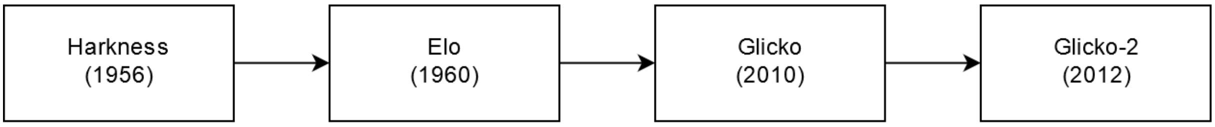
1. Introduction

Rating systems along the history were different, but all of them were trying to compare two unknown people in one field. Probably the most famous attempt to that problem was Elo rating system[1], widely used in chess games. But Elo rating system was improved and introduced as Glicko rating system[2]. The full improvements history can be shown as:



2. Power of a player

It is known that two players in the game can differ a lot with their power in the competition. For them it can be even several orders of magnitude of power difference (like 1 over 1000). But the pointing systems above are handling with that. How? As we can read in Wikipedia[1]:

$$Q = 10^{\frac{R}{400}}$$

Where Q - power of a player, R - his/her rating points

That means if there are two players with a difference of 400 points, the stronger one should be exactly 10 times stronger than the weaker one.
And since 10^(1/400) ≈ 1.005773063, the equation can be simplified to:

$$Q = 1.005773063^R$$
$$\beta = 1.005773063$$
$$Q = \beta^R$$

3. Two player teams

Since Widelands is not simple 1 vs 1 game only, we need to calculate the correct average of the players points. And since the power of the player is not linear, but exponential, the average power of two players is not a simple arithmetic mean.

To get proper equation, we need to determine our goal. For two players with the points R_1 and R_2, we need to find two exactly equal opponents with rating R_3 that their total power is equal to total power of R_1 and R_2:

$$Q_1 + Q_2 = Q_3 + Q_3 = 2 \cdot Q_3$$

Using equation above we get:

$$Q_1 + Q_2 = \beta^{R_1} + \beta^{R_2}$$
$$2 \cdot Q_3 = 2 \cdot \beta^{R_3}$$
$$\beta^{R_1} + \beta^{R_2} = 2 \cdot \beta^{R_3}$$
$$\frac{1}{2}(\beta^{R_1} + \beta^{R_2}) = \beta^{R_3}$$

Solving the equation is by using natural logarithm[3] ln on both sides:

$$\ln\left(\frac{1}{2}(\beta^{R_1} + \beta^{R_2})\right) = \ln(\beta^{R_3})$$
$$\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2) = R_3 \cdot \ln(\beta)$$

$$R_3 = \frac{\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)}$$

It can be also written as (assuming $R_1 < R_2$):

$$R_3 = R_1 + \frac{\ln(1 + \beta^{R_2 - R_1})}{\ln(\beta)} - \frac{\ln(2)}{\ln(\beta)}$$

4. Examples of usage

a) Same power

Let's assume that $R_1 = R_2$, then:

$$R_3 = \frac{\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)} = \frac{\ln(\beta^{R_1} + \beta^{R_1}) - \ln(2)}{\ln(\beta)} = \frac{\ln(2 \cdot \beta^{R_1}) - \ln(2)}{\ln(\beta)} =$$
$$= \frac{\ln(2) + \ln(\beta^{R_1}) - \ln(2)}{\ln(\beta)} = \frac{\ln(\beta^{R_1})}{\ln(\beta)} = \log_{\beta}(\beta^{R_1}) = R_1$$

So as we expected, the formula gets us to the same rating points as the initial players.

b) Difference of 200 (decent difference, but not huge one)

Let's assume that $R_2 = R_1 + 200$, then:

$$R_3 = R_1 + \frac{\ln(1 + \beta^{R_2 - R_1})}{\ln(\beta)} - \frac{\ln(2)}{\ln(\beta)} = R_1 + \frac{\ln(1 + \beta^{R_1 + 200 - R_1})}{\ln(\beta)} - \frac{\ln(2)}{\ln(\beta)} =$$
$$= R_1 + \frac{\ln(1 + \beta^{200})}{\ln(\beta)} - \frac{\ln(2)}{\ln(\beta)} = R_1 + \frac{\ln(1 + 3.162278\ldots)}{\ln(1.00577\ldots)} - \frac{\ln(2)}{\ln(1.00577\ldots)} =$$
$$= R_1 + 247.732\ldots - 120.412\ldots \approx R_1 + 127.3$$

So the result is that the expected power of two players is about 127 points higher than the weaker player or about 73 points lower than the stronger one.

c) Finding difference that doesn't affect the stronger player

Last part is about finding the rating that is not affecting the result rating of stronger player and the weaker one. In mathematical language can be written as:

$$R_1 + R_2 \approx 0 + R_2$$

To finish this equation we need to define approximately sign. From a rating system, two players are equal power when their amount of points differs 1 at most. Applying that:

$$2 \cdot R_3 = R_2 + 0$$
$$2 \cdot (R_3 + 1) = R_2 + R_1$$

So let's find out what R_3 equals to:

$$R_3 = \frac{\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)} = \frac{\ln(\beta^{0} + \beta^{R_2}) - \ln(2)}{\ln(\beta)} = \frac{\ln(1 + \beta^{R_2}) - \ln(2)}{\ln(\beta)}$$

Since R_2 is expected to be around 1000 or more, the value of β^{R_2} is expected to be much greater than 1:

$$\beta^{R_2} \gg 1$$

So applying that, we can calculate:

$$R_3 = \frac{\ln(1 + \beta^{R_2}) - \ln(2)}{\ln(\beta)} \approx \frac{\ln(\beta^{R_2}) - \ln(2)}{\ln(\beta)} = R_2 - 120.412\ldots$$

The result is that half of power of sufficient rating score is equal to score decreased by about 120 points. Going further, we apply those numbers to the second equation:

$$R_3 + 1 = R_2 - 120.412 + 1 = R_2 - 119.412 = R_2 - \gamma$$

$$\gamma = 119.412 \ldots$$
$$R_2 + R_1 = \frac{\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)} = R_2 - \gamma$$
$$\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2) = R_2 \cdot \ln(\beta) - \gamma \cdot \ln(\beta)$$
$$\ln(\beta^{R_1} + \beta^{R_2}) = R_2 \cdot \ln(\beta) - \gamma \cdot \ln(\beta) + \ln(2)$$
$$\beta^{R_1} + \beta^{R_2} = \exp(R_2 \cdot \ln(\beta) - \gamma \cdot \ln(\beta) + \ln(2))$$
$$\beta^{R_1} = \exp(R_2 \cdot \ln(\beta) - \gamma \cdot \ln(\beta) + \ln(2)) - \beta^{R_2}$$
$$R_1 = \frac{\ln(\exp(R_2 \cdot \ln(\beta) - \gamma \cdot \ln(\beta) + \ln(2)) - \beta^{R_2})}{\ln(\beta)}$$
$$= \frac{\ln\left(\frac{\exp(R_2 \cdot \ln(\beta))}{\exp(\gamma \cdot ln(\beta))} \cdot \exp(\ln(2)) - \beta^{R_2}\right)}{\ln(\beta)} =$$
$$= \frac{\ln\left(\frac{\beta^{R_2}}{\beta^{\gamma}} \cdot 2 - \beta^{R_2}\right)}{\ln(\beta)} = \frac{\ln\left(\beta^{R_2}\left(\frac{2}{\beta^{\gamma}} - 1\right)\right)}{\ln(\beta)} = \frac{\ln(\beta^{R_2}) + \ln\left(\frac{2}{\beta^{\gamma}} - 1\right)}{\ln(\beta)} =$$
$$= \frac{\ln(\beta^{R_2})}{\ln(\beta)} + \frac{\ln\left(\frac{2}{\beta^{\gamma}} - 1\right)}{\ln(\beta)} = \frac{\mathrm{R}_2 \cdot \ln(\beta)}{\ln(\beta)} + \frac{\ln\left(\frac{2}{\beta^{\gamma}} - 1\right)}{\ln(\beta)} = R_2 + \frac{\ln\left(\frac{2}{\beta^{\gamma}} - 1\right)}{\ln(\beta)}$$
$$R_1 \approx R_3 - 895.44$$

So if the weaker player has lower rating by about 900 points, this doesn't matter to global rating for both players.

d) Comparing two teams with ranks differed by exactly the same amount of points

Let's assume that there are two teams: R_1 and R_2 versus R_3 and R_4. And $R_3 = R_1 + d$, $R_4 = R_2 + d$. Expected result is that average R_1 and R_2 is exactly lower than average R_3 and R_4 by *d* points.

$$R_{12} = R_1 + R_2 = \frac{\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)}$$
$$R_{34} = R_3 + R_4 = \frac{\ln(\beta^{R_3} + \beta^{R_4}) - \ln(2)}{\ln(\beta)} = \frac{\ln(\beta^{R_1+d} + \beta^{R_2+d}) - \ln(2)}{\ln(\beta)} =$$
$$= \frac{\ln(\beta^{R_1} \cdot \beta^d + \beta^{R_2} \cdot \beta^d) - \ln(2)}{\ln(\beta)} = \frac{\ln\left(\beta^d(\beta^{R_1} + \beta^{R_2})\right) - \ln(2)}{\ln(\beta)} =$$
$$= \frac{\ln(\beta^d) + \ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)} = \frac{\ln(\beta^d)}{\ln(\beta)} + \frac{\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)} =$$
$$= \frac{\ln(\beta^d)}{\ln(\beta)} + R_{12} = d + R_{12}$$

So increasing rank of every player in a team by *d* points will increase average rank by *d* points.

5. Calculating Rating Deviation

Glicko and Glicko-2 rating systems contain two variables that define the rank position and player's strength: R - rating points and RD - rating deviation. The second one is a standard deviation of the first value.

To calculate standard deviation of a function of two variables we can use an equation[5]:

$$\sigma_t = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 \cdot \sigma_1^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 \cdot \sigma_2^2}$$

And in our case this formula can be written as:

$$RD_3 = \sqrt{\left(\frac{\partial R_3}{\partial R_1}\right)^2 \cdot RD_1^2 + \left(\frac{\partial R_3}{\partial R_2}\right)^2 \cdot RD_2^2}$$

The only problem is about calculating the derivatives:

$$\frac{\partial R_3}{\partial R_i} = \frac{\partial}{\partial R_i}\left(\frac{\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)}\right)$$
$$\frac{\partial R_3}{\partial R_i} = \frac{\beta^{R_i}}{\beta^{R_1} + \beta^{R_2}}$$
$$RD_3 = \sqrt{\left(\frac{\beta^{R_1}}{\beta^{R_1} + \beta^{R_2}}\right)^2 \cdot RD_1^2 + \left(\frac{\beta^{R_2}}{\beta^{R_1} + \beta^{R_2}}\right)^2 \cdot RD_2^2}$$

6. Example

Let's see what happened when we have: R_1 = 1700, RD_1 = 50, R_2 = 1300, RD_2 = 150.

$$R_3 = \frac{\ln(\beta^{R_1} + \beta^{R_2}) - \ln(2)}{\ln(\beta)}$$
$$R_3 = \frac{\ln(\beta^{1700} + \beta^{1300}) - \ln(2)}{\ln(\beta)}$$
$$R_3 \approx 1596$$

$$RD_3 = \sqrt{\left(\frac{\beta^{R_1}}{\beta^{R_1} + \beta^{R_2}}\right)^2 \cdot RD_1^2 + \left(\frac{\beta^{R_2}}{\beta^{R_1} + \beta^{R_2}}\right)^2 \cdot RD_2^2}$$
$$RD_3 = \sqrt{\left(\frac{\beta^{1700}}{\beta^{1700} + \beta^{1300}}\right)^2 \cdot 50^2 + \left(\frac{\beta^{1300}}{\beta^{1700} + \beta^{1300}}\right)^2 \cdot 150^2}$$
$$RD_3 \approx 47.5$$

So the average rating of two players like this is 1596 and their rating deviation is 47.5.

7. First generalization: more players in a team

Widelands contains games up to 8 players, so it can be up to 4 players in two average size teams. Also it can be played as 1 vs 7, so the largest possible team can contain up to 7 players. To calculate average ratings for such games we need general formula.

Sticking to point 3. way of thinking, we can calculate the average like this:

$$Q_1 + Q_2 + Q_3 + \ \dots + Q_n = \sum_{i=1}^{n} Q_i = n \cdot Q_t$$
$$Q_i = \beta^{R_i}$$

$$\sum_{i=1}^{n} \beta^{R_i} = n \cdot \beta^{R_t}$$

$$R_t = \frac{\ln(\sum_{i=1}^{n} \beta^{R_i}) - \ln(n)}{\ln(\beta)}$$

$$DR_t = \sqrt{\sum_{i=1}^{n} \left(\frac{\partial R_t}{\partial R_i}\right)^2 \cdot RD_i^2}$$

$$\frac{\partial R_t}{\partial R_i} = \frac{\beta^{R_i}}{\sum_{j=1}^{n} \beta^{R_j}}$$

$$DR_t = \sqrt{\sum_{i=1}^{n} \left(\frac{\beta^{R_i}}{\sum_{j=1}^{n} \beta^{R_j}}\right)^2 \cdot RD_i^2}$$

So in conclusion of this point, if there are two teams with the same number of players, the easiest way of thinking about their ranks it to calculate the average rank and rank deviation from formulas above. Then, after the game is won or lost, the expected gain or lose of rank points should be added to all of the players' ranks.

8. Second generalization: teams with different number of players

Considering a game where one player wants to compete with two other in a team is something possible in the Widelands. In point 5. c) it was calculated that if the team contains two players with a different in ranks over 900 points, the overall average is exactly the same as the stronger player would be in a team with larger difference in points. It doesn't matter to the average if the difference is 1000 or even 10 000. Also the calculated rank is lowered by 120 from the stronger players.

So with given R_1 = 1800 and R_2 = 900 (R_2 - R_1 ≥ 900), the calculated average R_3 = 1800-120 = 1680

a) Rank value calculation

The player who plays alone can be treated almost as a player in a team with a player who is doing nothing. It is suggesting that the situation should be considered as a player with a lowest valid rank. Unfortunately that will affect players that are close to the lowest rank and wants to play alone versus more players.

My proposal here is to use invalid value in terms of global rank, but valid in terms of mathematic: R_0=1. Then the lowest possible rank agreed by the community (f.e. R=600) will be still affected, but the affected range will be much smaller than if we set R_0 = 600.

It would be possible to set R_0 to values below zero (f.e. R=-900), but all the equations and derivations based on assumptions that all the values were real and above 1.

b) Rank Deviation value

Since R_0 value is small in comparison with players' R value, the DR_0 doesn't affect much the overall average deviation.

$$RD_3 = \sqrt{\left(\frac{\beta^{BIG}}{\beta^{BIG} + \beta^{SMALL}}\right)^2 \cdot RD_1^2 + \left(\frac{\beta^{SMALL}}{\beta^{BIG} + \beta^{SMALL}}\right)^2 \cdot RD_0^2}$$

$$RD_3 = \sqrt{\left(\frac{\beta^{BIG}}{\beta^{BIG}}\right)^2 \cdot RD_1^2 + \left(\frac{\beta^{SMALL}}{\beta^{BIG}}\right)^2 \cdot RD_0^2}$$

$$RD_3 = \sqrt{(\approx 1)^2 \cdot RD_1^2 + (\approx 0)^2 \cdot RD_0^2}$$

$$RD_3 \approx RD_1$$

The difference is smaller as a player's rank increases. It can be set as the agreed initial RD value for player's rank.

9. Third generalization: two teams with different number of players in each and at least 2 players in each

And to finish two teams matches, we need to consider also the situation where two teams wants to play and they are containing different number of players, but there are at least 2 players in each of the teams. For example in the first team we have 2 players, while in the second will be 3 players. Calculating average for 2 players in the first team and for 3 players in the second will provide inconsistency of the ranks: two players will be counted as 2 players, but they are facing 3 players.

Natural solution for that situation is to fill all the teams with null-players (those with agreed "0" ranks) to make it the same size as the bigger team, then calculate ranks and compare it with Glicko-2 rank system.

So in case 2 vs 3 team, the 2-players team should be counted as 2 players plus 1 null-player, while 3-players team should be counted as it is.

For extreme case of 3 vs 5 team, the 3-players team should be counted as 3 players plus 2 null-players, while 5-players team should be counted as it is.

10. References

[1] https://en.wikipedia.org/wiki/Elo_rating_system

[2] https://en.wikipedia.org/wiki/Glicko_rating_system

[3] https://en.wikipedia.org/wiki/Natural_logarithm

[4] https://www.widelands.org/forum/topic/4583/?page=3#post-28850

[5] https://www.itl.nist.gov/div898/handbook/mpc/section5/mpc55.htm

[6]