

Comparative Study on Voice Based Chat Bots

Anusha S¹, N Vignesh Karthik², Sampada K S^{3*}

¹Department of Computer Science & Engineering, RNS Institute of Technology, VTU, Bengaluru

²Department of Computer Science & Engineering, RNS Institute of Technology, VTU, Bengaluru

³Department of Computer Science & Engineering, RNS Institute of Technology, VTU, Bengaluru

*Corresponding Author: k.s.sampada@gmail.com, Tel.: +91 9845091271

Available online at: www.ijcseonline.org

Accepted: 09/Dec/2018, Published: 31/Dec/2018

Abstract— Natural dialog conversation between humans and machines is a flourishing topic of the modern world, now show off their presence across a range of daily use devices and digital assistants. This is tried to be achieved by using sub-field of Artificial Intelligence named Natural Language Processing. The task is to be practically implemented through a series of convenient and capable categories by some of the market leaders with their successful products named the Amazon ALEXA, Microsoft CORTANA, Google ASSISTANT, Apple SIRI to quote a few. A virtual assistant put to real world usage has to perfectly emulate a human dialogue, analyze the input given by a user, represent the same correctly, convert to its process able form and formulate a relevant and appropriate response. In this paper, we review a few market leading voice based chat bots, identifying their technical aspects along with their advantages and disadvantages during the practical usage.

Keywords— Chat bots, NLP (Natural Language Processing), CI (Conversational Interfaces), ALEXA, SIRI, Google ASSISTANT, Microsoft CORTANA)

I. INTRODUCTION

Tasks, earlier performed by secretaries are now getting digitized by virtual assistants. In order to increase and improve the ease of user interaction with any system, human and artifact collaboration is necessary. The assistants also called as chat bots can interact with humans via text, voice or images. A chat bot is a software designed to simulate an intelligent conversation with a human partner. A chat bot can be considered as a question-answer system where experts provide knowledge for solicitation of user. These chat bots are mostly supported by Artificial Intelligence. Some commonly available chat bots are Amazon Alexa, Apple Siri, Google Assistant, Microsoft Cortana, and many more. These assistants perform simple tasks such as scheduling appointments, placing calls, setting alarms, and many other common day-to-day tasks. Virtual assistants can quickly evolve into capable, responsible and valuable assets to the public. Cypher, a deep learning software is present in some of these devices to deliver high-end speech recognition and clarity. The crux of this paper is to prepare a comprehensive comparison of chat bot systems. We have studied the design and implementation of several chat bots and detailed survey of these systems along with their general characteristics.

Section I contains the Introduction of “*Comparative Study on Voice Based Chat Bots*”, Section II contains the Related work done for the paper, Section III contains the general

architecture of chat bots, Section IV contains the study of voice based chat bots, Section V contains applications, Section VI contains conclusion and future scope, Section VII contains references and Section VIII contains Author’s Profile.

II. RELATED WORK

The history of chatbots can be traced way back to 1950, when Alan Turing published his paper “*Computing Machinery and Intelligence*”. This paper is widely regarded as one of the basic foundations of Artificial Intelligence and the Turing Test he proposed in this paper can be considered as a benchmark for evaluating the intelligence of a computer system [1]. ELIZA was developed in 1966 at the MIT AI Laboratory. It was a simple, text based conversation between a human user and the computer representing the Rogerian psychotherapist. Weizenbaum’s main intention in creating ELIZA was to exhibit the superficiality of human-computer interaction. The first chat bots were not actually intelligent, but were programs that had a collection of predefined set responses corresponding to specific inputs. They used pattern matching and string processing to have conversation with end users, creating an illusion of intelligence of the computer

III. GENERAL ARCHITECTURE OF CHATBOTS

The figure 1. shows a generic flow of working of a chat bot. Once the user has entered the query, the chat bot sends it to

the machine learning NLP (Natural Language Processing) Engine. The NLP returns the entities in the phrase which are then used to find the relevant data. This data is given back to the chat bot and it is converted to an appropriate response to be given to the user [2].

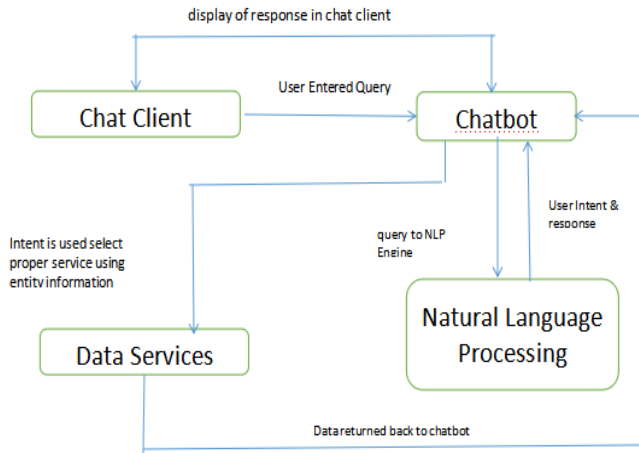


Figure 1. Generic Chat bot Work flow

The user message within the NLP engine is mainly sent into intent classification or entity classification. The working of NLP engine is as shown in figure2.

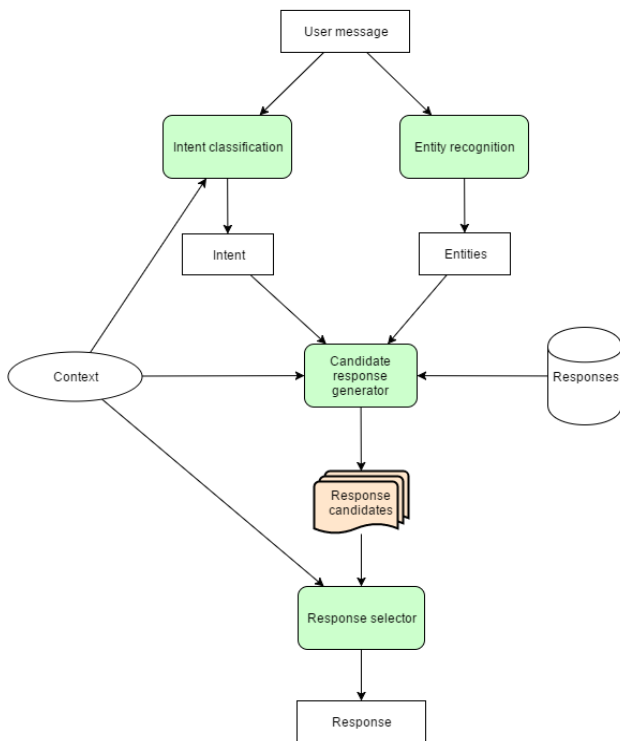


Figure 2. NLP engine

A. Intent Classification -Intent classification is basically selection of a particular intent from a set of predefined intents.

B. Entity Classification -Entity classification is majorly used for extraction of structured bits from the message.

C. Candidate Response Generator -Candidate response generator is used to do the domain specific calculations for the particular user input.

D. Response Candidates -The response of these calculations is the list of response candidates.0

E. Response Selector -The response selector is used to obtain the accurate response according to the domain. The response selector uses context of conversation, as well as intent and entities from the last user message.

IV. STUDY OF VOICE BASED CHATBOTS

There are various voice based chat bots available in the market. The task is practically implemented by some of the market leaders with their successful products named the amazon Alexa, Microsoft Cortana, Google assistant, apple SIRI.

A. Amazon ALEXA

Amazon Alexa is a virtual assistant from the Amazon available in the Amazon Echo range. The Alexa support to echo is available since 2014. Alexa is supported by the products of Amazon Web Services (AWS). Alexa collects the input from the user, locally filters the noise and sends the file to the AWS. The command issued is recognized, learnt and processed converting text to speech. Next, the file enters AWS platform and is run on the Amazon Lambda (a part of AWS that is event-driven, server-less computing platform to manage resources and run codes as a response to events). From here, required action is performed and reported back to the user.

Alexa is data-rich and well trained. Processing and understanding of text is better compared to its competitors. Alexa is a voice service inhabiting the Amazon Echo device. Alexa uses natural language processing algorithms for voice interaction. She uses these algorithms to receive, recognize and respond to voice commands. She is capable of music playback, making to-do lists, setting alarms, streaming pod casts, playing audio books, and providing weather, traffic, and other real time information. Alexa can also control several smart devices using itself as a home automation hub. The device can be woken up by an audio fingerprint, aiding it to identify natural wake-up call [3].

Alexa can be treated as step away from privacy. All voice, irrespective of its state is transmitted to the cloud. Data at cloud is freely accessible.

B. *Microsoft CORTANA*

Cortana is the voice-controlled virtual assistant for the devices built on Windows platform.

The project kicked-off in 2009 at Microsoft Speech Products team with the development of prototypes. Cortana is now available on the android and the iOS platforms too. It is integrated into the Microsoft edge (the bundled browser on windows 10) [4]. Cortana uses Microsoft Translator API (Application Program Interface) for cloud-based automated translation services. "Microsoft Bing" as the search engine. Signal Processing (Converts analogue signals to digital format), Speech Recognition (uses Dynamic Time Wrapping Algorithm and Sequence of Words for recognition of natural language), Semantic Interpreter (checks for sense making combination of words and uses "Command Mode"), Dialogue Management (Meaning extraction and Correction of Error) and Response generation (as either Text or Voice) as a part of Speech Recognition.. Cortana can learn from user's requests and interactions, personalized per user, set reminders, alarms, Refine and improve search, Compose e-mail, Multi-Sync, and many more.

On the contrary, Cortana cannot be trained by more than 1 user. It can be enabled on lock-screen enabling malicious access to the system. Cortana works at its best on Windows platform only.

C. *GOOGLE ASSISTANT*

Google assistant is the virtual assistant for the market from the Google, powered by artificial intelligence. The product was unveiled in 2016 and is available to the crowds since then. Later that year, support for the third party developers were provided. [5]. Present Google assistant differs from the latter with an increased functionality and reduced personality. Google Duplex is an extension of Google's assistant allowing natural conversations, mimicking the natural human voice. The software however, is in its development phase and is more prone to updates. Google assistant uses third party assistance to take on leading assistants, make phone calls, book appointments. It can be used on a wide range of globally accepted platforms. APIs are openly available that can be used by developers.

Google assistants consume a lot of data, and are not much user-friendly and productive as that of amazon devices.

D. *Apple SIRI*

Siri is a virtual assistant from the Apple Inc. available on most of its device Operating System. It's an off-spring of a project originally developed at SRI International Artificial Intelligence center. It is integrated into its products since 2011. Siri uses ASR (Automatic speech recognition) to translate human speech (which includes short utterances of commands, dictations or questions) into text. Using natural language processing (part of speech tagging, noun-phrase chunking, dependency and constituent parsing) it translates

transcribed text into "parsed text". Using question & intent analysis it analyzes parsed text, and detects user commands and actions. ("Schedule a meeting", "Set my alarm"). Third party web services like Open Table, Wolfram-Alpha are interfaced using data mash up technologies. They perform actions like search operations, and question answering. Speech that SIRI has identified as a question, but it cannot directly answer, is forwarded to more general question-answering services such as Wolfram Alpha.

Siri is a set of software and net-based services supported by variety of partners. The user-device is responsible for the collection of data as voice, with least noise and converting the analogue signals to digital files. The files is delegated to a voice recognition unit which uses context to understand text. The syntactical structure of text is parsed and the nouns, verbs, adjectives, etc. are filtered off. After this tedious, complicated process, the processed output is formatted and sent out as a digital file back to the host.

Siri is now available in 15 languages at hands-free mode. It has more realistic male and female voice. It can recognize songs using Shazam and Siri. It's AI driven and hence faster and smarter. Learning is personalized to users and hence, a greater user satisfaction can be obtained.

Siri can also be used for home automation and can also be used along with processors such as Raspberry PI, and what not, Siri can joke!! [6].

Siri is limited just to the apple's devices. It cannot be used on other leading platforms like the Windows, the Android and many more. Third-party software cannot be used reducing its usability. Siri violates the silent mode on its devices and still rings aloud irrespective of its modes. Internet access is a must for its functioning [7].

E. **ADVANTAGES AND DISADVANTAGES**

The table1. Summarizes the advantages and disadvantages of the digital personal assistants [8].

Table 1. Summary of Advantages and Disadvantages

Sl. No	Name	Technology	Advantages	Disadvantages
1	ALEXA	NLP+ CI	Better Understanding and processing, well trained data.	Data secrecy not available, data stored freely on cloud. Voice based
2	CORTANA	NLP	Learns well from users. Text+ voice	Permits malicious access, cannot be trained by more than one user.
3	GOOGLE ASSISTANT	NLP+ CI	Market leaders with wide range functionality, well developed APIs'. Text+ voice	Consume lot of data, not much user-friendly.

4	SIRI	NLP	Faster and smarter, available in 15+ languages Text+ voice	Dependent on IOS platform, third party software cannot be used.
---	------	-----	--	---

V. APPLICATIONS

With the rapid advancement of technology, chat bots have become increasingly important in various domains such as scientific, educational, commercial and educational. Chat bots can be implemented as intelligent personal assistants (also called virtual assistants) on mobile devices, as artificial tutors in the educational field as they can provide instant and personalized feedback to learners, and also in social networking domain for providing personalized marketing to customers.

Chat bots are a big step forward in enhancing human computer interactions. Some of the most notable applications of chat bots are as financial advisors [Credit Score Coach], providing free legal aid [DoNotPay], personalized stylist, and providing personal concierge services, offering preliminary medical advice, and many more. However the widest application of chat bots is in the field of e-commerce for automating customer service. Chat bots help to improve customer relations as well as drastically reduce human efforts.

VI. CONCLUSION AND FUTURE SCOPE

It has nearly been a decade since the virtual assistants have entered the commercial market, on the edge on reaching the 1.8 billion count mark by 2021. They are flexible and empowering. They are mostly employed to fetch information and perform tasks. The intelligence is connected and online. What not, it can even make children read too. But how safe is it to use them? Characters such as being internet connected and cloud usage pose risks for your privacy. Vital personal credentials could be available without any security. Its dual i.e. it can be used for the bad too.

It's so much useful for the people today. Though it has its own share of cons, it is still a flourishing product in the market and, whose development is a field of interest for many. Assistants are being developed to enhance its security, privacy and productivity.

Currently chatbots have limited language support. They do not support multiple languages, dialects and do not understand colloquial usage. Hence there is a great scope for removing such language barriers in future chatbots. Also, AIML templates could be improved to include more variations for the same input. Intelligent personal assistants integrate various chatbot services into one single platform and pave the way for a truly intelligent self-learning artificial entity. New approach called Ontology is proposed to model chatbots, also called OntBot. They use required mapping technology to transform ontologies to relational database and use the knowledge for further processing. This approach overcomes the need use chatbot specific language and complicated interferences [9].

REFERENCES

- [1] Turing, Alan M. (1950). "Computing machinery and intelligence." *Mind* 59(October):433-60. DOI:10.1093/mind/lix.236.433
- [2] A. M. Rahman, A. A. Mamun and A. Islam, "Programming challenges of chatbot: Current and future prospective," 2017 IEEE

- Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, 2017, pp. 75-78. DOI: 10.1109/R10-HTC.2017.8288910
- [3] Amanda Purington, Jessie G Taft, Shruthi Sannon, Natalya N Bazarova, Samuel Hardman Taylor, "Alexa is my new BFF": Social Roles, User Satisfaction, and Personification of the Amazon Echo CHI EA '17 Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, Pages 2853-2859, DOI: 10.1145/3027063.3053246
- [4] Zubairm, Paul; Bhat, Heenareyaz; Lone Tanveer, "CORTANA-INTELLIGENT PERSONAL DIGITAL ASSISTANT: A REVIEW". International Journal of Advanced Research in Computer Science; Udaipur Vol. 8, Iss. 7, (Jul 2017). DOI: <http://dx.doi.org/10.26483/ijarcs.v8i7.4225>
- [5] A. H. Michaely, X. Zhang, G. Simko, C. Parada and P. Aleksic, "Keyword spotting for Google assistant using contextual speech recognition," 2017 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), Okinawa, 2017, pp. 272-278. DOI: 10.1109/ASRU.2017.8268946
- [6] A. M. D. Celebre, A. Z. D. Dubouzet, I. B. A. Medina, A. N. M. Surposa and R. C. Gustilo, "Home automation using raspberry Pi through Siri enabled mobile devices", 2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Cebu City, 2015, pp. 1-6. doi:10.1109/HNICEM.2015.7393270
- [7] Assefi, Mehdi & Liu, Guangchi & Wittie, Mike & Izurieta, Clemente. (2015). "Experimental Evaluation of Apple Siri and Google Speech Recognition". October 2015, SEDE 2015, At San Diego, DOI: [10.1145/3027063.3053246](https://doi.org/10.1145/3027063.3053246)
- [8] V. Kępuska and G. Bohouta, "Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home)" 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, 2018, pp. 99-103. DOI: 10.1109/CCWC.2018.8301638
- [9] H. Al-Zubaide and A. A. Issa, "OntBot: Ontology based chatbot", International Symposium on Innovations in Information and Communications Technology, Amman, 2011, pp. 7-12. DOI: 10.1109/ISIICT.2011.6149594

Author's Profile

Ms Anusha S is pursuing Bachelor of Engineering (Computer Science and Engineering) from RNS Institute of Technology, Bengaluru.



Mr N Vignesh Karthik is pursuing Bachelor of Engineering (Computer Science and Engineering) from RNS Institute of Technology, Bengaluru.



Mrs K. S. Sampada has an overall 12 years of teaching and 3 years of industrial Experience. Awarded BE and Mtech Degree in specialization in Computer Science and Engineering. Currently pursuing PhD on Information retrieval and Text Mining in Visvesvaraya Technological University, Karnataka

