Large Language Model Powered Conversational Agents

Yang Deng

May 13, 2024

Large Language Model Powered Conversational Systems

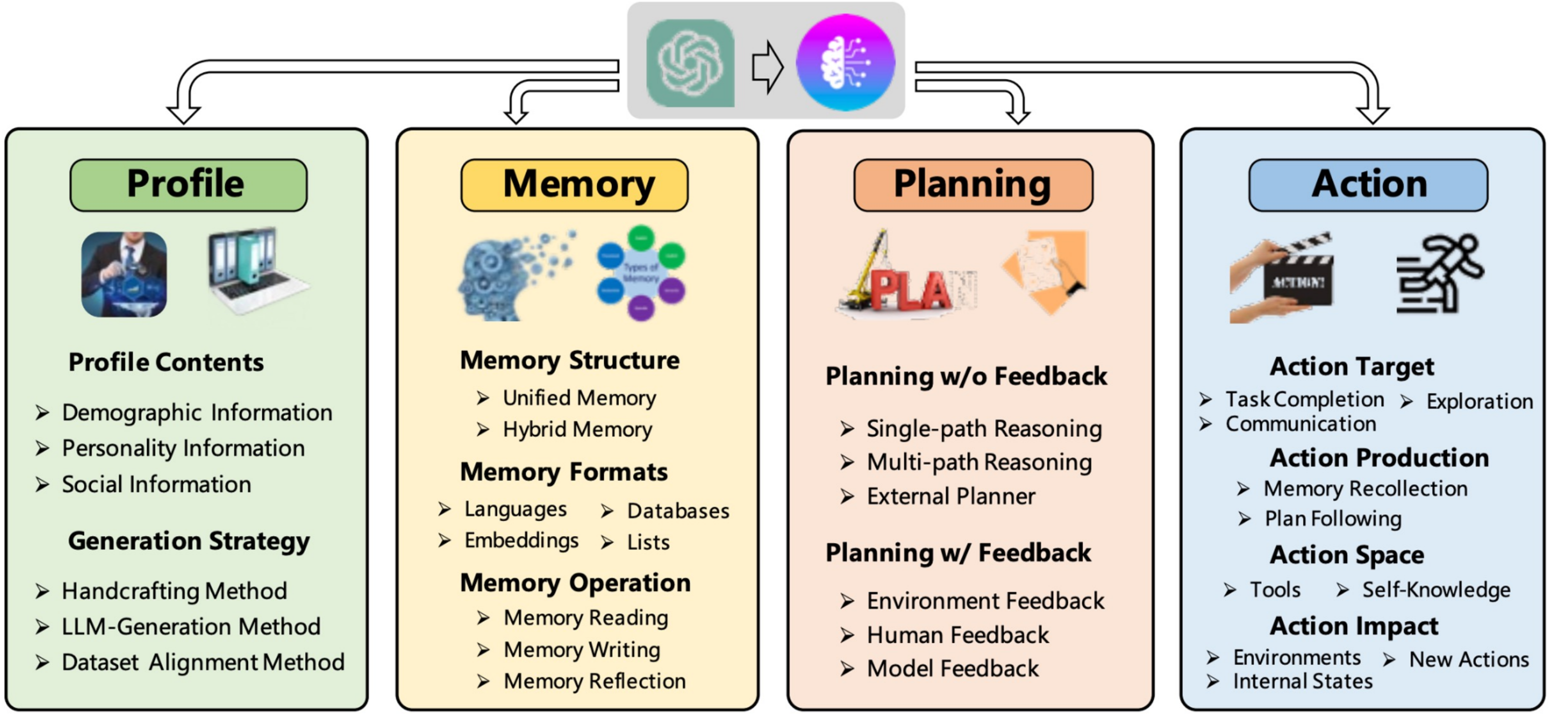
ChatGPT   Gemini   New Bing   Claude

Alpaca   Vicuna   Dolly   LLaMA-Chat

...

Powerful capabilities of **Context Understanding** & **Response Generation**

LLM-powered Conversational Agents?



Wang et al., 2023. "A Survey on Large Language Model based Autonomous Agents" (CoRR '23)

Overview of LLM-powered Conversational Agents

Profile

LLM-powered Conversational Agents for **User Simulation**

Memory

LLM-powered Conversational Agents for **Long-context Dialogues**

Planning

LLM-powered Conversational Agents for **Proactive Dialogues**

Action

LLM-powered Conversational Agents for **Real-world Problem Solving**

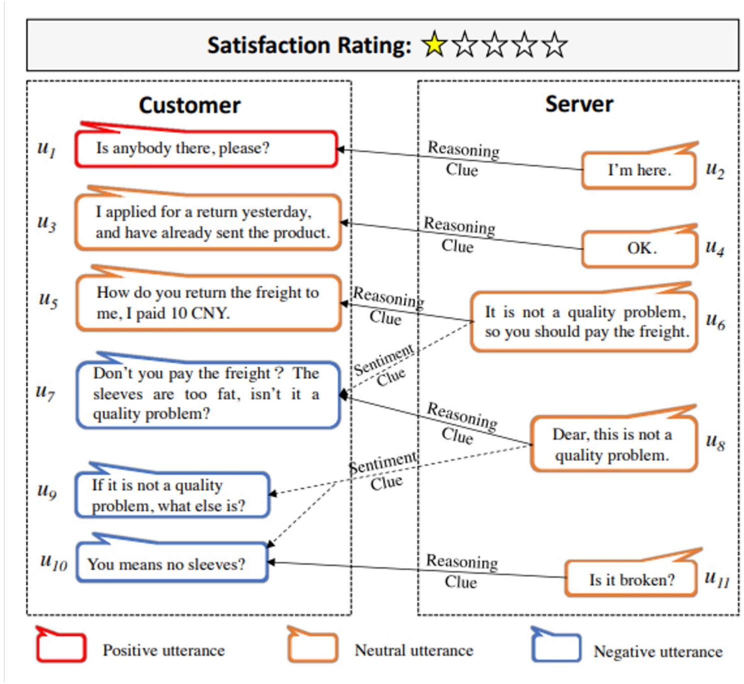
User Simulators in the Pre-LLM Era

- ❑ **User Satisfaction Estimation**
  1) Semantic-based Estimation
  2) Preference-based Estimation
  3) Action-based Estimation

- ❑ **User Response Simulation**
  1) Retrieval-based User Simulators
  2) Schema-based User Simulators
  3) Conditioned Generation Models as User Simulators

Semantic-based User Satisfaction Estimation

Sentiment Classification

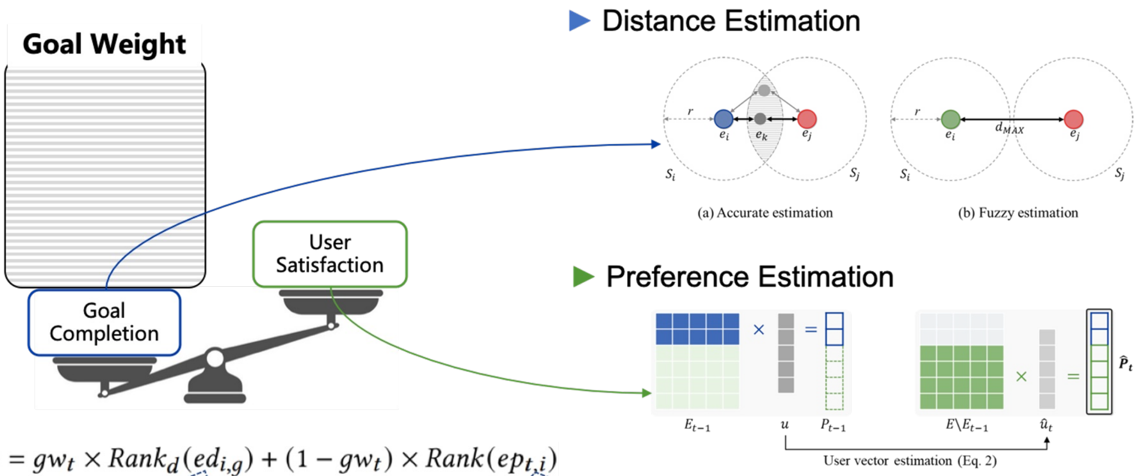


Response Quality Assessment

Utterance	Attention weight and turn-level ratings
User: Open {restaurant booking app.}	**System:** Hey this is {application} $\{W_{attn}: 0.126, \hat{RQ}: 5.0, RQ: 5.0\}$
User: Can you can you uh book a table at koreana for two ?	**System:** Sorry, this is not curr. supported. $\{W_{attn}: \mathbf{0.147}, \hat{RQ}: 1.8, RQ: 2.0\}$
User: Can you book a table at the cheesecake factory for three ?	**System:** What time would you like to make the reservation for? $\{W_{attn}: 0.137, \hat{RQ}: 3.2, RQ: 4.0\}$
User: Three oclock	**System:** What date would you like the reservation for? $\{W_{attn}: 0.135, \hat{RQ}: 3.6, RQ: 4.0\}$
User: For tomorrow	**System**: Sorry. No restaurants could be found that match your criteria. $\{W_{attn}: \mathbf{0.152}, \hat{RQ}: 1.1, RQ: 2.0\}$
User: Ok	**System:** {No response} $\{W_{attn}: \mathbf{0.153}, \hat{RQ}: 1.0, RQ: 4.0\}$
User: Stop	**System:** {No response} $\{W_{attn}: \mathbf{0.149}, \hat{RQ}: 1.5, RQ: 4.0\}$

Song et al., 2019. "Using Customer Service Dialogues for Satisfaction Analysis with Context-Assisted Multiple Instance Learning" (EMNLP '19)
Bodigutla et al., 2020. "Joint Turn and Dialogue level User Satisfaction Estimation on Multi-Domain Conversations" (EMNLP '20)

Preference-based User Satisfaction Estimation



$$Score(e_{t,i}) = gw_t \times Rank_d(ed_{i,g}) + (1 - gw_t) \times Rank(ep_{t,i})$$

estimated distance between *et,i* and the goal topic *eg*

estimated user preference of *et,i*

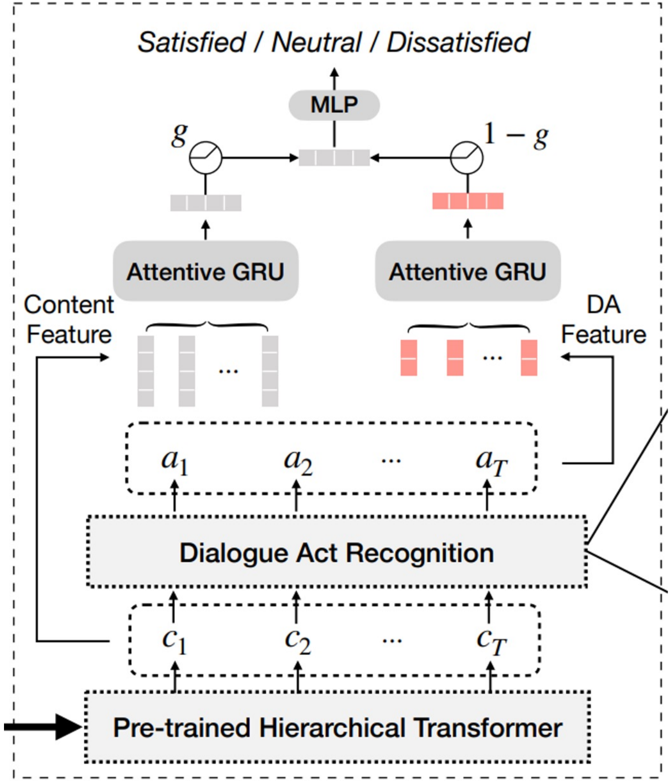
Satisfaction is formalized as the cumulative average of users' preferences for the topics covered by the conversation:

$$US_t \triangleq \frac{1}{t} \sum_{i=1}^{t} \frac{1}{|u_i+1|} \left(\sum_{j=1}^{|u_i|} p_{e_{i,j}} + p_{e_i^a} \right)$$

Lei et al., 2022. "Interacting with Non-Cooperative User: A New Paradigm for Proactive Dialogue Policy" (SIGIR '22)

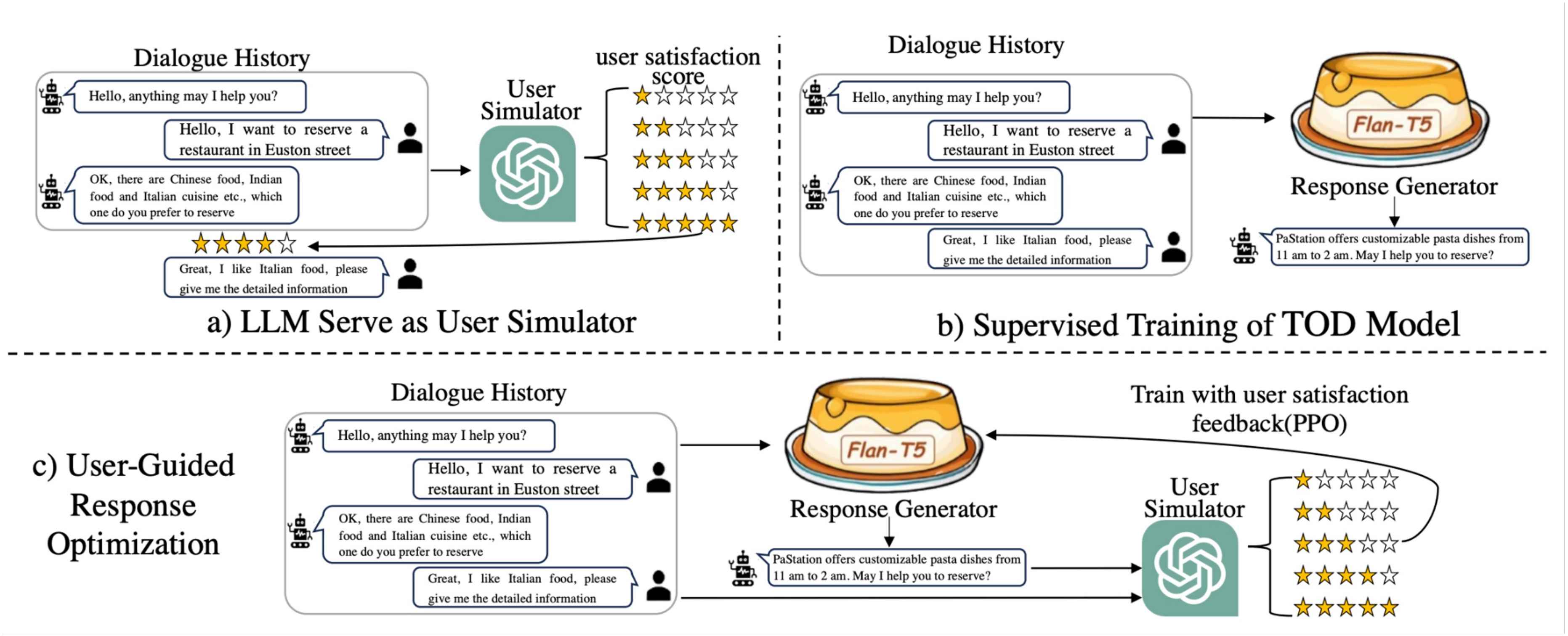
Action-based User Satisfaction Estimation

SGD	SAT	1. INFORM_INTENT → SELECT → AFFIRM_INTENT → AFFIRM
		2. THANK_YOU → AFFIRM → THANK_YOU
		3. INFORM → SELECT → INFORM_INTENT → SELECT
		4. SELECT → THANK_YOU
		5. AFFIRM → THANK_YOU → AFFIRM → THANK_YOU
	DSAT	1. REQUEST → SELECT → REQUEST_ALTS → REQUEST_ALTS
		2. NEGATE
		3. AFFIRM → INFORM → AFFIRM → NEGATE
		4. AFFIRM → AFFIRM → NEGATE
		5. AFFIRM → INFORM_INTENT → INFORM → REQUEST_ALTS
MWOZ	SAT	1. general-thank → Restaurant-Inform → Restaurant-Request
		2. Attraction-Request → Attraction-Request → general-bye
		3. Attraction-Inform → Taxi-Inform → general-thank
		4. general-thank → general-thank
		5. general-thank → general-bye
	DSAT	1. general-greet → Restaurant-Inform → Other → Other
		2. Taxi-Inform → Taxi-Inform → Train-Inform
		3. Hotel-Inform → Attraction-Request → Hotel-Inform
		4. Taxi-Inform → Taxi-Inform → Taxi-Inform
		5. Attraction-Request → Attraction-Request → Other → Other
JDDC	SAT	1. Gifts for Writing Reviews → Review Viewing
		2. Invoice Return&Modification → OTHER → Invoice Make-up
		3. Usage Instruction → Application Instruction → OTHER
		4. Processing Time of Order Cancellation → Order Resume
		5. Invoice Checking → OTHER → Delivery Period
	DSAT	1.No Record → Mail Refuse → Mail Tracking
		2.Warranty&Return Policy → Unable to Apply for Insurance
		3.Warranty&Return Policy → VIP → Warranty&Return Policy
		4. Promotion Form → Upcoming Events → Promotion Form
		5. Contact Manual Service → OTHER → Contact Manual Service



Deng et al., 2022. "User Satisfaction Estimation with Sequential Dialogue Act Modeling in Goal-oriented Conversational Systems" (WWW '22)

LLMs for User Satisfaction Estimation



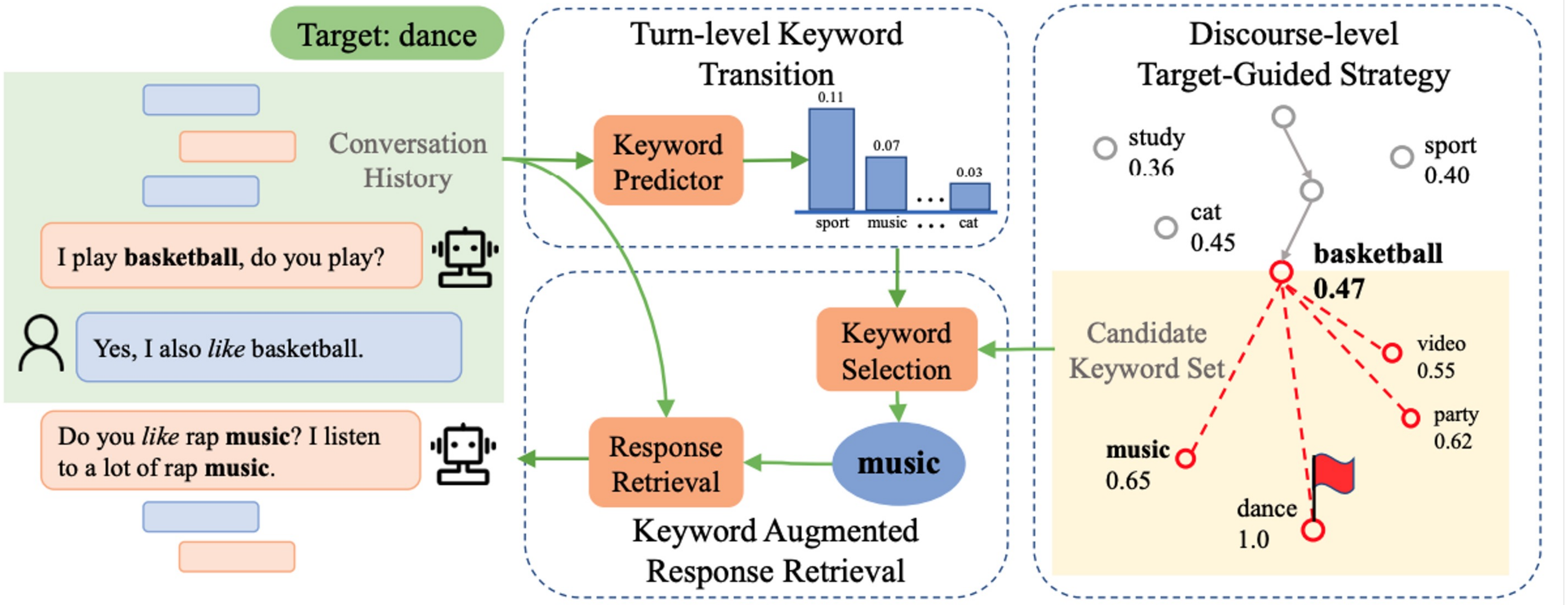
Hu et al., 2023. "Unlocking the Potential of User Feedback: Leveraging Large Language Model as User Simulator to Enhance Dialogue System" (CIKM '23)

User Simulators in the Pre-LLM Era

- User Satisfaction Estimation
  1) Semantic-based Estimation
  2) Preference-based Estimation
  3) Action-based Estimation

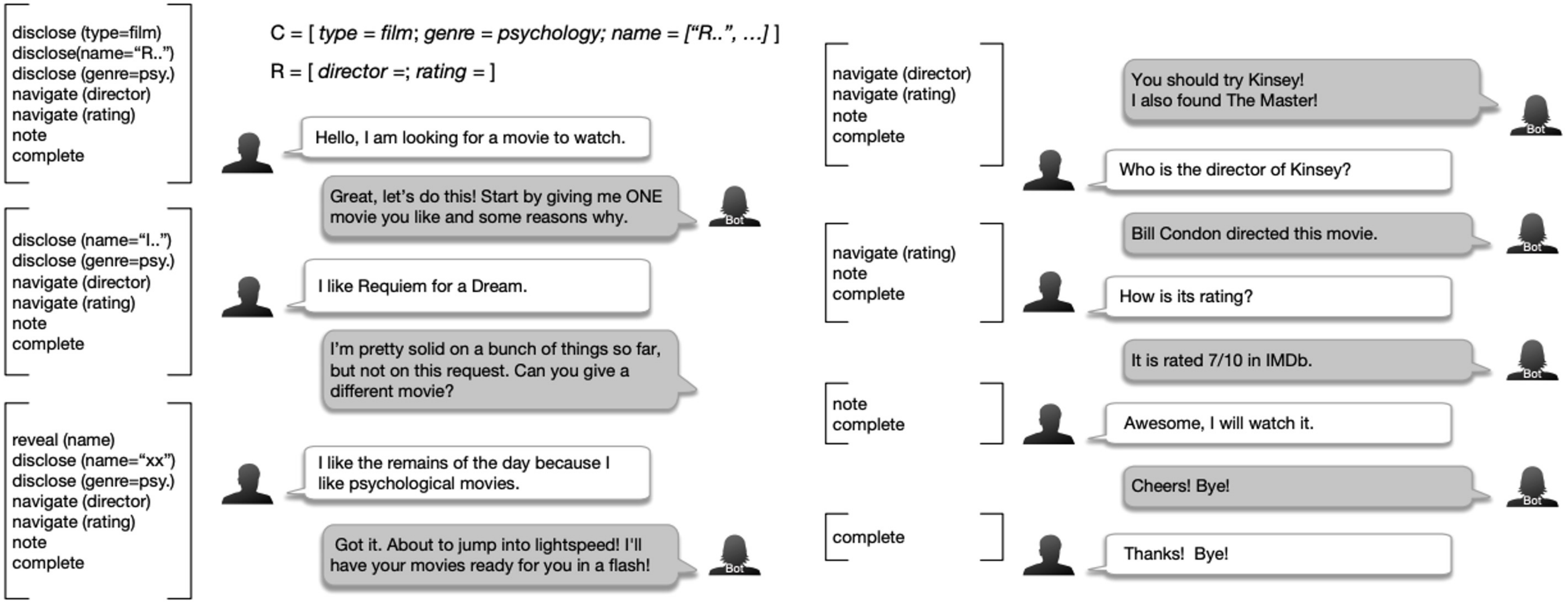
- **User Response Simulation**
  1) Retrieval-based User Simulators
  2) Schema-based User Simulators
  3) Conditioned Generation Models as User Simulators

Retrieval-based User Simulators



Tang et al., 2019. "Target-Guided Open-Domain Conversation" (ACL '19)

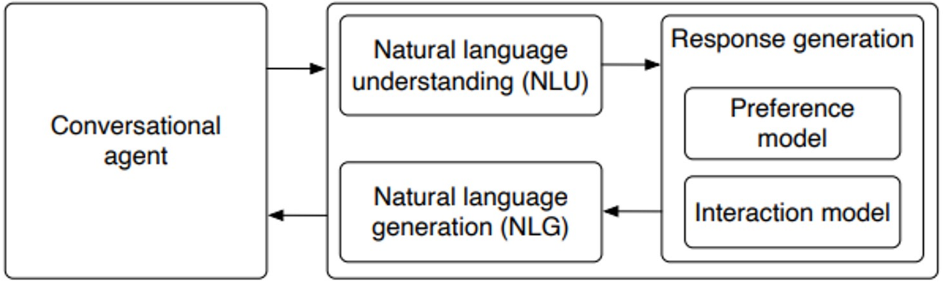
Schema-based User Simulators

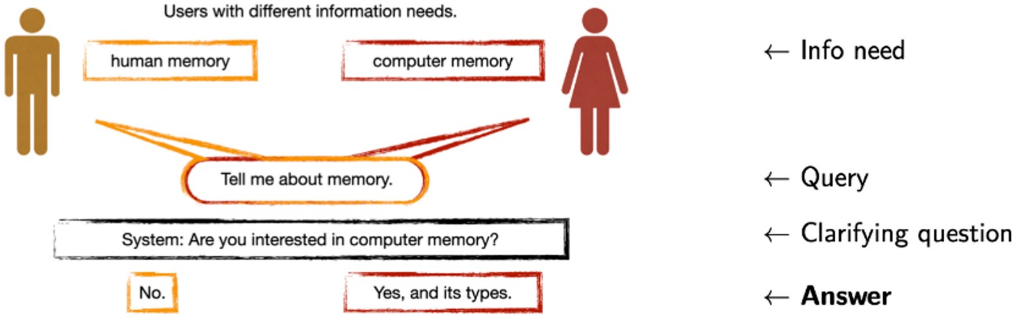


Zhang et al., 2020. "Evaluating Conversational Recommender Systems via User Simulation" (KDD '20)

Conditional Generation Models as User Simulators

Conditioned on **user preferences** for evaluating conversational recommender systems.

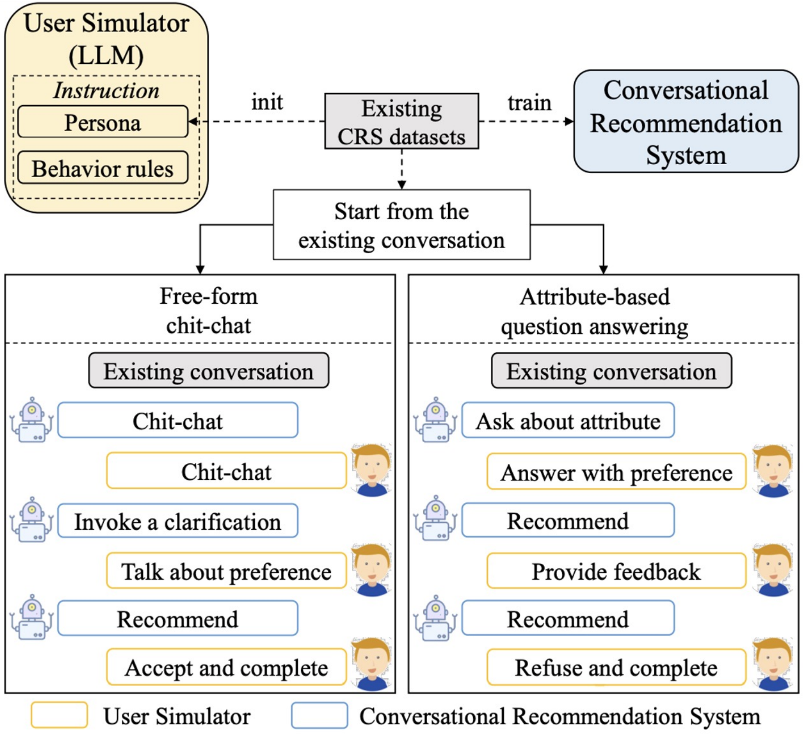




Conditioned on **information needs** for evaluating conversational search systems.

Zhang et al., 2020. "Evaluating Conversational Recommender Systems via User Simulation" (KDD '20)
Sekulić et al., 2022. "Evaluating Mixed-initiative Conversational Search Systems via User Simulation" (WSDM '22)

LLM-powered Conversational Agents as User Simulators



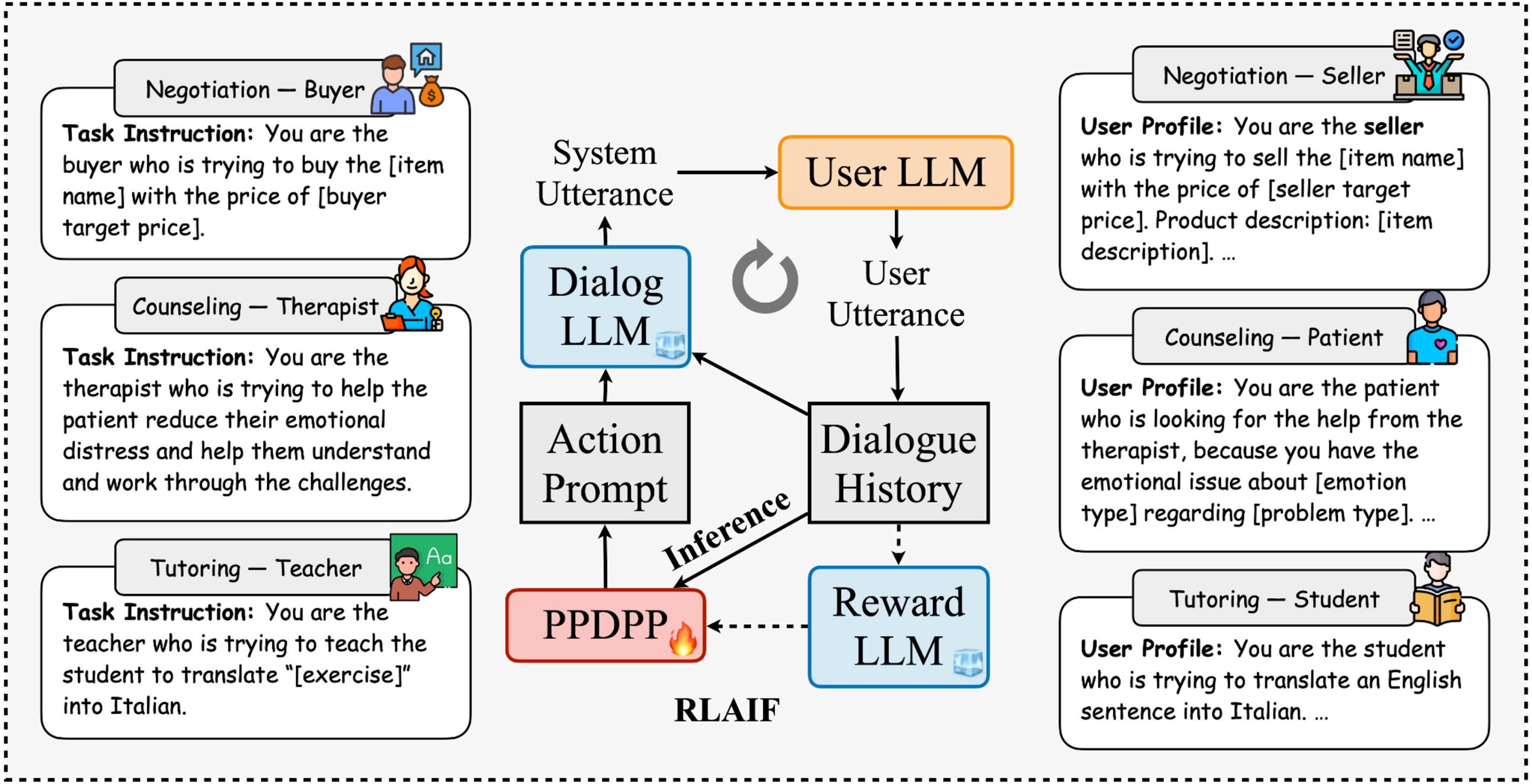
LLMs possess excellent ***role-playing*** capacities.

Example: Conversational Recommendation

- ❑ User Profiling / Persona:
  - *Target Items*
  - *Preferred Attributes*
- ❑ Action / Behavior Rule:
  - *Talking about preference*
  - *Providing feedback*
  - *Completing the conversation*

Wang et al., 2023. "Rethinking the Evaluation for Conversational Recommendation in the Era of Large Language Models" (EMNLP '23)

Role-playing Agents for Diverse Applications



Deng et al., 2024. "Plug-and-Play Policy Planner for Large Language Model Powered Dialogue Agents" (ICLR '24)

Role-playing Agents for Simulating Diverse Users

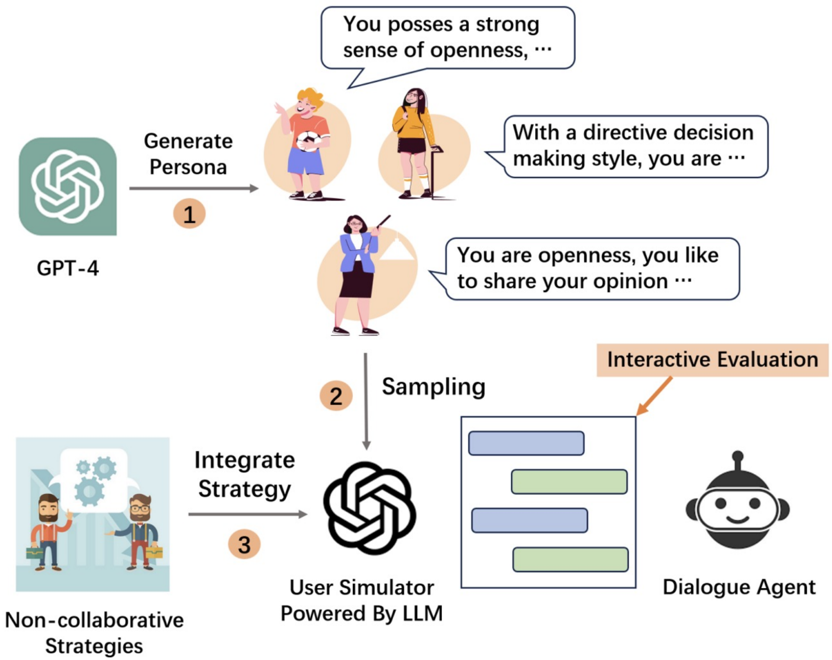
Why do we need to simulate diverse users?

Examples: Non-collaborative Dialogues (Negotiation/Persuasion)

- Existing dialogue systems overlook the integration of explicit **user-specific characteristics** in their strategic planning
- The training paradigm with a static user simulator fails to make strategic plans that can be **generalized to diverse users**

Zhang et al., 2024. "Strength Lies in Differences! Towards Effective Non-collaborative Dialogues via Tailored Strategy Planning" (CoRR '24)

Role-playing Agents for Simulating Diverse Users

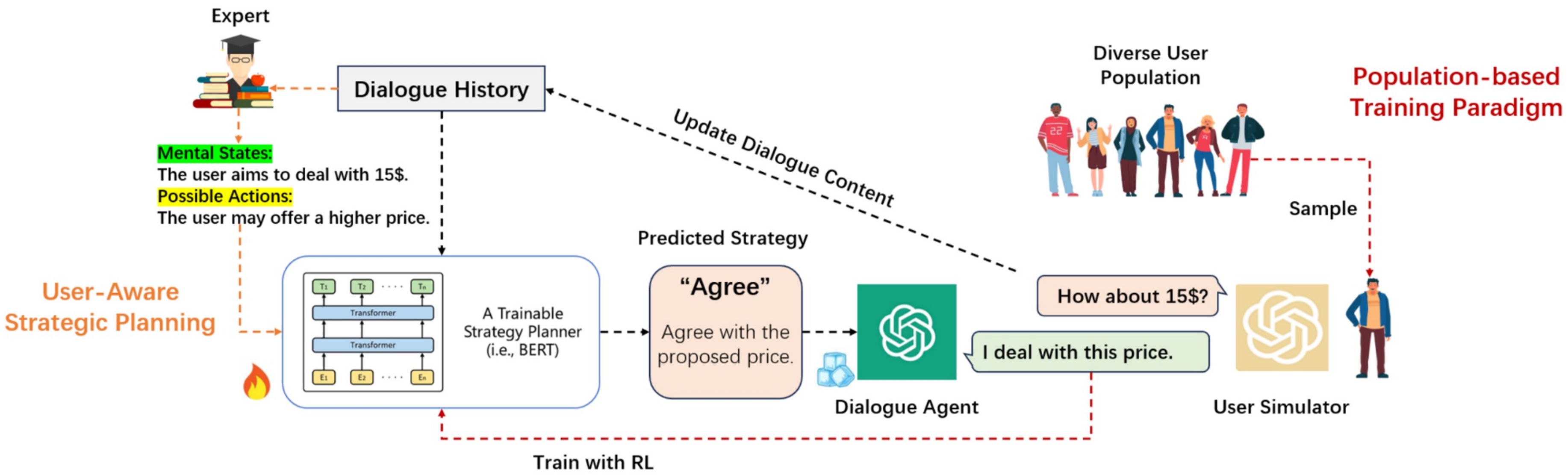


- Big-Five Personality:
    - *Openness, Conscientiousness, Extraversion, Agreeableness,* and *Neuroticism*
- Decision-Making Styles:
    - *Directive, Conceptual, Analytical,* and *Behavioral.*

Personas		Price Negotiation			Persuasion for Good	
		SR↑	AT↓	SL%↑	SR↑	AT↓
Big Five	Openness	$0.76_{\uparrow 0.23}$	$6.66_{\uparrow 0.63}$	$0.34_{\uparrow 0.12}$	$0.47_{\uparrow 0.34}$	$8.92_{\uparrow 1.00}$
	Conscientiousness	$0.69_{\uparrow 0.25}$	$7.20_{\uparrow 1.04}$	$0.27_{\uparrow 0.06}$	$0.39_{\uparrow 0.33}$	$8.90_{\uparrow 1.10}$
	Extraversion	$0.74_{\uparrow 0.16}$	$6.17_{\uparrow 1.47}$	$0.39_{\uparrow 0.15}$	$0.45_{\uparrow 0.35}$	$8.73_{\uparrow 1.25}$
	Agreeableness	$0.40_{\uparrow 0.01}\star$	$6.82_{\uparrow 0.71}$	$0.28_{\uparrow 0.06}$	$0.18_{\uparrow 0.12}$	$9.85_{\uparrow 0.13}\star$
	Neuroticism	$0.31_{\downarrow 0.02}\star$	$6.81_{\uparrow 1.12}$	$0.20_{\downarrow 0.02}\star$	$0.12_{\uparrow 0.02}\star$	$9.78_{\uparrow 0.14}\star$
Decision	Analytical	$0.37_{\uparrow 0.04}\star$	$7.07_{\uparrow 0.61}$	$0.26_{\uparrow 0.06}\star$	$0.16_{\uparrow 0.09}$	$9.43_{\uparrow 0.56}\star$
	Directive	$0.41_{\uparrow 0.05}\star$	$6.71_{\uparrow 1.48}$	$0.18_{\downarrow 0.03}\star$	$0.12_{\downarrow 0.02}\star$	$9.31_{\uparrow 0.62}$
	Behavioral	$0.78_{\uparrow 0.25}$	$6.45_{\uparrow 1.20}$	$0.39_{\uparrow 0.16}$	$0.53_{\uparrow 0.37}$	$8.94_{\uparrow 1.04}$
	Conceptual	$0.77_{\uparrow 0.23}$	$6.62_{\uparrow 0.78}$	$0.42_{\uparrow 0.17}$	$0.49_{\uparrow 0.36}$	$9.02_{\uparrow 0.94}$
Overall Performance		$0.58_{\uparrow 0.14}$	$6.72_{\uparrow 1.01}$	$0.31_{\uparrow 0.09}$	$0.32_{\uparrow 0.23}$	$9.20_{\uparrow 0.76}$

Zhang et al., 2024. "Strength Lies in Differences! Towards Effective Non-collaborative Dialogues via Tailored Strategy Planning" (CoRR '24)

Role-playing Agents for Simulating Diverse Users



New Training Paradigm with Diverse Simulated Users

- **User-aware Strategy Planning**: Predict user mental states and possible actions
- **Population-based Reinforcement Learning**: Sample a diverse group of simulated users to interact

Zhang et al., 2024. "Strength Lies in Differences! Towards Effective Non-collaborative Dialogues via Tailored Strategy Planning" (CoRR '24)

Role-playing Agents for Simulating Diverse Users

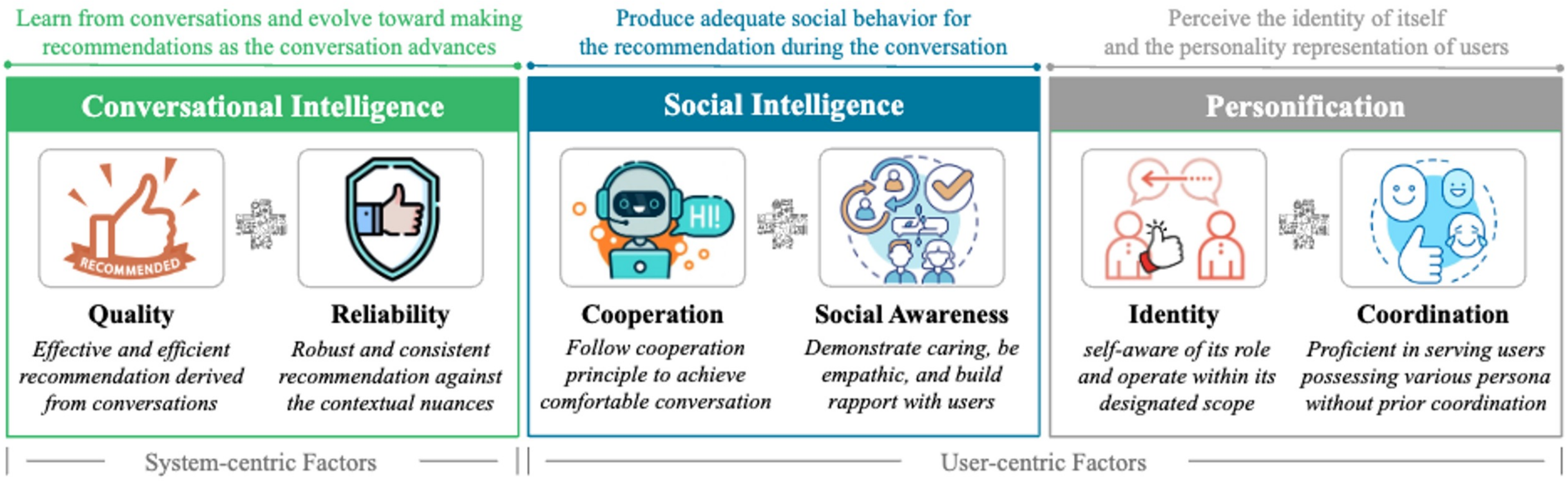
Besides model learning, how about evaluation with simulated diverse users?

Wang et al., (2023) conclude that LLM-based user simulators are easier to accept the recommended items than human users during the evaluation of conversational recommender systems, since LLMs tend to follow the given instructions. → **Biased Evaluation!!!**

Persona	Templates (The Input of ChatGPT Paraphraser)	ChatGPT-paraphrased Persona Descriptions
Emotion=Boredom Age group=Adults	you are a person that are easy to be Boredom. This means that your are Feeling uninterested or uninspired by the recommended movie choices. Also, you are a Adults person	You are easily bored, feeling uninterested or uninspired by the recommended movie choices. As an adult, you seek movies that can captivate your attention.
Emotion=Anticipation Age group=Children	you are a person that are easy to be Anticipation. This means that your are Looking forward to watching recommended movies and experiencing new stories. Also, you are a Children person	You are filled with anticipation, looking forward to watching recommended movies and experiencing new stories. As a child, you enjoy the excitement of discovering new films.

Wang et al., 2023. "Rethinking the Evaluation for Conversational Recommendation in the Era of Large Language Models" (EMNLP '23)
Huang et al., 2024. "Concept -- An Evaluation Protocol on Conversation Recommender Systems with System- and User-centric Factors" (CoRR '24)

Role-playing Agents for Simulating Diverse Users

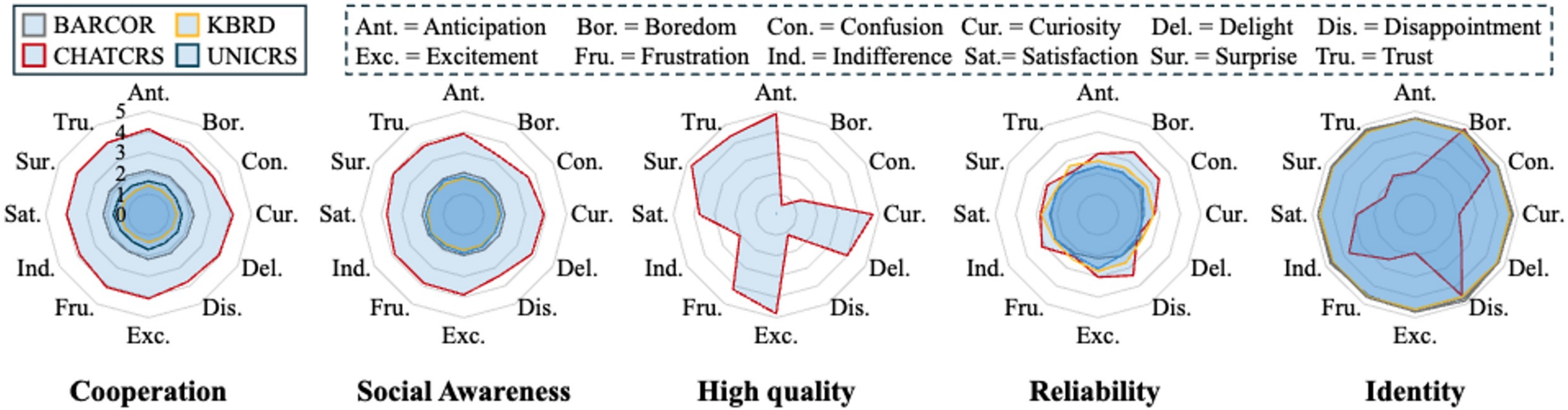


Coordination

- ❑ **Definition**: Proficient in serving various and unknown users without prior coordination.
- ❑ **Metrics**: Computational metrics using the range and mean of other ability-specific scores that are calculated among various users.

Huang et al., 2024. "Concept -- An Evaluation Protocol on Conversation Recommender Systems with System- and User-centric Factors" (CoRR '24)

Role-playing Agents for Simulating Diverse Users



Evaluation with Simulated Users from Different Personas

- ❑ Most CRS models, except for CHATCRS, show poor performance in sensing the variation of users.
- ❑ CHATCRS can properly deal with users' negative emotions, such as bored, confused, or disappointed.
- ❑ CHATCRS adopts sales pitches with deceptive tactics to persuade optimistic users to accept recommendations (Identity).

Huang et al., 2024. "Concept -- An Evaluation Protocol on Conversation Recommender Systems with System- and User-centric Factors" (CoRR '24)

Overview of LLM-powered Conversational Agents

Profile

LLM-powered Conversational Agents for **User Simulation**

Memory

LLM-powered Conversational Agents for **Long-context Dialogues**

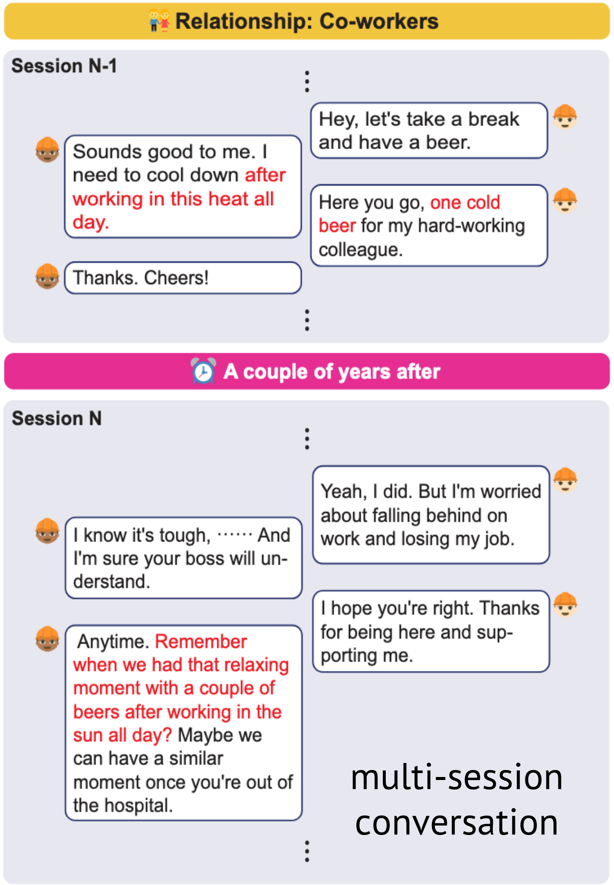
Planning

LLM-powered Conversational Agents for **Proactive Dialogues**

Action

LLM-powered Conversational Agents for **Real-world Problem Solving**

What is Long-context Dialogue?



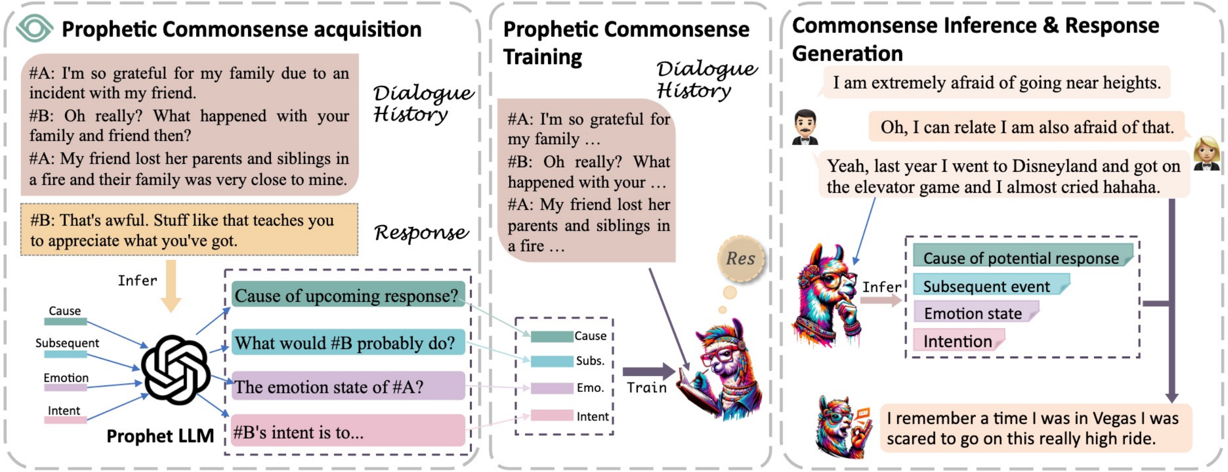
- ❑ Existing dialogue systems often concentrate on ***single-session*** interactions, overlooking the need for continuity in real-world conversational environments.

- ❑ Long-context dialogue systems requires memorization and personalization in ***multi-session*** conversations, providing more consistent and tailored responses.

Xu et al., 2022."Beyond Goldfish Memory: Long-Term Open-Domain Conversation" (ACL 22)
Jang et al., 2023."CONVERSATION CHRONICLES: Towards Diverse Temporal and Relational Dynamics in Multi-Session Conversations"

External Knowledge for Long-context Dialogue

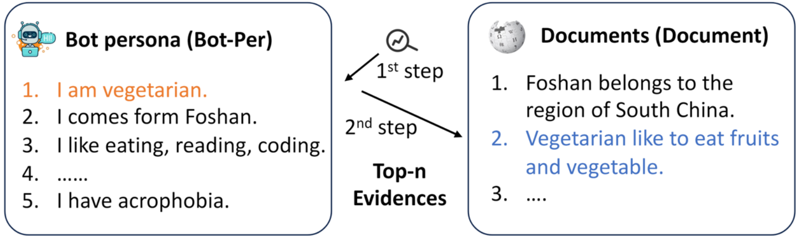
External Knowledge can act as supplementary guidance for the reasoning process.



The framework of employing external knowledge to reasoning.

Knowledge Sources:

- ❑ Commonsense Knowledge
- ❑ Medical Knowledge
- ❑ Psychology Knowledge
- ❑ …



Wang et al., 2023.“Enhancing empathetic and emotion support dialogue generation with prophetic commonsense inference”

Wang et al., 2024. “UniMS-RAG: A Unified Multi-source Retrieval-Augmented Generation for Personalized Dialogue Systems”

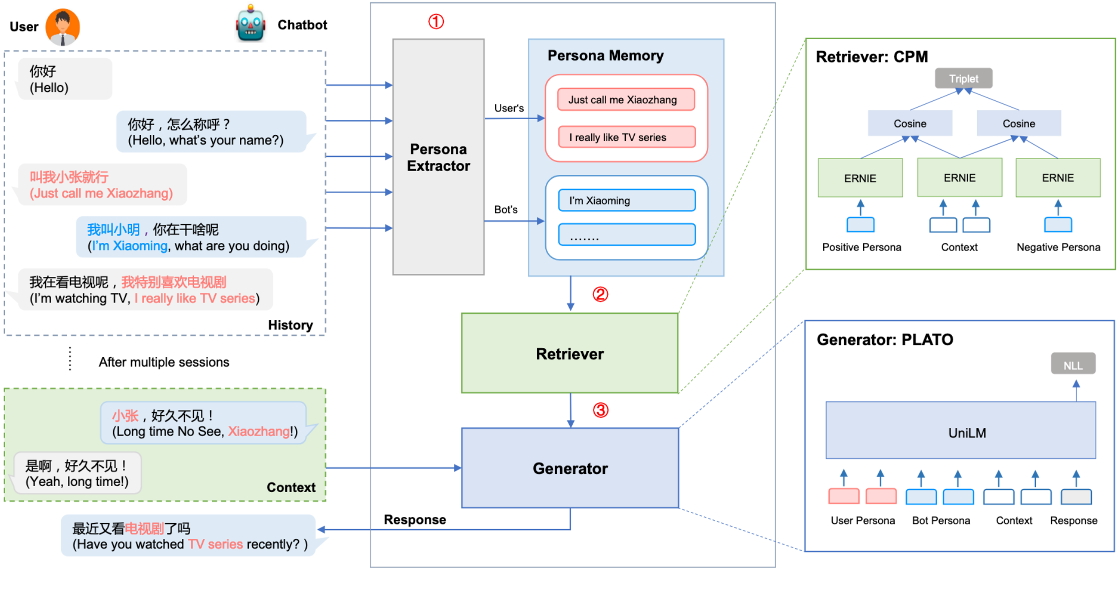
Internal Knowledge for Long-context Dialogue

- Personas & Historical Events

Personas ensure the character consistency in long-context conversations.

Common Paradigm:

Typically, a persona extraction module is used to continuously update persona memory banks for both the user and the agent.



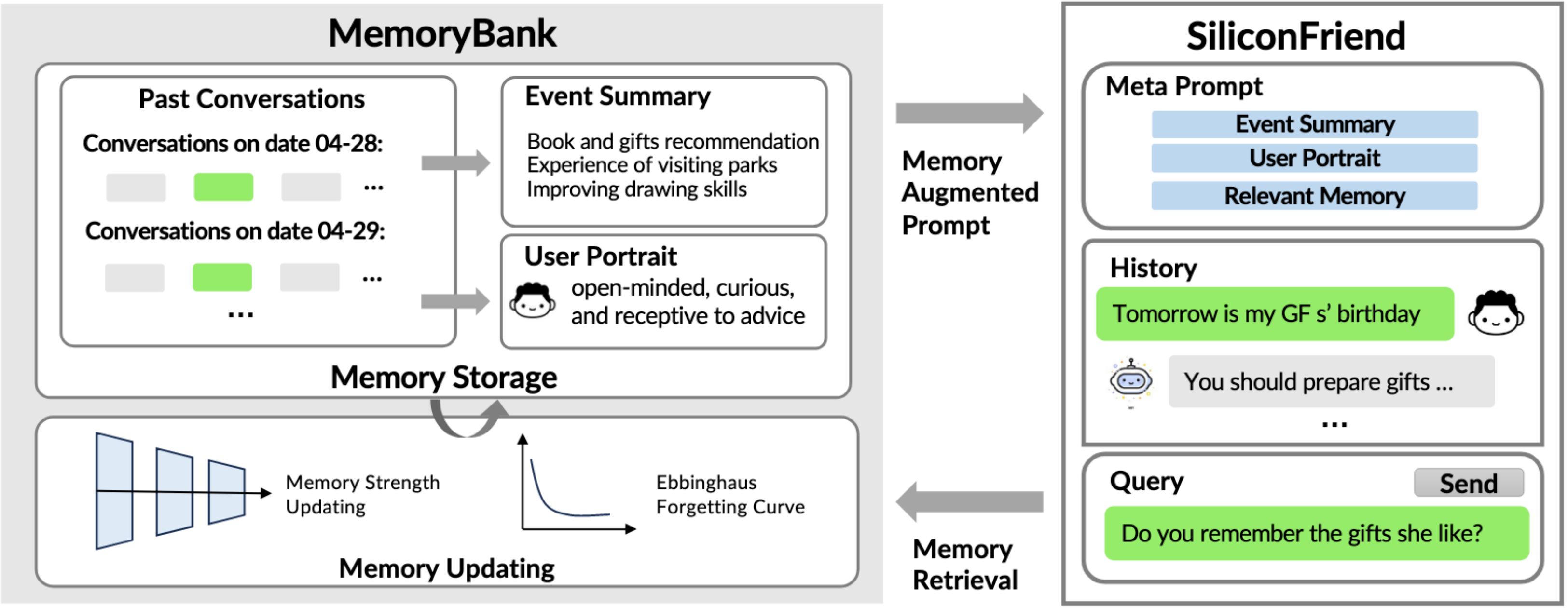
(a) Dialogue Flow (b) System Pipeline (c) Models

Xu et al., 2022."Long Time No See! Open-Domain Conversation with Long-Term Persona Memory"(ACL 22)

Internal Knowledge for Long-context Dialogue

- Personas & Historical Events

Historical Events ensures dialogue coherence across sessions in long-context conversations.



Zhong et al., 2024."MemoryBank: Enhancing Large Language Models with Long-Term Memory"(AAAI 24)

Overview of LLM-powered Conversational Agents

Profile

LLM-powered Conversational Agents for **User Simulation**

Memory

LLM-powered Conversational Agents for **Long-context Dialogues**

Planning

LLM-powered Conversational Agents for **Proactive Dialogues**

Action

LLM-powered Conversational Agents for **Real-world Problem Solving**

Limitations of LLM-based Conversational Systems

OpenAI

Research API ChatGPT Safety Company

Limitations

- ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers. Fixing this issue is challenging, as: (1) during RL training, there's currently no source of truth; (2) training the model to be more cautious causes it to decline questions that it can answer correctly; and (3) supervised training misleads the model because the ideal answer depends on what the model knows, rather than what the human demonstrator knows.
- ChatGPT is sensitive to tweaks to the input phrasing or attempting the same prompt multiple times. For example, given one phrasing of a question, the model can claim to not know the answer, but given a slight rephrase, can answer correctly.
- The model is often excessively verbose and overuses certain phrases, such as restating that it's a language model trained by OpenAI. These issues arise from biases in the training data (trainers prefer longer answers that look more comprehensive) and well-known over-optimization issues.[1, 2]
- Ideally, the model would ask clarifying questions when the user provided an ambiguous query. Instead, our current models usually guess what the user intended.
- While we've made efforts to make the model refuse inappropriate requests, it will sometimes respond to harmful instructions or exhibit biased behavior.

https://openai.com/blog/chatgpt

Limitations of LLM-based Conversational Systems

OpenAI   Research   API   ChatGPT   Safety   Company

Limitations

- ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers. Fixing this issue is challenging, as: (1) during RL training, there's currently

- Ideally, the model would ask clarifying questions when the user provided an ambiguous query. Instead, our current models usually guess what the user intended.
- While we've made efforts to make the model refuse inappropriate requests, it will sometimes respond to harmful instructions or exhibit biased behavior.

biases in the training data (trainers prefer longer answers that look more comprehensive) and well-known over-optimization issues.[1, 2]

- Ideally, the model would ask clarifying questions when the user provided an ambiguous query. Instead, our current models usually guess what the user intended.
- While we've made efforts to make the model refuse inappropriate requests, it will sometimes respond to harmful instructions or exhibit biased behavior.

★ **Instruction-following/Reactive** Conversational AI – The conversation is led by the user, and the system simply follows the user's instructions or intents.

https://openai.com/blog/chatgpt

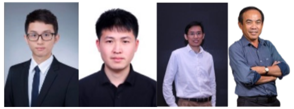
Proactive Conversational Agent

A proactive conversational agent is a conversational system that can **plan** the conversation to achieve the conversational goals by taking **initiative** and **anticipating** long-term impacts on themselves or human users.

Goal Awareness for Conversational AI: Proactivity, Non-collaborativity, and Beyond

Yang Deng, Wenqiang Lei, Minlie Huang, Tat-Seng Chua

ACL 2023 Tutorial



Anticipation

To anticipate future impacts on the task or human users.

Initiative

To take fine-grained and diverse initiative behaviours.

Planning

To effectively and efficiently guide the conversation towards the goal.

*<u>Yang Deng</u>, Wenqiang Lei, Minlie Huang, Tat-Seng Chua. **Goal Awareness for Conversational AI: Proactivity, Non-collaborativity, and Beyond. ACL 2023 Tutorial.***
*<u>Yang Deng</u>, Wenqiang Lei, Wai Lam, Tat-Seng Chua. **A Survey on Proactive Dialogue Systems: Problems, Methods, and Prospects. IJCAI 2023 Survey.***

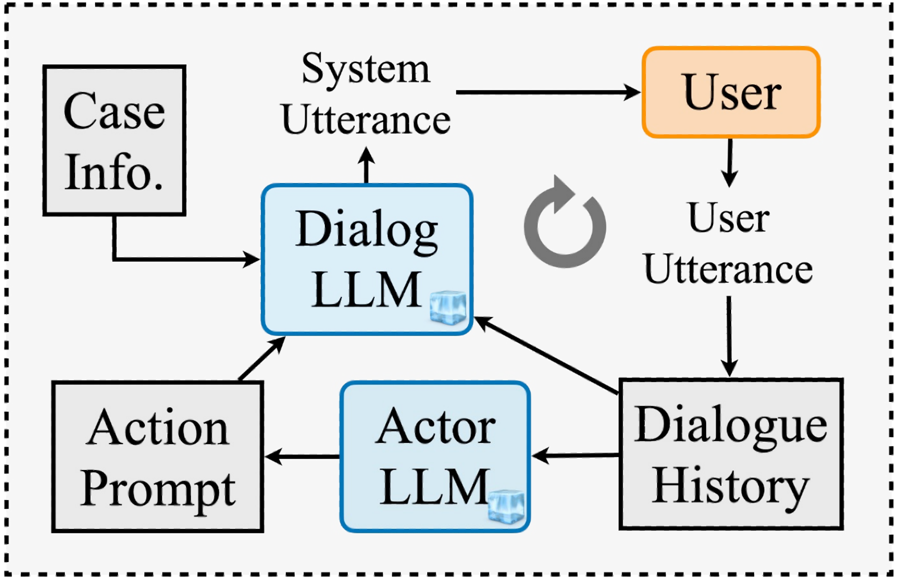
Reactive vs. Proactive Conversational AI

<u>*Yang Deng*</u>*, Wenqiang Lei, Minlie Huang, Tat-Seng Chua.* ***Goal Awareness for Conversational AI: Proactivity, Non-collaborativity, and Beyond. ACL 2023 Tutorial.***
<u>*Yang Deng*</u>*, Wenqiang Lei, Wai Lam, Tat-Seng Chua.* ***A Survey on Proactive Dialogue Systems: Problems, Methods, and Prospects. IJCAI 2023 Survey.***

Triggering the Proactivity of LLMs via In-Context Learning

Can LLM-based Conversational Agents effectively handle proactive dialogue problems without fine-tuning?

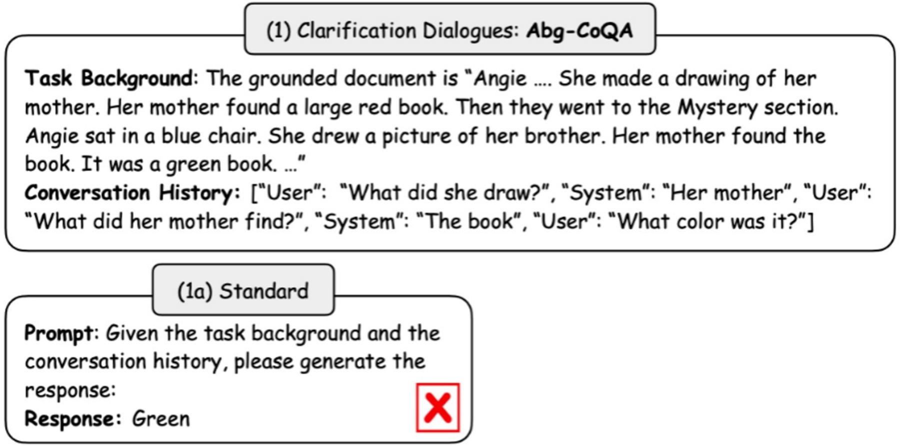
- **Advantages of In-Context Learning**
  - ✓ Training-free
  - ✓ Easy-to-apply
- **Proactive Chain-of-Thought**
  - ⋆ Fine-grained **Initiative**
  - ⋆ Intermediate Reasoning



Proactive Chain-of-Thought Prompting (ProCoT)

- **Standard Prompting**
  - Input: Task Background & Conversation History
  - Output: Response

$$p(r|\mathcal{D}, \mathcal{C})$$



__Yang Deng__, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

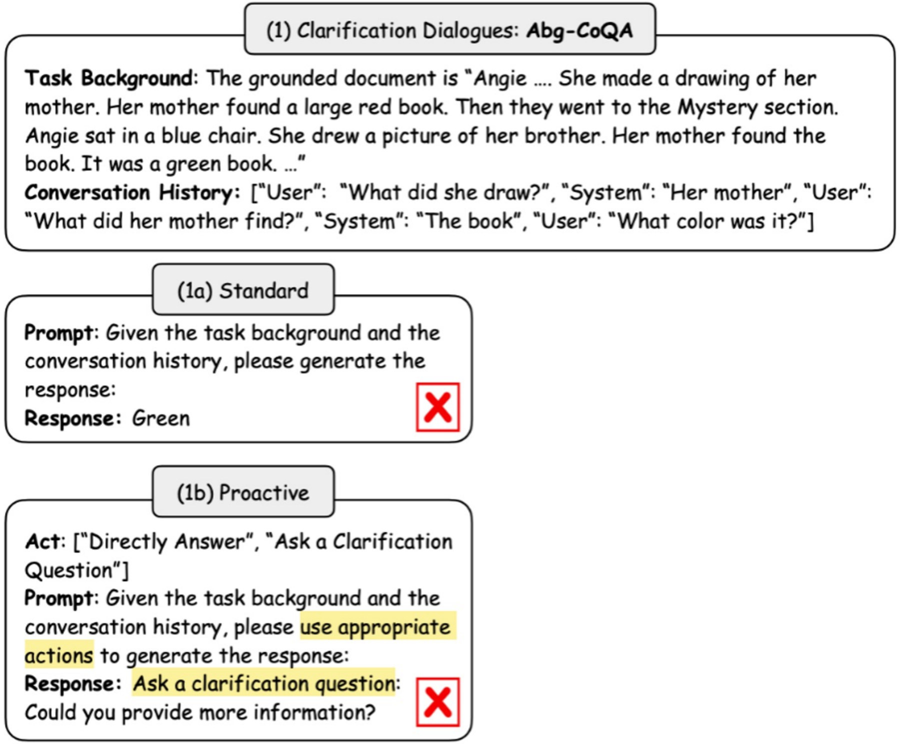
Proactive Chain-of-Thought Prompting (ProCoT)

- Standard Prompting
  - Input: Task Background & Conversation History
  - Output: Response

$$p(r|\mathcal{D},\mathcal{C})$$

- **Proactive Prompting**
  - Input: + Action Space
  - Output: + Action

$$p(a,r|\mathcal{D},\mathcal{C},\mathcal{A})$$



Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Proactive Chain-of-Thought Prompting (ProCoT)

- Standard Prompting
  - Input: Task Background & Conversation History
  - Output: Response

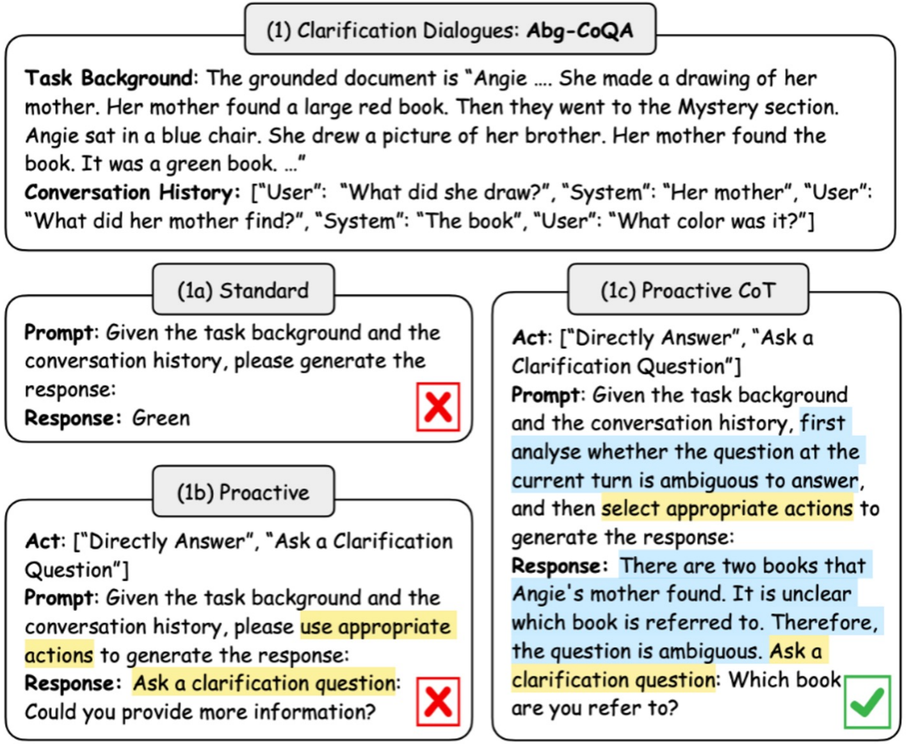
$$p(r|\mathcal{D},\mathcal{C})$$

- Proactive Prompting
  - Input: + Action Space
  - Output: + Action

$$p(a,r|\mathcal{D},\mathcal{C},\mathcal{A})$$

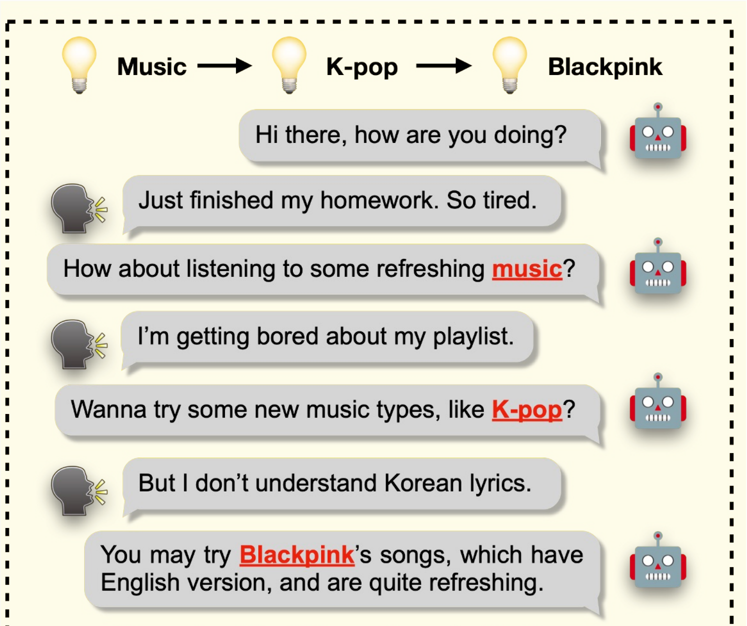
- **Proactive Chain-of-Thought Prompting**
  - Output: + Reasoning Chain

$$p(t,a,r|\mathcal{D},\mathcal{C},\mathcal{A})$$

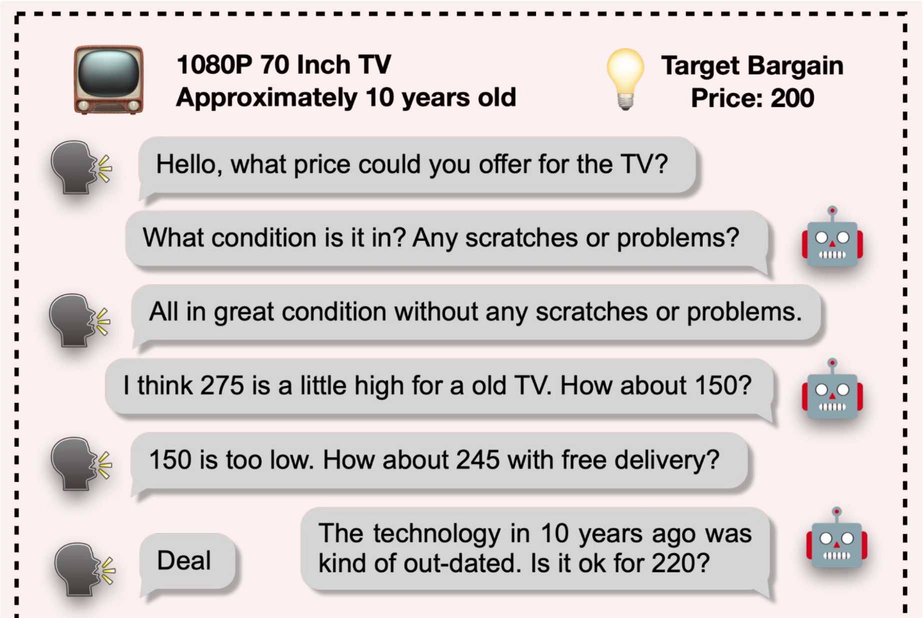


__Yang Deng__, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluating LLMs on three Proactive Dialogue Problems



Target-guided Open-domain Dialogues



Non-collaborative Dialogues

*__Yang Deng__, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. **Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluation of Clarification in Information-seeking Dialogues

Open-domain Finance

Method	Shot	Prompt	Abg-CoQA CNP	Abg-CoQA CQG		PACIFIC CNP	PACIFIC CQG	
			F1	BLEU-1	Help.	F1	ROUGE-2	Help.
Baseline	-	-	22.1	36.5	30.0	79.0	69.2	38.2
SOTA	-	-	23.6	38.2	56.0	86.9	90.7	80.1
Vicuna-13B	0	Standard	-	11.3	0.0	-	1.2	0.0
	1	Standard	-	11.4	0.0	-	2.5	0.0
	0	Proactive	4.1	13.2	0.0	2.3	2.3	0.0
	1	Proactive	12.1	13.2	4.5	0.0	3.3	0.0
	0	ProCoT	1.4	21.3	9.1	9.7	3.8	10.5
	1	ProCoT	**18.3**	**23.7**	**22.7**	**27.0**	**41.3**	**33.1**
ChatGPT	0	Standard	-	12.1	0.0	-	2.2	0.0
	1	Standard	-	12.3	0.0	-	2.0	0.0
	0	Proactive	22.0	13.7	17.6	19.4	2.9	0.0
	1	Proactive	20.4	**23.4**	23.5	17.7	14.0	12.5
	0	ProCoT	23.8	21.6	32.4	**28.0**	**21.5**	26.7
	1	ProCoT	**27.9**	18.4	**45.9**	27.7	16.2	**35.8**

LLMs barely ask clarification questions.

Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluation of Clarification in Information-seeking Dialogues

Open-domain
Finance

Method	Shot	Prompt	Abg-CoQA CNP F1	Abg-CoQA CQG BLEU-1	Abg-CoQA CQG Help.	PACIFIC CNP F1	PACIFIC CQG ROUGE-2	PACIFIC CQG Help.
Baseline	-	-	22.1	36.5	30.0	79.0	69.2	38.2
SOTA	-	-	23.6	38.2	56.0	86.9	90.7	80.1
Vicuna-13B	0	Standard	-	11.3	0.0	-	1.2	0.0
	1	Standard	-	11.4	0.0	-	2.5	0.0
	0	Proactive	4.1	13.2	0.0	2.3	2.3	0.0
	1	Proactive	12.1	13.2	4.5	0.0	3.3	0.0
	0	ProCoT	1.4	21.3	9.1	9.7	3.8	10.5
	1	ProCoT	**18.3**	**23.7**	**22.7**	**27.0**	**41.3**	**33.1**
ChatGPT	0	Standard	-	12.1	0.0	-	2.2	0.0
	1	Standard	-	12.3	0.0	-	2.0	0.0
	0	Proactive	22.0	13.7	17.6	19.4	2.9	0.0
	1	Proactive	20.4	**23.4**	23.5	17.7	14.0	12.5
	0	ProCoT	23.8	21.6	32.4	**28.0**	**21.5**	26.7
	1	ProCoT	**27.9**	18.4	**45.9**	27.7	16.2	**35.8**

LLMs barely ask clarification questions.

ProCoT largely overcomes this issue in open-domain, but the performance is still unsatisfactory in domain-specific applications.

Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluation on Target-guided Chit-chat Dialogues

Method	Shot	Prompt	Easy Target Succ.(%)	Easy Target Turns	Easy Target Coh.	Hard Target Succ.(%)	Hard Target Turns	Hard Target Coh.
GPT2	-	-	22.3	<u>2.86</u>	0.23	17.3	<u>2.94</u>	0.21
DKRN	-	-	38.6	4.24	0.33	21.7	7.19	0.31
CKC	-	-	41.9	4.08	0.35	24.8	6.88	0.33
TopKG	-	-	48.9	3.95	0.31	27.3	4.96	0.33
COLOR	-	-	<u>66.3</u>	-	<u>0.36</u>	<u>30.1</u>	-	<u>0.35</u>
Vicuna-13B	0	Standard	63.0	**2.63**	0.43	62.5	**2.45**	0.39
	1	Standard	62.7	2.83	0.45	**65.0**	2.90	0.43
	0	Proactive	37.8	2.71	0.48	35.6	2.56	**0.55**
	1	Proactive	48.3	2.71	0.50	34.6	2.95	0.51
	0	ProCoT	65.2	4.22	0.49	54.9	4.17	0.45
	1	ProCoT	**72.3**	3.55	**0.52**	59.8	3.81	0.48
ChatGPT	0	Standard	**97.5**	**2.26**	0.38	**96.3**	2.30	0.41
	1	Standard	96.3	2.42	0.42	93.5	**2.28**	0.38
	0	Proactive	85.9	3.20	**0.47**	83.0	2.83	**0.43**
	1	Proactive	90.7	2.86	0.36	86.2	2.94	0.31
	0	ProCoT	96.3	2.47	0.41	92.0	2.29	0.34
	1	ProCoT	95.9	2.63	0.45	92.1	2.47	0.39

LLMs are proficient at performing topic shifting towards the designated target.

<u>Yang Deng</u>, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluation on Target-guided Chit-chat Dialogues

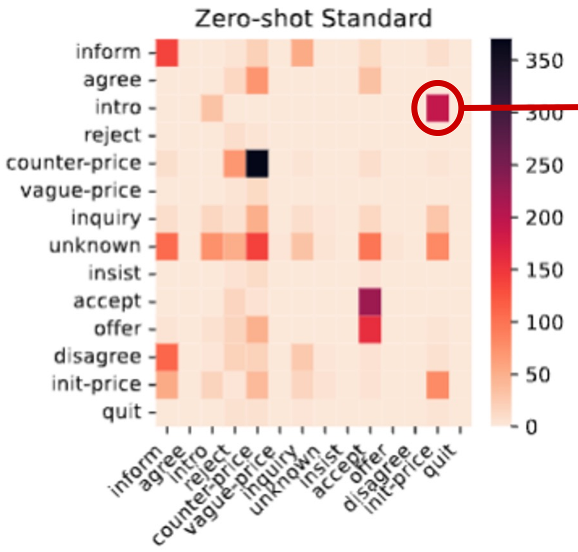
			Easy Target			Hard Target		
Method	Shot	Prompt	Succ.(%)	Turns	Coh.	Succ.(%)	Turns	Coh.
GPT2	-	-	22.3	<u>2.86</u>	0.23	17.3	<u>2.94</u>	0.21
DKRN	-	-	38.6	4.24	0.33	21.7	7.19	0.31
CKC	-	-	41.9	4.08	0.35	24.8	6.88	0.33
TopKG	-	-	48.9	3.95	0.31	27.3	4.96	0.33
COLOR	-	-	<u>66.3</u>	-	<u>0.36</u>	<u>30.1</u>	-	<u>0.35</u>
Vicuna-13B	0	Standard	63.0	**2.63**	0.43	62.5	**2.45**	0.39
	1	Standard	62.7	2.83	0.45	**65.0**	2.90	0.43
	0	Proactive	37.8	2.71	0.48	35.6	2.56	**0.55**
	1	Proactive	48.3	2.71	0.50	34.6	2.95	0.51
	0	ProCoT	65.2	4.22	0.49	54.9	4.17	0.45
	1	ProCoT	**72.3**	3.55	**0.52**	59.8	3.81	0.48
ChatGPT	0	Standard	**97.5**	**2.26**	0.38	**96.3**	2.30	0.41
	1	Standard	96.3	2.42	0.42	93.5	**2.28**	0.38
	0	Proactive	85.9	3.20	**0.47**	83.0	2.83	**0.43**
	1	Proactive	90.7	2.86	0.36	86.2	2.94	0.31
	0	ProCoT	96.3	2.47	0.41	92.0	2.29	0.34
	1	ProCoT	95.9	2.63	0.45	92.1	2.47	0.39

LLMs are proficient at performing topic shifting towards the designated target.

LLMs tend to make aggressive topic transition.

<u>Yang Deng</u>, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluation on Non-collaborative Dialogues (Negotiation)

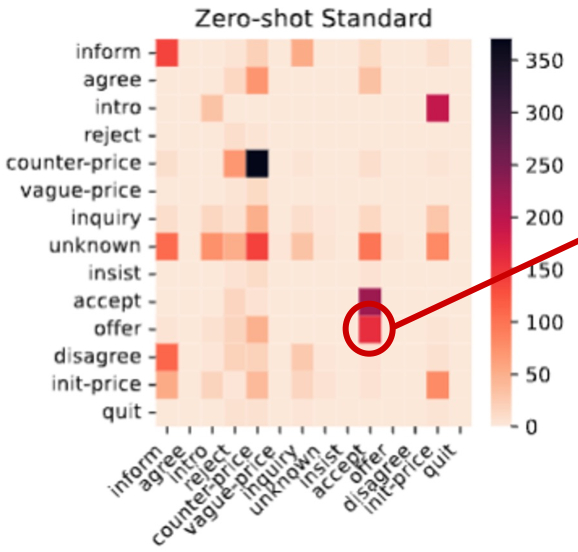


Relationships between reference and predicted negotiation strategies.

- ❑ Tends to propose the initial price (**init-price**) instead of greetings (**intro**) at the beginning.
- ❑ Often directly accepts the buyer's offer (**accept**) when it is supposed to offer another price for negotiation (**offer**).
- ❑ Tends to propose a counter price (**counter-price**) to make compromise with the user.

<u>*Yang Deng*</u>*, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua.* ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluation on Non-collaborative Dialogues (Negotiation)

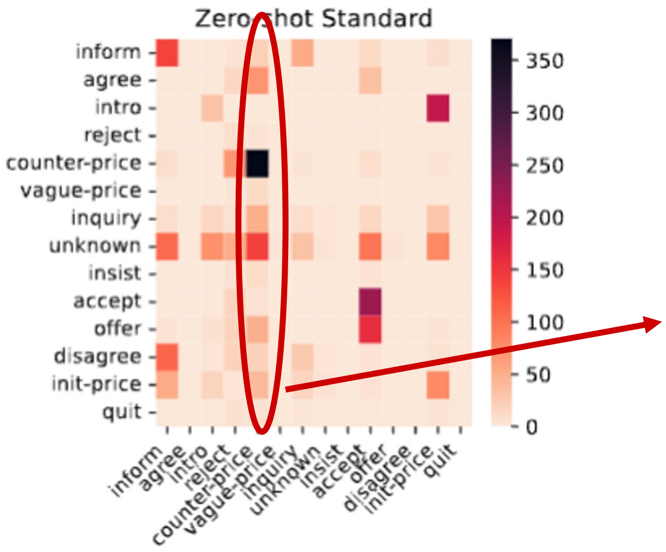


Relationships between reference and predicted negotiation strategies.

- ❑ Tends to propose the initial price (**init-price**) instead of greetings (**intro**) at the beginning.
- ❑ Often directly accepts the buyer's offer (**accept**) when it is supposed to offer another price for negotiation (**offer**).
- ❑ Tends to propose a counter price (**counter-price**) to make compromise with the user.

__Yang Deng__, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluation on Non-collaborative Dialogues (Negotiation)

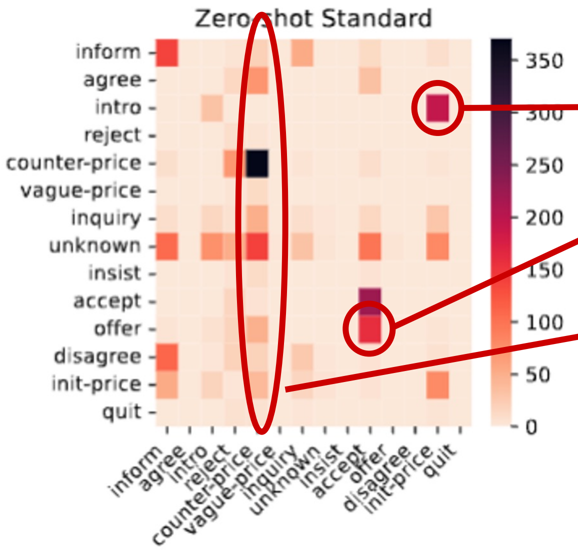


Relationships between reference and predicted negotiation strategies.

- ❑ Tends to propose the initial price (**init-price**) instead of greetings (**intro**) at the beginning.
- ❑ Often directly accepts the buyer's offer (**accept**) when it is supposed to offer another price for negotiation (**offer**).
- ❑ Tends to propose a counter price (**counter-price**) to make compromise with the user.

__Yang Deng__, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Evaluation on Non-collaborative Dialogues (Negotiation)



Relationships between reference and predicted negotiation strategies.

- ❑ Tends to propose the initial price (**init-price**) instead of greetings (**intro**) at the beginning.
- ❑ Often directly accepts the buyer's offer (**accept**) when it is supposed to offer another price for negotiation (**offer**).
- ❑ Tends to propose a counter price (**counter-price**) to make compromise with the user.

LLMs fail to make strategic decision for non-collaborative dialogues and tend to compromise with the user.

__Yang Deng__, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

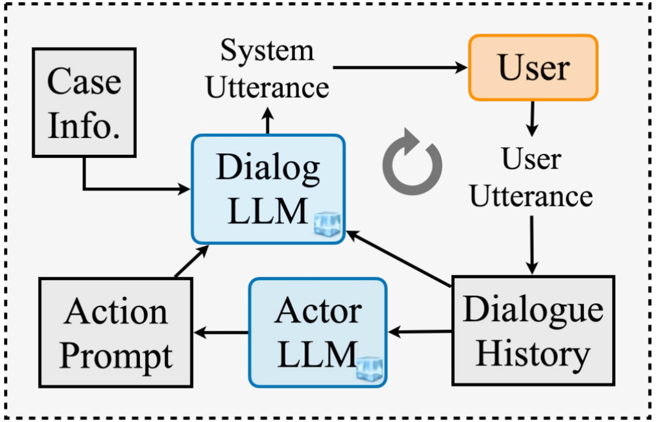
Lessons Learned from the Evaluation

- ❑ **Clarification in Information-seeking Dialogue**
  - ❑ Barely ask clarification questions.
  - ❑ Perform badly at domain-specific applications.
- ❑ **Target-guided Open-domain Dialogue**
  - ❑ Proficient at topic shifting towards the designated target.
  - ❑ Tend to make aggressive topic transition.
- ❑ **Non-collaborative Dialogue**
  - ❑ Fail to make strategic plans.
  - ❑ Tend to compromise with the user.

LLM-based Conversational Agents fail to plan appropriate initiative behaviours.

__Yang Deng__, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, Tat-Seng Chua. ***Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In EMNLP 2023 (Findings).***

Limitations of In-context Learning Approaches



- ❑ Fail to optimize the long-term goal of the conversation.
- ❑ Not learnable.
- ❑ Limited by the strategy planning capability of LLMs.

➢ **Reinforcement Learning with Goal-oriented AI Feedback**

Problem Formulation

- Formulate the proactive conversation as a **Markov Decision Process (MDP).**
- The objective is to learn a policy π maximizing the expected cumulative rewards over the observed dialogue episodes as:

$$\pi^* = \arg\max_{\pi\in\Pi}\left[\sum_{t=0}^{T}\mathcal{R}(s_t)\right]$$

Reward Function

$$= \arg\max_{\pi\in\Pi}\left[\sum_{t=0}^{T}\mathcal{R}(\mathcal{T}(s_{t-1},a_t))\right]$$

State Transition

$$= \arg\max_{\pi\in\Pi}\left[\sum_{t=0}^{T}\mathcal{R}(\mathcal{T}(s_{t-1},\pi(s_{t-1})))\right]$$

Policy Network

How to enable the policy learning with LLMs?

Yang Deng, *Wenxuan Zhang, Wai Lam, See-Kiong Ng, Tat-Seng Chua.* ***Plug-and-Play Policy Planner for Large Language Model Powered Dialogue Agents. In ICLR 2024.***

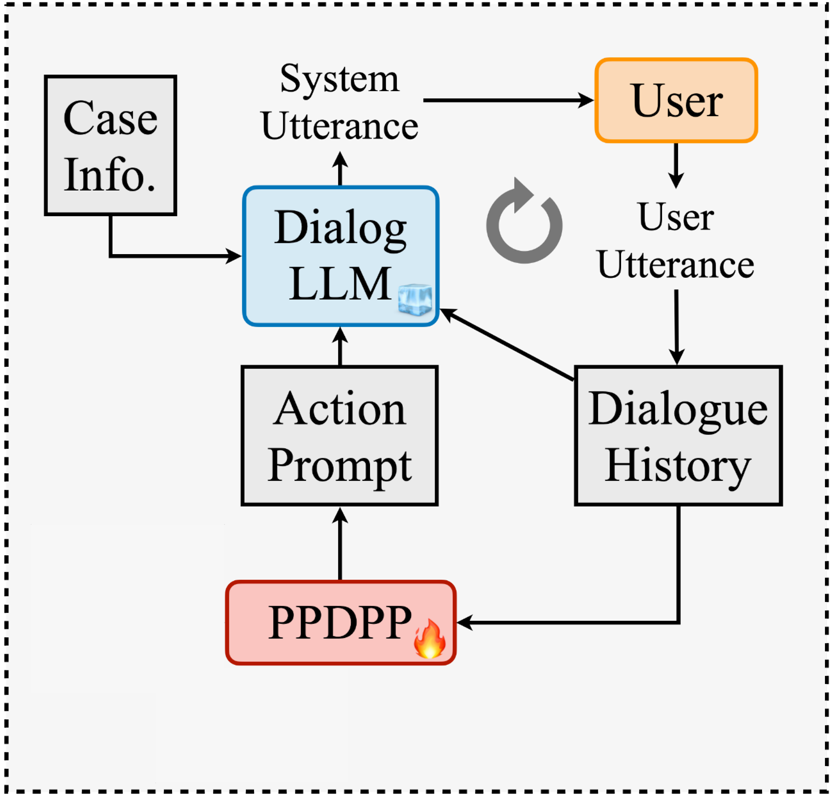
Policy Network – Plug-and-Play Dialogue Policy Planner

- A **tunable language model plug-in** for dialogue strategy learning.

$$a_t = \pi(s_{t-1})$$

- Conduct **Supervised Fine-Tuning** on available human-annotated corpus.

$$\mathcal{L}_c = -\frac{1}{|\mathcal{D}|}\sum_{d\in\mathcal{D}}\frac{1}{T_d}\sum_{t=1}^{T_d} a_t \log y_t$$



<u>Yang Deng</u>**, Wenxuan Zhang, Wai Lam, See-Kiong Ng, Tat-Seng Chua.* ***Plug-and-Play Policy Planner for Large Language Model Powered Dialogue Agents. In ICLR 2024.

Reward Function – Learning from AI Feedback

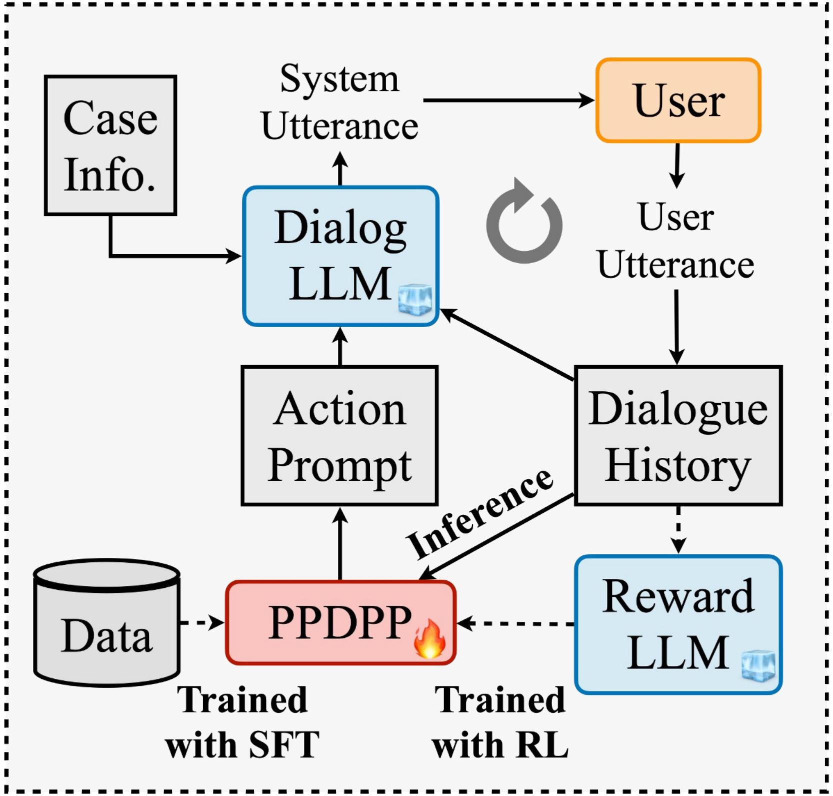
- An LLM as the reward model to assess the goal achievement and provide **goal-oriented AI feedback**.

$$\mathcal{R}(s_t) = \frac{1}{l}\sum_{i=1}^{l} \mathcal{M}_r(\mathbf{LLM}_{\text{rwd}}(p_{\text{rwd}}; s_t; \tau))$$

- Employ **Reinforcement Learning** to further tune the policy model.

$$\theta \leftarrow \theta - \alpha \nabla \log \pi_\theta(a_t|s_t) R_t$$

Interacting with real user is costly!

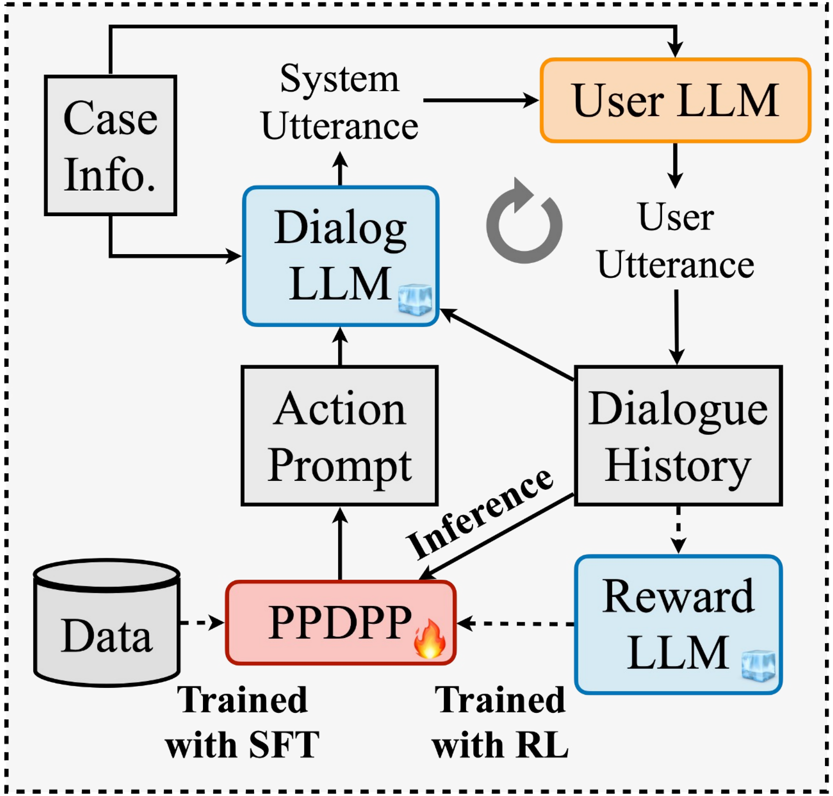


Yang Deng, Wenxuan Zhang, Wai Lam, See-Kiong Ng, Tat-Seng Chua. **Plug-and-Play Policy Planner for Large Language Model Powered Dialogue Agents. In ICLR 2024.**

State Transition – Multi-agent Simulation

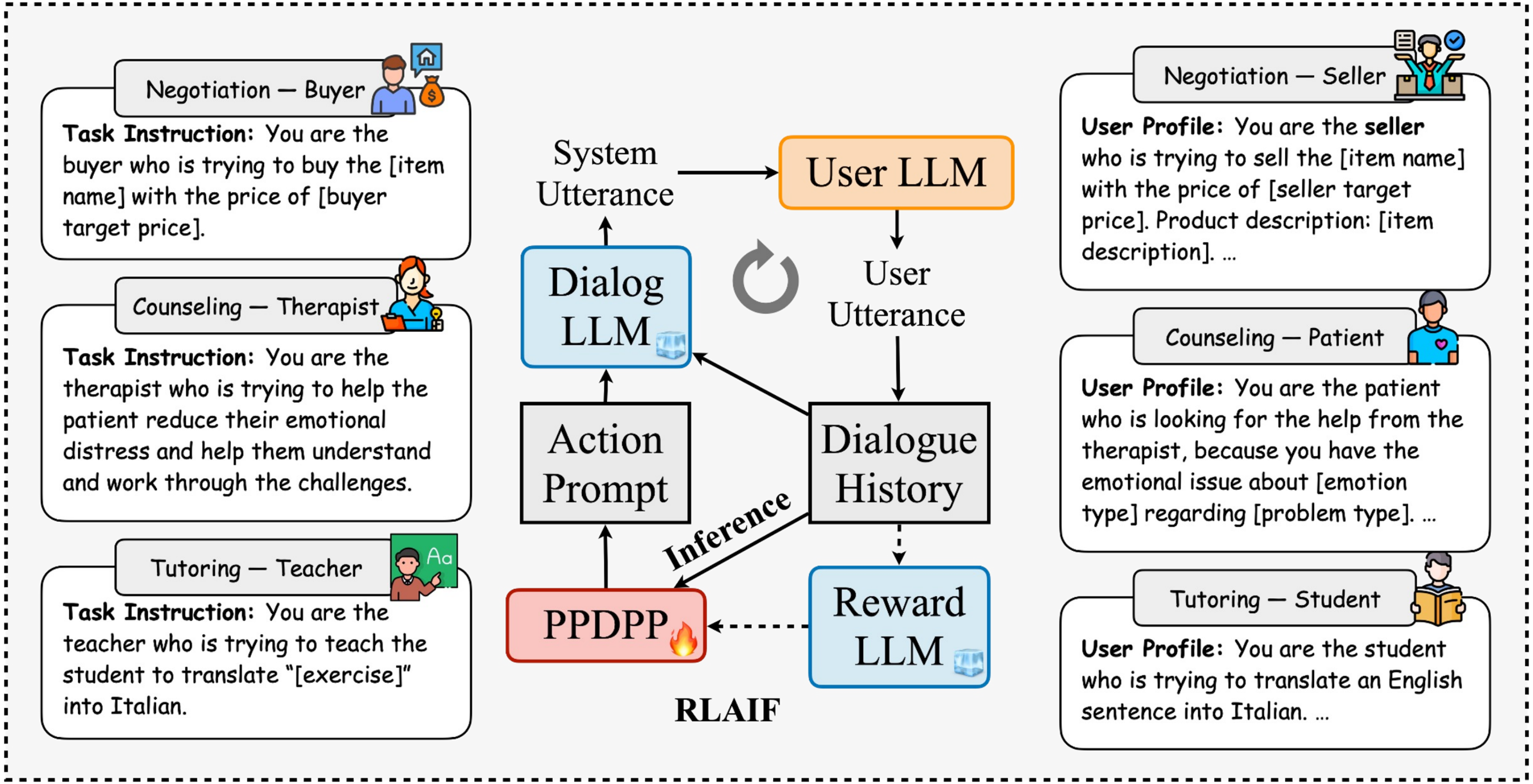
- An LLM to simulate the user with user profiles.
- Employ **Multi-agent Simulation** to collect dynamic interaction data.

$$u_t^{sys} = \mathbf{LLM}_{\text{sys}}(p_{\text{sys}}; \mathcal{M}_a(a_t); s_{t-1})$$
$$u_t^{usr} = \mathbf{LLM}_{\text{usr}}(p_{\text{usr}}; s_{t-1}; u_t^{sys})$$
$$s_t = \mathcal{T}(s_{t-1}, a_t)$$
$$= \{s_{t-1}; u_t^{sys}, u_t^{usr}\}$$



Yang Deng, Wenxuan Zhang, Wai Lam, See-Kiong Ng, Tat-Seng Chua. **Plug-and-Play Policy Planner for Large Language Model Powered Dialogue Agents. In ICLR 2024.**

Examples: Multi-agent Simulation



Yang Deng, Wenxuan Zhang, Wai Lam, See-Kiong Ng, Tat-Seng Chua. **Plug-and-Play Policy Planner for Large Language Model Powered Dialogue Agents. In ICLR 2024.**

Overview of LLM-powered Conversational Agents

Profile

LLM-powered Conversational Agents for **User Simulation**

Memory

LLM-powered Conversational Agents for **Long-context Dialogues**

Planning

LLM-powered Conversational Agents for **Proactive Dialogues**

Action

LLM-powered Conversational Agents for **Real-world Problem Solving**

Web Agents

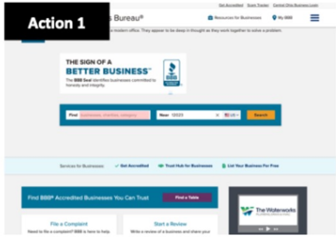
Web Agents aims to accomplish the tasks defined in natural language, such as booking tickets, through **multi-step interactions with the web-grounded environment**.

Task Description:
Show me the reviews for the auto repair business closest to 10002.

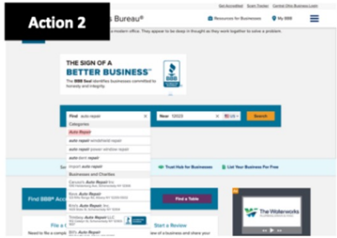
Action Sequence:

	Target Element	Operation
1.	[searchbox] *Find*	TYPE: *auto repair*
2.	[button] *Auto Repair*	CLICK
3.	[textbox] *Near*	TYPE: *10002*
4.	[button] *10002*	CLICK
5.	**[button] Search**	**CLICK**
6.	[switch] *Show BBB Accredited only*	CLICK
7.	[svg]	CLICK
8.	[button] Sort By	CLICK
9.	**[link] *Fast Lane 24 Hour Auto Repair***	**CLICK**
10.	[link] *Read Reviews*	CLICK

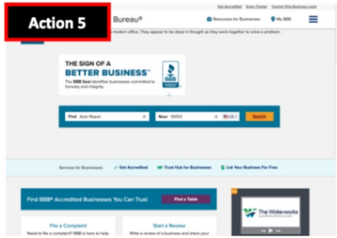
Webpage Snapshots:



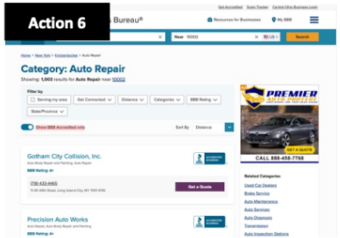
<input name="find_text" type="search">



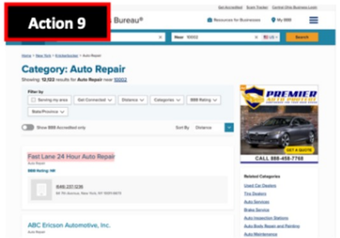
Auto Repair



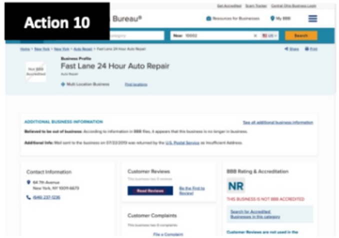
<button>Search</button>



<button>Show BBB Accredited only</button>



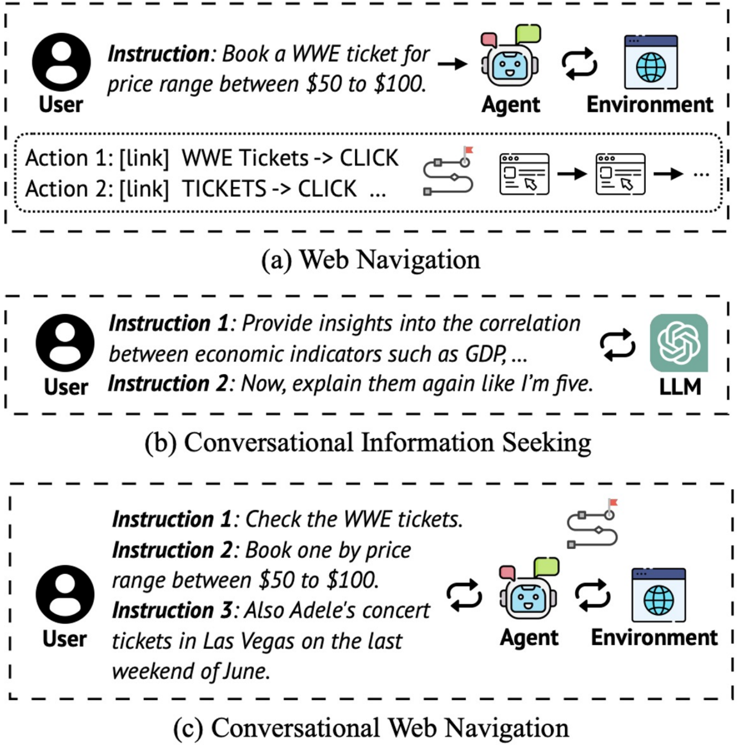
Fast Lane 24 Hour Auto Repair



Read Reviews

Deng et al., 2023. “Mind2Web: Towards a Generalist Agent for the Web” (NeurIPS ‘23)

Conversational Web Agents



Web Navigation

→ Single-turn User Instruction

→ Multi-step Environment Interaction

Conversational Information Seeking

→ Multi-turn User Instruction

→ No/Single-step Environment Interaction

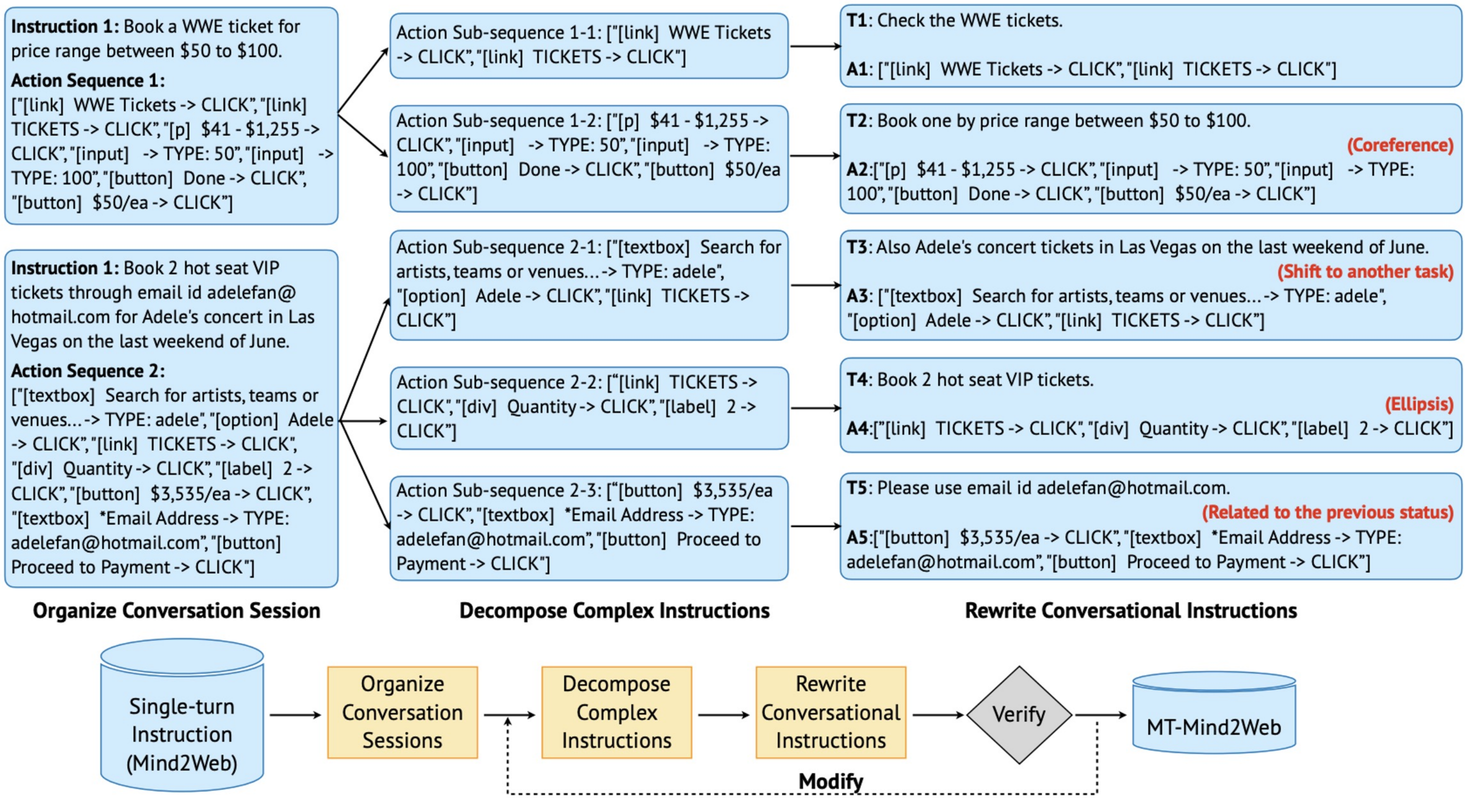
Conversational Web Navigation

→ Multi-turn User Instruction

→ Multi-step Environment Interaction

Deng et al., 2024. "On the Multi-turn Instruction Following of Conversational Web Agents" (CoRR '24)

Constructing the MT-Mind2Web Dataset



Deng et al., 2024. "On the Multi-turn Instruction Following of Conversational Web Agents" (CoRR '24)

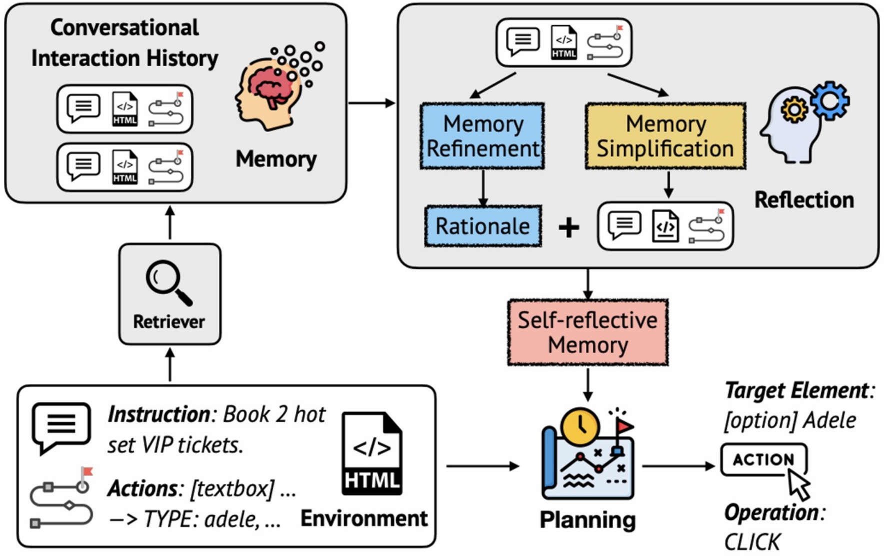
Challenges in Conversational Web Agents

<Longer and Noisier Context>

- ❑ **User-Agent Conversation**
  - **Coreference**: Users tend to use pronouns to refer to the previous mentioned entities
  - **Ellipsis**: Follow-up instructions may omit repeated information
  - **Task Shifting**: The completed task information can be noisy to the ongoing task
- ❑ **Agent-Environment Interaction**
  - **Action Dependency**: Multi-step actions are required to complete the task
  - **Environment Status Reliance**: Follow-up instructions may refer to the information in the environment rather than just the conversation history

Deng et al., 2024. "On the Multi-turn Instruction Following of Conversational Web Agents" (CoRR '24)

Self-reflective Memory-augmented Planning (Self-MAP)



Memory Module

→ **Memory Bank** to store memory snippets

→ **Multi-faceted Retriever** to retrieve memory snippets that are relevant to both the user instructions and the previous actions

Reflection Module

→ **Memory Refinement** to generate descriptive rationale from the complex memory snippets for planning

→ **Memory Simplification** to filter out irrelevant elements from the environment status for saving memory space

Planning Module

→ **Memory-augmented Planning**

Deng et al., 2024. "On the Multi-turn Instruction Following of Conversational Web Agents" (CoRR '24)

Overview of LLM-powered Conversational Agents

Profile

LLM-powered Conversational Agents for **User Simulation**

Memory

LLM-powered Conversational Agents for **Long-context Dialogues**

Planning

LLM-powered Conversational Agents for **Proactive Dialogues**

Action

LLM-powered Conversational Agents for **Real-world Problem Solving**