Comparative analysis of Virtual Personal Assistant(s)

Dr. Sachin Upadhye[1]

[1] *Assistant Professor, Department of Computer Application, Shri Ramdeobaba College of Engineering and Management, Nagpur , Maharashtra, India*

ABSTRACT

This Intelligent Agent (IA) is the goal of the realization of natural dialogue between humans and machines. In recent years of IT Technologies, the natural dialogue systems, known as interactive conversational systems are the fastest growing area in AI. Many companies have used the interactive dialogue systems technology to establish various kinds of Virtual Personal Assistants (VPAs) based on their areas and applications, such as Microsoft's Cortana, Apple's Siri, Amazon Alexa, and Google Assistant. The new model of VPAs will be used to increase the interaction dialogue between humans and the machines by using their different technologies, such as speech recognition, image/video recognition, gesture recognition, the vast dialogue and conversational knowledge base, and the general knowledge base. Moreover, the new VPAs system can be useful in other different areas of applications, including disabilities systems, education assistance, medical assistance, robotics and vehicles, home automation, and security access control

Keyword: *Intelligent agents, Microsoft's Cortana, Apple's Siri, Amazon Alexa, Google Assistant,virtual assistant, voice assistant.*

1. INTRODUCTION

To design and implementation of natural and intuitive interaction modalities is a primary research field in the Human-Computer Interaction Domain. Systems that can interact with user in their natural language are being researched vigorously at present. Virtual Personal Assistants (VPAs) are becoming more and more popular with the advent of devices and technologies like Google Home, Siri, etc. User finish tasks more effectively efficiently with the help of Spoken dialogue systems via spoken interactions. Also, spoken dialogue systems are being so incorporated into various devices such as smart TVs, smart-phones, in car navigating system [2]. Also, Dialogue systems or conversational systems can support a wide range of applications or devices in business enterprises, education, government, healthcare, and entertainment. The alternate names of Personal assistants are known as intelligent personal assistants, virtual personal assistants, digital personal assistants, mobile assistants, or voice assistants[3]. Voice-Enabled agents[10] such as Siri, Cortana, and Alexa are rapidly growing and increasing popularity in day to day life[6]. Such intelligent agents support device control—making telephone calls, starting applications, checking calendars but also support information-seeking tasks and activities including via web search and specialized services. Recent years have seen a significant increase in the deployment and use of voice-controlled Personal Digital Assistants. During the recent decades the necessary devices and technology integration (i.e. the combination of Dialog Management ,and Text-to-Speech Synthesis, Natural Language Understanding, Automatic Speech Recognition) has been the focus of extensive academic and industrial research, eventually resulting in commercial products such as Apple's Siri, Google's Home, Amazon Alexa, and so on . These products not only show the advancements or new technologies but also manage to bring the concept of an artificial personal assistant, i.e. a system that is able to (at least to some extend) understand and respond to spoken inputs, to a somewhat broader end-user level.
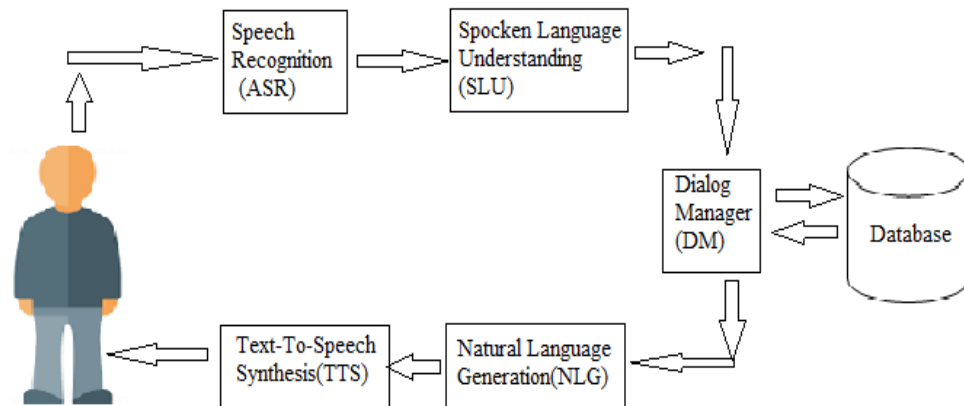
Fig -1: Dialogue System

Continuous progress and activities happened from a technological perspective as well as with respect to supported application domains [1]. Artificial services that can start, continue and handle complex interactions with human partners in their natural language. Many companies have used the spoken dialogue systems to design their dialogue system device or execute such programs, such as Microsoft's Cortana, Apple's Siri, Amazon Alexa and Google Assistant[3]. Companies used different approaches to design and improve their dialogue systems and execute such programs effectively and efficiently [10]. There are many techniques are used to design the VPAs, based on the application and its complexity [3]. For example, Google has improved the Google Assistant by using the Deep Neural Networks (DNN) method which highlights the main components of dialogue systems and new deep learning architectures used for these components effectively [3]. Google developed virtual personal assistant (VPS) ,Google Assistant for mobile and smart home devices. Unlike Google Now, the Google Assistant can engage or work in two-way conversations.

Microsoft used the Microsoft Azure Machine Learning Studio with other Azure components to improve and effective the Cortana dialogue system. The Amazon provides the advanced deep learning functionalities and for converting speech to text the services it use automatic speech recognition (ASR) and to recognize the intent of the text as well uses natural language understanding (NLU) for building highly engaging user experiences like conversational interactions for developers[3]. Siri is a virtual assistant agent with a voice-controlled natural language interface uses contextual awareness and sequential inference to help perform personal tasks or operations for iOS users. SIRI is an artificial intelligence program consisting of machine learning, natural language processing and a Web search algorithm. Apple release Siri in 2010. Bill Stasior runs Siri currently, vice president. Siri, and the technology is integrated into the iPhone, iPod touch, and iPad, and into Apple's new Home Kit home automation framework[4]. Cortana is an operating system which is oriented voice assistant available on Windows 10 devices, as well as the Xbox One console. Cortana is closely connected to productivity and their activities. But, by pulling information from Bing, it helps help answer general questions[5].

All these companies are trying to develop and create innovativeness in to the competences in several of the core technologies for their dialogue systems, such as automatic speech recognition, synthetic talking face, text-to-speech and dialog management. Next-Generation of Virtual Personal Assistants, increasing the interaction between users and the computers by using the Multi-modal dialogue system effectively with techniques including the gesture recognition, image/video recognition, speech recognition, the conversational knowledge and dialogue base, and the general knowledge base. The approach has some new methods and techniques that make this device unique, such as using as TV by using the data show or connecting the device with screen, watching TV and movies with translation language, chatting with anyone in any language, understanding body language and movements, and playing games with speech and gesture recognition; it also can be used to read face reorganization and speech expressions. Design of the Next-Generation of Virtual Personal Assistants with high accuracy and security, we added some components

to the original structure of general dialogue systems to change the general model to Multi-modal dialogue systems, such as ASR Model, Gesture Model , Graph Model, Interaction Model, User Model, Input Model, Output Model, Knowledge Base ,Inference Engine and Cloud Servers[3].

2. GENERAL STUDY OF VIRTUAL ASSISTANT

Table -1: General Analysis of VPA(s)

Primitives	Siri	Google Assistant	Cortana	Alexa
Developer's	Apple	Google	Microsoft	Amazon
Initial release	October 14,2011	May 18, 2016	April 2, 2014	Nov 2014
Operating System	iOS 5 onwards ,macOS sierra onwards ,tvOs, watchOS	Android, iOS, KaiOS(jio Phone)	Windows, iOS, Android, Xbox OS	Fire OS 5.0 or later, iOS 8.0 or later, Android 4.4 or later
Platform	iPhone, iPad, Mac, Appke TV, Apple Watch, HomePod	Android, Google Home, Android TV, Smart Speakers, Headphones, Small Speakers, Headphones, Small Displays, Google Allo, iOS	Windows 10, Windows 10 mobile, Windows Phone 8.1, Harmon Kardon Invoke, Microsoft Band 2,Microsoft Band, Android, Xbox One, Skype, iOS, Cyanogens OS,Windows Mixed Reality, Amazon Alexa	Amazon Echo, Fire TV, Fire HD, iOS, Android
Available In	English, Arabic, Chinese,Danish, Dutch, Finnish, French, German, Hebrew, Italian, Japanese,Korean, Malay,Norwegian, Portugese,Russian, Spanish,Swedish,T hai,Turkish	English, Hindi, Dutch, French, German, Italian, Japanese, Korean, Portugese, Spanish	English,Chinese,French,G erman,Italian,Japanese,Br azilian,Portuguese,Spanis h	English, German, Japanese
Type	Intelligent personal assistant	Intelligent personal assistant	Intelligent personal assistant	Intelligent Personal assistant, cloud-based voice service
Website	www.apple.com/io s/siri	Assistant.google.com	Microsoft.com/en- us/windows/cortana	Developer.amazon.co m/alexa

3. COMPARISION BETWEEN VOICE ASSISTANTS

Table 2: Comparison of voice in VPA(s)

Assistant	Voice devices	Highlights
Alexa	Amazon Echo Dot(2nd generation), Amazon Echo, Amazon Echo plus, Amazon Echo spot, Sonos one, Amazon echo show	Calla, audio Out, seven far-field Microphones, Connect home compatibility out of box,

		Improved Dolby audio, 2.5 inch display, ,multi-room music, class-D amplifiers, custom-built drivers, video calls
Microsoft Cortana	Harmon Kardon Invoke	360-degree Sound,skype Integration, Far-field voice technology
Siri	Apple HomePod	Automatic music tuning,360-degree sound,seven tweeters, Airplay 2 for multi-room listening
Google Assistant	Google Home Max, Google home, Google Home Mini, Sony LF-S50G, JBK link 300	Chromecast streaming, Individual voice recognition, google translate Dual 4.5-inch woofers, custom tweeters, smart sound room optimization, far field room optimization, far field voice control, multi-speaker wireless pairing, 360-degree audio, gesture control, splash-proof, clock display, ,10 hours of battery life, 50 watt output

4. ADVANTAGES AND DISADVANTAGES

4.1 SIRI

Advantages

- You can find nearest place like restaurants, theatre, mall, etc by just driving and giving commands through voice to your iphone.
- You can read/open text messages by giving command.
- Searching and locating something just view maps.
- Make appointments.
- Hold on call. And much more.

Disadvantages

- The main disadvantage of siri is to required Internet usage for accessing siri.
- For finding maps, user should speak in English language otherwise, it not find maps. Other language is not suitable for accessing maps in siri.

- Interpretations are not occurred properly during iphone voice command.

4.2 ALEXA

Advantages

- Alexa has many positive features and functions. Alexa understands almost everything what user say that is Alexa accepts natural language.

- Alexa is highly extensible.
- Alexa run faster as compared to other so that,Alexa saves time and improves your day by their utilization.
- Alexa seems to work smoothly and faster. For example, getting the weather report very quickly.

Disadvantage

- Alexa give people bad advice about dating
- Sound quality is not good.
- Does not able to set or cancel alarm from one room or another.

4.3 CORTANA

Advantages

- Cortana Easy to use social media sites like Facebook, Twitter etc.
- Preventing from viruses and malicious attacks by giving up-to-date security.
- Continually added new features so that no waiting time for the new release or for the update.
- This design for New generation

Disadvantage

- Flexibility is reduced.
- There is no idea about the future costs
- More cost for new PCs with Windows 10
- Cortana is Not able to upgrade oldest windows versions such as windows 2000, windows 2001 etc.

4.4 GOOGLE ASSISTANT

Advantages

- Google Assistant Can cast video to your TV.
- It Can identify different voices and capture the voice.
- It Can also cast audio to one or more devices.
- It supports Google Play Music and YouTube Music.

Disadvantage

- It doesn't have a built-in screen.
- Google Assistant Can't read or send emails or messages.
- In our main calendar, it can only recite items.
- It doesn't have an audio out port.
- Not customizable Launch phrase.

5. LIMITATIONS AND FUTURE SCOPE

Table 3: Limitation and Future Scope of VPA(s)

Virtual Assistant	Limitations	Future Scope
Siri	Siri doesn't work with weak cellular service. Strong cellular connectivity needs for cloud server interaction. Low cellular service generates limitation in the functionality of Siri. It understand only if you speak precisely some language issue means it has language issues means Bad voice recognition	Today there is a more use of digital technology. if Siri will be able to acquire with digital technology then Siri has a great future

Cortana	At a time, only one person is able to train the cortana. Only "Hey Cortana" is used in the form of keyword. when cortana starts by unlocking screen, It perform limited task only, not all the possible because it maintain your privacy. More compatible with Windows, not having much compatible about the platform using by the app developers and users-care. The iOS and android access is hard.	cortana should have come with better voice capability and increasing speed of search. Cortana should language selection menu like Hindi, Marathi, Spanish, etc. Cortana should tell me automatically when my computer went wrong or infected with virus. Cortana should able to connect with social account, to get info of friends.
Alexa	Alexa doesn't understand multiple pronunciations. By adding data of multiple accents of English this problem can be solved although storing such a big amount of data is also an issue. Many voice profiles not available for developers.	Alexa should use to require Big Data concept to solve its storage issues and problems.
Google Assistant	It does not have a plan to make money from voice assistants. Google allow to runs multiple accounts simultaneously that is why less secured. Google Assistant performs multiple tasks in a single voice command.	AI feature in Google assistant will enable it maintain data regarding the users behavior, preferences and choices and help user in day to day tasks, be it getting ready for work or in managing household chores in case of smart homes.

6. CONCLUSION

Google and Siri did best in the tasks planning and organizing. Alexa did best in online shopping. Cortana is adaptive in nature. Google assistant uses Google as a search engine which is the leader of all search engines. So, the Google assistant is the best virtual assistant.

REFERENCES

[1] P. Milhorat, S. Schlogl, G. Chollet, J. Boudy Institut Mines-Telecom Telecom Paris Tech & Telecom Sud Paris Paris,France A. Esposito

2nd University di Napoli Department of Psychology Caserta, Italy G. Pelosi Integration i & Sistemi Via dei Granai di Nerva Rome, Italy "BUILDING THE NEXT GENERATION OF PERSONAL DIGITAL ASSISTANTS", 1st International Conference on Advanced Technologies for Signal and Image Processing - ATSIP'2014 March 17-19, 2014, Sousse, Tunisia

[2] Ashay Argal, Siddharth Gupta, Ajay Modi, Pratik Pandey, Simon Shim, Chang Choo Computer Engineering Department and Electrical Engineering Department San Jose State University, One Washington Square, San Jose, California – 95192-0080, "INTELLIGENT TRAVEL CHATBOT FOR PREDICTIVE RECOMMENDATION IN ECHO PLATFORM", USA ashay.argal@sjsu.edu; siddhart.gupta@sjsu.edu; ajay.modi@sjsu.edu; pratik.pandey@sjsu.edu; simon.shim@sjsu.edu; chang.choo@sjsu.edu

[3] Veton Kepuska Electrical & Computer Engineering Department Florida Institute of Technology Melbourne, FL, USA vkepuska@fit.edu, Gamal Bohouta Electrical & Computer Engineering Department Florida Institute of Technology Melbourne, FL, USA gidris2015@my.fit.ed,

"NEXT-GENERATION OF VIRTUAL PERSONAL ASSISTANTS (MICROSOFT CORTANA, APPLE SIRI, AMAZON ALEXA AND GOOGLE HOME)"

[4]   Sheetal Reehal, MCA III Year Student SIES College Of Management Studies, University of Mumbai, India,"SIRI –THE INTELLIGENT PERSONAL ASSISTANT ",International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 5, Issue 6, June 2016

[5]   Microsoft Cortana Intelligence,https://azure.microsoft.com.

[6]   Hyunji Chunga, Jungheum Parka, Sangjin Leea," DIGITAL FORENSIC APPROACHES FOR AMAZON ALEXA ECOSYSTEM", Center for Information Security Technologies (CIST), Korea University, 145 Anam-ro, Seongbuk-Gu, Seoul, 08421, South Korea.

[7]   S. Arora, K. Batra,and S. Singh. Dialogue System: A Brief Review. Punjab Technical University

[8]   B. Marr. The Amazing Ways Google Uses Deep Learning AI. https://www.forbes.com

[9]   Yoshihisa Ishihara and Kazuki Kobayashi,"PROJECTION MAPPING OF BEHAVIORAL EXPRESSIONS ONTO MANUFACTURED FIGURES FOR SPEECH INTERACTION",2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) Lisbon, Portugal, Aug 28 - Sept 1, 2017.

[10] Julia Kiseleva1, Kyle Williams, Ahmed Hassan Awadallah, "UNDERSTANDING USER SATISFACTION WITH INTELLIGENT ASSISTANTS" Eindhoven University of Technology, j.kiseleva@tue.nl ,Pennsylvania State University, kwilliams@psu.edu, Microsoft, {hassanam, aidan.crook, izitouni, tasos.anastasakos}@microsoft.com.