Over a year ago the International Axe Throwing Federation (IATF) launched the Collins Rating System. This system is based on the Elo rating system widely used in chess, online gaming and other kinds of head-to-head competition. The system considers the level of stakes at hand as well as the value of one-on-one competition, and has been adapted by the IATF to accommodate the complexities of axe throwing. The Elo system was selected because it ranks players in head-to-head competition through a mathematical equation.

This system works well for the IATF. We value head-to-head competition, both in our sport and as sports fans, and see the value in a rating system that rates players according to skill and skill improvement. We operate in the spirit of fairness in competition; all venues operate differently, but still have an accurate system of rating for throwers across all IATF organizations.

Key Principles

- Head-to-head competition should be used to assess skills and determine ratings. (true since version 1.0)
- Beating lower rated players is less valuable than beating higher rated players. (true since version 1.0)
- Higher levels of competition should raise the stakes by allowing larger adjustments, if required to better reflect players' relative skill. ie. IATC has higher stakes than Regionals which in turn has higher stakes than league playoffs. (true since version 1.0)
- Throwing an 81 and <u>winning</u> the tiebreaker should never result in a downward adjustment (this became clear from version 1.0)
- Throwing more frequently should not disproportionately inflate ratings (true since version 1.0, however version 1.1 created a deviation from this principle, version 1.2 addresses the inflation from version 1.1)

Reviewing How it Works:

Each players' rating goes up or down at the end of a match based on the result of each round. Players have their rating compared to the rating of their opponent. If the higher rated player wins, as expected, adjustments are generally small. However, if the lower rated player wins, an adjustment in ratings moves both players' ratings - up for the lower rated player, and down for the higher rated player.

CRS 1.0 - The first version of CRS adjusted players rating based on winning or losing rounds within a match. These adjustments were affected by the ratings of their opponents as well as the expected result going into the match.

For example, in version 1.0 (example A): Alice and Bob are throwing a match. Alice is a higher rated player than Bob. When Alice wins the match, her rating goes up slightly and Bob's goes down slightly, by the same amount. However, if Bob (as the lower rated player) were to win the match, his rating would go up and Alice's rating would go down, again by the same amount.

Another example (example B): A highly rated player who ties a lower rated player at 81 and wins the tiebreaker. Based on their rating before the match, the lower rated player has exceeded their expectation by taking the match to a tie breaker, and needed an upward adjustment. Since in CRS 1.0 the adjustments were always equal, but opposite, the higher rated player's rating would be adjusted downward to match the lower rated player's increase after the match. The result of which was that the two players were now more closely rated, which was the desired result of the system.

Transition to Version 1.1 - After listening to feedback from the community and observing how the System worked in a live setting, it was clear that the rating calculation was not performing as desired in some match scenarios. It was determined that the cause was that ties at 27 were being treated in the same manner as other ties, which, since this is the maximum score, felt like a penalty for the higher rated player, as outlined in example B, described in the previous paragraph. The natural question was "How could I have done better? Why would my rating go down in that scenario?"

Version 1.1 - To address how ties at 27 were impacting rating adjustments, a modification to the calculation was made to treat these rounds as a win <u>in that round</u> for both players instead of a tie <u>in that round</u> for both players; the rationale being neither player could have done better, so this should count as a win. This meant that in matches with ties at 27, it was now possible for both players to be adjusted positively since both players threw as well as possible <u>in that round</u>. This effectively created new "bonus" rating points, adding them to the system. It also meant that the lower rated players' adjustments were greater than the higher rated players'.

This modification created a new issue wherein players who were tying at 27 more frequently, especially by throwing in multiple leagues, were receiving a disproportionate boost to their ratings since these "bonus" rating points threw off the original balance of the system. This effect was more pronounced for the lower rated players.

For example, in version 1.1 (example C): Alice and Bob are both rated 1600. Alice and Bob both throw 27s at the same percentage rate, however, Alice throws in one league and Bob throws in four leagues. Bob stands to gain four times the "bonus" rating points.

Following the release of Version 1.1, after listening to community feedback and observing actual match rating adjustments, it was clear that another modification would be needed to align the CRS to some key principles.

Version 1.2

A modification to the calculation was made that would treat ties at 27 as follows:

- The higher rated player's round is treated as their expected result in the round. This means that it is no longer possible to tie at 81 and win the tiebreaker and have the rating adjusted downward.
- The lower rated player's round is treated as a tie in the round (as with Version 1.0). It should be noted that a tie for the lower rated player is better than expectation. This means that throwing an 81 against a better player can result in a positive adjustment, even if the tie breaker is lost. However, it also avoids the inflated rating adjustments of Version 1.1.

For example, in version 1.2 (example D): Alice and Bob play a match. Alice is rated 1700 and Bob is rated 1500. Alice is expected to win. They both throw three rounds of 27 and tie overall at 81. Alice wins the tiebreaker.

The resulting rating adjustment for each of the CRS versions would be (see The Full Nerd Version below):

CRS Version	Alice	Bob
1.0	-3	+3
1.1	+5	+10
1.2	+1	+3

This particular case highlights the differences between the versions. CRS 1.2 produces adjustments that best align with the key principles behind the CRS.

Players should note that when version 1.2 is implemented there will be a shift in rankings. Multiple tests have been carried out to assess the accuracy of version 1.2 and our team has concluded that this latest version is the most balanced and fair while maintaining the key principles of head-to-head competition that we hold dear in our sport.

Version 1.2 will be launched on July 17, 2020. For those interested in a more detailed understanding of the rating calculation please see **The Full Nerd Version** below.

Remote Matches

We are excited to launch the IATF AxeScores app soon, which will include a remote challenge feature. The app will increase players ability to seek out higher levels of competition across all IATF players. Players who would normally not be able to geographically play against each other on a regular basis will now be able to throw heard-to-head matches via the app. These matches will count towards CRS ratings at half of the value of regular league play. This reflects the principles of competition outlined above wherein we value higher profile matches and tournament at higher rating multipliers ie. Playoffs have higher stakes than regular league play which in turn has higher stakes than remote matches. This also reflects the level of supervision during those matches. Remote matches will be regulated using tools within the app. Because there is less oversight and officiating, the matches' rating value will be less.

The Full Nerd Version - CRS 1.2

The procedure for calculating rating adjustments is as follows:

Let the initial rating of a player be p_i and the initial rating of the player's opponent be q_i.

Let the player's match expectation value be E. The expectation value is calculated as:

$$E = 2 / \left(1 + 10^{(q_i - p_i)/400}\right)$$

It follows that $p_i > q_i \Rightarrow 1 < E < 2$ and $p_i < q_i \Rightarrow 0 < E < 1$.

Each round in the match is evaluated to determine the result. A win is valued at 1, a loss at 0 and a tie at 0.5. In the case of a tie at 27, if $p_i > q_i$ the round is valued at half the expectation value ($E/2$).

The set of the player's round results is $\{g_1, g_2, ..., g_n\}$ where g_1 is the player's result in round 1 and g_n is the player's result in round n and n is the number of rounds.

Let the player's match result value be R. The match result value is calculated as follows:

$R = \frac{2}{n} \sum_{k=1}^{n} g_k$, where n is the number of rounds.

Observe $\because 0 < g_k < 1 \Rightarrow 0 < R < 2$.

Let the player's rating adjustment be A. The adjustment is calculated as follows:

$A = [f^* (R - E)]$, where $[\ ]$ denotes rounding to the nearest integer, and f^* is a swing factor[1].

Let the player's final rating be p_f. The final rating is calculated as follows[2]:

$$p_f = p_i + A$$

An example, for illustration:

Alice has a rating of 1700.
Bob has a rating of 1500.

Alice is the higher rated player, they are expected to win.

Alice's expectation value is $2 / \left(1 + 10^{(1500 - 1700) / 400}\right) \approx 1.52$.

Bob's expectation value is $2 / \left(1 + 10^{(1700 - 1500) / 400}\right) \approx 0.48$.

Alice's expectation value is higher than Bob's, since Alice is expected to win.

In round 1 Alice and Bob tie at 25.
In round 2 Alice wins 25-20.
In round 3 Alice wins 27-21.
Alice has won the match.

Alice's round results are: $\{0.5, 1, 1\}$.

Bob's round results are: $\{0.5, 0, 0\}$.

Alice's match result is: $\frac{2}{3} \times 2.5 \approx 1.67$.

Bob's match result is: $\frac{2}{3} \times 0.5 \approx 0.34$.

Alice's adjustment is: $10 \times (1.67 - 1.52) \approx\ +1$ (here 10 is the swing factor).

Bob's adjustment is: $10 \times (0.34 - 0.48) \approx\ -1$.

Alice's rating after the match is 1701.
Bob's rating after the match is 1499.

Since Alice won the match, and approximately by the expected amount, the adjustment is small.

Had the first round been a tie at 27 instead of 25, the results would follow:

Alice's round results are: $\{0.76, 1, 1\}$.
Bob's round results are: $\{0.5, 0, 0\}$.

Alice's match result is: $\frac{2}{3} \times 2.76 \approx 1.84$.
Bob's match result is: $\frac{2}{3} \times 0.5 \approx 0.34$.

Alice's adjustment is: $10 \times (1.84 - 1.52) \approx \ +3$.
Bob's adjustment is: $10 \times (0.34 - 0.48) \approx \ -1$.

Alice's rating after the match is 1703.
Bob's rating after the match is 1499.

Alice's performance is a little better and the adjustment is a little greater. The tie doesn't affect Bob differently though, a tie in this case is still better than their expectation.

Had Bob won rounds 2 & 3 (leaving the tie at 27), the results would follow:

Alice's round results are: $\{0.76, 0, 0\}$.
Bob's round results are: $\{0.5, 1, 1\}$.

Alice's match result is: $\frac{2}{3} \times 0.76 \approx 0.51$.
Bob's match result is: $\frac{2}{3} \times 2.5 \approx 1.67$.

Alice's adjustment is: $10 \times (0.51 - 1.52) \approx \ -10$.
Bob's adjustment is: $10 \times (1.67 - 0.48) \approx \ +12$.

Alice's rating after the match is 1690.
Bob's rating after the match is 1512.

Bob performed much better than expected and Alice performed worse. Their rating adjustments reflect the upset.

It should be noted that the first 28 matches a player completes in their career have an additional consideration, specifically, opponents' ratings in these 28 matches are not adjusted unless the match is also one of the opponents' first 28 matches. This prevents the uncertainty of new players' ratings from affecting their opponents' ratings in their first career season.

[1] The swing factor f^* is a parameter chosen to reflect the stakes of the type of match. The swing factor acts as a bounding factor on how large a rating adjustment can be for any given match of that type. Observe $|A| < 2f^*$. The swing factor generally increases with the level of competition or tournament profile. Regular league play has a swing factor of 10, larger tournaments can have swing factors of up to 20.

[2] The calculation can be expressed in terms of source input values as:

$$p_f = p_i + \left[2f^* \left(\frac{1}{n} \sum_{k=1}^{n} g_k - \left(1 + 10^{(q_i - p_i)/400} \right)^{-1} \right) \right]$$

Definitions

where:

p_f is the player's final rating,

p_i is the player's initial rating,

q_i is the opponent's initial rating,

f^* is a swing factor,

g_k is the player's result in round k,

n is the number of rounds,

and $[\ ]$ denotes rounding to the nearest integer.