

Team Description

TsinghuAeolus3D 2006

Liu Rujia, Xiao Qin, Wang Rongrong, Wu Shuo, Yang Fan, and Sun Zengqi

Department of Computer Science and Technology
Tsinghua University
Beijing 100084, China

Abstract. This paper describes the TsinghuAeolus3D team, from fundamental problems and solutions that are considered last year, to new ideas and methods employed recently. We first present a summary for our previous work, which can be seen as a short documentation for our source code distribution last year, then new research areas and advanced artificial intelligence methods are discussed, which are at the heart of our new team.

1 Introduction

The TsinghuAeolus3D have been participating in RoboCup Simulation 3D league since 2004. Being in the top sixteen teams in RobCup 2005, Osaka, we published our full source code online, which to our knowledge is the only source code distribution available to the public till now. The main research goal this year is to employ different methods to solve different problems in this domain, possibly by combining some of the together. A new decision making core discussed below allows us to do research from different perspectives, which pushes us forward to the goal.

One principle of TsinghuAeolus 2006 is to develop specified auxiliary tools and our own libraries and maximize their contribution to our research and software development. Purposes of some will be mentioned below.

The reset of the paper is organized as follows: section 2 is an introduction to our fundamental work, much of which are from our source code distribution, TsinghuAeolus3D 2005. Though immature, TsinghuAeolus3D 2005 provides frameworks and ideas, which are essential for understanding our new team. Section 3 gives a summary of our current research and briefly discusses some issues we're concentrating on, and what we may be researching in the future.

2 Fundamental work

World model and low-level individual skills are two most important parts in TsinghuAeolus. In fact, due to limited time and lack of human power, little work has been done besides them last year. Since a very simple decision system was used in TsinghuAeolus2005, we omit their details here.

2.1 World Model

To offer a foundation layer for higher skill and strategy layers, we build and maintain a world model of the game to provide a highly dependable simulation of the real soccer world, including the ball's and agents' positions and velocities. The world model also updates these status dynamically and makes proper prediction using a simplified physical model.

As noises are added every time when agents sense their coordinates, if the received data is used directly, it would cause a loss in accuracy and a series of inaccurate prediction. To deal with this problem, the Kalman filter uses a variance value to enhance the accuracy of the original data and we achieved a satisfactory result that the error is at most half the radius of the ball.

Another approach in noise reduction is about ball motion. By offline training and data regression, we reconstructed our physical model of ball's movement and made adjustment to kicking angles in order to have a better precision. This work greatly improved the accuracy of prediction and result in a more flexible kicking action.

2.2 Individual Skills

The training of individual skills aims at providing a basic building block for the high-level team play.

On one hand, we summarize the individual skills that are essential to high level agent control into two layers.

The first layer aims at agent motion control without ball. The main problem of this layer is to control the agent so that it moves along a pre-determined path with a predictable direction and speed. In order to achieve this goal as precisely as possible, we utilize the data from our experiments when building the world model framework for regression of the agent motion function and the agent control function.

The second layer aims at ball handling of a moving agent. This layer incorporates the ball's motion control into consideration. We design a set of experiments to extract the mapping information between the parameters of 'kick' action and the effect. By regression of such functions, we are able to control the agent to shoot the ball at a destination with very high accuracy. Combining the achievements from previous work, we are capable of providing higher level with a portable interface that can virtually achieve predictable agent/ball motion control.

On the other hand, basic skills are often arranged so that they form a complicated skill to achieve relatively independent goal. Skills such as dribbling and shooting are commonly used in soccer game. A naive approach is to implement these skills as basic skills. However, due to the complexity of coping with a dynamic competing environment at a low altitude, we choose to move these skills into a higher level to fully utilize the benefits of a broader view. The result is that we can learn both strategies and skills using high-level approaches.

2.3 Future Work

From our previous work, we understand how accuracy affects our agents' behaviors. In the future, we will center on improving the accuracy of our world model to provide a better foundation for strategic design. The following are some of the directions that we believe are promising in achieving our goals.

1. Designing automatic experiments to measure the real variance of the error of position instead of using ad hoc methods.
2. Using dynamic Kalman filter to achieve better noise reduction.
3. Making offline training and learning in world model and agent skills more systematic and automatic.
4. Using heuristics in probing the control space for optimal agent motion controlling.

3 New Research Areas

Based on our early work, we concentrate on some particular problems to solve. These problems are divided into three categories, but in fact there are a lot problems in each category. Some ideas below are currently not used in TsinghuaAeolus3D 2006, but we're planning to at least try some experiments on them.

3.1 Training and analyzing basic behaviors

Research on single-agent behavior is not as easy as its name might suggest. Because of its simplicity, we need more accurate analysis. In order to be used in decision making, the distribution of possible outcomes and their probabilities need to be analyzed. The possible outcomes (ball position and velocity, time elapsed when the ball is kicked) with their evaluations affect consecutive actions to be carried out, often considered in the tactic module, while the risk is used together with the team's mental status, which may change during the match, affected by the opponent. Given different requirement of action result, the actional generator produces different action sequences, while the risk and the possibility of changing objectives (which may greatly affect the action) are only considered in higher level.

Together with the action generator, a good tester, trainer and data analyzer are recommended. To make things simpler, the action generator is designed as a single input single output processor. In this way, predefined (or possibly joystick controllable) instructions can be used to test if the agent can carry out simple actions correctly.

To start a more complex research area, we chose some simple but typical multi-agent situations. These situation include shooting, 2v1 passing, corner kick and an offside trap. We apply reinforcement learning methods stated in [2] and [3], and some of them can have relatively good analytical or hard-coded solutions

3.2 Collaborative and adversarial behaviors

Once basic behaviors are learnt, we need to have good collaborative and adversarial considerations. [17] is a good starting point, [18] and [19] are also good references. Classical state-space search and real-time dynamic programming [6] are helpful, and real-time fuzzy decision tree [16] can be obtained via machine learning. In addition to this, rough sets make knowledge-based decision making more reasonable [20], and artificial immune system can be used in learning team behaviors [1]. Modular Q-learning [4] is also a good solution to some of the instances.

Communication is added this year. [12] provides a good paradigm, but there's still a lot to do to maximize the information retrieved via communication, helping to make decisions. A good reference for evolving communicating agents can be found in [9].

Opponent modelling is essential in finding different strategies for different opponents. Though it's very complex in nature, we made some effort in developing simple but useful opponent model and discussed some of the issues that might be considered in the future. The modelling is divided into two layers: parameter adjusting and action predicting. Parameters may affect everything from the way to generator basic actions, to the formation of the whole team. Action predicting is very hard so merely very few kinds are investigated, like predicting the passing target and dribbling directions. Currently a simple predictor is implemented, located in the situation manager discussed below. A special case of opponent modelling is described in [15].

3.3 Decision Making

This year, we developed a new decision making module, considering quite a few different kinds of options, from local perspective to the global one. This module can be seen as a set of suggestion maker, one evaluator, one decision making core and a decision decoder. The decoded decision will be simple enough for the low-level action generator to carry out.

There are three kinds of suggestion maker. The rule-based tactic manager (both off-line and on-line) gives some concrete action candidates to be evaluated and selected. Situation manager give warnings and instructions. For example it gives a list of "DON'Ts" (e.g don't be offside), and has the responsibility to keep a good formation. Parameter manager give additional adjustment, restrictions or requirements. First layer of opponent modelling is a part of it.

The evaluator sees if outputs of the suggestion makers are feasible adding special requirement needed to make the decision executed successfully. It also evaluates possible outcomes and corresponding risks.

The output of the evaluator is sent directly to the decision making core, which combines all the suggestions with the evaluations, to provide a single decision. It is worth mentioning that the core has its internal state, so the final decision does not depend on the world model only. Rather, the core records useful history

information to find a better decision. One of the possible construction of this adaptive action selector is described in [11].

When there are lot of plans that seems good, we have to consider actions interdependence[13] and the reliability of internal prediction and estimation [7]. But when there seems to be nothing good to do, heuristic evaluation of a small action set may help to make decision. The evaluation may be spatial information of the current match, while a good opponent model may increase the accuracy of the evaluation greatly. For both situations, heuristic planning [14] is a possible choice. In many situations we have to change our plan, so temporal abstraction is considered and used in reinforcement learning [8], to obtain better result.

3.4 Summary

In this paper, we briefly reviewed most important parts of TsinghuAeolus3D 2006 and relevant researches. While using largely computer science and techniques, we're still seeking for motivations from cognitive science [10] and real human decision making [21]. Though we're not using these ideas this year, we're most likely to research on this area in near future.

References

1. Colin Fyfe, Lakhmi Jain. Teams of intelligent agents which learn using artificial immune systems. *Journal of Network and Computer Applications* 29(2006) 147-159
2. Gerhard Weiß. Distributed reinforcement learning. *Robotics and Autonomous Systems* 15(1995) 135-142
3. R.Sun, T.Peterson. Multi-agent reinforcement learning: weighting and partitioning. *Neural Networks* 12(1999) 727-753
4. Kui-Hong Park, Yong-Jae Kim, Jong-Hwan Kim. Modular Q-learning based multi-agent cooperation for robot soccer. *Robotics and Autonomous Systems* 35(2001) 109-122
5. Chen K. Tham. Reinforcement learning of multiple tasks using a hierarchical CMAC architecture. *Robotics and Autonomous Systems* 15(1995) 247-274
6. Andrew G. Barto, Steven J. Bradtke, Satinder P. Singh. Learning to act using real-time dynamic programming. *Artificial Intelligence* 72(1995) 81-138
7. Yutaka Sakaguchi, Mitsuo Takano. Reliability of internal prediction/estimation and its application. I. Adaptive action selection reflecting reliability of value function. *Neural Networks* 17(2004) 935-952
8. Richard S.Sutton, Doina Precup, Satinder Singh. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence* 112(1999) 181-211
9. Hitoshi Iba. Evolutionary learning of communicating agents. *Journal of Information Sciences* 108(1998) 181-205
10. James S.Albus, Anthony J.Barbera. RCS: A cognitive architecture for intelligent multi-agent systems. *Annual Review in Control* 29(2005) 87-99
11. Satoshi Kurihara, Shigemi Aoyagi, Rikio Onai, Toshiharu Sugawara. Adaptive selection of reactive/deliberate planning for a dynamic environment. *Robotics and Autonomous Systems* 24(1998) 183-195

12. Philippe Pasquier, Brahim Chaib-draa. Agent communication pragmatics: the cognitive coherence approach. *Cognitive Systems Research* 6(2005) 364-395
13. Siamak Rajabi, D.Marc Kilgour, Keith W.Hipel. Modeling action-interdependence in multiple criteria decision making. *European Journal of Operational Research* 110(1998) 490-508
14. Josefina Sierra-Santibáñez. Heuristic planning: A declarative approach based on strategies for action selection. *Artificial Intelligence* 153(2004) 307-337
15. H.H.L.M.Donkers, J.W.H.M.Uiterwijk, H.J.van den Herik. Probabilistic opponent-model search. *information Sciences* 135(2001) 123-149
16. Han-Pang Huang, Chao-Chiun Liang. Strategy-based decision making of a soccer robot system using a real-time self-organizing fuzzy decision tree. *Fuzzy Sets and Systems* 127(2002) 49-64
17. Peter Stone, Manuela Veloso. Towards collaborative and adversarial learning: a case study in robotic soccer. *Int. J. Human-Computer Studies* 48(1998) 83-104
18. Barry Brian Werger. Cooperation without deliberation: A minimal behavior-based approach to multi-robot teams. *Artificial Intelligence* 110(1999) 293-320.
19. Milind Tambe, Jafar Adibi, Yaser Al-Onaizan, Ali Erdem, Gal A.Kaminka, Stacy C.Marsella, Ion Muslea. Building agent teams using an explicit teamwork model and learning. *Artificial Intelligence* 110(1999) 215-239
20. Zdzislaw Pawlak. Rough set approach to knowledge-based decision support. *European Journal of Operational Research* 99(1997) 48-57
21. Jean-Charles Pomerol. Artificial intelligence and human decision making. *European Journal of Operational Research* 99(1997) 3-25