

The level of play in CCRL and FIDE rating systems

Author: Erik Varend
erikvarend@hotmail.com

The aim of this paper is to give a brief comparison between two completely different and distinct rating systems and show how a certain increase in the strength of play is reflected by ratings in both rating systems: FIDE ELO and CCRL. The strength of play is represented by the accuracy of play. The methodology used to determine the accuracy of play is largely the same as described in this study with some minor modifications:

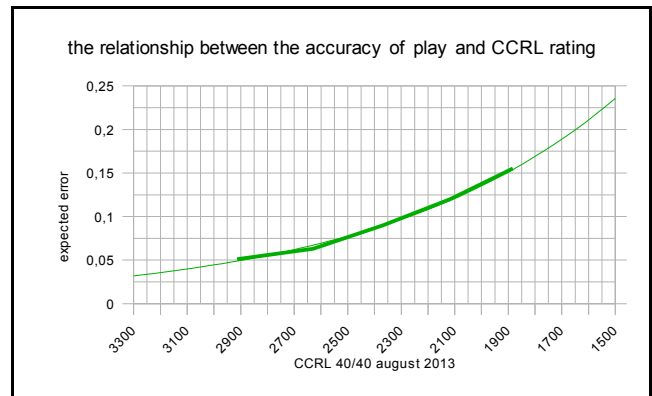
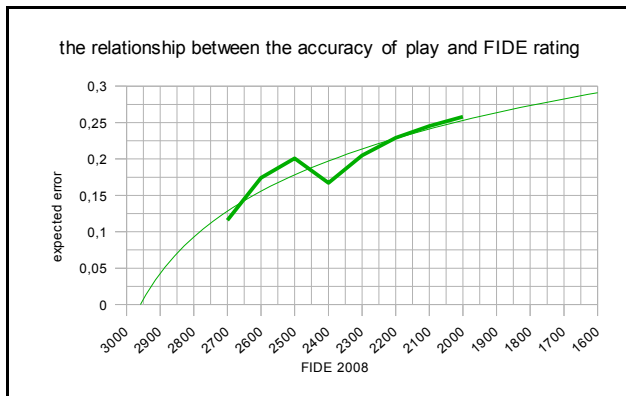
<http://www.chessanalysis.ee/a/%20study/%20on/%20chess/%20strength.pdf>

In a nutshell, the process can be divided into 3 phases:

1. Find out the actual average error. It is calculated by taking the eval of best move by Rybka and subtracting the evaluation of the move made by a player from it.
2. Next find out various difficulty parameters.
3. Derive the average expected error on the basis of difficulty factors and actual average error.

The name 'expected' means that it is not the 'real' value, but a hypothetical one in circumstances where the difficulty of positions is equal for all players so as to make valid comparisons.

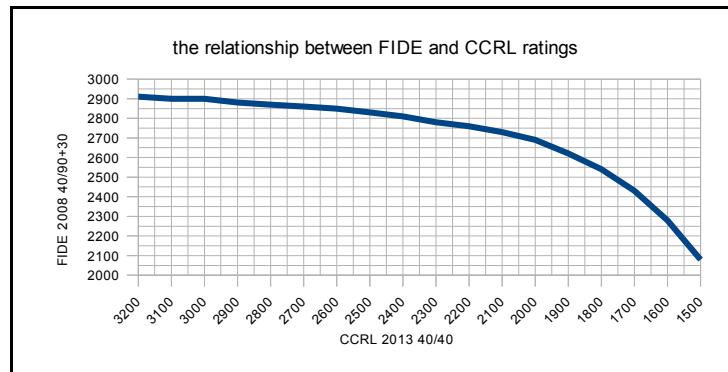
Data from 3778 positions and randomly picked 82 games by human players from 2008; and 1377 positions and 43 games from 2009 from CCRL 40/40 archive by 5 different engines (Micro-Max 4.8, Waxman 2008, Philou 2.8.0, Crafty 23.0 and Hiarc 12.1) were used to construct charts below. The average time control for human games was 158 seconds per move. The standard FIDE time control is 90/40 + 30", 165 seconds per move. Both values are very close, so further adjustments for thinking time were not necessary.



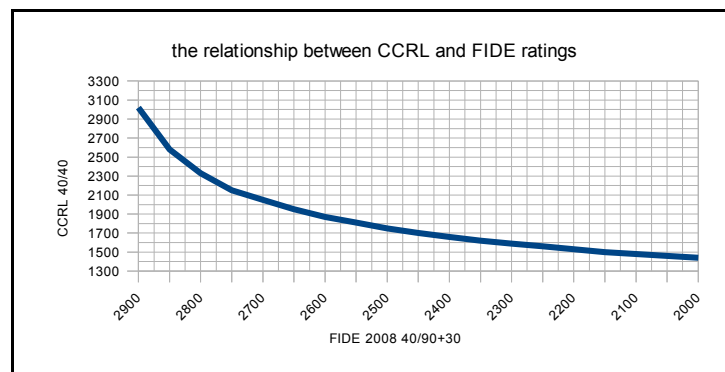
As we can see, perhaps somewhat surprisingly, we have opposite trend lines. Human players have a logarithmic relation to the rating system, whereas computers' accuracy of play decreases exponentially in relation to ratings, which is the inverse of logarithmic function. The differences in the accuracy of play gradually get smaller towards the bottom end of the rating scale on the FIDE list. In engine rating lists it's the upper end where differences get less distinguishable. It should be also noted how neatly engines' accuracy of play follows exponential curve, which demonstrates the preestablished fact - computers play with higher consistency than humans.

The two charts made it possible to establish conversion tables for finding strength equivalent for both rating systems. The tables and charts below show the relationship between FIDE and CCRL 40/40 ratings.

CCRL 40/40	FIDE 2008 40/90+30
3200	2910
3100	2900
3000	2900
2900	2880
2800	2870
2700	2860
2600	2850
2500	2830
2400	2810
2300	2780
2200	2760
2100	2730
2000	2690
1900	2620
1800	2540
1700	2430
1600	2280
1500	2080



FIDE 2008 40/90+30	CCRL 40/40
2900	3020
2850	2580
2800	2330
2750	2150
2700	2050
2650	1950
2600	1870
2550	1810
2500	1750
2450	1700
2400	1660
2350	1620
2300	1590
2250	1560
2200	1530
2150	1500
2100	1480
2050	1460
2000	1440



The FIDE equivalent of 3000+ ratings on CCRL may seem too low at first glance, but one must take into consideration that CCRL games are run on a relatively outdated hardware and at nearly 3 times as short time controls. The conclusions this paper offers are not definitive there are room for improvement, and numbers in the two conversion tables should be regarded as approximate values. But it nevertheless turns out that expectations that top engines on up-to-date desktop machines are supposed to perform 3100-3200 against humans are a myth.