

RoboCanes

RoboCup 3D Simulation League Team Description Paper

2017

Lloyd Beaufils, Chloe Arluck, Joseph Masterjohn, Pedro Peña, Kyle Poore,
Andreas Seekircher, and Ubbo Visser

Department of Computer Science
University of Miami
1365 Memorial Drive, Coral Gables, FL, 33146 USA
`robocanes@cs.miami.edu`

1 Introduction

The team, **RoboCanes**, was incepted in January 2010 at the University of Miami in the USA under the supervision of Dr. Ubbo Visser. The team has designed, developed, and implemented its autonomous agent framework and software from the ground up, and it has been evolved to a flexible research platform, that has contributed to many publications over the years. In 2017, the team is lead by Lloyd Beaufils, a new PhD student who started as a teammember in early 2017. The other team members are both part of the 3D Simulation Team but also our robotics team in SPL. However, each one of the team members also contributed to the Simulation League team.

RoboCanes team members are Lloyd Beaufils, Chloe Arluck, Joseph Masterjohn, Pedro Peña, Kyle Poore, Andreas Seekircher, and Ubbo Visser. Lloyd, Chloe, Joseph, Pedro and Kyle are PhD students, Andreas and Ubbo are senior team members. Andreas is a currently working in industry and Ubbo is a faculty member at the Computer Science Department of the University of Miami.

Lloyd is interested in reinforcement learning and AI in general and has extended experience in game development. He is currently working on new kicks and a new passing strategy. Chloe's research is focussed on vision in the SPL. Kyle focuses on team communication, especially audio communication on the NAO robot. Joseph is a member of the team working on algorithm development (also on team communication) and also contributes to the SPL. Pedro concentrates on motions, mainly in the SPL but also in simulation. He is currently developing a new kicking engine. Andreas comes from the SSL team B-Smart and has done studies for his MS and PhD Thesis on the physical NAO. He has published multiple papers on both simulated robots and physical robots. One paper about his MS thesis entitled "Entropy-based active vision for a humanoid soccer robot" received the best paper award at RoboCup 2010 in Singapore. He is interested in motions and motion learning on both the physical and the simulated NAO. Ubbo has participated with the RoboCup community in various functions and teams since 2000. He started with (and currently still is in) the Soccer Simulation League. He then founded more teams from the Bremen University in Germany (together with Thomas Röfer): The SSL team B-Smart and

the German Team in the 4LL (also together with H.-D. Burkhard, Humboldt University Berlin). Since 2008, he is affiliated with the University of Miami in the USA where he founded the RoboCanes team.

The rest of the paper is organized as follows. First, we describe our research interests and planned activities in Section 2. Second, in Section 3, we describe relevant work and a list of our contributions to the 3D Soccer Simulation League.

2 Research Interests and Planned Activities

Our main research activities are in the areas of: (1) behavior/situation recognition, (2) real-time knowledge representation, (3) signal processing, (4) prediction, (5) motion optimization, (6) control, and (7) real-time/approximate algorithm development. We divide these research areas into two groups: (1) the immediate/short-term activities that is planned to be addressed before RoboCup 2016, and (2) the long-term activities beyond the competition.

Besides working on the low-level skills that are described in section 2.1 we like to apply plan recognition methods in order to bring valuable knowledge into the behavior decision process. These efforts are presented in section 2.2. The real-time knowledge representation formalism is presented in section 2.3. The application of learning methods to interpret low-level skills as well as higher-level behaviors is another research direction addressed by our team presented in section 2.4.

2.1 Humanoid Walking Engine and Special Actions

Since our team participates in the 3D Soccer Simulation and the SPL, it is important to merge research efforts for the separate leagues. The 3D Soccer Simulation League can benefit from the experiences of the real robot humanoid league. Later on, a sufficiently realistic simulation (e.g. the Webots simulator that is tied with the physical NAO) can be used to ease certain aspects during the development of real robots by (pre-) learning some skills or testing different settings in the simulation that might be disadvantageous (and costly) for real robots. We use the same implementations for both leagues as far as possible and try to find methods to adapt motions to the behavior of different robots. This is an important step for controlling the different heterogeneous robot types in the 3D Soccer Simulation, as well as the slightly different physical robots (due to hardware tolerance, different calibration, etc.) in the SPL.

Another goal we pursue is creating a workflow for quickly generating reliable motions, preferably with inexpensive and accessible hardware. Our hypothesis is that using Microsofts Kinect sensor in combination with modern optimization algorithms can achieve this objective. We produced four complex and inherently unstable motions and then applied three contemporary optimization algorithms (CMA-ES, xNES, PSO) to make the motions robust; we performed 900 experiments with these motions on a 3D simulated NAO robot with full physics.

We described the motion mapping technique, compared the optimization algorithms, and discussed various basis functions and their impact on the learning performance. Our conclusion is that there is a straightforward process to achieve complex and stable motions in a short period of time [SSAV13, SV16]. These result can be used as seeds for the optimization on a physical NAO. We have used these methods successfully on a NAO humanoid robot to imitate motion trajectories from human motions.

We created several motions as so called “special actions”, such as “getting up” or “kicking the ball”. These motions are defined by a sequence of keyframes containing joint angles. We use the code that generates these motions in both leagues, the 3D simulation and SPL. However, the angles that define the motions need to be adapted for different robots. In order to create a motion for the simulated robot, we need a first version of the intended behavior. This can be slow and unreliable, but it provides the initial parameters for the fine tuning in a second phase. We applied automated optimization methods like genetic algorithms [Mit98, PLM08, Gol89] or reinforcement learning [Wil92, SB98] in order to identify good settings for the different actions.

For the RoboCup 2015, we implemented a new walking engine that replaces our previous walk that was based on the B-Human walking engine. Similar to the previous walk, the new walking engine using a linear inverted pendulum model as in [GR11]. The step-planning follows concepts similar to the capture steps used in [MB13]. Additionally, we use optimization to adjust parameters automatically for individual robots.

2.2 Behavior/Situation Recognition

A persistent research direction of our working group addresses the recognition of intentions and plans of agents. Such high-level functions cannot be used before a coordinated control of the agent is possible. Substantial advances have been made in past few years experimenting and developing various techniques such as logic-based approaches [WV11], approaches based on probabilistic theories [Rac08], and artificial neural networks [Sta08]. The results have been partly implemented in the current code. In the last years, improvements in the performance of SimSpark and faster hardware made it possible to run 11 vs 11 games with full physics in real-time. Therefore, the 3D Simulation League now offers a good environment for research in behavior recognition and prediction methods for a team of agents and we can follow this research approach as a short-term goal.

Our approach to plan recognition is based on a qualitative description of dynamic scenes (cf. [WSV03, WVH05, DFL⁺04, MVH04, MPJ⁺16]). The basic idea is to map the quantitative information perceived by the agent to qualitative facts that can be used for symbolic processing. Given a symbolic representation it is possible to define possible actions with their preconditions and consequences. In previous work real soccer tactical moves as, for instance, presented in Lucchesi [Luc01], have been formalized [Bog07]. As planning algorithms themselves are costly and thus hard to use in a demanding online scenario as robotic soccer,

previously generated generic plans are provided to the agent who then can select the best plan w.r.t. some performance measure out of the set of plans that can be applied to a situation. As the pre-defined plans take into account multi-agent settings it is possible to select a tactical move for a group of agents where different roles are assigned to various agents. In the 2D Soccer Simulation League and the previous server of the 3D Soccer Simulation League this approach has already been applied as behavior decision component in some test matches [WBE08, Bog07].

We propose a set of tools for spatio-temporal real-time analysis of dynamic scenes used in the 3D Soccer Simulation League to port and adapt it into the Standard Platform League. It is designed to improve the grounding situation of autonomous agents in (simulated) physical domains. We introduced a knowledge processing pipeline ranging from relevance-driven compilation of a qualitative scene description to a knowledge-based detection of complex event and action sequences, conceived as a spatio-temporal pattern matching problem. A methodology for the formalization of motion patterns and their inner composition is defined and applied to capture human expertise about domain-specific motion situations. It is important to note that the approach is not limited to robot soccer. Instead, it can also be applied in other fields such as experimental biology and logistics processes [WV11].

2.3 Real-time Knowledge Representation

Creating, maintaining, and deducing accurate world knowledge in a dynamic, complex, adversarial, and stochastic environment such as the RoboCup environment is a demanding task. Knowledge should be represented in real-time (i.e., within ms) and deduction from knowledge should be inferred within the same time constraints. In [AV15], we proposed an extended assertional formalism for an expressive $\mathcal{SROIQV}^D$ Description Logic to represent asserted entities in a lattice structure. This structure can represent temporal-like information. Since the computational complexity of the classes of description logic increases with its expressivity, the problem demands either a restriction in the expressivity or an empirical upper bound on the maximum number of axioms in the knowledge base. In this work, we assumed that the terminological/relational knowledge changes significantly slower than the assertional knowledge.

Henceforth, using a fixed terminological and relational formalisms and the proposed lattice structure, we empirically bound the size of the knowledge bases to find the best trade-off in order to achieve deduction capabilities of an existing description logic reasoner in real-time. The queries deduce instances using the equivalent class expressions defined in the terminology. The experiments were conducted in the RoboCup 3D Soccer Simulation League environment and provided justifications of the usefulness of the proposed assertional extension. We have shown the feasibility of our new approach under real-time constraints and conclude that a modified FaCT++ reasoner empirically outperforms other reasoners within the given class of complexity. Our next research objective is to use

our approach with incremental reasoning on a physical robot to model believes and interpret entities in uncertain environments.

2.4 Prediction and Control through Reinforcement Learning

Reinforcement learning (RL) is a popular method in the context of agents and learning where a reward is given to an agent in order to evaluate its performance and thus, learning an optimal or near-optimal policy for action selection [Wil92, SB98]. RL has been applied successfully in robotic soccer before by other teams (e.g., [MR02, RGH⁺06, KS04]). We have integrated a framework for reinforcement learning into our agent based on different variants of incremental, standard, and gradient temporal-difference learning algorithms (cf. [Wat89, WD92, SB98, Mae]).

We are planning to apply reinforcement learning at two levels: (1) the first level where we investigate how certain skills can be optimized by reinforcement learning, e.g., in order to walk faster or to stand up in shorter time, and (2) the second level where learning should be applied is located in the behavior decision process. If it is known which strategic moves are possible the selection of the preferable move should be learned by reinforcement learning methods. The set of possible actions is determined by the applicable plans. The reward is given with respect to the result of the plan execution, e.g. if it failed or if it could be finished successfully. The desired result would be an automatically optimized high-level behavior based on a set of pre-defined plans. Different experiments have to show how the performance of the team can be improved in matches with identical or varying opponent teams.

The recent learning tasks that have been carried out in the RL framework is based on linear function approximation, specially the penalty goal keep behavior. The RL framework is extended with gradient temporal-difference algorithms such as GTD(λ), GQ(λ), Greedy-GQ, Off-PAC algorithms [MSBS10, BBSE10, Mae]. These algorithms have been proven to converge with linear and non-linear function approximators and they have shown superior results in prediction and control problems.

3 Relevant Work

This section provides a brief introduction to a list of our contributions to the 3D Soccer Simulation League. In Subsection 3.1 we describes the visualization and debug tool that we have contributed to the league. Subsection 3.2 describes an efficient implementation of a reinforcement learning software, and Subsection 3.3 describes SimSpark and ODE improvements.

3.1 Monitor and Debugging Tool

A former RoboCanes member, Justin Stoecker, has invented a new 3D soccer server monitor (RoboViz) that runs platform independent. RoboViz is a software program designed to assess and develop agent behaviors in a multi-agent

system, the RoboCup 3D simulated soccer league. It is an interactive monitor that renders agent and world state information in a three-dimensional scene. In addition, RoboViz provides programmable drawing and debug functionality to agents that can communicate over a network.

The tool facilitates the real-time visualization of agents running concurrently on the SimSpark simulator, and provides higher-level analysis and visualization of agent behaviors (figure 1). We also use an adapted version of this tool for our SPL agents.



Fig. 1. RoboViz interface with debugging information and 2D bird view

Features include visualization and debugging (e.g. real-time debugging; direct communication with agents; selecting shapes to be rendered), interactivity and control (e.g. reposition of objects; switching game-play modes), enhanced graphics (e.g. stereoscopic 3D graphics on systems with support for quad-buffered OpenGL; effects such as soft shadows and bloom post-processing provide a visually enticing experience), easy use (e.g. simple controls, automatic connection to the server, platform independency), and other features (e.g. various scene perspectives, logfile viewing, playback with different speeds). A detailed description of RoboViz has been published as a paper for the RoboCup Symposium [SVar, SV14].

3.2 Reinforcement Learning Library for Robotic Platforms

Reinforcement Learning on robotics platforms need efficient implementation of the state-of-the-art algorithms. RLLib (<http://rllib.saminda.org>) is an implementation of incremental standard and gradient temporal-difference learning (GTDL) algorithms for robotics applications using C++ programming language. The implementation of this highly optimized and lightweight library is inspired

by the API of RLPark, which is a library of temporal-difference learning algorithms implemented in Java. The library is tested on the Robocup 3D simulator and on the NAO V4 humanoid robots with different configurations [AV16].

3.3 SimSpark and ODE improvements in 3D Simulation League

Sander van Dijk (Team Boldhearts) and our team RoboCanes have developed a new SimSpark and ODE version. This work is supported by a RoboCup Federation Grant and is focussed on the following goals: (1) Improve stability: fix bugs and increase robustness of simulator. (2) Enable starting multiple instances on a single machine or over a network: make it possible to easily run multiple simulations in parallel. The result has been at the Regional Opens in Germany and Iran in 2011 as well as used during the World Cup 2011 in Istanbul. (3) Enhance run-time control: give the possibility to alter any simulation detail at run-time, alleviating need to constantly restart the system. (4) Develop graphical utility tools: facilitate setting up a batch of experiments.

References

- [AV15] Saminda Abeyruwan and Ubbo Visser. A New Real-Time Algorithm to Extend DL Assertional Formalism to Represent and Deduce Entities in Robotic Soccer. In H. Akin Reinaldo, A. C. Bianchi, Subramanian Ramamoorthy, and Komei Sugiura, editors, *RoboCup 2014: Robot Soccer World Cup XVIII*, LNAI, to appear, 2015. Springer Berlin / Heidelberg.
- [AV16] Saminda Abeyruwan and Ubbo Visser. RLLib: C++ Library to Predict, Control, and Represent Learnable Knowledge using On/Off Policy. In Luis Almeida, Jianmin Ji, Gerald Steinbauer, and Sean Luke, editors, *RoboCup 2015: Robot Soccer World Cup XIX*. Springer Berlin / Heidelberg, 2016. To appear.
- [BBSE10] Lucian Busoniu, Robert Babuska, Bart De Schutter, and Damien Ernst. *Reinforcement Learning and Dynamic Programming Using Function Approximators*. CRC Press, Inc., Boca Raton, FL, USA, 1st edition, 2010.
- [Bog07] Tjorben Bogon. Effiziente abduktive Hypothesengenerierung zur Erkennung von taktischem und strategischem Verhalten im Bereich RoboCup. Master’s thesis, Universitaet Bremen, 2007.
- [DFL⁺04] F. Dylla, A. Ferrein, G. Lakemeyer, J. Murray, O. Obst, T. Röfer, F. Stolzenburg, U. Visser, and T. Wagner. Towards a League-Independent Qualitative Soccer Theory for RoboCup. In *RoboCup 2004: Robot Soccer World Cup VIII*. Springer, 2004.
- [Gol89] David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Kluwer Academic Publishers, Boston, MA., 1989.
- [GR11] Colin Graf and Thomas Röfer. A Center of Mass Observing 3D-LIPM Gait for the RoboCup Standard Platform League Humanoid. In *RoboCup*, pages 102–113, 2011.
- [KS04] G. Kuhlmann and P. Stone. Progress in 3 vs. 2 keepaway. In *RoboCup-2003: Robot Soccer World Cup VII*, pages 694 – 702. Springer Verlag, Berlin, 2004.
- [Luc01] M. Lucchesi. *Coaching the 3-4-1-2 and 4-2-3-1*. Reedswain Publishing, 2001.

- [Mae] Hamid Reza Maei. *Gradient Temporal-Difference Learning Algorithms*. PhD thesis.
- [MB13] Marcell Missura and Sven Behnke. Omnidirectional capture steps for bipedal walking. In *Proceedings of IEEE Int. Conf. on Humanoid Robots (Humanoids)*, 2013.
- [Mit98] M. Mitchell. *An introduction to genetic algorithms*. The MIT press, 1998.
- [MPJ⁺16] Joseph G. Masterjohn, Mihai Polceanu, Julian Jarrett, Andreas Seekircher, Cédric Buche, and Ubbo Visser. Regression and Mental Models for Decision Making on Robotic Biped Goalkeepers. In *RoboCup 2015: Robot Soccer World Cup XIX*. Springer Berlin / Heidelberg, 2016. To appear.
- [MR02] A. Merke and M. Riedmiller. Karlsruhe Brainstormers - a reinforcement learning way to robotic soccer. In *RoboCup 2001: Robot Soccer World Cup V*, pages 435–440. Springer, Berlin, 2002.
- [MSBS10] Hamid Reza Maei, Csaba Szepesvári, Shalabh Bhatnagar, and Richard S. Sutton. Toward off-policy learning control with function approximation. In *ICML*, pages 719–726, 2010.
- [MVH04] Andrea Miene, Ubbo Visser, and Otthein Herzog. Recognition and prediction of motion situations based on a qualitative motion description. In D. Polani, B. Browning, A. Bonarini, and K. Yoshida, editors, *RoboCup 2003: Robot Soccer World Cup VII, LNCS 3020*, pages 77–88. Springer, 2004.
- [PLM08] R. Poli, WB Langdon, and N.F. McPhee. *A field guide to genetic programming*. Lulu Enterprises Uk Ltd, 2008.
- [Rac08] Carsten Rachuy. Erstellen eines probabilistischen Modells zur Klassifikation und Prädiktion von Spielsituationen in der RoboCup 3d Simulationsliga. Master’s thesis, Universität Bremen, 2008.
- [RGH⁺06] M. Riedmiller, T. Gabel, R. Hafner, S. Lange, and M. Lauer. Die brainstormers: Entwurfsprinzipien lernfähiger autonomer roboter. *Informatik-Spektrum*, 29(3):175–190, June 2006.
- [SB98] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.
- [SSAV13] Andreas Seekircher, Justin Stoecker, Saminda Abeyruwan, and Ubbo Visser. Motion capture and contemporary optimization algorithms for robust and stable motions on simulated biped robots. In Xiaoping Chen, Peter Stone, Luis Enrique Sucar, and Tijn Van der Zant, editors, *RoboCup 2012: Robot Soccer World Cup XVII*. Springer Berlin / Heidelberg, to appear, Mexico City, 2013.
- [Sta08] Arne Stahlbock. Situationsbewertungsfunktionen zur Unterstützung der Aktionsauswahl in der 3D Simulationsliga des RoboCup. Master’s thesis, Universität Bremen, 2008.
- [SV14] Justin Stoecker and Ubbo Visser. Visualizing and Debugging Complex Multi-Agent Soccer Scenes in Real Time. In in print, editor, *RoboCup 2013: Robot Soccer World Cup XVII*. Springer Berlin / Heidelberg, 2014.
- [SV16] Andreas Seekircher and Ubbo Visser. An adaptive lipm-based dynamic walk using model parameter optimization on humanoid robots. *KI - Künstliche Intelligenz*, 30(3):233–244, 2016.
- [SVar] Justin Stoecker and Ubbo Visser. RoboViz: Programmable Visualization for Simulated Soccer. In Thomas Röfer, Norbert Michael Mayer, Jesus Savage, and Uluç Saranlı, editors, *RoboCup 2011: Robot Soccer World Cup XV*, 2012, to appear.

- [Wat89] Christopher J. C. H. Watkins. *Learning from delayed rewards*. PhD thesis, University of Cambridge, England, 1989.
- [WBE08] T. Wagner, T. Bogon, and C. Elfers. Incremental generation of abductive explanations for tactical behavior. *RoboCup 2007: Robot Soccer World Cup XI*, pages 401–408, 2008.
- [WD92] Christopher J. C. H. Watkins and Peter Dayan. Technical note: Q-learning. *Machine Learning*, 8(3-4):279–292, May 1992.
- [Wil92] R.J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3):229–256, 1992.
- [WSV03] T. Wagner, C. Schlieder, and U. Visser. An extended panorama: Efficient qualitative spatial knowledge representation for highly dynamic environments. In *Proceedings of the IJCAI-03 Workshop on Issues in Designing Physical Agents for Dynamic Real-Time Environments: World modelling, planning, learning, and communicating*, pages 109–116, 2003.
- [WV11] T. Warden and U. Visser. Real-time spatio-temporal analysis of dynamic scenes. *Knowledge and Information Systems*, pages 1–37, 2011.
- [WVH05] T. Wagner, U. Visser, and O. Herzog. Egocentric qualitative knowledge representation for physical robots. *Journal for Robotics and Autonomous Systems*, 2005.