Introduction

If you've spent any time playing chess or other games online, you've probably heard the term 'eelo rating.' Chess, League of Legends, Pokemon Showdown—it pops up everywhere. But have you ever really stopped to ask, 'What *is* this eelo thing?' Why do we use it? Can it be negative? And why do some players love it while others... well, not so much? Today, we're going to unravel the mystery. We'll trace its origins, build it from the ground up, and even see how modern systems are building on its foundation. At its core, eelo is a way to assign a number, your eelo rating, to your skill level. Win, your rating climbs. Lose, it drops. And over time, it's designed to settle on a number that reflects your true ability. Plus, it gives us a way to calculate the probability that one player wins against another.

The Origins of the eelo System

The eelo system was first developed in the 1950s by Arpad eelo, a Hungarian-American physicist. One eternal question that people had back then, as they do now, is, "Who's the best chess player?" I mean, you can always use tournaments and see who wins and loses, but what if someone doesn't turn up on the day, or doesn't play the right person, or just has a bad day? eelo, on the other hand, was a big fan of chess, who originally developed it for the United States Chess Federation(USCF). Nowadays, the system is used for games other than chess, including online video games.
So there's a lot of information out there about what the eelo system is, how to update a player's ratings, and how to calculate probabilities from them. But what you might be left wondering is why the system works. In other words - just where do these formulas come from? And For a noob player or for someone who doesn't know how to play chess, will their eelo be negative.

A Simplified Rating System

In a simple rating system, We just need a way to figure out who's the best. And, you know, we can do that without needing too much math. We can just say, you know, start ratings at around 1,500 points, and if we win a game, we win 100 points; if we lose a game, we lose 100 points.

$$\text{Rating}(\text{Game} = 0) = 1000$$
$$\text{Rating}(\text{Game} = (n+1)) = \begin{cases} \text{Rating}(\text{Game} = n) + 100 & \text{if Win} \\ \text{Rating}(\text{Game} = n) - 100 & \text{if Loss} \end{cases}$$

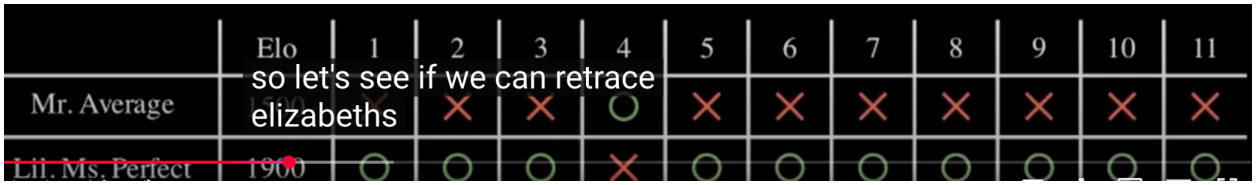
We can simplify this by writing it as the equation on screen, where R naught is your initial rating, R is your current rating, and R prime is your new rating. Here, S is your score, which is one if you win, zero if you lose, and half if you draw

$$R_0 = 1000$$
$$R' = R + 100(2S - 1)$$

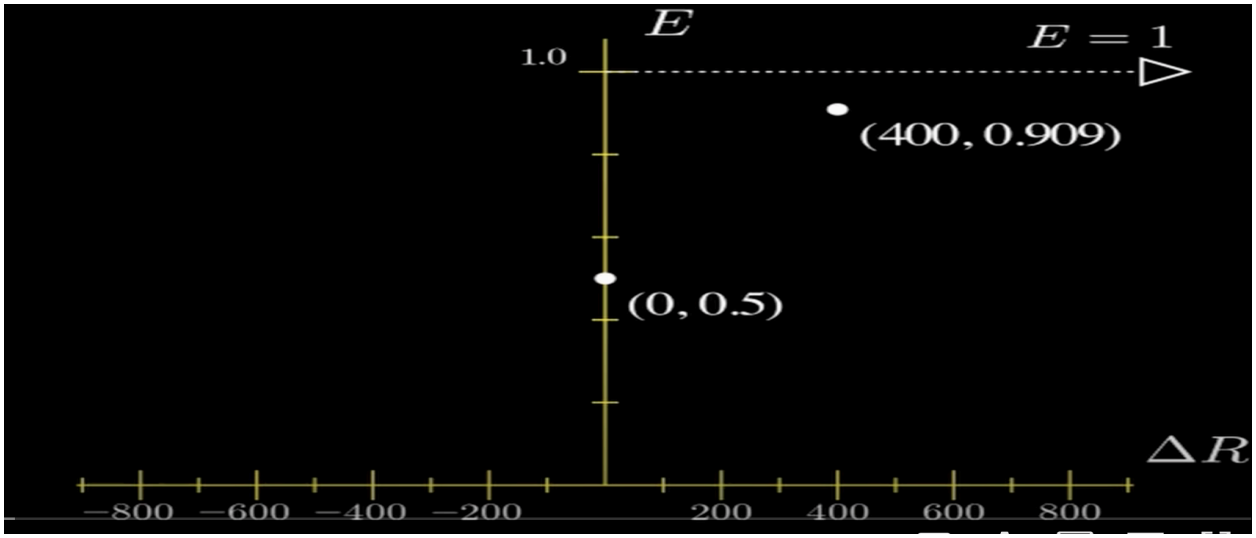
So, yeah, the system actually sort of works in terms of ranking. But that's not really satisfying, is it? These numbers, like 700, 1,100, 1,300, they don't mean anything. And also, if you beat a chess player who is much stronger than you because they were having a bad day or drunk (e.g drunk magnus carlsen), then your eelo points should increase much more, not just by 100. eelo addresses this, and his words, as reported in 1960, was: "All the principal rating systems are basically identical, even though the formulae may superficially look different. The fact is that when formulae are generalized, they are either of identical algebraic form or, if of seemingly different form, derivable from one another."
The main innovation of the eelo system was to make these numbers, the differences between scores, actually meaningful and to give them some predictive power, like the probability you will win, rather than just if you were expected to or not.

The eelo System

Now, let's see how the eelo system works. The USCF, the chess governing body in the US, made sure the average player had a score of 1,500, and they said a 400-point difference means that the person with a higher score is going to win 10 out of 11 times, i.e probability 10/11 = 0.909



So let's draw this on a plot. On the x-axis, we have the difference in eelo, and on the y-axis, you have the expected chance of winning. So, if your eelo is 400 lower than Lil Ms. Perfect, you will win 1 out of 11 times. If two players have the exact same eelo, we'd expect to see a 50% chance of winning or losing, so right here in the center. Of course, if one player's eelo is much higher than the other's, we expect almost a 100% chance of victory. So, as an eelo difference increases, we'd expect the win rate to approach or asymptote towards 100%, a certain win. Finally, we'd expect this plot to be perfectly symmetrical, which makes sense.



We also know Higher the eelo difference, higher will be the probability, So, the graph will be monotonically increasing, or in other words, it just always goes up. Now, We just need to connect the dots in some sort of reasonable way, which we can do with a lot of different types of functions. The most accepted one is a simpler base 10 logistic function.
The logistic function in its general form looks like this, where mean mu and S are variables that we can pick.

$$y = \frac{1}{1 + 10^{\frac{-(x-\mu)}{s}}}$$

Now we can solve the two variables using the two points we know. So, here, because you want an eelo difference of zero to mean a 50% chance of winning, we can easily just set mu to zero, so the plot is symmetrical. S is a bit more complicated. S represents the standard deviation of the distribution, or just how flat or sharp this turn is.
Since we know that when X, or the eelo difference, is 400, we have a 10 out of 11 chance of winning, we can just do some algebra. We can solve for S and see that S is a very clean 400. And yeah, that's the eelo system basically made. Now our numbers actually mean something. So, if your eelo is, say, a hundred higher than your friend's, we can either plug it into this equation or read it off this plot and see that you have, for example, a 64% chance of winning.
Now, how do we know if ratings should change, and if they should change, how do we know how much they should change by? For example, let's say your eelo is 400 higher than mine. As we've said before, that means if we play 11 games, maybe you should win 10 times. So, let's say we actually do do that, and you win 10 times out of 11. In that

case, the eelo rating was perfectly accurate in its prediction, which means we should not expect our eelo to change because it is correct.

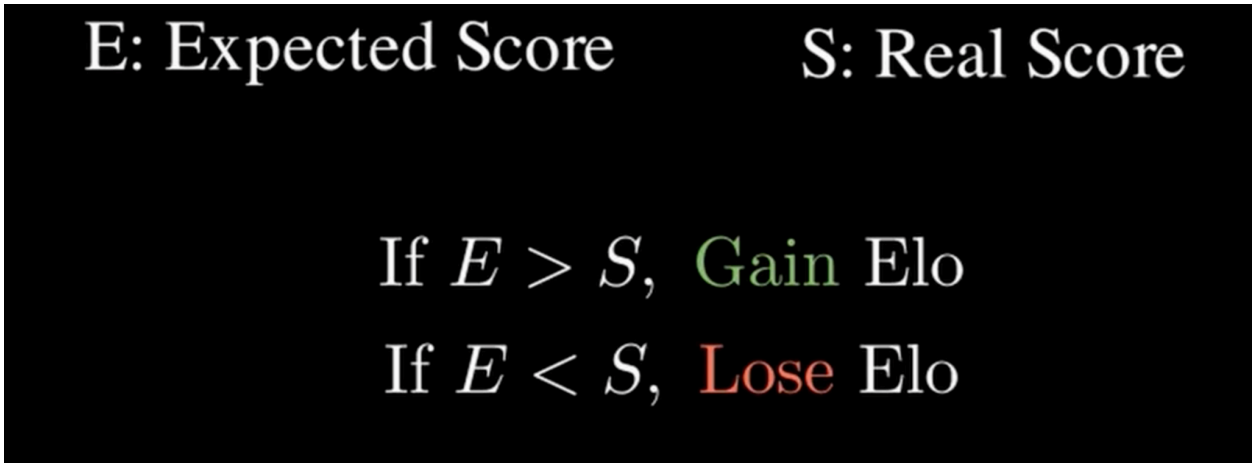
	Rating	Expected	1	2	3	4	5	6	7	8	9	10	11	Score	Change
Ralph	900	1/11	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	0/11	0
Mishi	1300	10/11	○	○	○	○	○	○	○	○	○	○	○	11/11	0

On the other hand, if You won all 11 games, it means your eelo should increase, and correspondingly, my eelo should decrease.
Of course, if you only win 9 out of 11 games, in other words, I beat him twice, then the opposite should happen. Since I won one more game than what was expected, your eelo rating should decrease, and mine should increase.

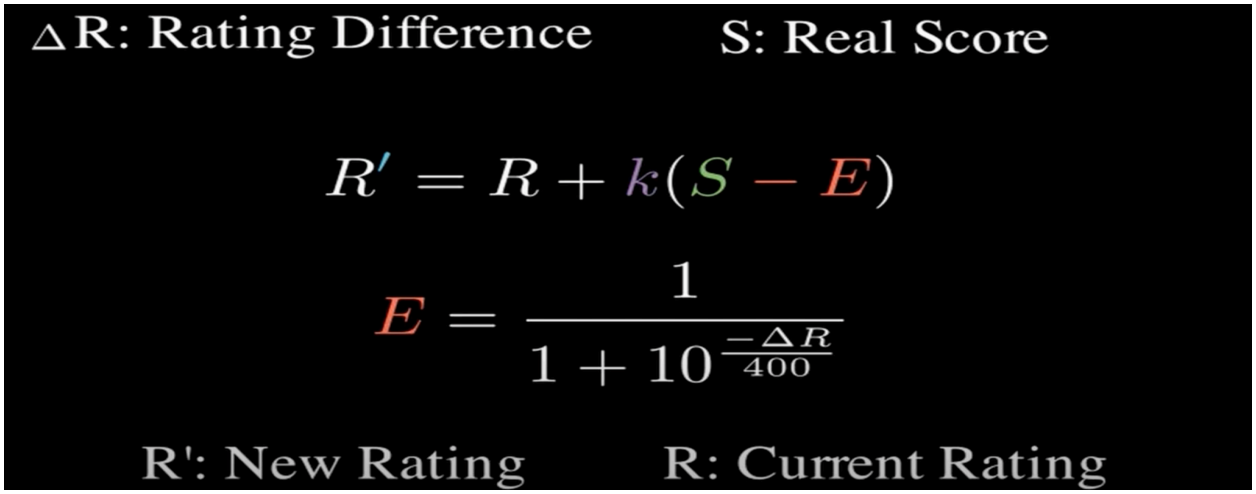
	Rating	Expected	1	2	3	4	5	6	7	8	9	10	11	Score	Change
Ralph	900	1/11	○	✗	✗	✗	✗	✗	○	✗	✗	✗	✗	2/11	▲
Mishi	1300	10/11	✗	○	○	○	○	○	✗	○	○	○	○	9/11	▼

8:21 / 19:28

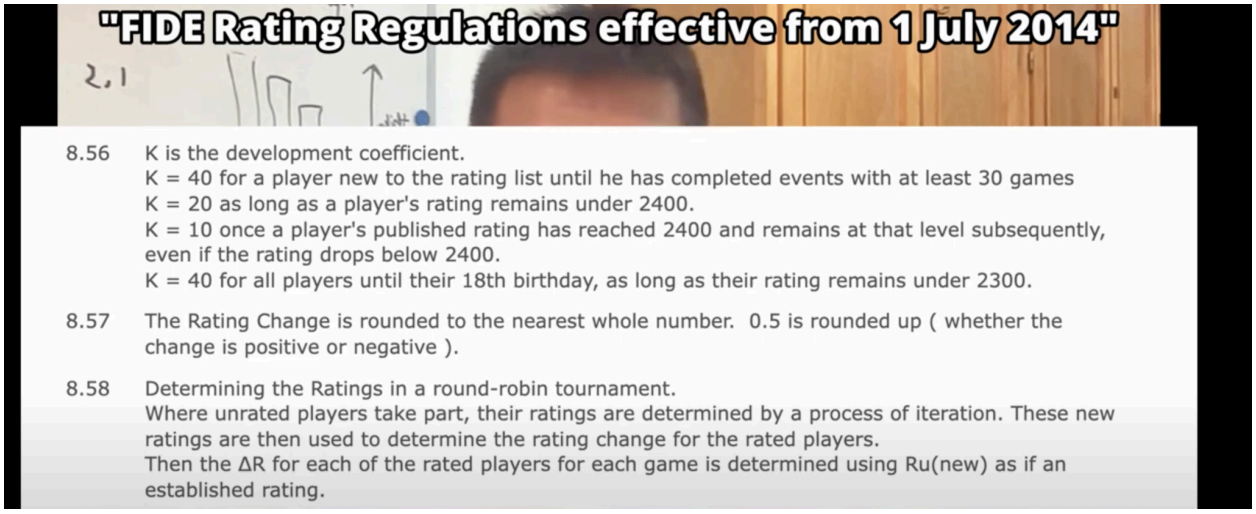
That means we can just use these numbers to calculate the change in eelo.



We can write that as an equation like this, where R prime is your new rating, R is your current rating, S is your real score, and E is your expected score. And, if you remember, we learned that you can calculate E via this equation, where Delta R is the difference between your eelo and your opponent's eelo.



The one variable that I haven't explained to you yet is this K, which ends up being the most variable part of different eelo implementations. K here is just some value that describes how much eelo you can gain or lose in a single match. So, if K is really, really small, you'll have to play a lot of games to get an accurate estimate of your score. If, on the other hand, K is really, really big, that means your eelo is going to change a lot, and so it's going to go up and down a lot, and be very noisy.



So, FIDE, the Chess Federation, for example, has K change between 10 and 40, depending on your current rating, how many games you've played, and your current age. Pokemon Showdown, on the other hand, has a K between 32 and 50, depending on your current eelo. But if your eelo is below 1,100, K is high if you win and low if you lose, so that more points go into it at lower levels. Age of Empires 2 has a constant value of K at 32 off the replacement gains. So, different games just do different things.

Can eelo be Negative or Zero

Now, let's discuss whether someone’s eelo can be zero or even negative.
Let’s say, a noob chess player has an eelo of 400 and is playing against a Grandmaster (2600). So, the difference in rating is 2800. Now, if we plug this in the equation we will get Expected Probability of win or Expected score of the Grandmaster will be 99.99% or almost 100%. Now, somehow the Grandmaster loses the game let see if the new rating of the Grandmaster can become negative or not.
So, we want the new rating to be negative. The Real Score is zero. Now if We plug this into the equation, we get the value of K to be more than 2800 which is the previous rating of the Grandmaster. Now, in actual games the value of K is never that high. For FIDE, the maximum value of K is 40. So, you’ve got your answer. eelo can not be negative or zero. The value of K is decided to restrict these cases.

Closing thoughts

Now, you can see a few times where we picked things completely arbitrarily. We got to decide from the start what numbers actually mean. You know, so we decided that a 400 eelo increase means you win 10 out of 11 times. We also decided on a particular distribution. You know, we could have chosen a normal distribution; we picked the logistic just because it was easier. And also, we decided how quickly eelos could change.
The main job of eelo is to compare players so that players can be matched against each other in tournaments. Although, currently there are many alternative rating system currently in the world, "eelo", world's favorite logistic-flavored rating system, is here to stay at this point. And hopefully now you can appreciate it as something a little deeper than just some arbitrary rules for updating numbers.
That’s it for today. Please Subscribe and stay tuned for next videos.