Learning Classifier System Equivalent with Reinforcement Learning with Function Approximation

Atsushi Wada
ATR NIS
2-2-2 Hikaridai, Seika-cho
Soraku-gun, Kyoto, Japan
wada@atr.jp

Keiki Takadama
Tokyo Institute of Technology
4259 Nagatsuta-cho
Midori-ku, Kanagawa, Japan
keiki@dis.titech.ac.jp

Katsunori Shimohara
ATR NIS
2-2-2 Hikaridai, Seika-cho
Soraku-gun, Kyoto, Japan
katsu@atr.jp

Abstract

We present an experimental comparison of the reinforcement process between Learning Classifier System (LCS) and Reinforcement Learning (RL) with function approximation (FA) method, regarding their generalization mechanisms. To validate our previous theoretical analysis that derived equivalence of reinforcement process between LCS and RL, we introduce a simple test environment named `Gridworld`, which can be applied to both LCS and RL with three different classes of generalization: (1) tabular representation; (2) state aggregation; and (3) linear approximation. From the simulation experiments comparing LCS with its GA-inactivated and corresponding RL method, all the cases regarding the class of generalization showed identical results with the criteria of performance and temporal difference (TD) error, thereby verifying the equivalence predicted from the theory.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning–Parameter learning.

General Terms

Algorithms, Design, Theory.

Keywords

Learning classifier systems, genetic-based machine learning, reinforcement learning, function approximation.

1. EXTENED ABSTRACT

Learning Classifier Systems (LCSs) are rule-based adaptive systems originally introduced by Holland [3], intended for a general framework to realize an intelligent behavior by combining two biologically inspired adaptive mechanisms – *learning* and *evolution* – with each essentially connecting to the fields of Reinforcement Learning (RL) and Evolutionary Computation (EC).

LCSs had long been regarded as weak in the theoretical sense due to their complicated mechanisms. That was until the invention of the Zeroth-level Classifier System (ZCS) [6] and the eXtended Classifier System (XCS) [7] ignited the recent development of LCSs. Especially for XCS, the rule discovery process has been heavily studied [1].

When focusing on the RL side, however, few works[4] have contributed to ground LCSs to the firm mathematical basis of RL[5], which derives mathematical convergence proof of learning or state generalization by means of the function approximation (FA) method. Applying such RL techniques to LCSs is expected to produce a variety of Genetic-based Machine Learning (GBML) methods ranging from the EC field to that RL field that would mutually potentate the adaptability of evolution and learning.

Towards our goal to build the foundations of LCS seamlessly connected to the basis of RL, in this paper, we first introduce our previous results deriving equivalence of the reinforcement process between an LCS and RL focusing particularly on the generalization mechanisms of both the LCS's rule condition generalization and RL's function approximation method. Next, we propose a simulation experiment that validates our theoretical results.

In detail, ZCS and XCS are compared with Q-learning with FA that derived: (a) the equivalence of the reinforcement process, and defines Reinforcement learning-based ZCS (RZCS), which is identical to ZCS but also satisfies the equivalence conditions; and (b) the inconsistency of the reinforcement process between XCS and Q-learning with FA, which derived Reinforcement learning-based XCS (RXCS) that is based on XCS but modified to be consistent with Q-learning with FA.

To validate this theoretical results, we conducted validation experiments to compare: (1) RZCS and Q-learning with FA; and (2) RXCS and Q-learning with FA, both regarding the approximated function designs of: (I) tabular representation; (II) state-aggregation; and (III) linear approximation. For these comparisons, the `Gridworld` environment was introduced, which is applicable to LCS with ternary representation and conformable to the approximated function design for Q-learning with FA. The empirical results showed identical performance and error between LCS and RL, thereby validating our previous work that derived the equivalence between LCS's reinforcement process and Q-learning with FA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
GECCO'05 June 25–29, 2005, Washington, DC, USA.
Copyright 2005 ACM 1-59593-097-3/05/0006 ...$5.00.

Acknowledgements

This research was conducted as part of 'Research on Human Communication' with funding from the National Institute of Information and Communications Technology, and supported in part by a Grant-in-Aid for Scientific Research (B) (No. 17360424) of Ministry of Education, Culture, Sports, Science and Technology (MEXT).

2. REFERENCES

[1] Butz, M., Kovacs, T., Lanzi, P.L., Wilson, S.W.: Toward a theory of generalization and learning in xcs. IEEE Transactions on Evolutionary Computation **8** (2004) 28–46

[2] M. V. Butz, D. E. Goldberg, and P. L. Lanzi. Bounding learning time in XCS. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2004)*, 2004.

[3] J. H. Holland. *Adaptation in Natural and Artifical Systems*. The University of Michigan Press, Michigan, 1975.

[4] P. L. Lanzi. Learning classifier systems from a reinforcement learning perspective. *Soft Computing*, 6:162–170, 2002.

[5] R. Sutton and A. Barto. *An introduction to reinforcement learning*. MIT Press, Cambridge, MA., 1998.

[6] S. W. Wilson. ZCS: A zeroth level classifier system. *Evolutionary Computation*, 2(1):1–18, 1994.

[7] S. W. Wilson. Classifier fitness based on accuracy. *Evolutionary Computation*, 3(2):149–175, 1995.