

# Knowledge-Based Reinforcement Learning for Data Mining

Daniel Kudenko and Marek Grzes

Department of Computer Science, University of York, York YO105DD, UK  
{kudenko, grzes}@cs.york.ac.uk

## 1 Extended Abstract

Data Mining is the process of extracting patterns from data. Two general avenues of research in the intersecting areas of agents and data mining can be distinguished. The first approach is concerned with mining an agent's observation data in order to extract patterns, categorize environment states, and/or make predictions of future states. In this setting, data is normally available as a batch, and the agent's actions and goals are often independent of the data mining task. The data collection is mainly considered as a side effect of the agent's activities. Machine learning techniques applied in such situations fall into the class of supervised learning. In contrast, the second scenario occurs where an agent is actively performing the data mining, and is responsible for the data collection itself. For example, a mobile network agent is acquiring and processing data (where the acquisition may incur a certain cost), or a mobile sensor agent is moving in a (perhaps hostile) environment, collecting and processing sensor readings. In these settings, the tasks of the agent and the data mining are highly intertwined and interdependent (or even identical). Supervised learning is not a suitable technique for these cases. Reinforcement Learning (RL) enables an agent to learn from experience (in form of reward and punishment for explorative actions) and adapt to new situations, without a teacher. RL is an ideal learning technique for these data mining scenarios, because it fits the agent paradigm of continuous sensing and acting, and the RL agent is able to learn to make decisions on the sampling of the environment which provides the data. Nevertheless, RL still suffers from scalability problems, which have prevented its successful use in many complex real-world domains. The more complex the tasks, the longer it takes a reinforcement learning algorithm to converge to a good solution. For many real-world tasks, human expert knowledge is available. For example, human experts have developed heuristics that help them in planning and scheduling resources in their work place. However, this domain knowledge is often rough and incomplete. When the domain knowledge is used directly by an automated expert system, the solutions are often sub-optimal, due to the incompleteness of the knowledge, the uncertainty of environments, and the possibility to encounter unexpected situations. RL, on the other hand, can overcome the weaknesses of the heuristic domain knowledge and produce optimal solutions. In the talk we propose two techniques, which represent first steps in

the area of knowledge-based RL (KBRL). The first technique [1] uses high-level STRIPS operator knowledge in reward shaping to focus the search for the optimal policy. Empirical results show that the plan-based reward shaping approach outperforms other RL techniques, including alternative manual and MDP-based reward shaping when it is used in its basic form. We showed that MDP-based reward shaping may fail and successful experiments with STRIPS-based shaping suggest modifications which can overcome encountered problems. The STRIPS-based method we propose allows expressing the same domain knowledge in a different way and the domain expert can choose whether to define an MDP or STRIPS planning task. We also evaluated the robustness of the proposed STRIPS-based technique to errors in the plan knowledge. In case that STRIPS knowledge is not available, we propose a second technique [2] that shapes the reward with hierarchical tile coding. Where the Q-function is represented with low-level tile coding, a V-function with coarser tile coding can be learned in parallel and used to approximate the potential for ground states. In the context of data mining, our KBRL approaches can also be used for any data collection task where the acquisition of data may incur considerable cost. In addition, observing the data collection agent in specific scenarios may lead to new insights into optimal data collection behaviour in the respective domains. In future work, we intend to demonstrate and evaluate our techniques on concrete real-world data mining applications.

## References

1. Grzes, M., Kudenko, D.: Plan-based Reward Shaping for Reinforcement Learning. In: Fourth International IEEE Conference on Intelligent Systems, vol. 2, pp. 22–29 (2008)
2. Grzes, M., Kudenko, D.: Learning potential for reward shaping in reinforcement learning with tile coding. In: Proceedings AAMAS 2008 Workshop on Adaptive and Learning Agents and Multi-Agent Systems (ALAMAS-ALAg 2008), pp. 17–23 (2008)