Statistical Analysis in MiroSot

J. L. Baxter[1], J. M. Garibaldi[1], E. K. Burke[1], M. Norman[2]
[1] Automated Scheduling, Optimization and Planning (ASAP) Research Group,
School of Computer Science and Information Technology,
University of Nottingham, Jubilee Campus, Wollaton Road,
Nottingham, NG8 1BB, United Kingdom.
{jlb,jmg,ekb}@cs.nott.ac.uk
[2] Merlin Systems Corp Ltd, ITTC Tamar Science
Plymouth, PL6 8BX, United Kingdom.
{marknorman}@merlinsystemscorp.co.uk

Abstract

The world of robot soccer is a subjective one due to the difficult nature of collecting data. In this paper we present a number of non-parametric statistical tests that can be used to analyze the performance of strategies without the need for a large sample size. We present results that show a "case-based" strategy performs significantly better (statistically) than a "potential field" strategy in simulation based on a sample of only ten games.

1 Introduction

When developing strategies for MiroSot Robot Soccer, comparing different strategies can be very subjective, e.g. *"Some teams dont even detect opponents... One of the successful elements in the defensive system was our ability to block opponents"* [1]. There is a need for more objective statistical analysis of the strategies. However, this can be difficult for a number of reasons:

- MiroSot games last ten minutes, once you add setup time and stoppage time having a large sample of games is problematic.
- The FIRA Simulator (Figure 1), like its real-world counterpart, requires a human referee, thus simulated matches cannot be run in batches.

This leaves researchers with the undesirable situation of having small sample sizes. This is a problem as parametric statistical tests include the following assumptions (taken from [2]):

- *"Observations are drawn from Normally distributed populations.*
- *These populations have the same variance."*

Figure 1: The FIRA Simulator

The larger the sample size the less important it becomes to validate these assumptions. However with such small sample sizes we cannot assume the above. Non-parametric statistical tests however do not specify conditions about the parameters of the population from which the sample is drawn. Some assumptions *are* made - for example, that the observations are independent - but these assumptions are fewer and weaker than those associated with parametric statistical tests. As they do not assume that the data under analysis is drawn from a population distributed in a certain fashion, the non-parametric statistical tests used in this paper negate the need to have large sample sizes.

The amount of statistical analysis in MiroSot is limited, as most teams rely upon subjective observations and competitive results in order to analyze their strategies, for example [1], or provide experiments with simplistic results. For example the data in [3] only consists of "goals scored by own team".

The following sections describe the statistical tests used, experiments undertaken and the results obtained. The paper will end with a discussion of the results and possible future work.

2 Experiments

In our experiments we were looking for statistical differences between two strategies (can we say statistically if one strategy is better than another). These strategies are; a *case-based* strategy and a *potential field* strategy. The analysis of the strategies was in two forms. Firstly the strategies were run against a number of *basic* strategies, each strategy being run against each basic strategy ten times, the goal difference (goals scored - goals conceded) was recorded for analysis (using the Kolmogrov-Smirnov test). This test shows how the strategies perform against different levels of basic strategies. Secondly each strategy played ten games against each other, with the number of wins/losses recorded for analysis (the Binomial test was in this instance). This test shows how well the strategies performed against each other. The Binomial test and the Kolmogrov-Smirnov test are now described.

2.1 Binomial Test

The Binomial test tells us if the frequencies we observe in a sample could have been drawn from a population having a specified value of P e.g. the proportion of expected wins for the potential field strategy and $Q = 1 - P$ e.g. the proportion of expected defeats for the potential field strategy. The probability of getting x objects in one classification and $N - x$ objects in another classification is given by Equation 1. To calculate the probability of obtaining the observed values and more extreme ones we use Equation 2. The null hypothesis (H_0) is that there is no difference between either strategy winning a match i.e. $P = Q = \frac{1}{2}$; the alternative hypothesis (H_1) is that the case-based strategy is better than the potential field strategy and therefore wins more matches, $P < Q$. As we speculate in advance that the case-based strategy will be better than the potential field strategy we will be using a one-tailed test.

$$p(x) = \binom{N}{x} P^x Q^{N-x}, \binom{N}{x} = \frac{N!}{x!(N-x)!} \quad (1)$$

$$\sum_{i=0}^{x} \binom{N}{i} P^i Q^{N-i} \quad (2)$$

If the result (p) obtained is equal to or less than the level of significance (α) chosen, defined as the probability value of a false rejection of the null hypothesis, we can reject H_0 in favor of H_1.

2.2 Kolmogrov-Smirnov Test

The Kolmogrov-Smirnov test is used to determine whether two independent samples have been drawn from the same population. A one-tailed test is used to decide if the values from one group are larger than that of the other group. As we are interested in which strategy is best when each is compared against the basic strategies, we will be using this one-tailed test. The null hypothesis (H_0) is that there is no difference between the goal differences. The alternative hypothesis (H_1) is the case-based strategy has a better goal difference than the potential field strategy.

To apply the test a cumulative frequency distribution for each sample of observations is made. For each interval one step function is subtracted from the other. $S_{n_1}(X)$ is the observed cumulative step function of one of the samples, that is $S_{n_1}(X) = K/n_1$, where K is the number of scores equal to or less than X. $S_{n_2}(X)$ is the observed cumulative step function for the other sample, that is $S_{n_2}(X) = K/n_2$. The test now focuses on Equation 3.

$$D = maximum[S_{n_1}(X) - S_{n_2}(X)] \quad (3)$$

If by computation we find that D is larger or equal to the *numerator of the largest distance between the two cumulative distributions* (K_D) we can reject H_0 in favor of H_1.

2.3 Basic Strategies

As stated earlier each strategy will play against a number of basic strategies. There are four such strategies, each strategy will play against each basic strategy in ten matches. These matches will obey the MiroSot middle sized league robot soccer rules. The following basic strategies are implemented:

- *Static Opponents* - opponent players are just static objects within the field of play.

- *Goal Keeper* - the opponent goal keeper is fully functional.

- *Defenders* - in addition to the goal keeper the opponent defenders (two) attempt to clear the ball from their half of the pitch.

- *Forwards* - the entire opponent team is active (two forwards added). In addition to attempting to clear the ball from their half the team attempts to score goals in the opponents half.

2.4 Results

Table 1 shows the goal differences for the case-based and potential field strategies against the *Static Opponents* basic strategy. Table 2 shows the cumulative frequency distribution of the goal difference obtained against the *Static Opponents* for both the case-based and potential field strategies. Tables 3-8 present the same as just described for the *Goal Keeper, Defenders* and *Forwards* respectively.

Table 2 below shows that the $Maximum[S_{n_1}(X) - S_{n_2}(X)]$ for the *Static Opponents* results is $\frac{3}{10}$. Table 4 shows that the $Maximum[S_{n_1}(X) - S_{n_2}(X)]$ for the *Goal Keeper* results is $\frac{6}{10}$. Table 6 shows that the $Maximum[S_{n_1}(X) - S_{n_2}(X)]$ for the *Defenders* results is $\frac{7}{10}$. Table 8 below shows that the $Maximum[S_{n_1}(X) - S_{n_2}(X)]$ for the *Forwards* results is $\frac{6}{10}$. As in each case $n_1 = n_2$ and n_1 and n_2 are less then forty, the K_D is 6 at an α of 0.05 (see Table 9). Therefore H_0 is rejected in favor of H_1 in three of the four cases (*Goal Keeper, Defenders, Forwards*) and it is shown that the case-based strategy has a proportionally higher goal difference. In the *Static Opponents* case H_0 is accepted and it is shown that both strategies have the same goal difference.

Table 10 shows the results of ten matches between the case-based strategy and the potential fields strategy. From the table it is observed that the potential field strategy won four matches i.e. x = 4. Substituting into Equations 1 and 2 the following was obtained:

$$\sum_{i=0}^{4} \binom{10}{i} 0.5^i 0.5^{10-i} = 0.376953125 \qquad (4)$$

This shows that there is a probability of 0.38 (2.d.p) that the potential field strategy would win four or less matches, if $P = Q = \frac{1}{2}$. Using $\alpha = 0.05$ we accept H_0. H_1 would only be accepted and H_0 rejected if the resultant probability was equal to or less than α.

Match	**Case-Based**	**Potential Field**
1	12	22
2	27	26
3	24	24
4	26	22
5	13	18
6	18	16
7	27	21
8	21	6
9	27	21
10	18	23

Table 1: Static Opponents: Goal differences for Case-Based and potential field strategies

	Goal difference			
	6-8	9-11	12-14	15-17
$S_{n_1}(X)$	$\frac{0}{10}$	$\frac{0}{10}$	$\frac{2}{10}$	$\frac{2}{10}$
$S_{n_2}(X)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{0}{10}$
	18-20	21-23	24-26	27-29
$S_{n_1}(X)$	$\frac{4}{10}$	$\frac{5}{10}$	$\frac{7}{10}$	$\frac{10}{10}$
$S_{n_2}(X)$	$\frac{3}{10}$	$\frac{8}{10}$	$\frac{10}{10}$	$\frac{10}{10}$
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{3}{10}$	0

Table 2: Static Opponents: Kolmogrov-Smirnov cumulative frequency distribution. $S_{n_1}(X)$ = potential fields step function, $S_{n_2}(X)$ = basic step function.

Match	**Case-Based**	**Potential Field**
1	16	4
2	16	11
3	11	11
4	19	13
5	8	4
6	14	9
7	8	5
8	13	8
9	16	13
10	16	4

Table 3: Goal Keeper: Goal differences for Case-Based and potential field strategies

	Goal difference			
	4-5	6-7	8-9	10-11
$S_{n_1}(X)$	$\frac{0}{10}$	$\frac{0}{10}$	$\frac{2}{10}$	$\frac{3}{10}$
$S_{n_2}(X)$	$\frac{4}{10}$	$\frac{4}{10}$	$\frac{6}{10}$	$\frac{8}{10}$
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{4}{10}$	$\frac{4}{10}$	$\frac{4}{10}$	$\frac{5}{10}$
	12-13	14-15	16-17	18-19
$S_{n_1}(X)$	$\frac{4}{10}$	$\frac{5}{10}$	$\frac{9}{10}$	$\frac{10}{10}$
$S_{n_2}(X)$	$\frac{10}{10}$	$\frac{10}{10}$	$\frac{10}{10}$	$\frac{10}{10}$
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{6}{10}$	$\frac{5}{10}$	$\frac{1}{10}$	0

Table 4: Goal Keeper: Kolmogrov-Smirnov cumulative frequency distribution. $S_{n_1}(X)$ = potential fields step function, $S_{n_2}(X)$ = basic step function.

	Goal difference			
	(-1)-1	2-4	5-7	8-10
$S_{n_1}(X)$	$\frac{0}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{3}{10}$
$S_{n_2}(X)$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{10}{10}$
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{7}{10}$
	11-13	14-16	17-19	
$S_{n_1}(X)$	$\frac{5}{10}$	$\frac{7}{10}$	$\frac{10}{10}$	
$S_{n_2}(X)$	$\frac{10}{10}$	$\frac{10}{10}$	$\frac{10}{10}$	
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{5}{10}$	$\frac{3}{10}$	0	

Table 6: Defenders: Kolmogrov-Smirnov cumulative frequency distribution. $S_{n_1}(X)$ = potential fields step function, $S_{n_2}(X)$ = basic step function.

Match	**Case-Based**	**Potential Field**
1	10	8
2	16	9
3	18	9
4	12	8
5	9	7
6	14	8
7	4	8
8	12	0
9	17	-1
10	19	5

Table 5: Defenders: Goal differences for Case-Based and potential field strategies

Match	**Case-Based**	**Potential Field**
1	6	0
2	1	-2
3	5	-4
4	7	-5
5	2	4
6	5	6
7	0	-2
8	5	2
9	10	-1
10	2	-7

Table 7: Forwards: Goal differences for Case-Based and potential field strategies

	Goal difference		
	(-7)-(-6)	(-5)-(-4)	(-3)-(-2)
$S_{n_1}(X)$	$\frac{0}{10}$	$\frac{0}{10}$	$\frac{0}{10}$
$S_{n_2}(X)$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{5}{10}$
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{5}{10}$
	(-1)-0	1-2	3-4
$S_{n_1}(X)$	$\frac{1}{10}$	$\frac{4}{10}$	$\frac{4}{10}$
$S_{n_2}(X)$	$\frac{7}{10}$	$\frac{8}{10}$	$\frac{9}{10}$
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{6}{10}$	$\frac{4}{10}$	$\frac{5}{10}$
	5-6	7-8	9-10
$S_{n_1}(X)$	$\frac{8}{10}$	$\frac{9}{10}$	$\frac{10}{10}$
$S_{n_2}(X)$	$\frac{10}{10}$	$\frac{10}{10}$	$\frac{10}{10}$
$S_{n_1}(X) - S_{n_2}(X)$	$\frac{2}{10}$	$\frac{1}{10}$	0

Table 8: Forwards: Kolmogrov-Smirnov cumulative frequency distribution. $S_{n_1}(X)$ = potential fields step function, $S_{n_2}(X)$ = basic step function.

	One-tailed test	
N	$\alpha = .05$	$\alpha = .01$
10	6	7

Table 9: Extract from Table of Critical Values of K_D in the Kolmogrov-Smirnov Two-Sample Test (Table L) in [2]. N = size of sample.

Match	**Case-Based v Potential Field**
1	8 - 9
2	11 - 8
3	4 - 5
4	12 - 3
5	8 - 4
6	4 - 7
7	15 - 5
8	8 - 7
9	3 - 4
10	15 - 6

Table 10: Case-Based v Potential: Results

3 Discussion

The results show that the probability of either strategy winning a match against one another is approximately even, despite the strategies being different in the application of coordination (by coordination we are referring to the method used to move the players around their environment and interact with objects). The case-based strategy used absolute and predicted positioning provided by the vision system, whereas the potential field strategy uses charged particle representation of the environment. However the underlying strategies were almost identical for example, both strategies required a goal keeper, defender, forward and two wingers. Perhaps therefore the results obtained are not so surprising.

The process of discovering *why* a strategy is better than another is still however a subjective one, as the statistics cannot tell you why a strategy performs as it does. However there are still dangers from conducting subjective observations. For example if an observer is looking for a group of robots showing specific behaviors, there is a tendency that these behaviors will be observed whether intentional or not. However, in our case, observations of the matches are a very useful tool for answering why the potential field strategy performed so badly in the simulation against the basic strategies. It was observed that the robots were suffering from the infamous problems of potential fields (taken from [4]):

- Trap Situations - players get stuck in local minima and perform cyclic actions.
- No passage between obstacles - players cannot move in between players and therefore must move around them.
- Oscillations in presence of obstacles - players oscillate between areas of similar charge.
- Oscillations in narrow passage - players in narrow passages oscillate between charges rather than moving forwards.

This resulted in players from the potential fields strategy being caught in local rather than global minima, causing them not to be in the positions necessary to defend or attack the ball. In this case local minima is defined as the players need to avoid obstacles and the global minima is defined as the players need to get the ball.

It is hoped that this approach to team analysis will be used to choose which strategy will be used at the

FIRA World Cup 2005 in Singapore. These statistical tools will also be used to analyze the performance of the strategy at the championship. As the Kolmogrov-Smirnov test can be utilized with a minimum sample size of three and $\alpha = 0.05$ (given in [2] Table L) is is possible to compare results against another team if both teams compete against the same opposition, in at least three matches. The Binomial test could also be used in theory on a sample size as small as 5 ([2] Table D) if $P = Q = \frac{1}{2}$. In addition to a tool for analyzing entire strategies it is hoped that in the future it will be possible to analyze the performance of individual players. This will enable the improvement of individual behaviors as well as group behaviors. The use of other data in statistical tests could also be investigated. For example, the possession time of each team could be recorded and the data examined to see the impact if any on the result of matches.

As the FIRA simulator does not model the Nottingham robot system e.g. the robots modeled in the simulation are the "Yujin Robots" [7] not the Miabot Pro robots [8] used by the Nottingham team, and it is not practical to complete all testing and analysis in the real-world (robots have limited battery life, as such running continuous samples is not feasible), the Nottingham team will have to develop its own simulator modeled upon its own system. Taking the view of Brooks [5, 6] on simulations, "toy environments lead to toy robots" statistical analysis will be carried out in the real world in the future. The correlation between data collected in simulation and the real world environments will be interesting to observe. This will be could be done using the Median test - have two samples been drawn from populations with the same median.

Acknowledgments

The Simulation software was provided by FIRA, http://www.fira.net/soccer/simurosot/overview.html.

References

[1] R. A. Seesink, *Artificial Intelligence in a multi agent robot soccer domain*, Masters thesis, University of Twente, 2003.

[2] S. Siegel, *Non-parametric Statistics*, McGraw-Hill, 1956.

[3] Yan-Duo Zhang & Feng Min, *Application of Reinforcement Learning Based on Artificial Neural network to Robot Soccer*, in Proc. of 2004 FIRA World Congress, Busan, Korea, 2004.

[4] Y. Koren and J. Borenstein, *Potential Field Methods and their Inherent Limitations in for Mobile Robot Navigation*, Proc. IEEE Conf. Robotics and Automation, April 1991, pp 1398-1404.

[5] R. A. Brooks, *A Robust Layered Control System for a Mobile Robot*, MIT AI Lab Memo No. 864, September 1985.

[6] R. A. Brooks, *Achieving Artificial Intelligence Through Building Robots*, MIT AI Lab Memo No. 899, May 1986.

[7] http://www.fira.net/soccer/simurosot/overview.html.

[8] http://www.merlinsystemscorp.com.