

Task 2: Analyzing Spoken Dialog Systems

For this task, you will be working with popular spoken dialog systems (sometimes also referred to as “voice assistants” or “conversational assistants/agents”).

Submission details (for folder “task2”): Text file with your responses (.txt, .doc, .docx, .pdf)

1. Ask any two of Google Assistant, Alexa, Cortana or Siri (or any other voice assistant!) for restaurant suggestions given a particular cuisine and neighborhood. Then get the phone number and restaurant hours. Use the same queries for both systems in order to directly compare the performance.
 - 1) Two Voice Assistants
: Google Assistant and Siri
 - 2) Ask for restaurant suggestion given a particular cuisine and neighborhood
: Asked for a fried chicken in Upper West Side
 - 3) Ask for phone number and restaurant hours
: Asked for the phone number and restaurant opening hours of the first recommendation.
2. Describe the experience that you had with each of the systems. What query or queries did you use, and how did the system respond? Was the system able to complete the task with the desired result? Was the dialogue efficient? Describe errors made by the system (if any) and how the system or user recovered from them. Which system did you prefer using, and why?

1) Queries Used

First, I asked, “Where can I eat fried chicken in Upper West Side?”. After receiving the answer, I asked, “What are the phone number and opening hours?”.

2) Responses

a. Siri

Me: Hey, Siri, where can I eat fried chicken in Upper West Side?

Siri: I found this on the web for ‘where can I eat fried chicken in Apple website’.

Me: Where can I eat fried chicken in Upper West Side?

Siri: One option is Chick Chick on Amsterdam Ave. Is that the one you’re looking for?

Me: What are the phone number and opening hours?

Siri: Here’s the phone number. Do you want that one?

Me: What are the opening hours?

Siri: Chick Chick on Amsterdam Ave is open today, from 4:30 PM to 10 PM. Want to try that one?

b. Google Assistant (GA)

Me: Ok, Google, where can I eat fried chicken in Upper West Side?

GA: I found a few places near you. The first one is Popeye’s Louisiana Kitchen at 2534 Broadway, New York. The second one is Texas Rotisserie & Grill at 2581 Broadway, New York. The third one is Jenkins Pickles at 509 Amsterdam Ave, New York.

Me: Ok, Google, what are the phone number and opening hours?

GA: You can contact Popeye’s Louisiana Kitchen at 917 409 3737.

Me: Ok, Google, what are the opening hours?

GA: Popeye’s Louisiana Kitchen at 2534 Broadway is open today from 10:30 AM to 10 PM.

3) Describe Experience with Each System

For Siri, I had to repeat the first question because it understood 'Upper West Side' as 'Apple website' (the letter 'A' was capitalized in the response screen). Probably this is because many people use Siri to ask things related to the company, Apple or products of Apple and they have a lot of data including the proper noun 'Apple'. After repeating the question, I was able to get the desired response. Then, I asked the second question to get the phone number and the opening hours, but I was only able to get the phone number information. So, I asked again just for the opening hours, and I got the correct answer. On the other hand, Google Assistant understood the first question correctly on the first try. It gave me three suggestions with each restaurant's information. Then I asked the second query for the phone number and the opening hours, and I only got the phone number of the restaurant; it made the same error as Siri for the second question. So, I asked for the opening hours and got the desired information.

Both systems were able to complete the task with the desired results. The dialogues weren't as efficient as I hoped, as I had to repeat some of the questions. Especially, both systems were not able to give the phone number and the opening hours in one response. The user should try to make a question simpler (limit the scope of the question to only one domain at a time) to get the desired outcome. Despite some errors, I think these systems are quite useful when we cannot use our hands to type and search for something, while they still have much room to improve.

Personally, I preferred using Siri, even though, it misunderstood my first question. Siri constantly asked if its answer was what I was looking for, so I had a chance to validate the answer right away and change the course of the conversation. Also, I didn't have to summon Siri every time when I wanted to continue asking a new question because Siri was automatically activated right after its response, so the overall dialog felt smoother compared to Google Assistant; with Google Assistant, I had to say "Ok, Google" before every question.