An Analysis of Feature Selection and Reward Function for Model-Based Reinforcement Learning

Shitian Shen, Chen Lin, Behrooz Mostafavi, Tiffany Barnes, and Min Chi

North Carolina State University, Raleigh, NC 27695, USA
{sshen,clin12,bzmostaf,tmbarnes,mchi}@ncsu.edu

1 Introduction

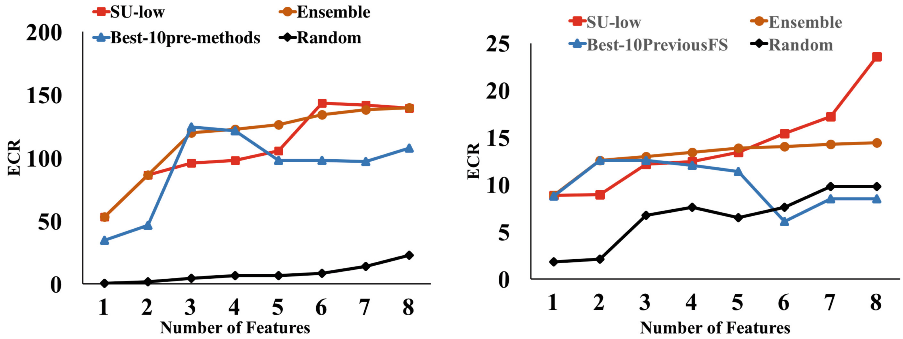
In this paper, we propose a series of correlation-based feature selection methods for dealing with high dimensionality in feature-rich environments for model-based Reinforcement Learning (RL). Real world RL tasks usually involve high-dimensional feature spaces where standard RL methods often perform badly. Our proposed approach adopts correlation among state features as a selection criterion. The effectiveness of the proposed methods are compared against previous methods referred as 10PreviousFS [2] using the data from an intelligent logic tutor called Deep Thought (DT) [1]. We evaluated the effectiveness of different feature selection methods by expected cumulative reward (ECR) [3], considering two types of reward: immediate and delayed. Our results show that our proposed methods significantly outperform previous feature selection methods with both types of rewards. Moreover, the "best" policy induced using immediate reward differs significantly from that induced from delayed reward.

2 Methodology, Experiments, and Results

Methodology: The proposed feature selection framework forwardly select the feature based on the correlation between the feature and current selected feature set. Particularly we applied three common used correlation metrics Chi-square (CHI), Information gain (IG) and Symmetrical uncertainty (SU) for measuring the correlation among features. When considering the three correlation metrics, we face a simple decision: should we select the next feature that is the most **correlated** or **uncorrelated** to the current selected features? However, for model-based RL, the answer is not straightforward. We used both high and low correlation on three correlation metrics thus resulted in six methods: CHI-high, IG-high, SU-high, CHI-low, IG-low, and SU-low. We will compare these six methods with 10PreviousFS in [2].

Experiments: We applied RL [3] to induce policy given the action set, rewards and state features. Here we focus on one simply tutorial decision: once a tutor determines the problem to be completed, should the tutor show the student how to solve the next problem directly–worked example (WE), or should it ask the

A. Micarelli et al. (Eds.): ITS 2016, LNCS 9684, pp. 504–505, 2016.
DOI: 10.1007/978-3-319-39583-8



(a) Four type of methods (b) Correlation-based method

Fig. 1. Feature selection methods for immediate reward.

student to solve the problem–problem solving (PS)? In addition, we designed two types of rewards: Immediate vs. Delayed. The former reflects students' performance level by level while The latter determines the students' performance across all levels. Besides, 134 state features were extracted representing a cumulative statistical measure of students' behavior and context information.

Result: Among the three correlation-based methods, our results showed that SU-based methods outperforms CHI and IG based ones. Moreover, Fig. 1a, b show that (1) SU-based > ensemble > the best of 10PreviousFS > random; and (2) the ECR of Immediate policies much higher than that of Delayed policies. This is most likely because of the issue of credit assignment. The more we delay success measures from a series of sequential decisions, the more difficult it becomes to identify which of the decision(s) in the sequence are responsible for our final success or failure.

3 Conclusion and Future Work

In this paper, we proposed six correlation-based feature selection methods for model-based RL and showed that they are more effective than the ensemble method and 10PreviousFS. In future work, we are applying correlation-based feature selection methods on other data sets. Currently we are implementing the optimal Immediate and Delayed policies into DT to experimentally evaluate their performance.

References

1. Mostafavi, B. , Liu, Z. , Barnes, T.: Data-Driven Proficiency Profiling. EDM (2015)
2. Chi, M., VanLehn, K., Litman, D., Jordan, P.: Inducing effective pedagogical strategies using learning context features. In: De Bra, P., Kobsa, A., Chin, D. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 147–158. Springer, Heidelberg (2010)
3. Tetreault, J.R., Litman, D.J.: A reinforcement learning approach to evaluating state representations in spoken dialogue systems. Speech Commun. **50**, 683–696 (2008)