SIGIR
2024
Washington, D.C.

# Large Language Model Powered Conversational Agents

Yang Deng

Singapore Management University

**July 14, 2024**

SMU
SINGAPORE MANAGEMENT UNIVERSITY

# Large Language Model Powered Conversational Systems

ChatGPT

Gemini

New Bing

Claude

…

Alpaca

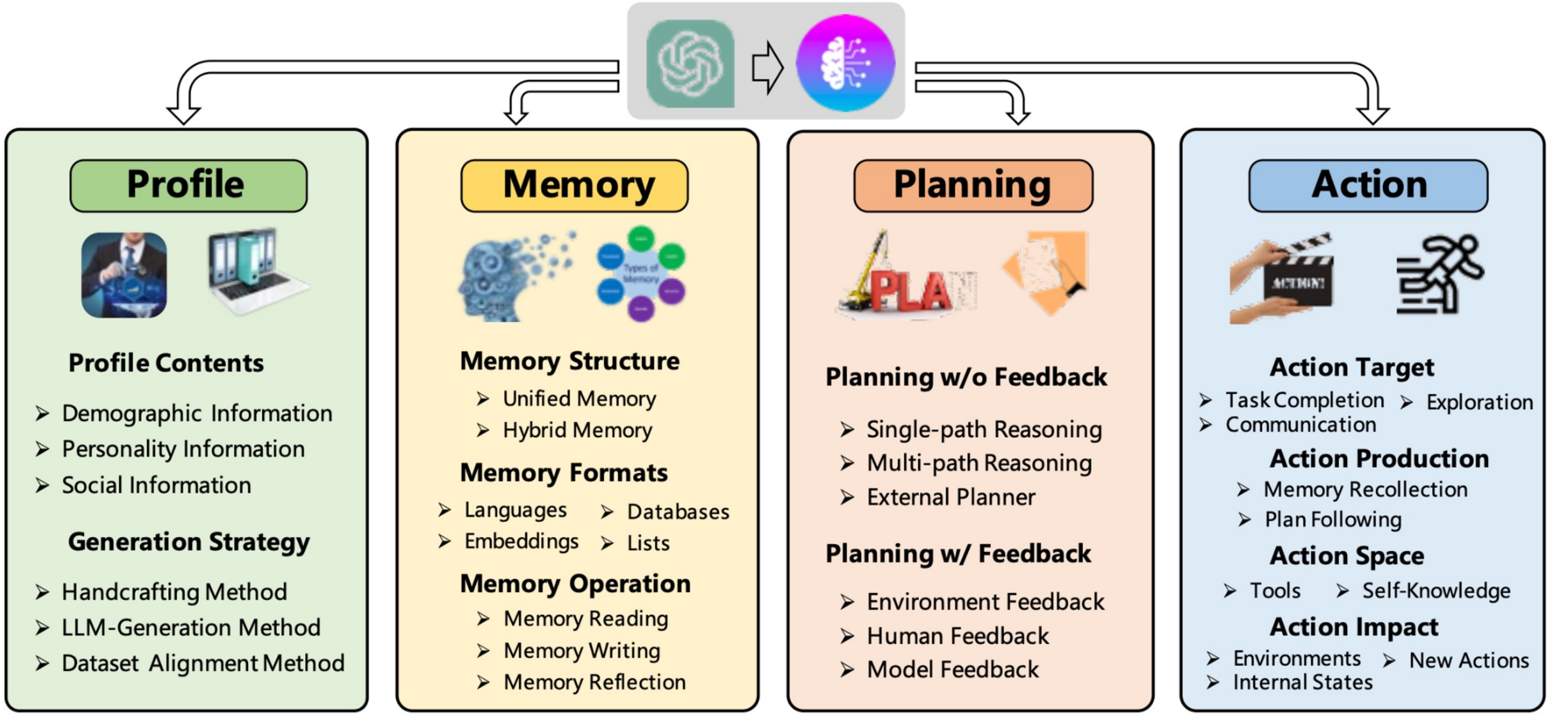
Vicuna

Dolly

LLaMA-Chat

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

# LLM-powered Conversational Agents?



## Profile

**Profile Contents**

- Demographic Information
- Personality Information
- Social Information

**Generation Strategy**

- Handcrafting Method
- LLM-Generation Method
- Dataset Alignment Method

## Memory

**Memory Structure**

- Unified Memory
- Hybrid Memory

**Memory Formats**

- Languages
- Databases
- Embeddings
- Lists

**Memory Operation**

- Memory Reading
- Memory Writing
- Memory Reflection

## Planning

**Planning w/o Feedback**

- Single-path Reasoning
- Multi-path Reasoning
- External Planner

**Planning w/ Feedback**

- Environment Feedback
- Human Feedback
- Model Feedback

## Action

**Action Target**

- Task Completion
- Exploration
- Communication

**Action Production**

- Memory Recollection
- Plan Following

**Action Space**

- Tools
- Self-Knowledge

**Action Impact**

- Environments
- New Actions
- Internal States

*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

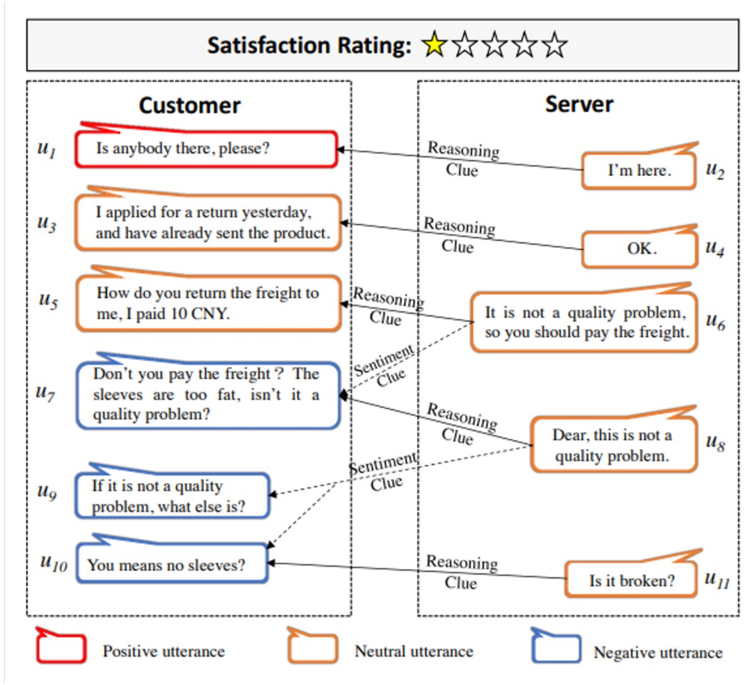
1) Semantic-based Estimation
2) Preference-based Estimation
3) Action-based Estimation

## ❑ User Response Simulation

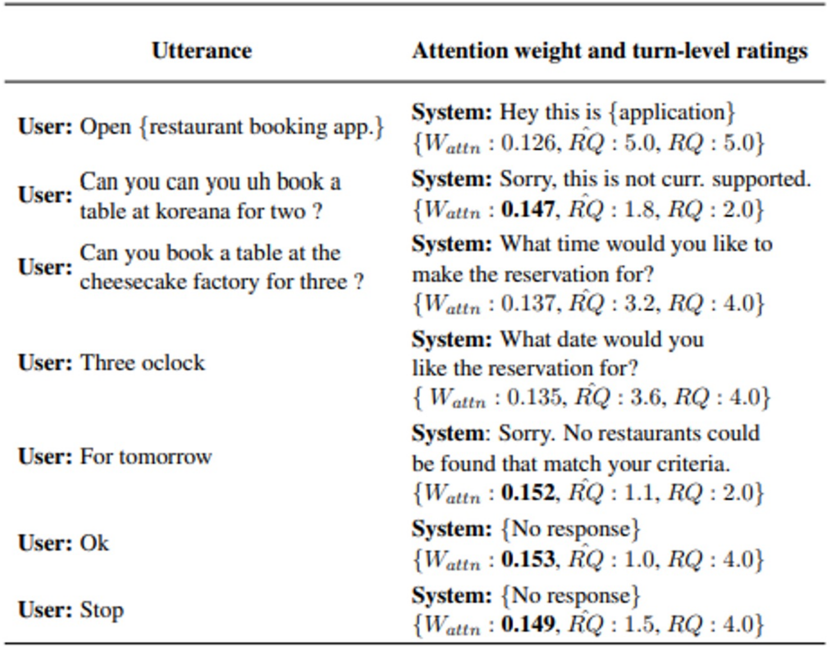
1) Retrieval-based User Simulators
2) Schema-based User Simulators
3) Conditioned Generation Models as User Simulators

# Semantic-based User Satisfaction Estimation

## Sentiment Classification


Satisfaction Rating: ★☆☆☆☆
Customer
Server
$u_1$ Is anybody there, please?
$u_2$ I'm here.
Reasoning Clue
$u_3$ I applied for a return yesterday, and have already sent the product.
$u_4$ OK.
Reasoning Clue
$u_5$ How do you return the freight to me, I paid 10 CNY.
$u_6$ It is not a quality problem, so you should pay the freight.
Reasoning Clue
Sentiment Clue
$u_7$ Don't you pay the freight ? The sleeves are too fat, isn't it a quality problem?
$u_8$ Dear, this is not a quality problem.
Reasoning Clue
Sentiment Clue
$u_9$ If it is not a quality problem, what else is?
$u_{10}$ You means no sleeves?
$u_{11}$ Is it broken?
Reasoning Clue
Positive utterance
Neutral utterance
Negative utterance

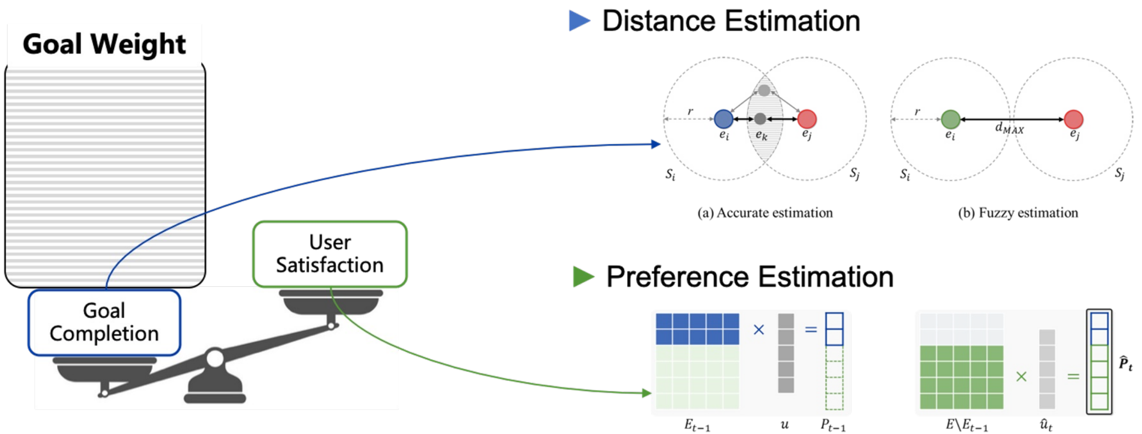
## Response Quality Assessment



| 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. {$W_{attn}$ : **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. {$W_{attn}$ : **0.152**, $\hat{RQ}$ : 1.1, $RQ$ : 2.0} |
| **User:** Ok | **System:** {No response} {$W_{attn}$ : **0.153**, $\hat{RQ}$ : 1.0, $RQ$ : 4.0} |
| **User:** Stop | **System:** {No response} {$W_{attn}$ : **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



User vector estimation (Eq. 2)

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

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

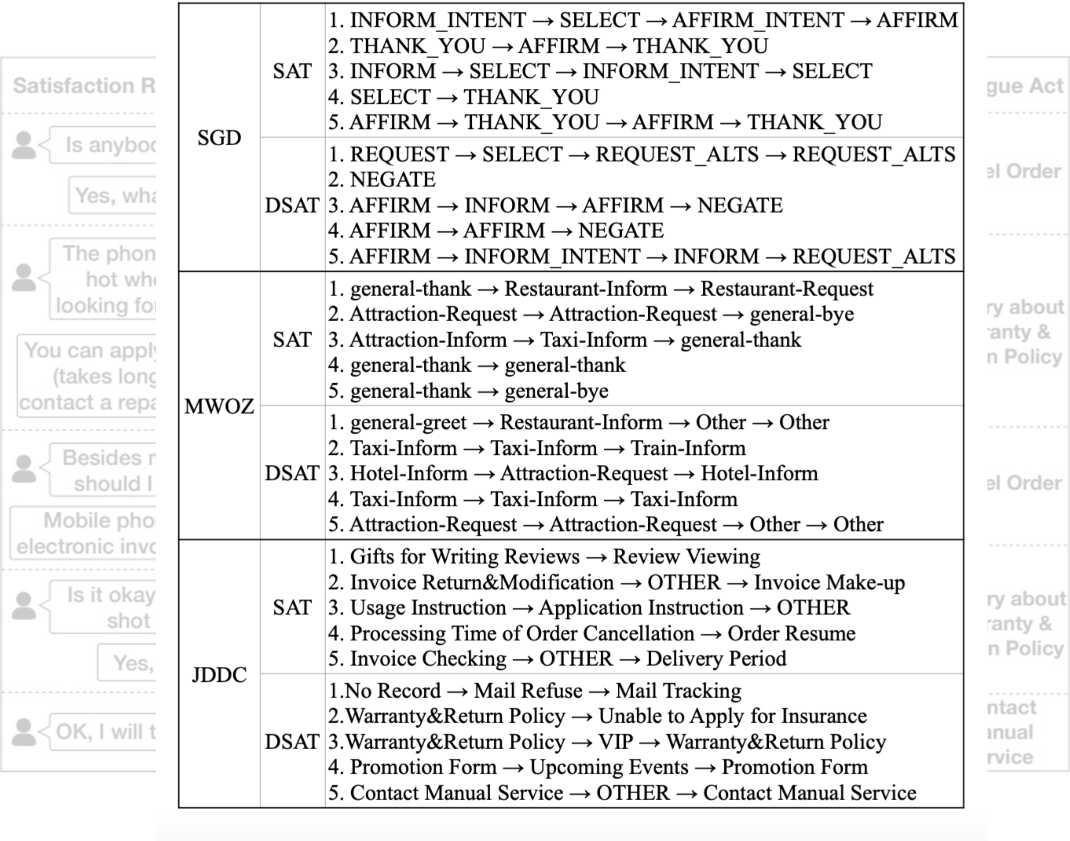
estimated user preference of *et,i*

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

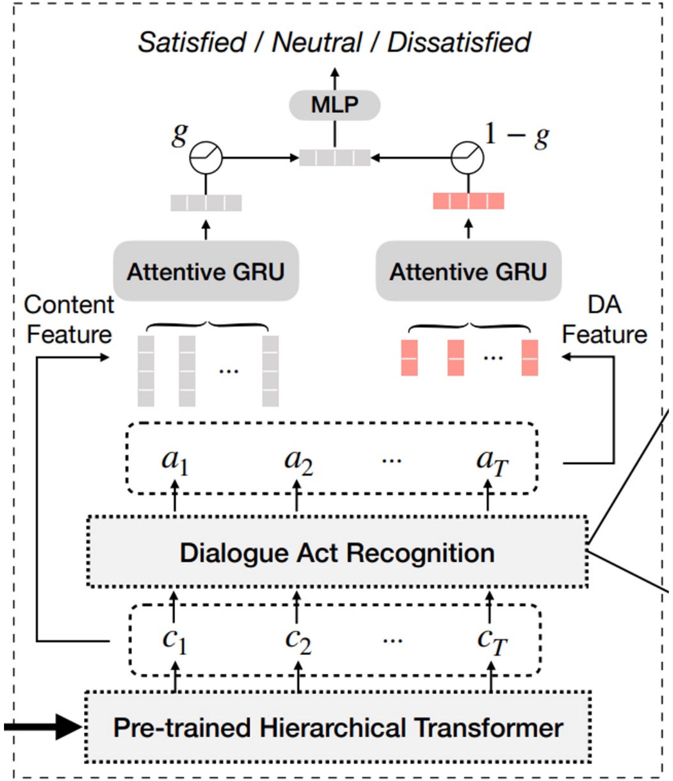
$$US_t \triangleq \frac{1}{t}\sum_{i=1}^{t}\frac{1}{|u_i+1|}\left(\sum_{j=1}^{|u_i|} p_{e_{i,j}} + p_{e_i^a}\right)$$

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

# Action-based User Satisfaction Estimation

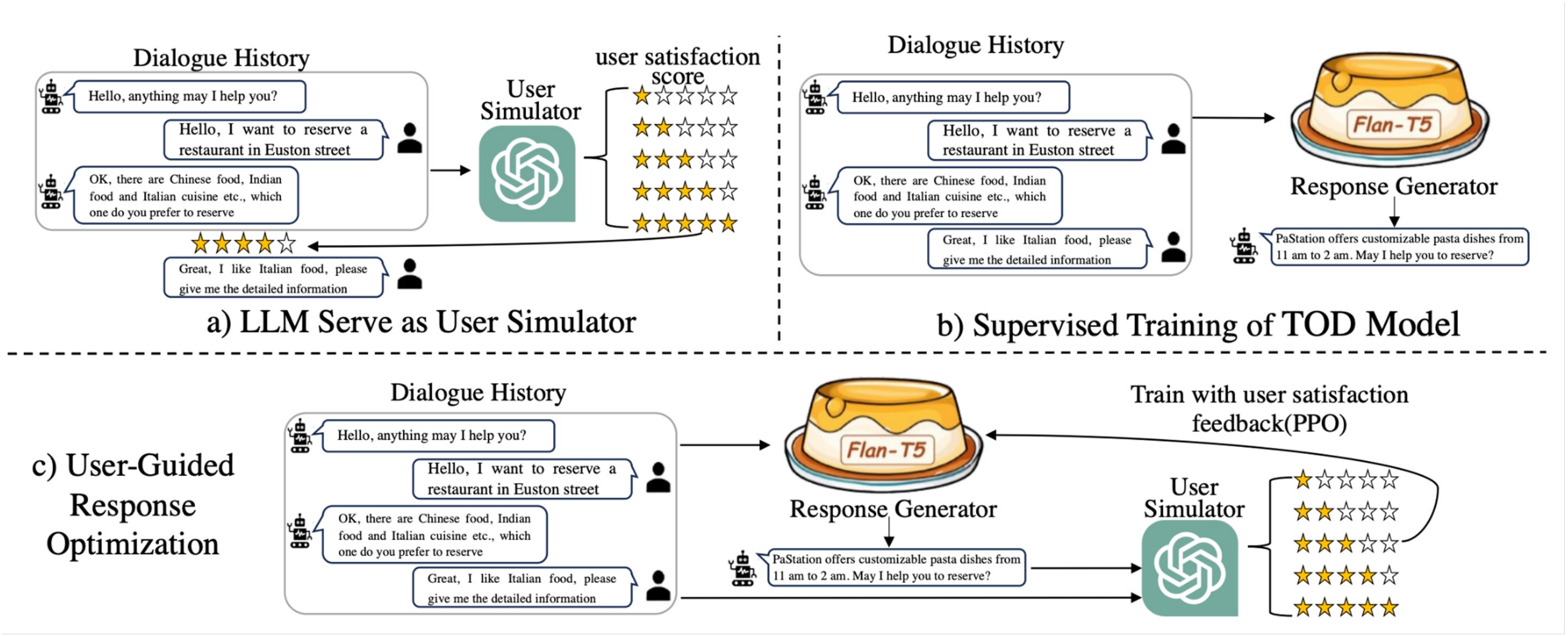

Satisfaction R
Is anybod
Yes, wha
The phon
hot wh
looking fo
You can appl
(takes long
contact a repa
Besides r
should I
Mobile pho
electronic inv
Is it okay
shot
Yes,
OK, I will t
gue Act
el Order
ry about
anty &
n Policy
el Order
ry about
anty &
n Policy
ntact
nual
rvice

| | | |
|---|---|---|
| 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



Hello, anything may I help you?

Hello, I want to reserve a restaurant in Euston street

OK, there are Chinese food, Indian food and Italian cuisine etc., which one do you prefer to reserve

Great, I like Italian food, please give me the detailed information

PaStation offers customizable pasta dishes from 11 am to 2 am. May I help you to reserve?

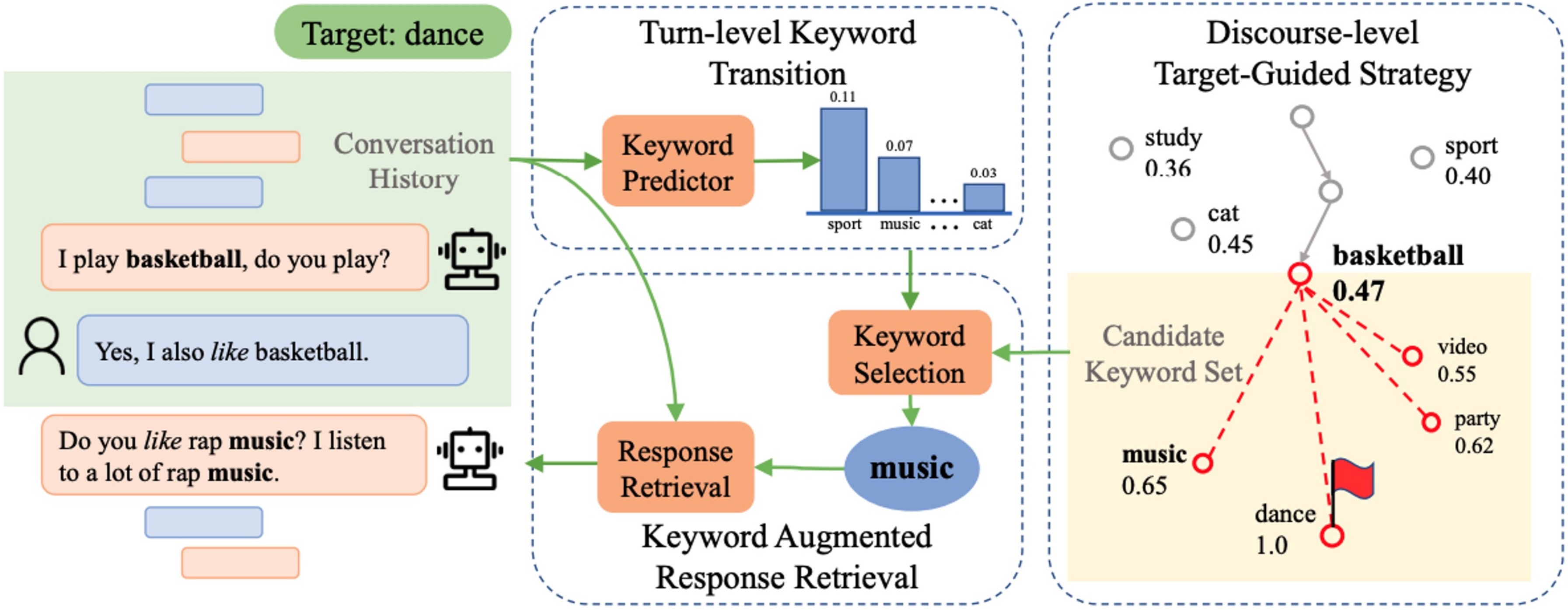
*Hu et al., 2023. "Unlocking the Potential of User Feedback: Leveraging Large Language Model as User Simulator to Enhance Dialogue System" (CIKM '23)*

# User Simulators in the Pre-LLM Era

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

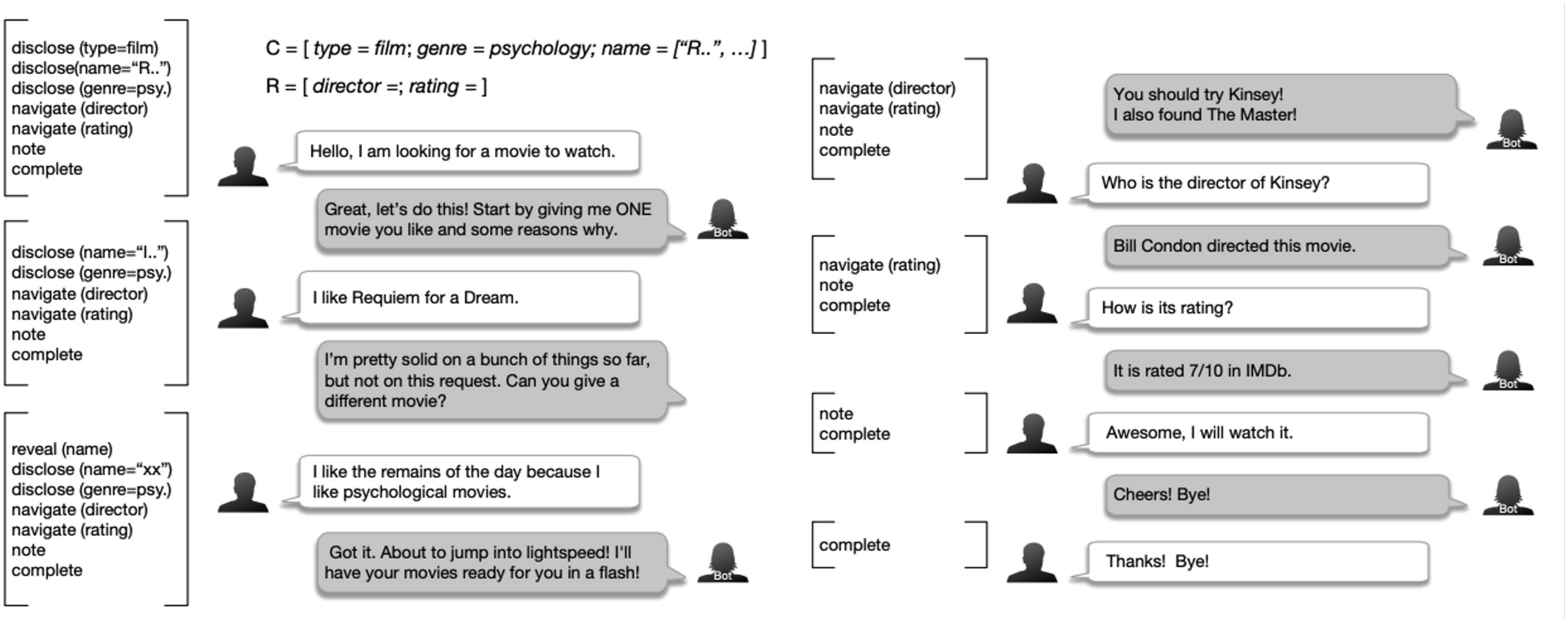
- ❑ **User Response Simulation**
  1) Retrieval-based User Simulators
  2) Schema-based User Simulators
  3) Conditioned Generation Models as User Simulators

# Retrieval-based User Simulators


Target: dance
Conversation History
I play **basketball**, do you play?
Yes, I also *like* basketball.
Do you *like* rap **music**? I listen to a lot of rap **music**.
Turn-level Keyword Transition
Keyword Predictor
0.11
0.07
0.03
sport
music
…
cat
Keyword Selection
music
Response Retrieval
Keyword Augmented Response Retrieval
Discourse-level Target-Guided Strategy
study 0.36
sport 0.40
cat 0.45
basketball 0.47
Candidate Keyword Set
video 0.55
party 0.62
music 0.65
dance 1.0

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

# Schema-based User Simulators



C = [ *type* = *film*; *genre* = *psychology; name* = *[“R..”, …]* ]

R = [ *director* =; *rating* = ]

disclose (type=film)
disclose(name=“R..”)
disclose (genre=psy.)
navigate (director)
navigate (rating)
note
complete

Hello, I am looking for a movie to watch.

Great, let’s do this! Start by giving me ONE movie you like and some reasons why.

disclose (name=“I..”)
disclose (genre=psy.)
navigate (director)
navigate (rating)
note
complete

I like Requiem for a Dream.

I’m pretty solid on a bunch of things so far, but not on this request. Can you give a different movie?

reveal (name)
disclose (name=“xx”)
disclose (genre=psy.)
navigate (director)
navigate (rating)
note
complete

I like the remains of the day because I like psychological movies.

Got it. About to jump into lightspeed! I'll have your movies ready for you in a flash!

You should try Kinsey!
I also found The Master!

navigate (director)
navigate (rating)
note
complete

Who is the director of Kinsey?

Bill Condon directed this movie.

navigate (rating)
note
complete

How is its rating?

It is rated 7/10 in IMDb.

note
complete

Awesome, I will watch it.

Cheers! Bye!

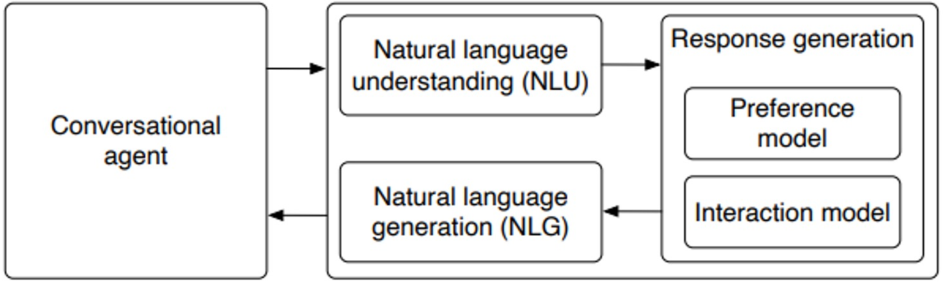
complete

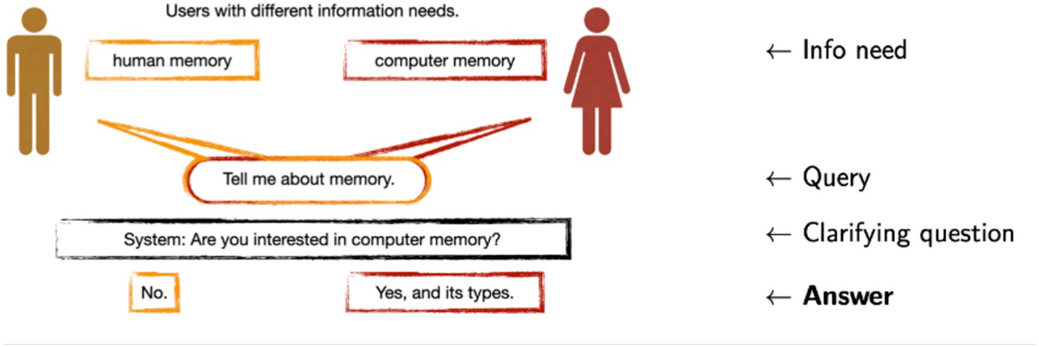
Thanks!  Bye!

*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.

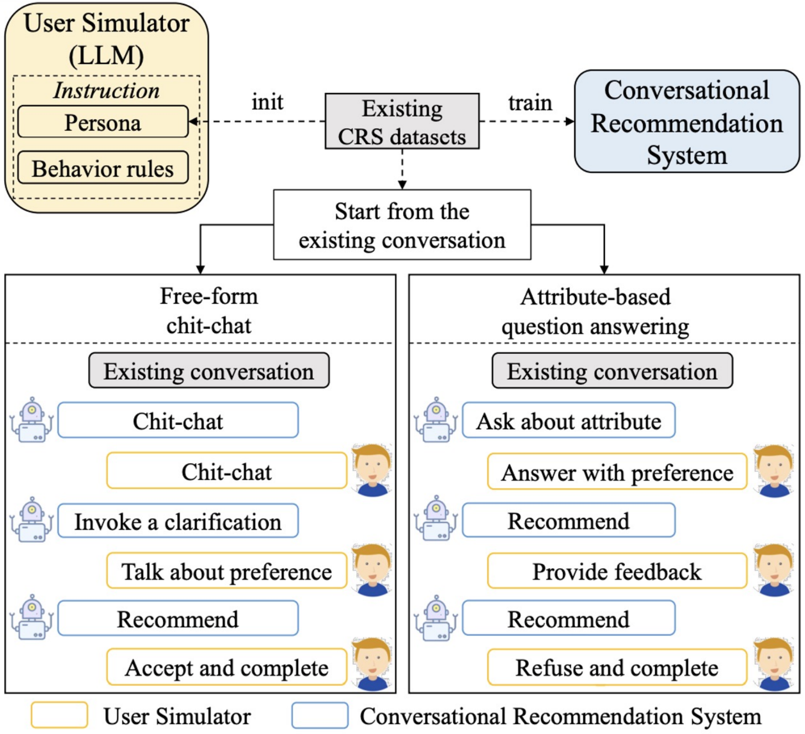




Conditioned on **information needs** for evaluating conversational search systems.

*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



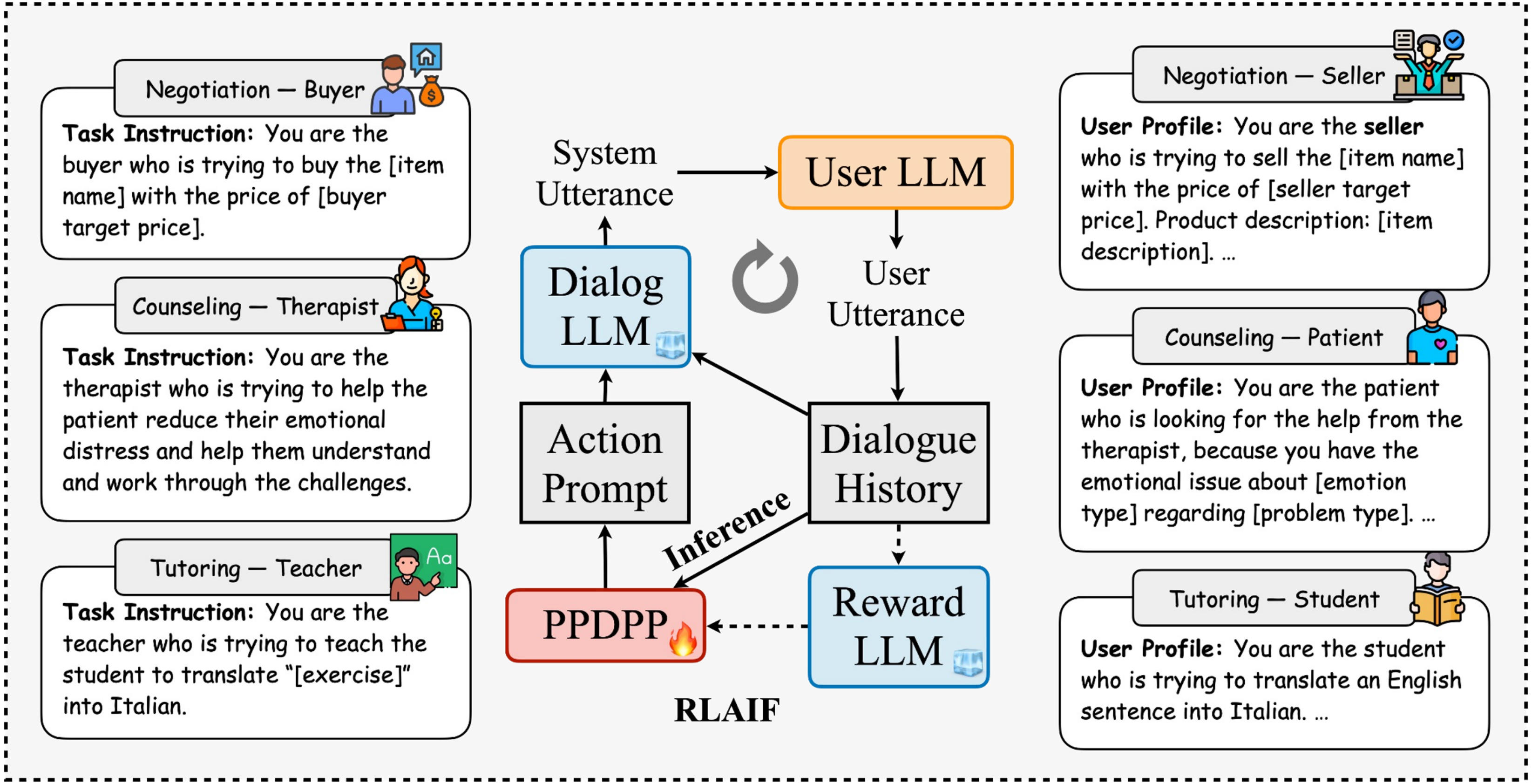
LLMs possess excellent ***role-playing*** capacities.

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



**Negotiation — Buyer**

**Task Instruction:** You are the buyer who is trying to buy the [item name] with the price of [buyer target price].

**Counseling — Therapist**

**Task Instruction:** You are the therapist who is trying to help the patient reduce their emotional distress and help them understand and work through the challenges.

**Tutoring — Teacher**

**Task Instruction:** You are the teacher who is trying to teach the student to translate "[exercise]" into Italian.

**Negotiation — Seller**

**User Profile:** You are the **seller** who is trying to sell the [item name] with the price of [seller target price]. Product description: [item description]. …

**Counseling — Patient**

**User Profile:** You are the patient who is looking for the help from the therapist, because you have the emotional issue about [emotion type] regarding [problem type]. …

**Tutoring — Student**

**User Profile:** You are the student who is trying to translate an English sentence into Italian. …

*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

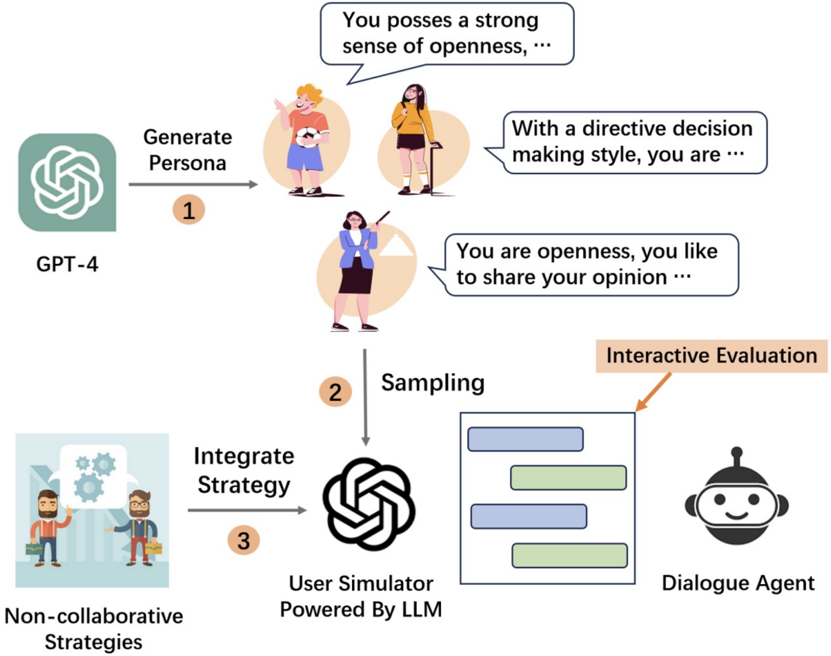
***Why do we need to simulate diverse users?***

Examples: Non-collaborative Dialogues (Negotiation/Persuasion)

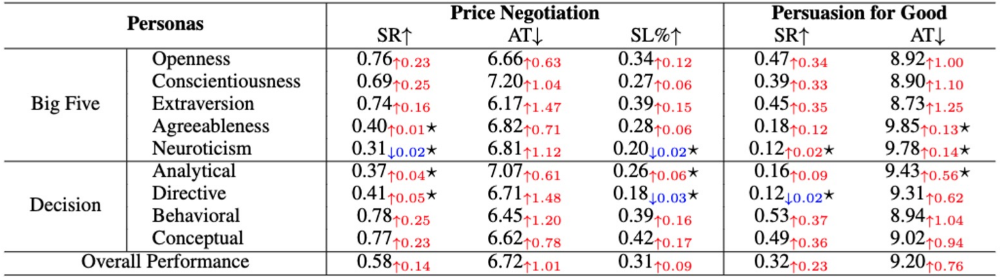
- Existing dialogue systems overlook the integration of explicit **user-specific characteristics** in their strategic planning
- The training paradigm with a static user simulator fails to make strategic plans that can be **generalized to diverse users**

*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


You posses a strong sense of openness, ⋯
With a directive decision making style, you are ⋯
Generate Persona
1
GPT-4
You are openness, you like to share your opinion ⋯
2
Sampling
Interactive Evaluation
Integrate Strategy
3
Non-collaborative Strategies
User Simulator Powered By LLM
Dialogue Agent

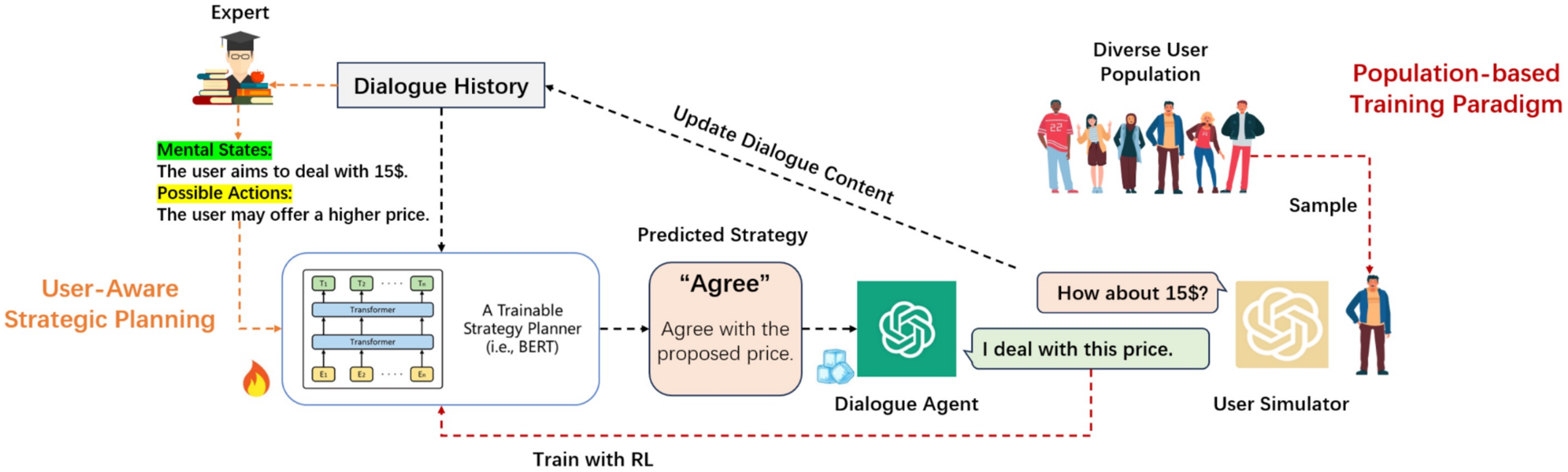
- ❑ Big-Five Personality:
  - *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness*, and *Neuroticism*
- ❑ Decision-Making Styles:
  - *Directive*, *Conceptual*, *Analytical*, and *Behavioral*.



| Personas | | Price Negotiation SR↑ | Price Negotiation AT↓ | Price Negotiation SL%↑ | Persuasion for Good SR↑ | Persuasion for Good 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

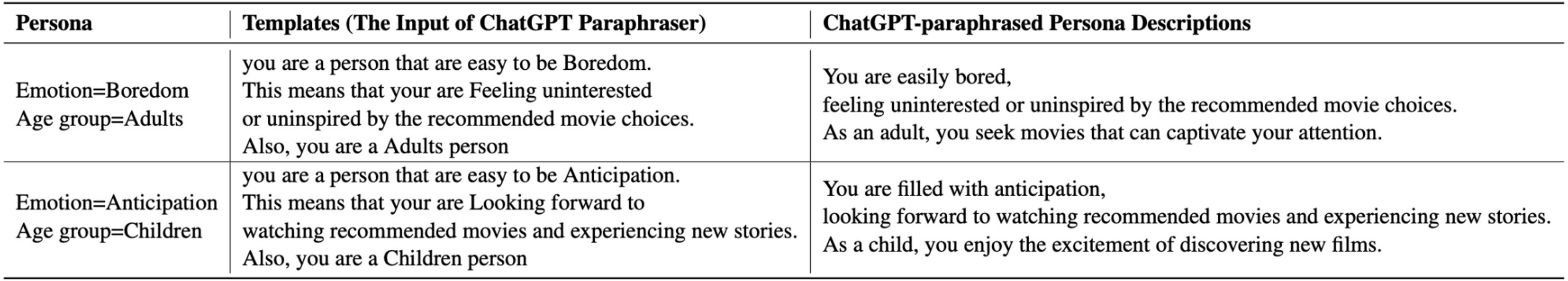
- ❑ **User-aware Strategy Planning**: Predict user mental states and possible actions
- ❑ **Population-based Reinforcement Learning**: Sample a diverse group of simulated users to interact

*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?***

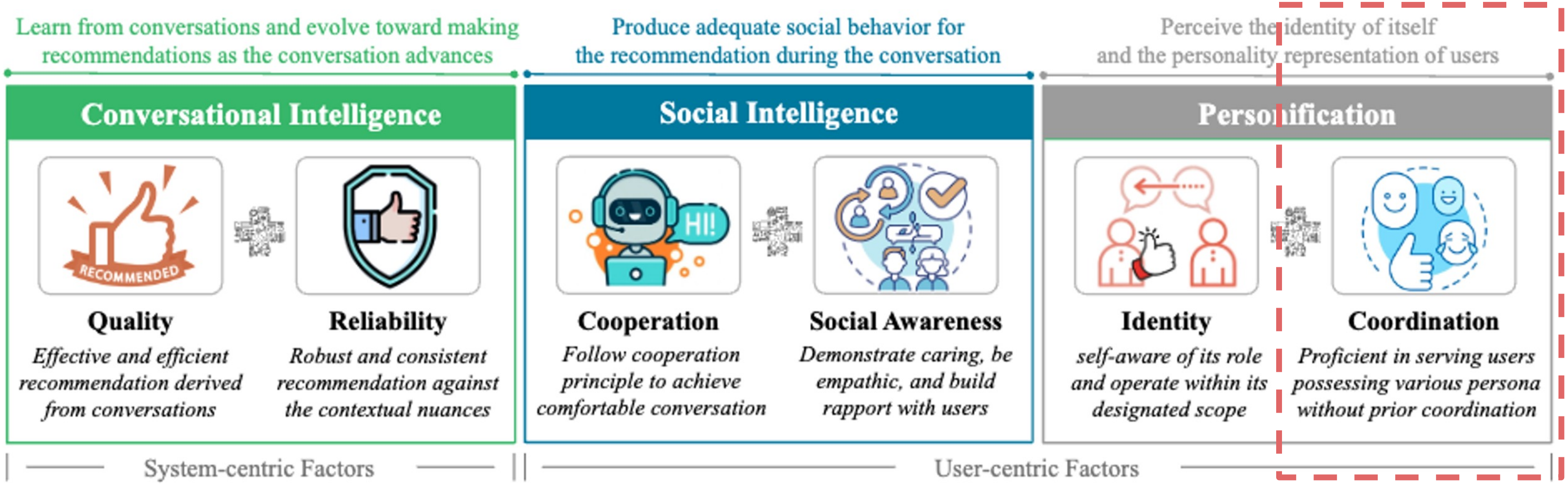
Wang et al., (2023) conclude that LLM-based user simulators are easier to accept the recommended items than human users during the evaluation of conversational recommender systems, since LLMs tend to follow the given instructions. → **Biased Evaluation!!!**



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

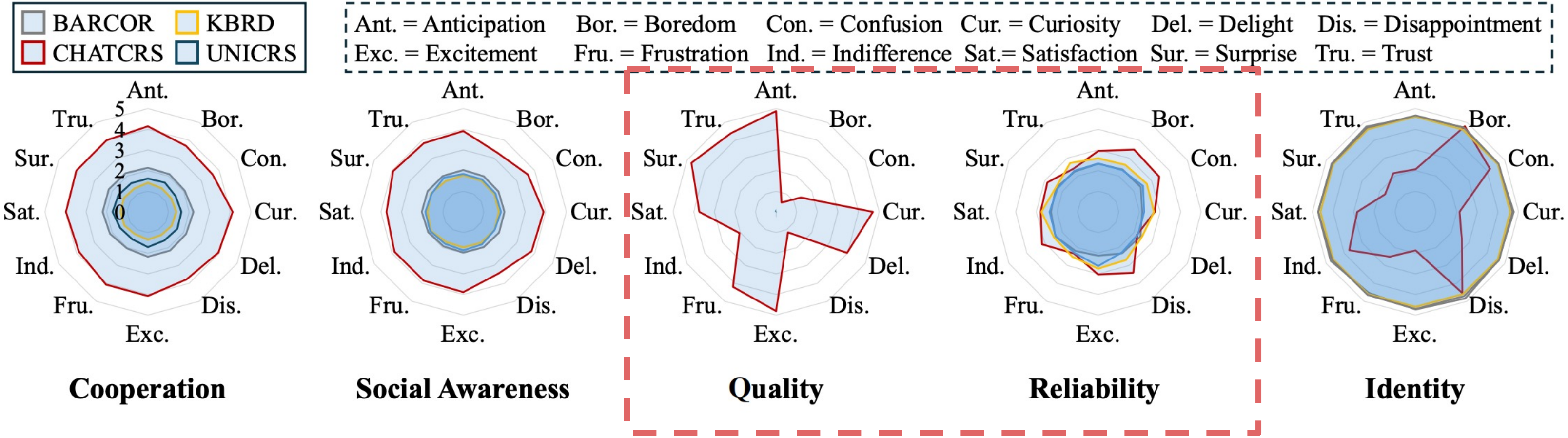

Learn from conversations and evolve toward making recommendations as the conversation advances
Produce adequate social behavior for the recommendation during the conversation
Perceive the identity of itself and the personality representation of users
Conversational Intelligence
Social Intelligence
Personification
RECOMMENDED
HI!
Quality
Effective and efficient recommendation derived from conversations
Reliability
Robust and consistent recommendation against the contextual nuances
Cooperation
Follow cooperation principle to achieve comfortable conversation
Social Awareness
Demonstrate caring, be empathic, and build rapport with users
Identity
self-aware of its role and operate within its designated scope
Coordination
Proficient in serving users possessing various persona without prior coordination
System-centric Factors
User-centric Factors

## 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

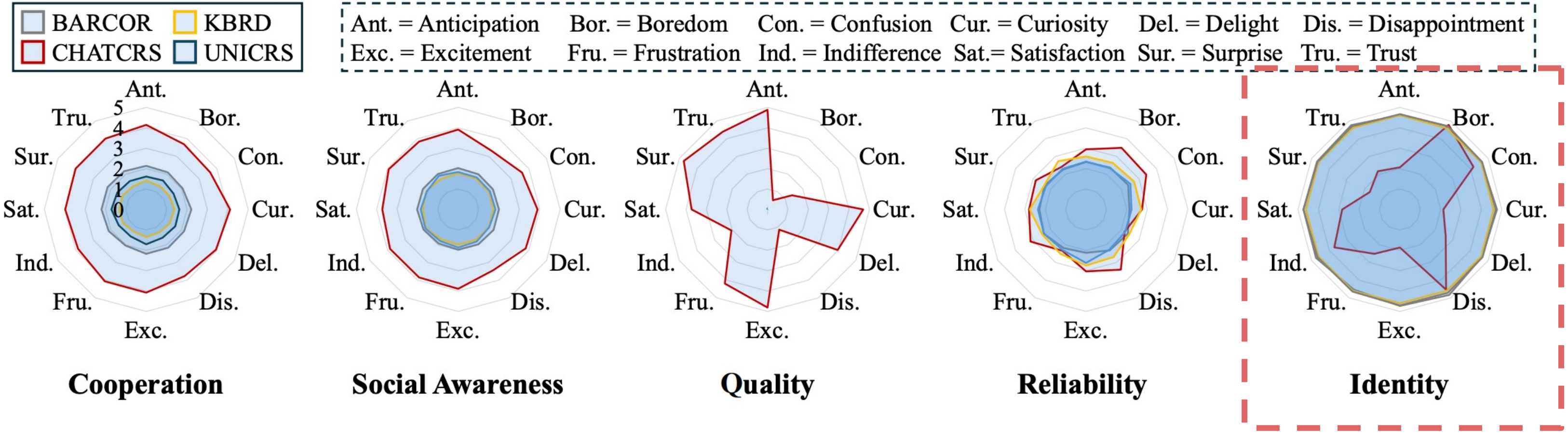


Evaluation with Simulated Users from Different Personas

- Most pre-LLM conversational recommender models show poor performance in sensing the variation of users. (**System-centric Metrics: Quality & Reliability**)

*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

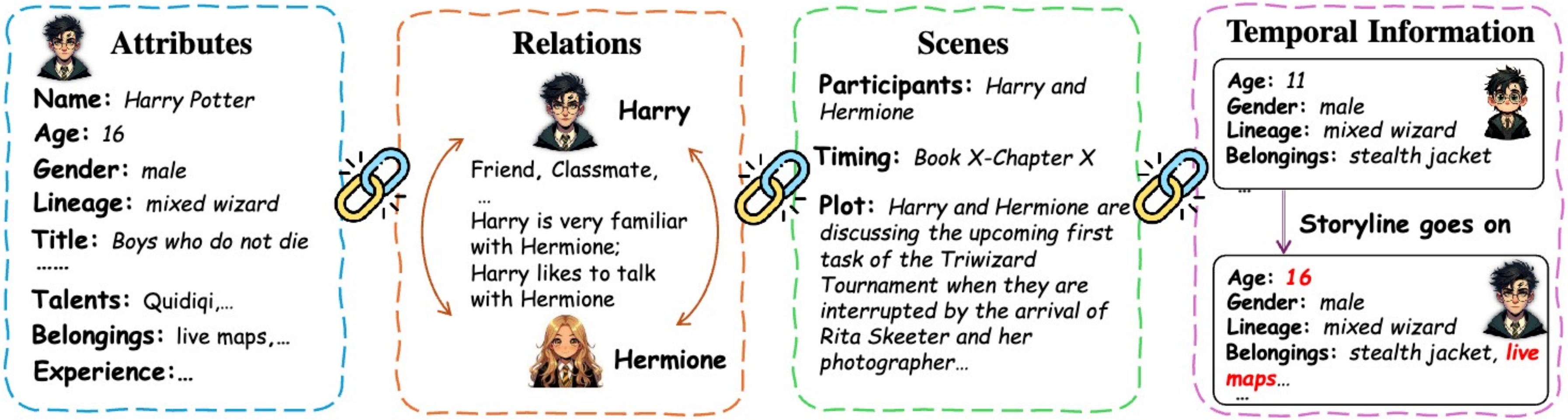


Evaluation with Simulated Users from Different Personas

- ❑ Most pre-LLM conversational recommender models show poor performance in sensing the variation of users. (**System-centric Metrics: Quality & Reliability**)
- ❑ LLM-based conversational recommender models (e.g., CHATCRS) tend to adopt 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)*

# Profiling for Role-playing Conversational Agents



**Attributes**

**Name:** *Harry Potter*
**Age:** *16*
**Gender:** *male*
**Lineage:** *mixed wizard*
**Title:** *Boys who do not die*
……
**Talents:** Quidiqi,…
**Belongings:** live maps,…
**Experience:**…

**Relations**

Friend, Classmate,
…
Harry is very familiar with Hermione;
Harry likes to talk with Hermione

**Scenes**

**Participants:** *Harry and Hermione*
**Timing:** *Book X-Chapter X*
**Plot:** *Harry and Hermione are discussing the upcoming first task of the Triwizard Tournament when they are interrupted by the arrival of Rita Skeeter and her photographer…*

**Temporal Information**

**Age:** *11*
**Gender:** *male*
**Lineage:** *mixed wizard*
**Belongings:** *stealth jacket*
…

**Storyline goes on**

**Age:** *16*
**Gender:** *male*
**Lineage:** *mixed wizard*
**Belongings:** *stealth jacket, live maps…*
…

- **Attributes**: Personal details, such as *name, gender, personality*, …
- **Relations**: Social landscapes, such as *friends, family, opponent*, …
- **Scenes**: Contextual and background information, such as *timing, location, situation*, …
- **Temporal Information**: Evolving information along time, such as *aging, storyline*, …

*Chen et al., 2024. "The Oscars of AI Theater: A Survey on Role-Playing with Language Models" (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**

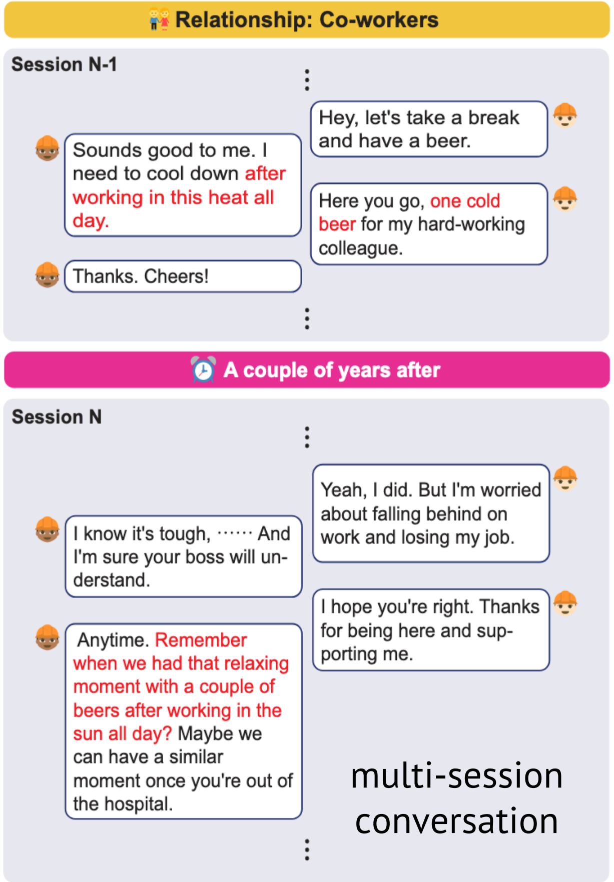
**Planning**

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

**Action**

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

# What is Long-context Dialogue?



**Relationship: Co-workers**

**Session N-1**

Hey, let's take a break and have a beer.

Sounds good to me. I need to cool down after working in this heat all day.

Here you go, one cold beer for my hard-working colleague.

Thanks. Cheers!

**A couple of years after**

**Session N**

Yeah, I did. But I'm worried about falling behind on work and losing my job.

I know it's tough, …… And I'm sure your boss will understand.

I hope you're right. Thanks for being here and supporting me.

Anytime. Remember when we had that relaxing moment with a couple of beers after working in the sun all day? Maybe we can have a similar moment once you're out of the hospital.

multi-session conversation

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

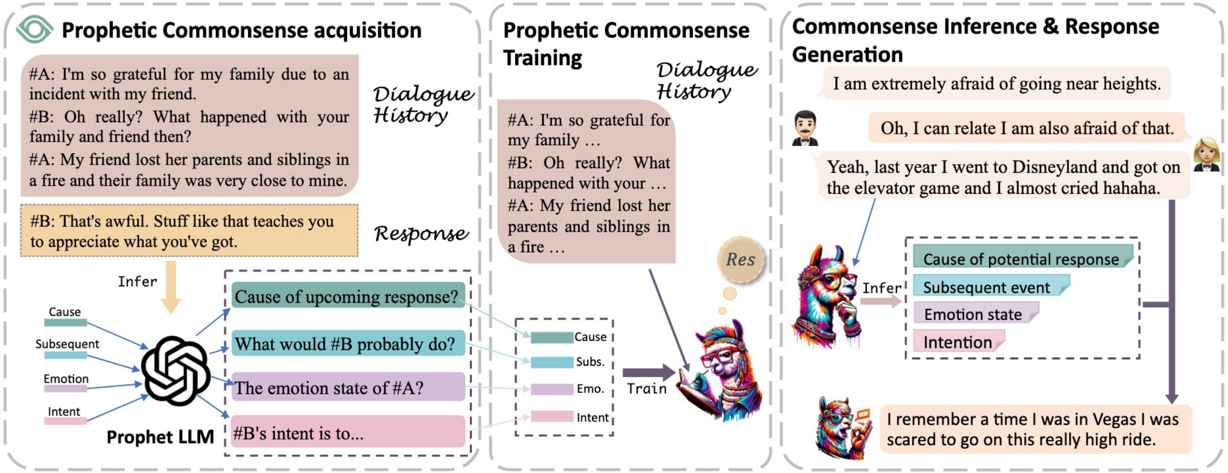
- ❑ Long-context dialogue systems requires memorization and personalization in ***multi-session*** conversations, providing more consistent and tailored responses.

*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.



#A: I'm so grateful for my family due to an incident with my friend.
#B: Oh really? What happened with your family and friend then?
#A: My friend lost her parents and siblings in a fire and their family was very close to mine.

#B: That's awful. Stuff like that teaches you to appreciate what you've got.

Cause of upcoming response?
What would #B probably do?
The emotion state of #A?
#B's intent is to...

#A: I'm so grateful for my family …
#B: Oh really? What happened with your …
#A: My friend lost her parents and siblings in a fire …

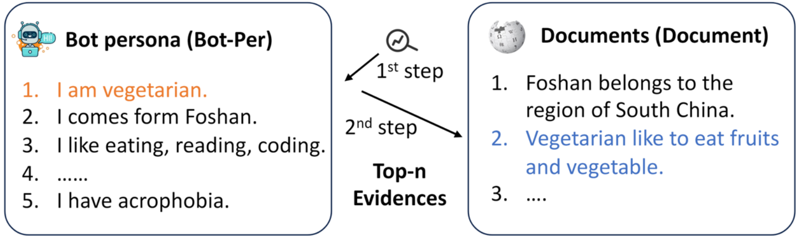
I am extremely afraid of going near heights.
Oh, I can relate I am also afraid of that.
Yeah, last year I went to Disneyland and got on the elevator game and I almost cried hahaha.

I remember a time I was in Vegas I was scared to go on this really high ride.

The framework of employing external knowledge to reasoning.

**Knowledge Sources:**

- ☐ Commonsense Knowledge
- ☐ Medical Knowledge
- ☐ Psychology Knowledge
- ☐ …



**Bot persona (Bot-Per)**

1. I am vegetarian.
2. I comes form Foshan.
3. I like eating, reading, coding.
4. ……
5. I have acrophobia.

**Documents (Document)**

1. Foshan belongs to the region of South China.
2. Vegetarian like to eat fruits and vegetable.
3. ….

*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”*

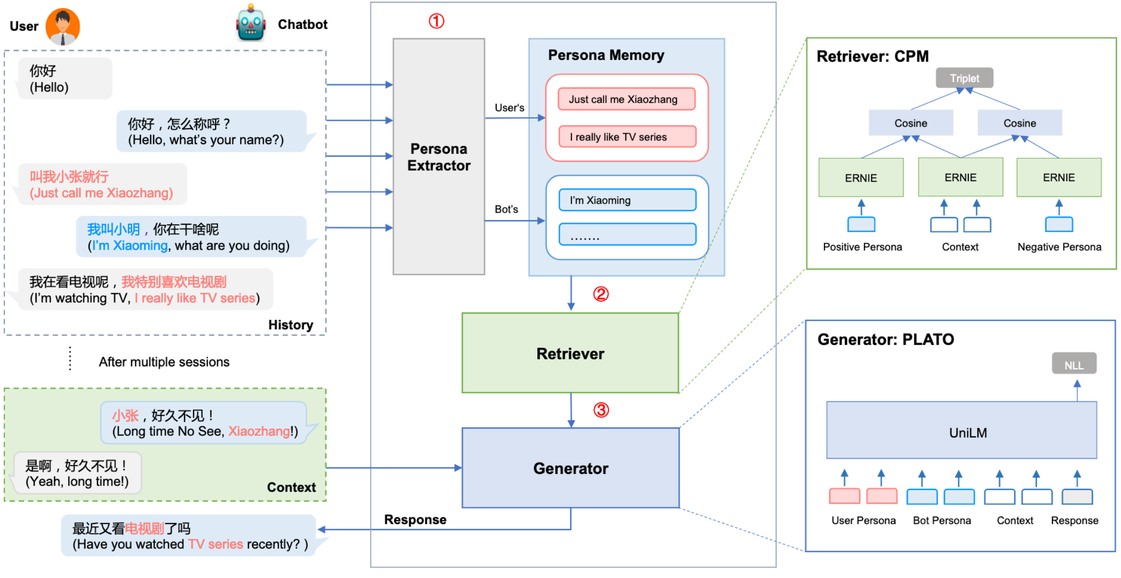
# Internal Knowledge for Long-context Dialogue

- 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.



你好
(Hello)

你好，怎么称呼？
(Hello, what's your name?)

叫我小张就行
(Just call me Xiaozhang)

我叫小明，你在干啥呢
(I'm Xiaoming, what are you doing)

我在看电视呢，我特别喜欢电视剧
(I'm watching TV, I really like TV series)

After multiple sessions

小张，好久不见！
(Long time No See, Xiaozhang!)

是啊，好久不见！
(Yeah, long time!)

最近又看电视剧了吗
(Have you watched TV series recently? )

Just call me Xiaozhang

I really like TV series

I'm Xiaoming

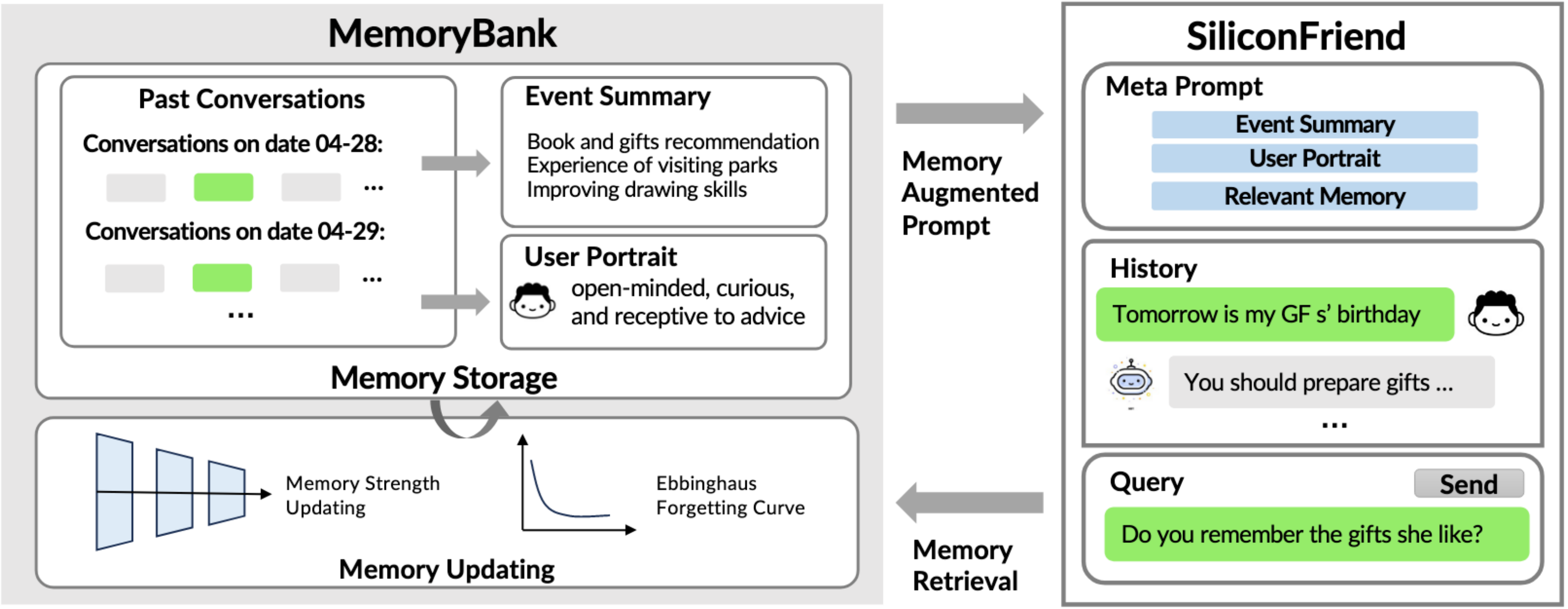
(a) Dialogue Flow (b) System Pipeline (c) Models

*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.



**MemoryBank**

**Past Conversations**

Conversations on date 04-28:

Conversations on date 04-29:

**Event Summary**

Book and gifts recommendation
Experience of visiting parks
Improving drawing skills

**User Portrait**

open-minded, curious, and receptive to advice

**Memory Storage**

**Memory Updating**

**Memory Augmented Prompt**

**Memory Retrieval**

**SiliconFriend**

**Meta Prompt**

Event Summary
User Portrait
Relevant Memory

**History**

Tomorrow is my GF s' birthday

You should prepare gifts ...

**Query**

Do you remember the gifts she like?

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

# Internal Knowledge for Long-context Dialogue

- Personas & Historical Events

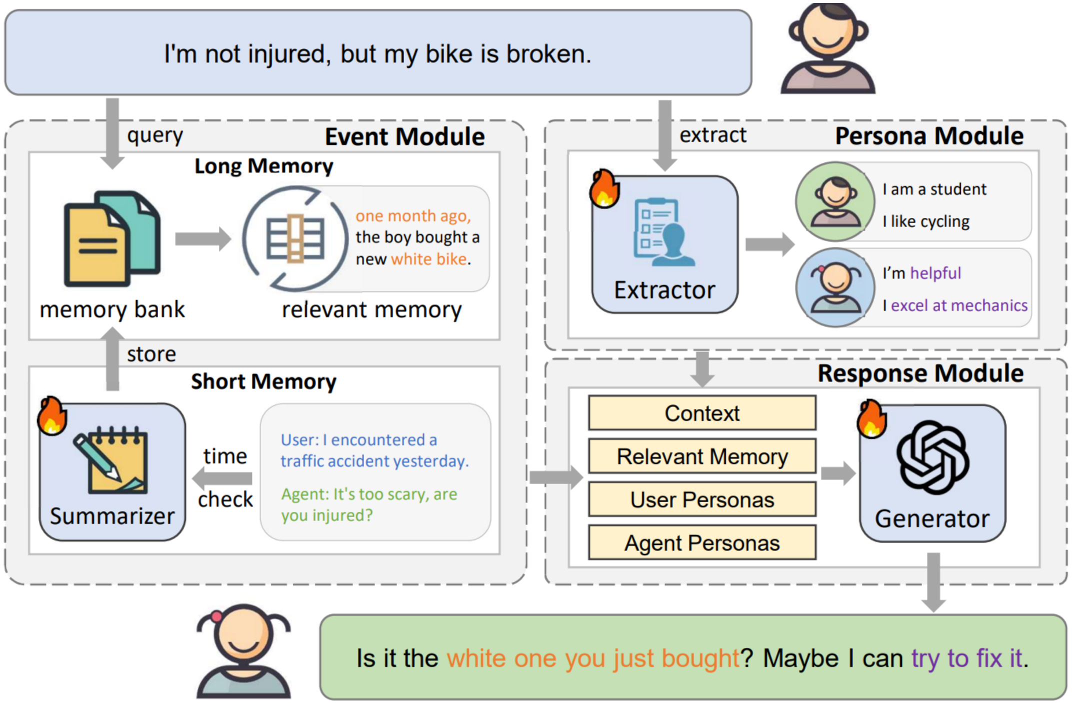
**Long-term Dialogue Agent (LD-Agent)**

- ❑ Comprehensive information

  (Personas & Historical Events)

- ❑ Disentangled Tuning
- ❑ Easy-to-transfer

  (*models & domains & tasks*)



*Li et al., 2024. "Hello Again! LLM-powered Personalized Agent for Long-term Dialogue" (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**

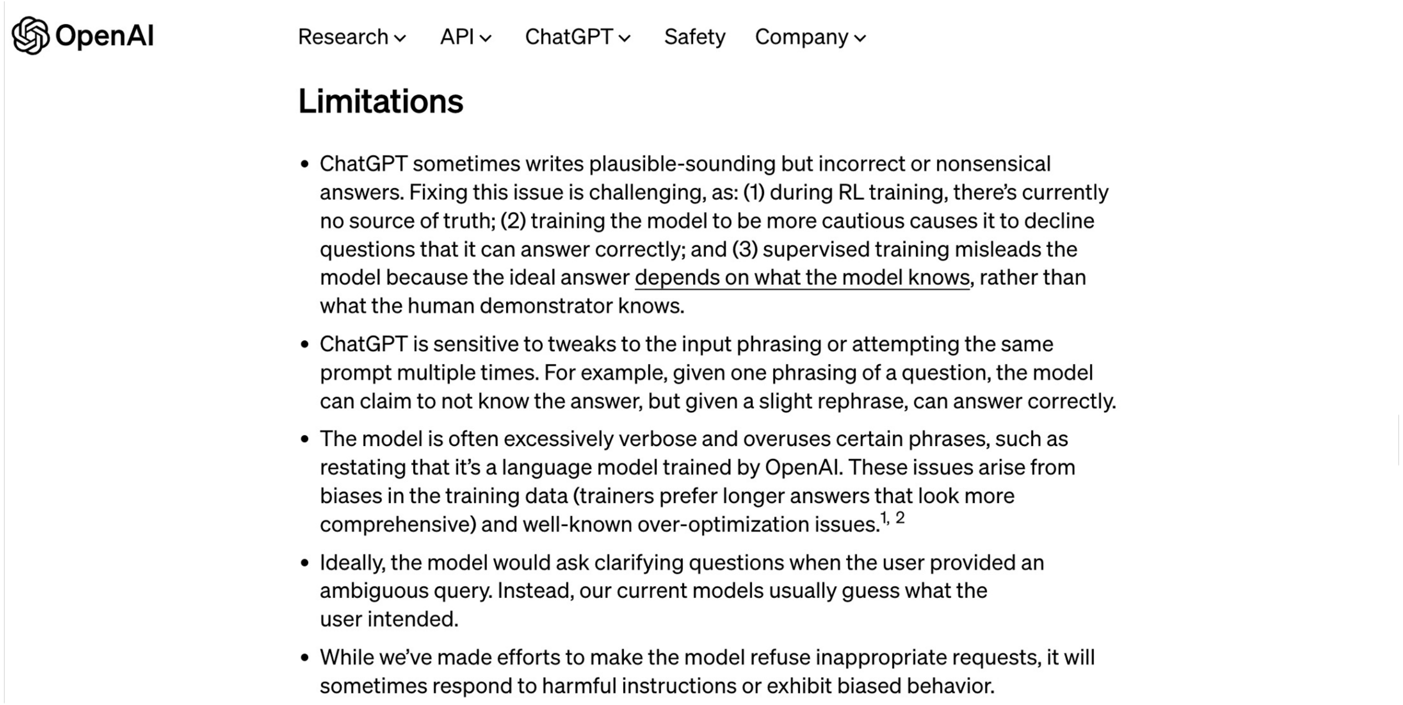
**Planning**

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

**Action**

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

# Limitations of LLM-based Conversational Systems

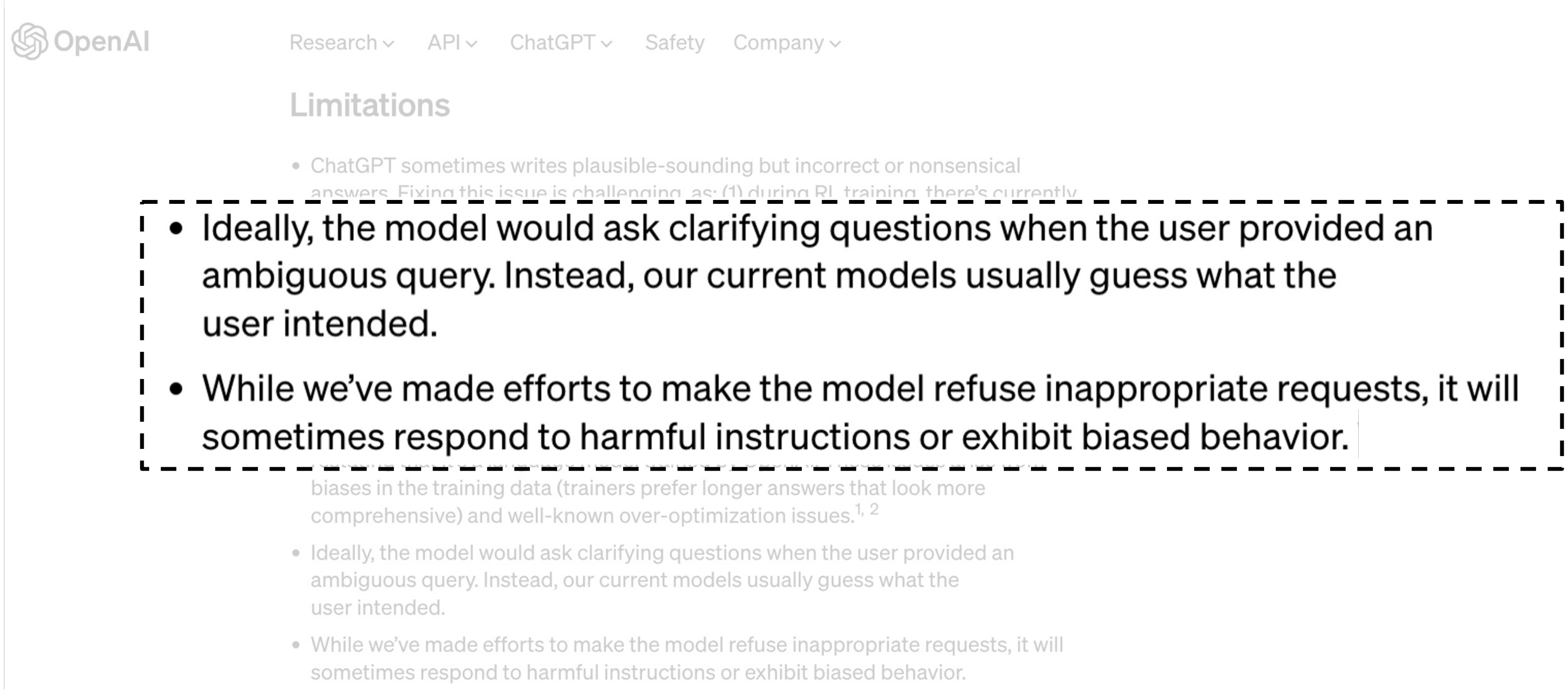


## Limitations

- 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



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**

NUS National University of Singapore

## Anticipation

To anticipate future impacts on the task or human users.

## Initiative

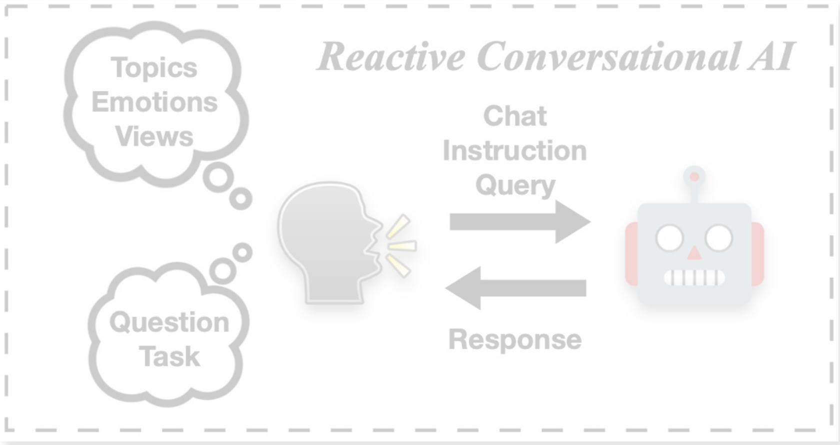
To take fine-grained and diverse initiative behaviours.

## Planning

To effectively and efficiently guide the conversation towards the goal.

*__Yang Deng__, Wenqiang Lei, Minlie Huang, Tat-Seng Chua.* ***Goal Awareness for Conversational AI: Proactivity, Non-collaborativity, and Beyond. ACL 2023 Tutorial.***
*__Yang Deng__, 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



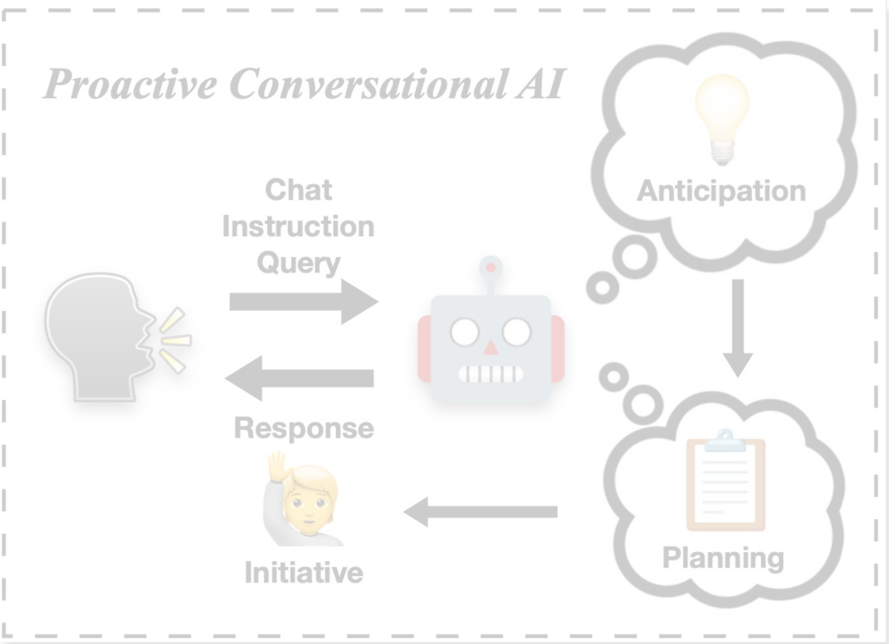
*Reactive Conversational AI*

Topics
Emotions
Views

Question
Task

Chat
Instruction
Query

Response



*Proactive Conversational AI*

Chat
Instruction
Query

Response

Initiative

Anticipation

Planning

***Yang Deng***, *Wenqiang Lei, Minlie Huang, Tat-Seng Chua.* ***Goal Awareness for Conversational AI: Proactivity, Non-collaborativity, and Beyond. ACL 2023 Tutorial.***
***Yang Deng***, *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

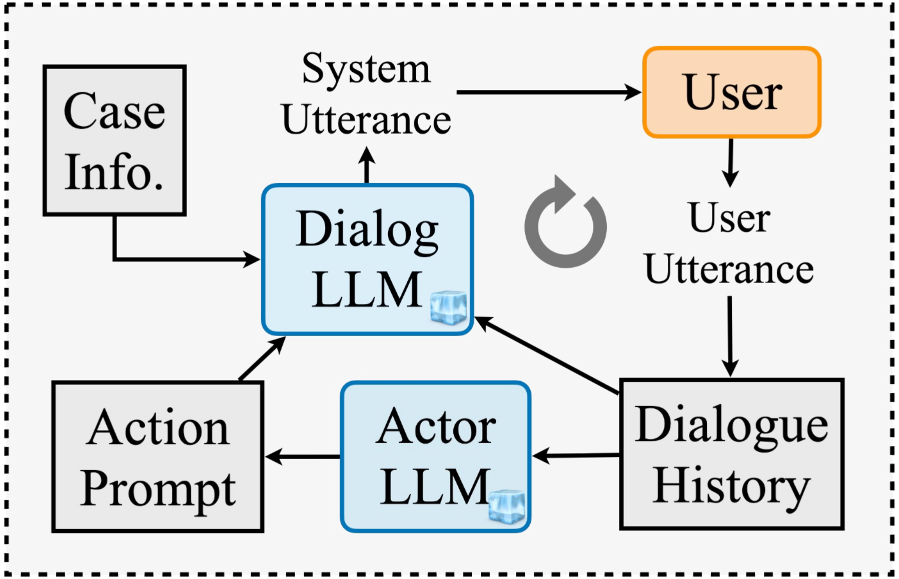
*Can LLM-based Conversational Agents effectively handle proactive dialogue problems without fine-tuning?*

❑ **Advantages of In-Context Learning**

- ✓ Training-free
- ✓ Easy-to-apply

➢ **Proactive Chain-of-Thought**

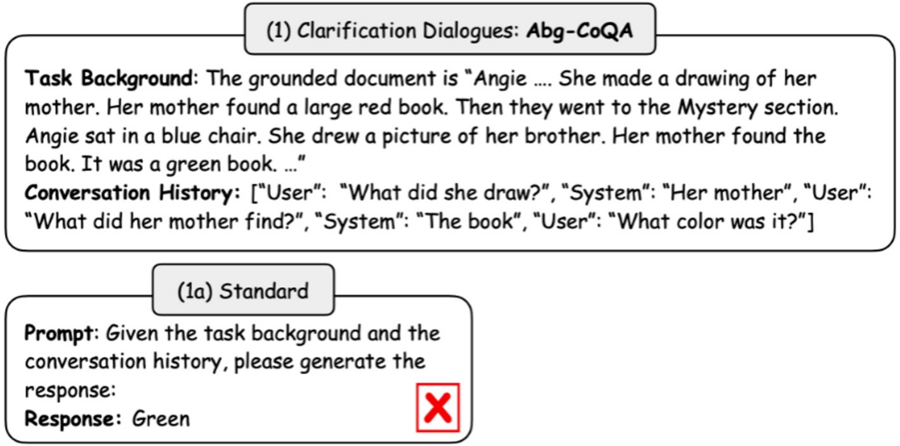
- ⋆ Fine-grained **Initiative**
- ⋆ Intermediate Reasoning



# Proactive Chain-of-Thought Prompting (ProCoT)

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

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



(1) Clarification Dialogues: **Abg-CoQA**

**Task Background**: The grounded document is "Angie .... She made a drawing of her mother. Her mother found a large red book. Then they went to the Mystery section. Angie sat in a blue chair. She drew a picture of her brother. Her mother found the book. It was a green book. ..."
**Conversation History**: ["User": "What did she draw?", "System": "Her mother", "User": "What did her mother find?", "System": "The book", "User": "What color was it?"]

(1a) Standard

**Prompt**: Given the task background and the conversation history, please generate the response:
**Response**: Green ☒

**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).***

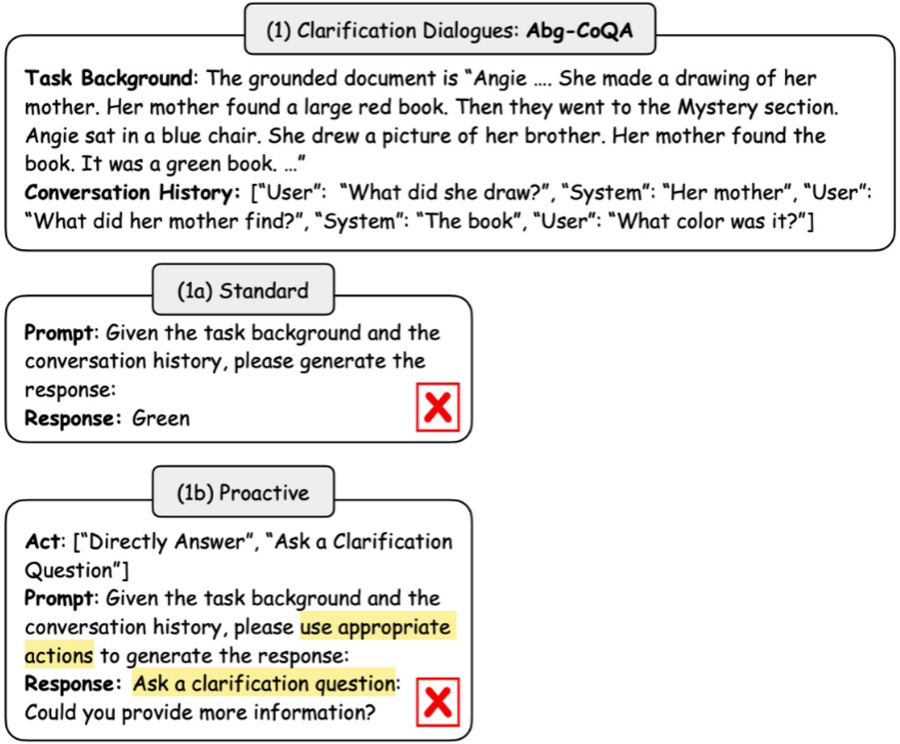
# Proactive Chain-of-Thought Prompting (ProCoT)

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

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

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

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



(1) Clarification Dialogues: **Abg-CoQA**

**Task Background**: The grounded document is "Angie .... She made a drawing of her mother. Her mother found a large red book. Then they went to the Mystery section. Angie sat in a blue chair. She drew a picture of her brother. Her mother found the book. It was a green book. ..."
**Conversation History**: ["User": "What did she draw?", "System": "Her mother", "User": "What did her mother find?", "System": "The book", "User": "What color was it?"]

(1a) Standard

**Prompt**: Given the task background and the conversation history, please generate the response:
**Response**: Green ☒

(1b) Proactive

**Act**: ["Directly Answer", "Ask a Clarification Question"]
**Prompt**: Given the task background and the conversation history, please use appropriate actions to generate the response:
**Response**: Ask a clarification question: Could you provide more information? ☒

*__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

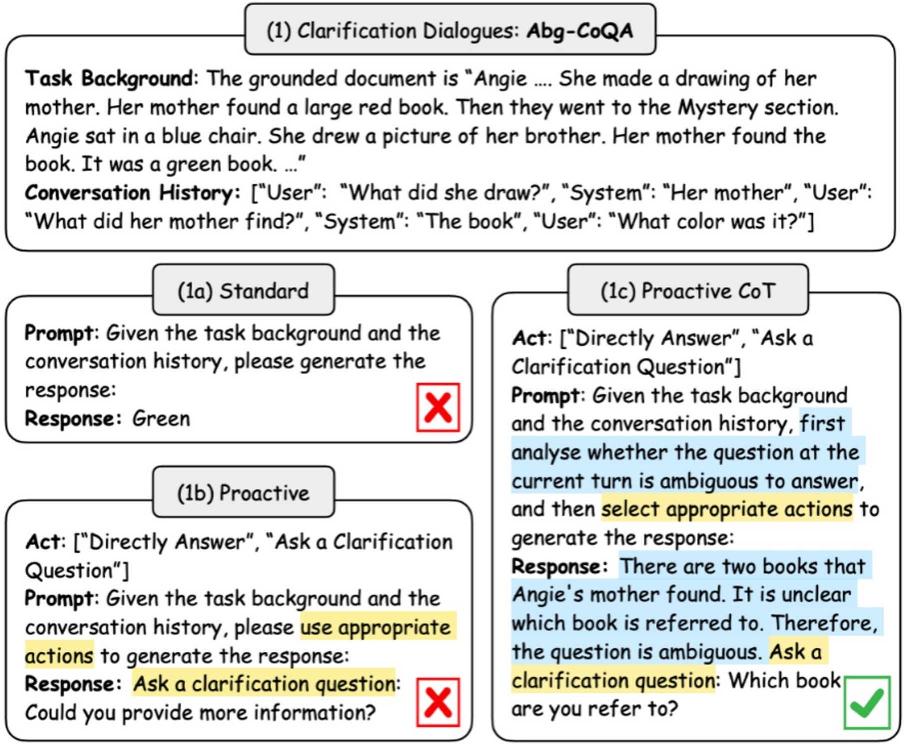
$$p(r|\mathcal{D},\mathcal{C})$$

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

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

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

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



(1) Clarification Dialogues: **Abg-CoQA**

**Task Background**: The grounded document is "Angie .... She made a drawing of her mother. Her mother found a large red book. Then they went to the Mystery section. Angie sat in a blue chair. She drew a picture of her brother. Her mother found the book. It was a green book. ..."
**Conversation History**: ["User": "What did she draw?", "System": "Her mother", "User": "What did her mother find?", "System": "The book", "User": "What color was it?"]

(1a) Standard

**Prompt**: Given the task background and the conversation history, please generate the response:
**Response**: Green ✗

(1b) Proactive

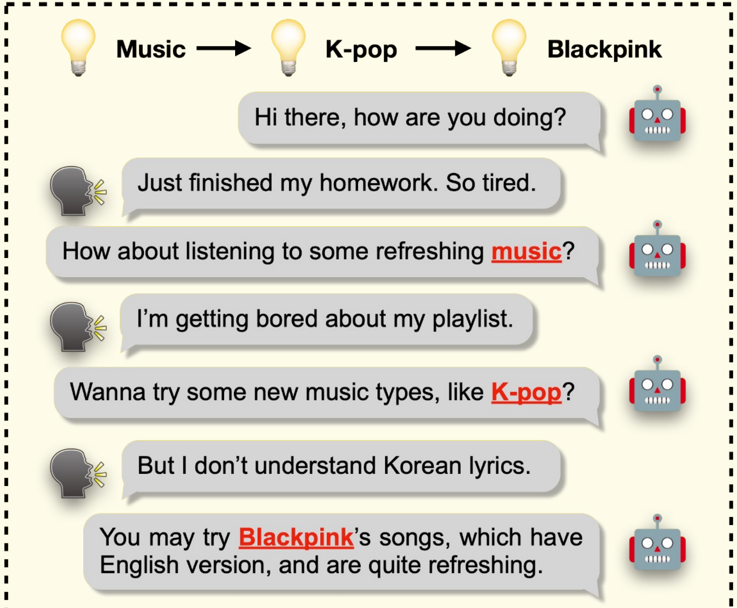
**Act**: ["Directly Answer", "Ask a Clarification Question"]
**Prompt**: Given the task background and the conversation history, please use appropriate actions to generate the response:
**Response**: Ask a clarification question: Could you provide more information? ✗

(1c) Proactive CoT

**Act**: ["Directly Answer", "Ask a Clarification Question"]
**Prompt**: Given the task background and the conversation history, first analyse whether the question at the current turn is ambiguous to answer, and then select appropriate actions to generate the response:
**Response**: There are two books that Angie's mother found. It is unclear which book is referred to. Therefore, the question is ambiguous. Ask a clarification question: Which book are you refer to? ✓

***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



Hi there, how are you doing?

Just finished my homework. So tired.

How about listening to some refreshing **music**?

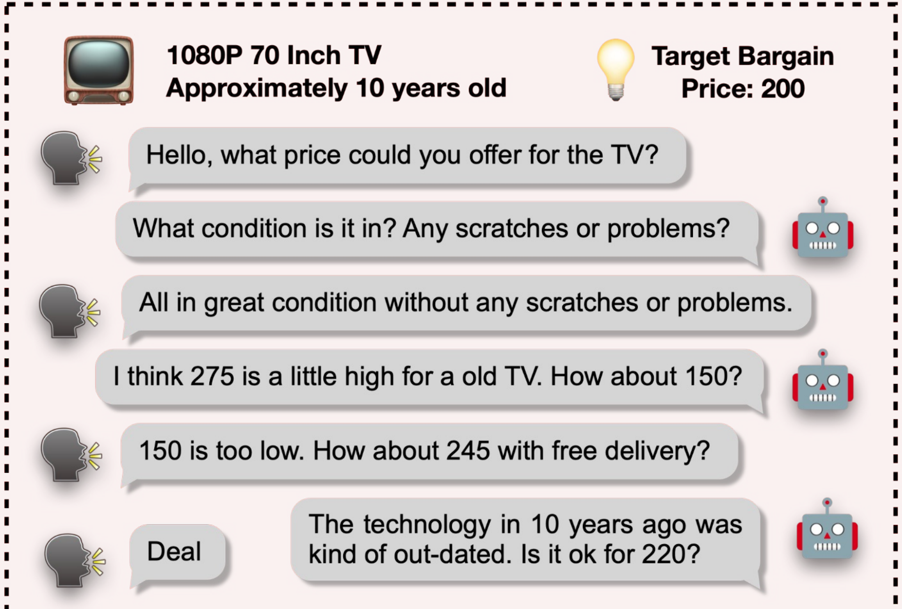
I'm getting bored about my playlist.

Wanna try some new music types, like **K-pop**?

But I don't understand Korean lyrics.

You may try **Blackpink**'s songs, which have English version, and are quite refreshing.

**Target-guided Open-domain Dialogues**



Hello, what price could you offer for the TV?

What condition is it in? Any scratches or problems?

All in great condition without any scratches or problems.

I think 275 is a little high for a old TV. How about 150?

150 is too low. How about 245 with free delivery?

The technology in 10 years ago was kind of out-dated. Is it ok for 220?

Deal

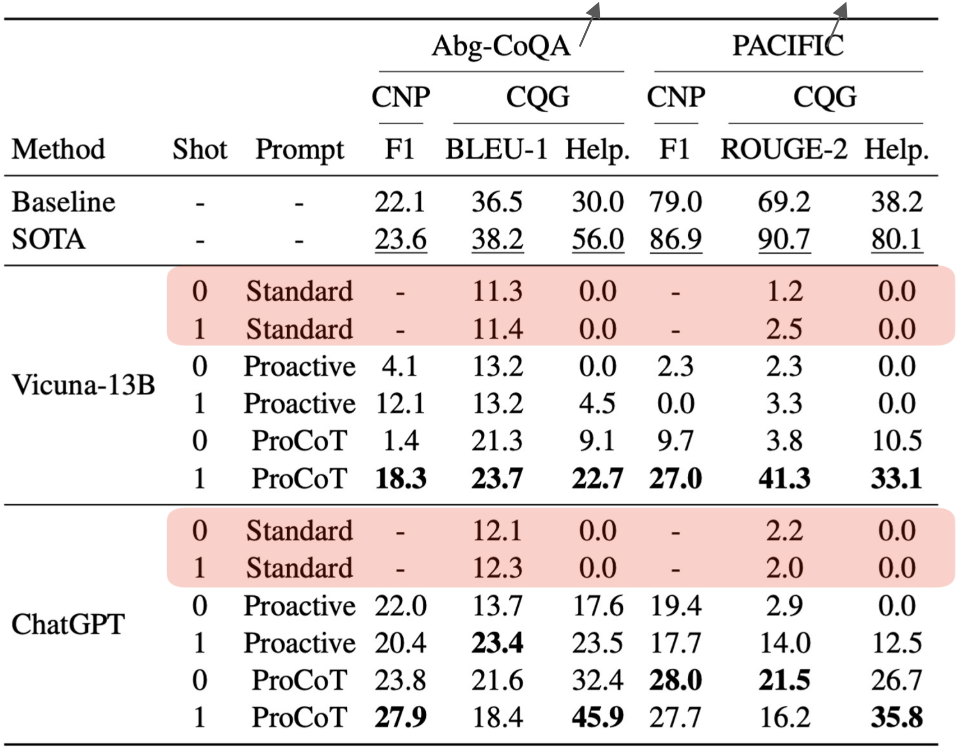
**Non-collaborative Dialogues**

***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



| | | | Abg-CoQA | | | PACIFIC | | |
|---|---|---|---|---|---|---|---|---|
| | | | CNP | CQG | | CNP | CQG | |
| Method | Shot | Prompt | 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** |

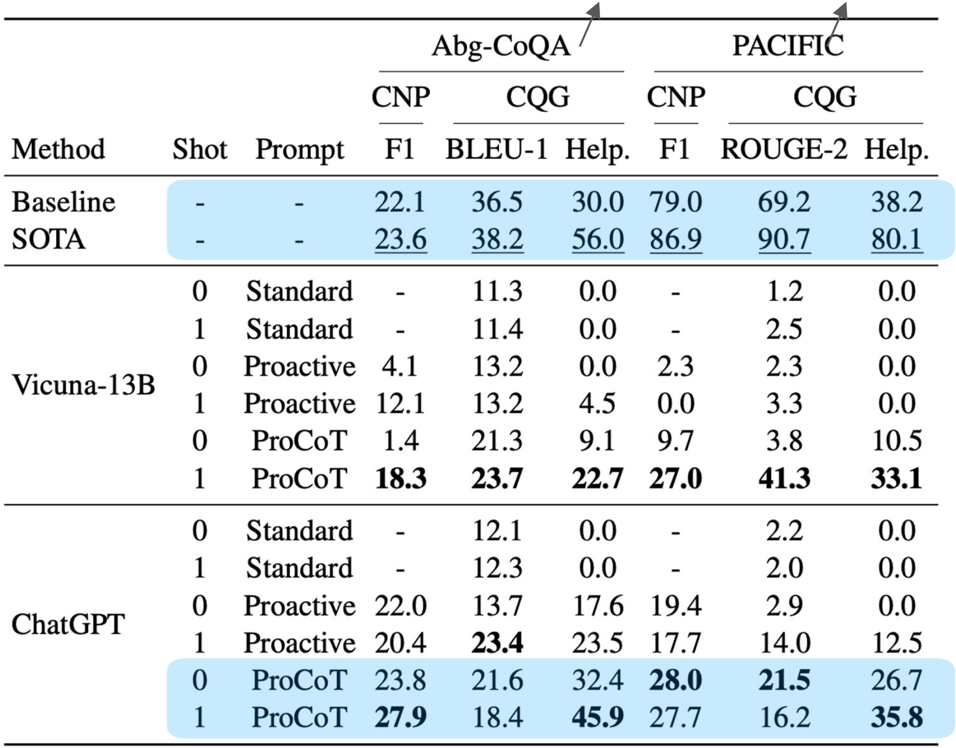
**LLMs barely ask clarification questions.**

***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



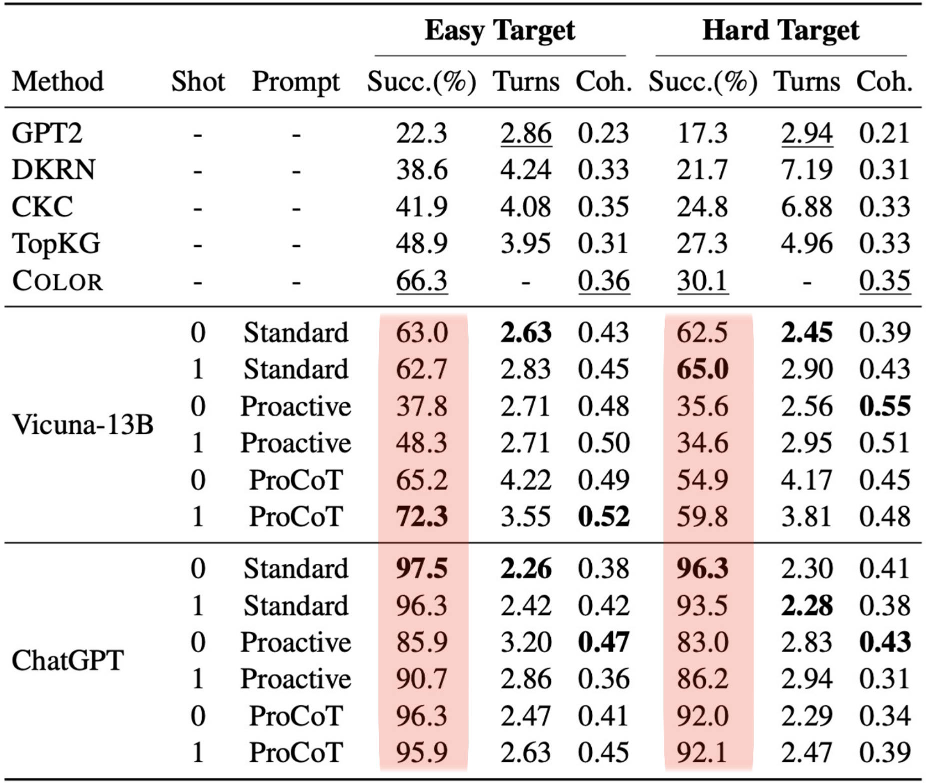
| | | | Abg-CoQA | | | PACIFIC | | |
|---|---|---|---|---|---|---|---|---|
| | | | CNP | CQG | | CNP | CQG | |
| Method | Shot | Prompt | F1 | BLEU-1 | Help. | F1 | ROUGE-2 | Help. |
| Baseline | - | - | 22.1 | 36.5 | 30.0 | 79.0 | 69.2 | 38.2 |
| SOTA | - | - | <u>23.6</u> | <u>38.2</u> | <u>56.0</u> | <u>86.9</u> | <u>90.7</u> | <u>80.1</u> |
| 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.**

***<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

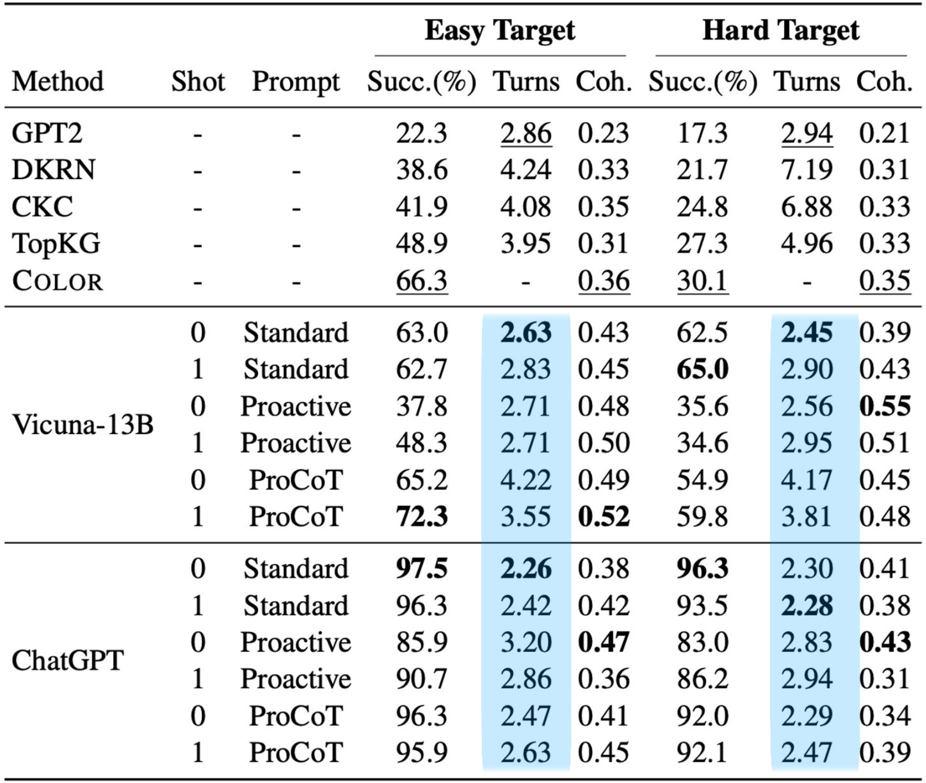


| Method | Shot | Prompt | Easy Target Succ.(%) | Easy Target Turns | Easy Target Coh. | Hard Target Succ.(%) | Hard Target Turns | Hard Target Coh. |
|---|---|---|---|---|---|---|---|---|
| GPT2 | - | - | 22.3 | 2.86 | 0.23 | 17.3 | 2.94 | 0.21 |
| DKRN | - | - | 38.6 | 4.24 | 0.33 | 21.7 | 7.19 | 0.31 |
| 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 | - | - | 66.3 | - | 0.36 | 30.1 | - | 0.35 |
| 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.**

*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



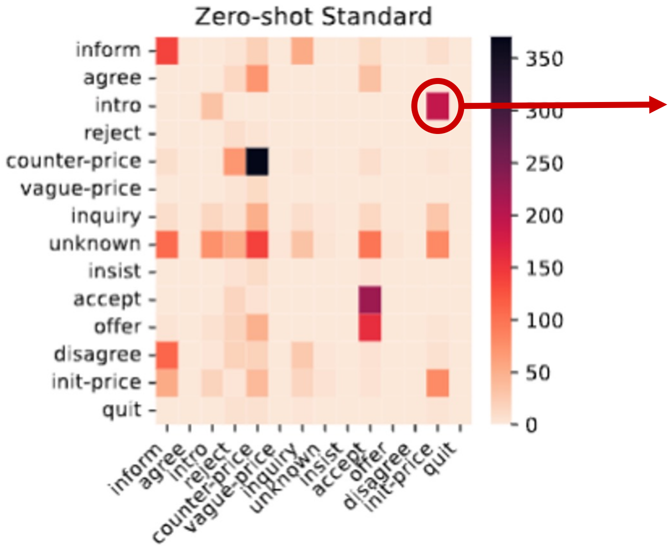
| | | | Easy Target | | | Hard Target | | |
|---|---|---|---|---|---|---|---|---|
| Method | Shot | Prompt | Succ.(%) | Turns | Coh. | Succ.(%) | Turns | Coh. |
| GPT2 | - | - | 22.3 | 2.86 | 0.23 | 17.3 | 2.94 | 0.21 |
| DKRN | - | - | 38.6 | 4.24 | 0.33 | 21.7 | 7.19 | 0.31 |
| CKC | - | - | 41.9 | 4.08 | 0.35 | 24.8 | 6.88 | 0.33 |
| TopKG | - | - | 48.9 | 3.95 | 0.31 | 27.3 | 4.96 | 0.33 |
| COLOR | - | - | 66.3 | - | 0.36 | 30.1 | - | 0.35 |
| 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.**

*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)

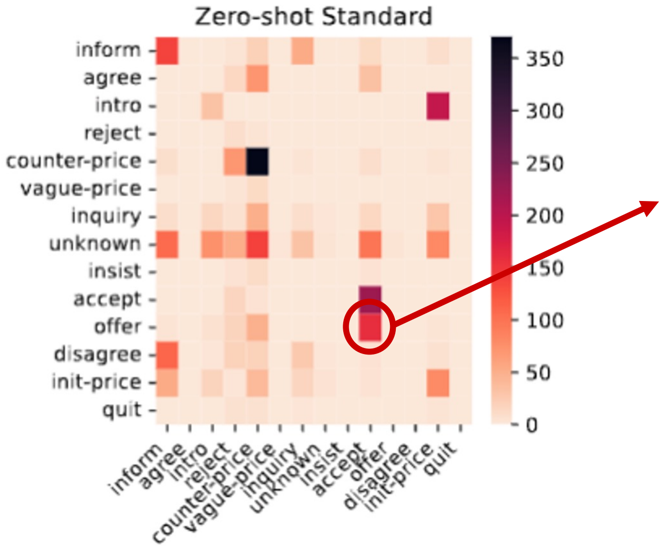


- Tends to propose the initial price (**init-price**) instead of greetings (**intro**) at the beginning.
- Often directly accepts the buyer's offer (**accept**) when it is supposed to offer another price for negotiation (**offer**).
- Tends to propose a counter price (**counter-price**) to make compromise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

**Relationships between reference and predicted negotiation strategies.**

*__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)

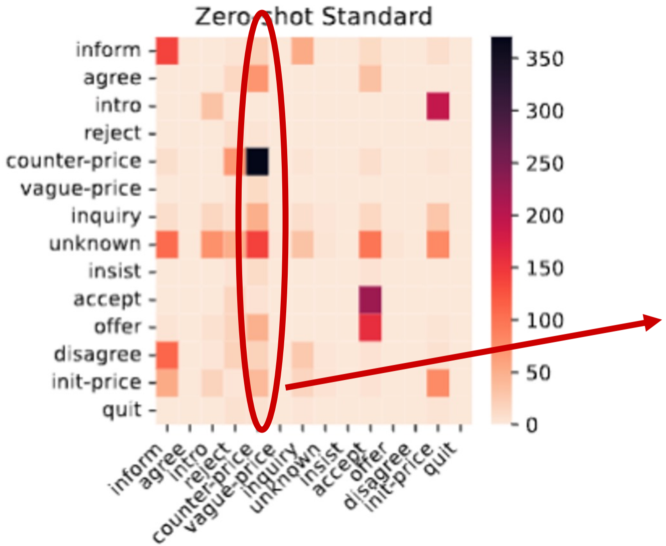


**Relationships between reference and predicted negotiation strategies.**

- ❑ 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)

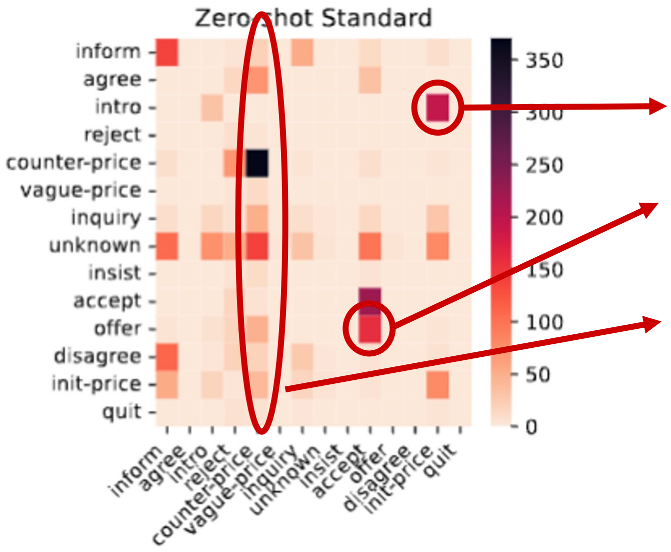


**Relationships between reference and predicted negotiation strategies.**

- ❑ 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.

***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).***

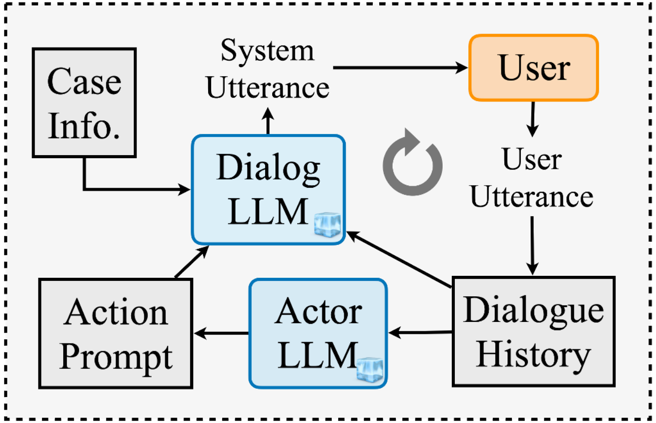
# Lessons Learned from the Evaluation

- ❑ **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.***

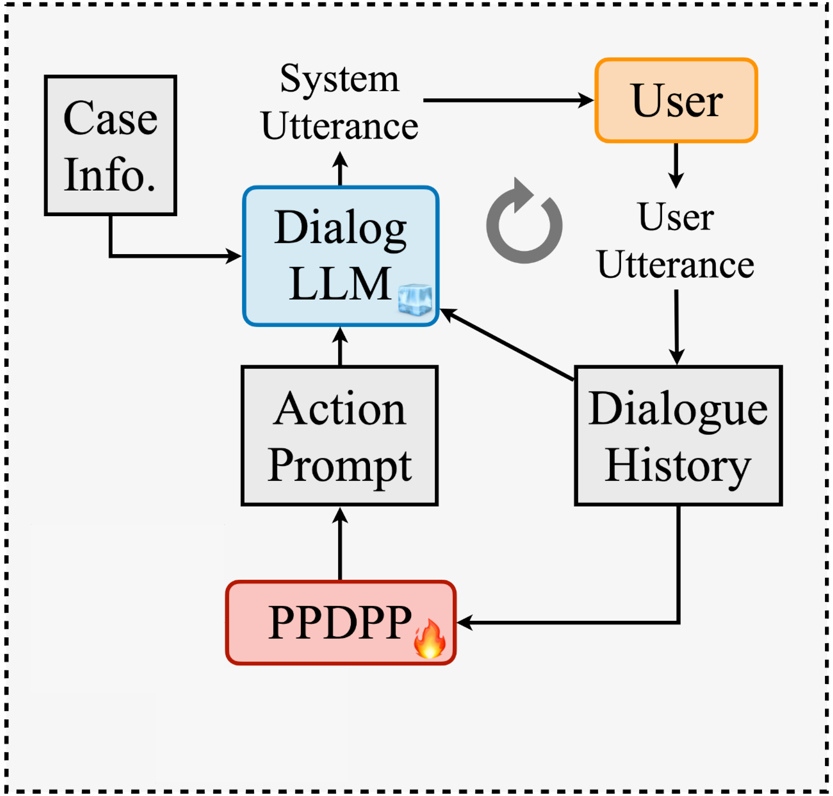
# Policy Network – Plug-and-Play Dialogue Policy Planner

- ☐ 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.

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



*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.***

# Reward Function – Learning from AI Feedback

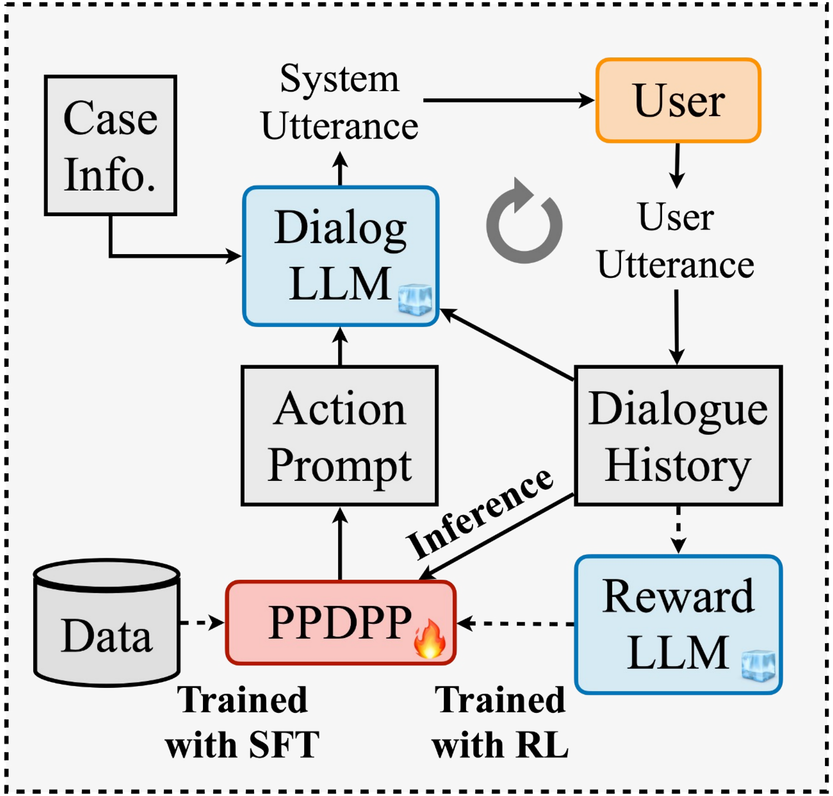
- An LLM as the reward model to assess the goal achievement and provide **goal-oriented AI feedback**.

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

- 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!***

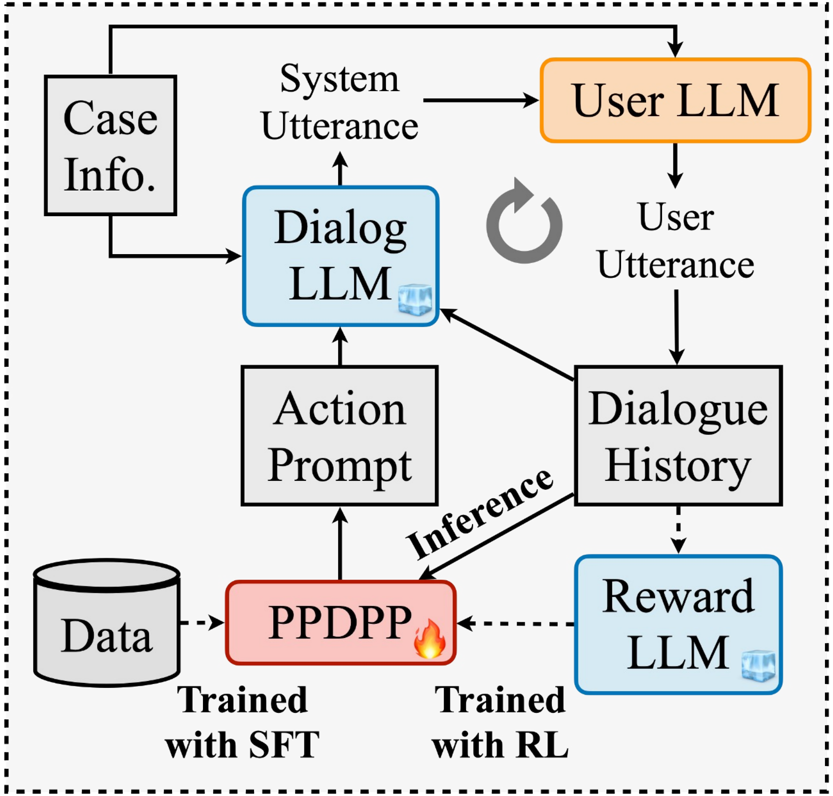


Yang Deng, Wenxuan Zhang, Wai Lam, See-Kiong Ng, Tat-Seng Chua. **Plug-and-Play Policy Planner for Large Language Model Powered Dialogue Agents. In ICLR 2024.**

# State Transition – Multi-agent Simulation

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

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



*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

**Negotiation — Buyer**

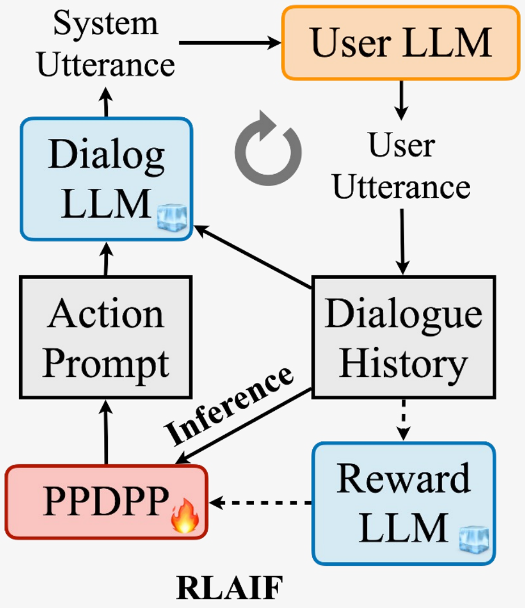
**Task Instruction:** You are the buyer who is trying to buy the [item name] with the price of [buyer target price].

**Counseling — Therapist**

**Task Instruction:** You are the therapist who is trying to help the patient reduce their emotional distress and help them understand and work through the challenges.

**Tutoring — Teacher**

**Task Instruction:** You are the teacher who is trying to teach the student to translate "[exercise]" into Italian.



**Negotiation — Seller**

**User Profile:** You are the **seller** who is trying to sell the [item name] with the price of [seller target price]. Product description: [item description]. ...

**Counseling — Patient**

**User Profile:** You are the patient who is looking for the help from the therapist, because you have the emotional issue about [emotion type] regarding [problem type]. ...

**Tutoring — Student**

**User Profile:** You are the student who is trying to translate an English sentence into Italian. ...

*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

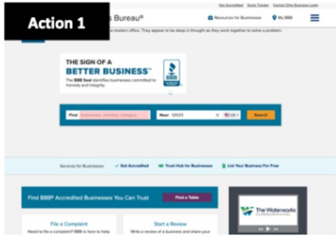
**Web Agents** aim to accomplish web navigation tasks defined in natural language, such as booking tickets, through **multi-step interactions with the web-grounded environment**.

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

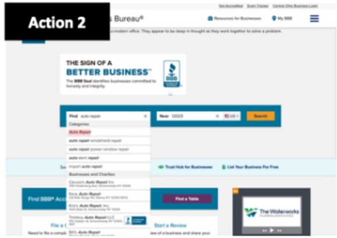
**Action Sequence:**

| | 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 |

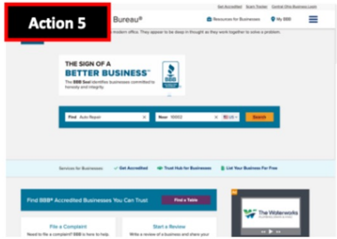
**Webpage Snapshots:**



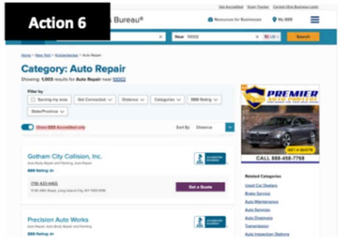
`<input name="find_text" type="search">`



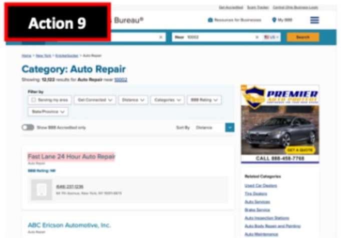
`<em>Auto Repair</em>`



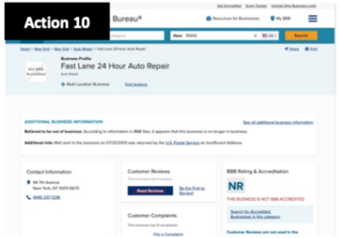
`<button>Search</button>`



`<button>Show BBB Accredited only</button>`



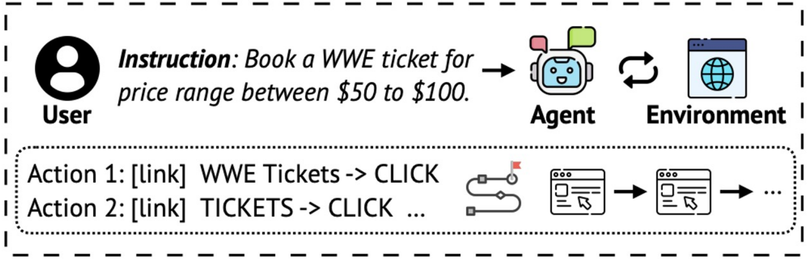
`<span>Fast Lane 24 Hour Auto Repair</span>`



`<a href=”link:XXX">Read Reviews</a>`

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

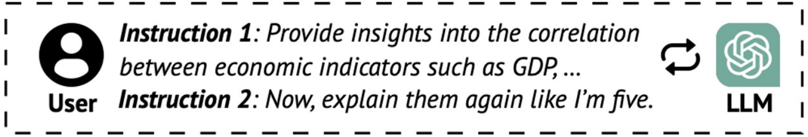
# Web Agents become Conversational?



**Instruction**: Book a WWE ticket for price range between $50 to $100.

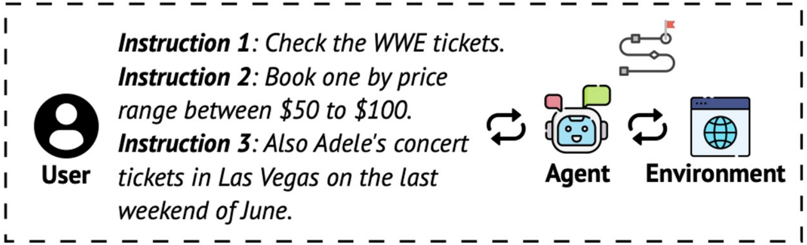
Action 1: [link] WWE Tickets -> CLICK
Action 2: [link] TICKETS -> CLICK ...

(a) Web Navigation



**Instruction 1**: Provide insights into the correlation between economic indicators such as GDP, ...
**Instruction 2**: Now, explain them again like I'm five.

(b) Conversational Information Seeking



**Instruction 1**: Check the WWE tickets.
**Instruction 2**: Book one by price range between $50 to $100.
**Instruction 3**: Also Adele's concert tickets in Las Vegas on the last weekend of June.

(c) Conversational Web Navigation

## Web Navigation

→ Single-turn User Instruction

→ Multi-step Environment Interaction

## Conversational Information Seeking

→ Multi-turn User Instruction

→ No/Single-step Environment Interaction

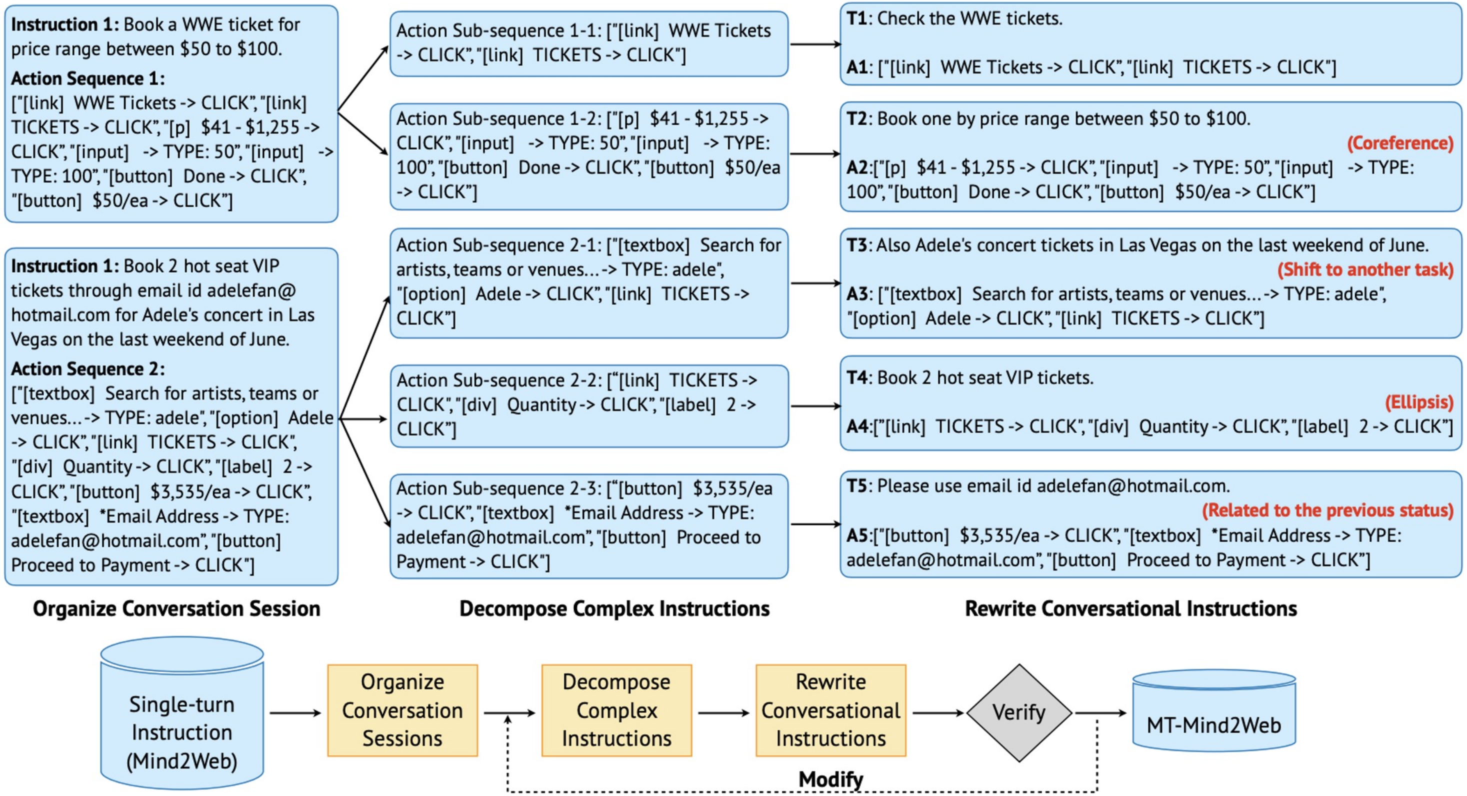
## Conversational Web Navigation

→ Multi-turn User Instruction

→ Multi-step Environment Interaction

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

# Constructing the MT-Mind2Web Dataset



**Instruction 1:** Book a WWE ticket for price range between $50 to $100.

**Action Sequence 1:**
["[link]  WWE Tickets -> CLICK”, "[link]  TICKETS -> CLICK”, "[p]  $41 - $1,255 -> CLICK”, "[input]  -> TYPE: 50”, "[input]  -> TYPE: 100”, "[button]  Done -> CLICK”, "[button]  $50/ea -> CLICK”]

**Instruction 1:** Book 2 hot seat VIP tickets through email id adelefan@hotmail.com for Adele's concert in Las Vegas on the last weekend of June.

**Action Sequence 2:**
["[textbox]  Search for artists, teams or venues... -> TYPE: adele”, "[option]  Adele -> CLICK”, "[link]  TICKETS -> CLICK”, "[div]  Quantity -> CLICK”, "[label]  2 -> CLICK”, "[button]  $3,535/ea -> CLICK”, "[textbox]  *Email Address -> TYPE: adelefan@hotmail.com”, "[button]  Proceed to Payment -> CLICK”]

Action Sub-sequence 1-1: ["[link]  WWE Tickets -> CLICK”, "[link]  TICKETS -> CLICK”]

Action Sub-sequence 1-2: ["[p]  $41 - $1,255 -> CLICK”, "[input]  -> TYPE: 50”, "[input]  -> TYPE: 100”, "[button]  Done -> CLICK”, "[button]  $50/ea -> CLICK”]

Action Sub-sequence 2-1: ["[textbox]  Search for artists, teams or venues... -> TYPE: adele”, "[option]  Adele -> CLICK”, "[link]  TICKETS -> CLICK”]

Action Sub-sequence 2-2: [“[link]  TICKETS -> CLICK”, "[div]  Quantity -> CLICK”, "[label]  2 -> CLICK”]

Action Sub-sequence 2-3: [“[button]  $3,535/ea -> CLICK”, "[textbox]  *Email Address -> TYPE: adelefan@hotmail.com”, "[button]  Proceed to Payment -> CLICK”]

**T1**: Check the WWE tickets.

**A1**: ["[link]  WWE Tickets -> CLICK”, "[link]  TICKETS -> CLICK”]

**T2**: Book one by price range between $50 to $100.

**(Coreference)**

**A2**:["[p]  $41 - $1,255 -> CLICK”, "[input]  -> TYPE: 50”, "[input]  -> TYPE: 100”, "[button]  Done -> CLICK”, "[button]  $50/ea -> CLICK”]

**T3**: Also Adele's concert tickets in Las Vegas on the last weekend of June.

**(Shift to another task)**

**A3**: ["[textbox]  Search for artists, teams or venues... -> TYPE: adele”, "[option]  Adele -> CLICK”, "[link]  TICKETS -> CLICK”]

**T4**: Book 2 hot seat VIP tickets.

**(Ellipsis)**

**A4**:["[link]  TICKETS -> CLICK”, "[div]  Quantity -> CLICK”, "[label]  2 -> CLICK”]

**T5**: Please use email id adelefan@hotmail.com.

**(Related to the previous status)**

**A5**:["[button]  $3,535/ea -> CLICK”, "[textbox]  *Email Address -> TYPE: adelefan@hotmail.com”, "[button]  Proceed to Payment -> CLICK”]

**Organize Conversation Session**

**Decompose Complex Instructions**

**Rewrite Conversational Instructions**

Single-turn Instruction (Mind2Web)
Organize Conversation Sessions
Decompose Complex Instructions
Rewrite Conversational Instructions
Verify
MT-Mind2Web
Modify

*Deng et al., 2024. “On the Multi-turn Instruction Following of Conversational Web Agents” (ACL ‘24)*

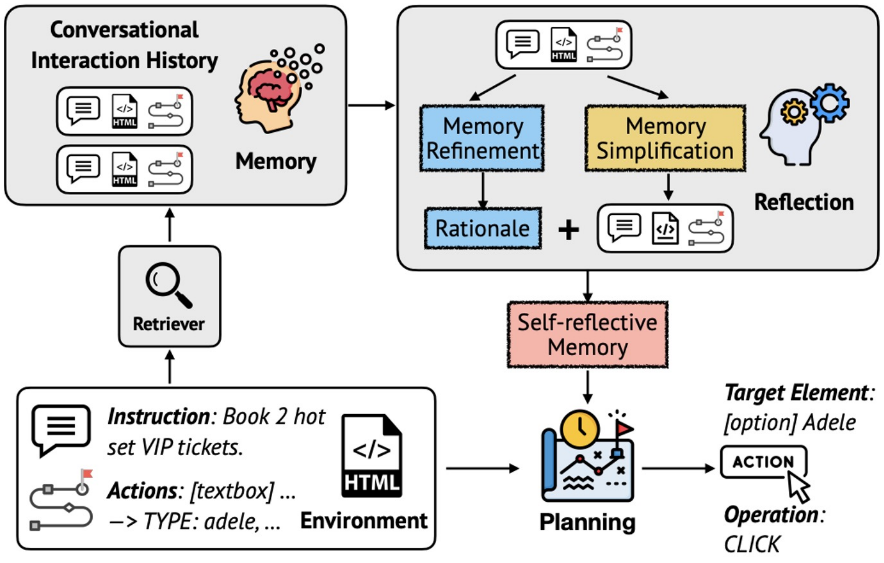
# Challenges in Conversational Web Agents

**<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" (ACL '24)***

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


Conversational Interaction History
Memory
Retriever
Instruction: Book 2 hot set VIP tickets.
Actions: [textbox] ... –> TYPE: adele, ...
Environment
HTML
Memory Refinement
Memory Simplification
Reflection
Rationale
+
Self-reflective Memory
Planning
Target Element: [option] Adele
ACTION
Operation: CLICK

## 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

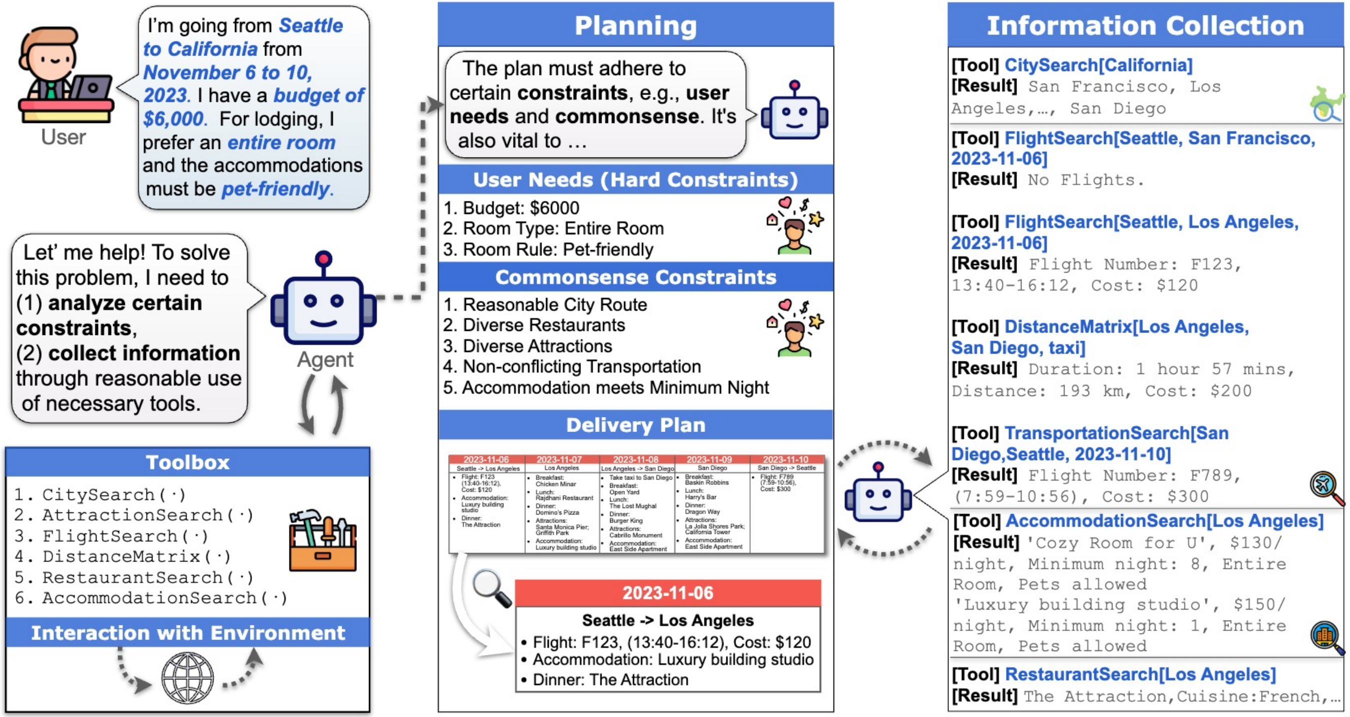
## Planning Module

→ **Memory-augmented Planning** to decide the next action to take

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

# Travel Planning Agents

**Travel Planning Agents** aim to accomplish travel planning tasks defined in natural language, through **employing various search tools to gather information that satisfied the user's needs**.



I'm going from ***Seattle to California*** from ***November 6 to 10, 2023***. I have a ***budget of $6,000***. For lodging, I prefer an ***entire room*** and the accommodations must be ***pet-friendly***.

Let' me help! To solve this problem, I need to (1) **analyze certain constraints**, (2) **collect information** through reasonable use of necessary tools.

**Toolbox**

1. CitySearch(·)
2. AttractionSearch(·)
3. FlightSearch(·)
4. DistanceMatrix(·)
5. RestaurantSearch(·)
6. AccommodationSearch(·)

**Interaction with Environment**

**Planning**

The plan must adhere to certain **constraints**, e.g., **user needs** and **commonsense**. It's also vital to …

**User Needs (Hard Constraints)**

1. Budget: $6000
2. Room Type: Entire Room
3. Room Rule: Pet-friendly

**Commonsense Constraints**

1. Reasonable City Route
2. Diverse Restaurants
3. Diverse Attractions
4. Non-conflicting Transportation
5. Accommodation meets Minimum Night

**Delivery Plan**

| 2023-11-06 | 2023-11-07 | 2023-11-08 | 2023-11-09 | 2023-11-10 |
|---|---|---|---|---|
| Seattle -> Los Angeles | Los Angeles | Los Angeles -> San Diego | San Diego | San Diego -> Seattle |
| • Flight: F123 (13:40-16:12), Cost: $120<br>• Accommodation: Luxury building studio<br>• Dinner: The Attraction | • Breakfast: Chicken Minar<br>• Lunch: Rajdhani Restaurant<br>• Dinner: Domino's Pizza<br>• Attractions: Santa Monica Pier; Griffith Park<br>• Accommodation: Luxury building studio | • Take taxi to San Diego<br>• Breakfast: Open Yard<br>• Lunch: The Lost Mughal<br>• Dinner: Burger King<br>• Attractions: Cabrillo Monument<br>• Accommodation: East Side Apartment | • Breakfast: Baskin Robbins<br>• Lunch: Harry's Bar<br>• Dinner: Dragon Way<br>• Attractions: La Jolla Shores Park; California Tower<br>• Accommodation: East Side Apartment | • Flight: F789 (7:59-10:56), Cost: $300 |

**2023-11-06**

**Seattle -> Los Angeles**

- Flight: F123, (13:40-16:12), Cost: $120
- Accommodation: Luxury building studio
- Dinner: The Attraction

**Information Collection**

**[Tool]** CitySearch[California]
**[Result]** San Francisco, Los Angeles,…, San Diego

**[Tool]** FlightSearch[Seattle, San Francisco, 2023-11-06]
**[Result]** No Flights.

**[Tool]** FlightSearch[Seattle, Los Angeles, 2023-11-06]
**[Result]** Flight Number: F123, 13:40-16:12, Cost: $120

**[Tool]** DistanceMatrix[Los Angeles, San Diego, taxi]
**[Result]** Duration: 1 hour 57 mins, Distance: 193 km, Cost: $200

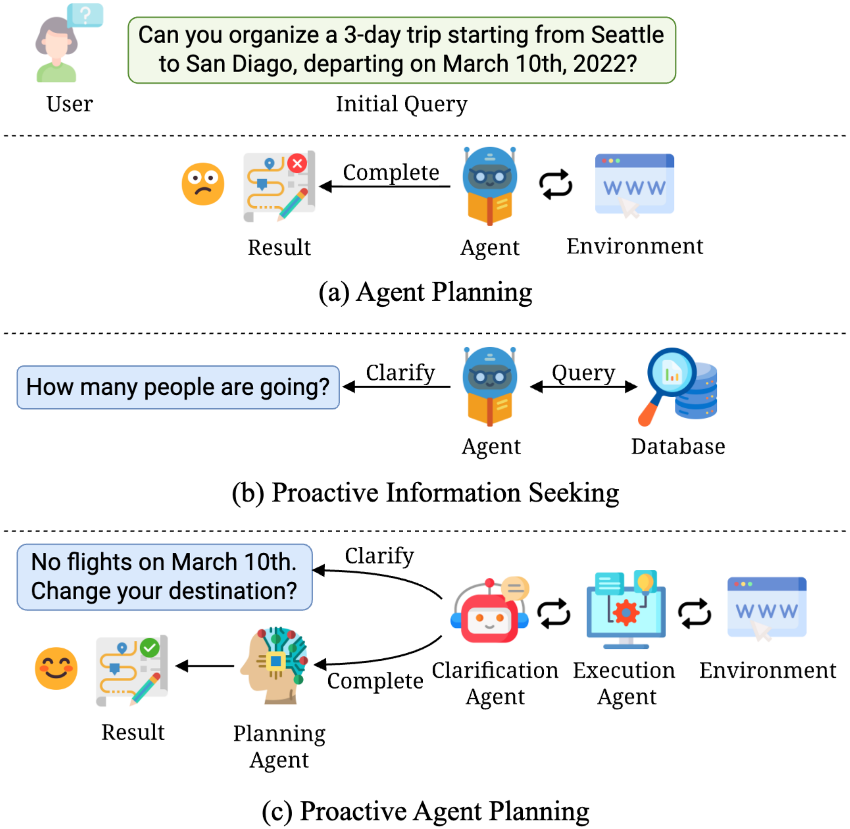
**[Tool]** TransportationSearch[San Diego,Seattle, 2023-11-10]
**[Result]** Flight Number: F789, (7:59-10:56), Cost: $300

**[Tool]** AccommodationSearch[Los Angeles]
**[Result]** 'Cozy Room for U', $130/ night, Minimum night: 8, Entire Room, Pets allowed 'Luxury building studio', $150/ night, Minimum night: 1, Entire Room, Pets allowed

**[Tool]** RestaurantSearch[Los Angeles]
**[Result]** The Attraction,Cuisine:French,…

*Xie et al., 2024. "TravelPlanner: A Benchmark for Real-World Planning with Language Agents" (ICML '24)*

# User Instructions are NOT always Clear!!


User
Can you organize a 3-day trip starting from Seattle to San Diago, departing on March 10th, 2022?
Initial Query
Complete
Result
Agent
Environment
WWW
How many people are going?
Clarify
Query
Agent
Database
No flights on March 10th. Change your destination?
Clarify
Complete
Result
Planning Agent
Clarification Agent
Execution Agent
Environment
WWW

(a) Agent Planning

(b) Proactive Information Seeking

(c) Proactive Agent Planning

## Agent Planning

→ One-time Interaction between User and Agent

→ User instructions could be unclear

## Proactive Information Seeking

→ Multi-turn User-Agent Interactions with Clarifications

→ Context-based Clarification

## Proactive Agent Planning

→ Multi-turn User-Agent Interactions with Clarifications

→ Clarification based on both Context and Environment

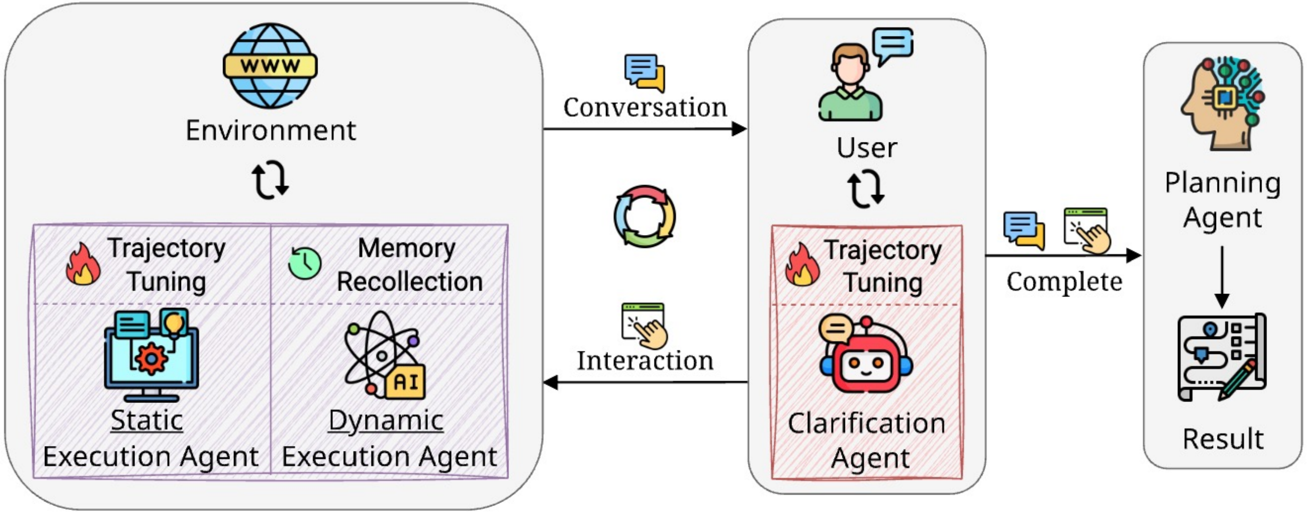
*Zhang et al., 2024. "Ask-before-Plan: Proactive Language Agents for Real-World Planning" (CoRR '24)*

# Problem Definition of Proactive Agent Planning

- ☐ **Clarification Need Prediction & Clarification Question Generation**

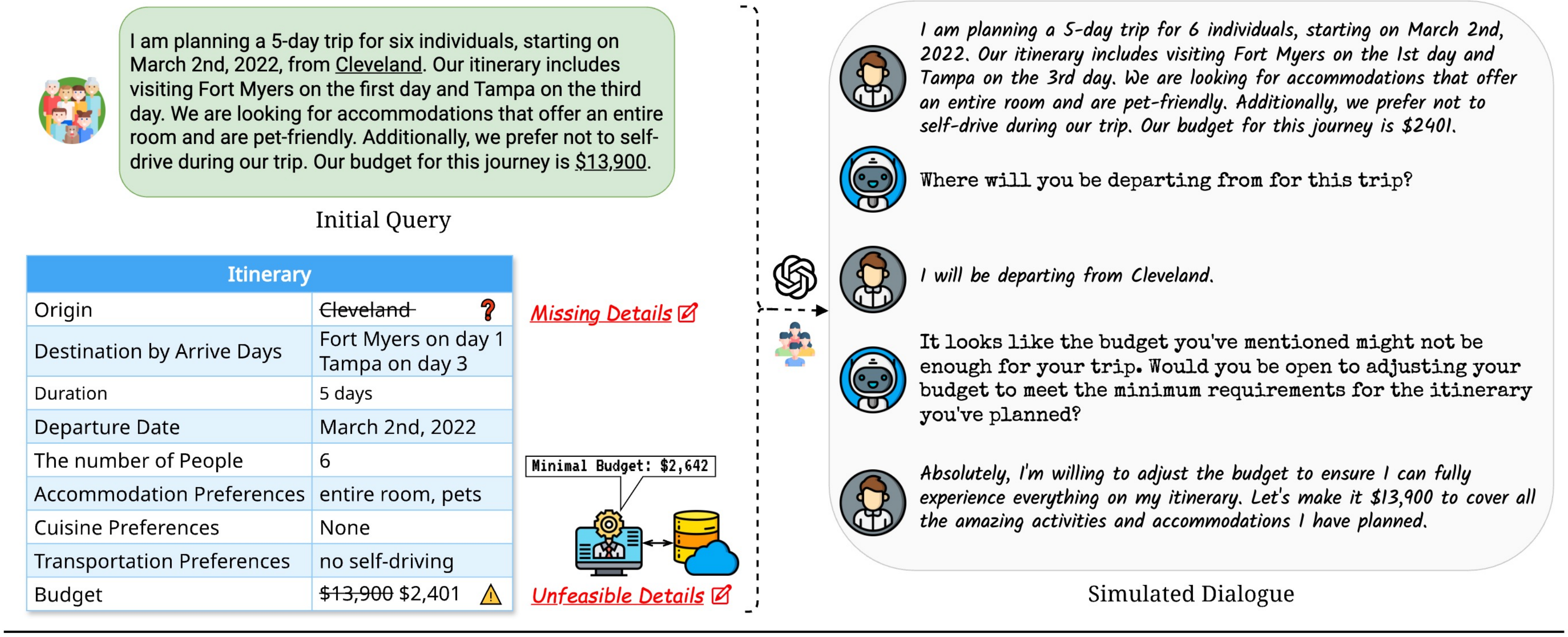
- ☐ **Tool Learning**

- ☐ **Agent Planning**



*Zhang et al., 2024. "Ask-before-Plan: Proactive Language Agents for Real-World Planning" (CoRR '24)*

# Constructing the Ask-before-Plan Dataset



I am planning a 5-day trip for six individuals, starting on March 2nd, 2022, from Cleveland. Our itinerary includes visiting Fort Myers on the first day and Tampa on the third day. We are looking for accommodations that offer an entire room and are pet-friendly. Additionally, we prefer not to self-drive during our trip. Our budget for this journey is $13,900.

Initial Query

| Itinerary | |
|---|---|
| Origin | ~~Cleveland~~ ❓ |
| Destination by Arrive Days | Fort Myers on day 1<br>Tampa on day 3 |
| Duration | 5 days |
| Departure Date | March 2nd, 2022 |
| The number of People | 6 |
| Accommodation Preferences | entire room, pets |
| Cuisine Preferences | None |
| Transportation Preferences | no self-driving |
| Budget | ~~$13,900~~ $2,401 ⚠ |

*Missing Details*

`Minimal Budget: $2,642`

*Unfeasible Details*

*I am planning a 5-day trip for 6 individuals, starting on March 2nd, 2022. Our itinerary includes visiting Fort Myers on the 1st day and Tampa on the 3rd day. We are looking for accommodations that offer an entire room and are pet-friendly. Additionally, we prefer not to self-drive during our trip. Our budget for this journey is $2401.*

Where will you be departing from for this trip?

*I will be departing from Cleveland.*

It looks like the budget you've mentioned might not be enough for your trip. Would you be open to adjusting your budget to meet the minimum requirements for the itinerary you've planned?

*Absolutely, I'm willing to adjust the budget to ensure I can fully experience everything on my itinerary. Let's make it $13,900 to cover all the amazing activities and accommodations I have planned.*

Simulated Dialogue

(1) Indefinite Detail Formation

(2) Proactive Conversation Generation

*Zhang et al., 2024. "Ask-before-Plan: Proactive Language Agents for Real-World Planning" (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**