Proceedings of the Seventeenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment
(AIIDE 2021)

A Case Study in Learning in Metagames: Super Smash Bros. Melee

Julien Codsi,[1] **Adrian Vetta**[2]
[1]Université de Montréal
[2]McGill University
julien.codsi@umontreal.ca adrian.vetta@mcgill.ca

Abstract

Imagine agents repeatedly playing a bimatrix game against opponents drawn from a population of assorted skill levels. This paper studies how agents strategize in such a metagame and the population distributions that result. Specifically, we investigate how an agent should adjust its strategy as it also learns to play the game, that is, as the agent improves its skills (from novice to expert) with repeated exposure to the game. To perform this task, we introduce a dynamic game-theoretic model of learning in metagames. We use it to explain the learning dynamics and character selection exhibited in data from the game Super Smash Bros. Melee. Indeed, the primary motivation behind this work is the application of game theoretic methods in video game balancing.

1 Introduction

The video games market is worth over 150 billions dollars per annum. Of interest here is the significant class of the multiplayer competitive games, a sector growing rapidly at both the general consumer level and the elite level. An illustration of this lies in the increasing popularity of e-sport events with respect to both the number of competitors and the number of viewers. For example, 60 million unique viewers watched the 2018 Mid-Season Invitational tournament, a *League of Legends* tournament.

A characteristic of e-sport events, such as fighting game competitions, is that the technical skill of a participant is not the only determinant of success: the success-rates of even the best players are greatly influenced by factors outside of the actual game itself. Indeed, the most important determinant is the exogenous choice of character. Consider the following two simple facts. One, a character may have a inherent advantage in a fight against another character. Two, some characters are intrinsically more popular than other characters. Together, these imply that, in order to have a chance of winning a tournament, the character selected by a player must have abilities that matchup favourably against the most popular characters. Such community-driven effects form what players call the *metagame*.

More formally, the term metagame in game theory was introduced by Howard (1971) in the study of arms control, where the purpose of the metagame was to endow the agents with more contextualized strategies extending beyond those implied simply by a *Prisoner's Dilemma* bimatrix. A standard modern interpretation is given in (Tekinbas and Zimmerman 2003) which defines a *metagame* as "the relationship between the game and outside elements, including everything from player attitudes and play styles to social reputations and social contexts in which the game is played"; see also (Boluk and Lemieux 2017).

From our perspective the most pertinent aspect of the metagame is that of character selection in an environment where agents repeatedly play a game against different opponents. This motivates the question: how should an agent learn to strategize optimally whilst, at the same time, learning to play and master the game ("learning-by-doing")? In this paper, we formulate a game theoretic model to study the dynamics of this learning process.

In addition to fighting games, our methodology gives insights into how agents may learn to play and strategize in more general games played by populations. Furthermore, for the specific application of gaming, our model has direct implications to the important practical issue of *game balancing*. As alluded to above, an essential aspect in creating a playable and enjoyable game is that the characters matchup well with each other – specifically, a wide range of characters should be competitive and certainly no character should be "*OP*" (*overpowered*). Moreover, this balance must apply at all skill levels including hobbyists and expert players. Furthermore, this does not just concern pair-wise character balance but also the more complex issue of balance at a population level. Indeed huge efforts are made by game developers before and after game release to correct imbalances. Interestingly, we will see that our model provides a systematic way to diagnose imbalances and to test the effectiveness of potential balance corrections, both before and after game release.

Overview of Paper We introduce a game-theoretic model of dynamic learning in metagames and test it empirically using data from the very popular game *Super Smash Bros. Melee*. Our work builds upon the ground-breaking model of Jaffe (2013); see also Jaffe et al. (2012), which is essentially a static model that studies strategy selection for expert game players. The format of the paper is as follows. In Section 2, we present the fighting game model of Jaffe (2013).

In Section 3, we apply his model using real data from Super Smash Bros. Melee. In doing so, we will see the necessity of formulating a dynamic model incorporating both time and variations in skill-levels. In Section 4 we build such a model by incorporate learning curves into the strategy selection dynamics. We then show how this model can explain strategic dynamics seen in the Super Smash Bros. Melee data and describe how it can be used in metagame balancing.

Related Work The dynamics of learning in metagames broadly relates to several fundamental areas of game theory. The concept of multiple agents playing a bimatrix game bears close resemblance to the field of evolutionary game theory instigated by Maynard Smith (Maynard Smith and Price 1973; Maynard Smith 1974, 1982), and the convergence of population distributions in animals to equilibria. There, rather than rationality, the primary driver of strategy (population) change is evolutionary fitness, and rather than a Nash equilibrium the standard stability concept is an evolutionary stable strategy (Bishop and Cannings 1976; Hines 1987; Apaloo et al. 2015). Replicator and other dynamics (Taylor and Jonker 1978; Schuster and Sigmund 1983; Samuelson 1988) have been studied to model the underlying convergence processes. The ideas from evolutionary game theory have also directly impacted computer science in the form of genetic algorithms (Goldberg 1989; Mitchell 1998). In addition to evolutionary game theory, the field of mean field games (Lasry, Lions, and Gueant 2010) also studies equilibria produced by populations, from an asymptotic perspective.

Learning in game theory is becoming of fundamental importance (Fudenberg and Levine 1998; Cesa-Bianchi and Lugosi 2006). The learning of equilibria dates back to the pioneering work of Brown and Von-Neumann (1950), Brown (1951) and Robinson (1951). Most notably the fictitious play method of Brown (1951) is the precursor to the ubiquitous multiplicative weights algorithm in machine learning (Grigoriadis and Khachiyan 1995; Arora, Hazan, and Kale 2012).

On the practical and experimental side, game-theoretic concepts have been widely studied in sports, not just e-sports! Of particular relevance is whether professional athletics use equilibrium strategies either as a group or individually. Examples include penalty kicks in soccer and serves in tennis (Chiappori, Levitt, and Groseclose 2002; Palacios-Huerta 2003; Walker and Wooders 2001; Yee, Rodríguez, and M. 2014).

We remark that there is a fundamental distinction between these works and our study of metagames. In the aforementioned areas, the primary task is to learn an equilibrium of the game, in particular, optimal strategy selection. For example, how should expert soccer players take and save penalty kicks against each other. In contrast, we have two objectives. In addition to learning optimal strategy selection, we are interested in how the agents learn to effectively use these strategies in the first place ("learning-by-doing"). Specifically, we model the dynamics arising as agents learn to play a game (from novice to expert) when repeatedly competing against a population of players.

2 A Game Theoretic Metagame Model

In innovative work, Jaffe (2013) presented a game theoretic model for the metagame of a competitive fighting game. This framework has subsequently become influential in game design; in particular, it proffers tools by which to *balance* a game by predicting how a community of comparable players (experts) should stabilize based on the relative strength of the characters. In this section, we present Jaffe's model. As we will see, his model is essentially a static model. The aim of this paper is to design a dynamic model that more accurately captures the evolution of a metagame, specifically, how the players learn and adapt over time. Our model is very general as it abstracts the complexity of a fighting game into a simple mathematical formulation while still being powerful enough to make predictions. In addition, our model is not game specific and can be applied beyond competitive gaming to more general applications where populations learn to play "games".

Character Selection and Nash Equilibria The most fundamental decision in a fighting game is character selection. Note this decision is made before a game begins and is thus part of the metagame. A natural way to model a fighting game is then to only consider the choice of character of both players and to consider their payoff as the *winning probability* of each character in the matchup. These probabilities depend on the options available to each character and how well they counter each other. These probabilities are commonly collected by the communities of players and presented in *matchup charts*. These charts are simply matrices containing information about each matchup. Moreover, these matrices induce payoff matrices for the game. Observe that the game is symmetric because both players in a fight choose from the same set of characters. The game is *constant-sum* (the winning probabilities always sum to one) and can be represented by a single payoff matrix M given by the matchup chart. Specifically, when x and y are the chosen strategies of players 1 and 2, respectively, the payoff functions are $\Pi_1(x,y) = x^T M y$ and $\Pi_2(x,y) = y^T M x = 1 - \Pi_1(x,y)$. We can then find a Nash equilibrium via the following linear program (LP):

$$\begin{aligned} \max\ \alpha \qquad s.t. \qquad \sum_r x_r &= 1 \\ x_r &\geq 0 \qquad \forall r \\ \sum_r M_{rc} \cdot x_r &\geq \alpha \qquad \forall c \end{aligned} \tag{1}$$

Because the game is symmetric, it must be that both players have a 50% win rate at the equilibrium. Furthermore, there exists a Nash equilibrium where both players play exactly the same strategy. An immediate objection to this approach is that the Nash equilibrium will typically be a mixed strategy and this is *not* representative of the reality in which players essentially use pure strategies. In particular, game players mostly practice a single character for years to learn all the intricacies of their "*main*". This practical observation is incompatible with mixing over a collection of characters. We remark that this observation, that a character takes time to master, has an important consequence: a player cannot be expected to obtain the win rates indicated in the matchup chart as soon as he starts playing a character. We will study in detail the implication of this, namely the learning process of the character, in later sections. For now we return to the issue of mixed Nash equilibria. We claim the Nash equilibrium of this game is meaningful not in analyzing a single two-player fight but rather in analyzing the entire *community of players*. Indeed, imagine the formation of a community of players as an iterative process where each player chooses a pure strategy (a single character) which maximizes his win rate against the pre-existing community and is then locked to this character from this point on. This process exactly simulates the classical *fictitious play method* for finding a Nash equilibrium introduced by Brown (1951). In this setup, a Nash equilibrium corresponds to an absorbent stationary point of the usage proportion of each character. Indeed, Robinson (1951) proved that this process always converges to a Nash equilibrium in finite constant-sum two-player games such as ours. This suggests the community distribution of players should resemble the Nash equilibrium found by LP (1).

Approximate Nash equilibria Naturally, it is unrealistic to expect the community distribution to mimic exactly the Nash equilibrium. For example, the matchup charts themselves cannot be taken as absolute truth: the matchup charts extracted from data are often noisy and the matchup charts issued by experts are approximations as they arise from qualitative discussions. As a result, it would be preferable to not only obtain the exact Nash equilibrium induced by the matchup chart, but also approximate equilibrium as well. For example, a strategy that allows a player to achieve a win rate of, say, 49.5% can be considered as potentially a realistic equilibrium of the game.

In particular, as described by Jaffe (2013), this viewpoint allows us to obtain constraints upon any approximately stable community. To wit, suppose we want to know the frequencies at which each character can appear in a stable community. It is then easy to design a linear program that can be used to determine the interval within which a character's frequency must appear if a strategy is to reach a targeted win rate. Specifically, to test if player 1 can successfully play character i with probability f_i we add the constraint $x_i = f_i$ into LP (1). Of course, previously a winning rate of $\alpha = 0.5$ was achievable. Now, by imposing the additional condition $x_i = f_i$ we restrain the strategies available for the player 1 and thus his win rate will necessarily be at most 0.5. The key point however is that this expands the set of strategies that are consider stable (provided their use is within a bounded frequency range) and allows for noise in the matchup charts and for small errors in character selection by the players. We remark that, throughout the paper, we will consider a win rate of $(50 \pm 0.5)\%$ as a stable win rate.

3 A Case Study: Super Smash Bros. Melee

As a case study, we consider *Super Smash Bros. Melee*, a game with several desirable properties. First, Super Smash Bros. Melee is a fighting game with one of the biggest communities, based on the number of tournament viewers. Second, the game is twenty years old. Consequently, there has been plenty of time for the players to learn and for the community to stabilize. Thirdly, although originally an offline game, last year a community-made online version of the game was developed. The database for the online game already contains a large number of matches, over 1.3 million. The model we develop is tested using this database.

To begin, we calculate the character distribution of the Melee community empirically. This distribution of the 26 characters is shown in Figure 1. Next, we find the empirical

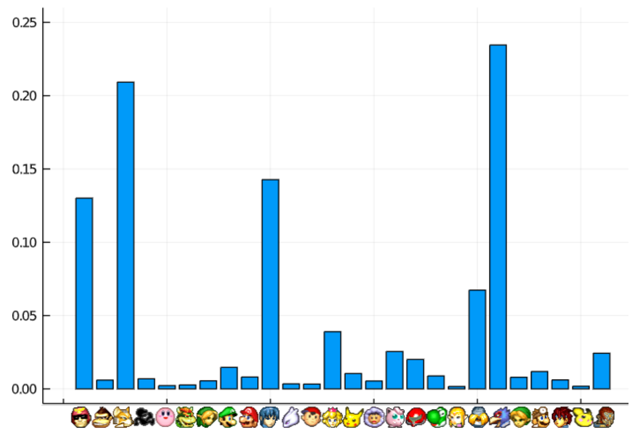

Figure 1: The Community Distribution

matchup chart using the 1.3 million matches. Figure 2 illustrates the resultant matchup chart using a heatmap. Here, the bluer the square the higher the probability the row player wins the fight; the redder the square the higher the probability the column player wins.

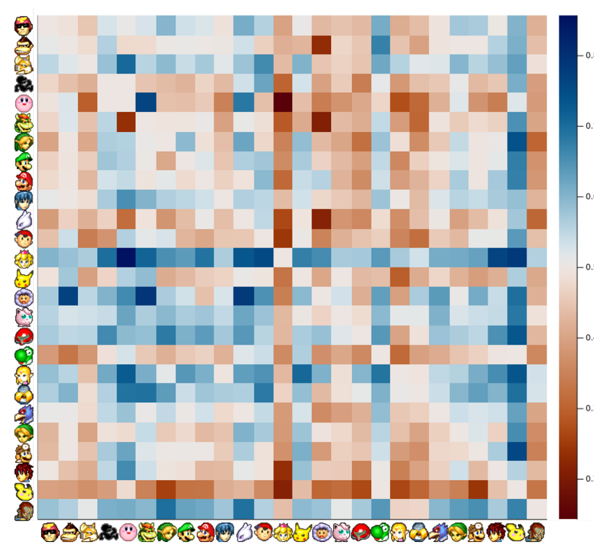

Figure 2: The Empirical Matchup Chart

The matchup chart immediately yields several interesting observations. Note that some characters are clearly very poor, for example, *Pichu* (Character #25). Thus, we should not expect Pichu to be played much in the Nash equilibrium. Indeed, we can see this in the community distribution of Figure 1. More remarkably, *Peach* (Character #13) wins every possible matchup; she has a greater than 50% win rate against every other character! But this means that if we find the Nash equilibrium using LP (1) for the empirical matchup chart it must be a pure strategy with Peach. In particular, we expect Peach to dominate the metagame and this is empirically not the case as shown by the community distribution. Nor does Peach dominate in higher level competitions of expert players.

This appears to render the game theoretic model of the metagame invalid. Let's investigate why this is not the case.

Data Filtering Despite the aforementioned advantages of Super Smash Bros. Melee as a case study, it does have some drawbacks. In particular, the online game consists predominantly of fights between players without regard to skill level. This creates noise in the data in two ways. First, in reality the *matchup chart* changes according to the general skill level of the participating players. Of particular relevance here is that *Peach* is known to be an excellent character choice among low-level players. This is primarily because her *down-smash* attack is hard to counter but easy to learn and apply in a versatile manner. Now, relating to the Pareto principle, a large portion of the players in Melee are not strong and therefore add a lot of noise to the data (Newman 2005). Second, a good player can beat a bad player regardless of their character matchup. Together, this means that the empirical matchup chart of Figure 2 does not accurately represent the actual chance of victory in any specific game.

To address these issues, we filtered the database to capture only good players. Specifically, if we consider only players with at least a 50% win rate the corresponding filtered community distribution is shown in Figure 3. This, in turn, in-

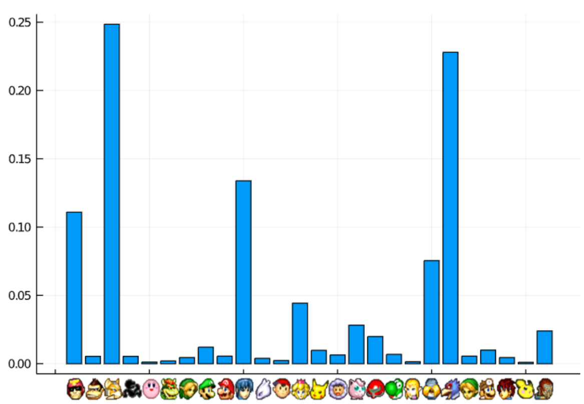

Figure 3: The Filtered Community

duces a matchup chart that is based only on games in which both players are good. This reduces the number of fights from 1.3 million to 150,000. The resultant matchup chart is shown in Figure 4. However, in the filtered matchup chart Peach still appears dominant. To verify this, we must calculate the stable frequency intervals for each character. The results are shown in Figure 5 (a distribution is considered vi-

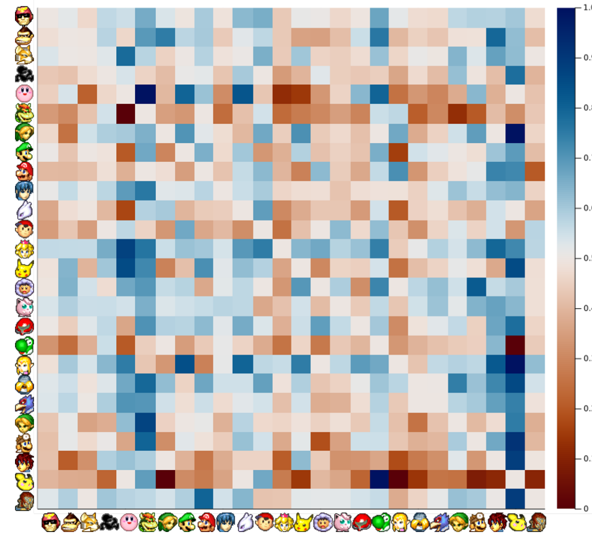

Figure 4: The Filtered Empirical Matchup Chart

able if it results in a stable win rate). Ergo, Peach dominates

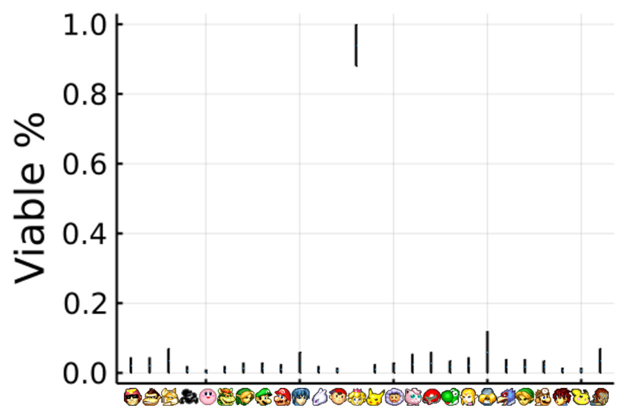

Figure 5: Viable Frequency Intervals for Stable Distributions using the Filtered Empirical Matchup Chart.

even in any approximate equilibrium for the metagame – she must be played at least 88% of the time in any approximate equilibrium! Thus, the same problem arises and the standard model of Jaffe is still unable to explain the community distribution.

Expert Advice Game balance is extremely important for all player standards, but is arguably most important at the novice level to encourage participation and at the expert level to drive community and competition interest. Furthermore, based upon their collective experience, a committee of expert players have created a *theoretical matchup table*. This table gives the matchup chart shown in Figure 6. Of course, we can now calculate a Nash equilibrium using this matchup chart. But now the Nash equilibrium is the pure strategy of playing Fox (#3). Again, this bears absolutely no resemblance to the community distribution. On the other hand, let's consider approximate equilibria. Using the theo-

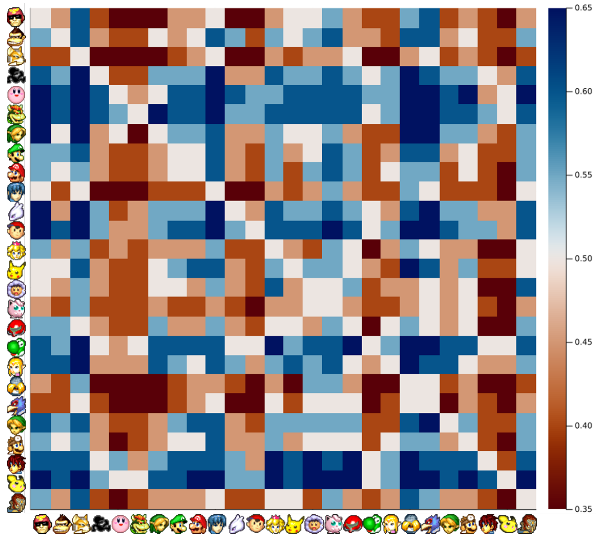

Figure 6: Theoretical Matchup Chart

retical matchup chart, the stable frequency intervals for each character are then:

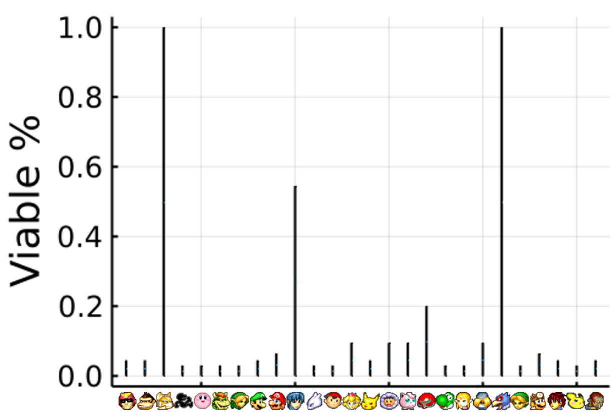

Figure 7: Viable Distributions with Theoretical Matchups

Observe that the frequency intervals shown in Figure 7 are very promising. Comparing with the empirical community distribution of Figure 3, we see that for 25 of the 26 characters their empirical usage lies within their stable frequency interval. Thus, by incorporating expert advice, we have, to a degree, successfully captured the general distribution of the community. We remark that the one character whose empirical usage lies outside its stable frequency interval is *Captain Falcon* (#1). Ability-wise, Captain Falcon is extremely fast with spectacular and fun to perform combos. Indeed, he is the most popular character to watch in e-competitions even through only one major tournament has been won by this character in the last decade. Consequently, it is also unsurprising that his usage, in practice, goes beyond that which could be expected based upon win-rate alone.

Still a degree of scepticism is justified here. We have seen that low skilled plays should select Peach and high skill players should select Fox. It is easy to verify this formally by calculating the best response of a player using the filtered empirical (low level) and theoretical (high level) matchup charts, respectively. How and when does this transition arise? Our study of approximate best responses, whilst providing a possible explanation of the community distribution does not answer this; it was based solely on the theoretical matchup charts of expert players and so provides no understanding of the community dynamics as players learn. We address this important issue by incorporating learning into the game theoretic model of the metagame.

4 A Learning Model

Our discussion so far implies that novices should take one action (namely, select Peach) and experts should take a different action (namely, select Fox). Implicit in this conclusion is that players learn. However, this argument relies on a neat division of players into novices and experts. Considering only those two levels is essentially equivalent to representing player skill as a bi-level step function. This of course is entirely unrealistic. But the fundamental idea of a transition from novice to expert is realistic. However, this is more likely to be a smooth progression based on the time invested playing a character – that is, the players *learn* to play the game!

Learning Curves Consequently, we now develop a learning model of metagames. Our first task is to incorporate learning curves. There are two structural properties these learning curves should possess. First, practical research on the learning of skills concludes that typically learning curves exhibit *diminishing returns*. This concept is very well captured by the *power law of practice* which has been studied in depth (Lee and Kirlik 2013; Crossman 1959). In particular, the skill progression obtained with a character on the one-thousandth game hour is not the same as in the opening hour. Second, different characters have different learning curves. Specifically, some characters are easier to master than others inducing steeper learning curves and lower skill barriers.

We can now model this phenomenon. Under the basic assumption that each player wishes to win as much as possible after playing the game for a fixed amount of time, say t thousand hours, the utility of our player is then: $U_1(x, y) = \sum_i Lr_i(x)(My)_i$. Here Lr_i is the *learning curve* of the character i, x is the *proportion* distribution of the t hours spent on each character, and y is the actual population distribution of the community. Unfortunately, the natural optimization problem induced by this choice is a non-linear program (NLP) because the objective function is not linear. To deal with this we transform the NLP into a mixed-integer linear program (MILP) which approximates it. Specifically, because the objective function is separable over its variables, we can simply approximate each Lr_i using univariate piecewise linear functions. To complete the learning model we incorporate our linearization for every character in the game.

Implementation Due to space constraints the technical details of our MILP are deferred to the full paper which can be found on arXiv. (The complete implementation of our model is also on github at https://github.com/Codsilla/MeleeMetagame). Importantly, the MILP can be solved almost instantly even on large data-sets. For our purposes, there is, however, a basic flaw in the Melee online data-set. As discussed, most of the players have up to 20 years experience in the offline version of the game. Thus it is impossible to deduce from the data-set the experience level of any specific player. But this information, is required to estimate the learning curves as described in the previous section.

In practice, we remark that game developers can easily obtain estimates of these curves even before game release through extensive game testing. Moreover, again applying expert advice, we can demonstrate proof of concept. Specifically, using learning curves estimated from community beliefs, our model illuminates how the transition from a low skill character (Peach) to high skill character (Fox) happens. So what are our estimated character learning curves? Based on the literature concerning the learning of learning curves (Leibowitz et al. 2010), it is natural to predict the learning curves are sigmoid functions. Specifically, let σ be a sigmoid function $\sigma(x) = (1 + e^{-x})^{-1}$. Now we know Peach is an easy character to learn. So we will give Peach a steeper learning curve than the other characters. In particular, for each character i except Peach we use the learning curve: $Lr_i(x) = Lr(x) = Lr = 2\sigma(3xt) - 1$. For the character Peach we use the learning curve. $Lr_{Peach}(x) = 2\sigma(6xt) - 1$. Thus, to model the low skill barrier in playing Peach, we simply use a learning rate that is twice as fast as the other characters. That is, one hour of practice with Peach is equivalent to two hours practice with another character.

We remark that the choice of the constant 3 in $Lr(x)$ is irrelevant in principle; its selection is simply to ensure that 1 unit of time, corresponding to 1000 training hours, is sufficient so that a player may learn enough to reach a very good (but not expert) level. In addition, for our learning model, we must linearize the learning curves. To do this, we used a dedicated package (Codsi, Gendron, and Ngueveu 2021) to enforce a maximum relative error of 1% between the linear-piecewise approximation and the learning curve. As relative errors are preserved through error propagation in our model, this suffices to ensure output solutions that are also within at most 1% of the optimal solution.

Our Results We can now present our results. Specifically, for the task of character selection, how should a player allocate a fixed amount of training time? Moreover, how does the choice of character selection vary as the player evolves from a novice all to way up to an expert? We can quantify this by making the assumption that a novice will have only a small amount of training time, an expert a large amount of training time, and intermediate players assorted amounts of time between the corresponding lower and upper bounds.

To do this we consider players with a minimum of 100 hours available training time (novice) up to a maximum of 15,000 available training time (expert). Using, our MILP model, we then found the optimal character selection strategies that should be used given the empirical matchup table.

The results, shown in Figure 8 are illuminating. It exhibits a natural transition from the prescription that a novice (Fig.8(a) 100 and (b) 250 hours) should only play Peach (#13) to the prescription that an expert (Fig.8(h) 15,000 hours) should only play Fox (#3). Even more interesting is the evolution we see between these two extremes. As the number of available training hours increases new characters begin to appear in the optimal strategy selection. For example, Fox and Falco (#21) should be used during 500 hours training (Fig.8(c)) and Jigglypuff (#16) should be used during 750 hours training (Fig.8(d)), etc. Moreover, as can be seen, at intermediate skill levels, the optimal strategy is not to play a single character as at the novice and expert levels. Rather, during the learning transition, it is optimal to train on multiple characters.

Most notable is Figure 8(e) with 1000 training hours. Compare this with the filtered community distribution shown in Figure 3. In particular, we see at 1000 hours the appearance in the optimal character selection of six (#3, #10, #13, #16, #20, #21) of the top-7 characters in the filtered community distribution. This comparison group is fair; recall the filtered group consists of players with at least a 50% win rate, that is, a mix of good players and above, whom one would expect have invested a reasonable amount, such as 1000 hours, of training time in the game.

So we have a near perfect match of characters. But what about the missing top-7 character. For the aforementioned reasons, the reader will not be surprised to see that this character is again Captain Falcon (#1). Indeed, our model predicts that Captain Falcon should appear slightly later. Specifically, it shows that Captain Falcon should be the ninth character to appear, as shown in Figure 8(f) at 1500 hours.

Recall, for these results, we have made only one major assumption: Peach has a fast learning rate. We have made no distinguishing assumptions for the other 25 characters. With just this simple assumption on *one* character, we obtain a plausible explanation of the evolution of strategy selection and of the population distribution across *all* characters in the metagame.

We do not claim this is an exact representation of reality. For example, our analysis explains how a player may most productively allocate time whilst learning to master the game and accurately predicts which characters are the most important. But the probability distribution for the characters generated in the model do not exactly match the true population distribution. Of course, this is unsurprising given we have made only one distinguishing learning assumption (for Peach). In the future we hope the database is enhanced to allow each learning curves to be learned directly from the data using the methods described. If so, we anticipate even stronger results may be obtainable.

Interestingly, the recommendation, inherent in Figure 8, that intermediate level players partition their training time among a group of characters counters the conventional wisdom that players focus solely on their main. This recommendation is reasonable when the group is reasonably small. But when the group is large, as in Figure 8(g), the recommendation is somewhat unrealistic. This is because additional factors (such as attachment to favourite characters, aversion to change, unmodelled fixed time-costs incurred in learning a new character, etc.) mean that one can only expect a player

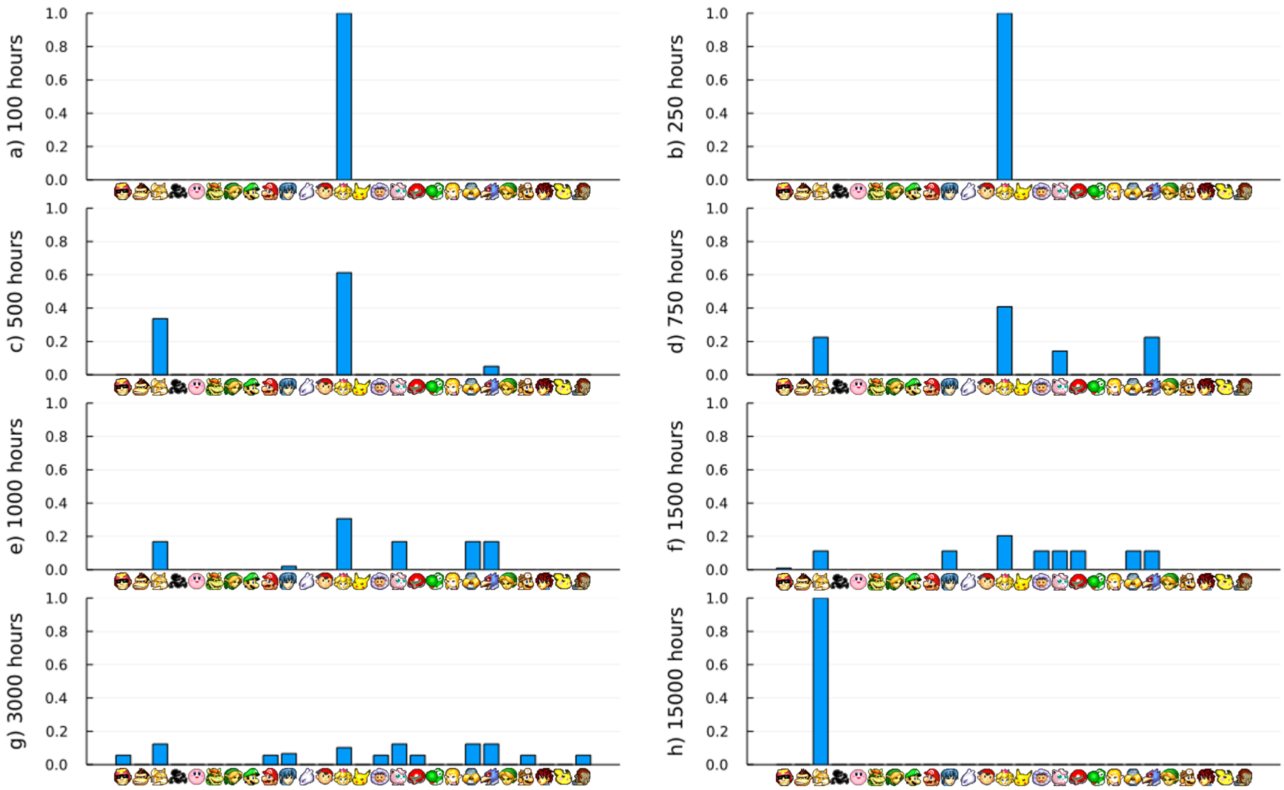

Figure 8: Evolution of the playtime distribution

to change their character during training a small number of times. On the other hand, applying a classical argument we may interpret a recommendation to play a group of many characters during training as being a recommendation made to a large number of players on entering the game. Assuming each of these players actually only selects a small number (say, 3 or 4) of the recommended characters to use while training, we see a natural way in which the relevant character distributions may arise. (Such an argument is analogous to that of evolutionary game theory, where a mixed strategy is induced by populations of agents playing only pure strategies.)

As stated, we estimated the learning curve on Peach based upon expert advice. There is one major benefit that arises from the approach taken here. Specifically, our results show that game balancing can be studied methodically *before* game release *without* an exact learning curve for each character. Theoretical matchup charts and conjectured learning curves created during testing by experts and game designers can successfully be used to predict and correct imbalances!

Finally, our model omits other learning effects that may be relevant in practice, such as group learning effects. For example, in training with Kirby (#5), say, one learns general games skills that will increase the learning speed with other characters. Similarly, training with Kirby (#5) will not only improve your win-rate with Kirby but also likely improve your win-rate in fights against Kirby. Incorporating such correlation effects into the learning model is straightforward in principle.

5 Conclusion

The model presented in this paper gives promising results on how game-theoretic techniques with learning can be used to make accurate short to long-term predictions for the metagame of a fighting game. This is particularly important in the context of game developers designing *balanced* games, where every character is viable to some extent.

Furthermore, in the near future, the additional feature of *rank-mode* will be made available for Super Smash Bros. Melee online. This will allow for less noisy data, as it will induce a player skill metric, and induce better estimates of the learning curves. In the long term, the most beneficial enhancement to the database would be a feature to track new players of Super Smash Bros. Melee. This would produce data that can be used to calculate the learning curves for each character precisely using the methodology described in the paper.

Acknowledgements

We are extremely grateful to Scott Norton for permission to use the Slippi Stats Online database.We thank Vlad Firoiu, Alex Jaffe and Clark Verbrugge for comments and advice, and Jonathan Dumas for introducing us to the topic of game balancing. Finally, we are grateful to the reviewers for detailed suggestions that much improved this paper.

References

Apaloo, J.; Brown, J.; McNickle, G.; Vincent, T.; and Vincent, T. 2015. ESS versus Nash: solving evolutionary games. *Evolutionary Ecology Research* 16(4): 293–314.

Arora, S.; Hazan, E.; and Kale, S. 2012. The Multiplicative Weights Update Method: A Meta-Algorithm and Applications. *Theory of Computing* 8: 121–164.

Bishop, T.; and Cannings, C. 1976. Models of animal conflict. *Advances in Applied Probability* 8(4): 616–621.

Boluk, S.; and Lemieux, P. 2017. *Metagaming: Playing, Competing, Spectating, Cheating, Trading, Making, and Breaking Videogames*. University of Minnesota Press.

Brown, G. 1951. Iterative solution of games by fictitious play. *Activity analysis of production and allocation* 13(1): 374–376.

Brown, G.; and Von-Neumann, J. 1950. Solutions of games by differential equations. *Annals of Mathematics Studies* 24: 73–79.

Cesa-Bianchi, N.; and Lugosi, G. 2006. *Prediction, Learning, and Games*. Cambridge University Press.

Chiappori, P.; Levitt, S.; and Groseclose, T. 2002. Testing Mixed-Strategy Equilibria When Players Are Heterogeneous: The Case of Penalty Kicks in Soccer. *American Economic Review* 92(4): 1138–1151.

Codsi, J.; Gendron, B.; and Ngueveu, S. 2021. LinA: A faster approach to piecewise linear approximations using corridors and it's application to mixed integer optimization. Technical report, CIRRELT and Université de Montréal.

Crossman, E. 1959. A theory of the acquisition of speed-skill. *Ergonomics* 2(2): 153–166.

Freund, P.; Levitt, S.; and Groseclose, T. 1999. Adaptive game playing using multiplicative weights. *Games and Economic Behavior* 29(4): 79–103.

Fudenberg, D.; and Levine, M. 1998. *The Theory of Learning in Games*. MIT Press.

Goldberg, D. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley.

Grigoriadis, M.; and Khachiyan, L. 1995. A sublinear-time randomized approximation algorithm for matrix games. *Operations Research Letters* 18: 53–58.

Hines, W. 1987. Evolutionary stable strategies: A review of basic theory. *Advances in Applied Probability* 31(2): 195–272.

Howard, N. 1971. *Paradoxes of Rationality: Theory of Metagames and Political Behavior*. MIT Press.

Jaffe, A. 2013. *Understanding Game Balance with Quantitative Methods*. Thesis, University of Washington. URL https://digital.lib.washington.edu:443/researchworks/handle/1773/22797.

Jaffe, A.; Miller, A.; Anderson, E.; Liu, Y.-E.; Karlin, A.; and Popovic, Z. 2012. Evaluating Competitive Game Balance with Restricted Play. In *Proceedings of the 8th Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, 26–31.

Lasry, J.-M.; Lions, P.-L.; and Gueant, . 2010. Mean field games and applications. *Paris-Princeton lectures on Mathematical Finance* 1–65.

Lee, J.; and Kirlik, A., eds. 2013. *The Oxford Handbook of Cognitive Engineering*. Oxford University Press.

Leibowitz, N.; Baum, B.; Enden, G.; and Karniel, A. 2010. The exponential learning equation as a function of successful trials results in sigmoid performance. *Journal of Mathematical Psychology* 54(3): 338–340.

Maynard Smith, J. 1974. The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* 47(1): 209–221.

Maynard Smith, J. 1982. *Evolution and the Theory of Games*. Cambridge University Press.

Maynard Smith, J.; and Price, G. 1973. The logic of animal conflict. *Nature* 246(5427): 15–18.

Mitchell, M. 1998. *Introduction to Genetic Algorithms*. MIT Press.

Newman, M. 2005. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46(5): 323–351.

Palacios-Huerta, I. 2003. Professionals Play Minimax. *Review of Economic Studies* 70: 395–415.

Robinson, J. 1951. An Iterative Method of Solving a Game. *Annals of Mathematics* 54(2): 296–301.

Samuelson, L. 1988. Evolutionary foundations of solution concepts for finite, two-player, normal-form games. In *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge (TARK)*, 211–225.

Schuster, P.; and Sigmund, K. 1983. Replicator dynamics. *Journal of Theoretical Biology* 100(3): 533–538.

Taylor, P.; and Jonker, L. 1978. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* 40(1): 145–156.

Tekinbas, K.; and Zimmerman, E. 2003. *Rules of Play: Game Design Fundamentals*. MIT Press.

Walker, M.; and Wooders, J. 2001. Minimax Play at Wimbledon. *American Economic Review* 91(5): 1521–1538.

Weibull, J. 1995. *Evolutionary Game Theory*. MIT Press.

Yee, A.; Rodríguez, R.; and M., A. 2014. Analysis of Strategies in American Football Using Nash Equilibrium. In *Proceedings of the 16th International Conference on Artificial Intelligence: Methodology, Systems, and Applications (AIMSA)*, 286–294.