RoboCanes
RoboCup 3D Simulation League Team Description Paper 2015

Saminda Abeyruwan, Nasir Laskar, Joseph Masterjohn, 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 2015, the team is lead by Saminda Abeyruwan, who has participated in RoboCup teams since 2010. Saminda is a final year PhD student and he contributes to RoboCanes teams in the 3D Soccer Simulation League and in the Standard Platform League (SPL).

RoboCanes team members are Saminda Abeyruwan, Nasir Laskar, Joseph Masterjohn, Kyle Poore, Andreas Seekircher, and Ubbo Visser. Saminda, Andreas, Nasir, Joseph, and Kyle are PhD students, and Ubbo is a faculty member at the Computer Science Department of the University of Miami.

Saminda focuses on knowledge representation, localization (robot, ball, opponents) and role formations, and has a good experience with filter techniques and reinforcement learning techniques. Nasir, Joseph, and Kyle are new member of the team working on algorithm development and also contribute to the SPL. Andreas comes from the SSL team B-Smart and has done studies for his MSc Thesis on the physical NAO. He has published a paper about his thesis entitled "Entropy-based active vision for a humanoid soccer robot" which received him 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 RoboCanes.

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 2015, and (2) the long-term activities beyond the competition.

Besides working on the low-level skills that are described in Subsection 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 Subsection 2.2. The real-time knowledge representation formalism is presented in Subsection 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 Subsection 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 seperate 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 heterogenious 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 pursuit 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 a modern optimization algorithm 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 [19]. Further optimizations are planned as described in Section 2.4.

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 robot types. 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 [14, 15, 8] or reinforcement learning [29, 22] in order to identify good settings for the different actions.

We plan to replace our current gait with a new walking engine that can also be optimized for different robot types. In contrast to directly modifying joint angles of a special action, we are planning to optimize the underlying models used to generate the gait dynamically.

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 [26], approaches based on probabilistic theories [16], and artificial neural networks [20]. The results have been partly implemented in the current code. For a big portion of last year, the 3D server settings and performance (especially for a larger number of robots) lowered the probability of a fully functional behavior recognition and prediction method for a team of agents. The latest implementation of SimSpark however has changed this situation significantly so that 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. [24, 25, 6, 13]). 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 [10], have been formalized [4]. 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 simulation league and the previous server of the 3D simulation league this approach has already been applied as behavior decision component in some test matches [23, 4].

We developed a set of tools for spatio-temporal real-time analysis of dynamic scenes that can be used in the 3D Simulation 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 [26].

Our research is partly an application of the concepts developed in the parallel project "Automatic Recognition of Plans and Intentions of Other Mobile Robots in Competitive, Dynamic Environments" (research project in the German Research Councils priority program "Cooperating Teams of Mobile Robots in Dynamic Environments"). It is necessary to identify a set of relevant strategic moves that can be either applied by the own team (if the probability for a successful move is high) or recognized from observing the behavior of the opponent team. The German Research Council (DFG) supported our research line between 2001 and 2007 and invited us to submit ideas for further long-term research ideas in that area. This clearly indicates the significance of our research efforts.

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 [3], we proposed an extended assertional formalism for an expressive $\mathcal{SROIQV}^{\mathcal{D}}$ 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 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, (hopefully) learning an optimal policy for action selection [29, 22]. Reinforcement learning has been applied successfully in robotic soccer before by other teams (e.g., [12, 17, 9]). We have integrated a framework for reinforcement learning into our agent where different variants like Q-Learning and SARSA have been used (cf. [27, 28, 22]). We have published our current work on one of the Humanoids 2011 and 2012 workshops on soccer playing humanoids [18, 1] and submitted a new paper for the AAMAS workshop ALA [2].

It is planned to apply reinforcement learning at two different levels: First of all, we want to investigate how certain skills can be optimized by reinforcement learning, e.g., in order to walk faster or to stand up in shorter time.

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 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 reinforcement learning framework is extended with GQ(λ), Greedy-GQ(λ), and Off-PAC algorithms [11, 5]. These algorithms have been proven to converge with linear function approximations and it is 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 not currently possible with existing tools (Figure 1).

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 [21].

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

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++ programing 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 with different configurations. Figure 2 shows the step time in

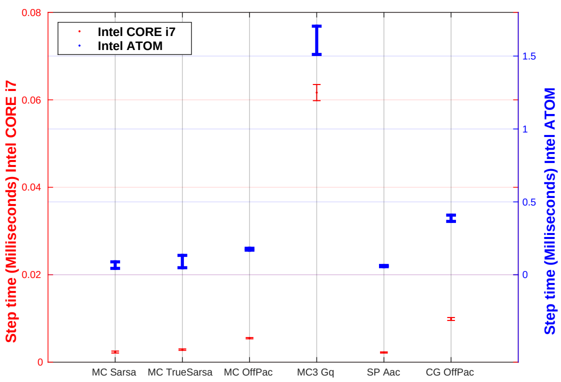

Fig. 2. Step update times in milliseconds. The thick error bars (blue) show the step time for Intel ATOM, while the thin error bars (red) show the step update time for Intel CORE–i7.

milliseconds, i.e., the time an algorithm requires to update its parameters from the observations, for a set of benchmark problems popular in RL literature [7]. We have considered two hardware platforms: 1) Intel CORE–i7 2.2GHz laptop; and 2) Intel ATOM 1.6GHz CPU available on NAO humanoid robot.

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.

Sander has announced some of the developments in the mailing list.

References

1. Abeyruwan, S., Seekircher, A., Visser, U.: Dynamic Role Assignment using General Value Functions. In: Behnke, S., Röfer, T., Visser, U. (eds.) IEEE Humanoid Robots, HRS workshop. IEEE, Osaka, Japan (2012), `http://www.humanoidsoccer.org/ws12/program.html`
2. Abeyruwan, S., Seekircher, A., Visser, U.: Robust and Dynamic Role Assignment in Simulated Soccer. In: AAMAS 2013, ALA Workshop (2013)
3. Abeyruwan, S., Visser, U.: A New Real-Time Algorithm to Extend DL Assertional Formalism to Represent and Deduce Entities in Robotic Soccer. In: Reinaldo, H.A., Bianchi, A.C., Ramamoorthy, S., Sugiura, K. (eds.) RoboCup 2014: Robot Soccer World Cup XVIII. LNAI, Springer Berlin / Heidelberg, to appear (2015)
4. Bogon, T.: Effiziente abduktive Hypothesengenerierung zur Erkennung von taktischem und strategischem Verhalten im Bereich RoboCup. Master's thesis, Universitaet Bremen (2007)
5. Busoniu, L., Babuska, R., Schutter, B.D., Ernst, D.: Reinforcement Learning and Dynamic Programming Using Function Approximators. CRC Press, Inc., Boca Raton, FL, USA, 1st edn. (2010)
6. Dylla, F., Ferrein, A., Lakemeyer, G., Murray, J., Obst, O., Röfer, T., Stolzenburg, F., Visser, U., Wagner, T.: Towards a League-Independent Qualitative Soccer Theory for RoboCup. In: RoboCup 2004: Robot Soccer World Cup VIII. Springer (2004)
7. Gatti, C.: Design of Experiments for Reinforcement Learning (Springer Theses). Springer (2015)
8. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Kluwer Academic Publishers, Boston, MA. (1989)
9. Kuhlmann, G., Stone, P.: Progress in 3 vs. 2 keepaway. In: RoboCup-2003: Robot Soccer World Cup VII. pp. 694 – 702. Springer Verlag, Berlin (2004)

10. Lucchesi, M.: Coaching the 3-4-1-2 and 4-2-3-1. Reedswain Publishing (2001)
11. Maei, H.R., Szepesvári, C., Bhatnagar, S., Sutton, R.S.: Toward off-policy learning control with function approximation. In: ICML. pp. 719–726 (2010)
12. Merke, A., Riedmiller, M.: Karlsruhe Brainstormers - a reinforcement learning way to robotic soccer. In: RoboCup 2001: Robot Soccer World Cup V. pp. 435–440. Springer, Berlin (2002)
13. Miene, A., Visser, U., Herzog, O.: Recognition and prediction of motion situations based on a qualitative motion description. In: Polani, D., Browning, B., Bonarini, A., Yoshida, K. (eds.) RoboCup 2003: Robot Soccer World Cup VII, LNCS 3020. pp. 77–88. Springer (2004)
14. Mitchell, M.: An introduction to genetic algorithms. The MIT press (1998)
15. Poli, R., Langdon, W., McPhee, N.: A field guide to genetic programming. Lulu Enterprises Uk Ltd (2008)
16. Rachuy, C.: Erstellen eines probabilistischen Modells zur Klassifikation und Prädiktion von Spielsituationen in der RoboCup 3D Simulationsliga. Master's thesis, Universität Bremen (2008)
17. Riedmiller, M., Gabel, T., Hafner, R., Lange, S., Lauer, M.: Die brainstormers: Entwurfsprinzipien lernfähiger autonomer roboter. Informatik-Spektrum 29(3), 175–190 (June 2006)
18. Seekircher, A., Saminda, A., Visser, U.: Accurate Ball Tracking with Extended Kalman Filters as a Prerequisite for a High-level Behavior with Reinforcement Learning. In: Humanoids 2011, 6th Workshop on Humanoid Soccer Robots (2012)
19. Seekircher, A., Stoecker, J., Abeyruwan, S., Visser, U.: Motion capture and contemporary optimization algorithms for robust and stable motions on simulated biped robots. In: Chen, X., Stone, P., Sucar, L.E., van der Zant, T. (eds.) RoboCup. Lecture Notes in Computer Science, vol. 7500, pp. 213–224. Springer (2012)
20. Stahlbock, A.: Sitationsbewertungsfunktionen zur Unterstützung der Aktionsauswahl in der 3D Simulationsliga des RoboCup. Master's thesis, Universität Bremen (2008)
21. Stoecker, J., Visser, U.: Roboviz: Programmable visualization for simulated soccer. In: Röfer, T., Mayer, N.M., Savage, J., Saranli, U. (eds.) RoboCup. Lecture Notes in Computer Science, vol. 7416, pp. 282–293. Springer (2011)
22. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA (1998)
23. Wagner, T., Bogon, T., Elfers, C.: Incremental generation of abductive explanations for tactical behavior. RoboCup 2007: Robot Soccer World Cup XI pp. 401–408 (2008)
24. Wagner, T., Schlieder, C., Visser, U.: 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. pp. 109–116 (2003)
25. Wagner, T., Visser, U., Herzog, O.: Egocentric qualitative knowledge representation for physical robots. Journal for Robotics and Autonomous Systems (2005)
26. Warden, T., Visser, U.: Real-time spatio-temporal analysis of dynamic scenes. Knowledge and Information Systems pp. 1–37 (2011)
27. Watkins, C.J.C.H.: Learning from delayed rewards. Ph.D. thesis, University of Cambridge, England (1989)
28. Watkins, C.J.C.H., Dayan, P.: Technical note: Q-learning. Machine Learning 8(3-4), 279–292 (May 1992)
29. Williams, R.: Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine Learning 8(3), 229–256 (1992)