
Towards Interactive Relational Reinforcement Learning of Concepts

Matthias Nickles
Department of Computer Science
Technical University of Munich
Boltzmannstr. 3, D-85748 Garching, Germany
nickles@cs.tum.edu

Achim Rettinger
Institute AIFB
Karlsruhe Institute of Technology
76128 Karlsruhe, Germany
rettinger@kit.edu

Abstract

We present a framework for the interactive machine learning of denotational concept semantics in communication between humans and artificial agents. The capability of software agents and robots to learn how to communicate verbally with human users is obviously highly useful in several real-world applications. Whereas the large majority of existing approaches to the machine learning of word sense and other language aspects focusses on learning using text corpora, our framework allows for the interactive learning of concepts in a dialog between human and agent, using recent advances in the area of Relational Reinforcement Learning.

Keywords: *Concept Learning, Statistical Relational Learning, Relational Reinforcement Learning, Natural Language Processing*

1 Introduction and Related Work

The large majority of existing approaches to the machine learning of human communication focusses on learning the semantics of words and other language constructs in a textual context, typically from large text corpora. These approaches can neither take into account the dynamic behavioral context of the word use nor do they learn in interaction with the human user who conceptualizes the respective word. Arguably, an important reason for this lack is that relevant approaches to machine learning (most important: Reinforcement Learning (RL)) typically still do not cope well with the complexity of high-level, symbolic interaction. Our approach aims at this issue by seamlessly combining a state-of-the-art approach to Relational Reinforcement Learning with rich yet computationally effective logic reasoning capabilities. While our implementation currently focusses “only” on the interactive learning of the denotational semantics of uninflected concept names, our approach already allows for the integration of complex bodies of formal rules (provided by domain experts) into the learning process. Furthermore, our approach aims at the learning of context-sensitive policies for the incremental *querying* of concepts from human users, i.e., learning of sequences of questions and answer suggestions with which the learning agent tries to determine the semantics of a word in a dialogue with a human user. Thus, our approach has also a strong pragmatic aspect. Initial empirical results indicate the applicability of our approach and directions of future research.

We are not aware of any other approaches to interactive Relational RL of concept semantics. However, various approaches to supervised and unsupervised disambiguation of word sense exist (see, e.g., [2, 9]), including approaches in the area of Statistical Relational Learning (e.g., [12]). Conceptually related to our approach are also approaches to grounded language acquisition, where the meaning of words is grounded in observable environment states or events (e.g., [8, 4]). Most of these do not combine grounding with a dialogic learning setup as we do, although some (e.g., [8]) enhance the learning process with ask/tell (but do not learn to act in such dialogs). Pioneering work in the area of computational emergence of language semantics in interactive settings was done by

Luc Steels [6]. Although technically only remotely related, his work had a significant influence on our and many other contemporary approaches to semantics learning. On the foundational machine learning level, in [18] a simulator of the environment is employed as a stochastic sample generator for Relational RL. However, this learning approach is goal-based and does not approximate a value function as in our case but learns policy representations. [14] proposes an approach where the performance of Relational Q-Learning is improved at runtime with plans generated using learned probabilistic models of the environment, and [7] combines Hierarchical Reinforcement Learning with planning. Both approaches are remotely related to our use of the Event Calculus in the learning process (as outlined below). [10, 11] and others integrate RL with programs in the Golog action language, which is based on the Situation Calculus (a close relative of the Event Calculus). [16, 15] provide logical formalisms for the representation of Markov Decision Processes.

2 Framework Outline

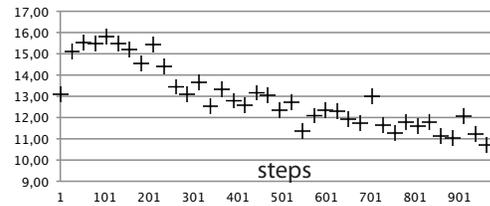
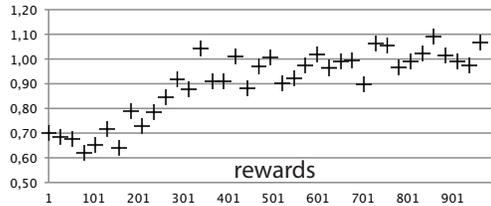
Being part of the increasingly popular area of Statistical Relational Learning (SRL) [17], Relational Reinforcement Learning [1] uses relational representations of Markov states and actions. This allows for a rich formal characterization of complex domains (like in NLP) whose complex structural properties would otherwise be inaccessible to RL. Various approaches to Relational RL exist, including [3, 14]. Our framework differs from existing Relational RL approaches mainly by its human-agent interaction component and by the use of a highly efficient approach to the employment of formal rules as background knowledge, namely an Answer-Set Programming (ASP) implementation of the Event Calculus (EC) [5]. The basic learning algorithm is a variant of Relational Q-Learning [13]. Besides the general benefit of Relational RL for NLP (i.e., the ability to represent and learn in structurally very complex domains), the main advantage of this hybrid approach is that it seamlessly integrates logical reasoning and RL. Concretely, the use of the EC significantly simplifies the modeling of logical conditions, context and effects for/of agent and human actions, and the use of ASP (instead of the traditional use of Prolog in Relational RL) facilitates a computationally very efficient computation of reasoning tasks in the Event Calculus.

3 Example Result

In the following, we outline a simple example instantiation of our framework, situated in a blocks world domain. In each learning episode, the agent asks a number of questions and the human user answers these questions (although not necessarily correctly). The agent’s goal is to find out the denotation of the term “nice building”, as conceptualized by the human user. In our initial experiments we simulated the behavior of the human user to reduce the manual training effort. The simulated human user wants the agent to build a “nice building” which for the user means arranging four blocks next to each other on the table. Every reinforcement learning epoch starts with blocks randomly placed on top of each other. Then, in each time step the agent can rearrange one block or in turn ask questions like `ask(isHigher)` (to build a higher building), `ask(isAonB)` (to put block A on B). Potential user responses can be `tell(yes)`, `tell(no)`, `tell(dontKnow)`, and others. For a positive answer the agent receives a small reward (0.1) and a large reward (1.0) if it built the desired structure. The maximum number of steps per epoch is 20. The overall goal is on the one hand to learn the right questions in the right sequence and on the other hand to learn the meaning of “nice building”.

We ran this setup for 6 trials with 1000 epochs each, and monitored the actions, rewards and number of steps performed in each epoch before “nice building” was found. The figure below shows the average over all trials after every 26 epochs for the *reward* (left) and the *steps performed* (right). The results show that in the end the agent learns to build a “nice building” in every epoch with a decreasing number of steps per epoch. The 95% confidence intervals for the reward were around ± 0.5 also indicating a statistically significant learning success. Analyzing the actions performed suggests, that the agents learns to avoid questions which are answered by `tell(dontKnow)`.

While the described experiment demonstrates basic functionality of our framework, our current work aims at significantly more complex tasks, in order to show that a generalized communication strategy for varying human users with different denotational concept semantics can be learned efficiently.



Acknowledgments

This work is partially supported by Deutsche Forschungsgemeinschaft (DFG), grant BR609151.

References

- [1] S. Dzeroski, L. De Raedt, H. Blockeel: Relational reinforcement learning. Procs. ICML'98. Morgan Kaufmann, 1998.
- [2] R. Navigli: Word Sense Disambiguation: A Survey. ACM Computing Surveys, 41(2), 2009.
- [3] K. Driessens: Relational Reinforcement Learning. PhD thesis, Department of Computer Science, Katholieke Universiteit Leuven, 2004.
- [4] D. L. Chen, R. J. Mooney: Panning for Gold: Finding Relevant Semantic Content for Grounded Language Learning. In Procs. MSLSP, 2011.
- [5] T.-W. Kim, J. Lee, R. Palla: Circumscriptive event calculus as answer set programming. Procs. IJCAI'09, 2009.
- [6] L. Steels: Grounding Symbols through Evolutionary Language Games. In A. Cangelosi, D. Parisi: Simulating the Evolution of Language. Springer, 2001.
- [7] M. R. K. Ryan: Hierarchical Reinforcement Learning: A Hybrid Approach. PhD thesis, University of New South Wales New South Wales, Australia, 2002.
- [8] W. Kerr, P. R. Cohen, Y.-H. Chang: Learning and Playing in Wubble World. Procs. AIIDE, 2008.
- [9] R. Mihalcea: Unsupervised Large Vocabulary Word. Sense Disambiguation with Graph-based. Algorithm for Sequence Data Labeling. Procs. of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT'05), 2005.
- [10] A. Finzi, T. Lukasiewicz: Adaptive multi-agent programming in GTGolog. Procs. of the 29th Annual German Conference on Artificial Intelligence (KI 2006), 2006.
- [11] D. Beck, G. Lakemeyer: Reinforcement Learning for Golog Programs. Procs. Workshop on Relational Approaches to Knowledge Representation and Learning, 2009.
- [12] L. Specia et al: Word Sense Disambiguation Using Inductive Logic Programming. In Selected papers from the 16th International Conference on Inductive Logic Programming, Springer, 2007.
- [13] Ch. Rodrigues, P. Gerard, C. Rouveirol: Relational TD Reinforcement Learning, Procs. EWRL'08, 2008.
- [14] T. Croonenborghs, J. Ramon, M. Bruynooghe: Towards informed reinforcement learning. Procs. of the Workshop on Relational Reinforcement Learning at ICML'04, 2004.
- [15] K. Kersting, L. De Raedt: Logical Markov decision programs. Procs. IJCAI'03 Workshop on Learning Statistical Models of Relational Data, 2003.
- [16] C. Boutilier, R. Reiter, B. Price: Symbolic dynamic programming for First-order MDP's. Procs. IJCAI-01, Morgan Kaufmann Publishers, 2001.
- [17] L. Getoor, B. Taskar, eds. Introduction to Statistical Relational Learning. MIT Press, 2007.
- [18] A. Fern, S. Yoon, R. Givan: Reinforcement Learning in Relational Domains: A Policy-Language Approach. In L. Getoor, B. Taskar, eds. Introduction to Statistical Relational Learning. MIT Press, 2007.