Enabling International Rating of Go Players

Chris Kirschner
American Go Association President
president@usgo.org

Paul G. Matthews, Ph.D.
American Go Association Ratings Coordinator
ratings@usgo.org

1. Overview

Many International Go Federation (IGF) members, including the American Go Association (AGA), maintain their own rating system. While most amateur players beyond the novice stage seem to know their strength to within a rank or two, the objective of the AGA rating system is to achieve greater accuracy, currently on the order of one tenth of a rank, based on data accumulated over the lifetime of each player; also, tournament directors commonly depend on accurate ratings to seed and pair players appropriately. Players from other countries who participate in AGA tournaments are rated, but currently these data are not shared in any regular way with other organizations, and the same is true for United States residents playing abroad. As a result, a player may have ratings from several systems, each based on different data, and no easy way for either the player or a tournament director to reconcile them. AGA members want to know how they compare with the rest of the world, but today much of the available evidence are anecdotes and not hard data.

The IGF could help in several ways, and play an even more vital role in the international go community. Because data are not widely shared, relationships between ratings from different systems are poorly understood. The IGF could take a leadership role in facilitating the publication of data needed for a proper analysis and mapping between rating scales. Also, the IGF could establish an international player register, assigning an international ID number that could be used to correlate a player's data across systems. These steps would enable a federated system of international ratings, where a player's strength could be known unambiguously based on data from all the participating rating systems. The AGA would strongly support an IGF initiative in this direction.

The following sections explain our approach, and propose actions that the IGF could undertake in the near term to get rolling.

2. Approach

The AGA has been involved in several past efforts to unify ratings in North America and Europe and on Internet go servers, at various times providing software, detailed algorithms, and rating services. These efforts have not succeeded as well as we had hoped for several reasons.

- Each rating authority has strong feelings of autonomy and a need for control to attend to the needs of their community. However, the quality and credibility of ratings depend on the extent of the player population networked together by playing rated games, which is an incentive toward rating data on a larger scale from more sources.

- Each system uses a rating scale and method of calculation that are at least slightly different, and each implements different policies for managing the deflation/inflation typically caused by improving players. Ratings also differ between systems due simply to different tournament input data. All these factors require detailed analyses to tease apart and resolve. However, the data required to perform such analyses are not commonly made available.

- Each system defines different data formats, which require extra programming effort just to translate between systems. This adds to the burden of sharing data.

No one will accept ratings from another authority, or understand what they mean, until they see how their own system would rate the same data, and can directly compare individual ratings. The same was true within the AGA in 1989 when the current AGA rating system was instituted: the new system was accepted only after actual rating lists were compared and judged to be reasonable.

Sharing of rating algorithms and theoretical ideas has been relatively easy to do, but has had little impact in achieving unified international ratings. Existing rating systems share many common principles and mathematical methods that make them more similar than different; for example, every rating system increases a winner's rating and decreases the loser's to achieve more plausible relationships between players, or else the system would be quickly abandoned. Technical specialists will continue to refine the theory and programming of rating systems, but such technical concerns are not pertinent to the main problems of international ratings. Rather than debate the merits of this or that rating system, or creating yet another rating system, we recommend focusing on how ratings from existing systems can be related.
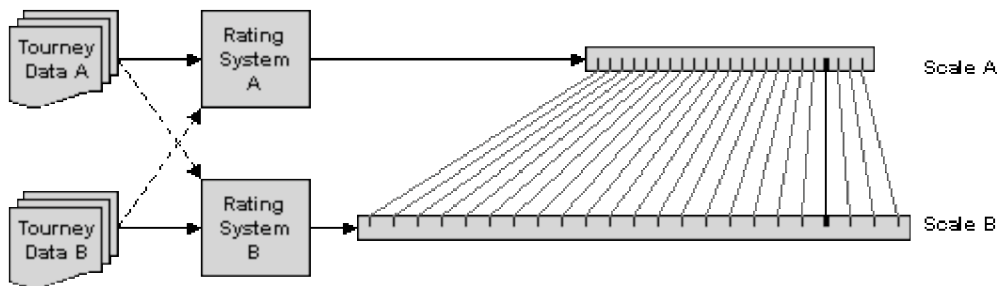
Figure 1 illustrates the core of our recommended approach.

Figure 1 – Mapping Between Rating Scales

To simplify the discussion, consider just two rating systems based on different continents. Today, each system rates a different set of tournament data, and only a small percentage of players are present in both systems. Analysis proceeds in two steps.

- First, all the data normally input to each system are also rated by the other system, to calculate for every player their system A and B ratings based on exactly the same data. The quantitative relationship between the two sets of ratings can be analyzed in several ways, but a simple scatter plot of players with the two rating scales as the axes would be a start, and there would be a large number of data points. We expect that a relationship between ratings in different systems can be described fairly well by a set of equations or tables to convert between them.

- Second, because rating scales generally describe relative differences between players but not absolute levels, the two scales still need to be aligned. For example, a player with the same relative position on each rating scale might be labeled "1 kyu" on one and "6 dan" on the other. For alignment purposes, we need actual ratings for at least a few players who are in both systems. Concentrating on just the stronger players may be sufficient, since if we can align the scales at the top, then the rest will follow mathematically given the first step of the analysis described above.

If accurate mappings between rating scales can be determined, then any player rated by any system would have, in effect, have an internationally recognized playing strength, without necessarily having to create yet another rating system. In addition, if every rating authority were to publish the tournament data it receives within its domain, and if players who compete across domains were assigned a commonly

recognized international player ID number, then these data could be gathered together as input for worldwide ratings, which could be provided by one or more rating systems to be determined later.

To make this vision a reality, the following technical enablers are practically necessary.

- An international standard for publishing tournament data sufficient for ratings is needed so that rating authorities can publish their data in a way that can be read by others and fed into automated ratings analysis programs. The Internet provides convenient platforms for publication, including XML which is widely used in international e-business to exchange structured data using a variety of character sets and languages.

- An international player register is needed that would assign an international player ID number to any player who wants to compete internationally, and would include sufficient data to enable each rating authority to identify that player in their own records. Otherwise, we cannot reliably correlate player records in different systems.

If these enablers are achieved, then international ratings will follow quickly.

3. Actions

The IGF has a golden opportunity to help meld ratings from around the world into a coherent whole. Following the approach outlined above, we recommend the following specific actions.

- Establish a working group to define an IGF sponsored standard for publication of tournament and other rating related data. This group would consider the current practices and needs of existing rating systems, and propose a set of common data formats that could be ratified by IGF members.

- Create an IGF player register that assigns a lifetime ID number to any player who competes internationally and agrees to be included. IGF members would match these ID numbers with the players in their respective rating systems. Using this register, all a player's games, from anywhere in the world, could be brought together for a complete analysis.

- Publish data from IGF sponsored tournaments, such as the WAGC and IWABC, on a web site using the player ID numbers and data formats proposed above. IGF member organizations should do likewise for the tournaments within their respective domains.

- When sufficient data are available, establish a working group to analyze the data with the objective of defining standard conversions between rating scales. Analyses may be performed independently by IGF members and contributed to the working group.

Eventually, the IGF may decide to calculate its own ratings, but this may not be necessary. A federated approach that respects the integrity of each rating authority, and enables them to analyze each other's data, will avoid many of the problems of past unification efforts.