

COMP 10261: Chat Bots

Sam Scott, Mohawk College, 2021.

CHAT BOT FUNDAMENTALS

A chat bot is an Artificial Intelligence application that “listens” on a voice or text channel for a user’s **utterances**. The chat bot’s job is to determine the **intent** behind each utterance, identify any relevant **entities**, and then produce a **response**.

DEFINITIONS

Term	Definition	Examples
Utterance	What the user said on a voice channel or typed on a text channel	“Good Morning” “How do I apply to Mohawk College?” “I’m trying to get to the Library”
Intent	What the user is trying to do	Greet the bot Obtain administrative information Get directions to a library
Entity	Something from the real world that the user mentioned in an utterance, or that the bot needs to ask the user about	Morning Mohawk College Hamilton Library, Main Branch
Response	The text that the chat bot sends back to the user on a text or voice channel	“Hello, how can I help you?” “Please visit our web site.” “The Hamilton Library’s main branch is at 55 York Boulevard.”

THE CHAT BOT PIPELINE

Natural Language Processing professionals often modularize complex NLP tasks by thinking in terms of components in an NLP **pipeline**. Each component performs a computation on its input and then passes its output on to the next component. A typical Chat Bot pipeline is shown below.



SPEECH RECOGNITION

This component is only present in a voice-activated chat bot. Its job is to turn a spoken **utterance** into a text string. The AI technologies behind speech recognition are quite mature and it is rare that a chat bot developer would create their own speech recognition system. Instead, an off-the-shelf tool would be used, perhaps tuned to the linguistic community the app is intended to serve. In this course, we will ignore this component and focus on NLP for text strings.

NATURAL LANGUAGE UNDERSTANDING

This component can perform a variety of tasks aimed at “understanding” the user’s **utterance**. This can include extracting **Entities** from the text, labelling words with their **Parts of Speech** (Noun, Verb, etc.), **Parsing** sentences to determine their grammatical structure, resolving **co-references** (words like “she”, “that”, etc.), and/or **classifying** sentences for sentiment (emotional tone), speech act (question, greeting, statement, etc.), or topic.

This is the pipeline component where we will be concentrating most of our efforts in this course. The techniques and technologies required for this component of the chat bot are most likely to generalize to other NLP tasks such as text summarization, information retrieval, question answering, word prediction, email filtering, and so on.

DIALOG MANAGEMENT

This component decides what is most important about the conversation, tracks the state of the conversation, and/or guides the flow of the conversation towards a goal. The dialog manager can also identify follow-up questions or **entities** that need to be identified in order to respond effectively. Not all chat bots will contain this component.

NATURAL LANGUAGE GENERATION

This component’s job is to generate a human-readable **response** in Natural Language to send back to the user. The response is based on the results of Natural Language Understanding and Dialog Management and the output is a text string that can be shown to the user or passed to a text-to-speech component. Responses can be simple stock answers (e.g., for a FAQ Bot), they can be generated from templates (e.g., an assistant fills in the underlined slots to say “Now playing Daddy’s Home by St. Vincent”), or they can be generated from scratch using a generative model of language.

TEXT-TO-SPEECH SYNTHESIS

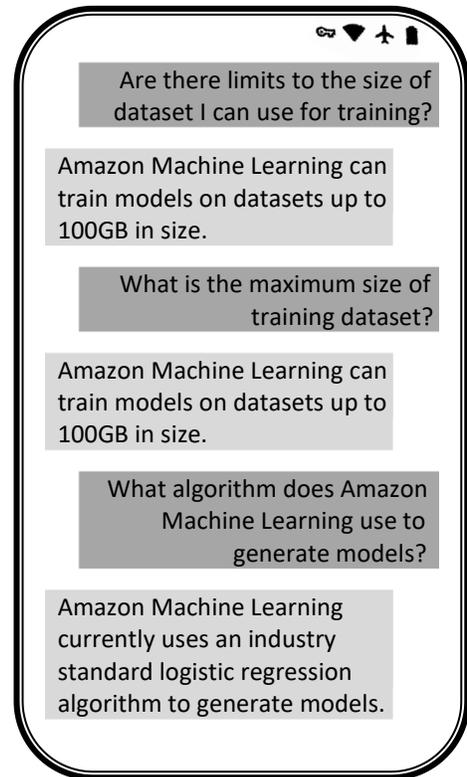
If the chat bot is on a voice channel, this component turns the text output from the Natural Language Generation component into audible speech. As with Speech Recognition, this is very close to a solved AI problem, so most chat bot developers would use an off-the-shelf solution. We will not consider this component any further for this course.

FAQ BOTS

An **FAQ Bot** (Frequently Asked Questions Bot) is a very simple chat bot that generates **fixed responses** based on its understanding of a user's **utterance**. Most of the work of a FAQ Bot is in the Natural Language Understanding component.

When an FAQ Bot receives an **utterance**, it determines the user's **intent** by matching that utterance to one of its stored question and answer pairs. If it succeeds in determining intent in this way, it uses the answer as its **response**. In the example on the right (from Vajjala et al.'s *Practical Natural Language Processing*) the FAQ Bot has determined that the first two utterances have the same intent and has responded with the same text in both cases.

If an FAQ Bot fails to determine intent, it usually outputs a standard message to let the user know that it does not know the answer.



FAQ BOT COMPONENTS

Component	Task	Technologies
Natural Language Understanding	Determine the user's intent . Are they asking a question? If so, what are they asking about?	String matching, pattern matching with regular grammars, Named Entity Recognition, Key Phrase extraction, speech act classification, etc.
Dialog Manager	Not applicable. Each utterance is treated separately from the others.	
Natural Language Generation	Return a stock response based on the determination of the user's intent .	Simple array lookup.

BASIC ARCHITECTURE

The file **faq_bot_skeleton.py** is an example of a very simple architecture for a FAQ bot that talks to a user in the Python shell. The **understand** and **generate** methods implement the Natural Language Understanding and Natural Language Generation components from the pipeline above. The bot only knows 2 questions and the user's utterance must exactly match the question to get a response. This is not very sophisticated NLP (yet).