**Summary of TrueSkill: How it Works and How to Use it**

**IN PROGRESS**

References:

http://trueskill.org/

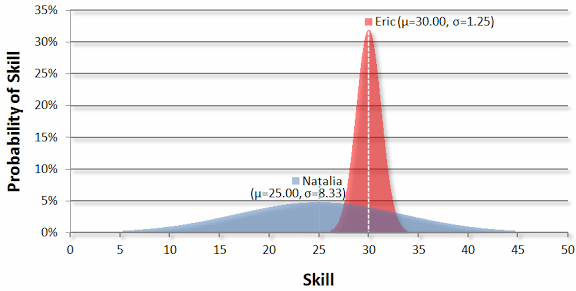
http://en.wikipedia.org/wiki/TrueSkill

http://research.microsoft.com/en-us/projects/trueskill/

http://www.moserware.com/2010/03/computing-your-skill.html

**How it works**

TrueSkills ranking system is based on two numbers: µ (mu) and (sigma). mu is a arbitrary value reflecting a players skill. Sigma is an arbitrary value reflecting TrueSkills confidence in the mu value. Together, it ranks a player. Loosly, mu is based on the win/loss ratio of a player, and sigma is based on the number of games played (although there is MUCH more to both numbers behind the scenes, which we do not need to know to implement TrueSkill on Pong Tracker)



Sigma is the confidence TrueSkill has in a players mu value, and has to do with probability. In the above figure, we have two players, Natalia and Eric. Natalia has a sigma value of 8.33 (which makes her graph wide). TrueSkill, in her case, is saying that her mu is 25. However, her high sigma value says that the probability of her actually having a mu of 25 is low, so her true skill could be in the range from 5 to 45 (to simplify the math), the likelihood becoming less the farther from mu.

Eric, on the other hand, has a low sigma. TrueSkill is saying that it has confidence that his actual true skill value, mu, is very close to 30, and could be in a range from about 26 to 34, the likelihood falling off very steeply the farther from mu.

These two players, with mu values close to each other, are actually ranked very far apart.

When two players play a game against each other, both values are updated. How much these values change depend on both players mu and sigma values. The amount of change in a players values depends on who defeats who, and how "surprising" the outcome is. A player with a high mu and low sigma value is not expected to defeat a player with a low mu and high sigma. If they did, it would be considered surprising, and the defeating players rank would see a large increase. Vice versa, an unsurprising defeat results in an almost negligible change.

**"TrueSkill" Number**

In order to rank players against each other, we need to have one number. Without going into the math too much, this number is calculated using the formula

The 3 in the equation is there because, as Microsoft has found, a players true rank is usually higher than what TrueSkill says. This is a rough way to compensate.

Therefore, with the standard starting values of mu = 25 and sigma = 8.33334 (sigma is an approximation),

**Implementation**

Each player starts off with a mu value of 25, and a sigma of 8.33334.

We start by giving each player a Rating using the mu and sigma values stored in the database:

r1 = Rating(); // player 1 Rating

r2 = Rating(); // player 2 Rating

r3 = Rating(); // player 3 Rating

r4 = Rating(); // player 4 Rating

Each Rating contains mu and sigma for a given player.

This then allows us to set up teams using lists:

t1 = [r1, r2]; // team 1

t2 = [r3, r4]; // team 2

When a game is complete and confirmed, new TrueSkill mu and sigma values are calculated for each player. To do this, we use the rate() function, and specify who defeated who using a ranks list:

(new\_r1, new\_r2), (new\_r3, new\_r4) = rate([t1, t2], ranks=[0, 1]);

// team 2 defeated team 1

In Pong Tracker, two teams can NOT draw, so ranks can either be [0, 1] or [1,0], where 0 is loss and 1 is win.

# Richard's Comments

I've been looking at the various resources listed on the TrueSkill page.

By far, this one <http://www.moserware.com/2010/03/computing-your-skill.html>

seems to be the most useful.

Skimmed through the "Math Behind TrueSkill" article. It seems to be about logical derivations (a.k.a. proofs) which won't really help us understand how to implement the library. (It explains 'why' it works, not 'how' it works)

So far, from what I understand, a player's TrueSkill is reported as being Mu - 3\*Sigma.

This is a very conservative estimate as for a Gaussian value there is more than a 99% chance

that a player's actual skill value will be greater than that.

I mentioned Confidence Intervals earlier. All that I mean by that is given Mu and Sigma,

theoretically there is a 95% chance that a player's actual skill value is in the interval

[Mu - 2\*Sigma, Mu + 2\*Sigma]

This is called a 95% Confidence Interval, and so a statistician may say things like "I have 95% confidence that a player's actual skill value is in that range of values"

Likewise, a 99% confidence interval would be

[Mu - 3\*Sigma, Mu + 3\*Sigma]

(This is true for Gaussians, other distributions will have different intervals.

See <http://en.wikipedia.org/wiki/68%E2%80%9395%E2%80%9399.7_rule> if you like pretty pictures)

Of course, if you want a 100% confidence interval, then that would be

[-Infinity, Infinity]

which is a nice example of a <http://tvtropes.org/pmwiki/pmwiki.php/Main/MathematiciansAnswer>