

NB: As the total filesize of my data and project code exceeds Blackboard’s upload limit, please note that it is publicly available on Github at URL <http://github.com/sverona/meleeWHR>. Replication instructions are available in the repository’s readme file.

1 Background and problem description

Super Smash Bros. Melee, a fighting game originally released in 2001, has recently experienced a resurgence in competitive play. Major tournament series such as Genesis and Evolution regularly attract upwards of 1,000 entrants [1], a scale otherwise unheard of for decade-old games incapable of online play.

Melee It On Me (*MIOM*), the competitive community’s primary governing body, produces annual ranking lists of the 100 most skilled players, established by panel vote. While the general accuracy of these lists is widely accepted, both the placement of individual players and the rankings’ general methodology are highly debated within the community at large. There have been many attempts at using existing rating systems such as Elo and Glicko to generate comparable rankings, but few are still publicly maintained, partially due to the scattered nature of the available data. Further, these rankings (like their MIOM counterparts) are produced at infrequent intervals from aggregated results. Thus, since the scene lacks a formal competitive circuit, the task of ranking the top players attending an upcoming tournament (e.g., for seeding purposes) falls mainly on that tournament’s organizers.

2 Objectives

The goals of this project were to

- i. consolidate the available tournament data, dating as far back as 2003-4, into a publicly maintained dataset;
- ii. use this dataset and a maximum-likelihood estimation method such as Whole-History Rating (*WHR*) [2] to reconstruct real-time ratings for the period comprising the earliest MLG tournaments in 2005-6 to the present day;
- iii. use D3.js [3] to visualize these time-series in a manner similar to [4].

3 Overview of methodology and research question

The primary question of interest is to derive a stable measure of player skill from the sequence of match records. One method of doing so that has been adopted by other e-sports communities is TrueSkill [5], which assumes that player skill before any given match follows a Gaussian distribution. The probability of one player defeating another is then roughly approximated by the probability that a randomly sampled value from the former’s distribution exceeds one from the latter’s; this gives rise to the *Bradley-Terry model*

$$P_t(i > j) \approx \frac{\mu_{it}}{\mu_{it} + \mu_{jt}},$$

where μ_{it}, μ_{jt} are the means of players i and j ’s distributions at time t . The TrueSkill algorithm, like many others, performs maximum-likelihood estimation of these parameters using this assumption. One goal of this project is to adapt this algorithm slightly to model some of the peculiarities of *SSBM*’s competitive environment, as demonstrated, e.g., in [6].

The WHR algorithm differs from TrueSkill in that it performs MLE over the player’s entire rating history (hence the name.)

4 Data cleaning

The data consists primarily of match metadata from tournament brackets, as shown in the sample below at left. This data was sourced from Liquipedia, using a scraper written in Python, and cleaned and reformatted into JSON, as in the sample below at right. It contains the match data from a set between top-level players at the tournament Shine 2017:

```
|l1m3p1=S2J |l1m3p1flag=us |l1m3p1score=3  
|l1m3p2=HugS |l1m3p2flag=us |l1m3p2score=1  
|l1m3win=1  
|l1m3p1char1=cf |l1m3p2char1=samus |l1m3p1stock1  
  ↳ =0 |l1m3p2stock1=1 |l1m3win1=2 |  
  ↳ l1m3stage1=Yoshi's Story  
|l1m3p1char2=cf |l1m3p2char2=samus |l1m3p1stock2  
  ↳ =1 |l1m3p2stock2=0 |l1m3win2=1 |  
  ↳ l1m3stage2=Pokémon Stadium  
|l1m3p1char3=cf |l1m3p2char3=samus |l1m3p1stock3  
  ↳ =2 |l1m3p2stock3=0 |l1m3win3=1 |  
  ↳ l1m3stage3=Yoshi's Story  
|l1m3p1char4=cf |l1m3p2char4=samus |l1m3p1stock4  
  ↳ =3 |l1m3p2stock4=0 |l1m3win4=1 |  
  ↳ l1m3stage4=Yoshi's Story  
|l1m3date=August 26, 2017  
|l1m3details={{BracketMatchDetails|reddit=|  
  ↳ comment=|vod=https://www.youtube.com/  
  ↳ watch?v=7zTSvNM-E1c}}
```

```
"l1m3": {  
  "date": "August 26, 2017",  
  "details": {  
    "comment": "",  
    "reddit": "",  
    "vod": "https://www.youtube.com/watch?  
      ↳ v=7zTSvNM-E1c"  
  },  
  "p1": "S2J",  
  "p1char1": "cf", "p1char2": "cf", "p1char3  
    ↳ ": "cf", "p1char4": "cf",  
  "p1flag": "us",  
  "p1score": "3",  
  "p1stock1": "0", "p1stock2": "1", "  
    ↳ p1stock3": "2", "p1stock4": "3",  
  
  "p2": "HugS",  
  "p2char1": "samus", "p2char2": "samus", "  
    ↳ p2char3": "samus", "p2char4": "  
    ↳ samus",  
  "p2flag": "us",  
  "p2score": "1",  
  "p2stock1": "1", "p2stock2": "0", "  
    ↳ p2stock3": "0", "p2stock4": "0",  
  
  "stage1": "Yoshi's Story", "stage2": "Pok\  
    ↳ u00e9mon Stadium", "stage3": "Yoshi  
    ↳ 's Story", "stage4": "Yoshi's Story  
    ↳ ",  
  
  "win": "1",  
  "win1": "2", "win2": "1", "win3": "1", "  
    ↳ win4": "1"  
}
```

5 Results

6 Directions for further improvement

References

- [1] https://www.ssbwiki.com/List_of_largest_Smash_tournaments
- [2] <https://www.remi-coulom.fr/WHR/WHR.pdf>
- [3] <https://d3js.org/>
- [4] <https://www.youtube.com/watch?v=z2DHPW79w0Y>
- [5] <https://papers.nips.cc/paper/3331-trueskill-through-time-revisiting-the-history-of-chess.pdf>
- [6] https://www.reddit.com/r/SSBM/comments/4pitia/an_objective_ranking_system_that_compensates_for/
- [7] http://liquipedia.net/smash/Main_Page
- [8] http://liquipedia.net/smash/index.php?title=Shine/2017/Melee/Singles_Bracket&action=edit§ion=4