An Experiment on Evaluating Game Regulations Using Artificial Intelligence Agents for Trading Card Games

line 1: 1st Shota Takahashi   
line 2:*Interdisciplinary Mathematical Science   
(of Frontier Media Science)*  
line 3: *Meiji Uinvercity*line 4:Nakano ,Tokyo, Japan  
line 5: ev180694@meiji.ac.jp

line 1: Motoki Ito  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: Meiji Univercityline 4: Nakano,Tokyo,Japan  
line 5:

line 1: 3rd Kazushi Ahara  
line 2: *Interdisciplinary Mathematical Science   
(of Frontier Media Science)*  
line 3: *Meiji Uinvercity*line 4: Nakano, Toyko, Japan  
line 5: email address or ORCID

*Abstract*— Trading card games (TCGs) have various rules and parameters, such as the initial number of cards in hand, the stats of each card, and unique cards that compensate for the disadvantage of a rear attacker. In general, they say that these values have a great deal to do with the appropriateness of winning and losing or game balance. However, only a few published research papers on the suitability of TCG game rules in general. The authors considered this research a genre of game informatics. They proposed a method to observe how intentional TCG rules and parameters affect the win rate, using an initially developed platform based on the famous TCG HearthStone. In this paper, we propose a method for observing how intentional changes in the rules and parameters of TCG affect the winning rate and conduct an experiment focusing on the initial number of mana granted and the initial number of cards in hand. This experiment discusses the quantification of the magnitude of the influence of the parameters governing the game regulations. (*Abstract*)

Keywords—TCG, AI, Non-perfect information game

# Introduction (*Heading 1*)

A trading card game (TCG) is a two-player non-complete information game using cards, with Magic: the Gathering (MTG) and Hearthstone being the most famous. TCGs such as Pokémon Cards, Shadowverse, and Yu-Gi-Oh are wide-played in Japan. TCGs have the following standard rules and regulations:

1. There is a set of cards to be used, called a pool card set, from which players make a specific number of cards, called a deck. The deck is shuffled face down and placed on the board, cards drawn from the deck make up the hand, and the players move their hand-cards onto the board. Each card has a description, which gives each card a unique effect. There are some victory conditions, and the player achieves the victory condition before the other player wins. It differs from board games using playing cards in three main ways.

1. Each card has its description, and one of the player's interests is to build a unique deck.

2. The combination of two or more cards can have a more significant effect called synergy. Finding synergy combinations is also an interest of the game.

3. The player should observe and imagine the opponent's deck characteristics based on their play and move cards accordingly in the board battles.

In line with the above characteristics, artificial intelligence (AI) research would include AI constructing decks, AI discovering synergies, and AI discovering the personality of the opponent player. Another interest of AI research is measuring game regulations' appropriateness by using AI to play.

In the current widely played TCGs, there is little public information on the appropriateness of the game regulations. We can only imagine that they (the game companies) divide players into advanced, intermediate, and beginner levels, analyze how players use each card in each class, and apply the results to architect game regulations. We can only imagine that they derive the appropriateness from big data analytics of users' massive amount of play data. Many articles on how AI makes these analyses, but the detailed information has not been open.

The authors constructed an original platform of hearthstone, where card descriptions are publicly available. We call it FireplaceAharalab[\*1], a platform with enhanced card implementation, improved a platform Fireplace[2] in python. Our goal is to test whether the Vector agent, an AI agent prototyped on this platform, can be used to measure the appropriateness of game regulations. Specifically, we investigated how changing two or more rules simultaneously in a TCG affects the win rate. The authors are convinced that we widely apply this study as a criterion for creating rules for TCGs and determining the appropriateness of regulations.

# previous studies

In their paper [\*3], Yamada and Ahara proposed a method to measure the synergy between two specific cards using fireplaceAharalab and conducted experiments on its effectiveness. Their attempt to quantify the synergy using simulation is similar to the idea of this research. They compared two cases (a) where the deck contains two specific cards in a row and (b) where the deck contains two specific cards at random. And they tested whether there is a significant difference in the win rate.

Daniel Whitehouse, Edward J. Powley , et al. showed in their paper [\*4] that Monte Carlo Tree Search (MCTS) and Determinization are effective methods for implementing AI agents in Magic: The Gathering when the deck composition is known to both players.

In their paper [\*5], Takahashi and Ahara investigated the change of winning rate by changing parameters on fireplaceAharalab. They discovered that they could quantify the tendency for players with many initial cards to have a higher win rate. This study corresponds to the preliminary research of this paper.

# experiment

To experiment, we used fireplaceAharalab, which is our modification of Fireplace, the simulation engine of the DTCG Hearthstone. The artificial intelligence player we used was the Vector agent, which is a one-move reading agent using Monte Carlo tree search. As described in detail below, we played 1000 games per simulation. It took about 1 hour and 30 minutes on an Intel(R) Core™ i7-8700 3.20GHz CPU.

In this paper, we experimented by changing two things: the initial number of mana granted and the initial number of cards in hand. Mana is a renewable resource on the number of cards you can use on your turn. Cards have a weighting called cost. You cannot combine cards that exceed the number of mana. We chose this initial number of cards in hand, and the initial number of mana granted because our preliminary experiments showed that these two rule changes significantly impacted the game. In the following, we will describe the method of the initial investigation, the standard experiments in this experiment, and the changes made in each experiment.

We conducted nine different experiments for each case to investigate how changing the initial number of mana granted and the initial number of cards in hand would affect the win rate. As a result, both significantly impact the game. The following is a detailed description of the experiment.

### **Preliminary Experiment 1**

First, let's change the initial number of mana granted to player A. Let's say that player A's initial mana is α and player B's initial mana is β. We assume that player A always starts the game first, and that α and β are integers satisfying 1<=α<=3 and 1<=β<=3, respectively (written in set notation). Each condition was played 10,000 times for a total of 90,000,000 times.

We calculate the average win rate and variance for the experiment results and look at the change in win rate.

### Preliminary Experiment 2

Next, we will discuss the case where the initial number of cards in hand is changed. The number of cards in A's initial hand is γ, and the number of cards in B's initial hand is Θ, where γ and Θ are natural numbers satisfying 3<=γ<=5 and 3<=Θ<=5, respectively.

### Main Experiment

After conducting preliminary experiments, we investigated how changing both the initial number of mana granted and the initial number of cards in hand would affect the game's outcome.

We experimented with two combinations of the initial number of mana given to the first player α, two varieties of the initial number of cards in the first player's hand γ, two combinations of the initial number of mana given to the second player β, and two varieties of the initial number of cards in the second player's hand Θ. The initial number of mana granted, α and β, is 1 or 2, and the initial number of cards in hand, γ, and Θ, is 3 or 4. Unlike the previous experiment, we changed one of the initial numbers of mana and the initial number of cards in hand and changed the initial number of cards in hand at the same time. For this purpose, we conducted 16 experiments .In the following, we denote each combination of conditions as (α, Θ, β, γ).

# Result

We describe the results of our preliminary and main experiments in detail below.

### Results of Preliminary Experiment 1

Table 1 shows the results of preliminary experiment 1. This table shows that the larger the number of mana, the more advantageous the game tends to be. Therefore, we formulated the null hypothesis that the increase or decrease of mana did not affect the winning rate of the game and conducted a chi-square test with the winning rate as the independent variable at a significance level of 0.05 for each combination of conditions. The varieties of each situation and the p-values are shown in table 2 below. As a result, the P-value was as close to zero as possible in all conditions. Therefore, the null hypothesis was rejected, and the increase in the initial mana granted affected the match.

Table 1 Number of start mana granted condition and A's win rate

|  |  |
| --- | --- |
| *α – β* | *A’s win rate* |
| *1-1(a)* | *0.5646* |
| *1-2(b)* | *0.1756* |
| *1-3(c)* | *0.13* |
| *2-1(d)* | *0.7828* |
| *2-2(e)* | *0.5872* |
| *2-3(f)* | *0.3817* |
| *3-1(g)* | *0.9027* |
| *3-2(h)* | *0.7745* |
| *3-3(i)* | *0.5984* |

Table 2 The varieties of each mana situation and the p-values

|  |  |  |
| --- | --- | --- |
| *Meaning of the combination* | The varieties of each situation | *P-value* |
| *Increase mana of B only and*  *A's mana is 1* | *(a)-(b)* | *0* |
| *(b)-(c)* | *0* |
| *(c)-(a)* | *0* |
| *Increase mana of B only and*  *A's mana is 2* | *(d)-(e)* | *0* |
| *(e)-(f)* | *0* |
| *(f)-(d)* | *0* |
| *Increase mana of B only and*  *A's mana is 3* | *(g)-(h)* | *0* |
| *(h)-(i)* | *0* |
| *(i)-(g)* | *0* |
| *Increase mana of A only and*  *B's mana is 1* | *(a)-(d)* | *0* |
| *(d)-(g)* | *0* |
| *(g)-(a)* | *0* |
| *Increase mana of A only and*  *B's mana is 2* | *(b)-(e)* | *0* |
| *(e)-(h)* | *0* |
| *(h)-(b)* | *0* |
| *Increase mana of A only and*  *B's mana is 3* | *(c)-(f)* | *0* |
| *(f)-(i)* | *0* |
| *(i)-(c)* | *0* |

### Results of Preliminary Experiment 2

Table 3 shows the results of the preliminary experiment 2. This table shows that the more cards you have in your first hand, the better your chances of winning. Therefore, we formulated a null hypothesis that "an increase in the number of cards in the initial hand does not affect the winning rate of the game" and conducted a chi-square test with the winning rate as the independent variable at a significance level of 0.05 for each combination of conditions. The varieties of each state are shown in table 4 below. The results showed that the p-value was below 0.05 in most cases. Therefore, the null hypothesis was rejected, and it was found that the increase in the number of initial cards in hand affected the game.

Table 3 Number of start hands granted condition and A's winning percentage

|  |  |
| --- | --- |
| γ -Θ | A's win rate |
| 3-3(a’) | 0.5719 |
| 3-4(b’) | 0.558 |
| 3-5(c’) | 0.5408 |
| 4-3(d’) | 0.5919 |
| 4-4(e’) | 0.5691 |
| 4-5(f’) | 0.5591 |
| 5-3(g’) | 0.6128 |
| 5-4(h’) | 0.5912 |
| 5-5(I’) | 0.5758 |

Table 4 The varieties of each start hand situation and the p-values

|  |  |  |
| --- | --- | --- |
| *Meaning of the combination* | The varieties of each situation | *P-value* |
| *Increase start hand of B only and*  *A's start hand is 3* | *(a’)-(b’)* | 0.047417176 |
| *(b’)-(c’)* | 0.014509046 |
| *(c’)-(a’)* | 0.00000958 |
| *Increase start hand of B only and*  *A's start hand is 4* | *(d’)-(e’)* | 0.001086832 |
| *(e’)-(f’)* | 0.153874036 |
| *(f’)-(d’)* | 0.000002699 |
| *Increase start hand of B only and*  *A's start hand is 5* | *(g’)-(h’)* | 0.001806554 |
| *(h’)-(I’)* | 0.027180983 |
| *(I’)-(g’)* | 0.00000099 |
| *Increase start hand of A only and*  *B's start hand is 3* | *(a’)-(d’)* | 0.004141839 |
| *(d’)-(g’)* | 0.000000386 |
| *(g’)-(a’)* | 0.002530685 |
| *Increase start hand of A only and*  *B's start hand is 4* | *(b’)-(e’)* | 0.113509885 |
| *(e’)-(h’)* | 0.000002051 |
| *(h’)-(b’)* | 0.001543606 |
| *Increase start hand of A only and*  *B's start hand is 5* | *(c’)-(f’)* | 0.009294636 |
| *(f’)-(i’)* | 0.000000627 |
| *(i’)-(c’)* | 0.01714777 |

### Results of main Expreiment

We describe the results of the main experiment below. Figure 1 shows the average win rate and 95% confidence interval for each condition. In this table, we can divide the results into four major groups according to the stage of the first attack victory rate. Group 1 consists of (α,γ,β,Θ) combinations (2,4,1,3), (2,5,1,4), (2,3,1,3), and (2,3,1,4) with a first-attack win rate of 70% or higher, and (α,γ,β,Θ) combinations (2,4,2,3), (2,4,2,4), (1,4,1,3), (2 (1,4,1,3), (2,4,2,4), (1,4,1,3), (2,3,2,3), (2,3,2,3), (2,3,2,3), and (2,3,2,3), Group 2 with (α,γ,β,Θ) combinations near 55% first-strike win rate, Group 3 with (1,4,1,4), (1,3,1,3), (2,3,2,4), and (1,3,1,4), and Group 4 with other first-strike win rates below 40%. From the results of Group 1 and Group 4, it is clear that players with more initial mana granted tend to win.

On the other hand, when the number of cards in the initial hand was increased and when it was not, the difference in winning rate was not as pronounced as when there was a difference in initial mana. Therefore, we formulated a null hypothesis that the difference in the number of cards in the initial hand does not affect the win rate under the combination of conditions shown in Table 4. We also conducted a chi-square test with the winning percentage as the independent variable at a significance level of 0.05 for each combination of groups. Therefore, the null hypothesis was rejected except for the above two varieties. Instead, the alternative hypothesis, "the difference in the initial number of cards in hand can affect the winning rate," was adopted.

From the above, we could observe the change in the win rate even when both the initial number of mana and the initial number of cards in hand were varied, and we could quantitatively measure the difference. We also observed that the number of initial mana had a more significant impact than the number of initial cards.

# **Discussion and Considerations**

##### We discuss the results of our experiment as follows.

##### First, it was verified that changing the parameters related to the rules has a non-linear effect on the win rate. We will discuss this from two points of view: the initial number of cards in hand and the initial number of mana. First, the number of cards in the initial hand seems to give a linear advantage to the player as the number of cards increases. What were the results of the experiment? Let's evaluate the change in the number of wins from (3,4) to (3,8) ((1,3,1,t) rules, where t=4,5,6,7,8) in Hearthstone's "with coins" and "starting at one mana" rules, using a 95% interval estimation. Neither the win rate (number of wins/match) nor the win ratio (number of wins/losses) can be considered linear, as shown in graph 2. It is not clear from this experiment alone what mathematical formula approximates this distribution.

##### Next, the effect of the initial number of mana is evaluated by estimating the change in the number of wins in the 95th percentile interval ((1,3,t,4), where t=1,2,3,4) for the number of mana from 1 to 4 with "coins" and "hand starting with three cards for the first attack and four cards for the second attack. The change in the number of wins is evaluated by 95% interval estimation in Figure \*\*. In this case, neither the win rate nor the win ratio is linear.

##### Second, we verified that changing the rule's parameters on the win ratio depends on the type of parameter. In the regulations with coins, the change in win ratio from the conditional combination (1,3,1,4) to (1,3,1,5) and the change in win ratio from (1,3,1,4) to (1,3,2,4) are not found to be identical by a 95 percent chi-square test. A quantitative experiment verified that the effect of mana increase on the game is significantly more significant than the increase in the number of cards in hand. Although it seems impossible to numerically express the comparison between the effects of the initial number of cards and the initial number of mana depending on the player's senses, we can say from this experiment that we could quantify it (if we consider the situation fixed). However, this does not guarantee that the impact of increased mana on the game is significantly more significant than the increased number of cards in hand in all situations.

##### Third, the impact of the two parameters on the rules was verified to be non-independent. The effect of increasing the initial number of cards in hand by one is different by the 95 percent test for the change in win rate and win ratio from condition (1,3,1,4) to (1,3,1,5) and from (1,3,2,4) to (1,3,2,5). Then, can we formulate some model equations for the change in win ratio induced by these two parameters? Although we cannot discuss this in this paper, this paper has a specific value because we were able to present a numerical solution that the solution of the model equation should satisfy.

##### Fourth, it can be seen that certain conclusions can be drawn by repeating the method of this experiment for changes in parameters related to the rules that were not treated in this experiment. For example, changes in deck size, specific card stats (nerfs), etc., are possible. Nevertheless, the number of possible combinations is so large that it would be impossible to draw generalizations, even if they could be measured in a

# Future works

Based on these considerations, we will discuss AI research issues in TCG.

First, we fixed the deck type and the artificial intelligence agent in this experiment and then changed the rules. Still, it is possible to use the same experimental method to change the deck type and the artificial intelligence agent after fixing the rules. This presents us with one major problem. Can we quantify the compatibility between deck type and play style? We distinguish between aggro decks and control decks and classify playstyles into aggro and control. The current artificial intelligence agents can provide play at the so-called beginner level. It is unlikely that artificial intelligence can give aggro or control players for human players. An approach to this simple yet profound problem is awaited.

Secondly, in a large-scale DTCG environment, there is big data analysis by recording the play of human players. We will study the consistency between these significant data analysis results and experiments' comments via artificial intelligence agents. We believe that this will lead to the development of TCGs with more appropriate and satisfactory rules.

Finally, I would like to mention the possibility of creating advanced agents using deep learning. If we assume that the deck used by the player is a heuristic deck (a deck that makes sense to the human eye), and the deck used by the opponent is one of several popular deck types, the search space for deep learning does not seem extremely large (compared to Shogi or Go). If we assume that the deck is one of several popular deck types, the search space for deep learning (compared to Shogi or Go) does not seem extremely large. Take Hearthstone as an example. There are roughly 20 cards used in a single deck (some cards can be summoned independently of the deck, but a small number). This is less than the 34 types in Mah-jongg (a turn-based card game with incomplete information). We are preparing a paper that will report significant conclusions about this.

##### particular environment.

# References

1. FireplaceAharalab