UTUtd 2007
3D Team Proposal

HesamAddin Torabi Dashti, Sahar Asadi, Meysam Bastani
, Fetemeh Miri Disfani, and Alireza Fotuhi Siahpirani

School of Math, Statistics, and Computer Science,
University College of Science, University of Tehran, Iran
{dashti, asadi, bastani, fmiri, fotuhi}@khayam.ut.ac.ir
http://www.fos.ut.ac.ir/robocup

Abstract. This is the proposal of UTUtd for RoboCup 2007 competitions in Soccer Simulation 3D. In the new version of simulator agents are humanoids, simulated version of Fujitsu robot HOAP-2. The agent has a large Degrees of Freedom (DOF) and control of its motion is very difficult. UTUtd plans for agent motion control is based on current research projects in which its members are involved: evlolutionary system for buildable block robot design and behavioral learning in multi-agent environment.

1 Introduction

Soccer Simulation 3D competitions was started in 2004 with the aim of providing a more realistic environment for developers to have the advantages of simulated environment while considering constraints of the real world in the development of basic skills. With the effort of providers of 3D soccer server remarkable improvements have been made ever since to decrease the gap between simulation and reality [2]. A great extent of these efforts were concentrated on modifying the structure design of the agents to resemble a humanoid robot. As a result HOAP-2 humanoid robot was simulated and is going to be used in the following competitions.

Along with 3D competitions, Soccer Simulation 3D Development Competitions started in 2005 to accelerate the improvement of the server and motivate developers to take part in this process.

UTUtd as an experienced team in 2D competitions, took the third place in the first year of 3D competitions. UTUtd team members have been actively involved in soccer simulation 3D related events since its formation. Being familiar with the structure of the server, UTUtd development team developed a coach to be integrated in the latest version of 2006.They also developed an advanced monitor and online debugger for the server which led to taking the first place in Development Competitions.

This year, together with the essential modifications in the server, UTUtd decided to reorganize its structure and apply some changes in team work in order to provide more productive results. Therefore members who are working on related subjects such as evolutionary robotics and distributed AI decided to put their knowledge together to find convenient solutions for the recent challenge. Using a model of HOAP-2 robot as simulated agents gives us the ability to implement our ideas in a practical environment.

In the following chapters UTUtds approach to develop a HOAP-2 agent is described.

2 Agent Specifications and Basic Skills

According to TC decision, although qualification is based on soccerbot agent but the focus of development in 3D simulator is on HOAP-2 agent which is generally very similar to soccerbot. So, our proposal is also focused on HOAP-2. HOAP-2 stands for Humanoid for Open Architecture Platform and is simulated based on Fujitsu HOAP-2 robot [2]. HOAP-2 is a robot with 50cm height, 7kg weight, and 25 DOF: 6 DOF in each foot, 4 in each arm, 1 in each hand and 2 in neck and one in the waist [4].

The first step toward controlling an agents behavior is to define and develop its low level skills. Comprehensive definition of each skill and proper implementation of it is a complicated task and requires considering many parameters such as each joints restrictions based on HOAP specifications. In the following is the list of low level skills and their definition.

– Walk : This skill includes walking in forward and backward direction and turning to left and right.
– Stop : In this skill agent tries to decrease its pace and stop while maintaining its balance.
– Stand up : In this skill agent tries to stand up from the lying position when it's lies on its stomach or on the back.

Regarding the fact that taking control of a HOAP agent is a complicated and time consuming process, development of high level skills such as decision making, passing the ball and so on can not be achieved for 2007 competitions. Nevertheless we can define some intermediate skills which are not as intricate as decision making but need more work than low level skills. Our far plan is to acquire such skills after reaching a certain level of flexibility in low level skills. Skills such as chasing the ball, positioning behind it and kicking, and making simple decisions based on its perception from the field are intermediate skills.

3 Motion Control Methods

Biped Walking is a complex and open problem. Since HOAP-2 has many DOF in its joints movement, checking all possible configurations or finding a mathematical solution to optimize the performance of the humanoid is exponential in time. In order to control agent's motion, many learning methods like Neural Networks, Evolutionary methods, and Reinforcement Learning are introduced.

There are introduced two approaches in studying of humanoid walking behavior: Zero Moment Point (ZMP) control and Inverted Pendulum model. ZMP-based methods usually require precise knowledge of the robots dynamics (e.g. mass, center of mass location, and inertia of each link) to generate the walking patterns. Hence, they are dependent on the accuracy of the models. In contrast, there are methods which work with only a little knowledge of system dynamic like CM and total angular momentum and they use a feedback control. These methods are usually termed as the inverted pendulum approach [1]. Both proposed methods in this paper are on the basis of inverted pendulum approach.

3.1 Central Pattern Generator Networks (CPG)

The real robot is able to learn movement using dynamically reconfigurable neural networks. Fujitsus technology is based on CPG learning system. CPG is a network with many sub-networks and interconnections. The idea behind CPG is to simulate neural oscillations sent to human body joints and muscles during walking motion. In order to set and check the configuration and weights of the neural network, Numerical Perturbation method has been used which can approximates nonlinear functions difficult in analyzing. Other methods are also introduced to tune networks parameters like Evolutionary Algorithms [2] and Reinforcement Learning (RL) [7] and [6].

3.2 Evolutionary Learning

Various implementation of evolutionary methods are available in controlling humanoid robot motion. Here are two evolutionary methods implemented by Daniel Hein in his Diploma Thesis [5]:

- In the fist method, Biped walking can be classified in the category of periodic motions where each joint repeats its motion after one period. Therefore a cyclic target angle function can be defined for each joint. One period of each function corresponds to one full step. Regarding the fact that sinusoid is a basic cyclic function we can use partial Fourier series to describe the motion of the agent. Parameters of Fourier series are optimized using evolutionary algorithms.
- The neural oscillator approach generates the core oscillation with the use of the discrete time dynamics of two neurons. A single Genotype is given by a real number parameter vector denoting the synaptic weights of the controllers network.

4 Proposed Methods

4.1 Evolutionary Buildable blocks

Based on researches on coupled evolution of controller and morphology [3], we decided to use variable length genomes to determine the effect of inclusion of different sets of joints in behavior of agent. In this method, we use angles of each joint to represent agent's situation. The first generation would be a random combination of genotypes. Mutating the first generation using unequal crossover can lead to child genomes with lengths differing from their parent genomes. Using this method can result in convenient combination of joints and excluding less effective ones in distinct actions.

4.2 Multi-Agent Cooperation and Reinfocement Learning

The idea is to consider the simulated humanoid as a multi-agent system in which each joint is called an agent. In this system, each agent with its own specification and constraints behaves independently and in corporation with other agents to achieve goals of the system. The general behavior of the system should provide expected values. The most important desired value is stability in center of mass (CM) position. It is desired to minimize the oscillations of CM during walking. The other expected value is walking distance which we want to maximize it before the robot falls down.

In order to learn this multi-agent system, the suggested method is Reinforcement Learning (RL) on a behavioral based model. A critic from outside of the system will give punishment and reward to the behavior of each agent. State of action space of humanoid walking system is very large and in order to use RL for on-line learning, this complexity should be managed. Finding optimal policy with typical TD or Q-learning makes the learning of the system with such a complexity slow. The idea is to use a feed-forward neural networks or Actor-Critic method but it needs more studying. Actor-Critic is a policy gradient method incorporating a value function [7].

References

1. Bacher, Y.: Developing Controllers for Biped Humanoid Locomotion. Diploma Thesis, University of Edinburgh, 2004
2. Boedecker, J., Mayer, N.M., Ogino, M., da Silva Guerra, R., Kikuchi, M., Asada, M.: Getting closer: How simulation and humanoid league can benet from each other. In Murase, K., Sekiyama, K., Kubota, N., Naniwa, T., Sitte, J., eds.: AMiRE, Springer (2005) 9398
3. Bongard, J.C.: Incremental Approaches to the Combined Evolution of a Robot's Body and Brain. PhD Thesis, University of Zurich, Zurich, 2003
4. Cominoli, P.: Development of physical simulation of a real humanoid robot. Diploma Thesis, EPFL, Swiss Federal Institute of Technology, 2005
5. Hein, D.: Simloid, Evolution of Biped Walking Using Physical Simulation. Diploma Thesis, University of Humboldt, Berlin, 2007
6. Katik, D., Vukobratovic, M.: Control Algorithm for Humanoid Walking Based on Fuzzy Reinforcement Learning. In: 4th Serbian-Hungarian Symposium on Intelligent Systems, 2006
7. Mori, T., Nakamura, Y., Sato, M., Ishii, S.: Reinforcement Learning for CPG-Driven Biped Robot. AAAI (2004) 623–630