Can Evolution Strategies Improve Learning Guidance in XCS? Design and Comparison with GA-based XCS

Sergio Morales-Ortigosa

Albert Orriols-Puig

Ester Bernadó-Mansilla

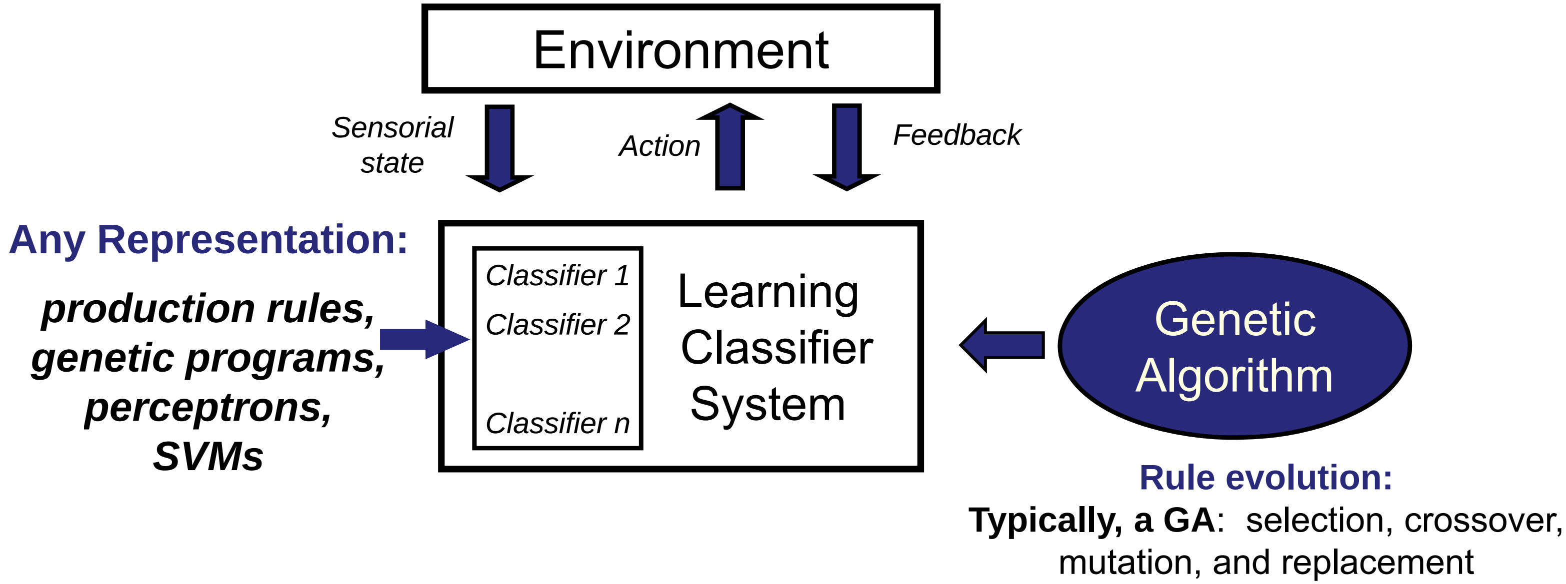
Enginyeria i Arquitectura La Salle

Universitat Ramon Llull

{is09767,aorriols,esterb}@salle.url.edu

Framework

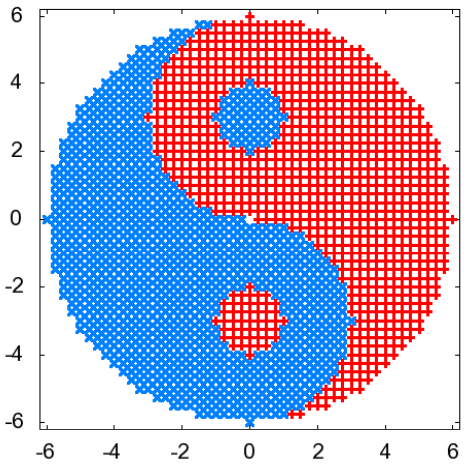
- Michigan-style LCSs (Holland, 1976) have reached maturity

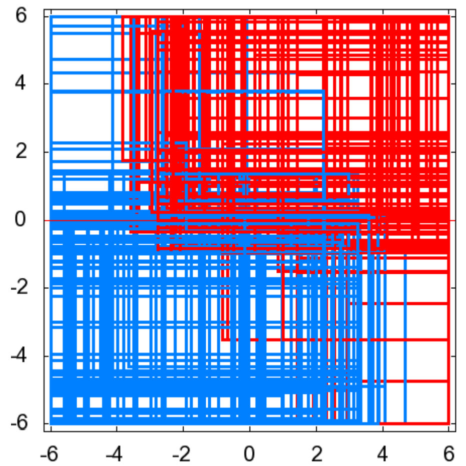


- Extended Classifier System - XCS (Wilson, 1995, 1998)
  - By far, the most influential LCS

Motivation

- **Problems with continuous attributes**
  - Interval-based representation (Wilson, 2001)
  - **IF** $v_1 \in [l_1, u_1]$ ***and*** $v_2 \in [l_2, u_2]$ ***and*** … ***and*** $v_n \in [l_n, u_n]$ **THEN** $class_i$





- **They yield competitive results, but we have little understanding of how they work!**

- **2-point crossover**
  - **Too disruptive?**
- **Mutation:** add a random uniform value
  - **Could we use more information?**

- **Could we design better genetic operators?**
  - Not exactly clear the impact of crossover and mutation
  - Systematic analysis
  - Creative analysis: propose new operators

Purpose of the Work

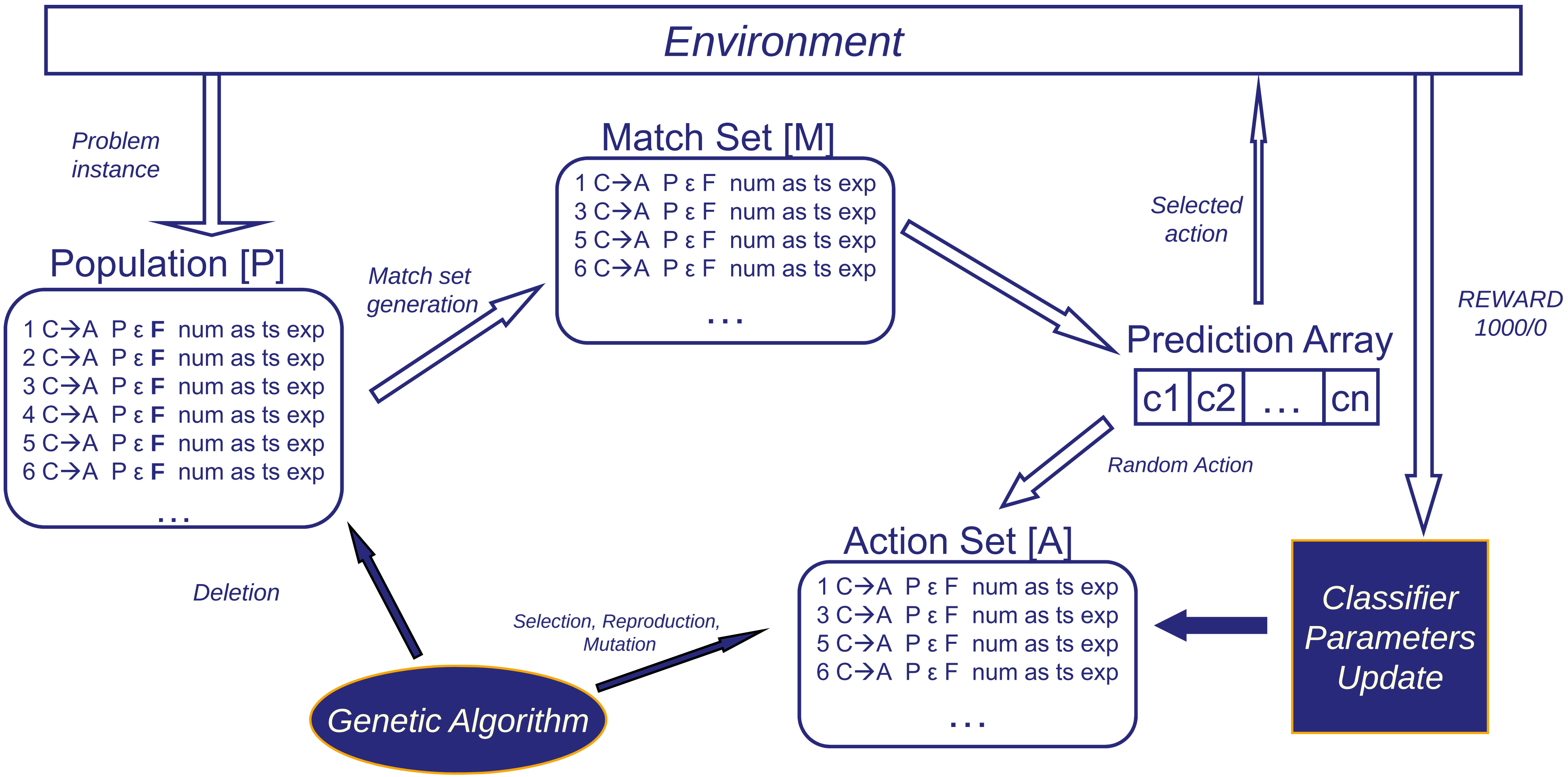
- **Looking at the continuous optimization realm**
  - Evolution strategies
  - Real-coded GAs

- **The purpose of this work is to**
  - Design an XCS based on evolution strategies (ES)
    - **Adapt classifier representation**
    - **Design ES mutation and crossover alike for XCS**
  - Analyze the role of Gaussian mutation
  - Compare whether ES-based XCS outperforms GA-based XCS

Outline

1. Description of XCS
2. Evolution Strategies in XCS
3. Experimental Methodology
4. Results
5. Conclusions and Further Work

Description of XCS



Genetic Operators

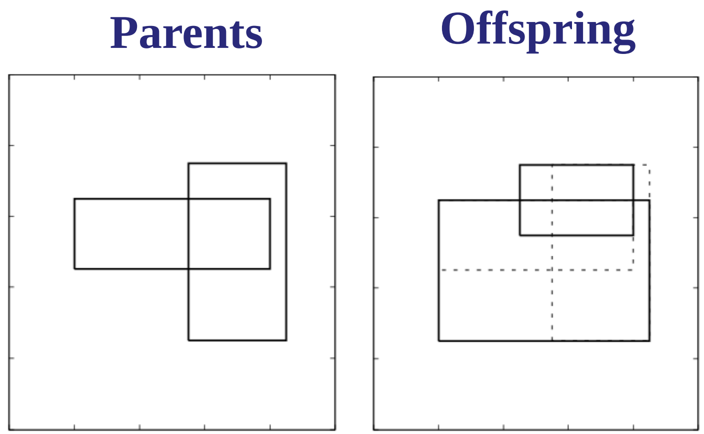
Selection

- Proportionate selection
- Tournament selection

Crossover:

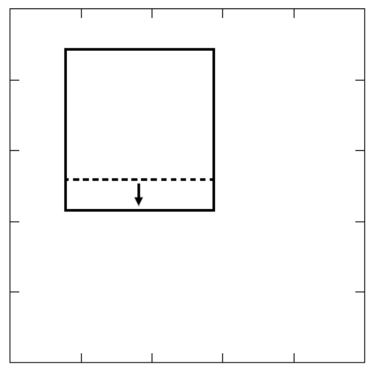
- Two-point crossover

$$\begin{array}{l} <[0.20,0.80],[0.45,0.55]> \\ <[0.60,0.85],[0.25,0.75]> \end{array} \Longrightarrow \begin{array}{l} <[0.20,0.85],[0.25,0.55]> \\ <[0.60,0.80],[0.45,0.75]> \end{array}$$



Mutation:

- **GA-based XCS:** Add a uniform random value



Outline

1. Description of XCS
2. Evolution Strategies in XCS
3. Experimental Methodology
4. Results
5. Conclusions and Further Work

GAs vs ESs Head to Head

- **Genetic Algorithms**
  - Initially used with binary representation
  - Key aspects:
    - **GAs process (mix & ensemble) building blocks**
    - **Crossover as primary search operator**
    - **Mutation as local search operator**

- **Evolution Strategies**
  - Initially designed for problems with continuous attributes
  - Key aspects:
    - **Search focuses little improvement/selection**
    - **Gaussian mutation is the search operator**
    - **Crossover included afterwards to resemble GAs**

ES-based XCS

- **Representation extended with a vector of strategy parameters**

IF $v_1 \in [l_1, u_1]$ ***and*** $v_2 \in [l_2, u_2]$ ***and*** … ***and*** $v_n \in [l_n, u_n]$ **THEN** $class_i$

$(\sigma_1, \sigma_2, \ldots, \sigma_n)$

  - The strategy parameters (SP) evolve with the representation
  - Genetic operators modified to deal with the new rep.

- **Mutation**
  - Intervals *i* mutated as:

$$l_i = l_i + \sigma_i N(0,1) \qquad u_i = u_i + \sigma_i N(0,1)$$

  - Strategy parameter vector mutated as:

$$\sigma_i' = e^{\tau_0 N_0(0,1)} e^{\tau \, N_i(0,1)}$$

where
$\tau_0 = 1/(2n)^{0.5}$ and $\tau = 1/(2n^{0.5})^{0.5}$

ES-based XCS

- **Crossover**
  - Discrete/dominant recombination for *object parameters*
    - **Each variable and SP are randomly selected from one parent**
  - Intermediate recombination for *strategy parameters*
    - **Calculates the center of mass of the parents**
    - **Pushes to the average value**

- **Selection**
  - Fitness proportionate selection
  - Tournament selection
  - Truncation selection

Outline

1. Description of XCS

2. Evolution Strategies in XCS

3. Experimental Methodology

4. Results

5. Conclusions and Further Work

Experimental Methodology

- **Analyze the effects of**
  - Selection + mutation (local search)
  - Selection + mutation + crossover (innovation)
- **Experiments run on 12 real-world data sets (UCI rep.)**
  - 10-fold cross-validation

Id.	dataset	#Inst	#Fea	#Re	#In	#No	#Cl	%MisInst
bal	Balance	625	4	4	0	0	3	0
bpa	Bupa	345	6	6	0	0	2	0
gls	Glass	214	9	9	0	0	6	0
h-s	Heart-s	270	13	13	0	0	2	0
irs	Iris	150	4	4	0	0	3	0
pim	Pima	768	8	8	0	0	2	0
tao	Tao	1888	2	2	0	0	2	0
thy	Thyroid	215	5	5	0	0	3	0
veh	Vehicle	846	18	18	0	0	4	0
wbcd	Wisc. breast-cancer	699	9	0	9	0	2	2.3
wdbc	Wisc. diagnose breast-cancer	569	30	30	0	0	2	0
wne	Wine	178	13	13	0	0	3	0

Experimental Methodology

- **Results statistically compared by means of**
  - The multicomparison Friedman test
  - The post-hoc Bonferroni-Dunn test for multiple comparisons
  - The Wilcoxon signed-ranks test for pairwise comparisons

- **XCS configured as:**
  - #iter=100000, N = 6400, θ_{GA} = 50, P_{cross} = 0.8, P_{mut} = 0.04, r_0 = 0.6, m_0 = 0.1

Outline

Analysis of Selection + Mutation

- Test Accuracy

	XCS-GA with proportionate	XCS-ES with proportionate	XCS-GA with tournament	XCS-ES with tournament	XCS-ES with truncation	XCS-ES weighted mutation
Dataset	**XCS$_{GA}$-ps**	**XCS$_{ES}$-ps**	**XCS$_{GA}$-ts**	**XCS$_{ES}$-ts**	**XCS$_{ES}$-tr**	**weig. XCS$_{ES}$**
pim	70.83 (4)	71.05 (2)	69.99 (6)	72.87 (1)	70.88 (3)	70.53 (5)
tao	89.32 (6)	92.90 (3)	89.79 (5)	93.80 (1)	93.01 (2)	89.90 (4)
thy	94.73 (6)	95.50 (4)	95.66 (2.5)	96.28 (1)	94.88 (5)	95.66 (2.5)
veh	65.52 (3)	66.00 (2)	64.50 (4)	67.26 (1)	63.83 (6)	64.34 (5)
wbcd	80.88 (6)	85.84 (1)	81.26 (5)	85.65 (2)	82.50 (4)	82.93 (3)
wdbc	78.68 (2)	75.28 (3)	80.20 (1)	74.93 (4)	67.60 (6)	69.13 (5)
wne	80.71 (5)	86.70 (1)	82.21 (2)	82.02 (3)	78.09 (6)	81.65 (4)
rnk.	4.38	2.67	3.54	2.00	4.50	3.92
pos.	**5**	**2**	**3**	**1**	**6**	**4**

According to a post-hoc Bonferroni-Dunn test:

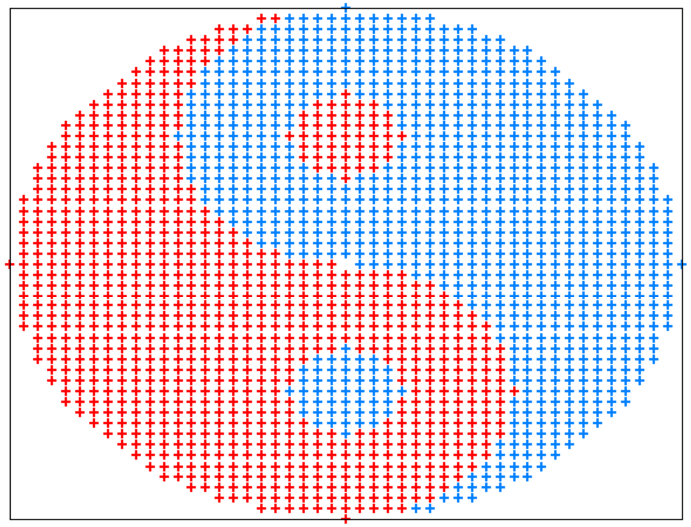
- **XCS-ES tourn.** significantly outperformed **XCS-GA** with both selection schemes
- **XCS-ES proportionate** significantly outperformed **XCS-GA proportionate**

Selection + Crossover + Mutation

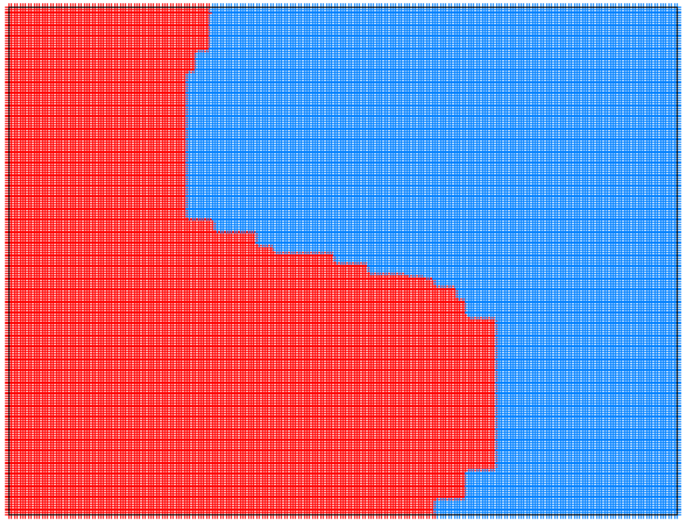
Dataset	XCS-GA with proportionate XCS_{GA}-ps	XCS-ES with proportionate XCS_{ES}-ps	XCS-GA with tournament XCS_{GA}-ts	XCS-ES with tournament XCS_{ES}-ts	XCS-ES with truncation XCS_{ES}-tr
bal	83.20 (1)	82.77 (2.5)	82.72 (4)	82.77 (2.5)	82.13 (5)
pim	72.53 (5)	74.43 (2)	73.39 (4)	73.83 (3)	74.74 (1)
tao	91.22 (4)	93.52 (3)	91.19 (5)	94.35 (1)	94.17 (2)
thy	95.81 (2)	95.50 (5)	96.43 (1)	95.66 (3.5)	95.66 (3.5)
veh	71.79 (2)	71.75 (3)	71.20 (4.5)	72.89 (1)	71.20 (4.5)
wbcd	94.85 (3)	95.47 (1.5)	93.51 (4)	95.47 (1.5)	92.61 (5)
wdbc	91.09 (4)	91.80 (3)	92.44 (2)	92.85 (1)	89.51 (5)
wne	95.50 (4)	96.25 (1.5)	95.69 (3)	96.25 (1.5)	91.38 (5)
rnk.	2.79	2.5	3.33	2.17	4.21
pos.	**4**	**2**	**3**	**1**	**5**

- **XCS-ES still is the best method**
- **But now, no significant differences**

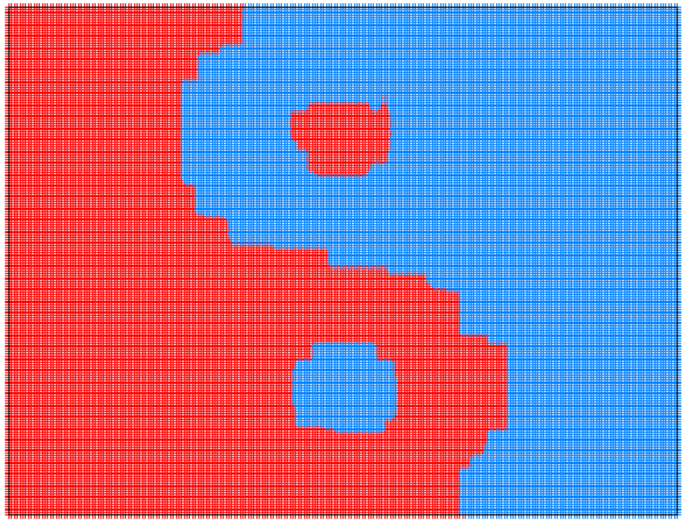
A Cool Example



Domain



XCS-GA with proportionate selection



XCS-ES with proportionate selection

Outline

1. Description of XCS

2. Evolution Strategies in XCS

3. Experimental Methodology

4. Results

5. Conclusions and Further Work

Conclusions

- **The analysis performed in this paper permitted**
  - To study the discovery component of XCS, especially focusing on the role of mutation.
  - Improve XCS to deal with problems with complex boundaries described by continuous attributes.
- **Two important observations:**
  - Gaussian mutation performs innovation tasks.
  - When crossover is included XCS-GA does not significantly outperform XCS-ES. But still, it wins.
- **The overall work clearly shows the importance of further researching on GA operators.**

Further Work

- **XCS-ES is good! But, always?**
  - On average, yes!
  - Specific problems may not benefit from ES operators

- **May evolution tell me when to use one type of search or another?**
  - Existing studies on self-adaptation mutation for ternary rules
  - Search for evolution signals
  - Combine different operators
  - Let classifiers decide which operator to use
  - Characterize learning domains

Can Evolution Strategies Improve Learning Guidance in XCS? Design and Comparison with GA-based XCS

Sergio Morales-Ortigosa

Albert Orriols-Puig

Ester Bernadó-Mansilla

Enginyeria i Arquitectura La Salle

Universitat Ramon Llull

{is09767,aorriols,esterb}@salle.url.edu