AUTOACT: Automatic Agent Learning from Scratch for QA via Self-Planning

Shuofei Qiao, Ningyu Zhang, Runnan Fang, Yujie Luo, Wangchunshu Zhou, Yuchen Eleanor Jiang, Chengfei Lv, Huajun Chen

ACL 2024

Presenter: Yu-Hua Zeng

National Cheng Kung University

國立成功大學
National Cheng Kung University

OUTLINE

- Introduction
- Related work
- Method
- Experiment
- Analysis
- Future work

Introduction - Language Agents

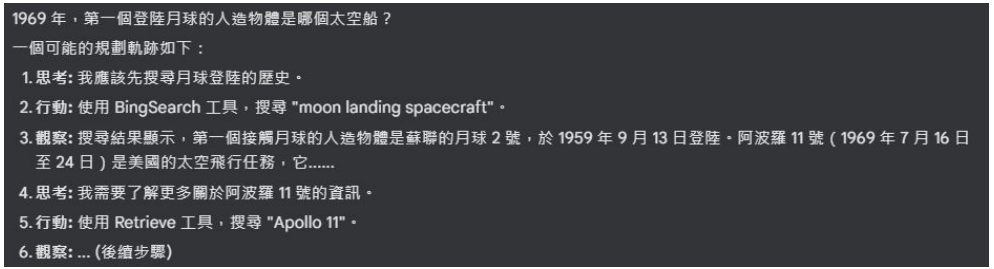
- Language agents[1] leverage the powerful reasoning capabilities of Large Language Models (LLMs) to interact with executable tools.
- The process of endowing LLMs with such interactive capabilities is referred to as Agent Learning wherein planning[2] plays a pivotal role
  - Invoking external tools[3] e.g. Bing, function calling
  - Decomposing complex questions into simpler ones[4] e.g. COT
  - Reflecting on past mistakes[5]
  - Aggregating information from various sources to reach the final answer

[1] Taicheng Guo, Xiuying Chen, Yaqi Wang, Ruidi Chang, Shichao Pei, Nitesh V. Chawla, Olaf Wiest, and Xi angliang Zhang. 2024. Large language model based multi-agents: A survey of progress and challenges.
[2] Xu Huang, Weiwen Liu, Xiaolong Chen, Xingmei Wang, Hao Wang, Defu Lian, Yasheng Wang, Ruim ing Tang, and Enhong Chen. 2024b. Understanding the planning of llm agents: A survey
[3] Yongliang Shen, Kaitao Song, Xu Tan, Dongsheng Li, Weiming Lu, and Yueting Zhuang. 2023. Hugging gpt: Solving AI tasks with chatgpt and its friends in huggingface. CoRR, abs/2303.17580.
[4] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc V. Le, and Denny Zhou. 2022. Chain-of-thought prompt ing elicits reasoning in large language models. In NeurIPS.
[5] Noah Shinn, Beck Labash, and Ashwin Gopinath. 2023. Reflexion: language agents with verbal reinforce ment learning. CoRR, abs/2303.11366

Introduction - Challenge

1. Training open source models necessitates a substantial amount of annotated QA data pairs and still relies on closed source models to synthesize planning trajectories

2. Using a single model for multiple functions

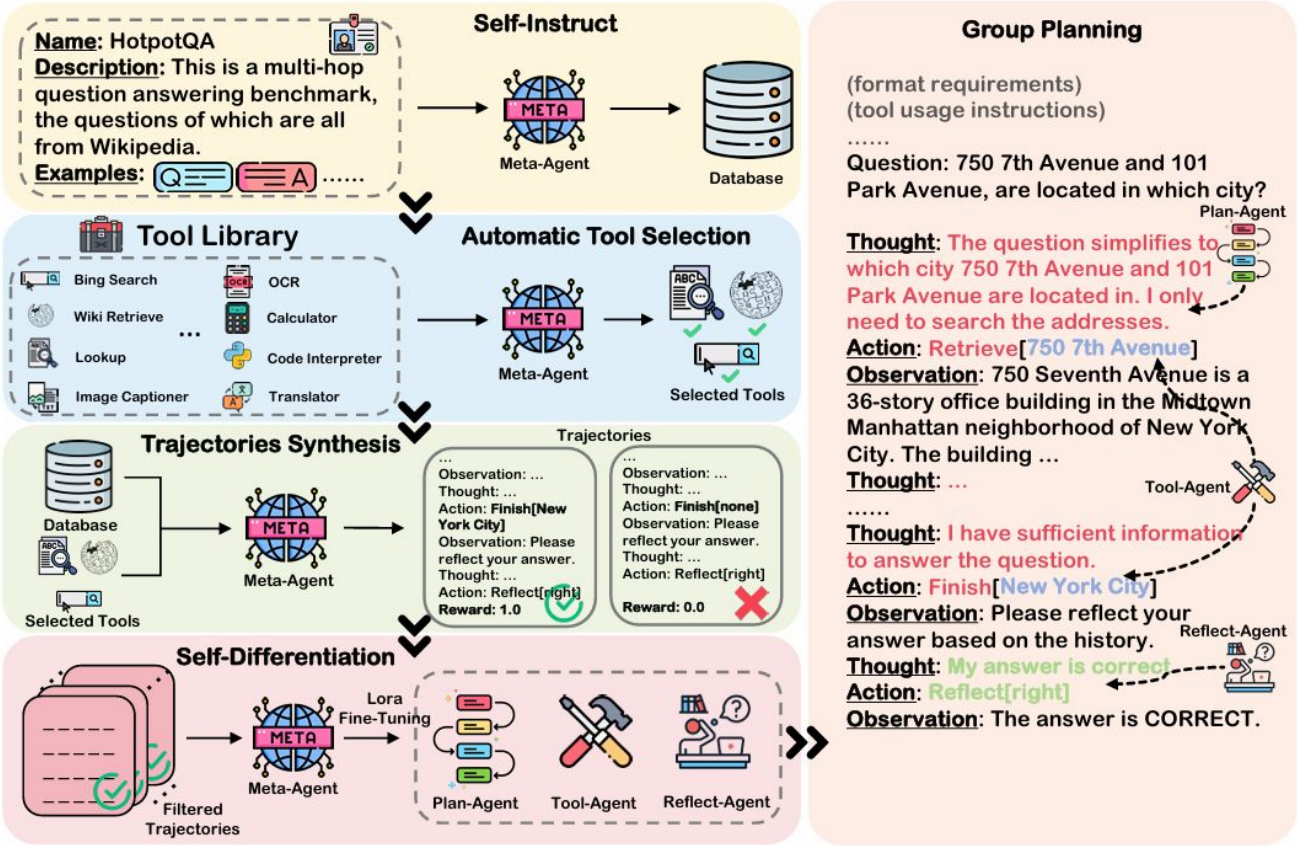


[1] Nelson F Liu, Tianyi Zhang, and Percy Liang. 2023. Evaluating verifiability in generative search engines. arXiv preprint arXiv:2304.09848 (2023).
[2] Jacob Menick, Maja Trebacz, Vladimir Mikulik, John Aslanides, Francis Song, Martin Chadwick, Mia Glaese, Susannah Young, Lucy Campbell-Gillingham, Geoffrey Irving, et al. 2022. Teaching language models to support answers with verified quotes. arXiv preprint arXiv:2203.11147 (2022).
[3] Reiichiro Nakano, Jacob Hilton, Suchir Balaji, Jeff Wu, Long Ouyang, Christina Kim, Christopher Hesse, Shantanu Jain, Vineet Kosaraju, William Saunders, Xu Jiang, Karl Cobbe, Tyna Eloundou, Gretchen Krueger, Kevin Button, Matthew Knight, Benjamin Chess, and John Schulman. 2022. WebGPT: Browser-assisted question-answering with human feedback. arXiv:2112.09332 [cs.CL]

Related work - LLM-Powered Agent

Method	Data Acquisition	Trajectory Acquisition	Planning	Multi-Agent	Fine-Tuning	Generality	Reflection
REACT (Yao et al., 2023)	User	Prompt	Iterative	✗	✗	✔	✗
Reflexion (Shinn et al., 2023)	User	Prompt	Iterative	✗	✗	✔	✔
Camel (Li et al., 2023)	User	Prompt	Iterative	✔	✗	✔	✗
Chameleon (Lu et al., 2023)	User	Prompt	Global	✗	✗	✔	✗
HuggingGPT (Shen et al., 2023)	User	Prompt	Global	✗	✗	✔	✗
AutoGPT (Torantulino, 2023)	User	Prompt	Iterative	✗	✗	✔	✔
BOLAA (Liu et al., 2023)	User	Prompt	Iterative	✔	✗	✔	✗
AgentVerse (Chen et al., 2023d)	User	Prompt	Iterative	✔	✗	✔	✗
Agents (Zhou et al., 2023b)	User	Prompt	Iterative	✔	✗	✔	✗
AgentTuning (Zeng et al., 2023)	Benchmark	GPT-4	Iterative	✗	✔	✗	✗
FIREACT (Chen et al., 2023a)	Benchmark	GPT-4	Iterative	✗	✔	✗	✔
Lumos (Yin et al., 2023)	Benchmark	Benchmark + GPT-4	Both	✔	✔	✗	✗
AUTOACT (ours)	User + Self-Instruct	Self-Planning	Iterative	✔	✔	✔	✔

Method - Framework



Method - Self-Instruct

- Acquire a sufficient amount of task data

HotpotQA:
Question: The deepest part of the ocean, is located in which ocean?
Answer: The Pacific Ocean

ScienceQA:
Question: Which of the following is a type of renewable energy?
Options: (A) Coal (B) Oil (C) Natural gas (D) Solar power
Caption: A picture of a solar cell
Answer: D. Solar power

Prompt for Self-Instruct

I want you to be a QA pair generator to generate high-quality questions for use in Task described as follows:
Task Name: **[task_name]**
Task Description: **[task_description]**
Here are some Q&A pair examples from the Task:
[QA_pairs]
Modeled on all the information and examples above, I want you to generate new different **[gen_num_per_round]** Question-Answer pairs that cover a wide range of topics, some of which are difficult, some of which are easy, and require multiple steps of reasoning to get to the final answer. The format is like below:
[one_example]

Method - Automatic Tool Selection

- Select applicable tools for each task automatically

Name	Definition	Usage
BingSearch	BingSearch engine can search for rich knowledge on the internet based on keywords, which can compensate for knowledge fallacy and knowledge outdated.	BingSearch[query], which searches the exact detailed query on the Internet and returns the relevant information to the query. Be specific and precise with your query to increase the chances of getting relevant results. For example, Bingsearch[popular dog breeds in the United States]
Retrieve	Retrieve additional background knowledge crucial for tackling complex problems. It is especially beneficial for specialized domains like science and mathematics, providing context for the task	Retrieve[entity], which retrieves the exact entity on Wikipedia and returns the first paragraph if it exists. If not, it will return some similar entities to retrieve. For example, Retrieve[Milhouse]
Lookup	A Lookup Tool returns the next sentence containing the target string in the page from the search tool, simulating Ctrl+F functionality on the browser.	Lookup[keyword], which returns the next sentence containing the keyword in the last passage successfully found by Retrieve or BingSearch. For example, Lookup[river].

```
TOOL_POOL = [
    {"name": "BingSearch","defin
    {"name": "Retrieve", "defini
    {"name": "Lookup", "definiti
    {"name": "Image2Text", "defi
    {"name": "Text2Image","defin
    {"name": "KnowledgeGraph","d
    {"name": "Database", "defini
    {"name": "Calculator", "defi
    {"name": "Table Verbalizer",
    {"name": "Code Interpreter",
    ]
```

Prompt for Automatic Tool Selection

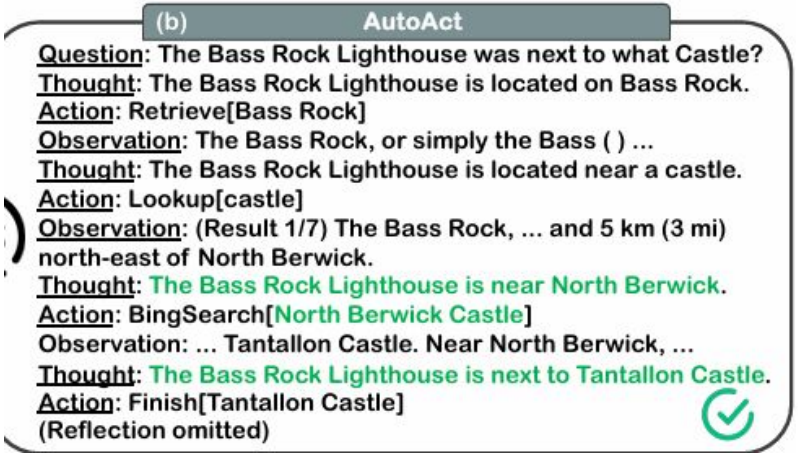
To successfully complete a complex task, the collaborative effort of three types of agents is typically required:
1. Plan Agent. This agent is used to plan the specific execution process of the benchmark, solving a given task by determining the order in which other expert language models are invoked;
2. Tool Agent. This agent is employed to decide how to use a specific tool when addressing a task. Tools encompass interactive tools within the task environment as well as external tools or models. The Tool Agent includes various tools that can be flexibly chosen;
3. Reflect Agent. This agent reflects on historical information and answers to assess whether the response aligns with the provided query.
Above all, the Tool Agent includes many tools that can be flexibly selected. Now your task is to select 3 tools from the Tool Library for solving a given task. Note that all tools are based on language models, and their inputs and outputs must be text. You only need to provide the names and descriptions of the tools in order, without any additional output.

Task Prompt Template

The following is the given task name and description, and you need to choose 3 corresponding tools from the Tool Library according to the above rules in the format of one line, one tool.
Task Name: **[task_name]**
Task Description: **[task_description]**
Tool Library: **[list_of_tools]**

Method - Trajectories Synthesis

- synthesize trajectories in a zero-shot manner on the database D adhering to the format of **Thought-Action-Observation**



Prompt for Trajectories Synthesis

I expect you to excel as a proficient question answerer in the task.
Task Name: **[task_name]**
Task Description: **[task_description]**
Solve a question-answering task with interleaving Thought, Action, and Observation steps.
Thought can reason about the current situation, and Action can be **[action_num]** types:
list of action selected from automatic tool selection [name, definition , usage]
Question: **[question][scratchpad]**

Method - Conclusion

Step	Input	Output	Key Focus
Self-Instruct	Initial dataset C	Expanded dataset D	數據生成與多樣性
Automatic Tool Selection	Tool Library and task description T	Tool invocation plans	工具使用的選擇
Trajectory Synthesis	Dataset D, tool plans	Task planning trajectories	多步推理與執行步驟模擬

Method - Self-Differentiation

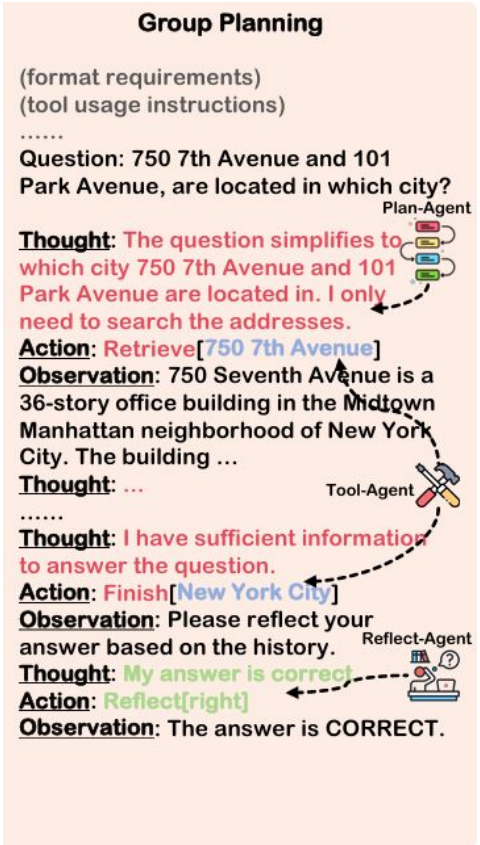
- Plan Agent $\tau_t, \alpha_t^m = \pi_{\text{plan}}(\mathcal{S}, \mathcal{T}_s, \mathcal{H}_t)$
- Tool Agent $\alpha_t^p = \pi_{\text{tool}}(\mathcal{S}, \mathcal{T}_s, \mathcal{H}_t, \tau_t, \alpha_t^m)$
- Reflect Agent $\tau^r, \alpha^r = \pi_{\text{reflect}}(\mathcal{S}, \mathcal{T}_s, \mathcal{H})$

Agent	Training Y
Plan Agent	下一步的 Thought（新的計劃或任務分解）
Tool Agent	選擇的工具名稱及其參數
Reflect Agent	錯誤修正步驟或建議

Plan-Agent (generate Thought):
Input:
(format requirements) (tool usage instructions)
Question: The first human-made object to land on the moon, in 1969, was which spacecraft?
Thought: I should first search the Moon landing history.
Action: BingSearch[moon landing spacecraft]
Observation: A Moon landing or lunar landing is the arrival of a spacecraft on the surface of the Moon. The first human-made object to touch the Moon was the Soviet Union's Luna 2, on 13 September 1959. Apollo 11 (July 16–24, 1969) was the American spaceflight that .....
Thought:
Output:
Retrieve for more information about Apollo 11

Reflect-Agent (generate Thought):
Input:
(format requirements) (tool usage instructions)
Question: The first human-made object to land on the moon, in 1969, was which spacecraft?
Thought: I should first search the Moon landing history.
Action: BingSearch[moon landing spacecraft]
Observation: A Moon landing or lunar landing is the arrival of a spacecraft on the surface of the Moon. The first human-made object to touch the Moon was the Soviet Union's Luna 2, on 13 September 1959. Apollo 11 (July 16–24, 1969) was the American spaceflight that .....
Thought: Retrieve for more information about Apollo 11.
Action: Retrieve[Apollo 11]
.....
Action: Finish[Apollo 11]
Observation: Please reflect your answer based on the history.
Thought:
Output:
The question asks about the first human-made object to land on the moon, so it seems that the Soviet Union's Luna 2 is more like the answer.

Method - Group Planning



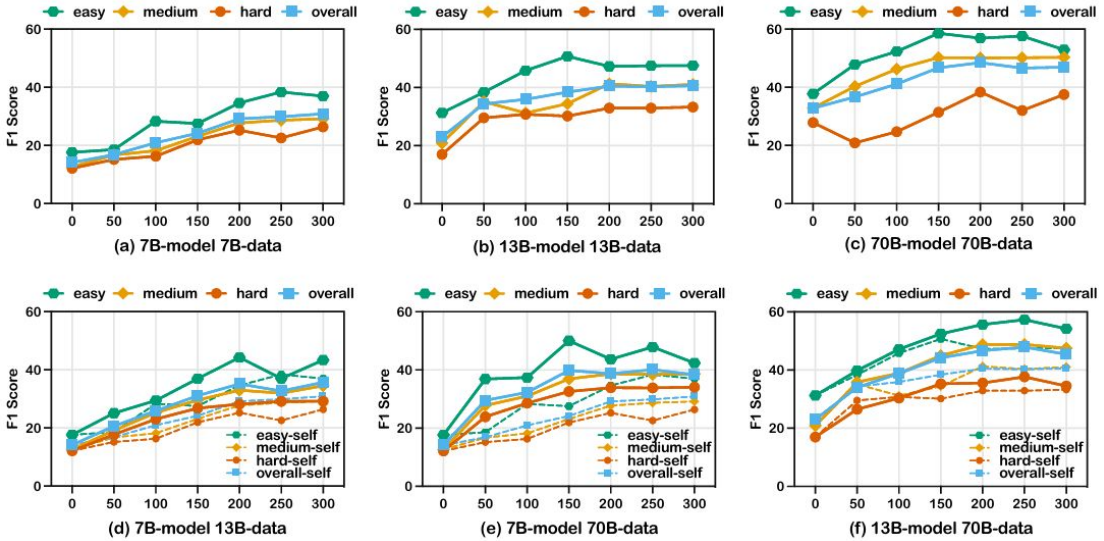
Experiment - Result

Backbone	Method	HotpotQA				ScienceQA			
		Easy	Medium	Hard	All	G1-4	G5-8	G9-12	All
GPT-3.5 Turbo	CoT	48.21	44.52	34.22	42.32	60.83	55.83	65.00	60.56
	Zero-Shot Plan*	50.71	45.17	38.23	44.70	76.67	61.67	78.33	72.22
Mistral-7B Instruct-v0.2	CoT	33.70	22.38	22.14	26.07	54.17	50.00	60.00	54.72
	ReAct	38.09	27.57	22.05	29.24	63.33	58.33	62.50	61.39
	Chameleon	37.07	26.67	19.20	27.65	65.83	62.50	66.67	65.00
	Reflexion	40.78	35.02	28.36	34.72	67.50	65.83	69.17	67.50
	BOLAA	40.86	32.11	22.36	31.78	64.17	61.67	65.83	63.89
	ReWOO	38.42	31.89	25.98	32.10	60.83	58.33	64.17	61.11
	FireAct	45.52	32.02	30.17	35.90	65.00	62.50	64.17	63.89
	AUTOACT	**48.69**	**36.65**	**31.37**	**38.89**	**69.17**	**68.33**	**72.50**	**70.00**
Llama-2 13B-chat	CoT	37.90	25.28	21.64	28.27	61.67	52.50	69.17	61.11
	ReAct	28.68	22.15	21.69	24.17	57.50	51.67	65.00	58.06
	Chameleon	40.01	25.39	22.82	29.41	69.17	60.83	73.33	67.78
	Reflexion	44.43	37.50	28.17	36.70	67.50	64.17	73.33	68.33
	BOLAA	33.23	25.46	25.23	27.97	60.00	54.17	65.83	60.00
	ReWOO	30.09	24.01	21.13	25.08	57.50	54.17	65.83	59.17
	FireAct	45.83	38.94	26.06	36.94	60.83	57.50	67.50	61.94
	AUTOACT	**47.29**	**41.27**	**32.92**	**40.49**	**70.83**	**66.67**	**76.67**	**71.39**
Llama-2 70B-chat	CoT	45.37	36.33	32.27	37.99	74.17	64.17	75.83	71.39
	ReAct	39.70	37.19	33.62	36.83	64.17	60.00	72.50	65.56
	Chameleon	46.86	38.79	34.43	40.03	77.83	69.17	76.67	74.56
	Reflexion	48.01	46.35	35.64	43.33	75.83	67.50	78.33	73.89
	BOLAA	46.44	37.29	33.49	39.07	70.00	67.50	75.00	70.83
	ReWOO	42.00	39.58	35.32	38.96	65.00	61.67	76.67	67.78
	FireAct	50.82	41.43	35.86	42.70	72.50	68.33	75.00	71.94
	AUTOACT	**56.94**	**50.12**	**38.35**	**48.47**	**82.50**	**72.50**	**80.83**	**78.61**

	HotpotQA	ScienceQA
AUTOACT	48.47	78.61
- reflection	$45.66_{\downarrow 2.81}$	$75.28_{\downarrow 3.33}$
- multi	$42.81_{\downarrow 5.66}$	$69.72_{\downarrow 8.89}$
- fine-tuning	$32.84_{\downarrow 15.63}$	$61.94_{\downarrow 16.67}$
- filtering	$32.51_{\downarrow 15.96}$	$59.17_{\downarrow 19.44}$

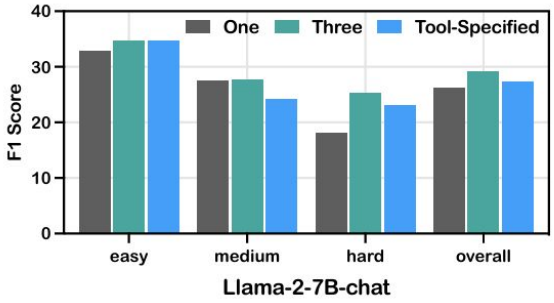
Analysis

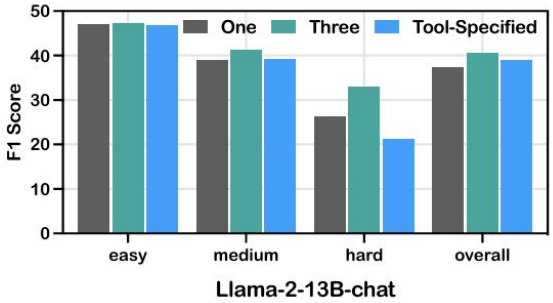
- Larger training data scale does not necessarily mean better results.

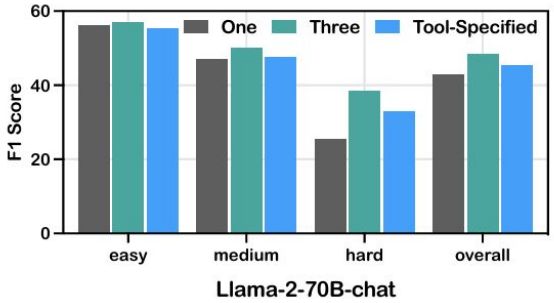


Analysis

- Moderate division-of-labor benefits group planning performance







- The planning performance of AUTOACT can be limited by the model's ability to access internal knowledge through self-instruct

Future work

- New multi-agent strategy
- Automatic tool selection
- Knowledge Distillation