Dual-structured Classifier System Mediating XCS and Gradient Descent based Update

Atsushi Wada
National Institute of Information and Communication Technology / ATR Cognitive Information Science Laboratories
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
Doshisha University / ATR Cognitive Information Science Laboratories
2-2-2 Hikaridai, Seika-cho
Soraku-gun, Kyoto, Japan
katsu@atr.jp

ABSTRACT

Aiming at building an LCS with reinforcement process consistent with the gradient descent update of RL while utilizing XCS's accuracy-based rule discovery process, this paper presents *Dual-structured Classifier System* (DCS), which processes the gradient descent based update and XCS style update in parallel. DCS is evaluated on the popular testbed problems for LCSs with three types of bucket-brigade algorithms, Q-bucket-brigade, implicit-bucket-brigade, and residual-bucket-brigade each combined with XCS's original bucket-brigade. Empirical results are provided that shows DCS performing better than ZCS with the same optimal level of XCS, while the consistency with RL's gradient descent update enables to apply the results of theoretical analyses in RL field including rigorous conditions for the convergence of DCS's reinforcement process.

Categories and Subject Descriptors

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

General Terms

Algorithms, Design.

Keywords

Learning classifier systems, XCS, reinforcement learning, function approximation.

1. EXTENDED ABSTRACT

The relationship between Learning Classifier System (LCS) [4] and Reinforcement Learning (RL) [8] has been regarded as one of the essential issue to be clarified, which identifies the nature of LCS's learning process and to strengthen its theoretical basis. Comparison between the *bucket-brigade* algorithms in LCS and the concept of *Temporal Difference* (TD) in RL were carried out [3, 7] due to the history of both sharing the concept of *reinforcement* that originates in psychology of animal learning.

The main concern in LCS field, in the meanwhile, had been *generalization* since XCS [14] was introduced, which became the currently mainstream model and was analyzed in theoretical aspect for its *accuracy-based* rule discovery process [1, 2, 5]. Formal analysis of LCS's reinforcement process, together with generalization issue has become essential in this context.

We have approached to this issue by focusing on Function Approximation method, a common generalization technique in RL, to be compared with LCS's reinforcement process with rule condition generalization. We firstly compared the rule reinforcement process of Zeroth-level Classifier System (ZCS) [13] and RL with function approximation method in [12]. This analysis revealed that ZCS's bucket-brigade is consistent with a popular gradient descent based update in RL with function approximation by suppressing the rule discovery and adjusting ZCS's parameters, and named such model Reinforcement learning based ZCS (RZCS). This consistency enabled to apply numerous theoretical results for the gradient descent update published in RL field. One of such example is to carry out a rigorous discussion on the conditions for convergence, which we presented in [11].

On the other hand, the analysis in the same manner applied to XCS showed that the accuracy-based update of XCS is critically inconsistent with the gradient descent update of RL, while alternatively introduced XCS's variant, RXCS, whose bucket-brigade modified to become consistent with the gradient descent update turned out to perform seriously worse than original XCS [10].

Therefore, our objective is to build an LCS with a reinforcement process consistent with the gradient descent update of RL while utilizing XCS's accuracy-based rule discovery process. In this paper, a newly designed LCS, *Dual-structured Classifier System* (DCS) is presented[1], which in-

GECCO'06 July 8–12, 2006, Seattle, Washington, USA.

[1]Preliminary idea of DCS is originally proposed in our previous work [9], which does not include any empirical results

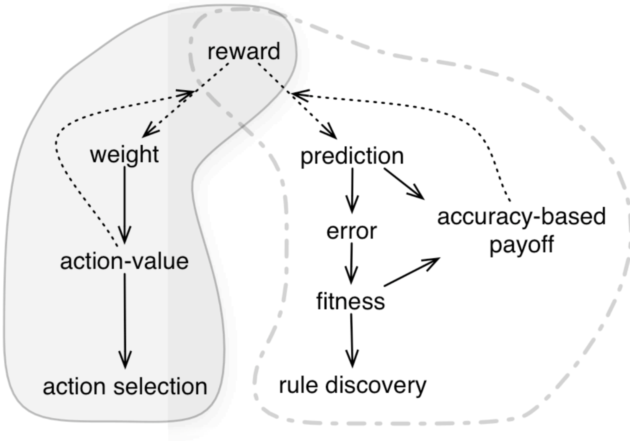

Figure 1: The relationship between classifier attributes in DCS.

cludes both the reinforcement process consistent with Q-learning with linear FA and the rule discovery process identical with XCS. This architecture enables to merge the strong theoretical basis of RL and the effective GA of XCS. Figure 1 describes the basic idea of merging the gradient-based update, the left side process, and XCS's accuracy-based process, the right side process[2].

Firstly, two candidates models, DCS(RZCS-XCS) and DCS(RXCS-XCS) are proposed, which merges the gradient descent based update of RZCS and RXCS with XCS's update to be processed in parallel. From the preliminary experiment, DCS(RXCS-XCS) showed better performance compared with DCS(RZCS-XCS). Next, DCS is evaluated on the popular test-bed problems for LCSs regarding three bucket-brigade algorithms: Q-bucket-brigade, implicit-bucket-brigade and residual-bucket-brigade. Compared with ZCS and XCS, DCS with residual-bucket-brigade on multi-step maze environments showed better performance than ZCS and is at the same optimal level with XCS. On the multiplexer problems, DCS showed competitive performance compared with XCS, which requires the accurate generalization of the classifier conditions.

These results showed that DCS not only assures the consistency with RL with FA for its reinforcement process, but also works effectively in practice regarding both the control problem and the concept learning problems. This shows that the accuracy-based rule discovery process is effectively working within the DCS framework.

2. ACKNOWLEDGMENTS

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).

nor discussions related to DCS(RZCS-XCS), the implicit-bucket-brigade, and the experiments on `Maze(4,5)`.

[2]This graphical form, which represents the relationship between the rule attributes is originally proposed by [6].

3. REFERENCES

[1] M. V. Butz, D. E. Goldberg, and P. L. Lanzi. Bounding learning time in XCS. In e. a. K. Deb, editor, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2004)*, pages 739–750. Springer, 2004.

[2] M. V. Butz, T. Kovacs, P. L. Lanzi, and S. W. Wilson. Toward a theory of generalization and learning in XCS. *IEEE Transactions on Evolutionary Computation*, 8(3):28–46, 2004.

[3] M. Dorigo and H. Bersini. A comparison of Q-learning and classifier systems. In D. Cliff, P. Husbands, J. A. Meyer, and S. W. Wilson, editors, *Proceedings of From Animals to Animats, Third International Conference on Simulation of Adaptive Behavior*, pages 248–255, Cambridge, MA., 1994. MIT Press.

[4] J. H. Holland. Escaping brittleness: the possibilities of general-purpose. *Machine Learning, an artificial intelligence approach*, 2:593–623, 1986.

[5] T. Kovacs. Evolving optimal populations with xcs classifier systems. Technical Report CSRP-96-17, University of Birmingham, School of Computer Science, October 1996.

[6] T. Kovacs. *Learning classifier systems, from foundations to applications*, chapter Strength or Accuracy? Fitness Calculation in Learning Classifier Systems, pages 143–160. Springer-Verlag, London, UK, 2000.

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

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

[9] A. Wada, K. Takadama, K. Shimohara, and O. Katai. DCS: Learning classifier system with convergence and generalization. In *5th Asia-Pasific Conference on Simulated Evolution and Learning (SEAL'04)*, pages CD–ROM # 199, 2004.

[10] A. Wada, K. Takadama, K. Shimohara, and O. Katai. Is gradient descent method effective for XCS? analysis of reinforcement process in XCSG. In *Seventh International Conference on Learning Classifier Systems (IWLCS2004)*, 2004.

[11] A. Wada, K. Takadama, K. Shimohara, and O. Katai. Counter example for q-bucket-brigade under prediction problem. In *The Eighth International Workshop on Learning Classifier Systems (IWLCS2005)*, pages 24–29, 2005.

[12] A. Wada, K. Takadama, K. Shimohara, and O. Katai. Learning classifier system equivalent with reinforcement learning with function approximation. In *The Eighth International Workshop on Learning Classifier Systems (IWLCS2005)*, pages 24–29, 2005.

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

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