CMμs 2018 Team Description

Philip Cooksey, Devin Schwab, Rui Silva, Rishub Jain, Yifeng Zhu, and Manuela Veloso

Carnegie Mellon University
{pcooksey,dschwab,yifengz2}@andrew.cmu.edu, {ruisilva,rishub}@cmu.edu, mmv@cs.cmu.edu

Abstract. In this paper, we present our team's recent work in expanding our team strategy to comply with the RoboCup Small-Size League's recent rule changes and in learning new policies for our robots. In regards to the new rules, we have expanded our offensive strategy to utilize all of the robots across the larger field size. Our defensive strategy was changed to handle the new rectangular defense area. In learning new policies, we focused on using machine learning to automate the process of creating different parts of the Skills, Tactics, and Plays hierarchy. Lastly, we discuss new robots and hardware.

1 Introduction

The CMμs 2018 team from Carnegie Mellon University (Figure 1) builds upon the CMDragons team, and its extensive research from previous years (1997–2010, 2013-2016 [1], [2]) in multiple areas, including path [3] and dribbling [4] planning, team architecture [5], among others. This year marked a big change for our team with the development of new robots and the commitment to a new research focus. Building upon our legacy, we will now drive our research from a more data and learning oriented perspective. This new research focus inspired the name change, with μ symbolizing one of the many parameters used in learning.

This paper presents the technical details of our most recent work, since the Team Description Paper of 2016 [2]. The overall architecture has remained largely unchanged since 2010 [6,7,5], and in this paper we focus on the novel contributions from this year. Our team website[1] provides a thorough description of the robots hardware and links to technical documents from previous years.

In the following sections, we present an overview of our overall team coordination architecture, and our most recent work related to improvements in our offense, defense, and hardware. We also provide an overview of our ongoing work in incorporating learning techniques into our team. Section 7 summarizes the contributions and discusses future work.

[1] http://www.cs.cmu.edu/~robosoccer/small/

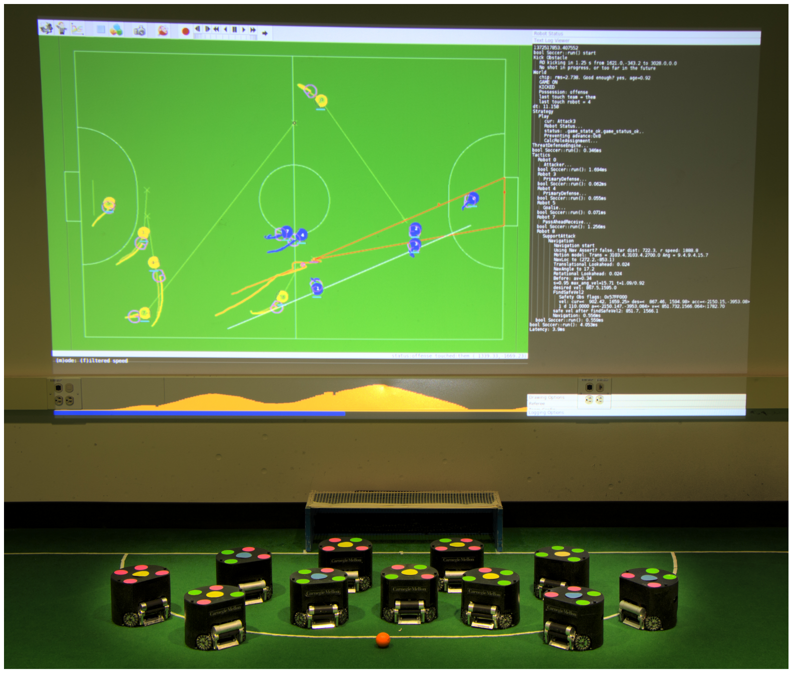

Fig. 1. CMμs team of soccer robots. In the background, our layered disclosure viewing and debugging tool [8].

2 Team Coordination Overview

At the core of controlling and coordinating our team is the Skills, Tactics and Plays (STP) architecture [5]. The main contribution of STP is the separation between single robot and team behavior, where team behavior results from executing a coordinated sequence of single robot behaviors for each team member. A *play* consists of a fixed team plan with a set of applicability and termination conditions, and a role for each team member. Each role defines a sequence of tactics, and is assigned dynamically at runtime through auctioning. A *tactic* encapsulates the behavior of a single robot. Specifically, a tactic determines the *skill state machine* to be executed by the robot, and sets the parameters to be used when using specific skills. Finally, a *skill* is a control policy for performing some complex action.

A key component to our team performance is our zone-based coordination, which dynamically divides the field in several zones to be assigned to different robots. This has advantages over position-based planning, since we can create and carry out plans independently of the opponents' starting positions [9]. Figures 2 and 3 depict an example of the dynamic zone-based coordination within a robot soccer match.

Fig. 2. Attacker shown in the red circle has the ball with the other two teammates located in their assigned dynamic zones.

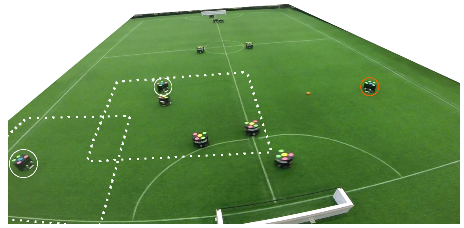

Fig. 3. Following Figure 2, a new dynamic zone is placed around the previous attacker and the teammate on the right is assigned the attacker role.

3 Offense

Following the rule changes made to the Small-Size League this year, we updated our offensive strategy to accommodate the new number of robots and field size, as well as the automatic ball placement requirement.

3.1 Increased Number of Robots and Field Size

In response to the 2018 RoboCup Small-Size League rules, we have put significant effort in updating our offensive team coordination strategy. In particular, we expanded our offense plays [5] to better take advantage of the increased number of robots and field size. The first consideration we made was balancing the number of robots on the offense and defense of our team. The second consideration regards the placement of the robots on the field, to efficiently make use of the increased field size. We also updated our zone-based coordination strategy [9], by creating new dynamic zones across the entire field.

Additionally, we are developing several new strategies for maintaining ball possession as our team moves the ball down the field towards the opponent's goal. Some of these strategies are described in more detail on the Learning section.

3.2 Automatic Ball Placement

Our offense can now handle the new rule on automatic ball placement. Taking advantage of existing skills and tactics, such as dribbling and safe navigation, our robots are now able to place the ball automatically on the field, at the referee's request.

4 Defense

Following the rule changes, our team has been overhauling our defense strategy. We started by updating our defensive plays to account for the new rectangular

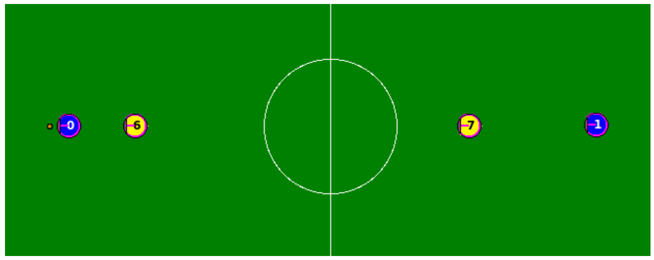

Fig. 4. Passing with Marking domain [10].

defense area. This entailed the specification of new positions for the defensive players, as well as a new coordinated movement strategy.

Additionally, the increased number of robots has also forced us to reevaluate our dynamic zone-based defense coordination scheme. In particular, we are working on making some of our defensive players more aggressive in regaining ball possession, and limiting the open angles to the goal.

5 Learning

As discussed before, our current architecture uses a hierarchical policy referred to as Skills, Tactics and Plays [5]. In previous years, all the components of this policy have been hand-coded using mostly human-designed heuristics. This year, as part of our new research focus, we are attempting to automate the creation of some parts of this hierarchical policy using machine learning and deep reinforcement learning techniques. Instead of designing and creating every low-level skill by hand, some of the skills will be learned directly by the robots. Similarly, some of the multi-agent coordination at the play level will be learned automatically.

We have also been attempting at using learning approaches for ball possession purposes. One of the shortcomings of human heuristics is at predicting the outcomes of using a tactic within the STP hierarchy. As such, we have been using deep learning techniques to predict the probability of success of *eventually* passing to a teammate when using different tactics, given the current state of the world [11]. The domain we used for testing is the *Passing with Marking* domain, a domain similar to keep-away, see Figure 4. In this domain, we have two blue and two yellow soccer robots. The objective is for the blue robot, that starts with the ball, to pass the other blue robot while the yellow robots attempt to steal and intercept the pass. To reiterate, we are attempting to learn the probability that the blue robot will successfully passing to its teammate using different tactics in this dynamic adversarial domain.

Additionally, we are utilizing state-of-the-art DeepRL algorithms such as DQN [12] and DDPG [13] to train tactics for the competition, and we have been testing these methods on a grid world Keepaway domain [14] where two teams play against each other.

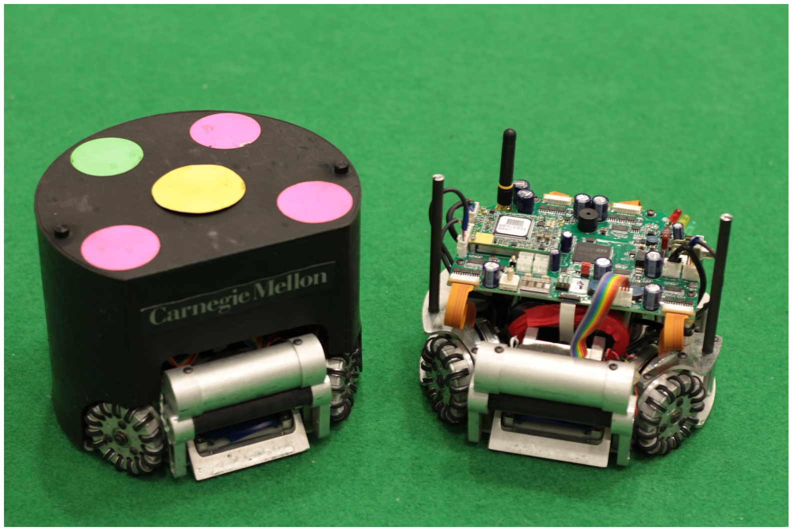

Fig. 5. A robot shown with and without protective cover.

6 Hardware

This year we are getting new robots for our team. These new robots will need to be modeled so that their physical abilities work properly with our previously developed tactics and skills. We will also explore enhancing the skills with the new robots' speed and acceleration. In particular, this will be important on our defense side and goal keeping.

We are also improving the way we control our robots over radio. Preliminary experiments seem to suggest that a team of eight robots puts too high of a strain on our existing radio hardware. We are buying new radio hardware in order to reliably control the eight robots on the field.

Regardless of the new robot hardware, we are still capable of competing with our current robots. A priority for our team is to ensure that we are backwards compatible with our existing robots. Figure 5 shows one of our current robots without the protective cover.

7 Conclusion

This paper presents the changes of the CMμs team since our last (CMDragons) Team Description Paper [2]. Our focus has been placed on updating our offense and defense strategy to handle the new 2018 RoboCup Small-Size League rules. Our new name is a symbol of the large changes we are making to our team with the development of new robots and a new research focus on applying learning methods to improve our team's performance.

References

1. J. Biswas, J.P. Mendoza, D. Zhu, S. Klee, and M. Veloso. CMDragons 2014 Extended Team Description Paper. In *Robocup 2014*.
2. Juan Pablo Mendoza, Joydeep Biswas, Danny Zhu, Richard Wang, Philip Cooksey, Steven Klee, and Manuela Veloso. CMDragons 2016 Extended Team Description Paper. In *Robocup 2016*.
3. M. Bruce, J.and Veloso. Real-time randomized path planning for robot navigation. In *IROS 2002*, pages 2383–2388.
4. S. Zickler and M. Veloso. Efficient physics-based planning: sampling search via non-deterministic tactics and skills. In *AAMAS 2009*, pages 27–33.
5. B. Browning, J. Bruce, M. Bowling, and M. Veloso. STP: Skills, tactics, and plays for multi-robot control in adversarial environments. *Proceedings of the IME, Part I: Journal of Systems and Control Engineering*, 219(1):33–52, 2005.
6. S. Zickler, J. Biswas, J. Bruce, M. Licitra, and M. Veloso. CMDragons 2010 Extended Team Description Paper. In *Robocup 2010*.
7. J. Biswas, J.P. Mendoza, D. Zhu, P.A. Etling, S. Klee, B. Choi, M. Licitra, and M. Veloso. CMDragons 2013 Team Description Paper. In *Robocup 2013*.
8. Patrick Riley, Peter Stone, and Manuela Veloso. Layered disclosure: Revealing agents' internals. In *ATAL-2000*, July 2000.
9. Juan Pablo Mendoza, Joydeep Biswas, Danny Zhu, Philip Cooksey, Richard Wang, Steven Klee, and Manuela Veloso. Selectively reactive coordination for a team of robot soccer champions. In *Proceedings of AAAI'16, the Thirtieth AAAI Conference on Artificial Intelligence*, February 2016.
10. Philip Cooksey, Juan Pablo Mendoza, and Manuela Veloso. Opponent-Aware Ball-Manipulation Skills for an Autonomous Soccer Robot. In *Proceedings of the RoboCup Symposium*, Leipzig, Germany, July 2016. Springer. Nominated for Best Paper Award.
11. Philip Cooksey and Manuela Veloso. Intra-robot replanning to enable team plan conditions. In *Proceedings of IROS'17, the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vancouver, Canada, September 2017. An earlier version was presented at the Workshop on Planning and Robotics, at ICAPS'17, June, 2017, Pittsburgh PA.
12. Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.
13. Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.
14. Peter Stone, Gregory Kuhlmann, Matthew E Taylor, and Yaxin Liu. Keepaway soccer: From machine learning testbed to benchmark. In *Robot Soccer World Cup*, pages 93–105. Springer, 2005.