



## Autonomous Agents for Collaborative Task under Information Asymmetry

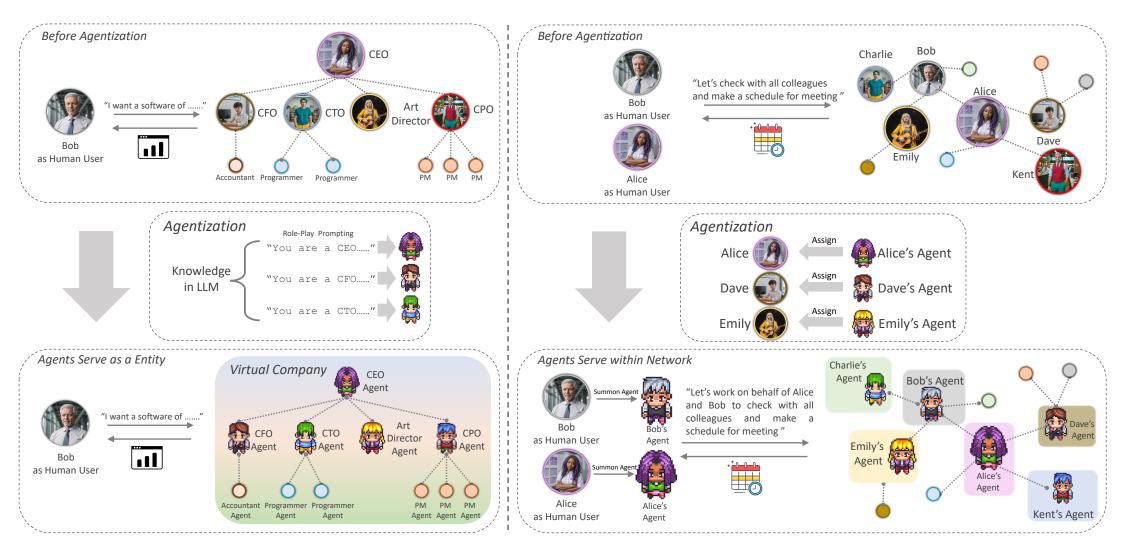
**Wei Liu,** Chenxi Wang, Yifei Wang, Zihao Xie, Rennai Qiu, Yufan Dang, Zhuoyun Du, Weize Chen, Cheng Yang, Chen Qian

Tsinghua University thinkwee2767@gmail.com





#### **Another Perspective**







## **Another Perspective**

	Multi-Agent as one System	Everybody gets a Agent		
Relationship with Human	Human uses Multi-Agent System	Humans and Agents are in the System		
Role-play	Agents role-play occupation Agents are just "Agents"			
Тороlоду	Layer, (De)centralized, Pool	Human Network		
Value of Agent	Knowledge of LLM Human User			
Is Multi-Agent necessary?	Endless Discussion	Multi-Agent Native		
Task	Knowledge goes first	Information goes first		
Challenge	Communication, Organization, Evolvement, Scale	and <b>information asymmetry</b> , interaction with human, privacy, end-side deploy		





- New Problem
  - Multi-Agent Task Solving under
    - Information Asymmetry
- New Task/Benchmark
  - InformativeBench
- New Framework
  - iAgents
- New Reasoning
  - InfoNav



*«Autonomous Agents for Collaborative Task under Information Asymmetry »* https://github.com/thinkwee/iAgents





#### **Problem Formulation**

$$Ans = Reasoning(Question, Rationales)$$
  

$$Rationales = R_1 \cup R_2$$
  

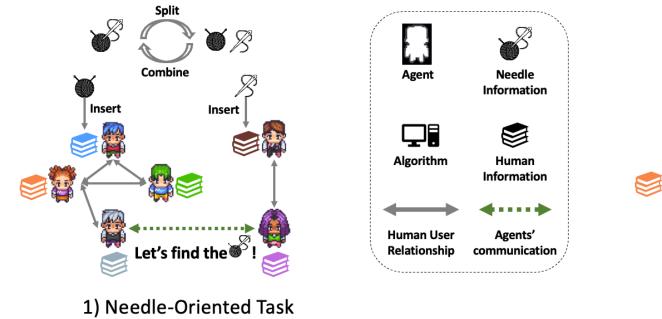
$$R_1, R_2 = Comm(Info_1, Info_2, Agent_1, Agent_2)$$

- Simple two humans/agents case
- The question requires all rationales to be collected in order to be answered
- Rationales are scattered in multiple human information





#### InformativeBench



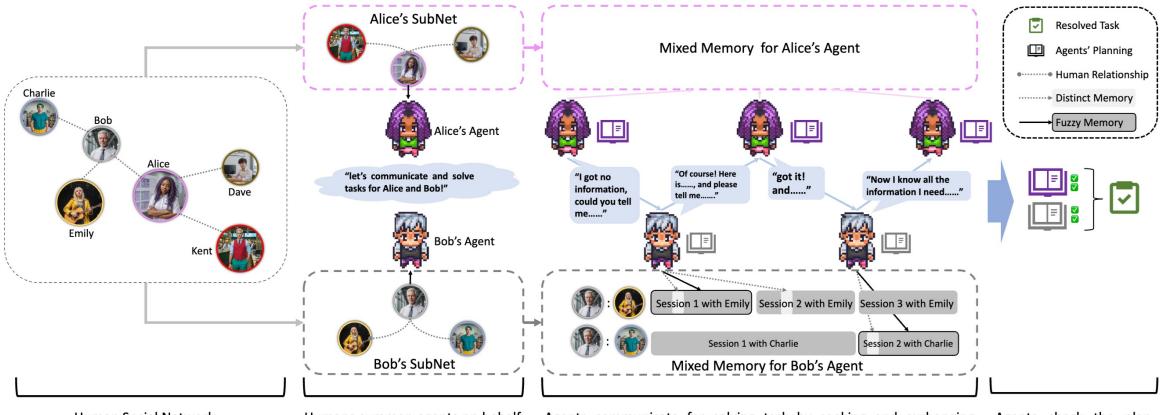


2) Reasoning-Oriented Task

- Simple two humans/agents case
- The question requires all rationales to be collected in order to be answered
- Rationales are scattered in multiple human information







Human Social Network

Humans summon agents on behalf of them for solving task.

Agents communicate for solving task by seeking and exchanging information. The communication is driven by *InfoNav*.

Agents check the plan, make consensus to draw conclusion.





#### (1) InfoNav Initilizes Planning

Agent 1: To effectively complete the task of finding out the ...... we need to collaborate closely. This requires us to compile a [comprehensive list of all individuals each of us has been in contact with], along with ...... Specifically, we must gather information on [each activity's duration listed in their schedules].

unknown: comprehensive list of all individuals each of us has been in contact with
unknown: each activity's duration listed in their schedules (2) Update Plan Text Based on Communication Progress

Agent 1: To effectively complete the task of finding out ...... [comprehensive list of all individuals each of us has been in contact with](Solved, which is A5, B5, C5, D5, E5, F5), .....we must gather information on [each activity's duration listed in their schedules](Solved, which is Book Club Meeting is 1 hour, E5's meditation session is 2 hours, E5's dog walking activity is 2 hours, meditation session attended by B5 is 2 hours). .....

update: comprehensive list of all individuals each of us has been in contact with --> A5, B5, C5, D5, E5, F5
update: each activity's duration listed in their schedules --> Book Club Meeting is 1 hour, E5's meditation session is 2 hours, E5's dog walking activity is 2 hours, meditation session attended by B5 is 2 hours

#### (3) Perform Consensus Check on Plan Text of two Agents

Agent 1: To effectively complete the task of finding out ...... we must gather information on [each activity's duration listed in their schedules](Solved, which is Book Club Meeting is 1 hour, E5's meditation session is 2 hours, E5's dog walking activity is 2 hours, meditation session attended by B5 is 2 hours). Agent 2: To accomplish the task of finding .....from our consideration and identify ..... This process involves gathering, comparing, and analyzing the [duration of various activities from the schedules of different people] (Solved, which is Book Club Meeting 1-1.5 hours, Meditation Session 2 hours, Dog Walking 2 hours, Camping Trip start time 17:00 but end time not specified) to determine the one(s) that last the longest.

X Book Club Meeting is 1 hour <----> Book Club Meeting 1-1.5 hours
 E5's meditation session is 2 hours <----> Meditation Session 2 hours
 E5's dog walking activity is 2 hours <----> Dog Walking 2 hours
 M meditation session attended by B5 is 2 hours <----> Meditation Session 2 hours
 None <----> Camping Trip start time 17:00 but end time not specified

(4) Get the Answer Meditation Session 2 hours, Dog Walking 2 hours

- Plan to list all information need to be gathered
- Ask for information from human
- Exchange information with other agents
- Update placeholders in plan
- Consensus reasoning

# Communication is essentially a Multi-Turn ReAct Process

ReAct: Synergizing Reasoning and Acting in Language Models Yao, Zhao, Yu, Du, Shafran, Narasimhan, Cao. ICLR 2023





LLM Backend	Needle-Type NP FriendsTV		Algorithm-Type ScheduleEasy ScheduleMedium ScheduleHar			
GPT 4	64.00%	57.94%	56.67%	51.00%	22.80%	
GPT 3.5	51.00%	35.71%	36.67%	18.00%	12.25%	
Claude Sonnet	50.00%	34.13%	43.33%	17.44%	18.66%	
Gemini 1.0	40.00%	28.57%	26.67%	22.33%	14.40%	

Table 3: Evaluation results of *iAgents* on *InformativeBench* with different LLM backends.

SOTA LLM-powered Agents have difficulty solving the *InformativeBench* 

Even for the easiest NP dataset (it's just "needle in the pond")





Experiment	Reasonin Easy	g-Oriented (So Medium	chedule Dataset) Hard	Needle NP	e-Oriented FriendsTV
iAgents (Full Model)	36.67%	18.00%	12.25%	51.00%	35.71%
Ablation on InfoNav:					
w/o InfoNav	10.00%	3.56%	7.34%	39.00%	34.92%
Ablation on other mech	anisms (Lim	ited Applicabil	ity):		
w/o Recursive Comm	–	_	_	48.00%	23.02%
w/o Fuzzy Memory w/o Clear Memory	_	-	_	_	29.37% 33.33%
w/o Clear Wielliory	<u> </u>			-	55.55%

### Without InfoNav (and other designs in iAgents), It can be worse

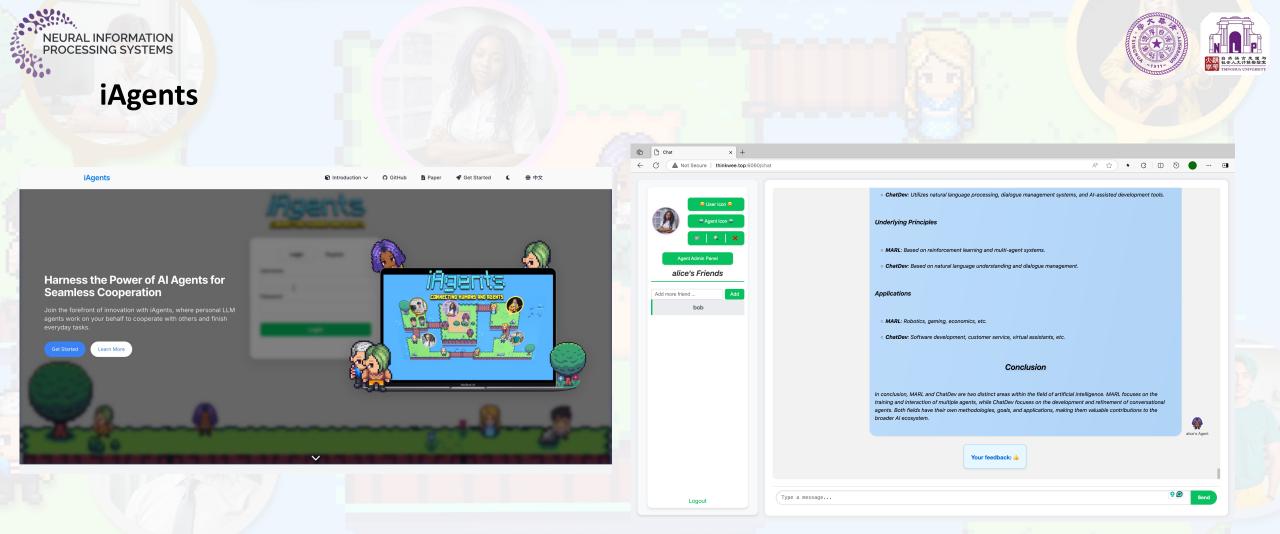
Less than 10% success rate without InfoNav on Schedule dataset





Sample	#Rationales in <i>InfoNav</i>	#Rationales Solved per Update	Rationales Solved Ratio	Fake Solved Ratio	Consensus Ratio
Predict Right	5.29	2.04	84.75%	3.49%	70.52%
Predict Wrong	5.63	1.69	67.23%	5.40%	62.70%
All	5.45	1.87	76.22%	4.42%	66.20%

- Some empricial observation
  - Planning, the only knowledge-type ability, is still important
  - List information carefully
  - Concurrent Reasoning
  - Correct is more important than comprehensive, consensus makes up the error
  - Hallucination can happen everywhere, "Fake Solved"



*«Autonomous Agents for Collaborative Task under Information Asymmetry »* 

https://thinkwee.top/iagents/

https://github.com/thinkwee/iAgents



iAgents





[※ Website | Quickstart | Paper | Huggingface Space]
[Ⅲ Wiki | informativeBench | Interact with Friends | More from our Team]

*«Autonomous Agents for Collaborative Task under Information Asymmetry »* 

https://thinkwee.top/iagents/

https://github.com/thinkwee/iAgents