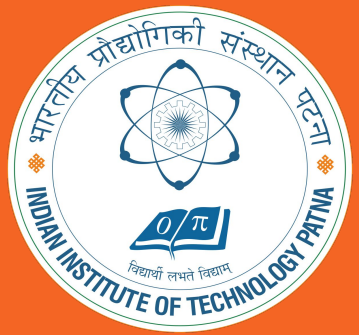
INTRODUCTION TO CONVERSATIONAL AI



Presented By: Kshitij Mishra (Ph.D. Scholar, CSE Department, IIT Patna)
Research Area: Personalization in Conversational AI

Supervised By: Dr. Asif Ekbal

Outline

- Conversational AI
- Properties of Conversation/Dialogue
- Dialogue: What Kinds of Problems?
- Architectures of Dialogue Systems
- Modular and Non-modular Dialogue Systems
  - Control structure for Task-oriented dialogue Systems
  - Determining Domain, Intent, and Slot fillers
  - Natural Language Generation module
- Dialogue State Architecture
  - Dialogue Acts
  - Slot Filling
  - Dialogue State Tracking
  - Natural language generation in the dialogue-state model
- Evaluating Dialogue Systems

Conversational AI

- Two words: *Conversation* and *AI*
  - AI: *intelligence* demonstrated by *machines.*
  - Conversation: *interactive communication* between *two or more people*.
- Conversational AI: a machine (demonstrating *intelligence* like a human) in an *interactive communication* with a human.
- Used for various purposes such as customer service, request routing, information gathering or for simple chit-chat.

Properties of Conversation/Dialogue

- **Utterance:** Smallest unit of speech. It is a continuous piece of speech beginning and ending with a clear pause.
- **Turns:** A **dialogue** is a ***sequence of turns*** (C1, A2, C3, and so on), each a single contribution from one speaker to the dialogue.
- **Speech Acts:** Each utterance in a dialogue is a kind of ***action*** being performed by the speaker. These actions are commonly called ***speech acts*** or dialog acts.
- **Grounding:** Grounding means ***acknowledging*** that the hearer has understood the speaker; like an ACK used to confirm receipt in data communications.

C_1:	...I need to travel in May.
A_2:	And, what day in May did you want to travel?
C_3:	OK uh I need to be there for a meeting that's from the 12th to the 15th.
A_4:	And you're flying into what city?
C_5:	Seattle.
A_6:	And what time would you like to leave Pittsburgh?
C_7:	Uh hmm I don't think there's many options for non-stop.
A_8:	Right. There's three non-stops today.
C_9:	What are they?
A_{10}:	The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.
C_{11}:	OK I'll take the 5ish flight on the night before on the 11th.
A_{12}:	On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
C_{13}:	OK.
A_{14}:	And you said returning on May 15th?
C_{15}:	Uh, yeah, at the end of the day.
A_{16}:	OK. There's #two non-stops ...#
C_{17}:	#Act...actually #, what day of the week is the 15th?
A_{18}:	It's a Friday.
C_{19}:	Uh hmm. I would consider staying there an extra day til Sunday.
A_{20}:	OK...OK. On Sunday I have ...

Figure 1: Part of a phone conversation between a human travel agent (A) and human client (C).

Properties of Conversation/Dialogue

- **Ex:** One taxonomy consisting of 4 major classes:
  - **Constatives:** *committing the speaker to something's being the case* (answering, claiming, confirming, denying, disagreeing, stating)
  - **Directives:** *attempts by the speaker to get the addressee to do something* (advising, asking, forbidding, inviting, ordering, requesting)
  - **Commissives:** *commi4752630tting the speaker to some future course of action* (promising, planning, vowing, betting, opposing)
  - **Acknowledgments:** *express the speaker's attitude regarding the hearer with respect to some social action* (apologizing, greeting, thanking, accepting an acknowledgment)

C_1:	. . . I need to travel in May.
A_2:	And, what day in May did you want to travel?
C_3:	OK uh I need to be there for a meeting that's from the 12th to the 15th.
A_4:	And you're flying into what city?
C_5:	Seattle.
A_6:	And what time would you like to leave Pittsburgh?
C_7:	Uh hmm I don't think there's many options for non-stop.
A_8:	Right. There's three non-stops today.
C_9:	What are they?
A_{10}:	The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.
C_{11}:	OK I'll take the 5ish flight on the night before on the 11th.
A_{12}:	On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
C_{13}:	OK.
A_{14}:	And you said returning on May 15th?
C_{15}:	Uh, yeah, at the end of the day.
A_{16}:	OK. There's #two non-stops . . . #
C_{17}:	#Act. . . actually #, what day of the week is the 15th?
A_{18}:	It's a Friday.
C_{19}:	Uh hmm. I would consider staying there an extra day til Sunday.
A_{20}:	OK. . . OK. On Sunday I have . . .

Figure 1: Part of a phone conversation between a human travel agent (A) and human client (C).

Dialogue: What Kinds of Problems?

- **Task completion:** To accomplish user tasks ranging from restaurant reservation to meeting scheduling, and to business trip planning. (**task completion *bots***)
- **Social chat:** To converse seamlessly and appropriately with users — like a human as in the Turing test — and provide useful recommendations. (**social *chatbots***)
- **Question answering:** To provide concise, direct answers to user queries based on rich knowledge drawn from various data sources including text collections such as Web documents and pre-compiled knowledge bases such as sales and marketing datasets. (**QA *bots***)

★ Can be grouped into two categories, ***task-oriented*** and ***chitchat***.

Dialogue: What Kinds of Problems?...

Turn	Dialogue
1	*usr:* Good morning!
	agt: Good morning, I hope you had a great weekend.
2	*usr:* not really.
	agt: You sound depressed. What can I do for you?
3	*usr:* I'd like to know where sales are lagging behind our forecast?
	agt: The worst region is [country], where sales are 15% below projections.
4	*usr:* Do you know why?
	agt: The forecast for [product] growth was overly optimistic.
5	*usr:* How can we turn this around?
	agt: Here are the 10 customers in [country] with the most growth potential, per our CRM model.
6	*usr:* Can you set up a meeting with the CTO of [company]?
	agt: Yes, I've set up a meeting with [person name] for next month when you are in [location].
7	*usr:* Thanks!

Figure 2: A human-agent dialogue during the process of making a business decision. (**usr: user, agt: agent**). The dialogue consists of *multiple segments of different types*. **Turns 1 and 2** are a **social chat** segment. **Turns 3 to 5** are a **QA** segment. **Turns 6 and 7** are a task-completion segment.

Architectures of Dialogue Systems

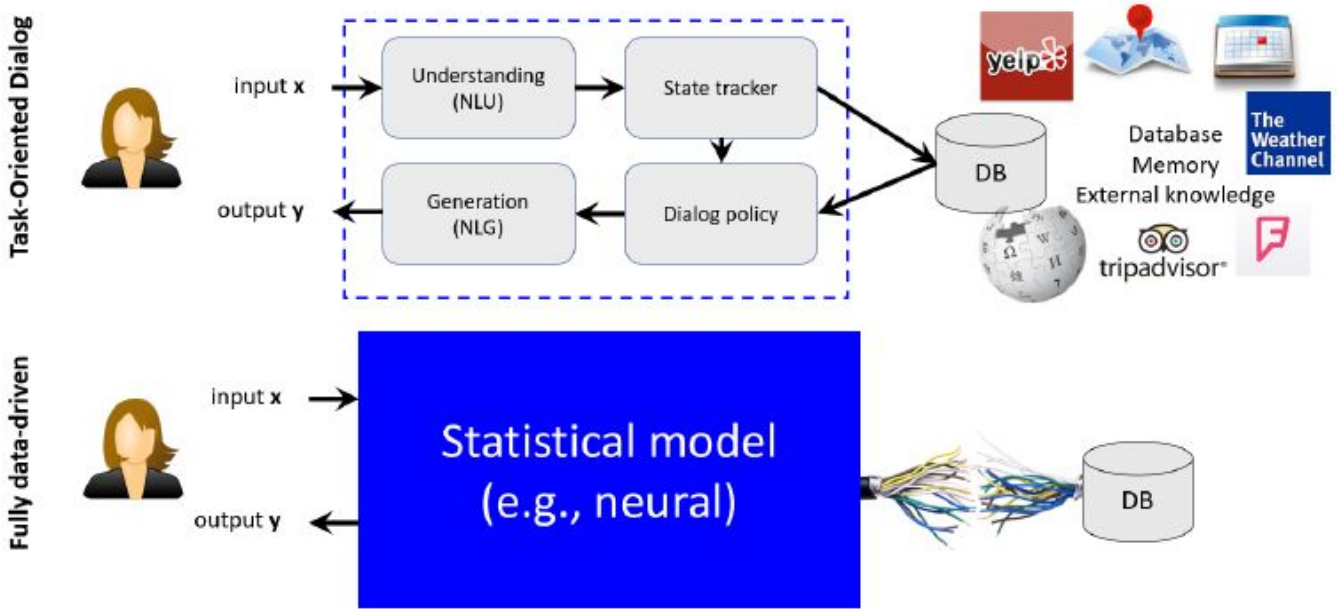


Figure 3: Two architectures of dialogue systems for **(Top)** traditional **task-oriented dialogue** and **(Bottom) fully data-driven dialogue**.

Modular and Non-modular Dialogue Systems

- A typical task-oriented dialogue agent is composed of four modules:
  - A Natural Language Understanding (NLU) module to identify **user intents**
  - A **state tracker** for tracking the **dialogue state**
  - A **dialogue policy** that selects the next action based on the current state
  - A **Natural Language Generation (NLG)** module to **generate responses** in natural language

★ *Most **task-oriented** bots are implemented **using a modular system**,*
★ ***Social chatbots**, on the other hand, are often implemented **using a unitary (non-modular) system**.*

Modular based Dialogue Systems

Task Oriented Dialogue Systems

Task-oriented Dialogue Systems

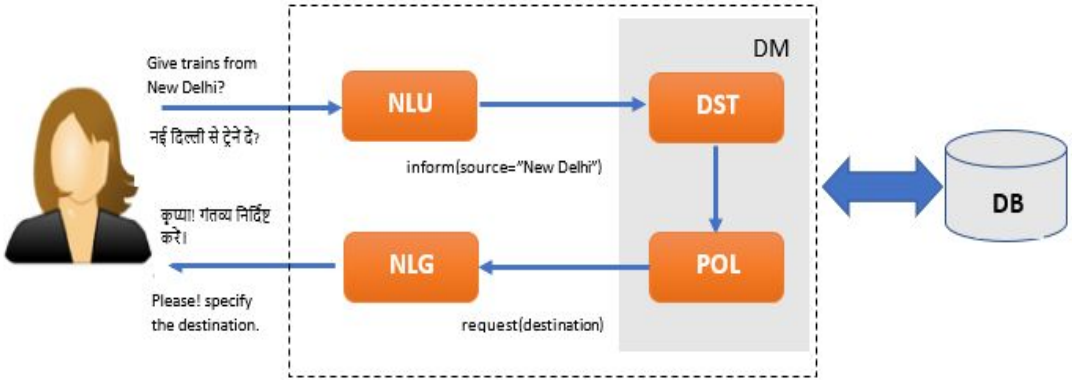
- **Task-oriented dialogue systems:** A dialogue system has the goal of helping a user solve some task like making an airplane reservation or buying a product.
- **GUS architecture:**
  - quite old, but long lived.
  - underlies most if not all modern commercial digital assistants.
  - based around frames.
  - **frame** is a kind of knowledge structure representing the kinds of ***intentions*** the system can extract from user **utterances**, and consists of a collection of ***slots***, each of which can take a set of possible values.
  - This set of frames is sometimes called a ***domain ontology.***

Task-oriented Dialogue Systems

Slot	Type	Question Template
ORIGIN CITY	city	“From what city are you leaving?”
DESTINATION CITY	city	“Where are you going?”
DEPARTURE TIME	time	“When would you like to leave?”
DEPARTURE DATE	date	“What day would you like to leave?”
ARRIVAL TIME	time	“When do you want to arrive?”
ARRIVAL DATE	date	“What day would you like to arrive?”

Control structure for Task-oriented dialogue Systems

- **Goal:** ***To fill the slots*** in the frame with the fillers the ***user intends***, and then ***perform the relevant action*** for the user (answering a question, or booking a flight).
  - system ***asks questions*** from the user (using pre-specified question templates associated with each slot of each frame, filling any slot that the user specifies.
  - once the system has enough information it performs the necessary action (like querying a database of flights) and ***returns the result to the user***.



Determining Domain, Intent, and Slot fillers

- The first aim of a task-oriented dialogue system is to extract three things from the user's utterance:
  - ***domain classification:*** talking about airlines servicers, restaurant booking or purchasing sports items.
  - ***intent determination:*** the general task or goal user is trying to accomplish.
  - ***slot filling:*** extract the particular slots and fillers that the user intends the system to understand from their utterance.
    - For example: *'Show me morning trains from Delhi to Patna on Tuesday'*

      a system might want to build a representation like:

```
DOMAIN: TRAIN-TRAVEL
INTENT: SHOW-TRAINS
ORIGIN-CITY: Delhi
ORIGIN-DATE: Tuesday
ORIGIN-TIME: morning
DEST-CITY: Patna
```

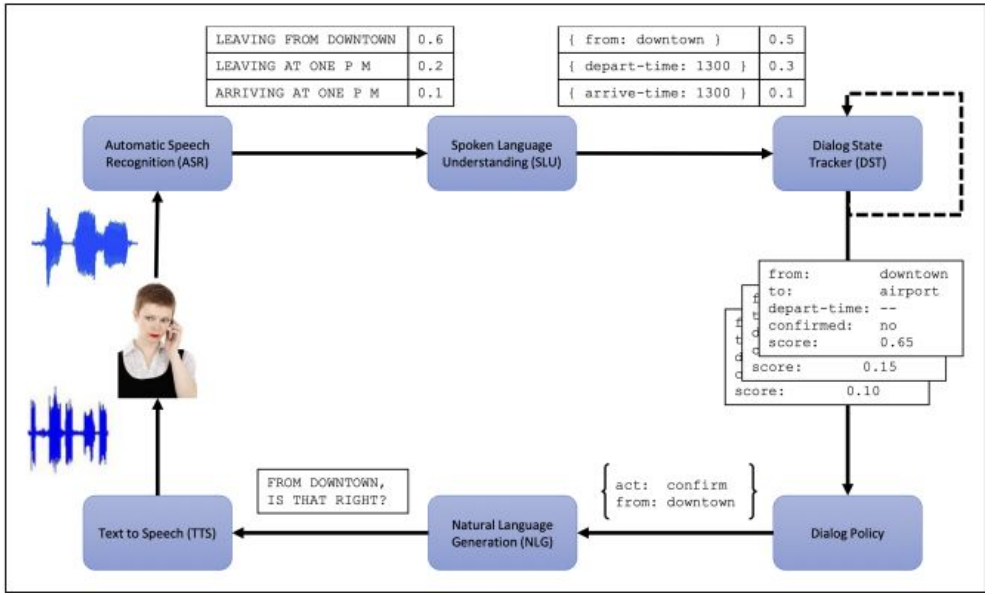
Natural Language Generation module

- Dialogue system produces the utterances that the system says to the user.
  - **template-based generation:** all or most of the words in the sentence be uttered to the user are pre-specified by the dialogue designer. (***prompts***)
  - Templates might be completely fixed (like *'Hello, how can I help you?'*), or ***can include some variables*** that are filled in by the generator. For example:
    - What time do you want to leave CITY-ORIG?
  - Possible to do some simple grounding. For example:

```
System: Did you want to review some more of
your personal profile?
Caller: No.
System: What's next?
System: Okay, What's next?
```

Dialogue State Architecture

- Modern research systems for task-based dialogue systems.
- More sophisticated version of the frame-based architecture called the dialogue-state or belief-state architecture.



Dialogue State Architecture

- **The dialogue state tracker** maintains the current state of the dialogue (which include the user's most recent dialogue act, plus the entire set of slot-filler constraints the user has expressed so far).
- **The dialogue policy** decides what the system should do or say next.
- A more sophisticated **natural language generation** component conditioning on the exact context to produce turns that seem much more natural.

★ most commercial system are architectural hybrids, based on GUS architecture augmented with some dialogue-state components.

Dialogue Acts

- **Dialogue acts** represent the interactive function of the ***turn*** or sentence, combining the idea of ***speech acts*** and ***grounding*** into a single representation.
- Different types of dialogue systems require labeling different kinds of acts, and so the tagset—defining what a dialogue act is exactly— tends to be designed for particular tasks.

Tag	Sys	User	Description
HELLO($a = x, b = y, \ldots$)	✓	✓	Open a dialogue and give info $a = x, b = y, \ldots$
INFORM($a = x, b = y, \ldots$)	✓	✓	Give info $a = x, b = y, \ldots$
REQUEST($a, b = x, \ldots$)	✓	✓	Request value for a given $b = x, \ldots$
REQALTS($a = x, \ldots$)	χ	✓	Request alternative with $a = x, \ldots$
CONFIRM($a = x, b = y, \ldots$)	✓	✓	Explicitly confirm $a = x, b = y, \ldots$
CONFREQ($a = x, \ldots, d$)	✓	χ	Implicitly confirm $a = x, \ldots$ and request value of d
SELECT($a = x, a = y$)	✓	χ	Implicitly confirm $a = x, \ldots$ and request value of d
AFFIRM($a = x, b = y, \ldots$)	✓	✓	Affirm and give further info $a = x, b = y, \ldots$
NEGATE($a = x$)	χ	✓	Negate and give corrected value $a = x$
DENY($a = x$)	χ	✓	Deny that $a = x$
BYE()	✓	✓	Close a dialogue

Dialogue Acts

- **Dialogue acts** represent the interactive function of the ***turn*** or sentence, combining the idea of ***speech acts*** and ***grounding*** into a single representation.
- Different types of dialogue systems require labeling different kinds of acts, and so the tagset—defining what a dialogue act is exactly— tends to be designed for particular tasks.

	Utterance	Dialogue act
U:	Hi, I am looking for somewhere to eat.	`hello(task = find,type=restaurant)`
S:	You are looking for a restaurant. What type of food do you like?	`confreq(type = restaurant, food)`
U:	I'd like an Italian somewhere near the museum.	`inform(food = Italian, near=museum)`
S:	Roma is a nice Italian restaurant near the museum.	`inform(name = "Roma", type = restaurant, food = Italian, near = museum)`
U:	Is it reasonably priced?	`confirm(pricerange = moderate)`
S:	Yes, Roma is in the moderate price range.	`affirm(name = "Roma", pricerange = moderate)`
U:	What is the phone number?	`request(phone)`
S:	The number of Roma is 385456.	`inform(name = "Roma", phone = "385456")`
U:	Ok, thank you goodbye.	`bye()`

Slot Filling

- The task of ***slot-filling***, and the simpler tasks of ***domain*** and ***intent classification*** can be treated as a supervised learning problem.
- We have a training set that ***associates each sentence*** with the ***correct set of slots, domain, and intent.***
- A simple method is to train a sequence model to map from input words representation to slot fillers, domain and intent. For example:

  `'I want to fly to San Francisco on Monday afternoon please'`

- Our training data is sentences paired with sequences of **BIO** labels:

```
O O    O  O   O    B-DES    I-DES      O     B-DEPTIME   I-DEPTIME     O
I want to fly to   San      Francisco  on    Monday      afternoon     please
```

Slot Filling