Leveraging Knowledge Graphs and LLMs for Zero-Shot Conversational Movie Recommendation

Derek Yao and Pranav Viswanathan

Motivation

The Problem:

Conversational Style Recommender Systems

- Traditional Recommender Systems rely on static user interaction signals while CRSs engage users in **dynamic, multi-turn dialogues**
- CRS must extract and infer **evolving user preferences** while generating **human-like responses**

Large Language Models

- Have been shown to outperform even fine-tuned traditional CRS models *even in a zero-shot capacity* due to their extensive pre-training and context-aware capabilities [1]
- However, **LLMs still struggle to incorporate domain-specific knowledge** and **maintain real-time awareness** of users' shifting preferences and relationships between multiple users

Our Proposed Solution:

Relational Knowledge Graph-Augmented Reasoning for Zero-Shot Conversational Movie Recommendation

Why Knowledge Graphs (KGs)?

- Have proven effective in providing a **structured, interconnected representation** of domain-specific entities and their relationships [2]
- Enhanced **contextual groundings** and **explainability** in recommendations

Our Approach

- Integrate relational, structured KGs with LLMs for **richer, context-aware reasoning**
- Facilitate **zero-shot reasoning** where prior knowledge of the specific task is minimal or absent **at each conversational turn**

[1] Zhankui He et al. https://doi.org/10.1145/3583780.3614949

[2] Zhangchi Qiu et al. https://arxiv.org/abs/2411.14459

Data Description

ReDIAL

- A canonical conversational dataset for evaluating conversational recommendation systems
- **Over 10,000 movie recommendations**, accompanied by the **conversational history**
- Created by using Amazon Mechanical Turk

IMDB Movie Metadata

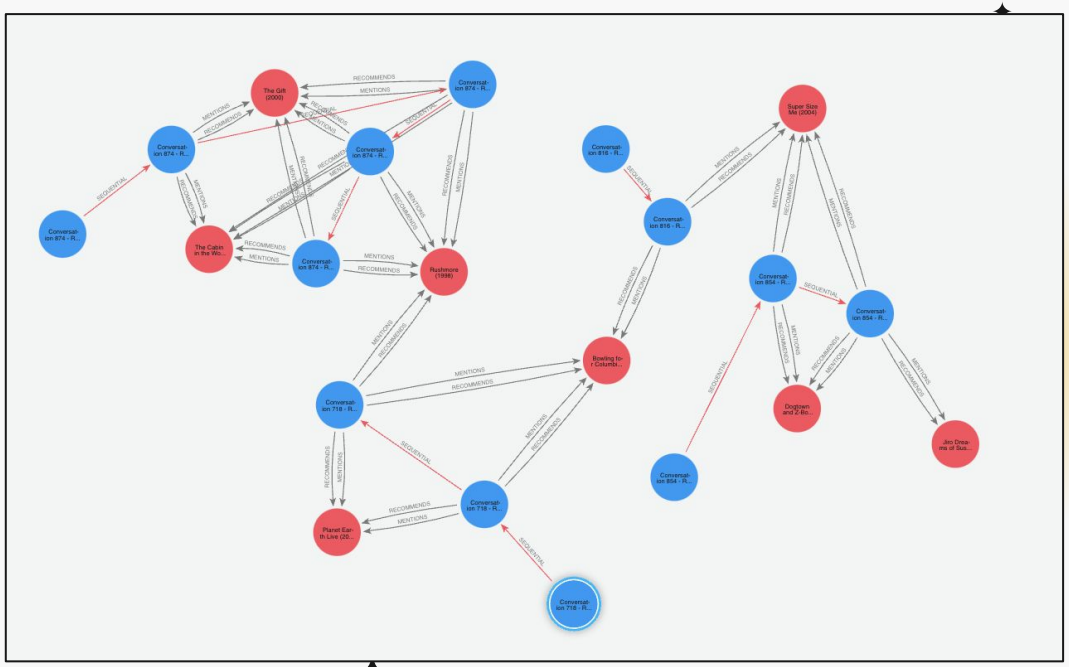
- Kaggle Dataset which provides **structured movie metadata** (genres, directors, actors, descriptions)

Example Conversation 1

- **Movie Mentions**:
  - `165710` : "The Boss Baby (2017)"
- **Respondent Questions**:
  - `165710` : Suggested: 1, Seen: 0, Liked: 1
- **Initiator Movie Mentions**:
  - `84779` : "The Triplets of Belleville (2003)"
  - `191602` : "Waking Life (2001)"
  - `122159` : "Mary and Max (2009)"
- **Messages**:
  - **0s**: Initiator: "Hi there, how are you? I'm looking for movie recommendations"
  - **15s**: Respondent: "I am doing okay. What kind of movies do you like?"
  - **66s**: Initiator: "I like animations like @84779 and @191602"
  - **86s**: Initiator: "I also enjoy @122159"
  - **95s**: Initiator: "Anything artistic"
  - **135s**: Respondent: "You might like @165710 that was a good movie."
- **Conversation ID**: 391
- **Respondent Worker ID**: 1
- **Initiator Worker ID**: 0
- **Recommendation Index**: 1

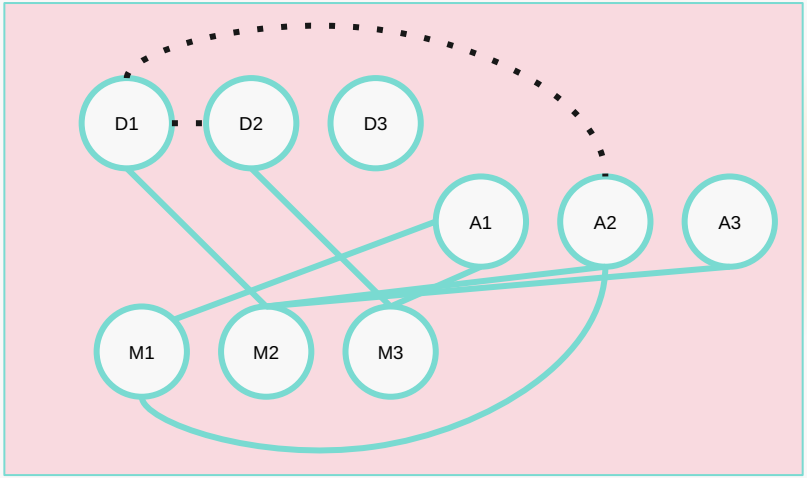
Knowledge Graph Creation

- **Primary Relationships**
  - Actor - Movie
  - Director - Movie
  - Genre - Movie
- **Scoring**
  - Count
- **Other Features**
  - Sequential Mapping of conversations



Hidden Relationships

- **We use the primary relationships to create further links between nodes**
- **“Hidden” Relationships**
  - Actor-[movie]-Actor
  - Actor-[genre]-Actor
  - Actor-[movie]-Director
  - Director-[genre]-Director
- **Scoring**
  - Count



LLM Connection

[User]: I love **Interstellar**, any recommendations?

You would love **The Martian**! : [System]

[User]: Who is in it?

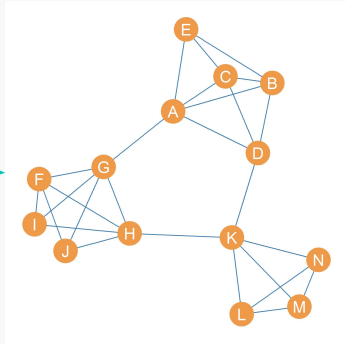
Matt Damon! : [System]

[User]: I'm looking for movies more like the ones directed by Christopher Nolan.

You might like **Arrival** then! [System]

[User]: I like this movie! Can you recommend me more sci-fi movies like the ones we've talked about?

[BLANK]: [System]



"""

Pretend you are a movie recommender system. I will give you a conversation between a user and you (a recommender system). Based on the conversation, you reply me with 5 recommendations without extra sentences. The recommendations must follow this format:

—

…

—

Here is the conversation:

—

…

—

All movies you can recommend (movie id and name):

—-

…

—-

Please provide 5 recommendations:

"""

LLM

1. Gravity
2. Oblivion
3. Blade Runner 2049
4. Elysium
5. Prometheus

Prompting

Processing

Results

Model	Recall @ 5	Hallucination ID's	Hallucination Movies
GPT-4o with KG Context	8.16%	31.33%	2.86%
GPT-4o	3.06%	16.16%	1.63%

Findings

- Higher recall @ K, and high hallucinations

Metric Flaws

- Recall @ K wants an exact match – not a measure of quality
- Hallucinations are not necessarily bad

Discussion & Future Work

Temporal Knowledge Graphs

- Adding a temporal aspect to the knowledge graph
  - Sequential Links between conversations, and incorporation of timestamps to track **evolving user preferences over time**

Alternative Models

- Smaller, fine tuned LLMs
- BAIZE, Vicuna are representative open-sourced LLMs fine-tuned on LLAMA-13B

Graph Entity Captioning

- Modality gap between KGs and LLMs
- Can instead fine-tune Transformer models to generate NLP captions