The RoboFlag SURF Competition: Results, Analysis, and Future Work

Julie A. Adams[1] and Adam T. Hayes[2]
[1] Department of Computer Science
Rochester Institute of Technology. Rochester, NY 14623 USA
adamsj@ieee.org
[2] Department of Computer Science
California Institute of Technology, Pasadena, CA 91125, USA
athayes@caltech.edu

Abstract

The culmination of the 2002 RoboFlag Summer Undergraduate Research Fellowship program, jointly operated between California Institute of Technology and Cornell University, was a final competition between two teams of three undergraduate researchers. After ten weeks of preparation, Team Pasadena defeated Team Ithaca in two of the three final games. This paper provides the detailed results of the competition, an analysis of the competition, and reviews the future work.

1. Introduction

As autonomous and semi-autonomous vehicles operating in uncontrolled environments work their way into practical use, the need arises to build mechanisms through which humans can simultaneously interact with multiple individual units. The RoboFlag concept, which pits two semi-autonomous teams of robots against each other in a game similar to capture-the-flag [1], is being developed to examine the influence of different interaction and control techniques on task performance. During the summer of 2002, two teams of three undergraduates participated in the California Institute of Technology's Summer Undergraduate Research Fellowship program [2]. The teams were given the task of designing initial robot controllers and human-robotic interfaces for RoboFlag, with the goal being to emerge victorious in a head-to-head competition held at the end of the summer [3, 4].

2. Competition Results

The final competition between team Pasadena and team Ithaca consisted of three games, each composed of two twelve-minute halves. In each of these games, the teams employed human operators to guide groups of six robots in a capture-the-flag type of competition. The game half time was fifteen minutes. There was a short break between games one and two, and a two-hour break between games two and three. The teams were permitted to modify their systems during half time as well as the game breaks.

There are several ways to score points in RoboFlag, see [1] for a full description of the game rules. Briefly, tagging an opponent is worth one point, an opponent arbiter tag is worth ten points, each inactive opponent is worth ten points, a flag capture scores five points, and returning the captured flag to the home zone is worth twenty-five points. The following tables show the scoring breakdown for the three games.

As the Table 1 indicates, team Pasadena won the first game by a total score of 1054 to team Ithaca's 506 points. The first half of the game was very close,

Table 1. Game one results.

	First Half		Second Half	
	Pas.	Ithaca	Pas.	Ithaca
Tagged Opponents	5	9	14	8
Opponent Arbiter Tags	2	2	7	7
Inactive Opponents	4	5	4	2
Flag Captures	10	11	23	10
Flag Returns	7	6	21	3
Total Points	290	284	764	223
Final Score	**1054**	**506**		

Table 2. Game two results.

	First Half		Second Half	
	Pas.	Ithaca	Pas.	Ithaca
Tagged Opponents	9	6	6	5
Opponent Arbiter Tags	1	3	2	1
Inactive Opponents	5	3	4	3
Flag Captures	12	9	16	6
Flag Returns	8	4	11	3
Total Points	329	211	421	150
Final Score	**750**	**361**		

but team Ithaca lost the majority of its robots (four out of six) in the second half. This situation allowed team Pasadena to clearly out capture and return the flag over team Ithaca. Team Pasadena had a flag capture to return ratio of 91% in the second half. This was the highest capture to return ratio during the entire competition. In total, team Pasadena captured the flag thirty-three times with twenty-seven successful flag returns while team Ithaca had twenty-one captures and nine successful flag returns. Team Pasadena tagged team Ithaca a total of nineteen times, twice more than team Ithaca tagged team Pasadena. Team Ithaca concluded both halves with four inactive robots. All but one of team Pasadena's robots were inactive at the end of the first half, but during the second half, they lost only two robots.

Team Pasadena again defeated team Ithaca in the second game by a score of 750 to 361 points. As Table 2 indicates, once again the first half of the game was fairly close. It was during the second half that team Pasadena substantially pulled ahead of team Ithaca. Team Pasadena tagged team Ithaca four times more than team Ithaca tagged team Pasadena. The number of inactive robots for each team at the end of the halves is similar to game one. Team Ithaca captured the flag only fifteen times during the entire game. Team Ithaca did slightly improve their capture to return ratio from 43% in the first game to 46% in this game. At the same time, team Pasadena captured the flag a total of twenty-eight times returning the flag to their home zone nineteen times for a capture to return ratio of 68%.

Table 3. Game three results.

	First Half		Second Half	
	Pas.	Ithaca	Pas.	Ithaca
Tagged Opponents	11	18	11	14
Opponent Arbiter Tags	3	11	7	4
Inactive Opponents	2	3	0	5
Flag Captures	8	18	1	25
Flag Returns	2	12	1	17
Total Points	151	548	111	654
Final Score	**262**	**1202**		

Team Ithaca obtained revenge in the final game of the competition. They handedly defeated team Pasadena, out

scoring team Pasadena by almost 1000 points. The final game score was 1202 points for team Ithaca and 262 points for team Pasadena. As indicated by Table 3, team Ithaca captured the flag a total of forty-three times, the highest number of captures by either team for a single game. Team Ithaca improved their capture to return ratio to 67%, for twenty-nine flags returned to their home zone. Team Ithaca also tagged team Pasadena a total of thirty-two times compared to team Pasadena's twenty-two tags of the team Ithaca robots. One significant difference can be seen in the second half. Team Pasadena finished the half with only one active robot while team Ithaca finished with all their robots active. This was a clear advantage to team Ithaca and allowed team Ithaca to out capture team Pasadena twenty-five to one in the second half.

Team Pasadena won the competition two games to one, but team Ithaca clearly demonstrated their capabilities in the third game. Note that the winning team in each game also scored the most points during each half of the game.

A factor that attributed to team Ithaca's demise during the first two matches was the team's addition of "one last behavior" the evening before the competition without fully testing it. The behavior was designed to allow their robots to successfully navigate the field at almost 1.0 m/s. This change would have provided an enormous advantage over team Pasadena, whose navigation algorithms functioned reliably at less than 0.5 m/s. During the first game, however, team Ithaca found the high-speed behavior interfered with some of their other behaviors, causing their robots to crash into obstacles and become inactive. They unsuccessfully modified the behavior between games one and two. During game two, team Ithaca found that their recent modifications exacerbated the problem. Team Pasadena was able to assess the unfortunate state of its opponent early on in each game, and thus was able to steer a conservative course to victory. This strategy mostly entailed waiting until team Ithaca had lost enough of their active robots to properly defend their flag.

There were two primary reasons that enabled team Ithaca to soundly trounce team Pasadena in the third game. First, during the break between the second and third games, team Ithaca removed the high-speed behavior, ensuring that their robots stayed active longer. The second reason team Ithaca did so well was that team Pasadena, realizing that they could not simply wait for their opponent to crash and burn, decided to implement their own special strategy that involved the purposeful placement of inactive robots on the opponent's side of the field to restrict defensive movement. However, again due to a lack of preparation (and resulting poor execution of robot placement), team Pasadena's strategy backfired and simply resulted in most of the team becoming inactive. Since there were fewer opponents playing defense, team Ithaca was able to significantly outscore their opponent. The portion of team Ithaca's score attributed to flag captures and returns accounted for 940 of the team's 1202 points.

3. Final Competition Analysis

Overall, the competition showed that the teams were able to develop systems that provided an exciting competition in only ten weeks. Throughout the SURF program, modifications to the RoboFlag rules were

made as detailed in [1]. Slight modifications were made just days before the competition, including a reduction in the number of robots on each team as well as the number of obstacles on the field. These changes were motivated by practice games that the teams played in the days leading up to the final competition.

During the final competition, a clear pattern emerged. Even though the number of total robots on the field had been reduced from twenty-four to sixteen (six robots on each team plus four obstacles), the center of the field quickly became congested with obstacles and inactive robots. While the obstacles move in a random manner, they tended to gravitate towards the center of the field, and as a result, the majority of inactive robots were positioned in the center of the field. The placement of inactive robots greatly influenced the development of the game, as they could either act as a defensive blockade, preventing attacking robots from reaching the flag (see Figure 1a), or as a defender blockade, preventing defensive robots from reaching the attackers path (see Figure 1b). At the same time, the refueling task could also be complicated by the placement of inactive robots, possibly preventing robots from refueling, which would render them inactive (and further cluttering the playing arena). As discussed earlier, the initial stages of the games were largely defensive, and as the blockades of inactive robots developed (often largely by accident), the team with the favorable configuration could exploit that configuration to score a large number of points in the waning minutes of the half.

It is clear from this competition that the concepts of tagged robots and the need to refuel robots affect the game play. If a team has the majority, if not all, of its

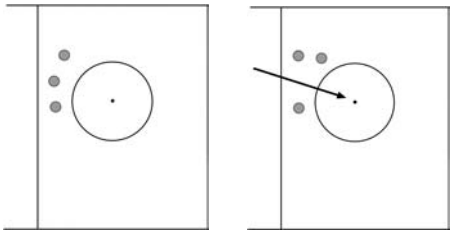

Figure 1: Inactive robots (shown in gray) can either (a) serve as defensive blockades or (b) shield defenders from approaching the attack lanes.

active robots tagged or inactive, then the opponent is usually able to take extreme advantage of the situation. It was not uncommon to observe the opponent repeatedly capturing the flag and returning the flag to its home zone during the final stages of each half. The result is a rapid change in the game score.

The need to refuel the robots also provides the opportunity for an opponent to quickly accumulate points. If a team appropriately manages fuel usage, they can maintain an offensive or defensive strategy with little detriment to their strategy when robots need to refuel. On the other hand, if the team's robots all require fuel at approximately the same time, then all the robots will head for their home zone. This action tends to be a detriment to the current strategy.

4. Future Work

The SURF program and final competition provided a plethora of information that will guide modifications to the RoboFlag game as well as additional research topics.

Throughout the entire SURF program the teams felt that the number of robots (eight per team, and eight obstacle robots) was too many given the field size. As a result, the final competition used six

robots per team and four obstacle robots. Future work will include determining the appropriate team size, and number of obstacles. After the final competition, both teams felt that the size and placement of the defense zones contributed to the large number of inactive robots positioned in the center of the field. Suggestions are to assess shrinking the defense zone size as well as moving the defense zones further away from the team's home zone and the mid-field line, thus creating more separation between the two defense zones and opening up the field of play. The SURF program and the final competition have highlighted other possible rule changes. An effort is underway to assess the impact of various rule changes on the dynamic and parallel nature of game play. The result will be an updated RoboFlag rule set. For the most current version of the rules and competition videos please see [5].

Acknowledgements

The work described in this paper was supported in part by: the Defense Advance Research Projects Agency under cooperative agreement F30602-01-2-0577; the Air Force Research Laboratory, Information Directorate, PECASE award F49620-02-0388; and the Caltech SURF Program.

References

[1] D'Andrea, R.; and Murray, R.M. 2003. The RoboFlag Competition. 2003 American Controls Conference. Denver, CO.
[2] Caltech SURF Program: http://www.its.caltech.edu/~surf
[3] Zigoris, P., Siu, J., Wang, O., and Hayes, A.T. 2003. Balancing automated behavior and human control in multi-agent systems: a case study in RoboFlag. 2003 American Controls Conference, Denver, CO.
[4] Chamberlain, L., Tang, J., Watugala, M., Adams, J.A., and Babish, M. 2003. A behavioral architecture for strategy execution in the RoboFlag Game. 2003 American Controls Conference, Denver, CO.
[5] The RoboFlag Homepage: http://roboflag.mae.cornell.edu/