

## The Go Ranking System of Arpad Elo

Franco Pratesi

In the last issue we looked at the ranking system introduced by Bruno R uger. In this and subsequent articles some developments of Go ranking systems – up to the present day – will be outlined. This time we look at the system devised by Arpad Elo and originally applied to Chess.

The traditional ranking system of Go players (of Japanese provenance, from about 9 dan pro down to 1 dan amateur and then from it to about 35 kyu) and the European ranking system (from 1, or 0, continuously down to 60, or a higher limit set for beginners) are both based on the number of handicap stones, be they integers or fractions. As such, they are typical of Go and cannot be transferred to other games.

There is however another ranking system that is in principle suitable for any strategic game and sport. It is based on the rating system implemented by Arpad Elo and fully described in his book *The Rating of Chess Players*, published in London by B. T. Batsford in 1978. In particular, what we get thereby is essentially a rating number, associated with every player and changed either after each tournament, or at fixed times, by the various federations.

Together with the rating numbers we obtain a direct way to rank the players, by selecting a suitable interval of rating values for distinguishing subsequent ranks. This interval of ratings is commonly selected so that a player has 76 percent probability to win a game against a player one rank below. The specific 0.76 value derives from the assumption that game results follow a normal distribution, but using different statistical laws would lead to similar values.

The fundamental concept is that any Elo scale is an ‘open-ended floating’ one. This implies that there are no fixed reference points; it is the differences between rating values that are important and not their absolute values, which are set by one or another arbitrary selection.

In applying the Elo system to Chess, the rating interval between ranks was chosen to correspond to 200 points, for historical reasons, in order to fit an existing system. The ratings traditionally employed for Chess increase as known up to about 2000 for the strongest club players,

2400 for international masters and 2700 for candidates to the world championship – only ten Chess players now pass this value. The lower limit is rather undefined, with beginners at or somewhat below 1000. Thus, the total number of ranks required, at 200 point steps in rating until 2800, corresponding to the world champion, is about ten.

Apart from the rating values chosen, which may variously agreed upon according to requirements of the specific application, the whole concept of Elo rating (and ranking!) is suitable for most games and sports and has been officially adopted, in particular, by federations of Chess, Draughts, Othello-Reversi, and other board games.

Unfortunately, the detail of the system and its parameters as adopted for the various games usually are somewhat different, so that comparing player ratings for different games does not ensure that similar rating values correspond to similarly strong players. Moreover, there are intrinsic reasons why rating numbers cannot be exactly comparable for different games: each game actually has its own complexity and so the number of ranks required for ranging the various strengths of the players must be smaller for easy games and larger for complex ones.

For instance, one can check at the World Football Elo Ratings web site: [www.elaboratings.net](http://www.elaboratings.net), the application to Football, where the following ratings may be found for October 2000: 1 France 2026, 2 Brazil 1943, 12 England 1795, 38 Scotland 1672, 100 Northern Ireland 1440, 101 Wales 1438, 214 and last, Anguilla 907.

Thus, for Football, the number of ranks required for ranging all the national teams of the world turns out to be about five. Compared to Chess, this indicates a much lower complexity of the game – if one is allowed to compare on these terms a team game mainly played by a score of feet with one played between individuals, mainly using their minds.

What about Go? Several studies have indicated that Go has a remarkably greater overall complexity compared with the other traditional board games, also with Chess, not to mention Draughts, or Othello-Reversi.

A confirmation may derive from the much lower progress in playing skill reached by the corresponding computer programs.

As noted at the beginning, ranking of Go players traditionally occurs in terms of stone handicaps. A specific Grading System, based on promotion points, has been used for many years by the BGA for ranking

dan players and will soon be introduced for kyu players (see *BGJ 115*, 1999, pp. 12-15).

Other countries have adopted Elo-type ratings, but – as far as I know – always modified in order to obtain some fitting with the established stone-handicap ranks. For instance, the Kommission Go of the German Democratic Republic (apparently, not the best candidate coming to our minds for accepting a suggestion from the USA) officially adopted the Elo system in 1989. The same normal distribution curve in use for Chess was applied and the 200-point interval still kept the meaning of 0.76 winning probability. However, a rating interval of 100 was empirically found to be better suitable for fitting the already existing stone-handicap based ranks. Therefore, the system worked by setting its rating values between 0 at 20k and 2300 at 4d, regularly stepped so that 1000 corresponded to 10k, 2000 to 1d, and so on.

At present, the involvement of European Go players with Elo systems is increasingly gaining a widespread acceptance. In particular, since 1998, the European Go Federation itself has officially adopted a Go rating system of the Elo kind, introduced some months earlier by Ales Cieply for the Czech federation. The values of the European Official Ratings increase from 20k (here with the value of 100 assigned) to 7d at 2700 – again regularly stepped at 100 rating points intervals.

By applying this system to the ranks of practical occurrence for Go players, we eventually obtain Elo rating values reasonably similar to those commonly encountered in Chess. However, differently from Chess, here full use is made of lower ratings, down to zero. It would even be possible to shift further the zero setting in order to take weaker players into account, down to complete beginners. Compared with the series of about ten ranks for Chess, many more might be required for Go, an independent way to confirm its greater complexity.

Even if the general framework of these applications may be considered clear enough, I believe that the correspondence between handicap-stone and Elo ranks for Go players is worth of further study and hope soon to be able again to comment on it.