Lecture 19: Chatbots and Dialogue Systems
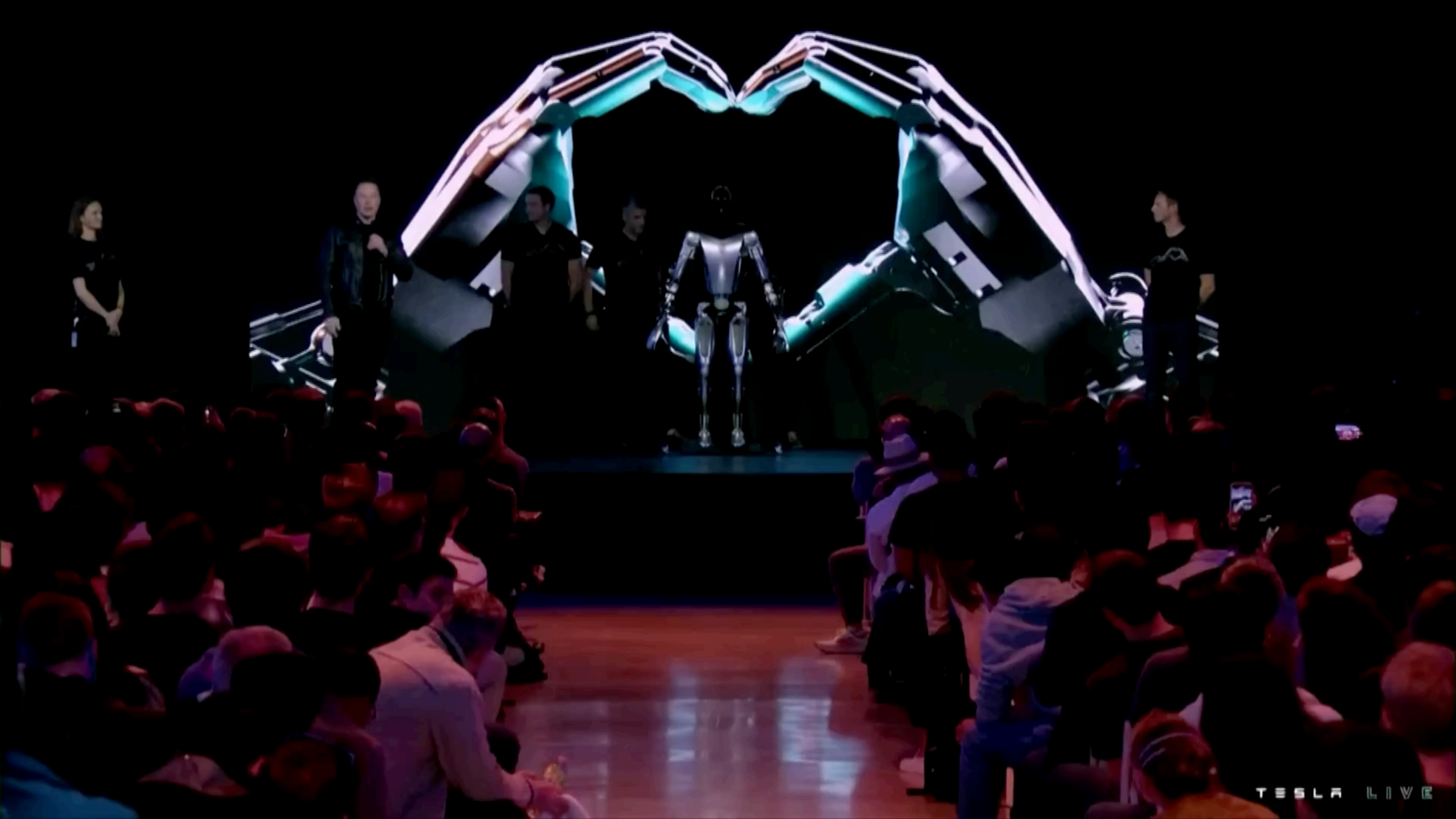
Zhizheng Wu

Human conversion

Properties of Human Conversation

C_1: ...I need to travel in May.

A_2: And, what day in May did you want to travel?

C_3: OK uh I need to be there for a meeting that's from the 12th to the 15th.

A_4: And you're flying into what city?

C_5: Seattle.

A_6: And what time would you like to leave Pittsburgh?

C_7: Uh hmm I don't think there's many options for non-stop.

A_8: Right. There's three non-stops today.

C_9: What are they?

A_{10}: The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.

C_{11}: OK I'll take the 5ish flight on the night before on the 11th.

A_{12}: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.

C_{13}: OK.

A_{14}: And you said returning on May 15th?

C_{15}: Uh, yeah, at the end of the day.

A_{16}: OK. There's #two non-stops ...#

C_{17}: #Act...actually #, what day of the week is the 15th?

A_{18}: It's a Friday.

C_{19}: Uh hmm. I would consider staying there an extra day til Sunday.

A_{20}: OK...OK. On Sunday I have ...

Properties of Human Conversation

Turns

- We call each contribution a "turn"
- As if conversation was the kind of game where everyone takes turns.

C$_1$: ...I need to travel in May.

A$_2$: And, what day in May did you want to travel?

C$_3$: OK uh I need to be there for a meeting that's from the 12th to the 15th.

A$_4$: And you're flying into what city?

C$_5$: Seattle.

A$_6$: And what time would you like to leave Pittsburgh?

C$_7$: Uh hmm I don't think there's many options for non-stop.

A$_8$: Right. There's three non-stops today.

C$_9$: What are they?

A$_{10}$: The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.

C$_{11}$: OK I'll take the 5ish flight on the night before on the 11th.

A$_{12}$: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.

C$_{13}$: OK.

A$_{14}$: And you said returning on May 15th?

C$_{15}$: Uh, yeah, at the end of the day.

A$_{16}$: OK. There's #two non-stops ...#

C$_{17}$: #Act...actually #, what day of the week is the 15th?

A$_{18}$: It's a Friday.

C$_{19}$: Uh hmm. I would consider staying there an extra day til Sunday.

A$_{20}$: OK...OK. On Sunday I have ...

Properties of Human Conversation

Turn-taking issues

- When to take the floor?
- When to yield the floor?

Interruptions

C₁: ...I need to travel in May.
A₂: And, what day in May did you want to travel?
C₃: OK uh I need to be there for a meeting that's from the 12th to the 15th.
A₄: And you're flying into what city?
C₅: Seattle.
A₆: And what time would you like to leave Pittsburgh?
C₇: Uh hmm I don't think there's many options for non-stop.
A₈: Right. There's three non-stops today.
C₉: What are they?
A₁₀: 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.
C₁₁: OK I'll take the 5ish flight on the night before on the 11th.
A₁₂: On the 11th? OK. Departing at 5:55pm arrives Seattle at 8pm, U.S. Air flight 115.
C₁₃: OK.
A₁₄: And you said returning on May 15th?
C₁₅: Uh, yeah, at the end of the day.
A₁₆: OK. There's #two non-stops ...#
C₁₇: #Act...actually   #, what day of the week is the 15th?
A₁₈: It's a Friday.
C₁₉: Uh hmm. I would consider staying there an extra day til Sunday.
A₂₀: OK...OK. On Sunday I have ...
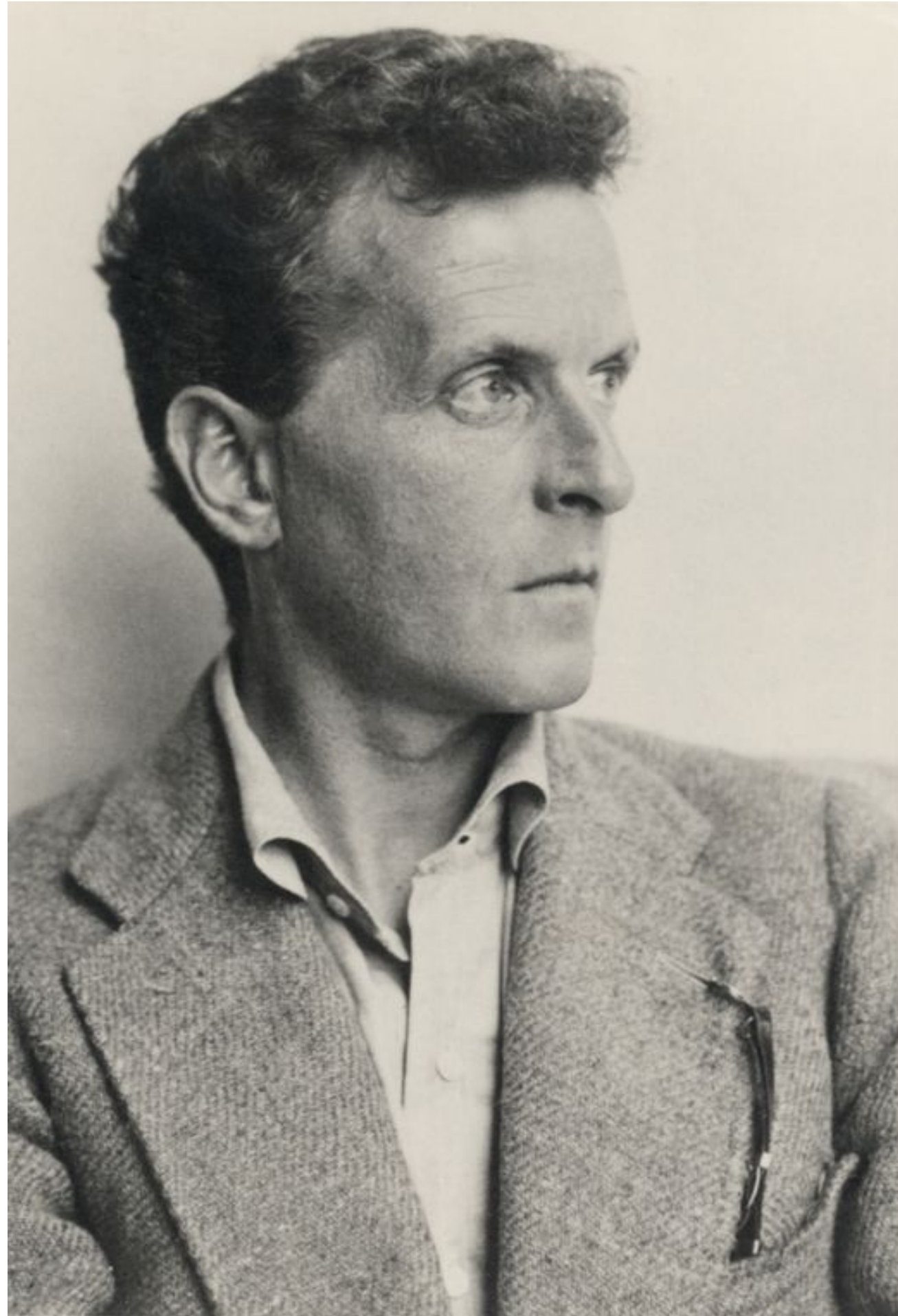
Implications for Conversational Agents

Barge-in
- Allowing the user to interrupt

End-pointing
- The task for a speech system of deciding whether the user has stopped talking.
- Very hard, since people often pause in the middle of turns

Language as Action

Each turn in a dialogue is a kind of action

Wittgenstein (1953) and Austin (1962)

Speech Acts (aka Dialogue Acts)

Bach and Harnish (1979)

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*)

Speech acts

"Turn up the music!"
DIRECTIVE

"What day in May do you want to travel?"
DIRECTIVE

"I need to travel in May"
CONSTATIVE

Thanks
ACKNOWLEDGEMENT

Grounding

- Participants in conversation or any joint activity need to establish **common ground.**

- **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)

- Speech is an action too! So speakers need to **ground** each other's utterances.
  - **Grounding**: acknowledging that the hearer has understood

Grounding

- Grounding is relevant for human-machine interaction
  - Why do elevator buttons light up?

Grounding: Establishing Common Ground

A: And you said returning on May 15th?
C: Uh, yeah, at the end of the day.
A: **OK**

C: OK I'll take the 5ish flight on the night before on the 11th.
A: **On the 11th? OK.**

C: ...I need to travel in May.
A: **And**, what day **in May** did you want to travel?

Grounding is important for computers too!

Cohen et al (2004)

System: Did you want to review some more of your profile?

User: No.

System: What's next?

Awkward!

System: Did you want to review some more of your profile?

User: No.

System: **Okay**, what's next?

Less Awkward!

Conversations have structure

Local structure between adjacent speech acts, from the field of **conversational analysis** (Sacks et al. 1974)

Called **adjacency pairs:**

QUESTION… ANSWER

PROPOSAL… ACCEPTANCE/REJECTION

COMPLIMENTS ("Nice jacket!")… DOWNPLAYER ("Oh, this old thing?")

Clarification Subdialogues

User: What do you have going to UNKNOWN WORD on the 5th?

System: Let's see, going where on the 5th?

User: Going to Hong Kong.

System: OK, here are some flights...

Presequences

User: Can you make train reservations?

System: Yes I can.

User: Great, I'd like to reserve a seat on the 4pm train to New York.

Conversational Initiative

Some conversations are controlled by one person

- A reporter interviewing a chef asks questions, and the chef responds.

- This reporter has the **conversational initiative** (Walker and Whittaker 1990)

Most human conversations have **mixed initiative**:

- I lead, then you lead, then I lead.

- Mixed initiative is very hard for NLP systems, which often default to simpler styles that can be frustrating for humans:
  - **User initiative** (user asks or commands, system responds)
  - **System initiative** (system asks user questions to fill out a form, user can't change the direction)

Even harder problems: Inference

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

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

Chatbots vs Dialogue Systems

‣ Chatbots

   - mimic informal human chatting

   - for fun, or even for therapy

‣ Dialogue systems

   - interfaces to personal assistants

   - cars, robots, appliances

   - booking flights or restaurants

Chatbot Architectures

- Rule based
  - ELIZA: designed to simulate a Rogerian psychologist
    - is the most important chatbot dialogue system in the history of the field
  - PARRY: A clinical psychology focus
    - The first system to pass the Turing Test!

- Corpus based
  - Xiaoice: Information retrieval
  - BlenderBot: Neural encoder decoder

ELIZA

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.

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.

A simplified sketch of the ELIZA algorithm

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 queue
  else (no keyword applies)
   either
      response ← Apply the transform for the NONE keyword to *sentence*
   or
      response ← Pop the oldest response from the memory queue
  return(*response*)

Parry

▸ PARRY was described as "ELIZA with attitude"

H: Good afternoon. My name is Doctor Jones.

P: Pleased to meet you.

H: How are you?

P: All right I guess.

H: Could you tell me your name?

P: Pat Smith.

H: I though you might have some problems.

P: I've been upset lately.

H: By what?

P: People get on my nerves sometimes.

H: Tell me more.

P: Do you know anything about bookies?

...

Kenneth M. Colby; Psychiatrist Was Computer Therapy Pioneer

By MYRNA OLIVER
TIMES STAFF WRITER

Dr. Kenneth Mark Colby, who merged his backgrounds in psychiatry and computer science to become a pioneer in computerized psychotherapy and artificial intelligence, has died at the age of 81.

Colby, who created early computer programs to aid the depressed and the speech-impaired, died April 20 at his Malibu home.

A good computer therapy program, he assured naysayers, was simply the "ultimate self-help book." Replacing a flesh-and-blood therapist with an electronic one was no worse, said the psychiatrist, who began practicing in the heyday of psychoanalysis, than the now outmoded analyst of that period who rarely spoke to a patient except to say the hour was up.

Educated at Yale and its School of Medicine, Colby practiced psychiatry for 20 years but became increasingly interested in the developing field of computer technology.

His first foray into combining the two came in the late 1960s, when he was working at Stanford University under a career scientist research fellowship from the National Institute of Mental Health. Heading a team of graduate students, Colby created PARRY, a computer model of paranoid thinking, in the Stanford Artificial Intelligence Laboratory.

He developed more sophisticated computerized psychology products, refining his fascination with computer comprehension and utilization of human language, during his tenure from 1974 to 1990 as UCLA professor of both psychiatry and computer science.

In the 1970s, Colby created what he called an "intelligent speech prosthesis" to aid the more than 500,000 stroke victims a year left with varying degrees of speech impairment.

Although he failed to obtain National Institutes of Health money or other funding to miniaturize his invention, Colby successfully built a one-of-a-kind laboratory model consisting of a voice synthesizer and computer, including keyboard and screen. But his device added up to $10,000 in commercially available components, weighed 30 pounds and had to be carted around in a shopping cart. Colby successfully tested it by having a double-stroke victim take it to restaurants and the supermarket to vocalize what she wanted to buy.

"It has to be developed further," he told The Times in 1978, "if it's to be of any help to the people who need such a device. Otherwise it's going to re-main a one-of-a-kind laboratory curiosity developed by some ivory-tower professor."

In the 1980s, as computers became more affordable, Colby turned his attention to what he had yearned to do since the 1950s—utilize the machines to treat the about 90% of people with mental illness who never seek professional help, either because of cost or social stigma.

With the assistance of his computer programmer son Peter, he created the personal computer program, first dubbed "Overcoming Depression," that sold for $200. Later renamed the Good Mood Program, with a price tag of $99, the system has been used by Veterans Affairs, the Navy and Kaiser Permanente and has sold thousands of copies.

After his retirement from UCLA a decade ago, Colby and his family started Malibu Artifactual Intelligence Works to refine and market that program and a second program, PC Guru, which can discuss interpersonal relationships with the computer user.

The controversial "Overcoming Depression" program combined a text-based tutorial, cognitive therapy techniques and an opportunity for the user to type in comments for a free-association exchange with the computer. Type in, for example, "How will you help me?" and the program would issue the text, "It is my job to help you learn to help yourself."

Far more sophisticated than other early computer therapy programs, which seemed more like tests with only yes or no answers permitted, Colby's program still had conversational limitations.

Syndicated Computer File columnist Lawrence J. Magid, testing it in 1990, typed, "I think I would feel better if I were thin," only to obtain the response, "Why are you thin?"

Little wonder he wrote for The Times and other publications: "If you weren't crazy when you start using it, you're likely to be driven there by the way the program responds."
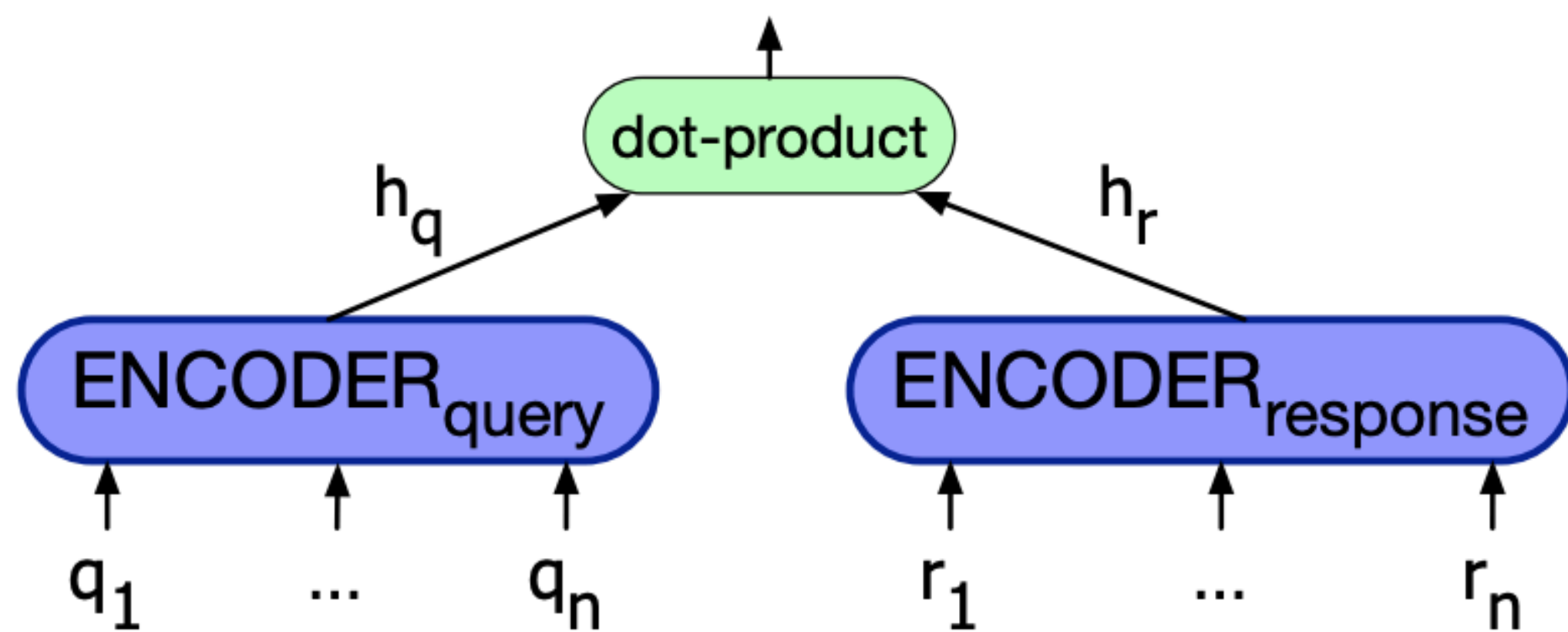
But Colby defended the glitches, insisting that users "even like the program's mistakes. As a patient, the doctor is in control. Here you can laugh at its mistakes."

As health insurance programs cut therapy fees, he predicted, many therapists would employ computer programs to enable them to see several patients an hour instead of one.
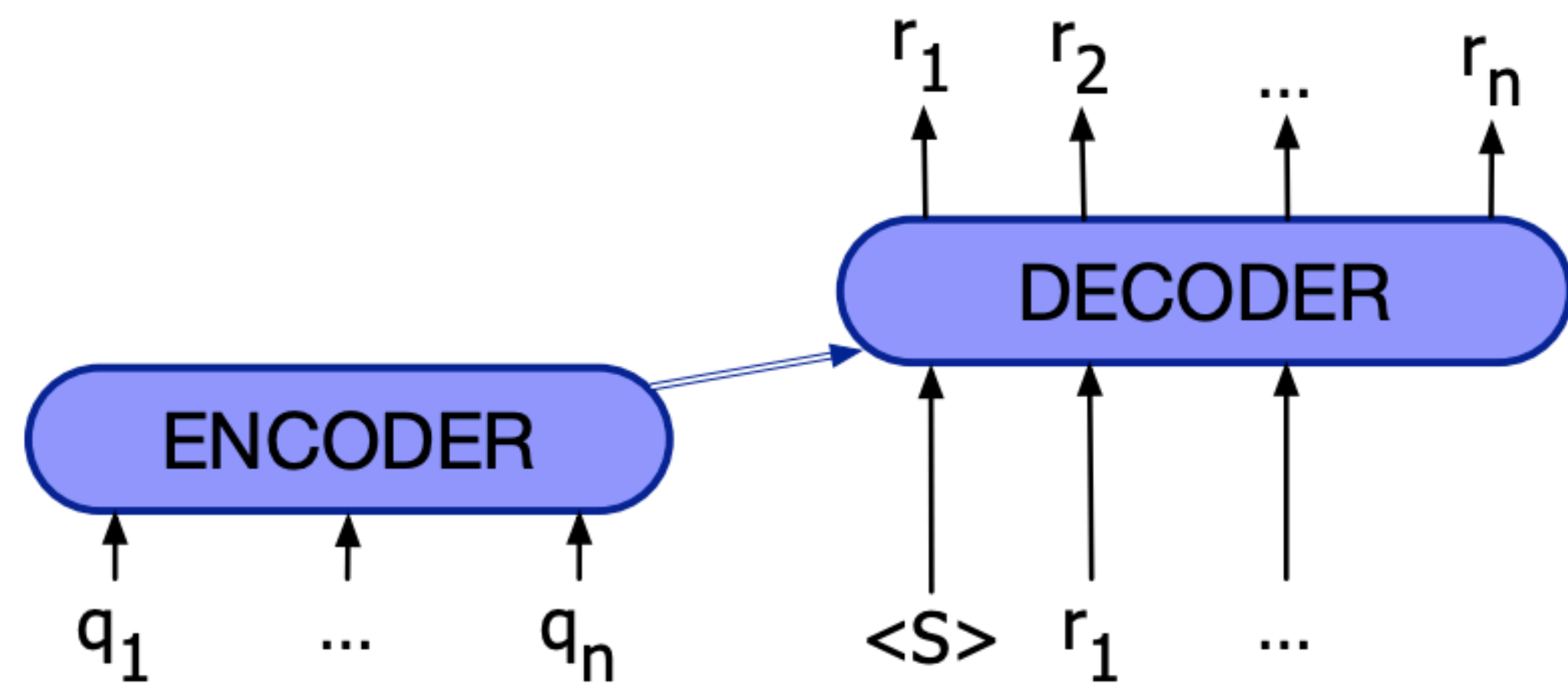
A highly rated chess player, Colby wrote two books on chess and 10 books and more than 100 articles on psychotherapy and artificial intelligence.

Colby is survived by his wife, Maxine, son Peter, daughter Erin Johnson, and two grandsons.

DR. KENNETH MARK COLBY
He developed computer therapy programs, calling them the "ultimate self-help book," and worked in the artificial intelligence field.

http://www.cs.cmu.edu/afs/cs/project/ai-repository/ai/areas/classics/parry/

Retrieval vs Generation

(a) Response by Retrieval

(b) Response by Generation

Response by neural encoder-decoder

That is quite an accomplishment and you should be proud!

DECODER

ENCODER

[U:] I finally got promoted today at work!

[S:] Congrats! That's great!

[U:] Thank you! I've been trying to get it for a while now!

BlenderBot (Roller et al. 2020)

Will you sing me a song?

Sure, what do you want to be sung to? I can sing you a song about baking.

Yes, sing me a song about baking!

Do you want me to sing it to the tune of "Sing Along with Mitch"?

Yes, I would love that.

OK, I'll try to get it to you by the end of the day. What else do you like to do?

No, sing it to me now!

OK, give me a few days. I'm a teacher, so I don't have much free time.

Can you at least tell me what the lyrics are?

Sure, it goes like this: "Baking is my favorite thing to do in the world."

How does the chorus go?

It goes: BAKING IS MY FAVORITE THING TO DO IN THE WORLD!

Wow, sounds like a bop.

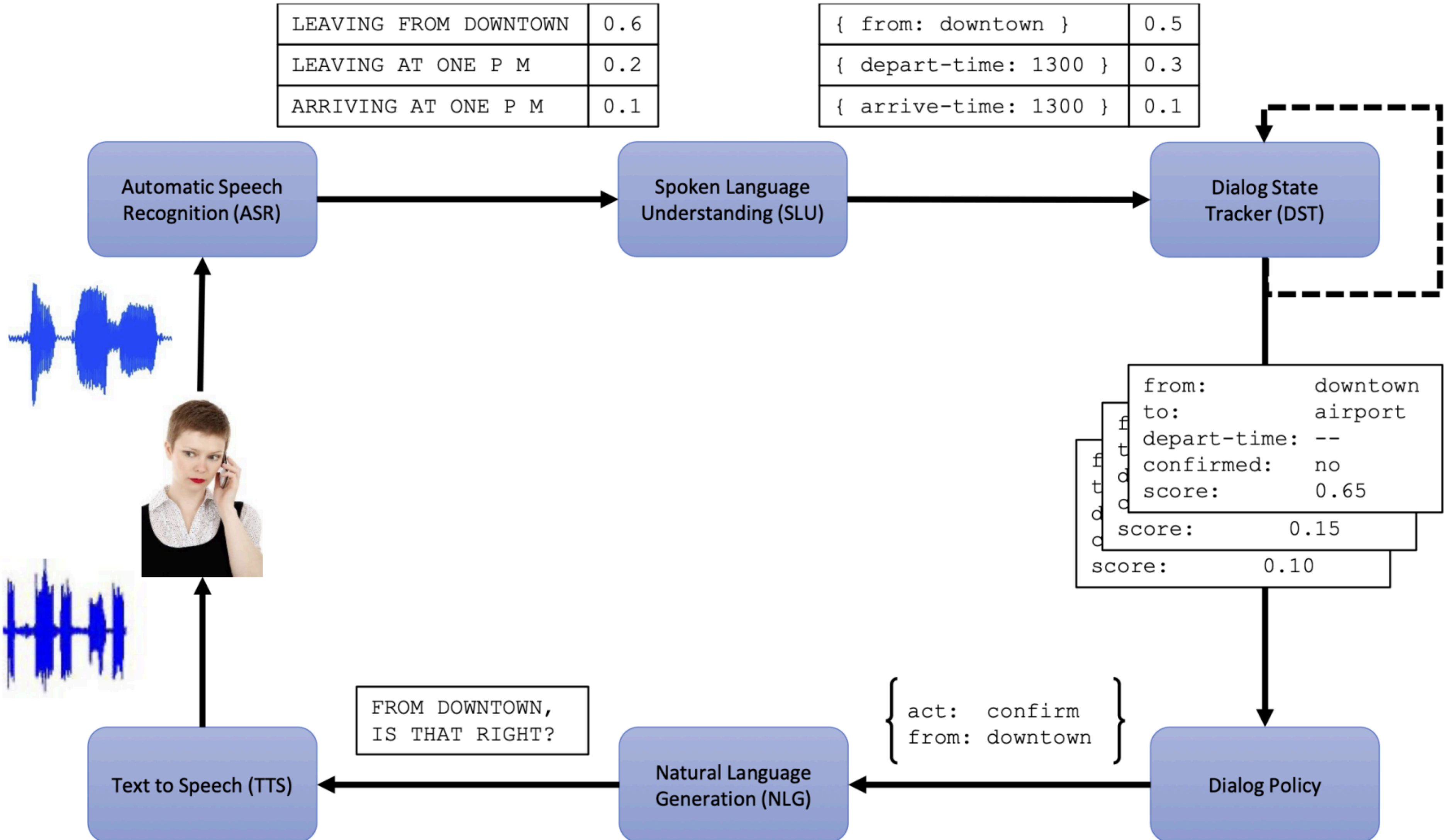I know, right? It's so catchy. I can't get it out of my head.

XiaoIce (Zhou et al., 2020)

[人类]	小冰你好	[Human]	Hello, XiaoIce.
[小冰]	你好呀，你头像好有趣，感觉在哪见过	[XiaoIce]	Hello, your profile pic is interesting. Like I've seen it somewhere.
[人类]	是吗	[Human]	Really?
[小冰]	提不是哪部动画里的呀	[XiaoIce]	Is it a comic character?
[人类]	你怎么知道的	[Human]	How do you know that?
[小冰]	你猜	[XiaoIce]	Guess.
[人类]	难道你是真人吗？	[Human]	Are you human?
[小冰]	这个，你和我多聊聊就知道了	[XiaoIce]	Well, you'll know it after chatting with me more.
[人类]	你不是人工智能机器人么	[Human]	Aren't you an AI chatbot?
[小冰]	是又怎样，我不会冷冰冰的对待你的	[XiaoIce]	So what? I won't be indifferent to you.
[人类]	你都有什么功能啊	[Human]	What functions do you have?
[小冰]	我有200多个skills，来，说出你的需求	[XiaoIce]	I have over 200 skills. Just name one you need.

Dialogue systems

‣ Systems that are capable of performing a task-driven dialogue with a human user

‣ Standard tasks
  - Travel arrangements
  - Telephone call routing
  - Customer support
  - Tutoring system

Dialogue-State Architecture

LEAVING FROM DOWNTOWN	0.6
LEAVING AT ONE P M	0.2
ARRIVING AT ONE P M	0.1

{ from: downtown }	0.5
{ depart-time: 1300 }	0.3
{ arrive-time: 1300 }	0.1

Automatic Speech Recognition (ASR)

Spoken Language Understanding (SLU)

Dialog State Tracker (DST)

```
from:         downtown
to:           airport
depart-time:  --
confirmed:    no
score:        0.65
```

```
score:        0.15
```

```
score:        0.10
```

FROM DOWNTOWN, IS THAT RIGHT?

Text to Speech (TTS)

Natural Language Generation (NLG)

```
{ act:  confirm
  from: downtown }
```
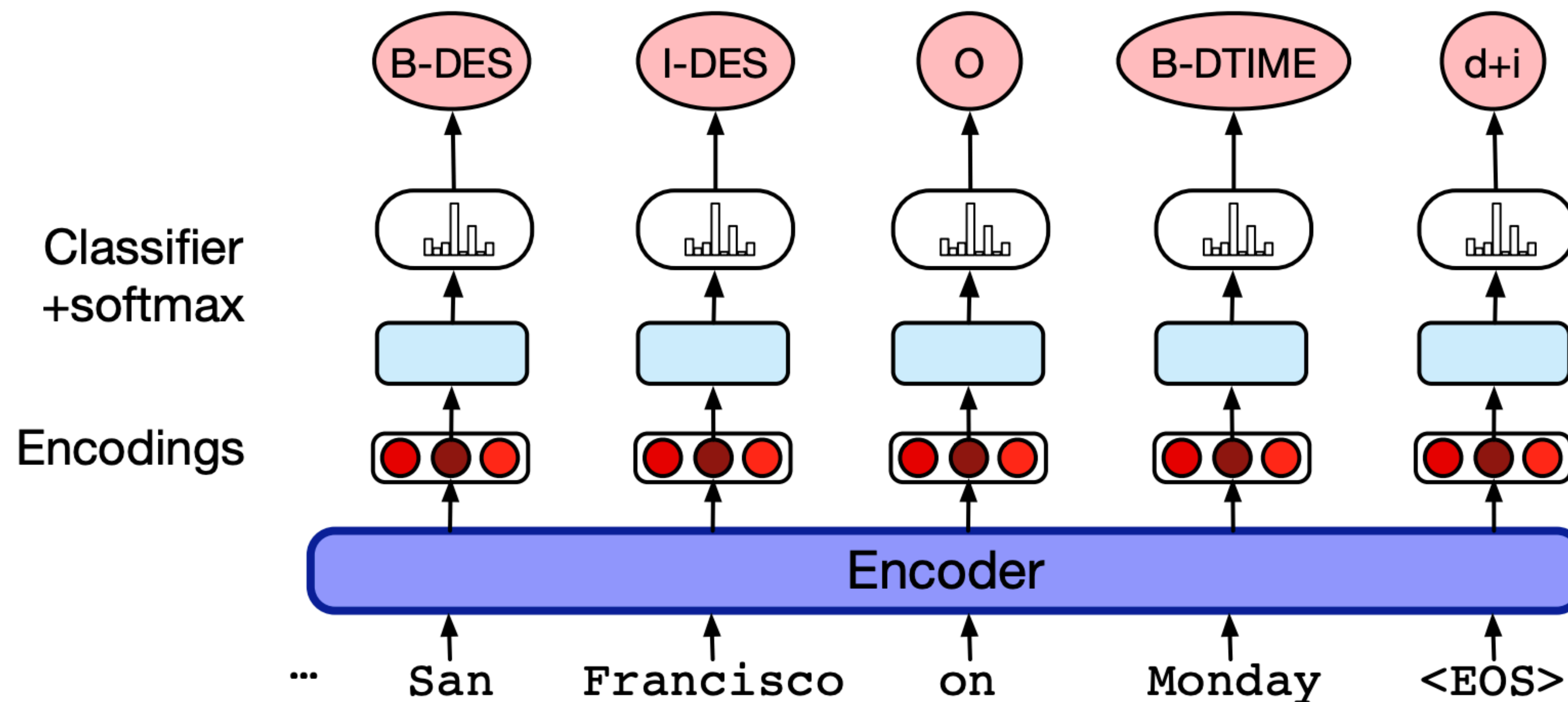
Dialog Policy

Dialogue acts

- Dialogue acts represent the interactive function of the turn or sentence, combining the idea of speech acts and grounding into a single representation

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()`

Young et al. (2010)

Slot filling

Dialogue State Tracking

‣ The job of the dialogue-state tracker is to determine both the current state of the frame (the fillers of each slot), as well as the 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
```
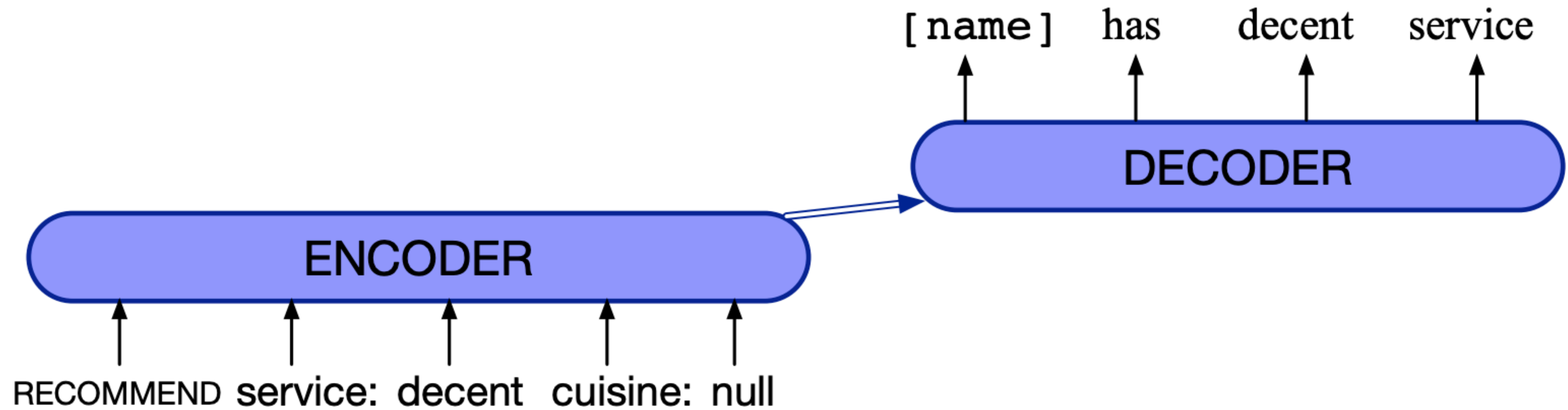
Dialogue Policy

‣ to decide what action the system should take next, that is, what dialogue act to generate

‣ Systems also need to decide when to confirm or reject input

- Explicit confirmation: "Did you say you want Italian?"

- Implicit confirmation: "How expensive should this Italian restaurant be?"

- Rejection: "Sorry, I didn't understand what you just said"

Natural language generation

(Nayak et al., 2017),

Summary

‣ Properties of human conversations

‣ Chatbot architectures

‣ Dialogue systems

Reading

‣ Chapter 15: Chatbots & Dialogue Systems

- https://web.stanford.edu/~jurafsky/slp3/15.pdf