For office use only	Team Control Number	For office use only
T1 ________________	**2021TH_ _**	F1 ________________
T2 ________________		F2 ________________
T3 ________________		F3 ________________
T4 ________________		F4 ________________

2021

The International Mathematical Modeling Challenge (IM^{2}C) Summary Sheet

(Your team's summary should be included as the first page of your electronic submission.)

In this entire world, there are many types of sports and determination of their rankings. Some of them don't work efficiently especially the tournament system in which players meet only a few competitors. Elo system, one of the widely used systems, is a good rating protocol for long competition with massive matchmaking. A performance evaluation may not just stop at the win-lose result of the match when it comes to tournament-based matches. In this report, we have illustrated our development of the model to find the Greatest Tennis Player in 2018 (Women's Tennis) and GOAT for Men's Tennis using score difference of matches (match toughness) with accumulated rating system over tournament each year.

For the model to find the Greatest Women Tennis Player in 2018, firstly we gather full data of match occurred during those big 4 tournaments. Then we find a point estimator of probability for a player to win each game, after that we calculate the probability for a player to win a match, given that we know the probability that a player will win a game independently. Then we define rating point as a value that can be compared to see how likely a player will win a match against another player. We believed that player performance distributed normally, so we use a normal distribution curve to help in defining the rating point. Then we assign a rating point for every player in every tournament, accumulate over all four tournaments in 2018, and conclude the result to find the greatest women's tennis player in 2018.

For the result of our model, we obtained that Angelique Kerber is the Greatest Tennis Player in 2018, she won 1 championship and pass through the semifinal and quarterfinal on the other two tournaments. Even though some of the results seem surprising at first, we found that the ones who got high ratings can beat lots of pro players, even in beginning round.

There are lots of benefits on this model, it prevents good players from losing point at the very beginning because of the tough match, also, prevent newbies from gaining rating by luck. The performance of each match really does matter. However, some limitations also occur, but errors normally about the score of losers so it shouldn't have any severe problem on Greatest Player.

Furthermore, we improve this model to find GOAT for Men's Tennis, representative of Tournament matchmaking sports. All Men's Tennis results from 2000-2016 in Grand Slam Tournament are chosen as representative for calculation. All of the ratings are assigned to each individual player in the tournament. Then, we assume that the performance accumulated in player decrease over time. So, the only way that a player can become GOAT is that they must have a high rating and also consistent over a period of time.

For the result of the GOAT model, we got Federer Roger as the Greatest Of All Time of Men's Tennis. Other explanations and discussion of this result can be seen further in the paper.

Lastly, we discussed a potential way to improve our model to be applicable to team sport by combining individual impact and team performance to get the individual performance and rated according to the model.

Contents

1 Introduction

Sports tournament competition has its unique character: each player will not face every other participant yet only some of them. This creates the problem of comparing the players that have never faced each other, or to combine result of multiple elimination tournaments. People usually rank the player by focusing on the Win-Lose results of each match that occur along with those tournaments which are ungeneralized. Hence, many mathematics modelers have designed a rating system to rank the player each year even from different tournaments.

The Elo rating system (Elo) is one of the most widely used rating systems. However, it is only viable in an intense, dense competition with sufficient amount of matches between players in a long season of competition, which occur a match almost all the time through the entire rival.

In this report, we will develop a mathematical model to find the greatest player based on statistics and scores in each match rather than just Win-Lose results for the Women's Tennis Tournament. Additionally, we also develop a model to find the greatest player of all time (G.O.A.T. player) for a Men's Tennis Tournament. We also provide discussion toward using this model in team sports.

1.1 Restatement of Problem

The main task is splited into 3 smaller tasks, which are

1. According to the Grand slam tournament results, find the best women tennis player in 2018.
2. Find the G.O.A.T. player of a particular individual sport by combining data of all time, which we chose to work with Men's Tennis.
3. Determine the adjustments necessary to the model to make the model applicable to a team sport

1.2 General Assumption

In order for our mathematical model to work correctly, it needs some prior assumptions including:

1. In each tournament, a player's performance is at the same level throughout the whole tournament.
   Justification: In a tournament, the time frame is small enough that the performance of each person will not vary by a lot in different matches.
2. The closer the result of a match is, the more chances that both player are at the same skill level.
   Justification: If the loser have scored closer to the winner, it means that both players have about the same potential to score points against each other then the skill point can be considered as closer.

3. Winning Probability of one player to another specific player of each game is constant and independent, and it is assumed to be the same as the probability of winning a tiebreaker.
   Justification: Though there are factors such as tiredness and comeback potential, the winning probability is largely more dependent with the actual skill level of the player. Other factors are negligible especially on a broad scale. Also, on average, the game length of tiebreaker and the normal game are close enough that it can be considered the same.

4. Champion of each tournament has the same level of skill (rating point)
   Justification: All of the grand slam tournament have very similar level of competitiveness, therefore, the winner of each tournaments is considered to have the same level of performance / rating point.

5. In each tournament, players performance disitrbuted normally
   Justification: From the central limit theorem, no matter what distribution it originally is, the mean of samples will tends to distribute normally.

6. Every player plays at their fullest potential.
   Justification: We can assume that all player want the best position they could possibly get, so by this assumption, players should play at their fullest potential to achieve their goal.

7. There is a correlation between player skill and probability that a player will win. A better player must have a higher chance of winning a game against the same opponent.
   Justification: In a long run, if a player has more skill the opponent, then it is more likely that player will win more games against the opponent.

1.3 Crucial factors

The crucial factors in finding the greatest tennis player from 4 grand slam tournaments of women's tennis 2018 are the following:

1. Games won by the player in every set in every match

2. Match result

We use the number of games won and lose by a player in specific match to determine winning and losing ratio, which helps in estimating probability of a player winning a game, which will later be used to determine probability of a player winning a set and a match. The probability of winning a match is assumed to have a close relationship with player's skill as mentioned in assumption 7

Match result is used in addition to the game data to determine champion and to recursively define rating point for every player based on the winner's rating.

These factors will help us find the performance of the loser comparing to the winner to help calculating the rating of each player.

The crucial factors in finding the GOAT of any sports we believed as following:

1. The forms of the competition of each sport
2. Performance of the player
3. Collective performance of the player in the previous year

These factors will help us find the cumulative rating of the player from the different types of sports in the precedent years.

2 Model to Find the Greatest Tennis Player in 2018

2.1 Solutions

By focusing on Women's Tennis in 2018, we further gathered full data of 4 Grand Slam tournaments to determine the real greatest player in that year. Since the problem has been addressed prior that focusing on the only result of each match does not efficient, our idea, instead, works on the skill levels of those two competitors in the match and how close is the match (match toughness). If the match is closest, it means both of the competitors are on par.

Mainly we have two main steps to determine the rating of each competitors. First is to determine the match toughness by composing of the Game-Win Probability Estimator determination and the Expected Match-Win Probability calculation. Another step is Rating Calculation by probability of winning the match.

2.1.1 Determining Game-Win Probability

First, from assumption 3 that a player perform independently on each game, and each game has an equal independent win probability, we defined match toughness as the mentioned probability.

For the sake of calculation, we defined $\hat{p}_{game}$ as the independent probability that a player will win their specific opponent a game, which is calculate as the average of each point estimator of each set they played in the match. Each set resembles binomial distribution with parameter p as the winning probability since the probability of winning each game is constant and independent. The last game of each set that ends with score 6-0, 6-1, 6-2, 6-3, or 6-4 must be won by the set winner. Therefore, from Maximum Likelihood Estimation, we obtain the point estimator $\hat{p} = \dfrac{5}{5+n}$ for sets that ends with score 6-n (for n ranging from 0 to 4). In addition, for sets that ends with score 7-5 and 7-6, Maximum Likelihood Estimation method yeilds $\hat{p} = 0.5$

$$\hat{p}_{game} = \frac{\Sigma \hat{p}_i}{\text{Total Set Played}}$$

where $\hat{p}_i$ is the point estimator of set i.

2.1.2 Calculating Match-Win Probability

From the Game-Win Probability, we calculate the Set-Win Probability by spliting into cases. We consider situation where a set is won by scoring 6-0, 6-1, 6-2, 6-3, or 6-4, the situation that a set is won with score 7-5, and the situation that the set is extended to tiebreaker and ended with score 7-6 independently. Combining all cases, we get that the probability of winning a set is

$$\hat{p}_{set} = \sum_{n=0}^{4} \binom{5+n}{n} \hat{p}_{game}^{6}(1-\hat{p}_{game})^{n} + \binom{10}{5} \hat{p}_{game}^{7}(1-\hat{p}_{game})^{5} + 2\binom{10}{5} \hat{p}_{game}^{7}(1-\hat{p}_{game})^{6} \quad (1)$$

Then we apply the same method to calculation the probability that a player will win a match given the probability that they will win a set.

$$\hat{p}_{match} = \hat{p}_{set}^2 + 2\hat{p}_{set}^2(1-\hat{p}_{set}) \tag{2}$$

From equation 1 and 2, we can calculate the probability that a player will win a match against their opponent, given the independent probability that they will win a game against their particular opponent.

2.1.3 Defining Rating Point

After the Expected Match-Win Probability of one match has been defined, we calculate the rating of each player in each tournament by considering each tournament independently.

Firsly, from assumption 4, we set the rating of the champion to be 1000, then we find rating point of other players recursively.

From assumption 1, we argue that the rating point of a player should not depends on the round that the player got in, but should only depends on the match that they were eliminated from the tournament, which is the only match that they lose. With this assumption, we set rating point for every player based on their performance in their last match of the tournament. The rating point of a player will reflect the relative performance to the champion. Formally, a player will have rating equal to x if they have win probability equal to $P(X \leq x)$ where X is a normal distribution (Assumption 5) which parameter μ equals to the opponent's rating point and $\sigma = 1$

From the defined rating point, we showed that the rating point of a player will change according to this formula.

from the cumulative density function of normal distribution,

$$P(X \leq x) = \frac{1}{2} \cdot [1 + \text{erf}(\frac{x-\mu}{2\sigma})]$$

where erf denotes gauss error function. Then, we derived

$$\delta_{rating} = 2\sigma \cdot \text{erf}^{-1}(2\hat{p}_{match} - 1) \tag{3}$$

Furthermore, to prevent the case where players skills differ by so much and the rating change is too strong, we set the maximum rating change to be 140 and minimum rating change to be -140 inclusive. Therefore, the different of rating of the winner and loser in every match is not greater than 140.

$$R_{loser} = \begin{cases} R_{winner} + 140 & \text{if } \delta_{rating} > 140 \\ R_{winner} - 140 & \text{if } \delta_{rating} < -140 \\ R_{winner} + \delta_{rating} & \text{otherwise} \end{cases}, \tag{4}$$

The justification for number 140 is that there are 128 contestants, with maximum of $log_2 128 = 7$ match per player and $\frac{1000}{7}$ is just a little bit more than 140. Therefore, the least rating possible for every contestant in a tournament is equal to 20, which is $1000 - 7 \cdot 140$.

Rating of each player will be defined for each tournament and then accumulated for the whole year of 2018. The best player of 2018 is the player with the highest total rating.

2.1.4 Assignmment of rating point

Rating point will firstly be given to the tournament champion. The champion of each tournament will get equal rating at 1000. Then rating point of other player will be assign only when the player get eliminated from a round. The assignment of rating point will follow the definition stated in 2.1.3. This being said, each player will get a rating only one time per tournament relative to the winner of the match that the player got eliminated from the tournaments.

To implement this, we visualize tournament system as a tree and use topological ordering "Topological Sort and Graph Traversals" to sort the player and assign each player a rating in the appropiate time.

2.1.5 Result

According to our model, we simulate our process using computer program shown in Appendix and dataset provided by Sackmann. We found that Angelique Kerber was the greatest player considering all 4 grand slam tournaments. She scores 3958.71 rating points in total.

2.2 Analysis of Our Model

Player	Rating
Angelique Kerber	3958.7124
Naomi Osaka	3956.5190
Kiki Bertens	3947.9802
Simona Halep	3942.2001
Carla Suarez Navarro	3938.9465
Alison Van Uytvanck	3935.3274
Madison Keys	3933.5511
Caroline Garcia	3930.0393
Anett Kontaveit	3928.6231
Caroline Wozniacki	3927.4804

Table 1: The table shows the top 10 player ranked by the rating system developed in 2.1.3

From all the tournaments in 2018, Angelique Kerber won 1 championship of Wimbledon, 1 semifinal of Austrailian Open, and 1 quarterfinal of French open. She won against Naomi Osaka with score 6-2, 6-4 and Kiki Bertens with score 7-6, 7-6. Naomi Osaka and Kiki Bertens was ranked second and third respectively from our model, so Angelique Kerber, who won against both of them unsurprisingly ranked first, as seen from table 1. Even though she lost to Simona Halep twice, those two matches were a close match with $\hat{p}_{match} = 0.5971$ and 0.8880.

However, Kiki Bertens who won no championship in 2018 was ranked third according to our model. We argue that this is not a mistake because Kiki Bertens lose to Angelique Kerber once with a really close score of 7-6, 7-6. This match connotes that Kiki Bertens performance in the tournament were close to Angelique, the winner, and thus receive high rating. She also lost to Caroline Wozniacki early in round 3 in the Australian open where Caroline Wozniacki later won a championship.

2.3 Model Strengths

1. This model prevent the situation where good player have too low rating when they loe to a better player in early rounds.
2. Likewise, this model prevent bad players from having too high rating by luckily reach deeper rounds but then lose to the other player with a huge different.
3. Rating determination of players based mainly on player's performance in each of the matches.

2.4 Model Limitations

1. For the situation like 7-6, 6-7, 7-6, $\hat{p}_{game} = \hat{p}_{match} = 0.5$. Therefore, the rating of the both player will be equivalent. This means that this model will sometimes (very unlikely) be undecisive in finding the best player.
2. The model will disregard deuces. For examples, sets with score 7-5, 7-6, 8-6, 9-7, are equivalent.
3. At some specific score of the match, the model will count the close performance as same rating, such as loser in match 6-0, 6-0 and 6-1, 6-1 got rating equal to -140 for both.
4. Player's performance in matches that they win will not directly affect their rating, but only the matches that they lose will.

3 Model to find the GOAT of any sports

3.1 Solutions

There are many types of sport in the world, even considering only individual sports. Each type of sports couldn't adopt the same model to determine the greatest player easily. We divide all types of sports into 3 categories based on how the tournament is conducted.

1. Sudden Death Tournament/Elimination Tournament
2. Round Robin
3. One Player Sports

For our idea, we would like to improve from our prior model on Women's Tennis. We choose type 1 ”Elimination Tournament” with chosen sport ”Men's Tennis”. We rather focus on the idea to develop the model of decaying accumulated rating over time to determine the GOAT. of Men's Tennis, so, we integrate the Women's Greatest Tennis Player model as part of a model to find GOAT. We concern that if the rating was accumulated over a year without decreasing, the player who played in the league for a longer duration will advantageous. On other hand, if the rating was determined for only a very short period, a newbie will have a chance to win pro players by luck, resulting in the wrong accumulated rating system.

3.1.1 Model to Find the GOAT for Men's Tennis

Begins with the same model as the Women's Tennis Tournament Competition, we assign the rating to each player in every year they have participated. As we mentioned the method of rating accumulation over the year to be developed, we apply the decaying constant to rating on each year. This method illustrates the reduction of performance of a player, players who have the consistency rating, high peak performance, and play long enough are able to be the GOAT. The model of rating decaying will be as follow:

$$\lambda(\lambda(...(\lambda R_1 + R_2) + R_3) + ...) + R_n = \sum_{i=1}^{n} \lambda^{n-i} R_i \tag{5}$$

Where $0 < \lambda < 1$ is the rate decaying constant, and R_i is the rating of the i^{th} year the player played until the n^{th} retirement year. The model of finding peak performance will be as follow:

$$max(\sum_{i=1}^{j} \lambda^{j-i} R_i) \qquad ; \forall j \in 1,2,3,...,n \tag{6}$$

3.1.2 Result

From the dataset prepared by Goblet, we calculate using the computer program provided in the Appendix, and we got that the GOAT for Men's Tennis (Considering only on the period 2000 - 2016) is Roger Federer.

We hereby annotate that Roger Federer might not be the GOAT of Men's Tennis. but from our limited dataset, we can only test the model from 2000 to 2016, in which Roger Federer is the best player. However, after obtaining data of All-time Men's Tennis, it's possible to use the very same model and find the GOAT for Men's Tennis.

3.2 Analysis of Our Model

Player	Cumulative Rating
Federer R.	7930.7440
Mirnyi M.	7907.3510
Hrbaty D.	7906.3070
Djokovic N.	7903.0164
Verdasco F.	7897.5910

Table 2: The table shows the highest cumulative rating between 2000 - 2016 of the top 5 players.

From the table 2, Federer R. is the GOAT that our model chose with the cumulative rating of 7939.2381.

From the table 3, Federer R. won the most championships (17) from 2000 to 2016 while the 2^{nd} is Djokovic N. which won 12 championships. This is the main reason why Federer R. is the GOAT. However, rank 2 and 3: Mirnyi M. and Hrbaty D. are the two who never won any championships is a little odd. For Mirnyi M., he has competed 62 matches total, 2 of them won against Federer R. and Hrbaty D., and most of the matches he played are close games. For Hrbaty D., he has competed 79 matches in total, although he did not win against any top 5 players, most of his winning matches, they were landslide victories.

Note that we use the data from 2000 to 2016 only because there is no data before 2000. Actually, the input should be the data collected since the first Grand Slam Tournament. Thus, we are not sure that who is the actual GOAT but Federer R. is the GOAT if we consider only years 2000 to 2016.

3.2.1 Model Strengths

1. This model reward player who not only high rating players but also long consistency rating.
2. Do not let newbie take the ranking so easily by luck.
3. Pro players have a chance to drop from the tier list if they cannot maintain their performance.

Year	Australian Open	French Open	Wimbledon	US Open
2000	Agassi A.	Kuerten G.	Sampras P.	Safin M.
2001	Agassi A.	Kuerten G.	Ivanisevic G.	Hewitt L.
2002	Johansson T.	Costa A.	Hewitt L.	Sampras P.
2003	Agassi A.	Ferrero J.	Federer R.	Roddick A.
2004	Federer R.	Gaudio G.	Federer R.	Federer R.
2005	Safin M.	Nadal R.	Federer R.	Federer R.
2006	Federer R.	Nadal R.	Federer R.	Federer R.
2007	Federer R.	Nadal R.	Federer R.	Federer R.
2008	Djokovic N.	Nadal R.	Nadal R.	Federer R.
2009	Nadal R.	Federer R.	Federer R.	Martin J.
2010	Federer R.	Nadal R.	Nadal R.	Nadal R.
2011	Djokovic N.	Nadal R.	Djokovic N.	Djokovic N.
2012	Djokovic N.	Nadal R.	Federer R.	Murray A.
2013	Djokovic N.	Nadal R.	Murray A.	Nadal R.
2014	Wawrinka S.	Nadal R.	Djokovic N.	Cilic M.
2015	Djokovic N.	Wawrinka S.	Djokovic N.	Djokovic N.
2016	Djokovic N.	Djokovic N.	Murray A.	Wawrinka S.

Table 3: The table by "List of Grand Slam women's singles champions" lists the Grand Slam men's tennis singles champions from 2000 to 2016.

3.2.2 Model Limitations

1. Some players in the top tier may come from landslide victories rather than winning a great match with a tough competitor. However, it does not mean that they are not good players.

3.3 Discussion

There are many types of sports out there in the world with different sets of rules and tournaments organization. We categorized them into 3 sections based on the tournament organization method.

Since we have developed the model for tennis tournament sports competition, we will discuss the other type of sports competition: Round Robin, and One Player Sports, about how it would differ from the Elimination Tournament.

One idea that still works with other types is the decaying of performance over time to identify the GOAT as from the equation (5), however, the rating system needs to be specifically changed for each type of competition.

3.3.1 Mathematical Model for Round Robin

In sports that use round-robin in matchmaking such as chess, go, draughts or even tennis (in some other competition). Each individual will have a chance to compete with more competitors than in Tournament. As mentioned in the introduction, we can use the Elo system since it works well under denser competitions. Moreover, the Elo system can also be integrated with score differences to make it more precise.

3.3.2 Mathematical Model for One-Player Sports

One-player sports are sports that the presence of an opponent that contribute just a little or no effects to a player's performance, such as running, marathon, shooting. This is easier than the other two forms of sports competition because these are score-based/record-based sports. The score which comes from the performance of only a single player and not from both competitors can be ranked without bias about the toughness of the match. Directly, the player who gets a higher score should have a higher rating.

4 Team Sports Discussion

As for team sport, the types of matching, such as elimination tournament, round-robin, or one-player sports, are identical to individual sports. So, as we had discussed in 3.3, there are ranking systems that can be adapt for team sports. Our proposed model can be used to determine the rating that each team should get in a tournament. However, finding the GOAT requires additional information. The team performance is dependent on each member, and each member of the team has a different impact on team performance. This requires statistics of action in a match of each individual to determine the impact. So, the necessary adjustments to the model are to incorporate personal team impact rating, which would come from combining personal statistics that determine helpfulness to the team. Each individual statistic will be weighted according to its importance during the game and combined to get the personal team impact rating. This rating, when combined with the team rating, will result in the actual rating that each player should get. After that, we can go back to our previous model and determine the GOAT by combining rating points with the decaying function. For example, if we consider basketball, even though a team can only win by having a higher score than the opponent, there are many parts of the team that must function together in order to achieve maximum performance. So, it's not enough to only calculate field goal for every player in a team, but many other aspects including block, steal, rebound or even turnover are needed to calculate the impact each player contribute to a team. With the team impact model of each player in a team and a model to give a rating to each team on a tournament, it will be possible to determine the greatest player for any team sports.

5 Letter to the Director of Top Sport

Dear the Director of Top Sport,

As per your request to find the yearly greatest player and the greatest player of all time, we came up with a mathematical model that helps in determination. Firstly, we argue that just only win-lose results and round reached, which is being used nowadays, is not enough to determine the real performance of sports players. Instead, we came up with the idea of focusing more on the game result. Specifically, close victory and landslide victory also contribute to the performance of the player.

We estimate the probability of winning the game by using the number of games won and total games played. Then, we calculate the probability of winning the match. We believe that the probability of winning the match of the loser is essential to the determination of rating and since the performance of players is most likely distributed normally, we adjust the rating of the loser to a value according to the loser's probability of winning by using normal distribution curve. Lastly, to prevent over-gaining and over-losing points, we set the maximum difference of rating between the loser and winner of each match.

However, Increasing the rating of the player from the beginning of a tournament is not a good idea since the player that lost to the tournament champion in the first round might be better than the player that lost to the champion in the final round. To solve this, we calculate the rating reversely from the champion by setting the rating of the champion of each tournament to 1000, then compare the rating by the method stated above. Furthermore, The person who did not attend or walkaway or retired in the tournament will not be rated for that specific tournament.

After finding the greatest player, we determine the GOAT of Single Men's Tennis, by using the same rating system. We suggested that each player's rating should be accumulated over every tournament they played. However, the rating in the past was supposed to have less impact than the latest year. We think that there is some value of rating decay. Formally, at every end of the year, the rating of every player will get multiplied by a factor of λ. We believed that performance in the past plays a smaller role than the present performance but yet not negligible. For instance, a player that holds runner-up for years could be considered better than a 1 championship. After stacking rating over years, the rating value represents combined performance at that period.

The main advantage of this model is it considers not only the peak performance of players but also the consistency. From our model, even though the cumulative rating was reduced by some proportion over time, it will reflect if the player was good enough to maintain their performance over a while, which means that our result, GOAT of Men's Tennis, shows the best-skilled player that can also consistently hold their position.

Which come to our result, the greatest women tennis player that our model has selected is Angelique Kerber, and the men's GOAT tennis player is Roger Federer.

We hope that our model will please and fulfill your desire of finding the greatest of each year, and the GOAT. We are looking forward to cooperating with you further.

Yours faithfully,
Team Members.

6 Appendix

```
const { printSortedDict, unique } = require('../../utils')
const getTournamentsData = require('./getData')
const {nCr, erfinv } = require('../../utils')

const championScore = (match) => 1000

const getRatingTourney = (tourney) => {
  let deg = {}
  let graph = {}
  let rating = {}

  for (let row of tourney) {
    graph[row.winner] = []
    graph[row.loser] = []
    deg[row.winner] = 0
    deg[row.loser] = 0
  }

  for (let row of tourney) {
    graph[row.winner].push(row);
    deg[row.loser] += 1
  }

  const illegalPlayer = []

  const DFS = (player) => {
    for (let row of graph[player]) {
      if (row.comment != null) {
        illegalPlayer.push(row.loser)
      }
      rating[row.loser] = adjustRating(row, rating[row.winner]);
      deg[row.loser] -= 1;
      if (deg[row.loser] == 0) DFS(row.loser);
    }
  }

  for (let player in deg) {
    if (deg[player] == 0) {
      rating[player] = championScore(tourney.length);
      DFS(player);
      break;
    }
  }

  for (let player of illegalPlayer) {
    rating[player] = 0;
  }

  return rating
```

```
}

const adjustRating = (row, winnerRating) => {
  let sigma = 1;
  let expectedWinRatio = getExpectedWinRatio(row)
  let deltaRating = 2 * sigma * erfinv(2*expectedWinRatio - 1);
  deltaRating = Math.max(Math.min(deltaRating, 140), -140)
  return winnerRating + deltaRating
}

const GameToSet = (p) => {
  let ans = 0;
  for (let i = 0; i <= 4; i++) {
    ans += nCr(5+i, i) * Math.pow(p, 6) * Math.pow(1-p, i);
  }
  ans += nCr(10, 5) * Math.pow(p, 7) * Math.pow(1-p, 5)
  ans += nCr(10, 5) * 2 * Math.pow(p, 7) * Math.pow(1-p, 6)
  return ans
}

const SetToMatch = (p, bestOf) => {
  if (bestOf == 3) return p*p + 2*p*p*(1-p)
  if (bestOf == 5) return p*p*p + 3*p*p*p*(1-p) + 6*p*p*p*(1-p)*(1-p)
}

const getExpectedWinRatio = (row) => {
  let p = 0;
  let totalset = row.wonBy[0]+row.wonBy[1]
  for (let set = 1; set <= totalset; set++) {
    let win = row.winSets[set]
    let lose = row.loseSets[set]
    if (win == 6 && lose < 5) {
      p += lose/(5+lose)
    }
    else if (lose == 6 && win < 5) {
      p += 5/(5+win)
    }
    else {
      p += 0.5
    }
  }
  p /= totalset;

  if (row.comment != null) p=0.5

  return SetToMatch(GameToSet(p), row.bestOf)
}

const splitTourney = async () => {
  const data = await getTournamentsData()
  let tournament = {}
```

```
  for (let i in data) {
    if (!tournament[`${data[i].year}-${data[i].atp}`]) tournament[`${data[i].year}-${data[i].atp}`] = []
    tournament[`${data[i].year}-${data[i].atp}`].push(data[i])
  }
  let tourneys = Object.keys(tournament).map((key) => {
    return tournament[key];
  })
  tourneys.sort((first, second) => (second[0] - first[0]));
  return tourneys
}

const ratingDecay = async () => {
  const lambda = 0.5

  const tourneys = await splitTourney()

  let nowScore = {}
  let mxScore = {}
  let i = 0
  for (let tourney of tourneys) {
    let rating = getRatingTourney(tourney)
    for (let player in rating) {
      if (nowScore[player] == undefined) nowScore[player] = 0
      nowScore[player] += rating[player]
    }
    if (i % 4 == 3) {
      for (let player in nowScore) {
        if (mxScore[player] == undefined) mxScore[player] = nowScore[player]
        mxScore[player] = Math.max(mxScore[player], nowScore[player])
      }
      for (let player in nowScore) {
        nowScore[player] *= lambda
      }
    }
    i += 1
  }

  for (let player in mxScore) {
    if (isNaN(mxScore[player])) mxScore[player] = -1
  }
  await printSortedDict(mxScore, (k)=>(k), 6)
}
```

7 Reference and Attribution

Elo, Arpad E. *The Rating of Chessplayers, Past and Present*. New York: Arco Pub., 1978. ISBN: 0668047216 9780668047210. URL: http://www.amazon.com/Rating-Chess-Players-Past-Present/dp/0668047216.

Goblet, Jordan. *ATP Men's Tour*. URL: https://www.kaggle.com/jordangoblet/atp-tour-20002016.

List of Grand Slam women's singles champions. Feb. 2021. URL: https://en.wikipedia.org/wiki/List_of_Grand_Slam_women%27s_singles_champions.

Sackmann, Jeff. *WTA Tennis Rankings, Results, and Stats*. URL: https://github.com/JeffSackmann/tennis_wta/blob/master/wta_matches_2018.csv.

"Topological Sort and Graph Traversals". In: *Advanced Graph Theory and Combinatorics* (2016), pp. 69–83. DOI: 10.1002/9781119008989.ch4.