Conversational UIs

Spoken Language Processing

Topic 5: Spoken Dialogue Processing

Dr. Dong Minghui
Institute for Infocomm Research
Agency for Science, Technology and Research (A-Star)
Email: mhdong@i2r.a-star.edu.sg

PART 1. INTRODUCTION

Examples of Spoken Dialogue

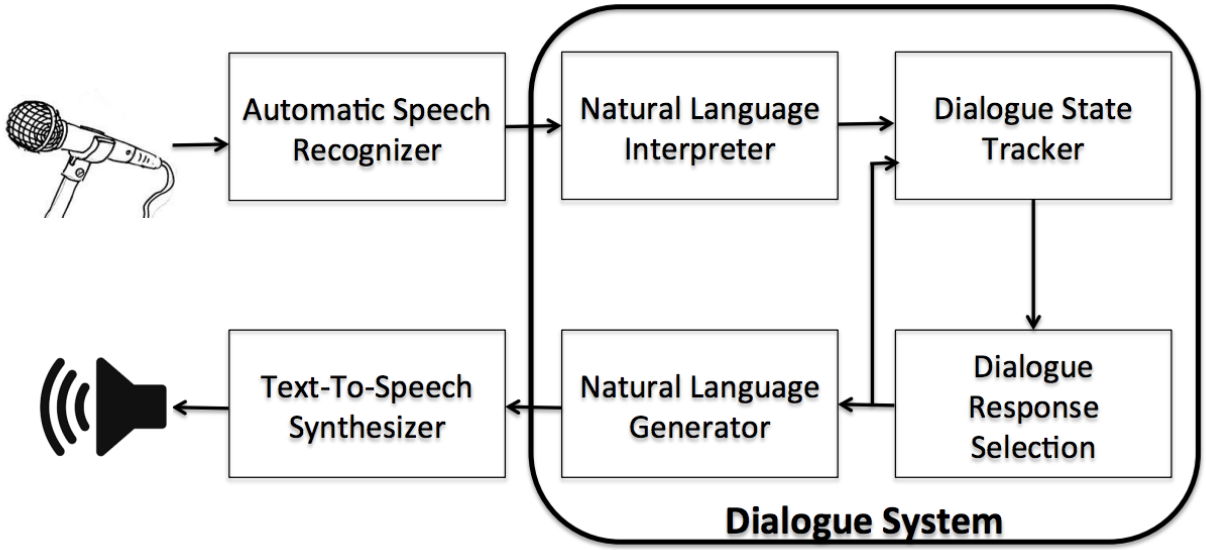
- Speech Assistant:
  - Amazon Echo
  - Google Home
  - Siri

- Robots
  - Social robot

- In Car
  - Navigation and information

Spoken Dialogue System (SDS)

- Speech recognizer: to listen to human voice
- Speech synthesizer: to speak to human
- Dialogue system: to understand the language and respond



Considerations in SDS

- **System design**
  - Be specific on task definition
  - Avoid using similar name for product/service
- **Speech recognition**
  - Make sure preprocessing component delivers correct speech signal.
  - The error patterns of speech recognition is different from keyboard typing, text based dialogue needs to further tuned for spoken input.
- **System synthesis**
  - Avoid using symbols/words that may be ambiguous.

Dialogflow

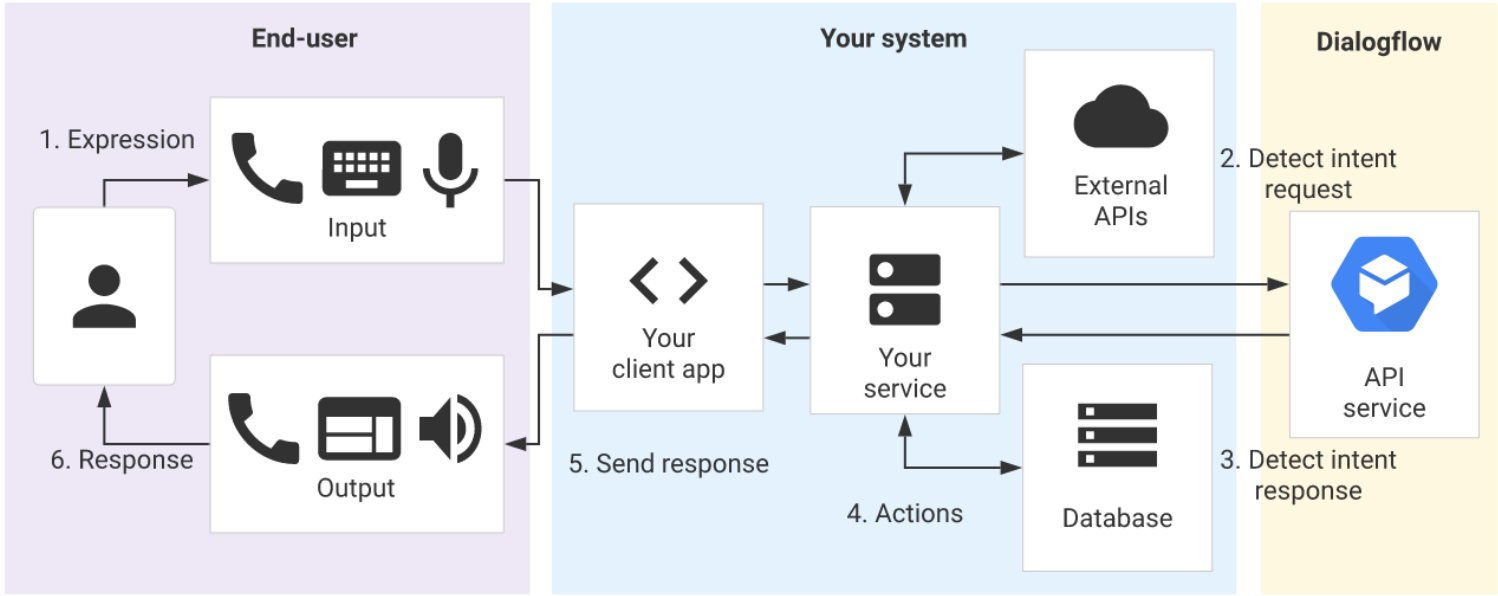
- **What is Dialogflow**
  - A Google Cloud product.
  - A natural language understating platform to build conversational user interface.
  - Both input and output can be text or speech.
- **Google Cloud**
  - It contains quite a lot of products such as NLP, Computer Vision, Machine Learning, Big data, etc.

Dialogflow concepts

- **Agent:**
  - a module to handle the conversation with end user.
- **Intent:**
  - to categorize end-user intentions for each conversation turn
- **Entity:**
  - to identify and extract specific data from end-user expressions. E.g. dates, times, colors, email addresses.

Dialogflow with Speech

- **Dialogflow supports speech input and output. Both input and output can be configured to use text or speech.**
- **DialogFlow support APIs.**
- **Two ways for speech input:**
  - Input file: pass the full recording to speech engine
  - Streaming: pass the recording little by little.



PART 2. WORK WITH DIALOGFLOW

Tasks

- **Create a dialog:**
  - To build a dialog demo that support booking a meeting room, the required information includes: Date, Time, Duration, Location, Room Name (A, B, C)
  - We will load the predefine dialogue from file.
- **Build spoken dialog with APIs.**
  - Using APIs to support speech recognition and speech synthesis.
- **Integration with Google Assistant**
  - To interact with the agent on smartphone quickly.

Make DialogFlow Work

- **On Google Cloud:**
  - Create the dialog agent, and create service account Key.
- **To use with API:**
  - Install package on computer: pip install dialogflow
  - Write your program.
- **To use with Google Assistant:**
  - Install Google assistant on smart phone.
  - Enable integration on web.

Dialogflow with Speech API

- **Create Agent**
  - dialogflow.cloud.google.com
  - You need to have a google account
  - Go to Dialogflow Console
  - Create an agent (Name, Language, Timezone)
  - Project automatically created. Keep a record of the Project ID
  - Import an existing project (RoomReservation.zip)

Create Intent

- **Create Intent**
  - If we would like to create an intent from scratch (instead of importing)
  - Please refer to https://cloud.google.com/dialogflow/cx/docs/quick/build-agent

Dialogflow with Speech API

- **Create Service Account Key**
  - https://console.cloud.google.com/apis/credentials/serviceaccountkey
  - Select the project
  - From the **Service account** list, select **Create service account.**
  - In the **Service account name** field, enter a name, click **Create and continue**
  - From the **Role** list, select **Project** > **Owner**, click **Continue**, then click **Done**
  - Click three dots, select **manage keys**.
  - Select **Add key**, select **Create new key,** Key type: select **JSON**
  - Click **Create** and save the generated file.

Dialogflow with Speech API

- **Samples:**
  - Speech input from file: test-asr.py

    https://cloud.google.com/dialogflow/es/docs/how/detect-intent-audio
  - Speech output to file: test_tts.py

    https://cloud.google.com/dialogflow/es/docs/how/detect-intent-tts
  - Speech input by Streaming: test-steam.py

    https://cloud.google.com/dialogflow/es/docs/how/detect-intent-stream

Dialogflow with Google Assistant

- **Integration**
  - Go to **Dialogflow Console**
  - Select your agent
  - Select **Integrations**
  - Select **Google Assistant** for integration
- **Integration Settings**
  - Explicit invocation: Default Welcome Intent
  - Implicit invocation: room.reservation
  - Click Test
- **Manage Assistant App**
  - Develop
  - Display name: My booking app
  - Save

Homework

- **In the examples, we provided**
  - A sample program for speech input (test-asr.py)
  - A sample program for speech output (test-tts.py)
- **Your task:**
  - Please write your own program to support both speech input and speech output.
  - Please submit it to the LumiNUS.

Thank you!

mhdong@i2r.a-star.edu.sg