

Analysis of Markov Chains to Predict Human Behavior in Fighting Games

Richard Fernando Heise
Ferreira
rfhferreira@inf.ufpr.br
Federal University of Paraná
Curitiba, Paraná, Brazil

Gabriel de Oliveira Pontarolo
gopontarolo@inf.ufpr.br
Federal University of Paraná
Curitiba, Paraná, Brazil

Vinicius Fulber-Garcia
vinicius@inf.ufpr.br
Federal University of Paraná
Curitiba, Paraná, Brazil

Abstract

This ongoing work explores the use of m -th order Markov chains to predict human behavior in fighting games, specifically during *okizeme* situations in *Street Fighter III: 3rd Strike*. Expanding on previous research that successfully applied Markov chains to predict human actions in simpler games like Rock, Paper, Scissors, this study investigates whether such statistical models can adapt to a more complex, asymmetric environment. A web-based simulation was developed to collect player data on offensive and defensive choices against an adaptive Multi-Markov chain AI. Preliminary results indicate a compelling dichotomy: while human players show a higher win rate in isolated attacking interactions (above 50%), the AI demonstrates a statistically significant advantage in Risk vs. Reward (RVR) optimization over time, leading to an overall round win rate exceeding 55%. These early findings suggest that, despite the complexity of fighting games, m -th order Markov chains can effectively model and predict player actions by exploiting cyclic behavioral patterns.

Keywords

Markov Chain, Stochastic models, Fighting Games

ACM Reference Format:

Richard Fernando Heise Ferreira, Gabriel de Oliveira Pontarolo, and Vinicius Fulber-Garcia. 2026. Analysis of Markov Chains to Predict Human Behavior in Fighting Games. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

1 Introduction

Fighting games involve complex decision-making and strategic guessing. A critical moment in these games is the *okizeme* (wake-up) situation, which occurs when a knocked-down player gets up while the attacker possesses a significant positional advantage. During this state, both players must evaluate their Risk vs. Reward (RVR) from a predefined set of offensive and defensive options. Predicting opponent actions in this scenario is notably difficult, raising the question of whether a statistical model can effectively anticipate these decisions.

One approach is to model this process using m -th order Markov chains, which calculate the probability of the next state based on the previous m states [2]. In the context of *okizeme*, a player's next move can be estimated by combining their last m wake-up choices and the associated RVR. Such models have successfully exploited cyclic human behavioral patterns in simultaneous games like Rock, Paper, Scissors [6, 7].

This ongoing study investigates the application of an m -th order Markov chain to predict human behavior during *okizeme*. Unlike Rock, Paper, Scissors, fighting games present a highly asymmetric

and dynamic environment. To collect data and test the model, a web-based simulation is being utilized where human players interact with an adaptive Markov-based AI.

Currently, our focus is on gathering player data to evaluate if the AI can leverage short-term behavioral patterns to achieve a measurable advantage in RVR evaluations over time. Future phases of this research aim to determine the model's predictive power and its limitations when applied to more complex, less restricted decision-making environments.

2 Background

2.1 Street Fighter III: 3rd Strike

Street Fighter III: 3rd Strike is a complex fighting game built on strategic guessing and risk-reward evaluation. While the game features numerous mechanics such as parries, blocks, and throws, this study isolates the *okizeme* (wake-up) scenario following a knockdown. In this state, the attacking player gains a temporary positional advantage to pressure the waking opponent. Both players are faced with a restricted but potent set of offensive and defensive options. The inherent volatility and asymmetrical advantage of 3rd Strike's *okizeme* interactions make it an ideal environment for modeling unpredictable human behavior and decision-making under pressure.

2.2 Markov Chains

Markov chains are stochastic models that capture uncertainty in state transitions. In an m -th order Markov chain, the probability of transitioning to the next state depends exclusively on the previous m states. The transition probability into a given state x_n is mathematically expressed by:

$$\begin{aligned} \Pr(X_n = x_n \mid X_{n-1}^1 = x_{n-1}^1) \\ = \Pr(X_n = x_n \mid X_{n-1}^{n-m} = x_{n-1}^{n-m}), \quad n > m \end{aligned} \quad (1)$$

Where $X_{n-1}, X_{n-2}, \dots, X_1$ are the random variables corresponding to the previous states. This study employs adaptive (non-stationary) Markov chains, meaning the state transition probabilities are not constant; they dynamically evolve and update based on the continuous data fed into the system by the players' actions.

3 Related Work

Predicting and exploiting human decision-making using stochastic models has shown promising results in simplified environments. Notably, Wang et al. [6] demonstrated that a Multi-AI architecture

based on m -th order Markov chains could effectively predict human actions in iterated Rock, Paper, Scissors (RPS). Their adaptive system utilized multiple single-length models to capture behavioral patterns across different memory depths.

Other research has broadened the scope of human behavior modeling, exploring predictive accuracy in simultaneous-move games [8] and utilizing Reinforcement Learning to simulate human-like opponents in fighting games [4].

Building on these foundations, this ongoing work investigates the applicability of Markov-based predictive models in a significantly more complex domain. Unlike RPS, which involves symmetric interactions and simultaneous discrete choices, *okizeme* scenarios introduce role asymmetry (attacker versus defender) and a broader, more nuanced action space, making behavioral prediction inherently more challenging.

4 Problem Modeling

4.1 Okizeme Simulator

The study models a Ken vs. Ken mirror match in *Street Fighter III: 3rd Strike*, utilizing a representative subset of mechanics. The attacker chooses from five options (Throw, Delayed Combo, Meaty High, Meaty Low, Shimmy), while the defender has seven options (Block, Guard-Jump, Button, Throw, Reversal, High Parry, Low Parry).

Interactions yield an attacker advantage, defender advantage, or damage. Results are normalized against a base throw damage (1.0). Players start with 8 Hit Points (HP); the first to reach 0 HP loses the round. Turn order is outcome-driven: successful attacks retain the turn, successful defenses steal it, and draws are decided by a coin toss. Table 1 summarizes the outcome matrix.

4.2 Markov-based AI

Following an architecture similar to [6], the AI employs a Multi-Markov system comprising M chains of varying orders (1 to M) to rapidly adapt to short-term behavioral patterns.

When an action a' occurs, the observation count n_{obs} for that transition in the current state s is incremented. A constant decay D is then applied to all possible actions $a \in A$, and the probability distribution $P(a|s)$ is recalculated:

$$\begin{aligned} n_{obs}(s, a') &= n_{obs}(s, a') + 1 \\ n_{obs}(s, a) &= n_{obs}(s, a) \times D, \quad \forall a \in A \\ P(a|s) &= \frac{n_{obs}(s, a)}{\sum_{a \in A} n_{obs}(s, a)} \end{aligned} \quad (2)$$

The inclusion of the decay factor D is a critical component of the AI's design, specifically tailored to handle human psychology. Unlike static systems, human players in fighting games constantly adapt, shifting their strategies when they realize they become predictable. By applying this decay, the Multi-Markov system effectively implements a "short-term memory" mechanism. It prioritizes the player's most recent behavioral patterns while gradually "forgetting" older habits that are no longer relevant, ensuring the AI remains responsive to mid-match adaptations rather than getting stuck on outdated data.

Each chain predicts the next action based on its highest transition probability. These predictions are weighted based on the chain's

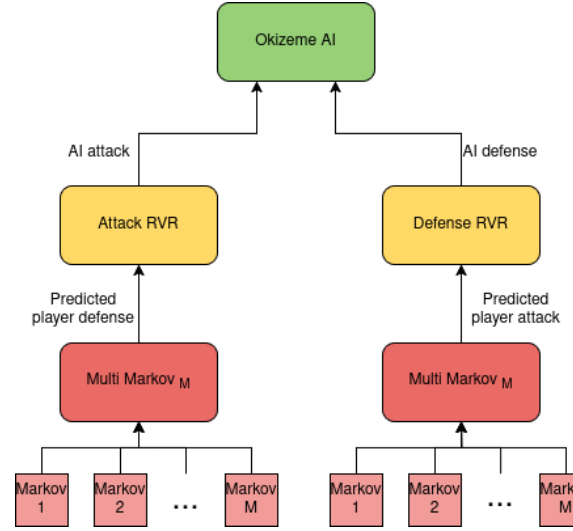


Figure 1: Okizeme AI decision-making diagram for choosing the response action.

recent accuracy (number of correct predictions in the last $\lceil M/2 \rceil$ steps). To select the optimal response action a_r , the AI calculates an expected damage reward $R(a_r)$ by summing the damage outcomes (from Table 1) against the top weighted predictions (S_{top}):

$$R(a_r) = \sum_{a_m \in S_{top}} \text{Damage}(a_m, a_r), \quad \forall a_r \in A \quad (3)$$

The AI selects the action maximizing $R(a_r)$. For our ongoing tests, parameters are set to $M = 7$ (maximum possible player actions) and $D = 0.9$. The system runs two parallel Multi-Markov Chains to independently evaluate offensive and defensive RVR, as illustrated in Figure 1.

5 Preliminary Results

To validate the proposed architecture, an online web-based game was developed to simulate *okizeme* scenarios. Preliminary data was collected from 24 valid sessions (out of 119 attempts), comprising *First to 5* and *First to 10* sets, which yielded approximately 150 decisions per set. The data collection framework is available at: <https://github.com/richardheise/okizeme-predictor>.

Early analysis presents an interesting dichotomy. As shown in Figures 2 and 3, human players achieved a win rate above 50% in isolated offensive interactions, effectively outperforming the AI on a turn-by-turn basis. Defensively, humans performed worse, winning only 22% of interactions.

Despite humans winning more individual offensive interactions, the AI won over 55% of the overall rounds, peaking at an average 58% win rate as sets progressed (Figure 4). This discrepancy is explained by the Risk vs. Reward (RVR) optimization. While humans won more often, the AI prioritized high-yield, safer decisions. A preliminary Wilcoxon rank-sum test comparing the AI's average attacking RVR (0.68) against the human average (0.10) yielded a p -value of 0.0002. This statistically significant difference suggests that the multi-Markov AI successfully exploits short-term behavioral

Table 1: Outcome matrix from the attacker's perspective for each interaction

Offense / Defense	Block	Guard Jump	Button	Throw (def)	Reversal	High Parry	Low Parry
Throw	Win (1.0 dmg)	Lose	Win (1.0 dmg)	Lose	Lose (1.4 dmg)	Win (1.0 dmg)	Win (1.0 dmg)
Delayed Combo	Lose	Win (3.5 dmg)	Lose (3.5 dmg)	Lose (1.0 dmg)	Win (3.0 dmg)	Lose (3.5 dmg)	Win (3.5 dmg)
Meaty High	Lose	Lose	Win (3.5 dmg)	Win (3.5 dmg)	Lose (1.4 dmg)	Lose (3.5 dmg)	Win (3.5 dmg)
Meaty Low	Lose	Lose	Win (2.5 dmg)	Win (2.5 dmg)	Lose (1.4 dmg)	Win (2.5 dmg)	Lose (3.5 dmg)
Shimmy	Lose	Lose	Lose (3.0 dmg)	Win (3.0 dmg)	Win (3.5 dmg)	Lose	Lose

patterns to make consistently better offensive decisions over time, leading to round victories despite losing individual interactions.

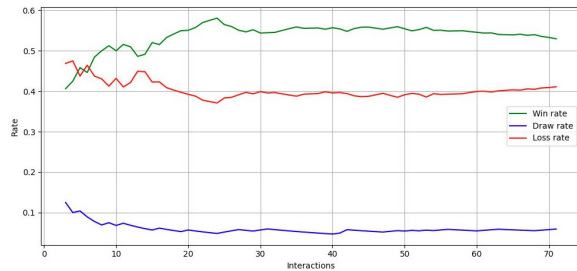


Figure 2: Average Human Win, Draw and Loss Rate when Attacking.

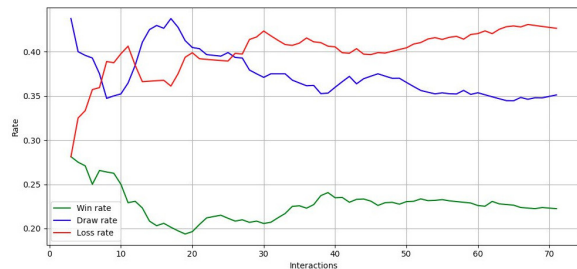


Figure 3: Average Human Win, Draw and Loss Rate when Defending.

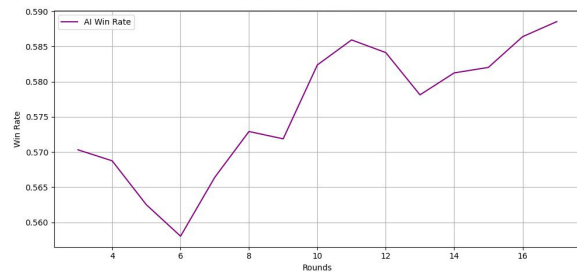


Figure 4: Average AI Win Rate per Round.

6 Conclusion and Future Work

This ongoing study investigates whether human behavior during *okizeme* scenarios in fighting games can be effectively predicted using m -th order Markov chains. Preliminary data suggests that, despite the asymmetric and dynamic nature of *Street Fighter III: 3rd Strike*, a Multi-Markov AI can successfully exploit cyclic behavioral patterns.

As a work in progress, several limitations are currently being addressed. The simplified simulator assumes idealized conditions, omitting execution errors, timing variability, and complex mechanics like meter management. The static reward matrix and restricted action set also limit strategic diversity.

Moving forward, our research will focus on expanding the simulation environment to incorporate these missing variables and conducting controlled experiments with a larger, more experienced player base. Furthermore, we intend to evaluate this Markov-based architecture as a baseline against more sophisticated predictive models, such as Hidden Markov Models (HMMs) [1], Reinforcement Learning [5], or Behavior Cloning [3], aiming for more robust and generalized behavioral predictions in complex game states.

References

- [1] Mario Benevides, Isaquie Lima, Rafael Nader, and Pedro Rougemont. 2014. Using hmm in strategic games. In *Proceedings of the Workshop on Developments in Computational Models (DCM 2013)* (EPTCS). Vol. 144, 73–84. doi: [10.4204/EPTC.S.144.6](https://doi.org/10.4204/EPTC.S.144.6).
- [2] Ka Chan, C. Lenard, and Terence Mills. 2012. An introduction to markov chains. (2012). doi: [10.13140/2.1.1833.8248](https://doi.org/10.13140/2.1.1833.8248).
- [3] José García, Carlos Castro, and Carlos Valle. 2025. Ensemble approach to adaptable behavior cloning for a fighting game ai. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (CIARP 2024)* (Lecture Notes in Computer Science). Ruber Hernández-García, R. J. Barrientos, and Sergio A. Velastin, editors. Vol. 15368. Springer, Cham, 16–30. doi: [10.1007/978-3-031-76607-7_2](https://doi.org/10.1007/978-3-031-76607-7_2).
- [4] Matheus R. F. Mendonça, Heder S. Bernardino, and Raul F. Neto. 2015. Simulating human behavior in fighting games using reinforcement learning and artificial neural networks. In *2015 14th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, 152–159. doi: [10.1109/SBGames.2015.25](https://doi.org/10.1109/SBGames.2015.25).
- [5] Inseok Oh, Seungeun Rho, Sangbin Moon, Seongho Son, Hoyil Lee, and Jinyun Chung. 2019. Creating pro-level ai for a real-time fighting game using deep reinforcement learning. *arXiv preprint arXiv:1904.03821*. doi: [10.48550/arXiv.1904.03821](https://doi.org/10.48550/arXiv.1904.03821).
- [6] Liheng Wang, Wei Huang, Yuzhe Li, et al. 2020. Multi-ai competing and winning against humans in iterated rock-paper-scissors game. *Scientific Reports*, 10, 13873. doi: [10.1038/s41598-020-70544-7](https://doi.org/10.1038/s41598-020-70544-7).
- [7] Zheng Wang, Bin Xu, and H. J. Zhou. 2014. Social cycling and conditional responses in the rock-paper-scissors game. *Scientific Reports*, 4, 5830. doi: [10.1038/srep05830](https://doi.org/10.1038/srep05830).
- [8] James R. Wright and Kevin Leyton-Brown. 2019. Level-0 models for predicting human behavior in games. *Journal of Artificial Intelligence Research* 64, 64, 1, (Jan. 2019), 357–383. doi: [10.1613/jair.1.11361](https://doi.org/10.1613/jair.1.11361).