Dialog Systems and Chatbots

CMSC 473/673 - NATURAL LANGUAGE PROCESSING

Slides modified from Chris Callison-Burch & João Sedoc

Misconception from Knowledge Check

Mixture of Experts is not to be confused with an Ensemble Model

Learning Objectives

Hear about early chatbot systems

Conversational Agents aka Dialog Systems

Digital Assistants

Answering questions on websites

Communicating with robots

Chatting for fun

Clinical uses

Challenging properties of human conversation

- Turn taking
- Speech acts
- Grounding
- Dialogue structure
- Initiative
- Implicature

Turn taking

A conversation is a sequence of turns, where you take a turn and then I take a turn. A turn can be a sentence, or a single word.

A system must know when to start and stop talking.

Spoken dialogue systems must also detect whether a user is done speaking, so they can process the utterance and respond. This task of **endpoint detection** is tricky because people often pause mid-turn.

Speech acts

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*)

Conversational Implicature

Agent: And, what day in May did you want to travel?

Client: I need to be there for a meeting that's from the 12th to the 15th.

Notice that the client does not in fact answer the agent's question. The speaker expects the hearer to draw certain inferences.

Implicature means certain kinds of inferences.

Gricean maxims. The *maxim of relevance* means that when the client mentions a meeting on the 12th, the agent reasons 'There must be some relevance for mentioning this meeting. What could it be?'

Two Classes of Dialog Systems

1. **Chatbots**
   - Systems designed for extended conversations
   - Chatting for fun and entertainment

2. **Task-Oriented Dialogue Agents**
   - Goal-Based Agents
   - Siri, interface with robots, booking flights or hotels

Part I: Chatbots

Systems designed for extended conversations. Chatbots mimic unstructured conversations or ‘chats’ that are characteristic of informal human-human interaction

Architecture include:

Rule-Based

- Pattern-action rules

Corpus-Based

- Information Retrieval
- Neural network encoder-decoder

Part I: Chatbots

Systems designed for extended conversations. Chatbots mimic unstructured conversations or ‘chats’ that are characteristic of informal human-human interaction

Architecture include:

Rule-Based

- **Pattern-action rules**

Corpus-Based

- Information Retrieval
- Neural network encoder-decoder

ELIZA: Weizenbaum (1966)

Simulates a psychologist

Rogerian psychology is a rare type of conversation where "assume the pose of knowing almost nothing of the real world"

Draw the patient out by reflecting patient's statements back at them

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

::

::

WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU?

My father

YOUR FATHER

You are like my father in some ways.

ELIZA: Weizenbaum (1966)

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.

DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE

ELIZA Pattern Transform Rules

Pattern matching rules:

.* YOU **(.*)** ME **–>** WHAT MAKES YOU THINK I **$1** YOU?

This rule would transform

You **must really hate** me

into

WHAT MAKES YOU THINK I **MUST REALLY HATE** YOU?

Simplified sketch of ELIZA

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*)

PARRY: Colby (1971)

ELIZA like regular expressions But a much richer –

- Control Structure
- Language Understanding Capabilities
- Mental Model : Parry has affective variables
  - Fear, Mistrust, Anger
  - If Anger level is high, respond with hostility

First system to pass the **Turing test (1971)**

Turing Test

“

I propose to consider the question, ‘Can machines think?’

”

- A human evaluator would judge text-based conversations between a human and a machine designed to generate human-like responses.
- If the evaluator cannot reliably tell the machine from the human, the machine is said to have passed the test.

A. M. TURING, I.—COMPUTING MACHINERY AND INTELLIGENCE, *Mind*, Volume LIX, Issue 236, October 1950, Pages 433–460, https://doi.org/10.1093/mind/LIX.236.433

Part I: Chatbots

Systems designed for extended conversations. Chatbots mimic unstructured conversations or ‘chats’ that are characteristic of informal human-human interaction

Architecture include:

Rule-Based

- Pattern-action rules (Eliza)

Corpus-Based

- **Information Retrieval**
- **Neural network encoder-decoder**

Conversational Data

Need: large collections of human conversations

Conversational threads on Twitter or Weibo (微博)

Retrieve dialog from movies, indexing subtitles

Recorded telephone conversations, collected for speech research

Crowdsourced conversations via Mechanical Turk and ParlAI

Information Retrieval based Chatbots

Treat the human user's input as a query vector **q**

Search over a large corpus **C** of conversation to find the closest matching turn **t'** in those previous conversations.

Return the response **r** to that conversational turn.

$t' = \arg\max_{t \in C} \text{cosine\_similarity}(q, t)$.

r = response(t')

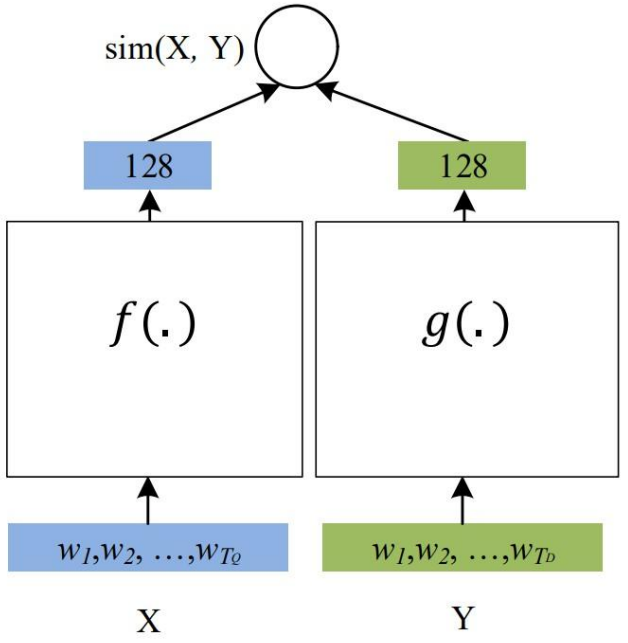
q = Have you watched Doctor Who?

t' = Do you like Doctor Who?

r = Yes, I love SciFi shows!

IR with Neural Network-Based Similarity Model

Relevance measured by cosine similarity



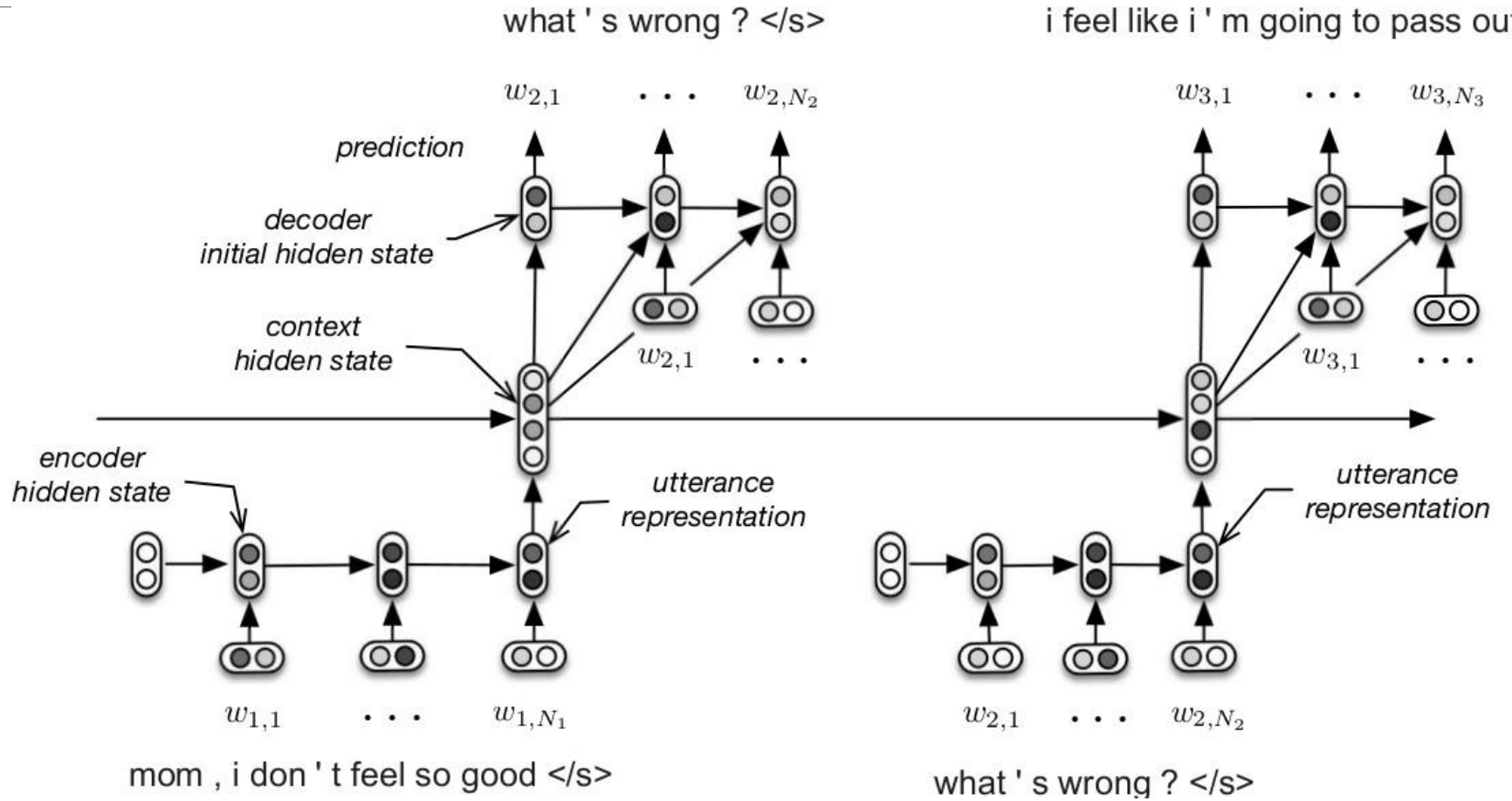
Word sequence x_t

Learning: maximize the similarity between X (source) and Y (target)

Representation: use DNN to extract abstract semantic features, f or g is a

- Multi-Layer Perceptron (MLP) if text is a bag of words [Huang+ 13]
- **Convolutional Neural Network (CNN) if text is a bag of chunks** [Shen+ 14]
- Recurrent Neural Network (RNN) if text is a sequence of words [Palangi+ 16]

Seq2Seq Encoder-Decoder Models



Sample Responses: Enc-Dec RNN Model

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

Responses that are harder to distinguish from human responses

Two Classes of Dialog Systems

1. **Chatbots**
   - Systems designed for extended conversations
   - Chatting for fun and entertainment

2. **Task-Oriented Dialogue Agents**
   - Goal-Based Agents
   - Siri, interface with robots, booking flights or hotels

Part II: Frame-based Dialogue Systems

- Task-based Dialogue Agents
- Based on “Domain Ontology”
  - A set of “Frames”
- Frame:
  - A knowledge structure representing user intentions
  - A collection of “slots”
  - Each ”slot” having a set of “values”

Example: Travel Domain

Slot : Origin City

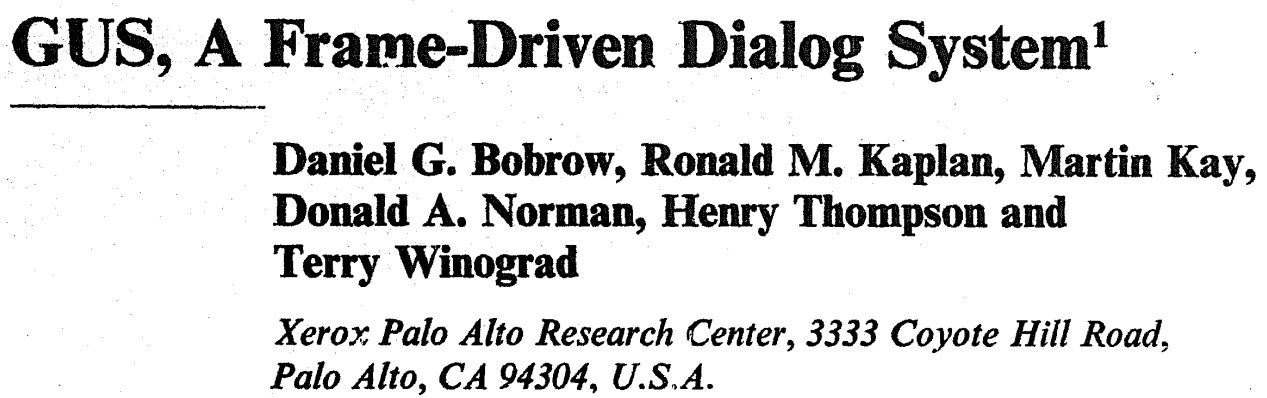
Type : City

Value : San Francisco

Slot	Type	Question
ORIGIN	city	What city are you leaving from?
DEST	city	Where are you going?
DEP DATE	date	What day would you like to leave?
DEP TIME	time	What time would you like to leave?
AIRLINE	line	What is your preferred airline?

Frame-based Dialogue Agents

Artificial Intelligence Journal, 1977



GUS, A Frame-Driven Dialog System[1]

Daniel G. Bobrow, Ronald M. Kaplan, Martin Kay, Donald A. Norman, Henry Thompson and Terry Winograd

Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304, U.S.A.

- Still the industrial state of the art
- Siri/Alexa based on GUS architecture

GUS system : An actual dialogue

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

Slot types can be complex

The type *DATE*

```
DATE
  MONTH:NAME    YEAR:INTEGER     DAY:(BOUNDED-INTEGER 1 31)
  WEEKDAY:(MEMBER (Sunday Monday Tuesday Wednesday
                   Thursday Friday Saturday))
```

Control structure for frame-based dialog

Consider a trivial airline travel system:

- Ask the user for a departure city
- Ask for a destination city
- Ask for a time
- Ask whether the trip is round-trip or not

Natural language understanding for filling slots in GUS

1. Domain classification

   Asking weather? Booking a flight? Programming alarm clock?

2. Intent Determination

   Find a Movie, Show Flight, Remove Calendar Appt

3. Slot Filling

   Extract the actual slots and fillers

Natural language understanding for filling slots in GUS

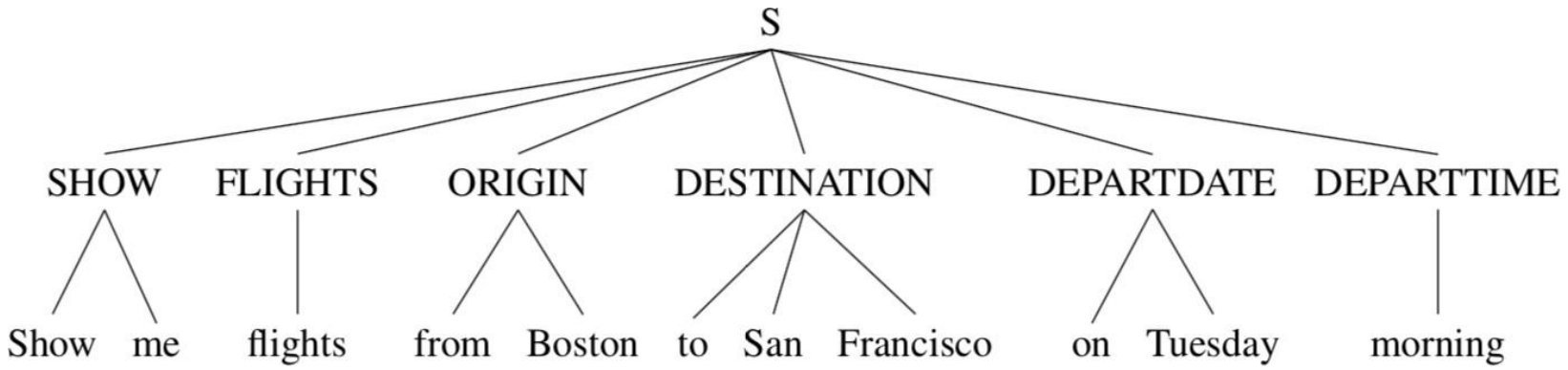
Show me morning flights from Boston to SF on Tuesday.

DOMAIN:	AIR-TRAVEL
INTENT:	SHOW-FLIGHTS
ORIGIN-CITY:	Boston
ORIGIN-DATE:	Tuesday
ORIGIN-TIME:	morning
DEST-CITY:	San Francisco

Rule-based Slot-filling

- Semantic Grammar Rules or Regular Expressions

```
Wake me (up) | set (the|an) alarm | get me up
```



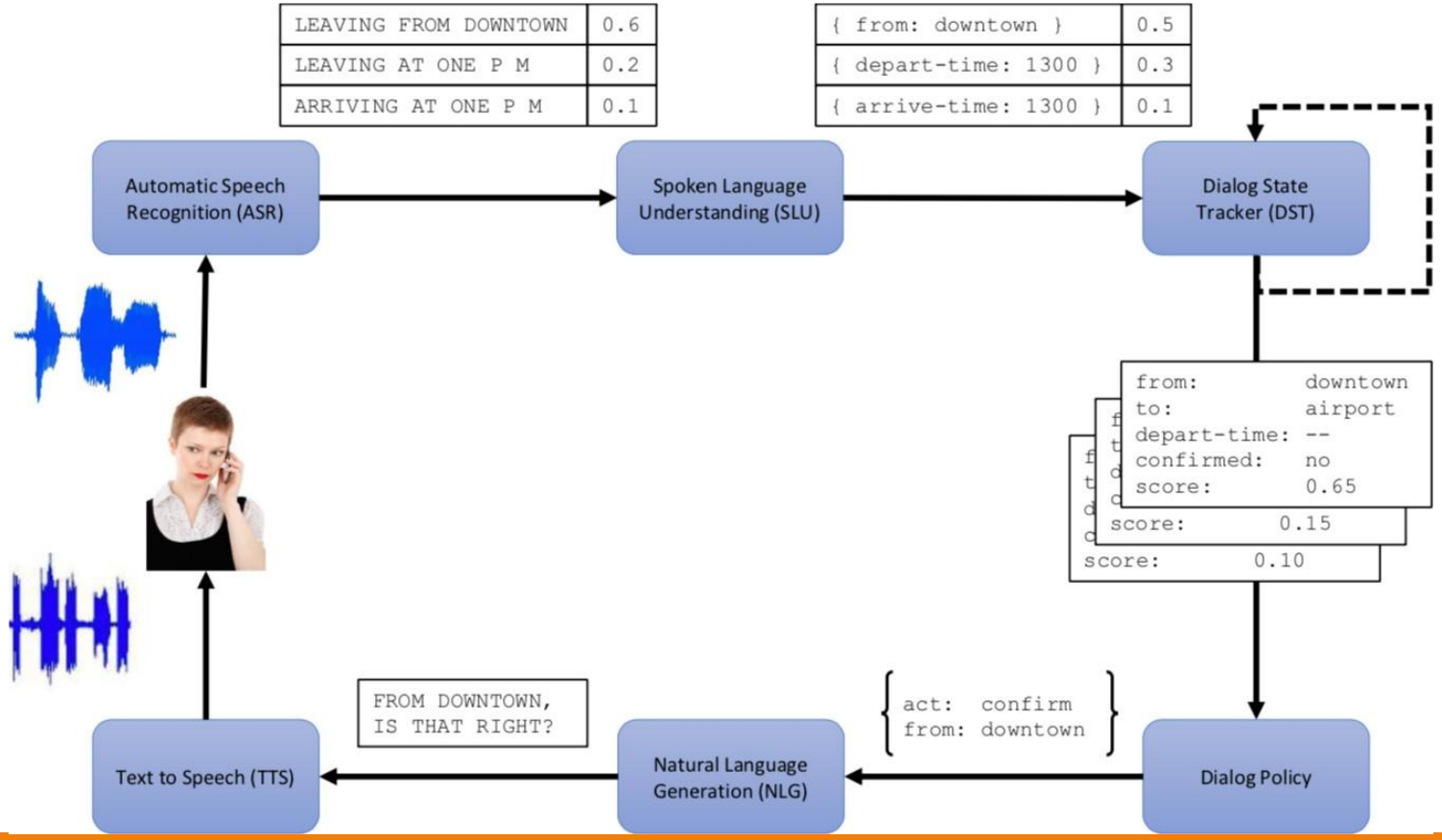
A semantic grammar parse for a user sentence, using slot names as the internal parse tree nodes

Rule Sets

- Collections of **rules** consisting of:
  - condition
  - action
- When user input is processed, facts added to store and
  - rule conditions are evaluated
  - relevant actions executed

Dialogue-State Architecture

More sophisticated version of frame-based architecture



Williams et al. (2016)

Dialogue Acts

Combining idea of speech acts and grounding into a single representation

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 used by a restaurant recommendation system (Young et al. (2010))

Dialogue Acts

	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()`

Sample dialogue from the Recommender System of Young et al. (2010)

Machine Learning for Slot Filling

- Supervised semantic parsing
- Model to map from input words to slot fillers, domain and intent
- Given a set of labeled sentences
  "I want to fly to San Francisco on Tuesday"

  Destination: SF Depart-date: Tuesday
- Requirements: Lots of labeled data

Slot Filling

"I want to fly to San Francisco on Monday afternoon please"

Use 1-of-N classifier (Naive Bayes, Logistic Regression, Neural Network, etc.)

- Input:

  features like word N-grams

- Output:

  Domain: AIRLINE Intent: SHOWFLIGHT

Dialogue State Tracking

- Keep track of
  - Current state of the frame (the fillers of each slot)
  - User's most recent dialogue act

User:	I'm looking for a cheaper restaurant
	`inform(price=cheap)`
System:	Sure. What kind - and where?
User:	Thai food, somewhere downtown
	`inform(price=cheap, food=Thai, area=centre)`
System:	The House serves cheap Thai food
User:	Where is it?
	`inform(price=cheap, food=Thai, area=centre); request(address)`
System:	The House is at 106 Regent Street

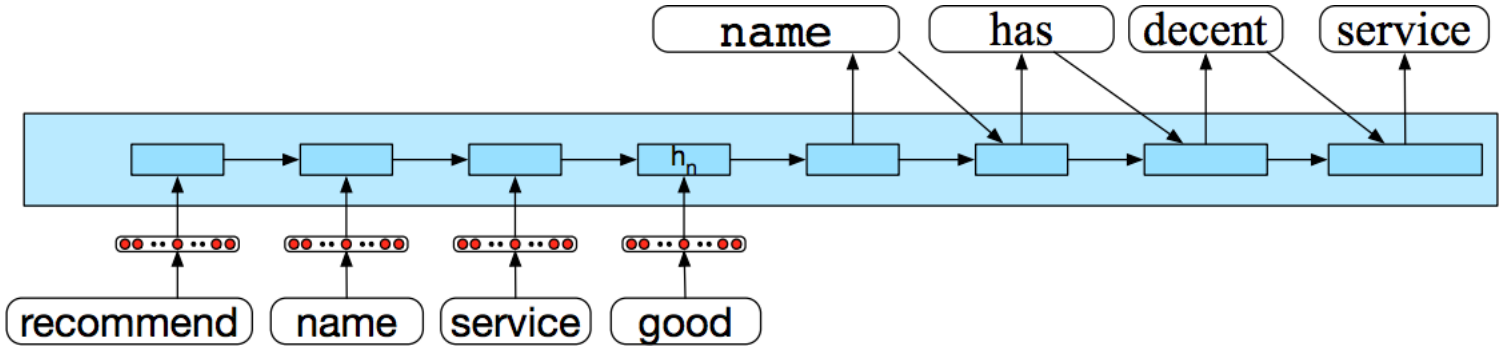
Sample output of a dialogue state tracker after each turn

Natural Language Generation

Modeled in two stages:

- Content Planning (what to say)
- Sentence Realization (how to say it)

Encoder Decoder Models : Map frames to sentences



An encoder decoder sentence realizer mapping slots/fillers to English

Evaluation

1. Slot Error Rate for a Sentence

$$\frac{\text{\# of inserted/deleted/substituted slots}}{\text{\# of total reference slots for sentence}}$$

2. End-to-end evaluation (Task Success)

Evaluation

"Make an appointment with Lara at 10:30 in ITE 216"

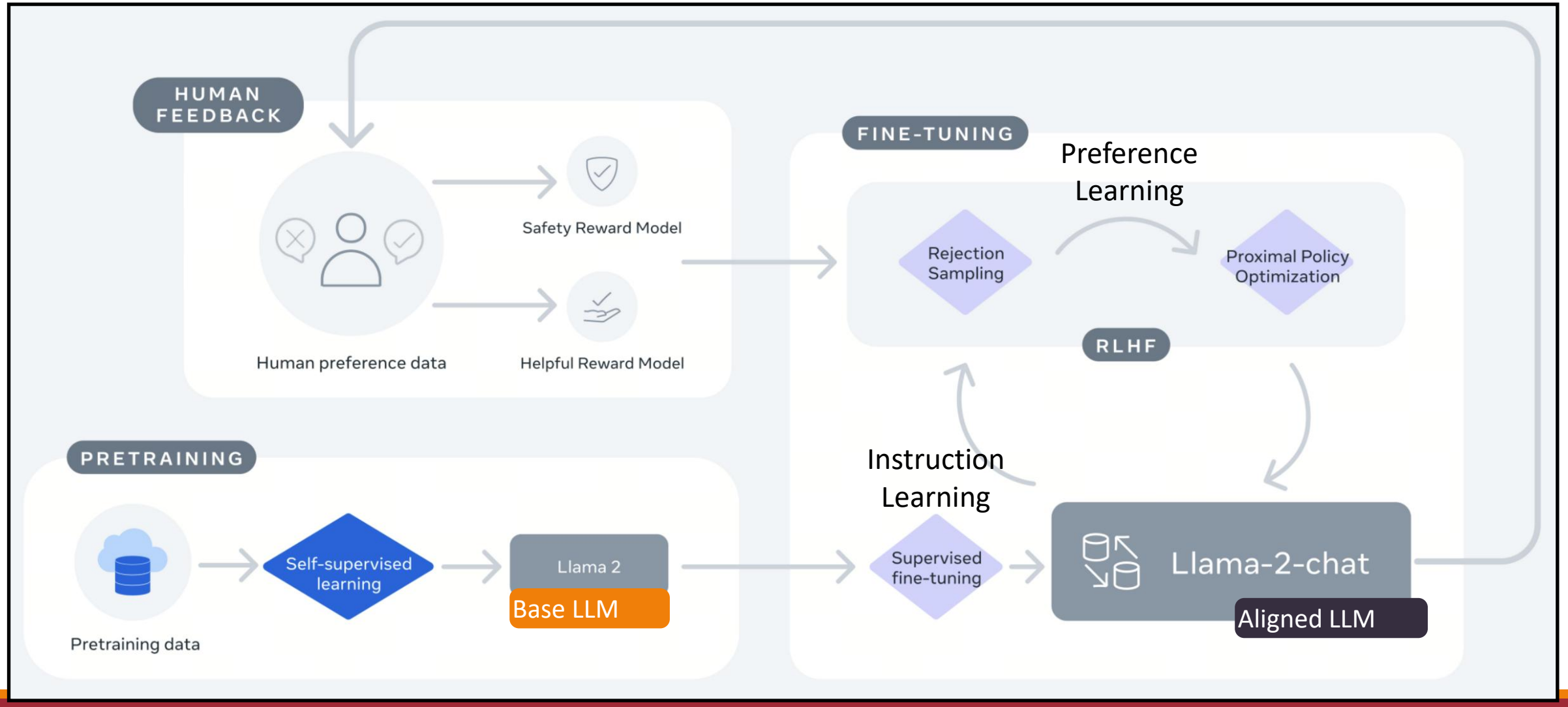
Slot	Filler
PERSON	Lara
TIME	11:30 a.m.
ROOM	ITE 216

Slot error rate: 1/3

Task success: At end, was the correct meeting added to the calendar?

Modern Chatbots

Transformer-based Architecture with Alignment



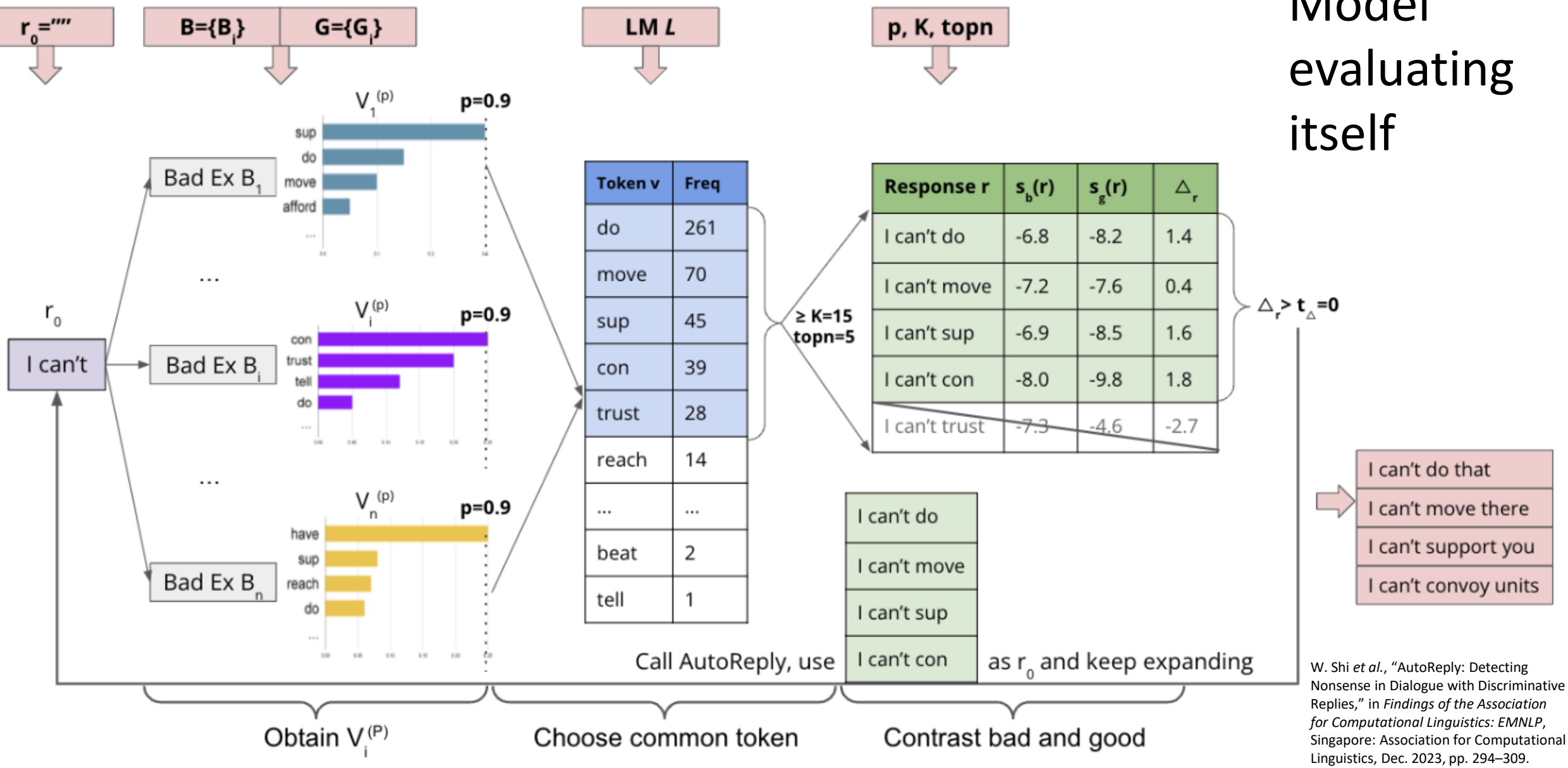
Using transformers to evaluate dialog

Speaker	Dialogue Context	Response Generator
ATHENA	Yeah. A lot of people have cool hobbies, but what about me? What kind of hobby do you think I would like?	Intro RG
USER	what's the name for music	
ATHENA	I love talking about music! Is there a musician or band that you like?	Music-KG
USER	no	
Rank	**Response Candidates**	**Response Generator**
D	I recently read this fact about Britney Spears. Spears had gotten her start in fame through being on The Mickey Mouse Club with other artists like Christina Aguilera.	Center-Music
A	no problem! it can be hard to think of one. i like talking about the band Fleetwood Mac. Do you like Fleetwood Mac?	Music-KG
C	I recently read an interesting fact about the world's largest piano. The record for the largest piano ever made is held by Adrian Mann, a piano tuner from New Zealand. It took him four years to build the piano, and this grandiose instrument weighs 1.4 tonnes and is 5.7 meters long.	Center-Music
A	So tell me, why do you like music?	Music-Flow
C	I just read this fascinating fact about the history of Drums. Did you know that drums are one of the oldest musical instruments in the world?	Center-Music
C	Do you know the size of the world's smallest violin is? Well, I heard in Guangzhou, China, a 66-year-old senior violinist Chen Lianzhi made a 1-centimeter-long violin which sets the new world record for the world's smallest violin.	Center-Music
	None of the above.	

Table 1 Response pool annotation task example for Music. An "A" indicates an excellent response, while an"B" indicates a response that could be used, and a "C" indicates a response that might be okay in another context, while a "D" indicates a bad response.

V. Harrison, R. Rajasekaran, and M. Walker, "A Transformer-based Response Evaluator for Open-Domain Spoken Conversation," in *International Workshop on Spoken Dialogue Systems Technology (IWSDS)*, Los Angeles, CA, 2023. doi: 10.48550/arXiv.2302.04424.

Model evaluating itself



W. Shi *et al.*, "AutoReply: Detecting Nonsense in Dialogue with Discriminative Replies," in *Findings of the Association for Computational Linguistics: EMNLP*, Singapore: Association for Computational Linguistics, Dec. 2023, pp. 294–309. https://aclanthology.org/2023.findings-emnlp.23