Tiebreak proposals for Swiss tournaments

(mei 2020) C.L. Pippel, IJmuiden

Discussions about tiebreaks goes back to as far as 1873 (Ahrends, p. 184). Vivid discussions arose when the famous 1914 Mannheim chess tournament was interrupted by World War One. At that point Alekhine was a half point ahead. But Vidmar did have met stronger opponents. The issue of defining rankings for tournaments has been studied in various fields like sports, psychology, decision theory, Condorcet voting. Many ranking methods based on different motivations have been defined.

Direct methods

Solkoff, Buchholz, and variants

The above methods have a practical disadvantage. If a player leaves the tournament early the Solkoff contribution will be flatted in a negative sense. His opponents are thereby disadvantaged. The FIDE has found a solution for this (Forlani). But this solution is complicated and artificial. All direct methods have a fundamental limitation. They do take into account the direct opponents, but not the opponents of the opponents. And the opponents thereof. Four alternative indirect tiebreak methods are presented.

Relative Elo Ratings

The normal distribution function

"From general experience in sports we know that the stronger player does not invariably outperform the weaker. A player has good days and bad, good tournaments and bad. By and large at any point in his career, a player will perform around some average level. Deviations from this level occur, large deviations less frequently than small ones. These facts suggest the basic assumption of the Elo system. It is best stated in the formal terms of statistics:"

The many performances of an individual wil be normally distributed, when evaluated on an appropriate scale (Elo, 1978, p. 19).

Development of the Percentage Expectancy Table

The normal probabilities may be taken directly from the standard tables of the areas under the normal curve, when the difference D in ratings is expressed as a z score. The z value of a difference equals D / sigma, where sigma = 2000 / 7. For example, let D = 160. Then z = 160 * 7 / 2000 = 0,56. The table / formula gives ,7123 and ,2877 as areas of the two portions under the cumulative curve. These numbers represents the expected scores of the two opponents (Elo, p. 158).

In a rating system (KNDB, FMJD) ratings are updated at regular time intervals. Relevant games are considered. Game by game the difference between actual and expected score is established, according to the percentage expectancy function. Then the existing rating is updated, proportional to the total difference between expected and actual score. However, when computing the relative Elo ratings, we do not assume any existing ratings prior to the tournament. All players in the tournament are assigned a rating, the relative Elo rating, such that for each player the expected score and the actual score are equal. This rating is calculated by a numeric iteration procedure, as no closed formula exists. The relative Elo rating of a tournament is unique, up to a constant.

Example Open de Sangmelima Cameroon 2014

At the start of the iterations (R0) one assumes all players have equal strength. In the subsequent iterations the rating of each player is updated according to -ln(We/W) * 2000 / 7. "We" is the expected score derived from the cumulative normal distribution, W is the actual score and ln() constitutes the natural logarithm function. The updates are normalized to set the sum of all updates to zero. The iteration is slow, but stable. We may choose at start any average rating level including zero. This does not affect the ordering of the players.

Pl Title	Name	Cn	Rating	+ =	- 1	N Pts	Whl	Wl	1	R0	We0	Up0	1	R1	We1	Up1	1	R2	We2	Up2	1	Rk	R28	
1 gmi	Leopold Kouomou	cm	2296	3 3	0	6 9	25	34	1	0,0	6,0	137,7		137,7	8,0	32,6	-	170,3	8,1	36,2	-	2	304,0	
2	Gerard Ngankou	cm	2265	3 3	0	5 9	25	34	1	0,0	6,0	137,7		137,7	7,7	42,9	-	180,6	8,0	38,3	1	1	312,7	
3	Mouanji Iliassou	cm		4 0	2	6 8	16	23	1	0,0	6,0	104,0		104,0	9,0	- 33 , 9	-	70,1	8,5	-10,6	-	6	18,3	
4	Armand Abouem	cm		3 1	2	5 7	30	39	-	0,0	6,0	65,9		65 , 9	6,0	41,7		107,6	6,5	25,6	-1	3	185,3	
5	Landry Nga	cm		1 5	0	5 7	28	37	-	0,0	6,0	65,9		65 , 9	6,3	31,1		97,0	6,6	20,5	-1	4	180,9	
6 mi	Bruno Fopa	cm	2262	2 3	1	5 7	21	28	-	0,0	6,0	65,9		65 , 9	7,7	-28,3		37,5	7,5	-12,8	-1	9	-26,6	
7	Desire Ghuendou	cm		2 3	1	5 7	18	25	-	0,0	6,0	65,9		65 , 9	8,2	-45,6		20,3	7,6	-17,6	-1	10	-63,0	
8	Maturin Tomi	be		2 2	2	6 6	26	35	1	0,0	6,0	21,8		21,8	6,4	-17,1	-	4,7	6,0	6,0	-	5	33,6	
9	Arnaud Foto	cm		2 2	2	6 6	25	32	1	0,0	6,0	21,8		21,8	6,3	-15, 9	-	6,0	6,2	-3,6	-	8	-19,4	
10	Patrick Akono	cm		1 3	2	5 5	31	40	1	0,0	6,0	-30,3		-30,3	4,4	36,2	-	5,9	5,2	-5,6	-	7	-16,3	
11	Bernard Mambo	cm		2 0	4	6 4	28	37	-	0,0	6,0	-94,0		-94,0	3,7	18,8		-75 , 3	4,2	-8,6	-1	11	102,5	
12	David Wabo Daco	cm		1 2	3	6 4	23	31	-	0,0	6,0	-94,0		-94,0	4,7	-45,7		-139,7	4,1	-3,3	- 1	12	177,2	
13	Barel Bimogo	cm		1 1	4	6 3	23	31	1	0,0	6,0	-176,2		-176,2	3,8	-66,9	-	-243,1	3,1	-5,4	-	13	299,4	
14	Prince Mvondo	cm		1 0	5	6 2	28	37	-	0,0	6,0	-292,1		-292,1	1,7	50,2	1	-241,8	2,5	-59,1	-	14	330,6	

It is easy to check the results of the above calculation:

Gerard Ngankou	313,0		Leopold Kouogueu Kouomou	304,0	
Opponents	R28	Diff We	Opponents	R28	Diff We
Patrick Akono	-16,3	329,3 1,75	Prince Arnaud Mvondo	-330,6	634,6 1,97
Landry Nga	180,9	132,1 1,36	Maturin Nyamsi Tomi	33,6	270,4 1,66
Maturin Nyamsi Tomi	33,6	279,4 1,67	Patrick Akono	-16,3	320,3 1,74
Bernard Mambo	-102,5	415,5 1,85	Armand Abouem	185,3	118,7 1,32
Leopold Kouogueu Kouomou	304,0	9,0 1,03	Gerard Ngankou	312,7	-8,7 0,98
Armand Abouem	185,3	127,7 1,35 +	Landry Nga	180,9	123,1 1,33 +
Expected Score		9,0 points	Expected score		9,0 points

Game file

```
Pl
                            Rtg R1
                                     R2
                                          R3
                                               R4
                                                         R6
                                                    R5
1 gmi Leopold Kouogueu Kouomou 2296 2/14b 1/8b 2/10w 2/4w 1/2b 1/5b
     Gerard Ngankou
                         2265 1/10b 1/5b 2/8w 2/11b 1/1w 2/4b
     Mouanji Iliassou 0/4w 0/10b 2/13b 2/14b 2/12w 2/11b 8
3
     Armand Abouem
                              2/3b 2/9b 2/11w 0/1b 1/5w 0/2w
                             1/8w 1/2w 1/12b 2/9w 1/4b 1/1w
5
     Landry Nga
                      2262 0/12w 1/7b 2/14w 2/8b 1/10b 1/9w
6 mi Bruno Fopa
7
     Desire Ghuendou
                             0/11w 1/6w 1/9b 1/13b 2/14w 2/10b 7
     Tomi Maturin Nyamsi 1/5b 1/1w 0/2b 0/6w 2/13w 2/14b 6
8
9
     Arnaud Foto
                              2/13b 0/4w 1/7w 0/5b 2/11w 1/6b
                         1/2w 2/3w 0/1b 1/12b 1/6w 0/7w
10
     Patrick Akono
                             2/7b 2/12w 0/4b 0/2w 0/9b 0/3w
11
     Bernard Mambo
                          2/6b 0/11b 1/5w 1/10w 0/3b 0/13w 4
12
     David Daco Wabo
                              0/9w 0/14b 0/3w 1/7w 0/8b 2/12b 3
13
     Barel Bimogo
     Prince Arnaud Mvondo 0/1w 2/13w 0/6b 0/3w 0/7b 0/8w 2
14
```

FMJD report

Domain

The iteration requires an indivisible (that is, strongly connected or irreducible) domain. In every possible partition of players into two nonempty subsets, some player of the second set has defeated at least one player in the first set (Glickman, p. 5). Or alternatively, there exists a directed path between any two players. Finding all maximal strongly connected components is a mathematical puzzle in its own right.

Remark: one can introduce a virtual opponent who draws to all other players to make the domain indivisible.

Skew-symmetric score systems

A score system R is skew-symmetric if for all possible outcomes (x, y) x + y = 0. Any traditional score system can be transformed to this form (Chebotarev, 1989, p. 1104):

where

- A represents the footbal score; win point = 3.
- R = (A transpose(A)) / 2.
- Divide by two to preserve the difference between own en opponent score.

Least squares method

Principle

If a player has a win, loss or draw his comparison outcome is set to 1, -1 or 0 respectively. The least squares method constructs a mean square approximation of the skew-symmetric comparison outcomes R by the differences of the desired least squares ratings (Cheboratev, 1999, p. 18):

Determine $q = (q_1,...,q_n)$ such that $\sum_i \sum_i (r_{ii} - (q_i - q_i)n_{ii})^2$ is a minimum (Gulliksen, 1956)

where

- n_{ii} is the number of games between player i and j, i ≠ j, zero otherwise (matches matrix).
- rii equals the sum of the results of player i against j.

Calculation

The LSM-ratings (q) are a solution of the set of linear equations:

- L.q = s.
- $\Sigma q = 0$.

where

- L = (Iii) is the Laplacian matrix of the tournament graph.
- Iii = Number of comparisons of player i.
- I_{ii} = Negative the number of games between player i and j, i ≠ j.
- $s = (s_i)$, $s_i = \Sigma_k r_{ik}$.
- $\bullet\,$ Skew-symmetric draughts scores r_{ik} are equivalent to (wins -/- losses).

Round-robin tournament

LSM-ratings q of a round-robin tournament:

 \bullet q = s / (N + r), where r = 1, 2, 3 for a single, double, triple round-robin tournament.

Note that if the percentage expectancy function of the relative Elo ratings is replaced by a linear relationship, for example the "algorithm of 400", then relative Elo ratings are identical to LS-ratings except for a scaling factor. Multiplying LS-ratings with a factor (400 / Rmax) gives a comparable rating in the Elo domain (FIDE Rating Regulations effective from 1 July 2017, table 8.1b).

Example Open de Sangmelima Cameroon 2014

Pl	Title	Name	Cn	Rating	+	Ν	s	Whl	Wl	Rk	LS-Rtg	r l
1	gmi	Leopold Kouomou	cm	2296	3	6	3	25	34	2	0,643	
2		Gerard Ngankou	cm	2265	3	6	3	25	34	1	0,708	
3		Mouanji Iliassou	cm		4	6	2	16	23	5	0,059	
4		Armand Abouem	cm		3	6	1	30	39	3	0,420	
5		Landry Nga	cm		1	6	1	28	37	4	0,392	
6	mi	Bruno Fopa	cm	2262	2	6	1	21	28	9	-0,042	
7		Desire Ghuendou	cm		2	6	1	18	25	11	-0,122	
8		Maturin Tomi	be		2	6	0	26	35	6	0,054	
9		Arnaud Foto	cm		2	6	0	25	32	110	-0,047	
10		Patrick Akono	cm		1	6	-1	31	40	8	-0,030	
11		Bernard Mambo	cm		2	6	-2	28	37	112	-0,234	
12		David Wabo Daco	cm		1	6	-2	23	31	13	-0,425	
13		Barel Bimogo	cm		1	6	-3	23	31	15	-0,694	
14		Prince Mvondo	cm		1	6	-4	28	37	14	-0,683	ı

It is easy to check the result of the above calculation as follows:

Gerard Ngankou	0,708	3	Leopold Kouogueu Kouomou	0,643	3
Opponents	LS-Rtg	Diff	Opponents	LS-Rtg	Diff
Patrick Akono	-0,030	0,738	Prince Arnaud Mvondo	-0,683	1,326
Landry Nga	0,392	0,316	Maturin Nyamsi Tomi	0,054	0,589
Maturin Nyamsi Tomi	0,054	0,654	Patrick Akono	-0,030	0,673
Bernard Mambo	-0,234	0,942	Armand Abouem	0,420	0,223
Leopold Kouogueu Kouomou	0,643	0,065	Gerard Ngankou	0,708	-0,065
Armand Abouem	0,420	0,288	Landry Nga	0,392	0,251 +
s1		3,00 (Wins minus losses)	s2		3,00 (Wins minus losses)

Recursive Buchholz

Let

Rb - Recursive Buchholz ratings

 $\begin{array}{ll} \text{RbOpp} & \text{- Average opponents recursive Buchholz ratings} \\ \text{s} & \text{- } (s_i) \text{ is sum of the skew-symmetric score points } r_{ik} \end{array}$

p - Average of s

Rb, RbOpp, s, p are n×1 column vectors indexed by player. The recursive Buchholz rating Rb is the result of the following iteration procedure: (González-Díaz et al., 2013, p. 6).

- Rb = RbOpp + p.
- ΣRb= 0.

Example Open de Sangmelima Cameroun 2014

Pl Tit	le Name	Cn	Rating -	+ = -	N E	Pts	Whl	Wl [,1]	[,2]	[,3]	[,4]	[, 5]	[,6]	[,7]	[,8]	[, 9]	[,10]	[,11]	[,12]	 [,16]
1 gmi	Leopold Kouomou	cm	2296	3 3 0	6	9	25	34	0	0.5000	0.5000	0.6250	0.6103	0.6416	0.6358	0.6436	0.6417	0.6436	0.6431	0.6436	 0.6435
2	Gerard Ngankou	cm	2265	3 3 0	6	9	25	34	0	0.5000	0.5556	0.6759	0.6728	0.7036	0.6993	0.7075	0.7056	0.7079	0.7071	0.7078	 0.7077
3	Mouanji Iliassou	cm	4	4 0 2	6	8	16	23	0	0.3333	0.0278	0.1157	0.0478	0.0729	0.0553	0.0626	0.0578	0.0600	0.0586	0.0593	 0.0590
4	Armand Abouem	cm		3 1 2	6	7	30	39	0	0.1667	0.3611	0.3611	0.4105	0.4070	0.4188	0.4173	0.4201	0.4196	0.4203	0.4201	 0.4202
5	Landry Nga	cm		1 5 0	6	7	28	37	0	0.1667	0.3056	0.3380	0.3781	0.3791	0.3907	0.3889	0.3924	0.3914	0.3925	0.3921	 0.3923
6 mi	Bruno Fopa	cm	2262 2	2 3 1	6	7	21	28	0	0.1667	0.0000	0.0046	-0.0324	-0.0328	-0.0397	-0.0404	-0.0414	-0.0419	-0.0418	-0.0421	 0.0421
7	Desire Ghuendou	cm	2	2 3 1	6	7	18	25	0	0.1667	-0.0833	-0.0509	-0.1173	-0.1048	-0.1222	-0.1174	-0.1223	-0.1206	-0.1220	-0.1214	 0.1217
8	Maturin Tomi	be	2	2 2 2	6	6	26	35	0	0.0000	0.0278	0.0370	0.0486	0.0507	0.0526	0.0536	0.0535	0.0540	0.0538	0.0540	 0.0540
9	Arnaud Foto	cm	2	2 2 2	6	6	25	32	0	0.0000	-0.0278	-0.0324	-0.0455	-0.0419	-0.0474	-0.0450	-0.0471	-0.0460	-0.0468	-0.0464	 0.0466
10	Patrick Akono	cm	:	1 3 2	6	5	31	40	0	-0.1667	0.0556	-0.0648	-0.0046	-0.0401	-0.0227	-0.0329	-0.0278	-0.0307	-0.0292	-0.0300	 0.0298
11	Bernard Mambo	cm	2	2 0 4	6	4	28	37	0	-0.3333	-0.1944	-0.2593	-0.2215	-0.2423	-0.2300	-0.2370	-0.2329	-0.2352	-0.2339	-0.2347	 0.2344
12	David Wabo Daco	cm	:	1 2 3	6	4	23	31	0	-0.3333	-0.3889	-0.3981	-0.4221	-0.4169	-0.4256	-0.4224	-0.4255	-0.4240	-0.4251	-0.4245	 0.4248
13	Barel Bimogo	cm	:	1 1 4	6	3	23	31	0	-0.5000	-0.5833	-0.6667	-0.6690	-0.6907	-0.6876	-0.6941	-0.6922	-0.6943	-0.6934	-0.6941	 0.6940
14	Prince Mvondo	cm		1 0 5	6	2	28	37	0	-0.6667	-0.5556	-0.6852	-0.6559	-0.6853	-0.6772	-0.6843	-0.6820	-0.6838	-0.6832	-0.6836	 0.6835

Note:

- The procedure converges iff the matches matrix is not bipartite (Csató, 2015, p. 8).
- Although the perspective of RB and LSM are quite different, both ratings coincide (González-Díaz et al., 2013, p. 10).

Generalized row sum method

Principle

Row sum scores (s) are calculated as the sum over k of the skew-symmetric outcomes r_{ik} . The generalized row sum method is an extension of the row sum method. The idea of the generalization is to complete the missing comparisons. Central to the statistical interpretation of the generalized row sum method is the assumption that the expected value $E(r_{ij})$ of a missing comparison is proportional to the difference of the scores x_i , x_i of the compared players i and j:

- $x_i = s_i + \Sigma_j E(r_{ij})$, where r_{ij} is missing, $s_i = \Sigma_k r_{ik}$ where r_{ik} is defined and
- $E(r_{ij}) = (x_i x_i) / \gamma$, for any positive factor γ , see Cheboratev, 1989, Chapter 3 Statistical model, p. 1105.

These relationships form a system of n linear equations in n unknowns x_i . For $y = m.n + \varepsilon^{-1}$ the system of equations is equivalent to:

• $x_i = \Sigma_k(r_{ik} + \varepsilon(x_k - x_i + r_{ik}n.m)), k = 1,...,n.$

where

- ε is a positive parameter.
- ε is well-chosen for given n and m if and only if $\varepsilon^{-1} \ge m(n-2)$.
- $r_{ik} + r_{ki} = 0$, the skew-symmetric outcome between players i, k.
- number of players n > 2, m = maximal number of games between any two players.

Monotonicity. Parameter ϵ is said to be well-chosen if for any outcome matrix R = (r_{ik}) the value of its contribution to x_i is nonnegative for a maximal win and nonpositive for a maximal loss (Chebotarev, 1997, Ch. 5, p. 9).

Calculation

The generalized row sums (x) are a solution of the set of lineair equations:

- $(L + \varepsilon^{-1}I)x = \gamma s$.
- $\Sigma x = 0$.

where

- L = (Iii) is the Laplacian matrix of the matches matrix, I is the identity matrix.
- $y = m.n + \varepsilon^{-1}$.

Note that LS-ratings can be obtained as a limit of the generalized row sum calculation by setting $\varepsilon^{-1} = 0$ and $\gamma = 1$.

Example Trainingsvierkamp Harm Wiersma Huizum 2005 (Blitz)

Statistical model

```
Check expected values xi == si + \Sigma(xi-xj)/\gamma, \gamma = 1/\epsilon + n.m, \gamma = 16 All players are expected to complete an equal number of games
```

Harm Wiersma	2,67	Jan Adema	-2,67
s1	2	s4	-2
Missing games		Missing games	
$(x1-x4)/\gamma$	0,333	(x4-x1)/y	-0,333
$(x1-x4)/\gamma$	0,333	(x4-x1)/y	-0,333
	+		+
s1 + missing gms	2,67	s1 + missing gms	-2,67

Direct method

- The statistical model is useful when the tournament is nearly complete.
- For sparse tournament matrices, the direct method is more practical.

Self-consistency (SC) and self-consistent monotonicity (SCM)

Self-consistency comes down to the following. Suppose we consider two alternatives (players), having the same number of comparisons, and the first alternative as compared to the second one

- · achieves the better scores against the "stronger" opponents or
- · achieves the better scores against the opponents of the same "strengths" or
- achieves the same scores against the "stronger" opponents.

Then the first alternative should be placed higher than the second one in the tiebreak ranking. In the above requirement, "stronger" signifies" is "placed higher in the ranking that is mentioned in the requirement".

To obtain *self-consistent monotonicity* we require that if the first alternative additionally achieves some extra "wins" and/or the second alternative has extra "losses", then the first alternative should remain higher than the second one in the tiebreak ranking. (Chebotarev, 1997, p. 2)

Direct methods as row sum score, Buchholz, Sonneborn-Berger (Neustadtl) are monotonic. Direct methods are rather indifferent to opponent strength and therefore incompatible with self-consistency. Recursive Buchholz and LSM satisfy self-consistency but contradicts monotonicity. Relative Elo ratings and generalized row sum ratings obey self-consistent monotonicity.

Consider the "algorithm of 400", the AVG400 rating system. If the rating difference between player and opponent exceeds 400, the expected plus score percentage is greater the 50%. However this is impossible to achieve. Therefore a strong player can lose points even with a perfect score and a weak player can gain by losing all his games (Elo, p. 144). However, in the context of a tie-break, this is not necessarily undesirable. It can be considered a correction of an overly easy match schedule.

References

W. Ahrends, Zur relativen Bewertung von Turnierpartien, pp. 181-192 in: Wiener Schachzeitung, Georg Marco (ed.), Organ des "Wiener Schach-Club", IV. Jahrgang., nr. 10/11. October-Novemb., Wien, 1901. ANNO

Mark Glickman, Introductory note to 1928. on-line

Harold Gulliksen, *A Least Squares Solution for Paired Comparisons With Incomplete Data*, RB-55-5, Educational Testing Service, Princeton New Jersey, 1955, DOI: 10.1002/j.2333-8504.1955.tb00052.x

Arpad E. Elo, The Rating of Chessplayers, Past&Present, ISHI Press International, Bronx NY 10453, 2008. ISBN 0-923891-27-7. First printing by Arco Pub., New York, 1978.

P. Yu. Chebotarev, Generalization of the Row Sum Method for Incomplete Paired Comparisons, Automation and Remote Control 50 (1989) 1103-1113. on-line

P. Yu. Chebotarev, E. Shamis, From Incomplete Preferences to Ranking via Optimization, arXiv: math/0602552v1, 24 Feb 2006.

A version of this paper was published as: P.Yu.Chebotarev, E.V.Shamis. Constructing an objective function for aggregating incomplete preferences, In: A.Tangian and J.Gruber, eds. Econometric Decision Models: Constructing Scalar-Valued Objective Functions. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, 1997, P.100-124

P. Yu. Chebotarev, E. Shamis, *Preference fusion when the number of alternatives exceeds two: indirect scoring procedures*, J. Frankl. Inst. **336(2)** (1999) 205–226, arXiv:math/0602171v3, 20 Apr 2006.

González-Díaz, J., Hendrickx, R. L. P., & Lohmann, E. R. M. A. *Paired Comparisons Analysis: An Axiomatic Approach to Rankings in Tournaments*, Social Choice and Welfare, **42** (2013) 139–169. on-line

László Csató, A graph interpretation of the least squares ranking method, Social Choice and Welfare, 44(1): 51-69, 2015. arXiv: 1508.06778v1 [cs.GT], 27 Aug 2015.

Csató, L. On the ranking of a Swiss system chess team tournament, *Annals of Operations Research* 254, 17-36 (2017). arXiv:1507.05045v5 [stat.AP] (p. 4)

Luigi Forlano, Treatment of unplayed games for Buchholz tie break: the FIDE rule. on-line