

# Personal Assistant Using Artificial Intelligence

Rohan Dhoyda<sup>1</sup>, Rushabh Kothari<sup>2</sup>, Shruti Iyengar<sup>3</sup>

<sup>1,2,3</sup> Dept of Electronics and Telecommunication

<sup>1,2,3</sup> K. J. Somaiya Institute of Engineering and Information Technology  
(KJSIET), Mumbai,

**Abstract-** This paper is about intelligent personal assistant- Jenny which is a proof of concept to great extent of Artificial Intelligence. Jenny is capable of recognizing speech with UN-EN culture and integration with Microsoft speech recognition. Jenny can control hardware processing unit with the use of Arduino. Jenny can be programmed to perform various tasks such as controlling Electronic Home appliances, sending SMS, control any .exe file of windows, retrieving websites from World Wide Web, suggesting medicines from observed symptoms. Jenny is like an interactive assistant with feature of Microsoft speech development kit (SDK) interfaced with medical kit to distinguish between medicines, setting up a reminder to patient regarding the intake of doses at right time.

**Keywords-** Speech Development Kit, Speech Recognition, Microsoft Visual Studio, C# language, Arduino Uno.

## I. INTRODUCTION

Intelligent personal assistant (IPA) is becoming a useful feature on mobile in day to day life. IPA are used to perform functions such as retrieving mails, calculation of numbers, stock market update, calendar management, weather update. Users can also browse on web through voice control and chat with assistant. The driving force behind the speech enabled IPA is the concept of Artificial Intelligence. The IPA uses the text to speech synthesis and various speech libraries like SDK, speech recognition system [1]. While these applications seem attractive for user, it is challenging to integrate personal assistant with hardware components such as controlling the fan, AC and several other electrical appliances. The personal assistant such as Siri, Bixby, Cortana, Google fails to control electrical appliances. The problem arises due to electrical appliances operating on 230V, 5A which cannot be interfaced with software. The personal assistant such as Siri, Cortana uses hybrid unit selection system where data recording occupies memory for command. Also, these personal assistants are not open source.

The front end of Siri can only be run on IOS platform which eliminates the option of running it on Windows and Android platform. The personal assistant -Siri is not able to monitor IOT sensors and control it.

The above problem can be solved with the technology module illustrated below.

Objectives: -

- To connect electronic appliances easily with the help of voice this is useful for physically disabled people.
- To diagnose symptoms by providing medical prescription.
- To program the personal assistant as per user convenience since Jenny is dedicated to open source libraries.
- To keep someone updated regarding daily news various advancement in science and technology.

## II. LITERATURE SURVEY

The market for people who use cellular interface is increasing day by day. It is evident that the use of mobile phones is not only limited to urban areas but rural areas as well. Thus, this fact can be used as an excuse to increase the level of technology in the department of Artificial Intelligence. Artificial Intelligence is an integration of human knowledge and machine intelligence to make people's life easy and comprehensive. The oldest technology in this field was virtual Chat-Box. Chat-Box HAL was created by Zabaware Inc [4]. Chat-box Kari basically operates like a virtual girlfriend [5]. DRAGON NATURALLY SPEAKING SOFTWARE has three roles: dictation, text to speech and command input. Syn Virtual Assistant was used for voice services which is an online platform. Briana communicates through Wi-Fi on a software application. There are various other Personal Assistance software like Siri, Google Now and Cortana. Our main aim is to add the functions of all such personal assistants combining our feature that includes controlling hardware.

## III. TECHNICAL OVERVIEW

The personal assistant – Jenny is implemented on software Microsoft Visual Studio 2017 which uses .net framework of 4.5.2 enabling speech recognition.

Microsoft Visual studio is the application-based software which is used to develop websites, apps for different organization.

The software provides tools necessary for the personal assistant to become an open source. The programming language C# is used on .net framework. C# is language similar to C++ with strength in creating web desktop applications. The Visual studio software enables the option of using four speech libraries such as Text To Speech synthesis, Speech Run time environment, Microsoft speech .DLL, SDK. The four speech libraries are embedded into visual studio package. Text-to-speech synthesis (TTS) is an essential component whose goal is to produce expressive, intelligible speech that is indistinguishable from human speech [2]. The quality of speech synthesizer depends on the confidence and similarity of the human voice. Speech recognition develops technologies that allow translation of speech into text by computers. Speech development kit generates the output speech of Jenny in number of voice and in accent of different countries. The project is created on software of visual studio and configuration files in it are edited to control hardware. To control hardware using Jenny, it is mandatory to interface the configuration of the program to serial port using Arduino and relay board.

The relay board consists of design which includes four relays on single chip that can be increased as per user convenience. The design of relay board is developed on software called Dip trace. The printing of design is done on PCB followed by etching, drilling and soldering. The relay board enables the control voltage of 230 V and turns it down to 5V which is interfaced with Arduino. The device Arduino is electronic platform which can direct both hardware and software. Arduino consists of microcontroller which is programmed to read inputs such as – turn on the light sensor and converts the output which can drive sensors and motors. Arduino is an open source tool available for extension by professional programmers. It consists of 14 Digital I/O Pins (6 Pins of which can be used as PWM Outputs). The Arduino Uno [3] is operated at 5V. The limited I/P voltage for Arduino is 620V. It has 6 Analog Input Pins. The DC Current per I/O pin is 40mA and that for 3.3V pin is 50mA. The Arduino UNO board also includes specification such as EEPROM-1Kb, SRAM2Kb, Flash Memory-32Kb. Arduino Uno operates at a clock speed of 16MHz. It contains everything needed to support the microcontroller directly to a computer with the USB cable or power it with DC/AC adapter to get started. The language can be expanded by adding more C# libraries. The advantage of Arduino such as cross platform, clear programming environment, and extensible hardware makes it popular to use

#### IV. IMPLEMENTATION AND WORKING

The software Microsoft visual studio is installed on the personal computer with all of its speech libraries. The features supported by Visual studio 2017 software such as speech synthesis and speech recognition are achieved. The coding is done in such a way that human speech delivered will be processed depending on the confidence level set in the code which is 0.75. The speech above this level will be intercepted by commands in the code. The command such as Speech On will be converted into text by speech recognition-engine.

The answer for these commands mentioned in the code will be cited. The output will be converted back in the form of speech by text to speech engine using the voice and accent of specific nation. This process consists of front end and back end.

##### A. Front End

The front end converts text containing symbols and abbreviations into words. This can be called as text normalization. The front end assigns transcriptions to each word and divides it into several units like sentences. Frontend of various platforms can be directly linked to Speech DLL and hence we can build this for any platform.

##### B. Back End

The back end often refers to the synthesizer that converts linguistic objects into sounds. The command responsible for hardware connection and management is converted using the above process. The waveform generated will be processed in Arduino through relay board. The command for the connection of electrical appliances requires 230 V which can destroy the processor. The high voltage and current can be cut down to 5V through relay board.



Fig. 1 Arduino Uno Board

The relay board is connected to Laptop through USB port. The appliances such as fan, light, mobile, computer relates to several ports of relay board. According to the connection established, the following appliances can be switched on and off. The user suffering from a disease having relevant symptoms with no knowledge about the medication can approach Jenny where it will act as doctor and depending on the command received from the user the proper medicine will be signalled through LED on the medicine board interconnected with relay board. Jenny can also suggest medicines depending on the symptoms observed. The final product implemented does not consist of Speech Development Kit. The Arduino Uno board is connected with Laptop on serial port 1 and 3. The SMS is sent to mobile through the use of SMS module connected with user computer indicated by the block Mobile in the following diagram presented below. The SMS command given by the user should specify the mobile number on which SMS is to send. The command will use the logic mentioned in the coding and will execute the relevant statement, pass the data to SMS module connected on a gateway allotted and accordingly SMS will be sent to desired number.

The charging and discharging of the mobile can also take place when the command “Charge the mobile” is spoken out via MIC. For the better sound quality, an additional Bluetooth speaker can be connected with laptop.

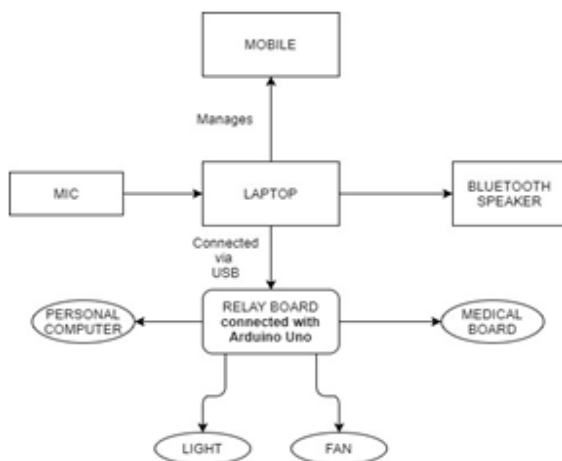


Fig. 2 Functional Block Diagram

The project of Jenny is implemented by several commands which are divided for different applications. Each command can be executed when spoken out in Mic with the application of Jenny running in the background. The command divided is classified as device control, online application, and chat box, weather report. The commands that are not properly pronounced and if additional words are introduced in the command will not lead to execution of desired task. More

commands can be executed by making changes in the coding pattern since due to Microsoft DLL extension file, the libraries are open source. The following table illustrates list of several such command.

TABLE 1  
TOP 5 DOMAINS WHICH INCLUDES SEVERAL COMMANDS OF JENNY

Chat Box	Device Control	Online Applications	Weather	Command
Speech On	Turn On Fan	Who is the Prime Minister	What's the temperature	Time and Date
How Are You	Charge the Mobile	Show My Mailbox	Atmospheric Pressure	When were you created
Hello Jenny	Turn On Pc	Show me Stock Market	What's the humidity	What is (no) plus (no)

V. RESULTS AND DISCUSSION

Whenever the user finds a reply to be incorrect, he may teach Jenny the correct reply for the same question. This is very useful in applications as each user might want to have different answers for questions depending on personal choices. For example, when the person is not in a meeting, he might not want his secretary to reveal the reason of his absence in that slot but as information is available with Jenny, she might disclose it to a person asking to schedule a meeting for the same time slot. At such times, the user could teach Jenny what to talk and what not to. Jenny replies averagely about 70% of the times. Humans are bound to make mistakes often. They misinterpret a statement or a question and may reply irrelevantly. Jenny is a piece of artificial intelligence, which offers 65-70% accuracy.

The various commands set are illustrated below:

1. Speech On: Jenny will start listening to the speech. The time taken by Jenny to respond is 2s with 93% accuracy.
2. Hello Jenny: This command will enable the voice recognition with an accuracy of 90%. The time taken by the software to perform the command is 2s.
3. Introduce yourself: This command will reveal the main purpose of Jenny with its use. It is 50% accurate with a time span of 4s.
4. How are you? : This command tells about the present state of Jenny. The ‘How are you’ command is 60% accurate which takes 3s.
5. How is weather today: This command gives the information about temperature, humidity and atmospheric pressure of the surrounding with an

- accuracy of 55%. The time taken by Jenny to execute this command is 3s.
6. Who is the prime minister of (choose any country): This command retrieves the data from WWW and displays it on Google Chrome at the rate of 5s. This command is 80% accurate.
  7. What is (no) plus (no): This command adds two numbers and displays the logic on screen with an accuracy of 90% at the rate of 2s.
  8. Show me the stock market: This command gives the stock market price of NSE and displays it on Chrome. This command takes 3s to show the results with an accuracy of 86%
  9. Show my mailbox: This command retrieves mails from GMAIL servers within 5s with an accuracy percentage 75%.
  10. Show time and date: This command gives the present date and time. This command is 95% accurate and it consumes 3s to execute the result.
  11. Turn on/off fan, light, AC, Mobile, and PC: This command will perform operations such as switching on and off the devices at the speed of 13s with 70% accurate results.
  12. Close it: This command closes all the tabs from Google Chrome within 1.5s. This command executes with 90% accuracy.
  13. Speech Of: This command will lead Jenny to disconnect speech recognition and synthesizer at an average rate of 11s with 65.5% accuracy
  14. Goodbye Jenny: This command will stop the execution of the program immediately within 2s. This command is 92% accurate.

The result helps in finally comparing the personal assistant Jenny with the other assistants which uses different platforms and different technologies for their implementation.

## V. CONCLUSION

Jenny is an intelligent personal assistant supporting multiple input and output languages with the great extent of future scope. People suffering from paralysis – except the one which restricts speaking – as well as blind people can use this application. In case of an emergency, the person can give commands to Jenny to send an SOS to its relatives. Anything and everything from the internet can be retrieved or found out with the help of an application. To know about the update and news of the stock market, the people working in such sectors can use this application. Household tasks such as controlling the household appliances like lights, fans, AC, washing machine, smart televisions can use this application. This will be very useful application as the time is being saved. Old

people, pregnant women and physically disabled people can do common household tasks with the help of this application.

## REFERENCES

- [1] H. Zen, K. Tokuda, and A.W. Black, "Statistical Parametric Speech Synthesis", *Speech Communication*, vol. 51, no. 11, pp. 1039-1064, 2009.
- [2] Tim Capes\*, Paul Coles, Alistair Conkie, Ladan Golipour and Quiong Hu. "Siri On-Device Deep Learning Guided Unit Selection Text-to-Speech System". *INTERSPEECH 2017*. August 2017, Stockholm, Sweden.
- [3] M. Banzi, *Getting Started with Arduino*. "O'Reilly Media, Inc.", 2009.
- [4] Zabaware Inc. (n.d.). Ultra-Hal can hold conversations with you. Available: <http://www.zabaware.com/assistant/index.html>. Last accessed 2nd March 2013.
- [5] Lhandslide Studios. (2012). *Advanced Virtual Girl with Artificial Intelligence*. Available: <http://www.karigirl.com/>. Last accessed 22nd April 2013.