

Elo Ratings for Structural Credit Assignment in Multiagent Systems

Logan Yliniemi and Kagan Tumer

Oregon State University
Corvallis, OR, USA

logan.yliniemi@engr.orst.edu, kagan.tumer@oregonstate.edu

Abstract

In this paper we investigate the applications of Elo ratings (originally designed for 2-player chess) to a heterogeneous nonlinear multiagent system to determine an agent’s overall impact on its team’s performance. Measuring this impact has been attempted in many different ways, including reward shaping; the generation of hierarchies, holarchies, and teams; mechanism design; and the creation of subgoals. We show that in a multiagent system, an Elo rating will accurately reflect an agent’s ability to contribute positively to a team’s success with no need for any other feedback than a repeated binary win/loss signal. The Elo rating not only measures “personal” success, but simultaneously success in assisting other agents to perform favorably.

Introduction

A key problem in the field of multiagent systems is that of credit assignment; that is, given a team of agents and their performance, which of the agents were responsible for the team’s success (Agogino and Tumer 2004)? Perhaps each agent contributed equally, but it is possible that one agent contributed significantly more than another. Calculating this contribution and rating agent successes is pursued in many works, and necessary for many techniques (Panait, Sullivan, and Luke 2006; Barto and Mahadevan 2003; Devlin and Kuznetsov 2011; Parkes and Duong 2007).

In professional chess, a simple calculation introduced by Arpad Elo was adopted in 1960 that calculated the skill of a player based on their wins and losses, and the calculated skill of their opponents (Elo 1978). Chess, however, is a competitive two-agent system where each agent’s performance is based solely on its skill. In a multiagent setting with static teams, it is a simple extension of the same math: the static team can be treated as a single unit, and the team as a whole deserves all credit for the win or loss.

In a situation where teams are no longer static, but must be reorganized, created, or adjusted in an ad-hoc manner to achieve a preset task, the credit assignment problem becomes key. In such a case, determining how an agent’s rating is based on its own skill and how it is based on the team’s skill is the critical credit assignment problem in multiagent systems. In a team with 5 agents, it might seem as though

the teammates would hold $\approx 80\%$ control over an agent’s Elo rating, invalidating the use of an Elo rating system.

In this work we show that even when participating with constantly-changing teammates, a pure Elo calculation works surprisingly well in correlating actual agent skill (known, in this case) to Elo rating, and is not overly sensitive to any tuning parameters explored. This sets the stage for future work which could use an Elo rating as a fitness calculation for many different learning or search algorithms.

Background & Experimental Setup

This section includes background on Elo ratings, and our Gaussian Draw Domain (GDD), in which two teams of five randomly-selected agents compete in a win-loss match.

This domain was selected for initial exploration because it is extremely quick to execute, but still offers some interesting non-linear relationships between agents.

Elo Calculation

The Elo system assumes each player has a skill that is drawn from a random distribution (sometimes a player will have a “good” game, sometimes a player will have a “bad” game); the Elo Rating attempts to find the center of that distribution and converge to that value.

We use a simple Elo Rating calculation after each match. In a game between two players A and B with current ratings of R_A and R_B respectively, the probability that player A will win is (Glickman and Jones 1999):

$$E_A = \frac{1}{1 + \alpha^{-(R_A - R_B)/\beta}} \quad (1)$$

Where $\alpha = 10$ and $\beta = 400$ are standard values. After receiving a win-loss result, the rating for each player is changed by:

$$R_{player} \leftarrow R_{player} + K(S - E_{player}) \quad (2)$$

where K is a pre-set factor with a given reduction schedule, S is the outcome (1 for victory, 0 for defeat), and E_{player} is the expected probability of victory for that player.

We chose the arbitrary initial K value of 55, and reduce it by 1 for every match an agent participates in, until it reaches a value of 25. Extreme K values do affect the speed of convergence, but as long as K is reduced, the Elo rating will converge in the single agent case. We found that in a multi-agent implementation, this convergence still occurs.

Agent Parameters and Team Organization

Once created, all 13 parameters used to describe an agent are held constant through all simulations. These parameters are: 5 skill mean and 5 skill variation values (one of each for each heterogeneous role in the GDD, used in Equations 3 and 5); a “persuasiveness” and a “stubbornness” rating, which affect the team’s self-organization; and a vector that describes which subteam the agent contributes to when holding role 5 (see Match Dynamics). Each agent has full knowledge of their mean skill levels, and can state their preference for each role, but has no knowledge of other agents’ skill levels.

Each match is formed by randomly selecting 10 agents within a specified Elo rating of each other for competition. These 10 agents are randomly split into two teams of 5, which we term “Red” and “Blue” teams. Each team then self-organizes such that each agent fills one of 5 heterogeneous roles. An agent always attempts to acquire their best-performing role, but in the case of multiple agents desiring the same role, their persuasiveness and stubbornness ratings are used to attempt to assign the role. In cases where the agents cannot agree, the role is randomly assigned.

Match Dynamics

We use a 2-phase simulation which models a system in which a team of robots search for an injured hiker. In the first phase, the team must work in sub teams to find the hiker:

$$T_{subteam} = \sum_{role \in subteam} Nor(\mu_{role}, \sigma_{role}) \quad (3)$$

In our 5-role GDD, three subteams are created: (roles 1 and 2) (role 3) and (role 4). Role 5 may contribute in full to one subteam, or may spread its contribution across multiple subteams (this choice is held constant on a per-agent basis). An “advantage” is carried forward to the second phase for the team that found the hiker more quickly (as this is a desirable behavior):

$$\delta_{subteam} = \phi \cdot (T_{subteam, Blue} - T_{subteam, Red})^2 \quad (4)$$

where ϕ is a parameter held constant ($\phi = 0.2$) through all trials. In phase 2, the entire team works as a whole to evacuate the injured hiker:

$$T_{team} = \sum_{subteams} I \cdot \delta_{subteam} + \sum_{roles} Nor(\mu_{role}, \sigma_{role}) \quad (5)$$

Where I is an indicator function which takes a value of 1 if the subteam won in phase 1, and 0 if the subteam lost. The team with the higher T value wins.

Results

Using this GDD, we carried out a series of large-scale experiments. We used 10,000 agents over 1 million matches: 9900 random agents as well as 100 specially designed agents to test hypotheses. To apply the Elo system, in a multiagent team, we aggregate the individual agent Elo ratings on each team by taking the mean Elo rating of all members. From the ending Elo distribution, we performed a Principle Component Analysis (PCA) on the agent parameters and attained Elo ratings. We discovered many interesting phenomena. Here, we discuss three:

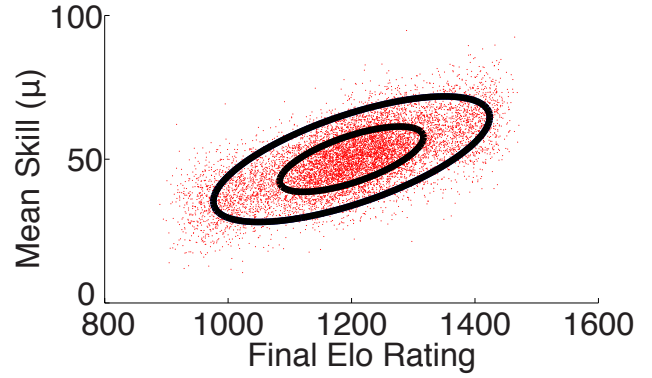


Figure 1: Elo rating vs. mean skill across all roles. The ellipsis presented are the equiprobable curves within which 68% of the population (inner) and 95% of the population (outer) reside. The positive slope of their major axis shows a strong positive correlation between Elo rating and mean skill.

1. High Elo ratings identify agents with strong mean skills across all roles.
2. Agents which are highly skilled in one role and unwilling to cede that role gain high Elo ratings. Agents which are evenly skilled across many roles and willing to take on any role gain high Elo ratings.
3. An agent whose Elo is artificially perturbed (lowered or raised) will re-converge to the same Elo.

Item (1) demonstrates that because Elo correlates with mean skill (shown in Figure 1): better performing agents on an individual level tend to earn a higher Elo rating. Item (2) demonstrates that it is not only the individual performance of the agent that affects their Elo rating, but also their cooperation with teammates: the credit or blame for the wins and losses that the agent participated in are embodied within their Elo rating. Item (3) demonstrates that the Elo ratings are not fragile to perturbances.

Conclusion and Future Work

In this work we show that an Elo calculation initially intended for 2-player chess shows a number of properties that are desirable in multiagent systems to address the structural credit assignment problem: Elo positively correlates with agent skill, and positively correlates with performing actions that are beneficial to the the team’s performance as a whole, while remaining robust to disturbances.

Selecting the the five agents that attained the highest Elo does not form the optimal hiker rescue team out of the 10,000 agents, but it does select a high-performing team. In domains that the optimal is not known or attainable, Elo could be a good way to select desirable agents. This allows us to use a search technique, like an evolutionary algorithm, with Elo as a fitness calculation. Additionally, tracking separate Elo ratings for performance on each role could lead to interesting comparisons with a single aggregate rating.

References

- Agogino, A. K., and Tumer, K. 2004. Unifying temporal and structural credit assignment problems. *AAMAS*.
- Barto, A. G., and Mahadevan, S. 2003. Recent advances in heirarchical reinforcement learning. *Discrete-Event Dynamic Systems: Theory and Applications* 13:41–77.
- Devlin, S., and Kudenko, D. 2011. Theoretical considerations of potential-based reward shaping for multi-agent systems. *AAMAS*.
- Elo, A. E. 1978. *The Rating of Chess Players, Past and Present*. New York: Arco Publishing.
- Glickman, M. E., and Jones, A. C. 1999. Rating the chess rating system. *Chance* 12(2):21–28.
- Panait, L.; Sullivan, K.; and Luke, S. 2006. Lenient learners in cooperative multiagent systems. *AAMAS*.
- Parkes, D. C., and Duong, Q. 2007. An ironing-based approach to adaptive online mechanism design in single-valued domains. *AAAI*.