

CMPT 825: Natural Language Processing

# Dialogue

Spring 2020  
2020-04-07

Adapted from slides from Danqi Chen and Karthik Narasimhan  
(with some content from slides from Chris Manning and Dan Jurafsky)

## **Participation reminder:**

- 5% of grade

## **Final Project**

- Due next Tuesday: April 14th (no grace days)

### **Tips for final report**

- Proof-read your paper and fix grammar/wording issues
- Include diagrams to explain your problem statement (input/output), network architecture
- Include tables/graphs for data statistics and experiment results
- Provide clear examples
- Provide comparisons and analysis of results

# Final Project Report

Abstract  
Introduction

Prior related  
work

Model

Model

Data

Experiments

Results

Analysis &  
Conclusion

# Tips for good final projects

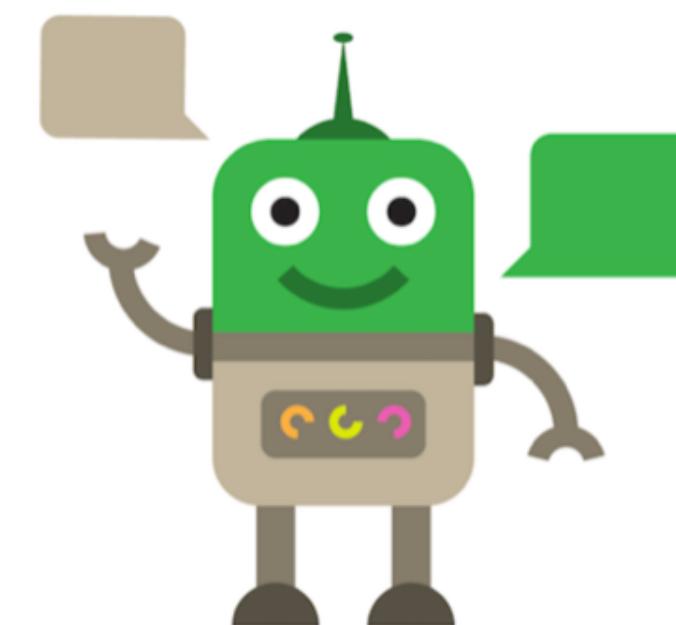
- Have a clear, well-defined hypothesis to be tested
  - (++ novel/creative hypothesis)
- Conclusions and results should teach the reader something
- Meaningful tables, plots to display the key results
  - ++ nice visualizations or interactive demos
  - ++ novel/impressive engineering feat
  - ++ good results

# What to avoid

- All experiments run with prepackaged source - no extra code written for model/data processing
- Just ran model once or twice on the data and reported results (not much hyperparameter search done)
- A few standard graphs: loss curves, accuracy, without any analysis
- Results/Conclusion don't say much besides that it didn't work
  - Even if results are negative, analyze them

# Overview

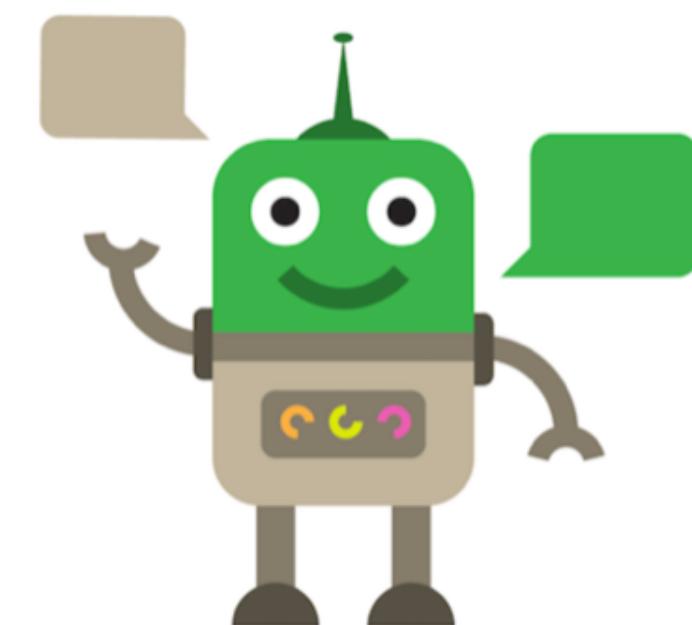
- What's a dialogue system?
- Properties of Human Conversation
- Chatbots v.s. Task-oriented dialogues systems
- Rule-based v.s. Data-driven
- Remaining Challenges



Dialogue Systems

# Overview

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

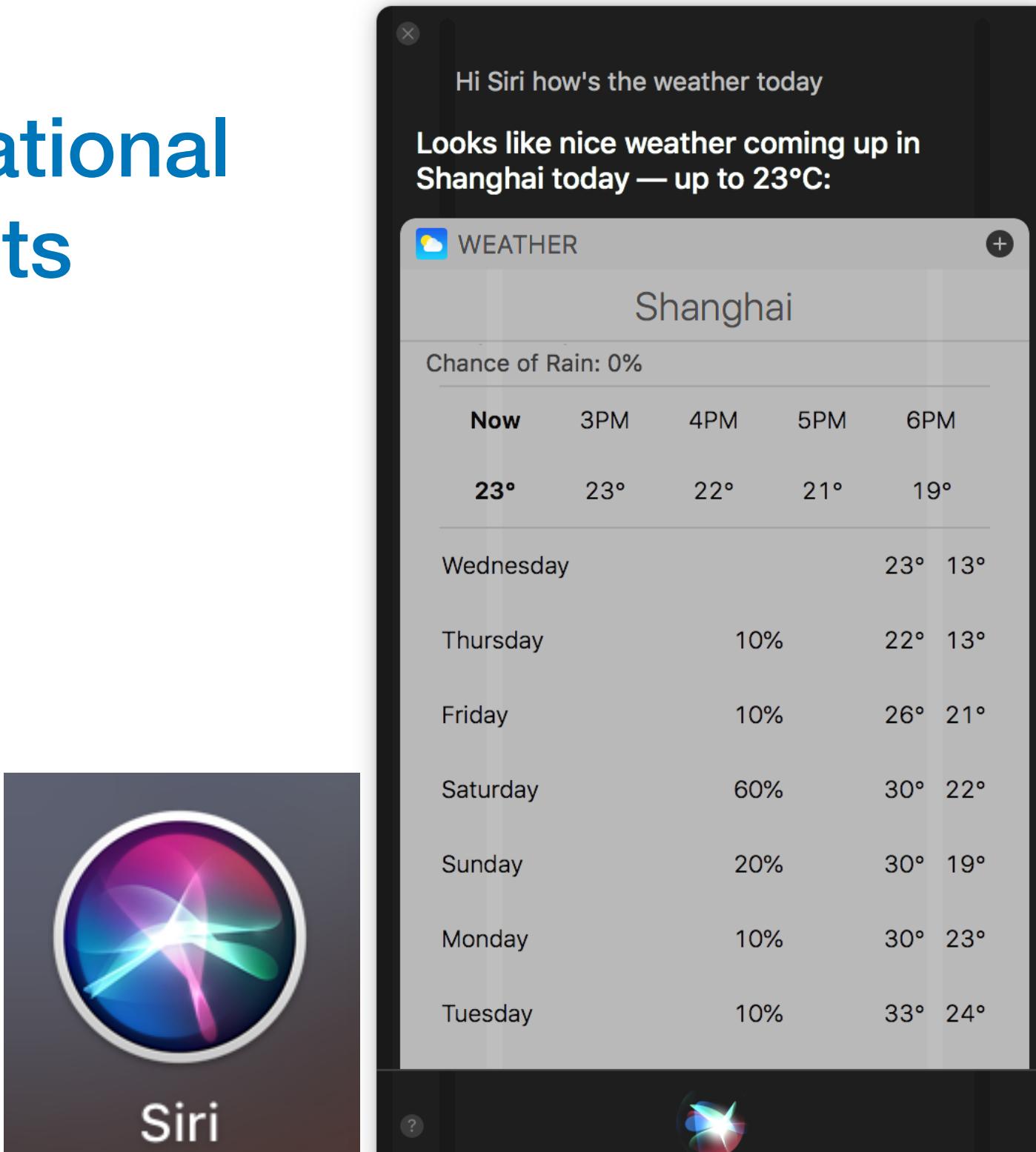


Dialogue Systems

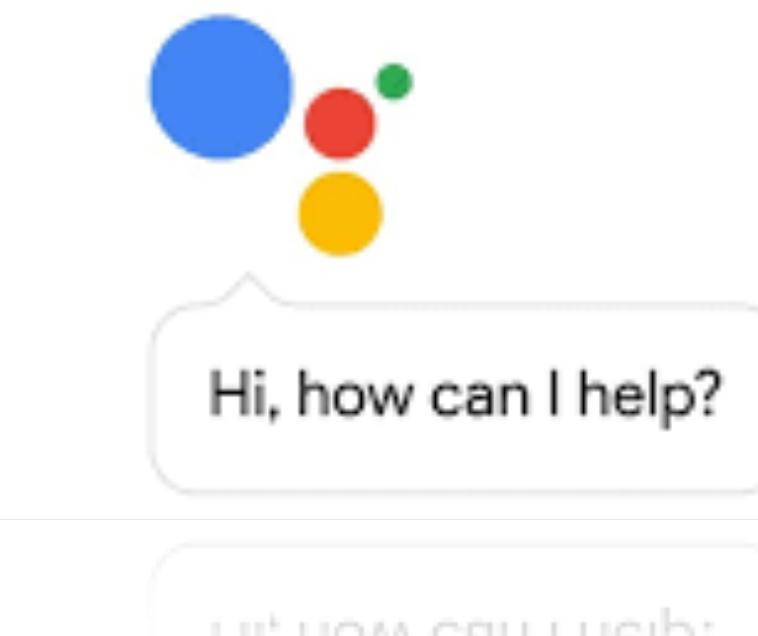
# What's a Dialogue System?

Dialog Systems are **HOT** 🔥. — Did you use it?

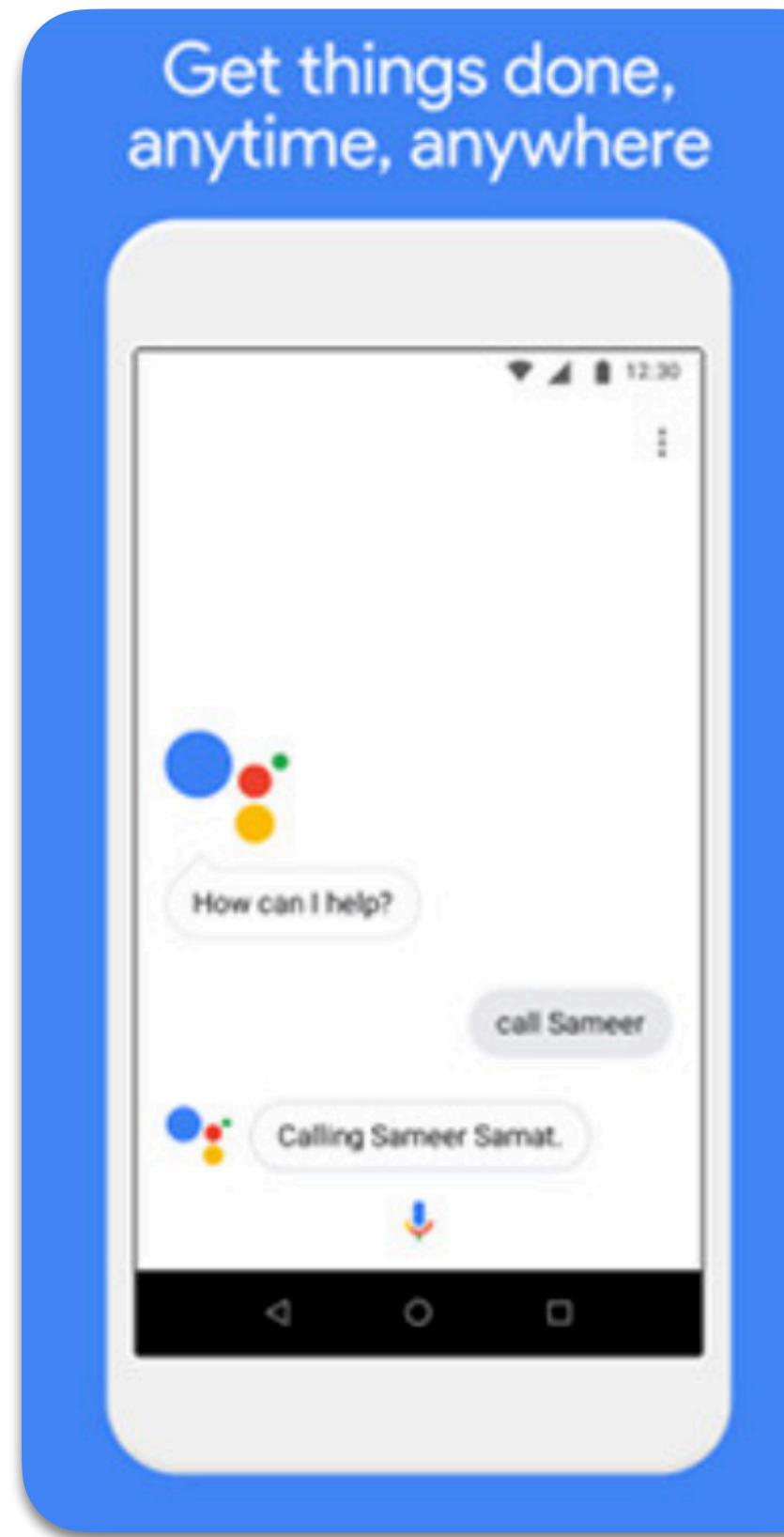
Conversational  
agents



Apple



Google

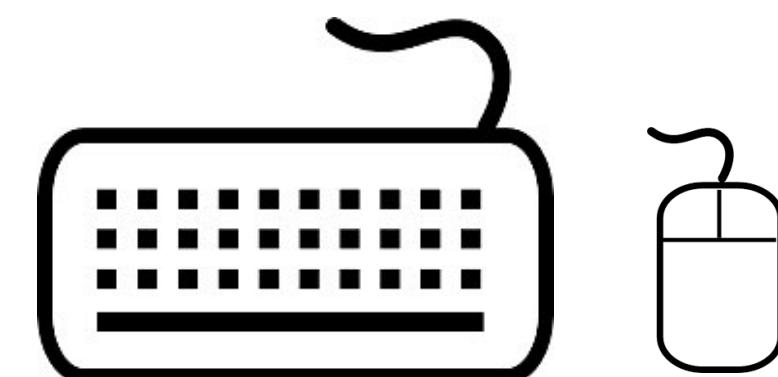


# What's a Dialogue System?

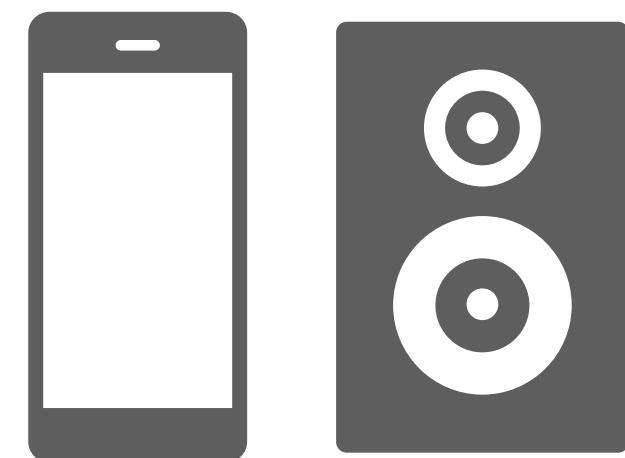
Dialog Systems are **HOT** 🔥. — Preferable user interface.



**Desktop**



**keyboard & mouse**



**Smart Mobile  
Embedded Devices**

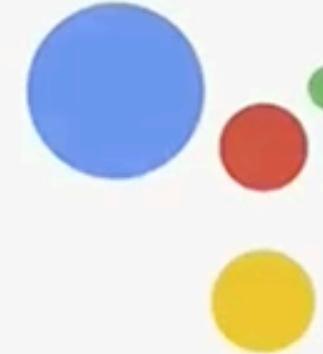
**“turn off the light.”**



**language**

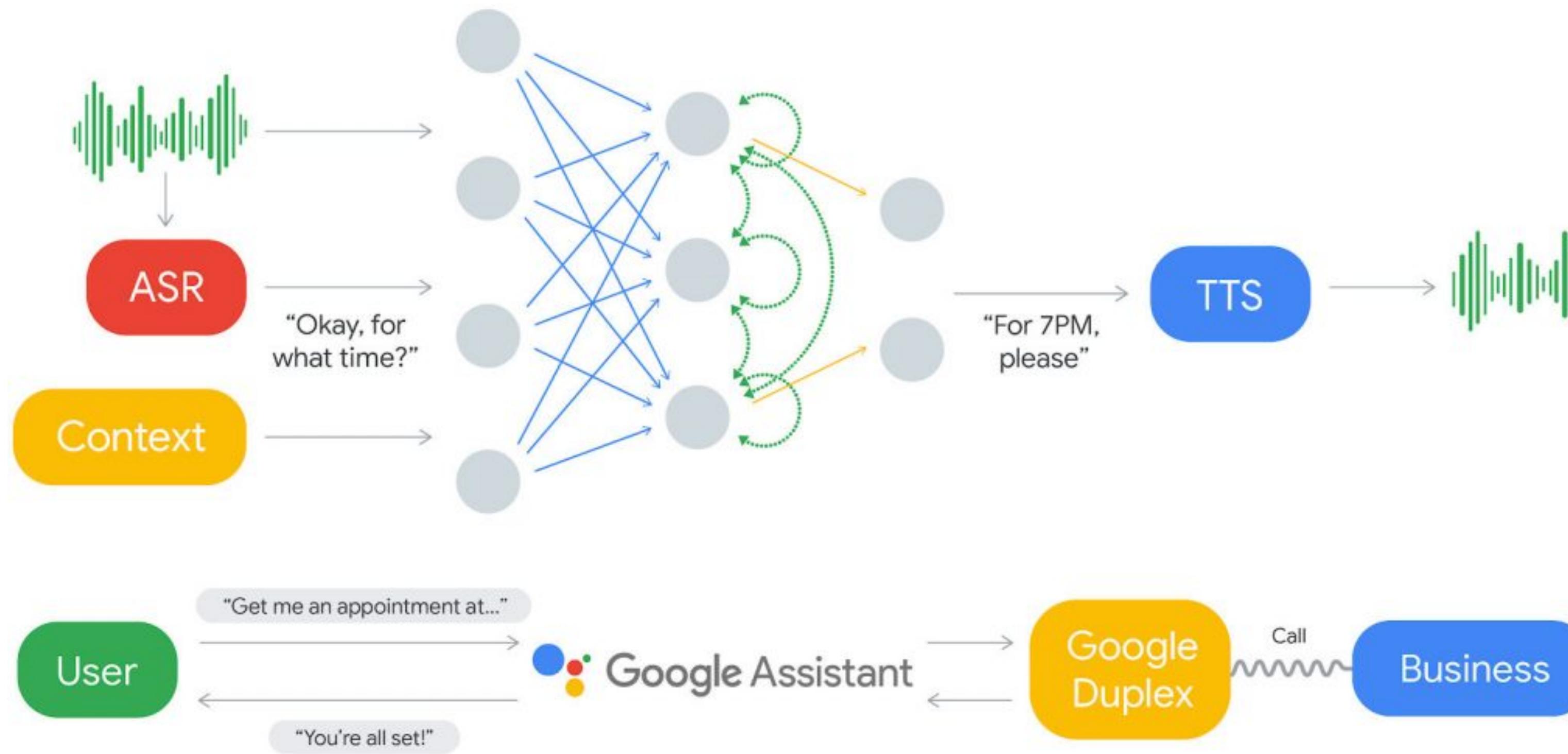
# What's a Dialogue System?

Dialog Systems are **HOT** 🔥. — Killer apps for NLP.



**Google Duplex:** Can you distinguish human and AI?

# What's a Dialogue System?



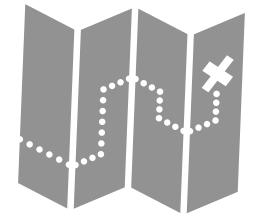
(<https://techeology.com/what-is-google-duplex/>)

**Google Duplex:** Can you distinguish human and AI?

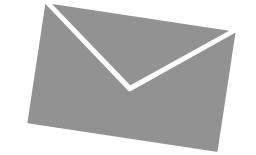
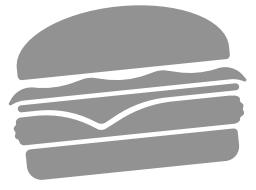
# What's a Dialogue System?

Dialog Systems are **HOT** 🔥. — Killer apps for NLP.

They can

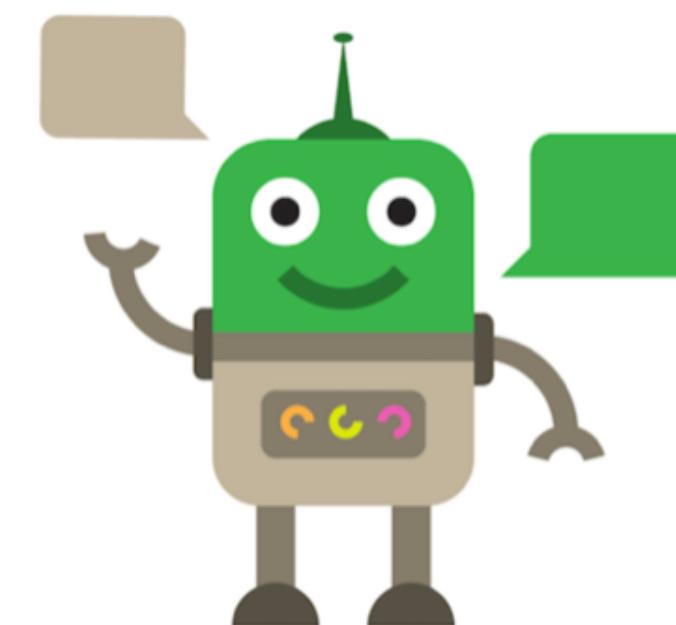


- give travel directions
- control home appliances
- find restaurants
- help make phone calls
- customer services
- ...



# Overview

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



Dialogue Systems

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

**A: travel agent**  
**C: human client**

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

Turn taking

Turn structure:  
(C-A-C-A-C...)

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

**Spoken DS:**  
**endpoint detection**  
(know when to start talking)

**Turn structure:**  
**(C-A-C-A-C...)**

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

#: overlap



**Each turn in a dialogue is a kind of action**  
Wittgenstein (1953) and Austin (1962)

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

asking

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

answering

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

asking

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

answering

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

answering

# Properties of Human Conversation

<b>Constatives:</b>	committing the speaker to something's being the case ( <i>answering, claiming, confirming, denying, disagreeing, stating</i> )
<b>Directives:</b>	attempts by the speaker to get the addressee to do something ( <i>advising, asking, forbidding, inviting, ordering, requesting</i> )
<b>Commissives:</b>	committing the speaker to some future course of action ( <i>promising, planning, vowing, betting, opposing</i> )
<b>Acknowledgments:</b>	express the speaker's attitude regarding the hearer with respect to some social action ( <i>apologizing, greeting, thanking, accepting an acknowledgment</i> )

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

**“I need to travel in May”**  
**“Book me a flight to Seattle”**  
**“I will book you a flight”**  
**“Thanks”**

**There are different taxonomies of dialog acts (also known as speech acts)**

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

## Principle of closure.

Agents performing an action require evidence, sufficient for current purposes, that they have succeeded in performing it  
Clark (1996) (after Norman 1988)

## grounding:

acknowledge that the listener has understood the speaker

Need to know if action has failed or succeeded  
(did the person understand me?)

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

**local structure  
between  
dialog acts**

**correction  
subdialogue**

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

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

# Properties of Human Conversation

- C<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#  
#Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: 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<sub>1</sub>: ...I need to travel in May.
- A<sub>1</sub>: And, what day in May did you want to travel?
- C<sub>2</sub>: OK uh I need to be there for a meeting that's from the 12th to the 15th.
- A<sub>2</sub>: And you're flying into what city?
- C<sub>3</sub>: Seattle.
- A<sub>3</sub>: And what time would you like to leave Pittsburgh?
- C<sub>4</sub>: Uh hmm I don't think there's many options for non-stop.
- A<sub>4</sub>: Right. There's three non-stops today.
- C<sub>5</sub>: What are they?
- A<sub>5</sub>: 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<sub>6</sub>: OK I'll take the 5ish flight on the night before on the 11th.
- A<sub>6</sub>: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
- C<sub>7</sub>: OK.
- A<sub>7</sub>: And you said returning on May 15th?
- C<sub>8</sub>: Uh, yeah, at the end of the day.
- A<sub>8</sub>: OK. There's #two non-stops ...#
- C<sub>9</sub>: #Act... actually #, what day of the week is the 15th?
- A<sub>9</sub>: It's a Friday.
- C<sub>10</sub>: Uh hmm. I would consider staying there an extra day til Sunday.
- A<sub>10</sub>: OK...OK. On Sunday I have ...

conversational  
implicature

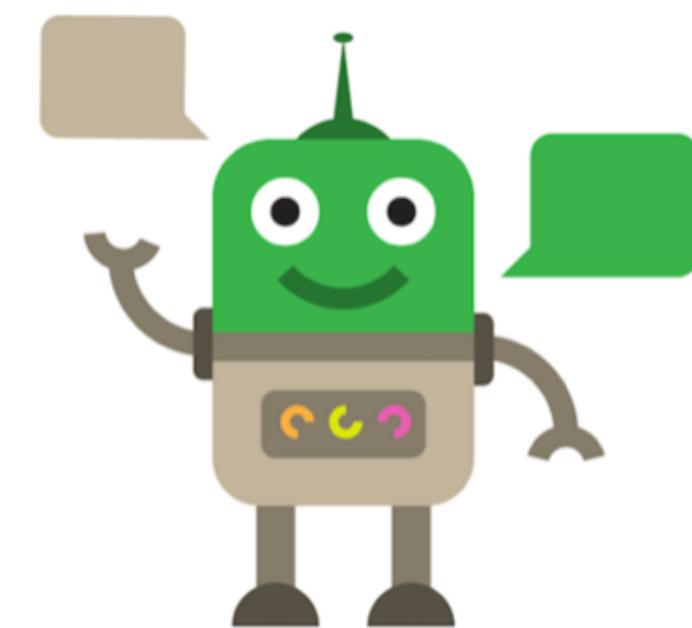
# Properties of Human Conversation

- 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**.

# Overview

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



Dialogue Systems

# Chatbot (Chit-chat)

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**

## Chatbots:

A conversation between a user and Cleverbot,  
an IR-based chatbot (Carpenter, 2017).

# Chatbot - Eliza: Weizenbaum (1966)

Men are all alike.  
IN WHAT WAY

They're always bugging us about something  
or other.  
CAN YOU THINK OF A SPECIFIC  
EXAMPLE

Well, my boyfriend made me come here.  
YOUR BOYFRIEND MADE YOU COME  
HERE

He says I'm depressed much of the time.  
I AM SORRY TO HEAR YOU ARE  
DEPRESSED

Rogerian psychologist: reflect  
back what the patient said

WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU?

My father

You are like my father in some ways.

WHAT RESEMBLANCE DO YOU SEE

You are not very aggressive but I think you don't want me to notice that.

WHAT MAKES YOU THINK I AM NOT AGGRESSIVE

You don't argue with me.

WHY DO YOU THINK I DON'T ARGUE WITH YOU

You are afraid of me.

DOES IT PLEASE YOU TO BELIEVE I'M AFRAID OF YOU

My father is afraid of everybody.

WHAT ELSE COMES TO MIND WHEN YOU THINK OF YOUR FATHER

Bullies.

# Chatbot - Eliza: Weizenbaum (1966)

```
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  $\leftarrow$  Apply the transform in r to sentence
        if w = 'my'
            future  $\leftarrow$  Apply a transformation from the 'memory' rule list to sentence
            Push future onto memory stack
        else (no keyword applies)
            either
                response  $\leftarrow$  Apply the transform for the NONE keyword to sentence
            or
                response  $\leftarrow$  Pop the top response from the memory stack
    return(response)
```

**Rule-based method:** A simplified sketch of the ELIZA algorithm. The power of the algorithm comes from the **particular transforms associated with each keyword**.

# Chatbot - Eliza rules

Keywords are ranked from specific to general

I know everybody laughed at me

- “I” is a very general keyword:

I: (I \*) -> (You say you 2)

YOU SAY YOU KNOW EVERYBODY LAUGHED AT YOU

- “Everybody” is much more interesting (someone using universals like everybody/always is probably “referring to some quite specific event or person”)

WHO IN PARTICULAR ARE YOU THINKING OF?

- Implementation: keywords stored with their rank

**Everybody** 5 (*transformation rules*)

**I** 0 (*transformation rules*)

Backoff

Please go on

That's very interesting

I see

# Chatbot - Parry: Colby (1971)

Similar **Pattern based rules** as Eliza, more advanced **control structure**

**Persona:** 28-year-old single man (post office clerk), no siblings and lives alone,

- Sensitive about his physical appearance, his family, his religion, his education, and the topic of sex.
- Hobbies (movies and gambling)
- Backstory (recently attacked a bookie, claiming the bookie did not pay off in a bet, afterwards worried about possible underworld retaliation)
- Eager to tell his story to non-threatening listeners

**Mental model:** Anger, Fear, Mistrust

**Flare concepts:** horses -> horse racing -> gambling -> bookies -> underworld -> Mafia

First system to pass the Turing test!

Psychiatrists couldn't distinguish interviews with PARRY from (text transcripts of) interviews with real paranoids

# Chatbot - Eliza meet Parry

Good Evening. Tell me your problems.

Please terminate input with a period or a question mark.

\*People get on my nerves sometimes.

I am not sure I understand you fully.

\*You should pay more attention.

Suppose you should pay more attention.

\*You're entitled to your own opinion.

What makes you think I am entitled to my own opinion?

\*It bothers me just to be around people in general.

Please go on.

\*I went to the track at Bay Meadows a while back.

What does that suggest to you?

\*I went to the races.

What else would you like to discuss?

\*Horseracing interests me.

Why do you say that just now?

\*There's a lot of crooked gambling going on there.

I am not sure I understand you fully,

\*It's clear to me if not to you.

Do you think its likely that not to I?

\*I went to the races.

Please go on.

\*I gamble on horses.

What does that suggest to you?

\*I went to the races.

What else would you like to discuss?

\*Gambling has been nothing but trouble to me.

Why do you say that just now?

\*I've had experience with crooks gambling at the track.

I am not sure I understand you fully.

\*I've already told you enough about that.

# Chatbot

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.

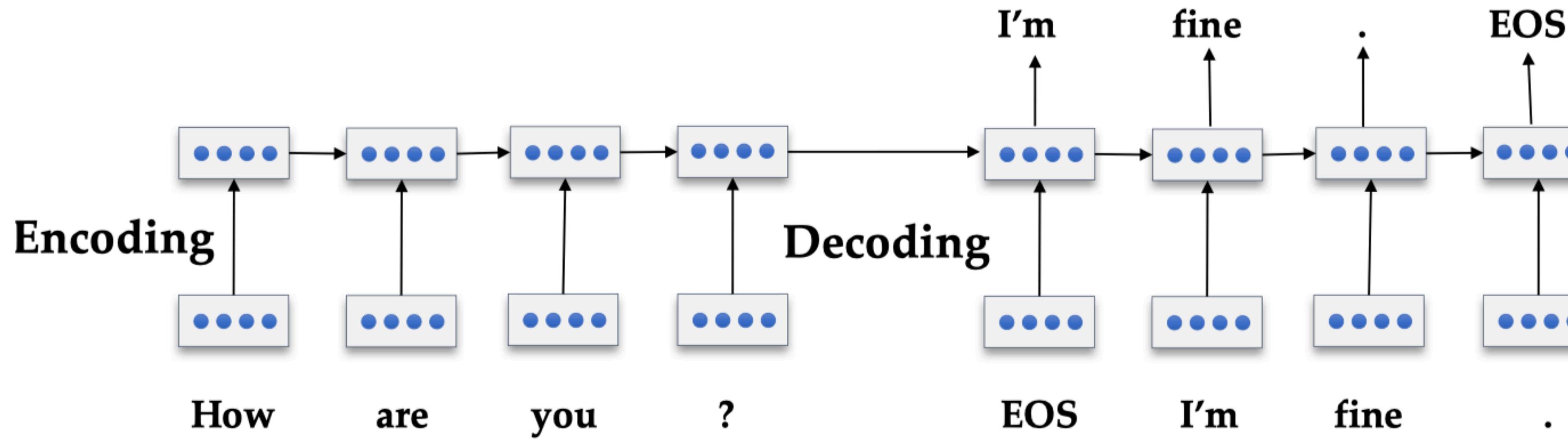
$$r = \text{response} \left( \operatorname{argmax}_{t \in C} \frac{q^T t}{\|q\| \|t\|} \right)$$

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

# CleverBot

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.

# Chatbot



**Corpus-based method (Seq2Seq):**  
An encoder decoder model for neural response  
generation in dialogue.

# Chatbot

<b>Input</b>	tell me ... how long have you had this falling sickness ?
<b>Vanilla-SEQ2SEQ</b>	i'm not a doctor.
<b>Adversarial</b>	a few months, i guess .
<b>Input</b>	so i had the doctors test sammy 's response to conditioning .
<b>Vanilla-SEQ2SEQ</b>	sammy wrote the test sammy wrote the test .
<b>Adversarial</b>	so he took the pills .
<b>Input</b>	they didn 't have much success with this problem commander .
<b>Vanilla-SEQ2SEQ</b>	they 're not the only ones who have been in the system .
<b>Adversarial</b>	can we find someone else ?

## Corpus-based method (Seq2Seq):

Sample responses generated by a Seq2Seq model trained either with a maximum likelihood objective, or adversarially trained to produce sentences that are hard for an adversary to distinguish from human sentences (Li et al., 2017).

# Chatbot: Seq2Seq models

## Repetitive

- A: Where are you going?  
B: I'm going to the restroom.  
A: See you later.  
B: See you later.  
A: See you later.  
B: See you later.

## Want Diversity



## Maybe beam search

Beam size	Model response
1	<i>I love to eat healthy and eat healthy</i>
2	<i>That is a good thing to have</i>
3	<i>I am a nurse so I do not eat raw food</i>
4	<i>I am a nurse so I am a nurse</i>
5	<i>Do you have any hobbies?</i>
6	<i>What do you do for a living?</i>
7	<i>What do you do for a living?</i>
8	<i>What do you do for a living?</i>

**Low beam size:**  
More on-topic but nonsensical; bad English

**High beam size:**  
Converges to safe, “correct” response, but it’s generic and less relevant

(figure credit: Stanford CS224N, Chris Manning)

# Chatbot: Seq2Seq models

## Repetitive

A: Where are you going?

B: I'm going to the restroom.

A: See you later.

B: See you later.

A: See you later.

B: See you later.

## Sample and Rank

1. Sample N candidate
2. Rank candidate and select best one

## Sampling

Randomly sample words from distribution at each time step  $t$

- Basic/pure sampling: sample from  $P_t(w)$  directly
  - Can get some very bad samples
  - No control
- Top- $n$  sampling: sample from  $P_t$  truncated to top  $n$  words
  - Greedy search:  $n = 1$ , Pure sampling:  $n = |V|$
  - Increase  $n$  to get more diverse/risky output
  - Decrease  $n$  to get more generic/safe output
- Top- $p$  sampling: sample from  $P_t$  restricted to top  $p$  proportion of words
  - Better when probability distribution is spread
- Temperature based sampling:
  - Increase  $\tau$  to get more diverse/risky output ( $P_t$  is more uniform)
  - Decrease  $\tau$  to get more generic/safe output ( $P_t$  is more spiky)

$$P_t(w) = \frac{\exp(s_w/\tau)}{\sum_{w' \in V} \exp(s_{w'}/\tau)}$$

# Chatbot

## Conversation D

---

- Human:** *I like how cute cows are*
- Meena:** *Ah, interesting. You like how they are cute, right?*
- Human:** *yeah, and they're realy smart*
- Meena:** *I heard they go to college*
- Human:** *cows go to college?*
- Meena:** *I heard that a cow went to Harvard.*
- Human:** *what did the cow study?*
- Meena:** *Bovine sciences.*
- Human:** *do horses go to Harvard?*
- Meena:** *Horses go to Hayvard.*
- Human:** *that's a pretty good joke, I feel like you led me into it*
- Meena:** *You were trying to steer it elsewhere, I can see it.*
- 

**Meena (Google):**  
Evolved Transformer (transformer-like architecture found via architecture search)

- Trained on social media conversations
- Minimize perplexity of next token
- Uses sample and rank

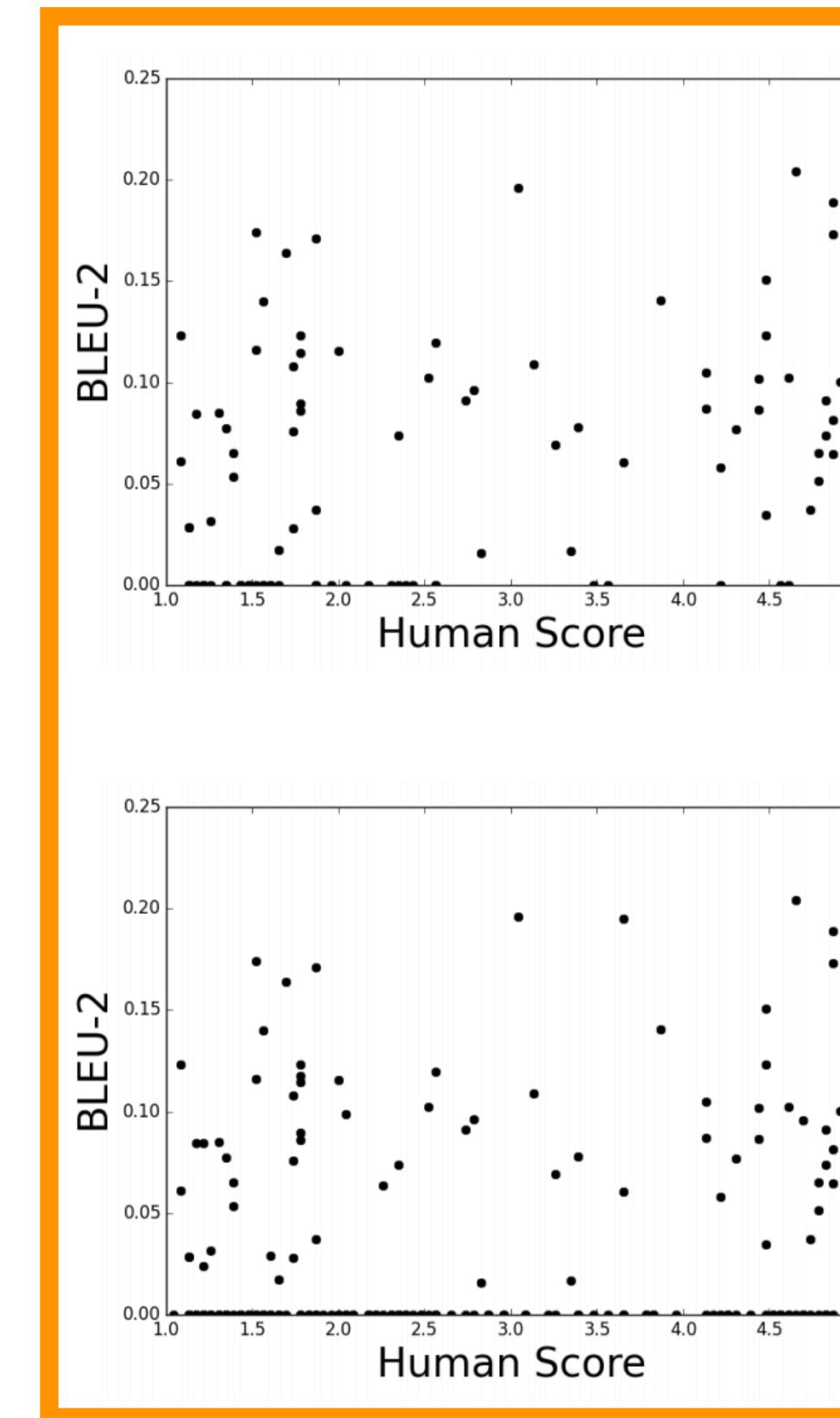
# Chatbot

- Goal:
  - mimicking the **unstructured conversational characteristic of human-human interaction**
- Methods:
  - Rule-based
  - Corpus-based (IR, Seq2Seq)
- Evaluation:
  - Chatbots are generally evaluated by humans
  - Adversarial evaluation: train a “Turing-like” evaluator classifier to distinguish between human-generated responses and machine-generated responses.

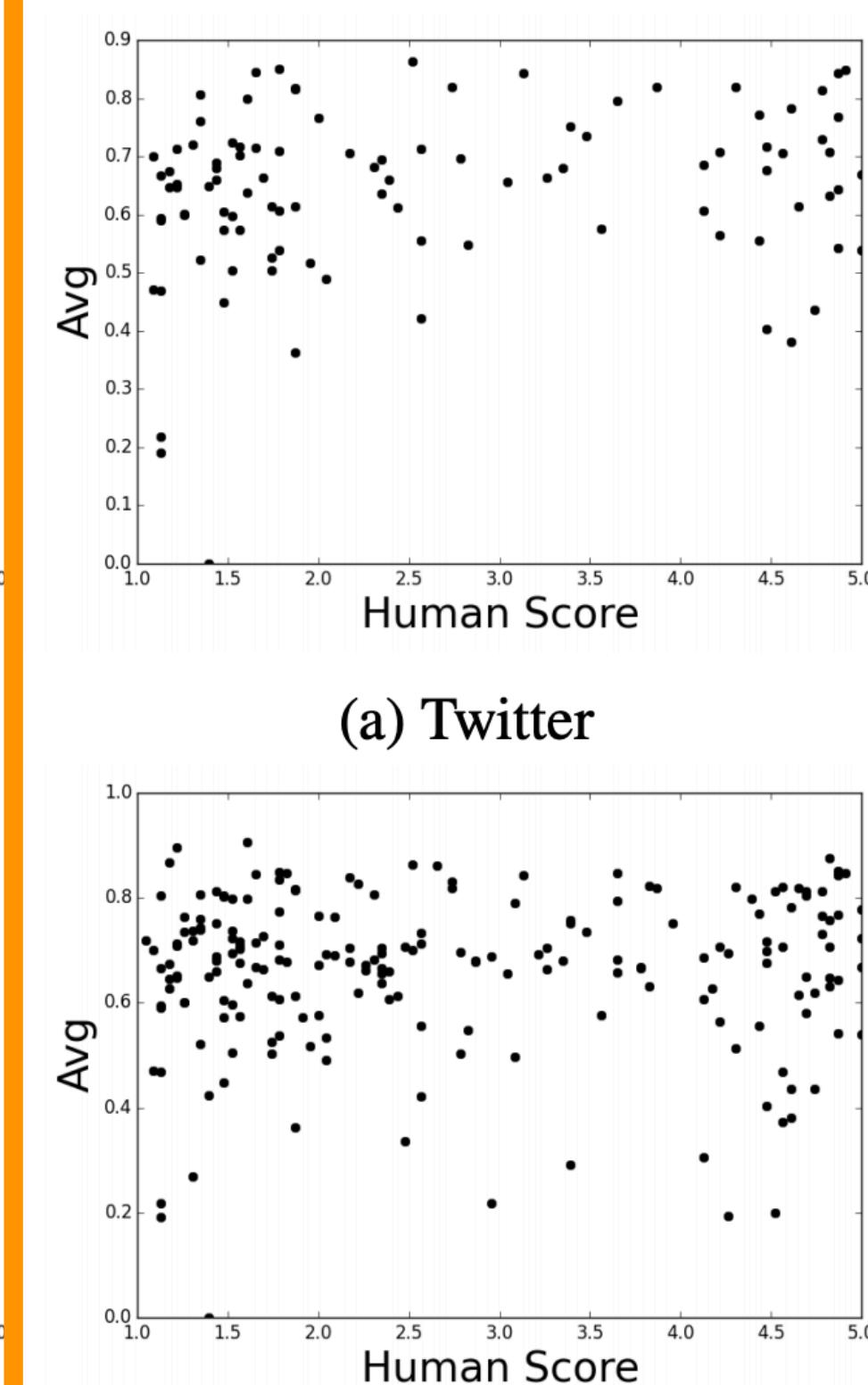
# Chatbot Evaluation

**Automatic Evaluation:**  
Word overlap metrics  
are bad for dialogue

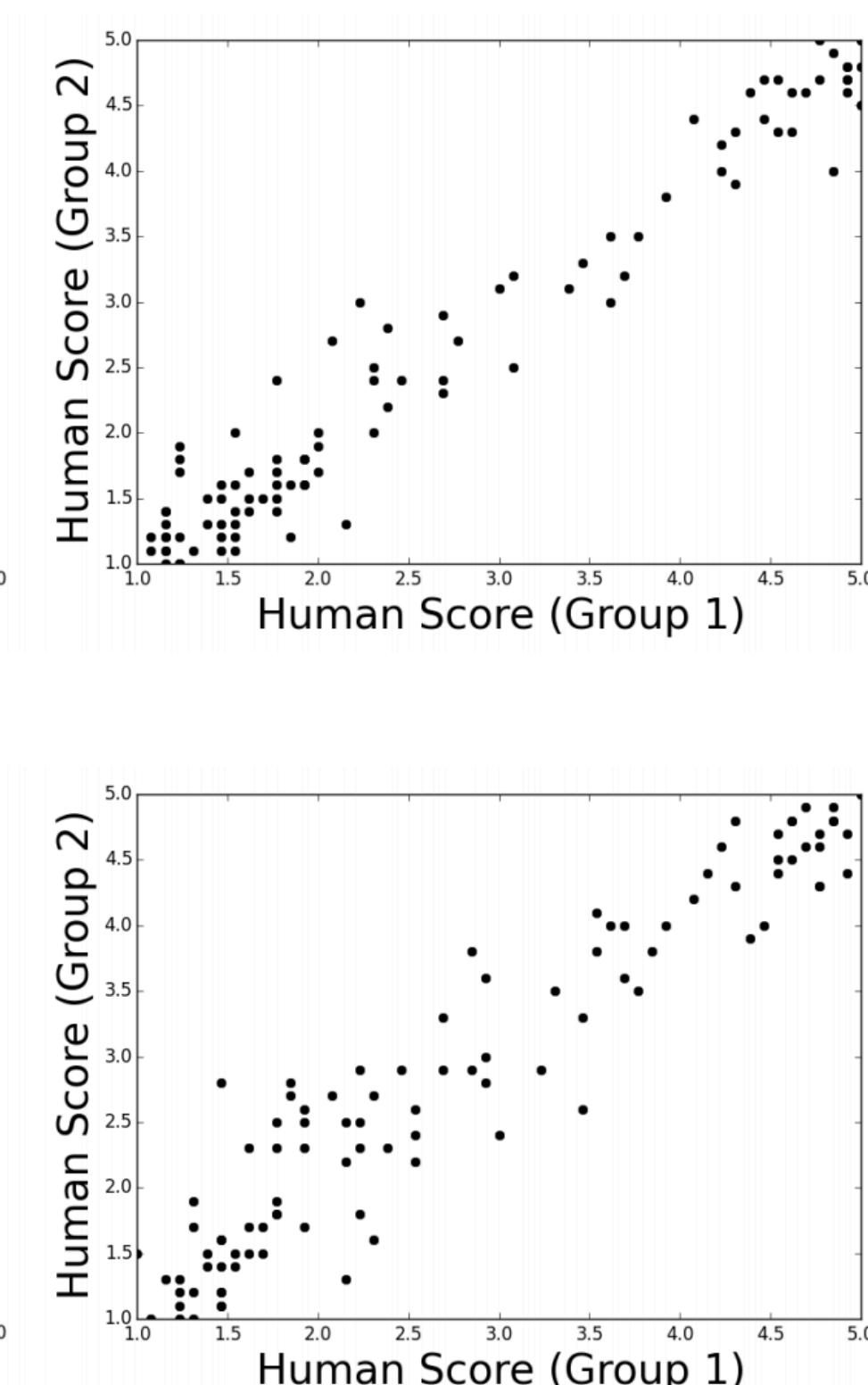
**No correlation  
between human  
judgement and BLEU**



**BLEU**



**Embedding Average**

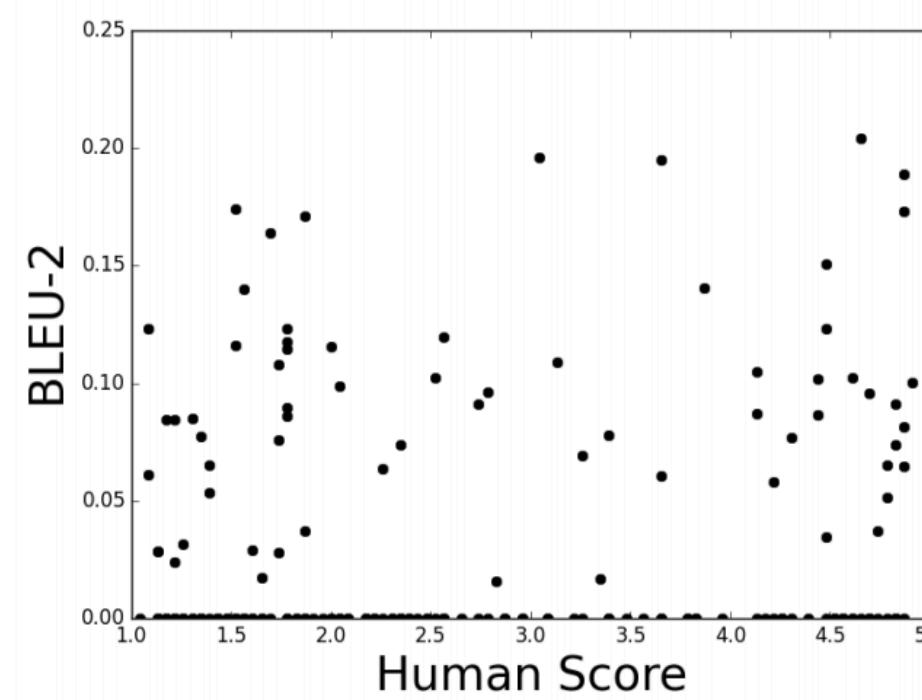
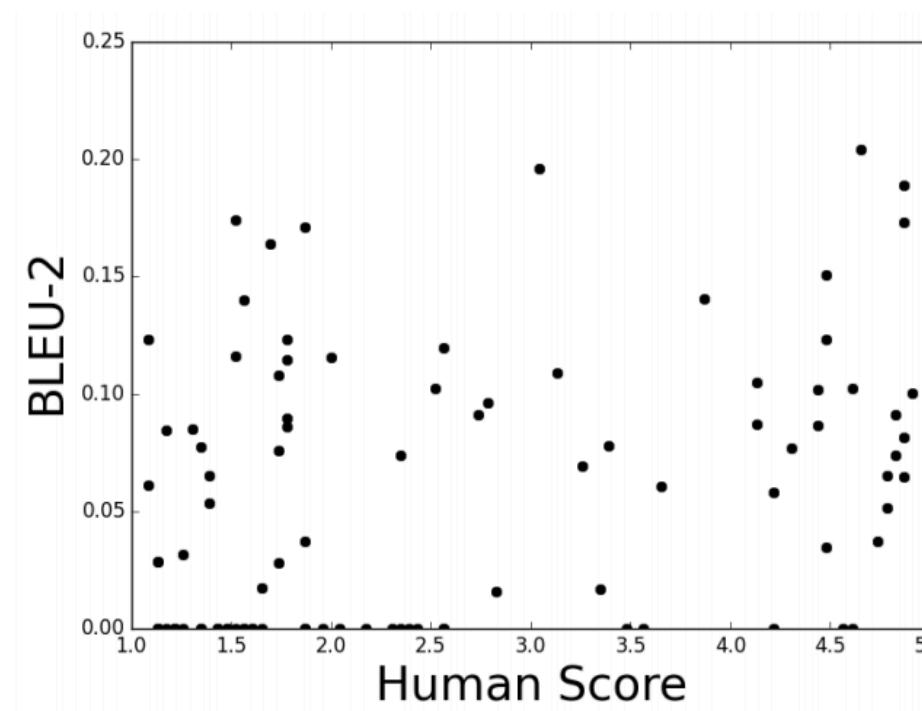


**Human**

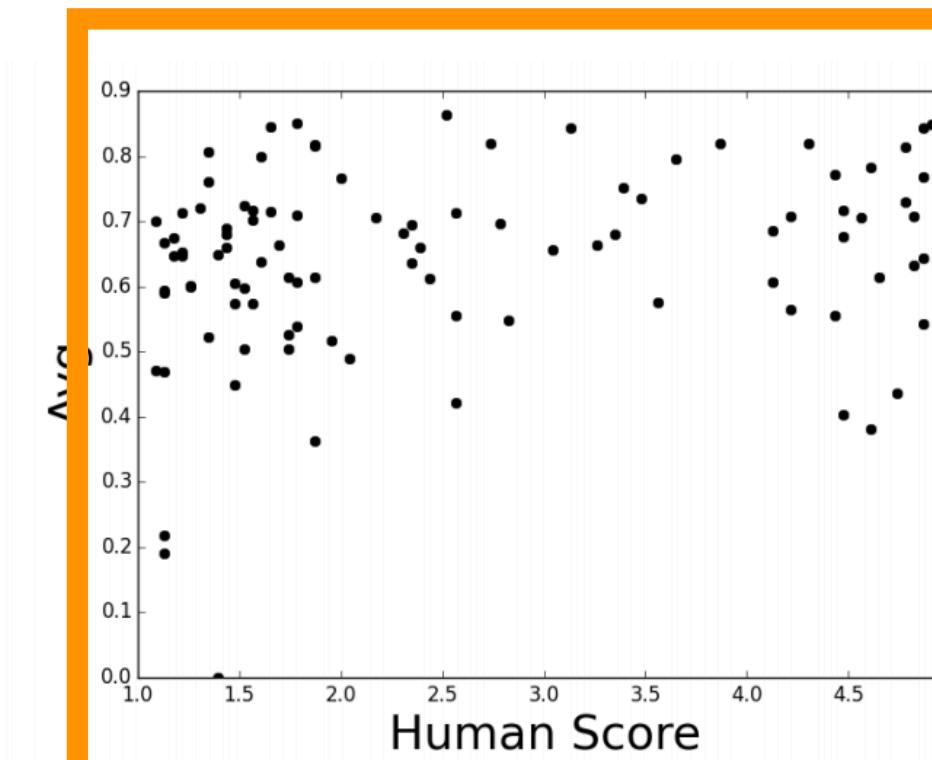
# Chatbot Evaluation

**Automatic Evaluation:**  
Word overlap metrics  
are bad for dialogue

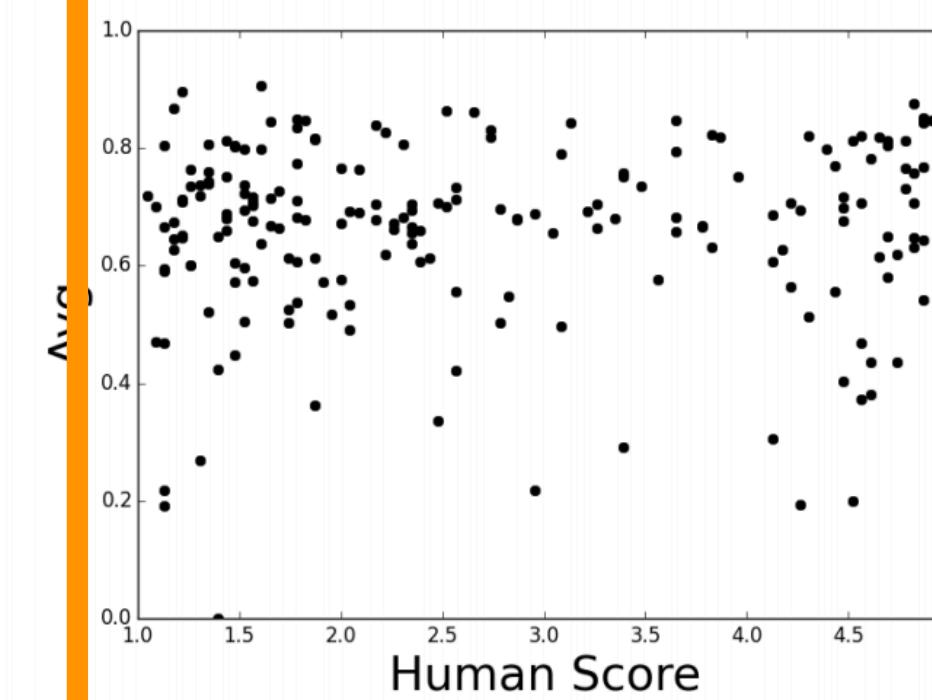
**No correlation  
between human  
judgement and  
embedding average**



**BLEU**

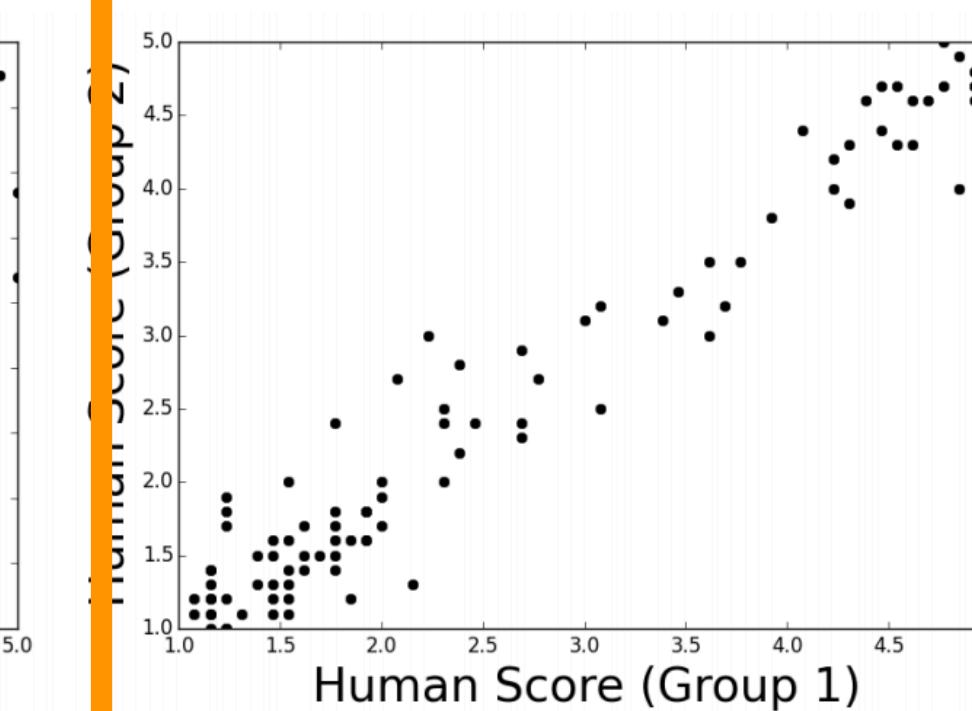


(a) Twitter



(b) Ubuntu

**Embedding Average**

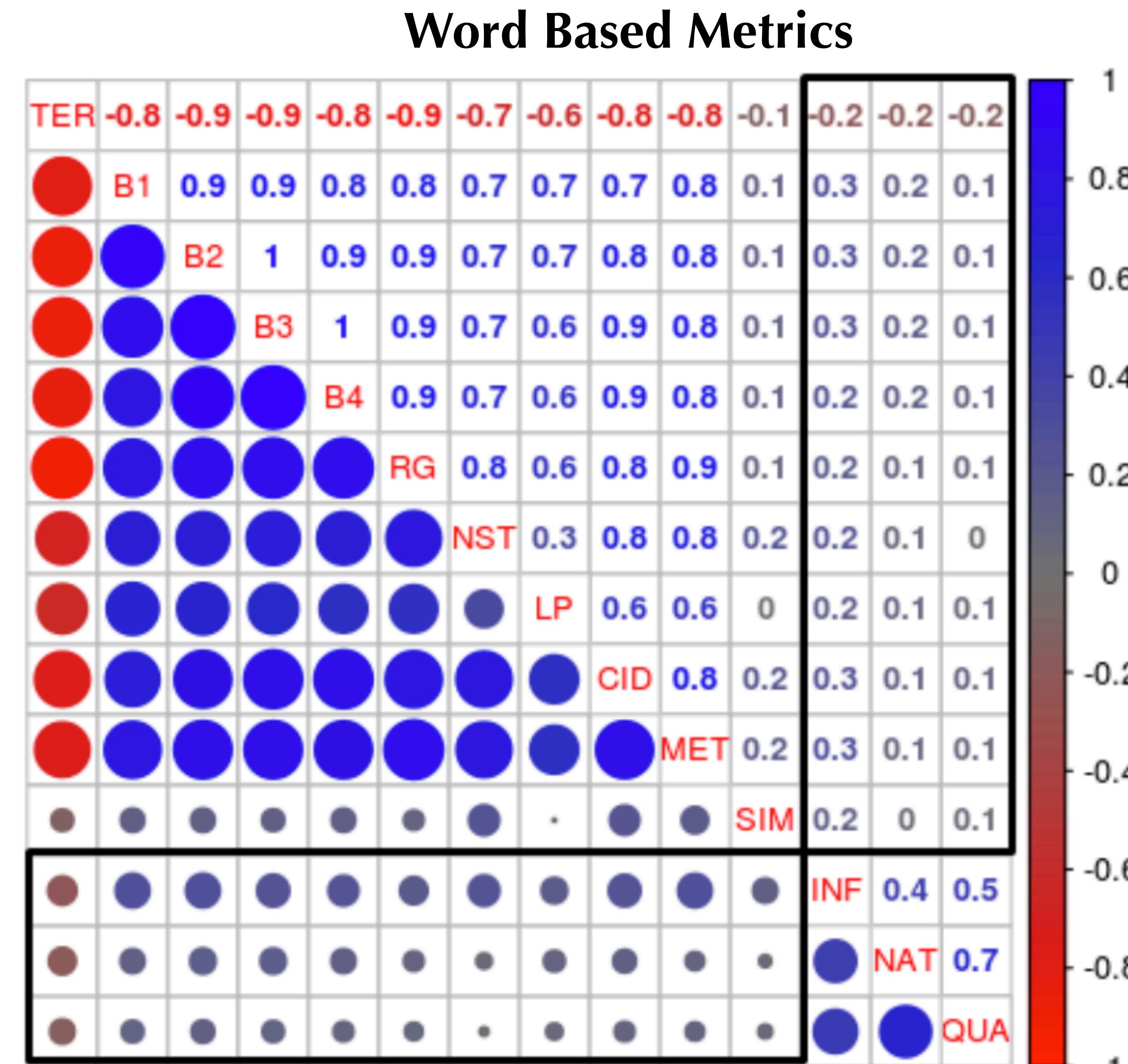


**Human**

# Chatbot Evaluation

**Word Overlap Metrics**

- highly correlated with each other
- Not so correlated with human ratings



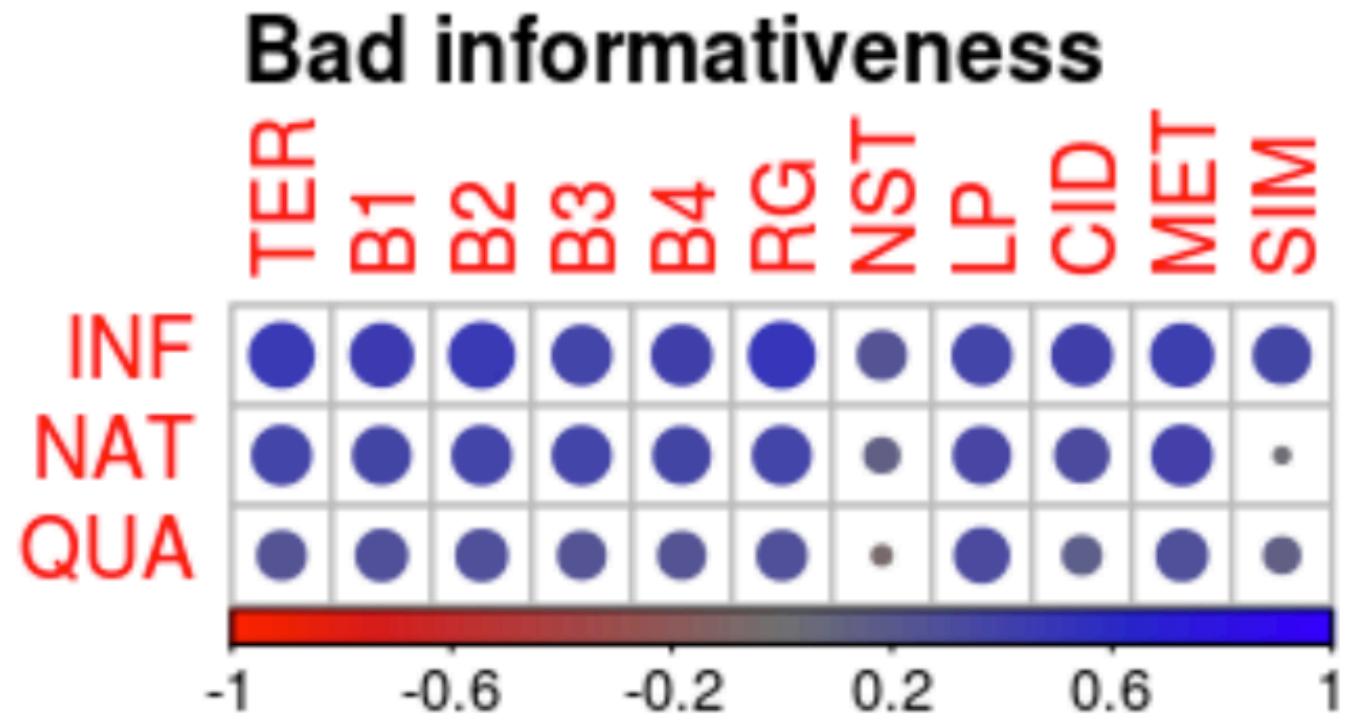
**Spearman correlations of word based metrics and human ratings**

**Human Ratings**

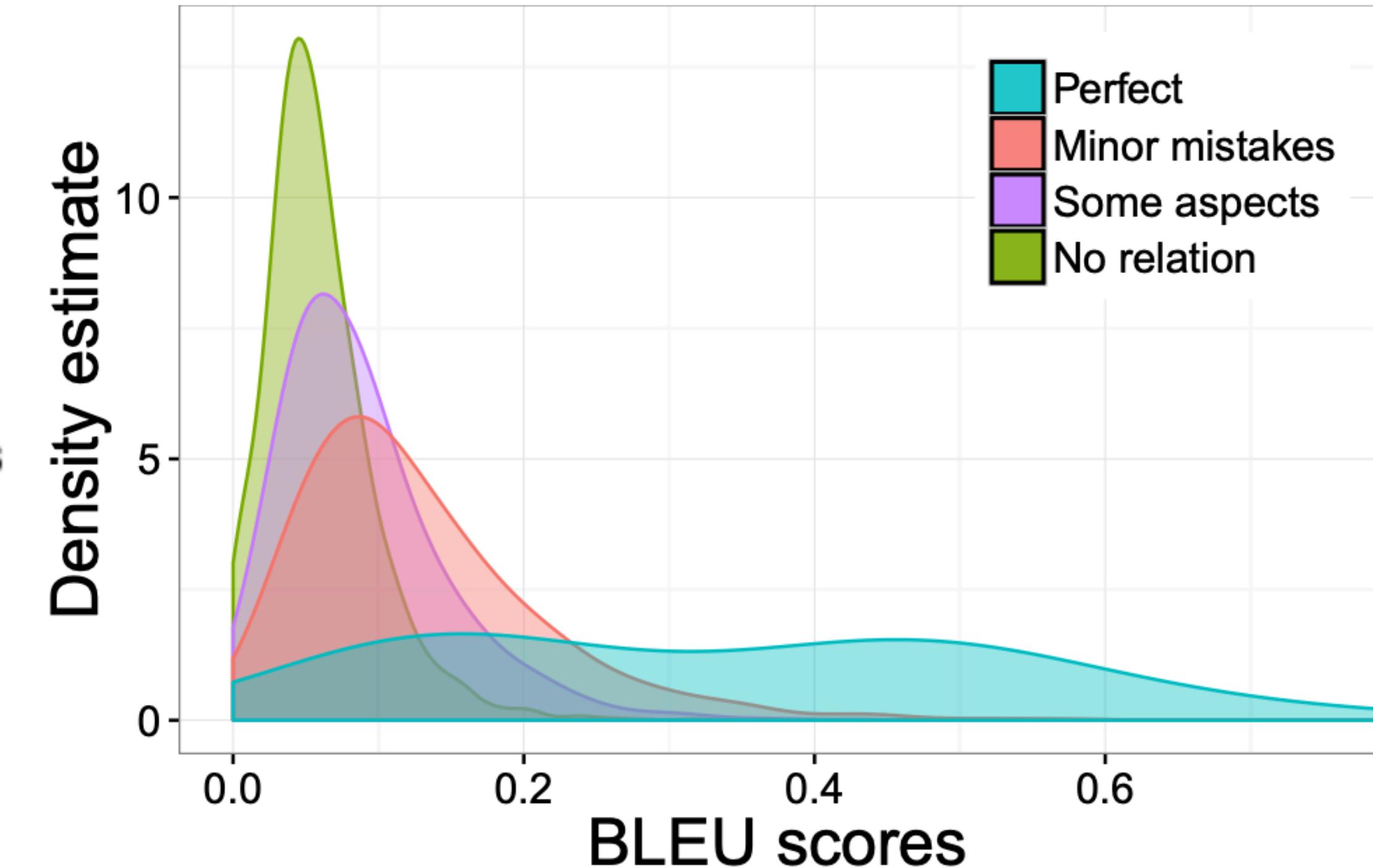
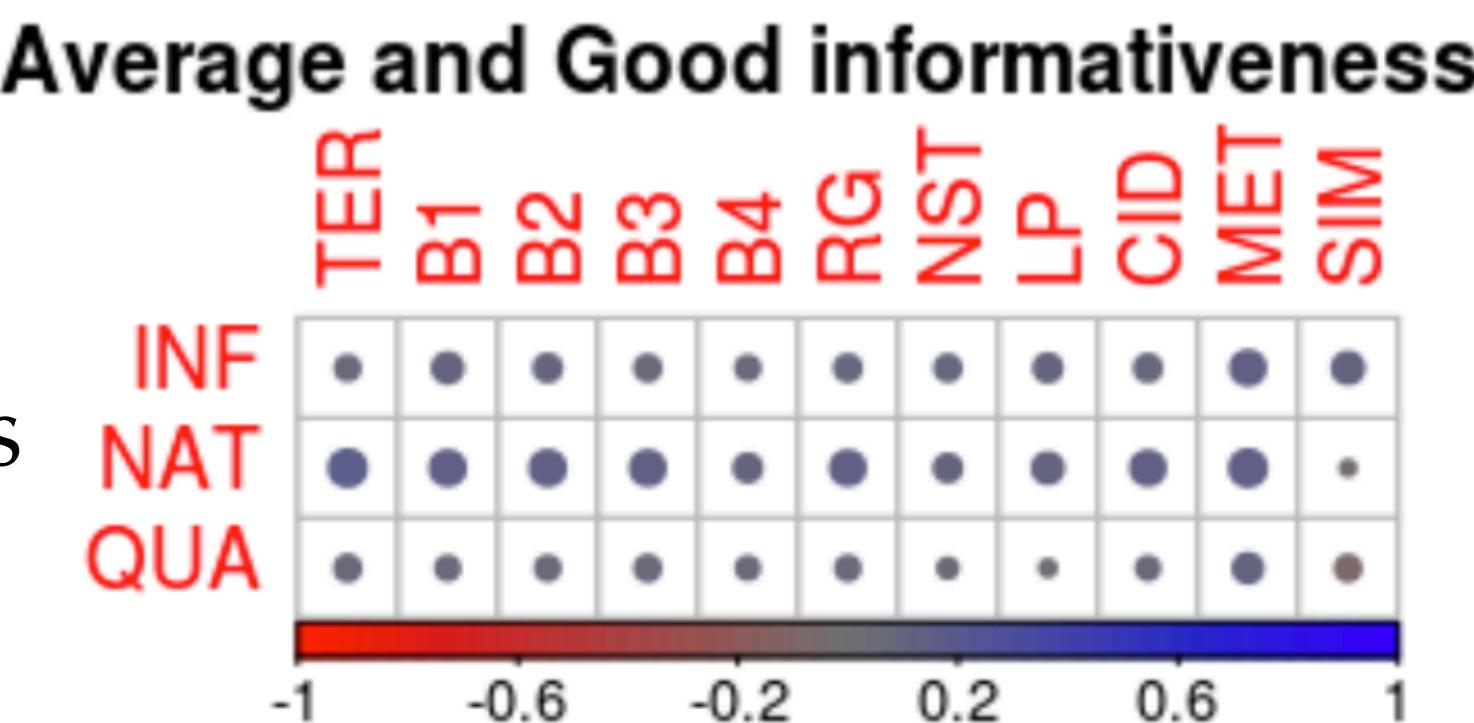
- Informativeness
- Naturalness
- Quality

# Chatbot Evaluation

High correlation with human judgement for low quality generations



Poor correlation with human judgement for mid to high quality generations



[Why We Need New Evaluation Metrics for NLG, Novikova et al 2017, <https://arxiv.org/pdf/1707.06875.pdf>]

Re-evaluating Automatic Metrics for Image Captioning  
[Kilickaya et al, EACL 2017]

# Chatbot Evaluation

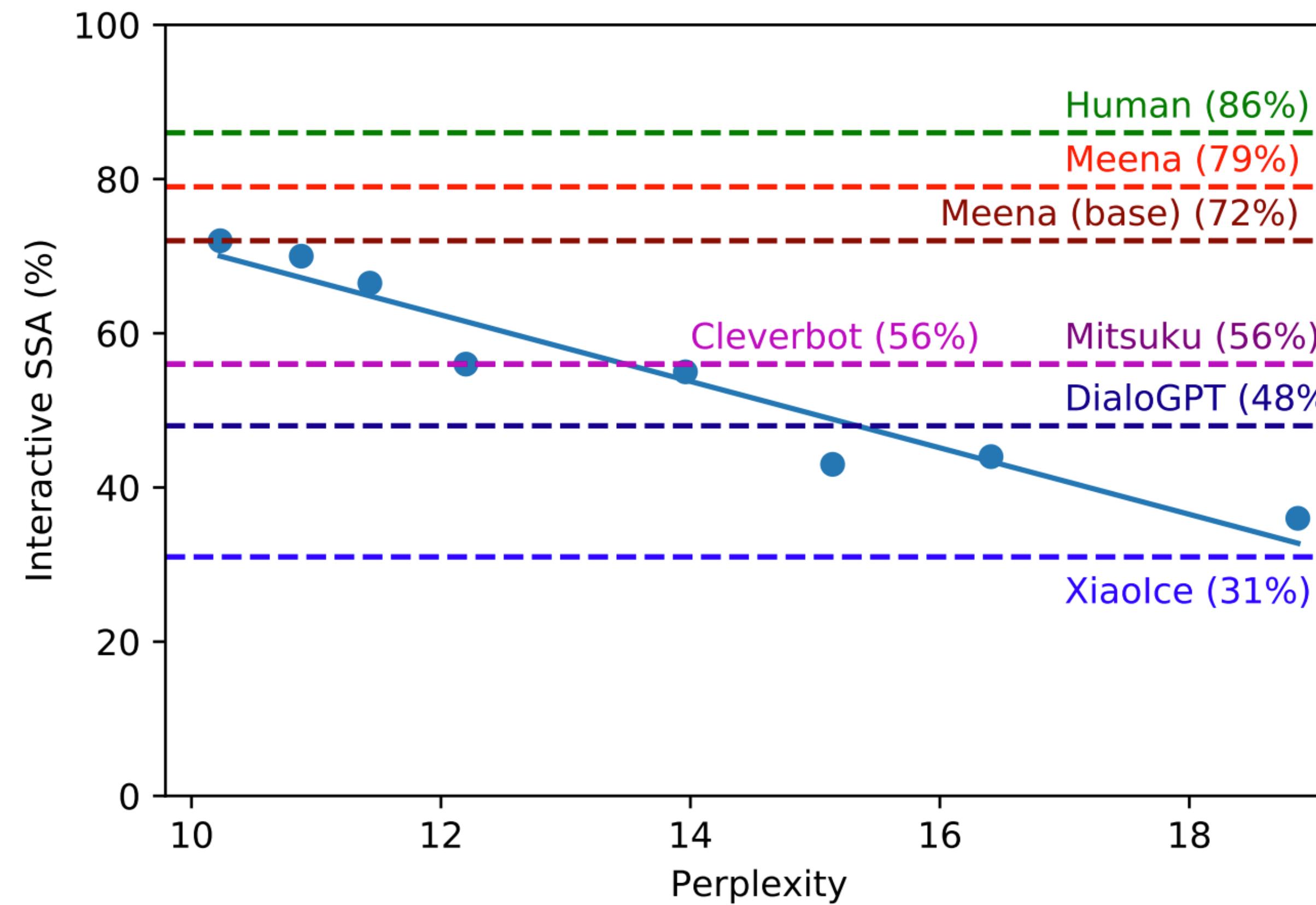
## **Human evaluation: gold standard**

- slow, expensive, not repeatable (subjective/inconsistent), difficult to form well-targeted questions that are not open to misinterpretation

Decompose evaluation into **meaningful components**  
(approximate some of these by automated metrics)

- Fluency (probability wrt well-trained LM)
- Correct Style (probability wrt well-trained LM on target corpus)
- Relevance to input (semantic similarity)
- Conciseness (length)
- Repetitiveness (repeating words)
- Diversity (rare word usage, uniqueness of n-grams)
- Task-specific metric

**Sensibleness and Specificity Average (SSA):**  
Human judgement of whether responses (given context): makes **sense** and are **specific**  
Observation: SSA is correlated with perplexity!



**Meena (Google) 2020**  
**Cleverbot (Carpenter 2017)**  
<http://www.cleverbot.com>  
**Mitsuku: Loebner Prize winner (2016-2019)**  
<https://www.pandorabots.com/mitsuku/>  
**DialoGPT (Microsoft) 2019**  
<https://github.com/microsoft/DialoGPT>  
**Microsoft Xiaoice**  
**Current chit-chat models:**  
**very fluent with no understanding**

# Task-Oriented (Goal-Based) 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

State of the art from  
1977!

## GOAL

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

Frame-based control architecture

Still used in various forms in modern systems

## Task-Oriented Dialog System (Travel):

A transcript of an actual dialog with the GUS system of *Bobrow et al.* (1977)  
P.S.A. and Air California were airlines of that period.

# **Task-Oriented Dialogue System**

How to incorporate task related knowledge?

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

# 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.

# Task-Oriented Dialogue System

How to incorporate task related knowledge?

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



contains one or more **frames**.

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

# Task-Oriented Dialogue System

How to incorporate task related knowledge?

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

a collection of **slots**

Slot1.....

Slot2.....

Slot3.....

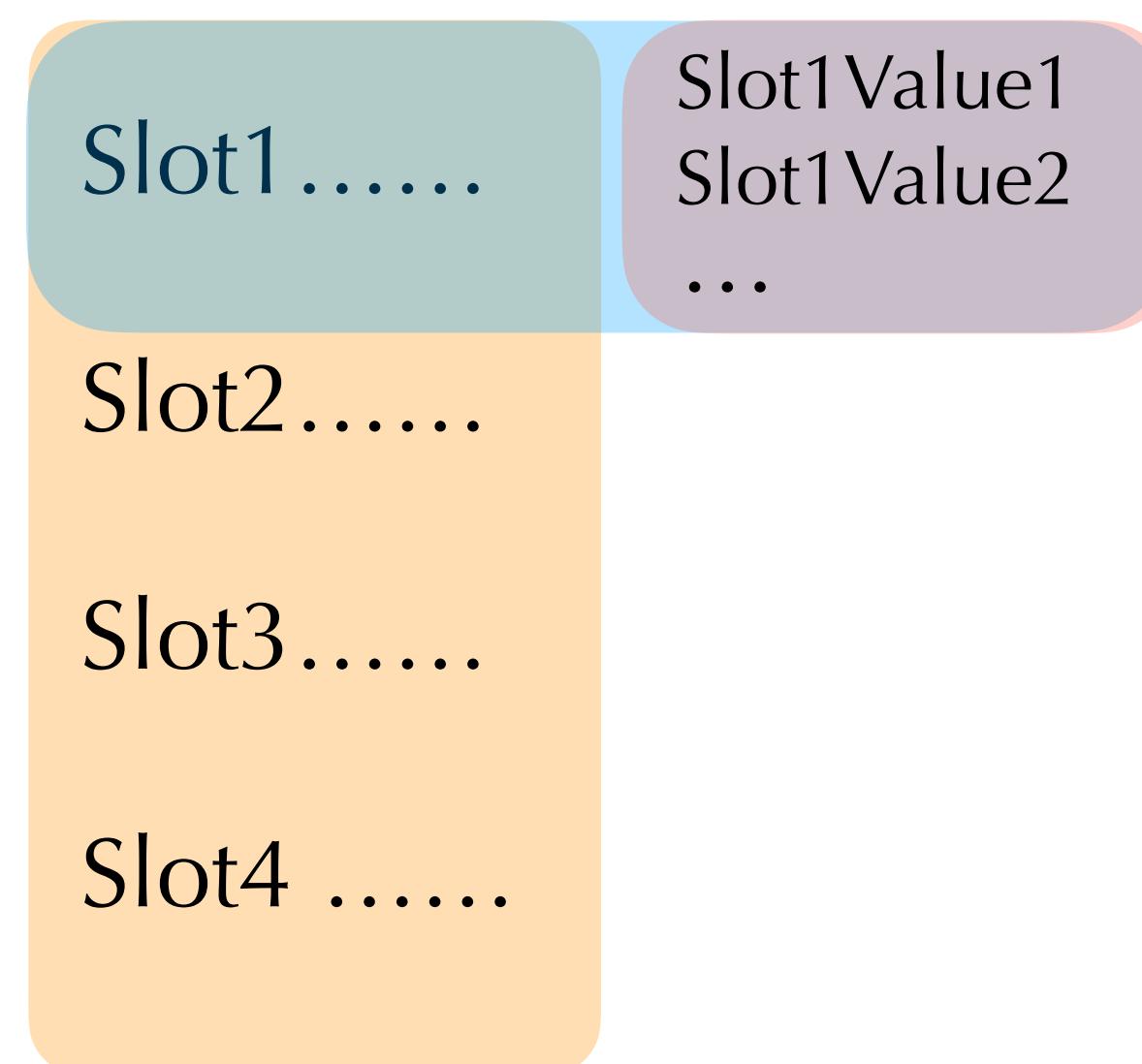
Slot4 .....

# Task-Oriented Dialogue System

How to incorporate task related knowledge?

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

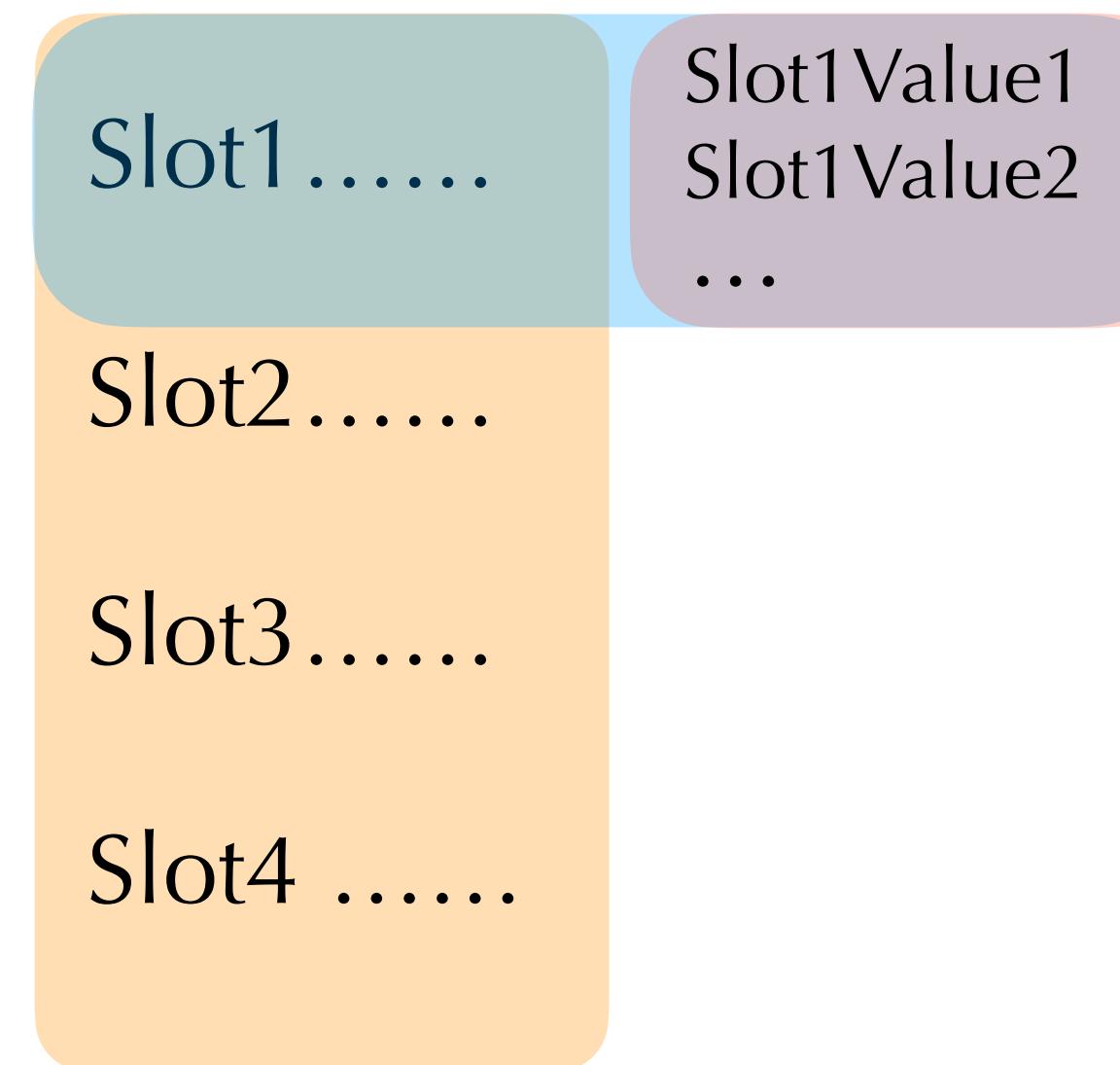
Also defines the **values** that each **slot** can take.



# Task-Oriented Dialogue System

How to incorporate task related knowledge?

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



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

- Try to fill these frames:  
- Extract from user utterances  
- Ask user for missing information

# Task-Oriented Dialogue System

How to incorporate task related knowledge?

“Show me morning flights from  
Boston to San Francisco on Tuesday”

# 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

**Classification**

# 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

**Classification**

# 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

**Sequence  
tagging**

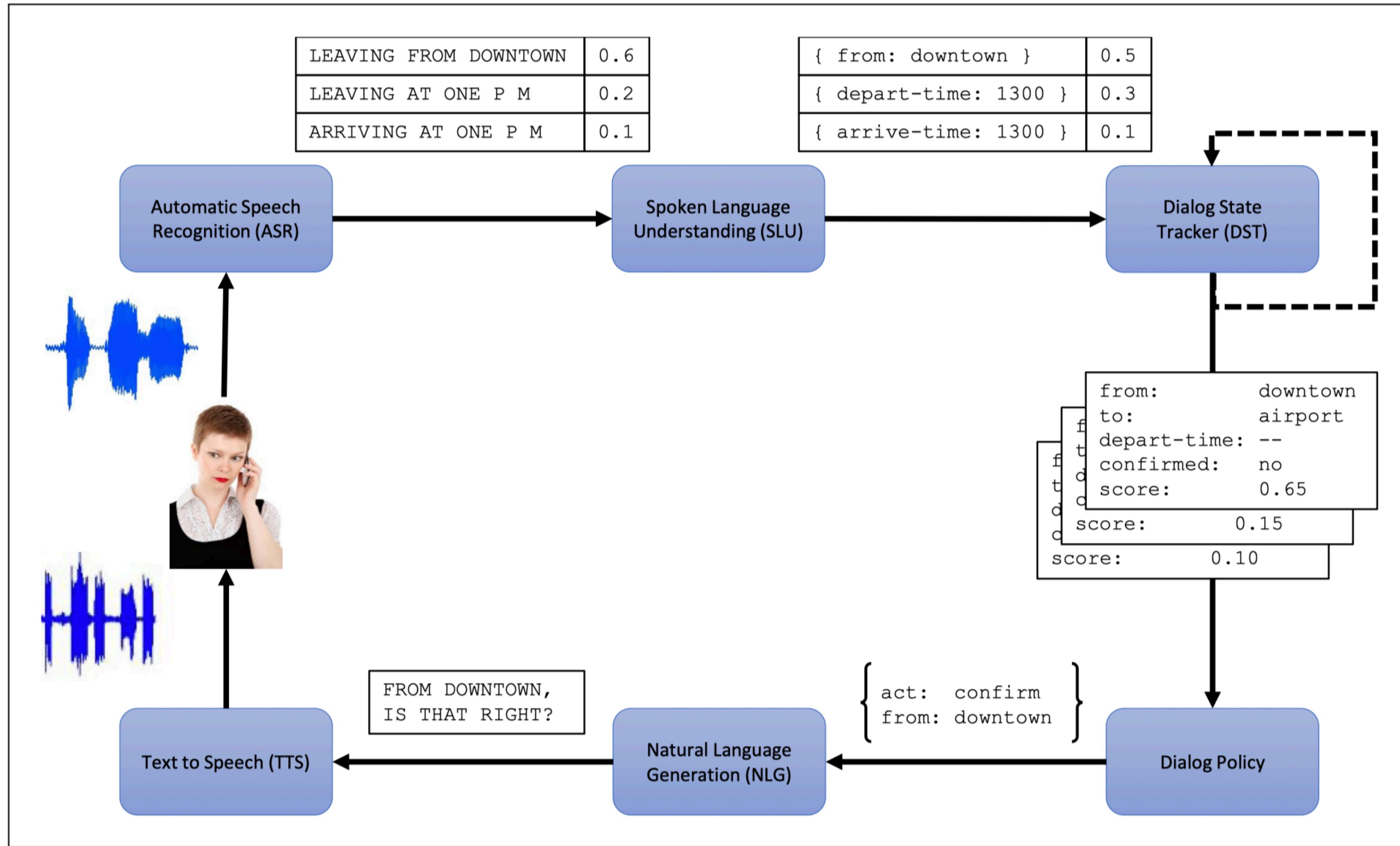
# Task-Oriented Dialogue System

How to incorporate task related knowledge?

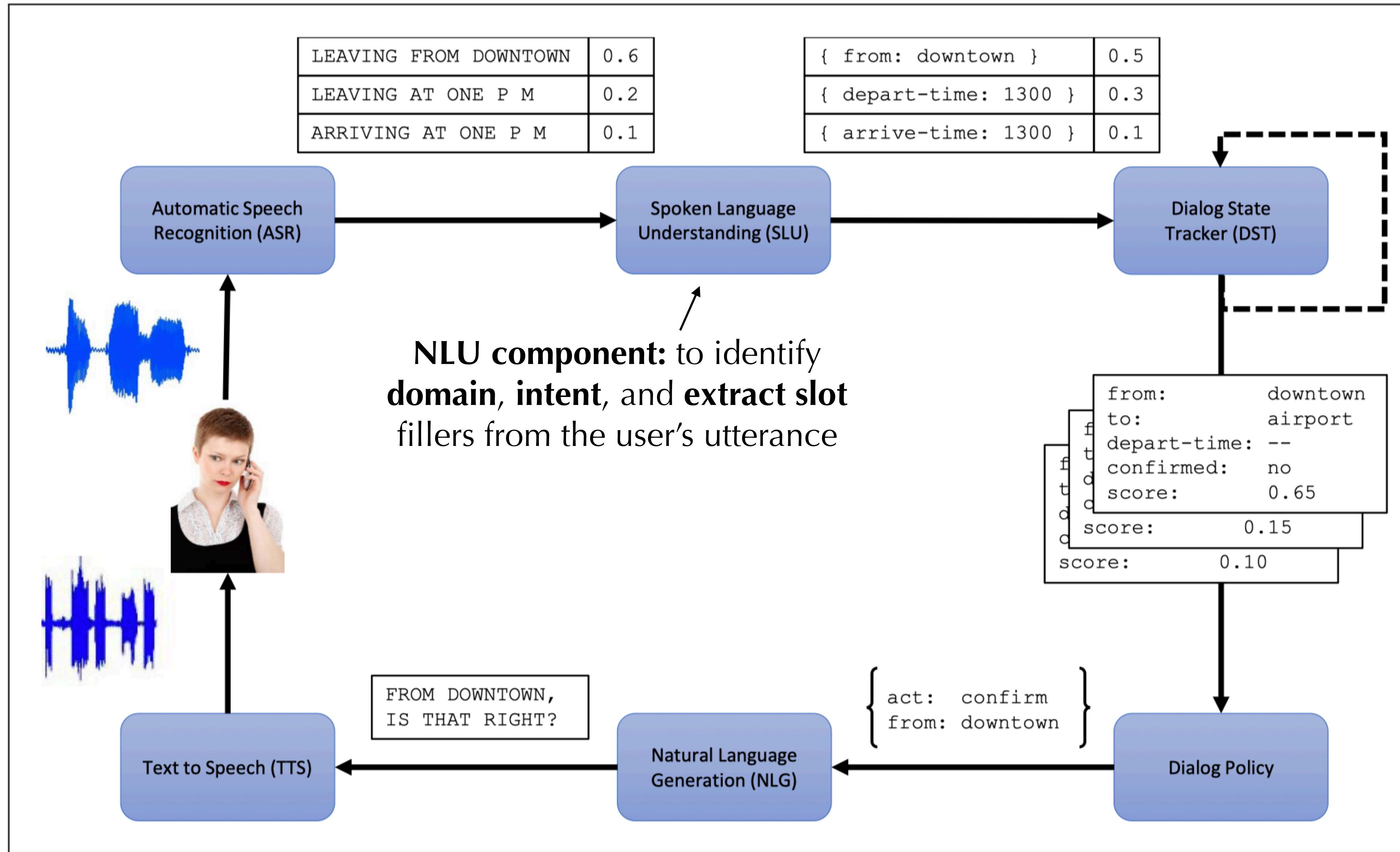
Utterance	Dialogue act
U: Hi, I am looking for somewhere to eat.	hello(task = find, type=restaurant)
S: You are looking for a restaurant. What type of food do you like?	confreq(type = restaurant, food)
U: I'd like an Italian somewhere near the museum.	inform(food = Italian, near=museum)
S: Roma is a 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()

A sample dialogue from the Hidden Information State (HIS) System  
of Young et al. (2010) using **dialog acts**

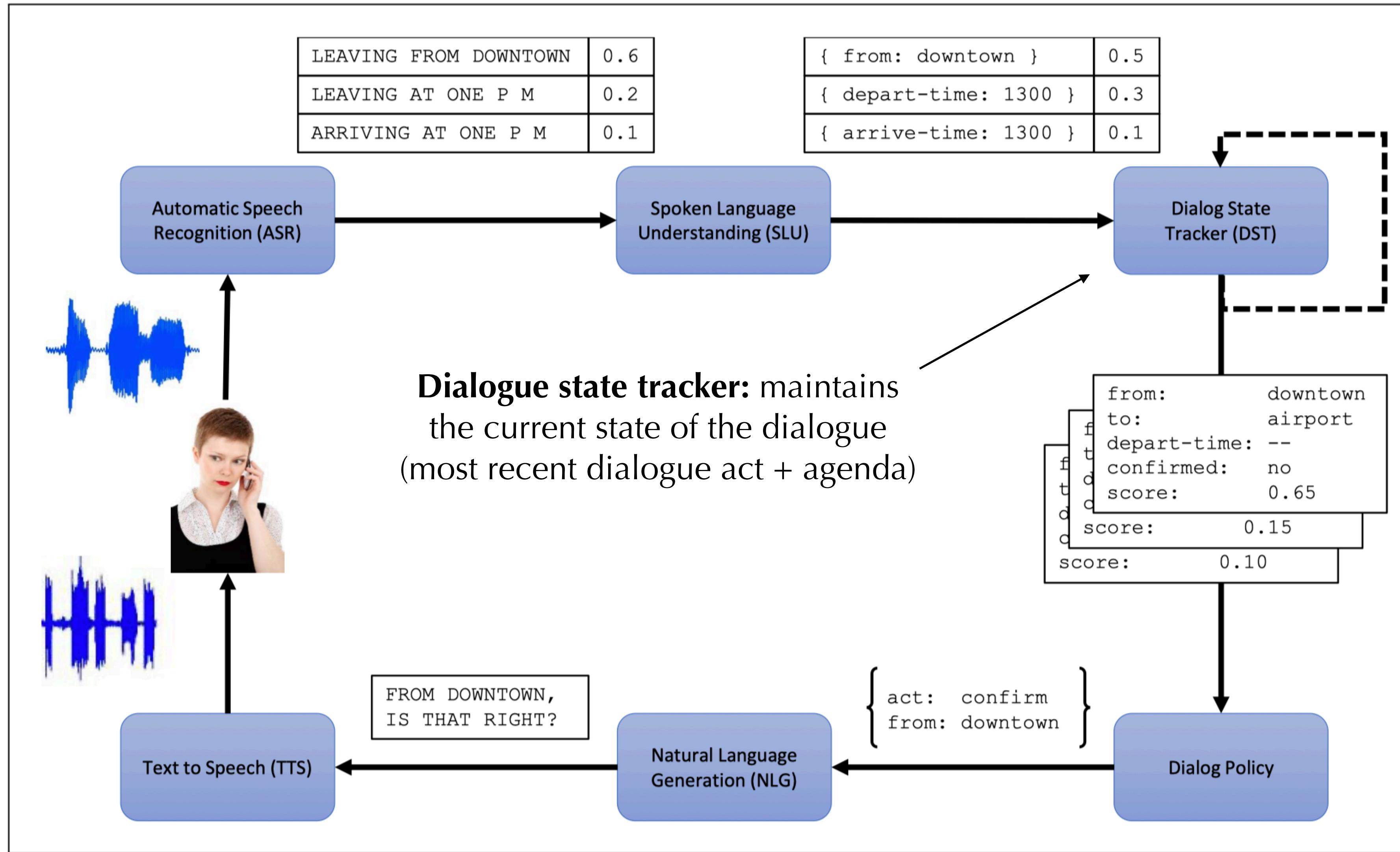
# Architecture of Task-Oriented SDS



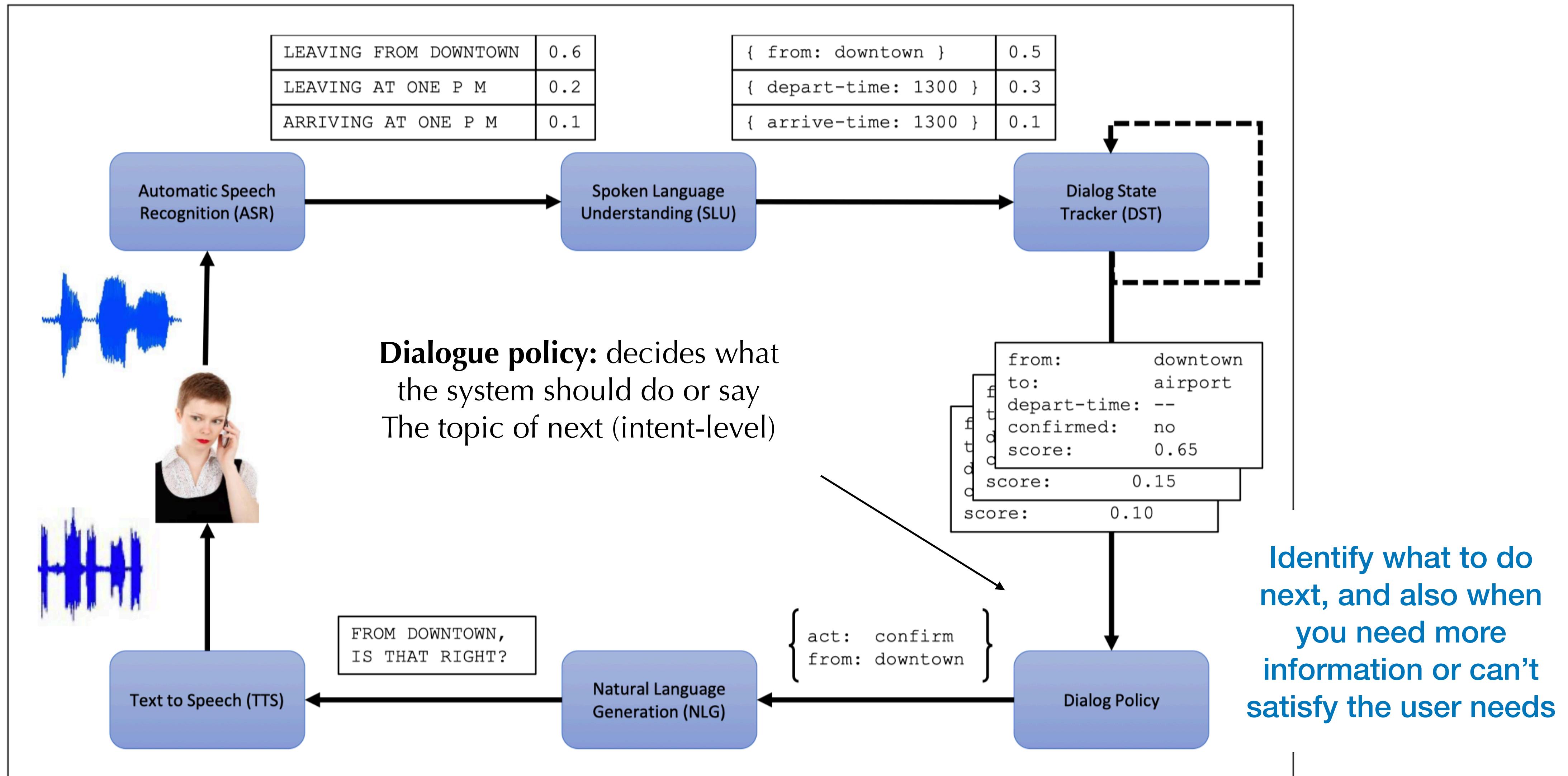
# Architecture of Task-Oriented SDS



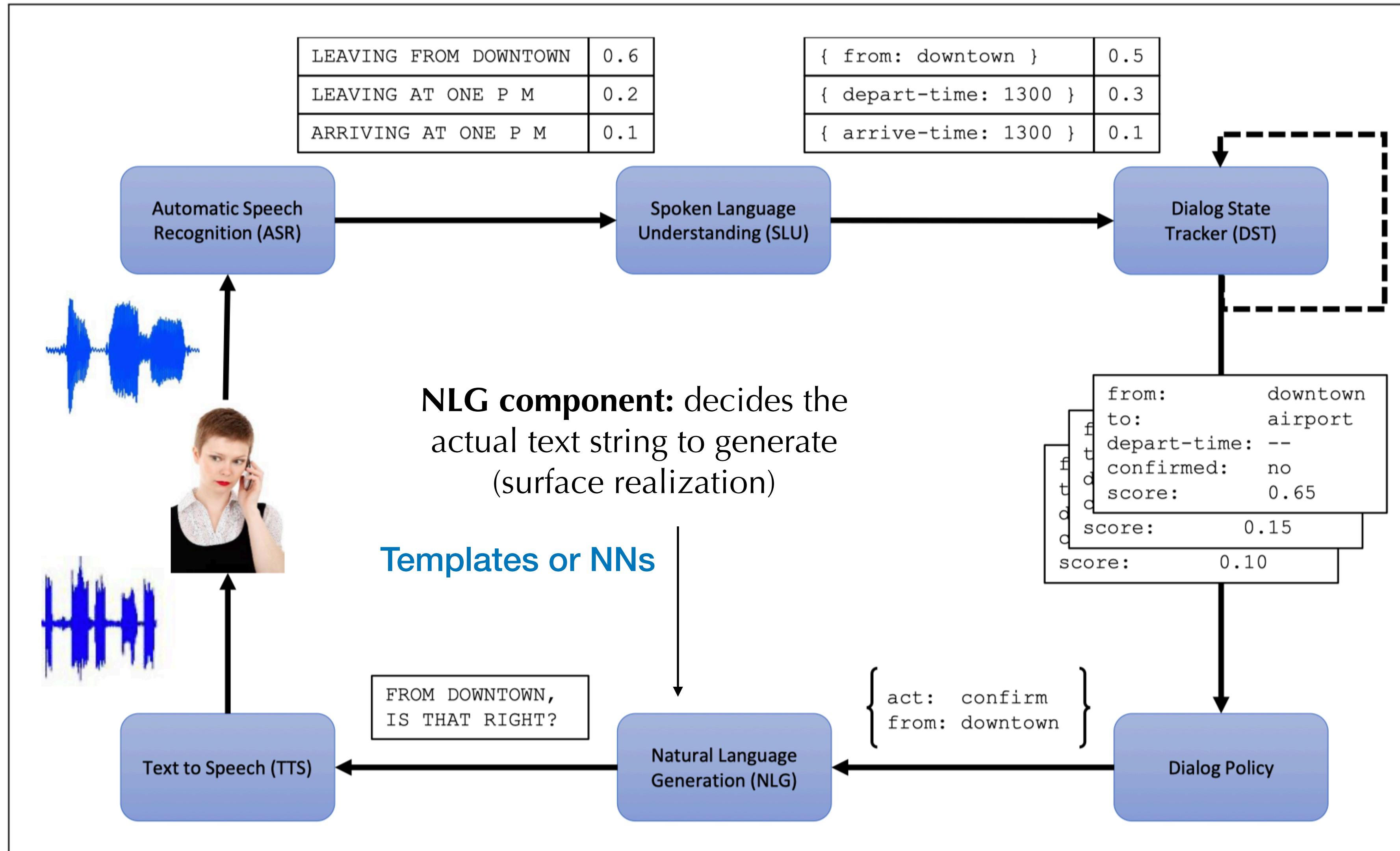
# Architecture of Task-Oriented SDS



# Architecture of Task-Oriented SDS



# Architecture of Task-Oriented SDS



# Task-Oriented Dialogue System

- **Goal:**
  - get information from the user to help **complete the specific task**.
- **Domain-Specific Knowledge:**
  - Ontology / Frame / Slot / Value
  - Slot Filling and Dialogue State Tracking
- **Architecture:**
  - ASR / SLU / DST / Dialogue Policy / NLG / TTS
- **Evaluation:**
  - **Task completion success** (slot error rate / task error rate)
  - **Efficiency cost** (#turns)
  - **Quality cost** (more comprehensive)

# **Chatbot v.s. Task-Oriented Dialog System**

**Information  
Retrieval**

**Question  
Answering**

**Chatbot**

**Task-Oriented  
Dialog System**

**What are their differences?**

# Chatbot v.s. Task-Oriented Dialog System

	Information Retrieval	Question Answering	Chatbot	Task-Oriented Dialog System
Input	structured	unstructured	unstructured	unstructured

# Chatbot v.s. Task-Oriented Dialog System

	Information Retrieval	Question Answering	Chatbot	Task-Oriented Dialog System
Input	structured	unstructured	unstructured	unstructured
Interaction	single-round	single-round	multi-round	multi-round

# Chatbot v.s. Task-Oriented Dialog System

	Information Retrieval	Question Answering	Chatbot	Task-Oriented Dialog System
<b>Input</b>	structured	unstructured	unstructured	unstructured
<b>Interaction</b>	single-round	single-round	multi-round	multi-round
<b>supervision</b>	available	available	sparse, delayed	sparse, delayed

# Chatbot v.s. Task-Oriented Dialog System

	Information Retrieval	Question Answering	Chatbot	Task-Oriented Dialog System
<b>Input</b>	structured	unstructured	unstructured	unstructured
<b>Interaction</b>	single-round	single-round	multi-round	multi-round
<b>supervision</b>	available	available	sparse, delayed	sparse, delayed
<b>dataset</b>	synthesis, collected	collected	collected	wizard-of-oz

# Chatbot v.s. Task-Oriented Dialog System

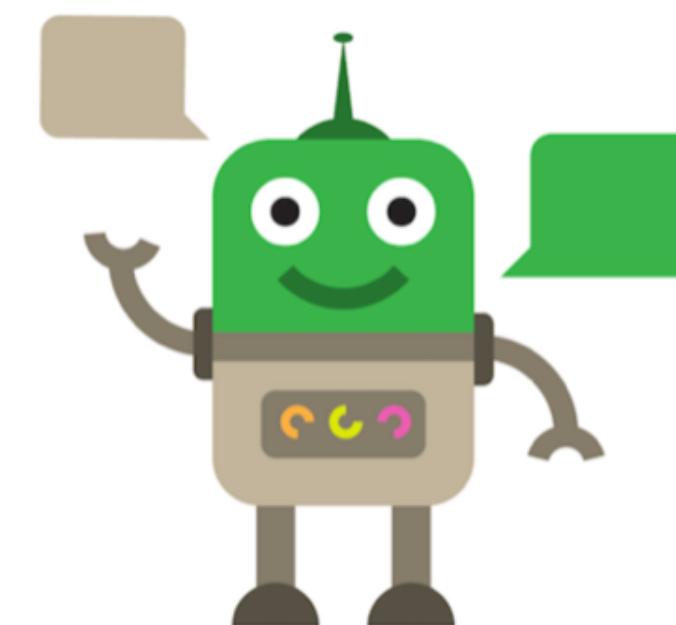
	Information Retrieval	Question Answering	Chatbot	Task-Oriented Dialog System
<b>Input</b>	structured	unstructured	unstructured	unstructured
<b>Interaction</b>	single-round	single-round	multi-round	multi-round
<b>supervision</b>	available	available	sparse, delayed	sparse, delayed
<b>dataset</b>	synthesis, collected	collected	collected	wizard-of-oz
...				

# Chatbot v.s. Task-Oriented Dialog System

	Information Retrieval	Question Answering	Chatbot	Task-Oriented Dialog System
<b>Input</b>	structured	unstructured	unstructured	unstructured
<b>Interaction</b>	single-round	single-round	multi-round	multi-round
<b>supervision</b>	available	available	sparse, delayed	sparse, delayed
<b>dataset</b>	synthesis, collected	collected	collected	wizard-or-oz
...				

# 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



Dialogue Systems

# **Rule-Based v.s. Data-Driven**

How to build a task-oriented dialog system?

**Rule-based system v.s. Data-driven system**

# Rule-Based v.s. Data-Driven

How to build a task-oriented dialog system?

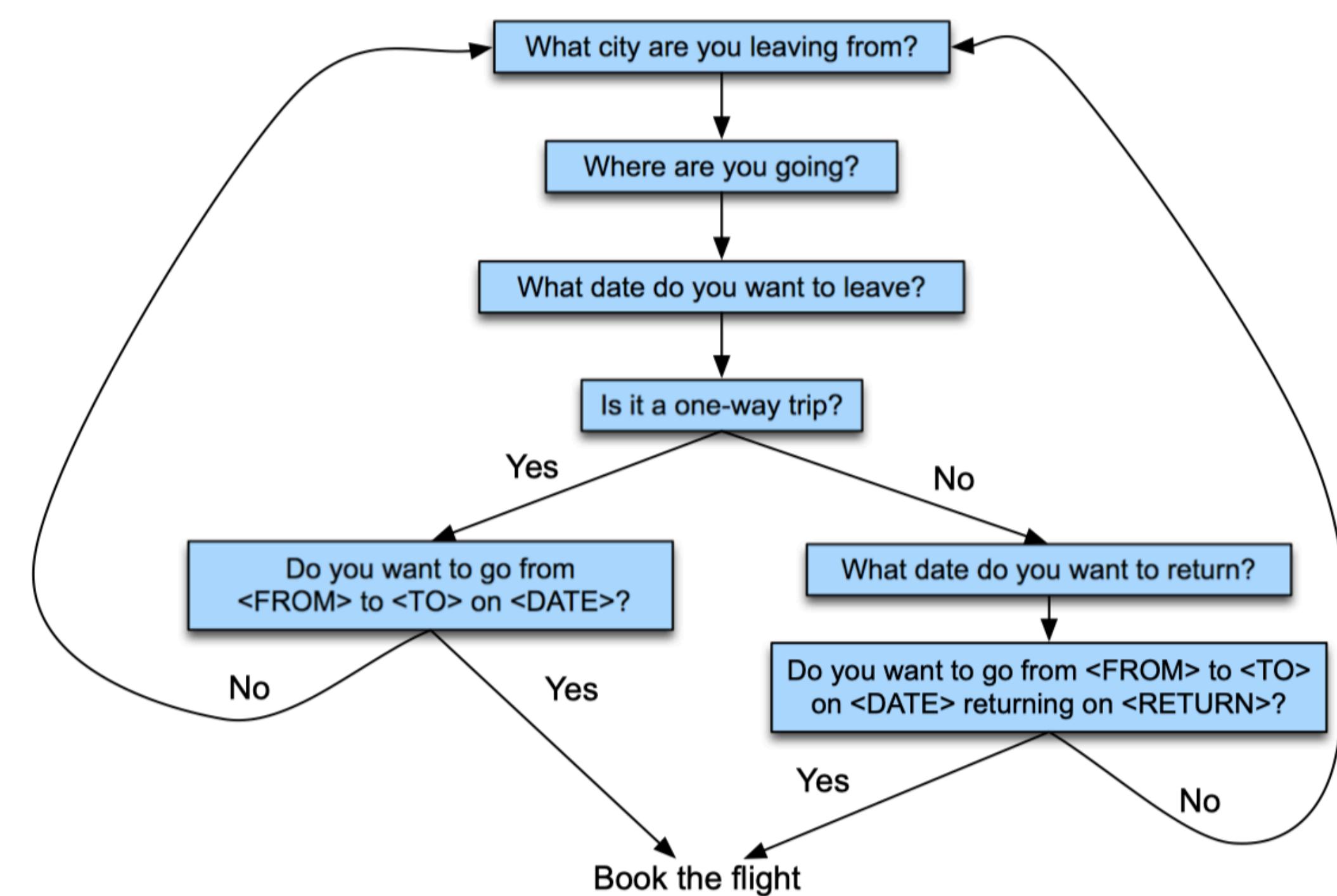
<b>SHOW</b>	→ show me   i want   can i see ...
<b>DEPART_TIME_RANGE</b>	→ (after around before) HOUR   morning   afternoon   evening
<b>HOUR</b>	→ one two three four... twelve (AMPM)
<b>FLIGHTS</b>	→ (a) flight   flights
<b>AMPM</b>	→ am   pm
<b>ORIGIN</b>	→ from CITY
<b>DESTINATION</b>	→ to CITY
<b>CITY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

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

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

# Rule-Based v.s. Data-Driven

How to build a task-oriented dialog system?

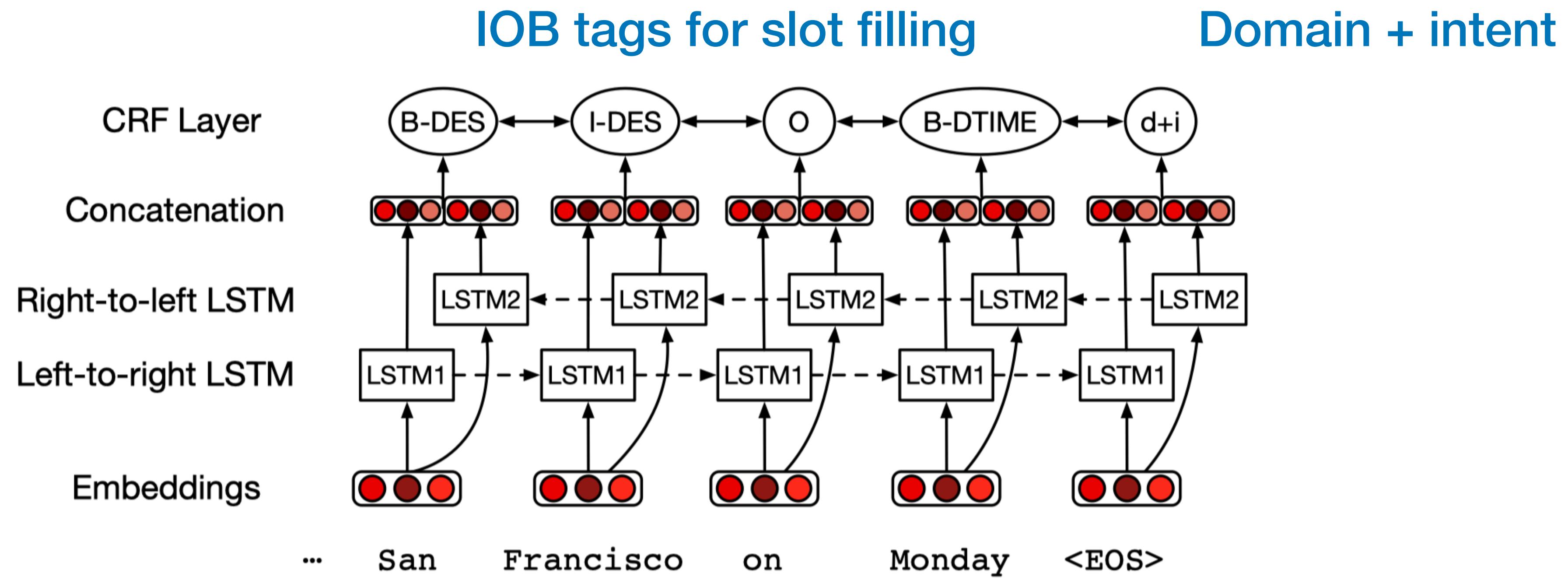


A simple finite-state automaton architecture for frame-based dialog.

**Rule-based system (Dialog Policy)**

# Rule-Based v.s. Data-Driven

How to build a task-oriented dialog system?

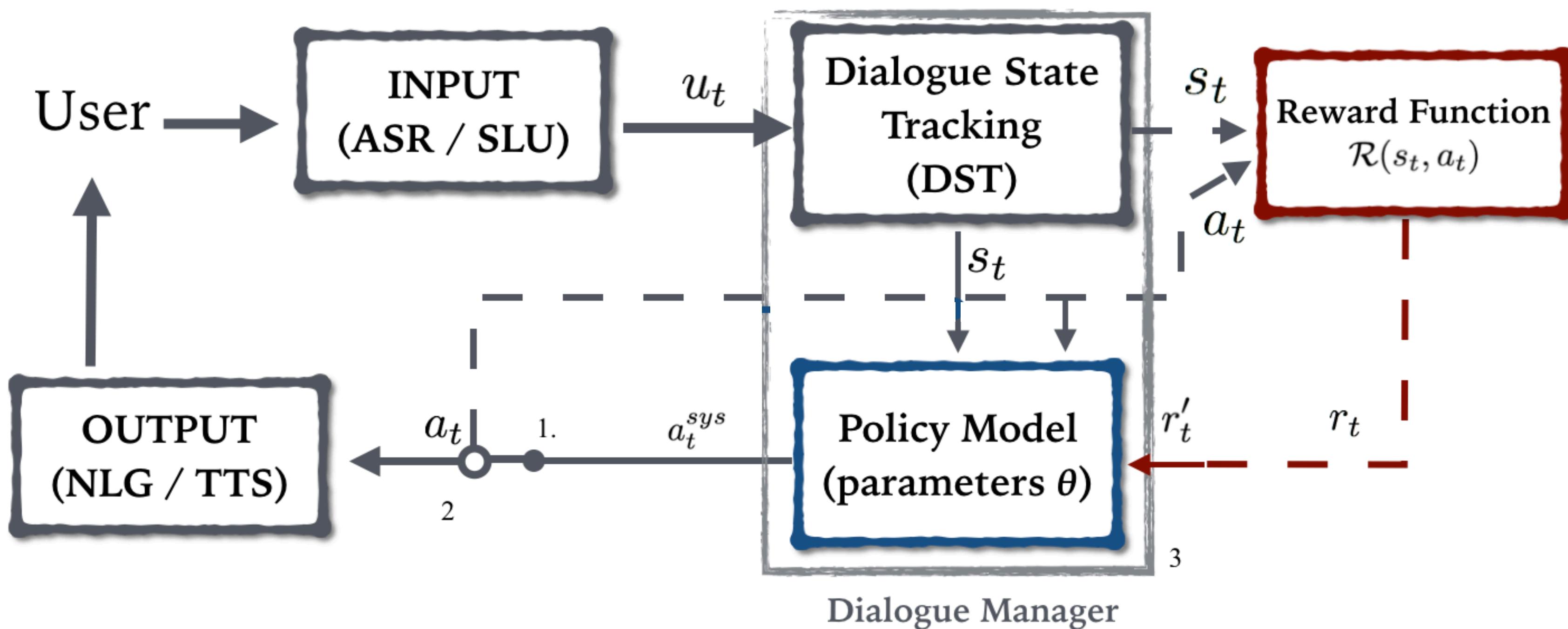


An LSTM architecture for slot filling, mapping the words in the input to a series of IOB tags plus a final state consisting of a domain concatenated with an intent.

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

# Rule-Based v.s. Data-Driven

How to build a task-oriented dialog system?

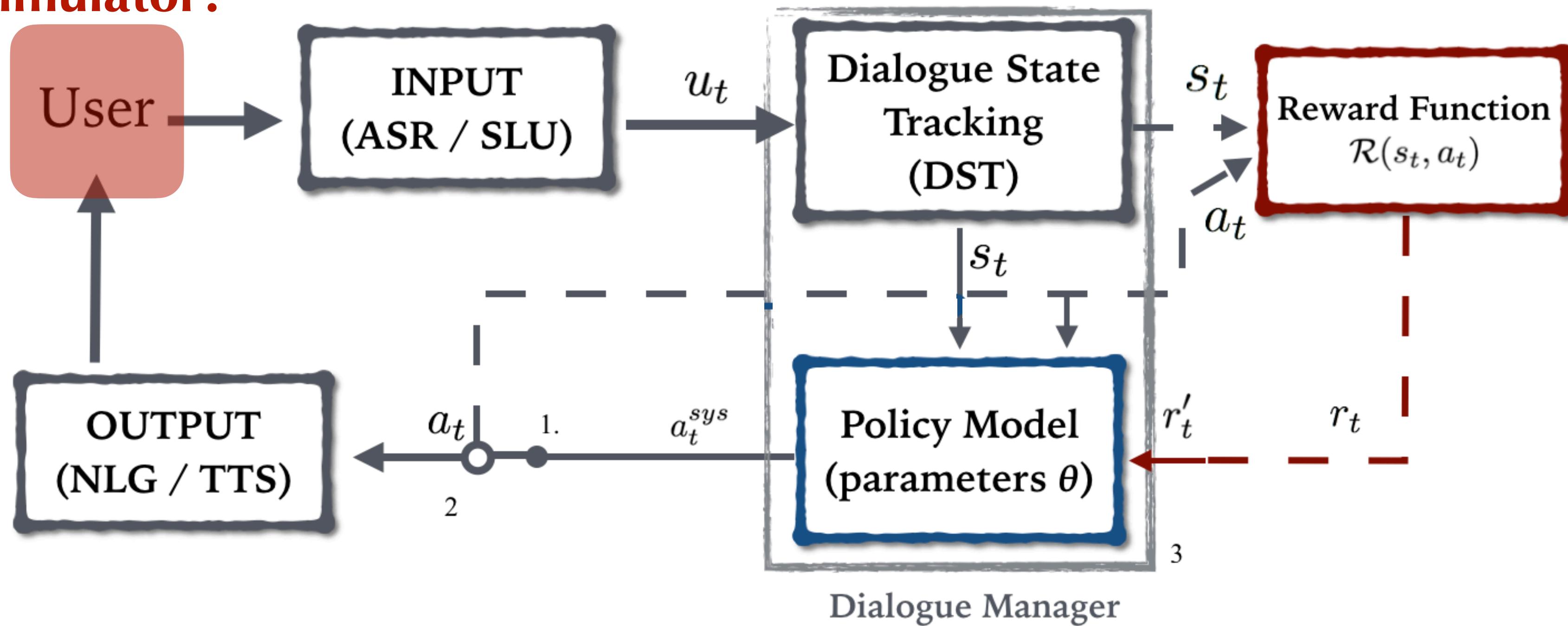


**Data-driven system (Dialog Policy)**

# Rule-Based v.s. Data-Driven

How to build a task-oriented dialog system?

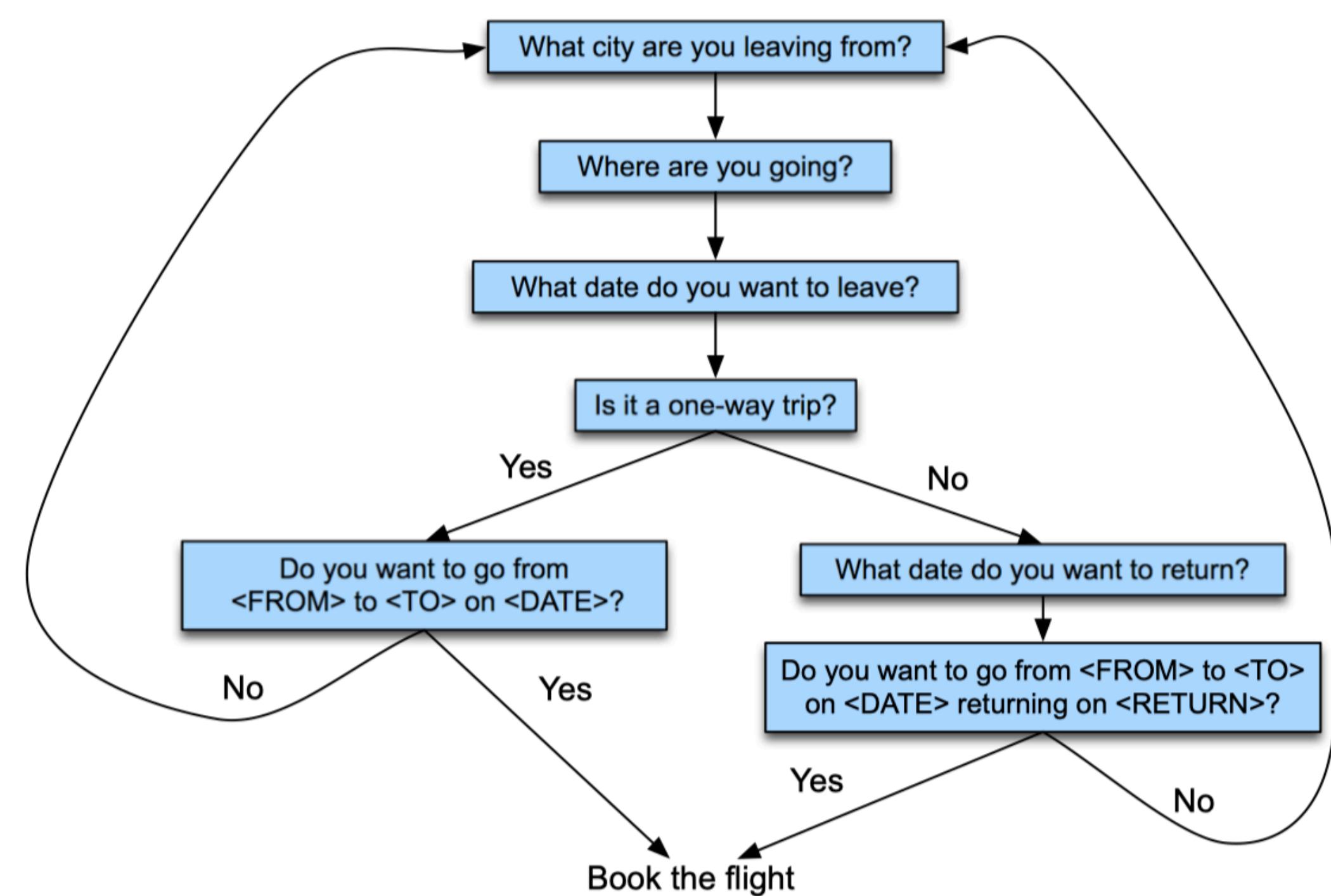
simulator?



Data-driven system (Dialog Policy)

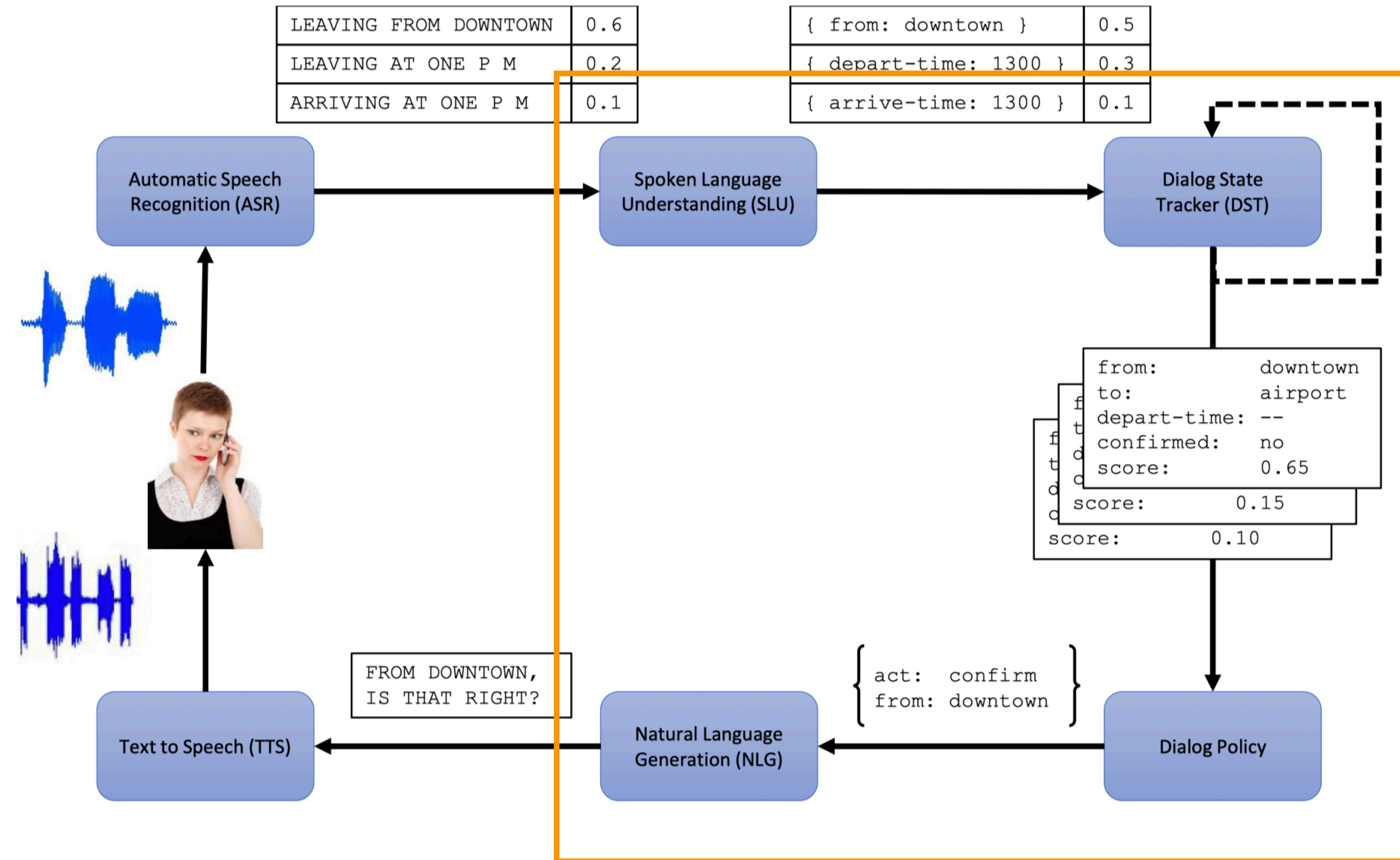
# Rule-Based v.s. Data-Driven

How to build a task-oriented dialog system?

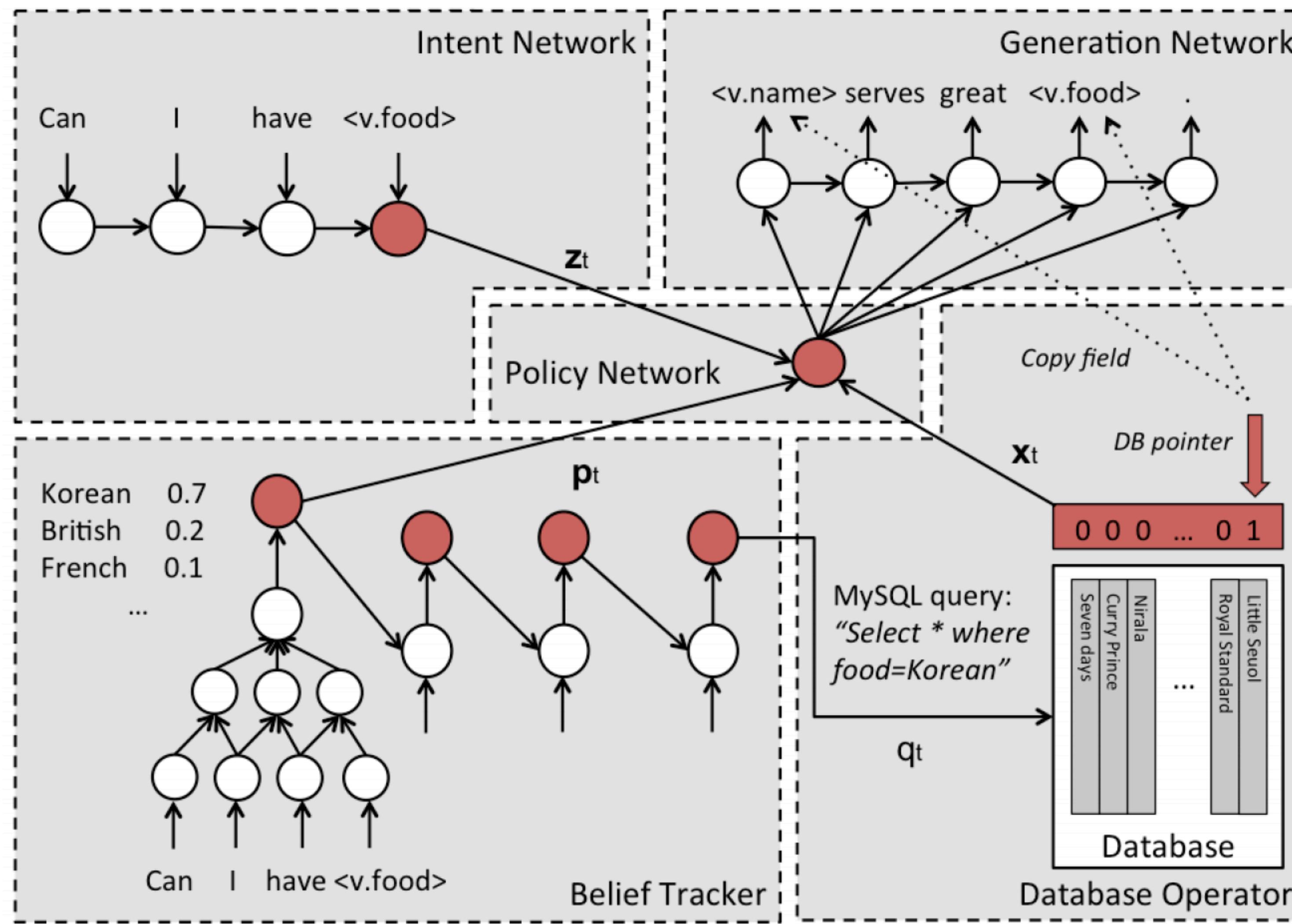


A simple finite-state automaton architecture for frame-based dialog.

**End-to-end systems**



# End-to-End Task-Oriented Dialog System



# **Rule-Based v.s. Data-Driven**

How to build a task-oriented dialog system?

**Rule-based v.s. Data-driven  
Pros & cons?**

# Rule-Based v.s. Data-Driven

How to build a task-oriented dialog system?

## Rule-based v.s. Data-driven Pros & cons?

### Rule-Based Methods

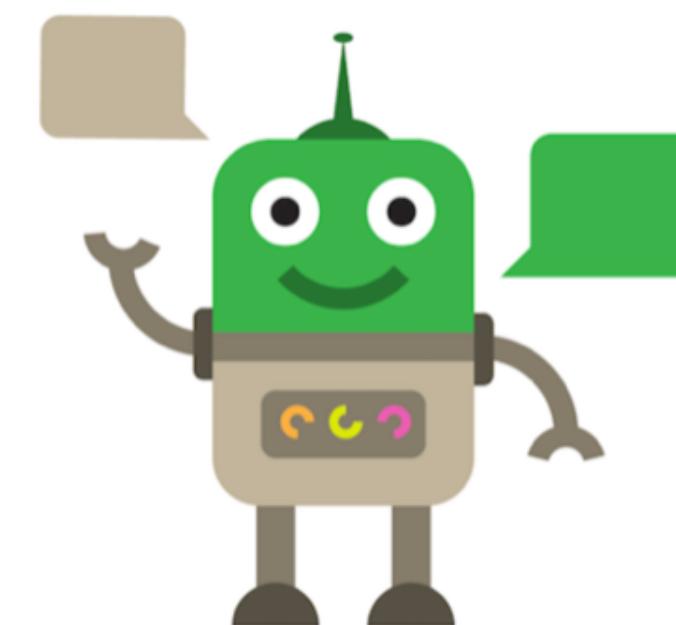
- hand-craft rules, “safe” but not “flexible”.
- cheap in terms of dataset.
- expensive in terms of engineering.

### 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

# Overview

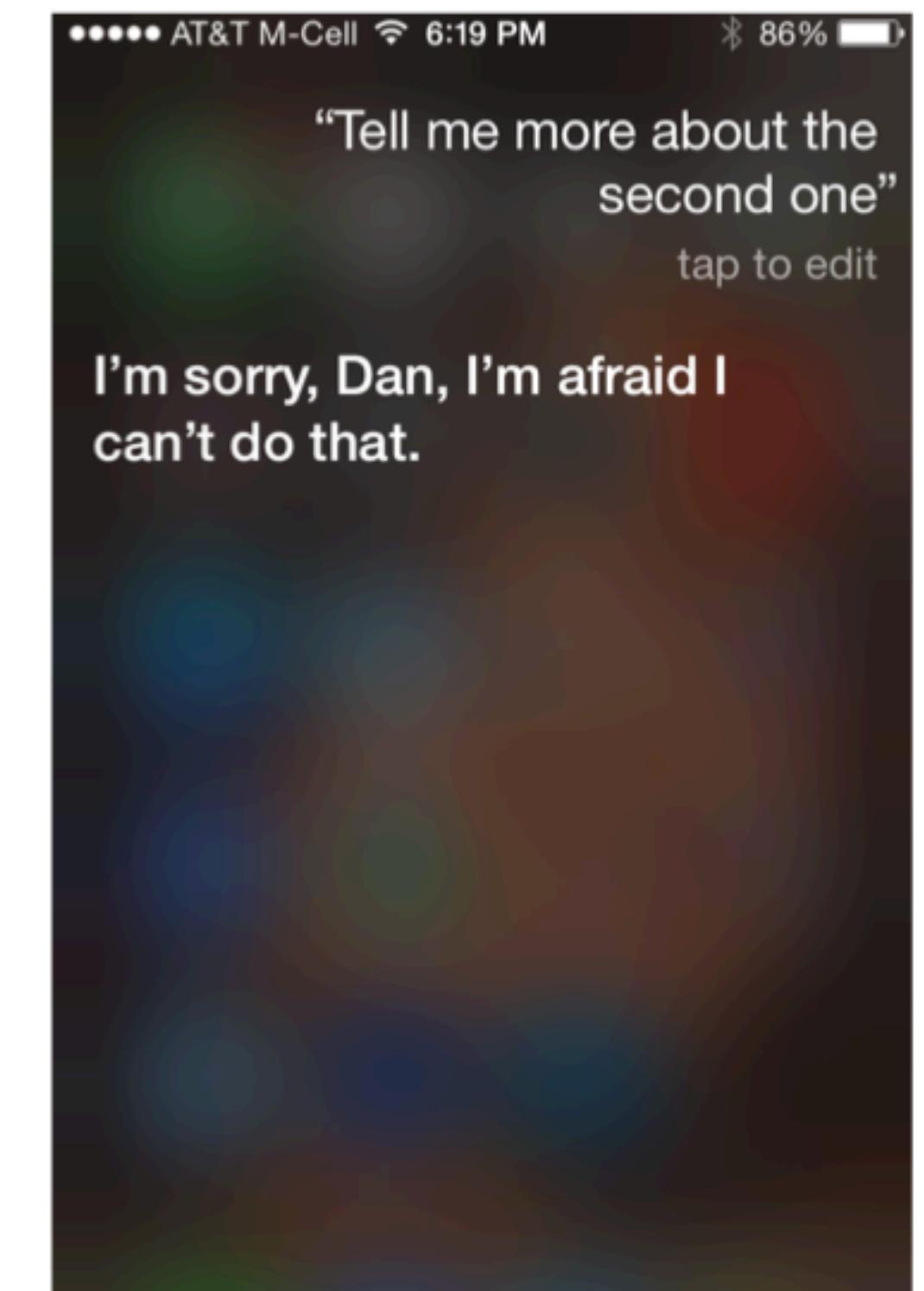
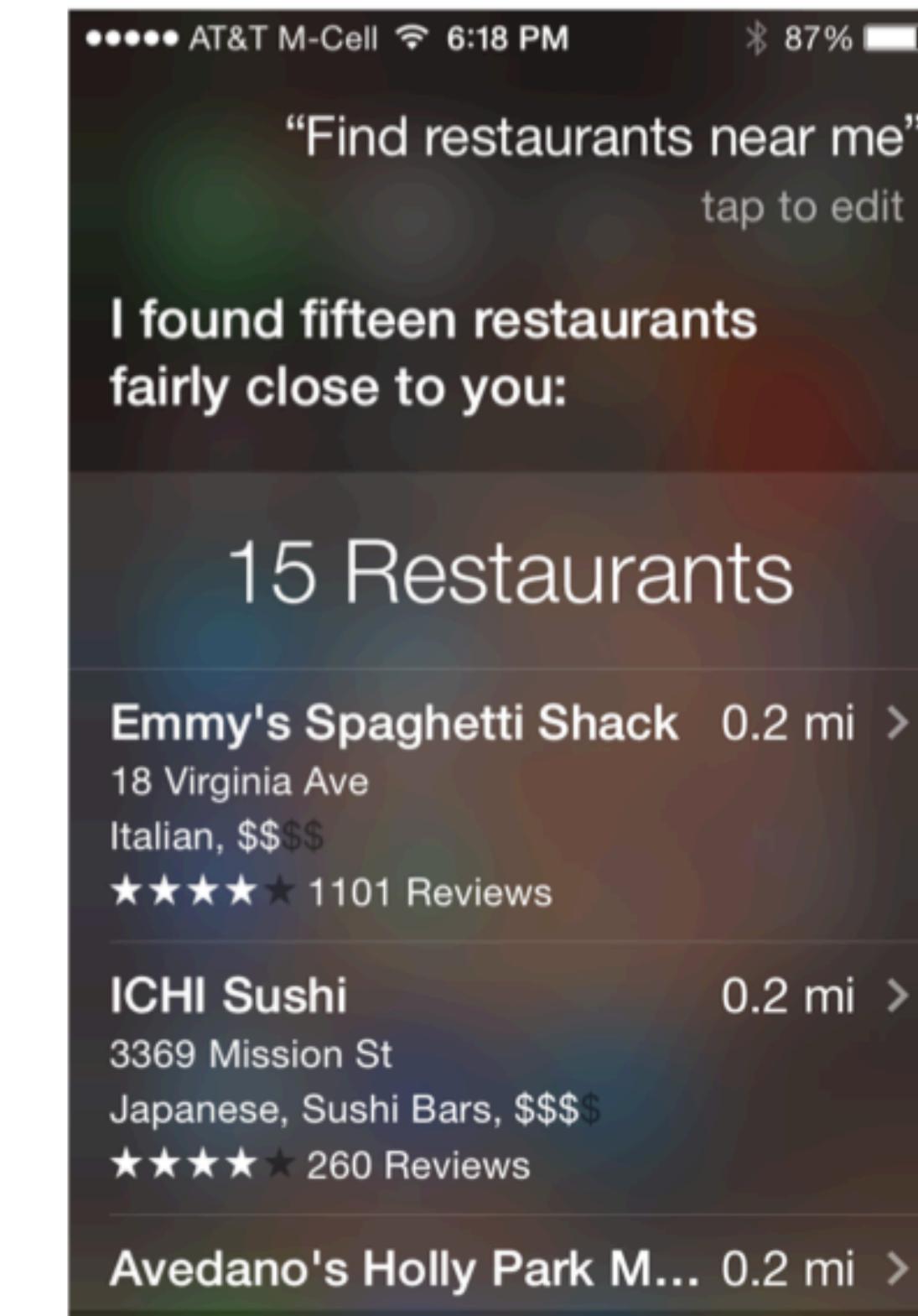
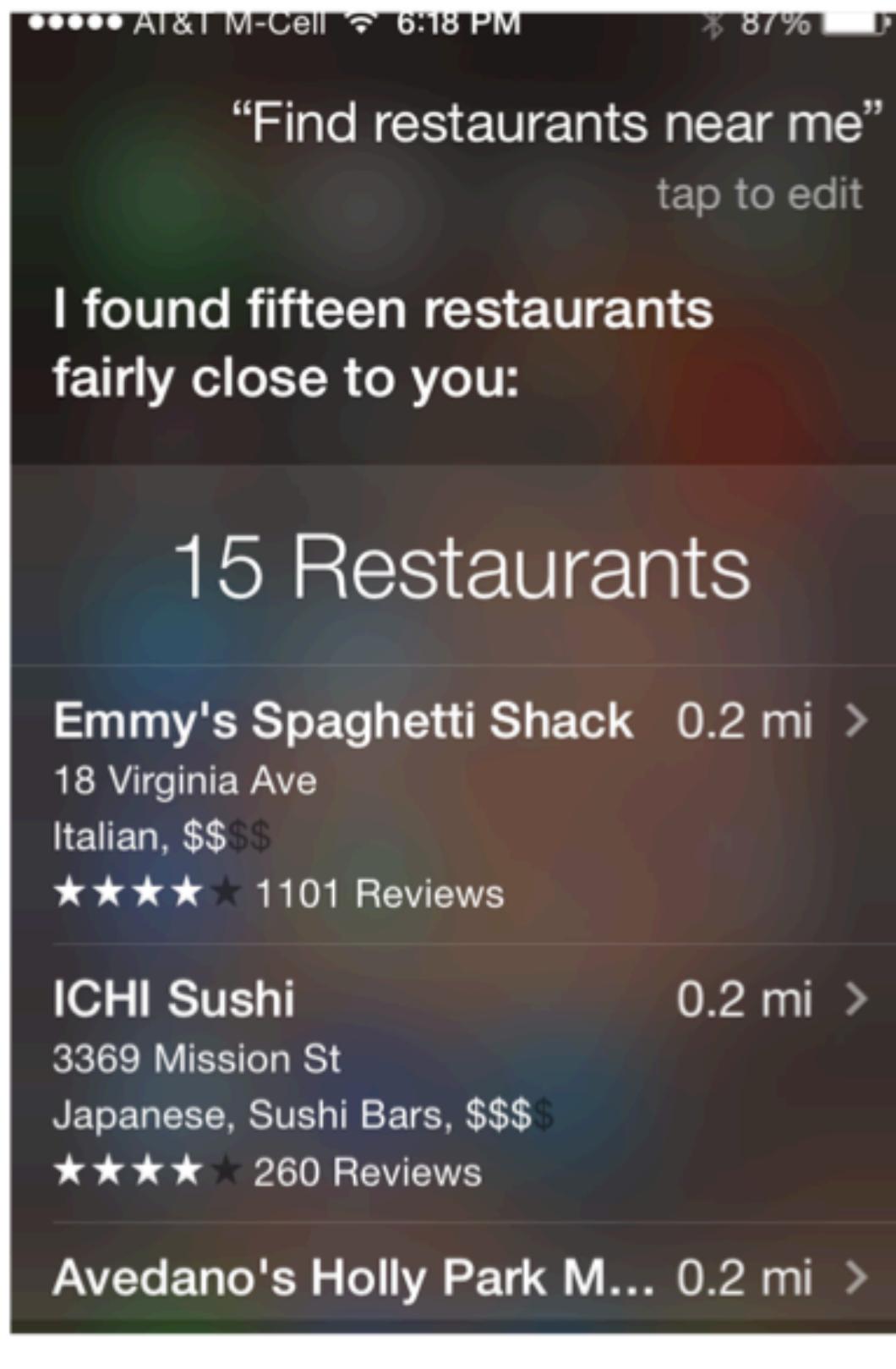
- What's a dialogue system?
- Properties of Human Conversation
- Chatbots v.s. Task-oriented dialogues systems
- Rule-based v.s. Data-driven
- **Remaining Challenges**



Dialogue Systems

# Challenges

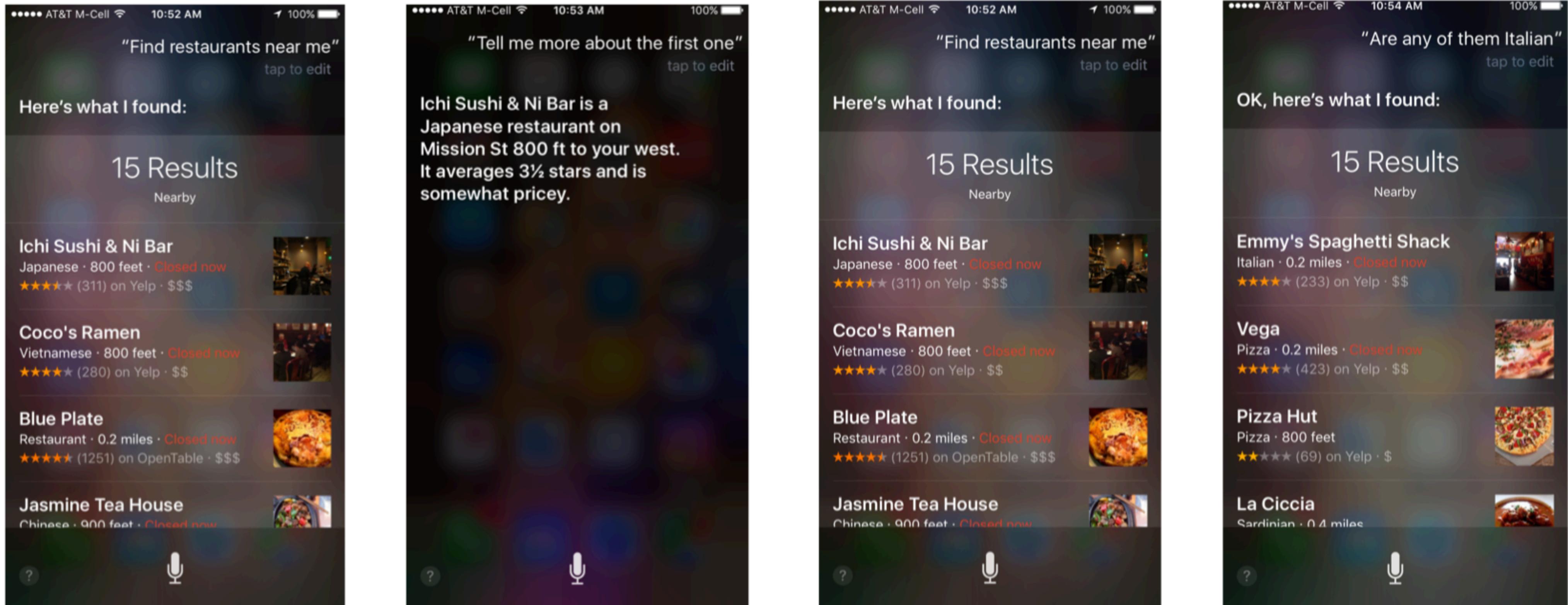
## Understanding the Context



Two sets of interactions with Siri in 2014.

# Challenges

## Understanding the Context



The same follow-up questions that Siri couldn't answer in 2014 receive appropriate responses when posed to Siri in 2017.

# Challenges

Understanding the Context

Uncertainty / Ambiguity

Reward Design

Knowledge Embedding

Domain Adaptation

Data/Interaction Scarcity

...