# CSE 4392 Special Topics
# Natural Language Processing

# Dialogue Systems



2024 Spring

# Overview

- What's a dialogue system?
- Properties of Human Conversation
- Chatbots v.s. Task-oriented dialogues systems
- Rule-based v.s. Data-driven
- Remaining Challenges

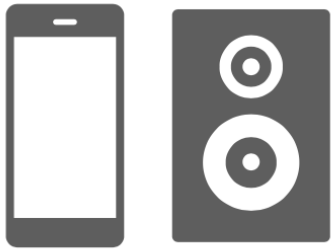
# What's a Dialogue System?

- Dialogue system is everywhere. Did you use it?


# What is Dialogue System?

- Dialogue is the preferred mode of interaction:

Desktop



Smart Mobile
Embedded Devices

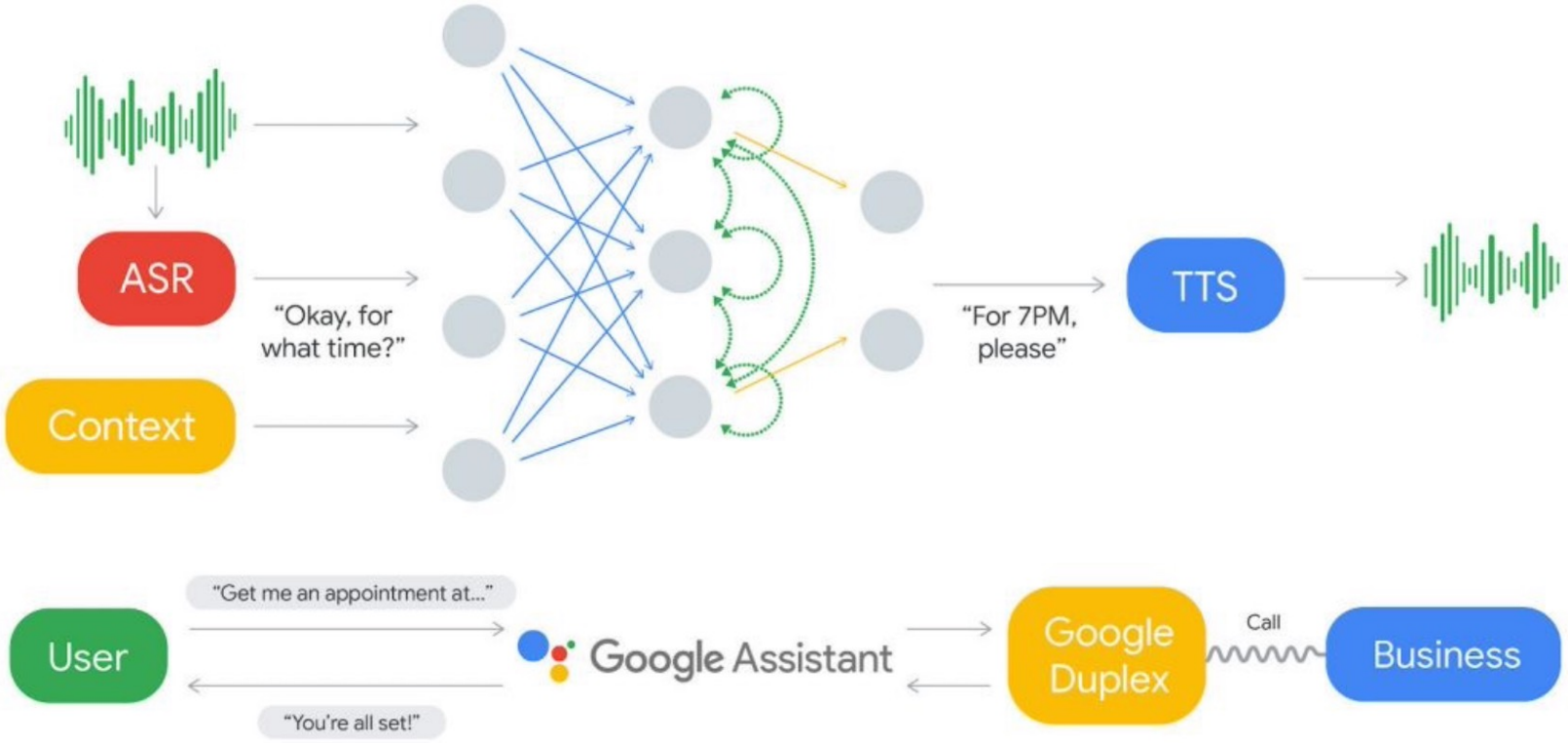
keyboard & mouse

"turn off the light."

language

# What is Dialogue System?

- Google Duplex: can you distinguish between human and AI?



# Dialogue System is a Killer App for NLP

**They can**

- give travel directions
- control home appliances
- find restaurants
- help make phone calls
- customer services
- …

# Properties of Human Conversation

| | |
|---|---|
| $C_1$: | . . . I need to travel in May. |
| $A_1$: | And, what day in May did you want to travel? |
| $C_2$: | OK uh I need to be there for a meeting that's from the 12th to the 15th. |
| $A_2$: | And you're flying into what city? |
| $C_3$: | Seattle. |
| $A_3$: | And what time would you like to leave Pittsburgh? |
| $C_4$: | Uh hmm I don't think there's many options for non-stop. |
| $A_4$: | Right. There's three non-stops today. |
| $C_5$: | What are they? |
| $A_5$: | 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. |
| $C_6$: | OK I'll take the 5ish flight on the night before on the 11th. |
| $A_6$: | On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115. |
| $C_7$: | OK. |
| $A_7$: | And you said returning on May 15th? |
| $C_8$: | Uh, yeah, at the end of the day. |
| $A_8$: | OK. There's #two non-stops . . . # |
| $C_9$: | #Act. . . actually #, what day of the week is the 15th? |
| $A_9$: | It's a Friday. |
| $C_{10}$: | Uh hmm. I would consider staying there an extra day til Sunday. |
| $A_{10}$: | OK. . . OK. On Sunday I have . . . |

Turn structure:
C-A-C-A-….

Spoken DS:
end-point detection
Know when to take the turn

# Properties of Human Conversation

| | |
|---|---|
| $C_1$: | . . . I need to travel in May. |
| $A_1$: | And, what day in May did you want to travel? |
| $C_2$: | OK uh I need to be there for a meeting that's from the 12th to the 15th. |
| $A_2$: | And you're flying into what city? |
| $C_3$: | Seattle. |
| $A_3$: | And what time would you like to leave Pittsburgh? |
| $C_4$: | Uh hmm I don't think there's many options for non-stop. |
| $A_4$: | Right. There's three non-stops today. |
| $C_5$: | What are they? |
| $A_5$: | 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. |
| $C_6$: | OK I'll take the 5ish flight on the night before on the 11th. |
| $A_6$: | On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115. |
| $C_7$: | OK. |
| $A_7$: | And you said returning on May 15th? |
| $C_8$: | Uh, yeah, at the end of the day. |
| $A_8$: | OK. There's #two non-stops . . . # |
| $C_9$: | #Act. . . actually #, what day of the week is the 15th? |
| $A_9$: | It's a Friday. |
| $C_{10}$: | Uh hmm. I would consider staying there an extra day til Sunday. |
| $A_{10}$: | OK. . . OK. On Sunday I have . . . |

# overlaps

# Properties of Human Conversation

| | |
|---|---|
| $C_1$: | . . . I need to travel in May. |
| $A_1$: | And, what day in May did you want to travel? |
| $C_2$: | OK uh I need to be there for a meeting that's from the 12th to the 15th. |
| $A_2$: | And you're flying into what city? |
| $C_3$: | Seattle. |
| $A_3$: | And what time would you like to leave Pittsburgh? |
| $C_4$: | Uh hmm I don't think there's many options for non-stop. |
| $A_4$: | Right. There's three non-stops today. |
| $C_5$: | What are they? |
| $A_5$: | 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. |
| $C_6$: | OK I'll take the 5ish flight on the night before on the 11th. |
| $A_6$: | On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115. |
| $C_7$: | OK. |
| $A_7$: | And you said returning on May 15th? |
| $C_8$: | Uh, yeah, at the end of the day. |
| $A_8$: | OK. There's #two non-stops . . . # |
| $C_9$: | #Act. . . actually #, what day of the week is the 15th? |
| $A_9$: | It's a Friday. |
| $C_{10}$: | Uh hmm. I would consider staying there an extra day til Sunday. |
| $A_{10}$: | OK. . . OK. On Sunday I have . . . |

Asking
Answering

# Properties of Human Conversation

| | |
|---|---|
| **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:** | committing 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*) |

A taxonomy of **dialog acts** (Bach and Harnish, 1979)

# Properties of Human Conversation

| | |
|---|---|
| $C_1$: | . . . I need to travel in May. |
| $A_1$: | And, what day in May did you want to travel? |
| $C_2$: | OK uh I need to be there for a meeting that's from the 12th to the 15th. |
| $A_2$: | And you're flying into what city? |
| $C_3$: | Seattle. |
| $A_3$: | And what time would you like to leave Pittsburgh? |
| $C_4$: | Uh hmm I don't think there's many options for non-stop. |
| $A_4$: | Right. There's three non-stops today. |
| $C_5$: | What are they? |
| $A_5$: | 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. |
| $C_6$: | OK I'll take the 5ish flight on the night before on the 11th. |
| $A_6$: | On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115. |
| $C_7$: | OK. |
| $A_7$: | And you said returning on May 15th? |
| $C_8$: | Uh, yeah, at the end of the day. |
| $A_8$: | OK. There's #two non-stops . . . # |
| $C_9$: | #Act. . . actually #, what day of the week is the 15th? |
| $A_9$: | It's a Friday. |
| $C_{10}$: | Uh hmm. I would consider staying there an extra day til Sunday. |
| $A_{10}$: | OK. . . OK. On Sunday I have . . . |

**grounding:** acknowledge that the listener has understood the speaker

# Properties of Human Conversation

| | |
|---|---|
| $C_1$: | . . . I need to travel in May. |
| $A_1$: | And, what day in May did you want to travel? |
| $C_2$: | OK uh I need to be there for a meeting that's from the 12th to the 15th. |
| $A_2$: | And you're flying into what city? |
| $C_3$: | Seattle. |
| $A_3$: | And what time would you like to leave Pittsburgh? |
| $C_4$: | Uh hmm I don't think there's many options for non-stop. |
| $A_4$: | Right. There's three non-stops today. |
| $C_5$: | What are they? |
| $A_5$: | 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. |
| $C_6$: | OK I'll take the 5ish flight on the night before on the 11th. |
| $A_6$: | On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115. |
| $C_7$: | OK. |
| $A_7$: | And you said returning on May 15th? |
| $C_8$: | Uh, yeah, at the end of the day. |
| $A_8$: | OK. There's #two non-stops . . . # |
| $C_9$: | #Act. . . actually #, what day of the week is the 15th? |
| $A_9$: | It's a Friday. |
| $C_{10}$: | Uh hmm. I would consider staying there an extra day til Sunday. |
| $A_{10}$: | OK. . . OK. On Sunday I have . . . |

local structure between dialog acts

correction subdialogue

# Properties of Human Conversation

| | |
|---|---|
| $C_1$: | . . . I need to travel in May. |
| $A_1$: | And, what day in May did you want to travel? |
| $C_2$: | OK uh I need to be there for a meeting that's from the 12th to the 15th. |
| $A_2$: | And you're flying into what city? |
| $C_3$: | Seattle. |
| $A_3$: | And what time would you like to leave Pittsburgh? |
| $C_4$: | Uh hmm I don't think there's many options for non-stop. |
| $A_4$: | Right. There's three non-stops today. |
| $C_5$: | What are they? |
| $A_5$: | 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. |
| $C_6$: | OK I'll take the 5ish flight on the night before on the 11th. |
| $A_6$: | On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115. |
| $C_7$: | OK. |
| $A_7$: | And you said returning on May 15th? |
| $C_8$: | Uh, yeah, at the end of the day. |
| $A_8$: | OK. There's #two non-stops . . . # |
| $C_9$: | #Act. . . actually #, what day of the week is the 15th? |
| $A_9$: | It's a Friday. |
| $C_{10}$: | Uh hmm. I would consider staying there an extra day til Sunday. |
| $A_{10}$: | OK. . . OK. On Sunday I have . . . |

**mixed initiative:** conversation is not controlled by one participant

**user-initiative v.s. system-initiative**

# Properties of Human Conversation

| | |
|---|---|
| $C_1$: | . . . I need to travel in May. |
| $A_1$: | And, what day in May did you want to travel? |
| $C_2$: | OK uh I need to be there for a meeting that's from the 12th to the 15th. |
| $A_2$: | And you're flying into what city? |
| $C_3$: | Seattle. |
| $A_3$: | And what time would you like to leave Pittsburgh? |
| $C_4$: | Uh hmm I don't think there's many options for non-stop. |
| $A_4$: | Right. There's three non-stops today. |
| $C_5$: | What are they? |
| $A_5$: | 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. |
| $C_6$: | OK I'll take the 5ish flight on the night before on the 11th. |
| $A_6$: | On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115. |
| $C_7$: | OK. |
| $A_7$: | And you said returning on May 15th? |
| $C_8$: | Uh, yeah, at the end of the day. |
| $A_8$: | OK. There's #two non-stops . . . # |
| $C_9$: | #Act. . . actually #, what day of the week is the 15th? |
| $A_9$: | It's a Friday. |
| $C_{10}$: | Uh hmm. I would consider staying there an extra day til Sunday. |
| $A_{10}$: | OK. . . OK. On Sunday I have . . . |

conversational implicature

# Properties of Human Conversation

- Turn structure
- Dialog acts
- Grounding
- Dialogue structure
- Initiative
- Implicature

- These **subtle characteristics of human conversations** are among the reasons it is difficult to build dialogue systems that can carry on **natural conversations with humans**.

# Chatbots vs. Task-oriented Dialogue Systems

## Chatbot

| | |
|---|---|
| User: | Do you like Doctor Who? |
| Cleverbot: | Yes, so funny. |
| User: | What's your favorite show? |
| Cleverbot: | I don't have one. |
| User: | What's your name? |
| Cleverbot: | Jason I already told you. |
| User: | No your real name. |
| Cleverbot: | Stephanie IS my real name. |
| User: | I thought it was Jason? |
| Cleverbot: | How did you come to that conclusion. |
| User: | Because you said so. |
| Cleverbot: | I have to do my homework. |

**Goal**: mimicking the **unstructured conversational** characteristic of **human-human interaction**

## Task-oriented Dialogue System

| | |
|---|---|
| GUS: | Hello. My name is GUS. I can help you plan a simple trip by air. Where do you want to go ? |
| Client: | I want to go to San Diego on May 28 |
| GUS: | What time do you want to leave? |
| Client: | I must be in San Diego before 10 am |
| GUS: | Would you like P.S.A. flight 102 that arrives at 9.15 am ? |
| Client: | What is the next flight ? |
| GUS: | Air California flight 310 that leaves at 8.30am and arrives at 10.00am |
| Client: | I'll take the first one |
| GUS: | In what name should I make a reservation, please? |
| Client: | Dan Bobrow |
| GUS: | I have confirmed the following flight: P.S.A. flight 102 on Wednesday May 28 from San Jose to San Diego that leaves at 7.30 am and arrives at 9.15 am. |
| GUS: | What date do you want to return on ? |
| Client: | On Friday in the evening. |
| GUS: | Would you like the flight that leaves at 7.45 pm ? |
| Client: | That's fine. |
| GUS: | I have confirmed the following flight: P.S.A. flight 307 on Friday May 30 from San Diego to San Jose that leaves at 7.45 pm and arrives at 9.30 pm Thank you for calling. Goodbye |

**Goal**: get information from the user to help **complete the specific task**.

# CHATBOTS: RULE-BASED

```
function ELIZA GENERATOR(user sentence) returns response

Find the word w in sentence that has the highest keyword rank
  if w exists
       Choose the highest ranked rule r for w that matches sentence
       response ← Apply the transform in r to sentence
       if w = 'my'
            future ← Apply a transformation from the 'memory' rule list to sentence
            Push future onto memory stack
   else (no keyword applies)
     either
          response ← Apply the transform for the NONE keyword to sentence
     or
          response ← Pop the top response from the memory stack
  return(response)
```

- A simplified sketch of the ELIZA algorithm. The power of the algorithm comes from the **particular transformations associated with each keyword**.

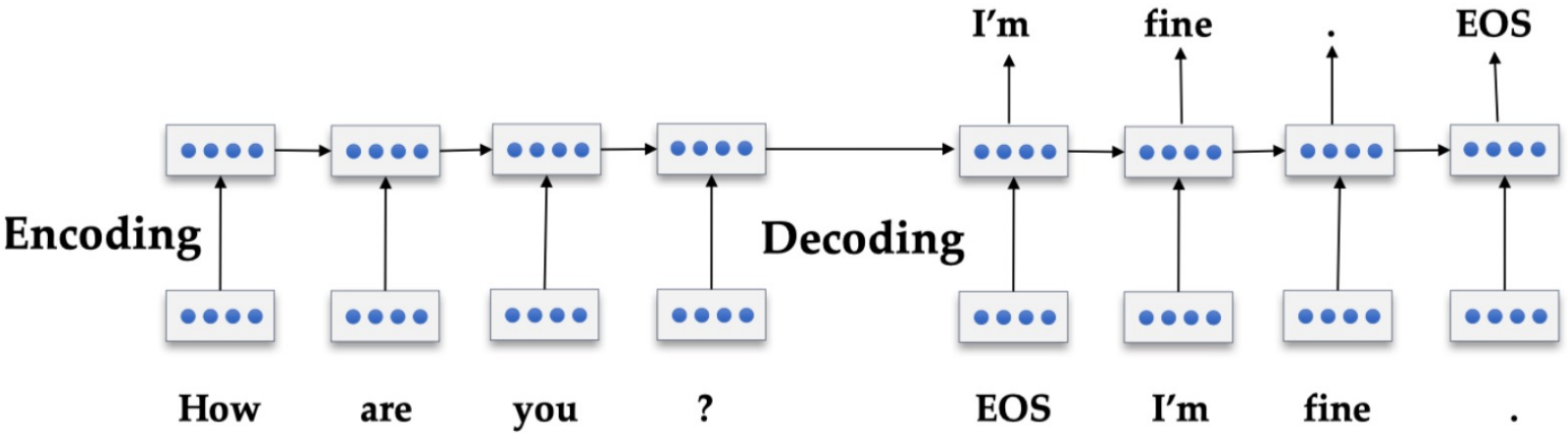
# Chatbots: Corpus-based

Given user query *q* and a conversational corpus *C*.
Find the turn *t* in *C* that is most similar to *q, and return the following turn.*

$$r = response\left(\underset{t \in C}{\text{argmax}} \frac{q^T t}{||q||t||}\right)$$

- **Corpus-based method (Information Retrieval):** Return the response to the most similar turn. (Jafarpour et al. 2009, Leuski and Traum 2011)

# Chatbots: Corpus-based




- **Corpus-based method (Seq2Seq):** An encoder decoder model for neural response generation in dialogue. RNN in the above diagram can be replaced with transformer.

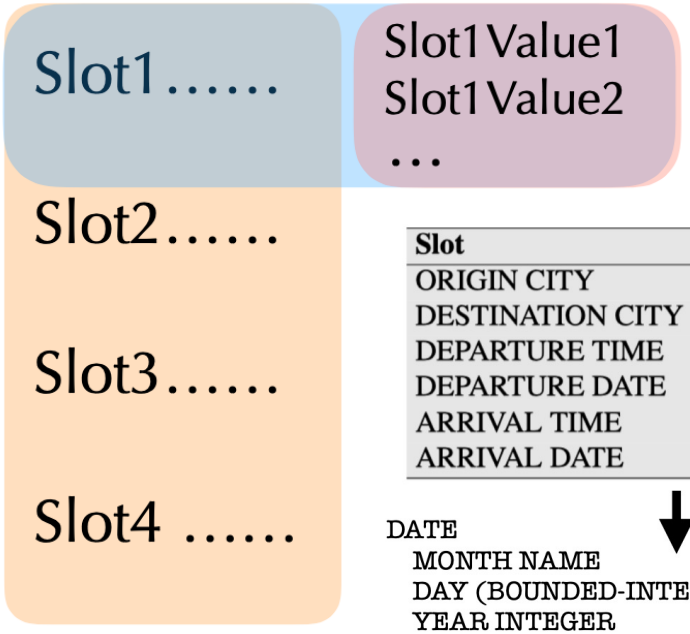
# Task-oriented Dialogue System

- How to incorporate task related knowledge?

**Domain-Specific Knowledge:**
**Ontology / Frame / Slot / Value**

a knowledge structure representing the kinds of intentions the system can extract from user sentences.

contains one or more frames.
A frame is a set of slot-value pairs to establish an intent.




| Slot | Type |
|---|---|
| ORIGIN CITY | city |
| DESTINATION CITY | city |
| DEPARTURE TIME | time |
| DEPARTURE DATE | date |
| ARRIVAL TIME | time |
| ARRIVAL DATE | date |

DATE
MONTH NAME
DAY (BOUNDED-INTEGER 1 31)
YEAR INTEGER

# Task-oriented Dialogue System

- How to incorporate task related knowledge?

**"Show me morning flights from Boston to San Francisco on Tuesday"**

**Step#1: domain classification** — DOMAIN: AIR-TRAVEL

**Step#2: intent determination** — INTENT: SHOW-FLIGHTS

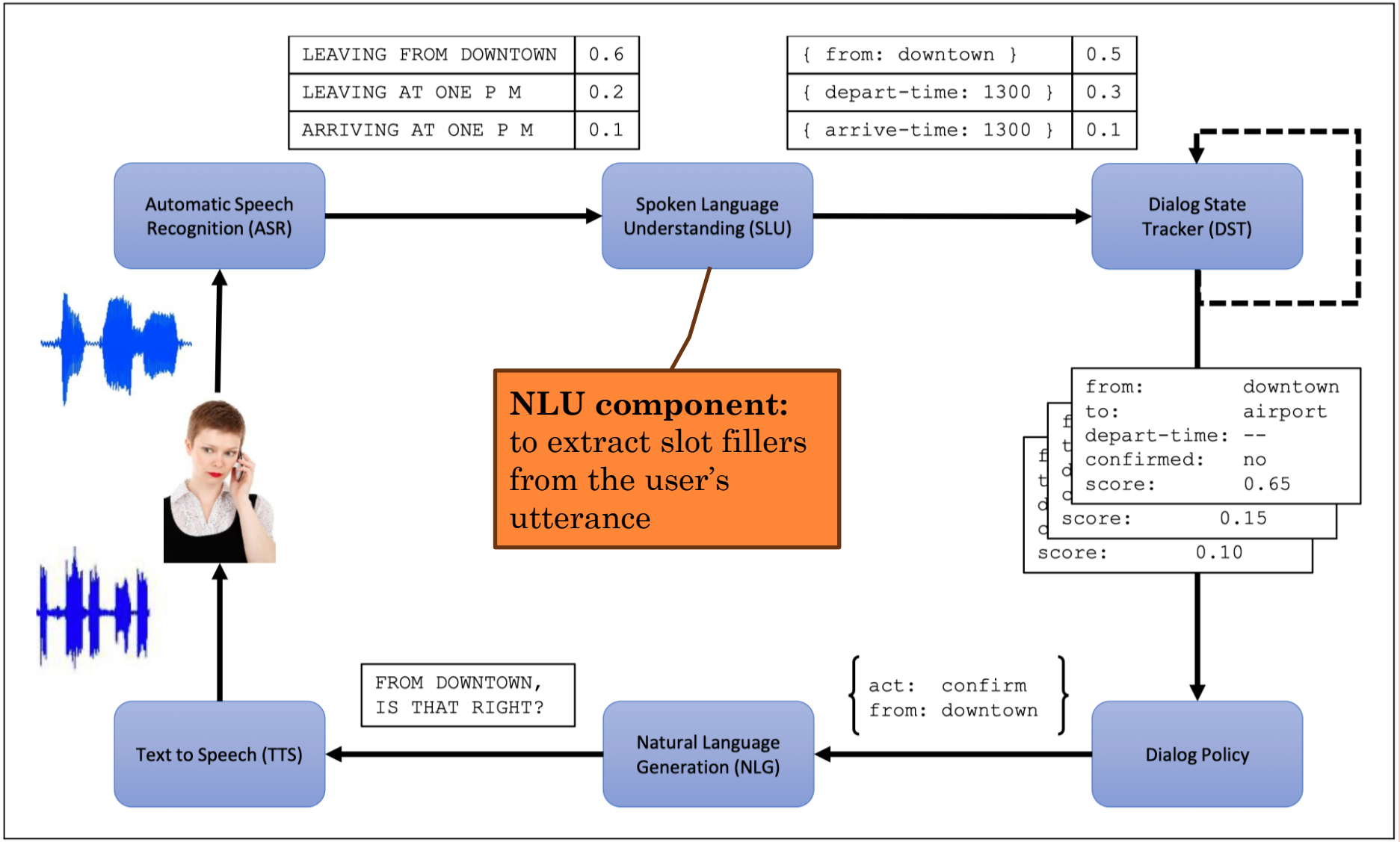
**Step#3: slot filling** — ORIGIN-CITY: Boston
ORIGIN-DATE: Tuesday
ORIGIN-TIME: morning
DEST-CITY: San Francisco

# Task-oriented Dialogue System

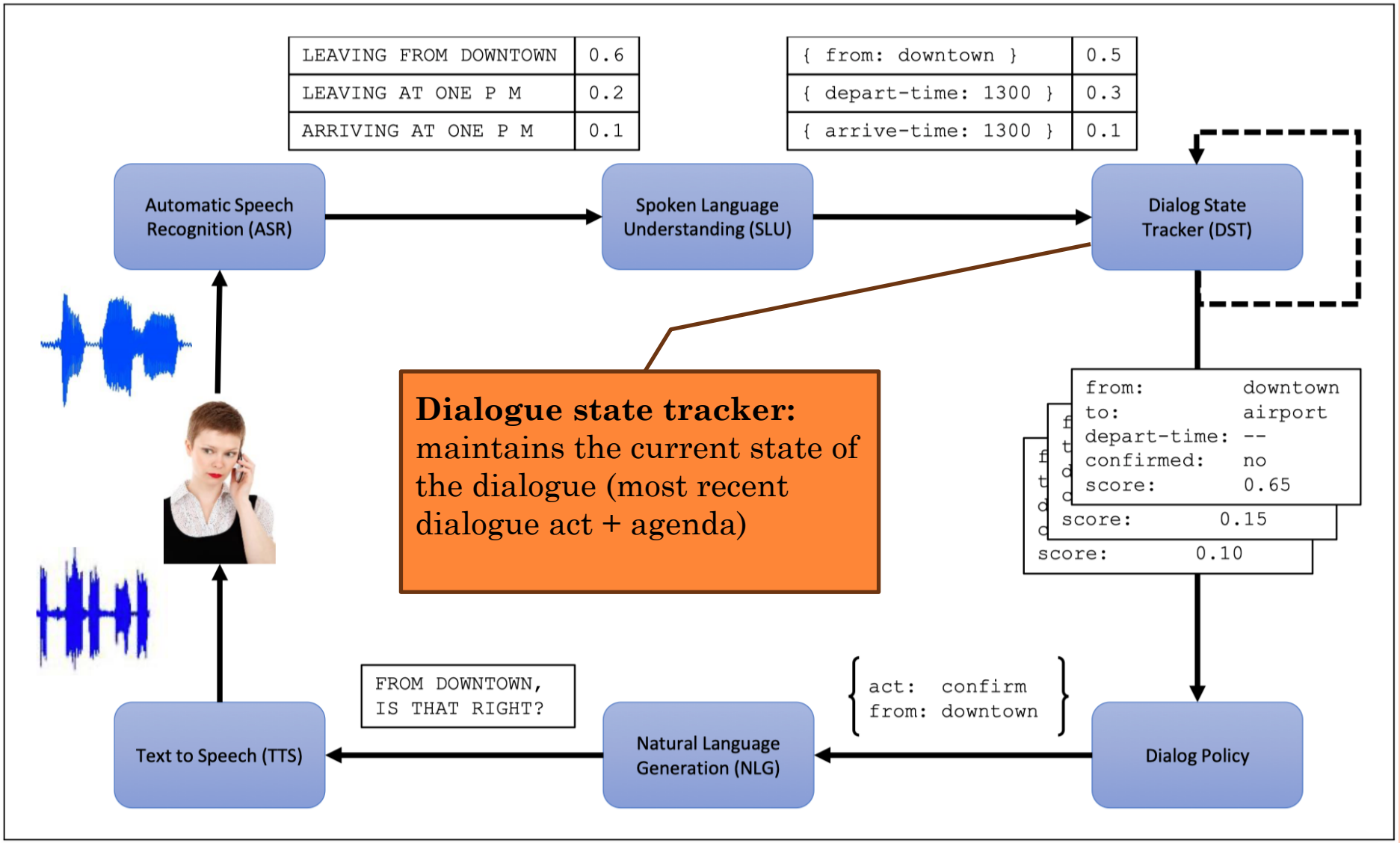
- How to incorporate task related knowledge?

| | 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? | `confirm(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()` |

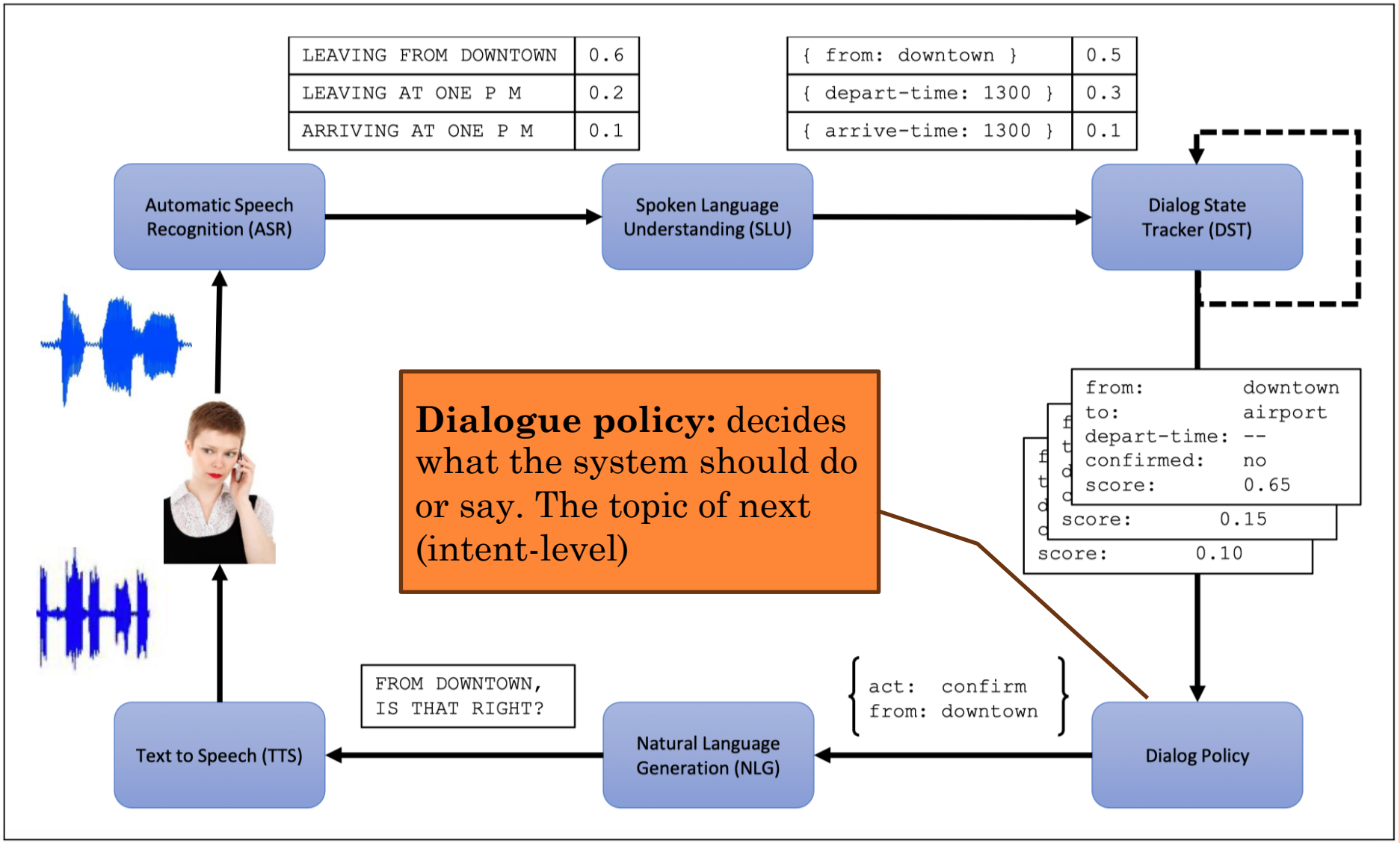
# Architecture of Task-oriented DS



# Architecture of Task-oriented DS



# Architecture of Task-oriented DS



# Rule-based vs. Data-driven

- How to build a task-oriented dialog system?

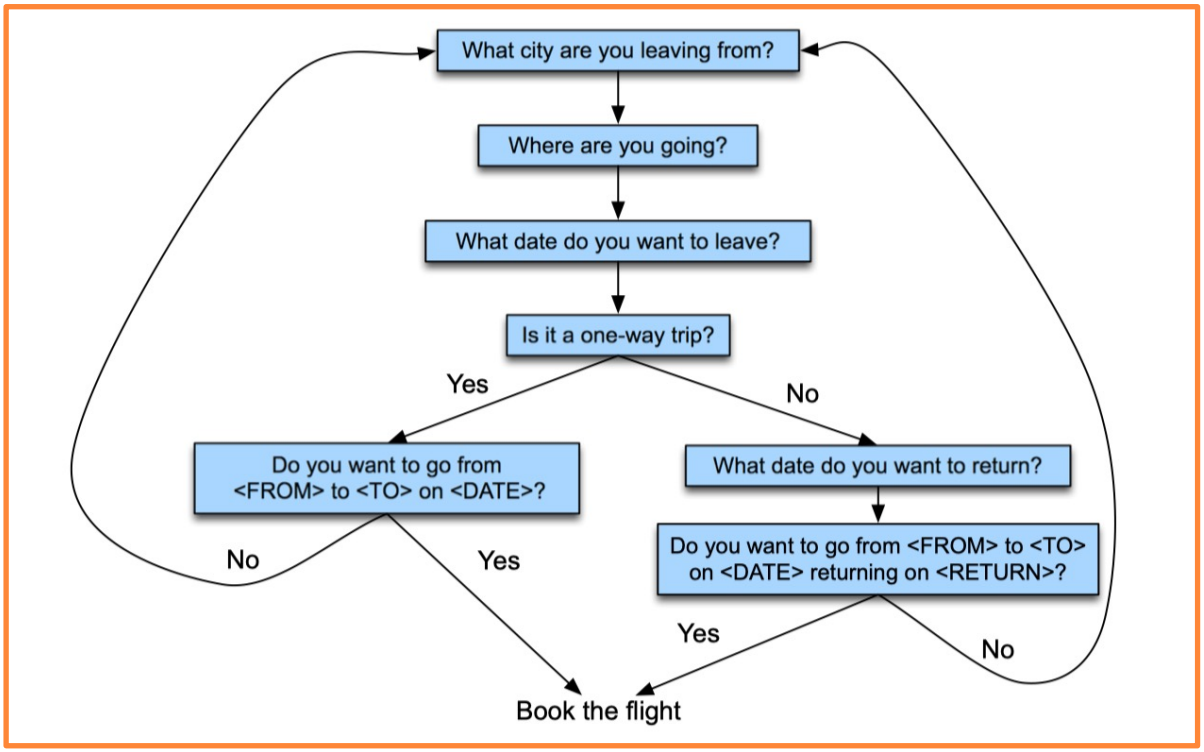
| | | |
|---|---|---|
| SHOW | → | show me \| i want \| can i see\|... |
| DEPART_TIME_RANGE | → | (after\|around\|before) HOUR \| morning \| afternoon \| evening |
| HOUR | → | one\|two\|three\|four...\|twelve (AMPM) |
| FLIGHTS | → | (a) flight \| flights |
| AMPM | → | am \| pm |
| ORIGIN | → | from CITY |
| DESTINATION | → | to CITY |
| CITY | → | Boston \| San Francisco \| Denver \| Washington |

Semantic grammars can be parsed by any Context-Free Grammar parsing algorithm.

**Rule-based system (SLU/DST)**

# Rule-based vs. Data-driven

- How to build a task-oriented dialog system?



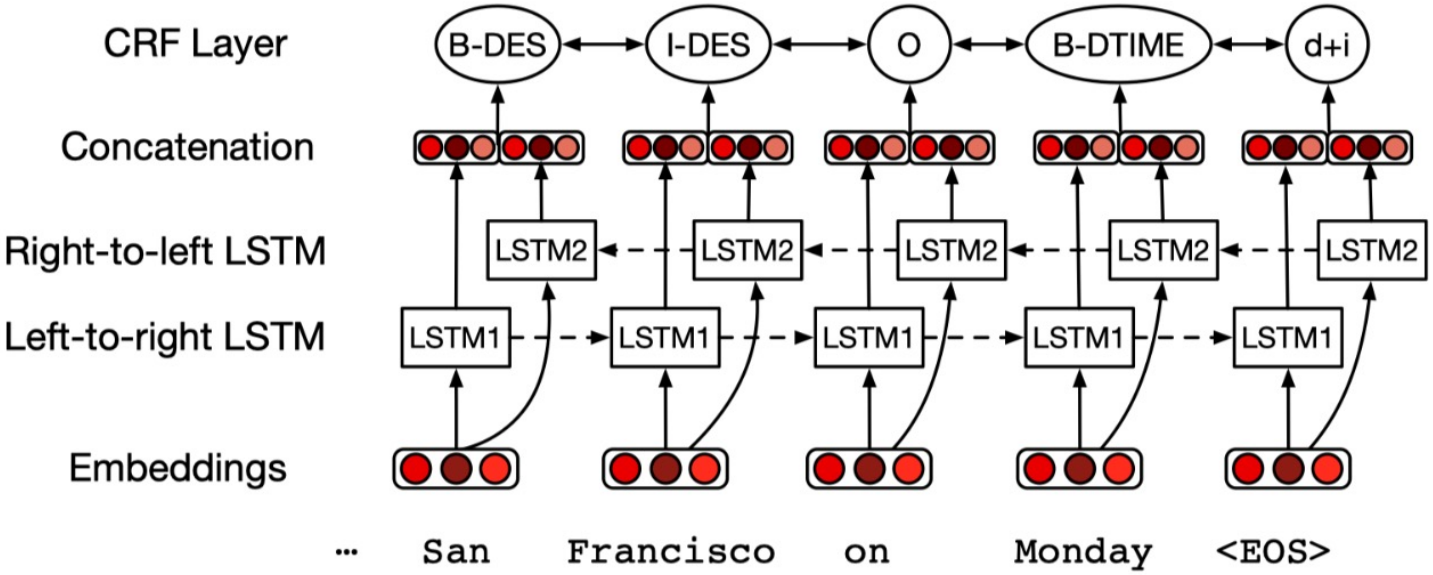

This is the frame!

A simple finite-state automaton architecture for frame-based dialog.

**Rule-based system (Dialogue policy)**

# Rule-based vs. Data-driven

- How to build a task-oriented dialog system?



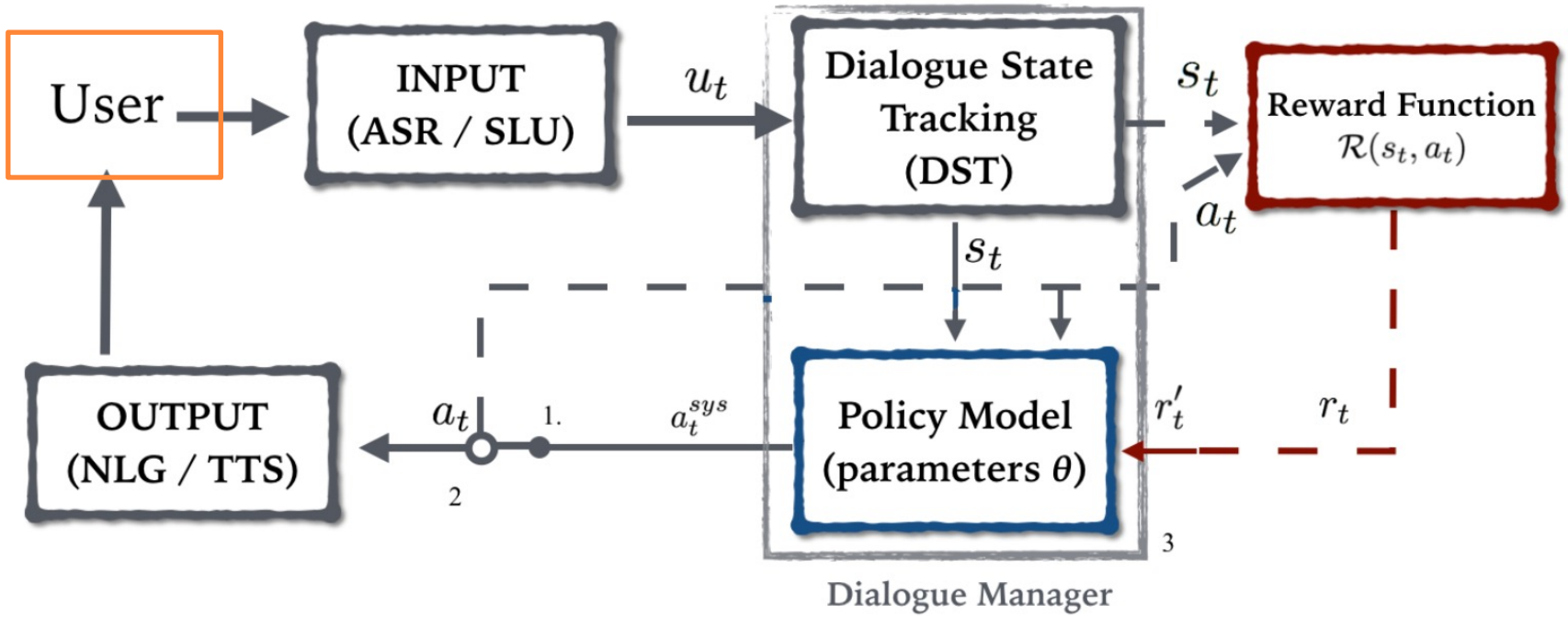

An LSTM architecture for slot filling, mapping the words in the input to a series of IOB tags plus a final state consisting of a domain concatenated with an intent.

**Data-driven system (SLU/DST)**

# Rule-based vs. Data-driven

- How to build a task-oriented dialog system?

**Simulator?**




**Data-driven system (Dialogue policy)**

# Rule-based vs. Data-driven

- Pros and Cons?

- **Rule-Based Methods**

  - hand-craft rules, “safe” but not “flexible”.

  - cheap in terms of dataset.

  - expensive in terms of engineering.

- **Data-Driven Methods**

  - learn from interactions, dialogue manager is **evolvable**.

  - uncontrolled behavior in unseen situation.

  - cheap in terms of engineering, but expensive in terms of data/interaction

# Remaining Challenges

- Understanding the context (better with LLMs permitting long inputs)
- Domain adaptation
- Data scarcity (chat data is very hard to obtain)
- Privacy protection
- Accuracy/Reliability (LLM's hullucination problem)
- Biases and toxicity