Proposals of the USCF Ratings Committee

August 1993

Mark E. Glickman, Ph.D.

Chairman, USCF Ratings Committee

Contents

Preface

The Ratings Committee was asked this year to propose a solution to the current implementation problems of the provisional rating system. Not only has the Ratings Committee worked out a solution, but we are proposing a general and simplifying *approach* to the rating system. The important features of our approach involve

- simplifying the rating algorithm so that only one formula, rather than two, governs rating update calculations,
- organizing tournament sections so that eligibility into a section is independent of minor fluctuations in rating,
- removing incentives for players to artificially manipulate their ratings (e.g., by sandbagging) so that ratings can become better forecasters of performance, and
- allowing relevant external information, such as age or tournament experience in non-USCF sanctioned tournaments, have an impact on an initial assignment of a player's USCF rating.

These general considerations underlie our three separate proposals in this report.

Ratings Committee Members

Harry Cohen Ph.D. in Operations Research, Massachusetts Institute of Technology (1975). Member of the USCF Ratings Committee 1988–present. Currently the principal in a transportation planning and management services consulting firm.

Mark Glickman Ph.D. in Statistics, Harvard University (1993). B.A. in Statistics, *Summa Cum Laude*, Princeton University (1986). Member of the USCF Ratings Committee 1985–present; chairman 1992–present. Currently a post-doctoral fellow at the Department of Health Care Policy, Harvard Medical School.

Bill Goichberg USCF Policy Board member, 1975–1978, and 1989–1992. Professional chess organizer and director.

Albyn Jones Ph.D. in Statistics, Yale University (1986). B.A. in Mathematics, UCLA (1978). Currently holds an Associate Professorship in Statistics at Reed College, Portland, Oregon.

Larry Kaufman B.A. in Economics, Massachusetts Institute of Technology (1968). Member of the USCF Ratings Committee 1976–present, chairman 1981–1986 and 1990–1992. Currently writes computer chess software.

Alan Losoff B.S. in Mathematics, Illinois Institute of Technology (1969). Proofreader for "The Ratings of Chessplayers" by Arpad Elo. Member of the USCF Ratings Committee 1991–present. Computer programmer working on mathematical modeling for a financial derivatives firm.

Kenneth Sloan Ph.D. in Computer and Information Science, University of Pennsylvania (1977). Sc.B. in Applied Mathematics, Brown University (1970). Currently holds an Associate Professorship at the University of Alabama in Birmingham.

Policy Board Liaison:

Frank Camaratta M.S. in Applied Mathematics, Applied Mechanics, and Aerospace Engineering, Drexel Institute of Technology (1968). Vice president of U.S. Chess Federation. Chairman of the Computer Rating Agency, 1985–present. Ratings Committee chairman, 1986–1990.

1 A New Provisional Rating System

1.1 The Unified Rating System

The currently implemented provisional rating system suffers from several problems, including the inability to rate tournament sections that involve all or mostly unrated players, and the fact that a positive result could contribute negatively toward one's rating (or positively from a negative result). The Ratings Committee proposes to unify the rating system so that only the established rating formula is used with a K-factor that depends on the number of tournament games in which a player has competed. Committee member Harry Cohen suggested adopting a system in which the established rating formula would be used in every situation. The key element to this system is that K would be large for players who have played only a small number of rated games. This approach provides continuity in the rating calculations, and allows for a smooth transition from provisional ratings to established ratings.

1.2 Outline of the Rating Procedure

Consider rating a tournament section consisting of an arbitrary number of unrated players, provisionally rated players, and established players. The following steps are to be performed sequentially:

1. Assign a pre-tournament rating to each unrated player in the section. We propose that this assignment is based on a player's age, and possibly other demographic information. In certain cases, a tournament director with appropriately high accreditation may be allowed to assign an initial rating.

2. Assign a K for each player, based on pre-tournament rating and the number of games played.

3. Update the ratings of each unrated player using the established rating formula with the values of K from step 2.

4. Update provisionally rated players using the established rating formula with the values of K from step 2.

5. Update the ratings of established players in the usual manner.

1.3 Assignment of Initial Ratings

The unified system requires an initial rating be assigned to each unrated player in an event. We discuss briefly the method of assignment.

1.3.1 For players with ratings from foreign rating systems:

If an unrated player has either a FIDE rating, a CFC rating, or a rating from any other foreign rating system, then the initial USCF rating assigned for that player would be computed using a conversion derived by the Ratings Committee. The Ratings Committee has already proposed a FIDE-to-USCF conversion (USCF = FIDE + 50), and plans to explore conversions for other rating systems.

1.3.2 For players without any rating:

The Ratings Committee proposes to assign ratings to unrated players according to age, and possibly other important demographic information. Though the Ratings Committee continues to explore the initial assignment, a possible assignment, though hypothetical, might be

$$\text{Initial Rating} = \begin{cases} \text{Age} \times 100 & \text{If Age} \leq 14 \\ 1500 & \text{If Age} > 14 \end{cases}$$

We emphasize that this assignment is entirely hypothetical; we do not suggest adopting this rule. The Ratings Committee plans to obtain information from the USCF office on provisionally rated players in order to construct a statistically appropriate rule for initial assignments.

In situations where an unrated player is suspected of possessing an ability that would not accurately be estimated by the above statistical procedure, because, for example, an unrated player may be known to have had tournament experience in a foreign country, the Ratings Committee considers allowing the tournament director to estimate an initial rating. We propose that tournament directors only with, say, national level accreditation be given the privilege to assign initial ratings to unrated players competing in their events.

1.4 The K schedule

The values of K in the rating update formula are proposed to be larger for a provisionally rated player or an unrated player than for an established player. If a player has already played a total of N rated games ($N < 20$), and then competes in an m-round event, we propose that the value of K to update his rating should be

$$K = \begin{cases} K_0 - \left(\frac{2N+m+1}{2}\right)\left(\frac{K_0-K_{20}}{20}\right) & \text{if } N+m \leq 20 \\ K_0 - \left(\frac{N+21}{2}\right)\left(\frac{K_0-K_{20}}{20}\right) + \left(\frac{N+m-20}{m}\right)K_{20} & \text{if } N+m > 20 \end{cases}$$

where K_{20} is the value of K for established players (32, 24 or 16 depending on the player's rating), and K_0 is the initial value of K for someone who has not yet played in a tournament. The value of K_0 in the formula will also depend on the player's current (or assigned, if unrated) rating; probably somewhere near 150–250 if the current rating is below 2100, 3/4 as small if the current rating is between 2100–2399, and 1/2 as small if the current rating is 2400 or higher. For example, if a player who has already played 10 rated games has a rating of 2200, then K_0 and K_{20} both acquire their "intermediate" values in the K-formula. The Ratings Committee is still investigating how to choose K_0.

The equation above can be interpreted in the following manner. Construct values $K_1, \ldots, K_{19}$ that are equally spaced between K_0 and K_{20}. For a player who has already played N games and competes in an m-round event, the value of K to use is the average of $K_{N+1}, K_{N+2}, \ldots, K_{N+m}$. If m games in an event results in more than 20 games, then for each game over 20 use K_{20} in the average.

Here are a two examples (assuming $K_0 = 200$):

1. A 10 year old unrated player is assigned a rating of 1000, and plays in a 4-round tournament. For this situation, $N = 0$ and $m = 4$. According to the formula, the value of K that should be used for updating his rating is $200 - (2(0) + 4 + 1)/2 \times (200 - 32)/20 = 179$.

2. For a provisionally rated player having played 17 games with a rating of 1673 who has competed in a 6-round tournament ($N = 17$, $m = 6$), the value of K to use should be $200 - (17 + 21)/2 \times (200 - 32)/20 + (17 + 6 - 20)/6 \times 32 = 56$.

The computation to update a player's rating when it moves across the 2100 or 2400 boundaries is analogous to the that of the established formulas. The values of K_0 and K_{20} change to reflect crossing these boundaries.

1.5 Conclusions

The proposed "unified" system is simply the established rating system augmented by a method of assigning initial ratings, and by a method of smoothly varying K from a very high initial value to the current values already in use for established players. The primary motivation for this proposal was the desire to be able to rate events involving an overwhelming number of previously unrated players. As a side effect, the rating system has been simplified. We find great appeal in using a single set of formulas to update ratings rather than using two separate formulas.

2 Conversion of FIDE Ratings to USCF Ratings

2.1 Introduction

The Ratings Committee has examined a method to convert FIDE ratings to USCF ratings. To the knowledge of this committee, no empirical work using appropriate statistical techniques has been performed to convert FIDE ratings to USCF ratings. Besides being of general interest to USCF membership, a conversion has several important practical consequences for FIDE-rated players who have not acquired USCF ratings.

- A conversion can help place a FIDE-rated player into an appropriate tournament section,
- A conversion can estimate a FIDE-rated player's rating for pairing purposes within a tournament section,
- A USCF rating that has been converted from a FIDE rating can be used in the updating formulas so that the opponents of the FIDE-rated player will be rated against the converted rating rather than against a provisional rating.

We describe below the analysis performed to obtain our conversion. The generality of our methods suggests that conversions from other chess ratings systems, not just from FIDE's system, can be performed analogously.

2.2 Analysis

We identified players that appeared in the January 1993 USCF Rating supplement who also appeared in the January 1993 FIDE Rating supplement. Because our goal was to convert from FIDE ratings to USCF ratings, we wanted to obtain a sample of players that could be treated as representative of the population of FIDE players who would play in USCF-rated tournaments. To this end, we only included players that had established USCF ratings and had played in at least 10 FIDE-rated games in the prior six months before the publication of the FIDE rating supplement. We also restricted our attention to players who had FIDE ratings of at least 2200 (players with FIDE ratings under 2200 may be viewed as a separate population). Imposing these restrictions resulted in a total of 165 players used in our analysis.

Figure 1 shows a plot of the USCF ratings against the FIDE ratings for the 165 players. Apart from some points corresponding to players with unusually low USCF ratings, the pattern of data appears linear. A robust linear regression fit using bisquare weighting[1] resulted in an estimated slope of 1.05. Because the USCF and FIDE updating formulas are virtually identical, we would expect the slope to be 1 (which is mathematically equivalent to expecting that a USCF rating would equal a FIDE rating plus a constant). The data provided no evidence to the contrary so we set the slope to 1. We estimated the constant by calculating a 10% trimmed mean difference between the FIDE and USCF ratings for the 165 players in the data set. A trimmed mean is calculated by computing all the USCF–FIDE differences, removing the top 5% and bottom 5% of the values, and calculating the mean of the remainder. Using a trimmed mean rather than the mean of all values prevents the unusually low USCF ratings seen in Figure 1 from inappropriately affecting the estimation process. The trimmed mean analysis resulted in a value of 51.3, which suggests that an appropriate conversion to estimate a USCF rating from a FIDE rating is

$$\text{USCF} = \text{FIDE} + 50.$$

This conversion is shown in Figure 2. The line appears to adequately summarize the relationship between USCF and FIDE ratings for this sample. Using trimming percents of 30%, 20%, or 5% does not change our conclusions substantially.

The amount of scatter around the line differs depending on the FIDE rating – the higher the FIDE rating, the closer the points appear to be near the line. This suggests that players with high FIDE ratings are more accurately estimated on the USCF scale than players with low FIDE ratings.

Figure 3 shows a plot of the difference in USCF and FIDE ratings for the 165 players with established ratings against their FIDE rating. The solid horizontal line represents an additive conversion of 50, while the dotted lines show conversions of 0 and 100. Conversions of 0 or 100 appear too extreme to be considered seriously.

2.3 Conclusions

The recommendation of the Ratings Committee is to adopt the 50 point conversion for estimating foreign FIDE-rated players for purposes of pairing within tournaments and for calculating opponents' rating updates.

[1]Mosteller, F., Tukey, J. W, *Data Analysis and Regression*, Addison-Wesley Publishing Company, Reading, Massachusetts (1977), pg. 356

3 Tournament Sections Based on Lifetime Titles

(Proposed by Ken Sloan)

3.1 Resolved

Tournament organizers are encouraged to experiment with Class Sections based on Lifetime Titles, rather than current Ratings.

3.2 Rationale

Class prizes and sections are, for better or worse, essential to USCF. They provide good, well balanced competition for players of all abilities. They are one of the more successful applications of the USCF rating system.

The fact that USCF ratings are used to divide players into sections has one major disadvantage: a small difference in rating can have a very large impact on whether or not a player will be eligible for a particular Class Prize. This difference is much smaller than the normal amount of fluctuation in rating. Worse, the player himself can manipulate his rating to remain (permanently) below the threshold marking a Class boundary. Thus, the fact that ratings are used to set class boundaries provides a motivation to manipulate the ratings. This means that class boundaries become suspect, and the original purpose of the rating scale is subverted.

Even without the possibility of manipulation, it is not good that small, expected fluctuations in rating can have large, catastrophic changes in section eligibility. This is not a large problem for players who are improving, and moving steadily up. It is a problem for players who reach a steady-state strength very close to a class boundary.

The USCF rating system is intended to track performance. The use in the setting of class boundaries is unnatural, and has the potential for distortion.

Lifetime Titles recognize sustained, confirmed performance. They are more stable than ratings, do not fluctuate, and are affected only by positive results. For all of these reasons, they are a better mechanism for awarding Class prizes and organizing Class sections.

In many other sports, a Class player is allowed a few tournament victories and then required to move up to a higher division. The Lifetime Title system can do the same thing. If you perform well above your current class (perhaps winning a class prize), you accumulate points towards the next class; eventually the next class title is awarded.

Lifetime Titles respond only to positive results. This removes the temptation to manipulate the system. It also means that rating fluctuations have no effect. A player's progress though the title system is slow, steady, and monotonic.

The above rationale supports the conclusion that it would be better to award class prizes, and organize class sections, on the basis of Lifetime Titles. For example, a tournament which formerly had two sections: Open and Reserve (under 1800) might instead have the same two sections, labeled Open and Reserve (Certified Class B and below). The only difference would be that a player with one recent good result, raising his rating*title to 1805*b7 would still be eligible for the Reserve Section.

Tournament organizers are already free to organize tournaments in this fashion. This proposal simply calls for USCF to encourage such sections, and develop a standard terminology for advertising and reporting purposes.

A Note on the Conversion from FIDE to USCF Ratings

Introduction

Section 2 of this report discusses the accurate conversion of a FIDE rating to a USCF rating for purposes of assigning an initial rating to unrated players. The Ratings Committee was asked to provide another conversion for which there would be only (for example) a 10% probability that the converted rating would be too high. This conversion is intended for use by tournament directors who want to protect USCF rated players from foreign players when they have reason to believe that the foreign players are in fact stronger than their FIDE rating would indicate. We emphasize that the conversion developed here is for *tournament sectioning only*, and not for initializing a player's actual rating.

The Protection Factor

The conversion formula we develop is of the form

$$\text{USCF}_{\text{practical}} = \text{FIDE} + 50 + \text{“protection factor.”}$$

The "protection factor" is an extra positive quantity added to the accurate conversion of FIDE + 50 to ensure a small probability that a player's true USCF rating is greater than the "practical converted rating."

To obtain the protection factor, we carry out the following procedure.

1. Determine the standard deviation of USCF ratings around the conversion line for a player's FIDE rating (see Figure 2). Because the variability of USCF ratings is wider for lower FIDE ratings, the standard deviation will be inversely related to FIDE rating.

2. Specify the "protection probability." We define the protection probability as the chance an unrated player's true USCF rating is higher than the rating assigned to him.

3. Obtain the protection factor by multiplying the standard deviation in (1) by a value which equals the number of standard deviations on a normal distribution beyond which the probability equals the "protection probability" from (2).

This analysis results in a simple formula easily used by a tournament director with a hand-held calculator.

The standard deviation of USCF ratings around the conversion line (in Figure 2) was found using the following technique. For each FIDE rating x, $x = 2200, 2210, 2220, \ldots, 2600$, we computed a robust estimate for the standard deviation of USCF values around the conversion[1] for FIDE ratings ranging from $x - 100$ to $x + 100$ for each x. So, for example, at $x = 2450$, we computed a robust estimate of the standard deviation of USCF values around the conversion line corresponding to FIDE ratings between 2350 and 2550. As in the analysis of the previous section, we ignored FIDE ratings below 2200. Figure 4 displays the estimates of the standard deviations for each x. The standard deviations appear roughly linear with respect to FIDE ratings, so we fit a least-squares regression line through the points to obtain a relationship between the standard deviation and FIDE rating. The equation we computed was

$$S = 66.9 - 0.0819(\text{FIDE} - 2200),$$

where S is the approximate standard deviation given a player's FIDE rating.

Choosing a protection probability involves the decision of how much protection is wanted for the rated tournament participants. The choice of a specific protection probability is not properly in the domain of the Ratings Committee. For the development of this section, however, we use a protection probability of 10% as an example. This is equivalent to allowing only a 10% chance for the player to have a true rating higher than the one which will be assigned to him. The number of standard deviations above the mean over a normal distribution beyond which 10% of the probability lies is equal to 1.28. Analogous values for different choices of protection probabilities can be found using a table of the normal distribution found in any elementary statistics textbook.

The protection factor corresponding to a protection probability of 10% is equal to $1.28S$, or

$$\text{"protection factor"} = 1.28 \times (66.9 - 0.0819(\text{FIDE} - 2200)) = 85.7 - 0.105(\text{FIDE} - 2200).$$

The overall practical conversion is therefore

$$\begin{aligned} \text{USCF}_{\text{practical}} &= \text{FIDE} + 50 + \text{"protection factor"} \\ &= \text{FIDE} + 50 + 85.7 - 0.105(\text{FIDE} - 2200) \\ &= \text{FIDE} + 135.7 - 0.105(\text{FIDE} - 2200), \end{aligned}$$

or even more simply, $0.895(\text{FIDE}) + 366.7$.

To show how this formula can be used, suppose an unrated player enters a tournament with a FIDE rating of 2310. The formula from Section 2 suggests that our best guess of his true USCF rating is

[1]We used the "median absolute deviation" multiplied by 1.4826. The resulting value is a "consistent" estimate of the standard deviation for approximately normally distributed data. A robust estimate is one that is not affected substantially by unusual data values, such as those that appear on Figure 2.

$2310 + 50 = 2360$ which would ordinarily entitle him to compete in an U2400 section. Using the formula above, we obtain a "practical" conversion equal to $2310 + 135.7 - 0.105(2310 - 2200) = 2434$. In this case, there is only a 10% chance that this player's true USCF rating is greater than 2434. Thus if we use the practical conversion because a tournament director is concerned that this player may be stronger than his FIDE rating indicates, we obtain a rating of 2434 which precludes him from competing in an U2400 section.