RoboCanes
RoboCup 3D Simulation League Team Description Paper 2012

Saminda Abeyruwan, Andreas Seekircher, Justin Stoecker, and Ubbo Visser

Department of Computer Science
University of Miami
Coral Gables, FL, USA
robocanes@cs.miami.edu

1 Introduction

Our team, *RoboCanes*, was formed in January 2010 for the 3D Soccer Simulation League. One of the team members was the former team leader of Virtual Werder 3D (VW3D) from the University of Bremen. His move to the University of Miami induced taking the original source code from VW3D and porting it to the new requirements of the soccer server and robot model. In Fall 2011, RoboCanes decided to develop its agent's software from scratch. We are almost finished with this step.

The team is lead by Ubbo Visser who is 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.

Among our team members are Saminda Abeyruwan, Andreas Seekircher, and Justin Stoecker, all PhD students at the Computer Science Department of the University of Miami. Saminda focuses on localization (robot, ball, opponents) and has a good experience with filter techniques. Andreas comes from the 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". He won the best paper award at RoboCup 2010 in Singapore. Justin's specialties are in graphics/visualization. His skills lead to the new 3D monitor RoboViz that is used for regional opens and the World Championship since 2011. The team is completed by various undergraduate and graduate students.

2 Research interests and planned activities

Our research activities are in the area of **behavior/situation recognition, prediction and control**. Our current activities can be divided into two parts:

short-term activities to be addressed before RoboCup 2012 and the mid-term activities beyond this competition.

Besides getting familiar with the physical robots and thus dealing with a lot of 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 application of learning methods for learning low-level skills as well as higher-level behaviors is another research direction addressed by our team presented in section 2.3.

2.1 Humanoid Walking Engine and Special Actions

The development of the robot's basic skills in the *RoboCanes* agent is based on the experiences and results of the Bremen humanoid team *B-Human* [RBF+07] (a follow-up from the *BreDoBrothers*, which was a joint team from the Universität Bremen and the Universität Dortmund [RFH+06]). This is an important step towards merging research efforts of two separate RoboCup leagues. The 3D soccer simulation league can benefit from the experiences of the real robot humanoid league. Later on, a sufficiently realistic simulation 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. In the first step, existing technologies of the *B-Human* team have been integrated into the *RoboCanes* agent. It is tested if and how these technologies can be used in the simulation league's environment. The first skill that has been implemented is the walking engine (as presented in the *RoboCanes* binary for the RoboCup qualification); for more information about the walking engine see [NRL07,LR06,RFH+06]. In order to use the walking engine, the dimensions and physical properties of the simulated agent have to be provided. Furthermore, the agent's status of the different joints must be passed to the walking engine and the resulting effector command have to be mapped to the corresponding effectors in the simulation.

The *B-Human* team has developed a number of further so called "special actions" like: 'getting up', 'walking backwards', 'walking left / right', 'kicking the ball' (with the left or right leg).

These special actions have also been tested on the simulated robot and adapted. Our idea is to benefit from the existing code in two different leagues. We expect that these special actions do not work out of the box. After some major parameter adaptions in order to create a first version of the intended behavior, fine tuning of the parameters has to be done in a second phase. We are planning to apply automated optimization methods like genetic algorithms [Mit98,PLM08,Gol89] or reinforcement learning [Wil92,SB98] in order to identify good settings for the different actions. Further optimizations are planned as described in section 2.3.

The experiences gained from the integration, adaption, and optimization of the actions in the simulation should then flow back to the real robot team in the next step, which hopefully can be helpful to improve the performance of the real 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. 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. [WSV03,WVH05,DFL$^+$04,MVH04]). 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 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 [WBE08,Bog07].

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 [WV11].

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 since 2001 (ended with move to US) and invited us to submit ideas for further long-term research ideas in that area. This clearly indicates the significance of our research efforts. Currently, several research proposals have been submitted or are in preparation (e.g. NSF, NIH, and internal UM proposals).

2.3 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 [Wil92,SB98]. Reinforcement learning 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 where different variants like Q-Learning and SARSA have been used (cf. [Wat89,WD92,SB98]). We have published our current work on one of the Humanoids 2011 workshops [SSVar] and preparing a new submission for the RoboCup 2012 Symposium.

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 w.r.t. 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) and Greedy-GQ algorithms [MSBS10,BBSE10]. These algorithms have been proven to converge with linear function approximators and it is shown superior results in prediction and control problems.

3 Past relevant work

3.1 Monitor and Debugging Tool

Justin Stoecker from our team RoboCanes 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).

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].

3.2 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

[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.

[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.

[LR06] T. Laue and T. Röfer. Getting upright: Migrating concepts and software from four-legged to humanoid soccer robots. In E. Menegatti E. Pagello, C. Zhou, editor, *Proceedings of the Workshop on Humanoid Soccer Robots in conjunction with the 2006 IEEE International Conference on Humanoid Robots*, 2006.

[Luc01] M. Lucchesi. *Coaching the 3-4-1-2 and 4-2-3-1.* Reedswain Publishing, 2001.

[Mit98] M. Mitchell. *An introduction to genetic algorithms.* The MIT press, 1998.

[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.

[NRL07] C. Niehaus, T. Röfer, and T. Laue. Gait optimization on a humanoid robot using particle swarm optimization. In C. Zhou, E. Pagello, E. Menegatti, and S. Behnke, editors, *IEEE-RAS International Conference on Humanoid Robots*, 2007.

[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.

[RBF+07] Thomas Röfer, Christoph Budelmann, Martin Fritsche, Tim Laue, Judith Müller, Cord Niehaus, and Florian Penquitt. B-Human team description for RoboCup 2007, 2007.

[RFH+06] Thomas Röfer, Martin Fritsche, Matthias Hebbel, Thomas Kindler, Tim Laue, Cord Niehaus, Walter Nistico, and Philippe Schober. BreDoBrothers team description for RoboCup 2006, 2006.

[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.

[SSVar] Andreas Seekircher, Abeyruwan Saminda, and Ubbo Visser. 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, to appear.

[Sta08] Arne Stahlbock. Sitationsbewertungsfunktionen zur Unterstützung der Aktionsauswahl in der 3D Simulationsliga des RoboCup. Master's thesis, Universität Bremen, 2008.

[SVar] Justin Stoecker and Ubbo Visser. RoboViz: Programmable Visualization for Simulated Soccer. In Thomas Röfer, Norbert Michael Mayer, Jesus Savage, and Uluç Saranli, 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.