# MILCS: A Mutual Information Learning Classifier System

R. E. Smith
University College London
Department of Computer Science
+44 777 185 2565
robert.elliott.smith@gmail.com

Max Kun Jiang
University College London
Department of Computer Science
+44 20 7679 3715
m.jiang@cs.ucl.ac.uk

## ABSTRACT
This paper introduces a new variety of learning classifier system (LCS), called MILCS, with mutual information fitness.

## Categories and Subject Descriptors
I.2.6 [**Artificial Intelligence**]: Learning: Induction

## General Terms
Algorithms

## Keywords
Learning classifier systems, LCS, XCS, mutual information, information theory, structural learning

## 1. INTRODUCTION
This paper presents a new form of learning classifier system (LCS) that uses mutual information as its primary fitness feedback, in supervised learning settings. This system is called the mutual information learning classifier system (MILCS, pronounced “my LCS”).

## 2. THE ROLE OF MUTUAL INFORMATION
Imagine that the existing error in a classification system is an input signal to a communication channel. In this case, any additional element should play the role of an encoder for this signal, so that the error can be cancelled (similar to CCN [2]). Therefore, we find a firm theoretical foundation for using the mutual information (MI) as the fitness of this new element, through Shannon’s theorems [3].

## 3. THE MILCS PROCESS
MILCS operates in a fashion similar to XCS [1]. Given the above considerations, MILCS evaluates fitness and selects actions as follows.

1) a) Repeat the following for all rules:
   - i) Remove a rule from the population
   - ii) Update counters necessary for the calculation of the MI between that rule’s matching status, and the errors that would result from using the remaining rules
   - iii) Add that rule back to the population

   b) Do action selection based on the prediction values of all ‘mature’ rules.
2) Update actions (output) and prediction values of all rules based on the reward and previous actions (supervised learning).
3) Calculate and update the fitness of action rule set based on MI counters.

Copyright is held by the author/owner(s).
*GECCO’07*, July 7–11, 2007, London, England, United Kingdom
ACM 978-1-59593-697-4/07/0007.

## 4. FINAL COMMENTS
Preliminary results on the 6, 11, and 20 multiplexer problems with MILCS are promising, with respect to accuracy, speed, and explanatory power. While MILCS seems to scale slightly worse than XCS, this may not be an entirely fair comparison, since our preliminary results show that MILCS finds a smaller, more explanatory rule set (see Figure 1, which projects all rules in the final rule set that are used in action selection onto a plane, with generality-proportional size and approximate proportional overlap). We believe this effect is to be expected, given the theoretic basis of the mutual information fitness function.

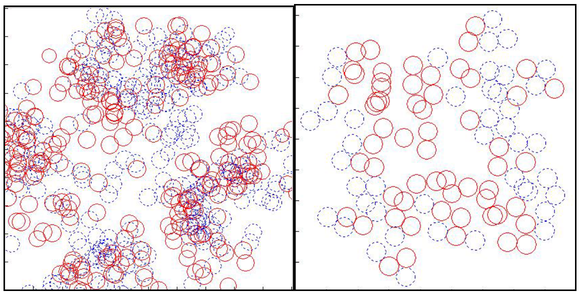

**Figure 1: Visualization of the final rule sets developed by XCS (left) and MILCS (right) on the 20 multiplexer.**

## 5. REFERENCES
[1] Butz, M. V.(203) Documentation of XCS+TS C-Code 1.2. *IlliGAL report 2003023*, University of Illinois at Urbana-Champaign.

[2] Fahlman, S. E., and Lebiere, C. (1990). The Cascade-Correlation learning algorithm. In *Advances in Neural Information Processing Systems 2*. Morgan Kaufmann.

[3] Shannon, C.E. (1948). A Mathematical Theory of Communication, *Bell System Technical Journal*, 27, pp. 379–423 & 623–656, July & October, 1948

[4] Smith, R. E. and Jiang, M. K. (submitted). A Learning Classifier System with Mutual-Information-Based Fitness