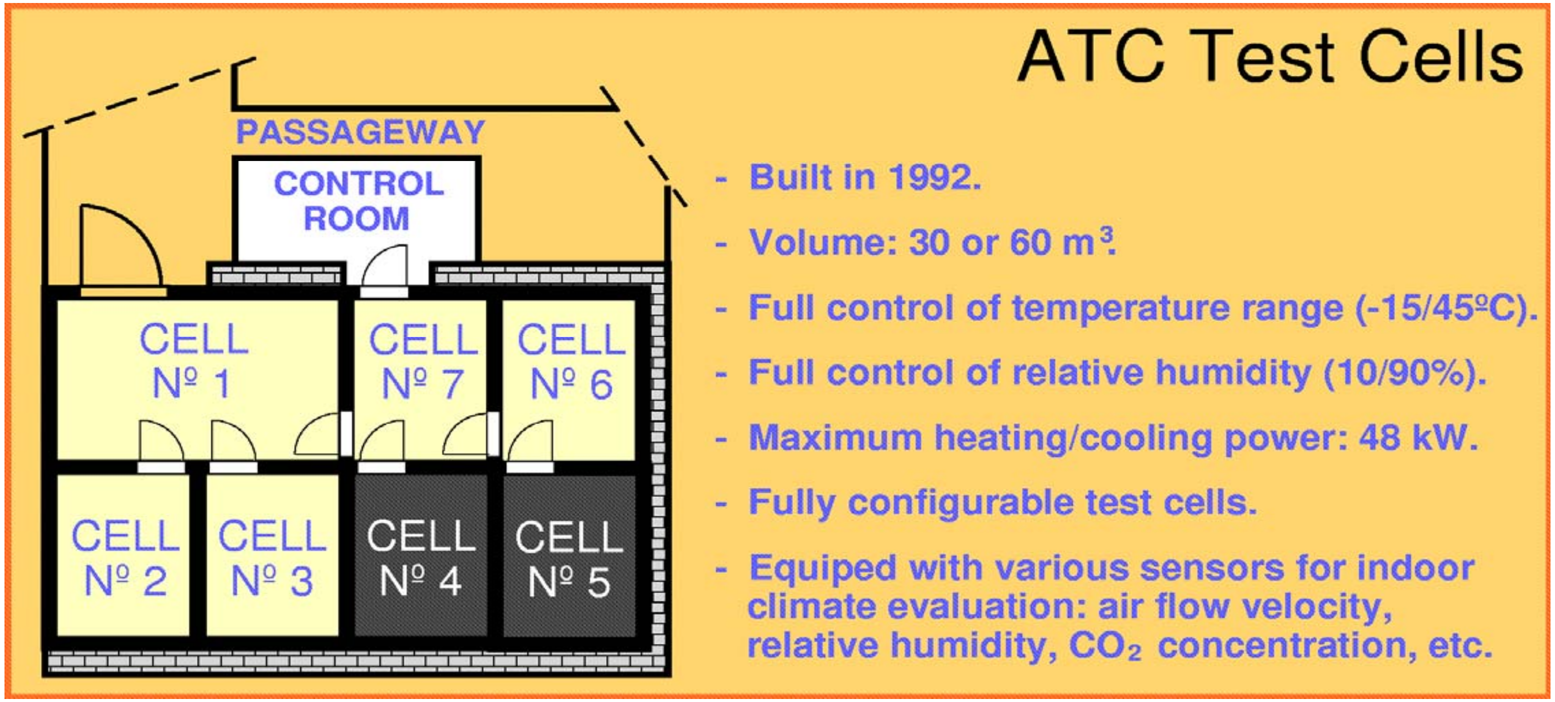
172 Rules



V1: PMV
V2: PMV t-1
V3: Thermal preference
V4: Tout-Tin
V5: Required heat
V6: CO_2
V7: dCO_2/dt
V8: Air quality preference
V9: Integral of PMV
V10: Integral of energy consumption
V11: Thermal/Energy priority
V12: Ventilation/Energy priority
V13: Valve old position
V14: Valve new position
V15: Fan coil speed
V16: Old extract Fan Speed
V17: New extract fan speed
M-1a
M-1a2
M-1b
M-2
M-3a
M-3b
Layer 1: System Demands
Layer 2: Priorities
Layer 3: Control Decisions
Module $1a_1$: Thermal Demands
Module $1a_2$: Thermal Preference
Module 1b: Air Quality Demands
Module 2: Energy Priorities
Module 3a: Required HVAC System Status
Module 3b: Required Ventilation System Status

Fuzzy Logic Controllers for Energy Efficiency Consumption in Buildings

Representation of the Test Cells



- **Two adjacent twin cells were available**

- **A calibrated and validated model of this site was developed to evaluate each FLC**

Fuzzy Logic Controllers for Energy Efficiency Consumption in Buildings

- **Goal**: **multi-criteria optimization of an expert FLC for an HVAC system: reduction of the energy consumption but maintaining the required indoor comfort levels**

O_1 Upper thermal comfort limit [3]: $if\ PMV > 0.5, O_1 = O_1 + (PMV - 0.5)$.

O_2 Lower thermal comfort limit: $if\ PMV < -0.5, O_2 = O_2 + (-PMV - 0.5)$.

O_3 IAQ requirement: $if\ CO_2\ conc. > 800ppm, O_3 = O_3 + (CO_2 - 800)$.

O_4 Energy consumption: $O_4 = O_4 +$ Power at time t.

O_5 System stability: $C_5 = C_5 +$ System change from time t to $(t-1)$.

- **INITIAL RESULTS**

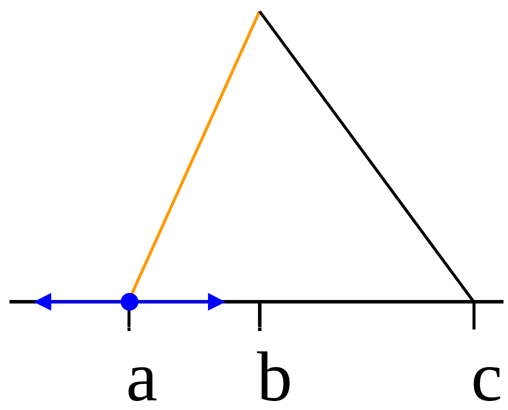
MODELS	#R	PMV>0.5 0_1		PMV<-0.5 0_2		$C0_2$ 0_3		ENERGY 0_4	%	STABILITY 0_5	%
ON-OFF	-	0,0		0		0		3206400	-	1136	-
FLC	172	0,0		0		0		2901686	9,50	1505	-32,48

GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

Improving the FLC Performance

The main objective was the reduction of the energy consumption (10%), improving the stability of the controller, maintaining the required indoor comfort levels

- **Genetic tuning of the Data Base**
  - **Local modification of the membership function definition points**



GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

Objectives (to be minimized)

O_1 **Upper thermal comfort limit** [3]: $if\ PMV > 0.5, O_1 = O_1 + (PMV - 0.5)$.

O_2 **Lower thermal comfort limit:** $if\ PMV < -0.5, O_2 = O_2 + (-PMV - 0.5)$.

O_3 **IAQ requirement:** $if\ CO_2\ conc. > 800ppm, O_3 = O_3 + (CO_2 - 800)$.

O_4 **Energy consumption:** $O_4 = O_4 +$ Power at time t.

O_5 **System stability:** $C_5 = C_5 +$ System change from time t to $(t - 1)$.

- Expert knowledge as objective weights:

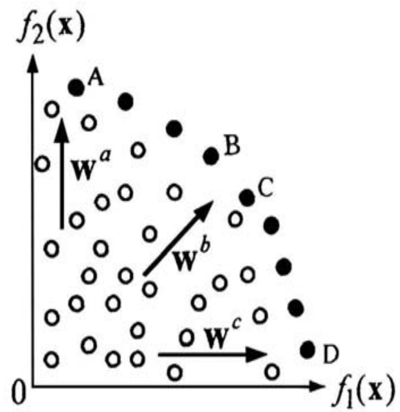
$$w_1^o = w_2^o = 0.0041511\ ;\ w_3^o = 0.0000022833$$

$$w_4^o = 0.0000017832\ ;\ w_5^o = 0.000761667$$

GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

Problem Restrictions

- **Different-Criteria-Based Evaluation**
  - Multiple Criteria Algorithms:
    - Multi-objective approach
    - Aggregation approach



$$F(x) = w_1 \cdot f_1(x) + \ldots + w_n \cdot f_n(x)$$

$$\sum w_i = 1, \qquad 0 \le w_i \le 1, \qquad i = \{1, \ldots, n\}$$

Since trusted weights exist:

- The problem solving is easier
- Quicker algorithms can be designed

GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

Problem Restrictions

- **The controller accuracy is assessed by means of simulations which approximately take 3-4 minutes**
  - Efficient tuning methodologies:
    - Local adjustment of each tuned parameter
    - Steady-State Genetic Algorithms: quick convergence

      2000 evaluations $\Rightarrow$ 1 run takes approximately 4 days
    - Considering a small population (31 individuals)

GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

Genetic Tuning of the Data Base

- **A real coded steady-state genetic algorithm for local tuning of the membership function definition points.**
  - **Two individuals are selected to be crossed and four descendents are obtained**
  - **The two best offspring are included in the population replacing the two worst individuals if they are better adapted than the latter**
  - **A restarting approach is considered if the population converges**

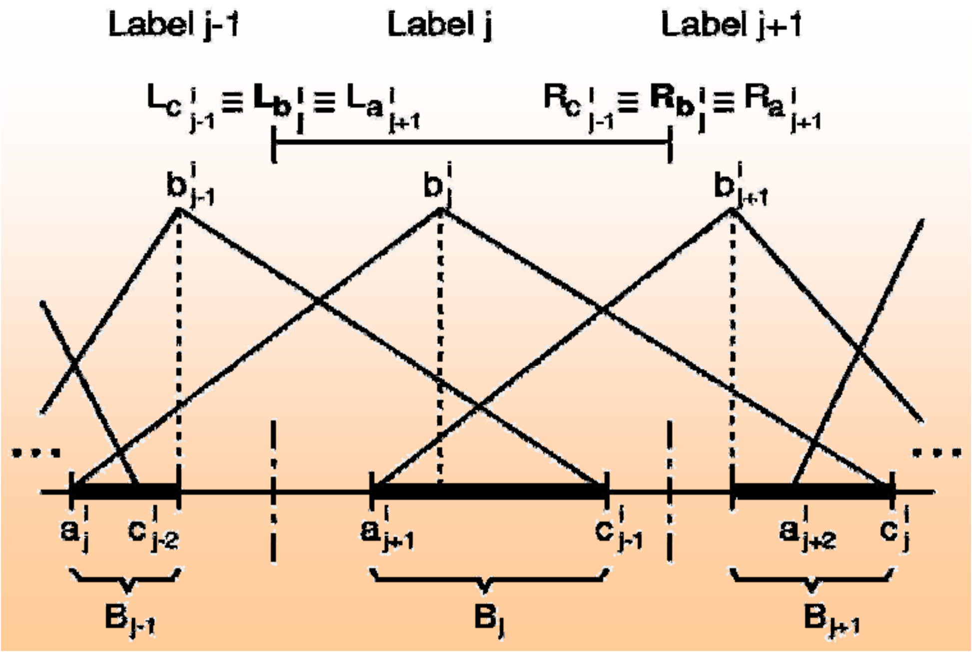
GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

Data Base Tuning: Algorithm (1)

- **Coding Scheme** (with *n* variables and L_i labels)**:**

$$C_i = \left(a_1^i, b_1^i, c_1^i, \ldots, a_{L_i}^i, b_{L_i}^i, c_{L_i}^i\right),\ i = 1, \ldots, n$$

$$C = C_1 C_2 \ldots C_n$$



GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

Data Base Tuning: Algorithm (2)

- **Genetic operators:**

$$\begin{aligned} C^v &= (c_1, \ldots, c_k, \ldots, c_H) \\ C^w &= (c'_1, \ldots, c'_k, \ldots, c'_H) \end{aligned}$$

  - The max-min-arithmetical crossover. From parents C^v and C^w, four offspring are obtained:

$$\begin{aligned} &C^{1'} = aC^w + (1-a)C^v \\ &C^{2'} = aC^v + (1-a)C^w \\ &C^{3'} \text{ with } c_{3k} = \min\{c_k, c'_k\} \\ &C^{4'} \text{ with } c_{4k} = \max\{c_k, c'_k\} \end{aligned}$$

  - Michalewicz's non-uniform mutation.

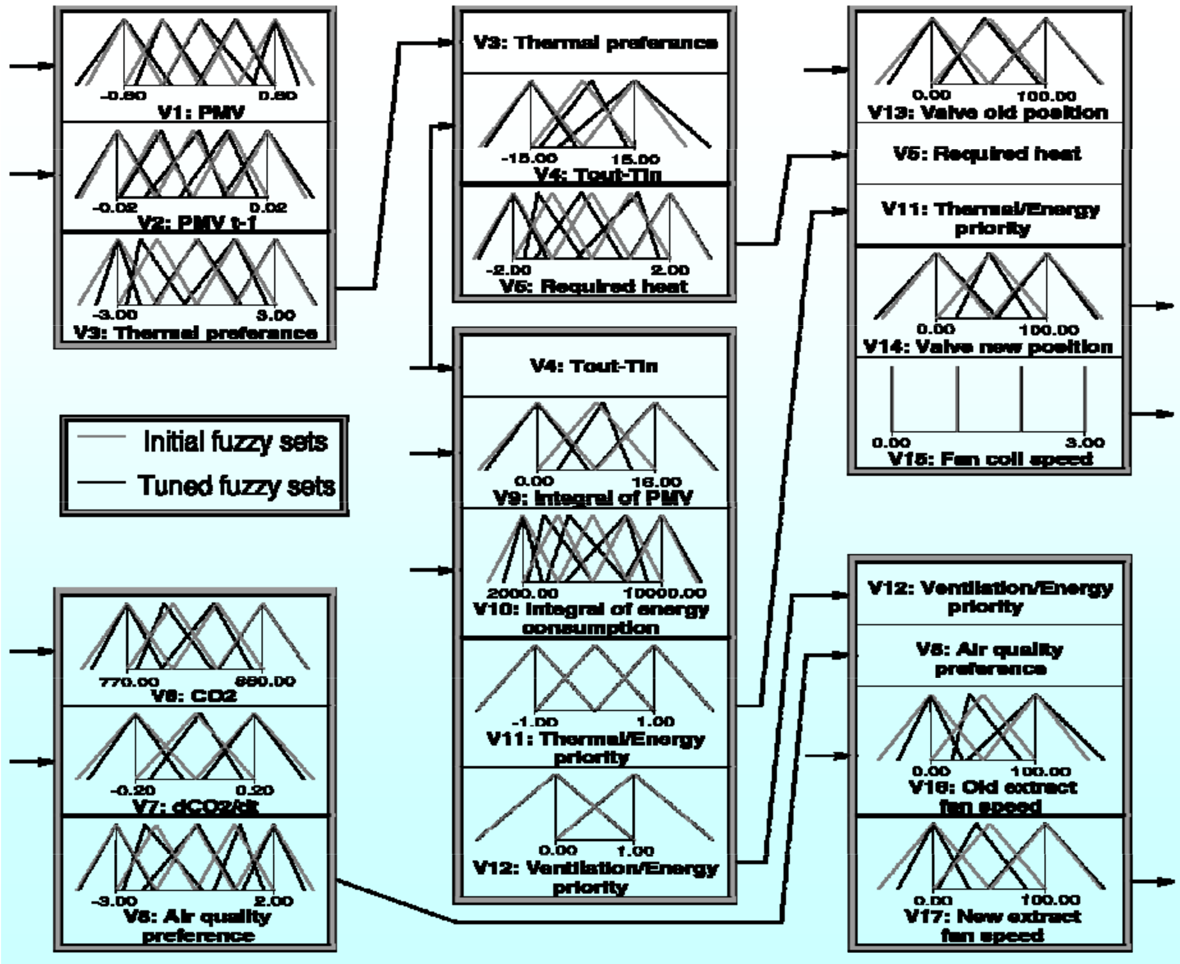
GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

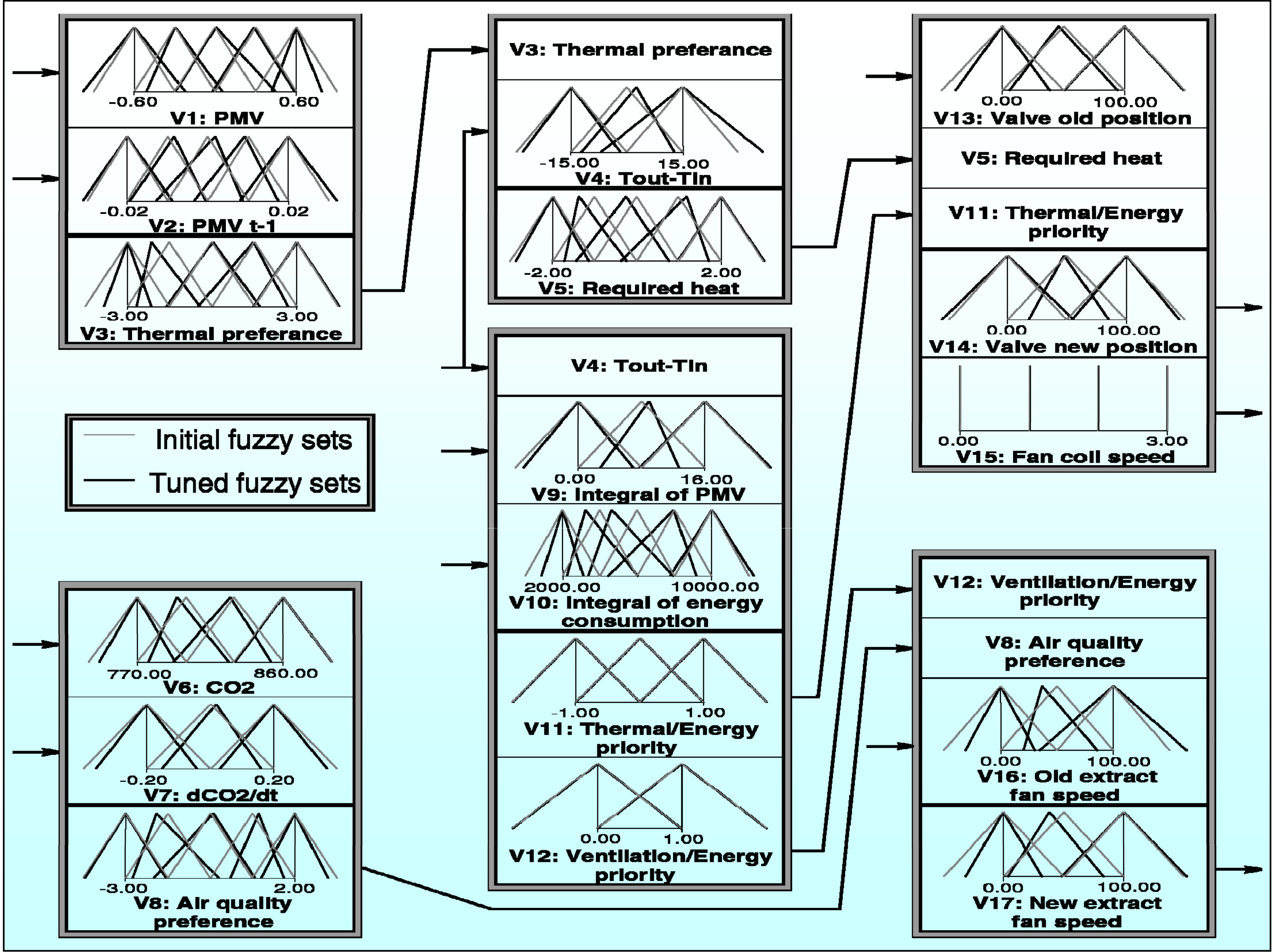
MODELS	#R	PMV>0.5 0_1		PMV<-0.5 0_2		$C0_2$ 0_3		ENERGY 0_4	ENERGY %	STABILITY 0_5	STABILITY %
CLASSICAL ON-OFF	-	0,0		0		0		3206400	-	1136	-
FLC	172	0,0		0		0		2901686	9,50	1505	-32,48
FLC TUNING	172	0,0		0		0		2596875	19,01	1051	7,48

R. Alcalá, J.M. Benítez, J. Casillas, O. Cordón, R. Perez, Fuzzy control of HVAC systems optimised by genetic algorithms, Appl. Intell. 18 (2003) 155–177

GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning

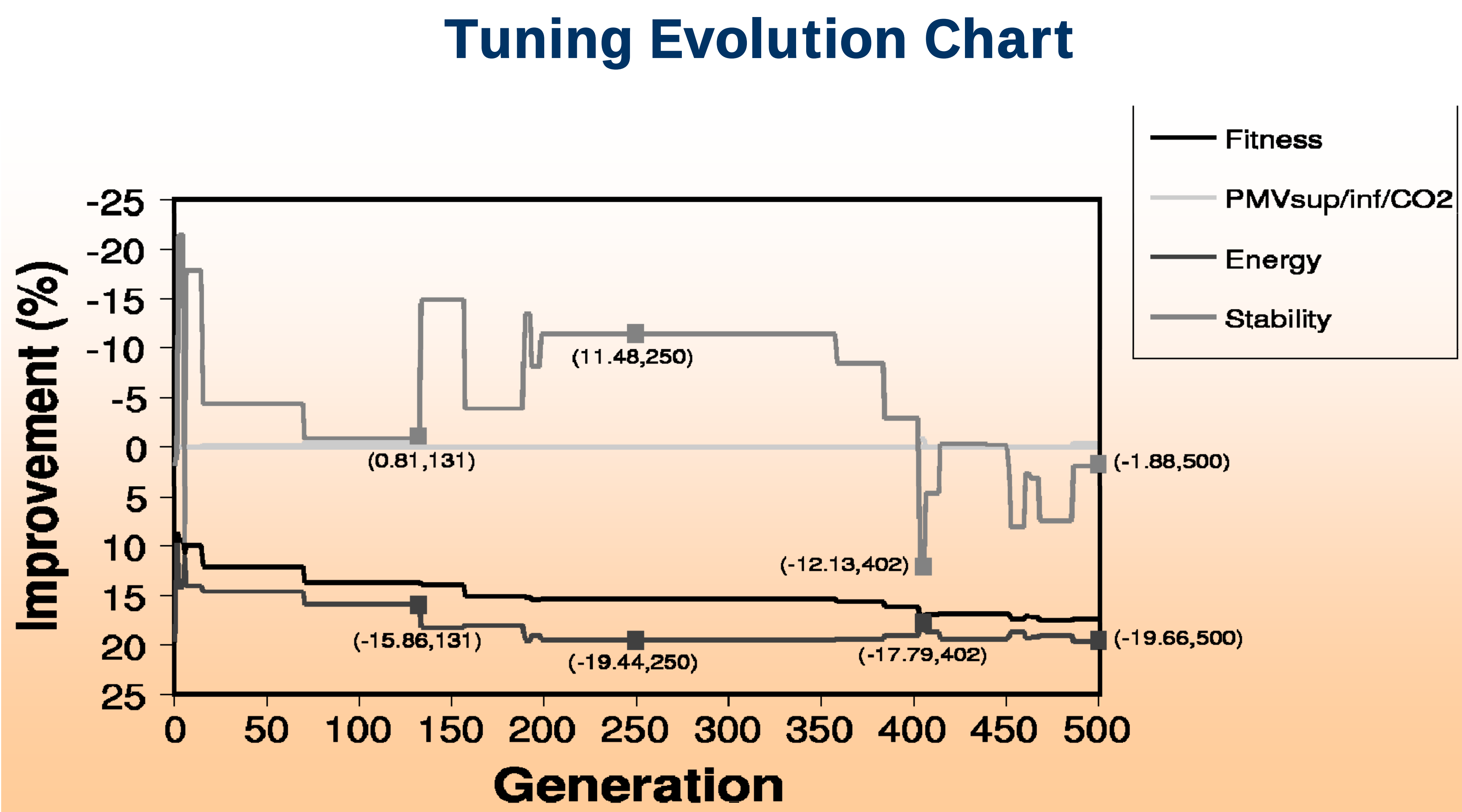
Tuned Data Base





-0.60 0.60
V1: PMV
-0.02 0.02
V2: PMV t-1
-3.00 3.00
V3: Thermal preferance
V3: Thermal preferance
-15.00 15.00
V4: Tout-Tin
-2.00 2.00
V5: Required heat
0.00 100.00
V13: Valve old position
V5: Required heat
V11: Thermal/Energy priority
0.00 100.00
V14: Valve new position
0.00 3.00
V15: Fan coil speed
Initial fuzzy sets
Tuned fuzzy sets
V4: Tout-Tin
0.00 16.00
V9: Integral of PMV
2000.00 10000.00
V10: Integral of energy consumption
-1.00 1.00
V11: Thermal/Energy priority
0.00 1.00
V12: Ventilation/Energy priority
770.00 860.00
V6: CO2
-0.20 0.20
V7: dCO2/dt
-3.00 2.00
V8: Air quality preference
V12: Ventilation/Energy priority
V8: Air quality preference
0.00 100.00
V16: Old extract fan speed
0.00 100.00
V17: New extract fan speed

GFS Models for Fuzzy Control of HVAC Systems: Genetic Tuning



GFS Models for Fuzzy Control of HVAC Systems: Genetic Rules Selection with Weights

GENETIC RULE WEIGHT DERIVATION AND RULE SELECTION

OBJECTIVE OF GETTING:

- **a subset of rules presenting good cooperation**
- **the weights associated to rules**

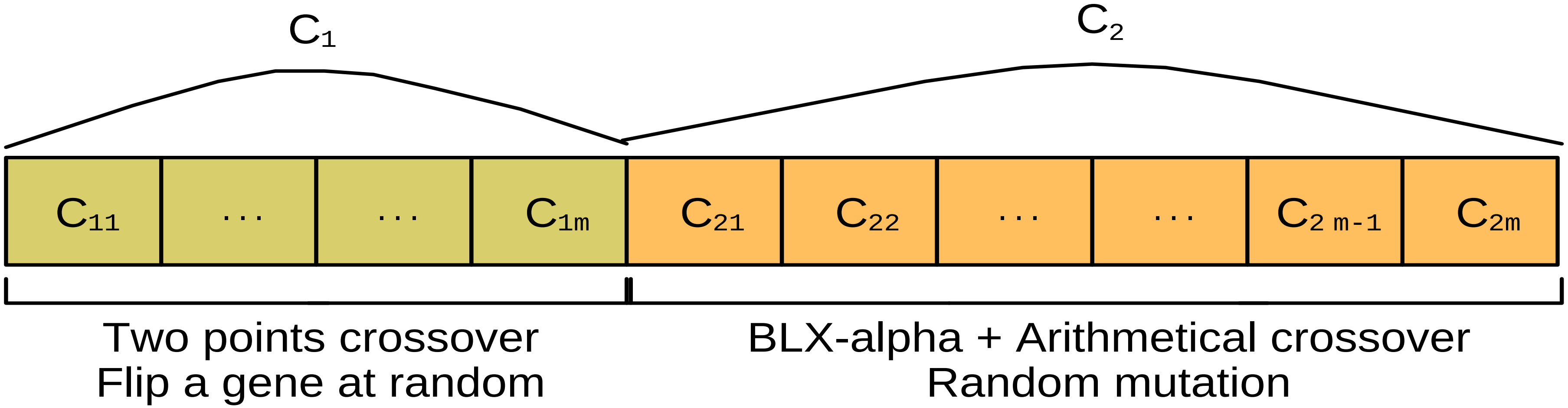
IF X_1 is A_1 and ... and X_n is A_n THEM Y is B with [w], $w \in [0,1]$

We use a steady-state genetic algorithm with a double coding scheme.

GFS Models for Fuzzy Control of HVAC Systems: Genetic Rules Selection with Weights

Weight Learning: Algorithm

- **A double coding scheme (*$C=C_1+C_2$*):**
  - C_1: The coding scheme generates binary-coded strings of length m (number of single rules in the previously derived rule set):
  - C_2: The coding scheme generates real-coded strings of length m. Each gene represents the weight used in the corresponding rule



GFS Models for Fuzzy Control of HVAC Systems: Genetic Rules Selection with Weights

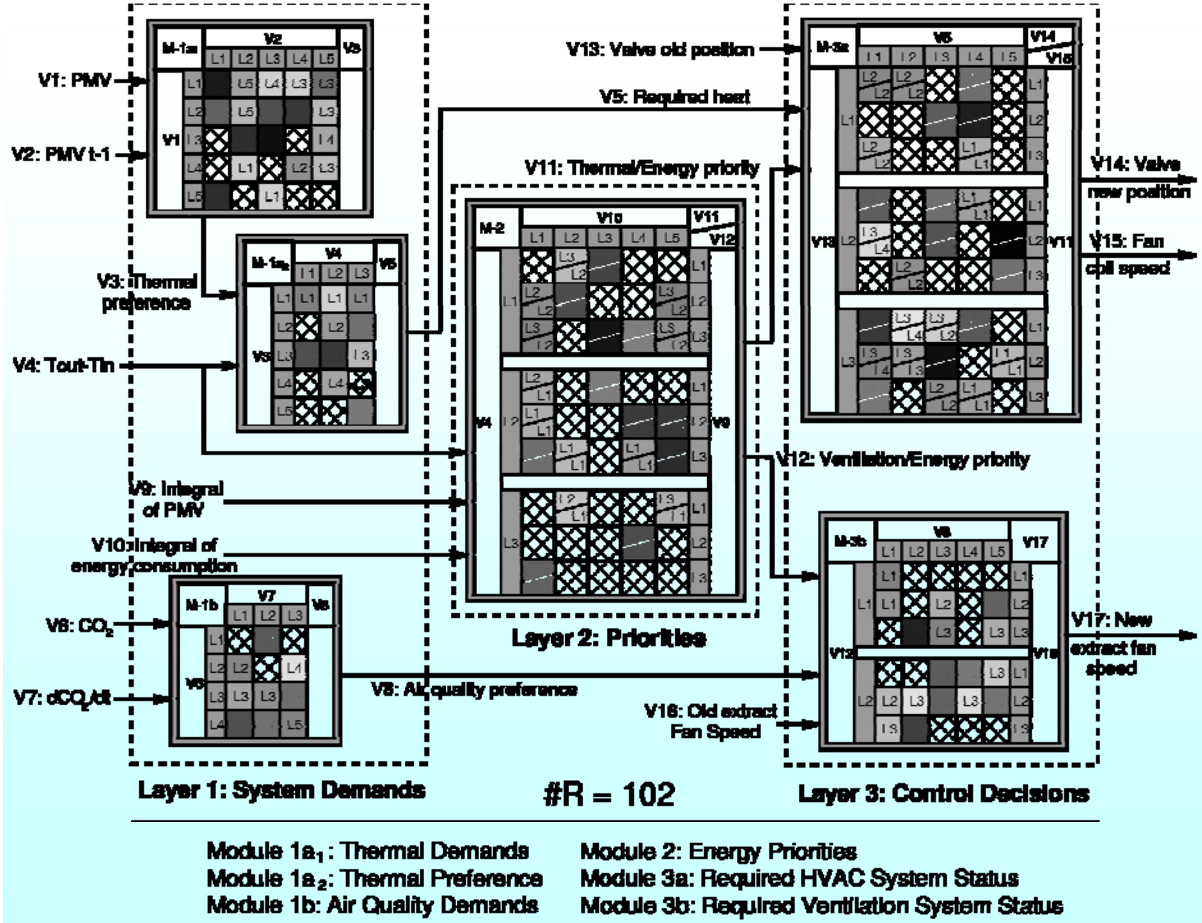
Obtained Results

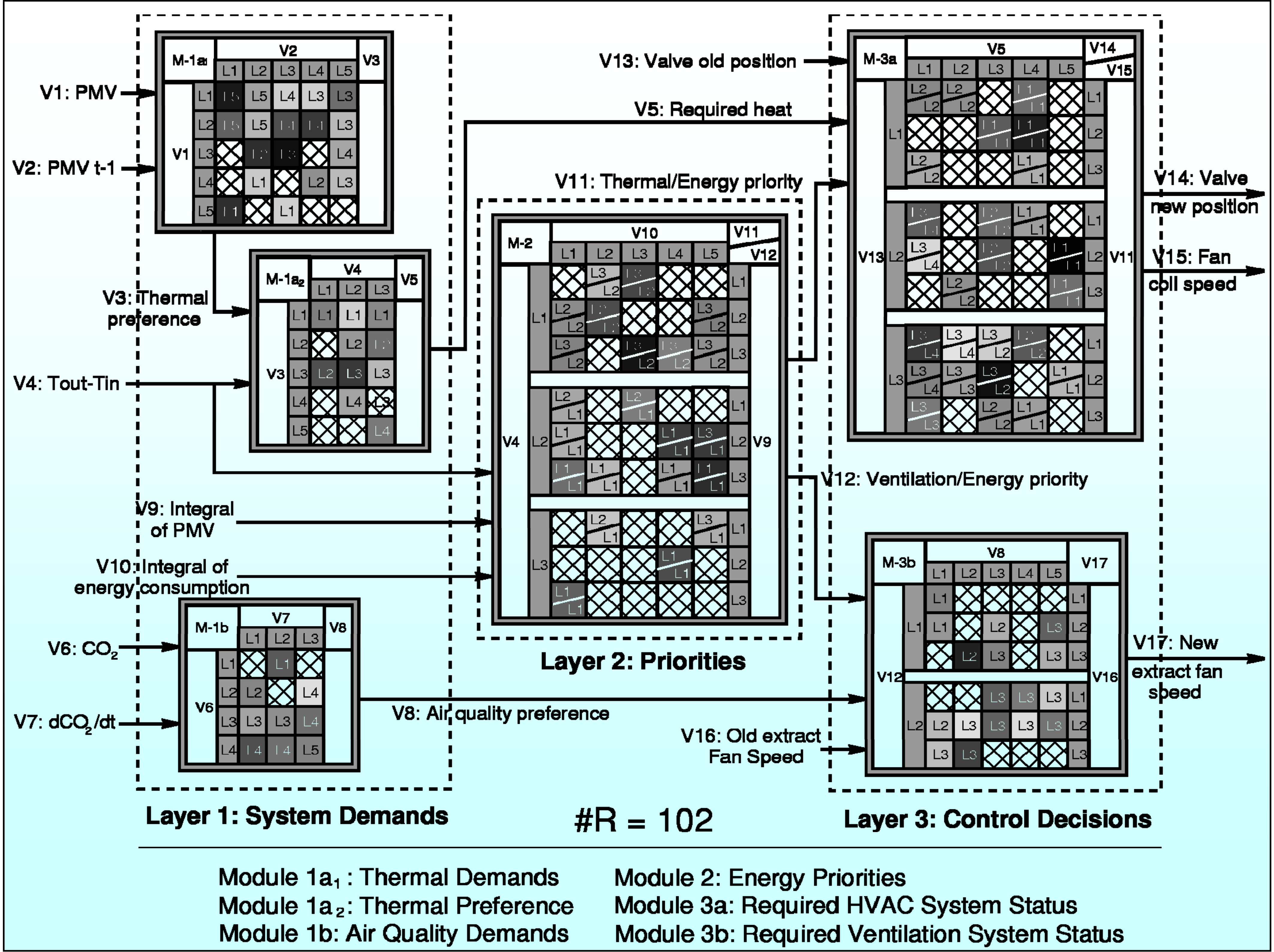
MODELS	#R	PMV>0.5 0_1	PMV<-0.5 0_2	$C0_2$ 0_3	ENERGY 0_4	%	STABILITY 0_5	%
ON-OFF	-	0,0	0	0	3206400	-	1136	-
FLC	172	0,0	0	0	2901686	9,50	1505	-32,48
TUNING	172	0,0	0	0	2596875	19,01	1051	7,48
SELECTION	147	0,2	0	0	2867692	10,56	991	12,76
SEL. + TUNING	109	0,1	0	0	2492462	22,27	989	12,94
SEL + WEIGTS	102	0,7	0	0	2731798	14,80	942	17,08

R. Alcalá, J. Casillas, O. Cordón, A. González, F. Herrera, **A Genetic Rule Weighting and Selection Process for Fuzzy Control of Heating, Ventilating and Air Conditioning Systems**. *Engineering Applications of Artificial Intelligence 18:3 (2005) 279-296*

GFS Models for Fuzzy Control of HVAC Systems: Genetic Rules Selection with Weights

Weighted Rule Base



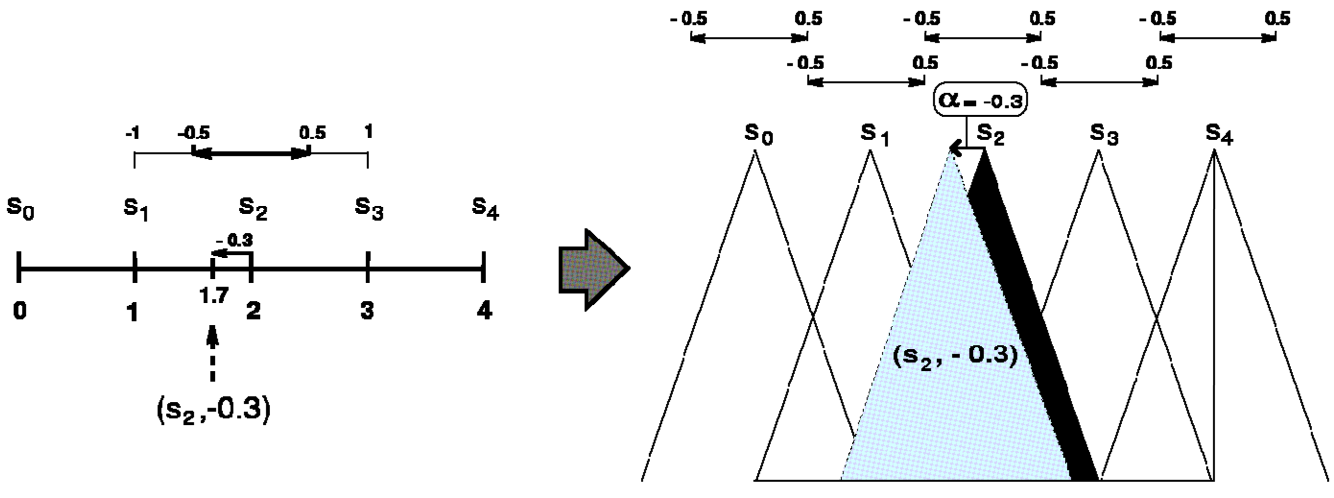

V1: PMV
V2: PMV t-1
V3: Thermal preference
V4: Tout-Tin
V9: Integral of PMV
V10: Integral of energy consumption
V6: CO_2
V7: dCO_2/dt
M-1a1
M-1a2
M-1b
M-2
M-3a
M-3b
V5: Required heat
V13: Valve old position
V11: Thermal/Energy priority
V12: Ventilation/Energy priority
V8: Air quality preference
V16: Old extract Fan Speed
V14: Valve new position
V15: Fan coil speed
V17: New extract fan speed
Layer 1: System Demands
Layer 2: Priorities
Layer 3: Control Decisions
#R = 102
Module $1a_1$: Thermal Demands
Module $1a_2$: Thermal Preference
Module 1b: Air Quality Demands
Module 2: Energy Priorities
Module 3a: Required HVAC System Status
Module 3b: Required Ventilation System Status

GFS Models for Fuzzy Control of HVAC Systems: Genetic Lateral Tuning and Rules Selection

New coding schemes: 2- and 3-tuples:

R. Alcalá, J. Alcalá-Fdez, M.J. Gacto, F. Herrera, Improving Fuzzy Logic Controllers Obtained by Experts: A Case Study in HVAC Systems. *Applied Intelligence, doi:10.1007/s10489-007-0107-6, in press (2008)*.

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



a) Simbolic Translation of a label b) Lateral Displacement of a Membership function

- New rule structure:

  IF X_1 IS (S^1_i, α_1) AND ... AND X_n IS (S^n_i, α_n) THEN Y IS (S^y_i, α_y)

GFS Models for Fuzzy Control of HVAC Systems: Genetic Lateral Tuning and Rules Selection

GENETIC LATERAL TUNING

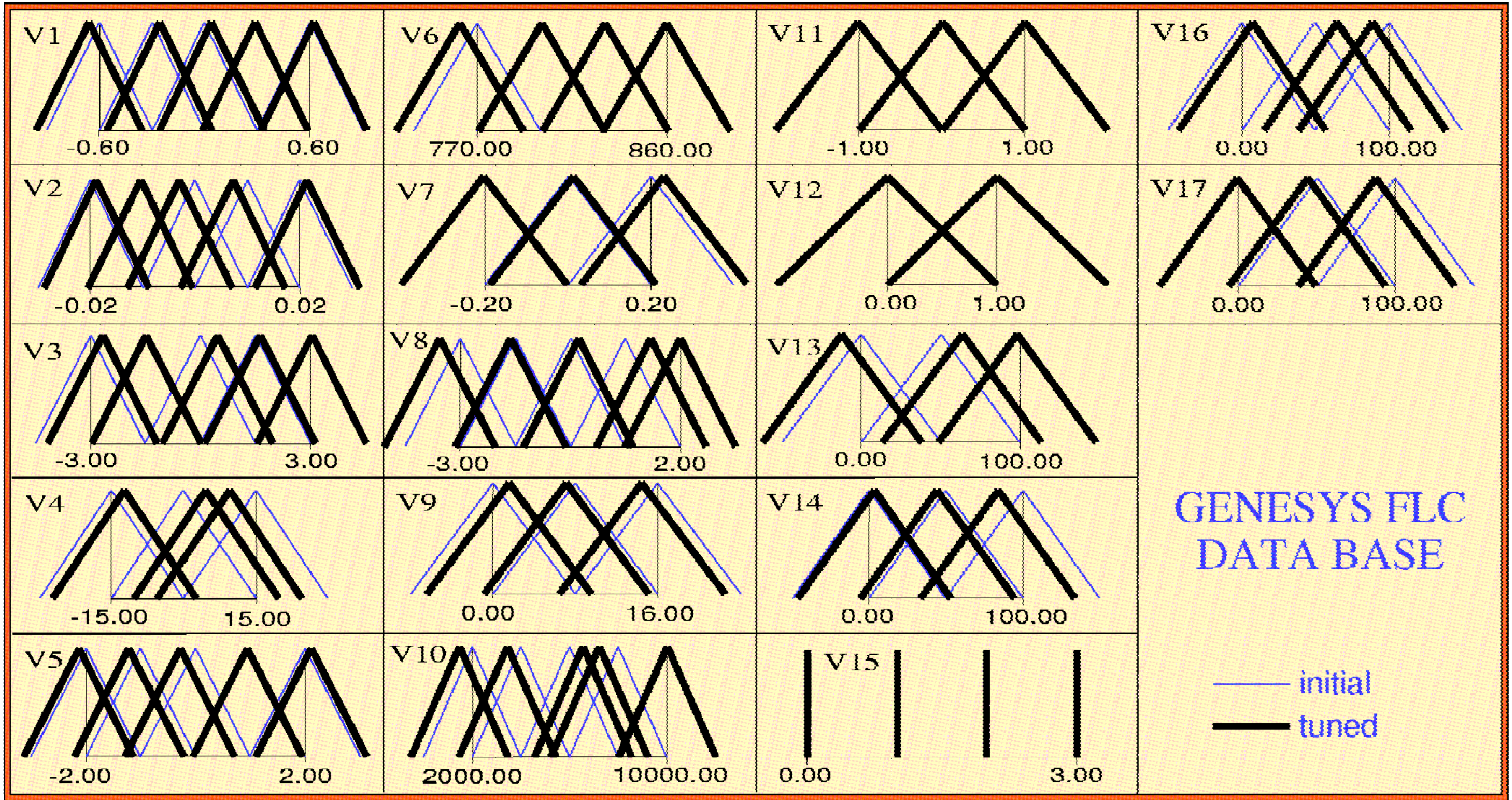
MODELS	#R	PMV>0.5 0_1	PMV<-0.5 0_2	$C0_2$ 0_3	ENERGY 0_4	ENERGY %	ESTABILITY 0_5	ESTABILITY %
ON-OFF	-	0,0	0	0	3206400	-	1136	-
FLC	172	0,0	0	0	2901686	9,50	1505	-32,48
TUNING	172	0,0	0	0	2596875	19,01	1051	7,48
SELECTION	147	0,2	0	0	2867692	10,56	991	12,76
SELEC. + TUNING	109	0,1	0	0	2492462	22,27	989	12,94
SEL + WEIGHTS	102	0,7	0	0	2731798	14,80	942	17,08
GL 1	172	0,7	0	0	2378784	25,81	1069	5,90
GL 2	172	1,0	0	0	2327806	27,40	1066	6,16
GL 3	172	0,9	0	0	2268689	29,25	1080	4,93
LL 1	172	0,9	0	0	2386033	25,59	896	21,13
LL 2	172	0,8	0	0	2343409	26,92	943	16,99
LL 3	172	0,3	0	0	2377596	25,85	938	17,43

GFS Models for Fuzzy Control of HVAC Systems: Genetic Lateral Tuning and Rule Selection

MODELS	#R	PMV>0.5 0_1	PMV<-0.5 0_2	$C0_2$ 0_3	ENERGY 0_4	%	ESTABILITY 0_5	%
ON-OFF	-	0,0	0	0	3206400	-	1136	-
FLC	172	0,0	0	0	2901686	9,50	1505	-32,48
TUNING	172	0,0	0	0	2596875	19,01	1051	7,48
SELECTION	147	0,2	0	0	2867692	10,56	991	12,76
SEL + TUNING	109	0,1	0	0	2492462	22,27	989	12,94
SEL + WEIGHTS	102	0,7	0	0	2731798	14,80	942	17,08
GL 2	172	0,9	0	0	2268689	29,25	1080	4,93
LL 1	172	0,9	0	0	2386033	25,59	896	21,13
GL - S 1	105	1,0	0	0	2218598	30,81	710	37,50
GL - S 2	115	0,4	0	0	2358405	26,45	818	27,99
GL - S 3	118	0,8	0	0	2286976	28,68	872	23,24
LL – S 1	133	0,5	0	0	2311986	27,90	788	30,63
LL – S 2	104	0,6	0	0	2388470	25,51	595	47,62
LL – S 3	93	0,5	0	0	2277807	28,96	1028	9,51

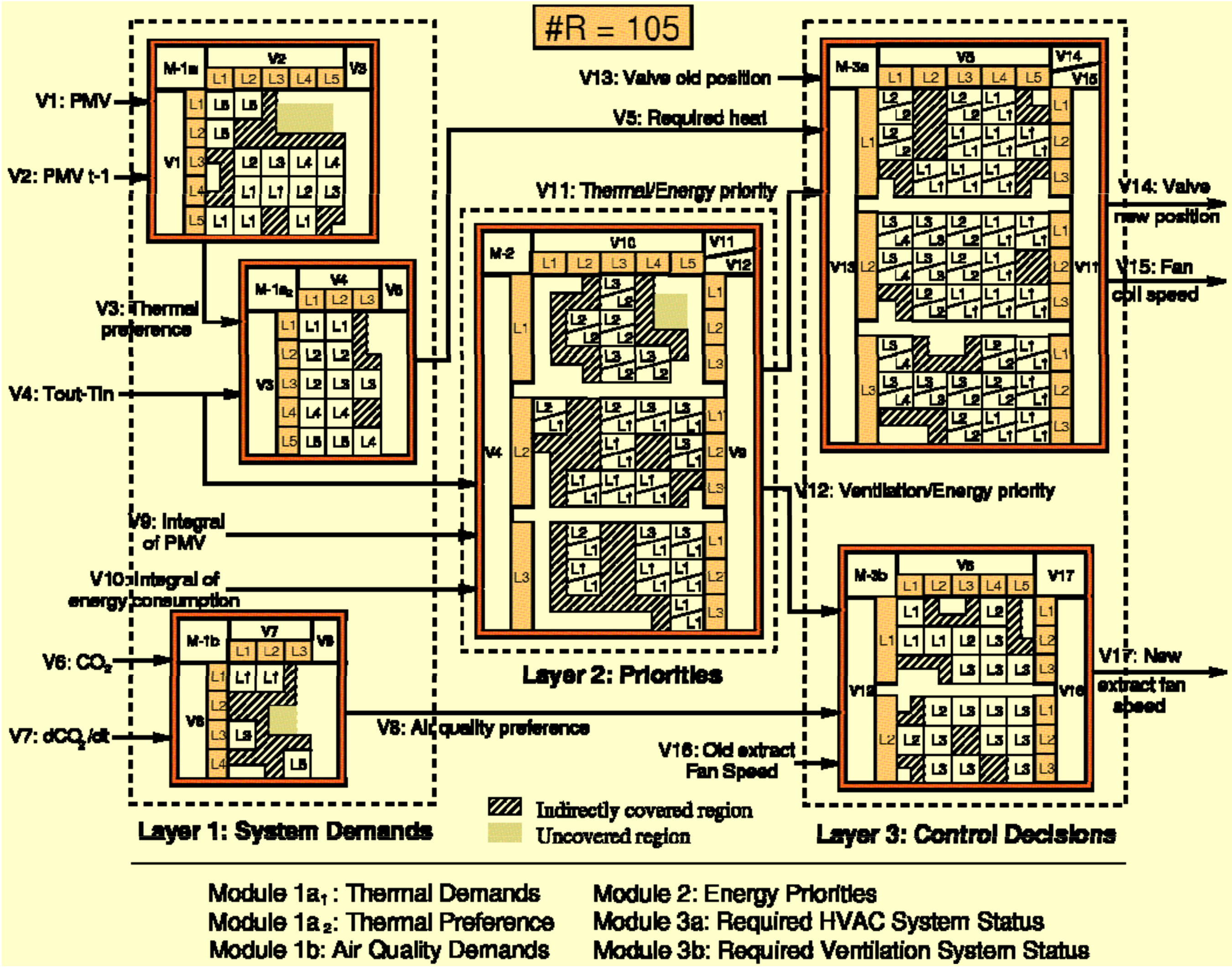
GFS Models for Fuzzy Control of HVAC Systems: Genetic Lateral Tuning and Rules Selection

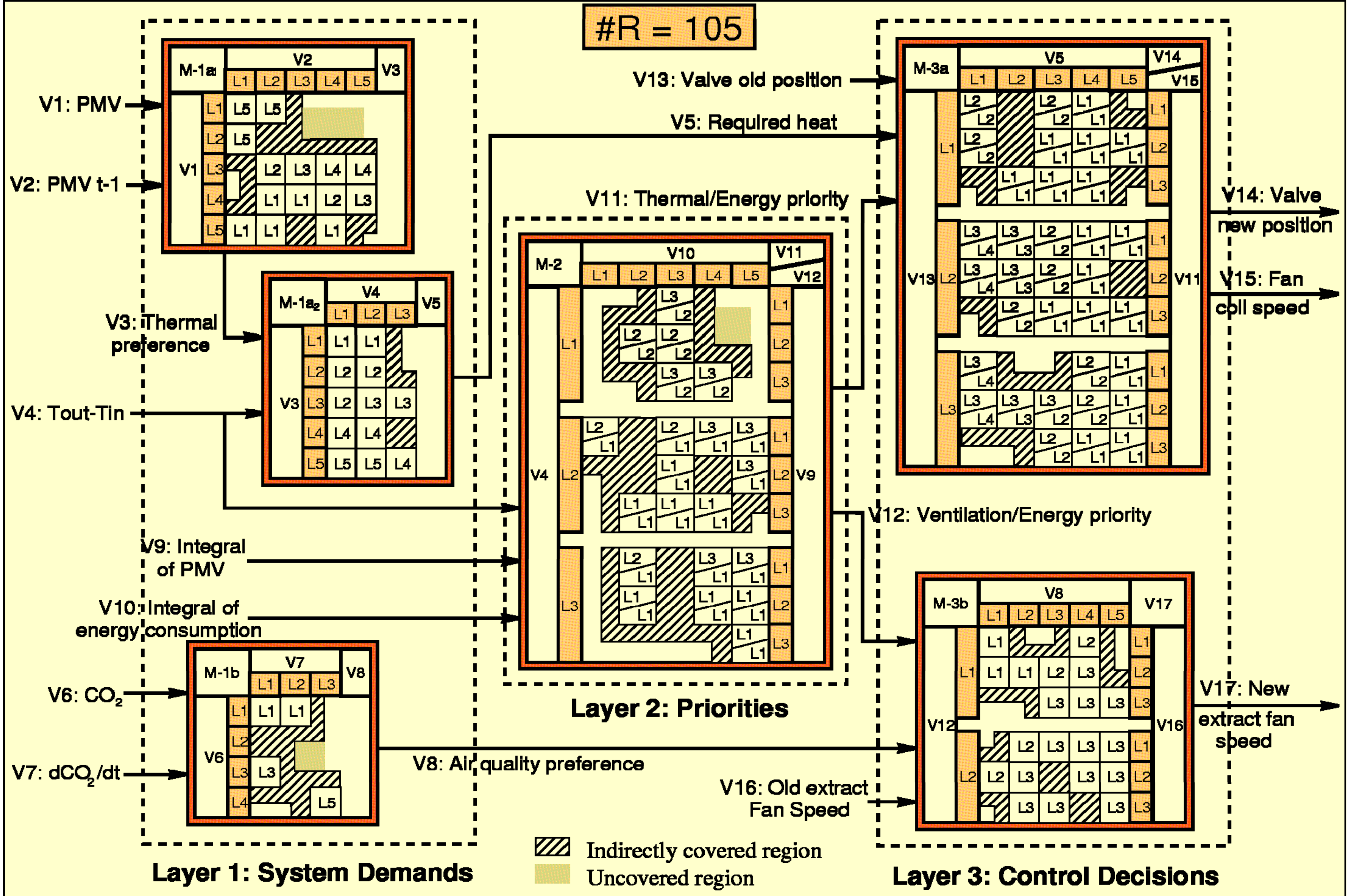
Tuned Data Base (GL-S_1):



GFS Models for Fuzzy Control of HVAC Systems: Genetic Lateral Tuning and Rules Selection

Selected Rule Base (GL-S_1):





Module $1a_1$: Thermal Demands

Module $1a_2$: Thermal Preference

Module 1b: Air Quality Demands

Module 2: Energy Priorities

Module 3a: Required HVAC System Status

Module 3b: Required Ventilation System Status

GFS Models for Fuzzy Control of HVAC Systems: Genetic Lateral Tuning and Rules Selection

⇨**The combination of lateral tuning (global and local) and rules selection allow us to eliminate redundant rules, tuning the parameters, and geting and high behaviour reducting the energy comsuption and with good stability.**

⇨ **¿What is the reason of the good behavior?**

The SBRDs tuning for an HVAC system is a large scale problem wiht 17 variables and a lot of parameters, and the use of 1 parameter per label allows us to reduce the search space, allowing to get a better optimal local than using 3 parameters per label.

GFS Models for Fuzzy Control of HVAC Systems

Bibliography

R. Alcalá, J.M. Benítez, J. Casillas, O. Cordón, R. Pérez, Fuzzy Control of HVAC Systems Optimized by Genetic Algorithms. *Applied Intelligence 18:2 (2003) 155-177*.

R. Alcalá, J. Casillas, O. Cordón, A. González, F. Herrera, A Genetic Rule Weighting and Selection Process for Fuzzy Control of Heating, Ventilating and Air Conditioning Systems. *Engineering Applications of Artificial Intelligence 18:3 (2005) 279-296*.

R. Alcalá, J. Alcalá-Fdez, M.J. Gacto, F. Herrera, Improving Fuzzy Logic Controllers Obtained by Experts: A Case Study in HVAC Systems. *Applied Intelligence, doi:10.1007/s10489-007-0107-6, in press (2008)*.

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

Currents Trends and Prospects

- Multiobjective genetic learning of FRBSs: interpretability-precision trade-off.
- GA-based techniques for mining fuzzy association rules and novel data mining approaches.
- Learning genetic models based on low quality data (noise data and vague data).
- Genetic learning of fuzzy partitions and context adaptation.
- Genetic adaptation of inference engine components.
- Revisiting the Michigan-style GFSs.

Currents Trends and Prospects

- Multiobjective genetic learning of FRBSs: interpretability-precision trade-off.

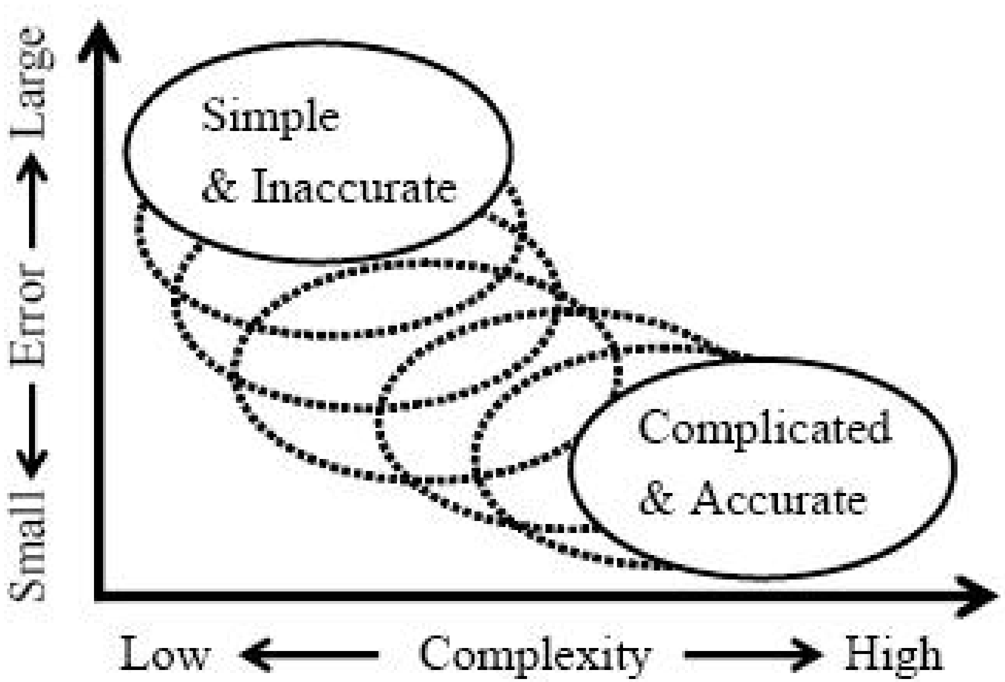


Fig. Non-dominated fuzzy systems

Currents Trends and Prospects

- **GA-based techniques for mining fuzzy association rules and novel data mining approaches.**
- **Genetic learning of fuzzy partitions and context adaptation.**

J. Alcalá-Fdez, R. Alcalá, M.J. Gacto, F. Herrera, **Learning the Membership Function Contexts for Mining Fuzzy Association Rules by Using Genetic Algorithms**. *Fuzzy Sets and Systems, doi:10.1016/j.fss.2008.05.012 ,* ***in press (2008)***.

Currents Trends and Prospects

- GA-based techniques for mining fuzzy association rules and novel data mining approaches.
- Genetic learning of fuzzy partitions and context adaptation.

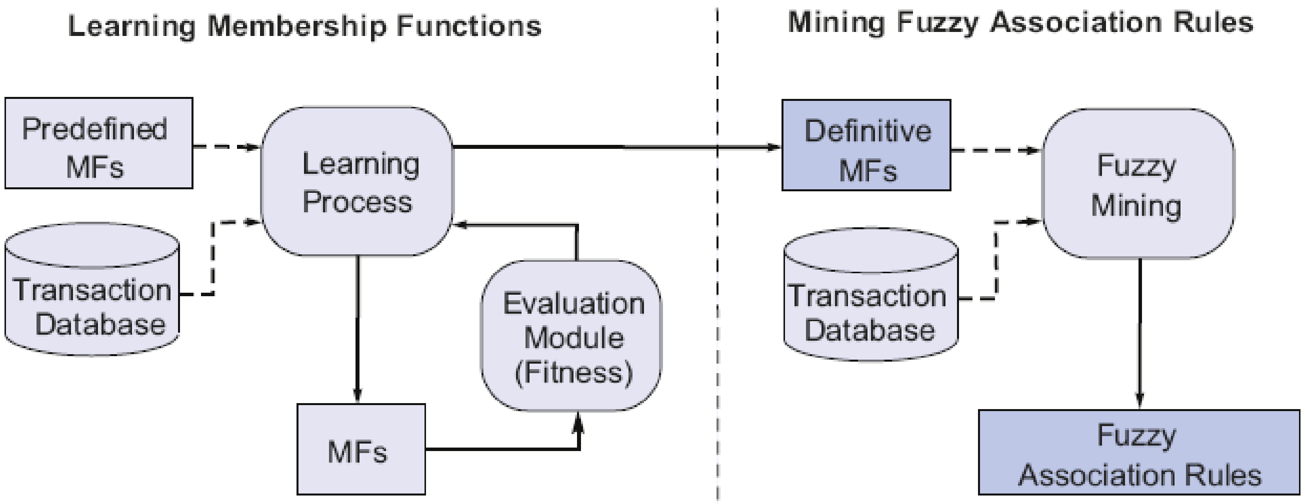


Fig. 1. Scheme for discovering both useful fuzzy association rules and suitable MFs.

J. Alcalá-Fdez, R. Alcalá, M.J. Gacto, F. Herrera, **Learning the Membership Function Contexts for Mining Fuzzy Association Rules by Using Genetic Algorithms**. *Fuzzy Sets and Systems,* *doi:10.1016/j.fss.2008.05.012* , ***in press (2008)***.

Currents Trends and Prospects

- GA-based techniques for mining fuzzy association rules and novel data mining approaches.
- Genetic learning of fuzzy partitions and context adaptation.

Census

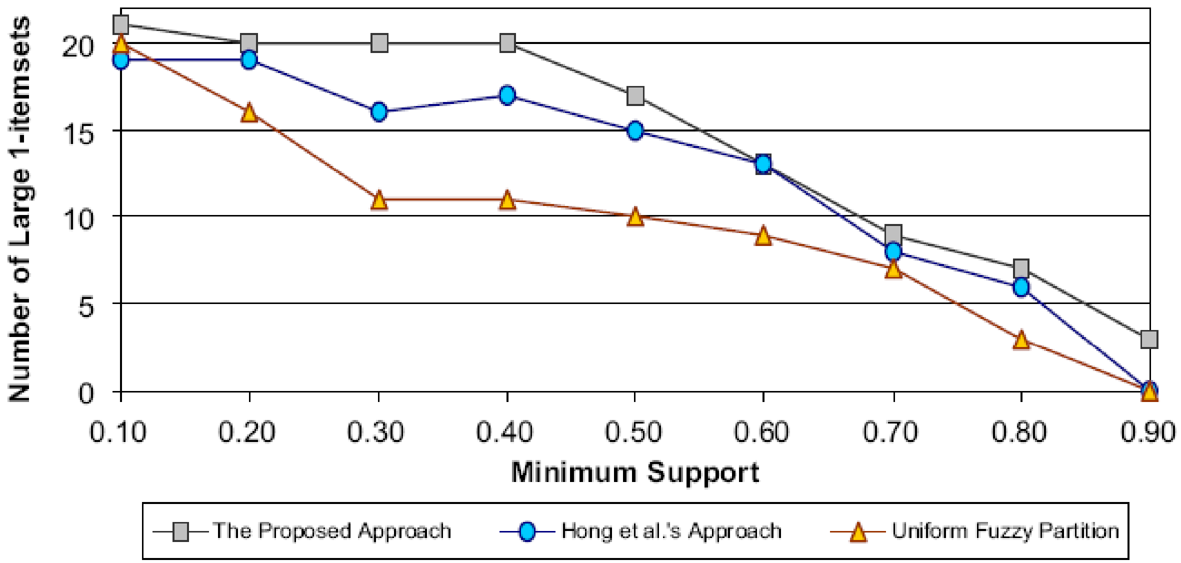


Fig. 9. Relationship between large 1-itemsets and minimum support.

J. Alcalá-Fdez, R. Alcalá, M.J. Gacto, F. Herrera, **Learning the Membership Function Contexts for Mining Fuzzy Association Rules by Using Genetic Algorithms**. *Fuzzy Sets and Systems,* *doi:10.1016/j.fss.2008.05.012* , ***in press (2008)***.

Currents Trends and Prospects

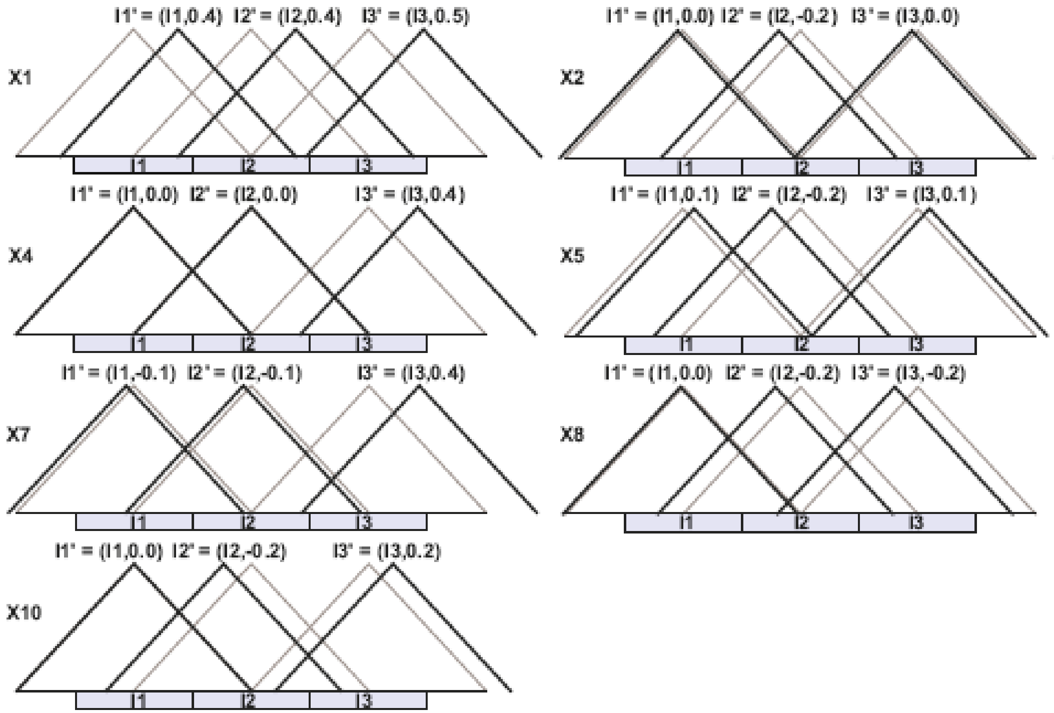


Fig. 10. MFs with/without lateral displacements (black/grey) and displacements of the MFs obtained by tl terms.

Classic Fuzzy Association Rule:

If **number of children** is Low and
hours head worked last week is Low
then **head's personal income** is Low
Factor of confidence 0.87

Rule with 2-Tuples Fuzzy Linguistic Representation:

If **number of children** is (Low, −0.16) and
hours head worked last week is (Low, −0.06)
then **head's personal income** is (Low, 0.1)
Factor of confidence 0.99

J. Alcalá-Fdez, R. Alcalá, M.J. Gacto, F. Herrera, **Learning the Membership Function Contexts for Mining Fuzzy Association Rules by Using Genetic Algorithms**. *Fuzzy Sets and Systems,* *doi:10.1016/j.fss.2008.05.012* , ***in press (2008)***.

Currents Trends and Prospects

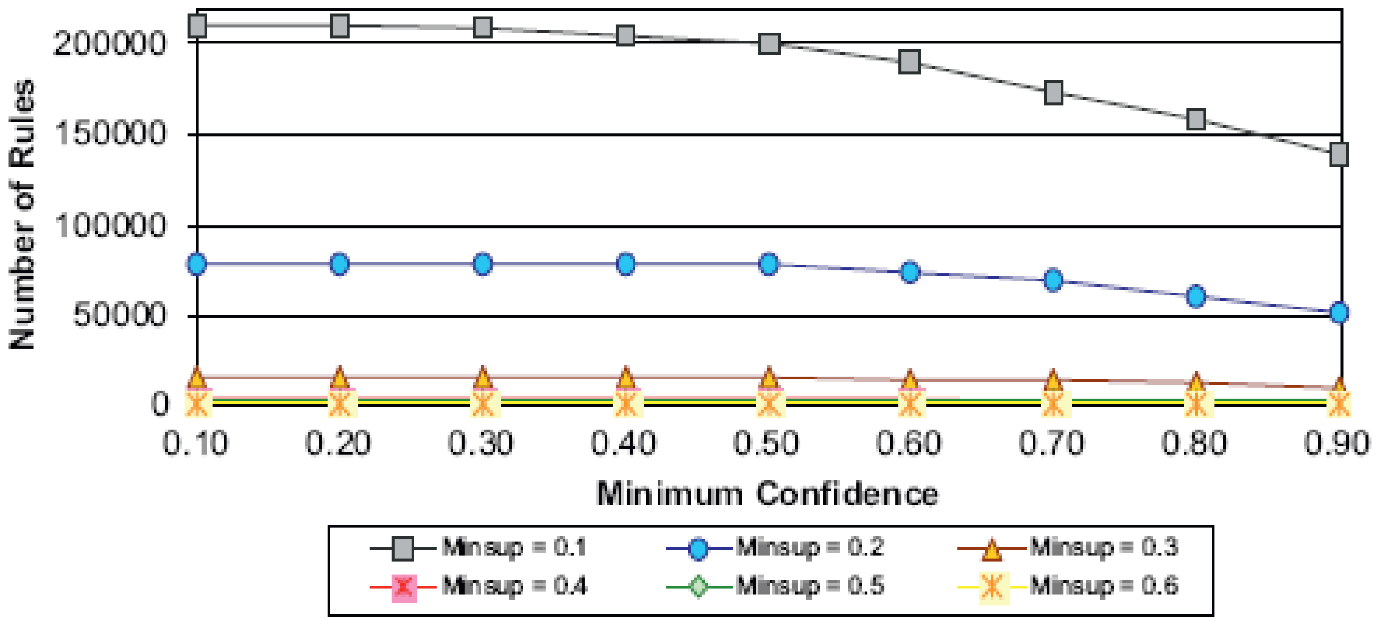


Fig. 15. Relationship between the number of fuzzy association rules and the confidence threshold along with different minimum supports.

J. Alcalá-Fdez, R. Alcalá, M.J. Gacto, F. Herrera, **Learning the Membership Function Contexts for Mining Fuzzy Association Rules by Using Genetic Algorithms**. *Fuzzy Sets and Systems,* *doi:10.1016/j.fss.2008.05.012* , ***in press (2008)***.

Currents Trends and Prospects

- Learning genetic models based on low quality data (noise data and vague data).
- Genetic adaptation of inference engine components.
- Revisiting the Michigan-style GFSs.

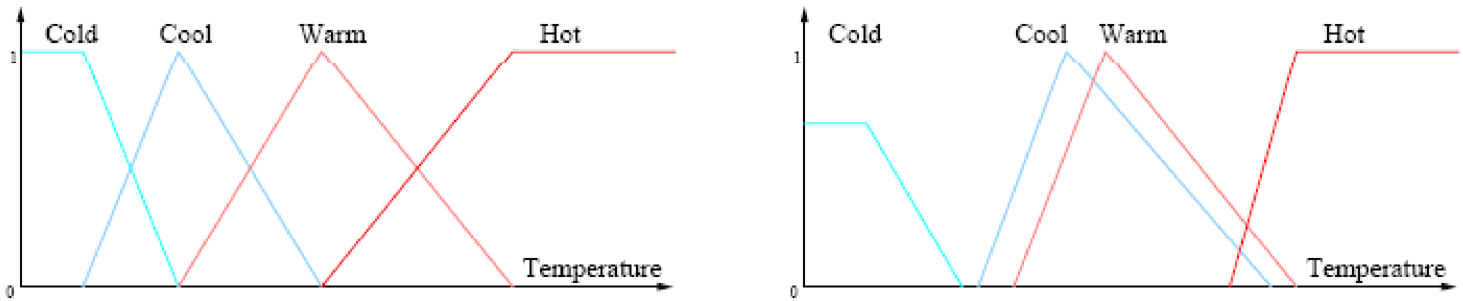
Currents Trends and Prospects

- **Interpretability quality:** associated to the meaning of the labels and the size of the rule base

Interpretability considerations: semantic criteria

Semantics: the study of meanings

- Distinguishability: Each linguistic label has semantic meaning
- Number of elements: Compatible with human capabilities
- Coverage: Any element belongs to at least one fuzzy set
- Normalization: At least one element has unitary membership
- Complementarity: For each element, the sum of memberships is one



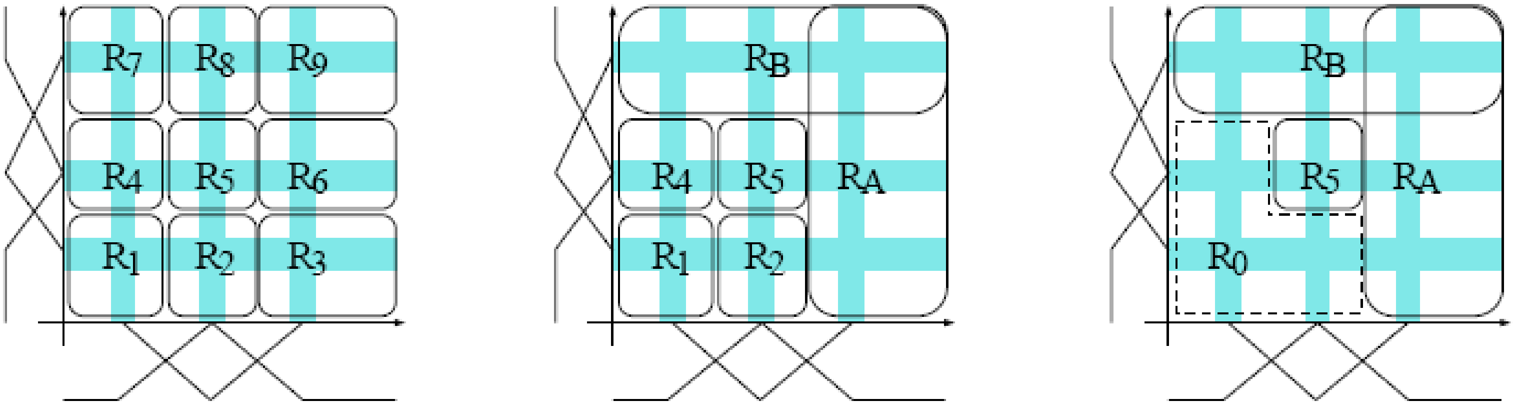
Currents Trends and Prospects

- **Interpretability quality:** associated to the meaning of the labels and the size of the rule base

Interpretability considerations: syntactic criteria

Syntax: the way in which linguistic elements are put together

- Completeness: for any input, at least one rule must fire
- Rule−base simplicity: Set of rules as small as possible
- Rule readability: small number of conditions in rule antecedents
- Consistency: rules firing simultaneously must have similar consequents

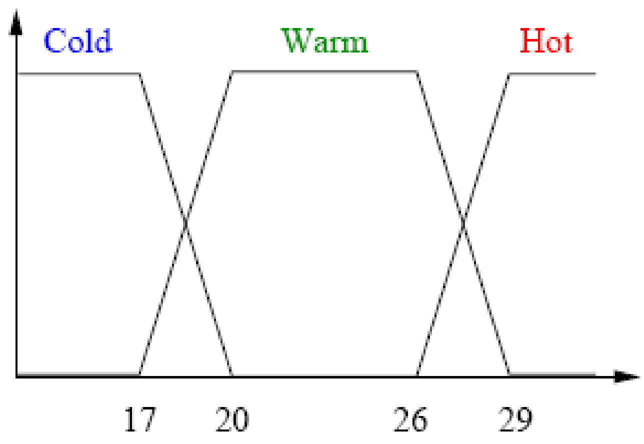


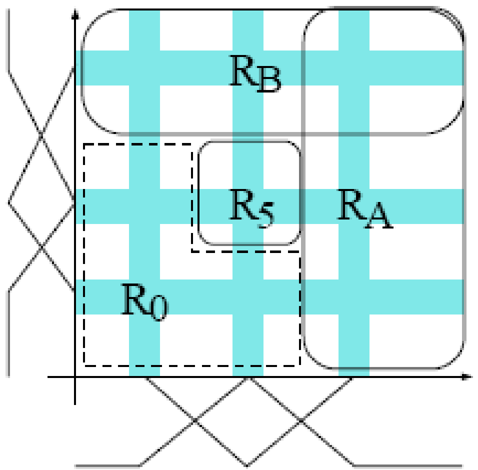
Currents Trends and Prospects

- **Interpretability quality:** associated to the meaning of the labels and the size of the rule base

Strategies to satisfy interpretability criteria

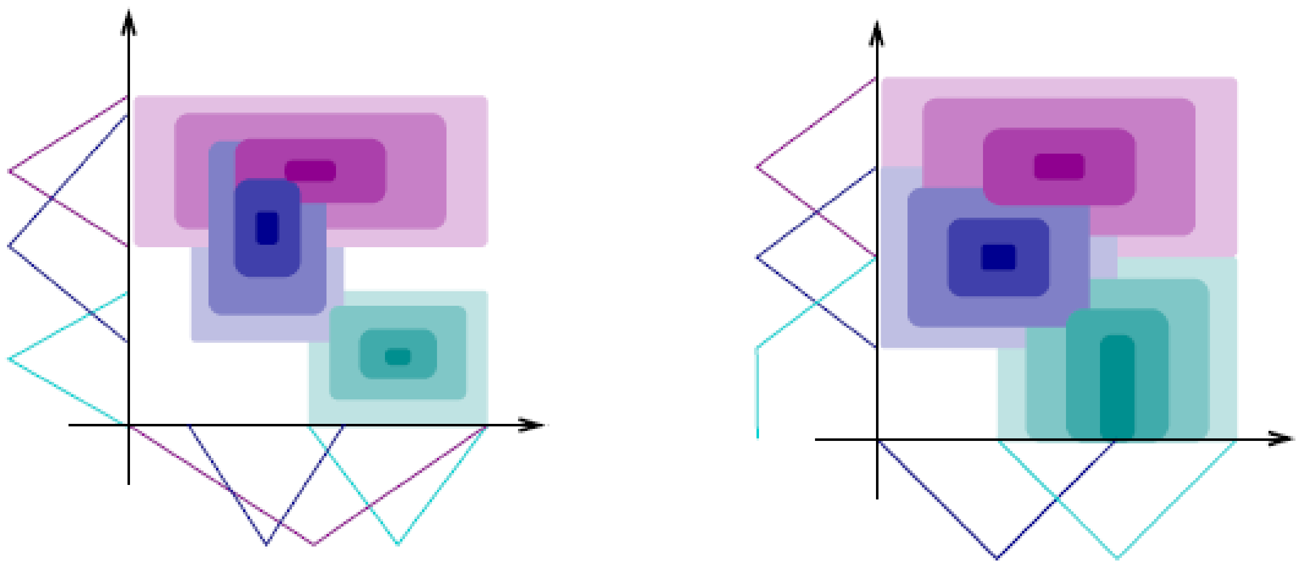
- Linguistic labels shared by all rules
- Normal, orthogonal membership functions
- Don't care conditions





Currents Trends and Prospects

- **Interpretability quality:**



What is the most interpretable rule base?

Currents Trends and Prospects

- **New data mining tasks: frequent and interesting pattern mining, mining data streams, etc**
- **Dealing with high dimensional data sets**

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

Concluding Remarks

The hybridization between fuzzy systems and GAs in GFSs became an important research area during the last decade. GAs allow us to represent different kinds of structures, such as weights, features together with rule parameters, etc., allowing us to code multiple models of knowledge representation. This provides a wide variety of approaches where it is necessary to design specific genetic components for evolving a specific representation.

Nowadays, it is a mature research area, where researchers need to reflect in order to advance towards strengths and distinctive features of the GFSs, providing useful advances in the fuzzy systems theory.

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.