

将棋の振り飛車の強化学習



山下 宏

2024年11月15日

GPW 箱根

Ver 1.01 (11月16日)

AlphaZero型の学習を振り飛車で

- AlphaZero 2017年
 - 相掛かりや角換わりが好きな居飛車党
- AobaZero 2019年
 - AlphaZeroの追試。ほぼ同じ棋風を確認
- Aoba駒落ち 2021年
 - 平手でなく駒落ちで。6枚落ちでアマ4段に勝つ
- Aoba振り飛車 → 今回のテーマ
 - 将棋のルールだけから振り飛車を学習

振り飛車の定義

- 定義は存在しない
- 11手目から24手目までで飛車が一番多く存在した筋をX筋飛車とする
- 先手後手とも7局面から
- 同数の時は手数短いものを採用
- 0なら居飛車とする
- 人間は6手目に飛車を振ることが多い

一般に使われる下段の用語を使う

9筋 8筋 7筋 6筋 5筋 4筋 3筋 2筋 1筋

9間 向 3間 4間 中 右4間 袖 居 1間

	9筋	8筋	7筋	6筋	5筋	4筋	3筋	2筋	1筋	
なし	香	桂	金				桂	香		一
駒		銀	王	金	銀		飛			二
持		歩	歩	歩	歩	歩	角			三
手	歩						歩			四
後		歩						歩		五
手	歩		歩	角			歩			六
口				歩	歩	歩	銀	歩	歩	七
			飛	銀	金		金			八
	香	桂			玉			桂	香	九

▲ 先手
持駒
なし

- 表は先手から見た場合
- 先手は3間飛車
- 後手は向飛車

そのままだと勝率が悪いので振らない

- 希望した筋に振れた場合にボーナスを

$$(\text{実際の勝敗}) \times (1 - a) + (\text{希望した筋に振れた}) \times a$$

- $a=0.3$

- 希望する筋に振れたら0.3勝とする
- この値を大きくすると負けても飛車を振る
- 1筋から9筋まですべての筋に振る
 - 先手、後手ともに振る。相振りも
 - NNの入力に自分の振る筋を与える。相手はなし

実際は24手以降はボーナスを下げる

$$\frac{\text{実際の勝敗} + \text{探索勝率}}{2} \times (1 - a) + (\text{希望した筋に振れた}) \times a$$

$$a = b - \frac{b}{c - 24} \times (\text{手数} - 24)$$

$$b = \text{ボーナス}, c = 40$$

- 24手目から徐々に下がり40手目で消える
 - 24手目までは $a=b$
 - MCTSの探索が深いと24手以上読むため

最終的なボーナスの値

9間	向	3間	4間	中	右4	袖	居	1間
0.2	0.18	0.14	0.15	0.13	0.15	0.13	0.09	0.14

- 最初はすべて0.3
- 実現率が9割を超えない範囲で手作業で調整
- 9間飛車や向飛車は実現率が低く、振らせるのが難しい
- 小さい値が好ましい
- 先手、後手とも同じ値

学習過程で出現した棋譜紹介

- 1手800playoutで乱数要素なしの自己対戦
 - 学習用でなく特別に生成
 - 同一の重みを使用
 - 振る位置を先手、後手とも1つだけ指定
- floodgateで3250程度
 - 藤井聡太7冠に近い強さ
- 1720万棋譜の学習から
- 棋譜は下で公開
 - http://www.yss-aya.com/furibisha/no_noise/sample12.html

ここで▲76歩が間に合う。3間飛車

▲3間
△居飛車

	9	8	7	6	5	4	3	2	1	
	香	桂	銀	金	王	金	銀	桂	香	一
		飛						角		二
	歩		歩	歩	歩	歩		歩		三
							歩		歩	四
		歩								五
し									歩	六
な	歩	歩	歩	歩	歩	歩	歩	歩		七
▽		角	飛				玉			八
	香	桂	銀	金		金	銀	桂	香	九

▲なし

初?

w0816

822万棋譜

(勝又さん指摘)

2008年に羽生さんが指していた

後手 羽生善治

9	8	7	6	5	4	3	2	1	
香	桂	銀	金		金	銀	桂	香	一
			王			飛	角		二
	歩	歩	歩	歩	歩		歩	歩	三
歩						歩			四
							歩		五
歩		歩							六
	歩		歩	歩	歩	歩		歩	七
	角						飛		八
香	桂	銀	金	玉	金	銀	桂	香	九

先手 深浦康市

2008-07-23 王位戦 深浦康市 vs. 羽生善治

<https://shogidb2.com/games/c0a550eb94974f2d9de8977ced09b15629f8ad7>

3間飛車は▲79金型を好んで指す

▲3間
△居飛車

	9	8	7	6	5	4	3	2	1	
	香	桂		金		金	銀	桂	香	一
			銀			王				二
	歩		歩	歩	歩	歩		歩		三
							歩		歩	四
										五
歩		飛	歩						歩	六
歩		桂	歩	歩	歩	歩	歩	歩		七
		飛	銀				玉			八
香		金				金	銀	桂	香	九

▲角歩

初?

w1687

1693万棋譜

(勝又さん指摘)

2024年に7月に古森5段が指していた



先手 阿部健治郎 七段
歩1

14手目 3一金(41)

論文の図12(5段4列)の
w1619, 1580万棋譜と
完全に同一局面

2024-07-09
阿部健治郎 vs.
古森悠太
順位戦C1組2回戦

ゴキゲン中飛車に超速37銀を再発見

▲居飛車
△中

	9	8	7	6	5	4	3	2	1	
	▲	桂		▲		▲		▲	▲	一
			王	銀		角				二
		歩	歩	歩		歩		歩	歩	三
	歩					銀	歩			四
						桂		歩		五
	歩		歩	銀	▲	銀	歩			六
歩		歩	歩			歩			歩	七
		角	玉					飛		八
	香	桂		金		金			香	九

▲歩

歩
▽

w1011
1017万棋譜

5筋を突かずに向飛車で反撃

▲居飛車
△中

	9	8	7	6	5	4	3	2	1	
	▲	桂	銀	金	飛				▲	一
		王				銀	金			二
		歩	歩	歩	歩	歩	桂		歩	三
	歩							飛		四
							歩		角	五
一	歩		歩							六
二		歩	桂	歩	歩	歩			歩	七
歩			玉		金	銀				八
△	香		銀	金				桂	香	九

▲角歩

初?

w1497

1503万棋譜

パックマンを△向飛車の変化で

▲居飛車
△向

	9	8	7	6	5	4	3	2	1	
	香	桂	銀	金	王			桂	香	一
						金	銀	角		二
	歩	歩	歩	歩				歩	歩	三
							歩			四
角										五
し		飛	歩				歩			六
な	歩	歩		歩	歩	歩		歩	歩	七
▽		銀						飛		八
	香	桂		金	玉	金	銀	桂	香	九

▲歩二

w0566
572万棋譜

ダイレクト向飛車の定跡を再発見

▲居飛車
△向

	9	8	7	6	5	4	3	2	1	
	香	桂	銀		王			桂	香	一
					金			飛		二 ▲
		歩	歩	歩	歩		銀	歩	歩	三 歩
	歩		角				歩			四
			金							五
	歩		歩							六
角		歩		歩	歩	歩	歩	歩	歩	七
△		銀		玉	金	銀		飛		八
	香	桂		金				桂	香	九

w1234
1240万棋譜

自陣角で△26歩を狙う

▲居飛車
△向

	9	8	7	6	5	4	3	2	1	
	香	桂		金				桂	香	一
		王	銀	角			金	飛		二
		歩	歩	歩	歩	歩	銀			三
歩							歩		歩	四
								歩		五
歩		歩			歩	銀			歩	六
	歩	銀	歩	歩		歩				七
歩		玉		金				飛		八
	香	桂		金				桂	香	九

▲
角

初?

w1619

1625万棋譜

▲向飛車は▲47銀から優雅に振る

▲向
△居飛車

	9	8	7	6	5	4	3	2	1	
	香	桂		金		金		桂	香	一
		飛	銀				王	銀		二
	歩		歩	歩	歩	歩		歩		三
		歩					歩		歩	四
										五
			歩		歩	歩			歩	六
毎	歩	歩		歩	歩	銀		歩		七
▽		飛								八
	香	桂	銀	金	玉	金		桂	香	九

▲角

w1563
1569万棋譜

天守閣美濃から4枚美濃に

▲居飛車
△4間

	9	8	7	6	5	4	3	2	1	
	香	桂		金				桂	香	一
		王	銀					飛		二
		歩		金		銀	角	歩	歩	三
	歩		歩	歩		歩	歩			四
					歩			歩		五
△	歩	歩	歩		歩		歩			六
△		玉	銀	歩		歩	桂		歩	七
△			銀		金			飛		八
	香	桂	角	金					香	九

▲なし

w1234
1240万棋譜

(勝又さん指摘)

1998年に7月に勝又さんの類似局

後手 日浦市郎

9	8	7	6	5	4	3	2	1
香	桂		金				桂	香
	王	銀					飛	
	歩		金		銀	角	歩	歩
歩		歩	歩		歩	歩		
				歩			歩	
歩	歩	歩	歩	歩		歩		
	玉	銀			歩			歩
		銀		金			飛	
香	桂	角	金				桂	香

先手 勝又清和

評価値 -

36手目 5五歩(54)

<https://shogidb2.com/games/2f9fe53231fd452989bc9b300c68d5ae9cc79fea>

▲37桂でなく▲66歩、だけが異なる。

1998-04-22
勝又清和 vs.
日浦一郎
王将戦

さらに松尾流穴熊に。ガチガチに固く

▲居飛車
△4間

	9	8	7	6	5	4	3	2	1	
	▲							桂	▲	一
		王	金		金	飛				二
		銀	桂	歩			角	歩	歩	三
	歩	歩	歩			銀	歩	歩		四
					歩	歩				五
し	歩		歩	歩	歩		歩			六
な		歩	角	金		歩			歩	七
▽	香	銀	金					飛		八
	玉	桂	銀					桂	香	九

▲なし

w1355
1361万棋譜

藤井システムに近い筋で反撃

▲4間
△居飛車

	9	8	7	6	5	4	3	2	1	
	香	桂				金	銀	桂		一
		飛						王	香	二
	歩			歩	銀	金	角	歩		三
			歩		歩	歩	歩		歩	四
		歩		歩				桂		五
┌	歩		歩			歩	歩		歩	六
└		歩	角	銀	歩			歩		七
▽				飛	金		銀			八
	香	桂				金	玉		香	九

▲なし

w1730
1736万棋譜

9間はここで△22飛と戻す。メリットなし

▲居飛車
△9間

	9	8	7	6	5	4	3	2	1	
	香	桂		金				桂	香	一
		王	銀						飛	二
				金		銀	角	歩	歩	三
	歩	歩	歩	歩	歩	歩	歩			四
								歩		五
┌	歩		歩	歩	歩		歩			六
└		歩	角	金	銀	歩			歩	七
▽	香		金					飛		八
	玉	桂	銀					桂	香	九

▲なし

w1498
1504万棋譜

1間は端を突き越せれば浮き飛車に

▲ 1間
△ 居飛車

	9	8	7	6	5	4	3	2	1	
	香	飛					銀	桂	香	一
				金	玉	金	角			二
			桂	銀	歩	歩	歩	歩	歩	三
	歩		歩	歩						四
						銀	歩		歩	五
一	歩								飛	六
二		歩		歩	歩	歩	桂	歩		七
三		角	金	銀		玉	金			八
△	香	桂							香	九

▲ なし

w0607
613万棋譜

相振りには2筋、8筋を攻める展開が多い

▲ 3間
△ 3間

	9	8	7	6	5	4	3	2	1	
	香	桂			王			桂	香	一
		銀	金		金	銀	飛	角		二
		歩	歩	歩	歩	歩				三
	歩						歩		歩	四
		歩						歩		五
し	歩		歩						歩	六
ひ				歩	歩	歩	歩	歩		七
▽		角	飛	銀	金	玉		銀		八
	香	桂				金		桂	香	九

▲ なし

w1425
1431万棋譜

どこに飛車を振っても勝率5割に調整

- 先手9個所、後手9個所で $9 \times 9 = 81$ 個の組み合わせ全てで
- これが必要だったかは不明
- 調整なしだと？
 - 9間飛車はハンデレートが200 ELOあり、ボーナスが0.2でなく、0.4ぐらいになると思われる

強さの調整の仕方

- Root ノードの着手選択確率の温度で調整
 - AlphaZeroで30手目までの乱数性を作る手法
- 温度1でレート差450、温度0でレート差0
 - 温度が高いと最善でない手を選びやすい
- ▲3間△向飛車の直近1000局の勝率が0.53 (+20 ELO)なら半分の+10 ELOをハンデレートに足す

ハンデレーティング

後手の飛車の筋。後手が居飛車。先手が居飛車

	9筋	8筋	7筋	6筋	5筋	4筋	3筋	2筋	1筋
先手9筋	24	-99	-39	-23	-1	-7	-32	-48	19
先手8筋	72	-94	39	51	74	62	42	16	107
先手7筋	97	-84	45	51	67	48	29	0	72
先手6筋	65	-97	-19	-2	40	8	11	-8	62
先手5筋	84	-119	-9	32	24	-3	-23	-23	26
先手4筋	71	-138	-16	13	4	23	-4	3	55
先手3筋	113	-67	16	75	54	76	4	35	88
先手2筋	195	23	132	183	168	136	145	144	152
先手1筋	23	-159	-74	-17	-16	-14	-51	-34	19

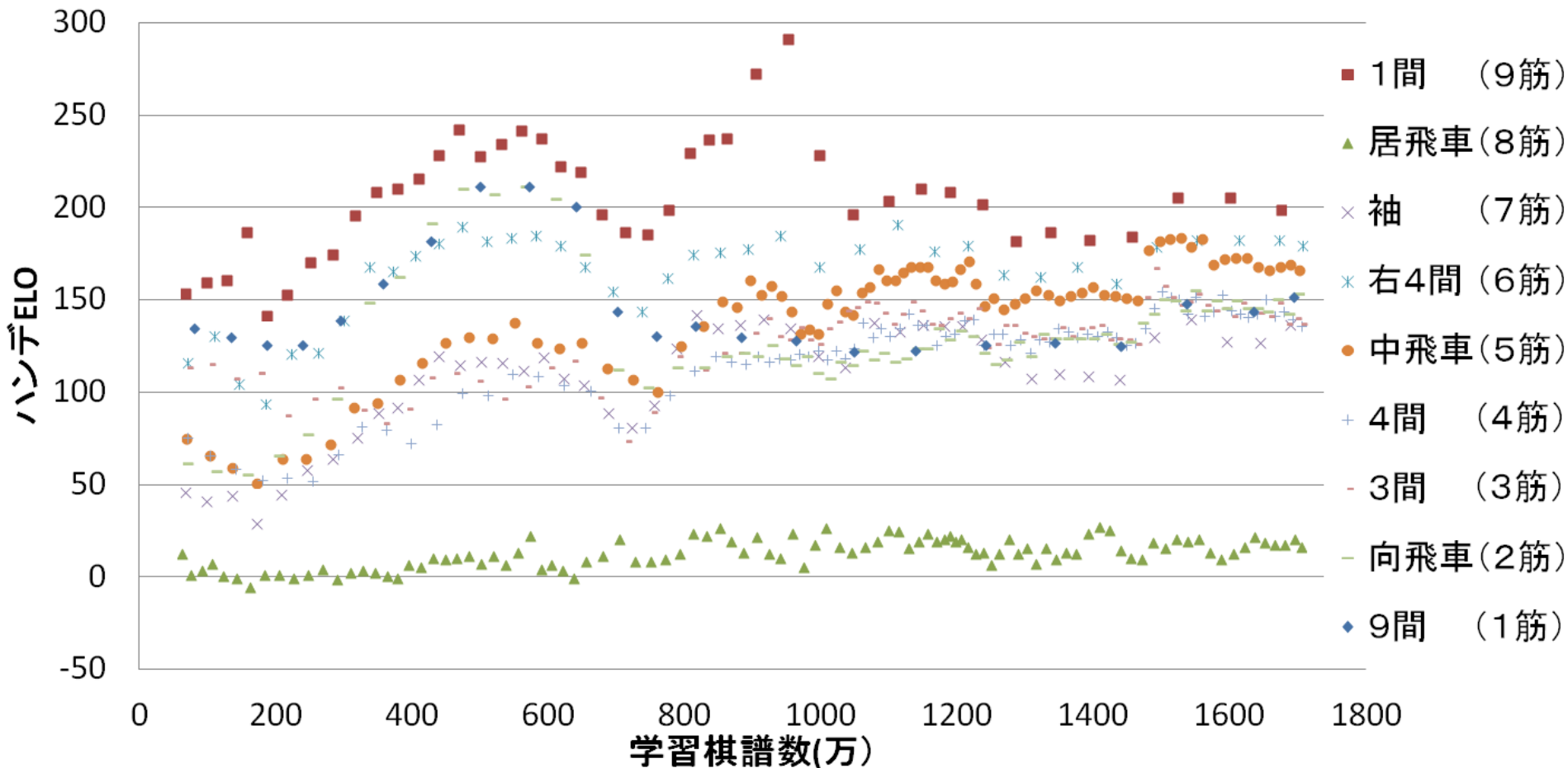
先手が居飛車で後手が中飛車のハンデは+168

飛車を振ることの損(居飛車と比べて)

	1間	居	袖	右4	中	4間	3間	向	9間
後手が居飛車	182	---	90	161	142	120	107	117	122
先手が居飛車	172	---	109	160	145	113	122	121	129

- 値が大きいほど損
- 先手も後手も振ることの損は同じくらい
- 対抗形では中飛車が一番不利
 - 3間、4間、向は大差なし
 - 20 ELO程度の差だが600万棋譜連続で差が安定
- 1間が一番損で、右4間が2番目に損

ハンデレーティングの推移



先手が居飛車の時の後手の位置のみ

振る筋の割合は人間が指定

- 対抗形、相振りを含め81種類
 - 3間または4間、など複数指定も
- 対抗形を多めに調整

生成される振り飛車の割合

	9筋	8筋	7筋	6筋	5筋	4筋	3筋	2筋	1筋
先手9筋	1638	3224	1666	2157	2067	1896	2150	1841	2338
先手8筋	1936	22409	1788	2406	4407	4996	7186	5892	2090
先手7筋	3433	28309	4010	3391	7657	6646	7619	8042	2640
先手6筋	3914	25275	3955	4181	6590	6102	6868	5649	2292
先手5筋	3254	21230	4036	2384	6568	6347	7409	5071	2339
先手4筋	1951	5693	1936	1496	3489	5201	4106	2197	1971
先手3筋	2358	7558	3999	1922	4800	4747	5113	1586	1529
先手2筋	5021	24312	7625	5868	22331	27161	27172	22745	3712
先手1筋	1885	4934	2257	1796	2954	3722	3758	1989	1808

直近の50万棋譜で。対抗形が4割を占める。

▲居飛車△四間飛車は27161局。勝率はすべてほぼ5割。

振る位置を指定したときの成功率

	9筋	8筋	7筋	6筋	5筋	4筋	3筋	2筋	1筋	合計
成功率	0.526	0.636	0.857	0.836	0.877	0.735	0.898	0.883	0.833	0.802
成功数	17174	45427	59275	56487	47654	19261	19924	104884	12752	382838
実現数	18772	51726	68029	62973	56404	26576	31877	137871	23412	477640
希望数	32650	71469	69160	67608	54329	26203	22182	118722	15317	477640

先手が振る場合のみ。単独指定


9間(9筋)、向飛車(8筋)は自分の角を動かす必要があり実現が難しい

振る位置の指定割合

	9筋	8筋	7筋	6筋	5筋	4筋	3筋	2筋	1筋
先手9筋	25	50	50	50	50	50	50	50	100
先手8筋	25	400	50	50	100	100	150	100	50
先手7筋	12	450	12	25	100	100	100	150	50
先手6筋	25	450	25	25	100	100	100	100	50
先手5筋	25	300	25	25	50	100	100	100	50
先手4筋	25	50	25	25	50	50	50	50	50
先手3筋	25	12	25	25	50	50	25	50	50
先手2筋	25	25	12	50	300	450	450	400	50
先手1筋	25	25	25	25	25	25	25	25	25
先手複数	25	50	50	50	50	25	25	25	12
後手複数	12	25	25	25	50	50	50	50	50
		2か所	3か所	4か所	5か所	6か所	7か所	8か所	9か所
複数指定		100	50	25	25	25	25	25	200

実際に実現した割合を考慮して手作業で調整

4つの失敗



自己対戦で先手後手を独立に探索

- 同じ重みで▲居△中、▲居△ 4 間で初手が変わる現象が
 - NNの入力は自分の希望する筋だけなのに
 - 自分は居飛車、相手は4間で探索させると読み筋、評価で相手の狙いがばれる
- Root手番が先手の時、後手は「どこに振ってもOK」に変更で解決
 - 9箇所全部を指定。ハッシュ表も分離。
- 相手モデルが大事だった

投了を常に禁止に

- AlphaZeroは10%の棋譜で投了を禁止
 - 投了の間違いが5%以下になるように投了の閾値を自動調整
- ▲9 間△1間のハンデレートが-352と極端に下がる現象が発生
 - この組み合わせは生成確率が0.5%なので10%の禁止だと検出できなかった
- AlphaZero形式の投了だと出現確率が低い局面が不当に低く評価されても修正されない

「希望する筋に振れたか」のNNの出力は役に立たなかった

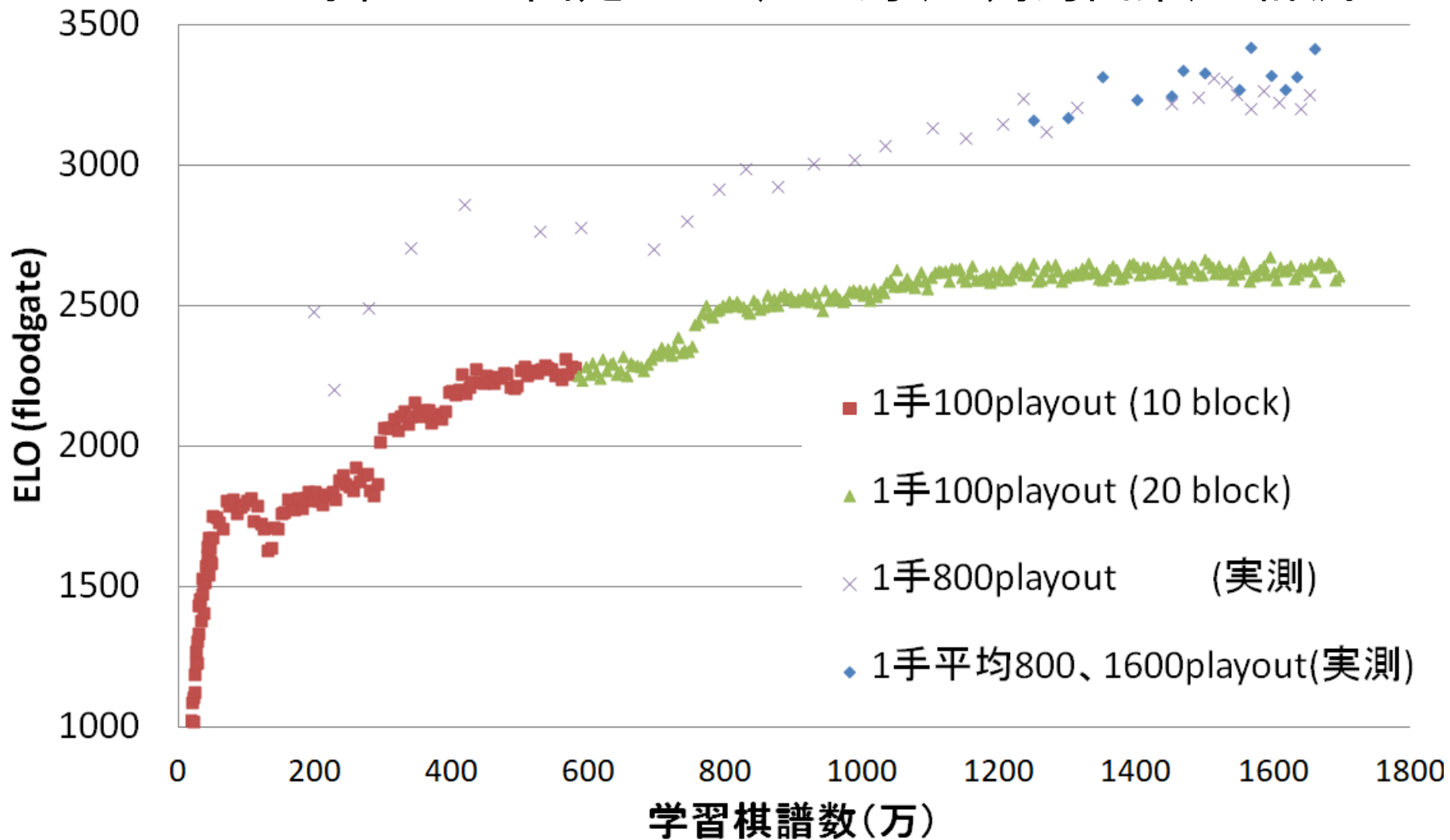
- Policy, Valueのように「希望する筋に振れたか」という出力を試した
 - まったく着手に影響を与えなかった
- Valueに直接影響を与えるボーナスに変更

個別ボーナスは「希望した筋」の平均で

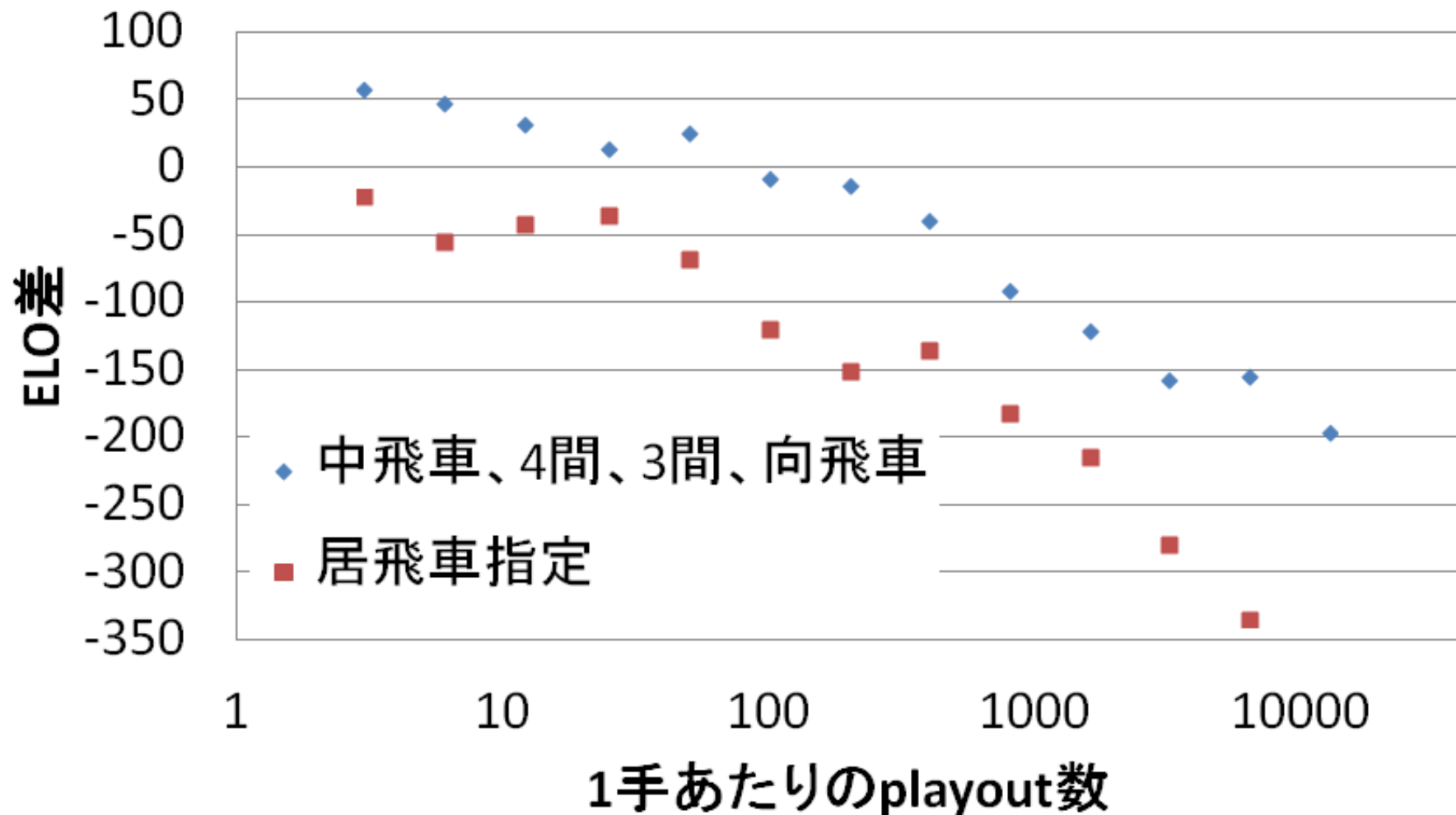
- 当初は1筋から9筋まで「希望した筋に振れた」場合に同一のボーナスを与えていた
- 個別の筋に対して「希望した筋に振れた」場合に、「実際に振った筋」のボーナスだと？
 - 成功率が大きく下がった
- 「希望した筋」の平均のボーナスで解決
 - 複数指定があるので

棋力の推移

水匠5との固定ノード、400局(互角局面集)で計測



1手のplayout数を固定した時のAobaZeroとの比較



Playout数を倍にした時の伸びがかなり悪い

感想など



- 対抗形は向飛車に振り直して飛車先を逆襲する展開が多い
- 振り飛車が穴熊に組む変化は出現しなかった
 - 理由は不明
- 振り飛車の棋譜は変化に富んで並べていて面白い

まとめ



- ゼロからの強化学習を振り飛車で行った
- 3間飛車で初と思われる指し方を見つけた
- ゴキゲン中飛車が超速37銀で衰退する過程を再現できた
- 対抗形では中飛車が一番不利
- 1間、9間飛車はそれを活かした指し方なし
- 80日間で1720万棋譜を作成

以下は補足資料



初手

	9間	向	3間	4間	中	右4	袖	居	1間
先手	▲76歩	▲76歩	▲78飛	▲76歩	▲58飛	▲76歩	▲38飛	▲76歩	▲16歩
後手	△34歩	△34歩	△94歩	△34歩	△34歩	△34歩	△94歩	△34歩	△94歩

(後手は1手目が▲76歩の場合)

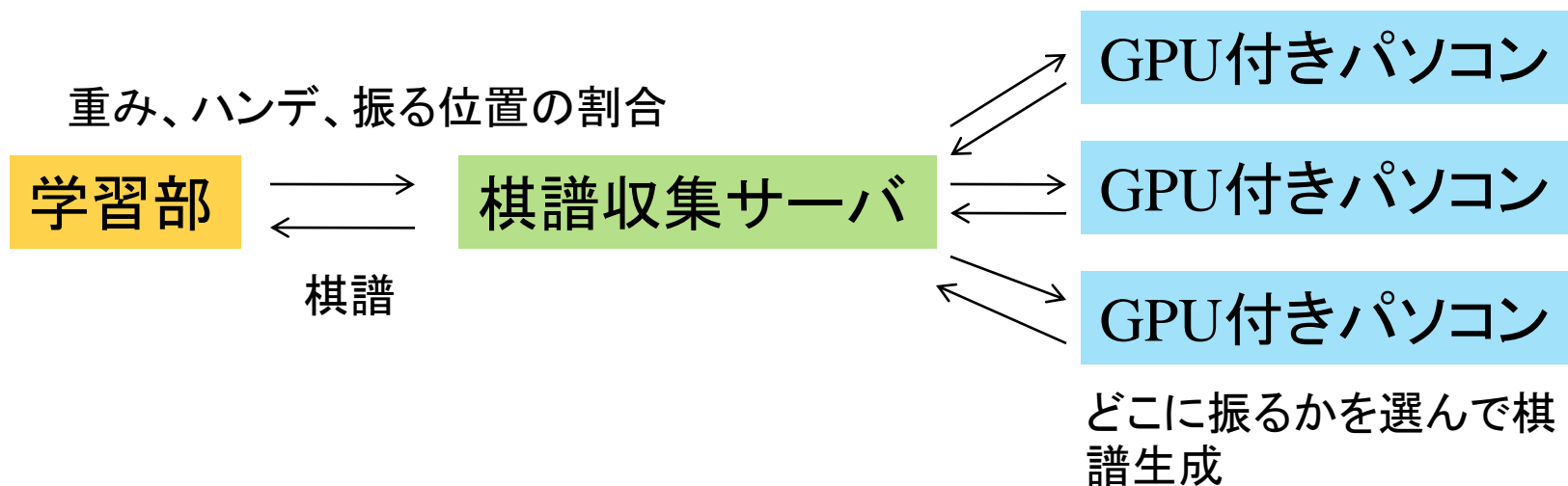
- 先手だと3間、袖はいきなりその位置に振る
- 後手は1手目が▲76歩の場合
- 1650万棋譜の時点で

計算資源



- Vast.aiを主に利用
 - 個人のGPUを貸し出すサービス
 - EC2のスポットインスタンスに似た形式がお得
- RTX 4090が1枚だと493日かかる
- 一般の方からの協力も
 - ありがとうございます

世界中から計算資源を募集



- 囲碁のLeelaZeroやチェスのLCZeroと同様
- オープンソース
 - AobaZeroを改造。BonanzaやLeelaZeroも利用
- 「Aoba振り飛車」で検索するとサイトがあります

ニューラルネットワークの入力

■ 駒の配置

- 過去の局面はなし

■ 同一局面になった回数

■ 手数、手番

■ 王手か

■ 駒の利きの数

■ 駒が動ける場所

■ 振る筋(9種類)

- 複数指定可能

将棋の盤面を「画像」として認識させている。
画像の大きさは9x9

常に下側が次の手番
(1手ごとに反転)

赤字はAlphaZeroにはない独自の特徴

→ 相手がどこに振るかは与えない

ニューラルネットワークの出力

- 次の指し手 (2187通り)
 - dlshogiと同じ。
 - あるマス目にどの方向から駒が移動するか
 - $(9 \times 9 \times 27) = 2187$ 通り
 - AlphaZeroは11259通り $(9 \times 9 \times 139)$
- 勝率 (1通り)

ニューラルネットワークの構造

■ ResNet

- 20ブロック
- 3x3のフィルターが256個
- パラメータの数は2340万個
- 基本はAlphaZeroと同じ
- 588万棋譜までは10ブロック、128個