

Mixing Expert Knowledge: Bring Human Thoughts Back To the Game of Go

Yichuan Ma

2025.10.19

Contents

1. Motivation
2. Method
3. Experiment
4. Conclusion



Part 1 | Motivation



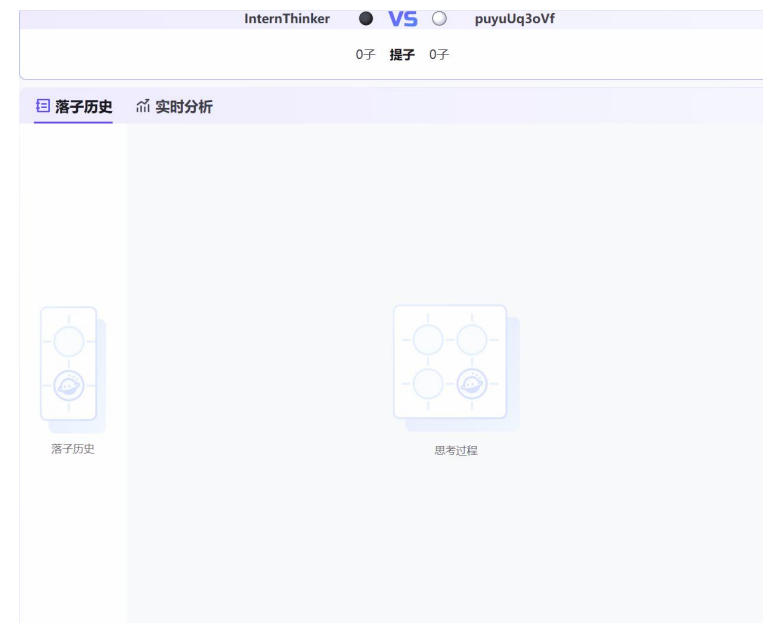
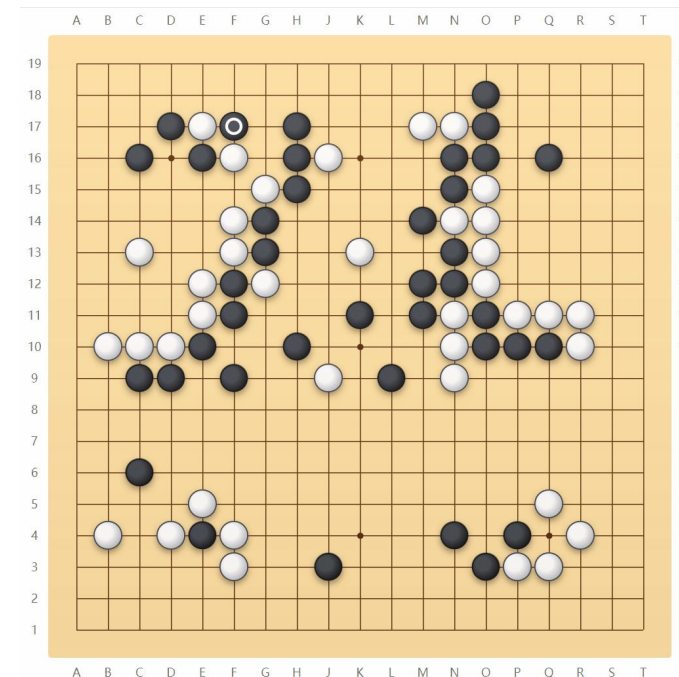
Motivation

General Reasoning LLMs

- Strong performance in **GENERAL** tasks(math, code)
- Lack data in specialized domains

Expert AI models

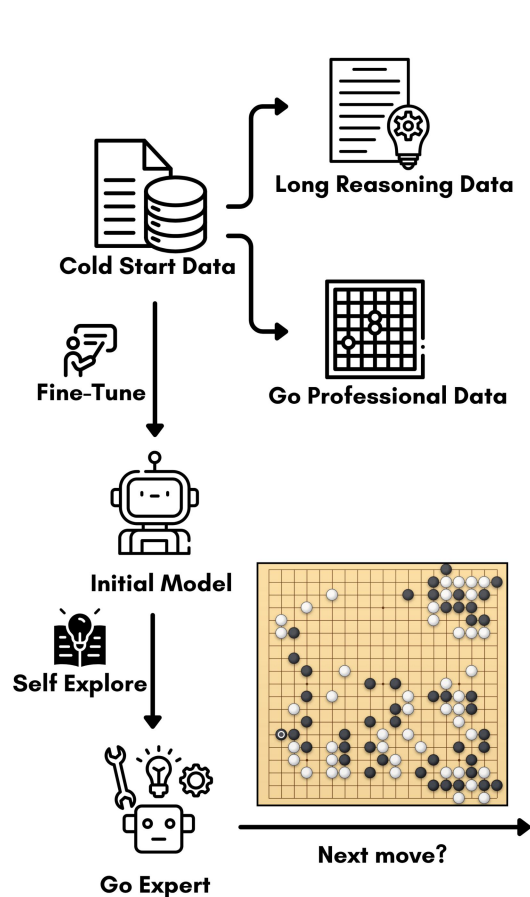
- Like AlphaGo for Go
- Expert level performance in **SPECIALIZED** tasks.
- unexplainable, no general ability(chat, explain, reason in natural language)



Part 2 | Method



Method



Query

You are a professional Go player. Your task is to analyze the board position based on the given game record, suggest several potential next moves with analysis, explore possible variations that could follow, provide thoughtful reasoning, and finally summarize by recommending the best next move. In the given game, "X" represents black stones and "O" represents white stones.

1.X-Q16 2.O-D16 3.X-R4 4.O-D3 5.X-D5 6.O-C5 7.X-C6 8.O-C4 9.X-D7 10.O-F3 11.X-C14 12.O-F17 13.X-D11 14.O-P3 15.X-Q5 16.O-R11 17.X-R9 18.O-R14 19.X-R15 20.O-Q14 21.X-O17 22.O-P10 23.X-O4 24.O-R17 25.X-S15 26.O-P18 27.X-Q17 28.O-Q18 29.X-S17 30.O-S18 31.X-R16 32.O-P17 33.X-P16 34.O-O16 35.X-O18 36.O-R18 37.X-T18 38.O-O15 39.X-M17 40.O-S14 41.X-P19 42.O-B15 43.X-N3 44.O-B14 45.X-C12 46.O-S3 47.X-R3 48.O-R2 49.X-Q2 50.O-R6 51.X-P9 52.O-Q7 53.X-S5 54.O-P8 55.X-O9 56.O-Q9 57.X-R8 58.O-Q8 59.X-S6 60.O-S1 61.X-T2 62.O-Q3 63.X-S2 64.O-Q1 65.X-R1 66.O-F9 67.X-D9 68.O-R2 69.X-O2 70.O-L3 71.X-P2 72.O-M9 73.X-J3 74.O-K5 75.X-G4 76.O-F4 77.X-G6 78.O-F5 79.X-G5 80.O-G3 81.X-K4 82.O-L4 83.X-J5 84.O-K6 85.X-J7 86.O-M6 87.X-J10 88.O-G11 89.X-L8 90.O-M8 91.X-L7 92.O-N5 93.X-L10 94.O-C8

Please predict and analyze the next move.

Response

First, I need to analyze the key features of the current position. White has just played at C8, which is clearly a move targeting Black's left-side formation. White's intention with this move is quite clear: to weaken Black's influence on the left side while potentially preparing for a future invasion into Black's territory. Let me consider several possible responses:

Firstly, Black playing at **B6** is a direct response, aimed at protecting the left-side space and preventing further White invasion. B6 can connect with Black's previous stone at C6, forming a relatively solid structure. If White chooses to counter at D8, Black can continue strengthening their position with E8, creating a powerful shape. Such a formation not only resists potential White invasions but also helps Black establish more influence on the left side.

B5 is a more direct choice, more directly confronting White's stones at C4 and C5. This approach might lead to more intense fighting, but could also help Black establish more structure on the left side. If White counters at D8, Black can use E8 to continue building their shape.

B7 represents a middle ground between B6 and B5, potentially offering a different balance point. However, if White counters at D8, Black might need additional steps to build and consolidate their structure.

Through calculation and comparison of the possible outcomes of these options, **B6** appears to be the most reasonable choice. It not only directly addresses White's threat and protects Black's space on the left side, but also helps Black establish more structure. According to my calculations, choosing B6 might give Black approximately a 51.8% win rate, which is a fairly good advantage in the current position.

Additionally, B6 has the benefit of simultaneously addressing multiple directions, increasing Black's overall flexibility. If White counters at D8, Black can continue building their shape with E8, then further adjust based on White's response. This flexibility is particularly valuable in the current position.

Pipeline Overview

Method: Modeling Go

1. Direct Modeling (Sec 2.1):

- Each move position is a combination of a letter and a number coordinate.

Given the Go game move list below:
1.X-Q16 2.O-D16 ... 141.X-J7
Please predict the next move.

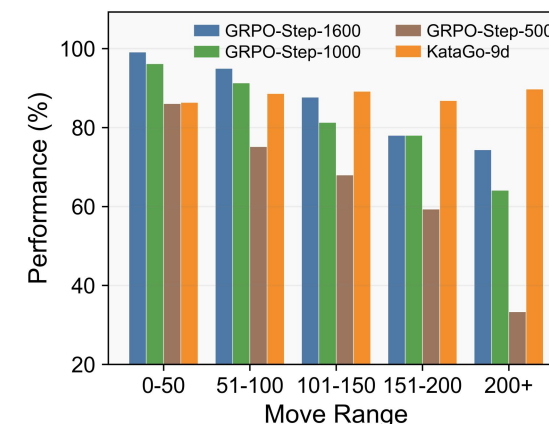
1.1 Direct Modeling

2. 2-D board rendering (Sec 4.3):

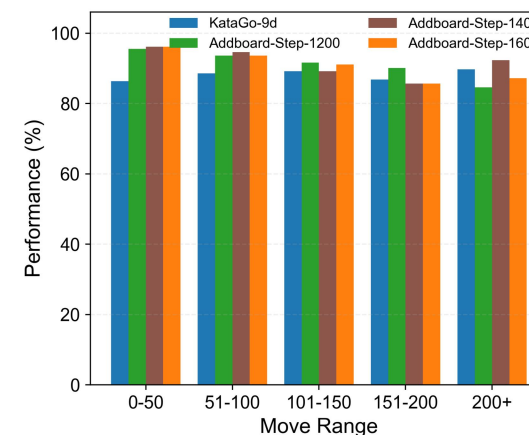
- This state is represented as a 19×19 2-D array, where the values 1, -1, and 0 denote a black stone, a white stone, and an empty intersection, respectively.

[illegible]

2.1 2-D Rendering



1.2 Context Curse



2.2 Better Performance

Part 3 | Experiments



Experiments

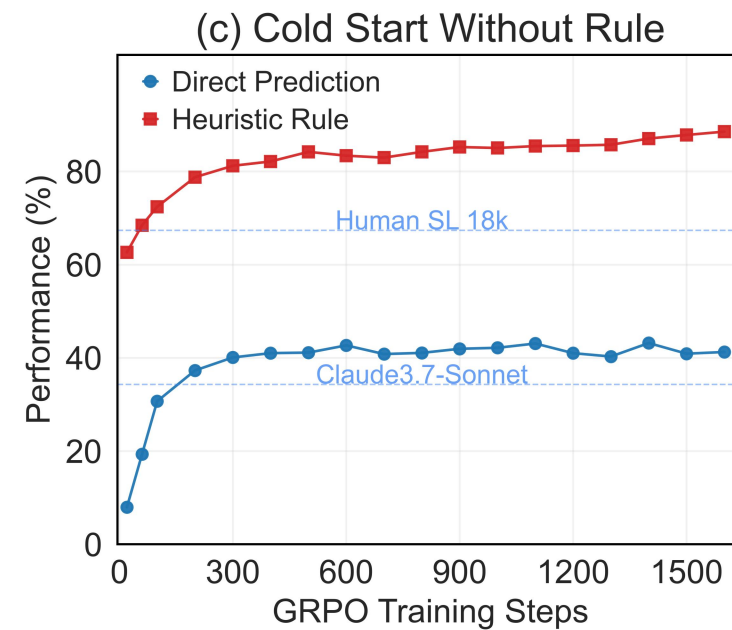
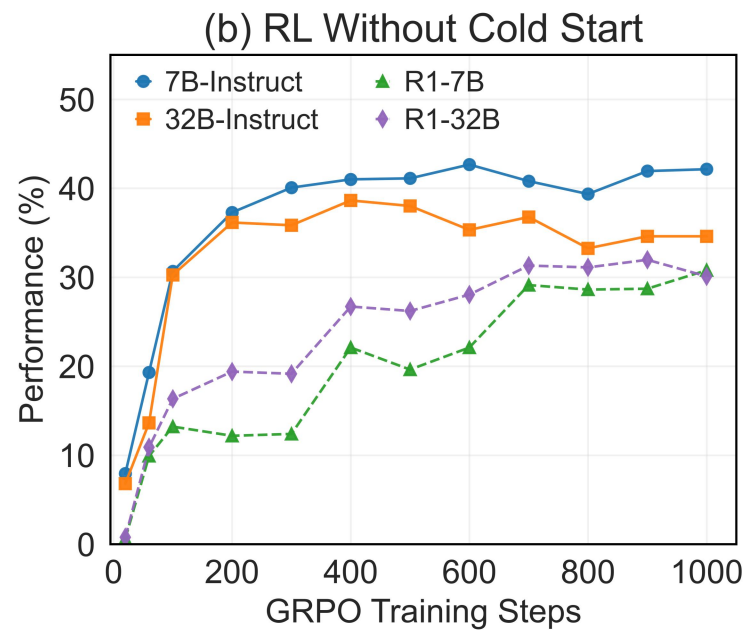
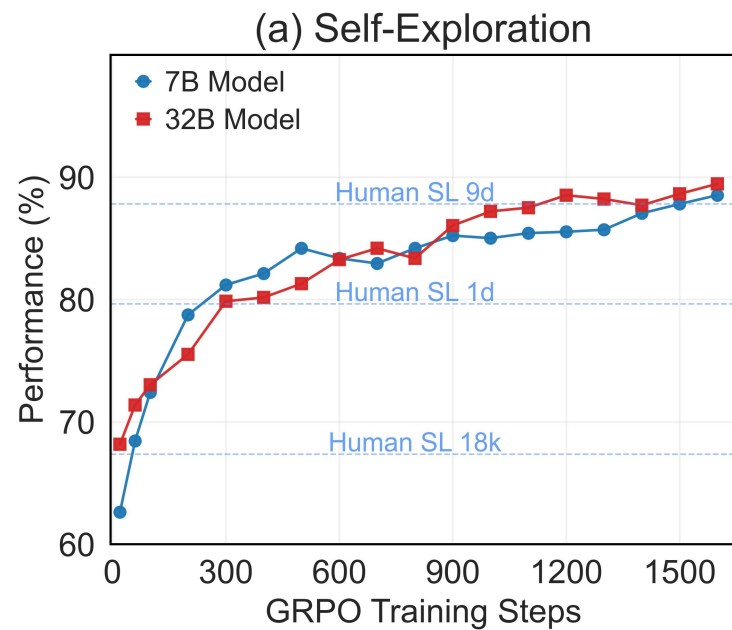
Strong General Performance

Expert-level Go Performance

Table 1: Main results on Go professional benchmark and general benchmarks. We evaluate and compare the models’ performance on KataGo-Bench-1K and several general reasoning benchmarks, with the evaluation results on general benchmarks obtained using Opencompass [Contributors, 2023]. The score on BBEH is calculated by averaging the results from each subset.

Model	KataGo-Bench	General Benchmark					
		GPQA Diamond	BBEH	KOR-Bench	AIME	MATH	LiveCodeBench
Closed-source Models							
DeepSeek-R1	17.6	69.7	44.5	78.3	86.7	97.6	83.8
o1-mini	27.3	61.1	-	70.16	56.7	95.0	75.0
Claude3.7-Sonnet	34.3	67.7	33.8	64.8	30.0	79.8	63.2
Open-source Models (7B)							
Qwen2.5-7B-Base	1.4	28.3	9.9	20.4	13.3	83.2	9.0
Qwen2.5-7B-Instruct	8.0	39.9	12.9	42.8	3.3	92.6	16.2
DeepSeek-R1-Distill-Qwen-7B	0.6	41.4	13.1	55.9	33.3	88.2	20.4
LoGos(7B)	88.1	37.9	22.1	65.7	40.0	93.2	23.4
Open-source Models (32B)							
Qwen2.5-32B-Base	1.5	35.9	15.6	42.8	10.0	90.7	13.8
Qwen2.5-32B-Instruct	6.8	46.0	18.0	58.1	20.0	95.2	29.9
DeepSeek-R1-Distill-Qwen-32B	4.7	56.1	27.4	70.0	46.7	94.5	36.5
LoGos(32B)	88.6	63.6	34.1	74.8	56.7	96.5	50.9
Go Professional Models							
KataGo-HumanSL-18k	67.4	-	-	-	-	-	-
KataGo-HumanSL-1d	79.7	-	-	-	-	-	-
KataGo-HumanSL-5d	85.5	-	-	-	-	-	-
KataGo-HumanSL-9d	87.8	-	-	-	-	-	-

Experiments: Reinforcement Learning



Part 4 | Conclusion



Conclusion

When AlphaGo Thinks

- Our method presents a viable solution for **general LLMs** to **master specialized tasks**.

A Possible Solution for more tasks with:

- Lack of natural language data
- scalable structured domain specific data
- a reward function

Limitations

- Incorrect Go terms
- Sometimes ambiguous COT
- Requires expert level human supervision



上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

Thank you for listening!



Code



Scan to play with our model online!



mayichuan@pjlab.org.cn

If you are interested in our work, feel free to contact us!