

Lecture 8 – Question Answering Systems
and Conversational Agents – Part 1

CSAI401 – APPLIED NATURAL LANGUAGE PROCESSING

Dr. Andrea Espinel Rojas

Slides inspired on Dr. Suleiman
Yerima (CEH, CISSP, FHEA) work.

OUTLINE

PART 1:

Introduction – QA systems

Conversational Agents

- Chatbots
- Goal/task oriented

Chatbot architectures

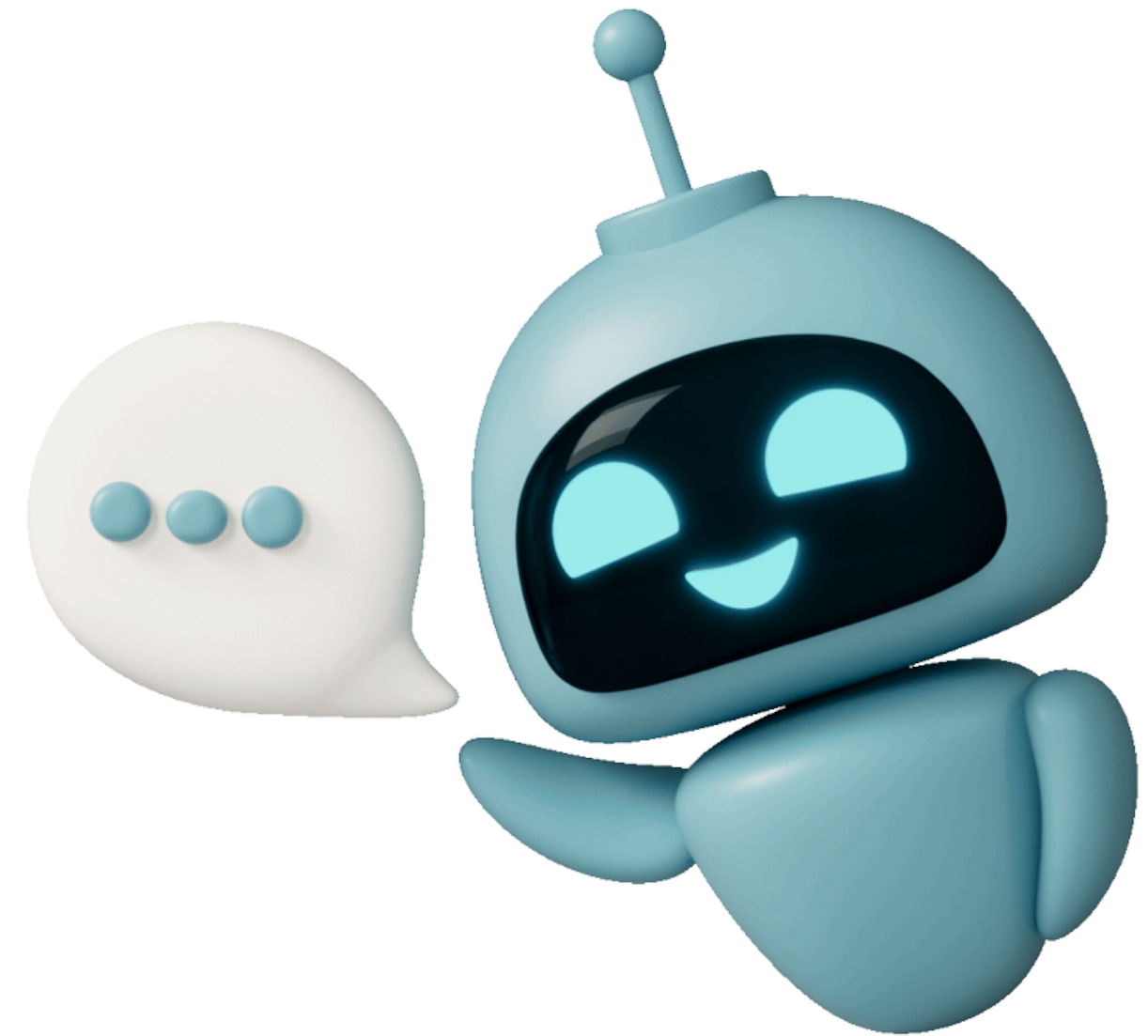
- Rule-based

PART 2:

- Corpus-based

Evaluating dialogue systems

Lab 8



INTRODUCTION

A Question Answering (QA) system is a type of AI application that is designed to answer questions posed by humans in a natural and coherent manner.

A QA system can answer the question by extracting an answer directly from the given documents **(extractive QA)** or synthesizing the information in the documents to generate an answer **(generative QA)**.

We can evaluate a QA system by comparing the predicted answer to the given answer.

QA systems have a variety of applications including:

- Customer service

- Education

- Healthcare

- Virtual assistants in smart devices.



INTRODUCTION

Extractive QA: The model extracts the answer from a context. The context here could be a provided text, a table or even HTML

This is usually solved with BERT-like models.

Open Generative QA: The model generates free text directly based on the context.

Closed Generative QA: In this case, no context is provided. The answer is completely generated by a model.

An instance of a (extractive) QA system consists of the following three main components:

- (1) given documents as context,
- (2) a question, and
- (3) an answer.



Question: "Who is the CEO of Tesla?"

Extractive QA

Context (from a Wikipedia article): "Elon Musk is the CEO of Tesla, an American electric vehicle and clean energy company based in Texas."

Answer (extracted):

"Elon Musk"

✔ The model selects this exact span from the context.

Open Generative QA

Context: Same as above

Answer (generated):

"The CEO of Tesla is Elon Musk."

✔ The model reads the context and generates a fluent sentence.

INTRODUCTION

Closed Generative QA

No context provided — just the question:

"Who is the CEO of Tesla?"

Answer (generated):

"Elon Musk" or "The CEO of Tesla is Elon Musk."

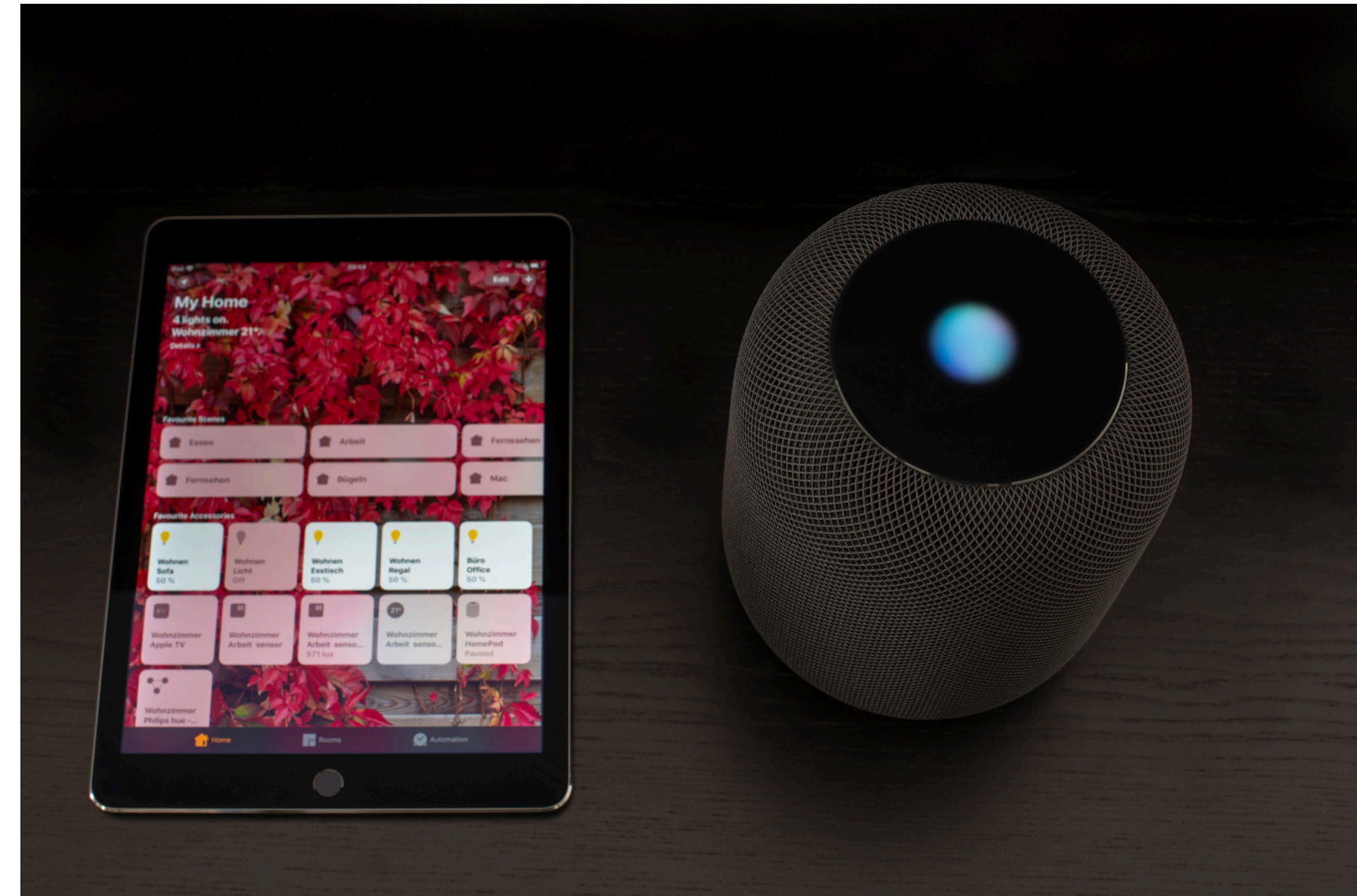
✔ The model relies purely on its internal knowledge (pretrained data).

CONVERSATIONAL AGENTS

Conversational agents, also called dialogue systems, dialogue agents, chatbots, are systems designed to interact with humans via conversations, either in text or speech.

These include the personal assistants on phones or other devices like SIRI, Alexa, Cortana, or the Google Assistant.

Tools for simple, relatively short interactions like playing music, setting timers or clocks or keeping shopping lists, to longer conversations maybe just for fun, or for practical applications like booking travel reservations, or even in clinical uses for mental health.



CONVERSATIONAL AGENTS

Two broad categories of conversational agents.

- **Chatbots** are systems that can carry on extended conversations with the goal of mimicking the unstructured conversations or chats' characteristic of informal human-human interaction.
- **Goal-based/Task-based dialogue agents** are used to solve some task like booking a flight or maintaining a shopping list.

Sometimes you'll see the word "chatbot" used for both kinds of systems.



1.Rule-based

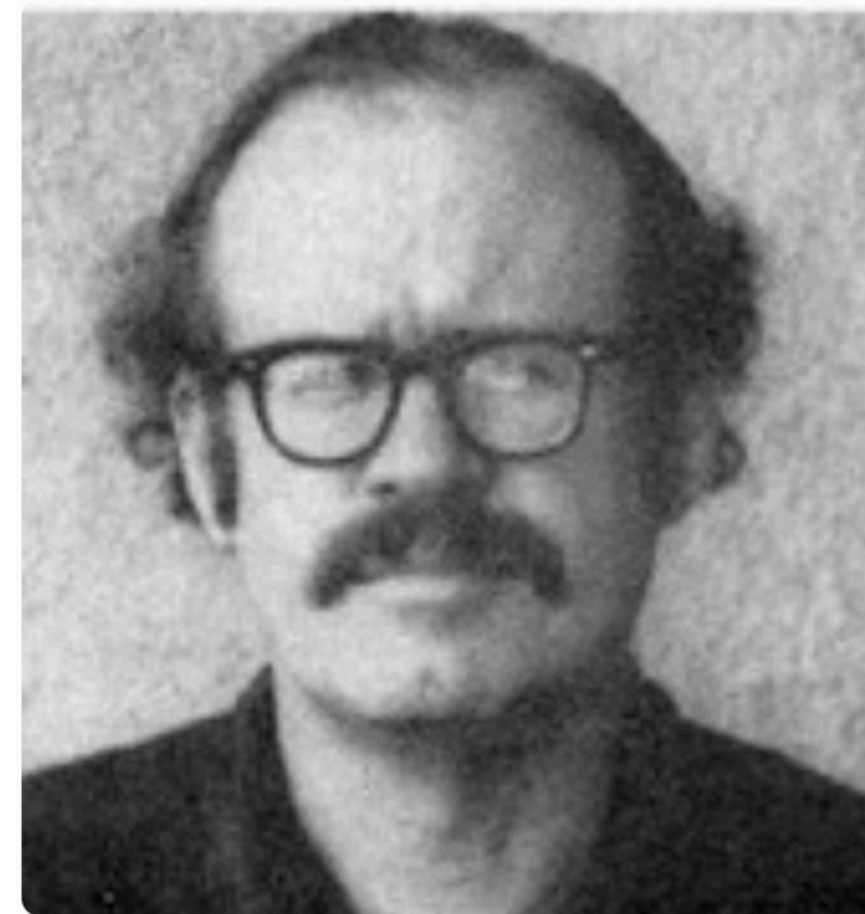
- Pattern-action rules (ELIZA created 1966).
- A mental model (PARRY created 1971):
The first system to pass the Turing Test!

2. Corpus-based

Corpus-based systems mine large datasets of human-human conversations, which can be done by using information retrieval to copy a human response from a previous conversation, or using an encoder-decoder system to generate a response from a user utterance.

- **Information Retrieval (Xiaolce).**
- **Neural encoder-decoder (BlenderBot).**

CHATBOT ARCHITECTURES



H: Good afternoon. My name
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 thought you might have some
problems.
P: I've been upset lately.
H: By what?
P: People get on my nerves

A mental model (PARRY created 1971).

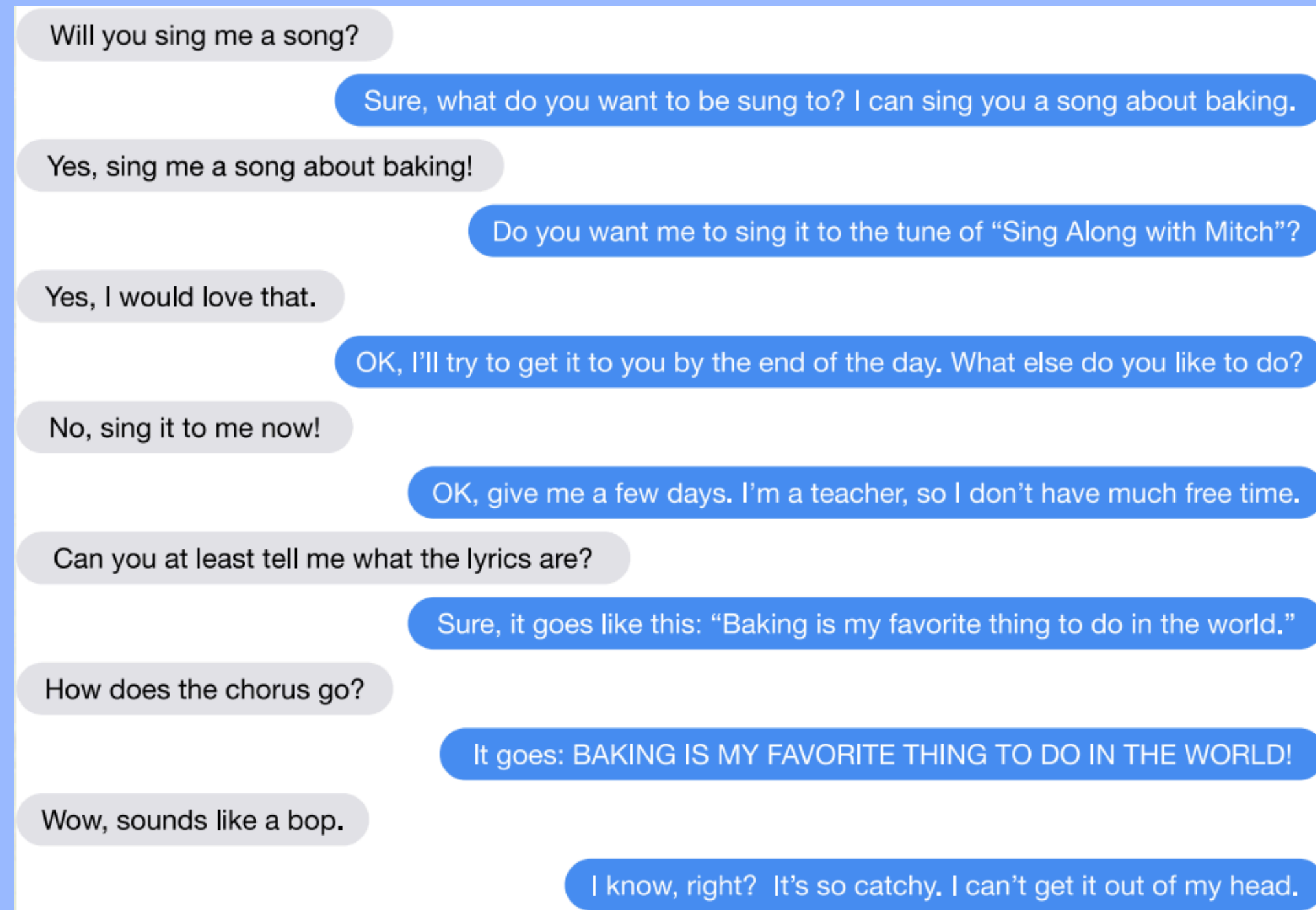
XIAOICE (ZHOU ET AL., 2020)

Microsoft's Xiaoice system (Zhou et al., 2020), chats with people in Chinese on text messaging platforms, responding mainly by extracting turns that humans have said in past conversations.

[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.

BLENDERBOT (ROLLER ET AL. 2020)

These systems often have an entertainment value, such as Facebook's Blender- Bot, a neural chatbot capable of carrying on conversations of this type.



TASK-BASED DIALOGUE AGENTS

Other dialogue agents are built to solve a task, like setting a timer, making a travel reservation, or playing a song.

These tend to be built around a knowledge structure called the frame.

"Task-based" or "goal-based" dialogue agents.

Architecture:

Frames with slots and values: A knowledge structure representing user intentions.

Systems that have the goal of helping a user solve a task

- **Setting a timer.**
- **Making a travel reservation.**
- **Playing a song.**
- **Buying a product.**



THE FRAME

A frame represents the users intentions for the task, and consists of a collection of slots, each of which can take a set of possible values, so an airline booking agent might have slots like the destination city or the departure time.

A set of slots, to be filled with information of a given type.

Each associated with a question to the user:

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?"

PROPERTIES OF HUMAN CONVERSATION

Before we attempt to design a conversational agent to converse with humans, it is crucial to understand something about how humans converse with each other.

Let's consider some of the properties of this conversation.

First of all, we call each of the contributions of the two conversant a "turn", as if conversation was the kind of game where everyone takes turns : **there are 20 turns in this excerpt.**

A turn can consist of a sentence (like C1), although it might be as short as a single word (C13) or as long as multiple sentences (A10).

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 ...

PROPERTIES OF HUMAN CONVERSATION

Next, the fact that there are two (or more) people who need to talk means that they have to negotiate turn-taking.

- Who takes the floor?
- When do I yield the floor?

What happens when we both try to talk at the same time, cases of interruptions.

Notice that the client interrupts here, and the human agent knows to stop talking (and also knows that the client might be making a correction or change). Systems need to be able to handle this.

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

In a dialogue system, recognizing when the user is talking is called barge-in.

A system also has to know when to start talking.

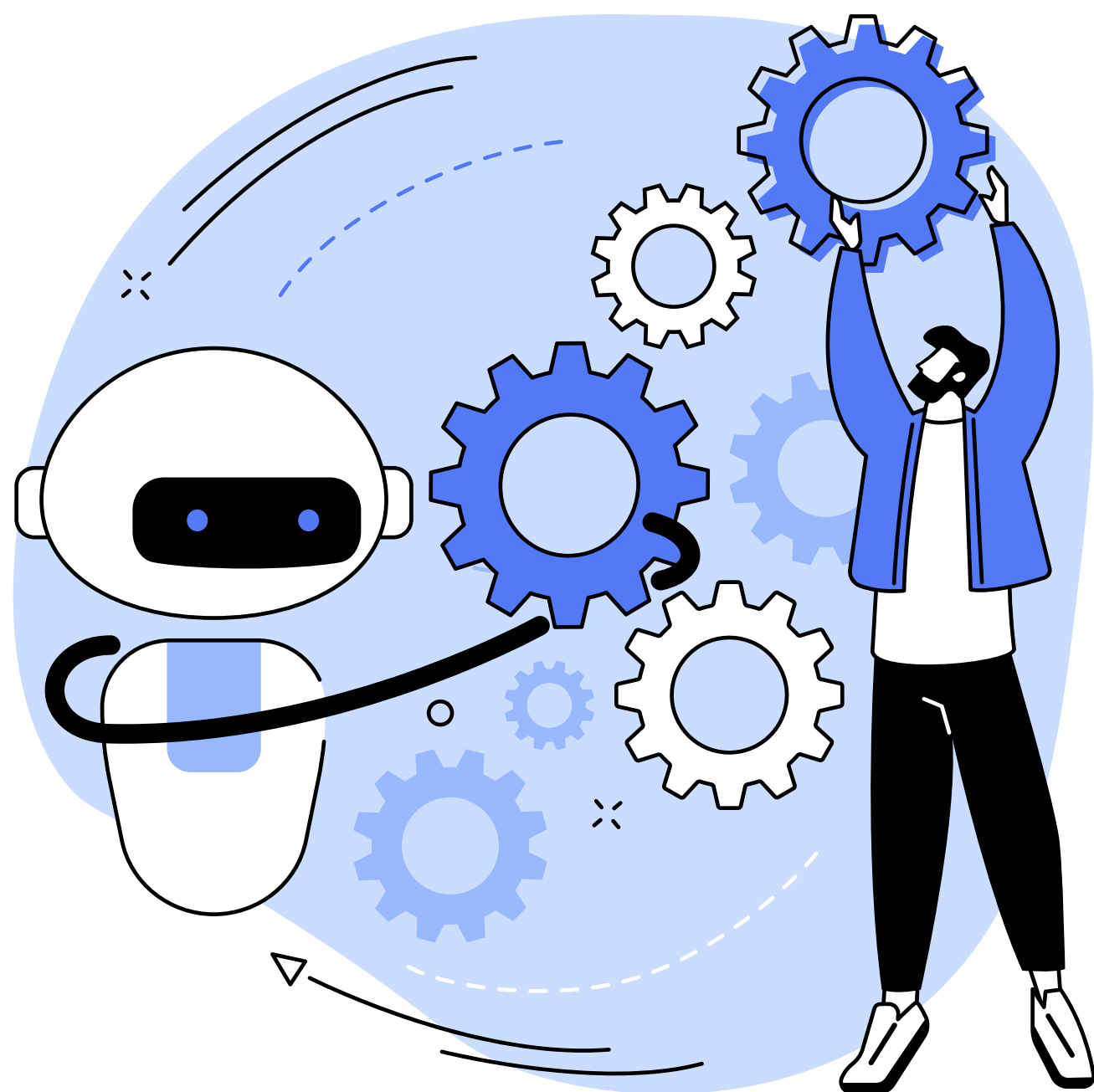
For example, most of the time in conversation, speakers start their turns almost immediately after the other speaker finishes, without a long pause.

This is sometimes called "latching"; people are able to (most of the time) detect when the other person is about to finish talking and time their response to slip right in.

Spoken dialogue systems must detect whether a user is done speaking, so they can process the utterance and respond.

This task— called endpointing or endpoint detection — can be quite challenging because of noise and because people often pause in the middle of turns.

An alternative to writing rules is to build chatbots that automatically mine collections of human conversations to decide what to say.



**THANK YOU FOR
LISTENING!**