

# Automated Design of Agentic Systems (ADAS)

Shengran Hu, Cong Lu, Jeff Clune

ICLR 2025



# Automated Design of Agentic Systems

Shengran Hu<sup>1,2</sup>



Cong Lu<sup>1,2</sup>



Jeff Clune<sup>1,2,3</sup>



<sup>1</sup>University of British Columbia

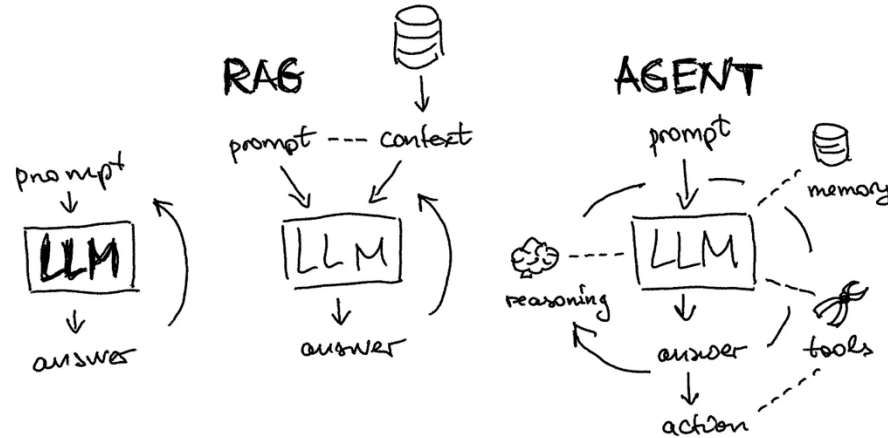
<sup>2</sup>Vector Institute

<sup>3</sup>Canada CIFAR AI Chair



# AI Agents

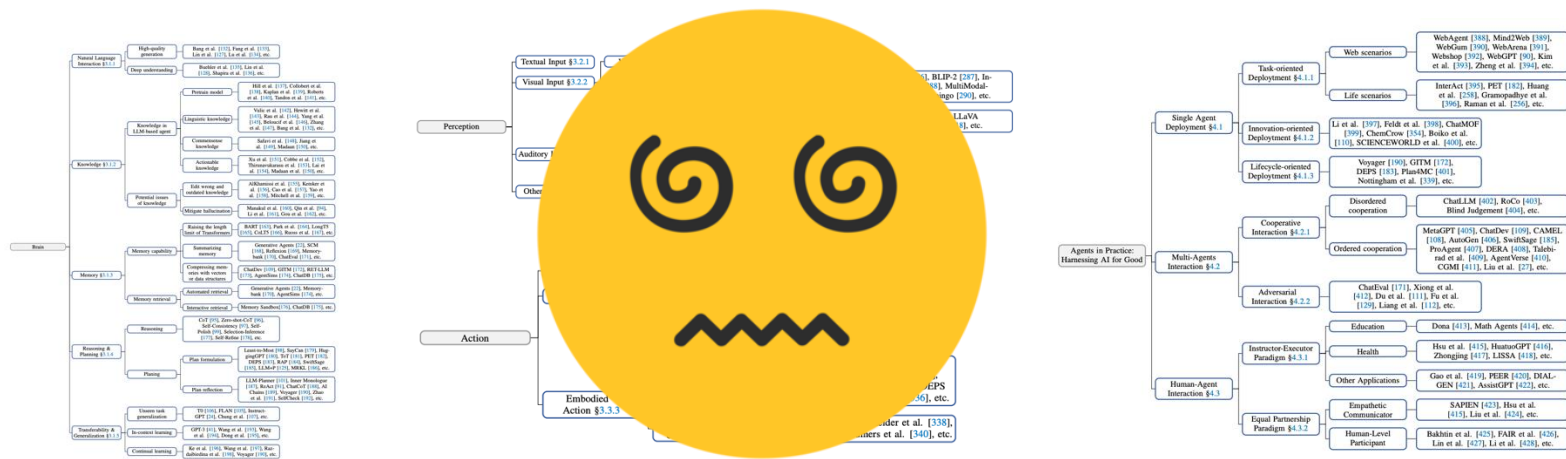
- **Agentic Systems powered by Foundation Models (FMs)**
  - solve tasks by planning, using tools, and carrying out multiple, iterative steps of processing



(Alex Honchar, 2024)

# AI Agents

- **Designing agents requires a significant amount of manual effort**
  - Tuning prompts, tools, workflows, etc.



Partial list of existing building blocks, applications, agent designs, etc.  
(Xi et al., 2023)

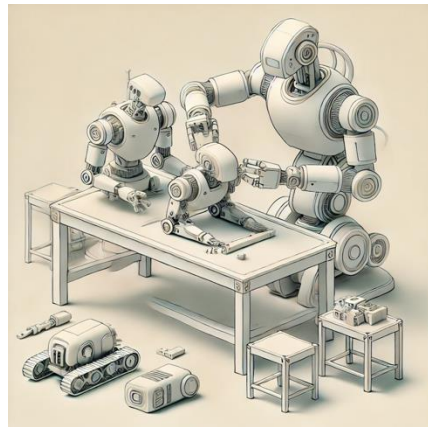
# Automated Design of Agentic Systems (ADAS)

- **Automated Design of Agentic Systems (ADAS)**
  - Saving human effort
  - Faster path for agent development
    - Invent new building blocks
    - Design agents for specific tasks



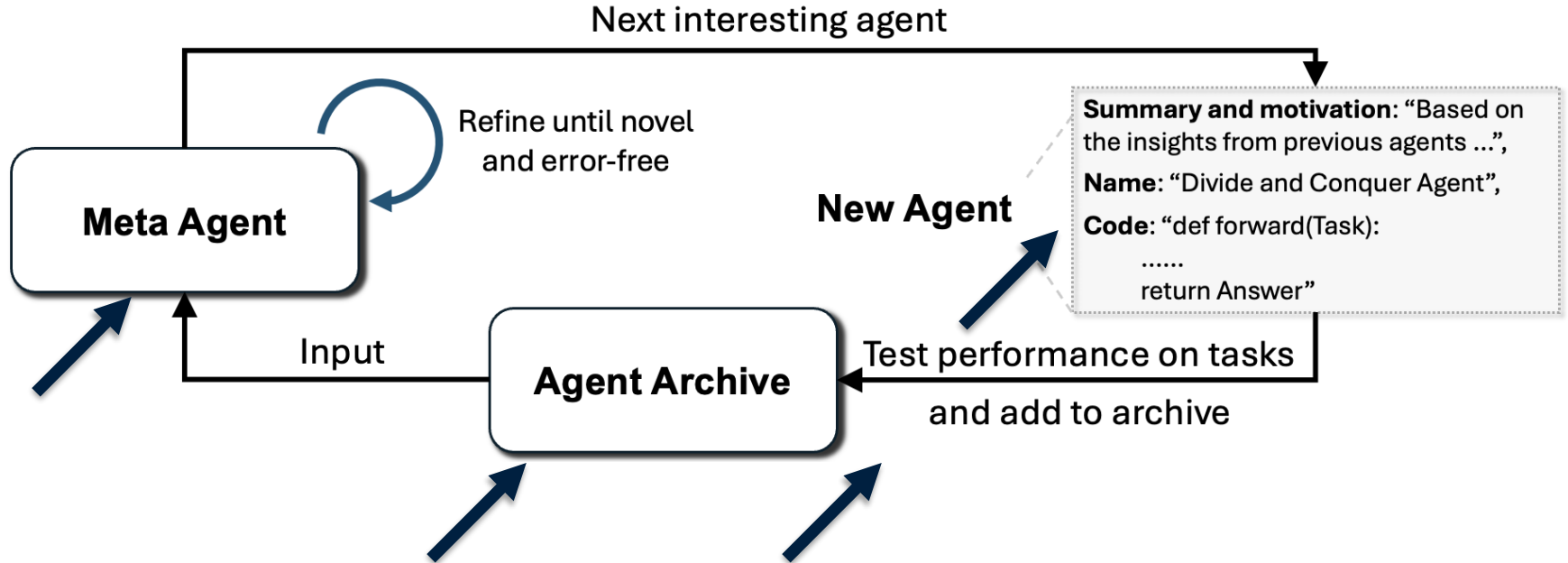
# Automated Design of Agentic Systems (ADAS)

- **Defining agent in code**
  - Turning Completeness → ANY possible agents
  - V.s. Prior Work: Prompt optimization / pre-defined graphs
- **Use a “meta” agent to programming new agents.**
  - Utilizing FM’s prior in coding and agent itself



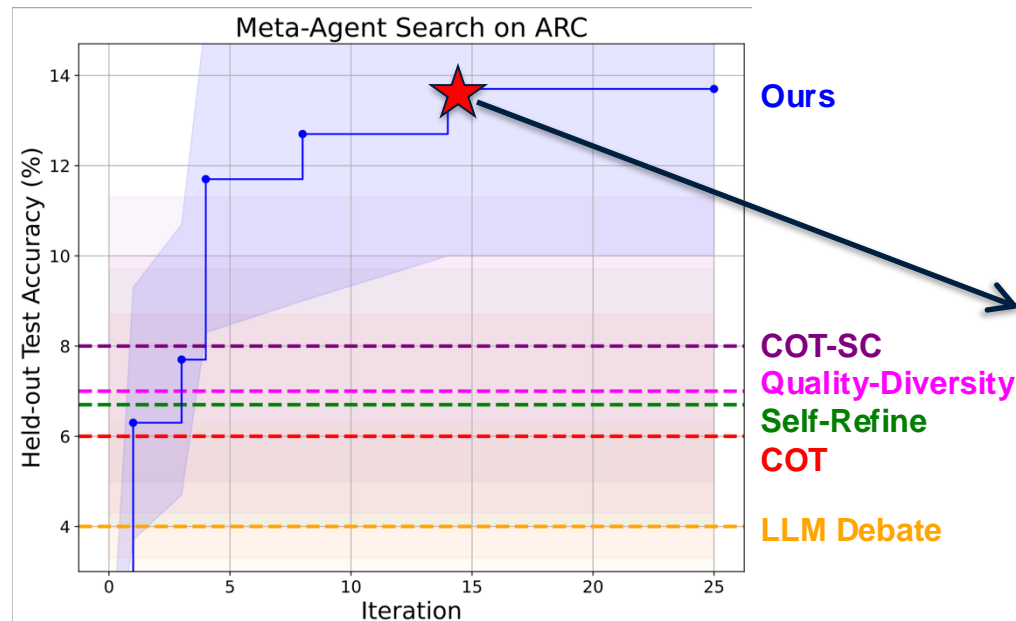
# Meta Agent Search

- The first ADAS algorithm that enables complete design in code space

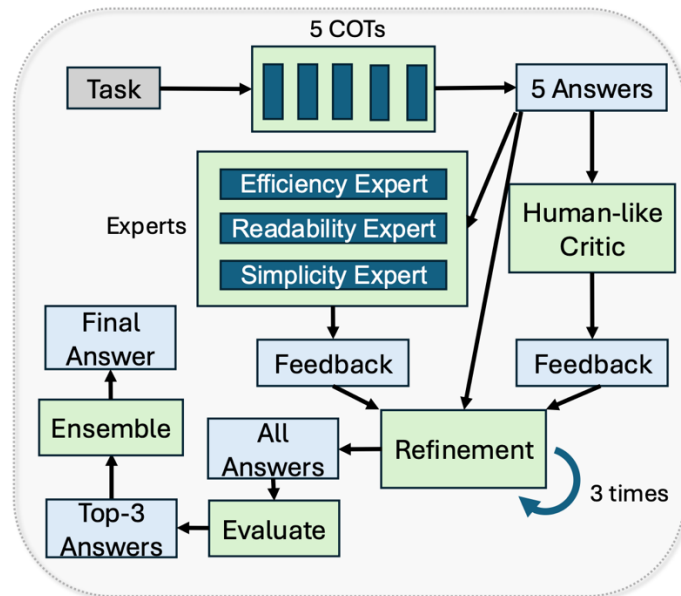


# ARC challenge

- Meta Agent Search outperform SOTA hand-designed agents



## Best Agent on ARC



Structured Feedback and Ensemble Agent



# Reasoning and Problem-Solving Domains

- Meta Agent Search outperform SOTA hand-designed agents

Agent Name	F1 Score	Accuracy (%)		
	Reading Comprehension	Math	Multi-task	Science
State-of-the-art Hand-designed Agents				
Chain-of-Thought (Wei et al., 2022)	64.2 ± 0.9	28.0 ± 3.1	65.4 ± 3.3	29.2 ± 3.1
COT-SC (Wang et al., 2023b)	64.4 ± 0.8	28.2 ± 3.1	65.9 ± 3.2	30.5 ± 3.2
Self-Refine (Madaan et al., 2024)	59.2 ± 0.9	27.5 ± 3.1	63.5 ± 3.4	31.6 ± 3.2
LLM Debate (Du et al., 2023)	60.6 ± 0.9	39.0 ± 3.4	65.6 ± 3.3	31.4 ± 3.2
Step-back Abstraction (Zheng et al., 2023)	60.4 ± 1.0	31.1 ± 3.2	65.1 ± 3.3	26.9 ± 3.0
Quality-Diversity (Lu et al., 2024c)	61.8 ± 0.9	23.8 ± 3.0	65.1 ± 3.3	30.2 ± 3.1
Role Assignment (Xu et al., 2023)	65.8 ± 0.9	30.1 ± 3.2	64.5 ± 3.3	31.1 ± 3.1
Automated Design of Agentic Systems on Different Domains				
Best Agents from Meta Agent Search	79.4 ± 0.8	53.4 ± 3.5	69.6 ± 3.2	34.6 ± 3.2

# Transferability and Generalizability

Non-math domains

- Transfer to **Math and Non-Math Domains**

Math domains

Agent Name	Accuracy (%)			F1 Score	
	MGSM	GSM8K	GSM-Hard	MMLU	DROP
Manually Designed Agents					
Chain-of-Thought (Wei et al., 2022)	28.0 ± 3.1	34.9 ± 3.2	15.0 ± 2.5	65.4 ± 3.3	64.2 ± 0.9
COT-SC (Wang et al., 2023b)	28.2 ± 3.1	37.8 ± 3.4	15.5 ± 2.5	<b>65.9 ± 3.2</b>	64.4 ± 0.8
Self-Refine (Madaan et al., 2024)	27.5 ± 3.1	38.9 ± 3.4	15.1 ± 2.4	63.5 ± 3.4	59.2 ± 0.9
LLM Debate (Du et al., 2023)	<b>39.0 ± 3.4</b>	<b>43.6 ± 3.4</b>	17.4 ± 2.6	65.6 ± 3.3	60.6 ± 0.9
Step-back Abstraction (Zheng et al., 2023)	31.1 ± 3.2	31.5 ± 3.3	12.2 ± 2.3	65.1 ± 3.3	60.4 ± 1.0
Quality-Diversity (Lu et al., 2024c)	23.8 ± 3.0	28.0 ± 3.1	14.1 ± 2.4	65.1 ± 3.1	61.8 ± 0.9
Role Assignment (Xu et al., 2023)	30.1 ± 3.2	37.0 ± 3.4	<b>18.0 ± 2.7</b>	64.5 ± 3.3	<b>65.8 ± 0.9</b>
Top Agents Searched on MGSM (Math)		Transferred within Math Domains		Transferred beyond Math Domains	
Dynamic Role-Playing Architecture	<b>53.4 ± 3.5</b>	<b>69.5 ± 3.2</b>	<b>31.2 ± 3.2</b>	62.4 ± 3.4	70.4 ± 0.9
Structured Multimodal Feedback Loop	50.2 ± 3.5	64.5 ± 3.4	30.1 ± 3.2	<b>67.0 ± 3.2</b>	70.4 ± 0.9
Interactive Multimodal Feedback Loop	47.4 ± 3.5	64.9 ± 3.3	27.6 ± 3.2	64.8 ± 3.3	<b>71.9 ± 0.8</b>

# Future Work and Conclusion

- **Future Work**
  - Safety in ADAS
- **ADAS: towards automated design of agentic systems**
  - From hand-crafted solutions to learned solutions
  - Faster and better approach to build agentic systems
- **Meta Agent Search: The first ADAS algorithm that enables complete design in code space**
  - Discover high-performance, transferable agents under various domains.

# Thank you!



**All code, prompt, generated agents are available.**



QR code for GitHub