Towards Exploiting Background Knowledge for Building Conversation Systems

Nikita Moghe, Siddhartha Arora, Suman Banerjee, Mitesh M. Khapra

Department of Computer Science and Engineering,
Indian Institute of Technology Madras, Chennai - 600036, India

RBCDSAI IITM

Introduction

- **Problem**: Existing **dialogue** datasets fail to incorporate background knowledge while conversing which has led to treating the problem as yet another **seq2seq** paradigm .
- **Proposed Direction**: We create a new dataset containing movie chats wherein each **response** is explicitly generated by **copying** and/or modifying sentences from unstructured **background knowledge** such as plots, comments and reviews about the movie.
- **Motivation**: This approach mimics the **human way of conversation** which relies on **knowledge** acquired through **reading** background articles/resources.
- **Baselines**: We evaluate the dataset using - (i) pure generation based model, (ii)copy-and-generate model and (iii)span-prediction model

Example: Conversation on Spider-Man

Plot: . . . Peter's science class takes a field trip to a genetics laboratory at Columbia University. The lab works on spiders and has even managed to create new species of spiders through genetic manipulation. While Peter is taking photographs of Mary Jane for the school newspaper, one of these new spiders lands on his hand and bites him ...

Review: I thoroughly enjoyed"Spider-Man" which I saw in a screening. I thought the movie very engrossing. Director Sam Raimi kept the action quotient high, but also emphasized the human element of the story. The casting was perfect. Tobey Maguire was very believable as the gawky teenager in the early part of the film and then, after his run-in with the radioactive

Speaker 1(N): Which is your favourite character in this?

Speaker 2(C): My favorite character was played by Tobey Maguire.

Speaker 1(N): I thought he did an excellent job as Peter Parker, I didn't see what it was that turned him into Spider-Man though.

Speaker 2(P): Well this happens while Peter is taking photographs of Mary Jane for the school newspaper, one of these new spiders lands on his hand and bites him.

Speaker 1 (N): I see. I was very excited to see this film and it did not disappoint!

Speaker 2(R): I agree, I thoroughly enjoyed "Spider-Man"

Speaker 1(N): I loved that they stayed pretty true to the comic.

Speaker 2(C): Yeah, it was a really great comic book adaptation

Speaker 1(N): The movie is a great life lesson on balancing power.

Speaker 2(F): That is my most favorite line in the movie, 'With great power comes great responsibility.'

Comments: Crazy attention to detail. My favorite character was played by Tobey Maguire. I can't get over the "I'm gonna kill you dead" line. No spoilers, but it does start to take itself more seriously towards the finale. It was too heavily reliant on constant light-hearted humor. However the constant joking around kinda was low. A really great comic book adaptation. ....

Collection	$403,706,375
Taglines	With great power comes great responsibility Get Ready For Ultimate Spin!
Awards	Golden Trailer Awards 2002
Similar Movies	Spider-Man 2 Iron Man

DataStats

9071 dialogues	91K utterances	921 movies	9278 resources	15.29 Avg words/turn	153.07 Avg words/chat

Data Collection Method

Step 1. Curate a list of popular movies
- 921 movies across 34 genres
- Collected using IMDb popular 250 list, top rated in every genre list, top 10 popular movies in the past five years and other lists

Step 2. Collect background knowledge
- Fetch top 2 popular **Reviews** per movie from IMDb
- Collect **Plot** and **Fact Table** data using Wikipedia pages and Infoboxes
- Use **Reddit** to collect Comments which are rich in informal interaction

Step 3. Crowdsource
- Collect opening statements for chats with a fixed starting template
- Collect complete chats by showing the user four background resources
- The Turker must produce coherent chats where alternate responses are copied from these resources while performing a role-play of both the speakers

Step 4. Manually Verify
- Check if the workers copied the spans and marked the respective documents
- Check if the chats were coherent and on topic

Is this a natural way of conversation?

Intelligible	Coherent	Grammar	Two-person-chat	On Topic
4.47	4.33	4.41	4.47	4.57

Table: Human evaluation on randomly chosen 500 collected chats (Scale 1-5)

Baseline Results and Analysis

Model	Method	F1	BLEU	Rouge-1	Rouge-2	Rouge-L
HRED[1]	-	-	5.23	24.55	7.61	18.87
GTTP[2]	oracle	-	13.92	30.32	17.78	25.67
GTTP [2]	mixed-long	-	7.51	23.20	9.91	17.35
BiDAF[3]	oracle	39.69	28.85	39.68	33.72	35.91
BiDAF[3]	mixed-short	45.73	32.95	45.69	40.18	43.46

Table: Performance of the proposed models on our dataset. (Single Reference)

Model	Method	F1	BLEU	Rouge-1	Rouge-2	Rouge-L
HRED[1]	-	-	5.38	25.38	8.35	19.67
GTTP[2]	oracle	-	16.46	31.6	21.21	27.83
GTTP [2]	mixed-long	-	8.73	21.55	10.42	18.12
BiDAF[3]	oracle	47.18	34.98	46.49	40.58	42.64
BiDAF[3]	mixed-short	51.35	39.39	50.73	45.01	46.95

Table: Performance of the proposed models on our dataset. (Multi Reference)

- Natural language **generation capabilities** of current generation based models are **far from being acceptable**.
- **Span prediction** models can exploit background knowledge much **better** than **generation** models.
- State-of-the-art **span prediction models** are complex and **expensive** which makes it **infeasible for long resources**.

Conclusion and Future Work

- Introduced a **new dataset** for building dialog systems that explicitly uses **background knowledge** associated with current context.
- Develop **hybrid** methods based on span prediction and generation.
- **Build less complex models** which can handle longer resources.

References

Iulian Vlad Serban, Alessandro Sordoni, Yoshua Bengio, Aaron C. Courville, and Joelle Pineau.
Building end-to-end dialogue systems using generative hierarchical neural network models.
In *AAAI*, pages 3776–3784, 2016.

Abigail See, Peter J. Liu, and Christopher D. Manning.
Get to the point: Summarization with pointer-generator networks.
In *ACL*, pages 1073–1083, 2017.

Min Joon Seo, Aniruddha Kembhavi, Ali Farhadi, and Hannaneh Hajishirzi.
Bidirectional attention flow for machine comprehension.
ICLR, 2017.