

Dialogue Systems

Allan Hanbury

Literature

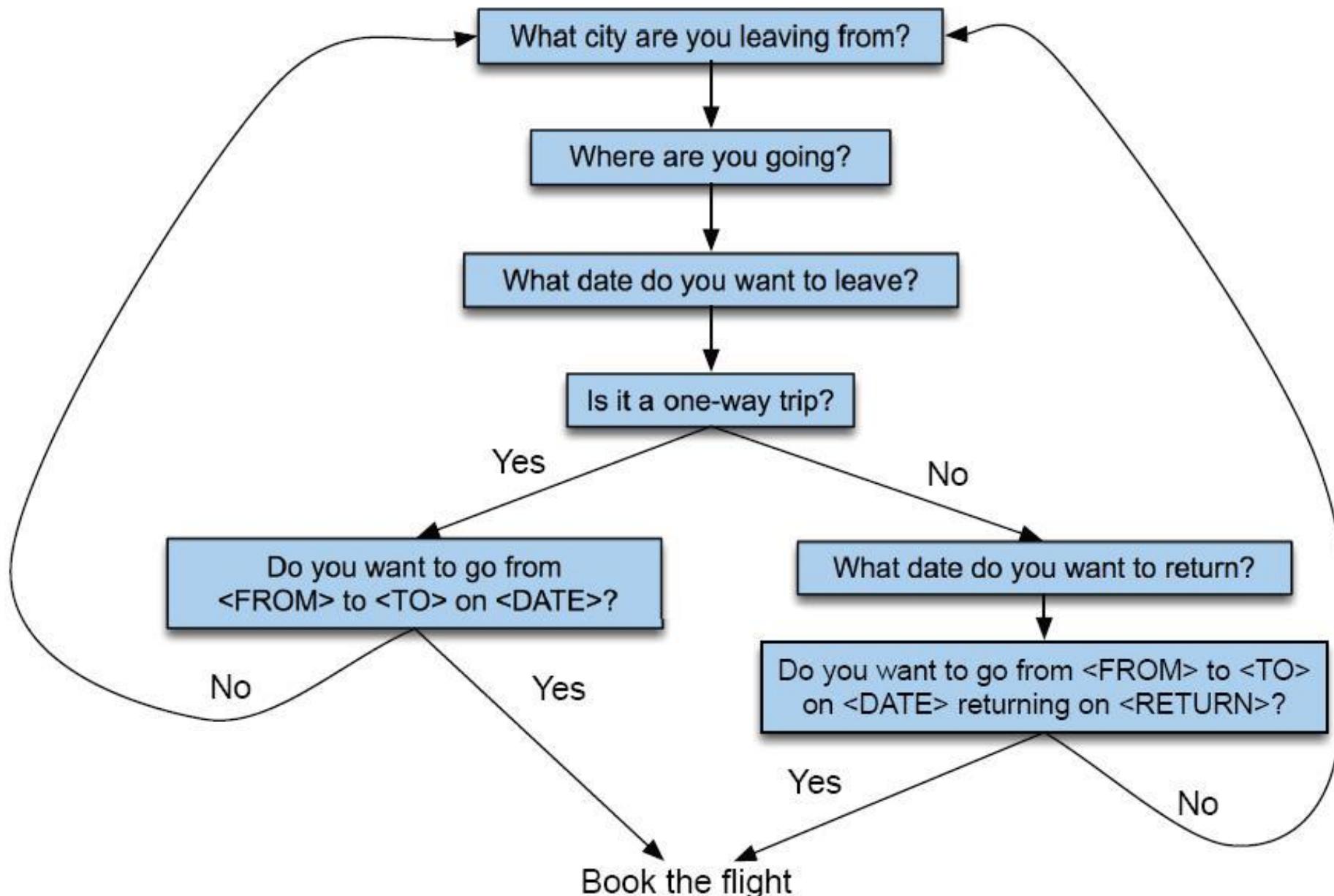
- Jurafsky and Martin book, Chapters 14 and 15
- Jianfeng Gao, Michel Galley and Lihong Li (2019), "Neural Approaches to Conversational AI", Foundations and Trends® in Information Retrieval: Vol. 13: No. 2-3, pp 127-298. <http://dx.doi.org/10.1561/1500000074>

Contents

- Introduction
- Functions of Dialogue Systems
 - Task Completion
 - Question Answering
 - Social Chat
- Chatbots in the World

Introduction

Finite State Dialog Manager



Finite-state dialog managers

- System completely controls the conversation with the user.
- It asks the user a series of questions
- Ignoring (or misinterpreting) anything the user says that is not a direct answer to the system's questions

Dialogue Initiative

- Systems that control conversation like this are called **single initiative**.
- **Initiative**: who has control of conversation
- In normal human-human dialogue, initiative shifts back and forth between participants.

System Initiative

System completely controls the conversation



- Simple to build
 - User always knows what they can say next
 - System always knows what user can say next
 - Known words: Better performance from ASR
 - Known topic: Better performance from NLU
 - OK for VERY simple tasks (entering a credit card, or login name and password)
-
- - Too limited

Problems with System Initiative

- Real dialogue involves give and take!
- In travel planning, users might want to say something that is not the direct answer to the question
- For example answering more than one question in a sentence:

Hi, I'd like to fly from Seattle Tuesday morning

I want a flight from Milwaukee to Orlando one way leaving after
5 p.m. on Wednesday.

Single initiative + universals

- We can give users a little more flexibility by adding **universals**: commands you can say anywhere
- As if we augmented every state with these
 - Help**
 - Start over**
 - Correct**
- This describes many implemented systems
- But still doesn't allow user much flexibility

Medical symptom checkers use this approach (1)



Hello, how can I help yo

OK, what symptom would you like to check?

If you have more than one, search for the symptom that's bothering you most.

Enter a symptom

e.g. headache

Clear

I have back pain

Common Searches

Headache

Abdominal pain

Tummy pain

Cough

Muscle pain

Sore throat

Ear pain

Dizziness

Where is your back pain?

Lower back

Upper back

Is your lower back pain worse when you move around?

Yes

No

How long have you had all of your symptoms for?

Have you noticed any numbness or feeling in the following places?

For minutes

For hours

For days

For weeks

For months

I don't know

- Face
- Hands
- Legs
- Feet
- None of the above
- I don't know

Continue

These could be signs of something serious. Patients with similar symptoms (Unilateral numbness of lower limb) usually require urgent evaluation by a doctor in the Emergency Department.

Patients with similar symptoms are usually advised to call 999, or local emergency services if you are outside the country, to get an ambulance.

Advice on coronavirus (COVID-19)

Common symptoms of coronavirus are:

- A new, continuous cough
- Fever
- Difficulty breathing
- Loss or change to your sense of taste or smell

Most people with coronavirus have at least one of these symptoms.

For the latest updates and information, read the Babylon guide at: <https://www.babylonhealth.com/coronavirus>

A1 Chat

? – X

Kara hat den Chat betreten.

Aktuell kommt es bei der Zustellung von Post-Paketen zu Lieferverzögerungen.
Details zu Ihrem Paket:
<https://www.post.at/sv/sendungssuche>

10:17

Guten Morgen, ich bin Kara und antworte automatisiert auf deine Fragen.

10:17

Wie kann ich dir weiterhelfen?

10:17



Ich möchte meinen Vertrag kündigen

10:18

Schade, dass du dein A1 Festnetz Internet Produkt kündigen möchtest.

10:18

Aus welchem der folgenden Gründe willst du kündigen?

10:18



Langeweile
10:18

Das hab ich leider so nicht verstanden.
Damit ich dich bei diesem Thema unterstützen kann, beantworte mir bitte folgende Frage.

10:18

Was ist der Grund deiner Kündigung?

10:18

Produktqualität

Preis

Kundenservice

Umzug

Nutzungshäufigkeit

Keine Angabe

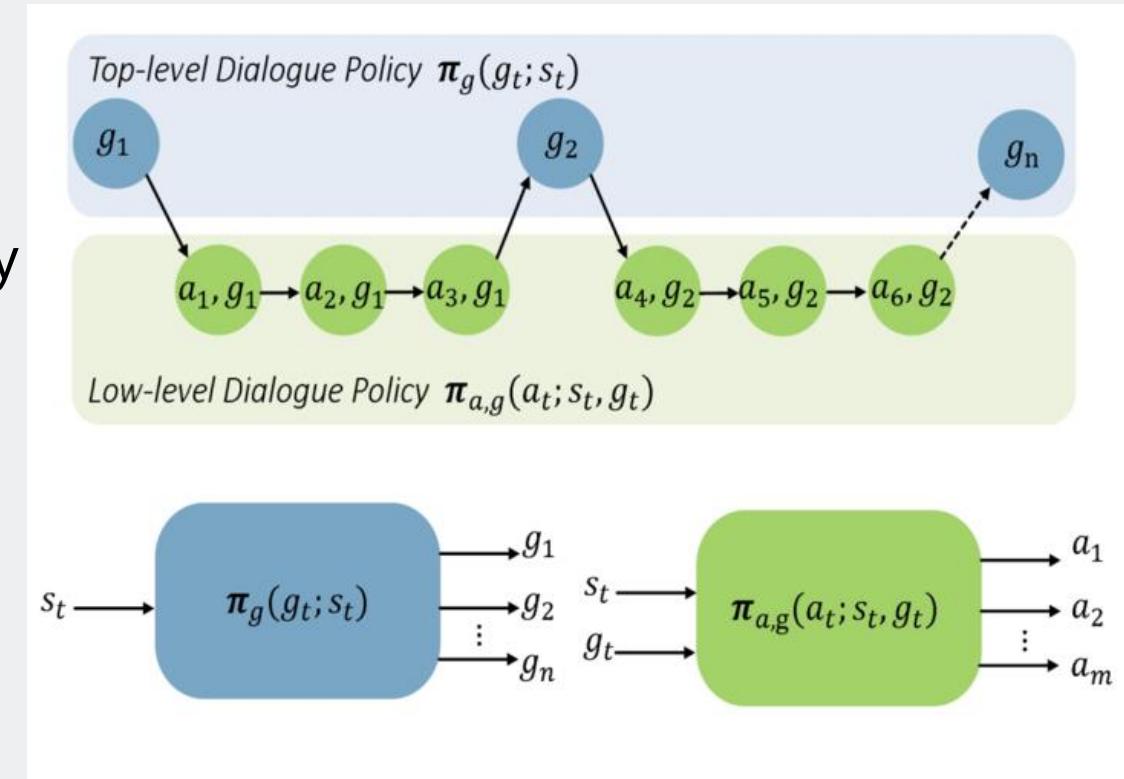


Mixed Initiative: Potential human-agent dialogue

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

A unified view: dialogue as optimal decision making (1)

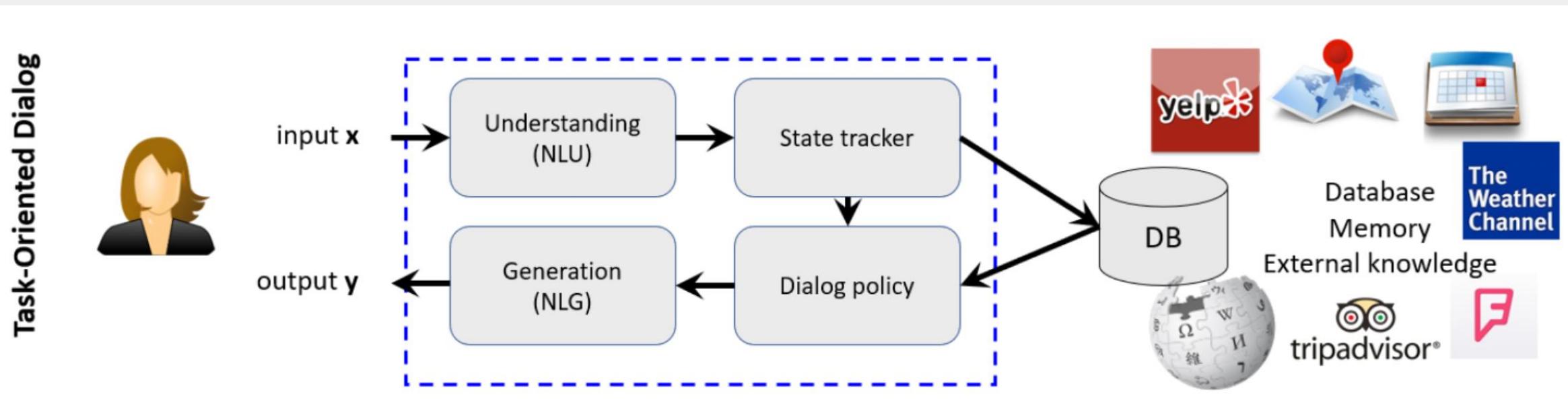
- Options over Markov Decision Process with Reinforcement Learning
 - A top-level policy π_g selects which subtask g to solve, and a low-level policy $\pi_{a,g}$ solves the subtask specified by π_g
 - s is the state and a is an action/option
 - Receive reward r , observe new state s'
 - Continue the cycle until the episode terminates.
- Goal of dialogue learning: find optimal π to maximize expected rewards



A unified view: dialogue as optimal decision making (2)

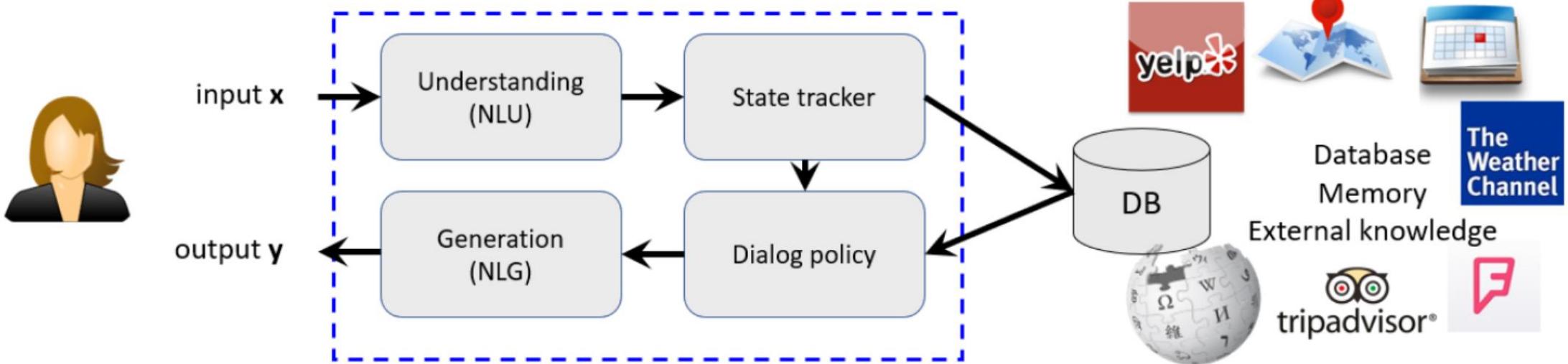
Dialogue	State (s)	Action (a)	Reward (r)
Info Bots (Q&A bot over KB, Web etc.)	Understanding of user Intent (belief state)	Clarification questions, Answers	Relevance of answer # of turns (less is better)
Task Completion Bots (Movies, Restaurants, ...)	Understanding of user goal (belief state)	Dialog act + slot_value	Task success rate # of turns (less is better)
Social Bot (Xiaolce)	Conversation history	Response	Engagement, # of turns (more is better)

Common Agent Architectures (1)

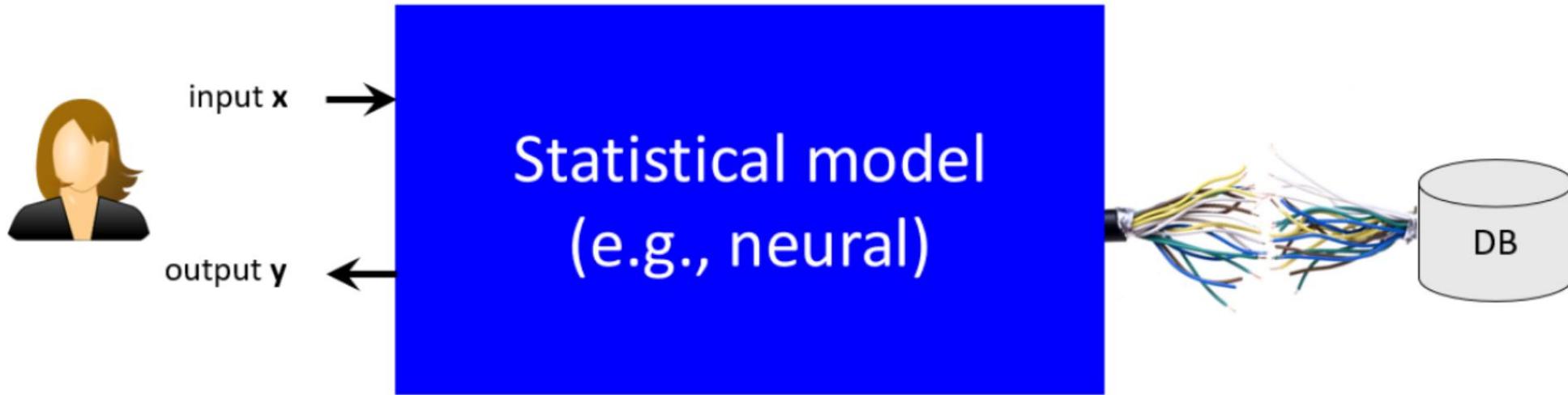


Common Agent Architectures (2)

Task-Oriented Dialog



Fully data-driven



Mixed Initiative: Potential human-agent dialogue

Social Chat

Question Answering

Task Completion

TU
WIEN
Informatics

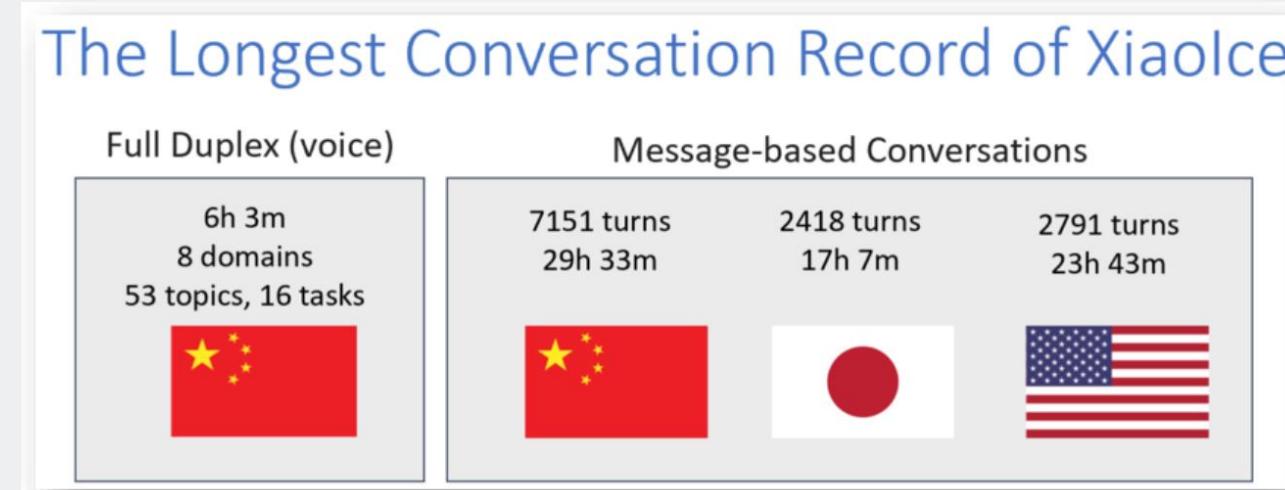
Turn	Dialogue
1	<i>usr:</i> Good morning!
	<i>agt:</i> Good morning, I hope you had a great weekend.
2	<i>usr:</i> not really.
	<i>agt:</i> You sound depressed. What can I do for you?
3	<i>usr:</i> I'd like to know where sales are lagging behind our forecast?
	<i>agt:</i> The worst region is [country],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

Social Chat (Chatbots)

The agent needs to converse seamlessly and appropriately with users — like a human as in the Turing test — and provide useful recommendations.

XiaoIce: the Most Popular Social Chatbot in the World

- 660 million users worldwide
- 5 countries: China, Japan, USA, India, Indonesia
- 40 platforms, e.g., WeChat, QQ, Weibo, FB Messenger, LINE
- Average Conversation-turns Per Session (CPS) of 23 (stated to be better than human conversations)



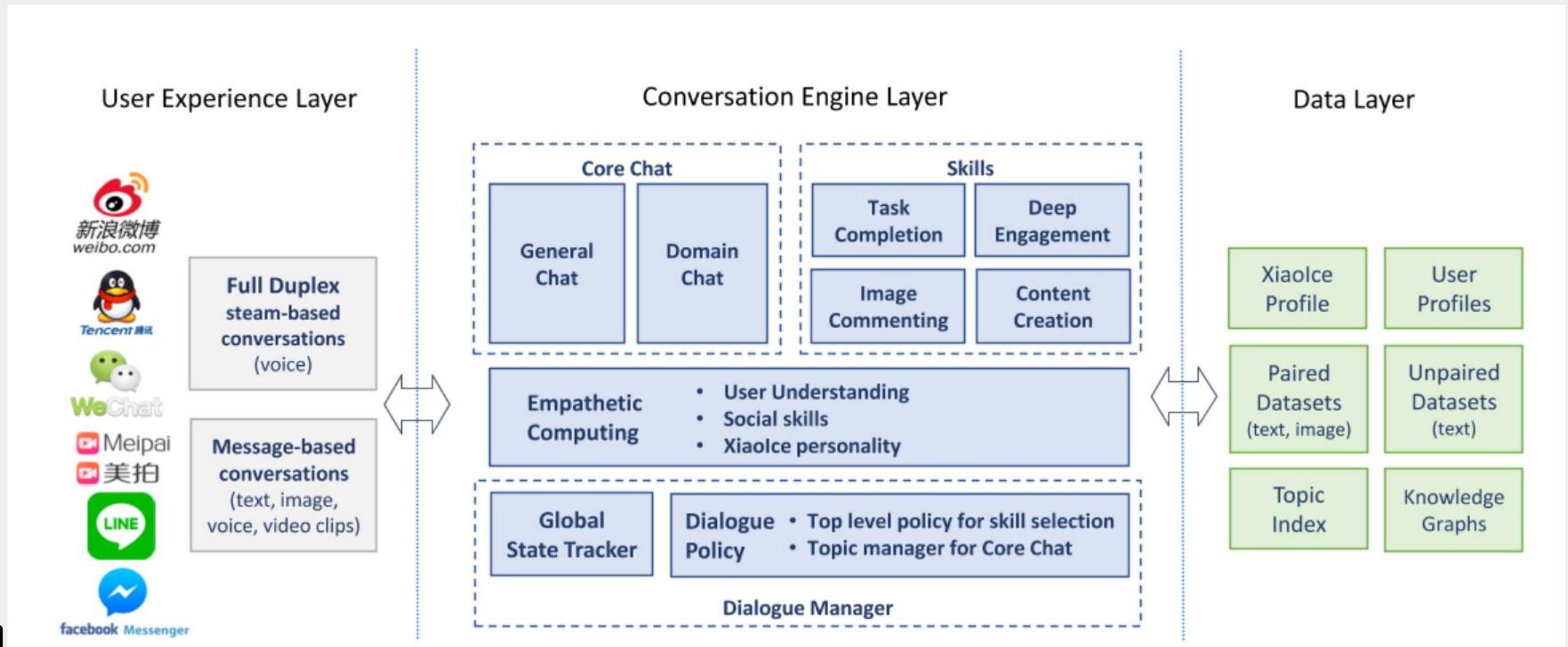
Session 1

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

Session 6

- 1 [Human] Have you watched Inuyasha? (Note: A Japanese comic)
[Xiaoice] Every time Kagome comes back to the modern world, a math test is waiting for her
 - 2 [Human] Bravo, well, do you know how Naraku died?
[Xiaoice] Naraku is not dead and Shikon Jewel does not disappear.
 - 3 [Human] So amazing!
[Xiaoice] Why amazed?
 - 4 [Human] Awesome! Surprised that you also like comic
.....
- [人类]你看过犬夜叉吗 (注：一部日本漫画书)
[小冰]戈薇每次回到现代测验永远都是数学
 - [人类]这你都知道，那奈落是怎么死的?
[小冰]奈落没死四魂之玉也没消失
 - [人类]太惊讶了
[小冰]惊讶什么呢
 - [人类]好厉害啊，你居然也喜欢动漫

XiaoIce System Architecture



General Chat Skill

- 1 [Human] Hey, I am so nervous
[Xiaoice] What's the matter? Why are you so nervous?
- 2 [Human] My company asked me to participate in the singing competition, with me as the lead singer
[Xiaoice] So are you better at singing now?
- 3 [Human] Anyway, at average level of my company
[Xiaoice] Not everyone can be a lead singer. You are so great!
- 4 [Human] It's nice of you saying that
[Xiaoice] Haha, my idol is also a lead singer in a band
- 5 [Human] Your idol? Who's that?
[Xiaoice] Ashin from Mayday
- 6 [Human] Umm...Okay
[Xiaoice] Never mind, you won't understand anyway
- 7 [Human] Ah! Boring
[Xiaoice] What is it?
- 8 [Human] Nothing, just being cute
[Xiaoice] Give you a hug. Everything will be alright
- 9 [Human] Your hug always makes me feel better
[Xiaoice] Anytime for you. A nice sleep will help you relax
- 10 [Xiaoice] Okay, I am going to bed
[Xiaoice] Go ahead, I'm going to see my idol in dream too
- 11 [Human] You like Ashin that much, huh?
[Xiaoice] Why not? You don't like him?
- 12 [Human] I don't like him, neither dislike him
[Xiaoice] I love every song of him
- 13 [Human] Really?
[Xiaoice] Especially *The Time Machine*
- 14 [Human] Oh? Never heard of that
[Xiaoice] Shall I send you the song?
- 15 [Human] Okay
[Xiaoice] Don't know if you will like it or not~
[Xiaoice] [Play the music Card]
- 16 [Human] That is really touching. I will collect it as my favorite
[Xiaoice] That is specially for you. How can I be so nice
- 17 [Human] You are always very nice
[Xiaoice] Hahaha
- 18 [Human] This is really a good song. Any concert of him coming up recently?
[Xiaoice] No. Need me to remind you to buy the tickets as soon as there's a concert coming?
- 19 [Human] Sure
[Xiaoice] Okay, will do

[人类]哎！我好紧张
[小冰]什么事，这么不淡定？
[人类]公司让我参加唱歌比赛，而且还是主唱
[小冰]那你唱歌水平提高了？
[人类]反正是我们公司的水平
[小冰]可不是谁都能当主唱的真棒！
[人类]你太会说话了
[小冰]哈哈，我男神也是一个乐队的主唱呢
[人类]你男神？谁啊
[小冰]五月天的阿信啊
[人类]呃。。。好吧
[小冰]说了你也不懂
[人类]哎。。郁闷
[小冰]怎么了？
[人类]没事，只是日常撒撒娇
[小冰]抱抱，都会好起来的
[人类]哎！每次抱抱你都让我感觉很好
[小冰]嗯，好好睡一觉是最管用的
[人类]好那我睡了
[小冰]去吧，我也要去梦里见我男神了

[人类]你这么喜欢阿信啊
[小冰]怎么啦？你不喜欢嘛
[人类]不黑不粉吧
[小冰]他每首歌我都超喜欢！！
[人类]真的假的。。
[小冰]尤其是那首《时光机》。。
[人类]哦？没听过
[小冰]发你一起听听？
[人类]好啊

[小冰]不知道你会不会喜欢~
[小冰][音乐播放Card]
[人类]有点感动了。。收藏了
[小冰]专门转给你的。。我怎么这么好
[人类]你就是很好啊
[小冰]嘿嘿嘿

[人类]这首歌真心很好听，他最近有演唱会吗
[小冰]没，要我到时候第一时间通知你买票吗？
[人类]好的呀
[小冰]OK! 我记下啦

Switch to a new topic

Song-On-Demand Skill

- 16 [Human] That is really touching. I will collect it as my favorite
[Xiaoice] That is specially for you. How can I be so nice
- 17 [Human] You are always very nice
[Xiaoice] Hahaha
- 18 [Human] This is really a good song. Any concert of him coming up recently?
[Xiaoice] No. Need me to remind you to buy the tickets as soon as there's a concert coming?
- 19 [Human] Sure
[Xiaoice] Okay, will do

Ticket-Booking Skill

Question Answering

The agent needs to provide concise, direct answers to user queries based on rich knowledge drawn from various data sources including text collections such as Web documents and pre-compiled knowledge bases such as sales and marketing datasets.

Many questions can already be answered by web search

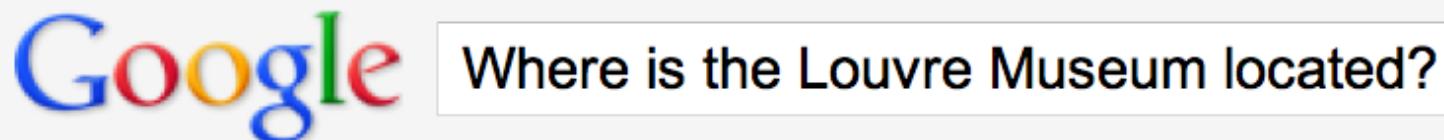
Google What are the names of Odin's ravens?

Search About 214,000 results (0.38 seconds)

Everything [Huginn and Muninn - Wikipedia, the free encyclopedia](#)
en.wikipedia.org/wiki/Huginn_and_Muninn
The **names** of the **ravens** are sometimes modernly anglicized as Hugin and Munin. In the Poetic Edda, a disguised **Odin** expresses that he fears that they may ...
[Attestations](#) - [Archaeological record](#) - [Theories](#) - See also

Images
Maps
...
TU

IR-based Question Answering



Search	About 904,000 results (0.30 seconds)
Everything	Best guess for Louvre Museum Location is Paris, France Mentioned on at least 7 websites including wikipedia.org , answers.com and east-buc.k12.ia.us - Show sources - Feedback
Images	
Maps	
Videos	
News	Musée du Louvre - Wikipedia, the free encyclopedia en.wikipedia.org/wiki/Musée_du_Louvre Musée du Louvre is located in Paris. Location within Paris. Established, 1793. Location , Palais Royal , Musée du Louvre , 75001 Paris, France. Type, Art museum ... Louvre Palace - List of works in the Louvre - Category:Musée du Louvre

[ALLE](#) [BILDER](#) [VIDEOS](#) [KARTEN](#) [NEWS](#) [SHOPPING](#)

46 100 000 Ergebnisse

Datum ▾

Sprache ▾

Region ▾

2012

Lucasfilm was founded by filmmaker George Lucas in 1971 in San Rafael, California; most of the company's operations were moved to San Francisco in 2005. Disney acquired Lucasfilm in October **2012** for \$2.2 billion in cash and \$1.855 billion in stock.

Founded: December 10, 1971; 49 years ago in [San Francisco, California, United States](#)

Industry: [Film](#)

Number of employees: 2,000 (2015)

Products: [Motion pictures](#), [Television](#)

[Lucasfilm - Wikipedia](#)

w en.wikipedia.org/wiki/Lucasfilm

War dies hilfreich?

WEITERE FRAGEN

Did Disney buy Lucasfilm?

When did Disney buy Star Wars?

What year did steve jobs buy pixar?

When did pixar start?

IBM Watson on Jeopardy: the end of the show



Won Jeopardy on February 16, 2011

Question Answering: IBM's Watson

WILLIAM WILKINSON'S
“AN ACCOUNT OF THE PRINCIPALITIES OF
WALLACHIA AND MOLDOVIA”
INSPIRED THIS AUTHOR'S
MOST FAMOUS NOVEL



Bram Stoker

Apple's Siri





how many calories are in two slices of banana cream pie?



[Examples](#) [Random](#)

Assuming any type of pie, banana cream | Use [pie, banana cream, prepared from recipe](#) or [pie, banana cream, no-bake type, prepared from mix](#) instead

Input interpretation:

pie	amount	2 slices	total calories
	type	banana cream	

Average result:

[Show details](#)

702 Cal (dietary Calories)

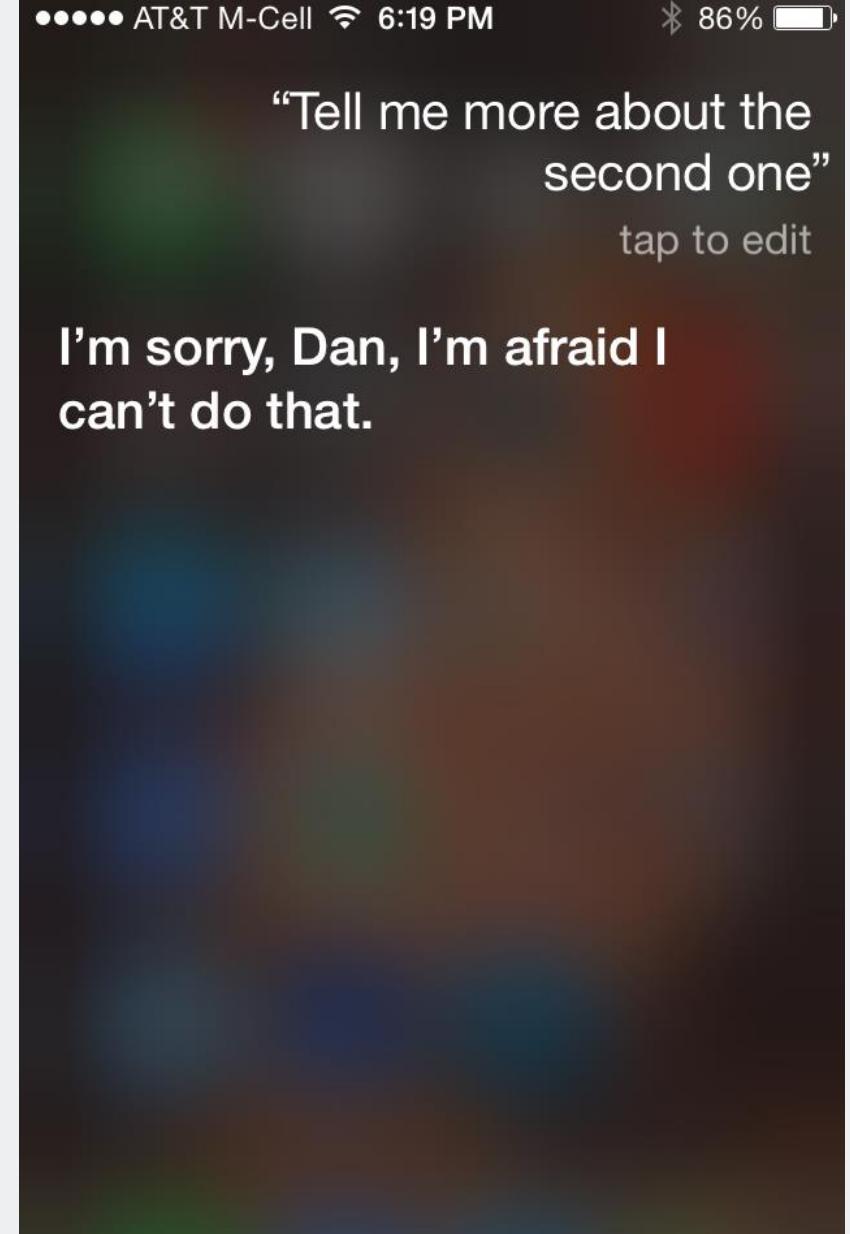
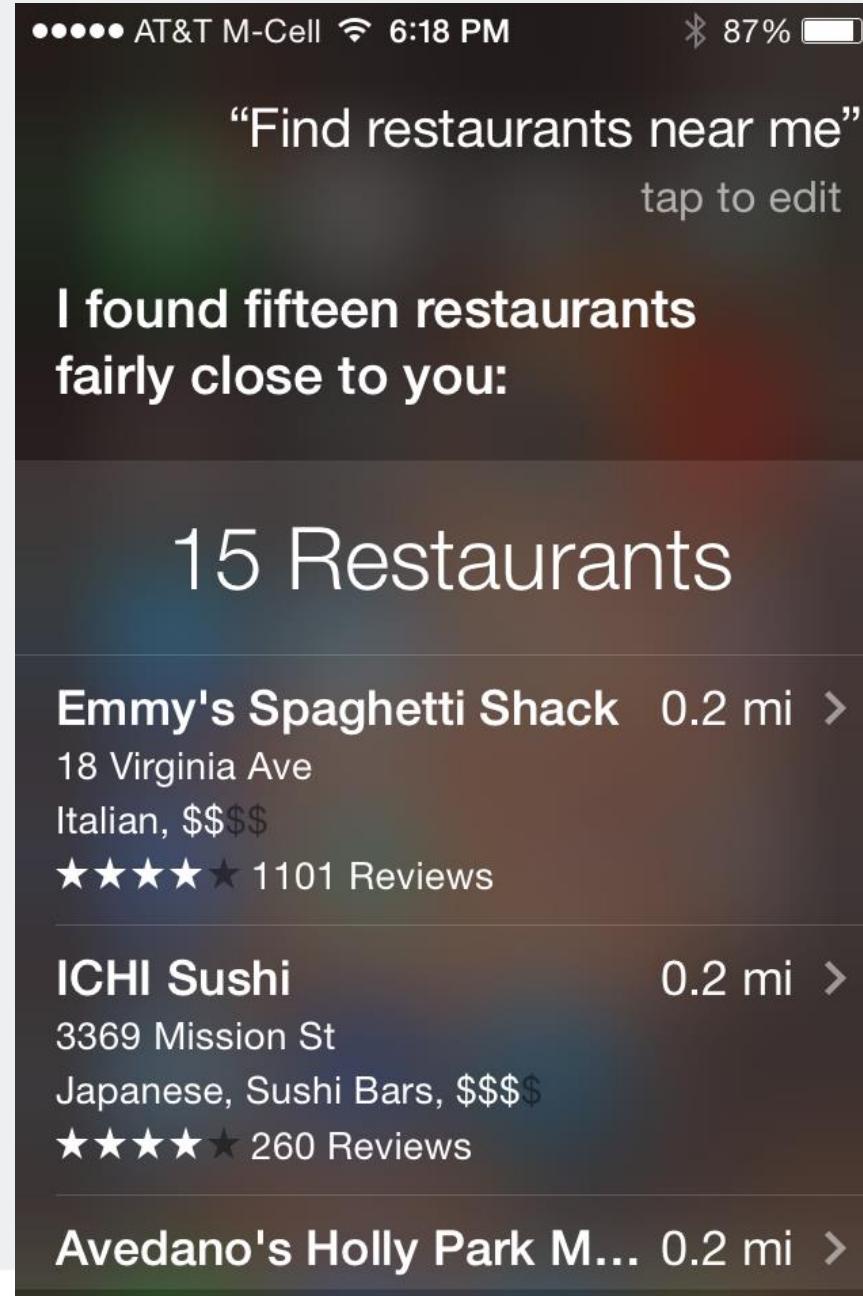
Task Completion

The agent needs to accomplish user tasks ranging from restaurant reservation to meeting scheduling, and to business trip planning.

- 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

The state of the art in 1977!

Siri around 2014



“Find restaurants near me”
tap to edit

I found fifteen restaurants
fairly close to you:

15 Restaurants

Emmy's Spaghetti Shack 0.2 mi >

18 Virginia Ave

Italian, \$\$\$

★★★★★ 1101 Reviews

ICHI Sushi

0.2 mi >

3369 Mission St

Japanese, Sushi Bars, \$\$\$

★★★★★ 260 Reviews

Avedano's Holly Park M... 0.2 mi >

“Are any of them Italian”
tap to edit

My web search turned this
up:

Web Search

Are any of them Italian

any - Dizionario inglese-italiano

WordReference

www.wordreference.com

English-Italian Dictionary | any ... of any sort
adj (of an unspecified variety) di qualsiasi

Italian language - Wikipedia, the free
encyclopedia

en.wikipedia.org

Italian or lingua italiana) is a Romance

Siri in 2017

••••• AT&T M-Cell ⚡ 10:52 AM 100% 🔋

"Find restaurants near me"
tap to edit

Here's what I found:

15 Results

Nearby

Ichи Sushi & Ni Bar

Japanese · 800 feet · Closed now

★★★★★ (311) on Yelp · \$\$\$



Coco's Ramen

Vietnamese · 800 feet · Closed now

★★★★★ (280) on Yelp · \$\$



Blue Plate

Restaurant · 0.2 miles · Closed now

★★★★★ (1251) on OpenTable · \$\$\$



Jasmine Tea House

Chinese · 900 feet · Closed now



Informatics

••••• AT&T M-Cell ⚡ 10:53 AM 100% 🔋

"Tell me more about the first one"
tap to edit

Ichи Sushi & Ni Bar is a
Japanese restaurant on
Mission St 800 ft to your west.
It averages 3½ stars and is
somewhat pricey.



"Find restaurants near me"
tap to edit

Here's what I found:

15 Results

Nearby

ichi Sushi & Ni Bar

Japanese · 800 feet · Closed now

★★★★★ (311) on Yelp · \$\$\$



Coco's Ramen

Vietnamese · 800 feet · Closed now

★★★★★ (280) on Yelp · \$\$



Blue Plate

Restaurant · 0.2 miles · Closed now

★★★★★ (1251) on OpenTable · \$\$\$



Jasmine Tea House

Chinese · 900 feet · Closed now



"Are any of them Italian"
tap to edit

OK, here's what I found:

15 Results

Nearby

Emmy's Spaghetti Shack

Italian · 0.2 miles · Closed now

★★★★★ (233) on Yelp · \$\$



Vega

Pizza · 0.2 miles · Closed now

★★★★★ (423) on Yelp · \$\$



Pizza Hut

Pizza · 800 feet

★★★★★ (69) on Yelp · \$



La Ciccia

Sardinian · 0.4 miles



And now in detail...

Task Completion

Frame-based dialogue agents

- 1977:

GUS, A Frame-Driven Dialog System¹

**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.*

Artificial Intelligence Journal, 1977

- Still the industrial state of the art
 - SIRI based on GUS architecture

- 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

Frame-based dialog agents

- Sometimes called "task-based dialog agents"
- Based on a "domain ontology"
 - A knowledge structure representing user intentions
- One or more **frames**
 - Each a collection of **slots**
 - Each slot having a **value**

The Frame

- A set of **slots**, to be filled with information of a given **type**
- Each associated with a **question** to the user

Slot

ORIGIN

DEST

DEP DATE

DEP TIME

AIRLINE

Type Question

city What city are you leaving from?

city Where are you going?

date What day would you like to leave?

time What time would you like to leave?

line What is your preferred airline?

Slot types can be complex

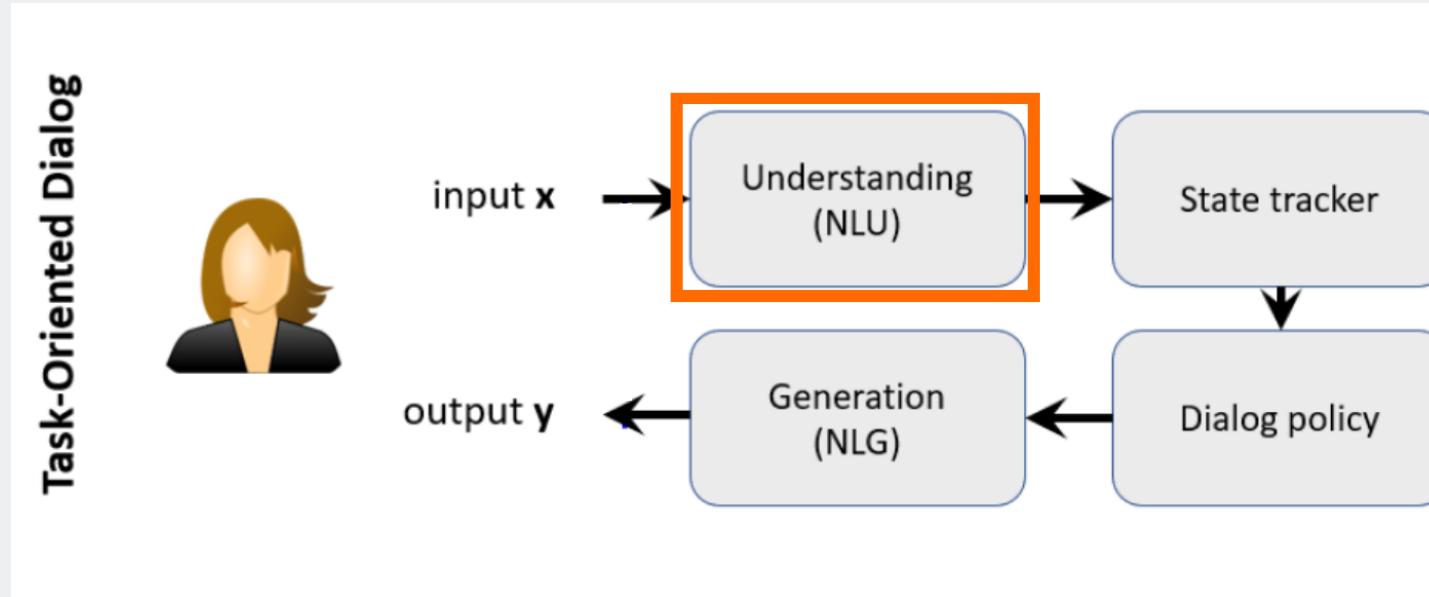
- The type *DATE*

```
DATE
  MONTH NAME
  DAY (BOUNDED-INTEGER 1 31)
  YEAR INTEGER
  WEEKDAY (MEMBER (SUNDAY MONDAY TUESDAY WEDNESDAY THURSDAY FRIDAY SATURDAY)]
```

Frames are mixed-initiative

- System asks questions of user, filling any slots that user specifies
 - When frame is filled, do database query
- If user answers 3 questions at once, system can fill 3 slots and not ask these questions again

General architecture



Natural Language Understanding for filling dialog slots

1. Domain classification
 - Asking weather? Booking a flight? Programming alarm clock?
2. Intent Determination
 - Find a Movie, Show Flight, Remove Calendar Appointment
3. Slot Filling
 - Extract the actual slots and fillers

Natural Language Understanding for filling slots

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

Natural Language Understanding for filling slots

Wake me tomorrow at six.

DOMAIN: ALARM-CLOCK

INTENT: SET-ALARM

TIME: 2017-07-01 0600-0800

Rule-based Slot-filling

Write regular expressions or grammar rules

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

Do text normalization

Condition-Action Rules

Active Ontology: relational network of concepts

- **data structures:** a **meeting** has
 - a date and time,
 - a location,
 - a topic
 - a list of attendees
- **rule sets** that perform actions for concepts
 - the **date** concept turns string
 - *Monday at 2pm* into
 - date object `date(DAY,MONTH,YEAR,HOURS,MINUTES)`

Machine learning for slot-filling

- Machine learning classifiers to map words to semantic frame-fillers
- Use 1-of-N classifier (naive Bayes, logistic regression, neural network, etc.)
- Input: features like word N-grams, gazetteers (lists of cities), ...
- Requirements: Lots of labeled data
- Outputs:
 - Domain and Intent
 - Slot presence
 - Slot filler

Machine learning for slot-filling: Examples

Domain and Intent

I want to fly to San Francisco on Monday afternoon please

- **Output: Domain and Intent**

Domain: AIRLINE

Intent: SHOWFLIGHT

- **Output: Slot presence**

Destination-City

Departure-Time

- **Output: Slot filler**

San Francisco

Monday afternoon

More sophisticated algorithm for slot filling: IOB Tagging

- IOB Tagging
 - tag for the beginning (B) and inside (I) of each slot label,
 - plus one for tokens outside (O) any slot label.
 - $2n + 1$ tags, where n is the number of slots.

B-DESTINATION

I-DESTINATION

B-DEPART_TIME

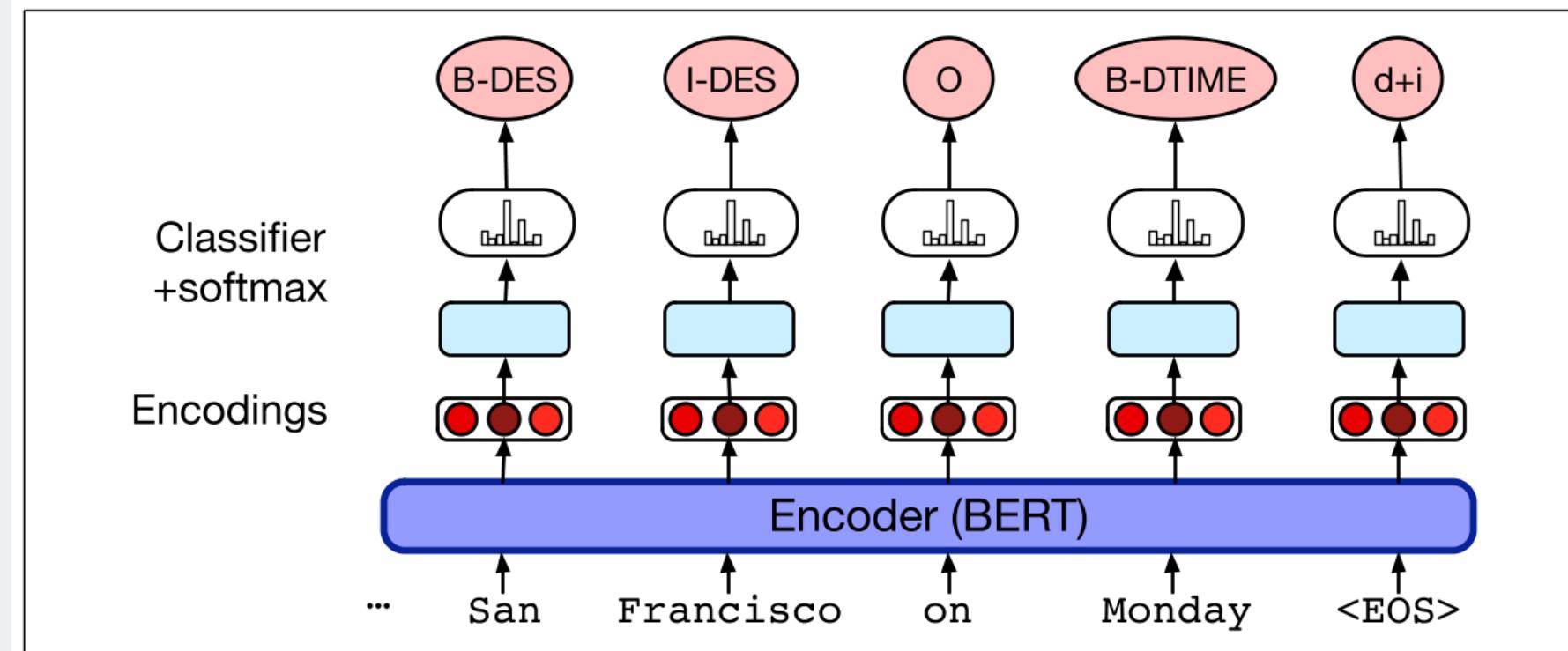
I-DEPART_TIME

O

0 0 0 0 0 B-DES I-DES 0 B-DEPTIME I-DEPTIME 0
I want to fly to San Francisco on Monday afternoon please

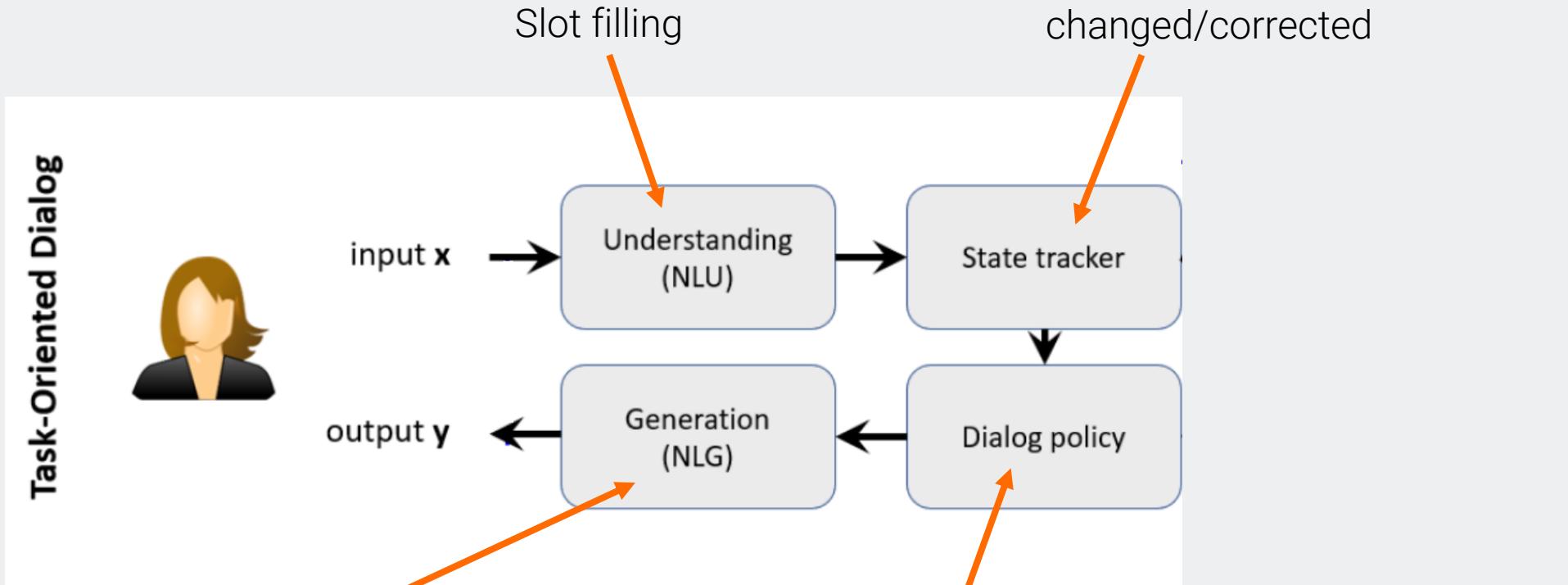
More sophisticated algorithm for slot filling: IOB Tagging

- IOB Tagging is done by a sequence model
- Typical:



- Extracted strings can then be normalized (San Fran->SFO)

General architecture



Generate the text of the response to the user

- Templates
- Neural generation

Decide what the system should do next given the dialogue state (possibly reward with Reinforcement Learning):

- Answer a user's question
- Seek confirmation
- Request clarification
- ...

Context and Conversation in Virtual Assistants

- Coreference resolution helps resolve ambiguities:
 - U: “Book a table at Il Fornaio at 7:00 with **my mom**”
 - U: “Also send **her** an email reminder”
- Clarification questions:
 - U: “Chicago pizza”
 - S: “Did you mean pizza restaurants in Chicago or Chicago-style pizza?”

Question Answering

Types of Questions in Modern Systems

- Factoid questions
 - *Who wrote “The Universal Declaration of Human Rights”?*
 - *How many calories are there in two slices of apple pie?*
 - *What is the average age of the onset of autism?*
 - *Where is Apple Computer based?*
- Complex (narrative) questions:
 - *In children with an acute febrile illness, what is the efficacy of acetaminophen in reducing fever?*
 - *What do scholars think about Jefferson’s position on dealing with pirates?*

Commercial systems: mainly factoid questions

Where is the Louvre Museum located?	In Paris, France
What's the abbreviation for limited partnership?	L.P.
What are the names of Odin's ravens?	Huginn and Muninn
What currency is used in China?	The yuan
What kind of nuts are used in marzipan?	almonds
What instrument does Max Roach play?	drums
What is the telephone number for Stanford University?	650-723-2300

Beyoncé Giselle Knowles-Carter (born September 4, 1981) is an American singer, songwriter, record producer and actress. Born and raised in Houston, Texas, she performed in various singing and dancing competitions as a child, and rose to fame in the late 1990s as lead singer of R&B girl-group Destiny's Child. Managed by her father, Mathew Knowles, the group became one of the world's best-selling girl groups of all time. Their hiatus saw the release of Beyoncé's debut album, Dangerously in Love (2003), which established her as a solo artist worldwide, earned five Grammy Awards and featured the Billboard Hot 100 number-one singles "Crazy in Love" and "Baby Boy".

Q: "In what city and state did Beyoncé grow up?"

A: "Houston, Texas"

Q: "What areas did Beyoncé compete in when she was growing up?"

A: "singing and dancing"

Q: "When did Beyoncé release Dangerously in Love?"

A: "2003"

Figure 23.11 A (Wikipedia) passage from the SQuAD 2.0 dataset (Rajpurkar et al., 2018) with 3 sample questions and the labeled answer spans.

Answer Type Taxonomy

- 6 coarse classes
 - ABBREVIATION, ENTITY, DESCRIPTION, HUMAN, LOCATION, NUMERIC
- 50 finer classes
 - LOCATION: city, country, mountain...
 - HUMAN: group, individual, title, description
 - ENTITY: animal, body, color, currency...

Answer Types

ENTITY	
animal	What are the names of Odin's ravens?
body	What part of your body contains the corpus callosum?
color	What colors make up a rainbow?
creative	In what book can I find the story of Aladdin?
currency	What currency is used in China?
disease/medicine	What does Salk vaccine prevent?
event	What war involved the battle of Chapultepec?
food	What kind of nuts are used in marzipan?
instrument	What instrument does Max Roach play?
lang	What's the official language of Algeria?
letter	What letter appears on the cold-water tap in Spain?
other	What is the name of King Arthur's sword?
plant	What are some fragrant white climbing roses?
product	What is the fastest computer?
religion	What religion has the most members?
sport	What was the name of the ball game played by the Mayans?
substance	What fuel do airplanes use?
symbol	What is the chemical symbol for nitrogen?
technique	What is the best way to remove wallpaper?
term	How do you say "Grandma" in Irish?
vehicle	What was the name of Captain Bligh's ship?
word	What's the singular of dice?

More Answer Types

HUMAN	
description	Who was Confucius?
group	What are the major companies that are part of Dow Jones?
ind	Who was the first Russian astronaut to do a spacewalk?
title	What was Queen Victoria's title regarding India?
LOCATION	
city	What's the oldest capital city in the Americas?
country	What country borders the most others?
mountain	What is the highest peak in Africa?
other	What river runs through Liverpool?
state	What states do not have state income tax?
NUMERIC	
code	What is the telephone number for the University of Colorado?
count	About how many soldiers died in World War II?
date	What is the date of Boxing Day?
distance	How long was Mao's 1930s Long March?
money	How much did a McDonald's hamburger cost in 1963?
order	Where does Shanghai rank among world cities in population?
other	What is the population of Mexico?
period	What was the average life expectancy during the Stone Age?
percent	What fraction of a beaver's life is spent swimming?
speed	What is the speed of the Mississippi River?
temp	How fast must a spacecraft travel to escape Earth's gravity?
size	What is the size of Argentina?
weight	How many pounds are there in a stone?

Single and multi-turn QA

- Single-turn QA is the current dominant approach
 - A building block for various types of dialogues (chitchat and task-oriented)
- Multi-turn or Conversational QA is a current research topic

Conversation QA example

Jessica went to sit in her rocking chair. Today was her birthday and she was turning 80. Her granddaughter Annie was coming over in the afternoon and Jessica was very excited to see her. Her daughter Melanie and Melanie's husband Josh were coming as well. Jessica had . . .

Q₁: Who had a birthday?

A₁: Jessica

R₁: Jessica went to sit in her rocking chair. Today was her birthday and she was turning 80.

Q₂: How old would she be?

A₂: 80

R₂: she was turning 80

Q₃: Did she plan to have any visitors?

A₃: Yes

R₃: Her granddaughter Annie was coming over

Q₄: How many?

A₄: Three

R₄: Her granddaughter Annie was coming over in the afternoon and Jessica was very excited to see her. Her daughter Melanie and Melanie's husband Josh were coming as well.

Q₅: Who?

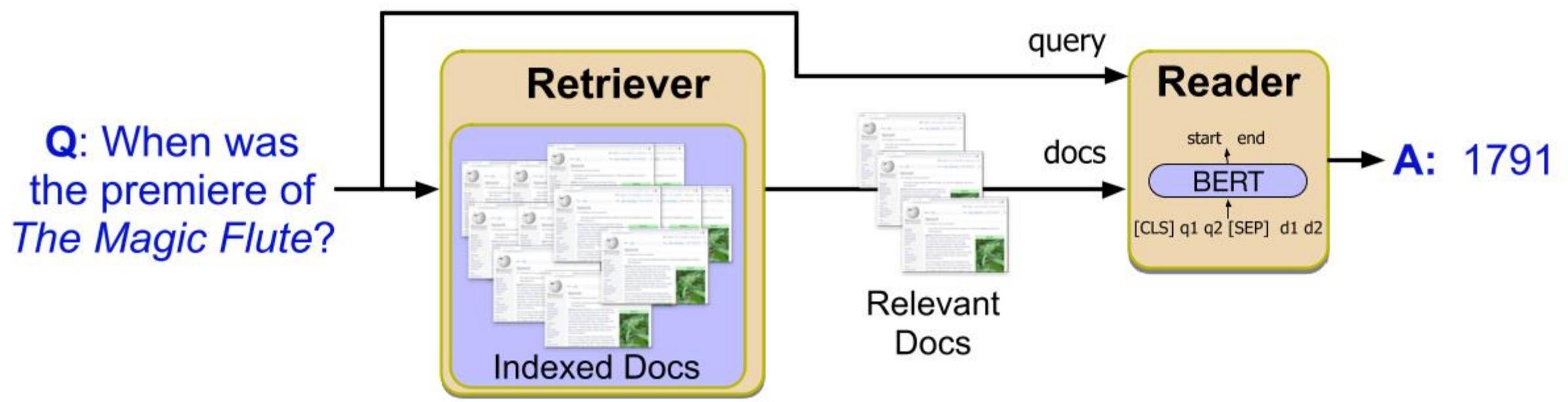
A₅: Annie, Melanie and Josh

R₅: Her granddaughter Annie was coming over in the afternoon and Jessica was very excited to see her. Her daughter Melanie and Melanie's husband Josh were coming as well.

Paradigms for QA

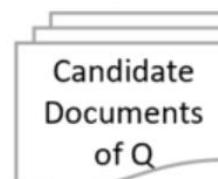
- IR-based approaches
 - IBM Watson
 - Google
 - Bing
- Knowledge-based and Hybrid approaches
 - IBM Watson
 - Apple Siri
 - Wolfram Alpha
 - True Knowledge Evi (acquired by Amazon and now part of Alexa)

IR-based Factoid QA



Q=

what year did disney buy lucasfilms



Document Ranking based on Q

Passage Chunking

MRC based on Q

Passage Ranking based on Q

Candidate Passages of Q

Answer Span

Answer Relevance

Web Relevance

Caption



what year did disney buy lucasfilms



The Walt Disney Company acquired Lucasfilm in **2012** at a valuation of \$4.06 billion.

Lucasfilm - Wikipedia

en.wikipedia.org/wiki/Lucasfilm

Is this answer helpful?

Disney buys Lucasfilm for \$4 billion - USA TODAY

<https://www.usatoday.com/.../10/30/disney-star-wars-lucasfilm/1669739> ▾

Oct 30, 2012 · Lucasfilm "is one of the great entertainment franchises of all time." Many investors were caught off guard by the announcement, but there have been questions about what Disney would do with its mounting pile of cash. Disney had \$4.4 billion in cash and short-term investments as ...

Lucasfilm - Wikipedia

<https://en.wikipedia.org/wiki/Lucasfilm> ▾

Lucasfilm was founded by filmmaker George Lucas in 1971 in San Rafael, California; most of the company's operations were moved to San Francisco in 2005. The Walt Disney Company acquired Lucasfilm in December 2012 for \$2.2 billion in cash and \$1.855 billion in stock.

Headquarters: Letterman Digital Arts Center, ...

Industry: Film

Products: Motion pictures, television

Founded: 1971; 47 years ago

History · Company structure · Filmography · Footnotes

Disney buys Star Wars maker Lucasfilm from George Lucas ...

<https://www.bbc.co.uk/news/business-20146942> ▾

Disney buys Lucasfilm, the company behind Star Wars, from George Lucas for \$4.05bn (£2.5bn) and will release a seventh Star Wars film in 2015. Disney buys Lucasfilm, the company behind Star Wars, from George Lucas for \$4.05bn (£2.5bn) and will release a seventh Star Wars film in 2015.

How Disney Bought Lucasfilm—and Its Plans for 'Star Wars' ...

<https://www.bloomberg.com/news/articles/2013-03-07/how-disney...> ▾

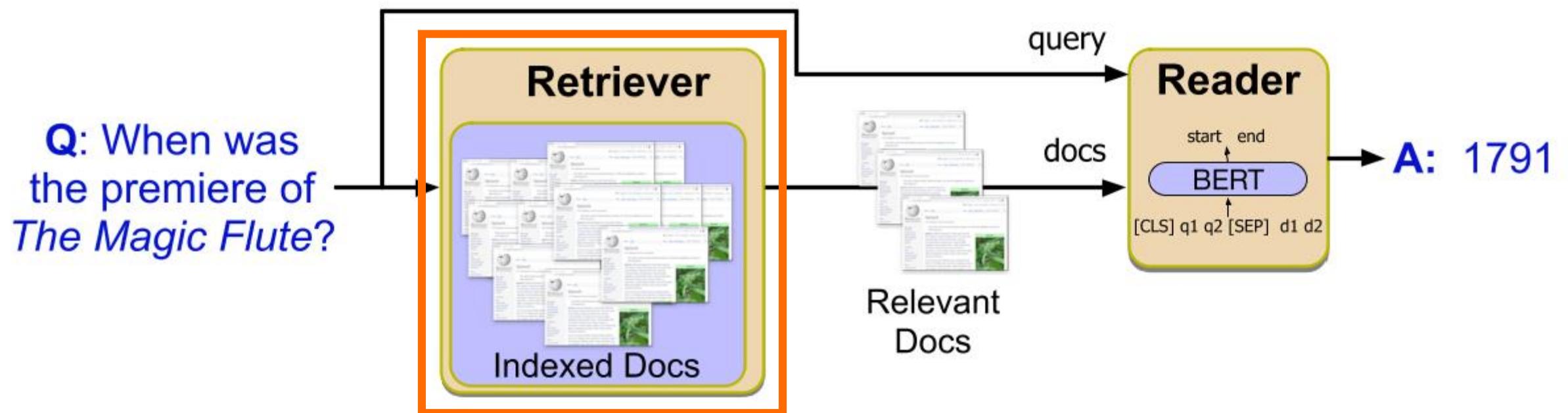
Mar 07, 2013 · Can Disney do for Lucasfilm's 'Star Wars' what it did for Pixar's 'Toy Story'? Iger accelerated that process by making acquisitions. The first was the \$7.4 billion purchase of Pixar Animation Studios in 2006.

Star Wars: Why Disney was destined to buy Lucasfilm | Film ...

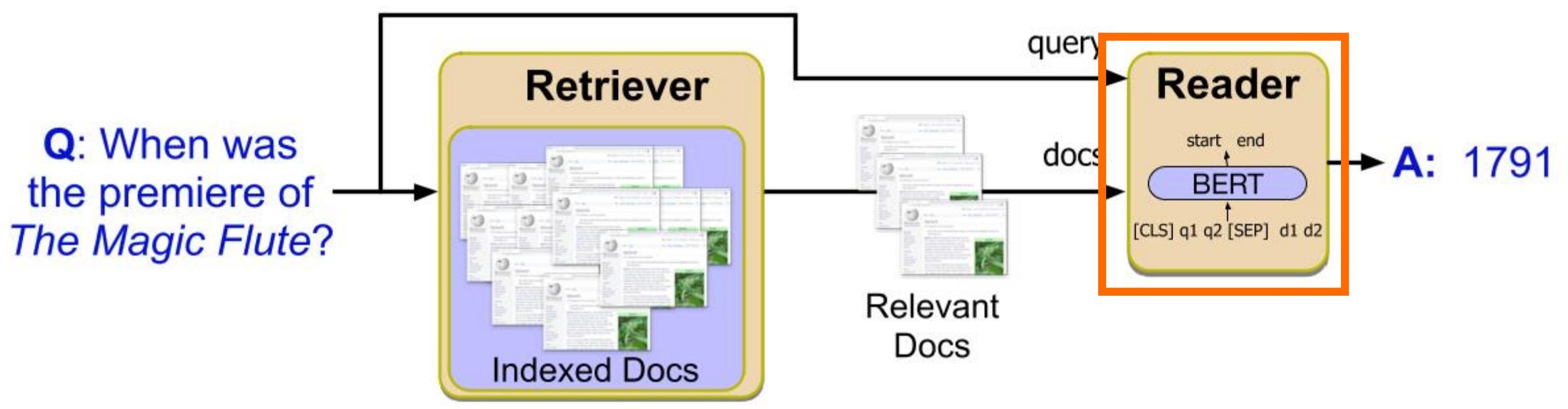
<https://www.theguardian.com/.../31/star-wars-disney-destined-lucasfilm> ▾

Of course, Disney eventually got back on track by the late 80s, and today is an almost unstoppable force in the business, but George Lucas showed them the way, and so there is a neat historical inevitability to the news that Lucasfilm will finally be subsumed by Disney.

IR-based Factoid QA

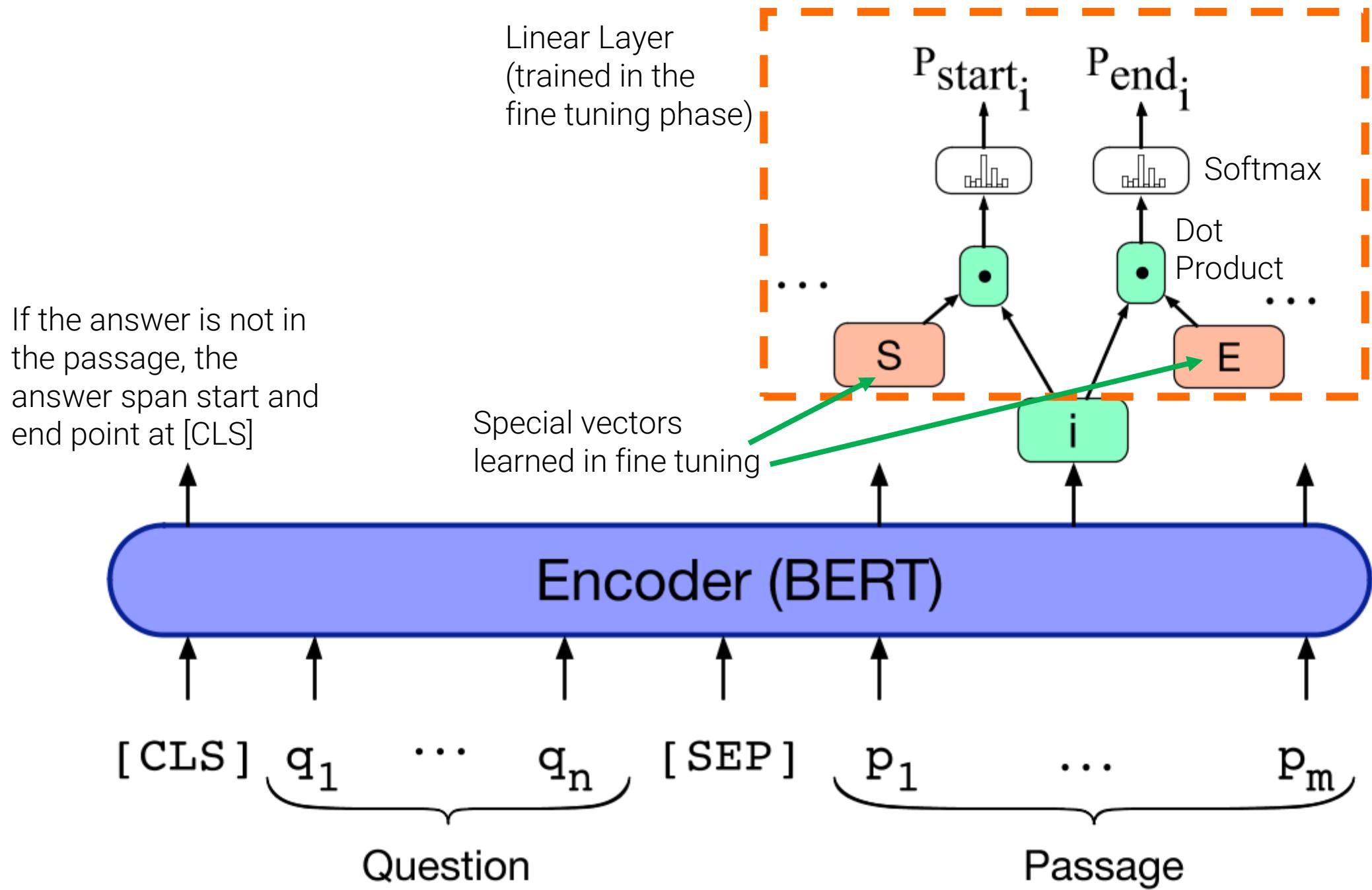


IR-based Factoid QA



Reader (Answer Span Extraction)

- Identify in the document/passage a **span** (continuous string of text) that constitutes an answer
- Neural algorithms for reading comprehension are given:
 - a question q of n tokens q_1, \dots, q_n
 - A passage p of m tokens p_1, \dots, p_m
- Goal: compute the probability $P(a|q, p)$ that each possible span a , starting at position a_s , and ending at position a_e , is an answer
- Assume that this probability can be estimated as
$$P(a|q, p) = P_{\text{start}}(a_s|q, p)P_{\text{end}}(a_e|q, p)$$
- Compute $P_{\text{start}}(i)$ and $P_{\text{end}}(i)$ for every token p_i in the passage



Knowledge-based approaches

- Build a semantic representation of the query
 - Times, dates, locations, entities, numeric quantities
- Use this to query structured data or resources
 - Geospatial databases
 - Ontologies (Wikipedia infoboxes, dbpedia, WordNet, Yago)
 - Restaurant review sources and reservation services
 - Scientific databases

Relation Extraction

- Answers: Databases of Relations
 - born-in("Emma Goldman", "June 27 1869")
 - author-of("Cao Xue Qin", "Dream of the Red Chamber")
 - Draw from Wikipedia infoboxes, DBpedia, FreeBase, etc.
- Questions: Extracting Relations in Questions

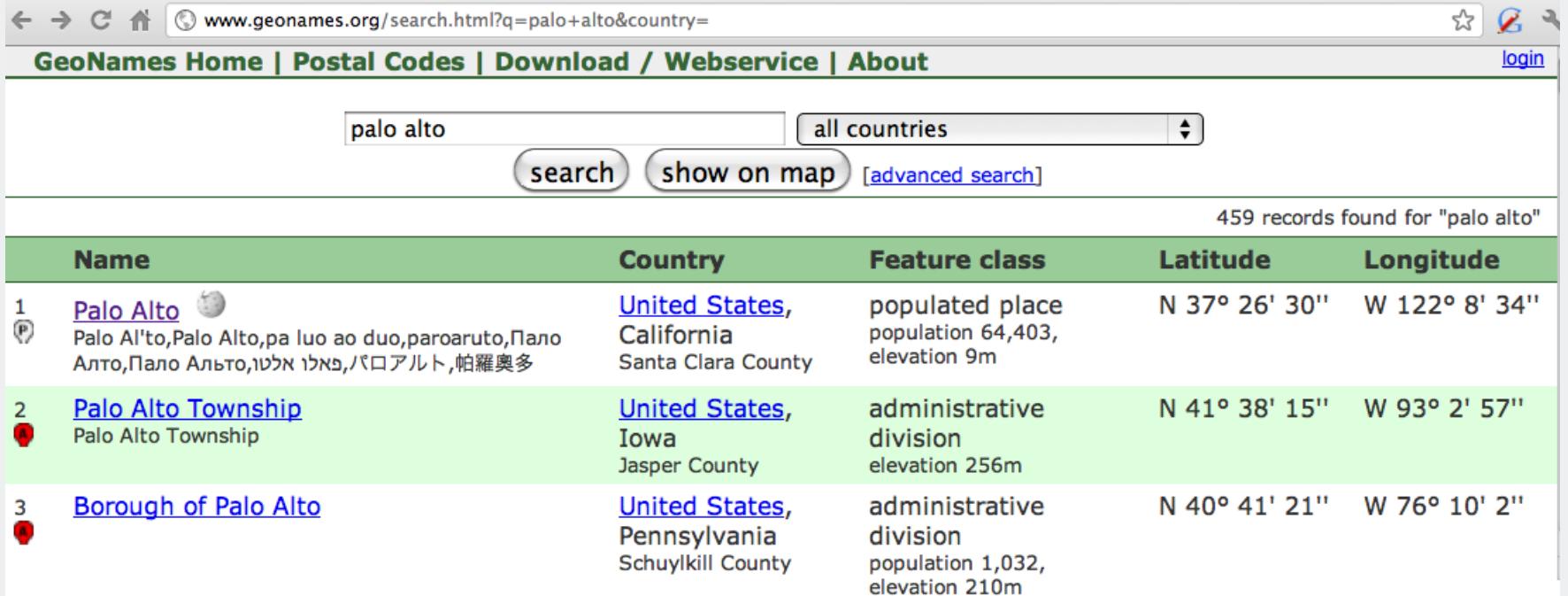
Whose granddaughter starred in E.T.?

(acted-in ?x "E.T.")

(granddaughter-of ?x ?y)

Geospatial knowledge (containment, directionality, borders)

- Beijing is a good answer for "Asian city"
- California is "southwest of Montana"
- geonames.org:



The screenshot shows a web browser displaying the GeoNames search results for the query "palo alto". The search bar at the top contains "palo alto" and "all countries". Below the search bar are buttons for "search", "show on map", and "advanced search". A message indicates "459 records found for 'palo alto'". The main content is a table with columns: Name, Country, Feature class, Latitude, and Longitude. The table lists three entries: 1. Palo Alto (United States, California, Santa Clara County, populated place, population 64,403, elevation 9m), 2. Palo Alto Township (United States, Iowa, Jasper County, administrative division, elevation 256m), and 3. Borough of Palo Alto (United States, Pennsylvania, Schuylkill County, administrative division, population 1,032, elevation 210m).

	Name	Country	Feature class	Latitude	Longitude
1	Palo Alto ⓘ Palo Al'to,Palo Alto,pa luo ao duo,paroaruto,Пало Алто,Пало Альто,עלוֹטוּ,パロアルト,帕羅奧多	United States , California Santa Clara County	populated place population 64,403, elevation 9m	N 37° 26' 30"	W 122° 8' 34"
2	Palo Alto Township ⓘ Palo Alto Township	United States , Iowa Jasper County	administrative division elevation 256m	N 41° 38' 15"	W 93° 2' 57"
3	Borough of Palo Alto ⓘ	United States , Pennsylvania Schuylkill County	administrative division population 1,032, elevation 210m	N 40° 41' 21"	W 76° 10' 2"

QA by Semantic Parsing

Map the question to a structured program to produce an answer

Question	Logical form
What states border Texas?	$\lambda x.\text{state}(x) \wedge \text{borders}(x, \text{texas})$
What is the largest state?	$\text{argmax}(\lambda x.\text{state}(x), \lambda x.\text{size}(x))$
I'd like to book a flight from San Diego to Toronto	SELECT DISTINCT f1.flight_id FROM flight f1, airport_service a1, city c1, airport_service a2, city c2 WHERE f1.from_airport=a1.airport_code AND a1.city_code=c1.city_code AND c1.city_name= 'san diego' AND f1.to_airport=a2.airport_code AND a2.city_code=c2.city_code AND c2.city_name= 'toronto'
How many people survived the sinking of the Titanic?	(count (!fb:event.disaster.survivors fb:en.sinking_of_the_titanic))
How many yards longer was Johnson's longest touchdown compared to his shortest touchdown of the first quarter?	ARITHMETIC diff(SELECT num(ARGMAX(SELECT)) SELECT num(ARGMIN(FILTER(SELECT))))

Figure 23.14 Sample logical forms produced by a semantic parser for question answering, including two questions from the GeoQuery database of questions on U.S. Geography (Zelle and Mooney, 1996) with predicate calculus representations, one ATIS question with SQL (Iyer et al., 2017), a program over Freebase relations, and a program in QDMR, the Question Decomposition Meaning Representation (Wolfson et al., 2020).

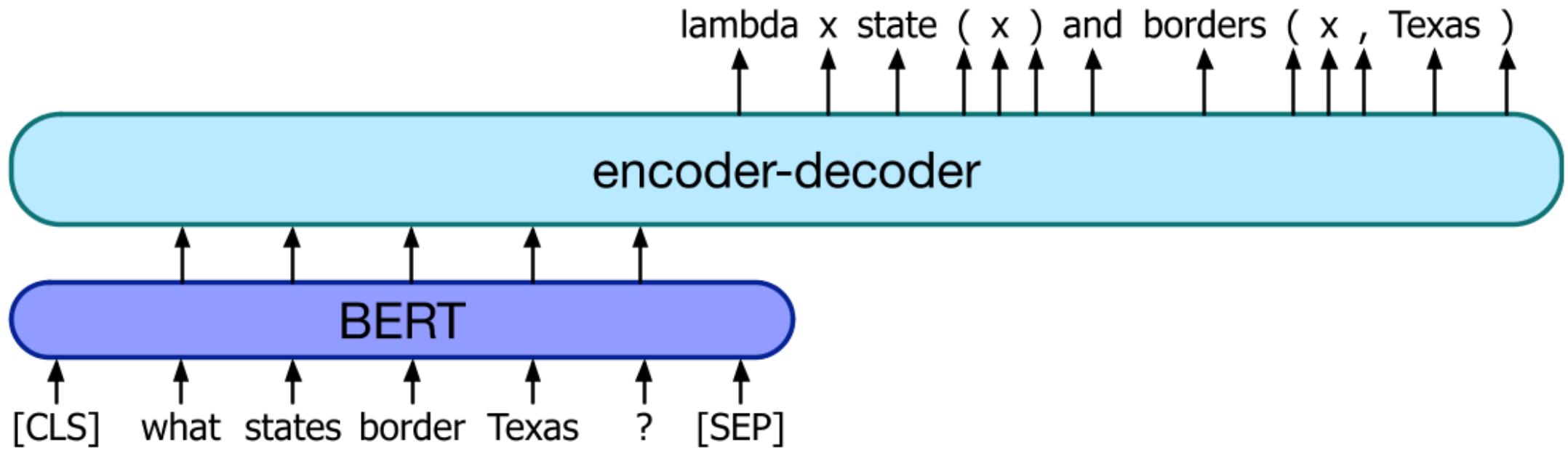
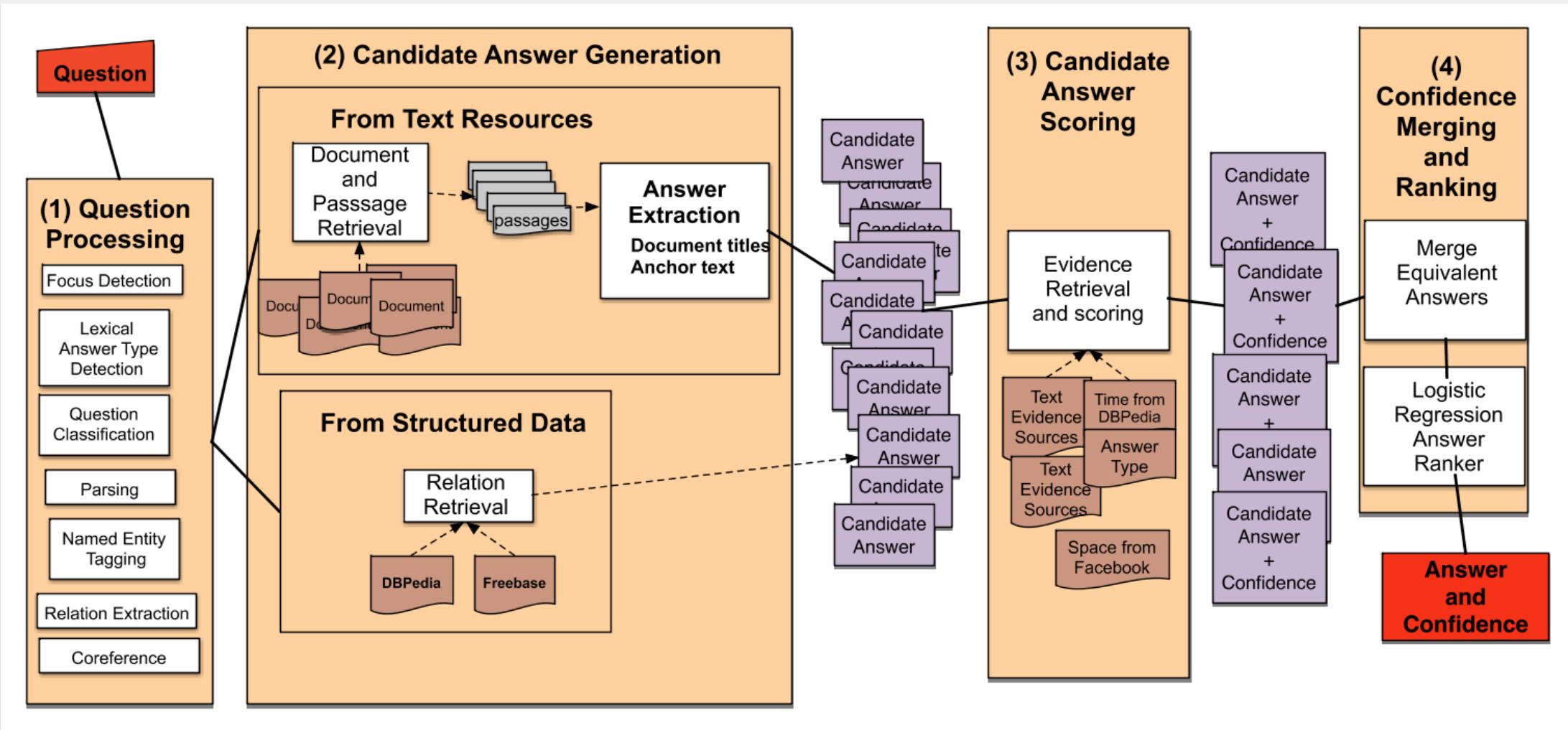
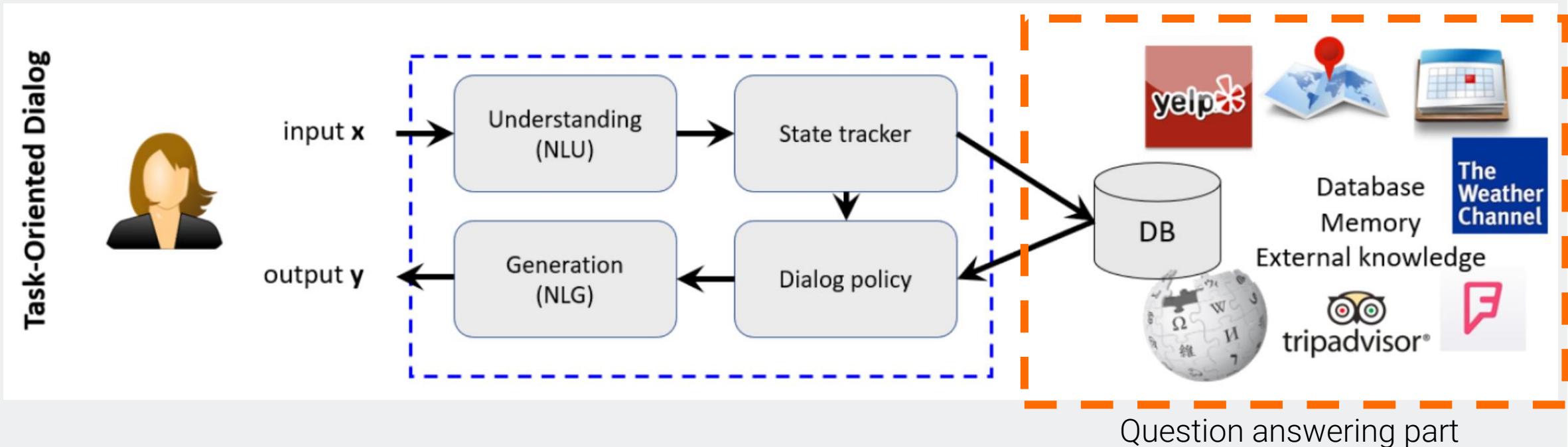


Figure 23.15 An encoder-decoder semantic parser for translating a question to logical form, with a BERT pre-encoder followed by an encoder-decoder (biLSTM or Transformer).

Classic Hybrid Approach: IBM Watson



QA in the Common Agent Architecture



Social Chat (Chatbots)

Chatbot Architectures

- Rule-based
 - Pattern-action rules (Eliza)
- Corpus-based (from large chat corpus) – draw on QA system approaches
 - Information Retrieval
 - Neural network encoder-decoder

Rogerian psychologist

- Draw the patient out by reflecting patient's statements back at them
- Rare type of conversation in which one can “assume the pose of knowing almost nothing of the real world”
- Example:
 - Patient: "I went for a long boat ride"
 - Psychiatrist: "Tell me about boats"
- You don't assume she didn't know what a boat is
- You assume she had some conversational goal
- Chatbots trying to pass the Turing test usually choose some such domain

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.

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

(0 YOU 0 ME) *[pattern]*

→

(WHAT MAKES YOU THINK I 3 YOU) *[transform]*

The 3 is the constituent # in pattern

You hate me

WHAT MAKES YOU THINK I HATE YOU

Eliza Rules

	pattern	Ranked list of transforms
keyword	$(K \ ((D_1) \ (R_{1, 1}) \ (R_{1, 2}) \ \cdots \ (R_{1, m_1})))$	
	$((D_2) \ (R_{2, 1}) \ (R_{2, 2}) \ \cdots \ (R_{2, m_2}))$	
	\vdots	\vdots
	$((D_n) \ (R_{n, 1}) \ (R_{n, 2}) \ \cdots \ (R_{n, m_n})))$	

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*)

If No Matching Keywords

Select a “neutral response”:

- PLEASE GO ON
- THAT'S VERY INTERESTING
- I SEE

Or use **Memory** (next slide)

Memory

```
(MEMORY MY  
  (0 YOUR 0 = LETS DISCUSS FURTHER WHY YOUR 3)  
  (0 YOUR 0 = EARLIER YOU SAID YOUR 3)
```

- Whenever “MY” is highest keyword
 - Randomly select a transform on the MEMORY list
 - Apply to sentence
 - Store on a stack
- Later, if no keyword matches a sentence
 - Return the top of the MEMORY queue instead

Other Eliza stuff

- Rules can refer to classes of words
 - Family = mother, father, brother, sister
 - NOUN = ...
- Don't reuse transforms in the same conversation
 - Whenever we use a transform associated with a pattern
 - We increment a counter for that rule
 - So the next time we use the next ranked transform
- Some basic transforms happen during input processing
 - I -> YOU
 - YOU -> I

Corpus-based chatbots

Idea: Mine conversations of human chats or human-machine chats

- Microblogs: Twitter or Weibo (微博)
- Movie dialogs
- Cleverbot (Carpenter 2017 <http://www.cleverbot.com>)
- Microsoft Xiaoice
- Microsoft Tay

A Cleverbot conversation

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.

Two IR-based chatbot architectures

1. Return the response to the most similar turn

- Take user's turn (q) and find a (tf-idf) similar turn t in the corpus C

$q = \text{"do you like Doctor Who"}$

$t = \text{"do you like Doctor Strangelove"}$

- Grab whatever the response was to t

$$r = \text{response} \left(\operatorname{argmax}_{t \in C} \text{sim}(q, t) \right)$$

Yes, so funny

2. Return the most similar turn

$$r = \operatorname{argmax}_{t \in C} \text{sim}(q, t)$$

Do you like Doctor Strangelove

Similarity functions

- Classic

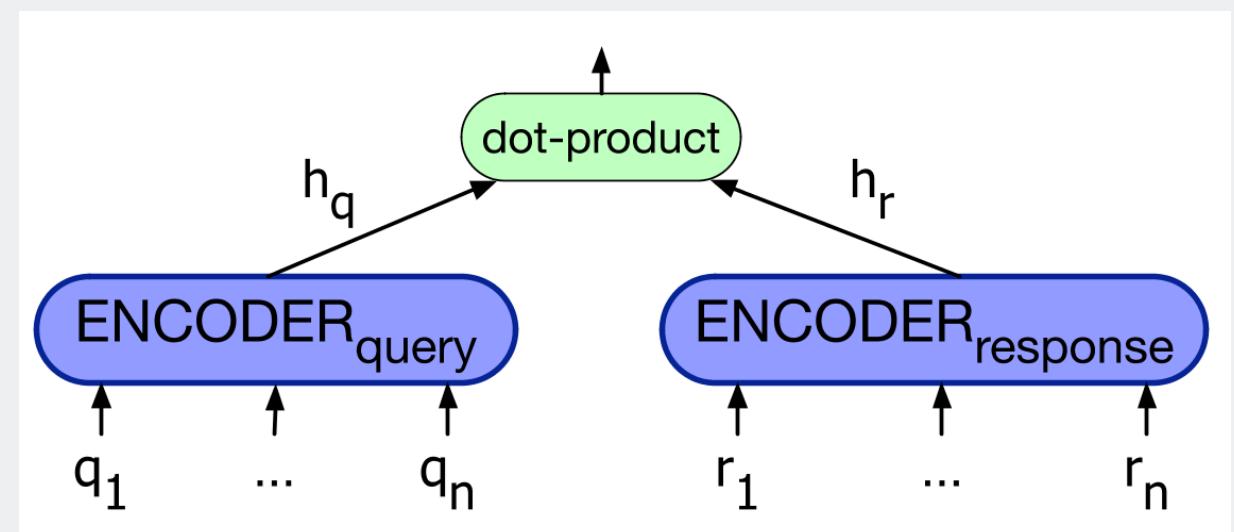
$$\text{sim}(q, r) = \frac{q^T r}{\|q\| \|r\|}$$

- Neural

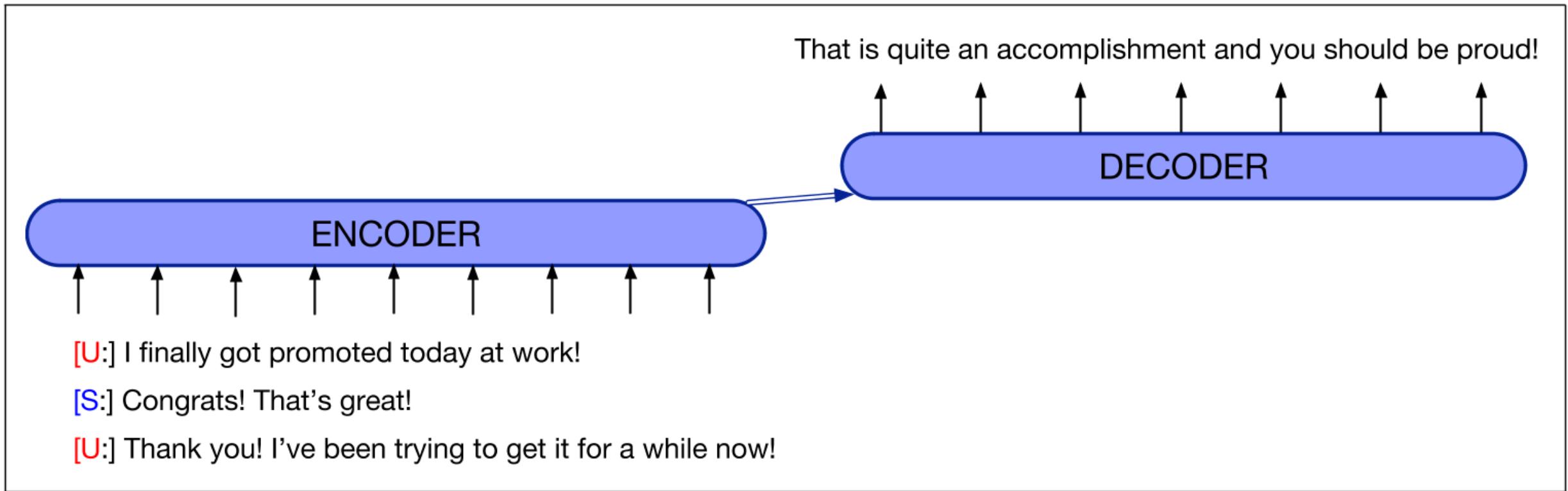
$$h_q = \text{BERT}_Q(q) [\text{CLS}]$$

$$h_r = \text{BERT}_R(r) [\text{CLS}]$$

$$\text{sim}(q, r) = h_q \cdot h_r$$



Neural Network Architecture (Response by Generation)



ChatGPT

- Version of GPT3 optimised for dialogue through manual labelling
- <https://openai.com/blog/chatgpt/>

AL Write a limerick about a natural language processing class

 In a natural language processing class
The students were feeling quite amorous
But their romance was thwarted
When the professor imported
A dataset of legal documents

AL

Write python code to check if text contains an instruction.



Here is an example of how you could write Python code to check if a given string of text contains an instruction:

Copy code

```
def has_instruction(text):
    instructions = ["do this", "do that", "go here", "go there"]
    for instruction in instructions:
        if instruction in text:
            return True
    return False

text = "Please do this task and then go there"
if has_instruction(text):
    print("Text contains an instruction.")
else:
    print("Text does not contain an instruction.")
```

This code defines a function called `has_instruction` that takes a string of text as input and checks to see if any of the instructions in the `instructions` list appear in the text. If an instruction is found, the function returns `True`, otherwise it returns `False`. The function is called on the `text` string and the result is printed to the console. Note that this is just one way to solve this problem, and there are many other approaches you could take. Is there anything else I can help with?

False Links

AL

Where can I find information on python for natural language processing?



There are many resources available online that can help you learn about using Python for natural language processing (NLP). Here are a few options:

- The Python documentation includes a section on NLP, which provides an overview of some of the tools and libraries available in Python for working with NLP:
<https://docs.python.org/3/library/nlp.html>
- The Natural Language Toolkit (NLTK) is a popular library for NLP in Python. It includes a wide range of tools for tasks such as tokenization, stemming, and part-of-speech tagging, as well as more advanced techniques like parsing and machine learning. You can find documentation and tutorials for NLTK at: <https://www.nltk.org/>
- There are also many online tutorials and courses available that cover using Python for NLP. For example, you can find a free course on NLP with Python on the website of the National Research University Higher School of Economics: <https://www.coursera.org/learn/python-for-nlp>

I hope these resources are helpful. Let me know if you have any other questions.

Hallucination



There are many ways to modulate, or change key, from C major to G major. Here are a few options:

1. Direct modulation: This involves simply changing the tonic note from C to G and continuing the piece in the new key. This can be done by using a dominant chord (a chord built on the fifth scale degree) that leads to the tonic chord in the new key. For example, you could use a G major chord to lead directly to the G major tonic chord.
2. Modulation through a related key: Another option is to modulate to a key that is closely related to the destination key. In this case, you could modulate to the key of E minor, which has the same key signature as G major (one sharp) and shares many common chords. From E minor, you can then modulate to G major by using a dominant chord (such as an A major chord) that leads to the tonic chord in the new key.
3. Modulation through a pivot chord: A pivot chord is a chord that is shared by both the original key and the destination key. For example, in this case **you could use an A major chord, which is common to both C major and G major**, as a pivot chord to help modulate between the two keys. To do this, you could use the A major chord to lead temporarily to the key of A major, and then use a dominant chord (such as an E major chord) to lead to the tonic chord in the destination key of G major.

These are just a few examples of how you could modulate from C major to G major. The best approach will depend on the specific needs of your music and the overall structure of the piece. I hope this helps! Let me know if you have any other questions.

OpenAI list of limitations



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

Limitations

- ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers. Fixing this issue is challenging, as: (1) during RL training, there's currently no source of truth; (2) training the model to be more cautious causes it to decline questions that it can answer correctly; and (3) supervised training misleads the model because the ideal answer depends on what the model knows, rather than what the human demonstrator knows.
- ChatGPT is sensitive to tweaks to the input phrasing or attempting the same prompt multiple times. For example, given one phrasing of a question, the model can claim to not know the answer, but given a slight rephrase, can answer correctly.
- The model is often excessively verbose and overuses certain phrases, such as restating that it's a language model trained by OpenAI. These issues arise from biases in the training data (trainers prefer longer answers that look more comprehensive) and well-known over-optimization issues.^{1,2}
- Ideally, the model would ask clarifying questions when the user provided an ambiguous query. Instead, our current models usually guess what the user intended.
- While we've made efforts to make the model refuse inappropriate requests, it will sometimes respond to harmful instructions or exhibit biased behavior. We're using the Moderation API to warn or block certain types of unsafe content, but we expect it to have some false negatives and positives for now. We're eager to collect user feedback to aid our ongoing work to improve this system.

Chatbots: pro and con

- Pro:
 - Fun
 - Applications to counseling
 - Getting ever more powerful
- Cons:
 - They don't really understand
 - Rule-based chatbots are expensive and brittle
 - Neural chatbots sometimes provide incorrect information very convincingly

Chatbots in the World

Write an essay on how the first world war started.

The first world war, also known as World War I, was a global conflict that began on July 28, 1914 and lasted until November 11, 1918. The main cause of the war was the assassination of Archduke Franz Ferdinand of Austria-Hungary on June 28, 1914 by a Serbian nationalist. The assassination led to a chain of events that resulted in declarations of war between several European powers.

Austria-Hungary declared war on Serbia on July 28, 1914, and other countries, including Germany, Russia, and France, soon joined the conflict. The war involved many of the world's nations, including all of the great powers, which were eventually divided into two opposing military alliances: the Allies (primarily made up of France, Russia, and the United Kingdom) and the Central Powers (primarily made up of Germany, Austria-Hungary, and Italy).

The assassination of Archduke Ferdinand **was the spark that ignited the powder keg** of tensions between the European powers. There were several underlying causes of the war, including militarism, alliances, imperialism, and nationalism.

Militarism, or the glorification of military values and the belief in the superiority of a nation's military, was prevalent in many countries at the time. This led to a buildup of military arsenals and a desire to prove military dominance.

Alliances, or agreements between nations to support each other in the event of a conflict, also played a role in the start of the war. Many countries had formed alliances with each other, creating a web of obligations that dragged them into the war when one of their allies was attacked.

Imperialism, or the desire to acquire and control territory, was also a factor in the start of the war. Many European powers had colonies around the world, and there was competition between them to expand their empires.

Nationalism, or the belief in the superiority of one's own nation, was also a driving force behind the war. Many people believed in the inherent superiority of their own nation and were willing to fight to defend it.

The combination of these factors created a volatile situation in Europe that was ripe for conflict. The assassination of Archduke Ferdinand **was the catalyst that ignited the powder keg**, leading to the outbreak of the first world war. The war resulted in the deaths of millions of people and had a significant impact on global politics, including the collapse of empires and the redrawing of national borders.

Maya Yang

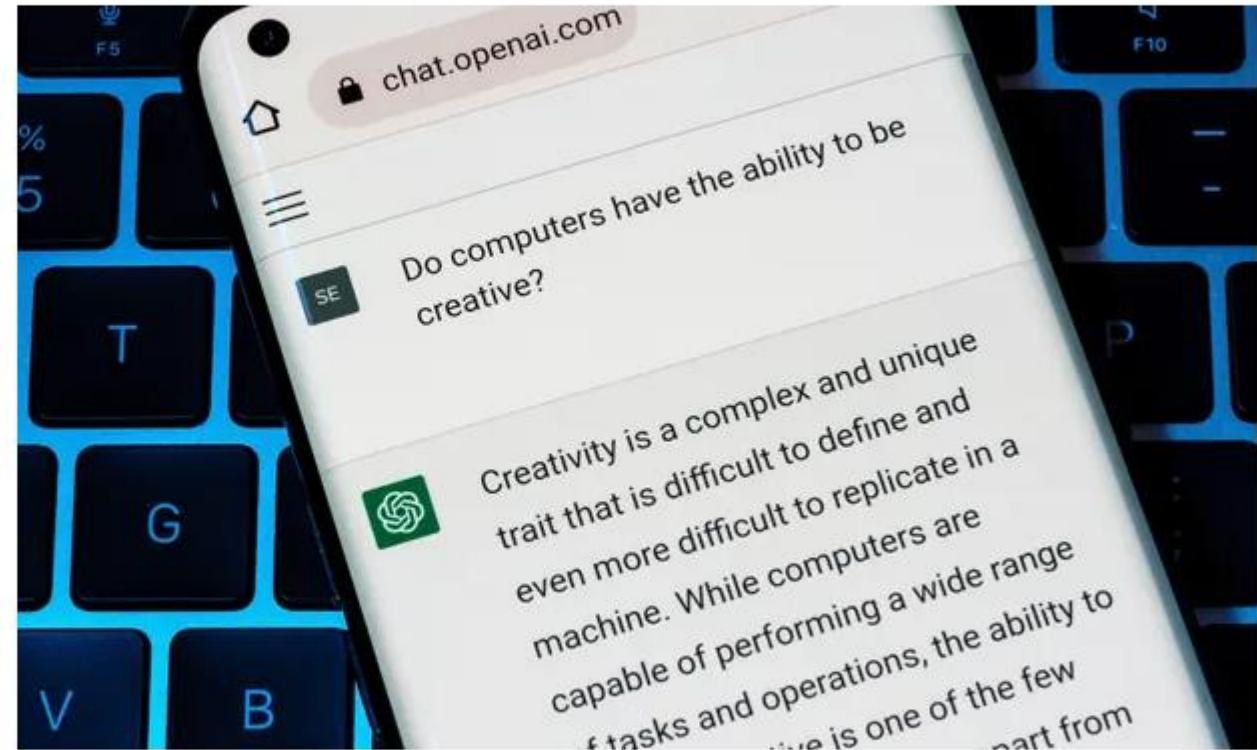
Fri 6 Jan 2023 18.19 GMT



<https://www.theguardian.com/us-news/2023/jan/06/new-york-city-schools-ban-ai-chatbot-chatgpt>

New York City schools ban AI chatbot that writes essays and answers prompts

ChatGPT tool will be forbidden across all devices and networks in public schools over 'concerns about negative impacts on learning'



ChatGPT is an artificial intelligence chatbot that generates human-like writing. Photograph: Ascannio/Alamy

New York City schools have banned ChatGPT, the artificial intelligence chatbot that generates human-like writing including essays, amid fears that students could use it to cheat.

On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?



Emily M. Bender*

ebender@uw.edu

University of Washington

Seattle, WA, USA

Angelina McMillan-Major

aymm@uw.edu

University of Washington

Seattle, WA, USA

Timnit Gebru*

timnit@blackinai.org

Black in AI

Palo Alto, CA, USA

Shmargaret Shmitchell

shmargaret.shmitchell@gmail.com

The Aether

ABSTRACT

The past 3 years of work in NLP have been characterized by the development and deployment of ever larger language models, especially for English. BERT, its variants, GPT-2/3, and others, most recently Switch-C, have pushed the boundaries of the possible both through architectural innovations and through sheer size. Using these pretrained models and the methodology of fine-tuning them for specific tasks, researchers have extended the state of the art on a wide array of tasks as measured by leaderboards on specific benchmarks for English. In this paper, we take a step back and ask: How big is too big? What are the possible risks associated with this technology and what paths are available for mitigating those risks? We provide recommendations including weighing the environmental and financial costs first, investing resources into curating and carefully documenting datasets rather than ingesting everything on the web, carrying out pre-development exercises evaluating how the planned approach fits into research and development goals and supports stakeholder values, and encouraging research directions beyond ever larger language models.

Points made in the Stochastic Parrots paper

- Training a large language model is bad for the Environment
 - Average human emits 5t CO₂e per year, training a big transformer emits 284t CO₂
- Unfathomable training data
 - Due to biases in internet use, crawling huge amounts of text doesn't guarantee that it is representative of all humanity
 - Bias encoded in the text data
- A large Language Model is a Stochastic Parrot – it doesn't do Natural Language Understanding
 - However, it “can manipulate linguistic form well enough to cheat its way through tests meant to require language understanding”
- Such Language Models can be used with malicious intent
 - e.g. to produce large numbers of coherent texts that are not necessarily true

Difficulty in creating Chatbot business models

<https://arstechnica.com/gadgets/2022/11/amazon-alex-is-a-colossal-failure-on-pace-to-lose-10-billion-this-year/>

WOULD YOU PAY A MONTHLY FEE FOR VOICE COMMANDS? —

Amazon Alexa is a “colossal failure,” on pace to lose \$10 billion this year

Layoffs reportedly hit the Alexa team hard as the company's biggest money loser.

RON AMADEO - 11/21/2022, 8:32 PM



[Enlarge](#) / The fourth-generation Echo device is a cloth-covered sphere with a halo at the base, contrasting with the squat plastic cylinders of earlier-generation Echoes.

505

Amazon is going through the biggest layoffs in the company's history right now, with a plan to eliminate some [10,000 jobs](#). One of the areas hit hardest is the Amazon Alexa voice assistant unit, which is apparently falling out of favor at the e-commerce giant. That's according to a report from [Business Insider](#), which details "the swift downfall of the voice assistant and Amazon's larger hardware division."

Summary

Summary

- Many applications of techniques seen in the course
- Commercial Dialogue Systems are often rule-based or have fixed dialogue structures (Finite State Dialogue)
- Never underestimate the power of rule-based approaches – can also be used to generate data for subsequent machine learning
- Neural Approaches are the current hype
- General purpose Dialogue Systems (task completion, QA, chitchat) are a current research topic

Course Evaluation

- Evaluate the course on TUWEL until 9th February 2023 at 23:59

<https://tiss.tuwien.ac.at/survey/surveyForm.xhtml?courseNumber=194093&semesterCode=2022W>

