Coevolution of a Fuzzy Rule Base for Classification Problems

Barbara Fusińska   Marek Kisiel-Dorohinicki   Edward Nawarecki

AGH University of Science and Technology, Kraków, Poland

RSEISP 2007

Outline

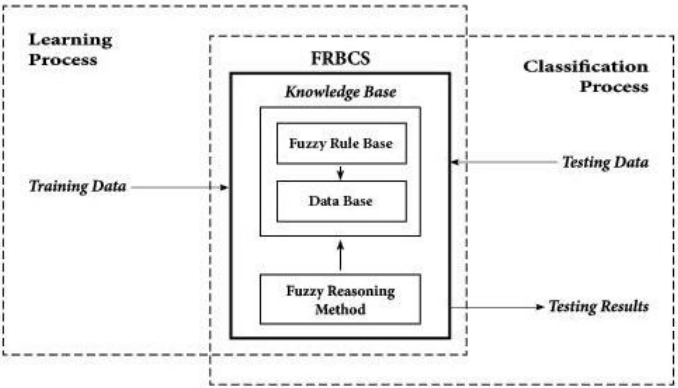
Fuzzy Rules for Classification

- **Classification** consists in assigning certain membership classes to objects.
- **Rules** represent the knowledge in a comprehensible form for those who will use the classification system.
- **Classification rules** can be based on **fuzzy logic**, which enables processing of imprecise or incomplete information, common in real classification problems.
- The systems that use fuzzy rules as knowledge representation are often called Fuzzy Rule-Based Classification Systems (FRBCS).

Fuzzy Rules representation

- Rule representation:
  IF x_1 **is** A_1 **AND** x_2 **is** A_2 **AND ... AND** x_n **is** A_n **THEN** C
- Conclusion mechanism:
  - After choosing proper rule class C becomes the response of the classifier.
  - Rule selection relies on membership values of classified object attributes.

The structure of FRBCS



Fuzzy Reasoning Method

Fuzzy Reasoning Method (FRM) - an inference procedure that derives a class to be assigned to object by applying the rules from the knowledge base to the read data.

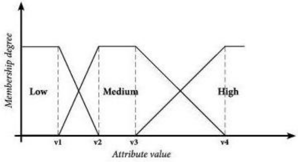
- In the first phase a matching degree is calculated for all the rules and the considered example.
- Then some rules are selected and used to draw the overall conclusion.

Evolving FRBCS

- Learning of ***fuzzy rules*** - generation of the rule base.
  - **Michigan approach** - in which the chromosomes represent single rules and the whole population an rule base.
  - **Pittsburg approach** - each chromosome encodes a whole rule base.
  - **Iterative Rule Learning approach** - each chromosome represents a single rule, but only the best individual (in iteration) is considered as the solution, discarding the remaining chromosomes in the population.
- Tuning the ***membership functions*** (data base) that describe the semantics associated to the linguistic labels used by the linguistic variables.

System description

- Each attribute value is assigned to one of linguistic labels (Low, Medium, High) that correspond to appropriate fuzzy sets defined separately for each attribute.



- Classification process according to the schema:
  - Compute membership values for every attribute in every rule.
  - Compute *matching degree* for the rule as a minimal value of membership values for all of its conditions.
  - The *winning rule* is the one with the highest matching degree.

Reproduction in system

Co-evolution of two populations:

- ***Rule bases*** - EMAS schema.
- ***Membership definitions*** - evolution strategy.

Fitness of data base is evaluated across to several rule bases individuals while of course the rule base needs membership function definition to calculate its own fitness.

Experimental studies

- Comparison tests of:
  - Iris set
  - Glass set
- Sets of data were divided to **learning** and **testing** parts according to divisions found in literature so that result can be properly compared.

Quality comparison for *iris data* - Phase 1

	Classifier quality for learning set [%]	Classifier quality for test set [%]
FCSOM	99.23	94.83
Nozaki	-	93.03
Umano	-	94.43
Our system	**98.07**	**93.48**

Quality comparison for *iris data* - Phase 2

	Classifier quality for learning set [%]	Classifier quality for test set [%]
C4.5	98.38	92.70
CN2	98.92	94.16
LVQ	98.55	95.72
FRBCS - FRM Classic	95.49	94.26
FRBCS - FRM WNS	97.47	94.36
WM-FRLP - FRM Classic	90.97	88.25
WM-FRLP - FRM WNS	98.56	94.38
Our system	**98.20**	**99.20**

Quality comparison for *glass kinds* data

	Classifier quality for learning set [%]	Classifier quality for test set [%]
LDA	73.74	83.33
SVN - linear	70.53	62.5
SVN - quad	73.68	75
SVN - RBF	86.84	37.5
CART - full tree	87.00	71.00
CART - best tree	81.00	67.00
Neural nets	80.95	75.00
Our system	**94.12**	**81.90**

Concluding remarks

- **Co-evolution** allows for fitness of a given set of membership function definitions to be evaluated across several fuzzy rule sets.
- A **multi-agent environment** allows for organization of an evaluation process.
- The reported preliminary results show **high classification quality** for the considered problems, as compared to the results of several other approaches found in the literature which used the same data sets.
- Future research
  - Adaptation of the system with high dimensional problems.
  - Influence of increasing the number fuzzy sets per attribute.