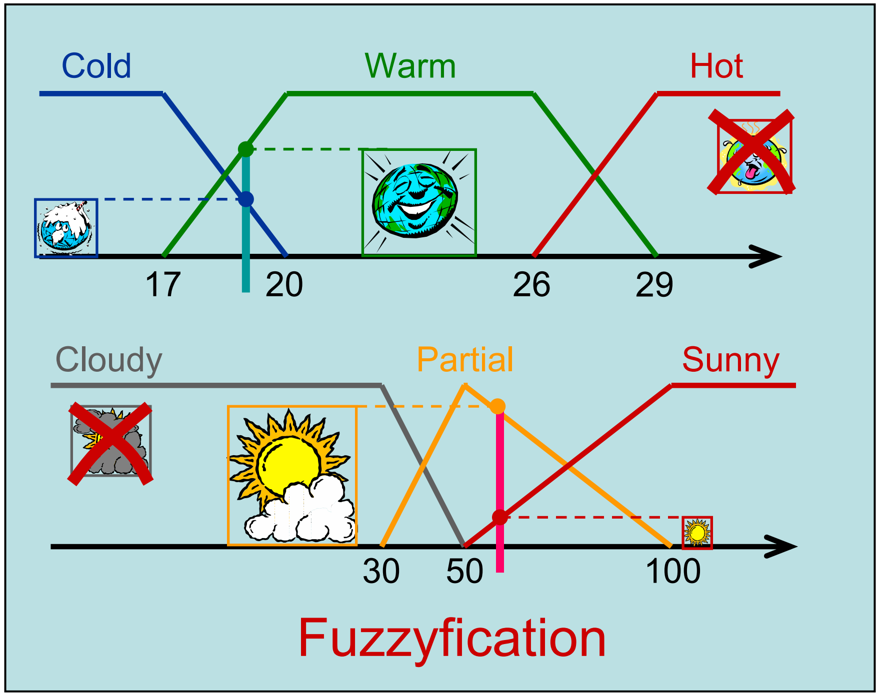
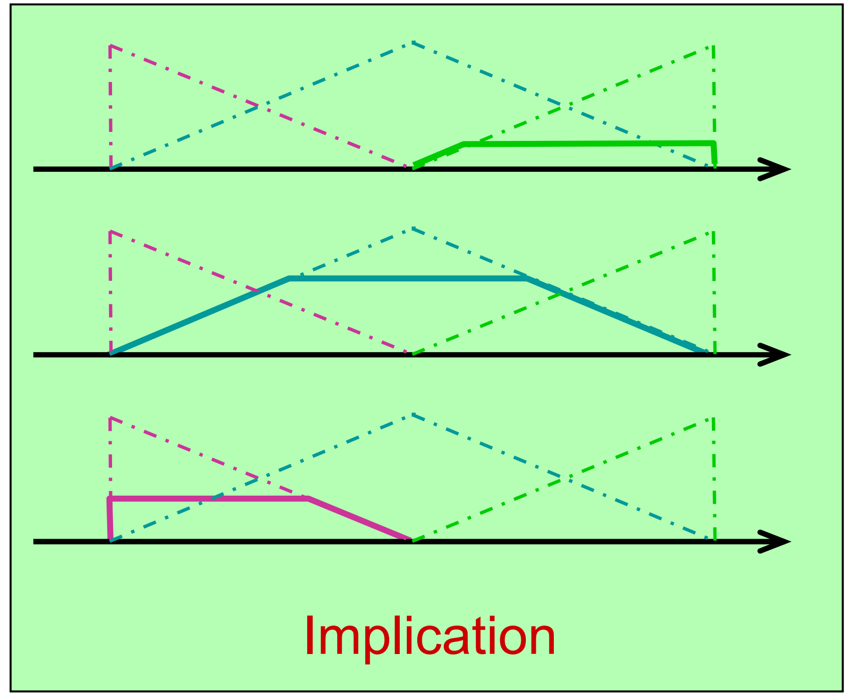
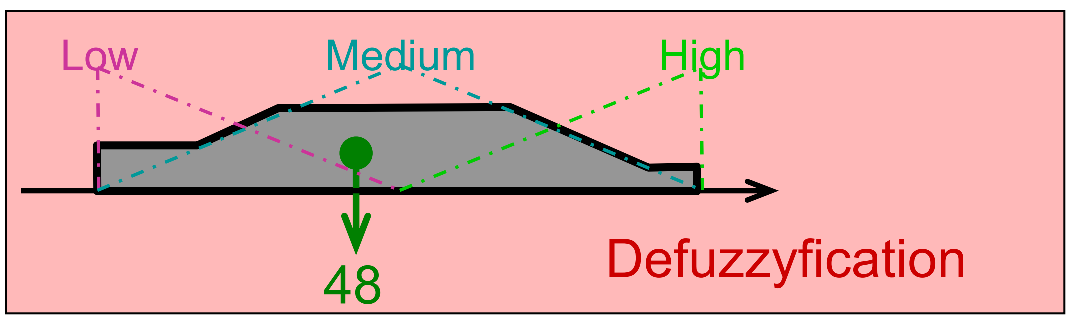
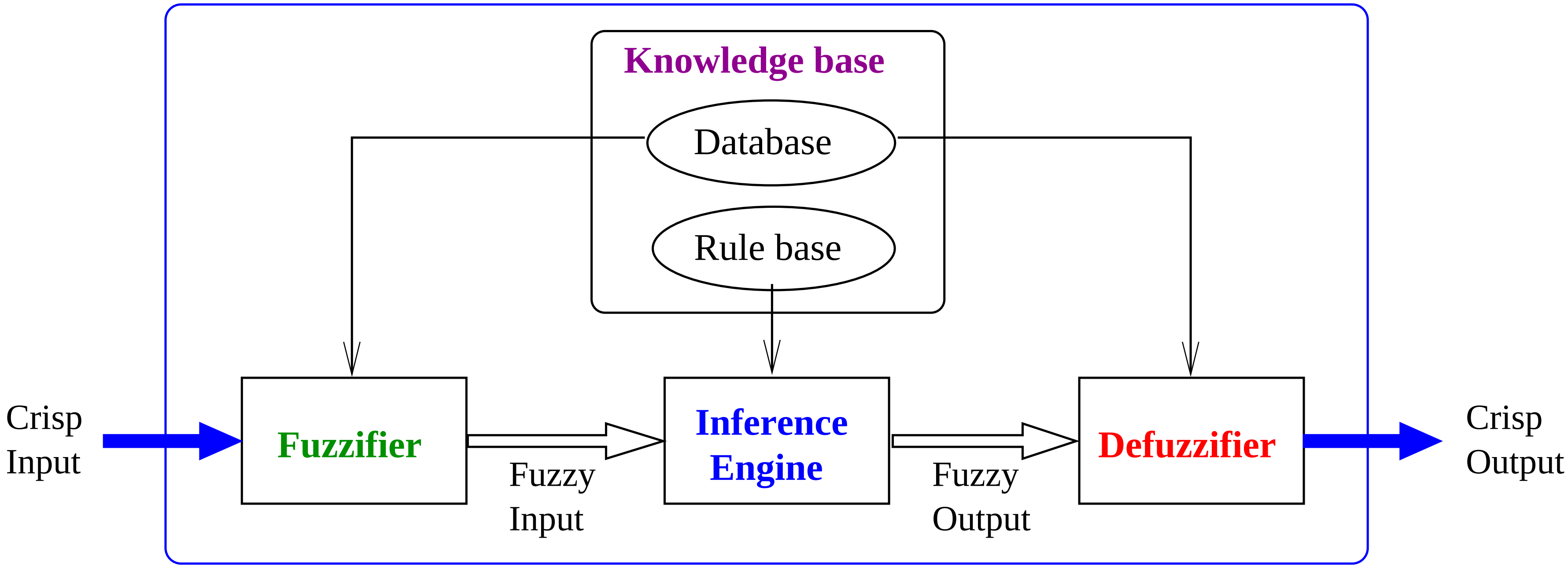
Fuzzy inference







Fuzzy inference systems



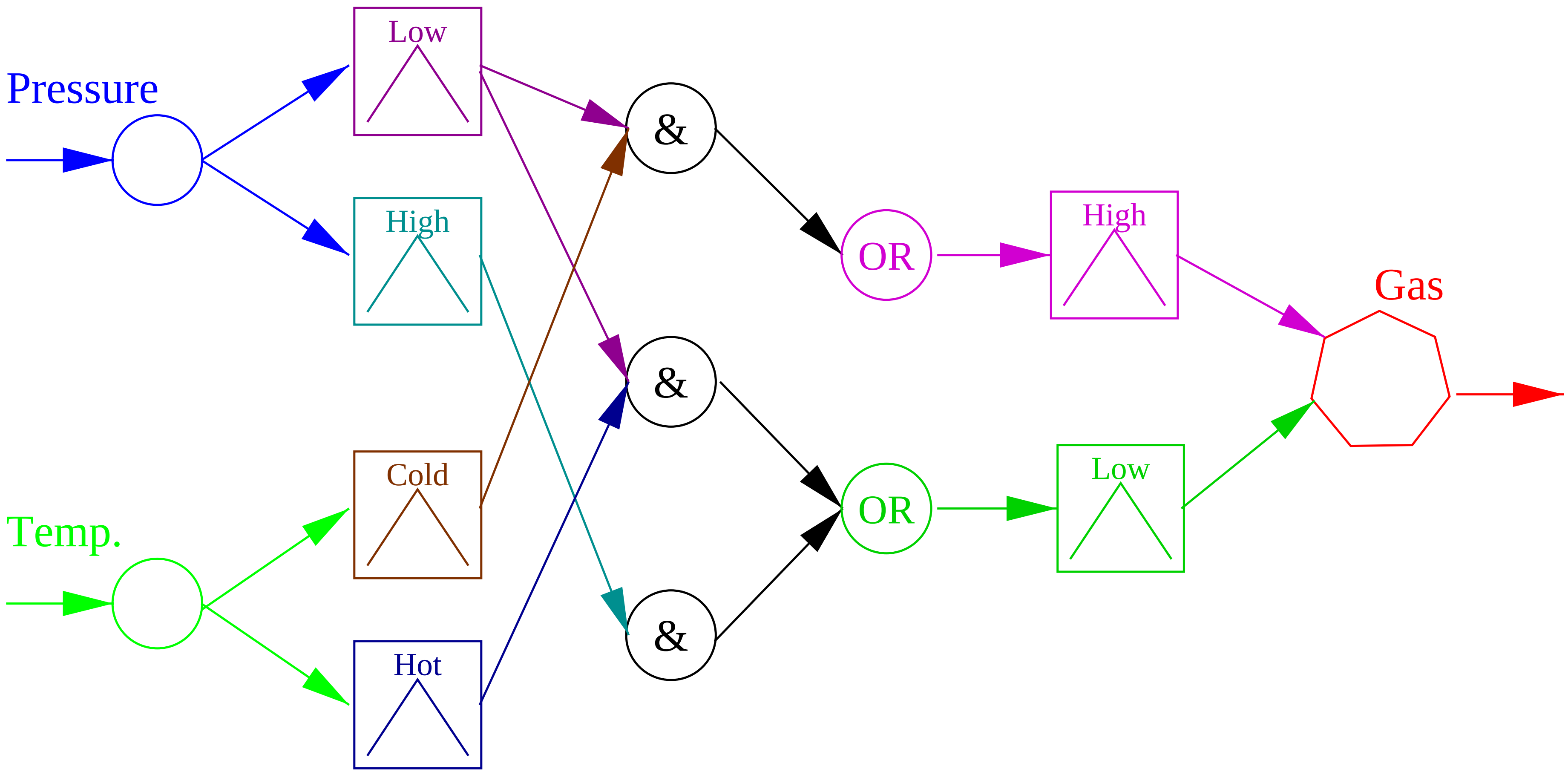
Fuzzyfier: translates crisp inputs into fuzzy values

Inference engine: applies reasoning to compute fuzzy outputs

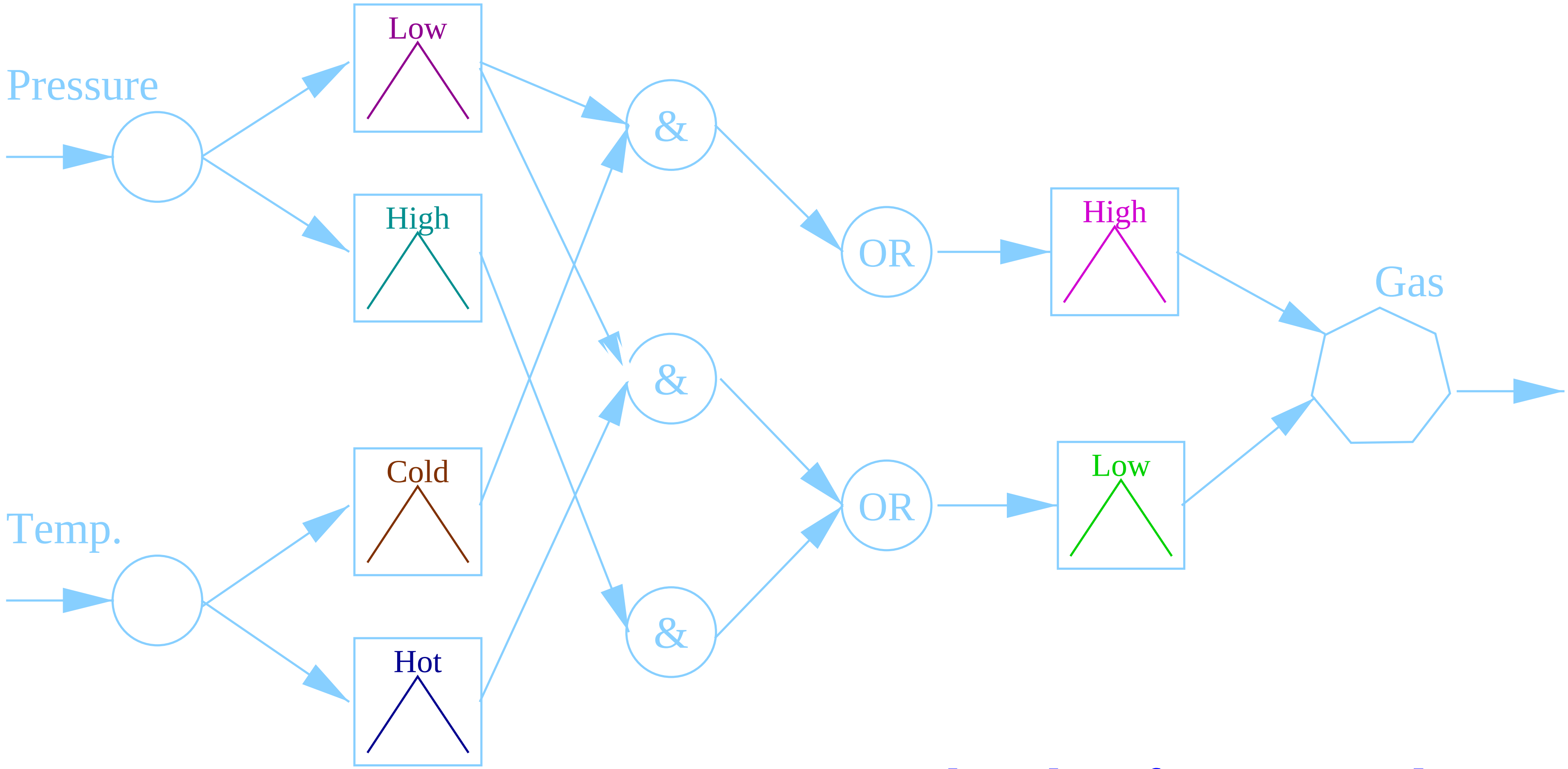
Defuzzyfier: translates fuzzy outputs into crisp values

Knowledge base: defines rules and membership functions

Network−like view of a fuzzy system

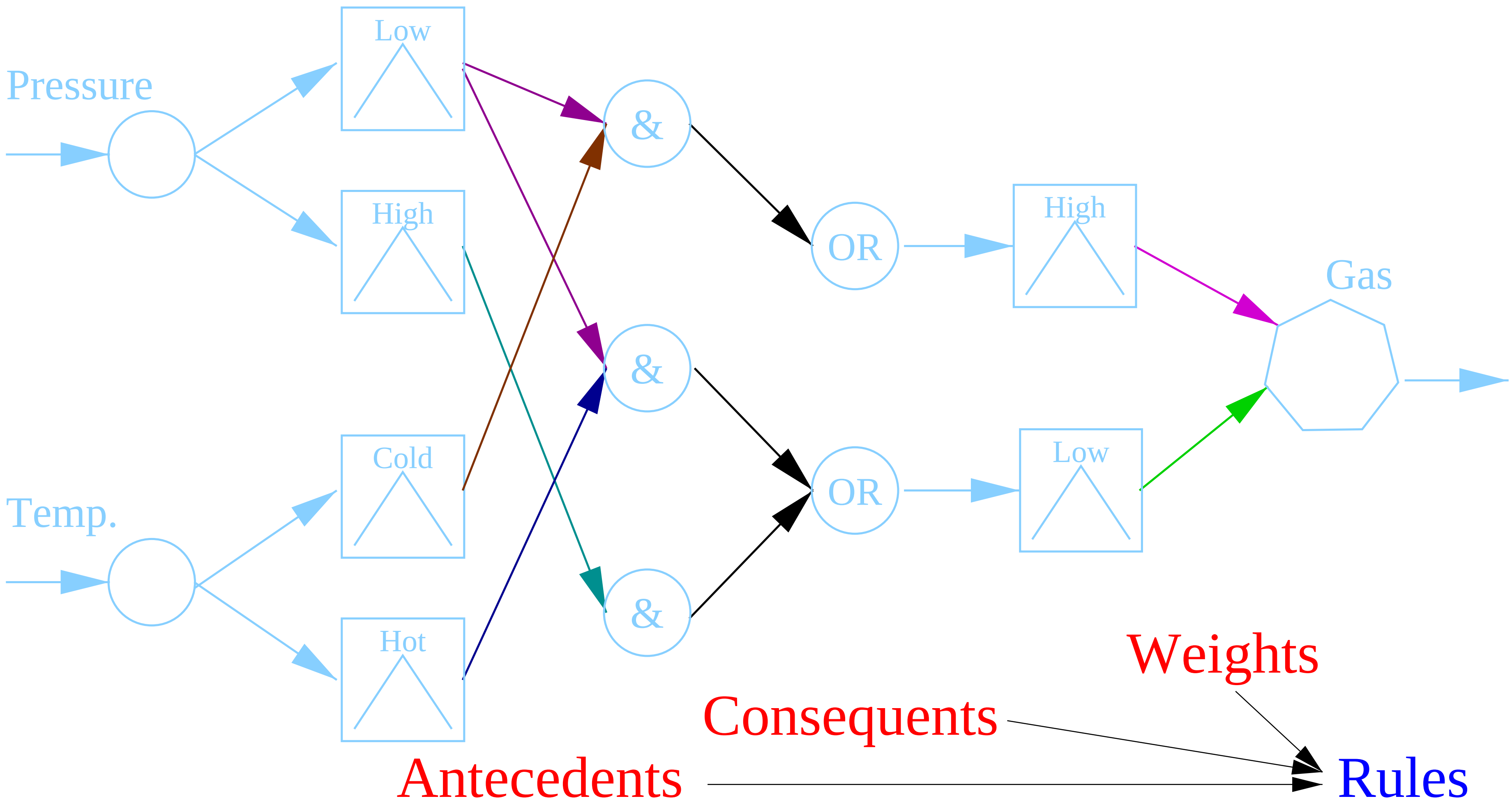


Operational parameters

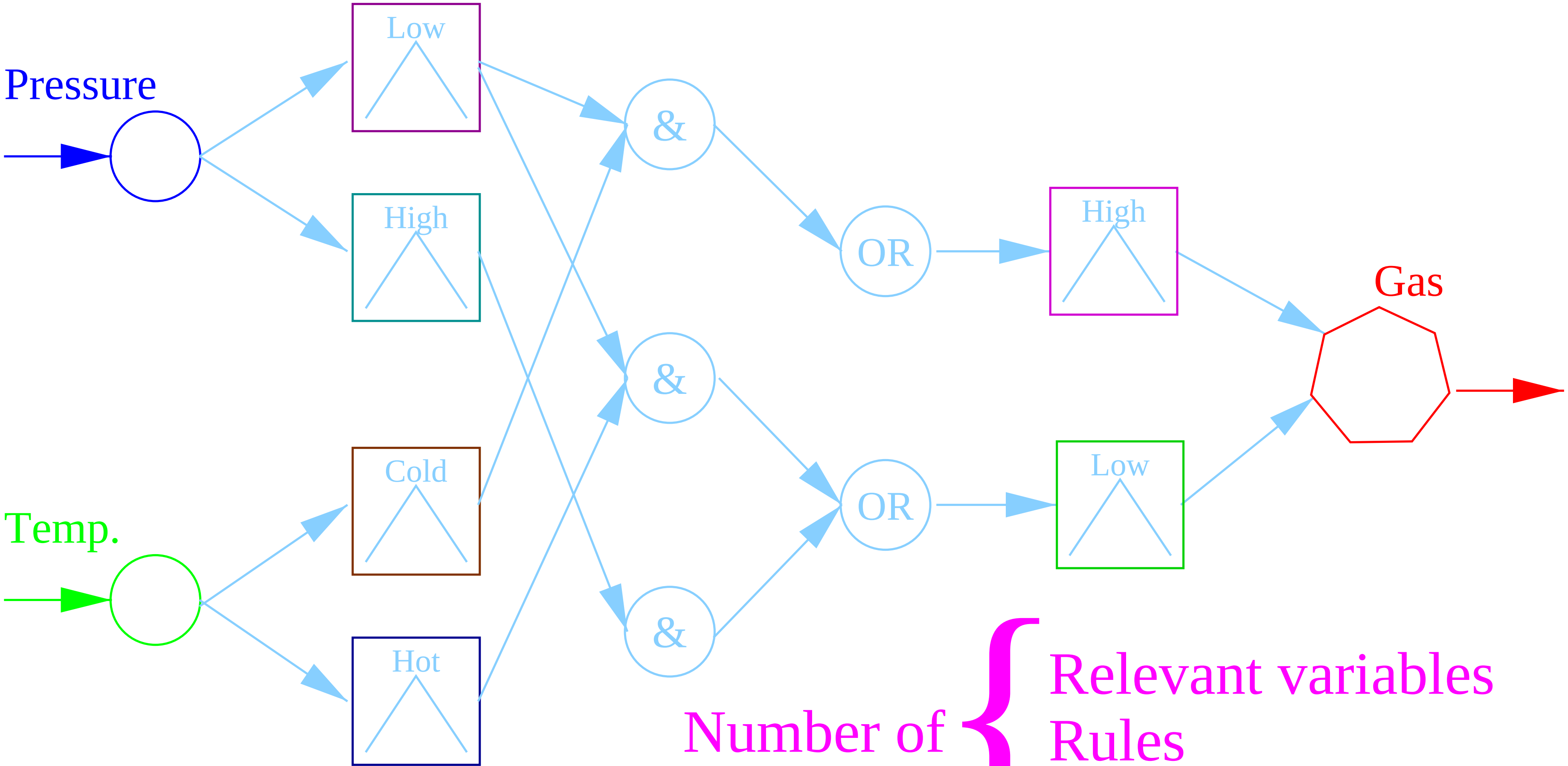


Membership function values

Connective parameters

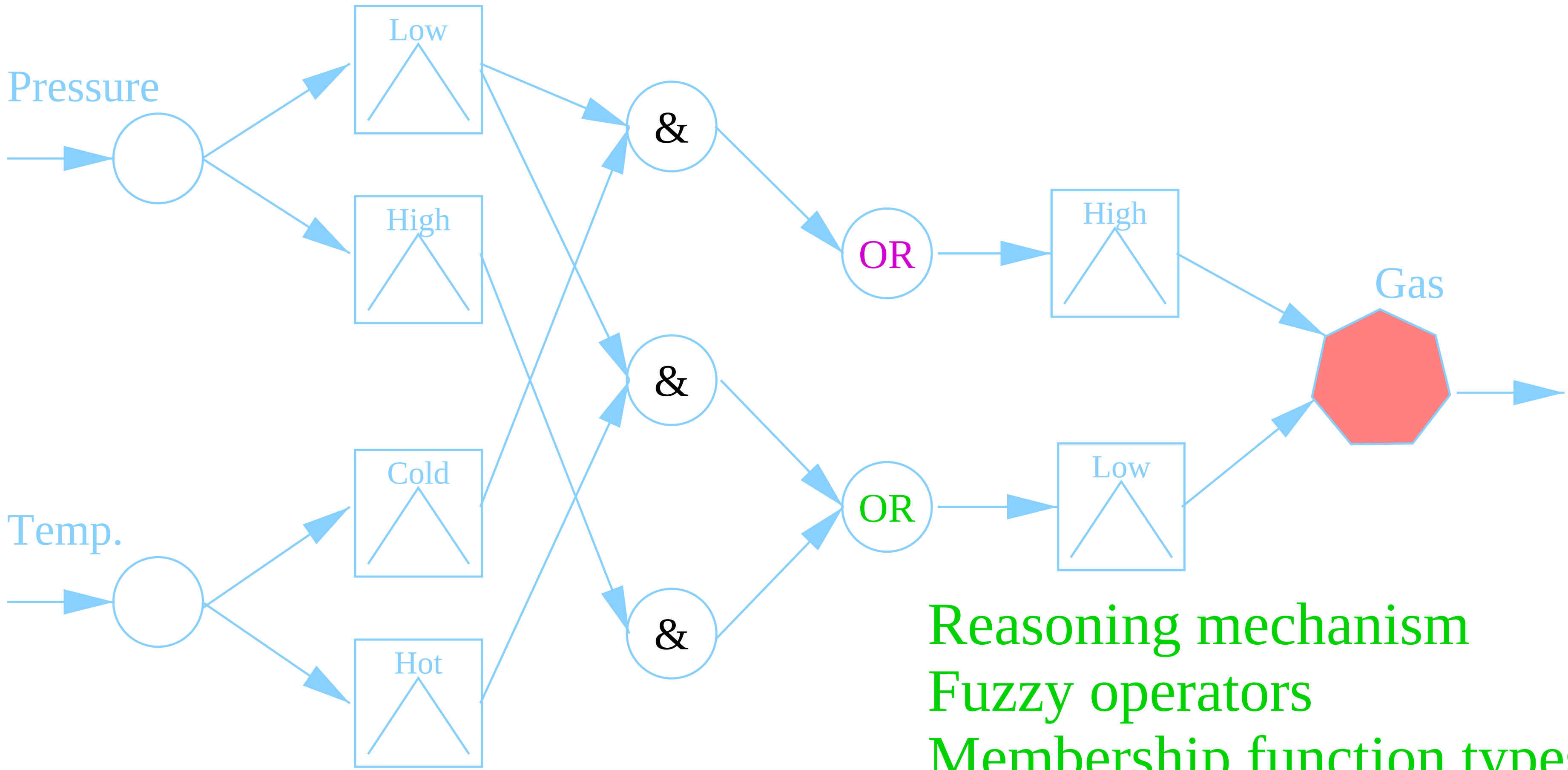


Structural parameters



Number of { Relevant variables / Rules / Membership functions

Logical parameters



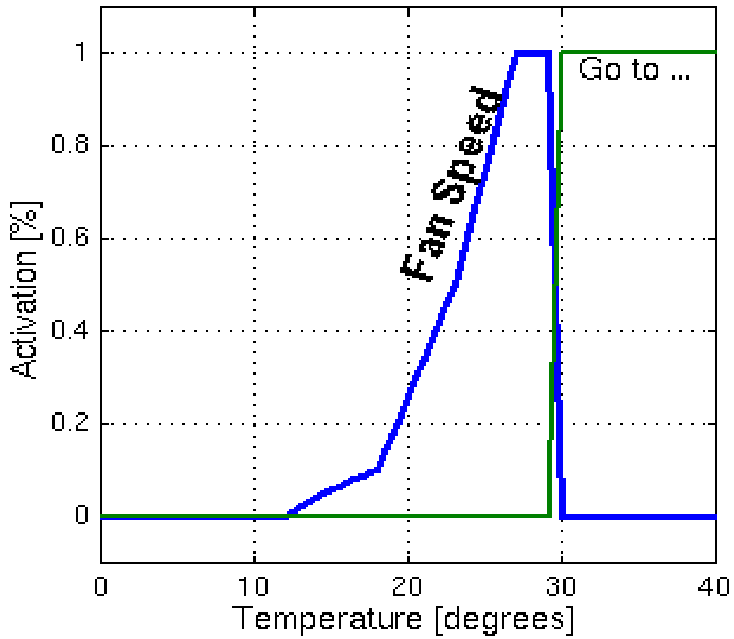
Reasoning mechanism
Fuzzy operators
Membership function types
Defuzzification method

Parameters of a fuzzy system

Class	Parameters	Component
Logic	Reasoning mechanism	Inference engine
	Fuzzy operators	
	Membership function types	Fuzzi- and defuzzifier
	Defuzzification method	Defuzzifier
Structural	Relevant variables	Knowledge base
	Number of membership functions	
	Number of rules	
Connection	Antecedents of rules	Rulebase
	Consequents of rules	
	Rule weights	
Operational	Membership function values	Database

Dual external nature

Numeric



Precision

Linguistic

If Temperature is COOL then Ventilator is Off

If Tempearture is WARM then Ventilator is Low

If Temperature is HOT then Ventilator is Medium

If Temperature is VERY−HOT then Ventilator is High

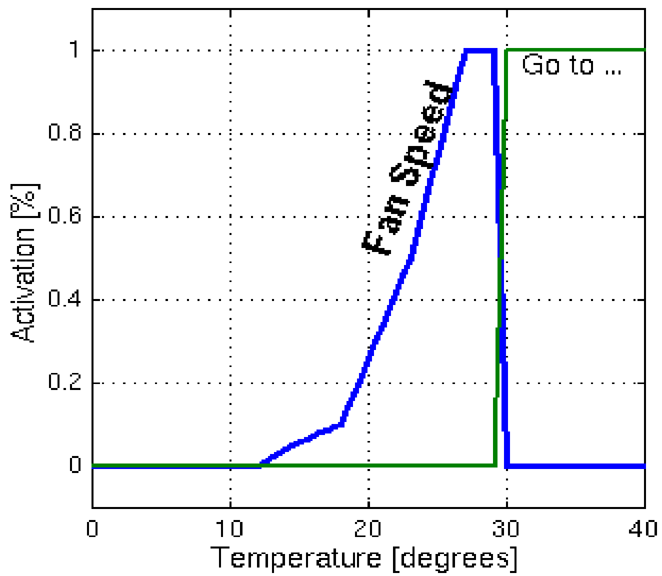
If Temperature is HELLISH then Ventilator is Off, and...

... Let's go to the lake!!!

Interpretability

Numeric issues

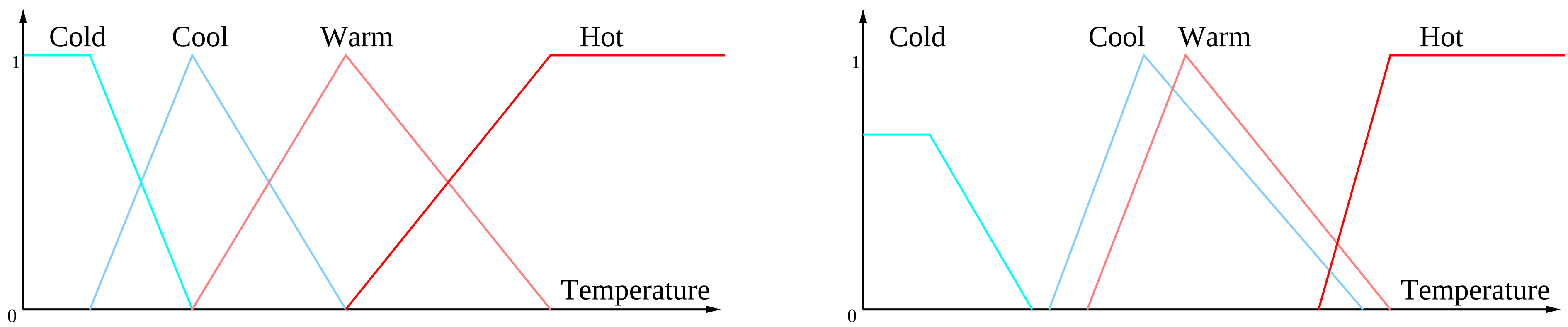
- Numeric mapping: Crisp inputs / Crisp outputs
- Nonlinear behavior, but linearity not excluded
- Universal approximator
- Uncertainty management: noise and low quality of data



Interpretability considerations: semantic criteria

Semantics: the study of meanings

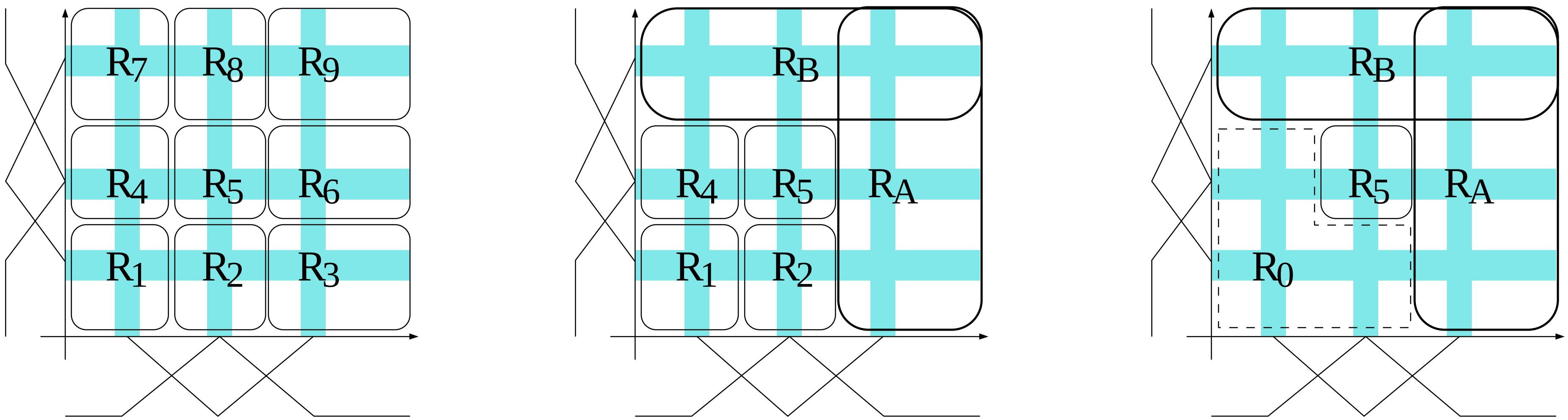
- Distinguishability: Each linguistic label has semantic meaning
- Number of elements: Compatible with human capabilities
- Coverage: Any element belongs to at least one fuzzy set
- Normalization: At least one element has unitary membership
- Complementarity: For each element, the sum of memberships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



Interpretability considerations: syntactic criteria

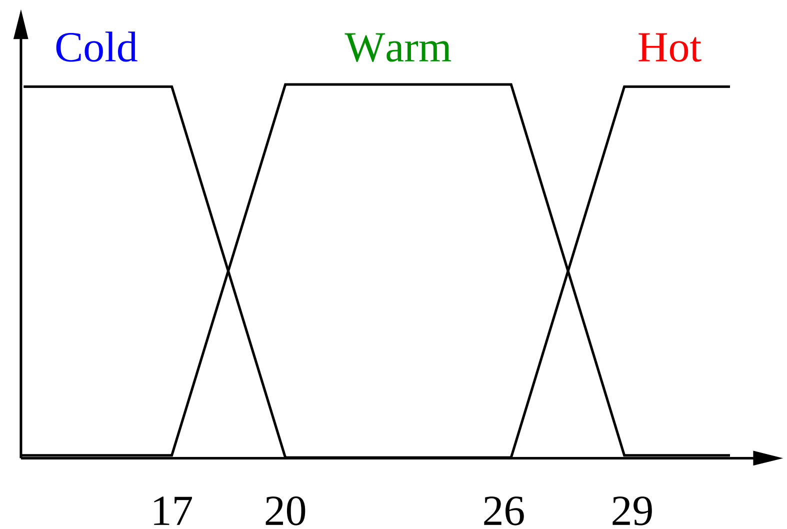
Syntax: the way in which linguistic elements are put together

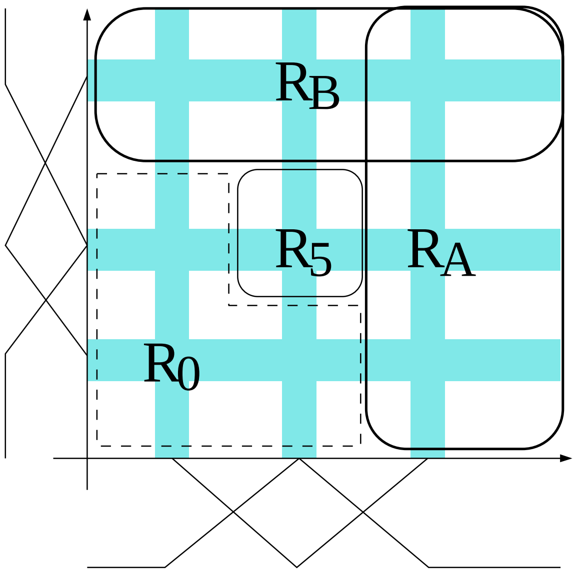
- Completeness: for any input, at least one rule must fire
- Rule−base simplicity: Set of rules as small as possible
- Rule readability: small number of conditions in rule antecedents
- Consistency: rules firing simultaneously must have similar consequents



Strategies to satisfy interpretability criteria

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





The general modeling problem

What do you know about the modeled system?

i.e. what is predefined and what looked for?

Search space

How do you search?

i.e. do a well suited and/or well known technique exists?

Search method

Have you preferences or restrictions to the model?

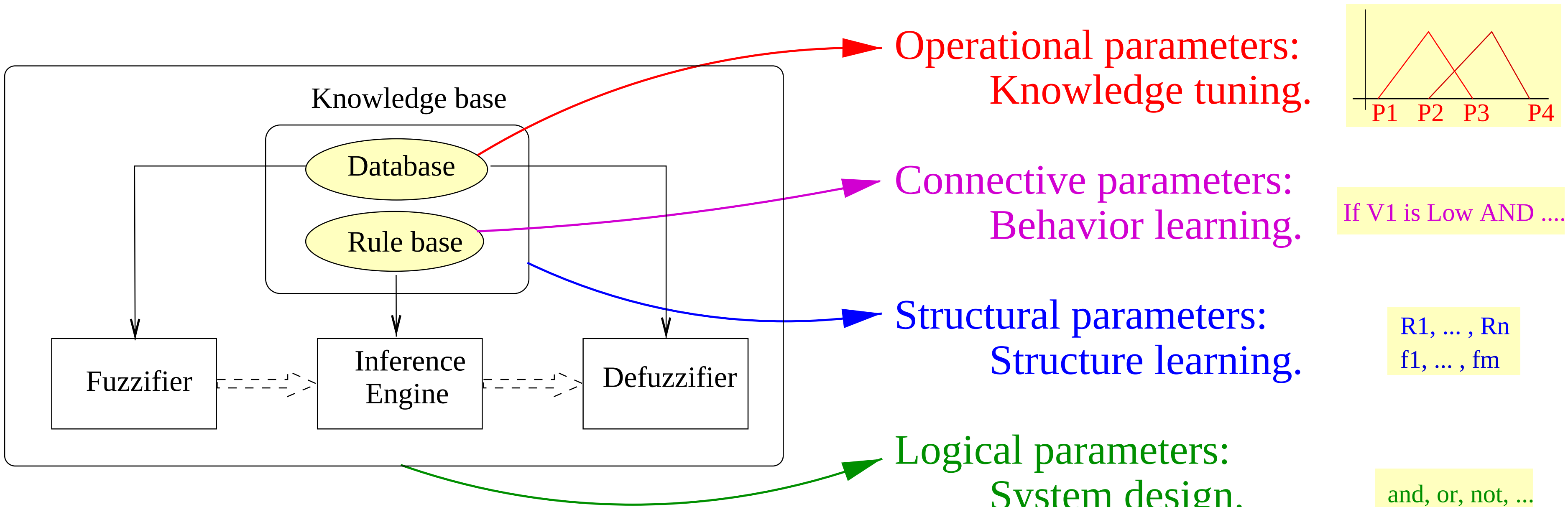
i.e. do issues like size, speed, or simplicity matter?

Constraints

Search space in fuzzy modeling

The number of parameters is too high to perform a full search, some parameter pre−definition is thus required.

According with the searched parameters we can have:



Search methods: Fuzzy modeling techniques

Knowledge engineering

"Classic" identification methods

Machine learning approaches

Neuro−fuzzy systems

Evolutionary fuzzy modeling techniques

Usual constraints in fuzzy modeling

What is the fuzzy system expected to do?

- Classification: Classification performance, quadratic error.
- Control: Dynamic response, adaptability, robustness, etc.
- Diagnostic: Overall performance, sensitivity, specificity
- Data mining: Completness, complexity.

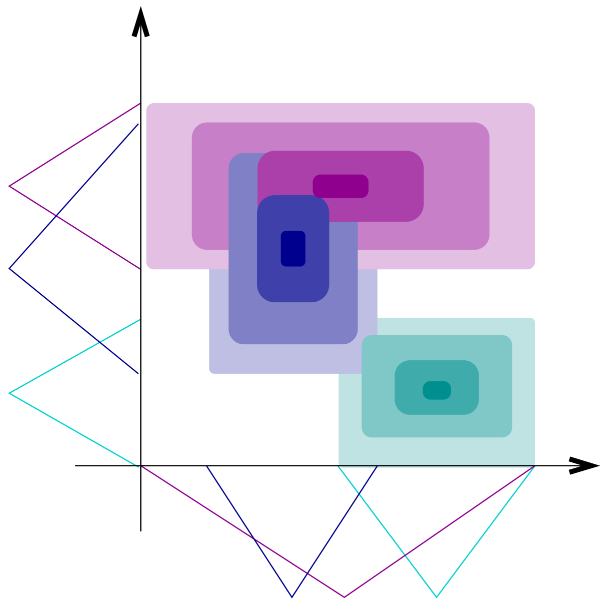
How is the system expected to do it?

- Speed: Real−time constraints, computing resources.
- Size: Available memory, computing platform.

Who is going to interact with the system?

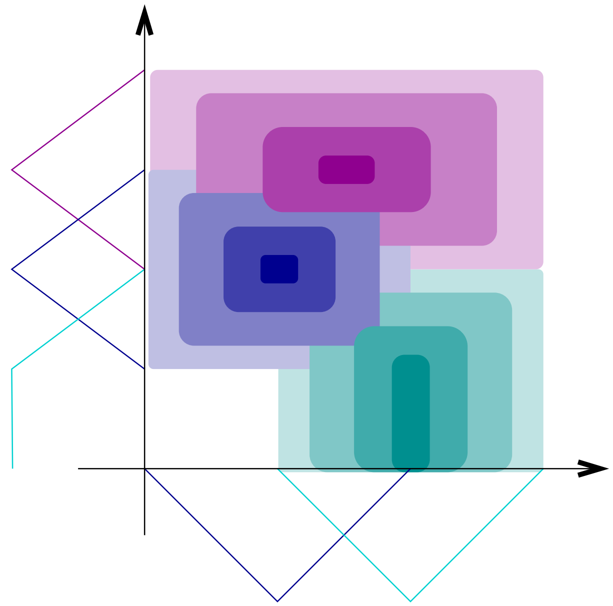
- Interpretability: Allowed complexity.
- Availability: Continuity of explanations (time to provide them)

Interpretability−related constraints



Rule−specific MFs are not allowed

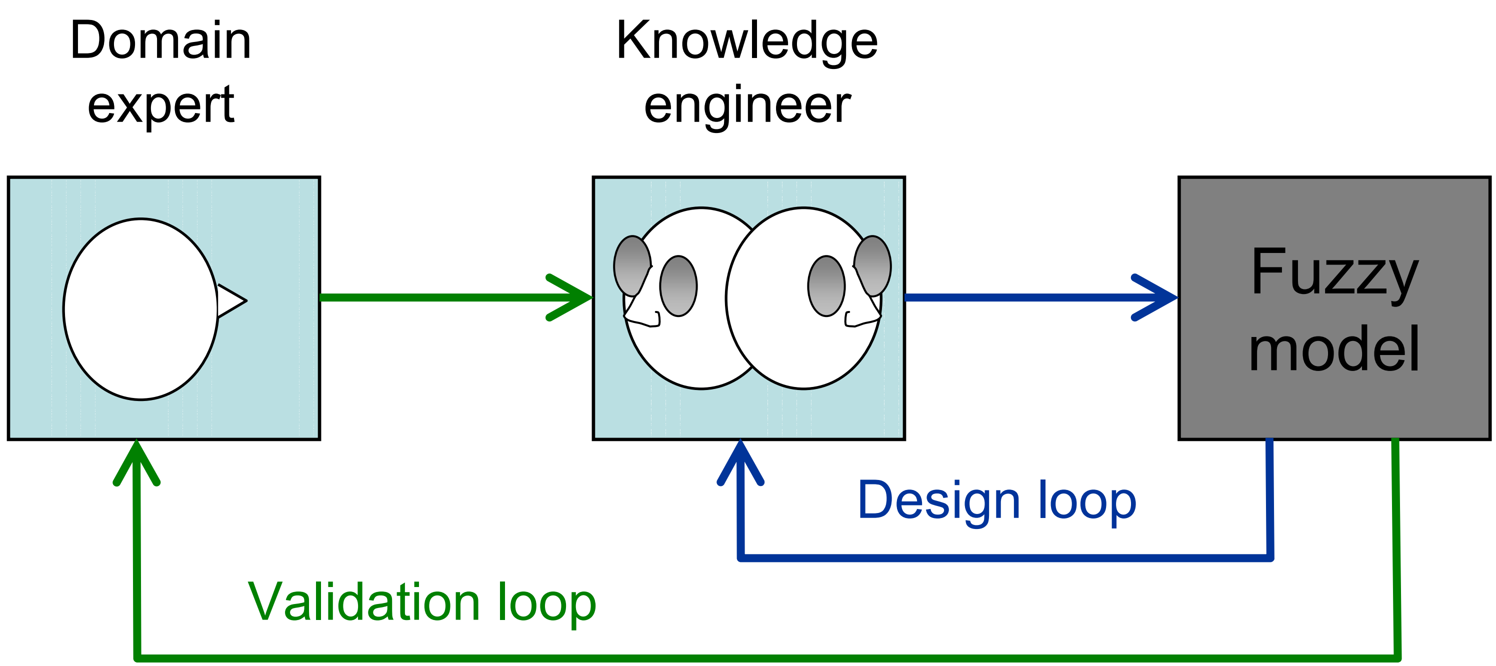
all rules share the same MFs



Orthogonal MFs with well defined

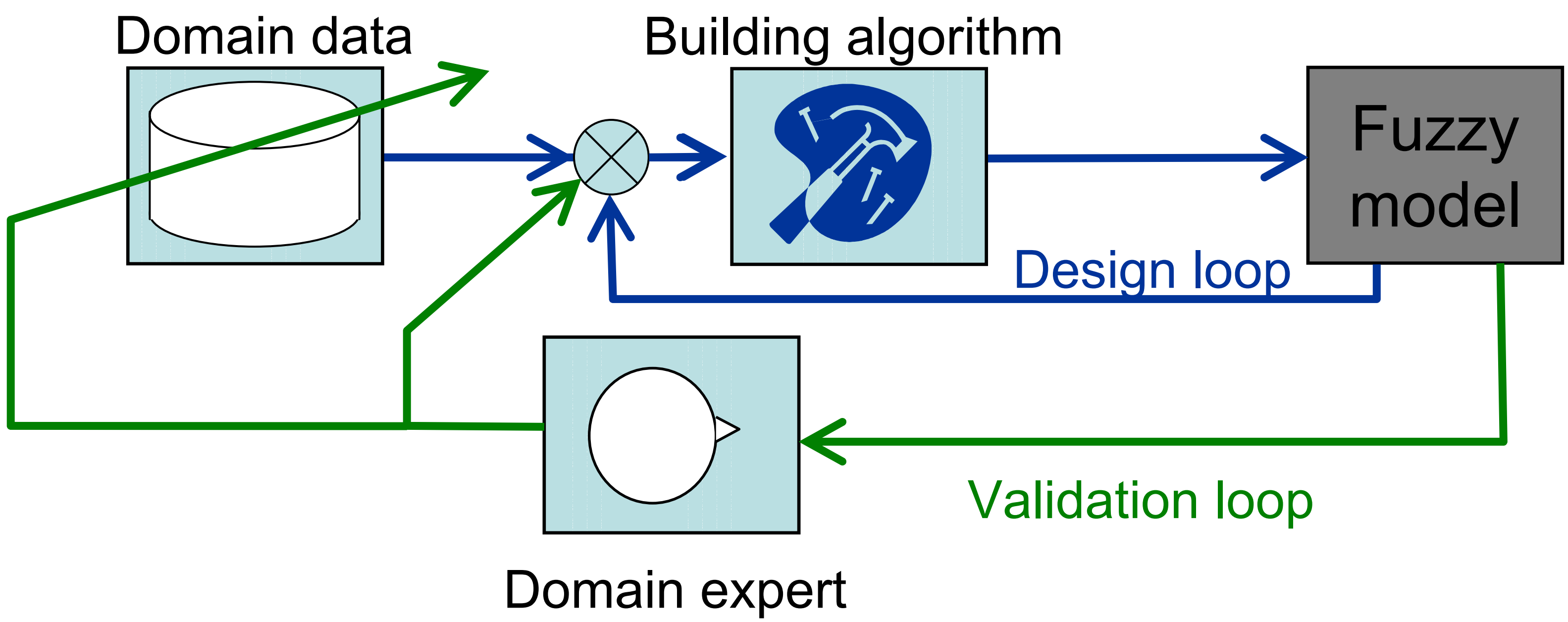
null and unity membership

Fuzzy modeling: direct approach



This approach is also called *knowledge engineering*

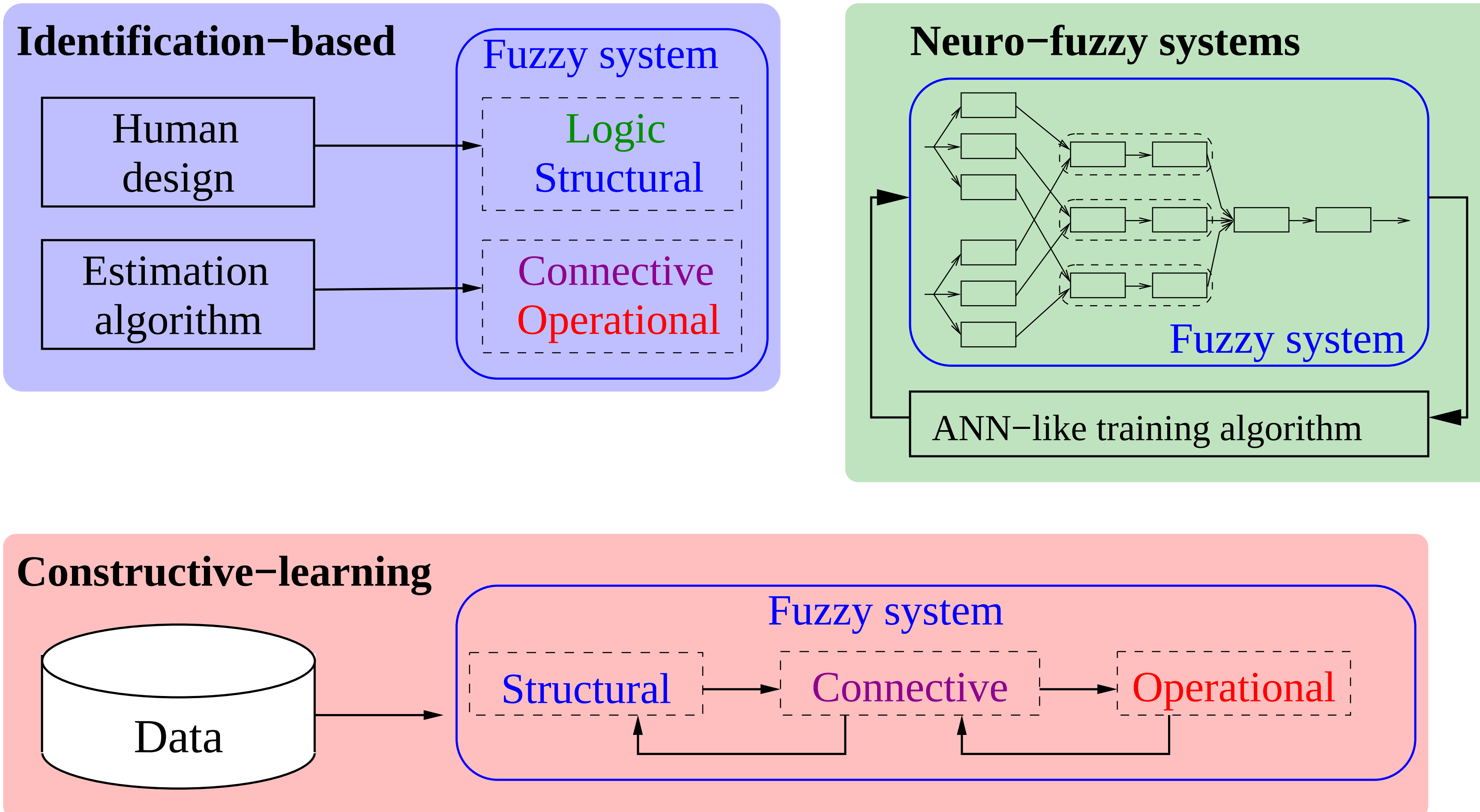
Fuzzy modeling: data-driven approaches



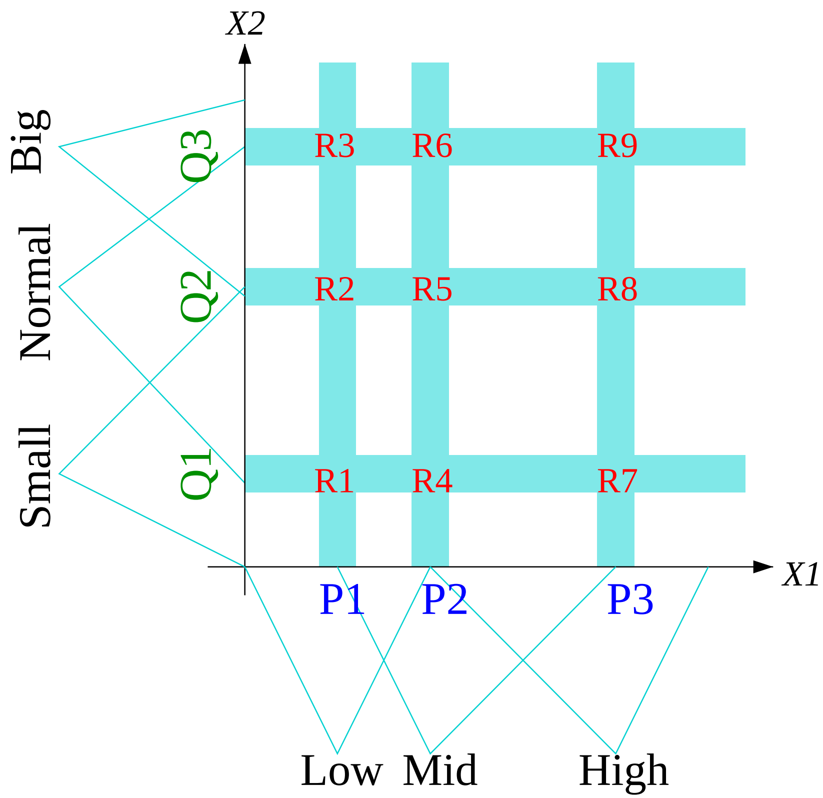
This approaches are also denominated

knowledge discovering

Fuzzy modeling: some data−driven approaches



Evolutionary knowledge tuning (database)



- Fixed rule base (completness)

- Rules of type:

  if X1=Low and X2=Normal then Output = Ci

- Knowledge is tuned by evolution,

  which searches for membership function values

- Genome encodes values P, Q, and C

  (3*P + 3*Q + 9*C) * 5 bits = 75 bits

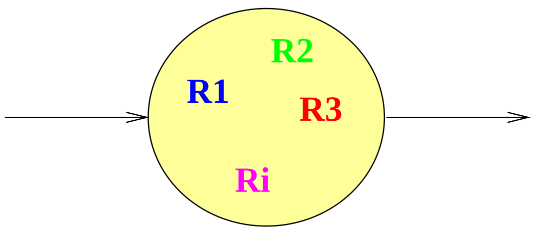
Evolutionary behavior learning (rule base)

- Number of rules exploses rapidly
- Two strategies for reducing this number:
  Don’t care conditions and default rule
- Evolution can be used to find a minimal (or fixed size) rule base,
- Three main approches to evolutionary behavior learning

Michigan

Individual = One rule

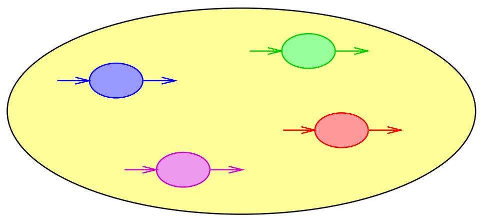
Population = Rule base
(i.e. fuzzy system)



Pittsburgh

Individual = Entire system
(rule base or knowledge base)

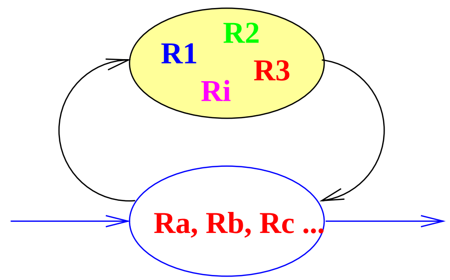
Population of systems



Iterative Rule Learning

Evolution finds the best rule

Incremental construction of
the knowledge base



Carlos Andrés Peña-Reyes

Logic Systems Laboratory - Swiss Federal Institute of Technology Lausanne

Evolutionary behavior learning: An example

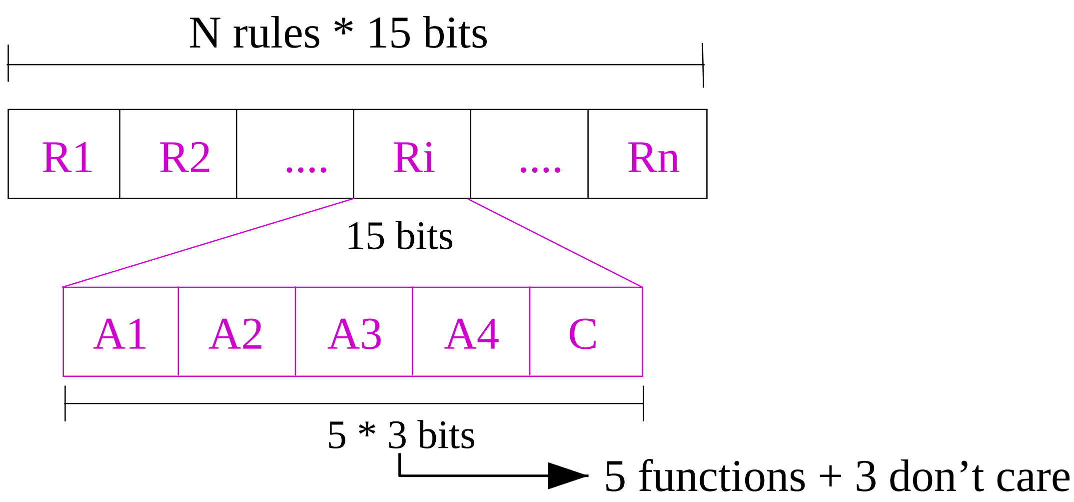
- 4 input variables, 5 membership functions per variable

  {Tiny, Small, Normal, Big, Huge}

- Space of 625 rules (1295 including don’t care conditions)

  IF V1 is Tiny AND ... AND V4 is Normal then Out = Huge

- Evolution searches for a subset of N rules (fixed by the designer)
- Genome encodes rules: Antecedents and consecuents



Evolutionary knowledge base learning

(Knowledge base = rule base + database)

Parameter class	Modeling type	Usual quantity	Type of values	Fuzzy system attribute
Connective	Behavior learning	10 - 1000	Symbolic	Rule base
Operational	Knowledge tuning	10 - 1000	Real-valued	Database

Critical issues for applying evolution:

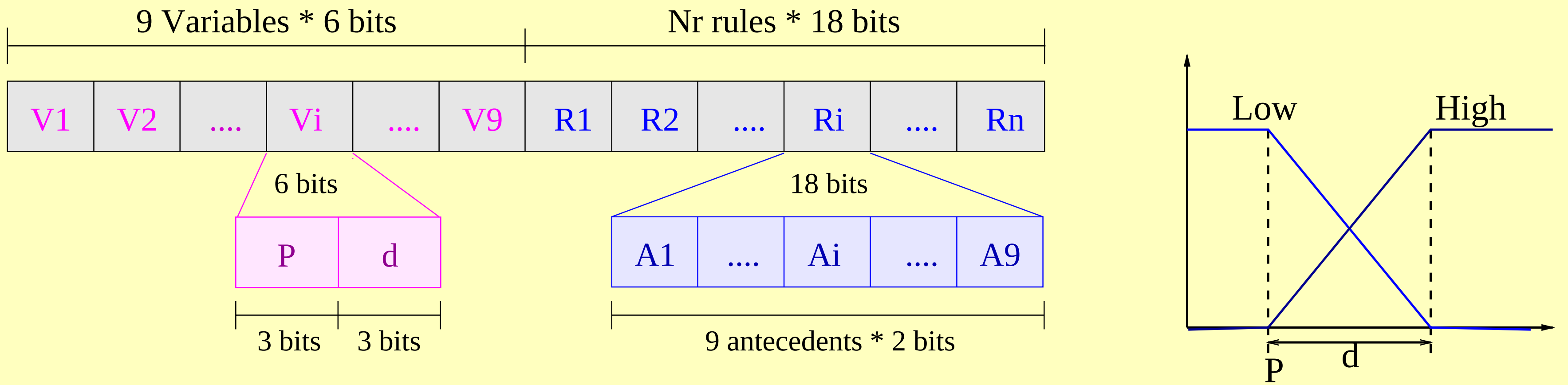
- Parameter representation
- Tight interdependency
- Size of the search space
- Computation requirements

Evolutionary knowledge base learning

- A basic approach: Single population, single evolution

Example: Breast cancer diagnosis problem (Peña and Sipper 99)

- 9 inputs, 1 output, 2 membership functions per variable
- A simple genetic algorithm searches for the knowledge base
- Genome encodes: Rule antecedents and membership function parameters



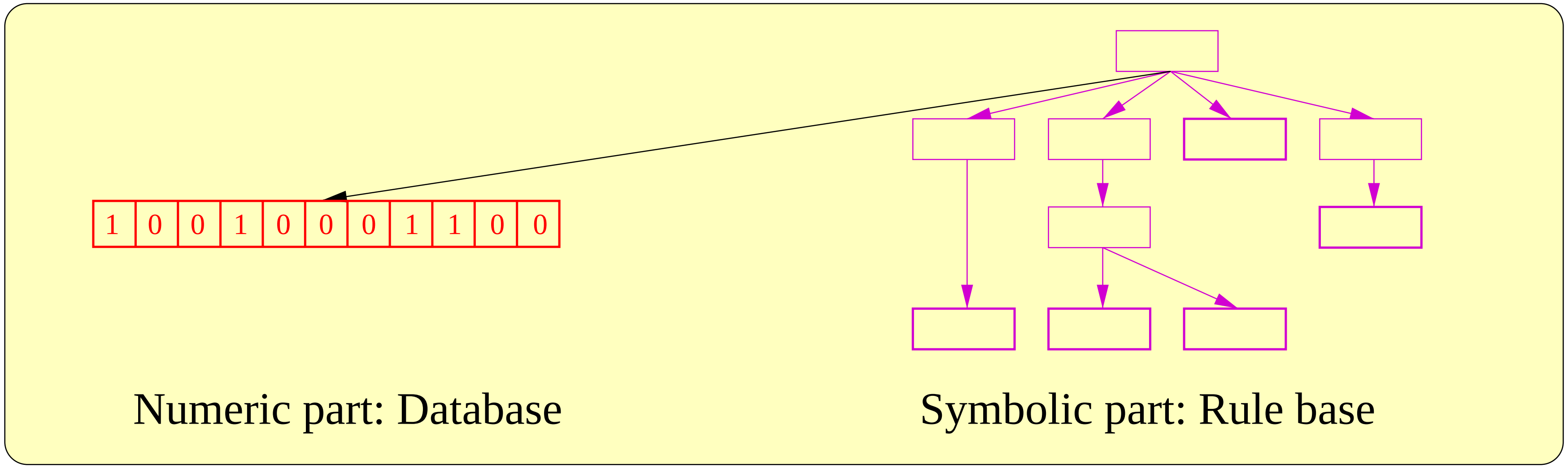
Sample ru;e: IF V1 is Low and V4 is High THEN Diagnostic is Benign

Evolutionary knowledge base learning

- A variation: Single population, double evolution

Example: Evolving fuzzy rule based classifiers with GA–P (García et al. 99)

- Genome encodes: Complete rule base and membership function parameters
- A simple genetic algorithm searches for the database
- The rule base is evolved using genetic programming

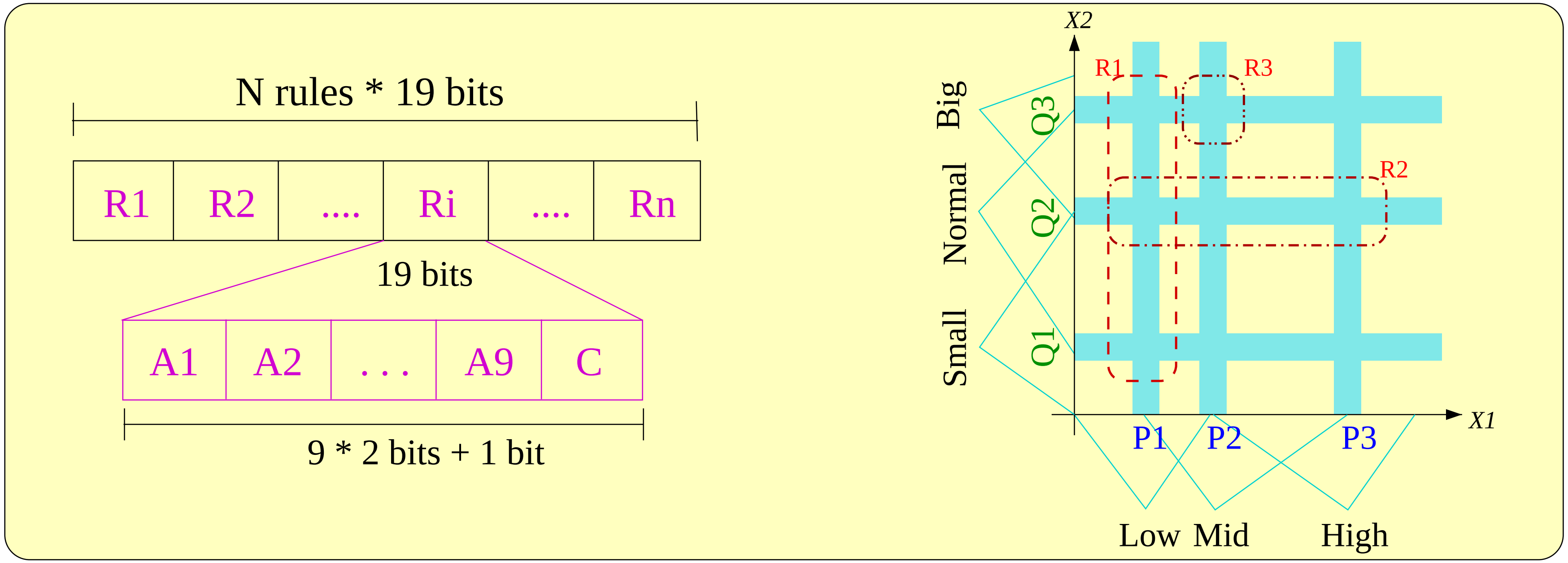


Evolutionary knowledge base learning

- Hybrid learning: Evolved rule base, learned database

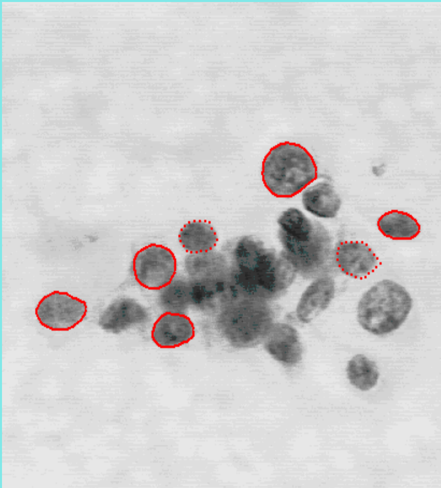
Example: Breast cancer diagnosis (J.-F. Philagor, student project SPG, 1999)

- Evolution searches for a fixed-size rule base
  - Genome encodes rules: Antecedents and consequent
- Database is tuned using a neuro-fuzzy approach
  - A fuzzy self-organizing map searches for P and Q values



The Wisconsin Breast Cancer Database

The test



The features

1. Clump Thickness (v_1)
2. Uniformity of Cell Size (v_2)
3. Uniformity of Cell Shape (v_3)
4. Marginal Adhesion (v_4)
5. Single Epithelial Cell Size (v_5)
6. Bare Nuclei (v_6)
7. Bland Chromatin (v_7)
8. Normal Nucleoli (v_8)
9. Mitosis (v_9)

The database

case	v_1	v_2	v_3	...	v_9	*diagnostic*
1	5	1	1	...	1	*Benign*
2	5	4	4	...	1	*Benign*
⋮	⋮	⋮	⋮		⋮	⋮
683	4	8	8	...	1	*Malignant*

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Carlos Andrés Peña-Reyes

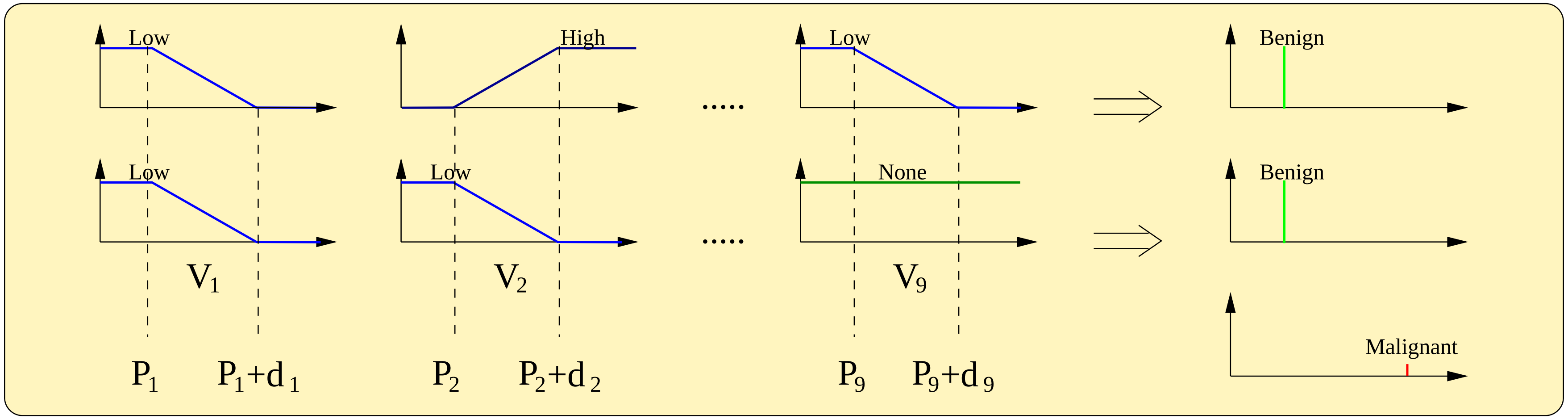
Logic Systems Laboratory - Swiss Federal Institute of Technology Lausanne

Logic Systems Laboratory

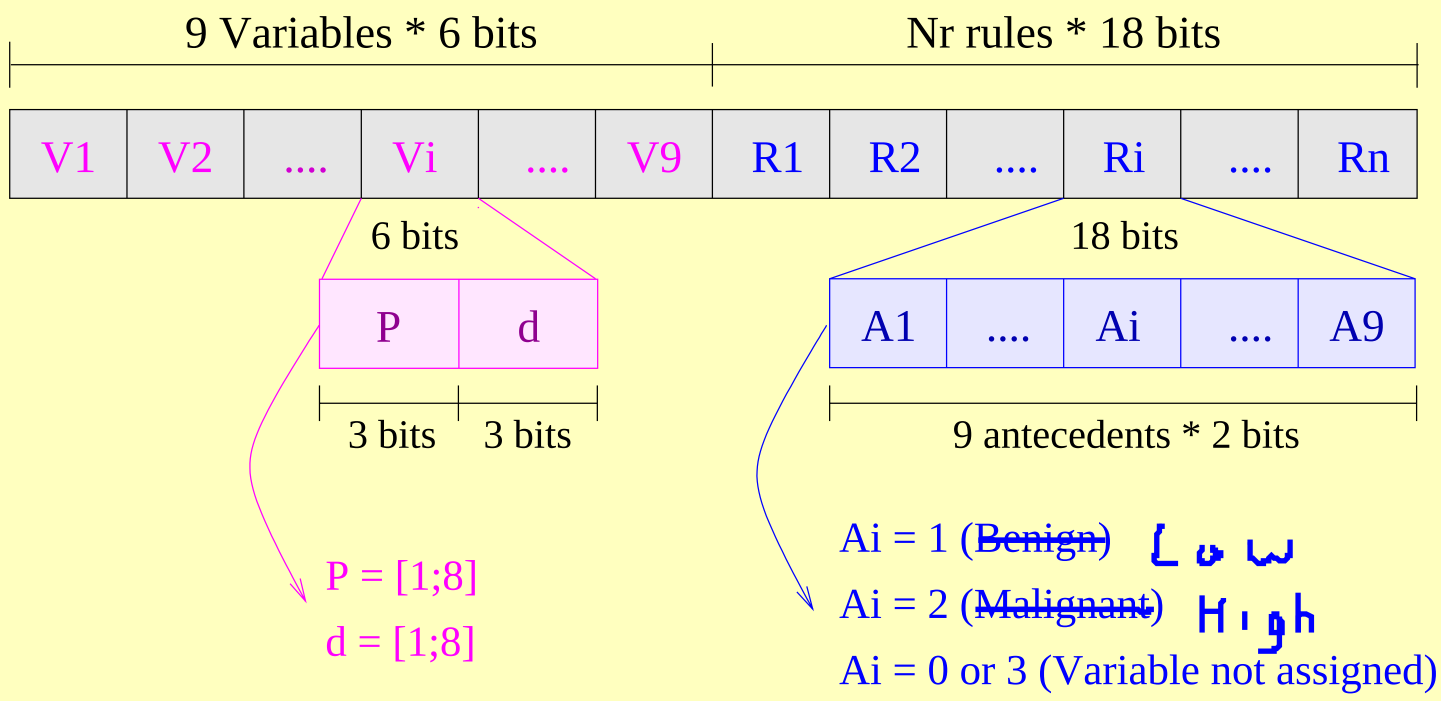
Laboratoire de Systèmes Logiques

Proposed Fuzzy System

R1: if (V1 is Low) and (V2 is High) and ... and (V9 is Low) then (output is Benign)
R2: if (V1 is Low) and (V2 is Low) and ... and (V9 is None) then (output is Benign)
⋮ ⋮ ⋮ ⋮ ⋮
else (Output is Malignant)



Genome encoding



Total genome length = 54 + 18Nr bits

Fitness function

F = Fc + a* Fv + b*Fe

Fc : Classification performance,
the most important performance measure

Fv : Number of variables
measures the interpretability

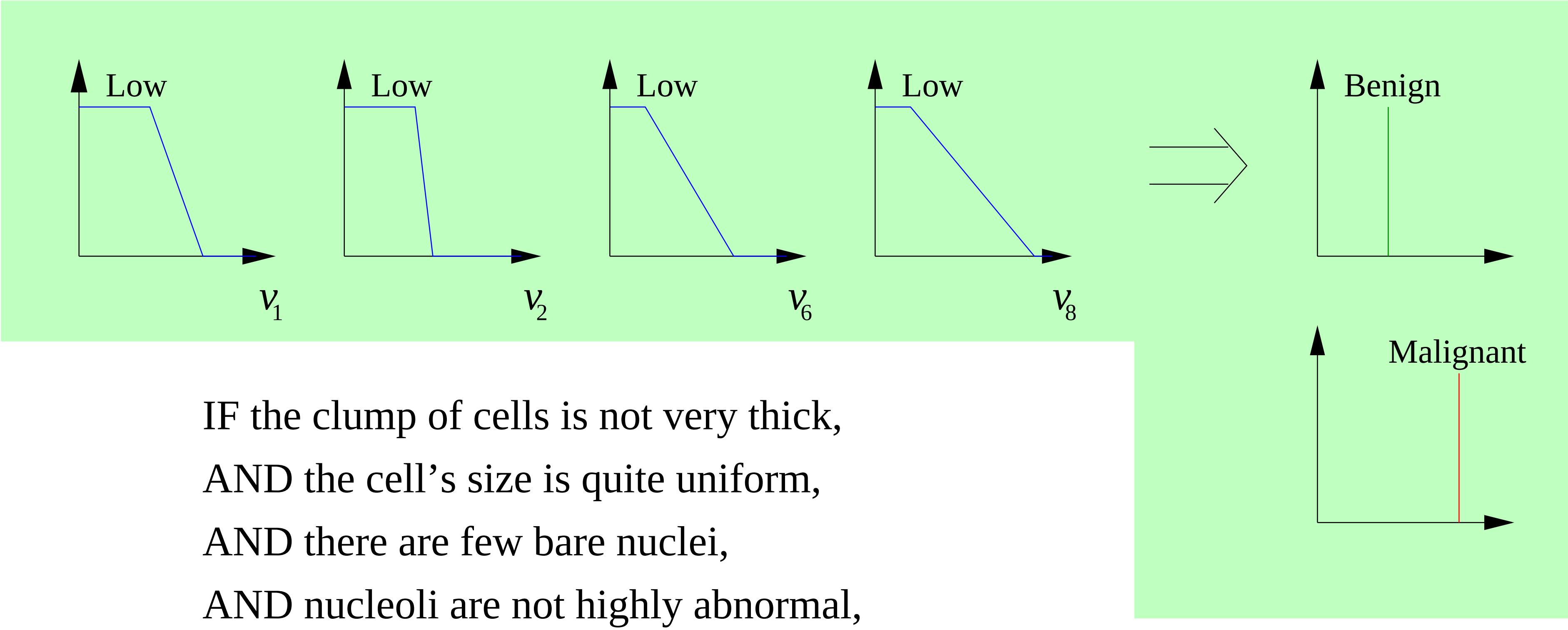
Fe : Quadratic error
selection pressure to fine tune parameters

Results: Classification performance (Number of variables)

Rules	Setiono (96)	Setiono Liu (96)	Taha Ghosh (97)	Peña Sipper (98)	This work (99)
1	95.42% (2)			96.35% (3)	**97.07% (4)**
2				96.65% (7)	97.36% (3)
3	97.14% (4)	97.21% (4)			**97.80% (4.7)**
4					97.80% (4.8)
5			96.19% (1.8)		97.51% (3.4)

$\underbrace{\text{Setiono (96), Setiono Liu (96), Taha Ghosh (97)}}_{\text{Learned Boolean rules}}$ $\underbrace{\text{Peña Sipper (98), This work (99)}}_{\text{Evolved fuzzy rules}}$

The best single-rule system



IF the clump of cells is not very thick,

AND the cell's size is quite uniform,

AND there are few bare nuclei,

AND nucleoli are not highly abnormal,

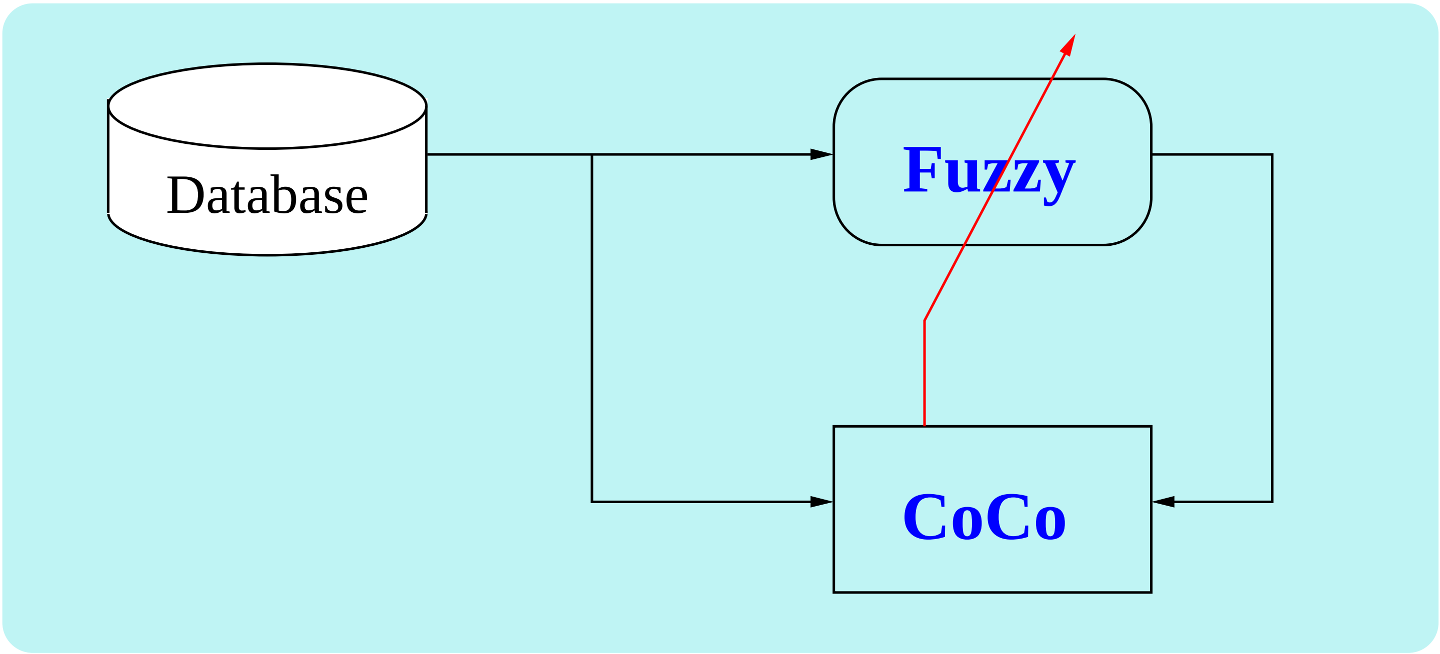
THEN the case is benign;

ELSE the case is malignant.

Proposed approach: two elements

1 A system model: **Fuzzy systems**

2 A building algorithm: **Cooperative coevolution**



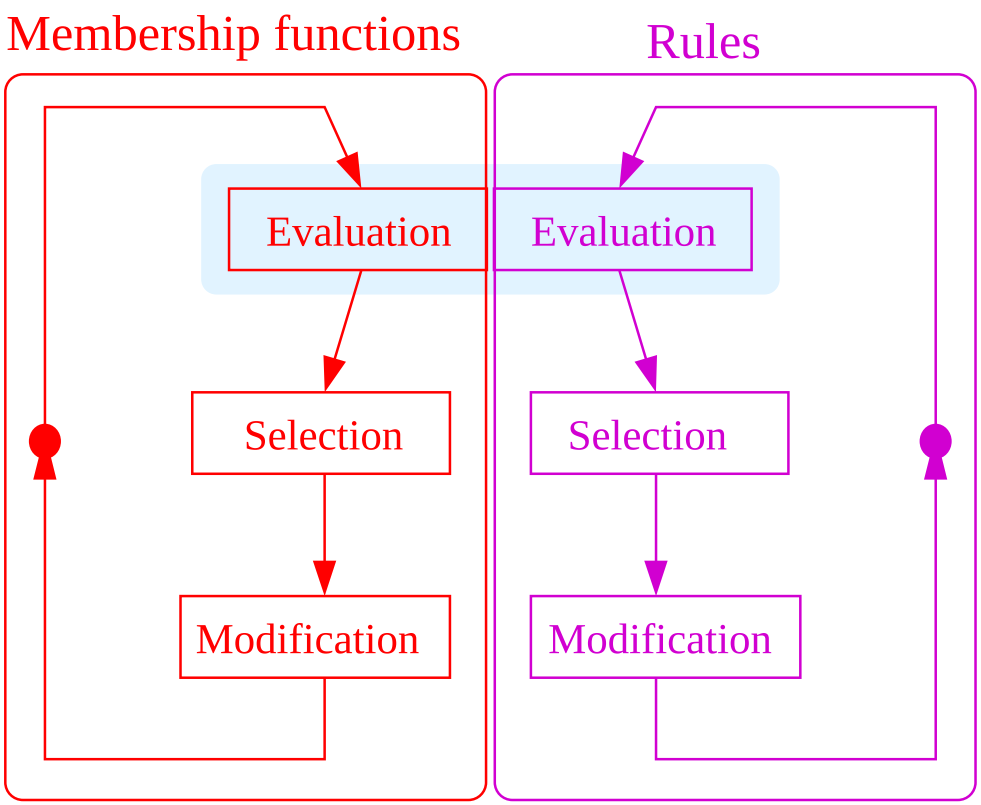
Fuzzy modeling: a coevolvable problem

Parameter class	Modeling type	Usual number	Type of values
Logical	System design	3 − 10	
Structural (size)	Structure learning	5 − 20	Integer
Connective (rules)	Behavior learning	10 − 1000	Symbolic
Operational (labels)	Knowledge tuning	10 − 1000	Real−valued

The required solutions can be very complex,
they can be decomposed in distinct components,
represented by different types of values,
and which are very interdependent.

These features render fuzzy modeling an adequate target for COOPERATIVE COEVOLUTION

Fuzzy CoCo:
A cooperative coevolutionary approach to fuzzy modeling



Two evolutionary algorithms searching for:

membership functions

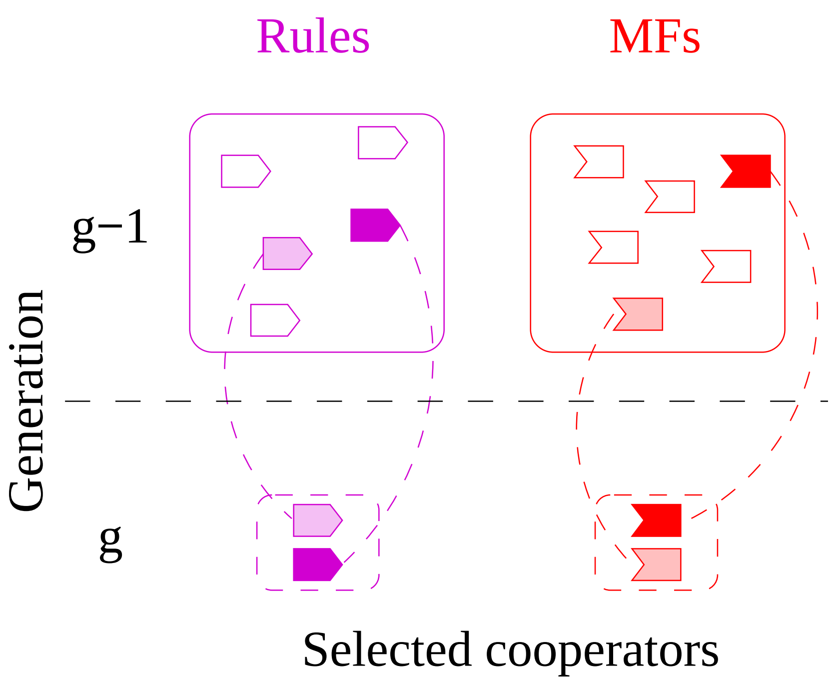
and rules.

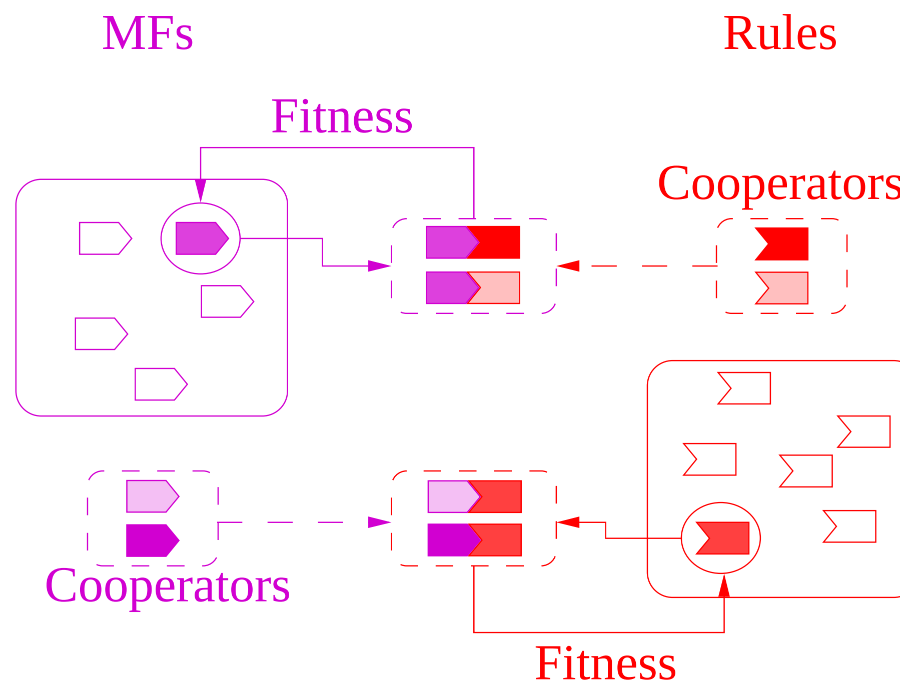
Advantages:

- Divide−and−conquer strategy
- Better search power
- Lesser computational cost
- More−flexible setup

Fitness evaluation in Fuzzy CoCo

1. Cooperators for generation g are selected from generation g−1 both fitness−dependent and randomly



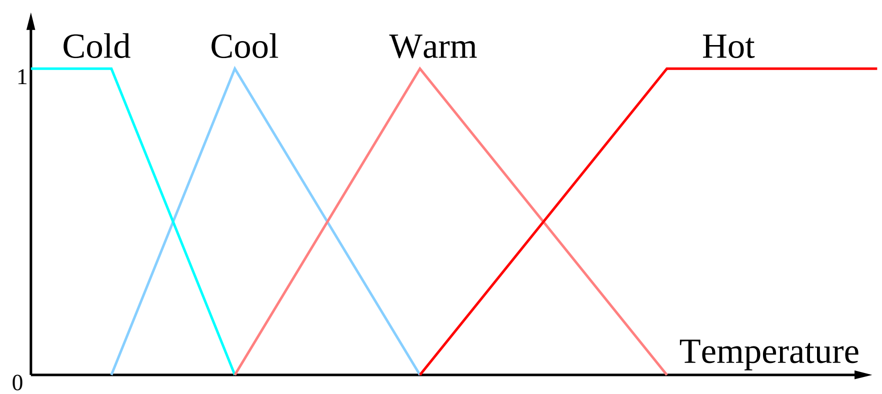


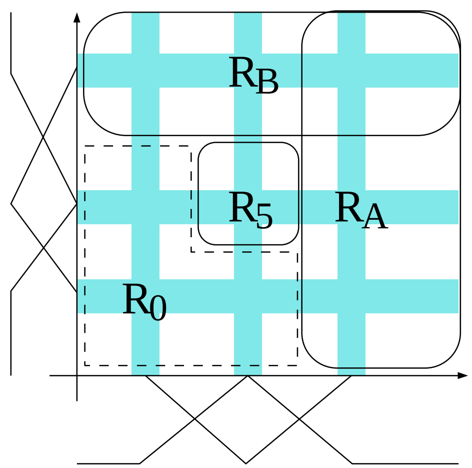
2. Individuals are combined with cooperators to form fuzzy systems.

3. These fuzzy systems are evaluated, and individual fitness is then calculated.

Interpretability strategies in Fuzzy CoCo

- Shared membership functions: reinforced by the existence of a separate species
- Normal, orthogonal membership functions: constrained representation
- Don’t care conditions: encourage shorter rules
- Default rule: guarantees complete coverage of the input space
- Linguistic fitness: when used, increases selective pressure for interpretability





Fisher's Iris Data

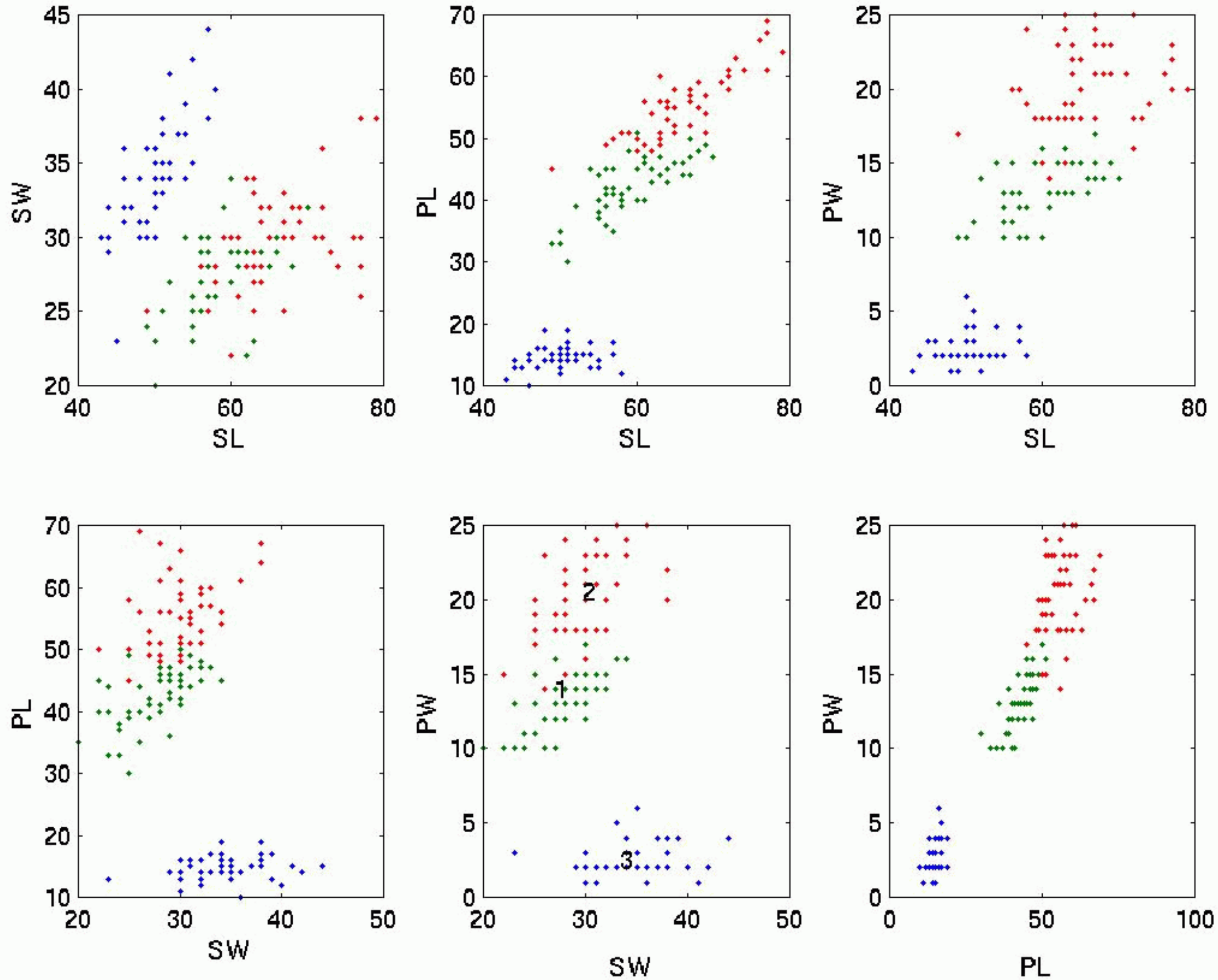
The variables

Features	**Classes**
(1) SL Sepal length	(1) setosa
(2) SW Sepal width	(2) versicolor
(3) PL Petal length	(3) virginica
(4) PW Petal width	

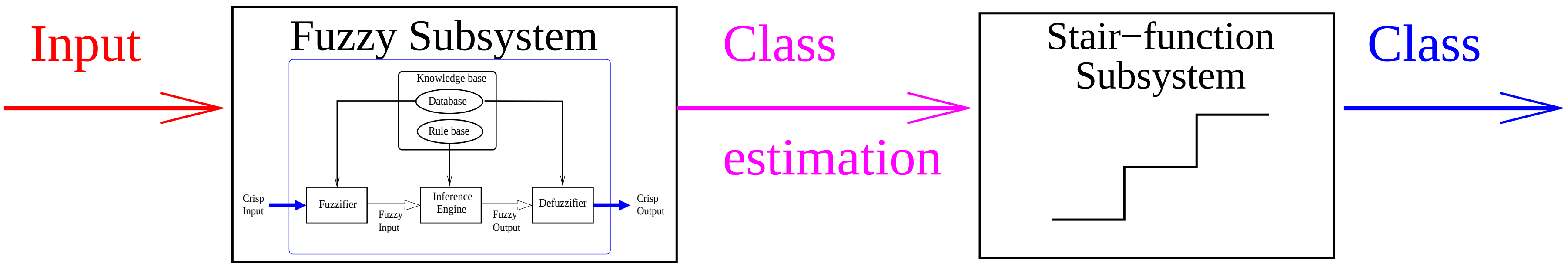
The database

Case	SL	SW	PL	PW	Class
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
150	5.9	3.0	5.1	1.8	Virginica

Iris: Variable analysis



Iris proposed solution: Controller−type



The fuzzy subsystem estimates a continuous "class" value

The selection unit approximates it to the nearest class

Iris controller−type: Proposed Fuzzy System

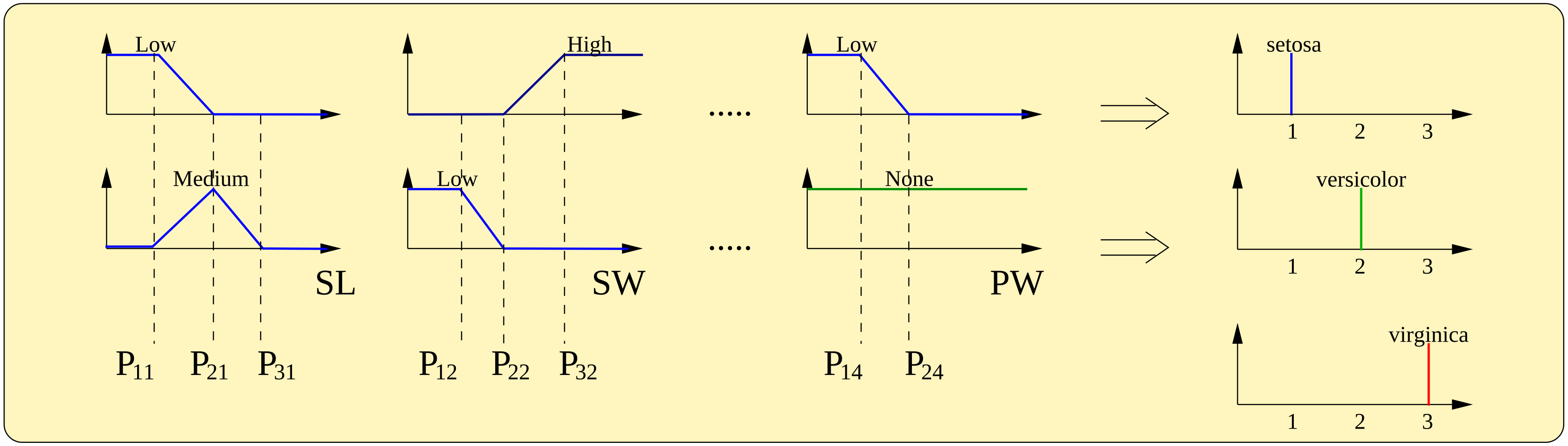
R1: if (SL is A11) and (SW is A12) and (PL is A13) and (PW is A14) then (output is Class1)

R2: if (SL is A21) and (SW is A22) and (PL is A23) and (PW is A24) then (output is Class2)

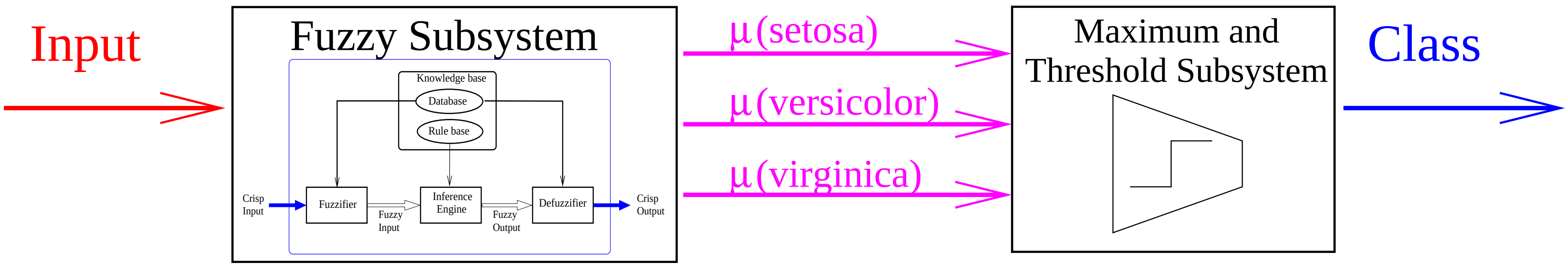
⋮ ⋮ ⋮ ⋮ ⋮

Rn: if (SL is An1) and (SW is An2) and (PL is An3) and (PW is An4) then (output is Classn)

else (Output is Class0)



Iris proposed solution: Classifier−type



The fuzzy subsystem estimates a continuous membership value for each class

The selection unit chooses the most active class

Iris classifier−type: Proposed Fuzzy System

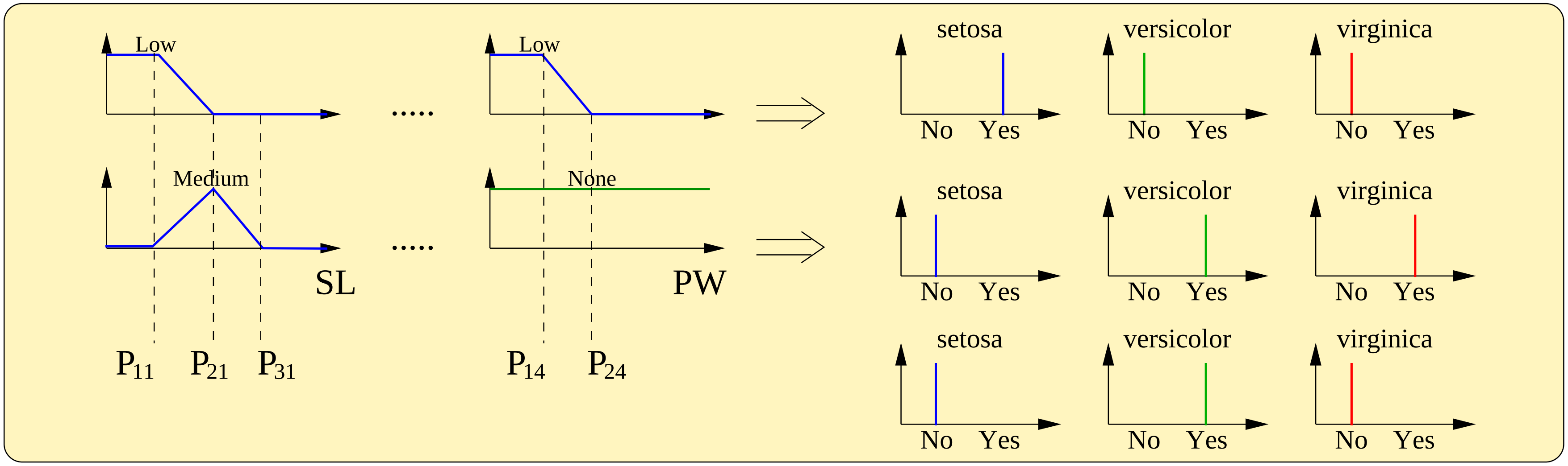
R1: if (SL is A11) and ... and (PW is A14) then (setosa is Yes),(versicolor is No),(virginica is No)

R2: if (SL is A21) and ... and (PW is A24) then (setosa is No),(versicolor is Yes),(virginica is Yes)

⋮ ⋮ ⋮ ⋮ ⋮

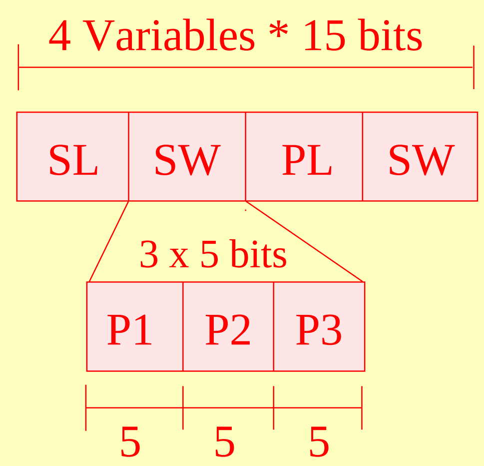
Rn: if (SL is An1) and ... and (PW is An4) then (setosa is No),(versicolor is No),(virginica is Yes)

else (setosa is No),(versicolor is Yes),(virginica is No)



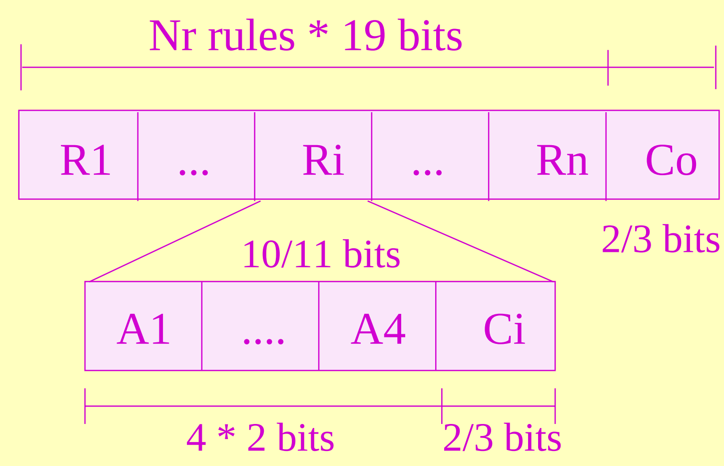
Iris: the genomes

Membership functions



Genome length = 60 bits

Rules (Controller/Classifier)



Genome length = 10/11*Nr + 2/3 bits

Iris: Fuzzy CoCo set−up

1. Fitness function

$$F = \begin{cases} Fc * Fm^b \\ (Fc + a*Fv) * Fm^b \end{cases}$$

Fc : Classification performance,
the most important

Fm : 1 − mse (mean square error)
encourages not−so−bad errors

Fv : Number of variables
measures the interpretability

2. Fuzzy CoCo parameters

Population size	{60, 70}
Maximum generations	500 + 100*Nr
Crossover probability	1
Mutation probability	{0.02, 0.05, 0.1}
Elitism rate	{0.1, 0.2}
"Fit" cooperators	1
Random cooperators	{1, 2}

Iris results: classification (average rule)

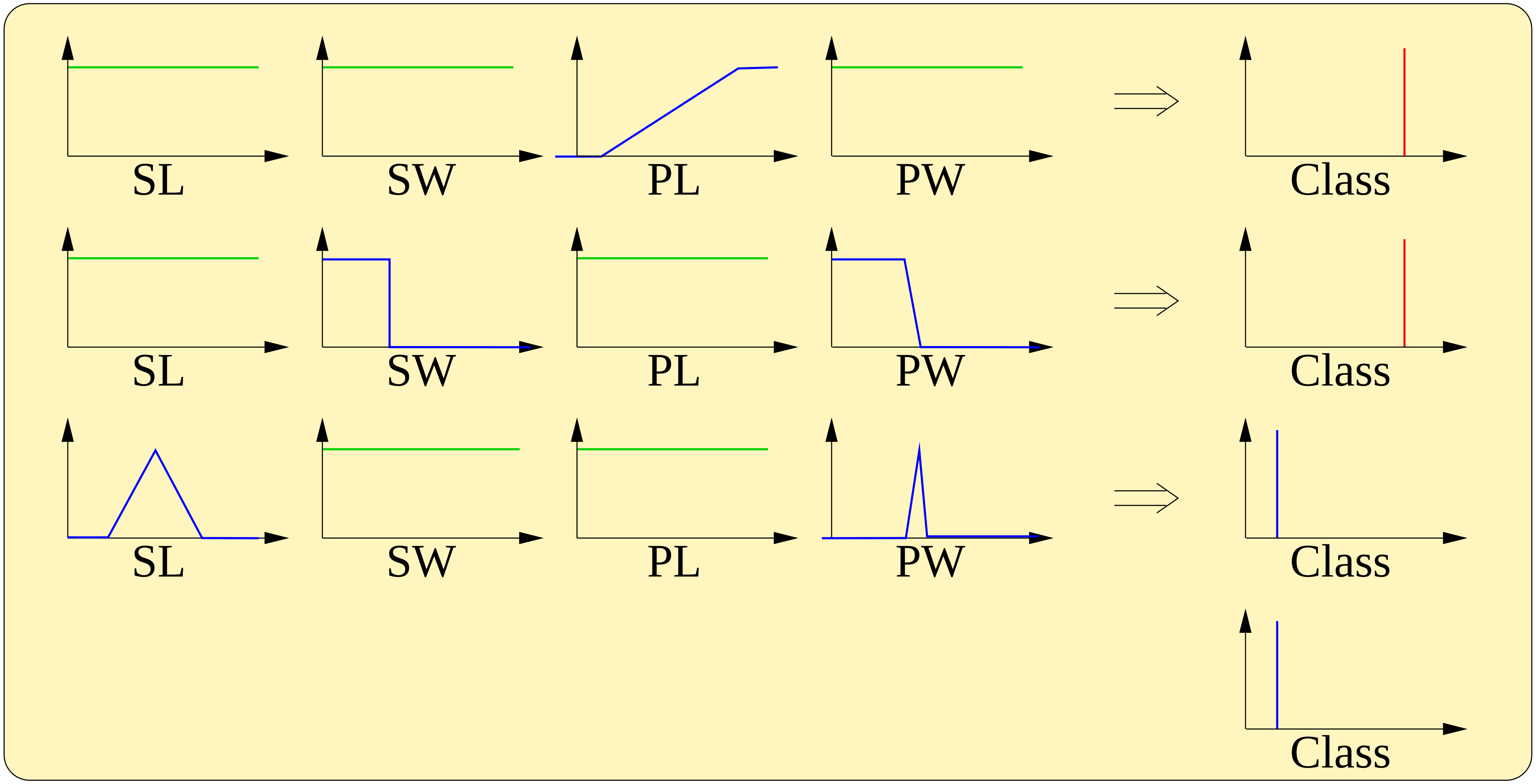
Controller

Rules	Simple GA Shi et al (1999)	FuGeNeSys Russo (1998)	Fuzzy CoCo ICANNGA'01
2			99.33% (2)
3			100 % (1.7)
4	98 % (2.6)		100 % (2.5)
5		100 % (3.3)	

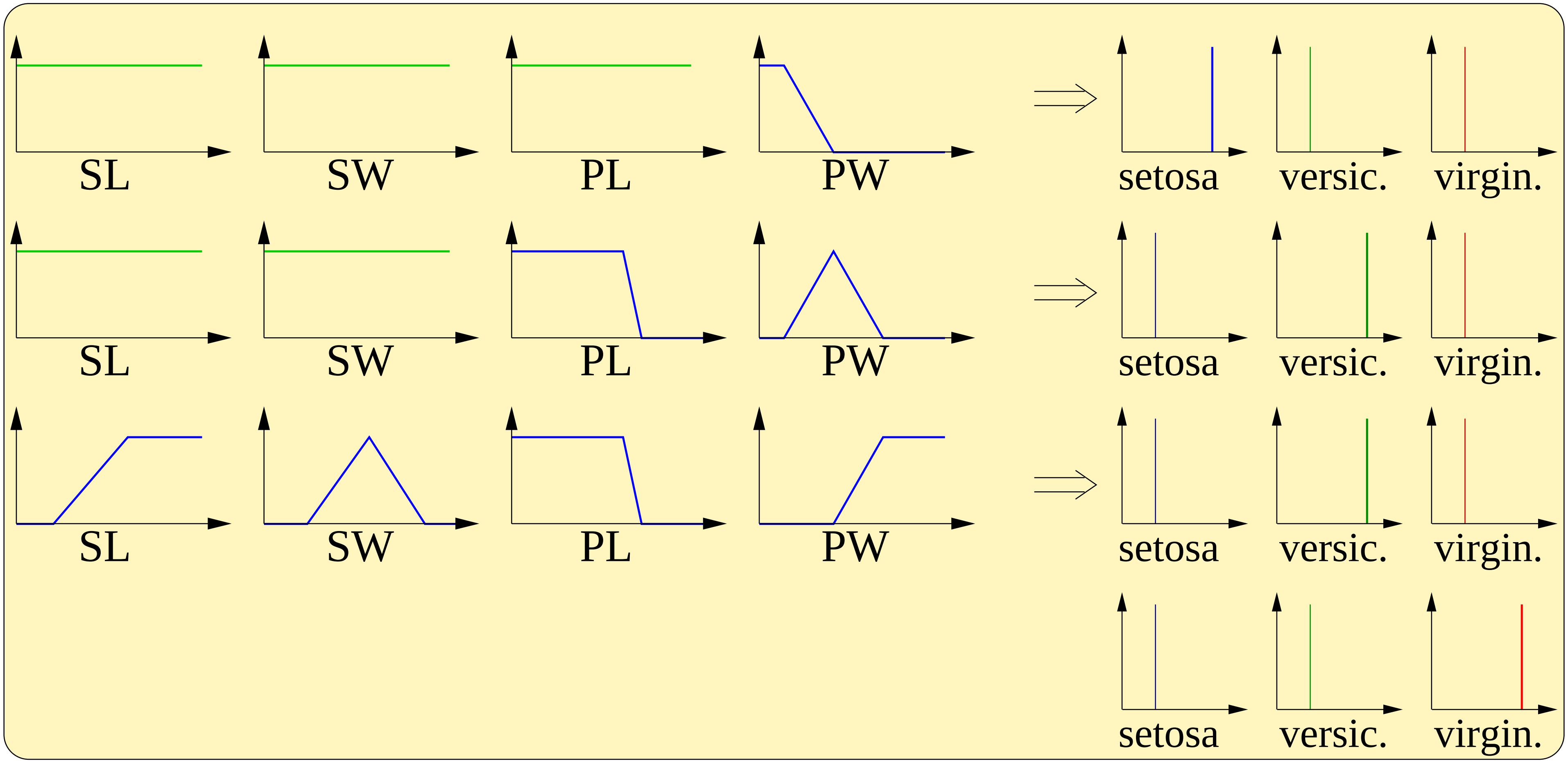
Classifier

Rules	Constructive Learning Methods		Neurofuzzy	Fuzzy CoCo
	Hong (00)	Wu (99)	Hung (99)	ICANNGA'01
2				98 % (1.5)
3		96.2 % (4)		99.33% (2.3)
4			97.4 % (4)	99.33% (2)
8	97.3 % (2)			

Iris controller−type: A three−rule system

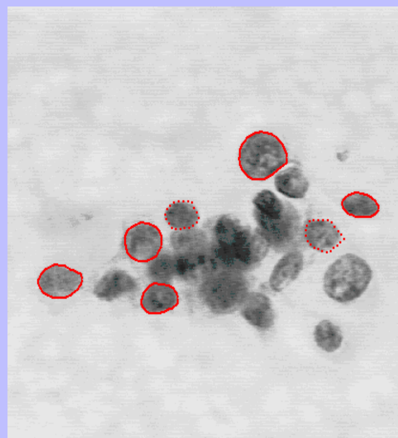


Iris classifier−type: A three−rule system



Breast cancer diagnosis: the WBCD problem

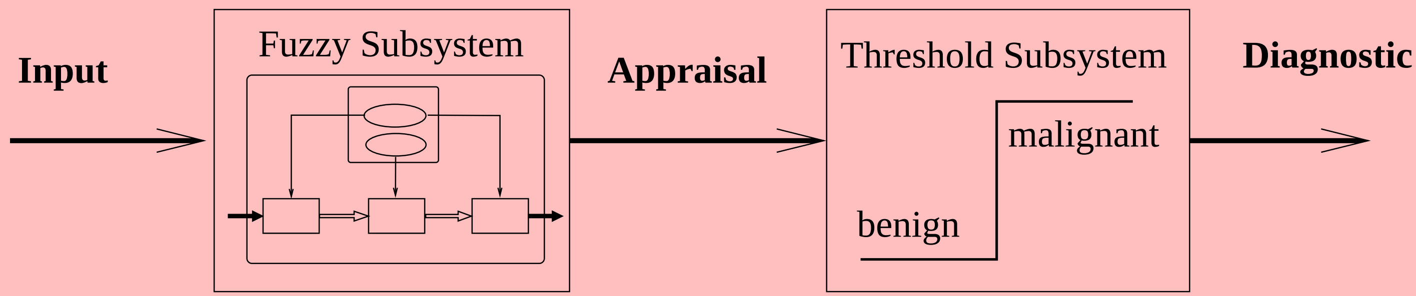
The test



The database

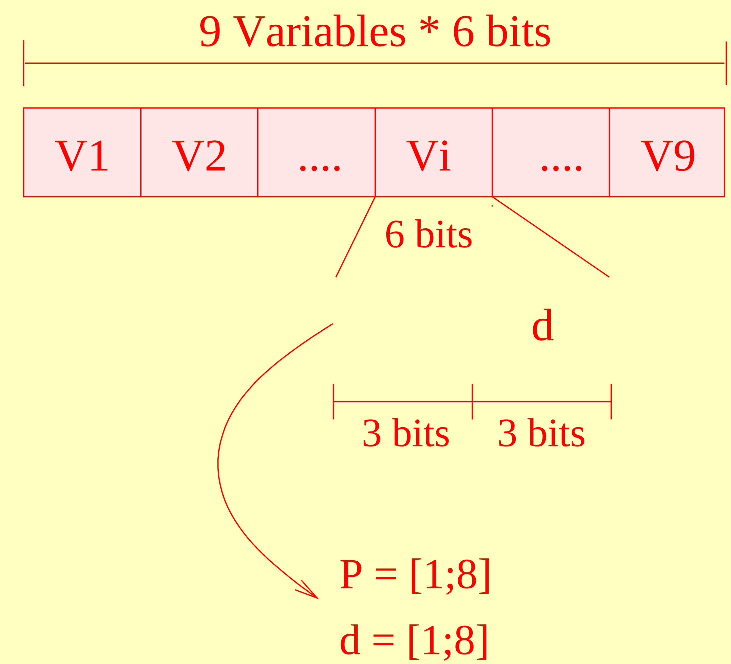
case	v_1	v_2	v_3	$\ldots$	v_9	*diagnostic*
1	5	1	1	$\ldots$	1	*Benign*
2	5	4	4	$\ldots$	1	*Benign*
:	:	:	:		:	:
683	4	8	8	$\ldots$	1	*Malignant*

Proposed solution



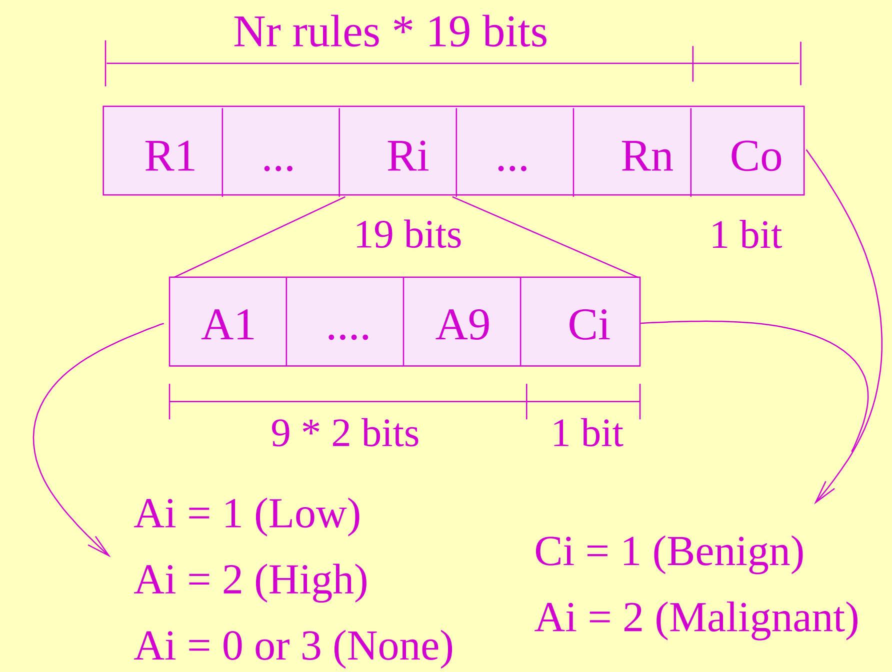
The genomes

Membership functions



Genome length = 54 bits

Rules



Genome length = 19*Nr + 1 bits

Fuzzy CoCo set−up

Fitness function

$$F = Fc - a^* \, Fv$$

Fc : Classification performance,
the most important performance measure

Fv : Number of variables
measures the interpretability

Fuzzy CoCo parameters

Population size	[30−90]
Maximum generations	1000 + 100*Nr
Crossover probability	1
Mutation probability	[0.02−0.3]
Elitism rate	{0.1−0.6]
"Fit" cooperators	1
Random cooperators	{1,2,3,4}

WBCD results: classification (longest rule)

Rules	NeuroRule Setiono (2000)	Fuzzy−genetic AIM 1999	Fuzzy CoCo − IEEE TFS 2001 Average	Best
1	97.36% (4)	97.07% (4)	97.36% (4.0)	97.36% (4)
2		97.36% (3)	97.73% (3.9)	98.54% (5)
3	98.10% (4)	97.80% (6)	97.91% (4.4)	98.54% (4)
4		97.80% (−)	98.12% (4.2)	98.68% (3)
5	98.24% (5)	97.51% (−)	98.18% (4.6)	98.83% (5)
6		98.10% (−)	98.18% (4.3)	98.83% (5)
7		97.95% (−)	98.25% (4.7)	98.98% (5)

Learned Boolean rules

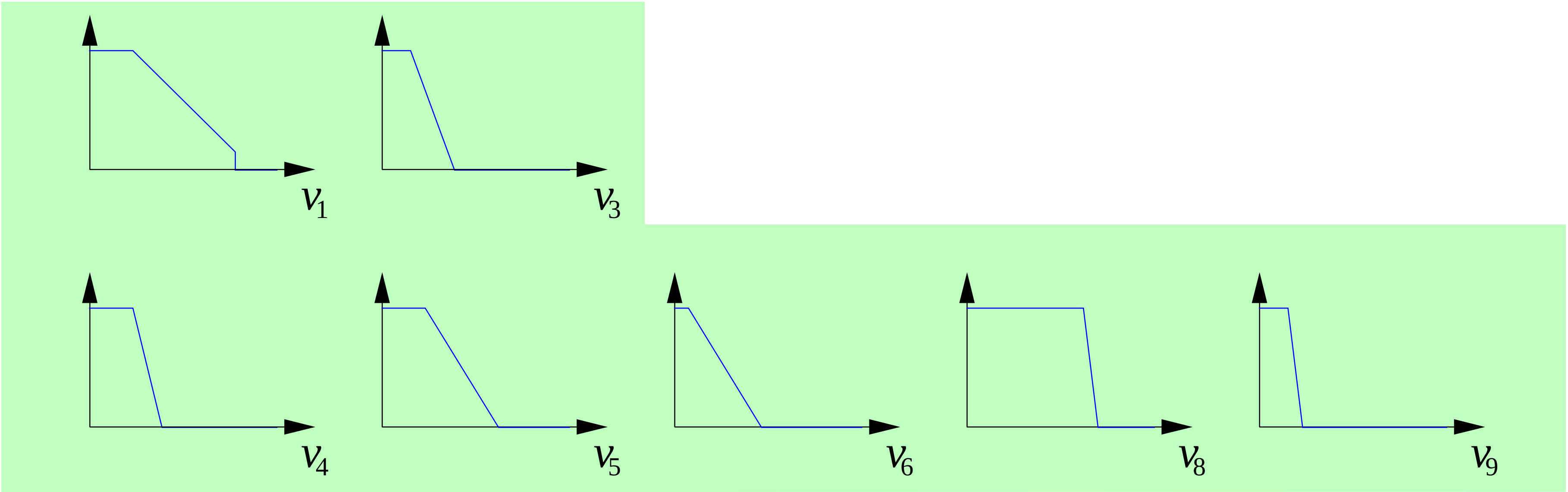
Evolved fuzzy rules

Two−rules evolved system

if (v1 is Low) and (v3 is Low) and (v5 is Low) then (output is Benign)

if (v1 is Low) and (v4 is Low) and (v6 is Low) and (v8 is Low) and (v9 is Low) then (output is Benign)

else (output is Malignant)



Classification rate = 98.54%

Computing requirements

Fuzzy GA: Single population (Peña & Sipper 99)

Number of fitness evaluations = Np * Gmax
200 * (2000 + 500Nr)

Single-rule systems: 500.000 fitness evaluations

Five-rule systems: 900.000 fitness evaluations

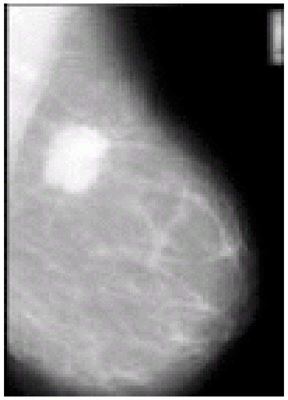
Fuzzy CoCo: Cooperative coevolution (CEC-2000)

Number of fitness evaluations = 2 * Np * Gmax * (Ncf + Ncr)
32.000 * (1000 + 100Nr) {worst case, Ncr=3}

Single-rule systems: 352.000 fitness evaluations

Five-rule systems: 480.000 fitness evaluations

The problem: mammography interpretation



mammogram reading protocol

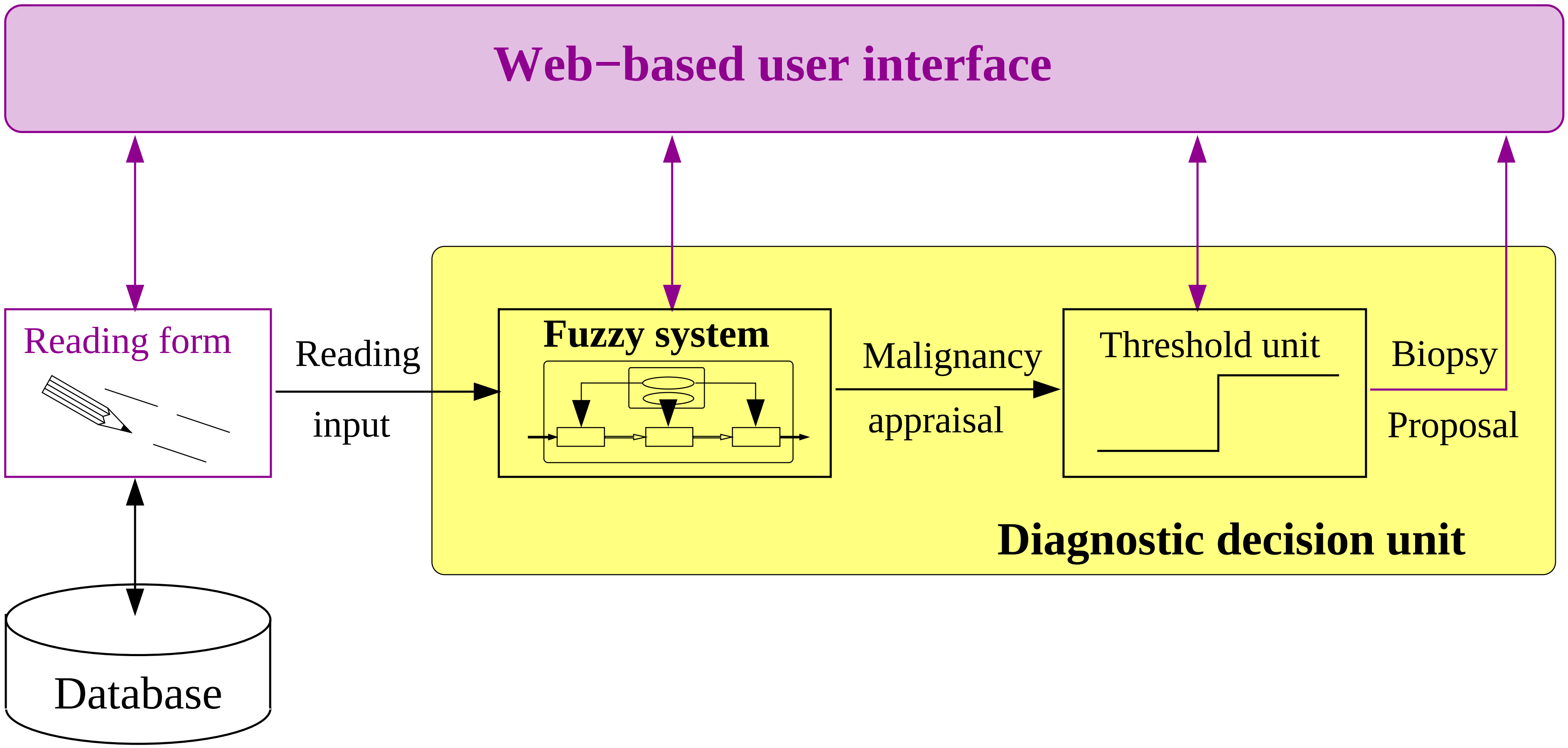
COBRA system: computer−assisted case interpretation

biopsy recommendation

Database

516 readings
- 187 malignant (pos)
- 329 benign (neg)

COBRA system: internal view



Understanding the database

CLINICAL DATA

Age :

Menopause : premenopausal

Antecedents : not available

Variable type	Number
Binary	4
Continuous	3
Discrete	8

MICROCALCIFICATION

Group shape : not available

Other group features : not available

Group diameter (mm) :

Number of micros : not available

Morphology of micros : not available

Size irregularity : not available

MASS

Mass morphology : not available

Mass margins : not available

Density bigger than parenchyma : not available

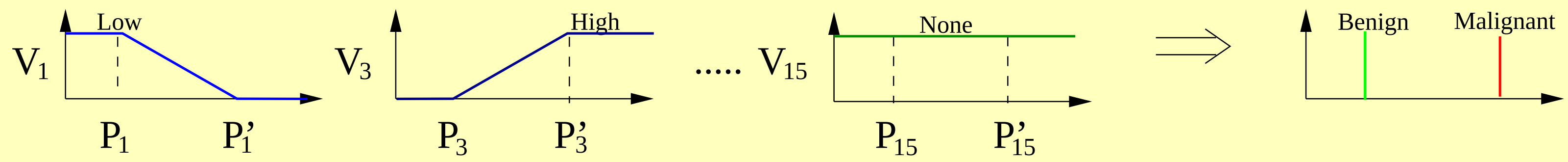
Focal distortion : not available

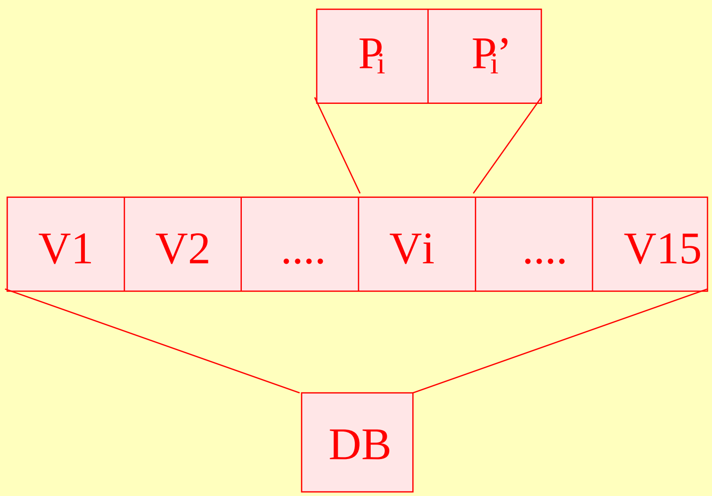
Focal symetry : not available

Focal Area (mm) :

Genome encoding for linguistic labels

Ri: if (v1 is Ai1) and (v2 is Ai2) and (v3 is Ai3) and ... and (v15 is Ai15) then (output is Ci)





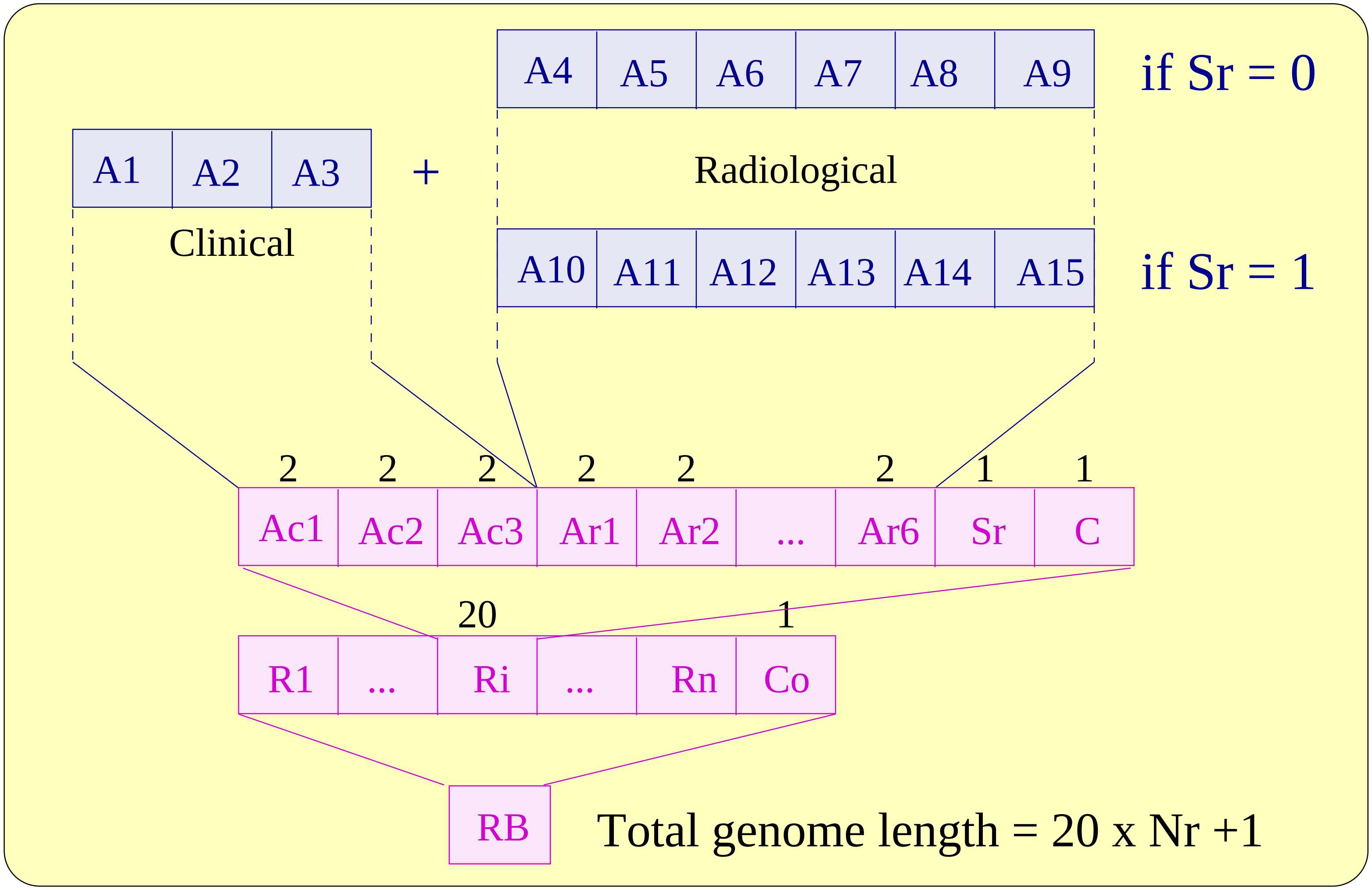
Binary variables (e.g., V2):
not encoded

Continuous variables (e.g., V1):
3 var. x 2 par. x 7 bits = 42 bits

Discrete variables (e.g., V3):
8 var. x 2 par. x 4 bits = 64 bits

Total genome length = 106 bits

Genome encoding for rules



Performance measures and fitness function

Sensitivity $\dfrac{TP}{TP + FN}$

Specificity $\dfrac{TN}{TN + FP}$

Accuracy $\dfrac{TP + TN}{TP+TN+FP+FN}$

PPV $\dfrac{TP}{TP + FN}$

Basic fitness (Fbase)

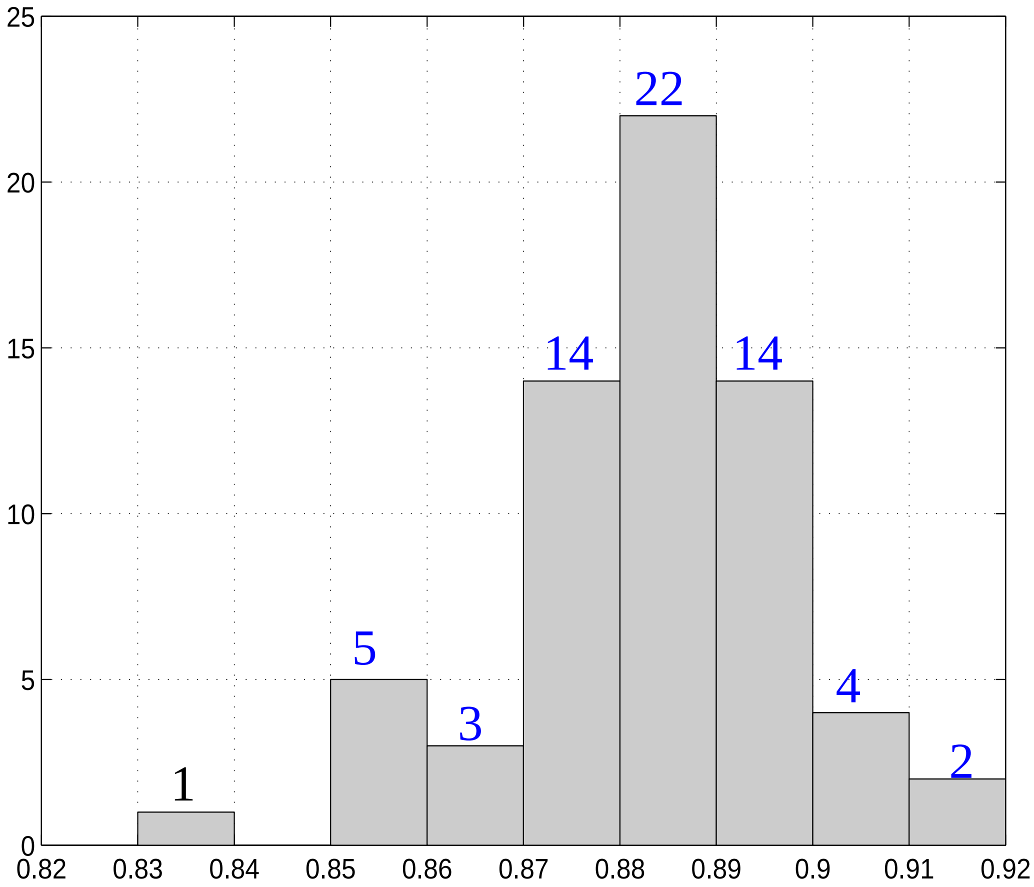
$$\frac{\text{Sensitivity} + \alpha\ \text{Specificity}}{1 + \alpha}$$

Accuracy reinforcement

$$\frac{\text{Fbase} + \beta\ \text{Accuracy}}{1 + \beta}$$

(note: done only if Accuracy > 0.7)

Fuzzy CoCo results on 65 runs



Average per class

Nr	Fitness	Reff	Vr
10	0.8754	9.17	2.52
15	0.8786	12.03	2.62
20	0.8934	14.15	2.59
25	0.8947	15.78	2.76

Best individual

Nr	Fitness	Reff	Vr
10	0.8910	9	2.22
15	0.8978	12	2.50
20	0.9109	17	2.41
25	0.9154	17	2.70

Performance of two selected systems

	17−rule		9−rule	
Measure	Figure	Ratio	Figure	Ratio
Sensitivity	99.47%	186/187	98.40%	184/187
Specificity	68.69%	226/329	64.13%	211/329
Accuracy	79.84%	412/516	76.55%	395/516
PPV	64.36%	186/289	60.93%	184/302

The 9−rule system with two different thresholds

	Threshold = 2		Threshold = 3	
Measure	Figure	Ratio	Figure	Ratio
Sensitivity	100.0%	187/187	98.40%	184/187
Specificity	63.22%	208/329	64.13%	211/329
Accuracy	76.55%	395/516	76.55%	395/516
PPV	60.71%	187/308	60.93%	184/302

COBRA system: reading form