Virtual Werder 3D
RoboCup 2007 Research Proposal

Andreas D. Lattner, Carsten Rachuy, Arne Stahlbock,
Ubbo Visser, and Tobias Warden

TZI, Universität Bremen
PO Box 330 440, D-28334 Bremen, Germany
grp-vw3d@tzi.de

1 Introduction

The intended research activities of *Virtual Werder 3D* [7] can be divided into two parts: The short-term activities to be addressed before the RoboCup 2007 and the mid-term activities beyond this competition. The recent changes in the 3D simulation league from the spheres-based simulation to the humanoid simulation forces the teams to work on basic skills first as it does not make sense to work on elaborated behavior if the agent cannot move at all. Thus, the short-term goals address the development of low-level skills as presented in Section 2. The mid-term goal of our team is to apply plan recognition methods in order to bring in valuable knowledge into the behavior decision process. These efforts are presented in Section 3. 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 4.

2 Humanoid Walking Engine and Special Actions

The development of the robot's basic skills in the *Virtual Werder 3D* agent is based on the experiences and results of the Bremen humanoid team *B-Human* [14] (a follow-up of the *BreDoBrothers* [15]). This is an important step towards merging research efforts of two related RoboCup leagues which has already proposed by Mayer et al. [10]. The 3D simulation league with the new server version 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 as well by (pre-) learning some skills or testing different settings in the simulation first that might be disadvantageous (and costly) for real robots (cf. [5]).

In the first step, existing technologies of the *B-Human* team are integrated into the *Virtual Werder 3D* agent. It is tested if and how these technologies can be used in the simulation league's environment. The first skill to be adapted is the walk skill using the *B-Human* walking engine [8, 15]. 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 commands have to be mapped to the corresponding effectors in the simulation. A successful integration of the *B-Human* walking engine into the *Virtual Werder 3D* agent could already be achieved for the RoboCup 2007 qualification binary. Parameter tuning of the walking engine is current work in progress.

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 are also to be tested on the simulated robot and adapted. It is not expected that these special actions work right away. After some major parameter adaptations 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 in order to identify good settings for the different actions [3]. Evolutionary algorithms have been applied successfully in order to learn (simulated) biped walking within the *Simloid* project [5, 4]. Further optimizations are planned as described in Section 4. 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, too.

3 Plan Recognition

A persistent research direction of our working group addresses the recognition of intentions and plans of agents. Of course, such high-level functions cannot be used before a coordinated control of the agent is possible. Nevertheless, we also address this research topic as a mid-term goal.

Our approach to plan recognition is based on a qualitative description of dynamic scenes (cf. [18, 2, 12]). 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 [9], have been formalized [1]. 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 [17, 1].

The intended research is to apply 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") to the new 3D server. It is necessary to identify 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) supports our research line since 2001 and invited us to submit ideas for further long-term research ideas in that area. This clearly indicates the significance of our research efforts.

4 Reinforcement Learning

In the context of agents and learning a popular method is reinforcement learning (RL) where a reward is given to an agent in order to evaluate its performance and thus, (hopefully) learning an optimal policy for action selection. RL has been applied successfully in robotic soccer before by other teams (e.g., [11, 13, 6]). In our previous agent we have integrated a framework for RL where different variants like Q-Learning and SARSA can be used (cf. [19, 16]). This RL framework will be integrated into our new agent.

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. It is focused to this kind of learning tasks until the RoboCup in Atlanta.

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 RL 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.

Acknowledgment
The work presented here was partially funded by the *Senator für Bildung und Wissenschaft, Freie Hansestadt Bremen* ("FIP RoboCup", Forschungsinfrastrukturplan). We would also like to thank the *B-Human* team – namely, Martin Fritsche, Tim Laue, Cord Niehaus, and Thomas Röfer – for providing their walking engine and the support integrating it into our agent.

References

1. T. Bogon. Effiziente abduktive Hypothesengenerierung zur Erkennung von taktischem und strategischem Verhalten im Bereich RoboCup. Master's thesis, Univer-

sitaet Bremen, 2007. Submitted.
2. 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.
3. D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Kluwer Academic Publishers, Boston, MA., 1989.
4. D. Hein. Simloid - evolution of biped walking using physical simulation. Master's thesis, Humboldt-Universität zu Berlin, 2007.
5. D. Hein and M. Hild. Simloid - research on biped robots controller, using physical simulation and machine learning algorithms. In *Concurrency, Specification and Programming CS&P'2006, vol. 1: Concurrency*, number 206 in Informatik-Berichte, pages 143–151. Humboldt-Universität zu Berlin, 2006.
6. 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.
7. A. D. Lattner, C. Rachuy, A. Stahlbock, U. Visser, and T. Warden. Virtual Werder 3D team documentation. Technical Report 36, TZI - Center for Computing Technologies, Universität Bremen, September 2006.
8. T. Laue and T. Röfer. Getting upright: Migrating concepts and software from four-legged to humanoid soccer robots. In E. M. 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.
9. M. Lucchesi. *Coaching the 3-4-1-2 and 4-2-3-1*. Reedswain Publishing, 2001.
10. N. M. Mayer, J. Boedecker, R. da Silva Guerra, O. Obst, and M. Asada. 3d2real: Simulation league finals in real robots. In *RoboCup 2006: Robot Soccer World Cup X Preproceedings*, 2007. To appear.
11. 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.
12. A. Miene, U. Visser, and O. 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.
13. 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.
14. T. Röfer, C. Budelmann, M. Fritsche, T. Laue, J. Müller, C. Niehaus, and F. Penquitt. B-Human team description for RoboCup 2007, 2007.
15. T. Röfer, M. Fritsche, M. Hebbel, T. Kindler, T. Laue, C. Niehaus, W. Nistico, and P. Schober. BreDoBrothers team description for RoboCup 2006, 2006.
16. R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT Press, Cambridge, MA, 1998.
17. T. Wagner, T. Bogon, and C. Elfers. Incremental generation of abductive explanations for tactical behavior. Submitted to the RoboCup International Symposium 2007, 2007.
18. T. Wagner, U. Visser, and O. Herzog. Egocentric qualitative knowledge representation for physical robots. *Journal for Robotics and Autonomous Systems*, 2005.
19. C. J. C. H. Watkins and P. Dayan. Technical note: Q-learning. *Machine Learning*, 8(3-4):279–292, May 1992.