Master' thesis

Can AI Resign in an Appropriate Position?

Pan Shize

Supervisor: Hiroyuki Iida

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
Information Science

February, 2025

Supervised by

Prof. Dr. Hiroyuki Iida

Reviewed by

Prof. Dr. Kokolo Ikeda

Prof. Dr. Kiyoaki Shirai

Prof. Dr. Shinobu Hasegawa

Abstract

This study presents two distinct research projects: the development of an AlphaZero-based Heian Shogi AI system and the optimization of a resignation mechanism in Shogi AI.

The first project focuses on reconstructing and analyzing Heian Shogi, an early form of Shogi, using reinforcement learning through self-play. Unlike modern Shogi, Heian Shogi features different board sizes, unique piece types, and alternative movement rules, making it an intriguing subject for AI-based historical game studies. To model strategic decision-making in this variant, the Heian Shogi AI employs deep residual networks and Monte Carlo Tree Search (MCTS), allowing it to learn optimal strategies from self-play without human game data. The effectiveness of this AI is evaluated using key performance metrics, including game length, branching factor, and strategic depth, contributing to a deeper understanding of how historical game mechanics influenced strategic evolution.

The second project investigates the integration of an optimized resignation mechanism into modern Shogi AI. Existing Shogi AI systems often struggle to determine appropriate resignation points, either resigning prematurely or prolonging lost positions. This study proposes a resignation threshold based on the maximum advantageous score achieved by the losing side, ensuring that AI resigns at a moment that aligns with strategic decision-making principles. The evaluation framework leverages the Suisho5-YaneuraOu engine and Game Refinement (GR) theory to analyze key factors such as game length, acceleration of uncertainty resolution, and strategic balance. Experimental results demonstrate that higher-skill AI players with the resignation mechanism exhibit more efficient decision-making, reducing unnecessary

game prolongation while maintaining competitive depth.

Through rigorous experimentation and analysis, this research contributes to AI development in strategic board games from both historical and modern perspectives. The Heian Shogi AI project enhances our understanding of ancient game strategies, while the resignation mechanism project improves AI decision-making in competitive play. Future research will explore the generalizability of these findings across other board games, further advancing AI applications in game theory, historical simulation, and strategic decision-making.

Keyword: *Shogi AI*; *Resignation Mechanism*; *Human and AI Resignation Behaviors*; *Game Refinement Theory*; *Machine Learning*

Acknowledgments

First of all, I would like to thank my supervisor, Professor Hiroyuki Iida. He has given me full guidance in my research and study. Professor Iida is knowledgeable, polite and courteous. He is always busy with work and has no time even for holidays, but he still often takes some spare time to guide me and give me some advice that inspires my research, for which I am deeply grateful. Professor Iida loves his work and treats research meticulously and strictly. I will cherish and treasure these guidance as precious memories of my life. Professor Iida also cares about my real life, like a father. What I learned from Professor Iida is not only knowledge, but also the attitude towards scientific and technological research. From the bottom of my heart is sincerely thankful for his kindness and guidance.

Secondly, I would like to thank my parents, who gave birth to me. Without their support, I would never be able to complete my studies. They gave me living expenses and tuition fees and guided me when I encountered difficulties in my research, which enabled me to maintain a good attitude to cope with various challenges in study and research.

I also thanks Professor Kokoro Ikeda. He asked me some profound questions related to my research and inspired me to complete it. Professor Ikeda is also very rigorous in his attitude towards research, and he also cares about students. I have learned many important things from Professor Ikeda, and I am very grateful to him and respect him in many ways.

Finally, I would like to thank all members of our lab. Thank you for your help and support in my research, thank you for everything. Every moment I spent with you was unforgettable and I want to cherish every emotion that dissipated here.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The rapid advancement of artificial intelligence (AI) has profoundly transformed multiple domains, ranging from image recognition and natural language processing to strategic decision making. Despite achieving superhuman performance in two-player perfect-information games such as Chess, Go, and Shogi [1], AI systems still struggle with tasks requiring human-like intuition and contextual reasoning, particularly in dynamic environments where decisions must balance short-term and long-term consequences [2].

A fundamental challenge in AI research is to develop decision-making mechanisms that align with human cognition. Traditional AI approaches rely on predefined evaluation functions and heuristic rules, limiting their adaptability in complex scenarios where strategic decisions cannot be strictly rule-based. Recent studies emphasize the need for AI to expand beyond pattern recognition and incorporate structured reasoning, including causal inference and learning from implicit human behaviors [3]. This shift is especially crucial in strategic environments, where AI must assess uncertain situations, anticipate opponent strategies, and determine optimal actions in real time.

Monte Carlo Tree Search (MCTS) has emerged as one of the most effective frameworks for AI decision-making, addressing the limitations of heuristic-based approaches. Since its introduction in 2006, MCTS has been successfully applied to

board games such as Go and Shogi, as well as modern video games, demonstrating its adaptability in different gaming environments [4]. By balancing exploration and exploitation through iterative simulations, MCTS has become a cornerstone of AI game research. However, while MCTS enables strong tactical play, it does not inherently model high-level strategic decisions such as resignation, leaving AI vulnerable to inefficient decision-making in lost positions.

Board games provide an ideal testbed for studying AI decision making due to their deterministic nature and well-defined strategic complexity [5]. Among them, Japanese Shogi stands out as particularly challenging due to its vast branching factor and unique piece-drop mechanism, which significantly expands the game's decision space. Unlike Chess and Go, where captured pieces are permanently removed, Shogi allows reintroduction of captured pieces, leading to longer and more complex game trajectories[6]. This distinct rule set presents additional challenges for AI, particularly in determining when a game is no longer winnable and resignation is the most strategic option.

Resignation plays a crucial role in high-level gameplay, as it reflects a player's ability to assess the board state and make a strategic decision to end the game efficiently rather than prolonging an inevitable loss. Human players typically resign based on a combination of positional evaluation, tactical foresight, and psychological factors. However, existing Shogi AI engines often struggle with resignation, either prolonging games in clearly lost positions or resigning prematurely due to rigid evaluation functions. Studies on the AlphaZero AI system, for example, have shown that its resignation behavior can be inconsistent, with instances of premature resignations in situations where human players might continue seeking counterplay [7].

AI-assisted decision-making has been extensively studied in various domains, with recent research emphasizing the need for AI to optimize human-centric objectives alongside decision accuracy [8]. Traditional decision-support AI systems focus on maximizing immediate outcomes, often overlooking long-term effects such as learning, cognitive engagement, and strategic growth[9]. This limitation is particularly evident in strategic games, where an AI's decision to resign should ideally reflect

not only immediate position evaluation but also the evolving strategic landscape of the game. While existing AI decision models have improved accuracy through reinforcement learning and policy optimization, they often fail to account for the broader context in which decisions occur. AI resignation mechanisms, for instance, still largely rely on static evaluation thresholds rather than dynamically adapting to the game's progression.

A promising direction for enhancing AI decision-making lies in augmenting human intelligence with AI-driven optimization techniques. Hybrid AI-human decision-making models, such as IndigoVX, aim to integrate human expertise with AI's data-driven insights to optimize decision-making in real time [10]. This approach leverages iterative feedback loops where AI refines its strategy based on human contextual knowledge, allowing for more adaptive and goal-oriented decision-making. Such hybrid models demonstrate significant potential in domains requiring strategic reasoning, including board games and business applications. However, their effectiveness in optimizing resignation mechanisms for Shogi remains largely unexplored.

Research on AI resignation mechanisms has explored multiple approaches to optimizing strategic disengagement. Bhatt and Sargeant proposed algorithmic resignation, which emphasizes strategic disengagement when further decision-making becomes inefficient [11]. Similarly, Wirth and Furnkranz demonstrated that integrating expert annotations into Chess AI systems improves resignation decisions by incorporating human knowledge [12]. While these methods provide valuable insights, most existing AI resignation strategies rely on predefined score-based thresholds, which fail to account for the dynamic nature of in-game evaluation.

Given these limitations, there is a growing need for a more refined resignation mechanism in Shogi AI that balances strategic awareness, efficiency, and game refinement principles. A well-designed resignation model should prevent AI from unnecessarily prolonging lost games while ensuring it does not resign prematurely. Moreover, an effective resignation mechanism must adapt dynamically to evolving game states, making decisions based on in-game factors rather than relying solely on static evaluation thresholds.

These challenges motivate the present study, which explores a systematic approach to AI resignation in Shogi. By leveraging game refinement theory, this research aims to establish a principled framework for resignation, ensuring that AI systems make decisions that are strategically sound, contextually appropriate, and aligned with human-like gameplay.

1.2 Problem Statement and Research Questions

A key strategic aspect of Shogi is the resignation, which represents a player's judgment of the board state and their anticipation of future [13]. In human play, resignation is often based on intuitive evaluation and situational dynamics, blending tactical assessments with psychological factors. However, existing Shogi AI systems predominantly rely on rigid, rule-based algorithms to trigger resignation, typically using fixed evaluation thresholds, such as predefined score differences. While these methods effectively automate decision-making processes, they fail to capture the nuanced reasoning exhibited by human players, often leading to resignation that appears abrupt, delayed, or strategically suboptimal.

The development of a robust resignation mechanism is critical to enhancing Shogi AI performance. By leveraging insights from game refinement theory and empirical analysis, this study aims to create a more strategic and human-like resignation framework. These advancements not only contribute to AI research in strategic decision-making but also have broader implications for AI applications in complex problem-solving environments.

As such, the research questions of this thesis explored upon can be summarized as follows:

- How to make AI Resign in an Appropriate Position like human beings?

1.3 Research Objectives

To address these limitations, this study explores a resignation mechanism based on the maximum advantageous score achieved by the losing player. By defining a dynamic threshold for resignation, the AI resigns when the score disadvantage surpasses this threshold, improving game efficiency and enhancing strategic engagement. Experimental results demonstrate that this mechanism significantly shortens game length and accelerates gameplay, particularly for higher-skill AI players. The study reveals that higher-skill players are more adept at leveraging resignation timing to avoid prolonged losing positions, thus enhancing the overall experience by maintaining a natural game flow.

Despite the effectiveness of the maximum score threshold approach, challenges remain in optimizing resignation behavior, especially in scenarios involving AI agents surpassing human expertise. The current method relies on static threshold definitions, which may not fully adapt to the evolving complexity of game states. Future research directions include incorporating additional contextual factors to refine the resignation mechanism, ensuring that AI players resign at moments that align more closely with human strategic decision-making.

Future research aims to refine the resignation mechanism by integrating an opponent-aware perspective. Effective resignation requires not only the recognition of a lack of winning opportunities but also an assessment of whether the opponent acknowledges their advantageous position. The resignation decision should ideally occur when both conditions are met, ensuring a more context-sensitive and human-like gaming experience.

Chapter 2

Building Shogi AI

2.1 History of Shogi

Shogi, often referred to as Japanese chess, is one of the oldest and most revered board games in Japan, with a rich history that spans over a thousand years. The origins of Shogi can be traced back to ancient India, where it evolved from a game called Chaturanga, which is widely regarded as the precursor to many modern chess variants [14]. As the game traveled through different regions, it underwent significant transformations before taking root in Japan, where it developed its unique characteristics and rules.

The earliest records of Shogi in Japan date back to the Heian period (794–1185), when it was primarily played among the aristocracy and the imperial court [15]. During this period, Shogi was considered more than just a game; it was a tool for strategic thinking and a form of entertainment that reflected the cultural and intellectual pursuits of the time. The version of Shogi played in the Heian period was different from the modern variant, with changes in the board size and the rules over the centuries.

The Edo period (1603–1868) marked a significant turning point in the history of Shogi [16]. Under the Tokugawa shogunate, the game became popular among the samurai class, and Shogi tournaments were held regularly. The establishment of the official title of Meijin (Master) further formalized the game and its competitive nature. It was during this period that the modern rules of Shogi began to take shape,

including the introduction of the drop rule, which allows captured pieces to be reused by the capturing player-a unique feature that distinguishes Shogi from other chess variants.

In the modern era, Shogi has evolved into a widely popular game across Japan and beyond. Professional Shogi leagues were established in the early 20th century, with players competing for prestigious titles such as Ryuo (Dragon King), Meijin (Master) and Kisei (Go Sage). These titles are fiercely contested, and Shogi professionals dedicate their lives to mastering the game's complexities and nuances.

Figure 2-1: Modern Shogi Pieces and Board (Image credit: licensed under Creative Commons Zero)

Technological advancements in the 21st century have further transformed the Shogi landscape. The development of Shogi AI programs has revolutionized the way the game is studied and played. AI engines, capable of analyzing millions of positions and predicting optimal moves, have become invaluable tools for both professional players and researchers. The advent of AI has also led to the emergence of new challenges and questions, particularly regarding the interplay between human intuition and machine precision in strategic decision-making.

2.2 Game Refinement (GR) Theory

Game Refinement (GR) theory provides a mathematical framework for assessing the level of engagement and challenge in strategic games. By leveraging the game progress model, it quantifies the balance between skill and uncertainty, helping researchers evaluate the entertainment value of different game structures. GR theory has been successfully applied to analyze the evolution of various board games, demonstrating how their structural adjustments contribute to sustained player engagement [17, 18]. Notably, sophisticated games typically fall within a refined range of game refinement values, specifically $GR \in [0.07, 0.08]$, indicating an optimal balance between predictability and strategic depth.

The concept of game refinement differs significantly from classical game theory. While traditional models, such as von Neumann's minimax theory[19], focus on optimizing decision-making processes and determining the probability of winning based on positional advantages, GR theory shifts the focus toward the overall appeal and entertainment value of a game. Instead of merely identifying optimal strategies for victory, GR theory aims to assess how well a game maintains player engagement. This approach provides a systematic method for classifying and analyzing games based on their refinement values, offering valuable insights into how different game mechanics influence player experience.

Moreover, GR theory serves as a powerful tool for exploring the historical development of games. It enables researchers to examine how traditional games have evolved over time to maintain their strategic depth and player appeal. By applying GR analysis to various game variants, it is possible to identify patterns in rule modifications and their impact on game dynamics. This perspective offers a broader and more structured understanding of how games adapt to player expectations while preserving core strategic elements.

Beyond board games, GR theory has significant implications in various other fields, including video games, business strategies, and education. Since engagement is a fundamental metric in evaluating interactive experiences, the principles of GR

theory can be used to measure participation levels in diverse contexts. The ability to quantify and refine engagement makes GR a valuable tool for optimizing decision-making processes, enhancing user experience in digital environments, and improving strategic planning methodologies. As computational models of game dynamics continue to advance, GR theory is expected to play an increasingly prominent role in shaping entertainment, decision-making, and learning frameworks. Its unique ability to bridge game mechanics with user experience highlights its importance in both theoretical and practical applications.

2.2.1 Game Refinement Theory Concept

The progression of game information is represented by a function $x(t)$, which quantifies the cumulative uncertainty resolved by time t. This function is normalized such that:

$$x(0) = 0 \quad \text{and} \quad x(T) = 1$$

where T represents the total length of the game, typically measured in moves. Normalizing the function ensures consistency across games of varying lengths and complexities, allowing meaningful comparisons of uncertainty resolution dynamics. The parameter n (where $1 \leq n \in N$) is the number of possible options.

The rate of uncertainty resolution during gameplay is modeled as:

$$x'(t) = \frac{n}{t}x(t) \tag{2.1}$$

where n denotes the number of feasible options available to the player at any stage of the game. This equation encapsulates the intuitive idea that the rate at which uncertainty is resolved depends both on the existing state of resolved uncertainty ($x(t)$) and the elapsed time (t). It captures the interplay between the complexity of decision-making and the progression of the game.

Solving the differential equation (2.1) yields the following solution:

$$x(t) = \left(\frac{t}{T}\right)^n \tag{2.2}$$

This expression describes how uncertainty is progressively resolved over time. The parameter n significantly influences the curve, with higher values of n corresponding to more complex games where uncertainty is resolved more gradually.

The acceleration of uncertainty resolution reflects changes in the rate at which information is acquired during gameplay. It is given by the second derivative of $x(t)$:

$$x''(t) = \frac{n(n-1)}{T^n} t^{n-2} \tag{2.3}$$

This formula describes how the pace of uncertainty resolution evolves throughout the game. At the conclusion of the game ($t = T$), the acceleration simplifies to:

$$a = \frac{n(n-1)}{T^2} \tag{2.4}$$

This acceleration metric serves as a key indicator of the dynamics of gameplay, providing insights into how rapidly information is resolved, particularly as the game progresses toward its conclusion.

Figure 3-3 illustrates a model of candidate selection for changes based on skill and chance. This illustration shows that skillful players would consider a set of fewer plausible candidates (say b) among all possible moves (say B) to find a move to play and that there is a core part of its original game with branching factor B. The core part is a stochastic game with a smaller branching factor b since it is assumed that each of the candidates b may be equally selected.

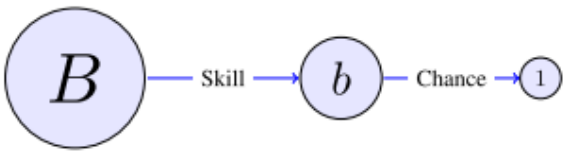

Figure 2-2: An illustration of move selection model based on skill and chance

The Game Refinement (GR) value is derived from the square root of the acceleration of uncertainty resolution:

$$GR = \sqrt{a} = \sqrt{\frac{n(n-1)}{T^2}} \tag{2.5}$$

Empirical studies have demonstrated that GR values for well-designed games typically fall within the range:

$$GR \in [0.07, 0.08]$$

This range reflects an optimal balance between skill and chance, ensuring that games remain engaging without being overly reliant on either extreme. Games with GR values outside this range may either feel overly simplistic (too reliant on chance) or excessively challenging (too reliant on skill).

2.2.2 Game Refinement Model in Board Game (Shogi)

In practical terms, the branching factor (B) represents the average number of feasible moves available at each turn, while the total game length (D) represents the typical number of moves in a complete game. Substituting B for n and D for T, the acceleration can be approximated as:

$$a = \frac{B}{D^2} \tag{2.6}$$

Consequently, the GR value becomes:

$$GR = \sqrt{\frac{B}{D^2}} = \frac{\sqrt{B}}{D} \tag{2.7}$$

This approximation simplifies the calculation of GR for real-world board games and facilitates comparisons across games with varying complexities and lengths.

The velocity of game progress, denoted as v, represents the rate at which uncertainty is resolved during gameplay. It is defined as:

$$v = \frac{1}{2}\frac{B}{D} \tag{2.8}$$

This measure provides insights into the tempo of gameplay, capturing how quickly players perceive progress toward resolving uncertainty. A higher velocity suggests a faster-paced game, while a lower velocity indicates a more deliberate, strategic experience.

The average number of moves required to achieve a single meaningful goal in the game is given by the reciprocal of the velocity:

$$N = \frac{1}{v} \tag{2.9}$$

This metric highlights the frequency of significant decisions made by players. Games with a smaller N value emphasize quick, impactful decisions, while larger N values allow for more extended strategic planning.

Popular board games such as Chess, Shogi, and Go demonstrate well-defined GR values, emphasizing their balanced design. Table 2.1 summarizes their branching factors (B), total game lengths (D), and corresponding GR values:

Table 2.1: Game Refinement Measures for Popular Board Games

Game	**Branching Factor** (B)	**Game Length** (D)	**GR** ($\sqrt{a}$)
Chess	35	80	0.074
Shogi	80	115	0.078
Go	250	208	0.076

The results validate the applicability of GR theory across diverse games, demonstrating its effectiveness in quantifying the delicate balance between challenge and engagement. By maintaining GR values within the optimal range, designers can ensure that games provide a compelling and balanced experience, fostering sustained interest and enjoyment.

For sophisticated board games such as Chess, Shogi, and Go, it is assumed that there exists a reasonable zone for acceleration (a), which is between $0.07 - 0.08$. This zone is very meaningful.

The intersection (Figure 2-3) of acceleration and jerk is the point where the maximum amount of achievement is greater than the discomfort (t_1), and after t_1, the

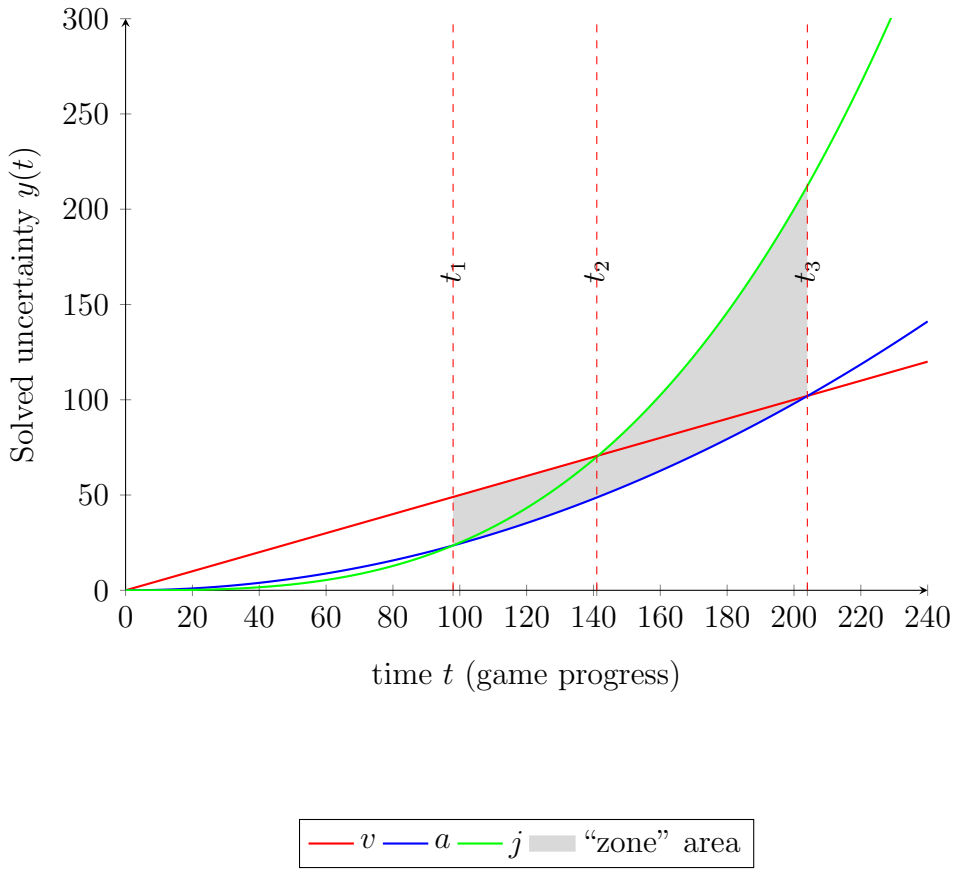

Figure 2-3: The cross point between the line with velocity v, curve with acceleration a and curve with jerk j. t_1; t_2 and t_3 represents the bound for effort, achievement, and discomfort, respectively.

discomfort will be greater than the achievement. The intersection of velocity and jerk is the point where effort is greater than discomfort (t_2), and the intersection of velocity and acceleration is the point where effort is more excellent than achievement(t_3). The intersection interval ensures a reasonable zone for game length [20].

2.3 Suisho AI

To generate game data for AI self-play simulations, this study utilizes the Suisho5-YaneuraOu engine, a highly regarded Shogi engine developed by Japanese programmer Hiroshi Yamashita. YaneuraOu has become a pivotal tool in the Shogi community due to its exceptional performance and the innovative application of deep

learning technology. The engine is capable of competing with top human players and has demonstrated outstanding results on various online platforms and Shogi competitions. Its significance extends beyond gameplay, as its open-source nature provides researchers with a powerful framework for exploring artificial intelligence and machine learning in Shogi. The emergence of YaneuraOu represents a major milestone in the advancement of computer Shogi and has garnered widespread attention and discussion within the field.

2.3.1 Evolution from AlphaGo to AlphaZero

The field of game-playing AI has witnessed rapid advancements, particularly with the development of deep reinforcement learning techniques. A pivotal breakthrough occurred with AlphaGo, the first AI system to defeat professional human Go players, leveraging a combination of deep neural networks and Monte Carlo Tree Search (MCTS) [21]. AlphaGo employed a supervised learning approach, initially trained on expert human gameplay before transitioning to self-play reinforcement learning to refine its strategy.

Following this success, AlphaGo Zero was introduced, eliminating reliance on human game data and instead learning purely through self-play [22]. This marked a significant departure from traditional AI approaches, demonstrating that AI could surpass human knowledge entirely through reinforcement learning alone. By training from scratch with only the game rules, AlphaGo Zero was able to outperform all previous iterations, proving that human data was not necessary for achieving superhuman performance.

The culmination of this research was AlphaZero, a generalized version of AlphaGo Zero that extended its learning capabilities to multiple board games, including Chess and Shogi, without requiring game-specific modifications [23]. Unlike previous iterations, AlphaZero learned all three games from scratch using the same training pipeline, refining its strategy purely through reinforcement learning and MCTS. The transition from AlphaGo to AlphaZero demonstrated the feasibility of using a single AI framework for mastering multiple strategic games, laying the foundation for

further advancements in game-playing AI.

2.3.2 Implementation of AlphaZero's Framework in Suisho5

Inspired by the success of AlphaZero, modern Shogi AI engines such as Suisho5-YaneuraOu have adopted similar methodologies. The Suisho5 engine employs deep neural networks alongside MCTS, faithfully implementing the reinforcement learning methodology outlined in the AlphaZero framework [23]. By replacing traditional handcrafted evaluation functions with deep residual neural networks, the engine effectively refines its decision-making process through continuous self-play, optimizing its strategy autonomously.

One of the primary advantages of Suisho5 is its exceptional speed in generating strategies and its computational efficiency. The engine's systematic evaluation of player strength rates its gameplay intensity at an impressive R4500, significantly surpassing the performance of human grandmasters. This remarkable capability makes Suisho5 an ideal tool for generating high-quality Shogi game data for research purposes.

Furthermore, Suisho5 implements a reinforcement learning pipeline similar to AlphaZero's training paradigm. The system alternates between self-play data generation, neural network training, and policy updates, ensuring that each iteration improves upon the last. This iterative optimization allows Suisho5 to develop sophisticated strategies without requiring any manually curated game data.

2.3.3 Visualization and Analysis

Since YaneuraOu itself lacks a graphical user interface (GUI), the software ShogiGUI is employed for visualization. ShogiGUI offers a comprehensive interface that displays various game-related information, including move scores, board visualizations, and other key metrics. This integration of Suisho5 with ShogiGUI enables researchers to conduct in-depth analyses of game dynamics, player strategies, and AI performance, facilitating a detailed exploration of Shogi gameplay.

2.4 Heian Shogi and Heian Shogi AI

Heian Shogi, an early form of Japanese Shogi, dates back to the Heian period (794–1185) and represents a significant stage in the evolution of the modern game [24]. As the predecessor of contemporary Shogi, Heian Shogi features distinct rules and piece types that have since evolved or been phased out. The study and simulation of Heian Shogi using artificial intelligence (AI) techniques offer valuable insights into the historical development of strategic board games and provide a deeper understanding of the cultural and strategic transformations that occurred over centuries.

The primary motivation for simulating Heian Shogi is to reconstruct its historical gameplay and analyze its strategic intricacies. Over time, the original rules and board configurations of Heian Shogi have been lost or replaced by modern conventions. Through AI-based simulations, it becomes possible to revive the gameplay dynamics of Heian Shogi, allowing researchers to explore how strategic decisions were made in the past and how the game gradually evolved into its present form. Moreover, by comparing Heian Shogi with modern Shogi, researchers can gain valuable insights into the impact of rule changes on the complexity, balance, and strategic depth of the game.

A key distinction between Heian Shogi and modern Shogi lies in the board size, piece types, and movement rules. While modern Shogi is played on a fixed 9×9 board, historical variants such as Heian Shogi utilized larger boards, sometimes extending beyond the standard dimensions. Additionally, the piece set in Heian Shogi included unique elements such as the *Drunken Elephant*, which was later removed in modern Shogi. Furthermore, the rules governing piece drops, a hallmark of modern Shogi, were not fully developed in Heian Shogi, leading to a different strategic landscape where captured pieces had more restricted usage. Victory conditions in Heian Shogi were also simpler, focusing on swift offensive strategies rather than the long-term positional play characteristic of modern Shogi.

The significance of studying Heian Shogi extends beyond historical curiosity; it provides meaningful contributions to various academic and technological fields. Ana-

lyzing the evolution of game rules can offer insights into how strategic decision-making has evolved over time, shedding light on fundamental aspects of cognitive science and game theory. The study of Heian Shogi also supports cultural heritage preservation by digitally archiving and simulating ancient gameplay, thereby ensuring that traditional Japanese board games are not lost to history. Additionally, the integration of AI in simulating Heian Shogi provides an opportunity to develop advanced algorithms capable of handling complex and dynamic strategic scenarios, which can be applied to broader fields such as decision science and artificial intelligence research.

Furthermore, the study of Heian Shogi contributes to educational and cultural outreach efforts by introducing traditional board games to a global audience. By developing interactive platforms that allow players to engage with historical Shogi variants, educational institutions and cultural organizations can promote an appreciation of Japanese culture and strategic thinking. Understanding the differences between Heian and modern Shogi offers a unique perspective on how game mechanics evolve to accommodate changing cultural and competitive landscapes.

In conclusion, the research and simulation of Heian Shogi using AI technologies represent an intersection of history, culture, and technology. Through the revival of ancient game mechanics and strategic principles, this study aims to bridge the gap between historical scholarship and modern computational approaches. Future research in this area will focus on refining AI algorithms to better simulate historical playstyles and expanding the application of these methods to other traditional games, contributing to a broader understanding of cultural evolution through strategic gameplay.

This project implements an AlphaZero-based Heian Shogi AI, utilizing reinforcement learning through self-play without relying on human game data. The entire system comprises several key components: game logic, Monte Carlo Tree Search (MCTS), policy-value network, and self-play data collection and training process.

In the game logic module, the board state is represented using a 2D list, with the initial configuration set according to Heian Shogi rules, where the red side is positioned at the top and the black side at the bottom. The game state is managed

by the ‘Board‘ class, which provides functionalities such as board initialization, legal move detection, state transition, and game termination checks. During self-play, the board states are stored in a ‘deque‘ data structure to maintain a history of recent moves, which helps in handling special rules such as perpetual check and repetition. Additionally, the system employs a mapping from piece strings to arrays, facilitating the conversion of board states into tensor formats suitable for neural network input.

2.4.1 Monte Carlo Tree Search (MCTS)

Monte Carlo Tree Search (MCTS) is the core decision-making module in this study, enabling the AI to efficiently navigate the vast game tree and identify optimal strategies[25]. MCTS operates by iteratively simulating potential game trajectories, balancing exploration of less-visited nodes and exploitation of high-value actions through the PUCT (Polynomial Upper Confidence Trees) formula. Each node in the search tree is associated with several key attributes: $N(s, a)$, representing the visit count of an action a at state s; $Q(s, a)$, the average value estimate based on previous simulations; $P(s, a)$, the prior probability of selecting action a, as predicted by the policy network; and $U(s, a)$, the upper confidence bound, which encourages exploration.

The action at each step is selected by maximizing the sum of the value estimate and the exploration term, as expressed in the following equation:

$$Q(s, a) + U(s, a), \tag{2.10}$$

where $U(s, a)$ is defined as:

$$U(s, a) = c_{\text{puct}} \cdot P(s, a) \cdot \frac{\sqrt{\sum_b N(s, b)}}{1 + N(s, a)}. \tag{2.11}$$

In this formula, c_{puct} is a hyperparameter controlling the balance between exploration and exploitation. The term $P(s, a)$ represents the prior probability of action a, encouraging the algorithm to initially favor actions deemed promising by the policy network. The visit count $N(s, a)$ indicates how often an action has been explored,

while $\sum_b N(s,b)$ is the total visit count for all actions at the current state s. The exploration term $U(s,a)$ scales with the square root of the total visit count, ensuring that less-explored actions receive greater weight during the search process.

The MCTS algorithm operates in four main phases. First, during the *selection phase*, the algorithm starts at the root node and recursively selects child nodes based on the PUCT formula until a leaf node is reached. This process ensures that the search prioritizes promising areas of the game tree while still allocating resources to underexplored branches. Upon reaching a leaf node, the *evaluation phase* begins, where the neural network evaluates the game state. The network outputs a policy vector $P(s,\cdot)$, which provides probabilities for all legal moves, and a value $V(s)$, which estimates the likelihood of winning from the current position. The evaluation results are then used in the *expansion phase*, where new child nodes are created for all legal actions at the leaf node, initializing their prior probabilities $P(s,a)$ using the policy vector. Finally, the *backpropagation phase* updates the visit counts and value estimates for all nodes along the path back to the root. The value $V(s)$ is recursively propagated upward, and the average value estimate $Q(s,a)$ is updated using the following equation:

$$Q(s,a) = \frac{\sum_{i=1}^{N(s,a)} V_i}{N(s,a)}, \tag{2.12}$$

where V_i represents the value of the i-th simulation that passed through the node.

Before making a move, MCTS performs a fixed number of simulations to thoroughly explore the game tree. The final action a^* is chosen based on the highest visit count:

$$a^* = \arg\max_a N(s,a). \tag{2.13}$$

To introduce diversity during self-play training, Dirichlet noise is added to the prior probabilities at the root node:

$$P'(s,a) = (1-\epsilon)\cdot P(s,a) + \epsilon \cdot \mathrm{Dir}(\alpha), \tag{2.14}$$

where ϵ determines the influence of the noise, and $\mathrm{Dir}(\alpha)$ is a Dirichlet distribution with parameter α. This adjustment ensures that the AI explores a wider range of strategies, promoting robust learning.

Through the iterative process of selection, evaluation, expansion, and backpropagation, MCTS dynamically adapts to the game state and improves the decision-making process. By integrating evaluations from the policy-value network, the algorithm identifies strategies that balance immediate gains and long-term objectives. This approach, combined with self-play training, allows the AI to continuously refine its strategies, achieving a high level of performance.

2.4.2 Policy-Value Network in Heian Shogi

The policy-value network is the core component of the Heian Shogi AI system, responsible for evaluating board states and predicting optimal moves. It follows a deep residual network architecture, designed to extract strategic and tactical information from the board efficiently. Residual networks (ResNet) have been widely adopted in deep learning due to their ability to address the vanishing gradient problem and improve feature extraction in deep models [26]. The network integrates both policy and value predictions, enabling it to make well-informed decisions during self-play and Monte Carlo Tree Search (MCTS) simulations.

The policy-value network serves two primary functions: first, it guides MCTS by providing a prior probability distribution over legal moves, improving the efficiency of the search process by prioritizing promising actions [23]. Second, it evaluates the board state by predicting the win probability from the current player's perspective. These dual outputs allow the AI to balance immediate tactical considerations with long-term strategic planning.

Training the network involves optimizing two key objectives: policy loss and value loss. The policy loss, computed using cross-entropy, measures the difference between the predicted move probabilities and the actual move distributions derived from MCTS simulations. The value loss, based on mean squared error (MSE), quantifies the discrepancy between the predicted board evaluation and the actual game out-

come. While MSE is a standard metric, recent studies suggest that the coefficient of determination (R^2) may provide a more informative evaluation for regression-based tasks in machine learning, as it accounts for variance explained by the model and provides a clearer interpretation of predictive performance [27]. L2 regularization is applied to prevent overfitting, ensuring stable training [28]. Regularization techniques play a crucial role in deep learning by mitigating overfitting and improving generalization, with various approaches such as dropout, weight decay, and batch normalization being extensively studied [28].

The Adam optimizer dynamically adjusts learning rates to accelerate convergence while maintaining robustness. By interacting with MCTS, the policy-value network enables effective decision-making. The policy head refines MCTS by providing well-calibrated prior probabilities, reducing unnecessary exploration, while the value head improves the accuracy of leaf node evaluations, enhancing the overall reliability of search outcomes. Through iterative self-play, the network continuously refines its understanding of optimal strategies, ultimately achieving strong performance in Heian Shogi.

2.4.3 Network Architecture and Input Representation

The Heian Shogi AI employs a deep residual network inspired by the AlphaZero framework. The network architecture consists of an initial convolutional block, multiple residual blocks, and two output heads: the policy head for move prediction and the value head for board evaluation[26]. The structure of the network is illustrated in Figure **??**.

Input Representation

The board state is encoded as a $9 \times 9 \times 12$ tensor, where the first two dimensions correspond to the 9×9 game board, and the third dimension consists of 12 feature planes capturing essential game dynamics. This structured input representation allows the network to process spatial relationships and game evolution effectively.

The first 10 feature planes encode piece positions, where each plane corresponds to a specific piece type. A value of 1 represents a red-side piece, -1 represents a black-side piece, and 0 indicates an empty square. These planes collectively store the board state just before the latest move, allowing the model to track piece distributions effectively.

The 11th feature plane records the opponent's most recent move. Initially, this plane is set to all zeros. When a move is made, the piece's previous position is marked as -1, and its new position is marked as 1. This feature provides the model with short-term transition information, helping it recognize recent changes in board dynamics.

The 12th feature plane encodes the turn indicator. If the current player is the first player (Sente), all values in this plane are set to 1; otherwise, they are set to 0. This feature ensures that the network correctly distinguishes between the two players, maintaining turn-based consistency throughout training and gameplay.

Residual Neural Network Architecture

The network begins with an initial convolutional block consisting of a 3×3 convolutional layer with 256 filters, followed by batch normalization and a ReLU activation function. This block extracts fundamental spatial features from the board state.

The core of the network comprises multiple residual blocks. Each residual block consists of two 3×3 convolutional layers with 256 filters, batch normalization, and ReLU activation. A skip connection is applied by adding the block's input to its output before applying the final activation function. These residual connections mitigate vanishing gradient issues and enable deeper learning of complex board states.

The policy head predicts a probability distribution over all legal moves. It consists of a 1×1 convolutional layer with 16 filters, batch normalization, ReLU activation, and a fully connected (FC) layer mapping the features to a 2038-dimensional vector (corresponding to all possible moves in Heian Shogi). A final softmax activation normalizes the outputs into a probability distribution.

The value head estimates the win probability of the current board state. It consists

of a 1×1 convolutional layer with 8 filters, batch normalization, ReLU activation, a fully connected layer reducing features to 256 neurons, another ReLU activation, and a final fully connected layer mapping to a scalar output. A tanh activation function constrains the output within the range $[-1, 1]$, where 1 indicates a certain win, -1 a certain loss, and 0 a balanced state.

This deep residual network architecture enables the Heian Shogi AI to model both short-term tactical decisions and long-term strategic planning. By leveraging reinforcement learning through self-play, the network refines its understanding of game strategies, leading to improved decision-making and overall performance.

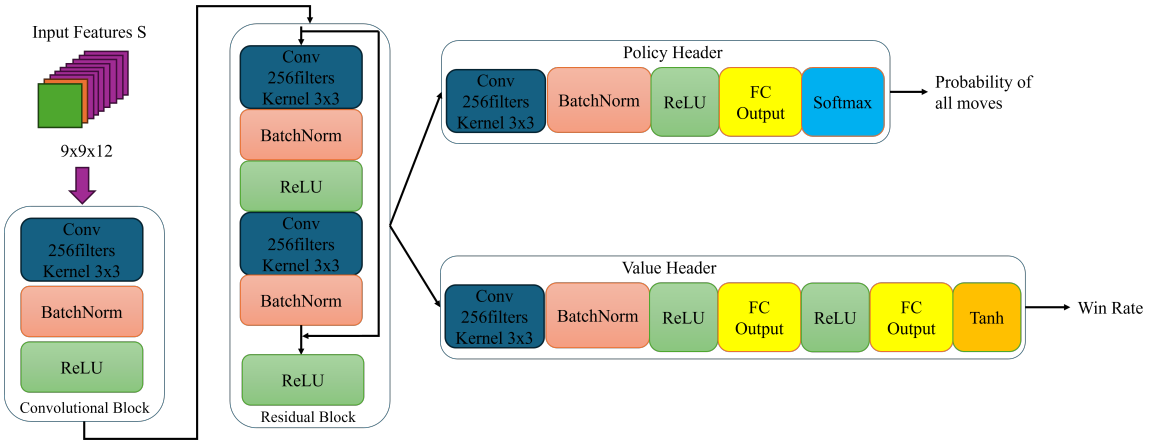

Figure 2-4: Heian Shogi network

The self-play data collection process is implemented by the ‘CollectPipeline‘ class, which handles game data generation and storage. During self-play, the AI sequentially makes moves while recording the board state, MCTS policy probabilities, and the final game outcome for each step. Self-play has been widely used in reinforcement learning-based game AI, as demonstrated in AlphaGo and AlphaZero [21, 22, 23]. The success of reinforcement learning (RL) in game-playing AI is largely attributed to the development of deep RL methods, particularly those that enable training without explicit human supervision [29].

To enhance data diversity, the system employs horizontal board flipping as a data augmentation technique, effectively doubling the dataset size. Data augmentation techniques, including board transformations, have been proven to be effective in im-

proving training efficiency in self-play AI systems [22, ?]. The collected self-play data is stored in an experience buffer and later sampled for neural network training, a technique inspired by experience replay methods in deep reinforcement learning [29]. Experience replay plays a crucial role in stabilizing RL training by breaking temporal correlations in data and improving sample efficiency. The use of such replay buffers has significantly improved performance in various deep RL applications, including board game AI and decision-making tasks [30].

Throughout the training process, the system periodically samples data from the buffer for model updates and saves the latest model for subsequent self-play sessions. This iterative training paradigm has been fundamental to modern RL-based AI systems, enabling the development of superhuman performance in Chess, Go, Shogi, and even Atari games [30]. The integration of MCTS with deep learning models has further enhanced AI's capability to evaluate board positions dynamically and adaptively improve its decision-making strategies over time.

The entire training process follows an iterative loop, consisting of self-play data generation, data storage, network training, and model updating. During training, the system dynamically adjusts the learning rate to prevent the model from converging to a local optimum. Additionally, metrics such as branch factor and leaf node value are recorded to monitor MCTS search efficiency and position evaluation stability.

In summary, this Heian Shogi AI system is built upon the AlphaZero framework, achieving end-to-end reinforcement learning through self-play. It integrates board logic, MCTS search, deep neural network evaluation, and data collection, enabling autonomous learning without human intervention and gradually improving its playing strength to a competitive level.

Chapter 3

Can AI Resign at the right Time in a Game of Shogi?

This chapter is based on the integration, update, and abridgment of the following publication:

- Pan, Shize, and Iida, Hiroyuki. "CAN AI RESIGN AT THE RIGHT TIME IN A GAME OF SHOGI?." Journal of Mathematical Sciences and Informatics 4.3 (2024).

3.1 Chapter Introduction

The ability of artificial intelligence (AI) to make strategic decisions has become a critical area of research in board games such as Shogi. While modern Shogi AI systems have achieved superhuman performance in gameplay, their resignation mechanisms remain suboptimal, often failing to emulate human-like decision-making. Existing AI engines frequently continue playing in losing positions where human players would logically resign, leading to prolonged and strategically unnecessary gameplay. This

discrepancy highlights the need for a refined resignation mechanism that aligns more closely with human intuition.

This chapter explores the integration of a dynamic resignation threshold within Shogi AI, aiming to enhance its decision-making capabilities. The study applies Game Refinement (GR) theory to analyze key gameplay metrics, such as game length and branching factors, in order to determine the impact of resignation on game dynamics. Additionally, Monte Carlo Tree Search (MCTS) and deep learning techniques are incorporated to optimize resignation decisions based on AI performance across different skill levels.

The chapter is structured as follows: Section **??** introduces the methodology used to assess Shogi gameplay, leveraging the Suisho5 engine and GR theory. Section **??** discusses the implementation of the CShogi engine as a computational tool for analyzing branching factors and game complexity. Section **??** presents the experimental setup, describing the classification of AI skill levels and the analysis of game length before and after applying the resignation mechanism. Finally, Section **??** summarizes the findings and discusses their implications for future AI development in board game strategy.

3.2 Assessment Methodology

In this study, the primary approach to evaluating Shogi games involves leveraging an opening source completion engine known as Suisho5. This engine provides a robust framework for analyzing the progression of Shogi games, focusing particularly on the strategic depth and decision-making processes. To complement this analysis, the Game Refinement (GR) theory is applied to the collected game data, enabling a quantitative assessment of key metrics such as game length, the diversity of strategies employed in each round, and the trends in acceleration of uncertainty resolution throughout the gameplay.

The evaluation framework is meticulously designed to integrate these elements, with a specific focus on two core aspects: the game length and the score dynamics

associated with legal moves. Game length serves as a critical metric, reflecting the overall pacing and balance of the game, while the score dynamics provide insight into the decision-making strategies of players, particularly the losing computer AI. By examining the score variations in each round, especially for losing players, this study aims to uncover patterns and trends that highlight the strategic opportunities and challenges inherent in Shogi gameplay.

Moreover, the GR theory is utilized to model the acceleration of uncertainty resolution across the course of a game, capturing the shifts in decision-making intensity as the game progresses. This approach allows for a nuanced understanding of the interplay between skill and chance, as well as the evolving complexity of strategic options. The integration of Suisho5 and GR theory thus offers a comprehensive evaluation framework, bridging qualitative and quantitative analyses to provide a holistic perspective on Shogi gameplay.

Through this methodology, the study seeks to identify the defining characteristics of optimal game design in Shogi, contributing to both the theoretical understanding of game refinement and practical insights into enhancing the AI's performance in strategic decision-making scenarios.

3.3 Analysis Tool: CShogi Engine

The CShogi engine, developed in C++, is an efficient tool for analyzing Shogi games. Compared to Python-based Shogi engines, CShogi offers significantly improved computational speed, making it highly suitable for large-scale analysis of game data. This performance enhancement stems from its efficient implementation in C++, which allows it to handle the computationally intensive task of generating and evaluating all possible legal move options for each turn in a Shogi game. These options encompass both optimal and suboptimal moves, providing a comprehensive dataset for game analysis.

One of the key advantages of CShogi is its ability to calculate the average number of legal moves (B) across the entire game. By analyzing this metric, researchers

can obtain precise data on the branching factor, which is critical for applying Game Refinement (GR) theory. GR theory relies on accurate measurements of game length (D) and branching factor (B) to model the complexity and engagement dynamics of board games like Shogi. The CShogi engine enables these calculations with a high degree of accuracy, making it an essential tool for quantitative studies in this domain.

In practical applications, the engine processes each move within a game to identify all feasible options, regardless of their strategic quality. This exhaustive evaluation ensures that the data captures the full range of decision-making possibilities available to players at each turn. After the game concludes, the average number of moves per turn is computed, yielding the branching factor (B) as a representative metric for game complexity. This approach allows researchers to explore patterns in decision-making and the evolution of strategies throughout the game.

The use of CShogi in this study offers a dual benefit: first, its computational efficiency enables the analysis of extensive datasets, ensuring scalability and reliability; second, its precise calculation of branching factors provides robust input for theoretical models like GR theory. By leveraging CShogi, this research gains a comprehensive understanding of the structural characteristics of Shogi games, paving the way for deeper insights into game complexity, player strategies, and AI performance in strategic decision-making scenarios.

Through its combination of speed, accuracy, and versatility, the CShogi engine establishes itself as a cornerstone for data-driven studies in the field of Shogi research. Its ability to seamlessly integrate with theoretical frameworks such as GR theory makes it an indispensable tool for analyzing the intricacies of Shogi gameplay.

3.4 Experimental Setup and Results

The experiment was divided into two parts. Initially, four distinct skill levels were defined: 20, 15, 10, and 5. In simpler terms, a lower skill level corresponds to a higher probability of selecting suboptimal moves for the second and subsequent actions in a strategy. However, even with reduced skill levels, the AI does not make extremely

poor decisions, as those are filtered out. This ensures that the effects of lowering skill levels remain within a reasonable range, avoiding excessively irrational moves. As the skill level decreases, the AI player's performance correspondingly declines.

The experiment involved two AI players at the same skill level competing against each other, with 300 games played for each skill level. During these matches, the AI was set to avoid resignation unless the winning rate reached 99.99%. In other words, the AI was not allowed to resign prematurely, ensuring that the games progressed fully. After each game, the moves were analyzed, and a score was assigned to every step.

The data collected from these experiments is summarized in Table 3.1. The results reflect the performance of AI players in scenarios without resignation. At the highest skill level (20), both the average branching factor and game length exceeded the benchmarks established by human master data, which are 80 and 115, respectively. As the skill level decreased, the average branching factor gradually approached the human master standard. However, the game length remained significantly longer than that of human players across all skill levels.

Table 3.1: Performance Data of AI Players Without Resignation

Skill Level	**Average Branching Factor (B)**	**Average Game Length (D)**
20	95	150
15	88	140
10	83	135
5	81	130

Through thorough analysis, we devised a resignation mechanism aimed at determining the optimal point at which a player should resign during a match. This mechanism is based on the principle that resignation occurs when the losing player reaches a state where they have no viable means to alter the outcome of the game. Specifically, the resignation threshold is defined by the maximum advantage score ever achieved by the losing side. Once the disadvantage score of the losing side exceeds this maximum advantage score, resignation is deemed appropriate.

To establish a reasonable maximum advantage score for the losing side, we ana-

lyzed data from 300 games across varying skill levels. Our findings revealed that the scores achieved by AI players varied significantly based on their skill levels. Notably, higher skill levels corresponded to lower maximum advantage scores for the losing side, while lower skill levels necessitated larger advantage scores to prevent potential comebacks. These observations led us to conclude that the resignation threshold should be based on the maximum advantage score achieved by the losing side under conditions of error-free gameplay.

Prior to conducting experiments, we hypothesized that more skilled AI players would be able to conclude games more quickly when applying the resignation mechanism. To test this hypothesis, we carried out experiments at different skill levels, analyzing the impact of the resignation mechanism on game length.

During the analysis, certain moves by the losing side were associated with unusually high scores. Further investigation revealed that these inflated scores were typically caused by mistakes made by the AI or by the adoption of suboptimal strategies. Such errors often allowed for dramatic reversals, with the losing side ultimately turning the game in their favor. This finding highlighted the importance of defining a sensible resignation threshold, grounded in the assumption that the AI performs flawlessly.

By integrating this hypothesis and conducting rigorous experiments, we aimed to verify whether higher-level AI players could significantly reduce game length through the resignation mechanism, thereby improving overall game efficiency. The results of the experiments are summarized in Table 3.2, which details the resignation threshold and the corresponding game length for each skill level.

Table 3.2: Resignation Threshold and Game Length

Skill Level	**Resignation Threshold**	**Game Length**
20	850	110
15	2950	112
10	6450	116
5	9999	121

In the experiment involving 300 games at skill level 20, the resignation threshold

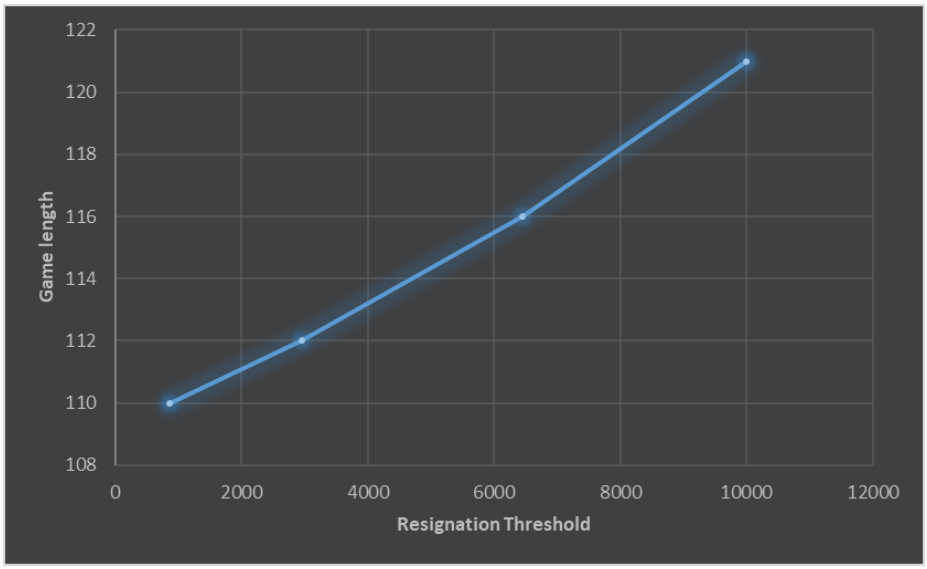

Figure 3-1: The relation between the resignation threshold and skill level

was defined as the point where the score advantage exceeded 850 points. At this threshold, the AI player was instructed to resign. Similarly, thresholds for skill levels 15, 10, and 5 were set at 2950, 6450, and 9999 points, respectively. Figure 3-1 illustrates the relationship between skill level and score threshold. Notably, as the AI's skill level increases, the score required to trigger resignation decreases. This indicates that more advanced AI players need a lower advantage score to determine the optimal resignation point.

The implementation of the resignation mechanism yielded significant reductions in game length, as detailed in Table 3.2. After applying the mechanism, the average game length at skill level 20 was reduced to 110 moves. At skill levels 15, 10, and 5, the average game lengths were shortened to 112, 116, and 121 moves, respectively. These results demonstrate that the resignation mechanism effectively accelerates gameplay and shortens the duration of games across different skill levels.

The application of the resignation mechanism led to a significant reduction in game length and an increase in game speed, as illustrated in Table 3.3. At skill level 20, the average game length decreased from 165.66 moves to 110 moves. Similarly, for skill levels 15, 10, and 5, the game lengths were reduced from 141.18, 133.97, and 121.17 moves to 112, 116, and 121 moves, respectively. This demonstrates the effectiveness of

the resignation mechanism in shortening game durations across different skill levels.

The resignation mechanism operates based on the advantage score of a player. If a player's score reaches or exceeds the resignation threshold, it indicates a decisive advantage, and the game concludes at that moment. This ensures that prolonged gameplay is avoided when the outcome becomes evident, thus enhancing the overall efficiency of the game.

Additionally, the analysis of gaming data revealed a notable relationship between skill level and the number of branching factors encountered during gameplay. This correlation highlights how the complexity of decision-making is influenced by the players' skill levels and further validates the effectiveness of the resignation mechanism in adapting to different game scenarios.

Table 3.3: Comparison of Game Length and Speed Before and After Applying the Resignation Mechanism

Skill Level	**B**	**D (Before)**	**D' (After)**	**V (Before)**	**V' (After)**	**Threshold**
20	92.41	165.66	110	0.28	0.42	850
15	82.31	141.18	112	0.29	0.37	2950
10	81.39	133.97	116	0.30	0.35	6450
5	78.74	121.17	121	0.33	0.33	9999

Figure 3-2 illustrates the relationship between the branching factor and skill level, compared with that of human masters. The data indicates that as player skill decreases, the complexity of decision points also diminishes, resulting in fewer branching factors for players with lower skill levels. This trend reflects an inherent dynamic in gameplay, where less skilled players encounter a more limited range of options at each decision point compared to their higher-skilled counterparts.

One possible explanation for this phenomenon lies in the narrower understanding of game mechanics and strategies among less skilled players. This limited understanding may lead to overlooking potential moves or failing to recognize all available options, thereby reducing the complexity of decision trees. As a result, gameplay for less skilled players tends to exhibit simpler decision points and lower branching factors.

The implications of this trend are significant for both game design and player experience. By leveraging these insights, game developers can design tailored gameplay experiences that align with the skill levels of their target audience. This approach not only makes games more engaging and enjoyable but also ensures accessibility for players of varying abilities. Additionally, this correlation between skill level and branching factors can inform strategies for player education and skill development. Specifically, helping less skilled players expand their understanding of available options can enhance their decision-making abilities and overall gameplay experience.

Notably, the data reveals that players at skill levels 20 and 15 demonstrate branching factors far exceeding those of human masters. Skill level 10 aligns closely with the performance of human masters, while skill level 5 corresponds to intermediate or novice-level gameplay.

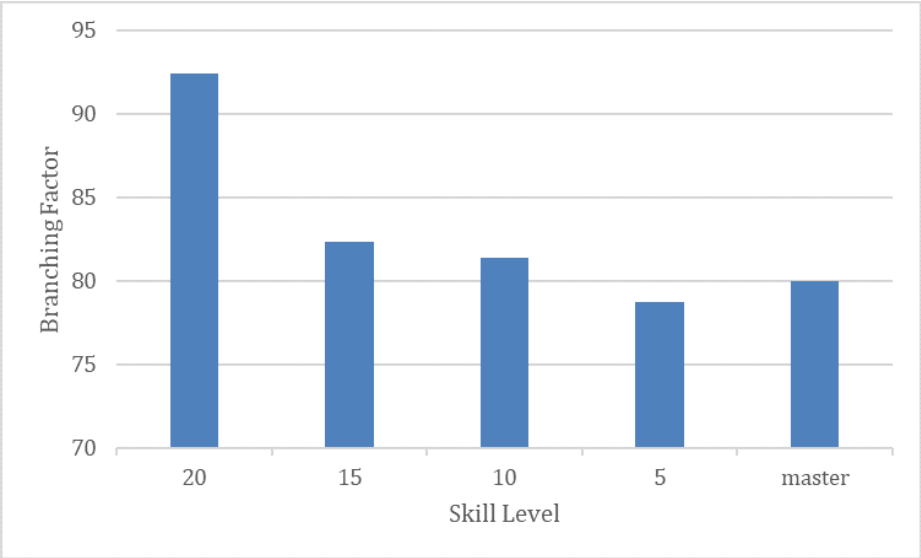

Figure 3-2: The relationship between branching factor and skill level compared to human masters

Figure 3-3 illustrates the impact of the resignation mechanism on game speed (velocity) for AI players across different skill levels. Before the resignation mechanism was applied, the game speed (v) decreased as the skill level of the AI increased. Specifically, the game speeds for skill levels 20, 15, 10, and 5 were 0.27, 0.29, 0.30, and 0.32, respectively. In contrast, after applying the resignation mechanism, the

game speed (v') exhibited a reverse trend, increasing with the AI's skill level. For the same skill levels, the game speeds increased to 0.42, 0.37, 0.35, and 0.33, respectively.

This shift indicates that the resignation mechanism significantly influenced the game dynamics, allowing higher-skilled AI players to conclude games more efficiently. The observed increase in game speed highlights the effectiveness of the resignation mechanism in streamlining gameplay and accelerating game conclusions, particularly for more advanced AI players.

The findings emphasize that while less skilled AI players initially exhibited faster game speeds due to simpler decision-making processes, the resignation mechanism effectively leveled this disparity by enabling skilled AI players to finish games more quickly. This dynamic demonstrates the utility of the resignation mechanism in optimizing gameplay efficiency.

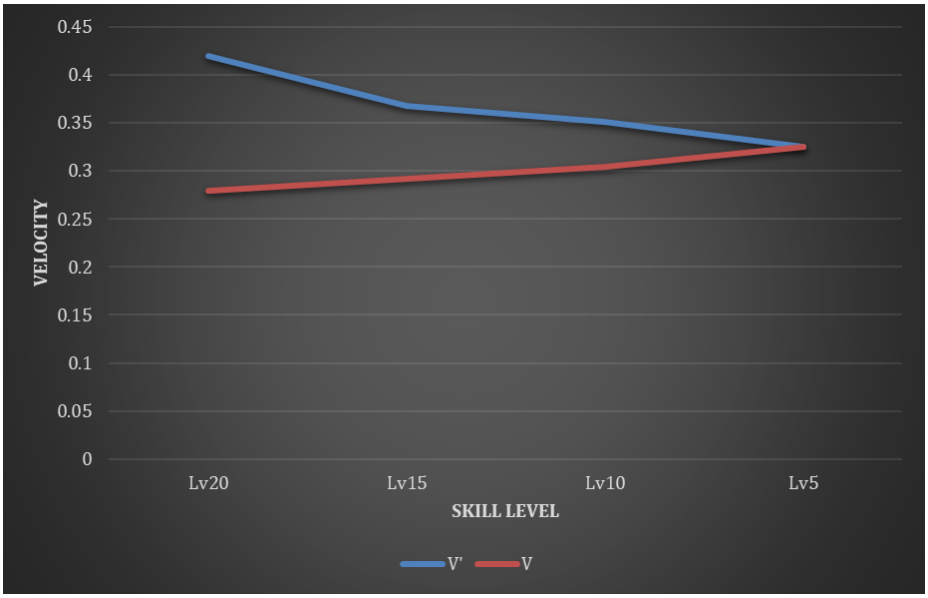

Figure 3-3: The relation between velocity of game and skill level

The findings of this study indicate that the opt-out mechanism has a more pronounced effect on game length among players with advanced skills. The correlation between skill level and the degree of game length reduction highlights the nuanced impact of the resignation mechanism on gameplay dynamics. Specifically, players with higher skill levels appear better equipped to strategically utilize the opt-out mechanism, resulting in more substantial reductions in game duration. In contrast,

less skilled players may not fully exploit this mechanism, leading to comparatively smaller decreases in game length.

These results have significant implications for game design and player experience. The opt-out mechanism proves to be an effective tool for regulating gameplay pace, particularly for advanced players who can derive the greatest benefit from its implementation. By facilitating quicker decision-making and streamlining gameplay, the mechanism contributes to a more dynamic and engaging gaming experience across varying skill levels.

Additionally, the implementation of the resignation mechanism has resulted in a noticeable acceleration of gameplay. This increase in game speed suggests that the mechanism not only reduces the duration of individual games but also enhances the overall tempo of gameplay sessions. Such acceleration further underscores the mechanism's role in optimizing gameplay efficiency and maintaining player engagement.

Based on the average number of branching factors observed for a human master player, which is approximately 80, we can infer the skill level of human masters to range between 5 and 10. This inference is grounded in the understanding that lower skill levels correspond to fewer branching factors, while higher skill levels involve a greater number of potential moves at each decision point.

Therefore, it is reasonable to conclude that the skill level of a human master player falls within the range of 5 to 10 in the context of the analyzed game. This estimation aligns with the observed branching factor averages and offers valuable insights into the proficiency levels of human players, further enhancing our understanding of player dynamics within the game environment.

3.5 Chapter Conclusion

This chapter investigated the effectiveness of a resignation threshold mechanism in Shogi AI, addressing the limitations of existing AI resignation strategies. By defining a dynamic resignation threshold based on the maximum advantage score of the losing player, this study introduced a more adaptable method for determining optimal res-

ignation timing. The results demonstrated that this mechanism significantly reduced game length across various AI skill levels, ensuring a more efficient and engaging gameplay experience.

The findings revealed a strong correlation between AI skill level and game length reduction, with higher-skilled AI players utilizing the resignation mechanism more effectively. Additionally, the analysis of branching factors and game speed confirmed that the resignation mechanism enhanced AI decision-making by preventing unnecessary prolongation of losing positions. The data further suggested that human master players likely fall within the skill level range of **5 to 10**, reinforcing the model's alignment with real-world gameplay.

Despite its strengths, the resignation mechanism has limitations. A critical drawback is its lack of opponent modeling, meaning it does not account for whether the opponent has recognized their winning position. This limitation may lead to premature resignations in cases where the opponent is unaware of their advantage. Future work will focus on enhancing opponent-aware decision-making, ensuring that resignation decisions are not only based on the losing player's position but also incorporate the strategic awareness of the opponent.

Chapter 4

Conclusion and Future Work

4.1 Conclusion

In this study, we introduced a resignation threshold mechanism specifically tailored for Shogi AI systems. The mechanism is defined as the maximum advantageous score attained by the losing player, augmented by a slight increment, ensuring that resignation occurs when the losing side perceives no viable opportunities to alter the game's outcome. Experimental results demonstrated that this threshold mechanism is effective within the context of Shogi, as the majority of analyzed matches revealed relatively low maximum advantageous scores for the losing player, with only a small fraction of games exhibiting higher scores. By establishing the resignation threshold as slightly exceeding these values, the mechanism was shown to enhance the strategic coherence of resignation decisions across various gameplay scenarios in Shogi.

One of the key observations from our experiments was the impact of the resignation mechanism on game length across AI skill levels. Higher-skilled AI players demonstrated a more significant reduction in game length, indicating their superior ability to recognize resignation opportunities earlier. This behavior underscores the mechanism's role in aligning AI decision-making more closely with human-like strategic intuition.

Beyond improving AI systems, the findings of this research also hold meaningful implications for human Shogi players. By emulating the decision-making patterns of

human masters, the resignation mechanism not only enhances the realism of AI gameplay but also serves as a valuable training tool for human players. The AI's improved strategic depth and decision-making capabilities provide insights into optimal resignation timing, offering a benchmark for human players to refine their understanding of when to concede a game. Additionally, the mechanism's ability to quantify the dynamics of strategic decision-making can aid human players in developing a more structured approach to evaluating game positions and predicting outcomes.

A significant contribution of this study is the development and implementation of a Heian Shogi AI system based on the AlphaZero framework. The proposed system integrates a Monte Carlo Tree Search (MCTS) algorithm with a deep residual policy-value network to evaluate board states and predict optimal moves. Unlike traditional Shogi AI engines, which rely on handcrafted evaluation functions, our system leverages reinforcement learning through self-play to autonomously refine its strategic understanding of Heian Shogi. By reconstructing the gameplay mechanics of Heian Shogi through AI simulation, this research provides a valuable framework for analyzing historical board games and understanding their evolutionary trajectory.

The Heian Shogi AI system not only demonstrates the adaptability of modern AI techniques to historical board games but also showcases how reinforcement learning can be applied to reconstruct and analyze lost game variants. The successful implementation of this AI contributes to both computational game theory and cultural heritage preservation by offering a systematic method for exploring ancient strategic games through self-play simulations.

The strengths of this study lie in its innovative integration of game refinement theory and motion-in-mind models to analyze the cognitive depth of gameplay. By introducing a resignation mechanism, the research has improved both the efficiency and realism of AI systems, enabling the collection of high-quality gameplay data. This data is invaluable for analyzing key metrics such as game length and branching factors, facilitating a deeper understanding of the strategic differences between human and AI players. Furthermore, the development of the Heian Shogi AI system highlights the potential of reinforcement learning in studying historical game evolution, opening

new avenues for AI-driven historical simulations.

However, the study is not without limitations. A critical drawback lies in the mechanism's lack of opponent modeling, which restricts its ability to evaluate the awareness of the opponent in determining resignation timing. Specifically, while the mechanism ensures resignation when the losing side has no winning opportunities, it does not account for whether the opponent has recognized their winning position. This limitation may result in premature resignation in scenarios where the opponent is unaware of their advantage, thereby overlooking potential opportunities to prolong the game strategically. Addressing this issue requires the development of an opponent-aware framework, where AI systems evaluate the opponent's skill level and their ability to identify a winning path. By incorporating such a perspective, future studies can enhance the contextual sensitivity of resignation mechanisms, ensuring that decisions are not only based on the losing side's perspective but also aligned with the opponent's understanding of the game state.

4.2 Future Work

Future work will expand on these findings by conducting experiments with different game engines and applying the resignation mechanism across a wider range of board games. This broader experimentation aims to develop generalized models capable of adapting to diverse gaming contexts, enabling a deeper understanding of how resignation dynamics vary across different rule sets and game structures. Additionally, further refinement of the Heian Shogi AI system can incorporate multi-task learning strategies to enhance both tactical and strategic decision-making in historical gameplay reconstruction. By addressing these challenges, this research will contribute to the advancement of intelligent and human-like AI systems, providing valuable tools for both academic exploration and practical applications in the field of game design, artificial intelligence, and historical game studies.

Bibliography

[1] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144, 2018.

[2] Corina Pelau, Dan-Cristian Dabija, and Irina Ene. What makes an ai device human-like? the role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Computers in Human Behavior*, 122:106855, 2021.

[3] John McCarthy. From here to human-level ai. *Artificial Intelligence*, 171(18):1174–1182, 2007.

[4] Guillaume Chaslot, Sander Bakkes, Istvan Szita, and Pieter Spronck. Monte-carlo tree search: A new framework for game ai. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 4, pages 216–217, 2008.

[5] Chengpeng Hu, Yunlong Zhao, Ziqi Wang, Haocheng Du, and Jialin Liu. Games for artificial intelligence research: A review and perspectives. *IEEE Transactions on Artificial Intelligence*, 2024.

[6] Steven Walczak. Knowledge-based search in competitive domains. *IEEE Transactions on Knowledge and Data Engineering*, 15(3):734–743, 2003.

[7] R. McIlroy-Young, S. Sen, J. Kleinberg, and A. Anderson. Aligning superhuman ai with human behavior: Chess as a model system. pages 1677–1687, 2020.

[8] Zana Buçinca, Siddharth Swaroop, Amanda E Paluch, Susan A Murphy, and Krzysztof Z Gajos. Towards optimizing human-centric objectives in ai-assisted decision-making with offline reinforcement learning. *arXiv preprint arXiv:2403.05911*, 2024.

[9] Xinyue Hao, Emrah Demir, and Daniel Eyers. Exploring collaborative decision-making: A quasi-experimental study of human and generative ai interaction. *Technology in Society*, 78:102662, 2024.

[10] Kais Dukes. Indigovx: Where human intelligence meets ai for optimal decision making. *arXiv preprint arXiv:2307.11516*, 2023.

[11] U. Bhatt and H. Sargeant. When should algorithms resign? *arXiv preprint arXiv:2402.18326*, 2024.

[12] C. Wirth and J. Fürnkranz. On learning from game annotations. *IEEE Transactions on Computational Intelligence and AI in Games*, 7(3):304–316, 2014.

[13] Hiroyuki Iida, Makoto Sakuta, and Jeff Rollason. Computer shogi. *Artificial Intelligence*, 134(1-2):121–144, 2002.

[14] The establishment of sho shogi. https://cir.nii.ac.jp/crid/1050864590245730176/. SHIMIZU, Yasuji.

[15] Reconsidering the introduction of shogi. https://www.academia.edu/6582978/. SHIMIZU, Yasuji.

[16] Gesine Kernchen. *History at play: How games reflect the socio-political situation in Edo Japan*. PhD thesis, 2021.

[17] Hiroyuki Iida, Kazutoshi Takahara, Jun Nagashima, Yoichiro Kajihara, and Tsuyoshi Hashimoto. An application of game-refinement theory to mah jong.

In Matthias Rauterberg, editor, *Entertainment Computing – ICEC 2004*, pages 333–338, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[18] Wu Yicong, Htun Pa Pa Aung, Mohd Nor Akmal Khalid, and Hiroyuki Iida. Evolution of games towards the discovery of noble uncertainty. In *2019 International Conference on Advanced Information Technologies (ICAIT)*, pages 72–77. IEEE, 2019.

[19] Ky Fan. Minimax theorems. *Proceedings of the National Academy of Sciences*, 39(1):42–47, 1953.

[20] Gao Naying, Gao Yuexian, Mohd Nor Akmal Khalid, and Hiroyuki Iida. A computational game experience analysis via game refinement theory. *Telematics and Informatics Reports*, 9:100039, 2023.

[21] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[22] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *nature*, 550(7676):354–359, 2017.

[23] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017.

[24] Elizabeth Kiritani. *Vanishing Japan: Traditions, Crafts & Culture*. Tuttle Publishing, 2012.

[25] Cameron B Browne, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon

Samothrakis, and Simon Colton. A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games*, 4(1):1–43, 2012.

[26] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[27] Davide Chicco, Matthijs J Warrens, and Giuseppe Jurman. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *Peerj computer science*, 7:e623, 2021.

[28] Jan Kukačka, Vladimir Golkov, and Daniel Cremers. Regularization for deep learning: A taxonomy. *arXiv preprint arXiv:1710.10686*, 2017.

[29] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *nature*, 518(7540):529–533, 2015.

[30] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, et al. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588(7839):604–609, 2020.