# The Go Ranking System of Walther Schmidt

Franco Pratesi

We have examined several systems and another couple will soon be described, selected because they suggest original methods for ranking Go players. All of them find a common obstacle in... the ancient tradition of the game, which often is strong enough to reject any scientific approach of nowadays. To begin with the top, it is hardly useful to find a scientific method for correctly ranking all the 9p masters into several sub ranks, before they are prepared to accept any such method!

We have limited our attention to ranking systems, but the conclusion would not be very different if we had examined other aspects and even the fundamental rules of the game. In the last years, several experts have brought useful contributions to a scientific approach to Go, which often requires a more rigorous definition of the game rules – just because any scientific treatment needs a solid basis. It is hard to use our scientific methods, for instance, when tradition states that, in cases of dubious issues, either the older player is right by definition or a player older than both contestants is searched out and charged with the decision. A general discussion of the scientific approach to Go and its basic rules would lead too far – let us better continue within the limits of the seemingly simple goal of ranking the players.

In our historical review of particular Go ranking systems, we have now to come back to 1968. In this memorable year – among many certainly more central and renowned events! – an article of 10 pages was published by Walther Schmidt (*Deutsche Go Zeitung*, Section 8.1). The author, a Doctor in chemistry, lived then in Timiçoara, Romania. His contribution to Go was not limited to theoretical study: in the 1960s, nothing less than a chapter of the Nihon Ki-in was present in Timiçoara, formed by a small group of Go players headed by Walther Schmidt and his brother. (I am not aware of any connection with the remarkable success obtained by Romanian players of today.)

In the article mentioned, Schmidt performed a statistical analysis on two representative sets: 45 games of Japanese pros, with strength 4-9p; 40 games from the 1962 and 1963 European championships, most by

1-2d players. Attention was given to evaluating the correct komi, found to be 4.5, a value that would appear too low by the standards of today. However, we are interested in another suggestion, a kind of side-product of the study, indicated there in a concise way.

Schmidt realised that the dispersion of game scores decreases with increasing player strength. (It is natural to suppose that others had remarked this earlier, but I did not find indication of this.) From his study, he could derive two values for the mean square deviation, 4.8 for the Japanese masters and 14.7 for the European top players. In addition to these experimental values, he proposed a third value, 0.5, as the estimated deviation in games among perfect players, considering the lowest possible score of one point. The most stimulating part of Schmidt's suggestion is that an absolute ranking system can be built on the basis of these mean deviations of game scores. The ranks thus obtained would allow correct comparison of strength among groups of players, even if belonging to different places and times.

The system was proposed by Schmidt for groups of Go players; nevertheless, it should not be too difficult to extend it to single players. If you meet a perfect Go player, then the dispersion of a few game scores against him automatically provides your correct rank. As often occurs in passing from theory to practice, it may be less complex to 42 search for a more complex situation: for instance, studying the dispersion behaviour among selected pairs of players, so that their strength is approximately the same within the pair but decreases stepwise among pairs.

By selecting a suitable unit, a new scale for playing strength can be defined, apparently different from both the Elo and the handicap stone systems. The new scale is no longer an interval scale, but an absolute ratio scale, the best that one can use for measuring any physical property, as we find for instance in the case of temperature when passing from Celsius or Fahrenheit degrees to Kelvin.

I do not know why Schmidt only suggested the new absolute scale without completing his proposal with the unit of the scale and its upper limit, the mean square deviation for complete beginners. I can imagine that he found difficulties; to begin with a rigorous definition of the game rules – about them he wrote a couple of unpublished papers. Left as such, the suggestion by Schmidt is only an indication for a new rating system, without yet any indication on ranks, which can only be defined after a suitable unit is applied to the scale.

Nevertheless, the idea of an absolute scale for measuring player strength looks very promising. Elo himself has discussed this basic question in his 1978 book, particularly on pages 138-143 and 148-149; more difficult is to find (my thanks are due to Theo van Ees) his fundamental article on the same topic, privately published in 1966, thus before the contribution by Schmidt described here.

Elo concludes that a ratio absolute scale can currently be used instead of his 'open ended floating' interval scale. For chess and similar games, using ratios would lead to so great numbers that one should better introduce a logarithmic scale. In particular, Elo suggests using the square root of 10 as the basis of the ratio scale, and substituting the normal distribution with the logistic one. Thus, the correspondence between Elo's interval and absolute scale – using decimal logarithms for the latter – should be as follows: 0=0, 400=1, 800=2, 1200=3, 1600=4, 2000 = 5, 2400 = 6, 2800 = 7, 3200 = 8, 3600=9, 4000=10.

However, unlike the 'open ended floating' interval scale, the absolute scale must be anchored. In Elo systems, it is natural to increase its values from zero at the complete beginner up to the strongest players. Actually, it would be hard to use this absolute scale beyond the strongest players in existence, as we do not have information on the ranks separating them from the theoretical perfect player. On the contrary, the scale section for weak players would be rather simple, owing to the logarithmic relation with handicap stone rankings; with only two or three Elo ratio ranks one can already reach an average kyu strength, thus reducing the uncertainties found with stone handicaps for consistently distinguishing players ranked from about 20k up to 35k and higher, or at 50 or 60 or 80 of the European scale.

Unfortunately, I do not know any game or sport in which absolute Elo scales have been adopted and in Go the tradition established is stronger than in other games. Thus, we shall soon come back to the framework of the ranking systems typical of Go, with its handicaps. Let us however note an essential difference in the two kinds of absolute scales suggested: if we use dispersion of game scores, or similar properties, it becomes obvious to apply the scale upside-down, fixing namely its zero at the perfect player and increasing its values with the dispersion of game scores, that is, with player weakness.

An approach that eventually completes the proposal by Schmidt – and may solve the problem of providing the scale with a suitable unit

of measurement – was suggested a few years later by Prof. Klaus Heine and will be described, if possible, in the next issue.