Understanding Boss Battles: A Case Study of Cuphead

Andrew Wood, Computer Science

Mentor: Dr. Adam Summerville

Kellogg Honors College Capstone Project

Key Terms

Attacks:
Any object on screen that will damage the player if it collides with the player.

Bosses/Boss Battles:
A particularly powerful enemy in a videogame, usually with iconic and unique attacks, design, and whose encounter lasts longer than other enemies.

Platforming:
In video games, platforming is the general act of moving/jumping between "floor" objects, usually where falling would punish the player in some way.

SHMUP (shoot-em-up):
A SHMUP is a genre of action video game where the player usually moves on a 2D plane, shooting enemies that appear and move around the screen, usually firing at the player as well.

Carrying:
In multiplayer video games, carrying is when a player of high skill level performs well enough to compensate for the other player's struggles.

Abstract

"Bosses" — powerful, difficult enemies — have been a part of video games for the majority of their existence. Despite their integral role in many games, they have rarely been the focus of study. Cuphead is a run-and-gun style 2D action game with a large pool of boss battles, making it an ideal game for comparing and contrasting bosses in order to determine what makes a compelling (or frustrating) boss battle. In this case study, we developed an ontology of 'shoot-em-up'-genre attacks and bosses. With this system for codifying a boss battle, we cluster the attacks using a Gaussian Mixture Model — which are then used to represent a boss as a "bag-of-attacks". We then use multinomial regression to predict the player experience of a boss given the parameterized boss.

Approach

Player Experience

To analyze what makes a **Boss Battle** considered good or bad, we first needed real player feedback. We chose to focus on 6 traits (we call them *aesthetics*) that we determined may have an impact on a player's experience. We asked players through online survey to name and rate both their favorite and least favorite bosses on the 6 aesthetics listed below to measure what attributes made bosses favorable or unfavorable.

AESTHETICS

1. Platforming Difficulty
2. Recognizing Attack Patterns/Cues
3. Avoiding Attacks
4. Fight Duration
5. Ability to Hit Boss
6. Number of Attacks

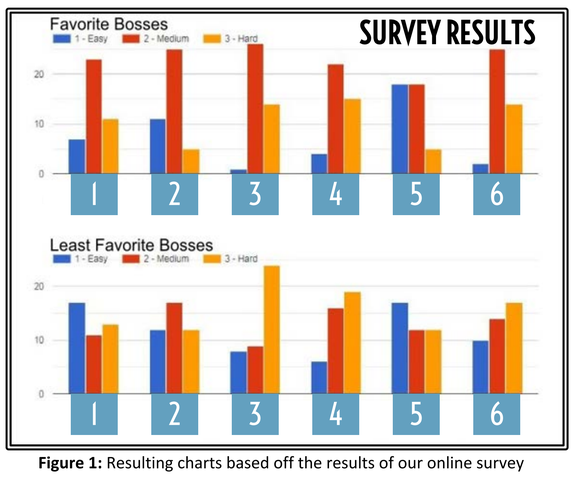

Figure 1: Resulting charts based off the results of our online survey

Ontology

We formulated a set of attributes that we decided would best capture the necessary information to determine how boss **attacks** affect player experience. The ontology is a set of attributes used to define an attack in a meaningful way. These attributes were selected because together they help define most of the characteristics of shmup-style attacks, thus limiting the use of our ontology to the **SHMUP** genre.

Value	Attribute
# px^2	Size
# px/second	Speed
# count	Concurrent Spawn Count
# count	Number of Cycles
0/1 binary	Infinite Spawn
# frames	Time Between Attack Waves
# frames	Attack Cue Duration
# frames	Lifespan
0/1 binary	On-Screen Until Dead
# count	Health
0/1 binary	Invincibility
0/1 binary	Parryable
0/1 one-hot * 4	Spawn Location

Table 1: Ontology of Attacks

Data Extraction

Data Source

We extracted data from YouTube videos of 4 boss encounters.

Extraction Tool

We developed a tool using Python for gathering static data about the bosses' attacks from the video, frame-by-frame (**Figure 2**). With the tool, we filled out the values defined in **Table 1** for each attack for each of the 4 bosses.

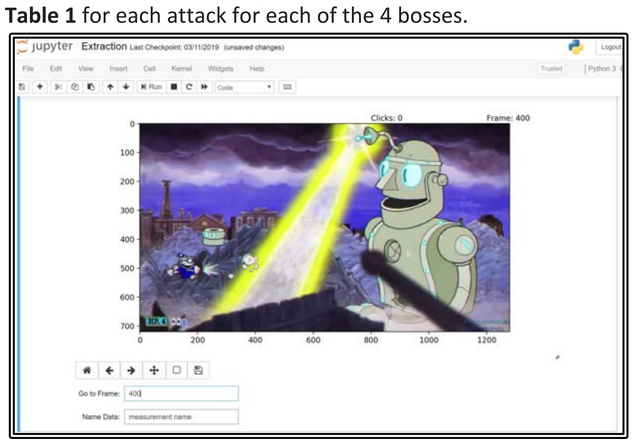

Figure 2: Screenshot of Data Extraction tool developed for gathering static data from screenshots

Clustering Attacks

With the data gathered to parameterize each of the attacks into a set of values, we used a Bayesian-Gaussian Mixture Model (BGMM) to cluster the attacks into groups, based on their various characteristics captured by the ontology. Our BGMM determined there to be 15 categories of attacks, each indicated by a different color in the below graph. We then treated each of the four bosses we examined as a "bag of attacks" defined by the categories they were clustered into.

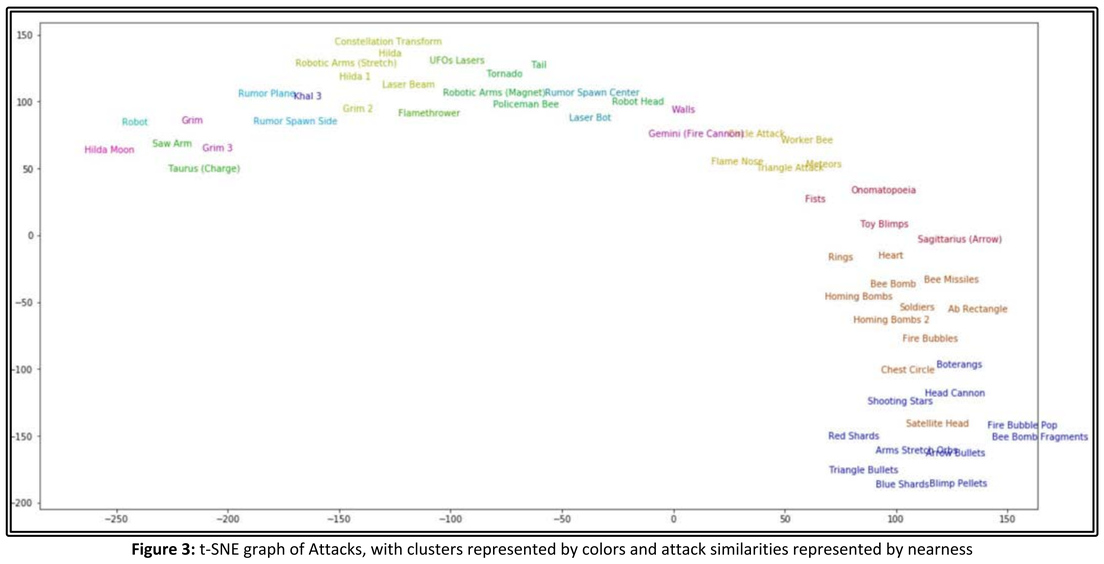

Figure 3: t-SNE graph of Attacks, with clusters represented by colors and attack similarities represented by nearness

Figure 3 above uses t-SNE (t-distributed Stochastic Neighbor Embedding) graphing in order to display "nearness and distance" between similar/disimilar points of data -- in this case, the attacks we measured.

Multinomial Regressions

With the 4 bosses represented as "bags of attacks", we ran a LASSO multinomial regression for each of the 6 aesthetics to create a model that would be used to correlate attack types and attributes with player experience ratings of the aesthetics. LASSO stands for Least Absolute Shrinkage and Selection Operator, a type of regression that allows the model to reduce the weights of input variables to near-zero, meaning that the regression determines which inputs (attack traits) have the most impact on the player experience (aesthetic ratings), and which have little-to-no correlation. **Table 2** below represents the resulting predictions of the 6 multinomial regressions for each of the 4 bosses. Predictions that matched the mode of the survey results for each bosses' aesthetic ratings are in bold, with incorrect predictions having the mode listed next to it in parentheses.

Table of Predictions for each Boss for each Regression

Boss \ Regression	Platforming Difficulty	Recognizing Attack Patterns/Cues	Avoiding Attacks	Fight Duration/Boss HP	Ability to Land Hits on boss (with any chosen attack)	Number of Simultaneous Boss Attacks
Rumor Honeybottoms	**Hard**	**Medium**	**Hard**	**Hard**	**Hard**	**Hard**
Hilda Berg	**Medium**	**Medium**	Hard (Medium)	Hard (Medium)	**Easy**	**Hard**
Grim Matchstick	**Hard**	**Medium**	Hard (Medium)	Hard (Medium)	**Medium**	**Medium**
Dr. Khal's Robot	**Medium**	**Medium**	**Hard**	**Hard**	**Hard**	**Hard**

Table 2: The six regression predictions for each of the 4 bosses aesthetic ratings. Bold text indicates a prediction matching the mode of the survey results for that boss' aesthetic ratings.

The models correctly predicted 20/24 ratings, an accuracy of 83.33%. All incorrect predictions were for the aesthetics "Avoiding Attacks" and "Fight Duration", and for only two of the bosses -- *Hilda Berg* and *Grimm Matchstick*. This indicates that the ontology is not capturing enough information to properly correlate boss features with those two aesthetics. However, the rest of the regressions performed fairly well.

Analysis and Conclusion

There are various improvements could be made to this study. For example, we realized after we gathered our boss data that there may have been more value in capturing the attributes "width" and "height" instead of "area". This is because an attack may span 80% of the screen vertically, which would make it relatively difficult to dodge if moving horizontally, but it may have low total area because of a small width. Therefore, "area" as a value doesn't weigh in to "Avoiding Attacks" as much as we believe "width" and "height" may have. Also, despite already having an 83% accuracy, we believe the regressions would perform even better with more data. The manual data extraction process was extremely tedius, and we believe having more data would have improved the accuracy of the regression models even more.

Over all, we see the results of this study as promising, and see value in continued research towards boss battles in video games. In the future, we could extract data for a 5th boss and predict the aesthetic ratings without feeding that boss' attributes and ratings into the models themselves. That way, we could potentially see how the regressions perform without input from the boss directly.

Though this research has somewhat conclusive results, it is ultimately a prediction based on correlations. There are many factors that may contribute to what makes a boss battle enjoyable or frustrating, and we simply cannot capture all the related factors. A player's impression is an opinion, and that is impossible to know based on boss data alone. However, we certainly see value in studying boss battles in this static form, and are hopeful for the future of this type of research.

Acknowledgements

Thank you to Dr. Adam Summerville, who **carried** me through this project. I would not have been able to complete this without his help and passion.