LLMs Can Simulate Standardized Patients via Agent Coevolution

Zhuoyun Du[1*], Lujie Zheng[1*], Renjun Hu[2*], Yuyang Xu[1], Xiawei Li[1], Ying Sun[3], Wei Chen[1], Wu Jian[1], Haolei Cai[1], Haochao Ying[1]

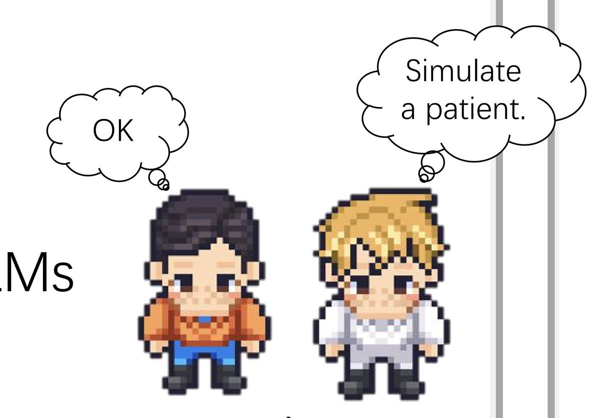
[1]Zhejang University & [2]East China Normal University & [3]Sun Yat-sen University Cancer Center

* Indicates equal contribution

ZHEJIANG UNIVERSITY
ACL 2025 VIENNA
JULY 27 - AUGUST 1

Introduction

SPs simulate real patients for education propose.
- They are use to enhance doctor's clinical skills.
- SPs incurs significant operational costs.
- Negtive impact on individuals.
- Requirements and data are hard to follow by LLMs

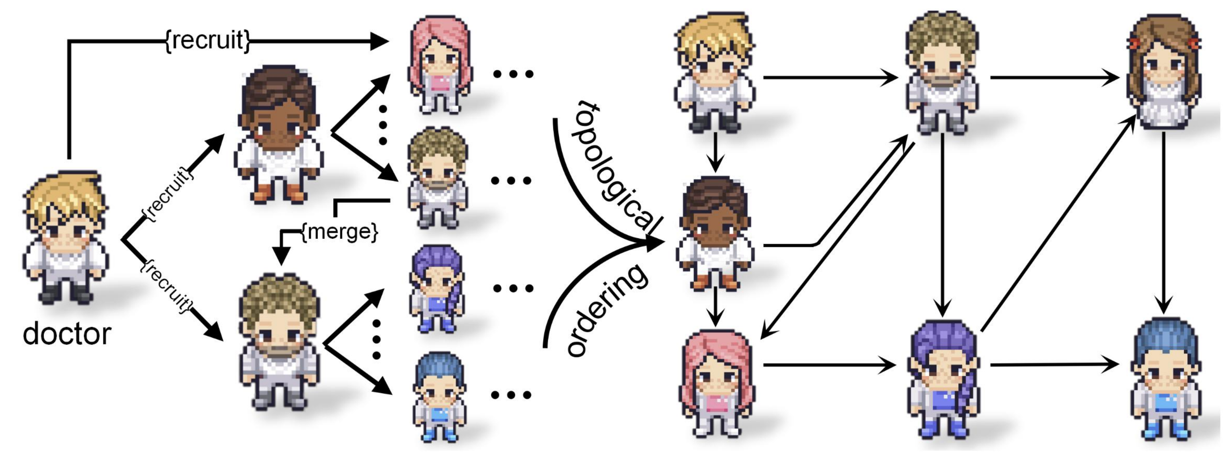


Previous solutions:
- Establish knowledge graphs. (No expresstion improvement)
- Elicit principles from human feedback. (Costy)

Method

We mirrors the diagnostic process into a series of phases. Then we assign an agent pair in the diagnosis process containing a patient agent with various profiles and RAG, doctor agents that can make dynamic recruitment.

Easy to design and prevents information backflow



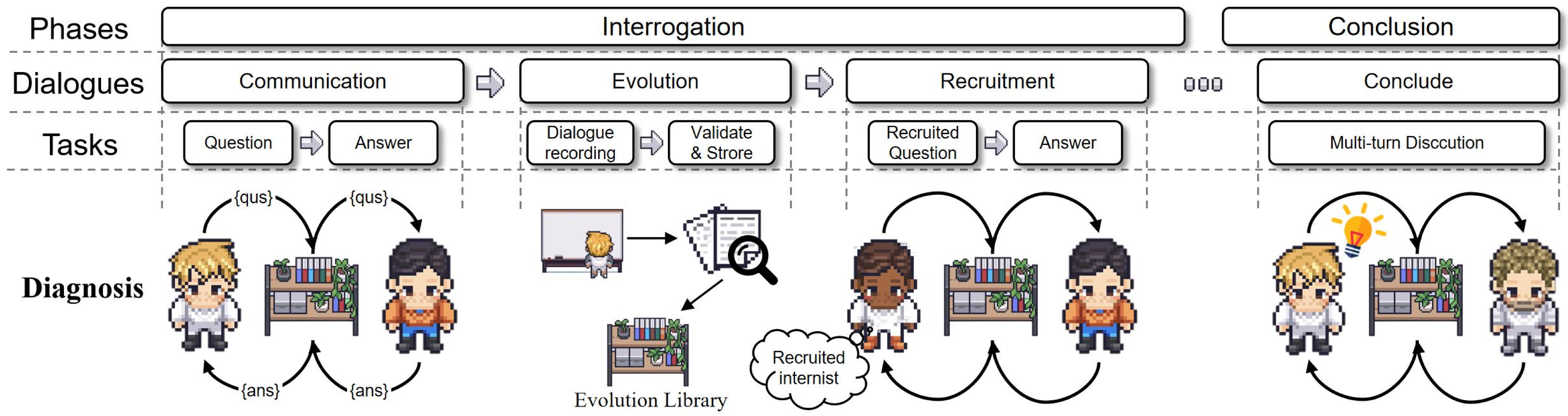
Attention Library (Patient side):
- If validated, the relevant information will be stored in the library in an organized quadruple of: <questions, records, answers, attention requirements>.
- Use as few shots for answer generation.

$$r_a, d = \mathbb{k}(sim(q, \mathcal{L})) \quad (\mathcal{P} \mid r_a, d) \rightarrow SP,$$

Trajectories Library (Doctor side):
- We validate and store high-quality dialogues series as a prediction-trajectories.

$$t_i = \{(q_{j-1}, a_{j-1}, q_j, a_j) \mid q \in \mathcal{Q}, a \in \mathcal{A}\}$$

By effectively using these libraries, we succesfully standardized agents in our framework.



Result

Overall Analysis

Overall performance of simulated SP methods

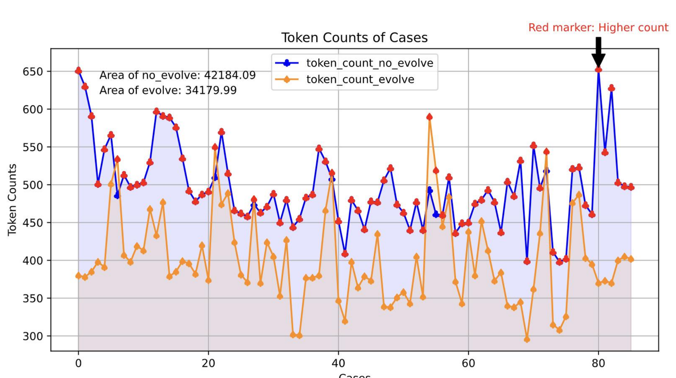
Method	Paradigm	Relevance	Faithfulness	Robustness	Ability
CoT		0.7157†	0.5571†	0.6714†	0.6481†
CoT-SC (3)		0.7337†	0.6123†	0.7002†	0.6821†
ToT		0.7469†	0.7143†	0.7714†	0.7442†
Self-Align		0.7205†	0.7273†	0.8148†	0.7542†
Few-shot (2)		0.7252†	0.7419†	0.8207†	0.7626†
Online Library		0.6903	0.7372†	0.7624†	0.7300†
EvoPatient		**0.7589**	**0.8786**	**0.9412**	**0.8597**

State-of-the-art performance

Good reliability on various question

Computational Analysis

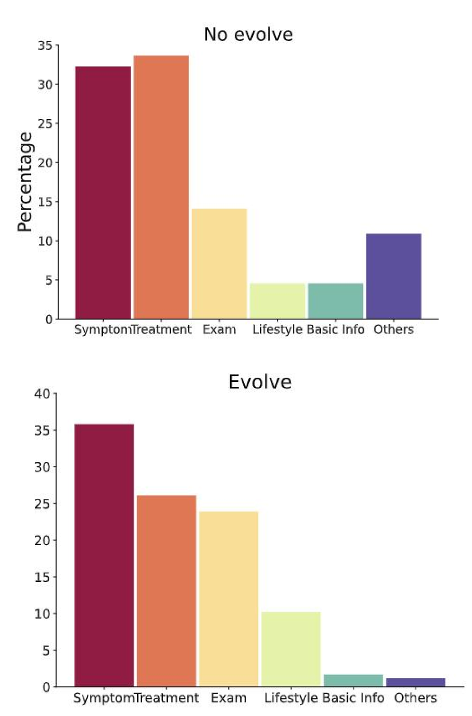
Method	Duration (s)	#Tokens	#Words
CoT	04.7500	0782.0571	45.7429
CoT-SC (3)	12.5559	5837.0286	49.8667
ToT	21.7040	2679.3428	38.9143
Self-Align	09.5146	1307.9435	51.0636
Few-shot (2)	**04.7182**	0959.4355	35.6334
(50) cases	06.7808	0445.3482	36.5571
EvoPatient	06.6922	**0401.5882**	**32.2432**
Δ compared to CoT	↑01.9422	↓0380.4689	↓13.4997



Good balance on computational complexity

Doctor Incremental Study

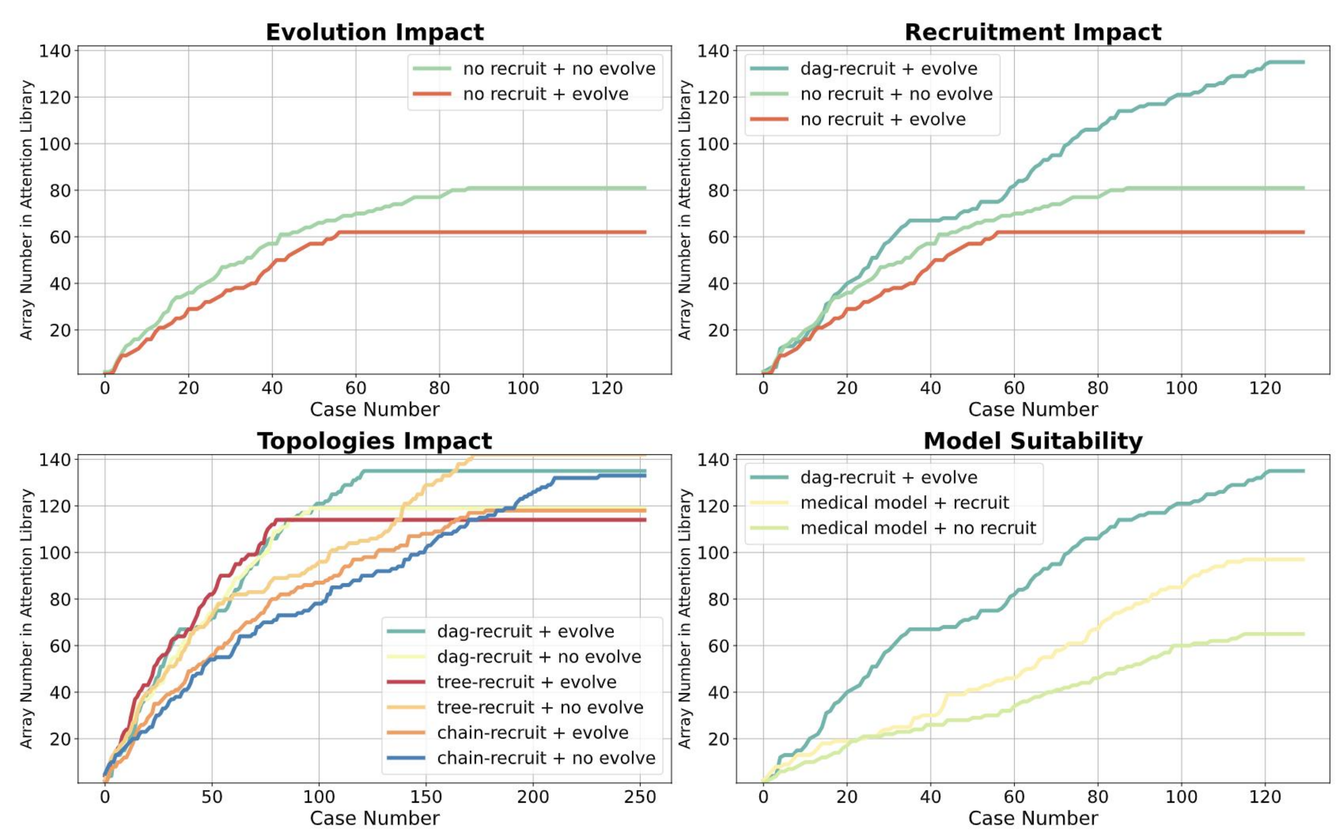
Method	Specificity	Targetedness	Professionalism	Quality
Doctor Agent	0.4713	0.2414	0.4904	0.4010
+ evolve	0.4725	0.2500	0.5650	0.4292
+ pool	0.5825	0.3200	0.5800	0.4942
+ profile	0.4148	0.3215	0.4952	0.4105
+ evolve + pool	0.4659	0.2079	0.7384	0.4707
+ evolve + profile	0.4884	0.3092	0.7023	0.5000
+ pool + profile	0.5925	0.3100	0.6450	0.5158
+ all component	**0.6275**	0.3100	**0.7625**	**0.5667**
Δ compared to Vanilla	+0.1562	+0.0686	+0.2721	+0.1657
Medical model doctor	0.5076	**0.4512**	0.6524	0.5371



Our doctor agent successfully forster patient agent into a SP agent

Discovering Visual Relations

Accumulation rate in the Attention Library



Recruitment improves questions diversity

DAG is a efficient recruitment topology

Method	Relevance	Faithfulness	Robustness	Ability
Doctor agent	0.7297	0.8000	0.8533	0.7943
+ dag-recruit	0.7455	0.8233	0.8733	0.8140
╲ designed recruit	-	-	-	-
╲ memory control	-	-	-	-
+ evolve	0.7311	0.8402	0.9100	0.8271
+ chain-recruit + evolve	0.7405	0.8424	0.8929	0.8253
+ tree-recruit + evolve	0.7488	0.8545	0.9101	0.8378
+ dag-recruit + evolve	**0.7573**	**0.8767**	**0.9333**	**0.8558**
Δ compared to Vanilla	+0.0276	+0.0767	+0.0800	+0.0615

Every components contributes positively

Recruitment in DAG topology boost the evolution process.

Contact us!

xiaodu_flying

duzy@zju.edu.cn
haochaoying@zju.edu.cn

Try our demo!

http://192.168.43.6:7513/

小红书 睿医实验室

https://github.com/ZJUMAI