A Review on Voice Browser

Akshay A. Zade[1], Lowlesh N. Yadav[2], Neehal B. Jiwane[3]

B.Tech Final Year Student, Computer Science and Engineering Department, Shri Sai College of Engineering And Technology, Bhadrawati, Maharashtra, India[1]

Assistant Professor, Computer Science and Engineering, Shri Sai College of Engineering and Technology, Bhadrawati, Maharashtra, India[2]

Assistant Professor, Computer Science and Engineering, Shri Sai College of Engineering and Technology, Bhadrawati, Maharashtra, India[3]

Abstract: A web browser is employed to show sites, navigate from one website to a different with the utilization of hyperlinks, and transfer any sort of information right from PDF files, displays, Word files to music, videos and pictures. Browsing is exploitation the mouse, keyboard and bit (in case of smartphone applications. however what concerning the incapacitated and the visually impaired users? however would they use the search engine? we've tried to come back up with an answer by creating a Voice primarily based Browser that's a totally hands-free search.

Keywords: Speech Recognition, Text to Speech, Web Browsing, Speech Synthesis, Web Crawler.

I. INTRODUCTION

A web browser is employed to show web content supported what the user has looked for and transfer any variety of data. Speech Recognition is quickly gaining pace in applications like Google currently, Siri, Cortana, etc. Speech Recognition is very important in building application for the handicapped and visually-impaired users. We have implemented this feature in our Voice based mostly Browser. With the assistance of Speech to Text feature, the user can speak out the words he/she would really like rummage around for on the browser. The Search Box can once more repeat the words using the Text to Speech feature in order that the user can apprehend the right words are expressed. Once the net pages square measure displayed, the hyperlinks would be spoken to the user with the Text to Speech feature. Thus, the disabled users will use and navigate through the browser simply.

II. RELATED THEORY

In several studies, algorithms and applications are implemented to facilitate the net browsing or to revolutionize the normal method of surfboarding the WWW in different angles. One application was enforced to display enumerated links within the browser window and to have conjointly a compass mouse with a curser positioned over a mouse-over pull-down menu by speech recognition.

In modern-day several users don't understand however they act with voice internet elements through VoiceXML often. There ar alternative standards that also are supported as a region of VoiceXML.

VoiceXml (VXML): A language that makes audio dialogs that use synthesized speech, digitized audio, recognition ofspeech, recorded speech and telephony;

Speech synchronic linguistics Recognition Specification (SRGS): A grammar knowledge document that's wont to specify words, phrases, patterns of words, sentences etc. within the manner they are needed to be listened by the listener/recognizer;

Semantic Interpretation for Speech Recognition (SISR): A data document that defines the foundations and protocols of grammar and language for extracting linguistics outputs from recognizer

Pronunciation Lexicon Specification (PLS): A document representing phonetic info to be utilized in speech recognition and synthesis;

Speech Synthesis nomenclature (SSML): it's a markup language for rendering a mix of pre- recorded speech and live speech and alternative audio files

III. PROPOSED THEORY

This application is enforced mistreatment MS Visual Studio 2015 that helps in changing Speech to Text additionally as Text to Speech mistreatment C# Language.

The commands like Back, Go, Refresh, Back, Forward, Home, Speak, Listen, Stop can perform functions as allotted in program.

The steps performed are:

I. Application takes voice input through electro-acoustic transducer and converts it into text. This helps the user to select the particular URL required.

II. The commands are matched with the command set wherever the particular commands with specific actions square measure keep e.g.: GO, etc.

III. If the match happens, then that individual action can be dead.

IV. By giving the voice command user will operate the web browser.

V. User are ready to open sites.

VI. The text showing on the URL bar can reborn back to speech, giving output in voice.

VII. Keywords square measure allotted for each link in URL.

VIII. Dialogue box with the keywords open and therefore the individual link opens.
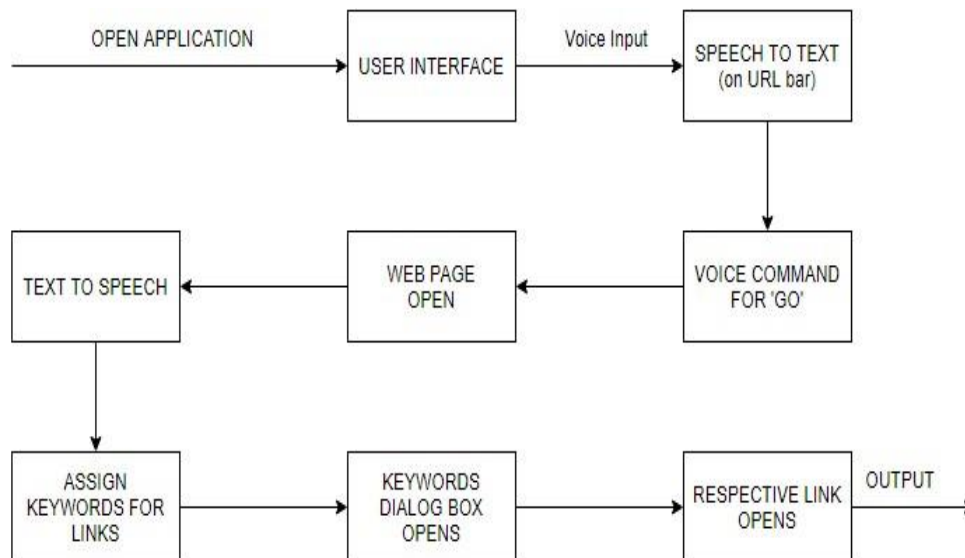
Description:

User: The user interacts with the browser by giving voice commands to the browser and conjointly reads text back to speech.

Net Browser: the net browser takes the command from the user and sends it to the speech to text device for text conversion. Speech to Text converter: This section converts the obtained voice commands to text and matches with the command set. Text to Speech converter: This section converts the obtained text to speech.

Command Set: This section contains all the commands user will offer to the net browser and given commands are matched here.

If it matches then specific action is performed else management is shipped back to browser. Performs action: This section performs the desired action and displays it on the browser.

These are the commands in the browser:

Table 1. Speech Command

Command	Description
Go	It will navigate to the entered URL
Back	It will open previous web page
Forward	It will open previous web page
Refresh	It will refresh the web page
Home	It will open home URL
Speak	It will take voice input
Listen	It will record the voice
Stop	It will terminate the process
Links	It will speak link titles of the web page

IV.   IMPLEMENTATION

Visual Studio 2015 are accustomed build the browser victimisation the C# programing language. Speech is a good manner of interacting with applications fully hands-free. It is useful for the unfit and therefore the visually impaired users. Here, the online browsing would be voice primarily based and a hands-free application for such users.

Benefits of victimisation the "System.Speech" framework in C# are:

1. Speech Recognition: The System.Speech.Recognition framework helps with the speech input. the advantages of Speech Recognition square measure as follows:

a) Speech Input: It takes any audio or voice as input for the application and notifies if there square measure any errors within the program

b) Grammar: victimisation the Grammar Builder and decisions classes we are able to give set of cases or alternatives we have a tendency to would like to own in our program. the alternatives given in the Choices category might then be enforced in switch cases or if-else statements.

c) Events: Events like recognizing speech is finished by the application. for instance the Speech Recognized raises events once speech is detected with correct volume and frequency levels.

d) Recognition Engines: Here you'll be able to work round the configuration of the input, alter and disable recognition or modification sure properties of the engine that affects the recognition method.

2. Speech Synthesis (Text-to-Speech): This feature converts the matter words on the screen to speech. The volume and speed of the speech may be regulated as per requirements.

a) produce TTS contents (or Prompts): regardless of the engine speaks may be a prompt. The Prompt Builder category may be used for text to speech and conjointly regulate speed and volume of the speech.

b) Manage the Speech Synthesizer: with this, the user will select a speaking voice, specify output, handlers for events, start, pause, record or continue the speech. The speaking may either be asynchronous victimization SpeakAsync().

c) management voice characteristics: The properties of the output speech may be controlled viz. the gender of the speaker, rate, volume of the speech, etc.
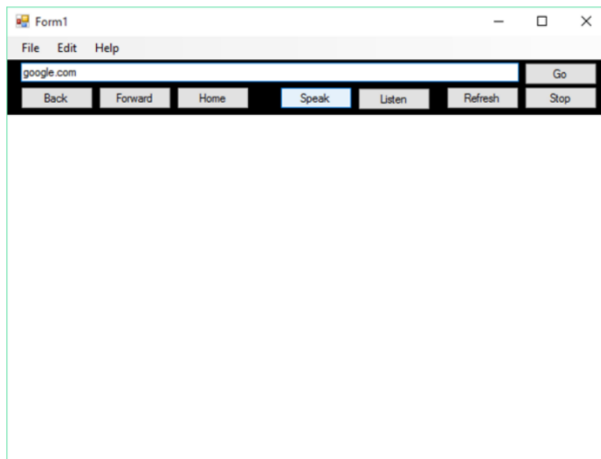
Fig 2: Voice based input to the browser
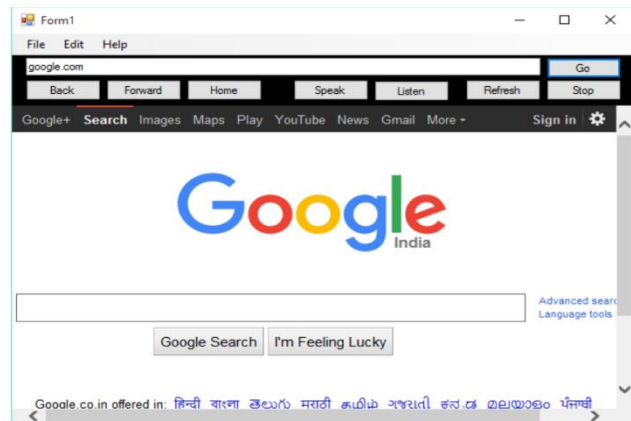
Fig 3: Opening of the web page on "Go" (speech command)

Problems associated with the voice interface With the superb tools and technologies out there like SALT (Speech Application Language tags) with ability to plug-in the correct places to develop speech enabled applications, a minimum of initially look look as simple as internet page creation. However, in spite of the many glorious tools out there these days, sites that square measure tough to know are laborious for developers to spell sure.

1. Organization before implementation

To construct any system or code, organization and analysis of its arrange is a crucial task. With considerations to be taken of user and purposeful requirements. Therefore, construct a road map to the existing applications by knowing the necessity of amendments in the current situation. info for the analysis might be out there within the web or is gathered manually. However it not very easy to create the speech primarily based applications, which suggests the atmosphere plays Associate in Nursing important role within the running of Associate in Nursing application. As the level of automation will increase in Associate in Nursing application the extent of accuracy falls, this marks the fundamental principle of speech technologies. This merely doesn't mean that the developer ought to disregard the principles of style connected to speech applications and their interfaces.

2. Prompt Clarity

Listening could be a troublesome task each for the user and therefore the machine. forward that the user are in Associate in Nursing setting where the noise around him/her are stripped and therefore the foreground sound signal are clear could be a mistake. Environment can not be outlined since it's dynamic in real world and changes randomly. therefore the system has to adapt to the dynamic environments within which the user is. For example, a locomotive conductor works in a wholly different setting than Associate in Nursing workplace employee. Likewise, a balance inquiry system for a bank can have a very different sense of urgency compared to a city's emergency line. By mistreatment totally different sentence frames, pausing between interval points, and maybe most significantly, by employing the correct voice character and voice talent, one will be able to develop resistance to prompts and background noise that actually facilitate listening.

3. Such a Big Amount of Things to mention

Listening could be a troublesome task at the best. it's natural that humans tend to talk a great deal and elaborate on everything they want to convey. therefore for the speech recognition machine, listening and recognizing long conversations and converting it to text would become not possible. Thus user must speak with acceptable speed with pauses permitting the recognition engine the time to convert the to this point listened speech to text.

Use of internet Crawler
A internet crawler (otherwise known as an online arachnid or web robot) could be a system or mechanized script that peruses the Internet in Associate in Nursing economical, computerised approach. This procedure is called internet crawl or creep.
Numerous honest to goodness destinations, specifically web indexes, use creep as a way for giving avantgarde info.Web crawlers ar for the foremost half wont to create a duplicate of all the visited pages for later getting ready by a web finder which will file the downloaded pages to offer quick quests.

Crawlers will likewise be used for computerizing upkeep errands on a website, for instance, checking interfaces or acceptive markup language code.

Likewise, crawlers may be used to accumulate explicit sorts of knowledge from website pages, for instance, aggregation email addresses (ordinarily for spam).

Here the net crawler is employed to reap URLs and URL titles from the net pages therefore on offer the collected knowledge to the prompt builder to talk the required URL titles for the blinds. To crawl through the webpages, use of normal expressions to match the required patterns of results is employed most of the days once operating with C#.

This concept also can be extended to extract downloadable files, images, phone/fax numbers etc.

1. Use of normal Expression

C# language will support regular expressions through various categories within the System.Text.Regular Expression namespace within the .NET framework. To be able to use the regular expression categories one should import victimization System.Text.RegularExpression; namespace within the supply file.

The Regex category within the namespac System.Text.RegularExpression permits to perform string matching and extracting helpful info from text with its interface of normal Expressions. Thus, here the text is that the markup language webpage supply out of which we tend to extract URLs and URL titles.

Here to search out all the URLs or URL titles, a world search or a complete search over the webpage ASCII text file must be done and accumulate the results. to try and do this a static method Regex.Matches() is employed to match the strings found and to gather all the matches MatchCollection will be used which might be iterated and processed over.
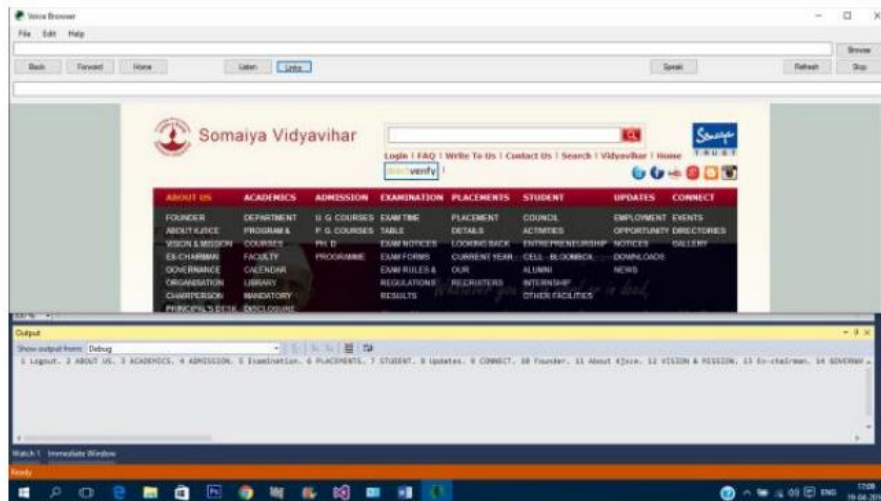
Fig 4: Extraction of URL titles using regular expression by voice command "Links"

V. CONCLUSION

Access to info has become a significant economic and social issue. Voice browsing technology could be a rapidlygrowing field. whether or not or not it proves to be ensuing web, it deserves a careful examination in its gift form, because the want for a straightforward and direct thanks to access the internet has become a requirement for several sorts of individuals, especially the unfit. From this, the thought of the project was raised so as to assist implement a voice based applications programme.

REFERENCES

[1] Voice Extensible Markup Language (VoiceXML) Version 2.0, [online]

[2] Huixiang Gu, Jianming Li, Ben Walter and Eric Chang, "Spoken Query for Web Search and Navigation", Best Student Poster Award 10th International World Wide Web Conference Poster Proceedings, pp. 2-3, May 1-5, 2001.

[3] Sam Rollins and Neel Sundaresan, "AVoN calling: AXL for voice-enabled Web navigation", Computer Networks, vol. 33, pp. 533-551, 2000.

[4] Dong Lin, Lin Bigin and Yuan Bao-Zong, "Using Chinese Spoken-Language Access to the WWW", Signal Processing Proceedings 2000. WCCC-ICSP 2000. 5th International Conference on, vol. 2, pp. 1321-1324, 2000.

[5] P.J. Danielsen, "The Promise of a Voice-Enabled Web", Computer, vol. 33, no. 8, pp. 104-106, Aug. 2000.

[6] Mohan, "A Strategy for Voice Browsing in 3G Wireless Networks", EUROCON'2001 Trends in Communications International 'Conference on, vol. 1, pp. 120-123, 2001. Araki Masahiro, Ono Tasuku, Ueda Kiyoshi, Nishimoto Takuya and Niimi Yasuhisa, "An Automatic Dialogue System Generator from the Internet Information Contents", Eurospeech2001, pp. 1743-1746.

[7] Stuart Goose, Mike Newman, Claus Schmidt and Laurent Hue, "Enhancing Web accessibility via the Vox Portal and a Web-hosted dynamic HTML↔VoxML converter", Computer Networks, vol. 33, pp. 583-592, 2000.

[8] Witawas Srisa-an, Chia-Tien Dan Lo and J. Morris Chang, Putting Voice into Wireless Communications, [online]

[9] John Geralds, "Industry Forum Approves VoiceXMLStandard", vnunet.com, Mar. 2000.

[10] Anne C. Lear, "New Spec Gives Voice to the Web", Computer, pp. 23, May 2000.

[11] Davis, S.B., Mermelstein, P. "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences", IEEE Trans. on Acoustic, Speech and Signal Processing, 28(4):357–366 (1980).

[12] Lakshmi Kanaka Venkateswarlu Revada, Vasantha Kumari Rambatla and Koti Verra Nagayya Ande,,,A Novel Approach to Speech Recognition by Using Generalized Regression Neural Networks" IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011

[13] Lingyun Gu and Stephen A. Zahorian, ,,A new robust algorithm for isolated word end point detection"ICASSP, page 4161. IEEE, (2002)

[14] Dr.R L K Venkateswarlu,Dr.VasanthaKumari,A K V Nagayya, ,,Efficient Speech Recognition by Using Modular Neural Network" International Journal on Computer Technology and Applications, Vol 2 (3), pp.463-470

[15] RabinarL, BingHwangJ." Fundamentals On Speech Recognition",Prentice Hall,1993

[16] .John Coleman, "Introducing Speech and language processing", Cambridge university press, 2005.
 Tzu-Chao Lin ,Ching –Yun Chang "A Survey of VQ Codebook Generation"Jorrnal of information hiding and Multimedia Signal Processing,Volume 1, No.3,July 2010

[17] Tzu-Chao Lin ; Pao-Ta Yu ,,A new unsupervised competitive learning algorithm for vector quantization" Neural Information Processing, 2002. ICONIP '02.Proceedings of the 9th International Conference on (Volume:2) 944 – 948,2002

[18] Abderrahmane Amrouche, and Jean Michel Rouvaen"Efficient System for Speech Recognition using General Regression Neural Network, International Journal of Electrical and Computer Engineering 1:6 2006

[19] . Rita Singh, Bhiksha Raj, and Richard M. Stern, Member, IEEE, "Automatic Generation of Sub word Units for Speech Recognition Systems", IEEE Transactions on speech and audio processing, VOL. 10, NO. 2, Feb. 2002.

[20] Dipmoy Gupta, Radha Mounima C. Navya Manjunath, Manoj PB" Isolated Word Speech Recognition Using Vector Quantization (VQ)" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 5, May 2012