

DOI:10.1145/3460349

AI is driving the next evolution of chess, giving players a glimpse into the game's future.

**BY NENAD TOMAŠEV, ULRICH PAQUET,
DEMIS HASSABIS, AND VLADIMIR KRAMNIK**

Reimagining Chess with AlphaZero

MODERN CHESS IS the culmination of centuries of experience, as well as an evolutionary sequence of rule adjustments from its inception in the 6th century to the modern rules we know today.¹⁷ While classical chess still captivates the minds of millions of players worldwide, the game is anything but static. Many variants have been proposed and played over the years by enthusiasts and theorists.^{8,20} They continue the evolutionary cycle by altering the board, piece placement, or the rules—offering players “something subtle, sparkling, or amusing which cannot be done in ordinary chess.”¹

Technological progress is the new driver of the evolutionary cycle. Chess engines increase in strength, and players have access to millions of computer games and volumes of opening theory. Consequently, the number of decisive games in super-tournaments has declined, and it takes longer for players to move from home preparation to playing original moves on the board.¹⁴ While classical chess remains a fascinating

game and is unlikely to ever fall out of fashion, alternative variants provide an avenue for more creative play. In Fischer random chess, the brainchild of former world champion Bobby Fischer, the initial position is randomized to counter the dominance of opening preparation in a game.⁷ One could consider not only entirely new ideas, but also reassess some of the newer additions to the game. For example, the “castling” move was only introduced in its current form in the 17th century. What would chess have been like had castling not been incorporated into the rules? Without recourse to repeating history, we reimagine chess and address such questions in silico with AlphaZero.²⁵

AlphaZero is a system that can learn superhuman chess strategies from scratch without any human supervision.^{19,22} It represents a milestone in artificial intelligence (AI), a field that has ventured down the corridors of chess more than once in search of challenges and inspiration. Throughout the history of computer chess, the focus was on creating systems that could spar with top human players over the board.³ Computer chess has progressed steadily since the 1950s, with better-tuned evaluation functions and enhanced search algorithms deployed on increasingly more computational

» key insights

- **Using neural networks together with advanced reinforcement learning algorithms can surpass human knowledge of chess and allow us to reimagine the game as we know it.**
- **We use AlphaZero, a system that learns to play chess from scratch, to explore what variations of chess would look like at superhuman level.**
- **Atomic changes to the rules of classical chess alter game balance, with the material value of the pieces being altered accordingly, resulting in either more draws or more decisive outcomes. The findings of our quantitative and qualitative analysis demonstrate the rich possibilities that lie beyond the rules of modern chess.**

resources.^{2,3,9,13,18,24} Alan Turing already envisioned more in 1953 by asking, “Could one make a machine to play chess, and to improve its play, game by game, profiting from its experience?”²⁷ Unlike its predecessors, AlphaZero learns its policy from scratch from repeated self-play games, answering the second part of Turing’s question. The result is a unique approach to playing classical chess²² and a new era in the development of chess engines, as spearheaded by Leela Chess Zero.¹⁵

AlphaZero’s ability to continually improve its understanding of the game, and reach superhuman playing strength in classical chess and Go,²⁵ lends itself to the question of assessing chess variants and potential variants of other board games in the future. Provided only with the implementation of the rules, it is possible to effectively simulate decades of human experience in a day, opening a window into top-level play of each variant. In doing so, computer chess completes the circle, from the early days of pitting *man vs. machine* to a collaborative present of *man with machine*, where AI can empower players to explore what chess is and what it could become.¹¹

Rule Alterations

There are many ways in which the rules of chess could be altered. In this work, we limit ourselves to atomic changes that do not involve changes to the starting position and keep the game as close as possible to classical chess. Some of the alterations that we consider are novel, while others have been previously discussed within the chess community but have yet to be widely adopted. The nine changes considered in this study are listed in Table 1. *No-castling* and *No-castling (10)* involve a full and partial restriction on the castling rule. *Pawn-one-square*, *Semi-torpedo*, *Torpedo*, *Pawn-back*, and *Pawn-sideways* involve changes to pawn mobility. *Self-capture* chess allows players to also capture

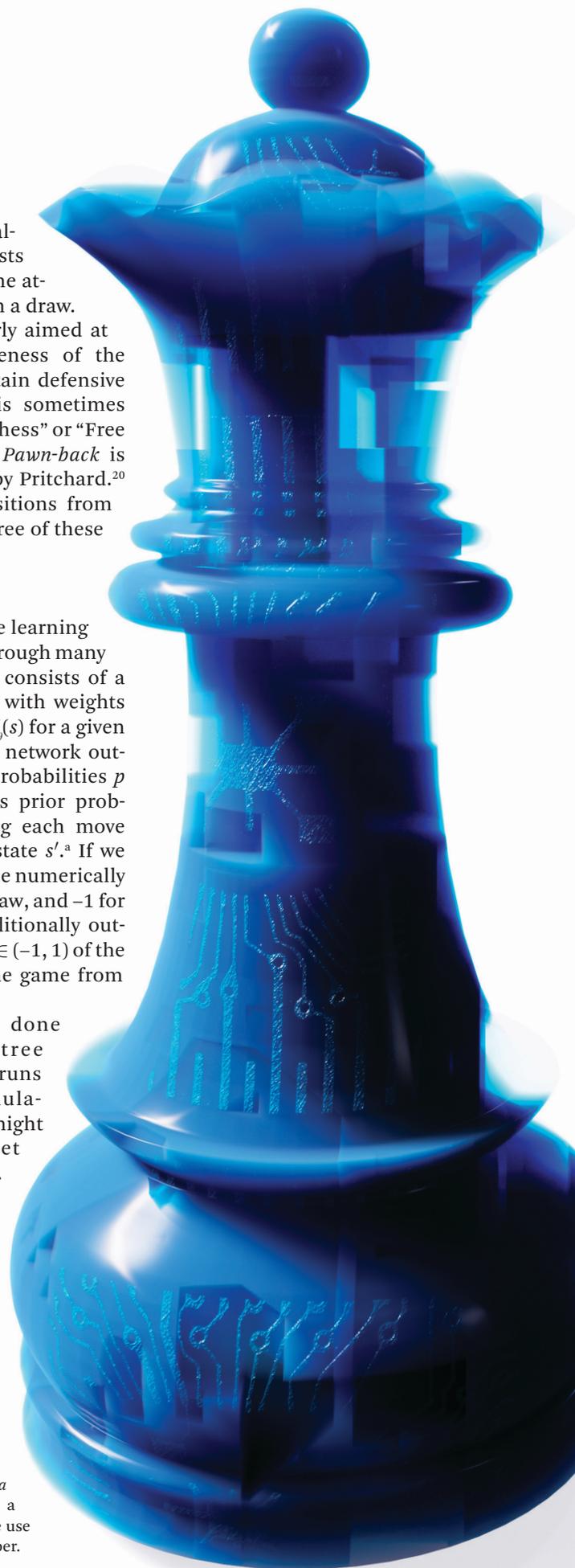
their own pieces. Finally, *Stalemate=win* recasts stalemate as a win for the attacking side, rather than a draw. As such, it is particularly aimed at increasing the decisiveness of the game, by removing certain defensive patterns. Self-capture is sometimes referred to as “Reform Chess” or “Free Capture Chess,” while *Pawn-back* is called “Wren’s Game” by Pritchard.²⁰ Figure 1 illustrates positions from AlphaZero games for three of these variants.

AlphaZero

AlphaZero is an adaptive learning system that improves through many rounds of self-play.²⁵ It consists of a deep neural network f_θ with weights θ that compute $(p, v) = f_\theta(s)$ for a given position or state s . The network outputs a vector of move probabilities p with elements $p(s'|s)$ as prior probabilities for considering each move and, hence, each next state s' .^a If we denote a game’s outcome numerically by +1 for a win, 0 for a draw, and -1 for a loss, the network additionally outputs a scalar estimate $v \in (-1, 1)$ of the expected outcome of the game from position s .

Move selection is done by Monte Carlo tree search (MCTS), which runs repeated search simulations of how the game might unfold up to a preset maximum ply depth. In one MCTS simulation, f_θ is recursively applied to a sequence of positions until a maximum-depth leaf node is reached. The sequence of moves in

^a We have suppressed notation somewhat; the probabilities are technically over actions or moves a in state s , but as each action a deterministically leads to a separate next position s' , we use the concise $p(s'|s)$ in this paper.



a simulation depends on an action selection criterion that is applied at each node along the path; it is this version of the PUCT algorithm²¹ that trades off exploration against revisiting more promising moves more frequently over consecutive simulations. The action selection criterion is such that before a state is first encountered in a simulation, its resulting prior vector p assigns weights to candidate

moves at a “first glance” of the board. When a leaf node is reached, its position’s evaluation v is “backed up” to the root, with each node along the path incrementing its visit count and including the leaf’s v in its action-value estimate. After a number of such MCTS simulations, the root move that was visited most is played.

Training $f_\theta(s)$ is done via gradient descent steps to let p and v predict the next move and final game outcome from s as closely as possible for a streaming sample of game positions. The sample is continually refreshed with self-play games that are generated using f_θ in MCTS as f_θ is updated. The result is a feedback loop that produces a streaming sample of games of increasing quality.

The starting point for our experimental setup is the input state representation, neural network architecture, MCTS depth and action selection criterion, AlphaZero training configuration, and hardware in Silver et al.²⁵ As the rule changes in Table 1 are atomic, we assume that a quality of play comparable to AlphaZero’s classical chess games²² would be reached for each of the variants under the same conditions. If the experimental setup is kept constant except for altering the legal moves list, we furthermore assume that differences in game outcomes would be informative of their *relative* decisiveness. We trained each variant in Table 1’s neural network from a random initialization for 1 million gradient descent steps using the configuration in Silver et al.²⁵

Self-Play Games for Each Chess Variant

For each chess variant, we used the resulting AlphaZero model to generate a diverse set of 10,000 self-play games at 1 sec per move and 1,000 self-play games at 1 min per move. In the absence of external stochasticity, each variant’s self-play games would be identical under the same time controls. To generate self-play games for analysis, we promoted

diversity by sampling the first 20 plies in each game proportional to the MCTS visit counts for each move. Game outcomes are presented in Figure 2. A selection of the self-play games is annotated and presented in our technical report, together with the qualitative assessments by Vladimir Kramnik.²⁶

Across all variants the percentage of drawn games increases with longer calculation times, suggesting that they may be theoretically drawn, as is believed to be the case for classical chess. However, when calculation is sliced at either 1 sec or 1 min, four variants consistently stand out with games that are more decisive than classical chess: Torpedo, Semi-torpedo, No-castling, and Stalemate=win. Conditioned on the games in Figure 2, the posterior probabilities that each of these four variants would yield a lower draw rate than classical chess is at least 99.9% at 1 sec calculation and 87% at 1 min calculation (see Tomašev et al.²⁶ for full analysis). Simply put, some variants may be harder to play, involving more calculation and richer patterns. Figure 2 adds data to a long-standing debate about whether letting stalemate count as a win would make top-level chess substantially more decisive.¹⁶ At 1 sec per move, there are 2.4% fewer Stalemate=win draws than under classical rules; at 1 min per move, this reduces to 0.8% fewer draws.

Utilization of Special Moves

The chess variants’ special moves play an important role in the arising dynamics of play, as evidenced by their frequency of use in the 1-min-per-move self-play games discussed in the previous section. The variants’ median game lengths ranged from 62 to 76 moves, while the median game under classical rules was 68 moves long. This suggests that none of the special moves would radically impact the time that a player spends at the board.

At least one torpedo move appeared in 83% of all Semi-torpedo games and 94% of Torpedo games. A pawn promotion with a torpedo move occurred in 21% of Torpedo games, highlighting the speed at which a passed pawn can be promoted to a queen. Backwards



Table 1. A list of considered alterations to the rules of chess.

Variant	Rule change
No-castling	Castling is disallowed.
No-castling (10)	Castling is disallowed for the first 10 moves (20 plies).
Pawn one square	Pawns can only move by one square.
Stalemate = win	Forcing stalemate is a win rather than a draw.
Torpedo	Pawns can move by one or two squares anywhere on the board and can be captured <i>en passant</i> after any two-square advance.
Semi-torpedo	Pawns can move by two squares both from the second and the third ranks, and they can be captured <i>en passant</i> after any two-square advance.
Pawn-back	Pawns can move backwards by one square but only back to the second/seventh rank for White/Black. Pawn moves do not count toward the 50-move rule.
Pawn-sideways	Pawns can also move sideways by one square. Sideways pawn moves do not count toward the 50-move rule.
Self-capture	It is possible to capture one's own pieces.

Figure 1. Examples by AlphaZero of three of the nine chess variants analyzed in this article. In Torpedo chess (left), White generates rapid counterplay with a torpedo move (b4-b6). ♖h1 is followed by yet another torpedo move, b6-b8=♖. In Pawn-sideways chess (center), Black plays a tactical sideways pawn move (f7-e7) after sacrificing a knight on f2 in the previous move, opening the f-file toward the White king. In Self-capture chess (right), White's self-capture move (♙xh4) generates threats against the Black king.

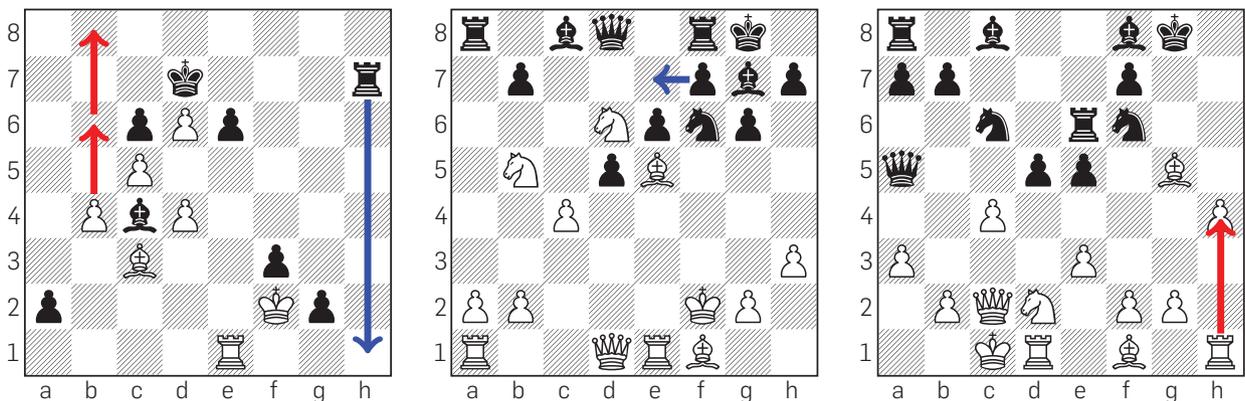
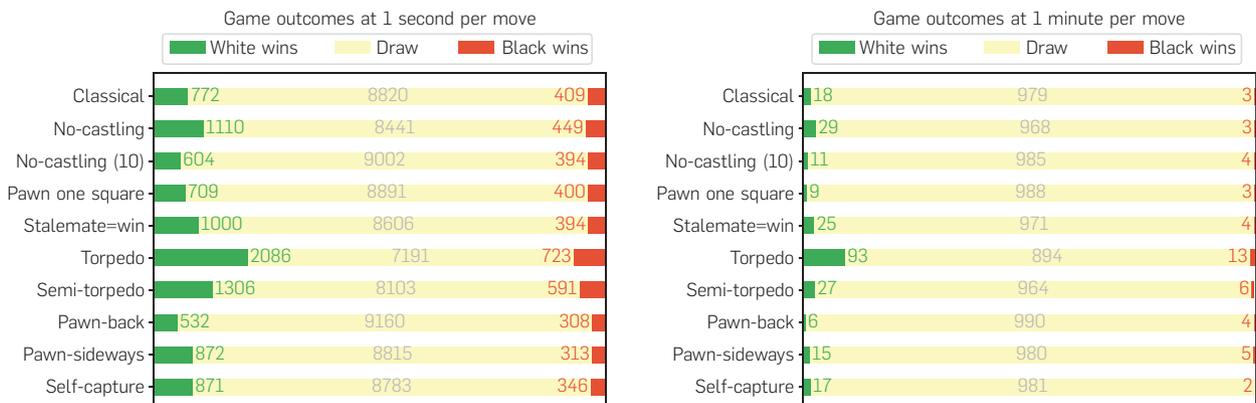


Figure 2. AlphaZero self-play game outcomes: for 10,000 games played at 1 sec per move (left) and for 1,000 games played at 1 min per move (right).



pawn moves occurred in 97% of Pawn-back games, while all Pawn-sideways games involved a sideways pawn move. A startling 12% of all moves in Pawn-sideways chess were sideways moves, demonstrating a high utilization of the newly introduced move. In Stalemate=win chess, the percentage of all decisive games that were won by stalemate rather than mate was 31%, although this number includes conclusions from endings such as ♔+♔ vs. ♔, where either conclusion is a valid win. In Self-capture chess, 42% of games featured self-capture moves, most commonly involving pawns (95%), while bishop (3%), knight (1%), and rook self-captures (1%) were rarer.

Diversity of Choices in Top-Level Play

Rule perturbations should not leave top-level players with only a few forcing lines of play and so diminish the richness of classical chess. Generally, if there are $M(s_t)$ legal moves in position or state s_t at ply t , then the number of candidate moves $m(s_t)$ —the number that a top player would realistically consider—is much smaller than $M(s_t)$. de Groot⁶ called $M(s_t)$ a player’s legal freedom of choice and $m(s_t)$ an objective freedom of choice. Iida et al.¹⁰ hypothesize that $m(s_t) \approx \sqrt{M(s_t)}$ on average, and we show that similar relationships hold across variants.

Choices in a single position. We

estimate the diversity of choice in a single position s_t via the entropy of the AlphaZero prior. The prior is a weighted list of possible moves to s_{t+1} from s_t that are used in AlphaZero’s MCTS search; it specifies candidates for consideration before MCTS calculation. The average information content, or **entropy**, $H(s_t) = -\sum_{s_{t+1}} p(s_{t+1} | s_t) \log p(s_{t+1} | s_t)$, represents the information content in the reasonable choices available in s_t . A more naturally interpretable number is the **average number of candidate moves** in the position, which we define as $m(s_t) = \exp(H(s_t))$. It is the number of *uniformly weighted* moves that could be encoded in the *same* number of nats as $p(s_{t+1} | s_t)$.^b These two quantities will be used to construct footprints for opening tree diversity.

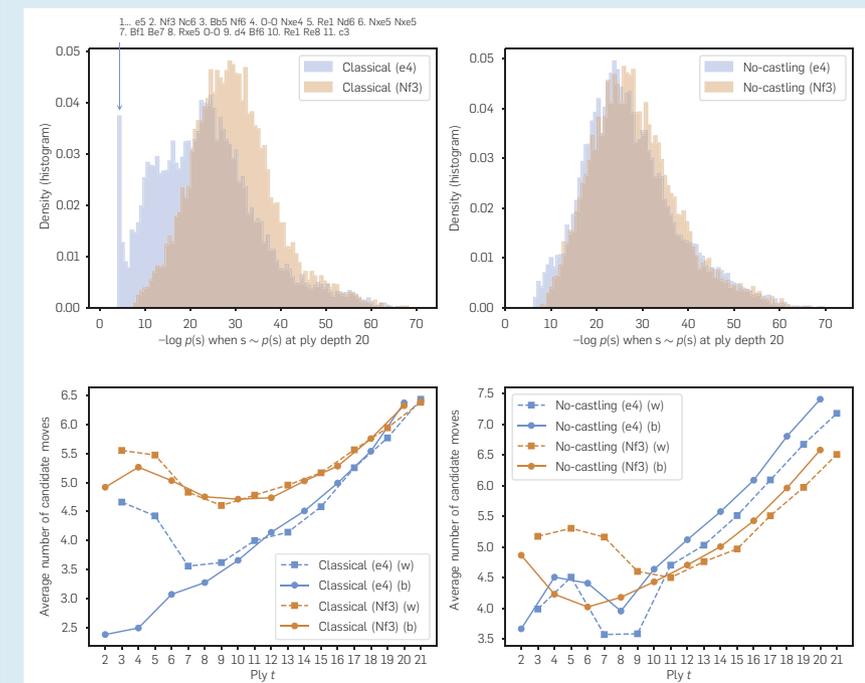
Opening tree diversity. We measure the diversity of choice with the entropy of the first T plies of play from each variant’s prior p from $f_\theta(s)$. If $s = [s_1, s_2, \dots, s_T]$ represents the sequence of states after T plies, the prior probability of s is $p(s) = \prod_{t=1}^T p(s_t | s_{t-1})$. The entropy in a sequence of T moves is hence $\mathcal{H}(T) = -\sum_s p(s) \log p(s) = \mathbb{E}_{s \sim p(s)}[-\log p(s)]$. An entropy $\mathcal{H}(T) = 0$ implies that, according to the prior, one and only one reasonable opening line could be considered by White and Black up to depth T , with all deviations from that line leading to substantially worse positions for the deviating side. A higher $\mathcal{H}(T)$ implies that we would *a priori* expect a wider opening tree of variations, and consequently a more diverse set of middlegame positions. $\mathcal{H}(T)$ contains an average over an exponential number of move sequences, which we approximate with a Monte Carlo estimate. Table 2 shows the estimated entropy of 20-ply opening trees for each variant. As one of the most drawish variants, players of Pawn-one-square chess will have substantially more playable candidate moves at their disposal, despite there being fewer legal moves than any other variant.

^b As an illustrative example, if the number of candidate moves is $m(s_t) = 3$ for some $p(s_{t+1} | s_t)$ that might put nonzero mass on all of its moves, then $m(s_t)$ is also equal to the number of candidate moves of a probability vector $p = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, \dots, 0]$ that puts equal nonzero mass on only three moves.

Table 2. The entropy (in nats) of the first 20 plies of the AlphaZero prior, with an estimate of the number of lines in the equivalent opening book.

Variant	Entropy	20-ply Games	Variant	Entropy	20-ply Games
No-castling	27.65	1.02×10^{12}	Stalemate = win	29.01	3.97×10^{12}
Torpedo	27.89	1.30×10^{12}	Semi-torpedo	31.63	5.45×10^{13}
Self-capture	27.94	1.36×10^{12}	Pawn-back	32.30	1.07×10^{14}
No-castling (10)	27.97	1.40×10^{12}	Pawn-sideways	34.16	6.85×10^{14}
Classical	28.58	2.58×10^{12}	Pawn one square	38.95	8.24×10^{16}
			Uniform random	64.96	1.63×10^{28}

Figure 3. Statistical footprints of the diversity of responses to 1. e4 and 1. ♖f3 in Classical and No-castling chess, as well as the average number of candidate moves available for White and Black at each ply.



Shannon’s source coding theorem states that we can compress a game sampled from $p(s)$ into just over $\mathcal{H}(T)$ nats. This is equivalent to assigning a unique number to each of $\exp(\mathcal{H}(T))$ games, and we take this as the size of a variant’s opening book: the number of plausible T -ply games in a variant.

The “uniform random” policy in Table 2 plays all legal moves in classical chess with equal probability. Its entropy is more than double that of classical chess; conversely, like the hypothesis by Iida et al.,¹⁰ the classical opening book is a little smaller than the square root of the number of all legal openings.

Classical vs. no-castling chess.

Both White and Black castle in most classical chess openings, and removing castling as an option profoundly changes the characteristics of the game.¹⁴ In this section, we explore the changes using one opening, the Berlin Defence, as an example. A tool is the decomposition of the entropy $\mathcal{H}(T)$ ’s statistical expectation, which can help identify the existence (or not) of defensive lines that equalize the game in an almost forceful way. The top row in Figure 3 shows histograms of $-\log p(s)$ when s is generated from the AlphaZero prior after 1. e4 and 1. ♘f3. The histograms provide a footprint of opening diversity.

Berlin defence. In classical chess, one equalizing defensive resource for Black in the Ruy Lopez (1. e4 e5 2. ♘f3 ♘c6 3. ♗b5) is the Berlin Defence, starting with 3... ♘f6. Vladimir Kramnik successfully used it as a defensive resource for Black during his world championship match against Garry Kasparov in 2000. Prior to the match, chess engines of the time evaluated the Berlin endgame at around +1 in White’s favor, but today it is considered to be very solid, with modern engines assessing most arising positions as being equal.²⁶

In Figure 3 (top left), samples of s that contribute to the high probability spike after 1. e4 correspond to AlphaZero’s strong preference for the Berlin Defence in classical chess. In the footprint of lines after 1. ♘f3, the spike disappears, indicating a wider range of possibilities for both sides. White’s main response to 3. ♘f6 is

Table 3. Estimated piece values from AlphaZero self-play games for each variant.

Variant	♔	♚	♙	♜	♞	Variant	♔	♚	♙	♜	♞
Classical	1	3.05	3.33	5.63	9.5	Self-capture	1	3.10	3.22	5.34	9.42
No-castling	1	2.97	3.13	5.02	9.49	Pawn-back	1	2.65	2.85	4.67	9.39
No-castling (10)	1	3.14	3.40	5.37	9.85	Semi-torpedo	1	2.72	2.95	4.69	8.3
Pawn one square	1	2.95	3.14	5.36	9.62	Torpedo	1	2.25	2.46	3.58	7.12
Stalemate = win	1	2.95	3.13	4.76	8.96	Pawn-sideways	1	1.8	1.98	2.99	5.92

4. O-O. Without the option to castle, the Berlin Defence and many other lines either disappear or become less prominent, where 1. e4 yields a similar variety of preferred lines of play as 1. ♘f3.

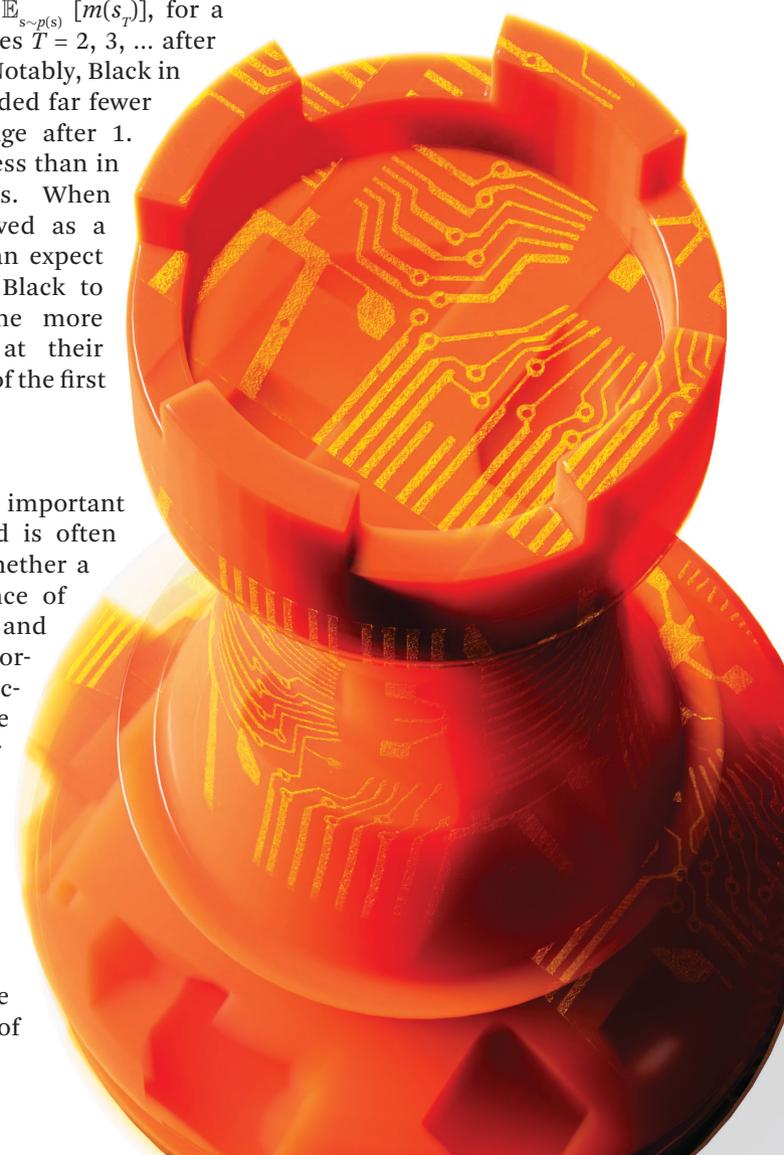
Average number of candidate moves. The entropy of a chess variant’s prior opening tree is an unwieldy number that does not immediately inform us how many move options are in each variant. Instead, Figure 3 (bottom row) also shows the average number of candidate moves at ply T , $\mathcal{M}(T) = \sum_s p(s) m(s_T) = \mathbb{E}_{s \sim p(s)} [m(s_T)]$, for a progression of plies $T = 2, 3, \dots$ after 1. e4 and 1. ♘f3. Notably, Black in particular is afforded far fewer options on average after 1. e4 in classical chess than in No-castling chess. When castling is removed as a legal move, we can expect both White and Black to have at least one more plausible move at their disposal for each of the first 20 plies.

Piece Value

Material plays an important role in chess and is often used to assess whether a particular sequence of piece exchanges and captures is favorable. Material sacrifices in chess are made either for concrete tactical reasons—for example, mating attacks—or as a trade-off for long-term positional strength. Understanding the material value of

pieces in chess helps players master the game and is one of the very first pieces of chess knowledge taught to beginners.

We approximate piece values via the weights of a linear model trained to predict game outcome from differences in piece counts for any given position, given as $d = [1, d_\Delta, d_\ominus, d_\blacktriangleleft, d_\blacktriangleright, d_\blacklozenge]$. We fit the model on fast-play AlphaZero games for each of the chess variants. We define g_w with weights $w \in \mathbb{R}^6$ as $g_w(s) = \tanh(w^T d)$. In the linear model, the weights w indicate relative



piece importance. If $(s, z) \sim \text{games}$ represent a sample of a position and final game outcome from a variant's self-play games, we minimize $\ell(w) = \mathbb{E}_{(s,z) \sim \text{games}} [(z - g_w(s))^2]$ empirically over w and normalize weights w by w_{Δ} to yield the relative piece values. The recovered piece values for each of the chess variants are given in Table 3.

Looking at the piece value estimates for classical chess, the method approximately recovers known material values^{4,12} and identifies bishops as more valuable than knights. Estimates of piece values in No-castling, No-castling (10), Pawn-one-square, Self-capture, and Stalemate = win variants look fairly similar, which is not surprising given the minor differences in piece mobility compared to the other variants. Variants involving an increase in pawn mobility result in lower relative values for other pieces, as can be seen in Pawn-back, Semi-torpedo, Torpedo, and Pawn-sideways. In Pawn-sideways chess, AlphaZero often sees the trade of a knight or bishop for two pawns as favorable, in accordance with this approximation. Such an exchange would normally be considered bad in classical chess. Material values may vary across different game stages and position types, and hence, the values in Table 3 are merely meant to help new players make sense of tactical exchanges in these chess variants.

From Reimagining to Reality

The combination of human curiosity and a powerful reinforcement learning system allowed us to reimagine what chess would have looked like if history had taken a slightly different course. When the statistical properties of top-level AlphaZero games are compared to classical chess, a number of more decisive variants appear, without impacting the diversity of plausible options available to a player. Aside from a mathematical evaluation, the actual games could be viewed through an aesthetic lens; this was done in Tomašev et al.²⁶ and in many online discussion forums.¹⁴

Taken together, the statistical properties and aesthetics provide evidence that some variants would lead to games that are at least as engaging as classical chess. Variants such as

Torpedo, Pawn-sideways, No-Castling, and Self-capture are now a reality, playable on a major chess portal such as chess.com.⁵ On the back of the initial evidence, the first No-castling tournament was held in Chennai in January 2020.²³ Chess's role in artificial intelligence research is far from over. The outcome of this article is a result of a *man with machine*, showing how AI can provide the evidence to take reimagining to reality.



Looking beyond chess, this article's contribution hinged on being able to learn a policy for an agent in an environment with known dynamics and then exploring changes in the environment to measure different emergent properties of agent behavior.²⁸ We believe that a similar approach could be used for auto-adjusting game mechanics in other types of games, including computer games, in cases where a sufficiently strong reinforcement learning system is available. **C**

References

1. Beasley, J. What can we expect from a new chess variant? *Variant Chess* 4, 29 (1998), 2.
2. Berliner, H. 1st U.S. Computer Chess Championship. *Chess Live* (November 1970), 638.
3. Campbell, M., Hoane Jr., J., and Hsu, F-h. Deep blue. *Artif. Intell.* 134, (2002), 57–83.
4. Capablanca, J.R. and de Firmian, N. *Chess Fundamentals: Completely Revised and Updated for the 21st Century*. Random House Puzzles & Games, (2006).
5. Chess.com variants. Chess.com. <http://chess.com/variants>.
6. de Groot, A.D. *Het Denken van den Schaker*. (Thought and Choice in Chess). Amsterdam University Press, 1946.

7. Gligorić, S. *Shall We Play Fischerandom Chess?* Batsford, 2002.
8. Gollon, J. *Chess Variations: Ancient, Regional and Modern*. Charles E. Tuttle Company, 1968.
9. Heinz, E.A. *Scalable Search in Computer Chess*. Vieweg+Teubner Verlag (2000).
10. Iida, H., Takeshita, N., and Yoshimura, J. A metric for entertainment of boardgames: Its implication for evolution of chess variants. In *Entertainment Computing: Technologies and Application*, R. Nakatsu and J. Hoshino, eds. (2003). 65–72.
11. Kasparov, G. *Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins*. John Murray, 2017.
12. Kaufman, L. The evaluation of material imbalances. *Chess Life* (1999).
13. Knuth, D.E. and Moore, R.W. An analysis of alpha-beta pruning. *Artif. Intell.* 6, 4 (1975), 293–326.
14. Kramnik, V. Kramnik and AlphaZero: How to rethink chess. *Chess.com* (December 2, 2019), <https://chess.com/article/view/no-castling-chess-kramnik-alphazero>.
15. LCZero Development Community. Leela Chess Zero. <https://lczero.org>.
16. Lillebo, P. Stalemate: The long and the short of it. *ChessBase* (Aug. 2, 2014), <https://en.chessbase.com/post/stalemate-the-long-and-the-short-of-it>.
17. Murray, H.J.R. *The History of Chess*. Oxford University Press, 1913.
18. Newborn, M. *Computer Chess*. Academic Press,
19. Nielsen, P.H. When Magnus met AlphaZero. *New Chess* (December 2019), 14–23.
20. Pritchard, D.B. *The Classified Encyclopedia of Chess Variants*. Games and Puzzles Publications (1994).
21. Rosin, C.D. Multi-armed bandits with episode context. *Ann. Math. Artif. Intell.* 61, 3 (2011), 203–230.
22. Sadler, M. and Regan, N. *Game Changer: AlphaZero's Groundbreaking Chess Strategies and the Promise of AI*. New In Chess (February 2019).
23. Shah, S. First ever "no-castling" tournament results in 89% decisive games! *ChessBase* (January 19, 2020), <https://en.chessbase.com/post/the-first-ever-no-castling-chess-tournament-results-in-89-decisive-games>.
24. Shannon, C.E. Programming a computer for playing chess. *Philos. Mag.* 41, 314 (1950), 256–275.
25. Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., and Hassabis, D. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362, 6419 (2018), 1140–1144.
26. Tomašev, N., Paquet, U., Hassabis, D., and Kramnik, V. Assessing game balance with AlphaZero: exploring alternative rule sets in chess (Sept. 2020). arXiv:cs.AI/2009.04374
27. Turing, A. Digital computers applied to games. In *Faster Than Thought: A Symposium on Digital Computing Machines*, B.V. Bowden ed. Pitman Publishing, London, 1953, 286–310.
28. Zhang, H., Wang, J., Zhou, Z., Zhang, W., Wen, Y., Yu, Y., and Li, W. Learning to design games: Strategic environments in reinforcement learning. In *Proceedings of the 27th Intern. Joint Conf. on Artificial Intelligence (2018)*, J. Lang ed, 3068–3074.

Nenad Tomašev is a research scientist at DeepMind Technologies Ltd.

Ulrich Paquet is a research scientist at DeepMind Technologies Ltd.

Demis Hassabis is the founder and CEO of DeepMind Technologies Ltd.

Vladimir Kramnik is a former World Chess Champion (2000–2007).

Copyright held by authors/owners.



Watch the authors discuss this work in the exclusive *Communications* video. <https://cacm.acm.org/videos/reimagining-chess-alphazero>