How the Rating System Works
乒乓球积分系统的工作原理

The rating system uses standard concepts and techniques from probability and statistics. This makes the rating system quite different from almost all other rating and ranking systems. The following is a nontechnical explanation of how the rating system works.

本积分系统使用概率论和统计学的标准概念与方法，这使得它与其他大多数积分和排名系统有着明显的不同。下面将对这个积分系统的工作原理作一非技术性的阐述。

Basic Concepts

We presume that each player has a *playing strength*, i.e., a number that quantifies how strong the player is. The playing strength of a player does not change during a single event, but may change over time, as the player gets better or worse. (An *event* is a collection of matches, e.g., from a tournament, that an event director submits to Ratings Central as a group.)

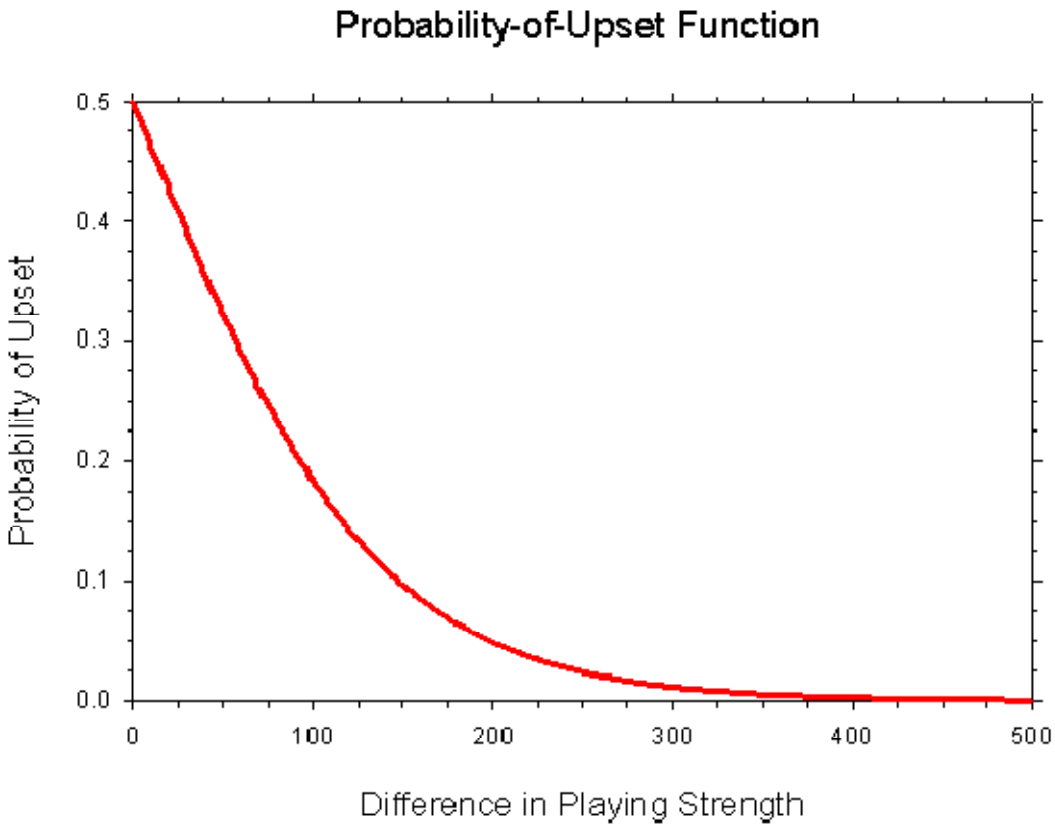
Even if we knew the playing strengths of two players, we would not know for certain which player would win, since a weaker player will sometimes beat a stronger player. A match is an *upset* if the player with the lower playing strength wins. We presume that the probability that a match will be an upset is determined solely by the difference in playing strengths of the two players. The larger the difference in the playing strengths, the more likely it is that the stronger player will win. The *probability-of-upset function* quantifies this.

基本概念

假设每个选手都有一个水平强度，即可以用来量化他/她水平的一个数值。一名选手的水平强度可能会随时间变化，比如选手的水平提高或者降低了，但是在单次赛事中并不发生变化。(一次赛事是指一次赛事中所有比赛的集合，赛事主管将所有比赛结果视为一个集合提交给计分中心。)

即使我们事先知道两名选手的水平强度，我们也不能断定谁将获胜，因为较弱的选手有时候也可能击败较强的选手。如果较弱的选手击败较强的选手，这场比赛就称为爆冷。我们假设一场比赛出现爆冷的概率仅仅由两名选手的水平强度之差别所决定，如果两者水平强度差别越大的话，则水平强度较高者就越有可能获得胜利。这可以用下图中的爆冷概率函数来加以定量描述。

图（1）Figure 1



There are two kinds of probability. There is the probability that one player will defeat another. (This probability is determined by the players' playing strengths.) There is also the probability that a player's playing strength is a certain value (e.g., 1106). The first probability is a property of the players, while the second probability is a property of the rating system.

这里有两种不同概念的概率。一种是上面刚刚提到的选手之间胜负的概率（这种概率与两名选手水平强度有关）；另一种是指一名选手的水平强度可以用一个特定数值（比如 1106）来衡量的概率。前一种是选手本身能力的特性参数，后一种则是本积分系统本身的一个特性参数。

Laws

The rating system does not know the playing strengths of players. It only sees match results. The rating system keeps track of what it knows about each player by constructing a law to describe the player's playing strength.

A *law* is a probability distribution (more or less). The rating system assigns a law to each player. The player's law describes the rating system's knowledge of the playing strength of the player. This knowledge is derived from all the match results. The player's law changes with every match the rating system processes (because the rating system's knowledge of the playing strength of the player changes

with every match). From the law, we may determine the probability that the player's playing strength is a certain value (e.g., 1106).

规律

积分系统本身事先并不知道一名选手的水平强度，它只参考比赛结果。积分系统通过记录它所知道的该选手所有比赛成绩来构建合理的规律性函数从而描述一名选手的水平强度。

一个规律性函数或多或少就是一个概率分布。积分系统给每一名选手选定一个规律性的特征函数，这个函数描述了该积分系统对这名选手水平强度的认识，该认识则是来源于该选手在系统中的所有比赛结果。积分系统处理选手的比赛结果后，该选手的特征函数会发生变化。（因为积分系统对一名选手水平强度的了解认识会随着每一场比赛发生变化）基于这个变化的特征函数，我们可以确定一名选手水平强度是某一个特定数值（比如 1106）的概率。

The *mean* of a law is the location of the center of the law (more or less). The mean of a player's law is the rating system's best estimate of the player's playing strength (because it is the center of the rating system's knowledge of the player's playing strength). The mean of a player's law is the *rating* that the rating system outputs for the player.

The *standard deviation* measures the spread (width) of a law. The greater the standard deviation of a player's law, the less certain the rating system is of the player's playing strength. The probability that a player's playing strength is within 1 standard deviation of the mean of the player's law is 68% (more or less). The probability that it is within 2 standard deviations is 95% (more or less). The probability that it is within 3 standard deviations is 99.7% (more or less).

If the meaning of a sentence like "The probability that a player's playing strength is within 2 standard deviations is 95%." isn't clear, here is another way of saying the same thing: There is a 95% probability that the player's playing strength is between the mean minus twice the standard deviation and the mean plus twice the standard deviation. For example, if the mean is 1106 and the standard deviation is 42, then:

- There is a 68% probability that the player's playing strength is between 1064 and 1148.
- There is a 95% probability that the player's playing strength is between 1022 and 1190.
- There is a 99.7% probability that the player's playing strength is between 980 and 1232.

特征函数的中值基本上是该特征函数的中间点，是积分系统对选手水平强度的最佳估计。（因为它也是积分系统对选手水平强度认知的中间点。）选手的特征函数的中值就是积分系统给这名选手定的积分。

标准偏差度量特征函数的分布宽度。选手特征函数的标准偏差越大，积分系统对选手的水平强度的衡量越不确定。（通常可以将选手的特征函数视为正态分布）选手的真实水平强度在其特征函数中值左右一个标准偏差以内的可能性是 68%，在中值左右两个标准偏差以内的可能性是 95%，在中值左右三个标准偏差以内的可能性是 99.7%。

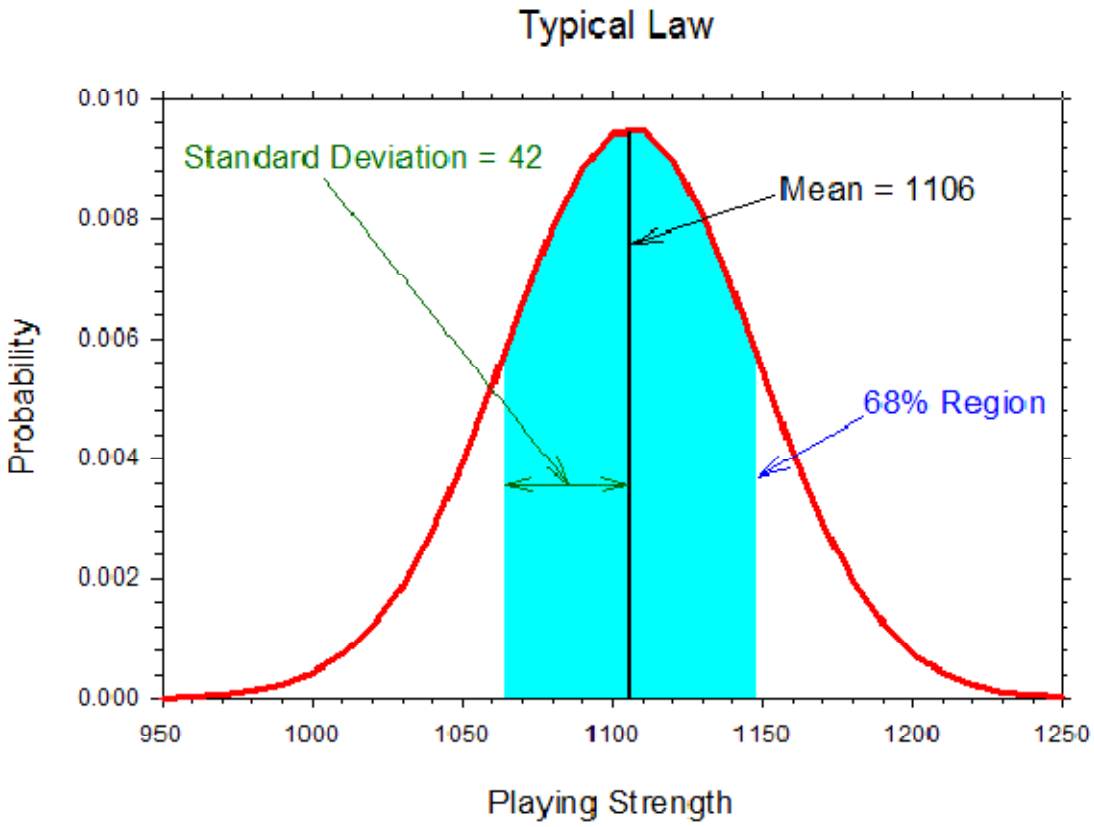
上面的陈述“水平强度在中值左右两个标准偏差以内的可能性是 95%”可能还不够清楚明了，可以换一种说法：给定一个范围，这个范围的上限是特征函数的中值减去两个标准偏差，下限是中值加上两个标准偏差，则这名选手的真实水平强度落在这个范围之内的可能性（也即概率）是 95%。例如，假设一名选手特征函数的中值是 1106，标准偏差是 42，那么：

该选手的水平强度在 1064 和 1148 之间的概率是 68%。

该选手的水平强度在 1022 和 1190 之间的概率是 95%。

该选手的水平强度在 980 和 1232 之间的概率是 99.7%。

图（2）Figure 2



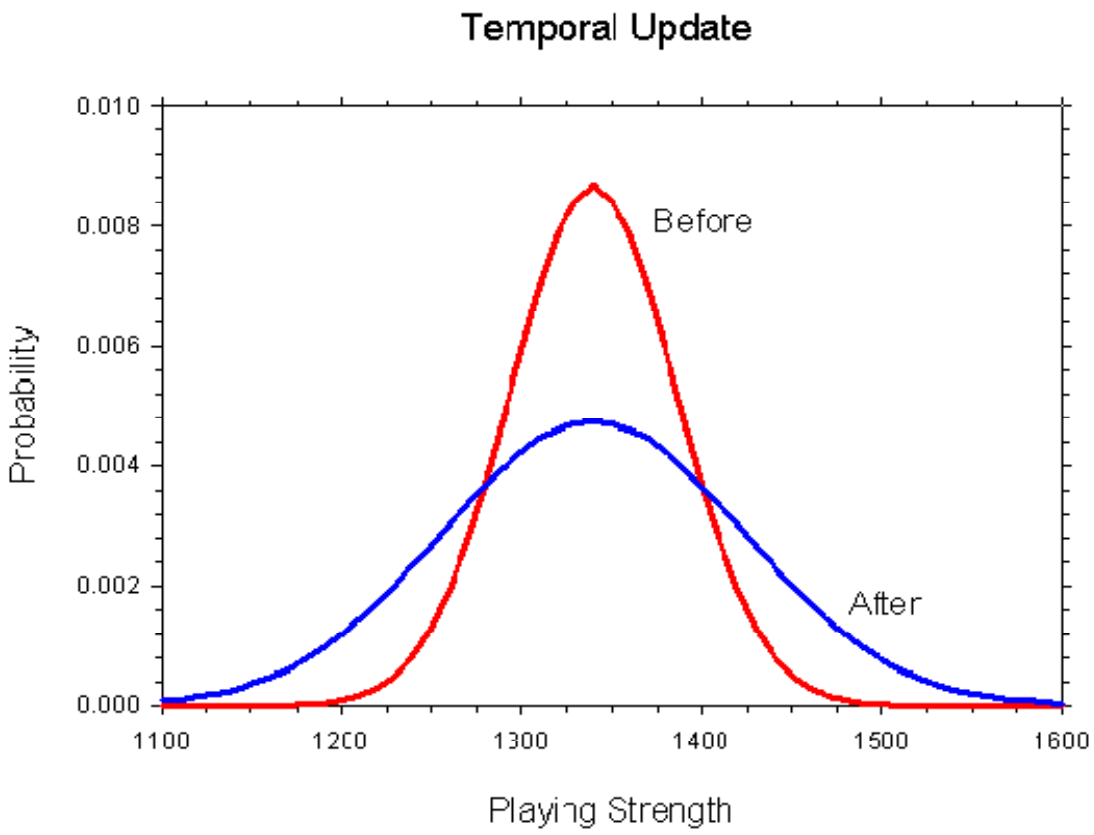
Updates

A player's playing strength may change over time, as the player gets better or worse. Because of this, the more time that has passed since the last event that the player played, the less certain we are of the player's playing strength. The process of updating a player's law to take into account the passage of time is the *temporal update*. The temporal update makes a law more spread out, but doesn't change the mean. For example, if the rating system has not seen a player for a year, the standard deviation of the player's law will increase by at most 70 points (to be precise, the variance increases by exactly 70^2 points per year).

积分更新

选手的水平强度可能会随时间变化，比如水平提高或者降低。因此选手上一次参加赛事的时间离现在越久，我们对他/她当前的水平强度就越不确定。为了考虑这个时间因素，我们采用一个时间校正来更新该选手的特征函数。这个时间校正只增大特征函数的分布宽度，但并不改变特征函数的中值。例如，如果积分制系统发现一名选手在不超过一年的时间内未参加赛事，这名选手特征函数的标准偏差将最多增加 70 分，精确地说，这个校正是标准偏差每年增加正好 70 分。

图（3）Figure 3

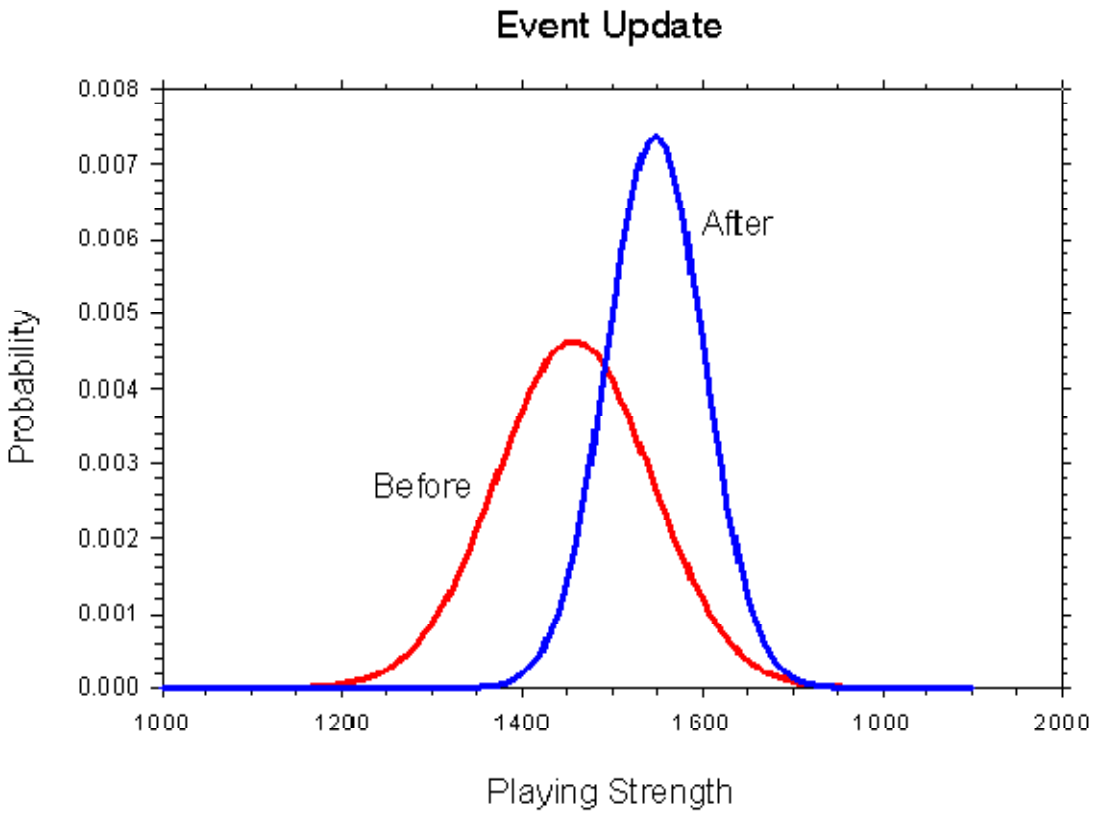


The process of updating a player's law to take into account the player's results in a single event is the *event update*. In theory, we should process all the results of an event as a single group. However, we need to do something that a computer can handle. When doing the event update, the rating system only looks at a player's results and the results of each of the player's opponents. This is similar to what you might do at a tournament: Suppose you lose a match and think your opponent is better than his rating. The way you might check is to look at the draw sheets to see how your opponent did against other players. These are the same results that the rating system looks at when updating your rating.

在选手参加一次赛事后，根据比赛结果对他/她的特征函数进行校正，这个过程称为赛事更新。理论上我们应该把一次赛事的所有比赛结果作为一个集合进行统一处理，但是这样的处理在实际上计算机并不好操作。因此，积分系统在对一名选手进行赛事更新的时候，它仅仅考虑这名选手以

及每一个与这名选手交过手的对手的所有比赛结果。这跟常理一致，就好比你输了一场比赛但你觉得对手要比他/她的积分更强，这时候你可能会去查找抽签对阵表，看看这名对手与其他选手的比赛结果，而这正是积分系统在进行赛事更新时所要做的。

图（4）Figure 4



A small standard deviation makes it harder both to gain and to lose points. If a player with a small standard deviation plays a player with a large standard deviation, then the former's rating will change less than the latter's.

The rating system only cares about who wins a match, not what the score is.

较小的标准偏差意味着相对而言积分更难以发生变化，无论增减都是如此。如果一个具有较小标准偏差的选手与另一个具有较大标准偏差的选手交锋，则前者的积分变化会小于后者的积分变化。

对于任一场比赛，积分系统只考虑比赛双方谁最终获胜，而不关心这场比赛的具体比分是多少。

New Players

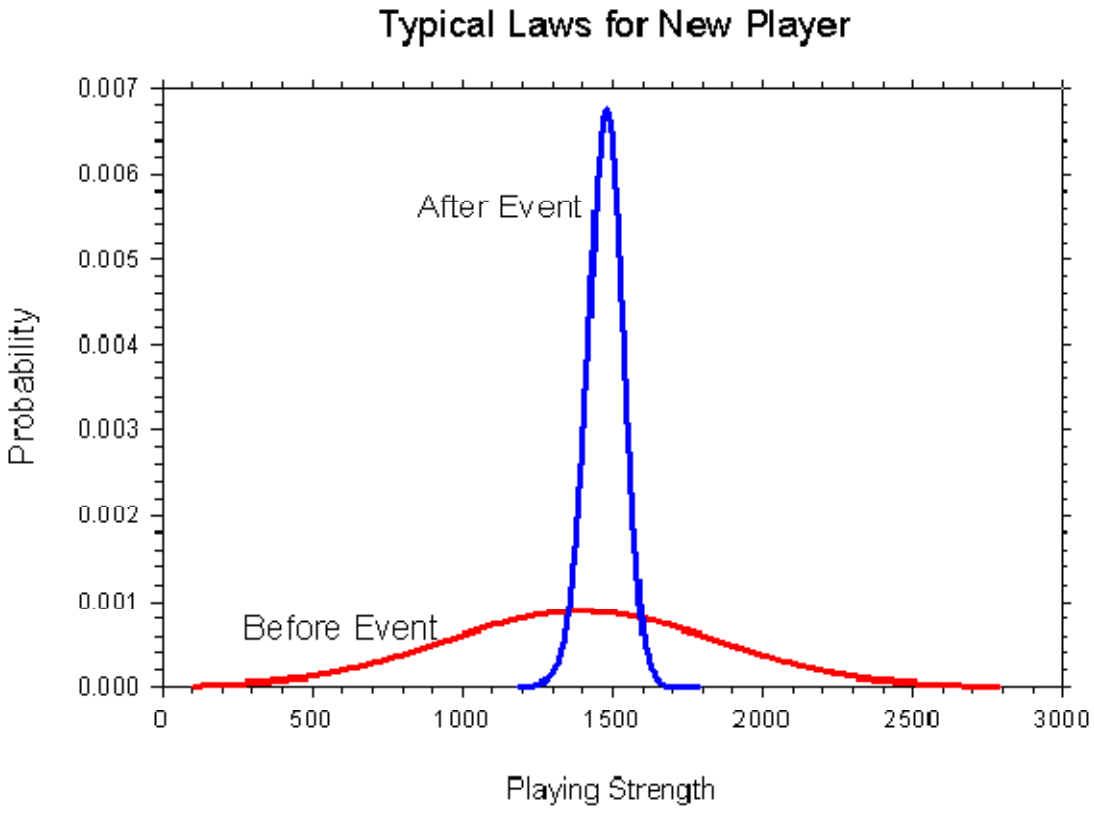
The rating system assigns a law to new (unrated) players. This law reflects what we expect the playing strength of a new player in the particular event to be. Generally, event directors tell us what mean and standard deviation we should use for their event based on their experience. Normally, the standard deviation will be large (several hundred) reflecting the range in playing strength of unrated players that

may enter the event. After playing an event or two, the standard deviation of a new player's law should drop significantly, say, to less than 100. How quickly this happens depends on how many matches the player plays, the outcome of the matches, and the laws of the player's opponents.

新选手

积分系统给没有积分的新选手先选定一个特征函数，这个特征函数只是一个初步预计，它反映了这名新选手在某一赛事中发挥的水平强度。一般来说，赛事主管根据自己的经验向计分中心建议新选手的中值和标准偏差。正常情况下这个标准偏差都相当大（几百）以涵盖新选手水平强度所在的范围，在参加一到两次赛事之后，这名新选手特征函数的标准偏差将明显缩小，减至比如说 100 以下。这个变化过程发生的快慢依赖于多个因素：该选手参加的比赛场数，这些比赛的结果，以及所有与该选手交过手的对手的特征函数。

图（5）Figure 5



Event Processing

Here are the steps the rating system goes through when processing an event:

1. Assign a law to each unrated player.
2. Retrieve the law of each rated player from the database and apply the temporal update.
3. For each player:

a. Calculate an adjusted law (defined below) for each of the player's opponents.
b. Update the player's law for each of the player's matches using in each case the adjusted law for the player's opponent.

The *adjusted law* is the opponent's law updated for all of the opponent's matches except for the matches with the current player. The adjusted law depends on both the player and the opponent. So, the same opponent will have different adjusted laws when different players are being processed.

赛事处理

下面是积分系统处理任一赛事的具体步骤：

1. 给每一名没有积分的参赛选手选定一个特征函数。

2. 从数据库中检索出每一名已经有积分的参赛选手的特征函数，并对其运用时间校正。

3. 对每一名选手：

a. 计算其每一名对手的调整特征函数（详见以下解释）。

b. 对这名选手的每一场比赛，使用该选手这场比赛的对手的调整特征函数来更新该选手的特征函数。

前面已经谈到，在对一名选手进行赛事更新的时候，积分系统仅仅考虑这名选手以及每一个与这名选手交过手的对手的全部比赛结果，而调整特征函数正是为此而引入的。为了对选手甲进行赛事更新，按照上边的步骤 3，首先要计算选手甲的每一名对手(例如选手乙)的调整特征函数，而为了计算选手乙的调整特征函数，又需要根据选手乙和所有与其交过手的选手(但选手甲除外！)之间的比赛结果来作更新。所以说，调整特征函数虽然是一个关于一名选手的对手的函数，但它既跟这名选手又跟这个对手有关。同样的道理，在对两名不同的选手进行赛事更新的时候，两者的同一个对手将有着不同的调整特征函数。也就是说，对于选手甲和选手乙而言，他/她们共同的对手比如选手丙的调整特征函数是不一样的。

Summary Report

Here is a sample from a summary report for an event:

摘要报告

下边的表（1）是从某次赛事的摘要报告中摘取的一段示例：

表（1）Table 1

ID	Name	Initial Rating	Point Change	Final Rating
5766	Bulatao, Jose G.	1797±58	−4	1793±52
5568	Cembura, Julianne	1500±450	−539	961±246
7355	Ching, Joe T.	1984±38	+2	1986±36
6655	Chiu, David	2050±66	+20	2070±49
5925	Collamore, Gil	1121±95	−126	995±59
5184	Conley, Denny	1463±38	+19	1482±34
5044	Cortesi, Tony	1139±90	−58	1081±54

The numbers after the plus/minus signs are the standard deviations of the laws. The initial-rating column contains the rating and standard deviation the player had at the beginning of the event. For unrated players, this is from the law the player is assigned. For rated players, it is from the result of applying the temporal update to the player's final law from their previous event. The final-rating column contains the rating and standard deviation for the player after processing all the matches in the event. The value in the point-change column is the final rating minus the initial rating.

表中位于±号之后的数字表示选手特征函数的标准偏差，表示初始积分的一列包括选手在该次赛事开始前的积分(也即中值)和特征函数的标准偏差。对于没有积分的新选手，他/她的初始积分来自于(赛事主管)给选定的特征函数；对于已经有积分的选手，这个初始积分就是对他/她上一次参加赛事之后的最终特征函数运用时间校正的结果。表示最终积分的一列包括在该次赛事的全部比赛都经过积分制系统处理完毕之后选手的积分和标准偏差，积分变化所在的一列表示选手的最终积分(中值)减去初始积分(中值)的差。

Detailed Report

Here is a sample from a detailed report for an event:

详细报告

表（2，3，4）是从某次赛事的详细报告中摘取的示例：

表（2）（3）（4）Tables 2, 3, 4

Boulard, Claude				**Rating Change 1701±53 + 90 = 1791±41**			
Wins				**Losses**			
Point Change	Opponent's Rating	Opponent	Score	Point Change	Opponent's Rating	Opponent	Score
+30	1915±50	Chen, Wei Teng	−8 7 −6 8 9	0*	1812±34	Bhatia, Sonu	8 9 4
+20*	1812±34	Bhatia, Sonu	−8 −9 6 9 8	−1	2016±48	Maitra, Subhajit	7 −8 4 7
+12*	1785±61	Landsman, Alex M.	−8 8 19 −5 11	0	2189±40	Wang, Yin	8 −6 8 5
+12*	1785±61	Landsman, Alex M.	−3 11 9 −9 9				
+10	1750±34	Marczak, Slawomir	−10 11 8 9				
+5	1629±52	Jordan, Kip	4 −6 3 9				
+2	1587±67	Warrier, Sunil	10 7 7				
0	1366±53	Sharma, Rajeev	4 −7 4 −9 9				

Landsman, Alex M.				**Rating Change 1776±64 − 35 = 1741±54**			
Wins				**Losses**			
Point Change	Opponent's Rating	Opponent	Score	Point Change	Opponent's Rating	Opponent	Score
+8	1630±49	Baird, Jim	8 8 5	−21*	1761±43	Boulard, Claude	−3 11 9 −9 9
				−21*	1761±43	Boulard, Claude	−8 8 19 −5 11
				0	2170±29	Chui, Lim Ming	5 7 8

Marczak, Slawomir				**Rating Change 1752±37 − 7 = 1745±33**			
Wins				**Losses**			
Point Change	Opponent's Rating	Opponent	Score	Point Change	Opponent's Rating	Opponent	Score
+9	1753±50	Baylies, Michael	−5 3 9 8	−7	1771±42	Boulard, Claude	−10 11 8 9
+3	1625±52	Jordan, Kip	9 5 5	−7	1798±33	Bhatia, Sonu	3 −6 11 9
				−5	1811±42	Massarsky, Lev	−10 9 6 9
				0	2015±48	Maitra, Subhajit	6 4 6
				0	2189±40	Wang, Yin	5 9 8

The top left of each table contains the player's name. In the top right of each table under the line "Rating Change" is the player's initial rating and standard deviation, the point change for the player for the event, and, after the equals sign, the player's final rating and standard deviation. Below this, the player's wins are listed on the left and the player's losses are listed on the right.

The value in the “Opponent's Rating” column is the mean and standard deviation of the opponent's adjusted law. As we mentioned above, the rating system uses different adjusted laws for the same opponent when processing different players, e.g., Claude Boulard's adjusted rating is 1761±43 when he played Alex Landsman, but 1771±42 when he played Slawomir Marczak.

The value in the “Point Change” column is the point change for the player for that result. The rating system processes multiple matches between the same two players as a unit. In such a case, the point change shown is the average point change, i.e., the total point change for all the matches between the two players divided by the number of matches. If the two players played more than one match with each other, there will be an asterisk after the point change value. For example, Claude Boulard gained 20 points total for his one win and one loss to Sonu Bhatia.

The point change per match depends on the order that the rating system processes the matches. This order is neither recorded nor shown. So, the values reported as the point change per match are merely suggestive. However, the sum of the per-match point changes equals the total point change for the player for the event, and this total does not depend on the order that the rating system processes the matches.

The dependence of the point change per match on the processing order makes intuitive sense: Suppose we see a 2000 player defeat a 2200 player. We will significantly increase our estimate of the rating of the 2000 player. Now, suppose we see the same player defeat another 2200 player. We will again increase our estimate of the player's rating, but not by as much as we did before.

The number of points gained by the winner of a match will hardly ever equal the number of points lost by the loser of a match. For example, Claude Boulard gained a total of 24 points for his two wins over Alex Landsman, but Alex lost 42 points for the same two matches. In this case, Alex lost more points because Alex's standard deviation is larger and Alex and Claude's means are similar after processing the other matches they played.

Since the point change values are rounded to the nearest integer for displaying, occasionally the sum of the per-match point changes will not equal the total point change for the player for the event. If there is a discrepancy, it will usually be only a point.

在每个表格的左上方是选手姓名，右上方在“积分变化”下面依次列出的是该选手的初始积分(包括中值和标准偏差)，该选手的在这次赛事中的积分变化，在等于号之后是该选手的最终积分(包括中值和标准偏差)。在选手姓名和积分变化之下又分别列出了该选手胜利和失利的比赛情况。

“对手积分”一列给出了这个对手的调整特征函数的中值和标准偏差。正象我们在上边提到的那样，在处理不同选手的时候，积分制对于同一个对手也将使用不同的调整特征函数。例如，在表（3）和（4）中，“同一个对手”ClaudeBoulard 的调整特征函数对选手 Alex Landsman 而言是 1761±43，但是对选手 Slawomir Marczak 而言，Claude Boulard 的调整特征函数却是 1771±42。

“积分得失”一列给出的是选手与某对手在某一场比赛中交手之后该选手积分的增减量。如果两名选手相互之间进行过多次交手，积分制将对这两名选手之间的所有比赛结果进行整体处理。在

这种情况下，积分得失只是一个平均的积分增减量，也即：基于这两名选手之间所有比赛结果的总积分得失，除以他/她们之间的比赛场数。如果这两名选手之间多次交手，系统会在相应的积分得失后面加一“*”号提示。例如，Claude Boulard 在与 Sonu Bhatia 打的两场比赛中，赢的一场得到 20 分，而输的一场不得分。

选手在一场比赛中的积分得失当然会跟积分系统处理不同比赛的顺序有关，这个顺序既无记录也没被给出，所以一场比赛中的积分得失仅可当作参考。但是尽管如此，选手每一场比赛中的积分得失的总和是个定数，它一定等于该选手在该次赛事中的积分变化，与积分系统处理比赛的顺序完全无关。

一场比赛中的积分得失跟比赛的处理顺序有关，这在直觉上也是合理的。假设一名 2000 分的选手击败了一名 2200 分的选手，我们将大大提高对这名 2000 分选手积分的预计。假设这同一名 2000 分的选手又击败了另一名 2200 分的选手，我们将再次进一步提高对这名选手积分的预计，但是这次提高的程度就不象第一次那样高了。

对于一场比赛，胜者赢得的积分几乎总是不等于负者失去的积分。例如，在表（2）和（3）中，Claude Boulard 两次击败 Alex Landsman，共赢得积分 24 分；而同样是这两场比赛，Alex 却失去积分共计 42 分。这是因为 Alex 的标准偏差比 Claude 的更大，另外这也是由于在处理完两者的其他比赛结果后 Alex 和 Claude 的中值相当接近的缘故。

为了显示的方便，积分得失的计算值总是四舍五入到最接近的整数，所以在有些时候，一名选手在每一场比赛中的积分得失的总和不一定恰好等于该选手在该次赛事中的积分变化。如果出现这种情况，误差也往往只不过是 1 分而已。

References

参考文献

Marcus, D. J. (2001), New Table-Tennis Rating System. Journal of the Royal Statistical Society: Series D (The Statistician), 50: 191–208. doi: 10.1111/1467-9884.00271
http://www.ratingscentral.com/Doc/NewTTRS.pdf