Human Rating Methods on Multi-agent Systems

Chairi Kiourt[1,2(✉)], Dimitris Kalles[1], and George Pavlidis[2]

[1] School of Science and Technology,
Hellenic Open University, 26335 Patras, Greece
{chairik, kalles}@eap.gr
[2] "Athena" Research Centre,
University Campus at Kimmeria, 67100 Xanthi, Greece
gpavlid@ceti.gr

Abstract. Modern artificial intelligence approaches study game playing agents in multi agent social environments, in order to better simulate the real world playing behaviors; these approaches have already produced promising results. In this paper we present the results of applying human rating systems for com petitive games with social activity, to evaluate synthetic agents' performance in multi agent systems. The widely used *Elo* and *Glicko* rating systems are tested in large scale synthetic multi agent game playing social events, and their rating outcome is presented and analyzed.

Keywords: Multi agent systems · Rating systems · Game playing

1 Introduction

Since complex problems began to be studied as Multi-Agent Systems (MAS), the study of Social Learning (SL) has become more exciting [1, 2]. Diverse scientific areas such as sociology, economics, computer science, mathematics and marketing use social learning as an Artificial Intelligence (AI) tool for developing MAS [2]. Ferber [3] shows that the two extremes of the Social Organizations (SO), *cooperation* and *competition*, may be studied autonomously or as a combined social organization, which depends on the case study. As it is quite usual in such cases, the social environments are being populated with game playing agents [4]. For a game agent, social environment is represented by a game with all its components and entities [3, 4]. Learning in a game is said to occur when an agent changes a strategy choice in response to new information, and thus mimicking human behavior [4 6]. All those studies and many others support that the simulation of complex social environments and the analysis of their data become an intractable problem of agent social learning mechanisms. In addition, due to the continuous evolution of the dynamic systems that attempt to better simulate the human behaviors and habits, there have been some attempts to apply human rating systems to evaluate virtual agents and assess their performance [7]. Among the most widely used human rating systems are *Elo* [8] and *Glicko* [9]. Generally, different rating systems may disagree about players' absolute performance but could report similar ranking results with small deviations on specific events, like tournaments, for example [10].

The contribution of the study presented in this paper is to demonstrate how these two human rating systems perform in multi-agent systems which try to enhance the potential of the social events for the purposes of learning in unknown environments. Our study shows that although these rating systems seem to be adequate and useful for MAS evaluation, they also tend to not always be consistent. Also, it should be highlighted that the simulation of human behavior with synthetic agents is far from being accurate. By comparing the selected rating systems in MAS, it was found that they do not agree in several agents' ratings. Since Glicko (*v2*) was introduced as an improvement of Elo it was expected that the two ratings would be fairly similar, but, as it turned out, various inconsistencies have been recorded in the experimentations, such the large deviations in various agents' ratings players, which presents an ambiguity for their effectiveness in multi agents systems.

The rest of this paper is structured in three sections. The next two sections provide a brief background of the selected rating systems, the game used for the experiments and the game-playing social environment. The fourth section describes our experimentation on multi agent systems, also highlights the comparison of the rating systems. The last section presents our conclusions and our scheduled future work.

2 Performance Rating

Rating systems were first used in chess to calculate an estimate the strength of a player, based on player's performance against the opponent. The Ingo and Harkness system was the first chess player rating system [11]. It was first used to allow the members of the United States Chess Federation (USCF) to track their individual progress in terms other than tournament wins and losses [11].

The *Elo rating system* was first introduced by Arpad Elo in 1960 as a simple skill calculation of players, based on their wins and losses, and of their opponents in chess [8]. Chess, however, is a competitive two-agent system, where each agent's performance is based solely on its skill. In multi-agent systems, it is used as a calculation of fitness for many different learning or search algorithms, with promising results [7].

The Elo system assumes that each player has a skill that is drawn from a random distribution (an agent may have a "good" game or may have a "bad" game); it attempts to find the center of that distribution and converge to that value. The calculation is performed after each match, in a game between two agents A and B. Each agent has a current rating, R_A for agent A and R_B for agent B. Unrated players, generally start with a rating of *800 Elo*, which is associated to bad playing or a beginner level. Rating also depends on the tournament type and the players' attributes.

The *Glicko rating system* was first introduced by Mark Glickman in 1995 as an improvement of the Elo rating system [9]. The Glicko (*v2*) rating system is a method for assessing a player's strength in games of skill, such as *chess* and *go*. The main contribution of this measurement method is "ratings reliability", the so-called *ratings deviation* (RD). RD measures the accuracy of players rating. After a game, the amount of the ratings change depends on the RD: the change is smaller when the players' RD is low, and also when their opponents' RD is high. The RD itself decreases after playing a game, but increases slowly over time of inactivity.

The Glicko rating system was improved by its inventor and was named Glicko-2. This newer version introduces the *rating volatility* σ [9]. A slightly modified version of the Glicko-2 rating system is used by the Australian Chess Federation.

In the Glicko rating systems, unrated players start with rating set to 1500 and RD set to 350. A player's most recent rating is used to calculate the new RD from the previous with a specific set of formulas provided by the Glicko rating systems.

3 The Game-Based Multi-agent System

Our workbench, *RLGame* [12], was initially presented as a purely competitive test environment. It is a tool for studying multi-agent systems via its tournament version, *RLGTournament* [4, 15] that implements a *round-robin tournament* scheme (combinations, repetitions not allowed) to pair participants against each other. RLGTournament fits the description both of an autonomous organization [3] and of a social environment [2, 3].

The RLGame board game is played on an $n \times n$ square board by two players and their pawns. Two $a \times a$ square bases on opposite board corners are initially populated by β pawns for each player, with the white player starting from the lower left base and the black player starting from the upper right. The goal for each player is to move a pawn into the opponent's base or to force all opponent pawns out of the board (it is the player and not the pawn that acts as an agent in this scenario). The base is considered to be a single square, therefore a pawn can move out of the base to any adjacent free square. Players take turns and pawns move one at a time, with the white player moving first. A pawn can move vertically or horizontally to an adjacent free square, provided that the maximum distance from its base is not decreased (so, backward moves are not allowed).

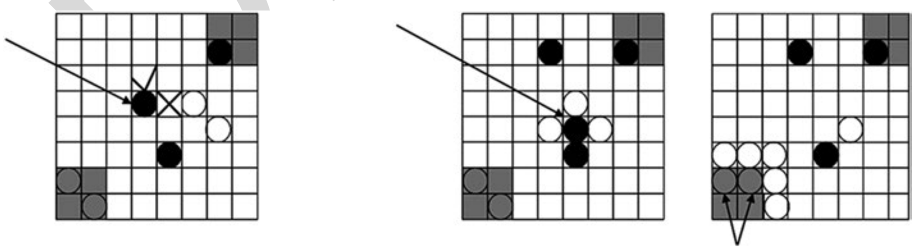

Fig. 1. Examples of game rules application.

The rightmost boards demonstrate the loss of pawns, with arrows showing pawn casualties. A "trapped" pawn automatically draws away from the game; thus, when there is no free square next to the base, the rest of the pawns of the base are lost. The leftmost board in Fig. 1 demonstrates a legal ("tick") and an illegal ("cross") move for the pawn pointed to by the arrow, the illegal move being due to the rule that does not allow decreasing the distance from the home (black) base.

Each agent is an autonomous system, which acts according to its characteristics and knowledge. The learning mechanism of each agent is based on approximating its (reinforcement-learning-inspired) value function with a neural network [2, 3], with similar techniques already documented in the field [13]. Each autonomous (back propagation) [14] neural network is trained depending on its customization and the next possible moves. The board positions for the next possible move along with some flags on overall board coverage are used as input-layer nodes. The hidden layer consists of half as many hidden nodes. A single node in the output layer denotes the extent of the expectation to win when one starts from a specific game-board configuration and then makes a specific move.

RLGame was transformed into a tool for studying multi-agent systems via its tournament version, RLGTournament. *RLGTournament fits the description both of an autonomous organization and of a social environment* [3]. Depending on the number of the agents, social categories can be split into sub-categories of micro-social environment, environment composed of agent groups and global societies, which are the next level of the cooperation and competition extremes of the social organizations [2, 3].

4 Experimentations and Results

In order to study the human rating systems Elo and Glicko applied onto MAS and analyze the performance and learning rate of the agents using as many reliable data as possible, a large scale tournament was configured as follows: 126 agents, all with different characteristics, were used in a round-robin tournament with 100 games per match (each match was repeated 100 times). Each agent played 125 matches against different agents, resulting in a total number of

$$\binom{126}{2} \times 100 = \frac{126!}{124! \times 2!} \times 100 = 787.500$$

experiments, which have been repeated twice. Both experiments are identical in terms of agent configurations and flow of execution.

A first comparison between the Elo and Glicko ratings obtained by the experiments is shown in Fig. 2, which shows the graph of the *Elo-Glicko signed difference*[1] (top) and the corresponding *histogram of signed differences* (bottom). The signed difference is simply

$$d_i = R_i^E - R_i^G \qquad i = 1 \ldots 126$$

where R_i^E is the rank of the *i-th* agent according to the Elo rating and R_i^G is the rank of the same agent according to the Glicko rating. It turned out the in our experiments these

[1] We are not considering here the L_1 distance but rather the simple subtraction of the rankings.

differences fall within [94, 68] and, apparently, the two rating systems “disagree” in how they rank the agents in most cases, with a “strong disagreement” in many cases. The histogram shows that the distribution of this “disagreement” is similar to a normal distribution with zero mean (denoted ‘m’ in the graphs) and large variance. The standard deviation (denoted in the graphs as ‘s’) is about 33 ($\sigma = 32.716$), which means that most of the Elo-Glicko rank differences can be expected to fall within a region that spans a range of about 66 rank positions. This spanning range is quite high (more than half of the total range) if one considers that there are 126 total rank positions, which is a strong indication that the two ranking systems treat the experiments in a quire different way and they are not expected to produce consistent rankings.

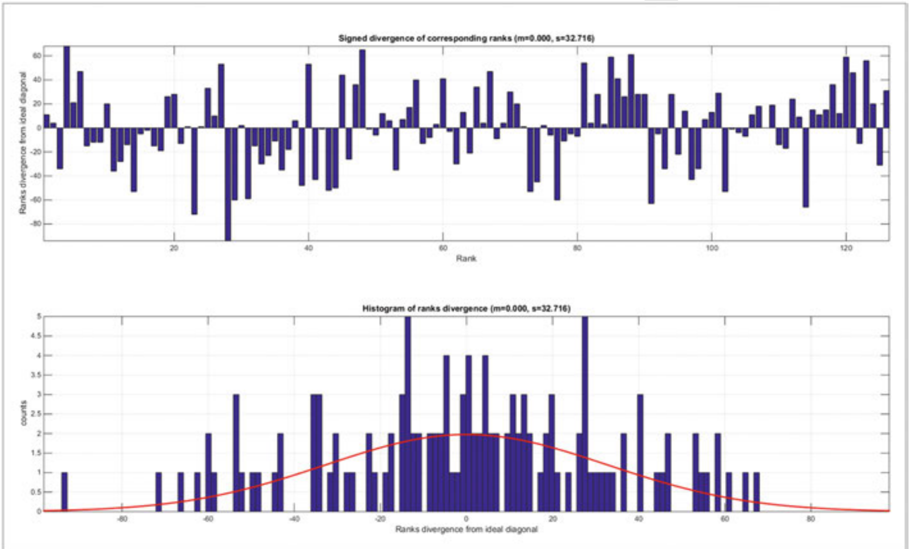

Fig. 2. Divergence between Elo and Glick ratings and the corresponding histogram

Further study of the results included the usage of the *Spearman's rank correlation coefficient* (ρ) [16], which measures the statistical dependence between two variables, and is specifically efficient at capturing the monotonic (non-linear, in general) correlation on ranks. As known, the range of the coefficient falls within $[-1, 1]$, with high negative values representing strong negative correlation, low absolute values representing small or no correlation and high positive values representing strong positive correlation. In our experiments it was estimated that

$$\rho = 0.5987$$

which indicates a typical positive correlation, which is not strong enough to support a consistent behavior of the two ranking methods.

Figure 3 presents a graph of the Elo rankings vs. the Glicko rankings. Ideally, if the two ranking systems would agree, we would expect to find all point on the diagonal. In our case we see that there are many agents off-diagonal. If we adopt a scenario that we are error-tolerant (i.e. we accept rank differences with specific limits) we may define various *Special Zones* centered around the diagonal that would represent various zones of ranking “agreement”. These Zones are presented in various shades of gray in Fig. 3. If for example agent X was ranked in the 26th position by the Glicko system and in the 36th position by the Elo system, then the error (disagreement) is ten positions, which corresponds to about 8 % error relative to the total range of 126. In addition, the green and red lines make a heuristic distinction of “good”, “moderate” and “bad” playing, by simply dividing the total ranking scale to three regions of equal lengths. Table 1 reports the number of agents (and corresponding percentage) in various Special Zones that accept absolute rank differences in the set {2, 5, 10, 20, 40}.

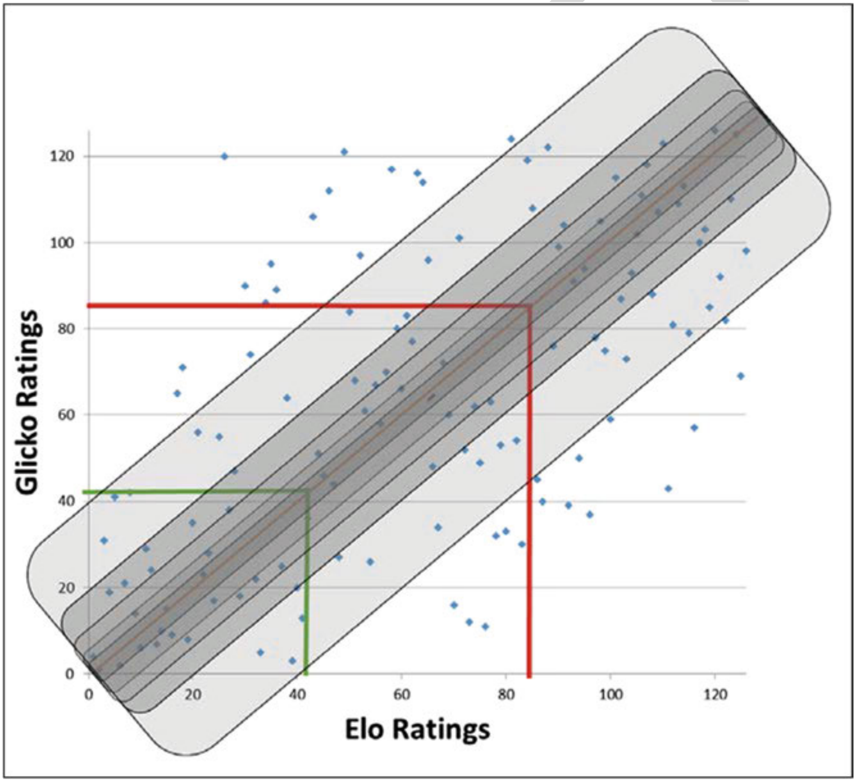

Fig. 3. Elo vs. Glicko ranking and the special zones of tolerance

It is evident that even one adopts a fault-tolerant approach the two ranking methods produce consistent rankings only for a very small number of agents.

Table 1. Various special zones of Elo Glicko absolute rank difference tolerance

Rank difference	#agents	% agents
2 (1.67%)	10	7.94
5 (4.03%)	21	16.67
10 (8.06%)	33	26.19
20 (16.13%)	64	50.79
40 (32.26%)	95	75.39

Table 2 presents a more detailed view of the difference in the ranks obtained by the two rating methods for the identified Special Zones and playing performance. Specifically, it shows how many of the agents fall within a Special Zone either for both methods (rows "Agree for") or for just one of the methods (rows "Disagree for"); it also presents those results using the heuristic classification as "Good", "Moderate" or "Bad" produced by uniformly dividing the total rank scale in three equal parts.

Table 2. Agreement and disagreement of the two methods within the special zones

Playing performance	Agreement / disagreement of methods	2 ranks difference #agents	5 ranks difference #agents	10 ranks difference #agents	20 ranks difference #agents	40 ranks difference #agents
Good	Agree for	4	7	9	21	31
	Disagree for	38	35	33	21	11
Moderate	Agree for	4	10	18	23	33
	Disagree for	38	32	24	21	9
Bad	Agree for	2	4	7	20	31
	Disagree for	40	38	35	20	11
SUM	Agree for	10 (7.94%)	21 (16.67%)	34 (26.98%)	64 (50.79%)	95 (75.39%)
	Disagree for	116 (92.06%)	105 (83.33%)	92 (73.02%)	62 (49.21%)	31 (24.61%)

5 Conclusions

As more and more multi-agent systems with social organization and advanced learning mechanisms are being studied, synthetic agent-rating mechanisms are starting to be applied and tested. Among the choices for agent rating there are the rating methods for human performance, such as Elo and Glicko, which have initially been developed to rank human players performance on games like chess and go. Extensive experiments have shown that the rankings produced be the two methods show excessive ranking inconsistencies and rise doubts on the applications of both methods in synthetic worlds.

By developing a simpler method as a substitute to existing rating methods, we also hope to use it as a benchmark to calculate the extent to which these two methods differ as regards score calculation in a series of multi-agent competitions. Such simpler methods may rely on just adding up the number of wins, maybe discounted over time, and still provide adequate information as to the quality of individual agents. In our future work we are planning to develop and compare more suitable rating mechanisms that would be efficient in assessing the performance of synthetic agents in multi-agent systems with social organization.

References

1. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist, 2nd edn. Open University Press, Maidenhead (2005)
2. Shoham, Y., Leyton, K.B.: Multiagent Systems: Algorithmic, Game Theoretic and Logical Foundations. Cambridge University Press, New York (2009)
3. Ferber, J.: Multi agent Systems: An Introduction to Distributed Artificial Intelligence. Addison Wesley, Boston (1999)
4. Kiourt, C., Kalles, D.: Social reinforcement learning in game playing. In: IEEE International Conference on Tools with Artificial Intelligence, Athens, Greece, pp. 322 326 (2012)
5. Marivate, V.N.: Social learning methods in board game agents. In: IEEE Symposium Computational Intelligence and Games, Perth, Australia, pp. 323 328 (2008)
6. Caballero, A., Botia, J., Gomez Skarmeta, A.: Using cognitive agents in social simulations. Eng. Appl. Artif. Intell. **24**(7), 1098 1109 (2011)
7. Logan, Y., Kagan, T.: Elo ratings for structural credit assignment in multiagent systems. In: Twenty Seventh AAAI Conference on Artificial Intelligence (Late Breaking Developments) (2013)
8. Elo, A.E.: The Rating of Chess Players, Past and Present. Arco Publishing, New York (1978)
9. Glickman, E., Albyn, J.C.: Rating the chess rating system. Chance **12**(2), 21 28 (1999)
10. Glickman, M.E.: A comprehensive guide to chess ratings. Am. Chess J. **3**, 59 102 (1995)
11. Harkness, K.: Official Chess Handbook. McKay, New York (1967)
12. Kalles, D., Kanellopoulos, P.: On verifying game design and playing strategies using reinforcement learning. In: Proceedings of ACM Symposium on Applied Computing, Special Track on Artificial Intelligence and Computation Logic, Las Vegas (2001)
13. Tesauro, G.: Practical issues in temporal difference learning. Mach. Learn. **4**, 257 277 (1992)
14. Sutton, R., Barto, A.: Reinforcement Learning An Introduction. MIT Press, Cambridge (1998)
15. Kiourt, C., Kalles, D.: Building a social multi agent system simulation management toolbox. In: 6th Balkan Conference in Informatics, Thessaloniki, Greece, pp. 66 70 (2003)
16. Spearman, C.: The proof and measurement of association between two things. Am. J. Psychol. **15**(1), 72 101 (1904)