Rating Systems for Gameplayers, and Learning

Warren D. Smith*
Email: wds@research.NJ.NEC.COM

July 28, 1994

Abstract — **This report studies rating systems: systems that produce quantitative measures, called "ratings," of the ability of players in a league, based on game results. By "quantitative", it is meant that win odds for a game between two players in the league may be estimated from their ratings. We consider both 'static' and 'dynamic' systems. The latter update the ratings after each game. Attention is given to noise in rating systems and to the distribution of ratings in the player population. Some real-world data is also included. This subject may be of interest to gamblers, gaming leagues, psychologists, consumer groups, and industry.** *Keywords* — Game playing, learning, rating systems, pairwise comparison.

Contents

1 Introduction 1
1.1 Definition of "rating system" 2
1.2 Dynamic and static rating systems 2
1.3 A Problem to be Surmounted: Rating Triangles. 2
1.4 Sketch of the remainder of this report. 2

2 A Rating System is defined by its win probability function 3
2.1 Perfect rating systems? 3
2.2 Binary games 4
2.3 Binary Symmetric Games 4
2.4 Linear Rating Systems 4
2.5 Summary of assumptions 5

3 Some BSG Models and the Kinds of Rating Systems They Imply 5
3.1 The Timesharing Model 5
3.2 The WM model 6
3.3 The WM^n model 6
3.4 Markov game models 7
3.5 Zermelo model 7
3.6 Mosteller model 8
3.7 Minimatch consistency conditions 8

4 Static Rating Systems 9
4.1 Probability distribution of ratings R, given tournament results T 9
4.2 Making the rating assignments R 10

5 Models of Learning and the Distribution of Ratings in the Population 10
5.1 Fluid equation 10
5.2 Model of learning 11
5.3 Solution of learning model 11
5.4 Learning model when "God is fallible" 12

6 Dynamic Rating Systems 12
6.1 Conditions on the update function dyn intended to assure "convergence" 12
6.2 Noise in dynamic rating systems 13
6.3 A convergence claim 14

7 The Internet Chess server's rating system 15

8 Conclusions: How to Design Real Rating Systems 17
8.1 Iterational method for static rating assignments 17
8.2 Kludges 17
8.3 Rating inflation 17

9 Acknowledgements 18

10 References 18

1 Introduction

The best way to introduce the reader to the idea of rating systems is to describe some actual rating systems. The most widely used rating system is the Elo system for rating chess players [elo78]. (A very similar system is used by the USTTA to rate ping-pong players.) Based on win/loss tournament results, the Elo system assigns "ratings" (numbers) to each player in the league; a player's rating is a quantitative measure of his chess ability. Thus a rating of 1300 means the player is an average tournament player, a rating of 2200 is a master, and the world

*NEC, 4 Independence Way, Princeton, NJ 08540

champion has a rating of 2750 or so; the higher the rating, the better the player. Elo ratings are a quantitative measure of ability because, by comparing two players' ratings, one may deduce an approximate value of the probability that one player will beat the other at chess. (For example, a rating difference of 100 points implies that the higher ranked player will win about 2 times for every time that he loses.)

Other rating systems in use include the "computer ranking system" of the Association of Tennis Professionals, which produces a permutation ordering tennis players from the top down, and the "New York Times Computer Football Rankings" appearing in the New York Times. The latter system rates the teams in the college top 20 and NFL. It assigns the top team a rating of 1.000 and gives the other teams ratings between 0 and 1 in a way that "reflects strength...relative to the top team." Based on the description published with the rankings, the algorithm is sophisticated and highly football specific. (Unlike the Elo system, which is readily applicable to games other than chess.) Both of these systems have the disadvantage, relative to the Elo system, that there is no published way to generate game result odds based on ratings.

One could also imagine applying the ideas of rating systems to the design of quantitative psychological tests or to scholastic competitions.

1.1 *Definition of "rating system"*

The most general possible rating system would use as input data a chronologically ordered list of tournament results, where by "tournament result", I mean a three-tuple

(player1, player2, game result).

The rating system would produce as output a list of ratings (reals), one for each player in the league.

Further, each rating system would have associated with it a function which, given any two ratings as input, would output a probability distribution that any given game result will happen in a game between two players with the given ratings.

1.2 *Dynamic and static rating systems*

A "static" rating system inputs all tournament results in one gulp and deduces rating assignments from them. A "dynamic" system inputs individual game results in an "on-line" manner (i.e. it does not know what the remaining input will be) and updates the ratings after each game.

1.3 *A Problem to be Surmounted: Rating Triangles.*

"Rating triangles" are situations in which A beats B, B beats C, and C beats A consistently. We will see that it is theoretically possible to design rating systems which can predict and successfully handle rating triangles. However these systems are all impractical because they use far too much information and require the rating system to be too game-specific for our taste. We will specifically eliminate these systems from consideration in this paper after a brief discussion in §2.

More generally, if the probability that player A beats player B is a totally arbitrary function $f(A, B)$, then most rating systems are going to be worthless. The question is, what assumptions must we make, or are reasonable to make, which permit us to accomplish something?

Experience shows that for human players of real life games, rating triangles are not a significant problem. Probably the reason for this is that the existence of triangles depends on the existence of specific player strengths and weaknesses. Human players tend to compensate for perceived weaknesses and to try for well rounded play, thus avoiding rating triangles.

1.4 *Sketch of the remainder of this report.*

In §2, we will first discuss rating system functions and will quickly turn away from the theoretically possible, but practically unattainable ideal of the "perfect" rating system, which can exactly handle rating triangles without breaking stride, towards the more restricted notion of differentiable "linear" rating systems for "binary symmetric games." The form of the win probability function $\mathrm{wp}(R_1, R_2)$ is of central importance, and in §3, through several independent lines of reasoning - namely self-consistency arguments, various models of games and arbitrary-sounding simplifying assumptions - we will deduce (or perhaps proclaim) that the "right" form for this function is "logistic" (EQ 24) or at any rate must obey the weaker "exponential tail property" (EQ 17). For the first time, it is shown that perfect, and entirely self-consistent, linear rating systems do exist, for some fairly realistic models of games.

In §4, we will show how, once armed with a form for the win probability function wp, we may design "static rating systems." §5 concerns "learning" (one may define learning speed L as the derivative of one's rating with respect to time) and shows a relationship (EQ 51) between learning speed and the distribution of ratings within the player population, which enables deduction of one from the other. By then assuming a certain simple model of learning we may predict the form of the (assumed steady-state) rating distribution from first principles. It turns out (EQ 60) to be a type of "extreme value distribution" that was known previously and involves a doubly exponential falloff in the high rating tail.

§6 considers "dynamic rating systems." Various desiderata restrict the possible forms of the update function to a certain (still, very large) class (EQ 68) (EQ 77); the "fundamental convergence claim for dynamic rating systems" then says that within this class, an enjoyable probabilistic convergence property holds. Finally, we consider "noise" in dynamic rating systems, and give a formula with rather non-intuitive behavior estimating the typical rating noise. With this formula in hand, we show how to design the rating update system to minimize noise.

§7 contains analysis of some data from the internet chess server.

§8 summarizes our results, highlighting the ones of greatest significance from the standpoint of designing a real-life rating system, and argues that systematic "rating inflation" is entirely avoidable.

2 A Rating System is defined by its win probability function

The essential components of a rating system are

1. machinery for assigning ratings to each player based on tournament results,
2. a function which can be used to deduce approximate probabilities for results of games between any two rated players.

This section will be concerned with the second component.

Let a game have n possible results (scores). We desire a function

$$\mathrm{wp}(R_1, R_2) : \mathbf{R}^2 \to \text{probability distributions on } \{1, 2, \ldots n\} \quad (1)$$

which, given the ratings of two players, returns the probabilities of each possible result of a match between them.

We will say that the function wp "defines" the rating system. An assumption central to all of our results is

Independence assumption: All games are independent events.

2.1 Perfect rating systems?

Definition: A rating system will be called "perfect" if the game and the players and the rating system have the property that $\mathrm{wp}(R_1, R_2)$ *exactly* gives the probability distribution of game results in a match between two players with ratings R_1 and R_2, for all possible pairs of players in the league.

Is it possible for a perfect rating system to exist in the real world with human players and some game such as chess or tennis? The answer is yes – although such a rating system would be quite impractical – because everything about each player may be encoded in their (real number) rating. To make matters more concrete, let us consider the game tree of chess. (*Tennis* may also be thought of as possessing a game tree; given the positions, momenta, and internal quantum states of the ball and players [as nodes], the branches of the tree will consist of all the possible actions [state transitions] the players could take at that moment in time. This will be a very large game tree, but (according to quantum uncertainty principles) a finite one.

Now every branch of the tree may be labeled by the probability that a player would choose that branch in a game, assuming he started at the position at the branch's top node. Such a labeled tree completely describes a player. Further, Cantorian theory makes it clear that any such labeled tree may be completely encoded as *one* very high precision real number. (Even if the tree has a countably infinite number of nodes, in fact.) Suppose that each player was given his own labeled, encoded game tree as his rating! It is clear that a function $\mathrm{wp}(R_1, R_2)$ could be designed that would take two such encoded trees and return the probability that one tree won vs. the other.

Clearly the rating system defined by the function described above will be "perfect." Such systems are of some theoretical interest because they can easily handle rating triangles without any contradictions.

However this kind of system is incredibly impractical. First, the amount of information inside each rating would be staggering. Second, there is the slight obstacle of determining the labeled game tree for some player (perhaps by taking him apart?). Even were one to regard this as feasible, legal or ethical restrictions might require one to deduce ratings purely from game result information; taking players apart might be forbidden[1].

We conclude that it is essential to limit the amount of information in our rating systems. This is not as easy as one might expect.

From now on we will assume

Differentiability assumption: $\mathrm{wp}(R_1, R_2)$ *is differentiable.*

Is this enough to prevent tree-encoder systems? No! Indeed, one can make an infinitely differentiable curve which is *dense* in $\mathbf{R}^n$, for example, which includes every point with n rational coordinates.[2] This permits an arbitrarily accurate differentiable rating system.

Another condition that one would like $\mathrm{wp}(R_1, R_2)$ to obey is that, as R_1+R_2 remains constant and R_1-R_2 becomes greater, we would like $\mathrm{wp}(R_1, R_2)$ to reflect "better play" by the first player. If we associate real "desirability" values with each of the n possible game results S, then we might require that the expected desirability to be a monotonic function of $R_1 - R_2$ with $R_1 + R_2$ fixed.

This monotonicity assumption in general *does* prevent tree-encoder systems, because

Observation 1 *No curve exists which is dense in the unit square, and which goes continuously "upward," if the*

[1] Even this restriction does not prevent a tree-encoding system, it merely slows it down. It would seem that for infinitely long-lived chess players who play an unbounded number of games, eventually exploring the entire game tree of chess unboundedly often, one ought to be able to deduce, by fitting "hidden Markov models" to game result data (in a large but finite computation) the labeled game trees for each player...

[2] Hence one can encode the n branch probabilities in a tree as a real number R so that they depend differentiably on R, hence the win probability in a game between two such tree-players depends differentiably on R_1 and R_2. It is a theorem of Besicovitch, however, that there is no differentiable space-*filling* curve. Also, it is topologically impossible for there to be a bijection, continuous both ways, between $[0,1]^2$ and $[0,1]$.

(continuous) "height" function on the unit square looks like this:

```
         |
   -     |     +
         |
---------------------
         |
   +     |     -
         |
```

prevents a suitable encoding of a tree in a real number from existing.

2.2 Binary games

It is painful to continue to allow an abstract set S of possible game results. So from now on we will only consider "binary games" with two possible results: (from the viewpoint of the first player)[3]

"win" or "loss".

Then we may (re)define a function

$$\mathrm{wp}(R_1,\, R_2) : \mathbf{R}^2 \to \text{probability}, \tag{2}$$

which, given two ratings R_1 and R_2, gives an approximate probability that player #1 would beat player #2 in a match between them. (The letters "wp" stand for "win probability.")

2.3 Binary Symmetric Games

A binary symmetric game ("BSG") is a binary game in which one's winning chances are not affected by whether one is the first or second player. For example, "symmetrized chess", a game in which

> players are assigned colors at random and play chess; in the event of draws, they continue playing more games until one of them wins

is a BSG. If the rating system is to be applied to a BSG, then we may adopt the

$$\textit{BSG assumption} : \mathrm{wp}(R_1,\, R_2) + \mathrm{wp}(R_2,\, R_1) = 1. \tag{3}$$

2.4 Linear Rating Systems

Finally, in most of the remainder of this paper, we will be concerned with "linear" BSG rating systems, which are systems in which $\mathrm{wp}(R_1,\, R_2)$ depends only on $R_1 - R_2$ and not on $R_1 + R_2$.

Unfortunately one may exhibit functions $\mathrm{wp}(x, y)$ which *cannot* be "linearized."

[3] One way to handle "ternary" games with a possible "draw" result is to view a draw as a win and a loss, a win is viewed as a two wins and a loss as two losses. Good [good55] suggests another method. Elo, on the other hand, considers it 'inordinately laborious' to consider draws in real rating systems [8.91 Elo]. Another idea is assign each of the $n \geq 3$ possible game outcomes a real "desirability" value, and then redefine $\mathrm{wp}(R_1,\, R_2)$ to return, not a probability distribution, but instead the expected desirability of this distribution. This is in fact what we are about to do, if the desirability of "win" is 1 and of "loss" 0.

Theorem 2 *Let*

$$h(x, y) = \frac{1}{2} + \frac{1}{2}\tanh\left((x-y)[6 + \exp(\frac{-1}{6(x^2+y^2)})]\right). \tag{4}$$

Then there do not exist differentiable functions f, g so that

$$h(x, y) = g(f(x) - f(y)) \tag{5}$$

for all real x and y.

Proof. This follows immediately from the observation that if

$$h(x, y) = g(f(x) - f(y)) \tag{6}$$

then

$$(\frac{\partial h}{\partial y})f'(x) + (\frac{\partial h}{\partial x})f'(y) = 0. \tag{7}$$

Thus the quantity on the right hand side of

$$\frac{f'(y)}{f'(x)} = -\frac{(\frac{\partial h}{\partial y})}{(\frac{\partial h}{\partial x})} \tag{8}$$

must be the same function of y (up to an overall multiplicative factor) no matter what value of x is substituted into it. (Incidentally, this observation represents a general method of finding any solutions f and g to (EQ 6).) One directly verifies that, when $h(x, y)$ is given by (EQ 4), the function of y that results from substituting (for example) $x = 0$ into this expression is *not* proportional to the function of y that results from using $x = 1$. □

Remark. The function $h(x, y)$ in (EQ 4) obeys (EQ 3), is infinitely differentiable, and strictly monotone (this last since

$$\frac{\partial}{\partial x}\left[(x-y)(1 + e^{-1/(6[x^2+y^2])})\right] = \\ 6 + \left[1 + \frac{(x-y)x}{3(x^2+y^2)^2}\right] e^{-1/(6[x^2+y^2])} \tag{9}$$

is positive for all real x, y) in its arguments, tends to 1 when $(x - y) \to \infty$, and tends to 0 when $(y - x) \to \infty$, and hence is apparently an entirely plausible BSG wp function.

So I know of no way to rigorously justify the assumption that linear rating systems exist that will work as well, or almost as well, as nonlinear rating systems, for most real-world BSG's. Still, we will make this assumption in most of the rest of this paper. Certain plausibility arguments support it: The "Timesharing" and "WM^n" BSG models to be discussed later in this paper have perfect linear rating systems; if they are assumed to be widely applicable models of games, then it follows that linear rating systems are widely applicable.

We also note that a rather wide class of wp functions *are* "linearizable."

Theorem 3 *If $\mathrm{wp}(x, y)$ is a ratio of two homogenous polynomials of the same degree, then it may be written in the form*

$$\mathrm{wp}(x, y) = g(\exp(R_1 - R_2)) \tag{10}$$

where g *is some rational function and* $R_1 = \log x$, $R_2 = \log y$.

Proof. Divide the numerator and denominator of said ratio by x^{degree} to write $\text{wp}(x, y)$ as a function of $y/x = \exp(R_1 - R_2)$. □

Indeed, the homogenous polynomials need not be "polynomials," but can more generally be "homogenous algebraic functions"... a term whose definition will be left to the reader. We only mention the example

$$\text{wp}(x, y) = \frac{1}{2} + \frac{1}{2}\frac{x - y}{\sqrt{x^2 + y^2}} \tag{11}$$

which is linearized by the same substitution.

The restriction of our study to linear rating systems reaps great advantages in mathematical simplicity. Some of these advantages follow from the fact that certain equations become linear, others follow from the fact that only linear rating systems do not have a "special" zero rating; the ratings are isotropic. (Although, as a matter of fact, we will reintroduce the concept of the zero rating in §5.2 of this paper on rating distributions.)

2.5 Summary of assumptions

For linear BSG's, we may define the rating system with the real valued, real domained, differentiable function

$$\text{wp}(R_1 - R_2) = \text{probability that player rated } R_1 \text{ will beat player rated } R_2. \tag{12}$$

and require this function to obey:

$$\text{wp}'(x) > 0 \quad \forall x \in \mathbf{R} \tag{13}$$

$$\text{wp}(-\infty) = 0\,, \quad \text{wp}(+\infty) = 1, \tag{14}$$

$$\text{wp}(x) + \text{wp}(-x) = 1. \tag{15}$$

In §6 we will sometimes also use the assumption that $\text{wp}(x)$ is concave-∩ when $x > 0$.

3 Some BSG Models and the Kinds of Rating Systems They Imply

In this section, we will analyze several successively more general models of games and "ability." For each game model we examine, one and only one functional form for $\text{wp}(x)$ will be perfect. A different way to deduce a functional form for $\text{wp}(x)$ is the use of "minimatch consistency conditions," which we will examine in §3.7. Surprisingly, several independent approaches will lead us to the same conclusion that the following "logistic" form for $\text{wp}(x)$

$$\text{wp}(x) = \frac{1}{1 + \exp(-Qx)} \tag{16}$$

is the right one. (Here $Q > 0$ is a constant arising from choice of units.) It will thus seem even more reasonable to assume the weaker

Exponential tail principle: Any linear BSG rating system that will be satisfactory for application to most real world BSG's, must have a win probability function $\text{wp}(x)$ *obeying the "exponential tail property."*

A function $\text{wp}(x)$ obeying the assumptions of §2.5, obeys the "exponential tail property" if

$$\text{wp}(x) \sim \exp(Qx) \tag{17}$$

for some positive Q when x is large and negative[4].

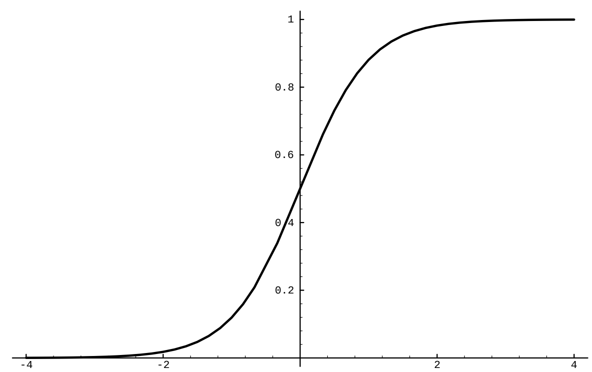

Figure 1: A function $\text{wp}(x)$ obeying the exponential tail property.

Conversely, any linear BSG rating system that does *not* obey the exponential tail property will not be satisfactory for most real world BSG's. This entire section may be thought of as argumentation in support of this principle.

3.1 The Timesharing Model

The first and simplest model of ability that we will consider will be the "Timesharing Model". The essential

[4] Note, the fact that the loss probability goes to 1 when your opponent is rated infinitely higher than you are, in no way makes rating systems obeying the exponential tail property inapplicable to games involving chance – such as backgammon or contract bridge – in which even "God" will lose rather frequently. Arbitrarily high ratings will simply not occur, that is all. Consideration of the "WM model" in §3.2, when right move probabilities are bounded by something below 1, should make this clearer.

point of the timesharing model will be as follows. Consider two players "X" and "Y". Let X beat Y with probability $P_{XY} > \frac{1}{2}$. Then we will model the situation by saying that during a fraction $2(1 - P_{XY})$ of the time X plays "on the same level" as Y. The rest of the time, X plays at an unbeatably higher level. ("At the same level" means wins $\frac{1}{2}$ the time, loses the other $\frac{1}{2}$.) Then let us consider three players "A", "B", and "C", listed in order of decreasing ability. Then

$$P_{AC} = 2P_{AB} - 1 + 2(1 - P_{AB})P_{BC} \quad \text{if } P_{AC}, P_{BC}, P_{AB} > \frac{1}{2} \tag{18}$$

Now let $R_A = A$'s rating, $R_B = B$'s rating, and so on. If we let $x = R_A - R_B$, $y = R_B - R_C$, then

$$\mathrm{wp}(x+y) = 2[\mathrm{wp}(x) + \mathrm{wp}(y) - \mathrm{wp}(x)\mathrm{wp}(y)] - 1 \quad \text{if } x, y > 0. \tag{19}$$

There is a unique function $\mathrm{wp}(\Delta)$ satisfying the above functional equation and the assumptions of §2.5. (The easy proof is omitted. It may be found in [smit84].) The result is

Theorem 4 *In the timesharing model, there is a unique perfect linear rating system, in which*

$$\mathrm{wp}(x) = \frac{1}{2}[1 - (2^{-|Qx|} - 1)\mathrm{sign}(Qx)] \tag{20}$$

where Q is a positive constant specifying choice of units.

3.2 The WM model

We will now consider a slightly more complicated BSG model, the "Wrong Move" ("WM") model. This model involves an imaginary game, the "WM game". The rules of WM are as follows:

> Players make moves simultaneously. There are two kinds of moves a player can make: "wrong" moves and "right" moves. If both players make the same kind of move, the game continues. If one player makes a "wrong" move while his opponent makes a "right" move, the "right" mover wins the game, and the game is over.

We will assume that each player X makes "right" moves with probability P_X, and that all moves are independent.

Then the probability that A beats B is (with $x = R_A - R_B = A$'s rating $-$ B's rating)

$$\mathrm{wp}(x) = \sum_{n=0}^{\infty} [P_A P_B + (1 - P_A)(1 - P_B)]^n P_A (1 - P_B)$$

$$= \frac{P_A(1 - P_B)}{P_A(1 - P_B) + P_B(1 - P_A)} \tag{21}$$

$$= \frac{P_A/(1 - P_A)}{P_A/(1 - P_A) + P_B/(1 - P_B)} \tag{22}$$

if we similarly set $y = R_B - R_C$, then it is a matter of algebra to show that

$$\mathrm{wp}(x + y) = \frac{\mathrm{wp}(x)\mathrm{wp}(y)}{\mathrm{wp}(x)\mathrm{wp}(y) + (1 - \mathrm{wp}(x))(1 - \mathrm{wp}(y))} \tag{23}$$

where $\mathrm{wp}(y) = \mathrm{Prob}(B \text{ beats } C)$ and $\mathrm{wp}(x + y) = \mathrm{Prob}(A \text{ beats } C)$.

Again this functional equation has a unique solution subject to the assumptions of §2.5. (Again the easy proof is omitted and may be found in [smit84].)

The result is

Theorem 5 *In the WM model, there is a unique perfect linear rating system, in which*

$$\mathrm{wp}(x) = \frac{1}{1 + \exp(-Qx)}. \tag{24}$$

(Here Q is an arbitrary positive constant.) The relationship between a player's rating R and right move probability p in this system is

$$p = \frac{1}{1 + \exp(-Q[R - R_{\frac{1}{2}}])}, \tag{25}$$

where $R_{\frac{1}{2}}$ is the rating of a player for whom $p = 1/2$.

The wrong move model might be thought to be a simplistic way to model mistakes in games. (In fact, it was suggested to the author that is may be a good model for the children's game "conkers" but not for any other!) But in fact there are models that are more "sophisticated" than the WM model, but which are based on the same fundamental idea, and which also yield theorem 5. See the next section.

Apparently the reason the WM model is so widely applicable is this. (1) Every real life game can be thought of as having a finite game tree (as discussed earlier); (2) Every move in this tree may be thought of as either "right" or "wrong;" (3) The WM model (and also the WM^n model to be discussed) then arise by thinking of the game tree as a homogeneous entity.

3.3 The WM^n model

An infinite class of game models, which include the WM model as the special case $n = 1$, are the "WM^n" models.

The WM^n model will involve the same imaginary game as the WM model, except that play will continue until one player has made n more mistakes than his opponent, for n some positive integer. The more mistaken player will then lose the game.

This model may be mathematically attacked by reducing it to a "restricted random walk" problem. We look upon the WM^n game as a chain of $2n - 1$ identical WM games.

```
(Leftward              START            (Rightward
 wins)  1-n ... -2  -1 * 1  2 ... n-1  wins)
```

The rules of WM^n may be restated as:

> There are two players, "rightward" and "leftward." They start play at the "START" node * of the chain. At each node, they play a game of WM. Both players then walk over to the neighboring node in the direction of the WM winner's name, and play continues. The game is over when a "win" node is reached.

Now let R be the probability that the rightward player will make nonwrong moves, and let L be the probability that the leftward player will make nonwrong moves. Then

$$p = \frac{R(1-L)}{R(1-L)+L(1-R)} \tag{26}$$

is the probability that the rightward player would beat the leftward in a game of WM^1, and let $q = 1-p$. The problem is clearly a restricted random walk of the type treated in [Hoel, Port, and Stone 1971, eqn 21 pg. 222], and we may immediately write

$$P_{RL} = \text{Prob(right beats left)} = \frac{1-(q/p)^n}{1-(q/p)^{2n}} = \frac{p^n}{p^n+q^n} \tag{27}$$

or

$$P_{RL} = \frac{[R(1-L)]^n}{[R(1-L)]^n+[L(1-R)]^n} \tag{28}$$

It is now a matter of algebra to verify that (where I am introducing a third player with right move probability K and all symbols have usual meaning)

$$P_{RK} = \frac{P_{RL}P_{LK}}{P_{RL}P_{LK}+(1-P_{RL})(1-P_{LK})} = \tag{29}$$

$$\frac{[L(1-K)R(1-L)]^n}{[L(1-K)R(1-L)]^n+[K(1-L)L(1-R)]^n} \tag{30}$$

Now if we set x =(R's rating$-L$'s) and y =(L's rating$-K$'s), (EQ 29) implies that (EQ 23) – the same functional equation as for the WM model – holds, with obviously the same solution (EQ 24).

Theorem 6 *In any WM^n model (for any fixed n), the unique perfect linear rating system is defined by (EQ 24).*

The relation between rating R and right move probability p is

$$p = \frac{1}{1+\exp(-Q[R-R_{\frac{1}{2}}]/n)} \tag{31}$$

which is the same as (EQ 25) except that Q is divided by n.

3.4 Markov game models

It is possible to imagine a BSG in which one needs (for example) to make two more mistakes than one's opponent, followed by a period of 5 consecutive correct moves by one's opponent, in order to lose.

More generally: consider all connected aperiodic Markov chains in which (1) there is a special "start" node, (2) there are exactly two "win" nodes[5], (3) in which all transition probabilities P_{ij} are ratios of two same-degree homogenous polynomials[6] of X and Y (possibly different functions for each (i,j)) where X is a quantity associated with the first player and Y with the second player, (4) in which exchanging $X \leftrightarrow Y$ is an automorphism that interchanges the two win nodes, (5) when $X \to \infty$ with Y fixed, the first player wins[7] with probability $\to 1$.

The WM^n models are very special cases of these very general "Markov game models." (As may be seen by defining X by $X = p/[1-p]$ where p is the WM-player's nonwrong move probability. Define Y similarly for the opposing player. Then the probability of winning a WM game (i.e. making a single transition) is $X/[X+Y]$, see (EQ 21), (EQ 22).)

Theorem 7 *There is a unique perfect linear rating system for any finite Markov game model. It always obeys the exponential tail property.*

Proof sketch. The winning probability for the Markov game may be expressed in terms of certain subdeterminants of the transition matrix. Substitutions $R = \log X$ then suffice to "linearize" this expression, i.e. cause it to become a function of $R_1 - R_2$, due to theorem 3. The exponential tail property is then evident. □

Another generalization of the WM model is called the "RL game." It was discussed in the appendix of [smit84]. The RL game is general enough to include the WM^n model as a special case, but sophisticated enough to have rating triangles.

3.5 Zermelo model

The "Zermelo model" [zerm26], which has also been called the "Bradley-Terry model" [brad76], has been rediscovered many times [ford57] [brad76] [jech83] [stob84]. This is a *nonlinear* rating system based on the function

$$\text{wp}(V_1, V_2) = \frac{V_1}{V_1+V_2}. \tag{32}$$

Although this function appears nonlinear, upon making the substitution

$$R_i = \log V_i \tag{33}$$

(as in the previous section) one obtains

$$\text{wp}(V_1, V_2) = \frac{1}{1+\exp(R_2-R_1)} \tag{34}$$

precisely the form (EQ 24) arising from the WM^n model. So the Zermelo model is isomorphic to (EQ 24), and we have rediscovered it yet again, though perhaps with more justification this time. (Yet another way to obtain the Zermelo model is the "minimatch consistency" argument below.)

[5] Outgoing transitions from "win" nodes are impossible.

[6] Again, the restriction to polynomials could be relaxed to allow certain algebraic functions, cf. the note after theorem 3, and the "same-degree" restriction can be relaxed so long as the determinantal expression in the proof of theorem 7 still has suitable numerator and denominator degrees...

[7] in fact, it will suffice if the win probability tends to any constant in $(0,1]$, not necessarily 1, in which case theorem 7 will still hold with a modified 'exponential tail property' in which win probability tends exponentially to constants which are not necessarily 0 or 1.

3.6 *Mosteller model*

There is one significant proposal in the literature of a linear rating system that does not obey the exponential tail property, namely the "Mosteller model" [most51][8]. This model postulates that every player i has a "true strength" R_i; his "playing strength" varies from game to game according to independent selections from a normal distribution with mean R_i and variance 1. The winner of a binary game is whoever has the higher playing strength during that game. In this model, $\mathrm{wp}(x)$ is a normal cumulative distribution function with variance 2 and zero mean – which of course does not obey the exponential tail property.

There is no real justification for the Mosteller model. Why should all the playing strength variances be the same? Why should strength perturbations be normally distributed? One might claim that normal distributions are somehow "natural" – maybe so, but why shouldn't the square root (say) of the strength perturbations be normally distributed, instead of the perturbations themselves?

The reason that exponential tailed wp-functions are recommended, is that we have shown that they actually generate internally consistent rating systems for a large class of fairly realistic models of games. There have been *no* previous similar demonstrations for any other form of the wp function[9].

3.7 *Minimatch consistency conditions*

Near the end of [stob84] is an interesting comment

> One way to refute a mathematical model is to use the argumentation given in its support to conclude the negation. So suppose the [Zermelo model] is substantially correct. The sport of tennis is a prime candidate for this method... It is reasonable to assume that... the results of a set are governed by the same model [as for the results of a match] (but possibly with different strengths). But it is easy to show that if strengths V_i exist so that the probabilities of winning a match are determined by [the Zermelo model], then it is impossible to find strengths so that the probabilities of winning a match are also given by [the Zermelo model]. This argument applies not only to tennis but also to many other games which are themselves composed of minimatches.

This comment is surprising in view of the fact that we have shown that (EQ 24) – which is isomorphic to the Zermelo model – *does* yield consistent rating systems for WM^n games for all values of n! The surprise is soon resolved upon noticing that Stob uses a different notion of "match, points, games, sets" than we do. Namely, our notion was: the first player to get n more "points" than one's opponent, wins the "match." Stob's notion is: the first player to get n points wins the match.

If the probability that (on any given attempt) A gets a point (while B does not) is p, and $p+q=1$, then with our notion of a match, the probability that A wins an n-point match is

$$\frac{p^n}{p^n+q^n}, \tag{35}$$

while with Stob's definition, we must use the (rather more complicated) expression

$$p^n\sum_{j=0}^{n-1}\binom{n+j-1}{j}q^j = p^n\sum_{k=0}^{n-1}\frac{n-k}{n+k}\binom{n}{k}\binom{2n-1}{n}(-p)^k. \tag{36}$$

In the simplest nontrivial case, namely with $n=2$, (EQ 36) becomes $p^2(1+2q)=p^2(3-2p)$. If there is some god-given functional form of the win-probability function, then this same functional form (only perhaps with some scaling κ of the rating units) should apply for both points and matches, leading to the self-consistency condition

$$\mathrm{wp}(\kappa x) = \mathrm{wp}(x)^2(3-2\mathrm{wp}(x)). \tag{37}$$

Defining $f(x)=\mathrm{wp}(x)-\frac{1}{2}$ and letting $g(y)$, $|y|<\frac{1}{2}$, be the inverse function of $f(x)$ (such an inverse always exists since we have assumed that wp is monotonic and $\mathrm{wp}(0)=\frac{1}{2}$), this functional equation is seen to be equivalent to

$$\kappa g(y) = g(y\frac{3-4y^2}{2}). \tag{38}$$

This is one of a class of functional equations known as "Schröder equations." By theorems in [Kucz90], a nontrivial solution exists if and only if $\kappa=3/2$ and then a unique solution exists that has specified $g'(0)>0$. It is infinitely differentiable for all real y, and analytic in the circle $|y|<1/2$,

When $|x|$ is small, write

$$\mathrm{wp}(x) = \frac{1}{2} + ax + bx^3 + cx^5 + dx^7 + \ldots \tag{39}$$

and then one may solve directly for the coefficients κ, b, c,... one after the other, obtaining

$$\kappa=\frac{3}{2},\quad b=\frac{-16a^3}{15},\quad c=\frac{1024a^5}{975},\quad d=\frac{-65536a^7}{77805},\ldots; \tag{40}$$

this series converges in some neighborhood of the origin, thanks to the previous analyticity statements, and $\mathrm{wp}(x)$ may be evaluated to high accuracy even when $|x|$ is not small by use of (EQ 37) to make it small.

As $x\to\infty$, when $\mathrm{wp}(-x)\to 0$, we may locally analyze the logarithm of $\mathrm{wp}(-x)$, obtaining

$$\log\mathrm{wp}(-x) \sim -A|x|^\alpha, \quad \alpha=\frac{\log 2}{\log(3/2)} \approx 1.709511. \tag{41}$$

[8] Arpad Elo's rating system [Elo78] used in chess is based on the Mosteller model.

[9] Possibly something like the Mosteller model *is* relevant to a sport such as high-jumping, in which head-to-head competition is irrelevant and performance is easily measured against some absolute standard, but of course then assigning ratings is trivial and the function $\mathrm{wp}(R_1,R_2)$ could be determined statistically.

Note that the parameter a is free (it corresponds to one's choice of rating units, and also this freedom arises from the fact that any solution g of (EQ 38) remains a solution when multiplied by a constant) while A is presumably dependent on a.

Note that in (EQ 41) we have $1 < 1.709511 < 2$, so that the "Stob$_2$-consistent" form for $\mathrm{wp}(x)$ is somehow intermediate between the Zermelo and Mosteller forms.

But if we had used an n-game match with $n \neq 2$, then a different expression for a self-consistent $\mathrm{wp}(x)$ would have arisen. In fact a different expression arises for each possible choice of n. In other words, it is possible to design a rating system that is consistent between n_1-point and n_2-point matches of Stob's type, but consistency can apparently only be achieved for *two* distinct values n_i and not more. This is rather disagreeable.

Furthermore, the $\mathrm{wp}(x)$ obtained do not obey the exponential tail property. But by theorem 7, the unique perfect rating system for Stob-type n-game matches ($n \geq 2$) composed of Markov games *does* obey the exponential tail property! Thus we see that not only is the Zermelo rating function (EQ 24) inconsistent for Stob-type matches, as Stob observed, but in fact the situation is far worse: in this huge class of game models, no perfect linear rating system exists which is self-consistent in matches of Stob's type.

Meanwhile, using n-game matches of *my* type, the situation is much more pleasant; the self-consistency equation

$$\mathrm{wp}(\kappa x) = \frac{\mathrm{wp}(x)^n}{\mathrm{wp}(x)^n + (1 - \mathrm{wp}(x))^n} \tag{42}$$

has the *unique* general solution $\kappa = n$ with $\mathrm{wp}(x)$ given by the Zermelo form (EQ 24). Consistency may be achieved for all values of n simultaneously, and of course the exponential tail property holds.

To summarize:

Theorem 8 *Consistency of a perfect linear rating system in n-game BSG matches of "Smith type" occurs, for all $n \geq 1$ simultaneously, if and only if $\mathrm{wp}(x)$ is given by the Zermelo form (EQ 24). Consistency in n-game BSG matches of "Stob type" is impossible to achieve if either: we require consistency for three 3 values of $n \geq 1$ simultaneously, or the rating system is perfect for a Markov game and $n > 1$.*

Bottom line: Linear rating systems can only stretch so far. The Zermelo form (EQ 24) for $\mathrm{wp}(R_1, R_2)$ remains the only one that is known to (1) yield a perfect linear rating system in a reasonable model of game playing, (2) obey an n-game match self-consistency condition.

4 Static Rating Systems

Suppose we have decided on a form for $\mathrm{wp}(x)$. We are still faced with the task of how to use it to compute ratings from tournament results. A static rating system uses

1. a win probability function $\mathrm{wp}(R_1, R_2)$ that defines the rating system.
2. a multiset of tournament results "T", assumed independent. (A "tournament result" is a 3-tuple (a, b, S) where $a \neq b$ are two players and S the result of a game between them.)

to assign ratings to each player.

4.1 Probability distribution of ratings R, given tournament results T

Let "R" represent an assignment of ratings to each player. Then (because our rating function is defined) we know $\mathrm{Prob}(T|R)$, the probability that given the match pairings of T between players rated with ratings R, the game results of T would happen,. (Strictly speaking, what I have called "$\mathrm{Prob}(T|R)$" should be written "$\mathrm{Prob}(T[\text{Results}] \mid T[\text{match pairings}] \text{ and } R)$".)

In fact, up to constants of proportionality (that reflect the many possible unknown time orderings T may have, and R normalization requirements; these constants being independent of R) we may immediately write down $\mathrm{Prob}(T|R)$:

$$\mathrm{Prob}(T|R) \propto \prod_{(a,b,S)\in T} \mathrm{wp}(R_a, R_b)_S \tag{43}$$

where wp is the function which defines the ranking system, i.e.

$$\mathrm{wp}(x, y)_S \tag{44}$$

is the probability that, if a player with rating x played with one with rating y, the result of the game would be S, (as in (EQ 1)) and where R_x is the rating of player x.

(EQ 43) follows immediately from the assumption that all games are independent.

Now $\mathrm{Prob}(R|T)$ may be written by use of Bayes' law:

$$\mathrm{Prob}(R|T) \propto \mathrm{Prob}(T|R)\mathrm{Prob}(R), \tag{45}$$

and $\mathrm{Prob}(R)$ may be written in terms of the probability density function $\rho(x)$ for ratings x, in the (assumed large) player population.

$$\mathrm{Prob}(R)\, d^{\#\text{players}}\vec{R} = \prod_x \rho(R_x)\, dx \tag{46}$$

The left-hand side of this equation represents the probability that the (vector) rating assignment will fall inside a differential (# players)-dimensional box.

It is *not* permissible (or even meaningful) to assume that all real ratings are equally likely, or equivalently that $\mathrm{Prob}(R|T) = \mathrm{Prob}(T|R)$. In particular, if this is done with the $\mathrm{wp}(x)$'s arising from the linear BSG models discussed in the last section, then any undefeated player's most likely and expected ratings will be infinite. Elo seems to have fallen into a trap something like this one. In [3.41 Elo], Elo proposes a method which, if taken literally, will assign infinite ratings to undefeated players.

Zermelo (and later Ford) proposed an iterative scheme with the same flaw. Certainly it is unreasonable that, after entering your first chess tournament with the result of 4 consecutive wins (losses), the rating system should regard you as the greatest chess player ever (worst ever) with a rating of $\pm\infty$. Instead it should make the reasonable assumption that you are somewhere in the tail of the known distribution of ratings in the population.

4.2 Making the rating assignments R

Once the probability distribution of $R|T$ is known, we may apply an arsenal of statistical methods to determine the "best" assignment of ratings R. Two favorite choices for the "best" assignment of ratings might be the expected rating assignment and the most likely rating assignment.

The "noise" in a static rating system may also be estimated by a variety of statistical methods. E.g. a measure of noise in a given rating might be the standard deviation in that rating when the other ratings are held constant.

If one is using a linear rating system one presumably would prefer expected ratings for assignments and the standard deviation for a noise measure. The use of expectation ratings with a fixed form for the rating distribution has the additional advantage that, when used as an "initiation routine" for linear dynamic rating systems (as will be discussed later §8.3) "rating inflation" will automatically be eliminated.

We will now specialize our discussion of static rating systems to BSG linear systems. In this case we may specify tournament results as a matrix

$$T_{i,j} = \text{number of times player } \#i \text{ beat player } \#j. \quad (47)$$

Then we may write:

$$\text{Prob}(R|T) \propto \prod_{i,j} \text{wp}(R_i - R_j)^{T_{i,j}} \prod_k \rho(R_k) \quad (48)$$

where $\text{wp}(x)$ is now defined as in (EQ 12), the rating of player #i is R_i, and i, j, and k range over all the players. The expectation values and variances of the R_i may be written

$$\mathbf{E}(R_i) = \frac{\int \text{Prob}(R|T) R_i d^{\#\text{players}} \vec{R}}{\int \text{Prob}(R|T) d^{\#\text{players}} \vec{R}} \quad (49)$$

$$\text{VAR}(R_i) = \frac{\int \text{Prob}(R|T)(R_i - \mathbf{E}(R_i))^2 d^{\#\text{players}} \vec{R}}{\int \text{Prob}(R|T) d^{\#\text{players}} \vec{R}}. \quad (50)$$

Given a definition for $\text{wp}(x)$, a form for $\rho(x)$ and T, (EQ 48), (EQ 49), and (EQ 50) are all we need to create a rating assignment complete with error estimates.

But it will be impossible to evaluate the integral expressions of (EQ 49) (EQ 50) unless one has a formula for the densities of ratings ρ. Even then, the integrals will probably have to be evaluated numerically[10].

It is possible to replace the numerical evaluation of high dimensional integrals by an iterative scheme consisting of repeated 1-dimensional integrations. (See §8.1.)

Meanwhile, an enjoyable unimodality property makes it easier to compute *most-likely* rating assignments.

Theorem 9 *If* $\text{wp}(x)$ *is of the form (EQ 25), and the logarithm of* $\rho(x)$ *is concave-∩, then the log of the entire expression (EQ 48) is concave-∩ in rating space. (And even if* $\log \rho(x)$ *is not convex, but instead is merely unimodal, then (EQ 48) still has the enjoyable property that it is unimodal along any line, so that there can be at most one maxima.)* □

This makes it rather easy to compute most likely rating assignments by numerically maximizing (EQ 48).

5 Models of Learning and the Distribution of Ratings in the Population

In this section we will show a relationship, called the "fluid equation," between "learning rate" and the distribution of ratings in the player population. By proposing a model of learning, we then attempt to determine the form of this distribution from first principles. There are some philosophical implications about the meaning of "learning," and the distribution of people's "ability," which we will leave for the reader to elucidate.

5.1 Fluid equation

We want to derive analytic expressions for the distribution of ratings. The primary tool for this purpose will be the "fluid equation" of learning.

To derive this equation, we will use the following postulates:

1. There is a differentiable[11] probability density function $\rho(x)$ for ratings x. This function does not depend on time, i.e. the rating distribution is in steady state.
2. New players ("beginners") are being introduced into the player pool at a constant ("birth") rate. All beginners have the same ability level, which we will WLOG (since we are only considering linear rating systems) say is the rating 0.
3. The birth rate is exactly balanced by the "death rate" $D(x)$ (fraction of players having rating x dying per unit time).

Now let $L(x)$ be the average learning rate among the players with rating x; a player whose "learning rate" is L has a rating R that is increasing with time t with speed $(dR/dt) = L$. I will postulate that $L(x) > 0$ for all x in an interval $[0, \epsilon)$, for some positive real ϵ, i.e. the ability level of beginners increases with time.

[10] Alternatively, instead of using a formula for $\rho(x)$, one could use a histogram of an experimental distribution and use numerical integration to evaluate (EQ 49) and (EQ 50).

[11] Actually, by invoking the notion of "weak solutions" of differential equations from fluid mechanics, the assumption of differentiability could be dropped. We will not dwell on the matter.

Then the "fluid equation" is

$$\frac{\partial\rho(x)}{\partial t} = 0 = -D(x)\rho(x) - \frac{\partial[L(x)\rho(x)]}{\partial x} \quad \text{for } x > 0, \rho(x) > 0, \tag{51}$$

with boundary condition

$$L(0)\rho(0) = \int_0^\infty D(x)\rho(x)dx = D. \tag{52}$$

I am calling this the "fluid equation" because of the analogy to the mass conservation equation in one dimensional fluid mechanics. Here, we are conserving the number of players. "Flow rate" of players along the rating axis is $L(x)\rho(x)$; $L(x)$ is analogous to a flow velocity.

The first term on the RHS of the fluid equation represents deaths. The second is due to buildup of rating "flow". The boundary condition forces the integrated death rate, which we call D, to balance the birth rate.

If $L(x)$ is known, this equation may be solved for $\rho(x)$:

$$\rho(x) = \frac{D}{L(x)} \exp\left(-\int_0^x \frac{D(u)}{L(u)} du\right). \tag{53}$$

If $\rho(x)$ is known, then $L(x)$ may be deduced:

$$L(x) = \frac{1}{\rho(x)} \int_x^\infty D(u)\rho(u)du. \tag{54}$$

The fluid equation may also be used to deduce rating vs. time $x(t)$ "development" curves for an "average" player. (Avg. P. always learns at the average rate for one with his rating, but never dies.) We write

$$x(0) = 0, \quad x'(t) = L(x) \tag{55}$$

so that

$$\int_0^{x(t)} \frac{dR}{L(R)} = t \tag{56}$$

This equation may be algebraically solved (sometimes) for $x(t)$.

5.2 *Model of learning*

We will now extend the WM game model to develop a "model of learning," with the purpose of deriving an analytic form for $L(x)$, and thence for $\rho(x)$. The postulates of this model are as follows.

1. Each player has personal repertoires of heuristics that he uses to determine whether a move will be correct.

2. Each additional heuristic added to a player's repertoire gives him knowledge about the correctness of a constant fraction c_1, $0 < c_1 < 1$ of the moves he did not know anything about before. (Indeed, c_1 should be quite small.) So if a player has h heuristics in his repertoire, he will make correct WM moves with probability

$$\frac{1}{1 + \exp(-[R - R_{\frac{1}{2}}])} = 1 - c_3(c_2)^h \tag{57}$$

   where $c_2 + c_1 = 1$, and I am

   - assuming that if one or more of a player's heuristics apply to some move decision, then he will move correctly with probability 1.
   - assuming that if none of a player's heuristics apply to a move decision, he will move correctly anyway (i.e. through sheer luck) with probability $1 - c_3$.
   - calling his rating "R".
   - The game is WM; The rating system is the unique perfect one defined by (EQ 24); I am using (EQ 25) to relate ratings and correct movement probabilities in WM and taking $Q = 1$ WLOG by choice of rating units.

3. I will assume that each player acts to increase his repertoire of heuristics ("learn") according to the following procedure.

   (a) Propose a heuristic. All proposed heuristics will have a fixed probability c_4 of being valid heuristics.

   (b) Confirm or disprove the heuristic through experiments conducted while playing (rated games of) WM. All players will be assumed to be playing WM games continuously at a rate of a one move per unit time, and it will be assumed that all hypotheses can be confirmed or disproved with a fixed number of experiments E. An "experiment" will mean any opportunity the player has to make a move decision in a game of WM in a situation in which none of his repertoire of accepted heuristics apply, but his tentative heuristic does. If the tentative heuristic is valid, add it to the repertoire. If it is invalid, discard it.

   This procedure attempts to embody the scientific method.

4. Finally, I'll also assume that the death rate $D(R)$ is asymptotically constant in the high-rating ("D_∞") and low rating ("D_0") tails.

A priori, there are quite a few free parameters (e.g. c_1, c_3, c_4/E, D_0, and D_∞) in our model. (Remember $c_2 + c_1 = 1$.) However, we will eventually wind up with much fewer free parameters.

5.3 *Solution of learning model*

It turns out that our model of learning may be almost exactly solved. We will need to make two tiny approximations to make solution possible.

The expectation value of the rate at which a player who knows h heuristics is adding additional heuristics to his repertoire is

$$\mathbf{E}(\dot{h}) = c_5(c_2)^h = \frac{c_6}{1 + \exp(R - R_{\frac{1}{2}})} \tag{58}$$

where $c_5 = c_4 c_1/E$ and $c_6 = c_5/c_3$ are positive constants. The first approximation we will make is to regard h as a continuous, rather than discrete, quantity, and thus pretend $\dot{h} = \mathbf{E}(\dot{h})$. This enables us to use (EQ 57) and the chain rule to write $L(R) = \dot{R} = -(1 + \exp(R_{\frac{1}{2}} - R)\log(c_2)\dot{h}$ so that

$$L(R) = c_7 e^{-R} \tag{59}$$

where $c_7 = \log(1/c_2) c_6 e^{R_{\frac{1}{2}}} > 0$.

The second approximation: we'll assume that the death rate $D(R)$ is independent of R, that is, $D(R) = D_0 = D_\infty = D$. (Alternatively, if we don't make this assumption, asymptotic solution in the high and low rating limits still will be possible.) This enables us to do the integral in (EQ 53) to obtain

$$\rho(R) = c_8 \exp\left(c_8 + R - c_8 e^R\right), \quad R \geq 0 \tag{60}$$

where $c_8 = D/c_7 > 0$ is a constant.

Note the *doubly* exponential dropoff as $R \to \infty$.

The time development curve is

$$x(t) = \log(1 + c_7 t). \tag{61}$$

Interestingly, the distribution (EQ 60) exemplifies the "first type" of "extreme value distribution" cf. [Gumb54, 3.16]. Also, a well known observation from mortality tables ("Gompertz's law" [gomp25]) is that the distribution of human lifespans also has a doubly exponential falloff.

5.4 *Learning model when "God is fallible"*

To model learning in games such as backgammon in which (see footnote 4) unbondedly large ratings are impossible, one may modify the learning model by replacing the "1" in the right hand side of (EQ 57) with some value $p_{\max}$ in $(0, 1)$. This leads to a learning curve which behaves, near the maximum possible rating $R_{\max}$, like

$$L(R) \approx (R_{\max} - R) c_9. \tag{62}$$

This leads to a time development curve like

$$x(t) = (1 - e^{-c_9 t}) R_{\max} \tag{63}$$

and the high tail of the rating distribution behaves like

$$\rho(x) \propto (R_{\max} - R)^{c_{10}}. \tag{64}$$

This again is an extreme value distribution, again arising quite without our requesting it to.

6 Dynamic Rating Systems

In a "dynamic" rating system, every player in the league starts out with an initial rating. After each game one plays, one's rating is adjusted to reflect the result of the game.

Dynamic rating systems for general games are thus described by an iteration

$$(R_1, R_2) \leftarrow \mathrm{dyn}(R_1, R_2, S), \tag{65}$$

where dyn() is a function with value being the new (adjusted) ratings R_1 and R_2 of two players 1 and 2 whose ratings were R_1 and R_2 respectively just before they played a game with result S.

We will now restrict ourselves to the BSG linear case as usual. BSG linear games may be described by the following (simpler) iteration:

$$\begin{aligned} R_2 &\leftarrow R_2 + \mathrm{dyn}(R_2 - R_1)/2 \\ R_1 &\leftarrow R_1 - \mathrm{dyn}(R_2 - R_1)/2 \end{aligned} \tag{66}$$

which is the rating adjustment to be made after player 2 beats player 1. (This is the most general possible iteration which affects and depends only on $R_2 - R_1$ and does not affect or depend on $R_2 + R_1$).

Note that this iteration conserves the sum of the ratings of all the players in the league.

6.1 *Conditions on the update function* dyn *intended to assure "convergence"*

The first problem that we will ask ourselves concerning dynamic ranking systems will be:

Given that our ranking system is defined by $\mathrm{wp}(x)$ – what conditions need $\mathrm{dyn}(x)$ satisfy in order to ensure convergence of ratings to accurate ratings? (And what do we mean by convergence?)

Our first necessary convergence condition will be what Good calls the "special principle of no incentives":

$$\mathrm{dyn}(x)\mathrm{wp}(x) = \mathrm{dyn}(-x)\mathrm{wp}(-x). \tag{67}$$

This states that one's expected rating increase after a game should be zero; i.e. correct ratings are stationary. The reason Good calls this a "principle of no incentives" is that it denies players the capability to abuse the rating system. If this condition were not met for some x, then one could play players who were rated x above one's own rating and thus expect to gain (or lose) rating points in a way not justified by one's actual ability. I will call it the "weak stability condition" ("WSC").

The WSC may be solved generally for $\mathrm{dyn}(x)$ to give

$$\mathrm{dyn}(x) = h(|x|)\mathrm{wp}(-x) \tag{68}$$

for some arbitrary function $h(|x|)$, which we would now like to restrict.

Another condition necessary for convergence is the "strong stability condition"

$$\mathrm{sign}(y - x) = \mathrm{sign}[\mathrm{dyn}(x)\mathrm{wp}(y) - \mathrm{dyn}(-x)\mathrm{wp}(-y)] \tag{69}$$

which says that if one's true ability is reflected by a rating y above one's opponent's, but the official rating difference is (erroneously) x, then the expected rating adjustment will be in the direction $y - x$. Note that the strong stability condition incorporates the weak one as the special case $y = x$, if $\mathrm{sign}(0) = 0$.

Using (EQ 68) in the strong stability condition yields

$$\mathrm{sign}(y - x) = \mathrm{sign}(h(|x|)[\mathrm{wp}(-x)\mathrm{wp}(y) - \mathrm{wp}(x)\mathrm{wp}(-y)]). \tag{70}$$

The argument of the sign function may be simplified, by using the symmetry condition (EQ 15), to

$$h(|x|)[\mathrm{wp}(y) - \mathrm{wp}(x)] \tag{71}$$

and then by the monotonic increase property of $\mathrm{wp}(x)$ (EQ 13), we see that the strong stability condition is merely equivalent to

$$h(|x|) > 0. \tag{72}$$

We might also require the "expected error decrease" condition:

$$|y-x| > |y-x-\mathrm{wp}(y)\mathrm{dyn}(x) + \mathrm{wp}(-y)\mathrm{dyn}(-x)| \tag{73}$$

which states that the magnitude of the expectation of the error in the official rating difference between two players (whose "true" rating difference should be y, but is x at the moment) will decrease after each game. Assuming that the strong stability condition holds, this is equivalent to saying that

$$\frac{\mathrm{wp}(y)\mathrm{dyn}(x) - \mathrm{wp}(-y)\mathrm{dyn}(-x)}{y-x} < 2 \tag{74}$$

which is the same as

$$\frac{\mathrm{wp}(y)(1-\mathrm{wp}(x)) - (1-\mathrm{wp}(y))\mathrm{wp}(x)}{y-x} h(|x|) < 2 \tag{75}$$

which simplifies to

$$0 \le \frac{\mathrm{wp}(y) - \mathrm{wp}(x)}{y-x} h(|x|) < 2. \tag{76}$$

We already know this quantity is ≥ 0 from the strong stability condition. The "< 2" prevents the correction from "overshooting" its goal by a factor of 2 or more. What we really would like is to bound this, not in the interval $[0, 2)$, but instead in an interval $[1-\epsilon, 1+\epsilon]$ but, considering what happens when $y - x \to \infty$, it is impossible to make the lower bound nonzero.

A necessary and sufficient condition for (EQ 76) to hold is that

$$0 < h(|x|) < \frac{2}{\max_y [\mathrm{wp}(y) - \mathrm{wp}(x)]/[y-x]}. \tag{77}$$

If wp obeys all the conditions of §2.5, including the convexity condition, then a sufficient condition for (EQ 77) is

$$0 < h(|x|) < \left|\frac{x}{\mathrm{wp}(x) - 1/2}\right|, \tag{78}$$

and regardless of convexity, it will suffice that

$$0 < h(|x|) < 2/\max_x(\mathrm{wp}'(x)). \tag{79}$$

All these conditions may always be met by the choice

$$h(|x|) = h = \text{constant}, \quad \text{for any positive } h < 2/\max(\mathrm{wp}'(x)). \tag{80}$$

This choice has 3 advantages: (1) its extreme simplicity, (2) it assures that the update function $\mathrm{dyn}(x)$ is monotonically increasing and (3) bounded by $|\mathrm{dyn}(x)| < h$. We will assume it in the rest of this section.

6.2 *Noise in dynamic rating systems*

Now the question arises: how reliable will the dynamic rating system we've just discussed be? How much "noise" will be inherent in the system?

I will address this question for $h(|x|) = h$ a constant. Some of my statements will be trivially generalizable to the $h(|x|)$ non-constant case.

If h is made small enough, it would seem that there will almost surely be arbitrarily small noise in anyone's estimated rating (after the estimated ratings have "converged" to the players's "true strengths"). On the other hand, the smaller h is, the longer the rating system will take to adjust to strength changes. This section will clarify this tradeoff.

First, there will be noise inherent in the system due to players whose ratings have not yet converged from their initial guess rating (which was given to them when they were "born") to a rating representing their actual ability. If the standard deviation of all the ratings in the league is σ, then the expected "noise" (RMS error) in a rating R due to this effect will be roughly

$$\frac{1}{\sqrt{3}} \frac{\sigma^2}{Gh} \tag{81}$$

where G is the number of games one plays in a typical lifetime, since it takes typically σ/h games before a typical player's rating will move from his initial to his final rating even if he is initialized at the mean of the rating distribution, and the fraction of players who have not yet played this many games is $\sigma/(hG)$, and $\sqrt{\int_0^1 x^2 dx} = 1/\sqrt{3}$.

Second, there will be noise due to statistical fluctuations in the ratings of players whose ratings have converged; every time you win or lose a game, your rating makes a small discontinuous jump. It is to be expected that this noise will have size h or larger, but it is not obvious exactly how much larger. (Namely, as $\sqrt{h}$ - read on.)

For the purpose of estimating the size of this noise, we will set up the following model.

Let the difference of two player's ratings be have equilibrium value R_0. Let the noise in this rating difference have value x, i.e. $R - R_0 = x$. Then if the two players only play each other, x obeys the following stochastic iteration:

$$x \leftarrow x + \begin{pmatrix} (1-f(x))h & \text{with probability } p = f(0) \\ -f(x)h & \text{with probability } q = 1 - f(0) \end{pmatrix} \tag{82}$$

where $f(x) = \mathrm{wp}(x + R_0)$.

Now let $\theta(x)$ be the probability density of x. (Presumably $\theta(x)$ is peaked near $x = 0$.) Then $\theta(x)$ obeys the difference equation

$$\theta(x) = \frac{p\theta(y)}{1 - hf'(y)} + \frac{q\theta(z)}{1 - hf'(z)} \tag{83}$$

where $x = y + h(1 - f(y)) = z - hf(z)$.

Now let us consider the above difference equation in the limit when x, y, z, and h are small. Then we may approximate $f(x)$ by its linear Maclaurin expansion

$$f(x) = f(0) + f'(0)x = p + sx. \tag{84}$$

Now we see that y and z are given by

$$y = \frac{x - hq}{1 - hs}; \quad z = \frac{x + hp}{1 - hs} \tag{85}$$

Then the difference equation (EQ 83) becomes (where I am calling $k = hs$ and I am using the identity $1/(1-k) = 1 + k + k^2 + \ldots$):

$$((1-k)\theta(x) = p\theta\left((x - hq)(1 + k + k^2 + \ldots)\right) + q\theta\left((x + hp)(1 + k + k^2 \ldots)\right) \tag{86}$$

Now we expand θ in a Taylor expansion about x:

$$0 = k\theta(x) + [p(kx - hq) + q(kx + hp)][1 + k + k^2 + \ldots]\theta'(x) + [p^2(kx - hq)^2 + q^2(kx + hp)][1 + k + k^2 + \ldots]^2\theta''(x)/2 + \ldots \tag{87}$$

Taking advantage of the fact that $p + q = 1$ gives

$$0 = k\theta(x) + \frac{kx}{1-k}\theta'(x) + \frac{(p^2 + q^2)(kx)^2 + 2(q - p)pqkxh + 2(pqh)^2}{(1-k)^2}\frac{\theta''(x)}{2} + \ldots \tag{88}$$

So far everything we have done has been exact. Now let us neglect all terms of order $O(h)$ and assume as we do so that x is of the same order of magnitude as the square root of h. (Then the nth derivative of $\theta(x)$ has the same order of magnitude as the $(-1-h)$th power of the square root of h.) (Also remember that $k = hs$.) We then divide the resulting equation by a common factor of h. Then:

$$s\theta(x) + sx\theta'(x) + h(pq)^2\theta''(x) = \mathbf{O}\left(\frac{x^2}{h} + 1\right). \tag{89}$$

This differential equation has solution

$$\theta(x) = \frac{s}{\sqrt{2\pi}h}\frac{1}{(pq)^2}\exp\left(\frac{-s}{2(pq)^2h}x^2\right) \tag{90}$$

A normal distribution! (Compare with [26.2.27 pg 933 HOMF].)

The standard deviation of this distribution is

$$\sqrt{h/\mathrm{wp}'(R_0)}pq. \tag{91}$$

The fact that x is about the same size as the square root of h justifies (admittedly circularly) the selection of negligible terms in (EQ 88)[12].

Therefore an estimate of the noise caused by statistical fluctuations of rating differences about their equilibrium value is

$$\frac{1}{\sqrt{2}}\sqrt{h/\mathrm{wp}'(\sigma)}\mathrm{wp}(\sigma)\mathrm{wp}(-\sigma) \tag{92}$$

where σ is the standard deviation (or some other measure of the typical rating difference) of the rating distribution.

Thirdly and finally, there is an additional source of dynamic rating system noise. This is noise due to the presence of rating triangles. In our treatment of statistical fluctuations so far we ignored this phenomenon by the device of considering two players who played only each other (or by pretending that a perfect wp function existed). However in a three player triangular universe, more noise is possible; in a triangle $A > B > C > A$ we can imagine A's rating shooting up as he plays B and dropping as he plays C.

Consideration of the game tree for games like chess, etc. shows that there is no limit inherent in the game itself to the triangularity possible. (It is easy to construct three chess-playing algorithms which beat one another in a triangular way with probability 1.)

It is therefore difficult or impossible to bound triangle noise, but one thing at least is clear: triangularity of players in a league should not depend on which dynamic rating system is used! We conclude that triangle noise is described by a typical noise amplitude

$$T(R). \tag{93}$$

Taking the square root of the sum of the squares of the various noise-amplitude contributions (EQ 81) (EQ 92) (EQ 93), we obtain

$$\sqrt{[T(R)]^2 + \frac{1}{2}[\mathrm{wp}(\sigma)\mathrm{wp}(-\sigma)\sqrt{h/\mathrm{wp}'(\sigma)}]^2 + \frac{1}{3}[\sigma^2/(2hG)]^2} \tag{94}$$

as an estimate[13](EQ 94)(EQ 95) have been derived under the fictitious assumption that one is always playing a "fixed " opponent whose true strength differs from yours by σ. In the event that one's imaginary fixed opponent is of equal strength, then there will be more noise; the relevant formulas are obtained from these by replacing all occurrences of $\mathrm{wp}(\sigma)$, $\mathrm{wp}(-\sigma)$, and $\mathrm{wp}'(\sigma)$ by $\frac{1}{2}$, $\frac{1}{2}$, and $\mathrm{wp}'(0)$ respectively. of the total rating noise in a dynamic BSG linear system. By setting the derivative of this with respect to h equal to zero, we may solve for an optimum value of h which minimizes the noise. This value is found to be approximately

$$h_{\text{opt}} \approx 3^{-1/3}\left(\frac{\sqrt{\mathrm{wp}'(\sigma)}\sigma^2}{\mathrm{wp}(\sigma)\mathrm{wp}(-\sigma)G}\right)^{2/3}. \tag{95}$$

6.3 A convergence claim

Now consider the "global L_2-error figure"

$$\sum_{i=1}^{N}(R_i - S_i)^2 \tag{96}$$

where R_i is the official rating, while S_i is the "true strength" of the ith player in an N-player league. (Assume that true strengths exist so that the win probabilities are exactly given by $\mathrm{wp}(S_i - S_j)$, i.e. a perfect rating

[12] This selection is also justifiable on the basis of numerical evidence (Monte Carlo experiments) which I have not presented here; also any other choice would have ended up, not circularly justifying itself, but instead leading to a contradiction.

[13] Do not take the numerical coefficients 1/2, 1/3 too seriously. Equations (EQ ??)eq900

system exists.) It is then a matter of straightforward algebra to see that, if $R_1 + R_2$ is held constant, then the quadratic form $(R_1 - S_1)^2 + (R_2 - S_2)^2$ achieves its unique minimum when $R_1 - R_2 = S_1 - S_2$. Therefore, decreasing the error in rating difference between two players will decrease the global error figure.

Considering the noise estimate (EQ 94), in which $T(R) = 0$ since we have assumed the rating function wp is perfect, and in which $G \to \infty$ since we assume that players live forever and play an unbounded number of games,

Claim 10 (Fundamental convergence claim for dynamic rating systems) *In an N-player league having a perfect BSG rating function* $\mathrm{wp}(x)$, *in which* $\sum_{i=1}^{N} R_i = \sum_{i=1}^{N} S_i$, *if we update the R_i's using a strongly stable dynamic rating system* $\mathrm{dyn}(x)$ *obeying (EQ 80), then after a sufficiently large number of statistically independent games (in which the games [=edges] form a multigraph over the players [=vertices] with sufficiently high edge connectivity) the global error figure will almost surely be $< N\sqrt{h}\kappa$ where*

$$\kappa = \max_x \ \mathrm{wp}(x)\mathrm{wp}(-x)/\sqrt{\mathrm{wp}'(x)} \qquad (97)$$

is a finite constant (if wp *obeys the assumptions of §2.5, including the concavity assumption, and the exponential tail property, κ will obviously exist) depending solely on the rating function* wp *itself, and not depending on the values of S_i nor the initial values of R_i.*

I have called this a "claim" rather than a "theorem" since I really do not have a rigorous proof of it. Such a proof should be possible to produce. It would have to be based upon making the noise estimate (EQ 94) rigorous, which would entail struggles similar to those encountered by mathematical physicists [kac47] trying to rigorize physics related to the Langevin stochastic differential equation [chan43]. Specifically, the plan would be to get rigorous bounds on the solution of the stochastic difference equation (EQ 82) by constructing a related finite Markov chain, bounding its solution, and proving that its solution provides an upper bound on the variance of the solution of (EQ 82).

7 The Internet Chess server's rating system

The *Internet chess server* ("ICS") is hardware/software which allows chessplayers from all over the world to play chess electronically over the internet. About 1500 games are played per day, and of these about 515 are "rated blitz" games. (Games between registered players are rated automatically as they happen, and various data is maintained concerning each registered player. There are about 2800 registered players.) Prof. Daniel Sleator of Carnegie-Mellon University, the author of ICS, kindly provided me with the blitz ratings list and win-loss records as they varied over the period June 13-August 28 1993.

Sleator, having seen an early draft of this paper, designed ICS's rating system along lines recommended in this paper. Specifically, it is a standard dynamic system (EQ 68) with $h(|x|) = 64$ and $\mathrm{wp(x)} = \frac{1}{1+10^{-x/400}}$. All games between "established" players (those who have played at least 20 games) are rated according to this system, but the "provisional" players are initialized into the system with a special procedure (we quote Sleator):

> A player is said to be provisional if he/she has played fewer than 20 games. Your rating during your provisional period is the average of a set of values, one for each game played. These values depend on whether the opponent is provisional, or is established. The value against a provisional player is the average of the two ratings (using 1600 if the player has never played) plus 200 times the outcome (which is -1, 0, 1 for loss, draw and win). The value against an established player is the opponent's rating plus 400 times the outcome. Some extra points are now added to the rating for the purpose of keeping the average rating of all established active players close to 1720. In particular, 1/10th of 1720 minus the current average is added to the rating.

When an established player plays a provisional player, the adjustment in the established rating after each game is $n/20$ times the usual value, assuming the provisional player has played n games before.

In the below, we will only consider registered players who have played at least one rated blitz game.

Most registered players are humans, but at ≈ 30 are computers. The average number of rated games played by each player is 159.5, but this number may give a biased picture since the distribution is highly skewed. Its median is only 35. (The most active player is the computer "gnufour," which has played 5988 blitz games. Of the 17 players with ≥ 2000 games, 10 are humans and 7 are computers.) Ratings vary from $R_{\min} = 450$ to $R_{\max} = 2869$. The mean rating is $R_{\mathrm{mean}} = 1622$ and the standard deviation is $\sigma = 324$. The "birth rate" of new registered players who've played a blitz game, averages 6.6 players per day, which in a league of ≈ 2300 players corresponds to a birth rate (measured in fraction of league per day) of 1/350. The "death rate" D is impossible to determine directly.

Assuming the number G of games played during a typical "lifetime" is 35 one finds from (EQ 95) that $h_{\mathrm{opt}} = 53.0$. This is close to Sleator's choice $h = 64$, so Sleator apparently knew what he was doing. (If we use the modified formula from footnote 13, then $h_{\mathrm{opt}} = 41.0$. The upper bound (EQ 79) is $h < 1390$.)

According to (EQ 94), using $h = 64$ and $G = 35$, the estimated noise (standard deviation) in ones rating is supposed to be

$$28.8 = \sqrt{T(R)^2 + 25.4^2 + 13.5^2} \qquad (98)$$

using $T(R) = 0$ so that triangle noise is not included in the "28.7." The use of $h = h_{\text{opt}}$ in place of $h = 64$ would have only decreased this noise estimate to 28.3. The figure 25.4, estimating noise arising from rating fluctuations under the assumption that one always is playing somebody σ different from one's own strength, rises to 37.3 if one is assumed always to be playing somebody of equal strength (as in footnote 13), in which case 28.8 rises to 39.7.

Meanwhile, the *measured* standard deviation (based on the sample standard deviation among measurements of players's ratings at 5 moments in time during a 2.5-month period; only the 136 established players who had played at least 8 games between measurements were used) is 53.3. The difference between this and the predictions (28.8 and 39.7) is, I suppose, due to triangle noise or to actual fluctuations in the playing strength of the players (which really is not the fault of the rating system) see the next paragraph. This noise figure is fairly constant across rating classes but may exhibit a downward trend for very high rated players, since of the 136 players surveyed, the 14 players who were rated over 2100, had a average noise figure of 40.4. This is to be expected from (EQ 92) since the high-rated players will probably often have to play opponents with considerably lower strength, corresponding to a larger value of "σ" in that equation.

The *learning rate* of players on the ICS ($d(\text{rating})/d(\text{time})$), may be estimated in two ways: directly by tracking players, and indirectly from the static ratings distribution by means of the solution (EQ 54) of the fluid equation. In the 136-player sample, the average number of rating points gained over the 2.5 month period was 21.1, which is statistically extremely significant (there is definitely learning going on, say 5 standard deviations worth of confidence). Plotting mean-points-increase versus 100-point-wide mean-rating bin, we observe

rating bin	bin popul.	pt.incr.
1200-1299	5	115.2
1300-1399	7	40.3
1400-1499	18	49.3
1500-1599	21	20.4
1600-1699	12	-3.9
1700-1799	15	-18.6
1800-1899	13	38.2
1900-1999	22	21.3
2000-2099	9	53.9
2100-2199	5	-9.0
2200-2299	5	-57.2
2300-2399	2	-52.0
2400-2499	2	2.5

that the lower rated 68 of these players gained an average of 29.9 points while the higher rated 68 gained an average of only 12.2 points. Therefore, there seems to be a definite trend that learning gets harder the better you already are, consistent with the prediction (EQ 59) of the model of learning in §5.2. Sleator's choice of units is such that (EQ 59) would predict that you learn $e \approx 2.71828$ times more slowly if you have 174 extra rating points. This seems rather inconsistent with the data for the two groups of 68 above, since the higher group is indeed learning a factor $\approx e$ more slowly, but it has an average rating that is about 450 points higher, not 173. But of course, the learning model (modeling chessplayers as simple minded creatures playing WM) was rather crude, so one should not expect a miracle. We are within a factor of 3.

Meanwhile, working from the entire rating distribution using the fluid equation and assuming a rating-independent death rate D, we get

ratg bin	binpop	L/D	log(L/D)
900-999	16	438.9 +- 117	6.0842 +- 0.2738
1000-99	72	96.9 +- 12.	4.5739 +- 0.1212
1100-99	195	35.1 +- 2.66	3.5582 +- 0.0739
1200-99	440	14.83 +- 0.75	2.6969 +- 0.0507
1300-99	614	9.77 +- 0.43	2.2795 +- 0.0441
1400-99	844	6.25 +- 0.25	1.8318 +- 0.0393
1500-99	869	5.08 +- 0.20	1.6253 +- 0.0399
1600-99	837	4.26 +- 0.18	1.4481 +- 0.0418
1700-99	850	3.20 +- 0.14	1.1624 +- 0.0438
1800-99	686	2.84 +- 0.14	1.0447 +- 0.0501
1900-99	535	2.50 +- 0.15	0.9178 +- 0.0586
2000-99	466	1.80 +- 0.12	0.5880 +- 0.0692
2100-99	281	1.66 +- 0.15	0.5048 +- 0.0919
2200-99	176	1.35 +- 0.17	0.2976 +- 0.1259
2300-99	82	1.32 +- 0.25	0.2754 +- 0.1885
2400-99	42	1.10 +- 0.31	0.0910 +- 0.2929
2500-99	10	2.00 +- 1.01	0.6931 +- 0.5568
2600-99	15	0.83 +- 0.70	-0.231 +- 0.8

(Note, the distribution given is actually the sum of the distributions for established players at all 5 points in time.) The average learning rate among players rated 1200-2499 is, according to the table above, $4.98D$. According to the tracking data from the preceding table it was 20.1, so that an estimate of the "death rate" is $D \approx 4.06$, which, considering the tracking period was 77 days, leads to the conclusion that $D \approx 1/19$ per day – i.e. the entire league will die out in only 19 days! The reason this estimate is exceedingly silly is that the tracked players were a biased sample of players who were particularly active (in fact, essentially the 136 most active mebers of the league at the time!), and thus presumably have a learning rate far above the usual one, leading to an estimate for the death rate which is much too high. Consulting tables of the distribution of games-played-in-life, we see that the top 6% have played about 30 times more games than the median, so our estimate for the death rate should probably be revised downwards by a factor of 30 to $D \approx 1/600$ per day. This is about half the birth rate, consistent with the fact that ICS is still young and growing.

The fact that the learning rate figures in the preceding two tables *agree* (except for the factor 4.06 and statistical noise, whose magnitude is rather obvious and unfortunately is unavoidably rather large) rather well, is confirmation of the validity of the fluid equation. The fact that the $\log(L/D)$ behaves roughly linearly as a function of rating R, at least in the range $R > 1200$, is confirmation of the qualitative predictions of the learning model in that rating range, including the prediction of doubly-exponential falloff in the population density in the high rating tail[14]

8 Conclusions: How to Design Real Rating Systems

I would recommend the use of a static rating system involving (EQ 24), (EQ 47), (EQ 48), (EQ 49), and (EQ 50) for the purpose of initializing new players into a dynamic system involving (EQ 24), (EQ 66), (EQ 68), (EQ 80), (EQ 94), and (EQ 95). Either an analytic approximation like (EQ 60) or else a "bootstrapped" experimental form for the rating distribution may be used in (EQ 48). Note that (EQ 49) does not preclude the use of established ratings to help determine initial ratings of new players. (One could also use a most-likely rating assignment based on (EQ 48).)

8.1 Iterational method for static rating assignments

For the purpose of numerically evaluating the multidimensional integrals in (EQ 49): It is possible to evaluate (EQ 49) in a way which does not require the direct numerical evaluation of integrals in a large number of dimensions, but instead only requires the ability to do one dimensional integrals. This iterational method is:

1. Guess ratings for all players whose ratings we wish to determine, except for player i's.
2. Set player i's rating to its expected value from the *one* dimensional integral (EQ 49) with all ratings except i's fixed.
3. Cycle i to a new player and loop back to step 1 until satisfied that the ratings have converged.

8.2 Kludges

Dynamic rating systems must be updated after each game, which may represent an excessive computational labor. Thus the original Elo system was semidynamic: updated after each tournament rather than after each game, with the use of "average tournament ratings" instead of individual opponent ratings. It seems to me that remembering the time, date and opponents in each individual match is worth the trouble.

The Elo system as used by the US Chess federation [USCF92] includes several "kludges" that were required, possibly to correct imperfections in the whole system. These include "bonus feedback points" that one gets when playing "underrated" opponents, and rating "floors" that prevent one's rating from dropping more than about 200 Elo below one's best ever rating. Also, high rated players are rated using a different form for $\text{dyn}(\Delta R)$ which insures slower rating adjustments for higher rated players, but which does not conserve rating sum in matches between high and low rated players.

We have no fundamental objection to slower adjustments for higher rated players. The idea here is that higher rated players tend to play more games, so they can be rated more accurately by use of smaller adjustments. But a more sensible method, in view of (EQ 95), would be to use a value of h which shrinks proportionally to $(G_1 G_2)^{-1/3}$ when rating a game between players who have played G_1 and G_2 rated games in their lives so far. We *do* have a fundamental objection to not preserving rating sums. We also fundamentally object to "extra" adjustments for "drastically underrated" players – the normal rating system ought to take care of this sort of thing without the need for a kludge. Probably the fact that the USCF felt that this kludge was needed is related to the fact that their *wp* function does not obey the exponential tail property (cf §3.6).

My feeling is that all such kludges must be used with extreme caution. I belive that the USCF would be better off if said kludges were eliminated, in favor of the simple inflation fighting measure discussed in the next section, and also if their $\text{wp}(x)$ function were of form (EQ 24).

8.3 Rating inflation

One problem which has arisen is "rating inflation" in which various ratings systems become out of sync. For example it generally agreed that USCF ratings are equivalent to somewhat (100-200 elo) lower international (FIDE and foreign) chess ratings. The usual explanation for rating *deflation* is that players leaving the rating system tend to have higher ratings than players entering it – causing the mean rating (in a sum-preserving system) to fall. Rating inflation is then caused by overdoing meddling intended to correct for deflation.

The "learning models" of §5 model the way it would be *if* every player were magically always rated accurately (i.e. reflecting his true playing strength at any time). In such a situation, we claim that the rating distribution should rapidly converge to a steady state distribution and stay there. Imagine a box of gas into which are introduced new gas molecules with energy E. These rapidly thermally equilibrate with the rest of the gas. Meanwhile, once a minute, a demon removes half the gas molecules whose energy $> 2E$. We claim that the energy distribution among the gas molecules in the box rapidly converges to a fixed distribution, and the temperature of, and number of molecules in, the gas is asymptotically *constant.*

[14] The ratings of the 33609 chessplayers who played in tournaments organized by the United Stated Chess Federation during the last year, are available on floppy disks from the USCF for $15. I have not used this data because the USCF's rating system does not follow the recommendations of the present paper nearly as closely as does the ICS rating system, cf. §8.2. A fit to 1978 USCF data was done in [smit84], with the results that that rating distribution was fit considerably better by a learning model than by a normal or a Maxwell-Boltzmann distribution, and again the doubly exponential falloff in the high tail was visible in the linear behavior of $\log(L/D)$.

Similarly to this physical analogy, suppose we have a magically accurate rating system in a league with a large number of players. New players enter the league with ratings x assigned from a fixed distribution[15] $\rho_{\text{in}}(x)$, and players in the league having a rating R die (or quit) with a probability $D(R) > 0$ per unit time. Then (cf. (EQ 51)) the probability density $\rho(x)$ of ratings x obeys the partial differential equation

$$\frac{\partial \rho(x)}{\partial t} = \rho_{\text{in}}(x) - D(x)\rho(x) + \frac{\partial \rho(x) L(x)}{\partial x} \tag{99}$$

where $L(x)$ is the average learning rate of players with rating x. Then we claim that the rating distribution will, as t increases, approach the fixed (time invariant) limiting distribution

$$\rho(x) = \frac{1}{L(x)} \exp\left(-\int^{x} \frac{D(u)}{L(u)} du\right) \times \int_{-\infty}^{x} \frac{\exp(\int^{w}[D(v)/L(v)]dv)\rho_{\text{in}}}{L(w)} dw \tag{100}$$

(where by "$-\infty$" we mean the greatest lower bound on those x for which $\rho_{\text{in}}(x) = 0$) which is the unique time-invariant solution of (EQ 99). If $D(x)$ is uniformly bounded below by some positive constant δ, then in fact this approach will be exponential, with a time constant of $\leq 1/\delta$.

In other words, "rating inflation" will be nonexistent and the system will be self-correcting.

The problem with the above treatment is that the differential equation (EQ 99) has not enforced any global conservation of rating sum. To understand rating inflation in a *non*-magical sum-preserving dynamic rating system, consider the following scenario. We have N players in a league, all initially beginners having strength 1. They gradually learn, and a year later their strengths have increased to 2. However, their ratings are *still* 1, since sums were preserved. Then, N more players (beginners) join the league. Suppose they are initially given a rating of 1. Rapidly, a "population inversion" occurs, with the beginners sinking to rating 1/2 and the older players rising to rating 3/2. (Their "true strengths" are 1 and 2, of course.) Now the older players die and we are left with a league of N players, whose true strengths are 1 but whose ratings are 1/2. If these players then grow older and wiser, their true strengths will increase to 2 but their ratings will remain at 1/2. Then N new beginners with ratings 1 are introduced, and quickly the population inverts so that the young players have ratings 1/4 and the older ones 5/4. Continuing the cycle of adding beginners with rating 1, dying, and learning, we see that after g generations, the mean rating of the $2N$-player leagues will be $2^{-g} + \frac{1}{2}$. Thus we see that rating inflation will still eventually *stop*, since the mean rating in the league will eventually approach $\frac{1}{2}$. This virtue is only achieved, however, by the expedient of giving beginners the same rating (on average) as experienced players, so that beginners's ratings will usually fall rapidly at the beginning of their careers, later commencing a slow rise, until they die – at which point they will have the same rating (on average) as when they were born.

Rating-flation is thus entirely avoidable, even in sum-preserving systems with learning, but only if beginners are initially overrated. In sum preserving rating systems in which beginners are introduced into the rating system at low ratings (accurately reflecting their low playing strength) and in which players tend to learn before they die, being accurately rated the whole time, rating-flation is *unavoidable.*

But here is an inflation-fighting method that will entirely circumvent this dilemma. Suppose a player who was born at rating R_0, dies, at rating R_1. Simply distribute $R_1 - R_0$ rating points equally to the other players in the league each time this happens. (If a player turns out not to have really died, your error is correctable by undistributing the points.)

[15] We now allow the possibility of assigning new players ratings chosen from some fixed distribution, as opposed to the perhaps less accurate method of assigning them a single number R_A. This could be accomplished in practice by using a static rating system for initialization. Every month (or at some fixed intervals) new players would be initialized into the dynamic rating system based on their results during the previous month, as assessed by using a static rating system. To ensure that all the new players of the month actually did have a fixed mean rating, their initial ratings could be normalized by adding a constant.

9 Acknowledgements

This paper is based upon my Bachelor's thesis [smit84] which is available on microfilm from MIT libraries. My thesis supervisor was Prof. Richard P. Stanley of the MIT mathematics department. Later, Colin L. Mallows, Nick Littlestone and Daniel Sleator contributed useful comments.

10 References

[Batc79] Batchelder, W.H., and N.J. Bershad: The statistical analysis of a Thurstonian model for rating chessplayers, J. Mathematical Psychology 19 (1979) 39-60.

[Brad76] Bradley, R. A.: Science, statistics, and paired comparisons, Biometrics 32 (1976) 213-232.

[Chan43] Chandrasekhar, S.: Stochastic problems in Physics and Astronomy, Rev. Mod. Phys. 15,1 (January 1943) 1-87.

[David63] David, H. A.: The method of paired comparisons, Griffin, London 1963.

[Elo78] Elo, Arpad, E.: The Rating of Chessplayers, Past and Present, Arco Publishing Inc., N.Y., N.Y. 1978.

[Ford57] Ford, L.R. Jr.: Solution of a ranking problem from binary comparisons, Amer. Math. Monthly 64 (October 1957, part II) 28-33

[Gomp25] B. Gompertz: On the nature of the function of the law of human mortality, Philos. Trans. Royal Soc. 115 (1825) 513-585.

[Good55] Good, I. J.: On the Marking of Chessplayers, Mathematical Gazette, 39, (1955) 292-296.

[Hoel, Port, Stone 1971] Hoel, Port & Stone: Introduction to Probability Theory, Houghton Mifflin, Boston, Mass. 1971.

[HOMF] Abramowitz, Milton (Ed.): Handbook of Mathematical Functions, Tenth Printing, NBS, December 1972.

[Gumb54] Gumbel, Emil J.: Statistical theory of Extreme value distributions and some practical applications, National Bureau of Standards Applied Math. series #33, February 1954.

[Jech83] Jech, Thomas: The ranking of incomplete tournaments, Amer. Math. Monthly 90 (1983) 246-266

[Kac47] Kac, Marc: Random walk and the theory of Brownian motion, Ammer. Math. Monthly 54,7 (September 1947) 369-391

[Kucz90] Kuczma, M., B. Choczewski, and R. Ger: Iterative functional equations, Cambridge University Press (Encycl. Math. and its Applics. vol. 32) 1990.

[Most51] Mosteller, F.: Remarks on the method of paired comparisons I: The least squares solution assuming equal standard deviations and equal corrections, Psychometrica 16 (1951) 3-9

[Smit84] Smith, Warren D.: Rating systems and Learning, Bachelor's thesis, Massachusetts Institute of Technology, January 1984.

[Stob84] Stob, Michael: A supplement to "A mathematician's guide to popular sports", Amer. Math. Monthly (May 1984) 277-282.

[Thom75] Thompson, M.: On any given sunday: fair comparison orderings with maximum likelyhood methods, J. Amer. Statist. Assoc. 70 (1975) 536-541

[USCF92] USCF rating guide, available from USCF 186 Route 9W, New Windsor, New York 12553.

[Zerm26] Zermelo, E.: Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Warscheinlichkeitberechnung, Math. Zeitschrift 29 (1926) 436-460