

A New Approach to Fuzzy LCSs in Two-Dimensional Continuous Multistep Environment with Continuous Vector Actions

José A. Ramírez-Ruiz
Center for Intelligent Systems
Tecnológico de Monterrey
64849 Monterrey, N.L., Mexico
a00792432@itesm.mx

Manuel
Valenzuela-Rendón
Center for Intelligent Systems
Tecnológico de Monterrey
64849 Monterrey, N.L., Mexico
valenzuela@itesm.mx

Hugo Terashima-Marín
Center for Intelligent Systems
Tecnológico de Monterrey
64849 Monterrey, N.L., Mexico
terashima@itesm.mx

ABSTRACT

This paper introduces a new Fuzzy Learning Classifier System that can learn a continuous action vector field from two-dimensional continuous multistep environments. Results where this approach is applied with success to two simple two-dimensional problems are presented.

Categories and Subject Descriptors

I.2.6 [Artificial Intelligence]: Learning—*Induction and Knowledge acquisition*; I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search—*Dynamic programming, Plan execution, formation, and generation*

General Terms

Algorithms, Experimentation

Keywords

Learning Classifier Systems, Fuzzy Classifier Systems, Fuzzy Logic, Genetic Algorithm, Induction Theory

1. INTRODUCTION

In the last years, many techniques have risen in the Machine Learning field; among others are Neural Networks (off-line learning), Reinforcement Learning (on-line method), Genetic Algorithms (off-line method), Learning Classifier Systems (on-line method), etc. Learning Classifier Systems (LCSs) combine a Genetic Algorithm (GA) with a production system to learn a transition function of the environment using only rewards; these systems have been applied in many contexts, but they were explored in simple grid environments in their beginnings. Fuzzy LCSs (FCSs) were developed to deal with continuous environments; two LCSs [5, 6, 4] that use fuzzy logic (FL) were introduced to learn

functions from rewards in one step environments. Bonarini introduced a FCS [1] where he used a scheme of competition and cooperation among rules to learn an action function in multistep environments with continuous reward. This system produced continuous actions from a continuous perception but with an exponential number of fuzzy states to represent the perception states. On the other hand, XCSF by Lanzi et al. [3] does not use FL and was able to act in multistep, one and two dimensional, continuous spaces but with a set of discrete actions. Kovacs et al. [2] presented three different architectures that used combinations of two XCSFs to deal with continuous actions in one step, one dimensional environments.

2. THE FUZZY LCS

This section presents an overview of the system proposed, the environments where the new approach was tested, and the results obtained.

2.1 Description of the Fuzzy LCS

In our FLCS, Each rule R_i contains a small fuzzy system with four fuzzy rules, a matrix that contains the expected prediction p_{uv}^i , and a square subregion X_i of activation over the input space (x_1, x_2) . Rules can only act over their subregion of activation. In that subregion of activation, each fuzzy system proposes a continuous vector field $\vec{y}_i = (y_1, y_2)$ by defuzzification. In that way, when an input vector \vec{x} enters the FLCS, rules cooperate to produce the real vector output \vec{y}_{Output} using the vector addition of the output vectors \vec{y}_i proposed by rules through their fuzzy systems weighted by their expected predictions. A modified Q-learning algorithm is used to learn the task from the environment, i.e., the learning of the continuous vector action function. This algorithm is used to change the values of the matrices of the expected predictions of each rule in the FLCS. A GA is applied to evolve rules based on their average expected values. This GA only evolves the action parts of the fuzzy systems.

2.2 Description of the Testing Environments

The environment is a finite two-dimensional space where the FLCS lives and moves. The environment has a goal that is defined as a region. The environment can have obstacles that are defined as regions on the space where the FLCS cannot move to. The FLCS can perceive its position in that environment through a position vector \vec{x} . The FLCS can

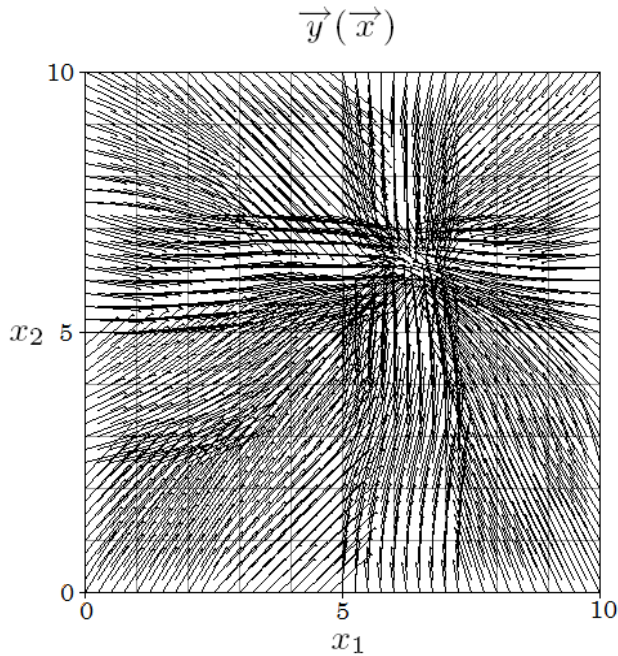


Figure 1: Action vector field learned by FLCs in a space without obstacles and with one goal = $\{(x_1, x_2) | 6 \leq x_1 \leq 7, 6 \leq x_2 \leq 7\}$

move in the environment with an action vector that moves it a distance of $\Delta \vec{x}$ where $\Delta \vec{x} = \vec{y}_{\text{Output}}$. When the FLCs reaches the limits of the environment (or the edge of an obstacle), the parallel component of the action vector to the normal vector to the perimeter of the corresponding limit of the environment (or of the obstacle) is eliminated. In this way, the FLCs cannot leave the environment and cannot move into any obstacle. A reward function $R(\vec{x})$ defines the task the FLCs has to learn; it defines the amount of reward given to the FLCs.

2.3 Results

In our experiments we have used 800 rules and we have divided the input space into 4×4 different uniform squared subregions where the rules can act. Rules are randomly distributed over all the subregions to have approximately 50 rules per subregion. p_{uv}^i was a 5×5 matrix and initially set at random in $[0, p_I]$. Figures 1 and 2 show some examples of the vector action fields learned after approximately 300,000 trails. The system was given a reward of 10 when it reached the goal, and 0, otherwise.

3. CONCLUSIONS

This paper presented a new approach to fuzzy LCSs. The new FLCs improves on previous efforts because it can deal with multistep environments where it must produce continuous outputs, and reward is not received at every step. We have shown that our approach works well in two-dimensional environments that contain obstacles.

4. ACKNOWLEDGMENTS

This research was supported by the Tecnológico de Monterrey (ITESM) under the Research Chair CAT-010.

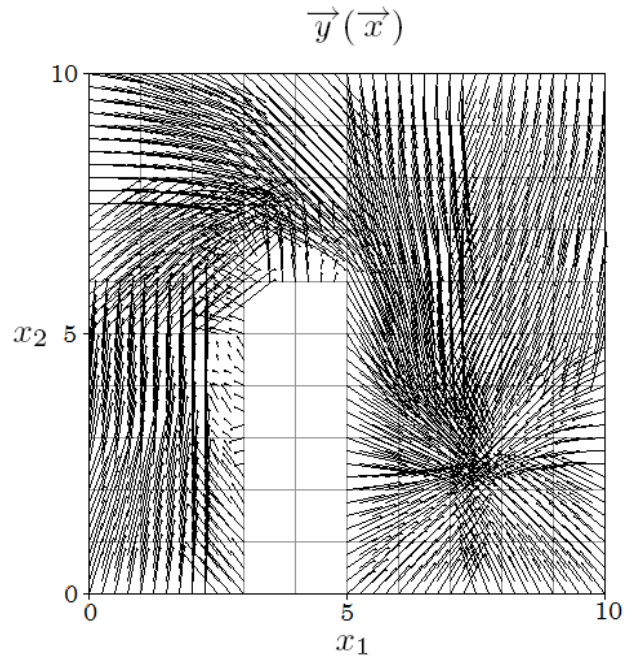


Figure 2: Action vector field learned by FLCs in a space with one obstacle = $\{(x_1, x_2) | 3 \leq x_1 \leq 5, 0 \leq x_2 \leq 6\}$ and with one goal = $\{(x_1, x_2) | 7 \leq x_1 \leq 8, 2 \leq x_2 \leq 3\}$

5. REFERENCES

- [1] A. Bonarini. Evolutionary learning of fuzzy rules: Competition and cooperation. *Fuzzy Modeling: Paradigms and Practice*, pages 265–284, 1996.
- [2] T. Kovacs, X. Llorà, K. Takadama, P. L. Lanzi, W. Stolzmann, and S. W. Wilson. Three architectures for continuous action. In *Lecture Notes in Artificial Intelligence (LNAI-4399)*, pages 239–257. Springer-Verlag, 2007.
- [3] P. L. Lanzi, D. Loiacono, S. W. Wilson, and D. Goldberg. XCS with computable prediction in continuous multistep environments. ILLiGAL Report 2005018, Genetic Algorithms Laboratory, University of Illinois at Urbana-Champaign, May 2005.
- [4] A. Parodi and P. Bonelli. A new approach to fuzzy classifier system. In *Proceedings of the Fifth International Conference on Genetic Algorithms*, pages 223–230, 1993.
- [5] M. Valenzuela-Rendón. The fuzzy classifier system: A classifier system for continuously varying variables. In *Proceedings of the Fourth International Conference in Genetic Algorithms*, pages 346–353, 1991.
- [6] M. Valenzuela-Rendón. Reinforcement learning in the fuzzy classifier system. *Expert Systems with Applications*, 14:237–247, 1998.