BOT-O- PEDIA - Learning Simplified

S.Sobitha Ahila, D.Sivakumar, Ashwin Suganthan, Arunkumar.M , Manish Kumar S

Abstract: ***A standalone teacher Bot that can be interacted through the user's voice. The Bot works as a teaching assistant which has the ability to respond for the voice input. The Bot has features like recognizing the voice, reading the text and searching in Wikipedia. The 21st-century people are in the era of intelligence and the communication between humans and Bot goes to the next level through intelligent technologies. The foundation of human-computer interaction is robot teaching. It can help the user in doing tasks like querying database, searching Wikipedia etc.. A customized teaching assistant is not implemented so far for primary kids and visually challenged people. Any visually challenged or a primary school kid can interact with the Bot The bot has the ability to understand the Human language and it is coded with Language Processing. The traditional teaching methods for physically challenged people can be made much more effective through Bot teaching methods. Bots have the ability to remember, to learn and to perceive which helps the intelligent system in teaching [8]. There are three main directions; teaching based on recognizing the speech, searching for the content from the local database as well as in Wikipedia, converting text into a voice output. The user interacts with the system through voice command, the system in-turn returns the output through the speakers attached to the system. The core concepts used in the project are speech-to-text conversion (for understanding user input) and text-to-speech conversion (for giving output to the user), reading the document in the local database, connecting with Wikipedia. The system can search Wikipedia and give the results through voice. It can help the visually challenged to connect with the world by giving them access to Wikipedia through their voice. A raspberry pi module is chipped with android things operating system and is also connected with a mic and a speaker for output. The raspberry pi is paired with the mobile application of android things and the module is modified in such a way that it continuously waits for the user's voice input to respond accordingly[6]. The voice-controlled Bot can be used for visually challenged students in primary school.***

Index Terms: Bot, Natural Language Processing, Customized Teaching Assistant

I. INTRODUCTION

The evolution of humans increases the expectation of getting computers even more smarter. With the development in the teaching industry, teaching assistant Bot is intelligent, multi-functional and collaborative. Even then, the ability of robots is insufficient when compared with human workers. However, efforts are being made to improvise the robot efficiency every day. An extraordinary child being naturally intelligent can still improvise his intelligence with the assistance of teaching Bot. The use of Robots is widely needed in case of human absence. They can guide users and can execute a simple task, like starting a web search.

The traditional teaching strategies for physically challenged individuals like blind, deaf, disabled is not any longer effective since the method of teaching takes a protracted method and may be created rather more effective through larva teaching strategies. Bots have the power to recollect, learn and understand by that a larva will simply teach and reach the goal. the primary Chatbot was developed in 1966 by Joseph Weizenbaum; most Catboats still use easy pattern matching and supply pre-prepared answers. Natural human-computer interaction is a vital interface to form a friendly collaboration with intelligent robots and humans. mechanism teaching through voice recognition could be a new method of teaching associate degreed it conjointly provides an interactive interface to the user [7][9]. The mechanism management instructions are often extracted by employing a text understanding technique. Easy teaching tool of a mechanism manipulator is one in all the essential technologies for economical human-robot collaboration.

Since we expect intelligent robots to participate widely in the society of the near future, effective interaction between them and humans will be essential. The Robot can also take the role of a tutor to give a lecture or to guide learning activities through pre-programmed code. As an educational tool, the robot teaching assistant has been researched, e.g. responding to student queries via email by posting short prompts and confirmations in an online course.

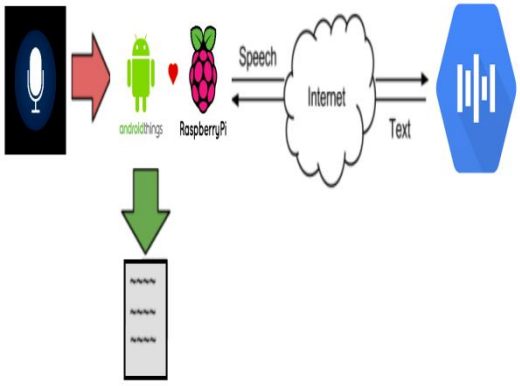


Fig 1 Cloud Speech to Text Architecture

Revised Manuscript Received on December 22, 2018.

S.Sobitha Ahila*Associate Professor, Computer Science Department, Easwari Engineering College, Chennai, India.*
sobitha.ooviya@gmail.com

D.Sivakumar*Professor, Electronics and communication department, Easwari Engineering college, chenai, india*dgsivakumar@gmail.com
suganthanmak@gmail.com,manishrockz161197@gmail.com

Ashwin Suganthan , Arunkumar.M ,Manish Kumar S,*UG Student, Computer Science and Engineering, Easwari Engineering College, Chennai, India* arunaaron85@gmail.com

Retrieval Number:F13370476S519/19©BEIESP

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The Bot or a standalone teaching assistant has the ability to respond for the voice input. The Bot has features like recognizing the voice, reading the text and searching in Wikipedia. There are three main directions; teaching based on recognizing the speech, searching for the content from the local database as well as in Wikipedia, converting text into a voice output. Cloud speech to text api is used (Fig 1.1) to convert the voice into text .

It can help the visually challenged to connect with the world by giving them access to Wikipedia through their voice. The voice-controlled Bot can be used for visually challenged students and primary school kids. The robot as a tool has been most widely applied to robotics education in which students learn engineering and technology by building and programming robots. We define this mechanism of providing support, assistance and active engagement over time as personalization. The robot teaching assistant is more expected to work in a variety of assignments, especially when it is difficult for a teacher to do it by himself/herself.

Educational robots are often associate degree great tool in early and education. in keeping with a journal on new views in science education, instructional robots will facilitate to develop skills that promote autonomy and assist their integration into society[10]. Social and private skills may be developed through instructional robots. As a substitute for college lecturers, Bots were able to work with old faculty youngsters so as to develop programs and improve the children's social and social skills. in addition, problem-solving skills and creative thinking were improved through the information that the larva has in these various fields. This will result in associate degree exaggerated ability in drawback resolution, reasoning and coming up with in usually developing educational institution youngsters. This study suggests that instructional robots within the room will result in associate degree improvement in visuospatial memory and mental coming up with still.

II. LITERATURE SURVEY

Guanglong Du et al [2018] planned, ”Online automaton Teaching with Natural Human-robot Interaction”, a natural method of online teaching exploitation robots. This online teaching automaton acknowledges the human hand motion exploitation kinetic motion detector and mechanical phenomenon activity unit. These sensors help to capture the gestures that a person's makes. Moreover, the speech may be processed through electro-acoustic transducer and therefore the speech gets reborn to text through Microsoft SDK. Interval Kalman filter and Improved Particle Filter area unit wont to estimate the position and therefore the orientation of the hand. The Audio-Video -Fusion supported Text (AVFT) helps to integrate each speech and gesture[14]. the most formula known as most Entropy formula is employed to convert the text to directions. The complicated text may be created perceive by the robots by fusing the speech text and therefore the gesture text. Firstly, the thorough text command management corpus is inputted. Then, a good automaton management command is intended. Finally, they thought-about the transformation of human’s linguistic communication text into automaton management commands. The achieved result's that, with static gestures and basic speech commands that deals with fine teaching have a lot of accuracy. Whereas with dynamic gestures and basic speech commands that deals with rough teaching has less accuracy as a result of the automaton finds it tough to form action as per the user’s directions. Kalman Filter and Particle Filter can estimate the states of linear models and upset extremely nonlinear mathematician models severally, it might be a decent methodology to integrate them to scale back the procedure complexness[1]. However, the quality Kalman Filter cannot upset such things with applied mathematics parameters and inaccurate dynamics. Moreover, the quality Particle Filter might cause severe particle degeneracy and dramatic loss of diversity in particles. the problems during this plan is that through static input to the automaton, it produces correct results however with the dynamic input, the accuracy to perform a selected task decreases .

Tokuda et.al [2017] proposed, ”An intelligent Virtual Assistant Using Raspberry Pi”, a Voice Command System as an Intelligent Virtual Assistant (IVA) that can perform numerous tasks or services for an individual. First half is liable for changing numbers and abbreviations to a word format. This is also referred to as normalization of text. Second part involves the signal to be processed into an understandable one. Initially, the speech is given as an input through microphone where the speech is processed to text through NLP (Natural Language Processing). It then picks the hotword from the text string and if the name of the command matches with the keyword then the actions are performed as per the given instructions. Finally, the output is provided as speech through text-to-speech converter using Optical Speech Recognizer (OCR). This output is transmitted via the speakers that area unit connected to the audio jack of the Raspberry Pi. The achieved result is that, word error rate is reduced by 10% to 17% which leads to utmost high accuracy. It can help the visually impaired to connect with the world by giving them access to Wikipedia, Calculator etc., all through their voice. The issue involved here is that, based on the slang inputted by the user might vary with the universal pronunciation and the results might differ accordingly.

Joe Saunders et.al[2016] proposed, “Teach Me–Show Me”, End User Personalization of a Smart Home and Companion Robot**,** the use of assistive robots smart home environments. The paper mainly focuses in involving the robots to provide motivation and support in terms of reablement defined as Support people ‘to do’ rather than ‘doing to / for people’ and co-learning—working together to achieve a particular goal. This paper uses a commercially available robot Care-o-bot. Initially, the house Ontology is instantiated in a MySQL database. The robot actions are defined through procedural memory where the post and pre-conditions are given priorly. The rules are encoded as SQL statements. It uses Robot Operating System navigation using its laser range-finders to update a map of the house in real time and can thus navigate to any given location whilst avoiding obstacles and by replanning routes. High-level commands are sent via the ROS “script server” mechanism

and interpreted into low-level commands using the robot software. The preconditions would be automatically encoded by the teaching system as SQL statements. The achieved result is that, out of 20 participants of the age over 45 ,17 people were comfortable with the assistive robots which understands the behaviours of the user and then suggests the actions with the post conditions. The SUS scores for the "Show Me" interface ranged from 67.5 to 80.The issues in this proposed idea is that, when the assistive robots are trained with typically partial actions such as "when it's time for bed do..." or "if I'm making dinner do ....",it waits for the user to complete and does not provide any post condition results.

Ingo Lutkepohl et.al [2017] planned, "The Curious golem – Structuring Interactive golem Learning", focuses on targeting undisciplined users, United Nations agency area unit supported by mixed-initiative interaction mistreatment verbal and non-verbal modalities. System development follows associate degree interactive analysis approach associate degreed it presents each an protractile and event primarily based interaction design to appreciate mixed-initiative and analysis results supported a video-study of the system. The hardware used for the interaction situation may be a bi-manual, fastened setup that affords advanced grasping and manipulation capabilities. to attain a golem capable of interacting with an individual's in an exceedingly natural approach, variety of perception and dialog parts area unit required additionally to the golem management software system[3]. The system consists of 3 major parts: sensory activity analysis, task generation and dialog-oriented task execution. The dialog shaping strategy happens within the sensory activity and initiative components of the "system-level" lane. Visual analysis creates events describing fascinating regions that area unit then hierarchical and planned as new dialog actions. The most purpose here is that coordination between numerous parts and progress in subtasks is achieved through the Active Memory, that stores task descriptions and notifies collaborating parts after they area unit updated throughout execution. Sensory activity analysis is multi-model, together with speech, vision and interception. Speech is recognized by the 'Esmeralda' speech recognizer, with speaker freelance acoustic model and a settled speech understanding part. The analysis has been performed as a video-study, wherever associate degree interaction with associate degree full-fledged check person was recorded and shown to the check subjects. Throughout the video, questions area unit posed to the topics. The queries area unit asked once the golem has acted however before the instant wherever the recorded person associate degreeswers to ensure an unbiased answer. The distinction within the responses for the assorted things will offer US insight on the overall effectiveness of the steerage and also the variability within the responses indicates whether or not the constraints increase sure thing or not.

Hyun Min Do et. al [2017] projected, "User-friendly Teaching Tool for a automaton Manipulator in Human automaton Collaboration", associate intuitive teaching tool which will be connected to the automaton and may accurately teach motions while not being laid low with device noise. The device consists of jog-dial and buttons for implementation. so as to avoid the device noise from the external forces, they projected associate intuitive teaching tool that uses the higher than parts to maneuver the automaton wherever the human needs to maneuver in an exceedingly explicit direction. This teaching tool consists of 3 components particularly motion setting unit, motion operation unit, automaton standing show unit[11]. The setup is employed to change between the automaton teaching mode and operation mode and thru the motion operation unit, the intuitive jog motions of the automaton like position and orientation teaching may be created perceive as per human's intention. The achieved result's that the device is side to the automaton and effective teaching is completed block the distortion from the analog signals made by the sensors. The problem is that the protection functions of the automaton aren't developed with high exactness rate. The automatons connected with this teaching tool have to be compelled to be increased so as to form it an advertisement learning tool.

Jiansheng Liu et.al [2017] planned, "An Intelligent Personal Assistant Robot: BoBi Secretary", AN intelligent personal mechanism named BoBi secretary. Owner will decision BoBi to remodel from a box to a movable mechanism and it'll perform several functions like humans, like moving, talking, singing, dancing, conversing with folks to create folks happy, enhance people's lives, have a good time with folks, connect folks with the surface world ANd assist and support folks as an intelligent personal assistant. It will do all works that a secretary is doing as well as programming of works, schedule reminders, causing emails, occupation phones etc. BoBi has 3 main functions: intelligent meeting recording, bilingual interpretation and reading papers. Within the service mechanism systems it's necessary to supply AN intelligent voice interaction perform that's ready to assist ANd support folks as an intelligent personal assistant. it's 3 main functions: intelligent meeting recording, bilingual interpretation and reading papers. BoBi records folks's voice from electro-acoustic transducer once people speaks with BoBi, and so sends the voice knowledge to a speech recognition server via internet API and obtains a recognition result text. we tend to apply a Voice Activity Detection (VAD) formula to find the speech and therefore the non-speech frames, and find the stop of the people's speech by the non-speech frames. once the stop time of the people's speech is longer than a threshold, BoBi stops the voice recording and sends the voice knowledge to the Baidu speech recognition server to get a result text[4]. at the moment BoBi replies the obtained answer by synthesizing the solution text into somebody's sounding speech victimization the Microsoft Windows Phone Speech Synthesizer and taking part in the synthesized speech.

III. SYSTEM DESIGN

The goal is to achieve a Bot that can recognize human language (in this project english is the only language that is used for processing) and respond to the user like reading any document, answering for predefined questions etc.. The recognized human voice is then converted into the syllables for a better understanding of the input voice. The

Retrieval Number:F13370476S519/19©BEIESP

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Bot is handy and can be carried anywhere. The Bot has the capability of working offline but still for refined search on the internet it can also be connected with the internet. The bot has given the capability of searching in Wikipedia and taking out the required contents which are then read by the bot. The main application of the bot is to make learning simplified for primary school kids as well as helping the visually challenged people to access the internet through Wikipedia.

A. SYSTEM ARCHITECTURE

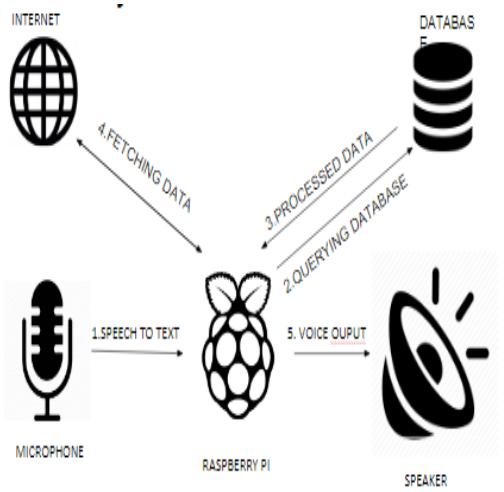


Fig. 2 System architecture of Bot-O-Pedia

The system design includes an electro-acoustic transducer, raspberry pi, speaker, internet association and database. All of these parts are bridged along to simulate a learning assistant that is run victimization associate automaton OS known as “Android Things”. The Android Things software is installed in an Raspberry Pi 3 module. We use an ARM processor which is specifically developed for Raspberry Pi module in order to achieve a Bot that works with efficiency. It includes four superior and high-performance ARM Cortex-A53 processing cores running at 1.2GHz with 32kB Level 1 and 512kB Level 2 cache memory**.** Initially, speech is given as input via electro-acoustic transducer and this speech is regenerated to text that is recognized by the Cloud API. In line with the given input, the Bot queries the database and also the processed information is received from same. If the question needs the utilization of the Internet, it fetches the information from the web. The fetched knowledge is scanned as output with the assistance of text to speech API that converts text to voice.

B. FUNCTIONAL ARCHITECTURE

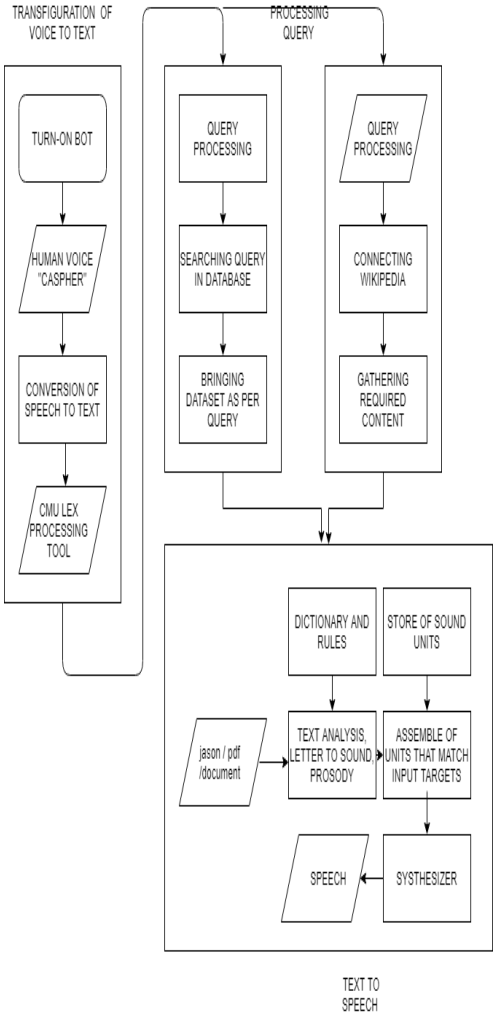


Fig 3 Functional architecture of proposed system

Before the Bot begins to process our queries, it has to be booted up. The input ‘Caspher’ is given in order to wake up the module. Once the Bot is absolutely setup, it begins with the boot up line “I’m ready”. This means that the Bot is prepared for performing any given operations. Once the input or question is given to the Bot, its recognized by Cloud speech API and it converts it into text using lex tool. The output obtained from lex tool is fed as the input for processing the question. Bot searches the local database and responds to the question consequently. It additionally permits the user to connect to the web if necessary which can be done using the module. The user would then get the privilege to access the web and procure outputs for his or her queries. The respective output fetched by the Bot is delivered to the user as a voice note.

The Hardware and Software components used are as follows:

C. Hardware:

1. ARM Processor

The aim is to develop a teaching assistant that has the flexibility to grasp the human question and answer consequently. this can be performed using raspberry pi 3 model B.. The module acts as the basic hardware to perform the above functions with efficiency. The microcontroller is connected with a speaker and a mic so as to obtain the input from the user as well as to respond to the query of the user.

The chip on the microcontroller includes 4 high performances ARM. The hardware is basically is constructed with ARM processor running at 1.2 GHz with 32kb level 1 and 512kb at level 2 cache. This advanced architecture provides efficient use of Raspberry Pi, thus enabling the Bot to work with high productivity.

2. Additional Hardware Components

The mic which is referred to as the electro-acoustic transducer, works as an input device that helps in capturing the human voice. It detects any input that is given in the natural language (i.e English in this case). This output is transmitted via the speaker which is connected to the audio jack of the Raspberry Pi.

D. Software:

1. Android Studio

Android Studio plays a vital role in developing the application. With Android Things being the OS used here, all Android applications are supported. Android Studio is based on the IntelliJ IDEA, a Java integrated development environment for software and incorporates its code editing and developer tools.

2. Cloud API

Cloud Text-to-Speech API permits developers to form a natural-sounding, synthetic human speech as playable audio. The Text-to-Speech API converts text or Speech Synthesis Markup Language input into audio data like MP3 or WAV files.

3. Pocketsphinx

CMU Sphinx, also called Sphinx in short, is the general term to describe a group of speech recognition systems. Sphinx featured feasibility of continuous-speech, speaker-independent large-vocabulary recognition.

IV. SYSTEM IMPLEMENTATION AND PERFORMANCE ANALYSIS

A standalone Bot consists of a Raspberry pi, mic, and a speaker. The raspberry pi is installed with Android things operating system. It is then paired with the mobile application for android things. This allows the user to make any changes in the application rather than changing it in the Raspberry pi. The raspberry pi is configured in such a way that the mic which is connected perpetually waits for the input. The Bot waits for the user to call it through its name "Casper". Once the Bot wakes up on hearing "Casper", it waits for the user to give the input. The Bot acknowledges human voice which is achieved by using Cloud api. The user has the privilege to give various commands like reading the pdf or a document. The user can also raise queries which the Bot will be able to fetch the answers for, if it is provided with internet access. This output is displayed to the user in the form of voice (i.e) the Bot reads out the content to the user.

A. IMPLEMENTATION OF MODULES

The proposed system consists of the following modules.

1. Transfiguration of Voice to Text.
2. Processing query using Local Database.
   Processing query using Internet.
3. Text to Voice.

B. Transfiguration of Voice to text

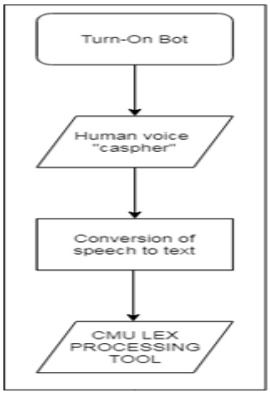


Fig 4 Work flow of recognizing the voice input

The BOT is initiated with android things and the steps to be processed are carried over. Once the Bot gets initiated it waits for the user to give the command - "Caspher". Figure 4.1 show how the bot goes on its initialization phase followed voice recognition. The word Caspher is employed to wake up the Bot. This activates the Bot and it is the user's way of instructing the Bot to be prepared for performing certain functions. Once it wakes up, it waits for the user to give the input, so that further processing like reading Documents and other actions can be processed. The bot can perceive human language. In this model, the bot can interpret English. It converts the input which is given in the form of speech into text form with the assistance of Cloud API. For better recognition of voice, the Bot is coded with CMU lex tool which converts the voice into syllables. Lex is commonly used with the Yacc parser generator. Lex, originally written by Mike Lesk and Eric Schmidt and delineated in 1975, is the standard lexical analyzer generator on many Unix systems and an equivalent

Retrieval Number:F13370476S519/19©BEIESP
Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

tool is specified as part of the POSIX standard. This module is coded with CMU Lex tool. The CMU lex tool has a dictionary that holds many words which are used in converting recognized text into syllables. It acknowledges the given input, which is now segregated into syllables using CMU Lex tool. The tool distinguishes the input and it is conjointly used as a query to search the local database. The user can ask the Bot to read any Document that is given in the local database.

C. Text Understanding

This section principally establishes a transparent plan on however the larva understands complicated text, which incorporates speech text. Firstly, we'd like to instigate an intensive text management command corpus. Secondly, we'd like to style a collection of effective question process codes. Thirdly, we'd like to understand a way to rework the intentions of the user's language text into the corresponding commands that may be understood by the larva. Supported the 3 issues higher than, a framework of larva is meant. The framework consists of 3 serial layers: input layer, method layer and output layer. The task of the larva within the input layer is to get the user's info, principally voice input and therefore the info are going to be given to the process layer to perform operations. This transfiguration of voice to text is coordinated within the input layer.

D. Processing query using Local Database and Internet

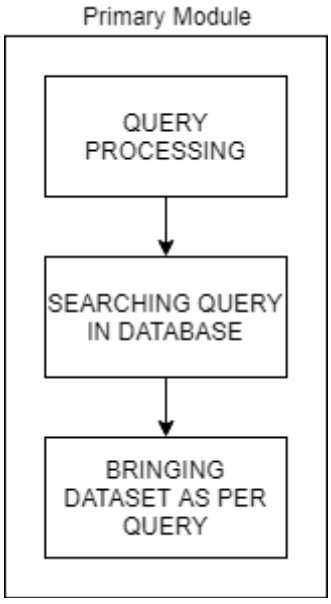


Fig 5 Work flow of Primary query processing

The processing layer is subcategorized into two modules: processing query using local database and processing query using Internet. The primary module (i.e) processing query using local database, permits the Bot to access the files within the user's device in order to process the query. The Bot determines if the query requires the use of web or not and performs the corresponding operation. If the initiated query does not require the use of web and it simply needs the Bot to fetch a particular data from the local database, the Bot would perform an intensive search in that database to obtain the file. The Bot, by itself, can locate the Document which is given as the input and fetch the same from the local database. The Figure 4.2 shows the sequential method by which Bot processes the query, searches the query in the local database and delivers the relative output. All of these processes can be done offline and will not require the use of internet even in the case of receiving the input from the user. To obtain query when the user is offline, an unique voice recognition tool called Pocketsphinx is used. It is a continuous-speech, speaker-independent and large-vocabulary recognition tool. If complex input is given, it does not have the ability to respond with a precise output.

E. Processing query using Internet

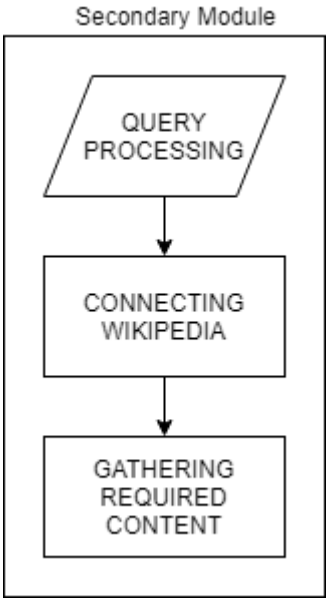


Fig 6 Work flow of secondary query processing

The secondary module (i.e) processing query using Internet allows the Bot to access the Internet to process the user's query. The Bot can intensively search the web and fetch correlated results from Wikipedia. The Figure 4.3 establishes a clear picture on how the Bot identifies that the query has to be processed online. It then initiates a connection with the Internet. It connects to Wikipedia and processes the query. Once Wikipedia displays the search results, Bot gathers these results and provides the output to the user. Suppose no results are found in Wikipedia, it returns 'Answer not found'. In case the user is not connected to the internet, it does not return any output. Cloud API is a distinctive voice recognition tool which is used to process those queries which requires internet access. It first records the query and stores it in Cloud Storage. The call speech API points to the audio file. The recognized text is sent to Natural Language API. Bot receives this processed query and performs a Wikipedia Custom Search. Cloud API is specifically used for queries that need to be

processed online because it can comprehend aggregated and complex inputs and provide explicit results for them. It is clearly apparent that the Bot is efficient in exercising the queries both online and offline. This makes the Bot user-friendly and capable for use by elementary school youngsters and blind individuals.

F. Text to Speech

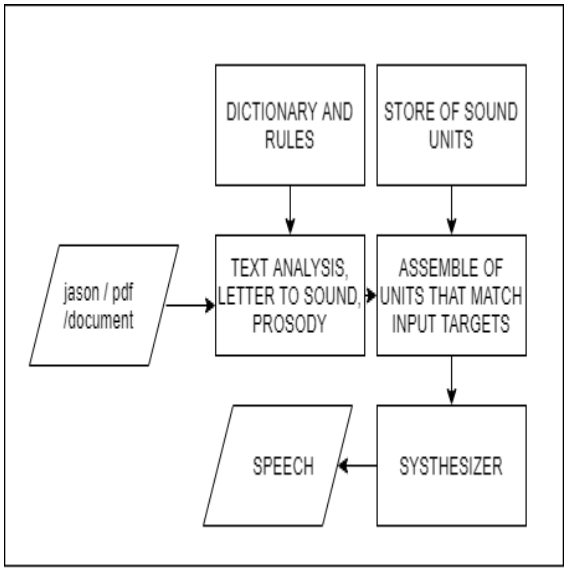
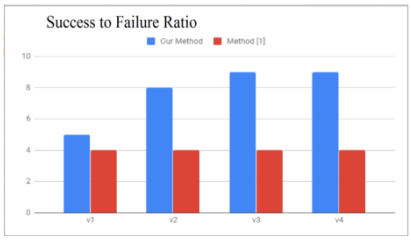


Fig 7 Work flow of text to speech

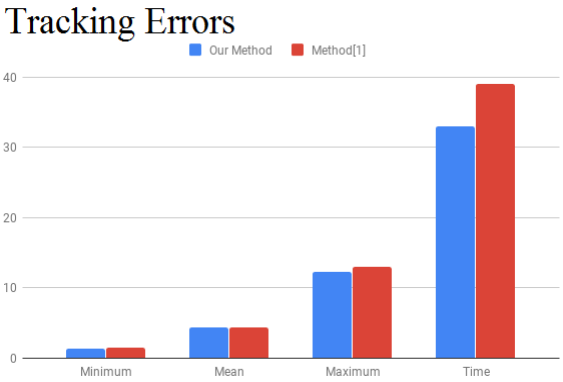
In the preface of this module, the output has been processed either from the local database or from Wikipedia. In order to convey the output to the user, the Bot follows some orderly procedure. The Figure 4.4 explains the procedure by which the outputs in text format are converted into voice. Firstly, with the use of certain vocabularies from its dictionary and predefined rules, it performs a text analysis where the output which is currently in the form of text are converted to sound prosody. At the end of this process, the Bot has recognized the output word-wise. Secondly, using the phonics that are already available in the text-to-speech library, the Bot structures these individual words into a sentence. The sound units contain profuse phonics which help the Bot in framing the sentences from individual and distinctive words. Lastly, these sentences are put together in a chain of sequence using the synthesizer and converted into speech. If the output has been yielded online (i.e) from Wikipedia, the files are temporarily stored in a local database in json format. The Bot will then read the results that are fetched as the output for the inputted query to the user.



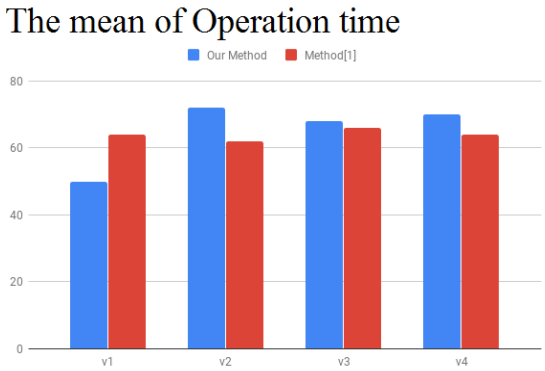
GRAPH I [THE SUCCESS TO FAILURE RATIO]

The blue bold lines represent the number of success

ratio out of 10 attempts. The red bold lines represent the number of success ration in the method [1]. The experiment was calculated on three conditions where the environments so as to calculate the exact success to failure ratio. The Graph I clearly shows the performance of the Bot. The tasks were barely done on voice recognition and search analysis in the Bot. It also has been checked for voice output.



GRAPH II
Tracking Errors



GRAPH III [The result of the working Bot.]The mean of Operation time.

The Graphs II and III represent the errors that are half-track throughout the implementation of the larva, moreover because the operation time and its mean values area unit calculated. The red daring lines area unit the values of methodology [1] and also the blue daring lines represents our methodology. The mean errors of the 2 tests victimization our methodology and methods[1] were four.29 millimeter and four.59 mm, 3.59 millimeter and three.64 mm, 5.70 millimeter and five.93 mm, severally, which suggests methodology [1] is of the best exactness. The tasks were barely enforced by victimization methods [1] for the dearth of speech management that created it arduous for users to perform the tasks. though there have been 2 volunteers with success acting the task once

Retrieval Number:F13370476S519/19©BEIESP

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

victimization method[1], they spent an excessive amount of time to acknowledge the voice and convey out the desired that caused the mean operation time of the task comparatively long. Our methodology is a lot of correct than [1] and once fine teaching with low noise from external factors.

V. RESULT AND DISCUSSION

The communication between humans and Bot goes to the next level through intelligent technologies. This Bot stands as the base work for the future teaching Bots. The module is specially designed with the goal of helping blind people and the primary school students to extensively use the information given in the local database as well as the information available in the internet. This module can talk with the user through voice output, can read any document and can also search for contents in Wikipedia. The Bot gives access to the internet for visually challenged people. The Bots can be implemented in rural areas where people have on some predefined input.

Poor teaching facilities. These Bots are extremely accurate and provide intact results for any given input. The art of teaching with the help of Bots can be increased in the future and this module can be taken as an example for teaching Bots. It is discussed that Bot can be used by many visually challenged since it is economical. It is also very simple to handle and can thus be used by people of all age groups.

VI. CONCLUSION AND FUTURE WORK

This paper presents an interactive teaching system that uses the human voice as input. The input voice is recognized by google CMUsphinx API. The voice is then converted into text. CMUsphinx tools play an important role in recognition as well as the conversion of voice into text. The device waits for the input from the user. Once the input is recognized by the device, it responds according to the corresponding query by the code. If the users subject or query is found in the local DB it is returned to the user through voice output. If not, the query is taken to the internet to search in Wikipedia. From Wikipedia, the related content is taken and also returned as a voice output to the user. The device is connected to the internet either in a wireless network or in a wired network. The text from Wikipedia or local database is read by the same google API which is used in recognizing the input voice for querying the Database in the local DB. The future work can be improvised with more features like face detection or voice verification so as to use this as a personal assistant only to its boss. Further, it can be as a developed as a humanoid robot integrated with artificial intelligence to work upon his/her master's command and suggest the work to master based.

REFERENCES

[1] Caibing Liu;Qichang Hu ; Mingxuan Chen ; Bo Zhang;, March (2018),"Online Robot Teaching with Natural Human-Robot Interaction", IEEE Transactions on Industrial Electronics , Volume: 65, Issue: *12*, 1002 - 1014.

[2] Dong-Sun Kim, Member, IEEE, Sang-Seol Lee, Member, IEEE, and Byeong-Ho Choi[2017],"A Real-Time Stereo Depth Extraction Hardware for Intelligent Home Assistant Robot" IEEE Transactions on Consumer Electronics Vol. 56, No. 3.

[3] H. M. Do, T.-Y. Choi and J. H. Kyung,[2017], "Automation of cell production system for cellular phones using dual-arm robots," Int. J. Advanced Manufacturing Technology, Vol. 83, pp. 1349-1360.

[4] Hyung-Don Kim, Jinsung Kim, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno[2016],"Target Speech Detection and Separation for Humanoid Robots in Sparse Dialogue with Noisy Home Environments"IEEE Transactions on Intelligent Robot Systems Vol.3 issue no.11 pp.110-121

[5] J-M. Valin, J. Rouat, and F. Michaud, "Enhanced Robot Audition Based on Microphone Array Source Separation with Post-Filter," IEEE Transactions on Intelligent Robots and Systems vol.5 pp. 2123-2128.

[6] J. Du, C. Mouser, W. Sheng,[2016] "Design and evaluation of a teleoperated robotic 3-D mapping system using an RGB-D sensor," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 46, no. 5,pp. 718-24, 2016.

[7] K. Livescu, F. Rudzicz, E. Fosler-Lussier,[2017] "Speech Production in Speech Technologies: Introduction to the CSL Special Issue,"Computer Speech and Language, vol. 36, pp. 165-172.

[8] L. Lu, H. J. Zhang, and H. Jiang[2016],, "Content Analysis for Audio Classification and Segmentation," IEEE Trans. on Speech and Audio Processing, vol. 10, no 7, pp. 504-516

[9] Lee, Chin-Hui, Frank K. Soong, and Kuldip Paliwal,(2017), " Automatic speech and speaker recognition",advanced topics. Vol. 355. Springer Science & Business Media,pp.48-65.

[10] O. Mubin, C. J. Stevens, S. Shahid, A. A. Mahmud, and J. J. Dong,[2017], "A review of the applicability of robots in education," Journal of Technology in Education and Learning, vol. 1,pp. 1-7.

[11] S. Wang, F. Xu, L. Chen, F. Zou, B. Li,(2016), "Industrial Robot Components Assembly Based on Machine Vision Technology" Modular Machine Tool & Automatic Manufacturing Technique, vol. 8, pp. 107-110.

[12] S. P. Won, W. W. Melek, and F. Golnaraghi, (2016),"A Kalman/particle filter-based position and orientation estimation method using a position sensor/inertial measurement unit hybrid system," IEEE Trans. Ind. Electron., vol. 57, no. 5, pp. 1787–1798.

[13] Schultz, Tanja, Ngoc Thang Vu, and Tim Schlippe,(October 2017), "GlobalPhone: A multilingual text & speech database in 20 languages." Acoustics, Speech and Signal Processing (ICASSP) IEEE transactions on. Industrial Informatics ,vol.3, pp.45-67.

[14] S. P. Won, W. W. Melek, F. Golnaraghi,[2018] "A fastening tool tracking system using an IMU and a position sensor with Kalman filters and a fuzzy expert system," IEEE Trans. Ind. Electron., vol. 56, no. 5, pp.1782-1792.