January 2023

Virtual Assistant That Provides Answers Based On Past Conversations

Kushal Lakhotia

Denis Savenkov

Sonal Gupta

Virtual Assistant That Provides Answers Based On Past Conversations

ABSTRACT

It is often the case that participants in a conversation, e.g., a chat conversation via a messaging app or an audio/video call, do not remember the details of the conversation. This disclosure describes the use of machine learning techniques to automatically answer questions related to a past conversation by use of machine reading comprehensive models. The models are trained using conversation transcripts and provide answers based on a conversation between users, or between a user and a virtual assistant. The models can be incorporated in a virtual assistant that can reason over any number of conversations to arrive at an answer.

KEYWORDS

- Conversation log
- Conversation history
- Chat transcript
- Call transcript
- Virtual assistant
- Natural language query
- Conversational turn

BACKGROUND

It is often the case that participants in a conversation, e.g., a chat conversation via a messaging app or an audio/video call, do not remember the details of the conversation. Participants in a conversation remember the gist of the conversation but not answers to specific questions, which can be important, e.g., a commitment made, a time or a place to meet, an agreement, etc. For example, consider the conversation:

A: Do you want to grab something for lunch?

B: Yes, I'm starving.

A: Great, let's meet at Ayenkar's Bakery at 12:15 PM.

B: Sounds good.

It can be the case that the participant B recollects that a lunch appointment was made, but forgets the venue. A natural question that B would then pose is: "Where did I agree to meet for lunch with A?"

The conversation can have several participants. For example, a group of friends may have jointly discussed plans for a party. Later one of the participants might want to recollect forgotten details about the party, e.g., the venue, gifts to be given, etc., that were discussed.

DESCRIPTION

This disclosure describes the use of machine learning techniques to automatically answer questions, e.g., posed as natural language queries, related to a past conversation by use of machine reading comprehensive models. The models are trained using conversational data and provide answers based on a conversation between users, or between a user and a virtual assistant. The models utilize transcripts of the conversation, obtained, e.g., from a messaging application, and provide answers to user queries posed in a natural language. The models can be incorporated in a virtual assistant that can reason over any number of conversations to arrive at an answer.

The models are trained by having human participants sift through transcripts of sample conversations and, based on the transcripts, generate natural language queries and corresponding answers. The conversational transcripts and the question-answer pairs serve as training data for the models. The trained models thus effectively extend a user's ability to memorize and retrieve information from computer-mediated conversations.

Dialog 102

M:	I must admit that I'm not really looking forward to it.
W:	No?
M:	Well, I've been to their parties before, and I don't suppose this one will be any better. All the same people are going on and on about their jobs.
W:	Aha.
M:	They all think they are so important. Honestly, I just can't be bothered with them. I never know what to say. Still, I've been invited, so I suppose I'd better go.

Generation of Training Question-Answer Pairs 104

Imagine you were one of the participants in the dialog above. Who is it? (type the alias without ":"). 104a

W

Confirm

Pretend you forgot some detail mentioned in the dialog (either by you, or by somebody else). Write a question you could ask to recall this detail: 104b

Is M going to the party?

Confirm

What is the correct answer to your question? (Please type the phrase exactly as it appears in the dialog, or copy-paste it from an utterance in the dialog.) 104c

I'd better go.

Confirm

Fig. 1: Tool to generate training data for a conversational assistant

Fig. 1 illustrates an example user interface of a tool to generate training data for a conversational assistant, per the techniques of this disclosure. Human volunteers, serving as conversation annotators, are provided such a tool in order to generate annotated training data for conversational assistants.

A section of the user interface includes a conversational transcript (102), e.g., a dialog with several turns. Another section of the tool includes questions for the human volunteer that generate training data (104), e.g., question-answer pairs, to train the conversational assistant. Questions asked of the human volunteers include requesting the volunteer to imagine they are one of the participants in the conversation (104a); requesting the volunteer to write a question pertaining to a detail in the conversation (104b); requesting the volunteer to write a natural-language answer to the question pertaining to the detail in the conversation (104c); etc.

The volunteer studies the dialog, picks the alias of a dialog participant, formulates a question about a detail of the dialog, and formulates an answer to that question. As illustrated, the tool provides volunteers the ability to feed in and confirm their entries. Provisions are made within the tool to enable a volunteer to add multiple question-answer pairs, or none (e.g., if the conversation does not include details that can be formulated as a question-answer pair).

To obtain high quality training data, the volunteers are encouraged to avoid generic questions, e.g., “what did we discuss with A?” “what did B tell me?” etc. Rather, they are instructed to isolate specific details within the conversation transcript in order of importance, or details that human participants in a conversation are likely to forget. Volunteers are guided to frame questions in enough detail that they yield unambiguous answers as far as possible, e.g., with lower likelihood of multiple valid answers. Volunteers are advised to write as short an answer as possible, and as far as possible copy-paste the answer from the conversation transcript. Volunteers are also advised to isolate dialogs from each other, even if names match across dialogs, e.g., information from a previously annotated conversation does not ordinarily carry over to a new conversation. Volunteers are instructed to assume certain relationships between the participants in the conversation if such a relation is implied. Thus, such question-answer pairs

such as ("who did my *boss* go out with yesterday?", "Bob") are acceptable even if the other party in the conversation is not explicitly identified as "*boss*." Volunteers are to restrict question-answer pairs to material within the conversation.

Fig. 2 illustrates another example of a conversation transcript and questions used to generate training data for a conversational assistant.

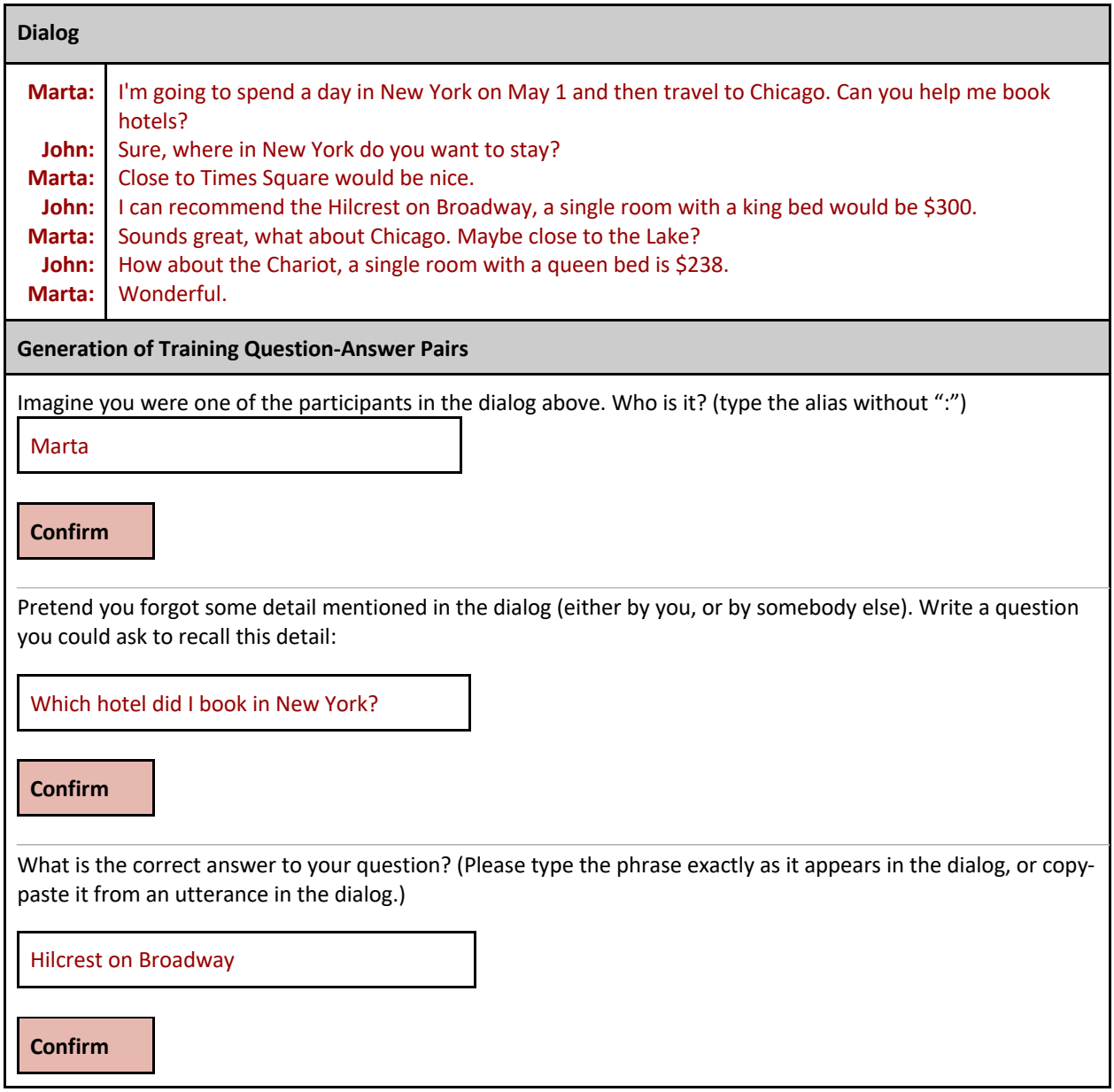

Fig. 2: Another example of a conversation transcript and questions used to generate training data for a conversational assistant

For more complex conversations, e.g., as illustrated in Fig. 3, several types of question-answer pairs can be composed by the volunteer, e.g., simple, participant, ambiguous, indirect, comparison, etc., as described below.

Amy: I am looking for places to go, is there a college I could see?

Benjamin: There are actually 18 colleges in town, did you have an area you wanted to stick to?

Amy: No, just the first one on the list is fine. Could you provide me with an address and an entrance fee?

Benjamin: The first on my list is Christopher's College on St. Andrew's Street; there is no entrance fee. Is this good?

Amy: Thank you! I am also looking for a place to eat. I am looking for a restaurant in the East that is moderate in price.

Benjamin: Can I recommend the Gol-maal, at C Farnwell Road Ben Ditton?

Amy: That's fine, book me a table for 3 for Sunday at 19:30 and give me a reference number.

Benjamin: I was unable to book under this criteria, is there another time of day you would consider?

Amy: Can you try a different restaurant?

Benjamin: How about Choco Slurry Prince? It is also located in the east.

Amy: That should work. Try that for 3 people at 19:30 on sunday.

Benjamin: You're all set, your reference number is C262TX42. Is there anything else I can help you with?

Amy: I would also like to book a taxi to commute between the two places, please.

Benjamin: I have booked you with a red sedan. The contact number is 01112223333. Can I assist you with anything else?

Amy: No, that is all I need, thank you for the help.

Benjamin: Happy to assist, take care!

Fig. 3: A complex conversation

***Simple questions*:** Simple questions are those that can be answered by looking at a single line in the original conversation and extracting some clearly stated value.

- Q: Where is Choco Slurry Prince located? A: The East.
- Q: What is the contact number for my taxi? A: 01112223333.

***Participant questions*:** Participant questions are those that require consideration of who said something, and/or whether they were making a statement, asking a question, etc.

- Q: What did Benjamin ask me to consider when the reservation at the Gol-maal didn't work? A: Another time of day.
- Q: What kind of place did Amy request information about before asking about restaurants? A: A college.
- Q: What area of town was Amy looking for a restaurant in? A: The East.

Participant questions often require consideration of content from several lines of the conversation, even if the quotation ultimately comes from only one. For example, to determine which restaurant was suggested first, the assistant needs to consider both restaurant suggestions and their order.

***Ambiguous Questions*:** Ambiguous questions may require combining multiple pieces of information in order to resolve a referent. For ambiguous questions, the answer is not uniquely found by looking at only part of the referring expression.

- Q: What is the number for the taxi? A: 01112223333.
    - In this question, considering just the phrase “the number” results in an ambiguity between the contact numbers for the taxi and the reference number for the restaurant reservation. To resolve this question, the assistant also considers the phrase “for the taxi.”
- Q: What was the restaurant that we tried first? A: The Gol-maal.
- Q: What was the location of the college Amy decided on? A: St. Andrew's Street.

***Indirect Questions*:** Indirect questions require analyzing multiple pieces of information to arrive at an answer, even though the answer itself can still be extracted from the conversation as a quote.

- Q: How many people are there in my reservation? A: 3.
- Q: Where will the taxi be picking me up? A: Christopher's College on St. Andrew's Street.

***Comparison Questions*:** Comparison questions require comparing or considering two or more elements in the original conversation to arrive at an answer. Generally, they will do one of the following.

- Present two or more options and pick between them based on a parameter or set of parameters:
    - Q: Is Christopher's College on St. Andrew's Street or C Farnwell Road Ben Ditton? A: St. Andrew's Street.
- Ask of a commonality between two or more things:
    - Q: What area are Choco Slurry Prince and the Gol-maal located in? A: The East

While the foregoing examples illustrate answers in the form of a phrase, the answer can be a phrase or a full sentence. For example, in response to the question "When am I meeting John?" the answer can be a full sentence, e.g., "You are meeting John at Central Park at 8 am."

In this manner, the techniques of this disclosure enable the collection of a dataset of question-answer pairs, where the questions and their corresponding answers are based on the transcript of a conversation that a person has had. The conversational transcripts and the question-answer pairs serve as training data for machine reading comprehensive models that are

trained to answer questions based on conversations, and may be utilized in a virtual assistant application. While the foregoing examples illustrate conversations between human users, the models can also provide answers based on past conversations between a user and a virtual assistant. The trained models can be incorporated in a virtual assistant to enable question answering features that effectively extend a user's ability to memorize and retrieve information.

CONCLUSION

This disclosure describes the use of machine learning techniques to automatically answer questions related to a past conversation by use of machine reading comprehensive models. The models are trained using conversation transcripts and provide answers based on a conversation between users, or between a user and a virtual assistant. Conversation transcripts that are used for training can include various types of questions that a user may ask naturally, obtained from human volunteers, and the corresponding answers. The models can be incorporated in a virtual assistant that can reason over any number of conversations to arrive at an answer.