Interactive Relational Reinforcement Learning of Concept Semantics (Extended Abstract)

Matthias Nickles[1], **Achim Rettinger**[2]
[1]INSIGHT Centre for Data Analytics and Discipline of Information Technology
National University of Ireland, Galway, Ireland
email:matthias.nickles@deri.org
[2]Institute AIFB Karlsruhe Institute of Technology, Karlsruhe, Germany
email:rettinger@kit.edu

Abstract

We propose a novel approach to the machine learning of formal word sense, learned in interaction with human users using a new form of Relational Reinforcement Learning. The envisaged main application area of our framework is human-machine communication, where a software agent or robot needs to understand concepts used by human users (e.g., in Natural Language Processing, HCI or Information Retrieval). In contrast to traditional approaches to the machine learning and disambiguation of word meaning, our framework focuses on the interactive learning of concepts in a dialogue with the user and on the integration of rich formal background knowledge and dynamically adapted policy constraints in the learning process, which makes our approach suitable for dynamic interaction environments with varying word usage contexts.

Keywords: *Concept Learning, Statistical Relational Learning, Planning, Word-Sense Disambiguation, Dialogue Systems*

Introduction

This extended abstract summarizes (Nickles and Rettinger 2014). In this article, we present a framework for the interactive (dialogical) machine learning of formal word semantics in communication between human users and artificial agents. The capability of software agents and robots to learn how to communicate verbally with human users, or, more specifically, to learn what a certain name used by the human interaction partner formally means, and to disambiguate word-sense if necessary, is obviously highly useful. However, most existing approaches to the machine learning of human communication semantics focus on learning or disambiguating the semantics of words in a static textual context (learning from large text corpora) only. Also, they do not learn in interaction with the human user who mentally conceptualizes the respective word. Arguably, an important reason for this lack is that relevant approaches to machine learning (most important in our context: Reinforcement Learning (RL)) typically still do not cope very well in complex symbolic and interactive environments. Our approach tackles this problem by combining *Relational Reinforcement Learning* (Dzeroski, Raedt, and Driessens 2001; Driessens and Deroski 2004; Christophe Rodrigues 2008) with computationally efficient logic reasoning (Answer Set Programming) and formal planning. Despite targeting only the interactive learning of plain concept name semantics (without considering full sentences), our approach already allows for the use of complex formal background knowledge (provided by domain experts) into the semantics discovery process. Most importantly, our approach learns optimal context-sensitive policies for the dialogical querying of word semantics, that is, policies consisting of questions and answer suggestions with which the learning agent determines word semantics in a dialog with the user.

Related approaches to the learning and disambiguation of word sense include, e.g., (Navigli 2009; Mihalcea 2005) and also approaches in the area of Statistical Relational Learning (Specia et al. 2007) and approaches to RL in dialogue systems (Cuayhuitl 2009; Rieser and Lemon 2011). Conceptually related to our approach are also approaches to language acquisition (e.g., (Kerr, Cohen, and Chang 2008; Chen and Mooney 2011)). Most of these do not combine grounding with a dialogic learning setup as we do, although some (e.g., (Kerr, Cohen, and Chang 2008)) enhance the learning process with ask/tell actions (but do not learn action policies in our sense). Pioneering work in the area of computational emergence of language semantics in interactive settings includes (Steels 2002). In (Fern, Yoon, and Givan 2007), an environment simulator is used as a sample generator for Relational RL. However, this approach does not approximate a value function as in our case but learns policy representations. (Croonenborghs, Ramon, and Bruynooghe 2004; Driessens and Deroski 2004) integrate forms of planning-like guidance into Relational RL and (Ryan 2002) combines Hierarchical Reinforcement Learning with planning. Both approaches are remotely related to our use of planning and the Event Calculus during the learning process. Hierarchical RL has also been successfully used in spoken dialogue systems (Cuayhuitl 2009). Like our approach (but in a non-relational setting), (Knox and Stone 2011) and others propose interactive forms of RL with human feedback.

Framework Overview

Relational RL (Dzeroski, Raedt, and Driessens 2001; Driessens and Deroski 2004; Christophe Rodrigues 2008) uses relational representations of Markov states and actions. This allows for a rich formal characterization of complex do-

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

mains (like in NLP) whose structural properties would otherwise be inaccessible to RL. Our framework differs from existing approaches in this field mainly by its human-agent interaction component and by the use of an efficient logical reasoning approach, namely Answer Set Programming (ASP) with a variant of the Event Calculus (EC) (Kim, Lee, and Palla 2009). The basic learning algorithm is a variant of Relational Q-Learning (Christophe Rodrigues 2008; Driessens and Deroski 2004). Besides the general benefit of Relational RL (i.e., the ability to represent and learn in structurally complex domains), the main advantage of this hybrid approach is that it seamlessly integrates logical reasoning and RL for the modeling of logical constraints, goals, context and effects for/of agent and human actions.

The agent's goal is to discover the formal meaning of words, as conceptualized by the human user. In each relational reinforcement learning episode, the agent presents the user a proposal for the meaning of the unknown/ambiguous word and asks the user questions until the correct formal meaning has been found. Possible questions and answers are constrained by the reasoner (ASP solver) and dynamically adapted during interaction with the user. The agent receives rewards depending on the answer (essentially, the user awards the agent). The overall learning goals are to learn the right questions in the right sequence (generalized over multiple episodes) and of course also to learn the formal meaning of each individual concept.

Our approach is closely related to planning. The learning agent generates policy constraints and planning subgoals from interactions with the human user. The ASP solver ensures that at the end of a learning episode, these sub-goals are reached (if possible), by constraining the set of feasible action policies. If this does not already discover an optimal action sequence, reinforcement learning finds the optimal policy among candidate policies. As a further contribution, we introduce so-called "profiles" which allow the algorithm to adapt on-line to different interaction partners. This way, we can deal with moving learning targets in form of multiple different interaction partners.
Our experimental evaluation (using simulations of human users) shows that our rather general approach successfully works in combinatorially challenging domains, i.e., that it learns optimal interaction policies (including in a non-stationary domain), and that it requires only a relatively low number of learning episodes to do so.

References

Chen, D. L., and Mooney, R. J. 2011. Mooney: Panning for gold: Finding relevant semantic content for grounded language learning. In *In Procs. MLSLP*.

Christophe Rodrigues, Pierre Gerard, C. R. 2008. Relational td reinforcement learning. In *EWRL'08*.

Croonenborghs, T.; Ramon, J.; and Bruynooghe, M. 2004. Towards informed reinforcement learning. In *Proceedings of the ICML2004 Workshop on Relational Reinforcement Learning*, 21–26.

Cuayhuitl, H. 2009. *Hierarchical Reinforcement Learning for Spoken Dialogue Systems*. Ph.D. Dissertation, School of Informatics, University of Edinburgh.

Driessens, K., and Deroski, S. 2004. Integrating guidance into relational reinforcement learning. *Machine Learning* 57(3):271–304.

Dzeroski, S.; Raedt, L. D.; and Driessens, K. 2001. Relational reinforcement learning. *Machine Learning* 43(1/2):7–52.

Fern, A.; Yoon, S.; and Givan, R. 2007. Reinforcement learning in relational domains: A policy-language approach. In L. Getoor, B. T., ed., *Introduction to Statistical Relational Learning*. MIT Press.

Kerr, W.; Cohen, P.; and Chang, Y. H. 2008. Learning and playing in wubble world. In *Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*.

Kim, T.-W.; Lee, J.; and Palla, R. 2009. Circumscriptive event calculus as answer set programming. In Boutilier, C., ed., *IJCAI*, 823–829.

Knox, W. B., and Stone, P. 2011. Augmenting reinforcement learning with human feedback. In *Procs. ICML 2011 Workshop on New Developments in Imitation Learning*.

Mihalcea, R. 2005. Unsupervised large-vocabulary word sense disambiguation with graph-based algorithms for sequence data labeling. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, HLT '05, 411–418. Stroudsburg, PA, USA: Association for Computational Linguistics.

Navigli, R. 2009. Word sense disambiguation: A survey. *ACM Comput. Surv.* 41(2):10:1–10:69.

Nickles, M., and Rettinger, A. 2014. Interactive relational reinforcement learning of concept semantics. *Machine Learning* 94(2):169–204.

Rieser, V., and Lemon, O. 2011. Reinforcement learning for adaptive dialogue systems: a data-driven methodology for dialogue management and natural language generation. Springer.

Ryan, M. R. K. 2002. *Hierarchical Reinforcement Learning: A Hybrid Approach*. Ph.D. Dissertation.

Specia, L.; Stevenson, M.; Graas, M.; and Nunes, V. 2007. Learning expressive models for words sense disambiguation. In *In Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL*, 41–48.

Steels, L. 2002. Grounding Symbols through Evolutionary Language Games. In Cangelosi, A., and Parisi, D., eds., *Simulating the Evolution of Language*. London: Springer. chapter 10, 211–226.