# Enhanced Rolling Horizon Evolution Algorithm With Opponent Model Learning: Results for the Fighting Game AI Competition

Kyle Keener

Information Systems, UMBC

IS 369 – Writings in Information Systems

Prof. Bill Ryan

May 1st, 2023

Tang, Z., Zhu, Y., Zhao, D., & Lucas, S. M. (2023). Enhanced Rolling Horizon Evolution Algorithm with Opponent Model Learning: Results for the Fighting Game AI Competition. *IEEE Transactions on Games, 15*(1), 5–15. https://doi.org/10.1109/TG.2020.3022698

I've decided to choose this particular article because I wanted to dive deeper into the research being done on the advancement of artificial neural networks and look into the development of more sophisticated AI for the purpose of improving multiplayer AI. Further still, I believed it would be a fascinating look into the development of real-time decision making and would provide some documentation on the algorithm for the process of real-time decision-making.

The main advancement into AI this article in particular was looking to research was the increase of interest in developing artificial intelligence for competitive, zero-sum multiplayer games. In the past, the most widely researched AI compatibility was done in turn-based multiplayer utilizing the Monte Carlo Tree Search algorithm combined with deep reinforcement learning. The MCTS managed to be proficient to excellent in beating real-world players at turn-based strategy games with very condensed rulesets such as poker, go, or chess. However, the MCTS required a large number of iterations in order to populate a move set that was competent enough to play an opponent, and thus, had very limited application to real-time decision making. This article proposes the use of a Rolling Horizon Evolution Algorithm which provides lower computational overhead, better memory of preferred actions, and most importantly can apply opponent selection to help decide the best population of move sets to implement. The advancements made in the RHEA will thus be put to the test against existing MCTS bots in a fighting game AI competition.

To take a closer look at the RHEA, it's an optimization process that evolves action sequences through a forward-planning model. What it does is it takes a population of individuals, with each individual containing a number of moves executed in a sequence. Each of these individuals were cross mutated with the previous iteration replacing one move in the move set by determining the overall fitness, hp score, action sequence, and forward model of the other individuals within the population. The major issue to this was that the RHEA by itself could not directly infer opponent actions alone. Thus, an opponent learning model was needed to enhance the effectiveness of RHEA's real-time decision making.

The OLM was applied to the RHEA in a number of different ways. First was the supervised-based model and the reinforcement-based models (RHEAOM-SL and RHEAOM-R, respectively). The former uses an online training model with the latest observable pattern saved for sample adaptation, the latter utilized cross-entropy, however the cross-entropy worked only for the RHEA, but not for the opponent. Likewise, there were two gradient models utilized with the RHEAOM, the Q-gradient model and the policy-gradient model (RHEAOM-Q and RHEAOM-PG, respectively), both utilize a training goal via an iterative n-step return. However, the Q-gradient updates its parameters with a descending gradient to minimize mean-square loss, while the policy-gradient directly optimizes its agent policy with an ascending gradient based on estimated total reward of mutated move sets, allowing for a possible reward calculation of expended health against the opponent's health.

The test itself was composed of three different subsections. Firstly, setup would pit a current bot AI known as FightingICE against the new RHEAOM with 56 potential actions done between 16.67 milliseconds to simulate typical player delay between actions with a 15-frame latency to balance move sets. The two bots would fight until one or the other would reach 0 health points or the round ends at 60s with whoever has the most health being the winner. The success of each bot would be calculated with the remaining health between each bot, the energy for special moves of each bot, the coordinates of each bot's location, the state of the bot's character (standing, crouching, in the air, or knocked down), and the distance between each other.

Before they can test the AI against other proven opponents, it began with a self-comparison challenge. At first, it utilized a non-opponent model to test all of the combinations on an opponent that would simply stand there as a punching bag, and then once again with an opponent that would move randomly. Once the RHEA was attuned with a random opponent, the numerous variations of opponent models were tested against each other until their win rates began to converge. And finally, they would test the RHEAOM bots against existing MCTSOM champions in the AI fighting game competition.

Within the competition, they entered the RHEAOM with the policy-gradient model as the RHEAPI as their bot of choice. Out of 10 contestants, the RHEAPI finished 2$^{nd}$. The competition was scored in two ways, in straightforward fighting, and fight speedrunning. However, as the competition progressed, the data found that the largest hp differences arrived at the lowest average frame cost per individual move set. With each character the bot utilized against the opponents, they found there was a negative correlation between the difference in health and the frame cost per movement, meaning that the RHEAOM had the capability of very aggressive, fast-acting decision making within the boundaries of a multiplayer game. The following year, an improved version of the RHEAOM-PG was introduced as the ERHEAPI with enriched opponent observation, enhanced reward design, an enlarged batch of state-action pair datasets, and introduced reliable forward model utilized by the previous year's winner, TeraThunder.  With the improvements to the opponent observation and reward design, the ERHEA won that year's competition.

This article's strengths came from its very comprehensive data definitions, with a wide variety of test subjects, live testing with other bots, and plenty of graphs with bot win percentages. On the topic of the bots, every bot was slightly different than the last, meaning that there were a large variety of opponents with varying algorithms to test against the RHEA bot. Furthermore, there were multiple ways to measure the proficiency of the bot in the areas of real-time decision making.  However, where the article struggled was in the definition of mathematical terms used to calculate the viability of strategies within a move set, and the descriptions of them did not clearly define all of the terms within a given formula.  Likewise, calculations were not utilized to decipher the results of individual population testing.  And though there were no shortage of competitors for the RHEA bot, the sample size of rounds competed in was relatively low.

In conclusion, this article was a presentation of the rolling horizon evolution algorithm AI and its capability to perform real-time decision making in a zero-sum multiplayer game.  It was hypothesized and proven to outperform state of the art, Monte Carlo Tree Search based fighting bots.  It was shown to efficiently find weaknesses within the opponents' move sets and even win in competitive tournaments against them.  However, while it managed to win, it still didn't completely defeat every opponent that was utilizing the MCTS, noting the potential for expanding the RHEA with a model based deep reinforcement learning algorithm.  And while the opponent model was effective against most MCTS bots, it relied heavily on opponent predictability, meaning it needed to expand its potential for long short-term memory, less restricted by real-time constraints.  And in the future, AI research would be keen to develop a more generalized approach to real-time decision making in a variety of different scenarios outside of a one-on-one fighting game.