Can we save near-dying games? An approach using advantage of initiative and game refinement measures

Htun Pa Pa Aung [1], Mohd Nor Akmal Khalid[2] and Hiroyuki Iida[3]
{ htun.pp.aung@jaist.ac.jp[1], akmal@jaist.ac.jp[2], iida@jaist.ac.jp[3] }

[123] School of Information Science, Japan Advanced Institute of Science and Technology, 923-1211 Ishikawa, Japan,

Abstract. Games are attractive and engaging due to the complexity they pose to the player. Some games are complex enough which made them attractive to play. However, popular games would lose their attractiveness due to the large advantage of initiative. High-performance AI like AlphaZero suggests from their actual games played that advantage of the first player would become larger as the performance level increases. This implies that games with a large advantage of the initiative would lose their attractiveness due to the unfairness. This paper explores an innovative way to make a game stay attractive. A link between the advantage of initiative and performance level is investigated while using Scrabble AI. Using two measures: the advantage of initiative and game refinement, possible treatments are considered. The experimental results with Scrabble AI suggest that reduction of search space from 15x15 to 13x13 board size is a possible enhancement.

Keywords: Refinement Games, Atificial Intelligence

1. Introduction

Artificial intelligence (AI) and games have a long history together. Many research in this direction were concerned with establishing computer program that capable of playing games, typically known as 'strong' AI player, with or without a learning component. Even since before AI was identified as a field, early pioneers of computer science had developed game-playing programs because they wanted to investigate whether computers could solve complex tasks that seemed to require "intelligence" [12, 17].

In classic board games, the game typically has two players where the full state of the game is visible to both players and the mechanics was a discrete turn-based game play. However, the interest of the research community had increasingly shifted towards the application of AI from board games to other types of game in the last decade and a half, particularly in video games. A large part of the research focuses mainly on developing AI agents for playing games; either as effectively as possible, or human-like (or a particular human), with respect to some other property [20].

However, recent advent of super-human performance of AI (level of performance that exceeds that of the strongest human players) had been demonstrated by the AlphaZero program [14]. Table 1 showed the results of the games played by Alpha Zero. While game engagement had been due to the effort and skills required from the player to complete them (or, in the case of puzzles, solve them), this super-human performance raised concerned whether it would provide engagement or even be fair to its opponent. As compared to the results of human players

collected from various Chess and Shogi tournament databases (Table 2), the winning rate as first player or second player had been roughly balanced ($\approx$ 50%) which gave the sense of engagement in the game. As such, if the performance of the AI program continues in this direction, the Chess game and Chess-like game may not be interesting or engaging anymore to the player, or even lose its attractive to the observer; thus, in the context of this paper, is called a "dying game."

Table 1. The results of Alpha Zero in Chess and Shogi [14]

	vs AI	Draw	Win Games		Win Rate (%)	
			White	Black	White	Black
Chess†	Stockfish 8	958	286	56	0.84?	0.16?
Shogi*	Elmo	0	55	45	0.55	0.45

?excluding the draw games.

†total games is 1300; *total games is 100.

Table 2. The statistical data of human players in Chess and Shogi

	Total	Draw	Win Games		Win Rate (%)	
			White	Black	White	Black
Chess†	898,173	252893	339149	306108	0.53?	0.47?
Shogi*	2,323	0	1167	1156	0.50	0.50

?excluding the draw games.

†the game covered for the years of 1475 to 2019 (545 years). *the game covered only on public matches.

The importance of fairness, called advantage of initiative (AoI), had been previously discussed and proved through a conclusive and elegant theorem on first-player wins over the second-player win [15]. The concept was further observed in the context of solved games where the concept of initiative seems to be a predominant notion under the requirement that the first player has sufficient space to fulfill the goals [19]. Also, another recent works also takes into account AoI in the context of the player strength, which revisited the concept and redefine AoI as the initiative action of both players in the first stage "to take the advantage" (i.e., to be the first winner in the earlier stages of the game) [4].

Previous works that addresses the AoI aspect had been investigated to determine the impact of the initiative on the game-theoretic value through a large number of k-in-a-row games and 200 Domineering games as a function of the board size [18]. Another study had adopted self-play experiments using AI in Scrabble games where player that established early opening advantage took higher winning expectancy [4]. Moreover, an interesting example with randomized initial position was observed using Quackle AI in Scrabble. Indeed, since AI in Scrabble showed the level of performance that exceeds the strongest human players, keeping fairness is becoming a difficult feat since the gap between AI and human performance is growing. Therefore, this study is driven towards a new challenge of maintaining the fairness, which in the current context, in the game of Scrabble.

2. Defining Near-Dying Games

2.1 From the perspective of AoI

To illustrates the AoI, different performance levels of the player have to be considered (see Figure 1). Based on the previous tournament data, it is reasonable to assume that human experts would be located somewhere between $\{h_l, h_u\}$ in the performance level, where h_l and h_u corresponds to lower bound and upper bound of human experts, respectively.
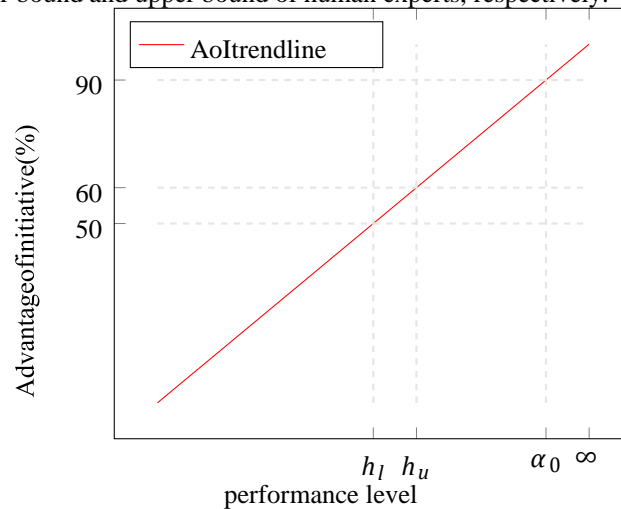

Fig. 1. Illustration of assumed advantage of initiative (AoI) over different performance levels

Meanwhile, AI such as AlphaZero (denoted as α_0) having performance level around 90% or more, and the perfect performance level corresponds to ∞. Suppose that the performance level of the player increases (as seen by the current trend of AI development), it can be conjectured that the AoI would also becomes greater. As such, any game under consideration would lose its attractiveness due to unfairness being far beyond the acceptable AoI margin. In essence, that game would approach "near-death" stage since it would not be attractive to play competitively or even for fun.

2.2 From the perspective of the game attractiveness

The general model of game refinement (GR) was proposed based on the concept of game information progress, where the gap between board games and sports games was bridged [16, 5]. Assuming the average length of games D and the average branching factor or the number of possible moves B played by human experts [7, 5, 6] and Alpha Zero [14], the GR value can be calculated which is given in Table 3 and illustrated in Figure 2.

Table 3. The game length and the GR value of Chess and Shogi games played human experts and AlphaZero

	Human	GR	Alpha Zero	GR
Chess	80	0.074	130	0.04
Shogi	115	0.073	204	0.04

The tendency of GR value typically converges towards comfortable zone ($GR \in [0.07, 0.08]$), which associated with the measures of game entertainment and sophistication, involving balance.
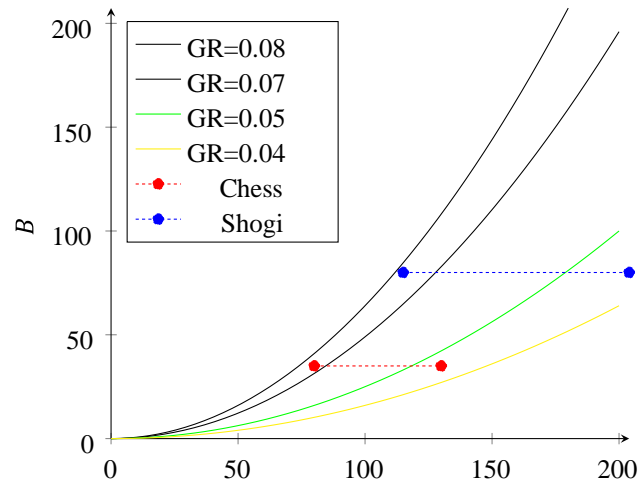

Fig. 2. The average number of Possible moves B and game length D in chess and shogi: human experts and Alpha Zero compared

Between level of skills and chance of the game [1, 3]. Based on this fact, entertaining experience had been demonstrated by the human experts based on the GR value that were in the comfortable zone. However, AlphaZero AI demonstrates serious experience due to low GR value. These contrasting situation implies that game play with super-human performance would not be enjoyable and less entertaining, thus loses its attractiveness (hence, "dying" in a sense that the game become predictable and causes frustrations instead of fun).

3. Related Works

Traditional board games usually have well-defined fixed rules, known in advance to the players. Many years ago, game designers and game developers primarily concentrated on having more realistic games and worked on implementing games with high-quality graphics (i.e., having high-resolution textures, a good measure of frames-per-second, etc.). Until recently, this policy guaranteed reasonable profits for the development company, in the last decade, it has not been enough to ensure the success of a game, as players now demand other features for their games. Many of these required features are associated with the resolution of problems that demand

knowledge from a wide set of research domains such as art, psychology, narrative, or music to name but a few.

In this heterogeneous context, one paradigm that is present in all (or in most) games is adoption of artificial intelligence (AI) [11]. The most traditional application of AI in games is to regulate the behavior of the non-player-characters (NPCs) in the game with the aim of enriching intelligence on the opponent and as a consequence increasing the satisfaction of the player who demands an opponent that reveals intelligent response.

In the transition from traditional games to games with AI players, it is hard to keep fairness when AI becomes stronger than human experts. Fairness is not simply so serious in human intelligence. But there is a challenging task in some games from the perspective of machine intelligence. Different definitions of fairness need to be considered. This forces us to investigate the original rules and elements of the game respective to the nature of the intelligence and level of play [2].

To quote Lucas et al. [10], this basically involves "restricting the information available to the AI player to the information a human player may be able to gather in the game, and restricting the actions of the AI player to human player actions." Additionally, the virtual player should pose a challenge to the human players independently of their skills and of the strategies they adopt to play [11].

The initial position plays an important role as observed from the artistic perspective. Therefore, it is a serious issue to determine the reasonable initial position among many potential candidates when creating a potentially sophisticated game. According to the definition of fairness, a two player game is fair if and only if the winning ratio for White and Black is statistically equal or nearly so [5].

For sufficiently strong players or even the omniscient players, the game-theoretic value of a game must be a draw to maintain fairness. Of course, it requires that a draw must exist in the game. But in Scrabble, a draw is rarely found as an outcome. Although according to the game-theoretic value prediction, [8] showed perfect play by both sides would lead to a draw, it is not right for all sophisticated games.

The previous work on Quackle and Maven showed that when the player is stronger than human experts and nearly perfect, a player has higher winning probability if he/she gets the early advantage [4]. The results of self-play experiments on Quackle is shown in Figure 3.
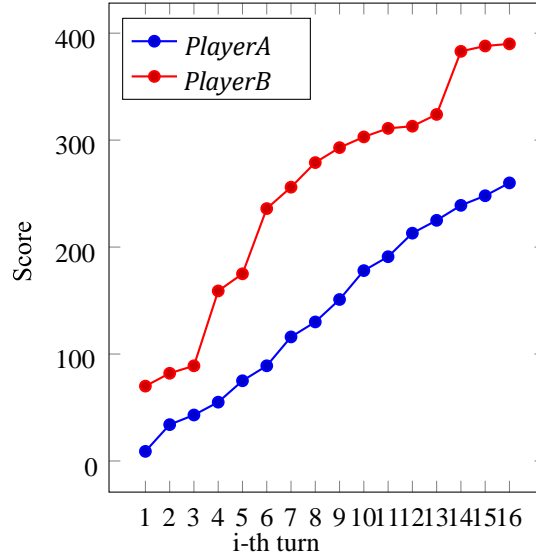
Fig. 3. A self-play experiment using Quackle: score at each turn

4. Related Works

Scrabble is a popular worldwide word game with numerous national and international Scrabble associations and a biennial world championship. Like chess and bridge, competitive play of the Scrabble game is very popular which continues to attract more players each year. The Scrabble game has also reached a new community of players in the digital age with digital versions of the game from Hasbro licensee, Electronic Arts [21].

On the other hand, computer AI players have already outperformed human opponents in competitive Scrabble. However, defeating a computer AI opponent requires complex and efficient heuristics. Currently, Maven and Quackle are the leading Scrabble AIs, where both have defeated the best human champions in tournaments [13].

Between the time when simulation became available as an analytical tool in 1990 and simulations were first used in competitive play in 1996, human players have improved their positional skills by studying simulation results [13]. However, Table 4 shows that Maven has maintained at least a slight superiority over human experts since its debut in 1986 according to Maven's tournament statistics where the total matches and tournament record is 3500 wins and 1500 losses against an average rating of 1975. This may be the earliest time at which a computer program achieved world-class status over human masters in a non-trivial game of skill [13].

Table 4. MAVEN and human experts compared [13]

	MAVEN	Human expert
Average Bingo per game	1.9	1.5
Average tiles played per game	4.762	4.348
Average turns per game	10.5	11.5
Chance to play Bingo if exists	100%	85%

The second AI, Quackle, is the strongest alternative Scrabble AI to Maven. Our previous work also showed that there is advantage of initiative issue in Scrabble AI [4]. However, the prospect of advancement in the Scrabble AI raises important questions:

- How will the game play among very strong players be?

- Is there any link between AoI and performance level of players or other game elements?

- What is the possible solution to maintain the fairness until the end of game?

Early work shows that the game is unfair when the players have equal strength and become stronger than the grandmaster level. The higher the player's level is, the greater the winning percentage he/she takes [4]. In this study, the initiative was redefined as an action of both players in the first stage "to take the advantage" (i.e., to be the first winner in the earlier stages of the game). Based on this definition, an experiment conducted in Scrabble game can be conducted by varying the performance level of the AI player and determining the appropriate solution to maintain fairness until the end of the game.

5. Experimental with Scrabble

5.1 Experimental Design
Scrabble AI program was built to simulate multiple Scrabble matches. 100 distinct games are simulated with 100 iterations each. To calculate the win ratio with respect to the AoI, the individual and total scores, the number of wins, and the number of losses, were collected. The performance of the AI player can be explored from several points of view.

In this paper, different strength levels of AI players was adopted based on various dictionary size (DS) and different performance level of the AI player (LV). The participants of the Scrabble matches were AI players who played Scrabble game with systematically varied levels of strength. It was found that the probability measure has a linear relation with LV when the player is much stronger than grandmaster LV (estimated at $LV = 0.7$) and the board size is standard 15×15[4].

5.1 Experimental Design
Table 5 and Figure 4 presents the win ratio of keeping the advantage of initiative for different dictionary size (0.5, 0.7 and 1.0) and its depiction in graph form. This reveals that the original setting of Scrabble has the advantage of initiative for the very strong players ($LV \geq 0.5$).

Table 5. Winning ratio of keeping advantage of initiative over different performance levels

Performance Level	Dictionary Size					
	0.5	0.6	0.7	0.8	0.9	1.0
0.1	0.5	0.5	0.5	0.5	0.5	0.5
0.2	0.5	0.5	0.5	0.5	0.5	0.5
0.3	0.5	0.5	0.5	0.5	0.5	0.5
0.4	0.5	0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.6	0.54	0.6	0.6	0.65	0.67	0.69
0.7	0.55	0.68	0.71	0.73	0.76	0.78

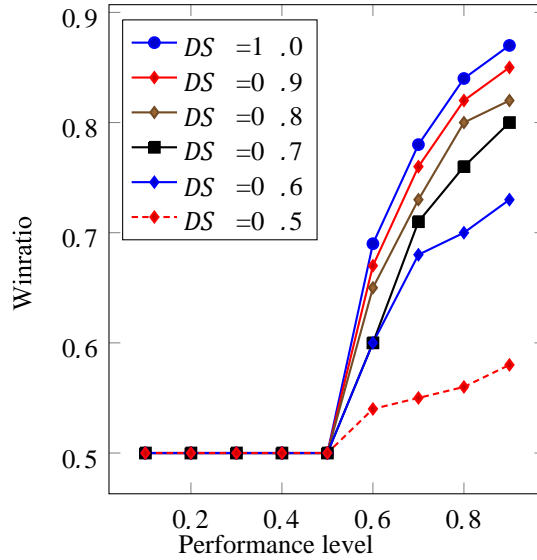| 0.8	0.56	0.7	0.76	0.8	0.82	0.84
0.9	0.58	0.73	0.8	0.82	0.85	0.7

Fig. 4. Win ratio for keeping the advantage of initiative over different performance levels for three different sizes of dictionary.

Increase in the dictionary size (DS) corresponds to the strength of the player. As such, higher DS is more likely to increase the winning ratio when reaching value greater than 0.5 (stronger than grandmaster level). Resetting of dictionary size (DS) on standard board size in Scrabble cannot maintain the fairness of the game outcome (in this context, win ratio) until the end of the game.

One possible enhancement is to reduce the board size (Figure 5). This results in keeping that the outcome of the game to be unclear, as shown in Figure 6. A question now arises: how large is the board size needed for the Scrabble for the player to earn winning advantage (AoI) in the early stage? With the focus on different levels of players (0 to 1), 15×15 Scrabble board turned out to be a winning criterion for the player who gets the advantage in the first stage under the condition if both players are much stronger than grandmaster level. However, fairness can be maintained by resetting the board size of Scrabble to 13×13 which keeps the probability of AoI of both players in each stage of the game.

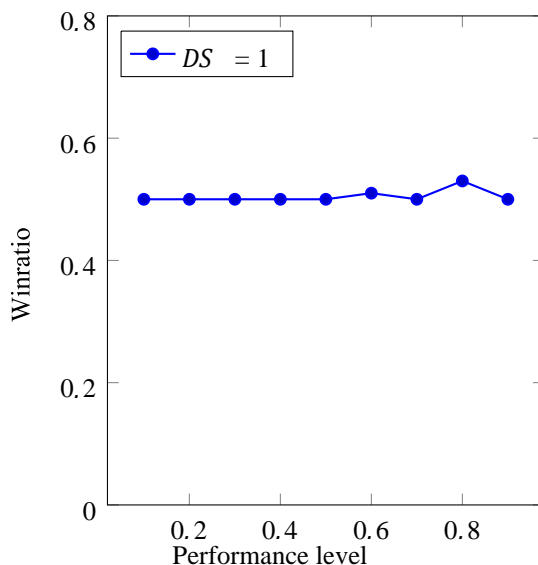| 2W			3L			3L				2W
	2W			2L		2L			2W	
		2W			2L			2W		
			2W				2W			
3L			3L			3L				3L
	2L			2L		2L			2L	
		2L			2W			2L		
	2L			2L		2L			2L	
3L			3L			3L				3L
			2W				2W			
		2W			2L			2W		
	2W			2L		2L			2W	
2W			3L			3L				2W

Fig. 5. 13x13 variant of Scrabble board

Fig. 6. Win ratio keeping the advantage of initiative over various performance levels on 13x13 board size

6. Concluding Remarks

While Artificial Intelligence is getting more powerful and about to beat human experts not only in some games but also in other areas, this fact causes fairness to be an important aspect of competitiveness.

Thus, the advantage of initiative (AoI) has been revisited using Scrabble as a test bed for different levels of AI players. This study focuses only on two aspects: the size of the search space (board size) and the quality of AI players (strength of players based on AI levels and dictionary size) of Scrabble from the viewpoint of the advantage of initiative.

First, the link between the search space and the quality of players was investigated. 13×13 board size is suitable enough to keep the balance of the information about the game outcome is not clear at the very end of the game. But for 15×15 board size, the player can maintain his advantage until the end of the game if he can establish an advantage in the first stage of the game.

Although further study is needed to determine the level of players from winning percentage of the game, one possible way is to establish new measurement by using the merit of Elo-rating system. Based on that measurement, we can reconsider the rules and elements of the original game in order that the players can play effectively for their specific purposes.Moreover, the advantage of initiative needs to be clarified by other domains as well.

References

[1] Claude E Shannon.Xxii. : Programming a computer for playing chess.The London,Edinburgh, and Dublin Philosophical Magazine and Journal of Science,41(314):256275, (1950).

[2] M Turing. : Digital computers applied to games.Faster than thought, 1953.

[3] Georgios N Yannakakis and Julian Togelius. : Artificial Intelligence and Games.Springer,(2018).

[4] David Silver,Thomas Hubert,Julian Schrittwieser,Ioan-nis Antonoglou,Matthew Lai,Arthur Guez, Marc Lanc-tot,Laurent Sifre,Dharshan Kumaran,Thore Graepel,et al. : Mastering chess and shogi by self-play with ageneral reinforcement learning algorithm.arXiv preprint arXiv:1712.01815, (2017).

[5] David Singmaster. : Almost all games are first person games.Eureka, 41:3337, 1981.

[6] H.Jaap van den Herik,Jos W.H.M.Uiterwijk,and Jackvan Rijswijck. Games solved : Now and in the future.Artificial Intelligence, 134(1):277311, (2002).

[7] Htun Pa Pa Aung, Hiroyuki Iida. : Advantage of initiative revisited: A case study using scrabble AI.In ICAIT2018, 11 (2018).

[8] J.W.H.M. Uiterwijk and H.J. van den Herik. : The advantage of the initiative.Information Sciences, 122(1):4358, (2000).

[9] Arie Pratama Sutiono, Ayu Purwarianti, and HiroyukiIida. : A mathematical model of game refinement. In International Conference on Intelligent Technologies for Interactive Entertainment, pages 148151. Springer, (2014).

[10] H.Iida. : On games and fairness. In 12th Game Programming Workshop in Japan, pages 1722, (2007).

[11] Hiroyuki Iida,Nobuo Takeshita,and Jin Yoshimura. : A metric for entertainment of boardgames:its implication for evolution of chess variants. In Entertainment Computing, pages 6572. Springer, (2003).

[12] Hiroyuki Iida. : Fairness, judges and thrill in games.Japan Advanced Institute of Science and Technology, Tech. Rep,28, (2008).

[13] Xiong S.,Zuo L.,Iida H. : Possible Interpretations for Game Refinement Measure. In:Munekata N., Kunita I., Hoshino J. (eds) Entertainment ComputingICEC 2017. ICEC Lecture Notes in Computer Science, vol 10507. Springer, Cham. (2017).

[14] Hiroyuki Iida. : Where is a line between work and play?IPSJ SIG Technical Report,2018-GI-39(13):16,February (2018).

[15] Mariela Nogueira Collazo,Carlos Cotta,and Antonio J.Fern an dez-Leiva. : Virtual player design using selflearning via competitive convolutionary algorithms.Natural Computing, 13(2):131144, Jun (2014).

[16] S.M.Lucas.Computational Intelligence and AI in Games : A New IEEE Transactions, in IEEE Transactions on Computational Intelligence and AI in Games, vol. 1, No. 1, pp. 1-3,March.doi:10.1109/TCIAIG.2009.2021433 (2009).

[17] Simon M Lucas, Michael Mateas,Mike Preuss, PieterSpronck,and Julian Togelius. : Artificial and computational intelligence in games (dagstuhl seminar 12191). In Dagstuhl Reports,volume 2.Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, (2012).

[18] H. Kita, A. Cincotti,and H. Iida. : Theoretical value prediction in game playing. Technical Report 87(2005-GI014), IPSJ, Sep (2005).

[19] About the game improve your knowledge.https://scrabble.hasbro.com/en-us/history.(Visited on 4/02/2019).

[20] B.Sheppard. : World-championship-caliber scrabble. Artificial Intelligence, 134(1):241 275, (2002).

[21] S.M.Lucas.Computational intelligence and AI in games:A new IEEE transactions.IEEE Transactions on Computational Intelligence and AI in Games,1(1):13, (2009).