ALPHA GO ZERO

DIE REVOLUTION DURCH SELBSTLERNENDE KI

Quelle: https://heise.cloudimg.io/width/1220/q50.png-lossy-50.webp-lossy-50.foil1/_www-heise-de_/imgs/18/2/3/0/0/7/5/2/Unbenannt-2-7276ff8ecdc85f34.png

GLIEDERUNG

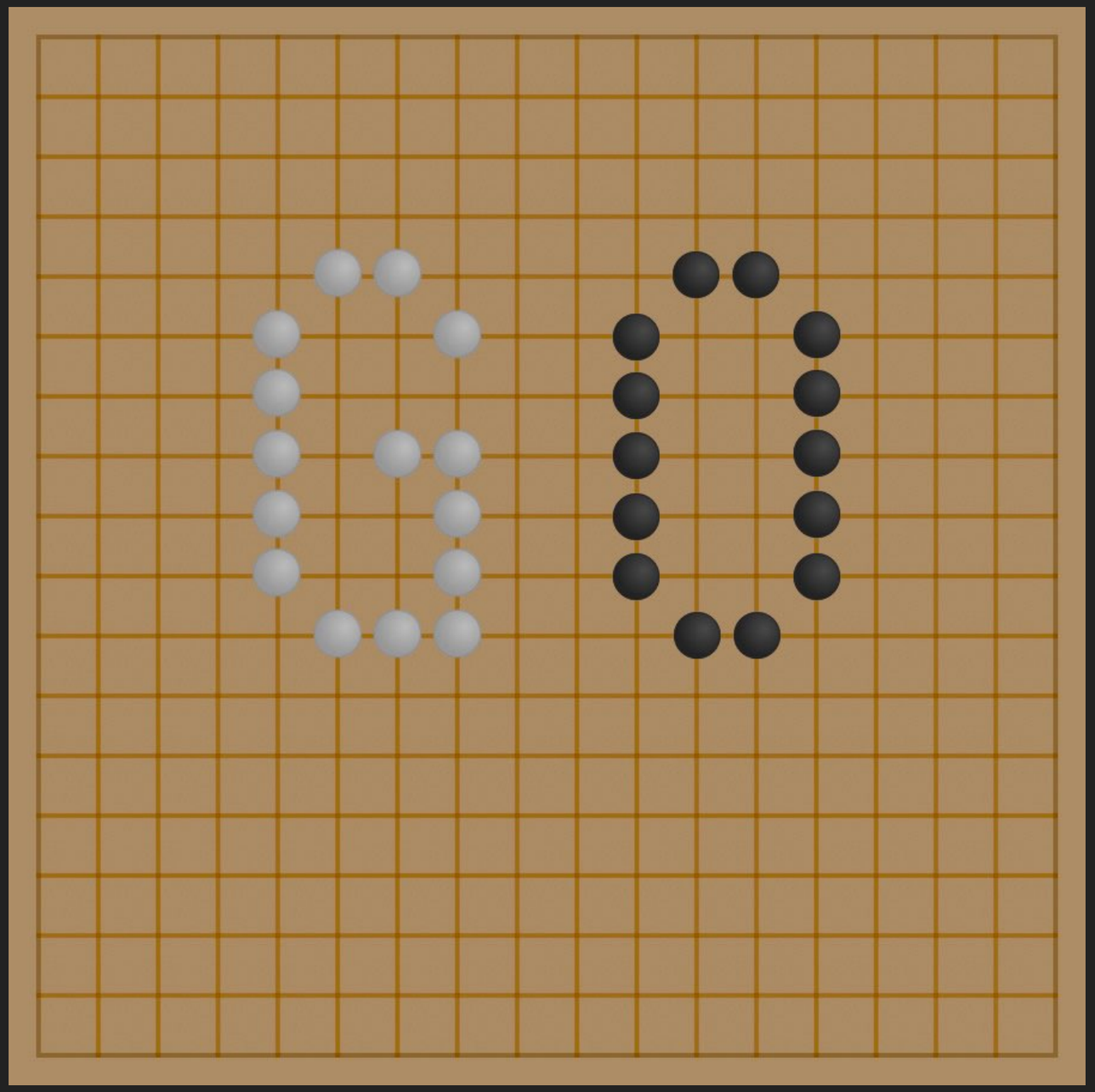
- Was ist Go?
- Was macht Go für Computer und KI so komplex?
- Historie der Go KIs von Google DeepMind
- Alpha Go Zero Algorithmus
- Ergebnisse und Performance

GLIEDERUNG

- Was ist Go?
  - Spielaufbau
  - Begriffserklärung: Freiheiten und Ketten
  - Grundlegende chinesische Regeln

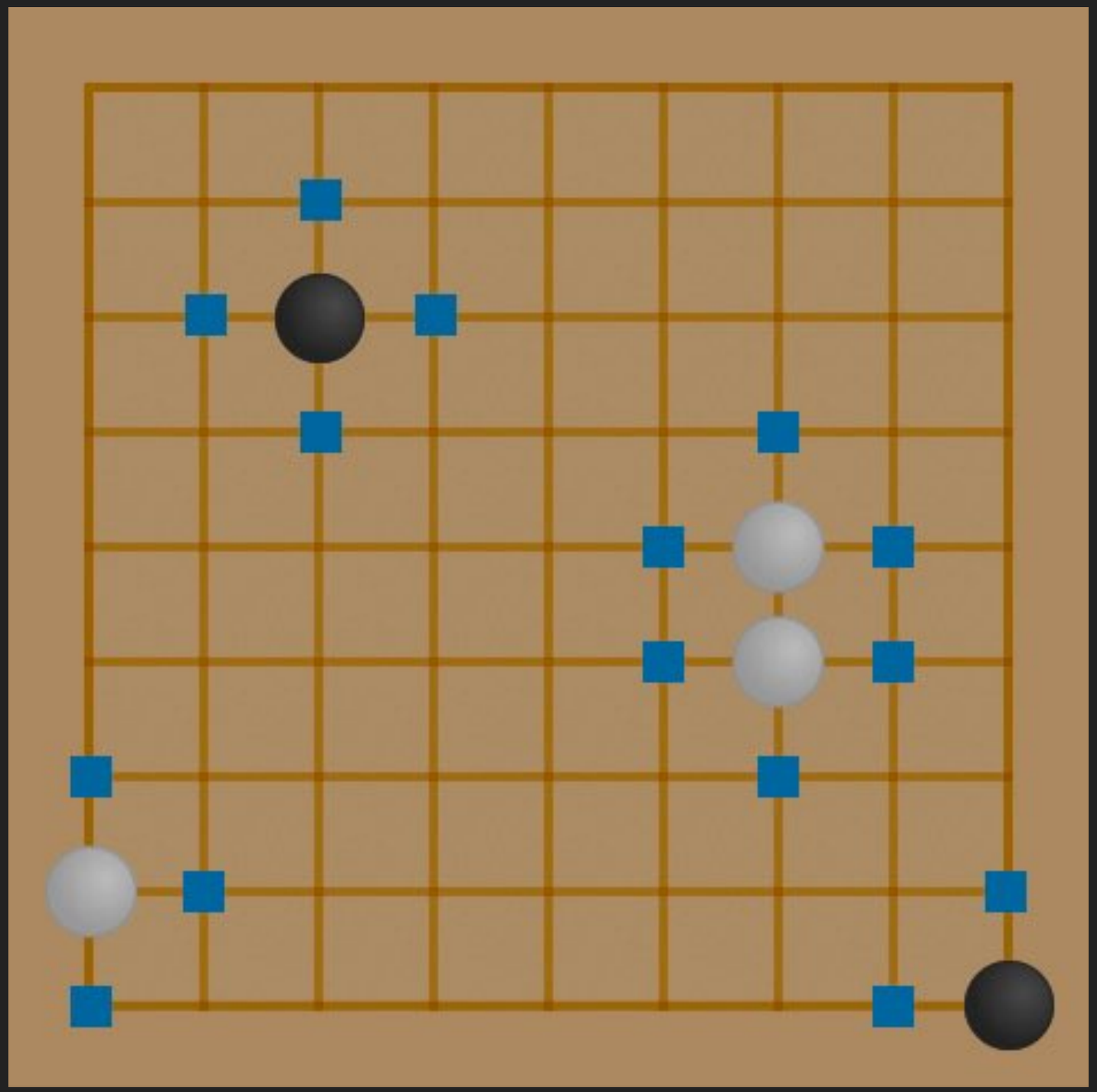
SPIELAUFBAU

- Spielbrett: 19 x 19 Gitter
- 2 Spieler: schwarz und weiß
- #Steine: unbegrenzt
- Steine werden auf die Schnittpunkte gesetzt



Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/spielflaeche.jpg

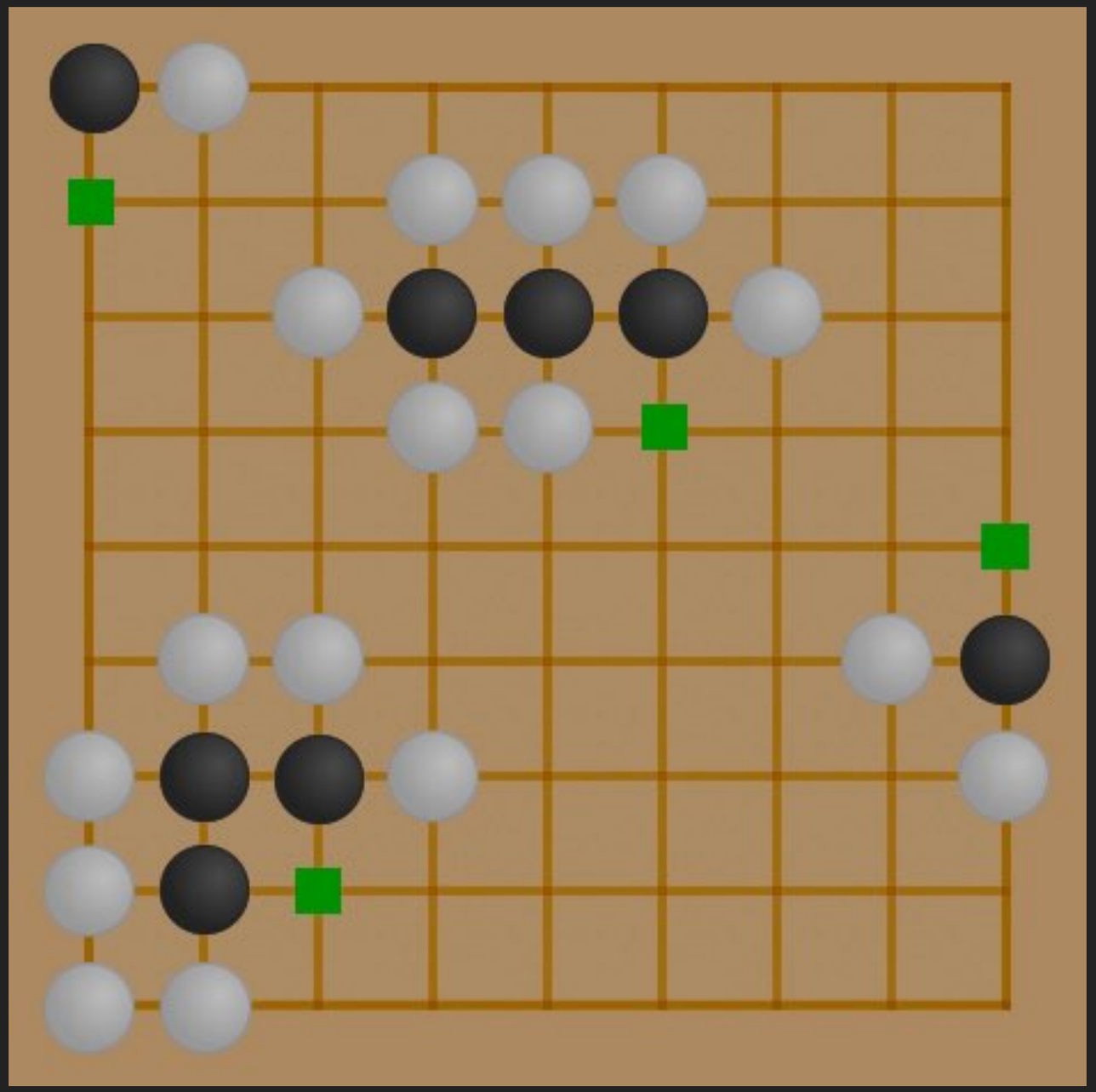
BEGRIFFSERKLÄRUNG: FREIHEITEN UND KETTEN



Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/ketten-und-freiheiten.jpg

GRUNDLEGENDE CHINESISCHE REGELN

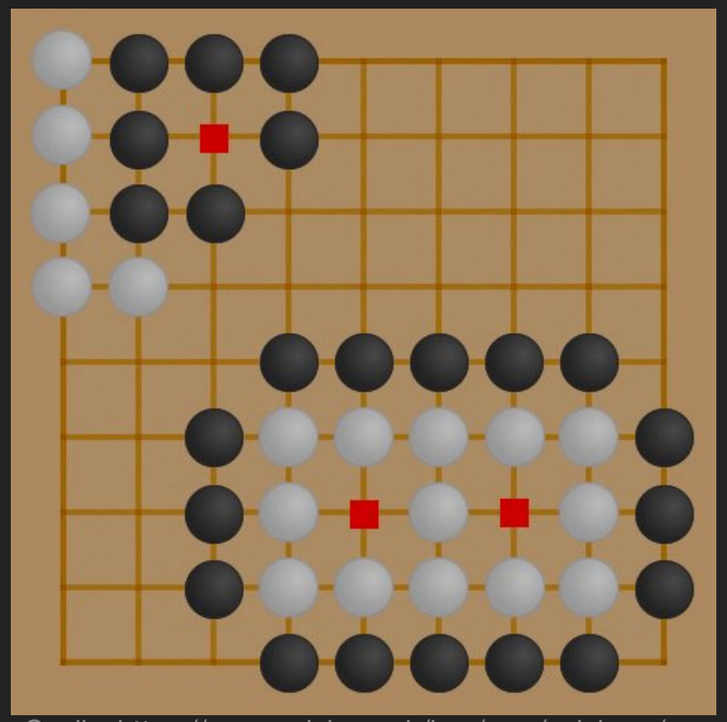
- Schwarz beginnt
- Spieler legen abwechselnd (oder passen)
- Steine ohne Freiheiten werden nach dem Setzen entfernt
- Schlagen



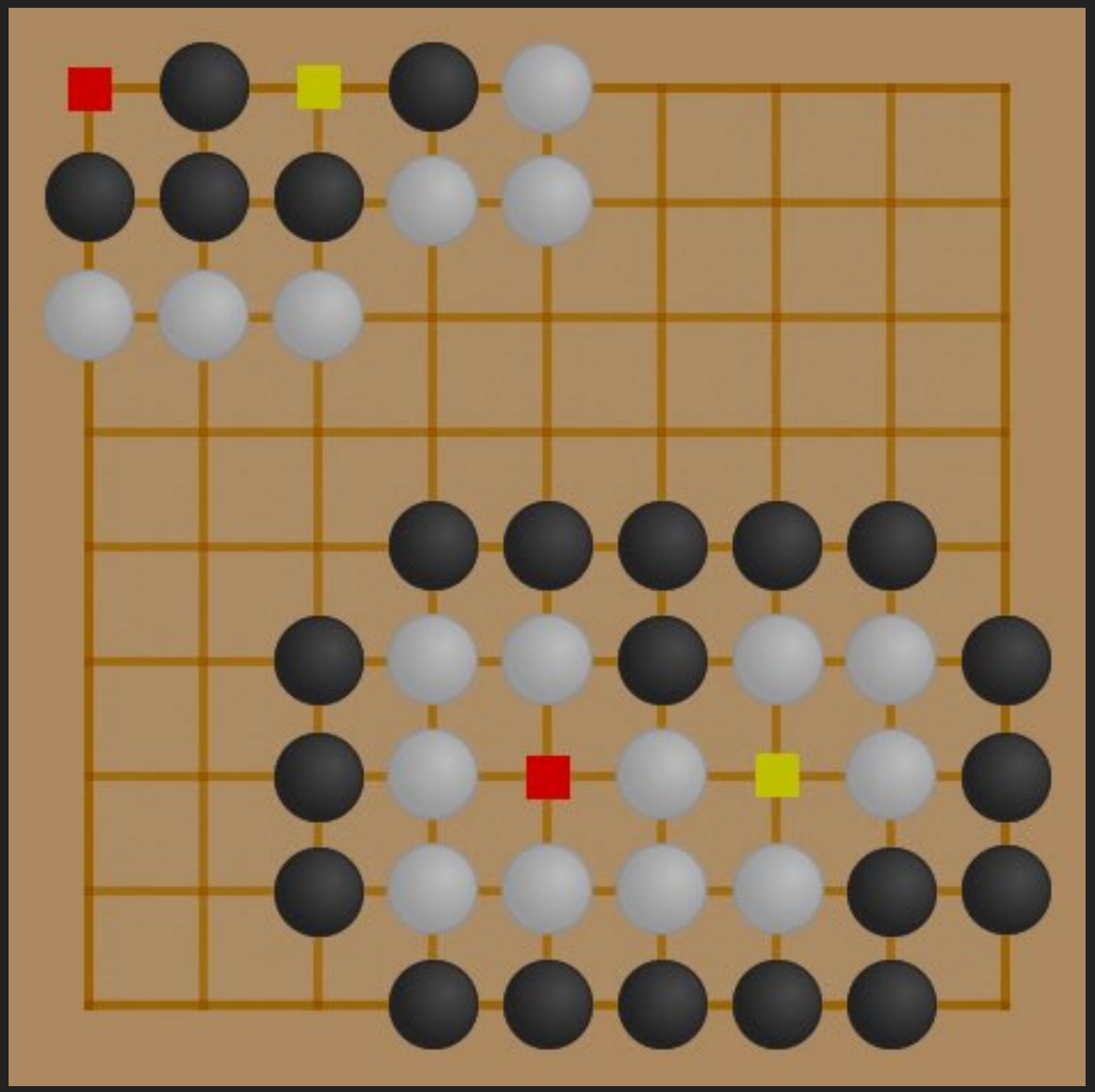
Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/schlagen.jpg

GRUNDLEGENDE CHINESISCHE REGELN

- Kein Spieler darf Steine ohne Freiheiten setzen
- Augen (rot)
- Unechte Augen (gelb)



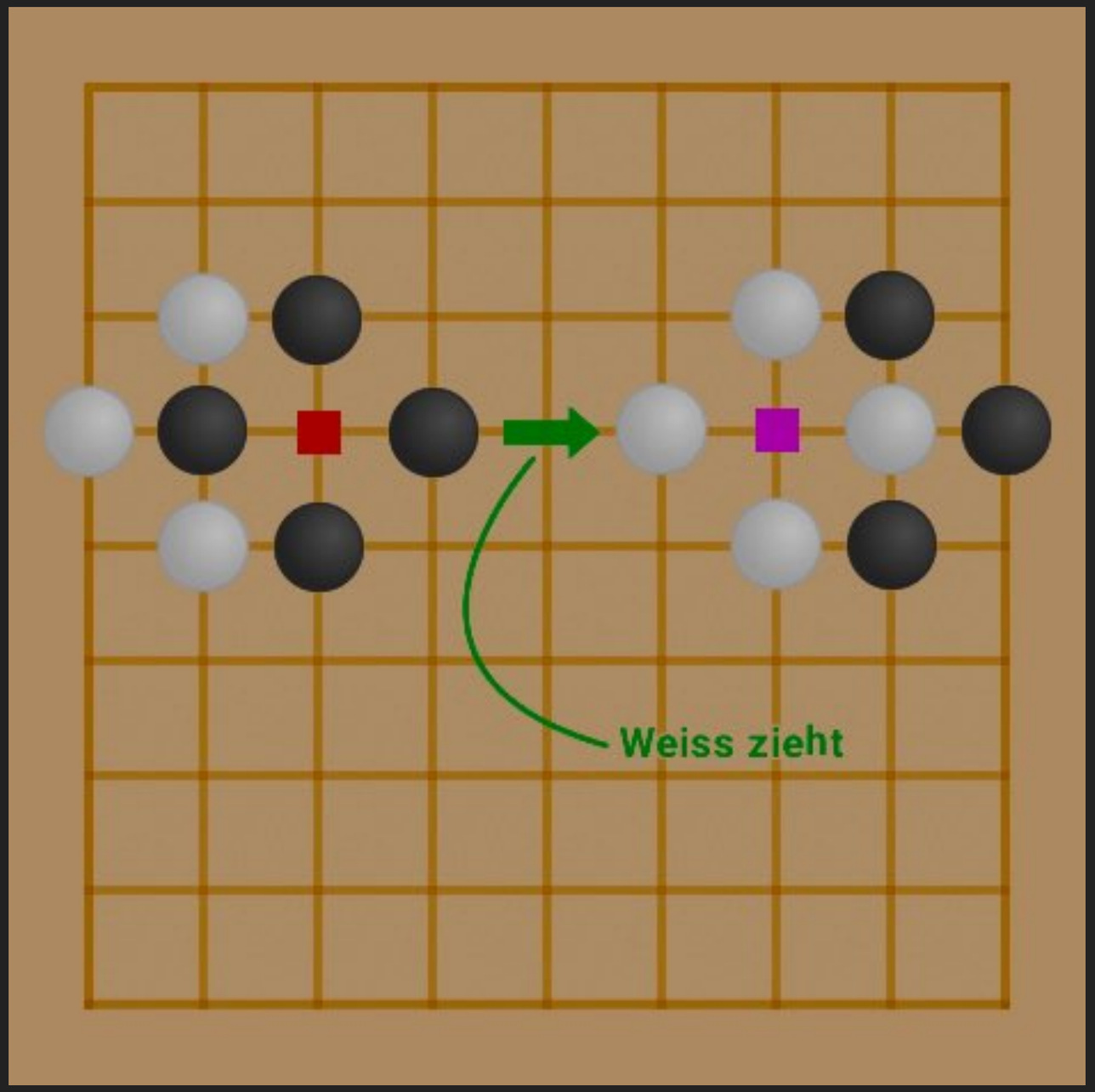
Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/augen.jpg



Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/unechte-augen.jpg

GRUNDLEGENDE CHINESISCHE REGELN

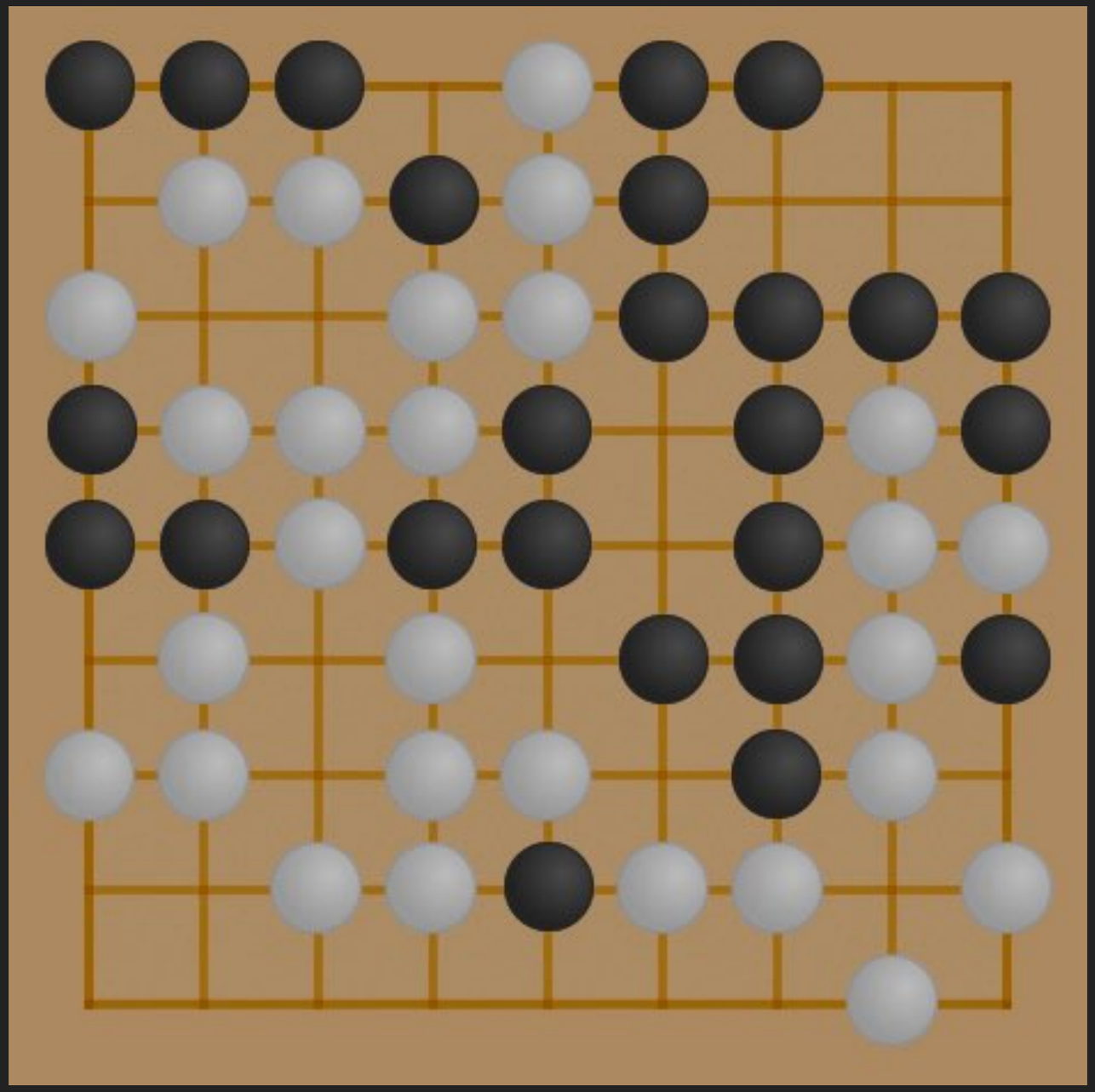
- Ko-Regeln:
  Wiederholungen von Brettpositionen sind verboten



Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/ko-regel.jpg

GRUNDLEGENDE CHINESISCHE REGELN

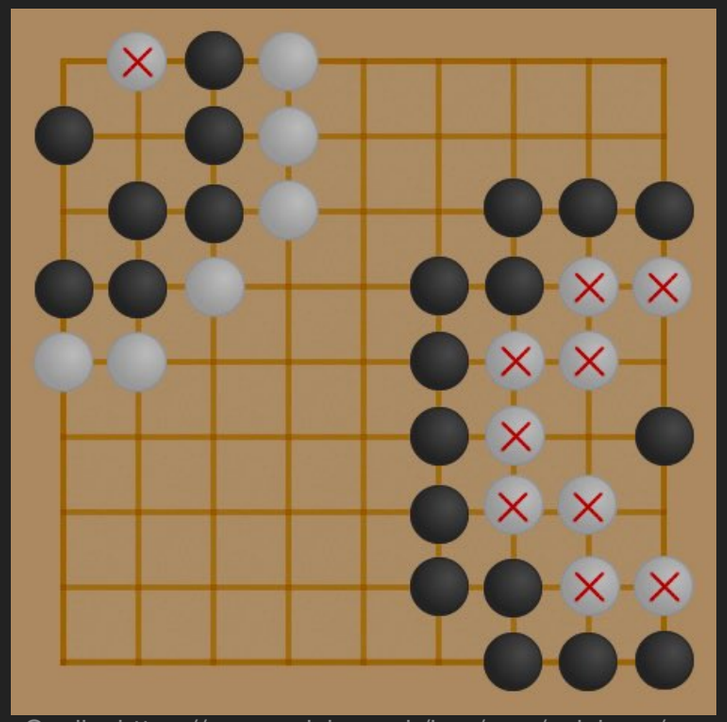
- Spielende:
  Beide Spieler stimmen überein, dass es keinen Zug mehr gibt und passen nacheinander



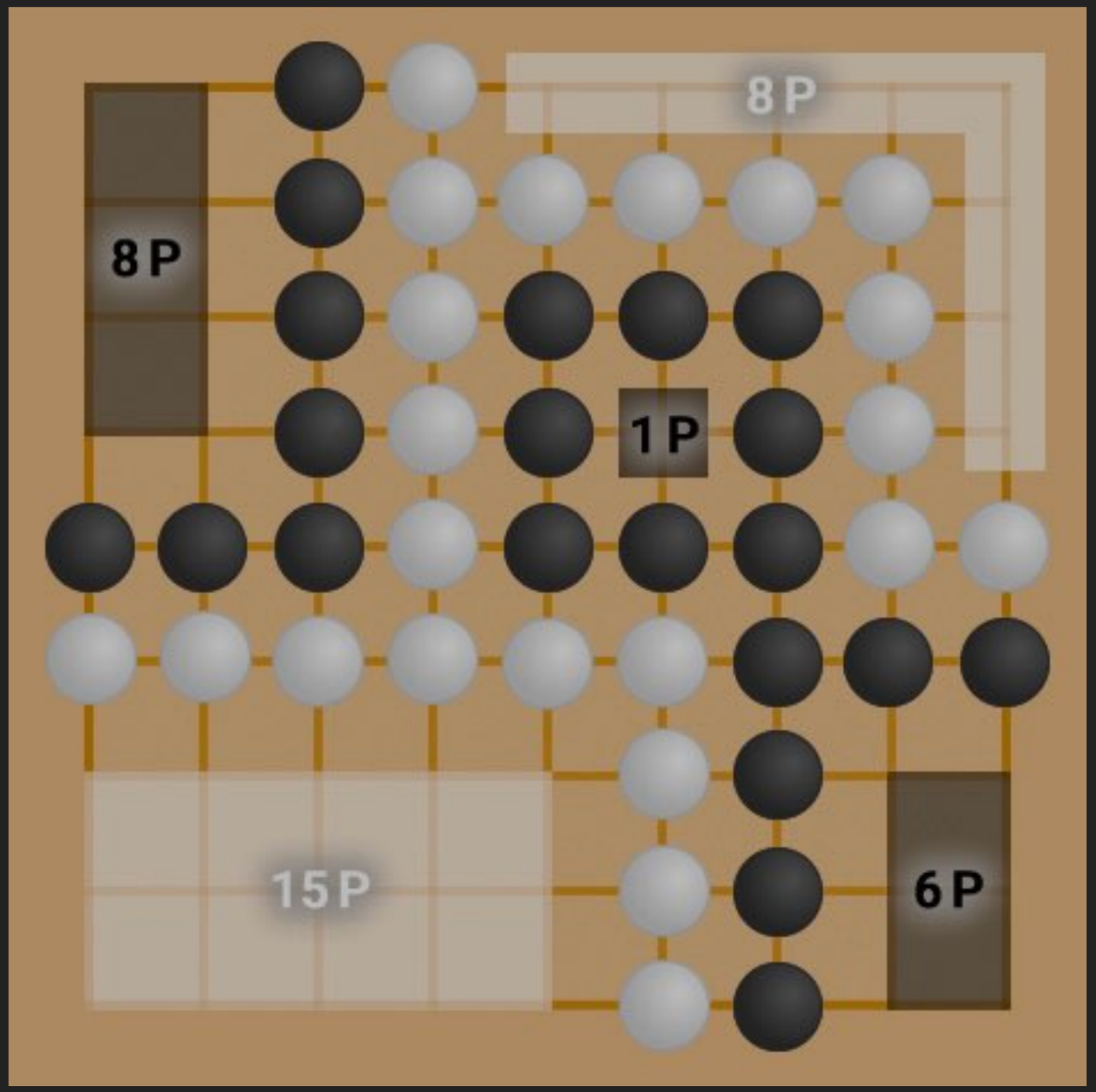
Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/spielende.jpg

GRUNDLEGENDE CHINESISCHE REGELN

- End-Wertung:
  - Tote Steine entfernen
  - Lebende Steine und Gebiete zählen
  - 181 Punkte zum Sieg



Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/tote-steine.jpg



Quelle: https://www.spielezar.ch/img/cms/spielezar/blog/artikel/brettspiele/go/umzingelte-gebiete.jpg

GLIEDERUNG

- Was macht Go für Computer und KI so komplex?

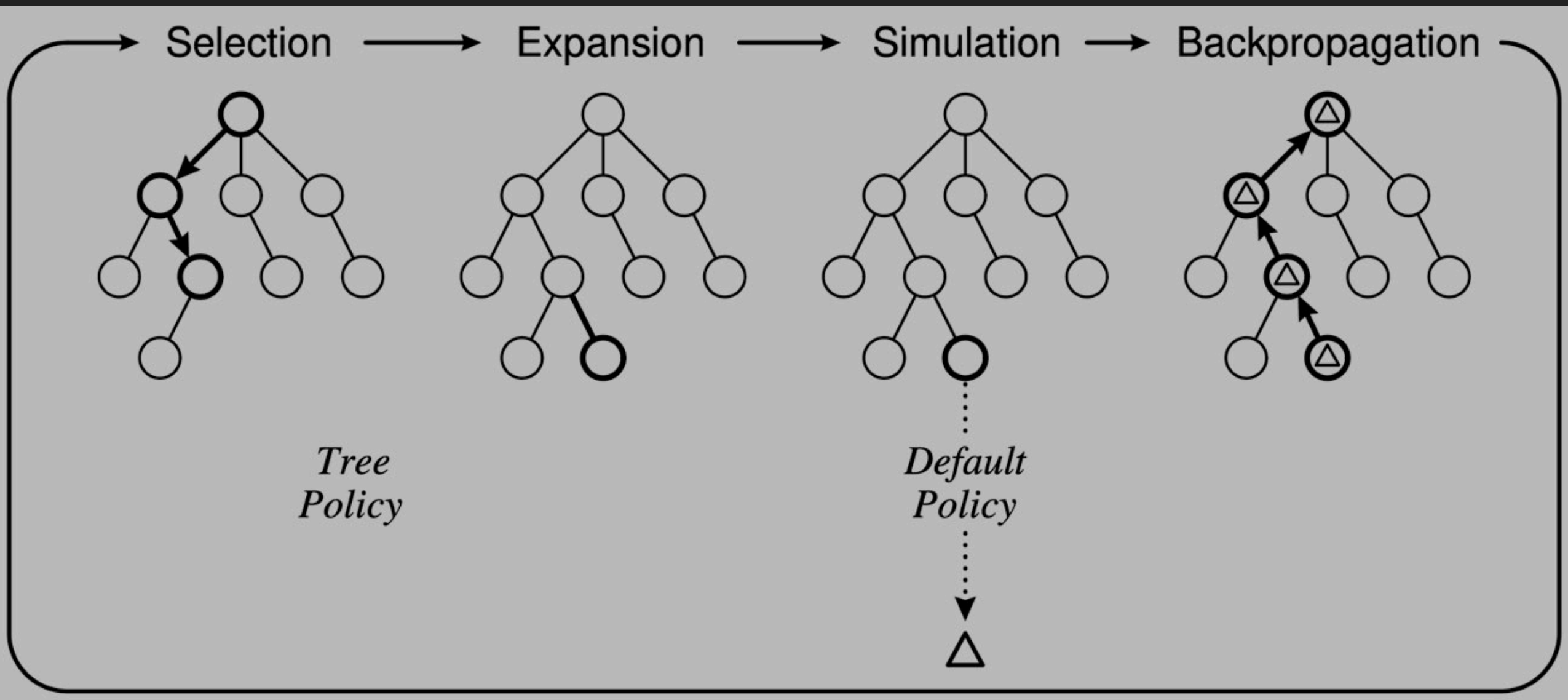
WAS MACHT GO FÜR COMPUTER UND KI SO KOMPLEX?

- Größte Herausforderung im Mind Game Bereich
- Viele Setzmöglichkeiten
- Statische Analyse langsam
- Stark musterbasiert
- Großer Suchbaum mit hohem Branchingfaktor
- $O(b^d) = O(250^{150})$

GLIEDERUNG

- Historie der Go KIs von Google DeepMind
  - Alpha Go Fan
  - Alpha Go Lee
  - Alpha Go Master

MONTE CARLO TREE SEARCH (MCTS)



Quelle: Browne, Cameron B., et al. "A survey of monte carlo tree search methods." IEEE Transactions on Computational Intelligence and AI in games 4.1 (2012): 1-43. Fig. 2

ALPHA GO FAN

- Oktober 2015 5-0 gegen Fan Hui (Europameister)
- Strategie-Netz und Bewertungs-Netz
- MCTS mit schneller Rolloutstrategie
- Ausführliche Features
- Supervised Learning und Self Play Reinforcement Learning (Strategie-Netz)
- Training auf Self Play Spieldaten des Strategie-Netzes (Bewertungs-Netz)

ALPHA GO LEE

- Ähnelt Alpha Go Fan
- März 2016 4-1 gegen Lee Sedol
- Tiefere Netze
- Training auf Self Play Spielen von Alpha Go (Bewertungs-Netz)

ALPHA GO MASTER

- Januar 2017 60-0 gegen die besten menschlichen Spieler
- Grundlegend gleicher Algorithmus wie bei Alpha Go Zero
- Unterschiede:
  - Lernen wird überwacht gestartet
  - Features und MCTS Rollouts werden wie bei Alpha Go Lee verwendet

GLIEDERUNG

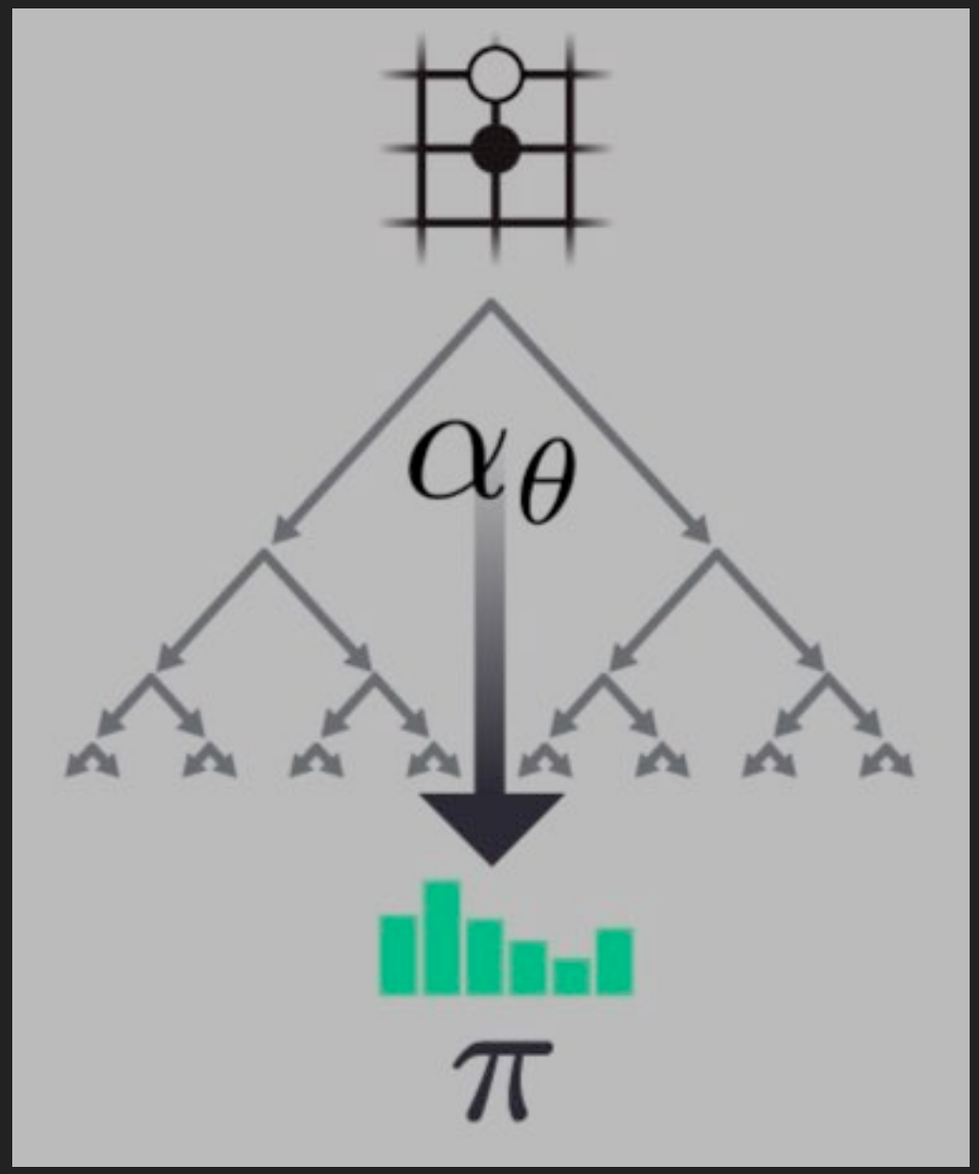
- Alpha Go Zero Algorithmus
  - Allgemein
  - Domänen Wissen
  - Tromp Taylor Scoring
  - Neuronales Netz
  - MCTS
  - Self Play Reinforcement Learning

ALLGEMEIN

- MCTS ohne Rollout
- Nur ein Neuronales Netz
- Kein überwachtes Lernen
- Kein menschliches Wissen über Spielzüge
- Nur echte Brettpositionen und Spielregeln

ALLGEMEIN

- Training: Self Play Reinforcement Learning eines Neuronalen Netzes
- Fertig trainierter Algorithmus: Angepasste MCTS berechnet nächsten Zug mit Unterstützung dieses Netzes



Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Fig. 2 d

DOMÄNEN WISSEN

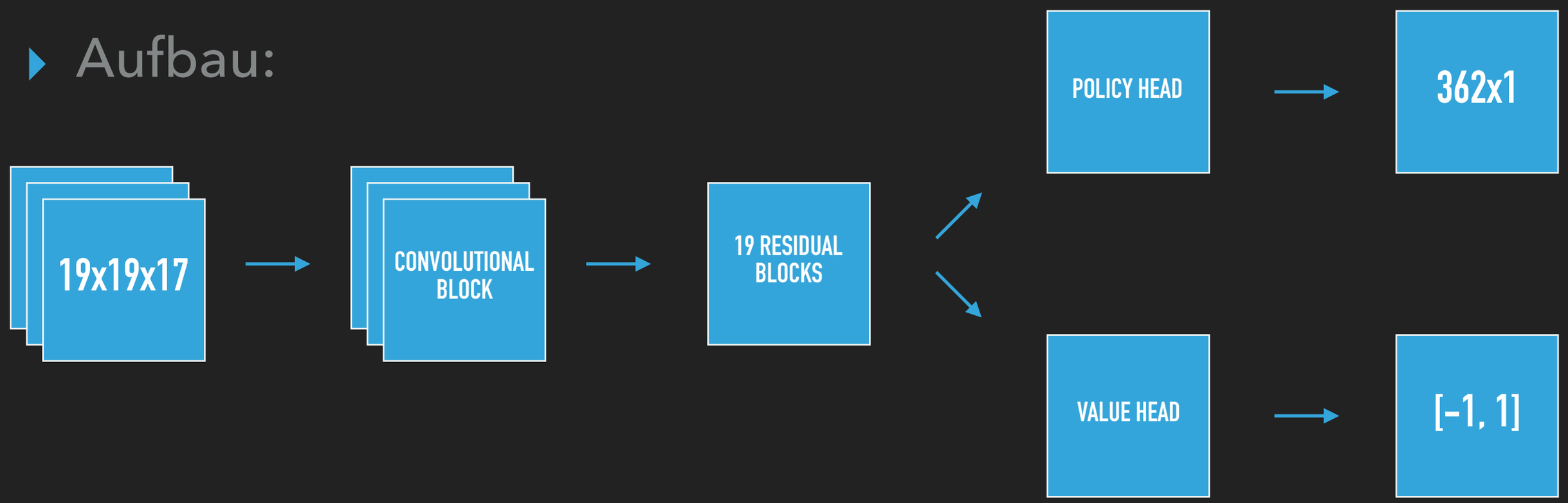
- Spielregeln (erlaubte Züge)
- Rotationen und Spiegelungen des Brettes verändern die Regeln und Möglichkeiten nicht
- Spielerfarbe hat keinen Einfluss auf die Regeln und Möglichkeiten
- Tromp Taylor Scoring zur Endabrechnung
- Kein weiteres menschliches Wissen

TROMP TAYLOR SCORING

- Vereinfachte Form der chinesischen Regeln
- Spielerendpunkte:
  Anzahl eigene Steine + eingeschlossene leere Felder
- Spieler mit höherer Punktzahl gewinnt

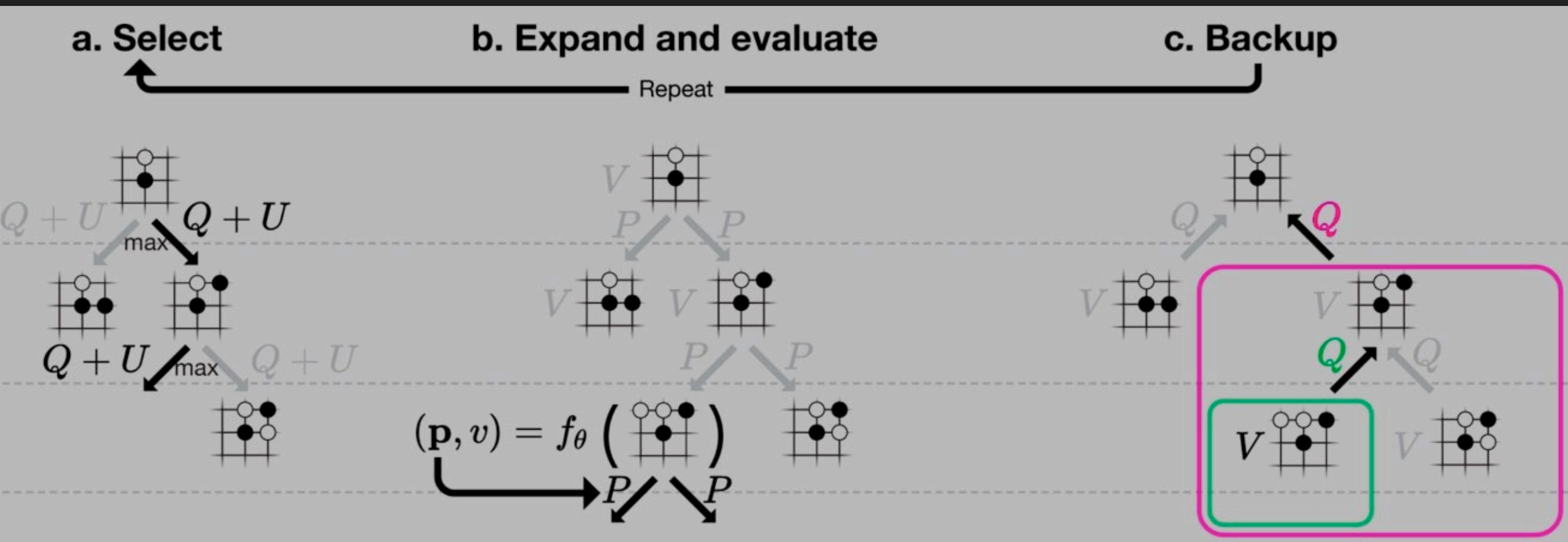
NEURONALES NETZ

- Kombiniertes Strategie- und Bewertungs-Netz
- Input: 19x19x17 mit binären Werten
- $[X_t, Y_t, X_{t-1}, Y_{t-1}, \ldots, X_{t-7}, Y_{t-7}, C]$
- Aufbau:



MCTS

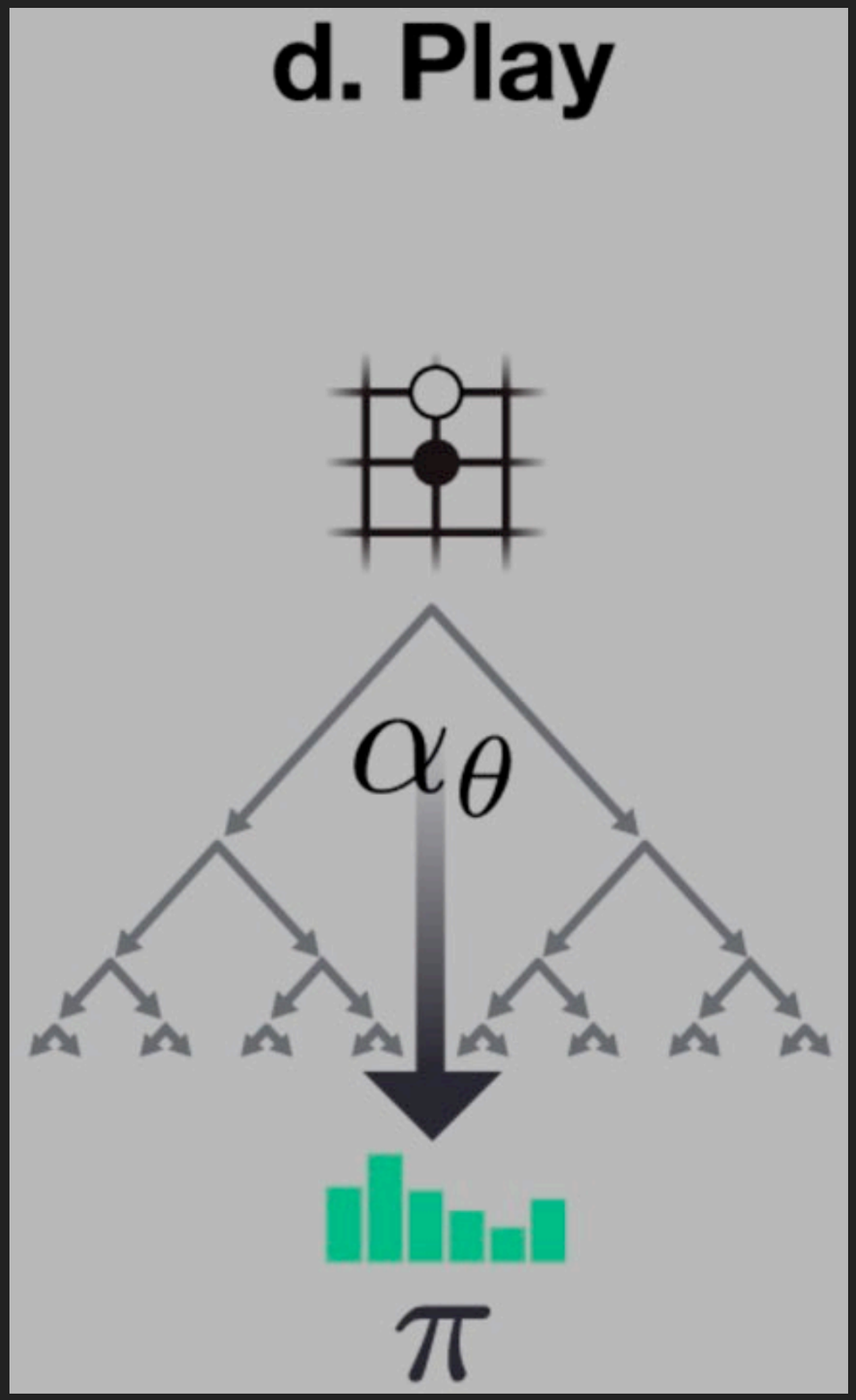
- Bekannt: Spielregeln (legale Züge zu Position)
- Kanteninformationen: {N(s,a), W(s,a), Q(s,a), P(s,a)}



Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Fig. 2

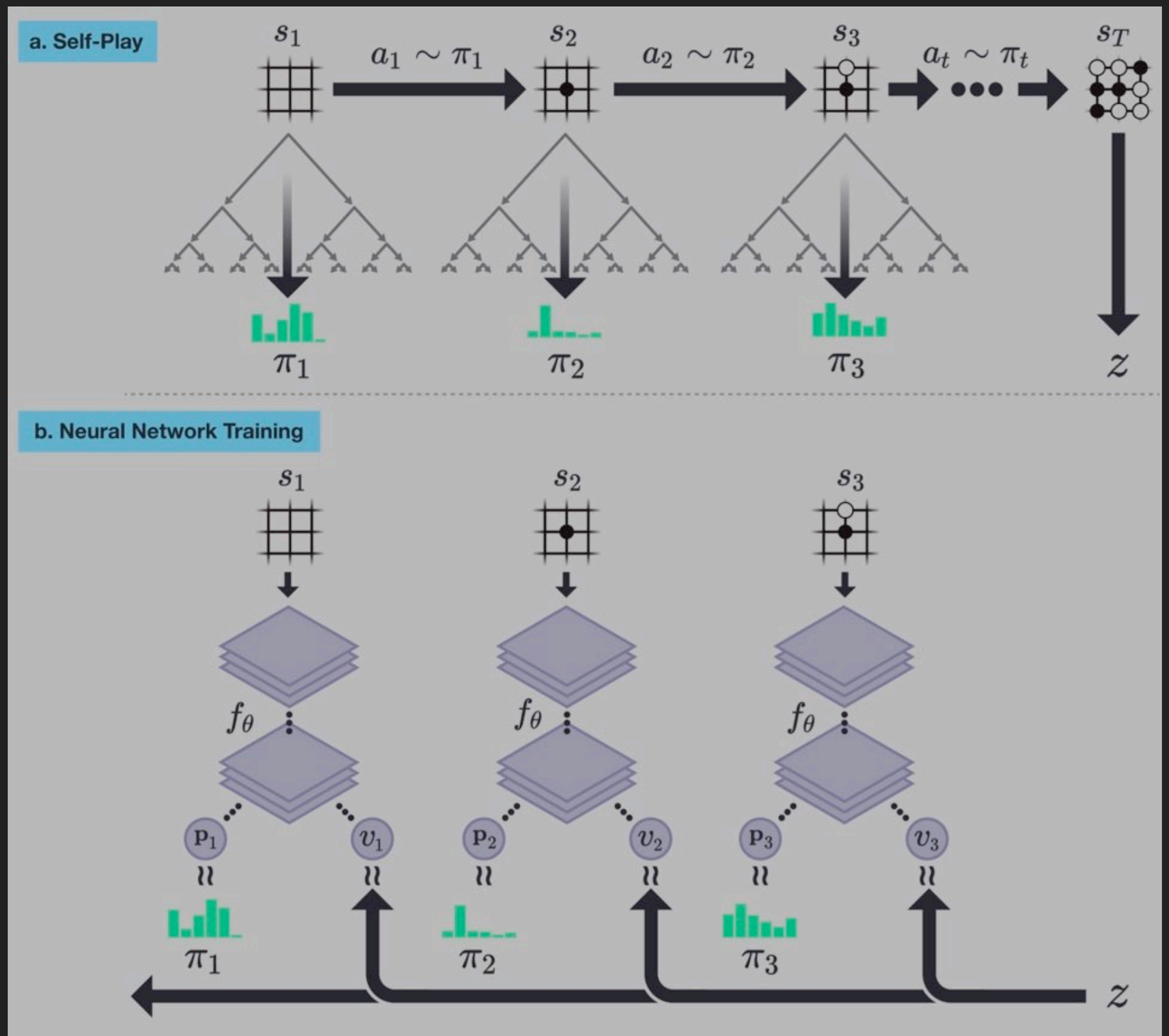
MCTS – PLAY

- Auswahl des besten Zuges:
  $\pi(a \mid s_0) = N(s_0, a)^{1/\tau} / \sum_b N(s_0, b)^{1/\tau}$
- τ ist ein Temperatur-Parameter
- Aufgabe bei Unterschreitung des Grenzwert v_{resign}



Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Fig. 2 d

SELF PLAY REINFORCEMENT LEARNING

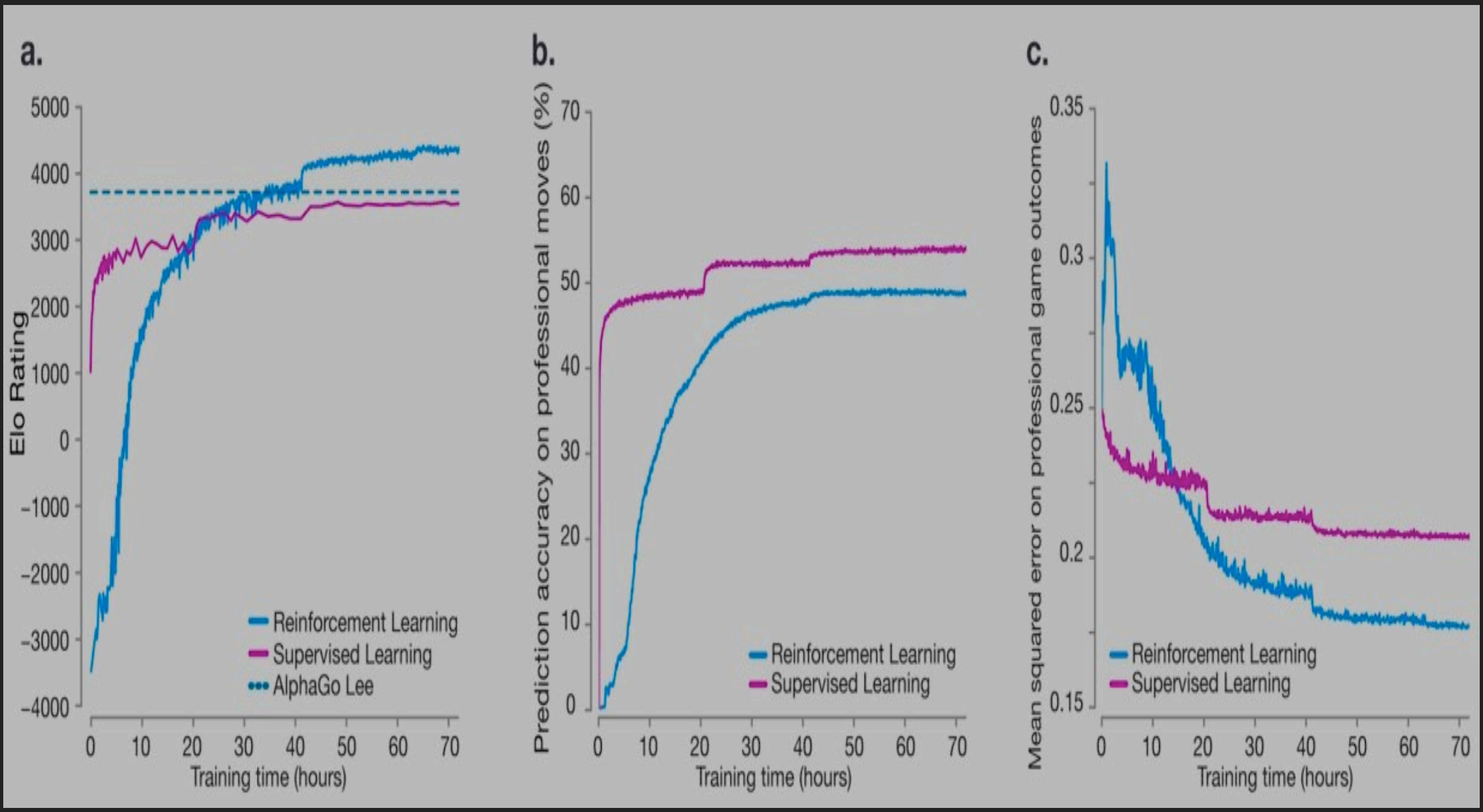


Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Fig. 1

GLIEDERUNG

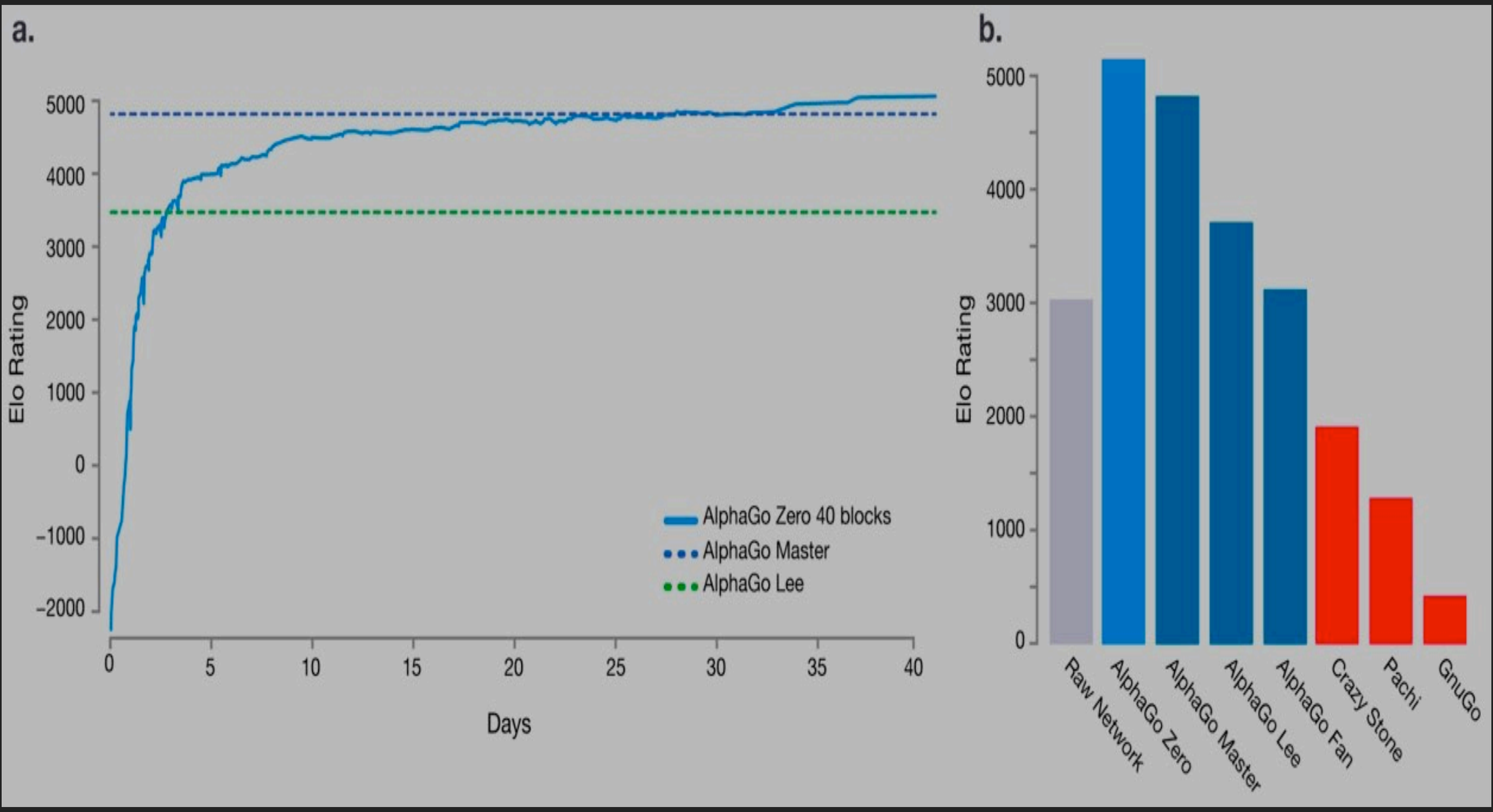
- Ergebnisse und Performance
  - Empirischer Vergleich
  - Performance gegen andere KIs
  - Gelerntes Wissen

EMPIRISCHER VERGLEICH



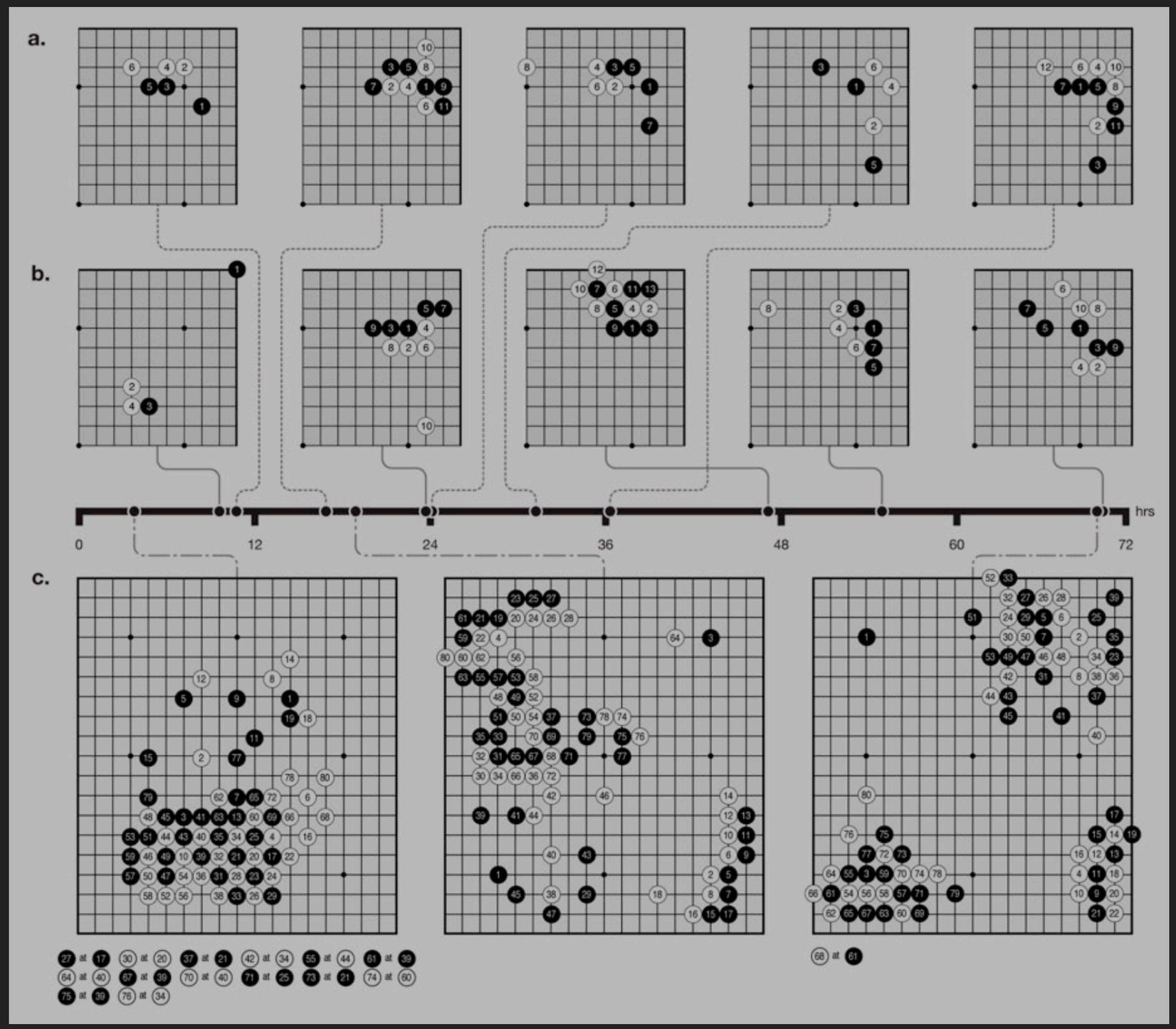
Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Fig. 3

PERFORMANCE GEGEN ANDERE KI



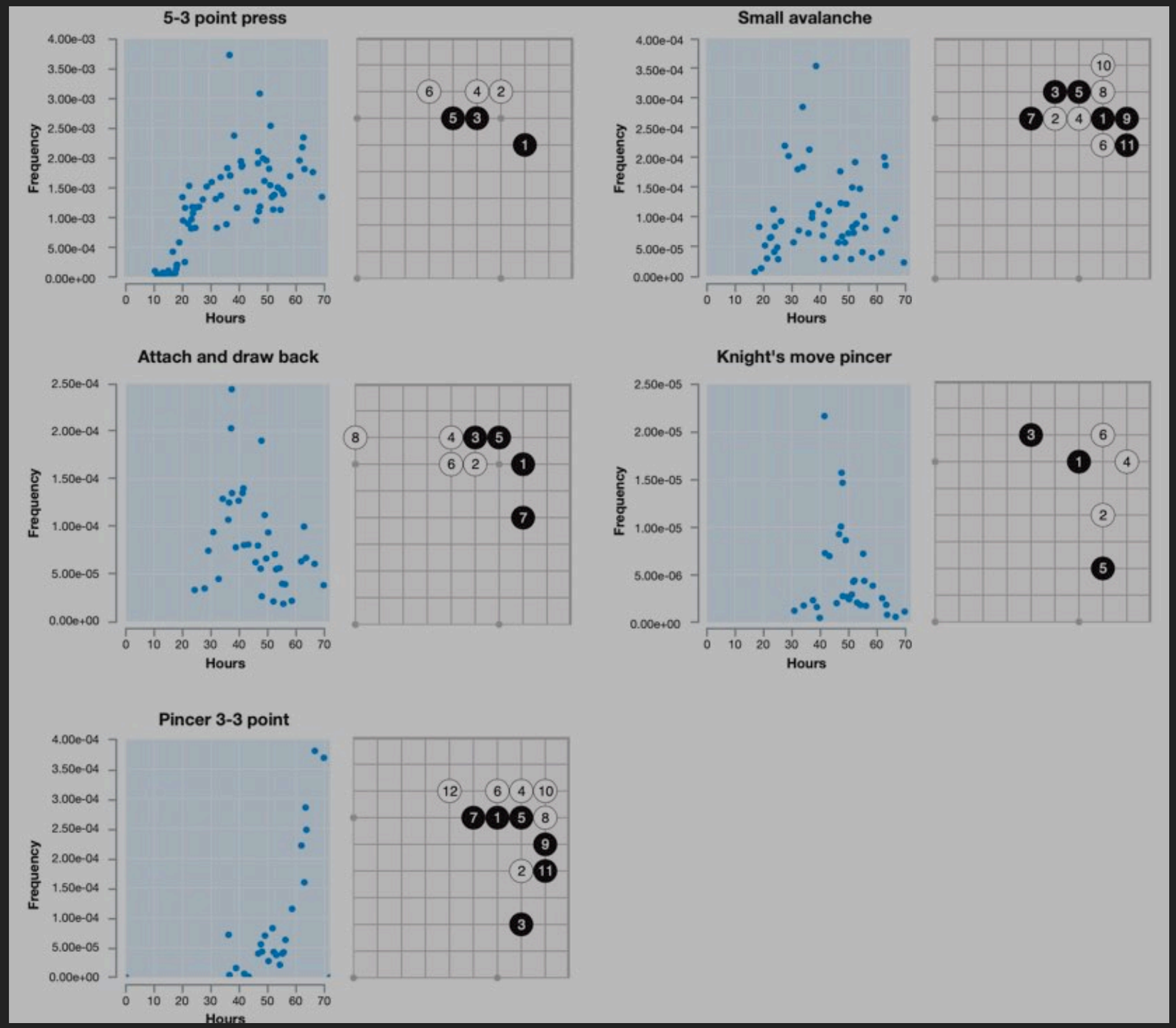
Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Fig. 6

GELERNTES WISSEN



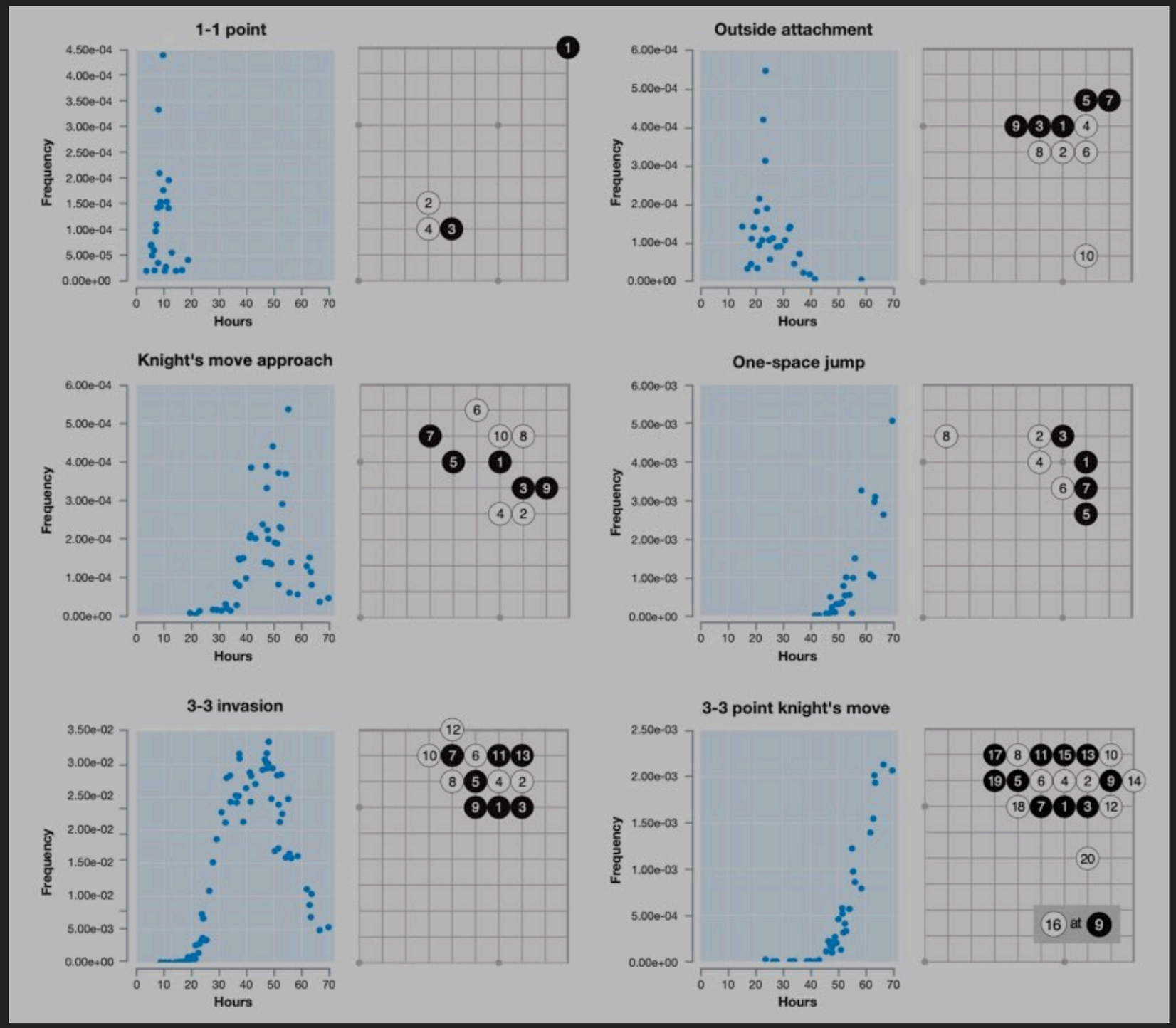
Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Fig. 5

GELERNTES WISSEN



Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Ext. Fig. 1

GELERNTES WISSEN



Quelle: Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354. Ext. Fig. 2

FRAGEN

QUELLEN

- Browne, Cameron B., et al. "A survey of monte carlo tree search methods." IEEE Transactions on Computational Intelligence and AI in games 4.1 (2012): 1-43.
- Silver, David, et al. "Mastering the game of go without human knowledge." *Nature* 550.7676 (2017): 354.
- Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." nature 529.7587 (2016): 484.
- https://www.cs.cmu.edu/~wjh/go/rules/Chinese.html
- http://tromp.github.io/go.html