

SIE: Speech Enabled Interface for E-Learning

Shikha

M.Tech Student Lovely Professional University, Phagwara , Punjab INDIA

ABSTRACT

In today's world, e-learning is very important and popular. E-learning provides a new approach of learning. E-learning means in which learner learns through the medium of internet, CDs or by installing the tools and software available. E-learning was earlier developed to teach math and reading to young children in elementary schools. But later on it was popular in organizations, where they use e-learning to train their trainees. Organizations spend more money on transporting and housing trainees than on actual training programs. E-learning has capacity to reduce this cost. As the expense of formal education in the educational institutes is increasing very fast, the expense of education via the internet is comparatively low and affordable. Although technology is growing up, however requirements of common people and disabled people are still not taking into account properly. For this a new concept needs to be embedded in the e-learning that is speech enabled e-learning. It provides full user interaction, in this user can ask the questions and reply back to the questions, can give commands to e-learning application.

1. INTRODUCTION

This proposal will provide the user a facility of evaluating himself through an e-test which will be speech enabled and will include objective questions. This will be done by matching the user input with the database and give the most relative and accurate result. This whole process will be speech enabled (input and output will be in the form of speech) . As we see, in the world there are lots of people who want to acquire knowledge but due to cost factor or some physical disability they are unable to acquire it, and they acquire the knowledge then they are unable to evaluate themselves effectively. So, the first objective is to make e-learning approachable for common people. To fulfill this objective I would like to propose an idea of computer based speech enabled e-learning. It will be cost-effective. User need not to pay for it. Just open the website/application to work and it will provide you the feature of speech. This provision of speech enabled e-learning is very much helpful for the disabled person who cannot see or who do not have hands. Because the disabled person faces many problems to acquire knowledge, Using web based learning it will provide a new way to e-learning and make it much easier than earlier learning systems. Disabled people do not need to hire an instructor; the website will itself act as instructor. This is about the easy availability of e-learning for common abled and disabled people. An e-test will be provided to user, questions will be in

audio as well as in text form, user will reply the answer, the answer will be checked accordingly and the results will be announced. In this, system should match the most relevant data with the database, match the necessary keywords and produce the result accordingly. Match of the keywords is necessary because it is not necessary that user will reply a particular question exactly like as it is stored in database. So if speech grammars will not be matched then effective results will not be displayed. The main feature is that it is **Speech enabled user interactive system**. No doubt lots of software and other products are available related to speech enabled e-learning, however there is no application which provides user interaction as well. User will only be able to learn by listening, cannot ask by speaking. This idea is about making such feature in applications/websites.

2. METHODOLOGY

2.1 Formulation of the hypotheses

This idea combines two areas. One is e-learning and second is speech synthesis and recognition. This application will be made on Microsoft ASP.net. Microsoft provides the features of speech synthesis and speech recognition. Speech synthesis means text to speech conversion and speech recognition means speech to text conversion. Both ideas will be combined.

In text to speech conversion, for speech natural language rules need to use, that means where is a need of high or low pitch, how to pronounce a particular word as its pronunciation changes according to the tense. To understand the text is one task. Now next task is to generate the sound of text understood. In this what will be happening, there will be a database which will store the sound segments which will be used accordingly. The effectiveness of text to speech conversion totally depends on how efficiently the sound segments are stored in the database.

In speech to text conversion, firstly the audio signal will be taken which includes strings, numbers and various pitch. This audio signal needs to check among the data stored in the database. For better results recognizer should care about only the required stuff. So for this purpose application will use a grammar which will enhance the efficiency of the search. This idea will be used to make an algorithm for searching the results of subjective questions from the database.

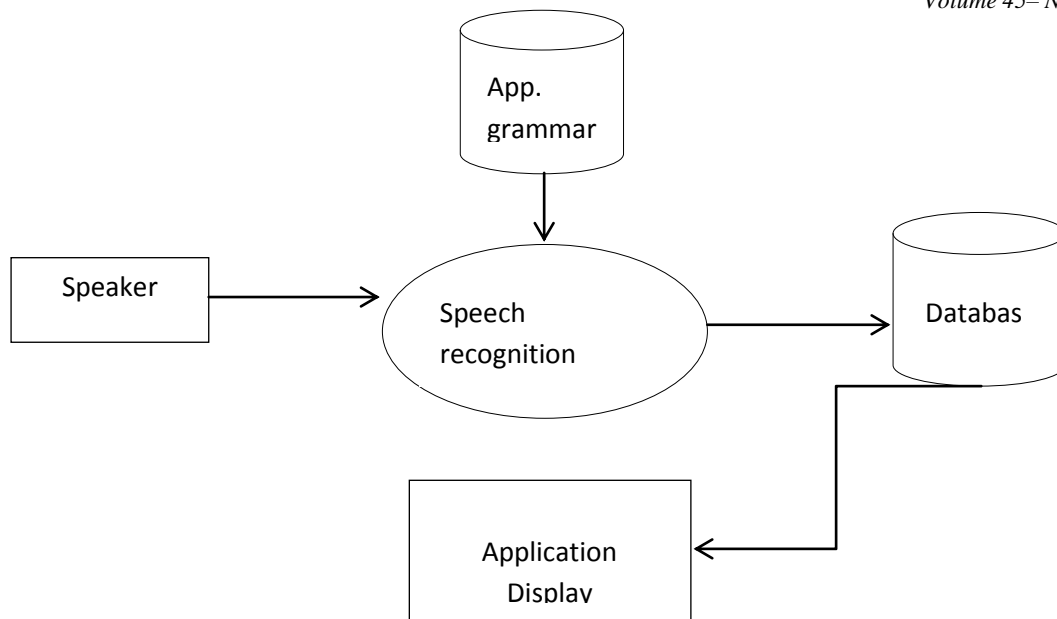


Fig1: comparison of spoken words

The above figure shows:

1. Speaker will speak the input like "a","b","c","d".
2. The word will be matched by speech recognition engine. It will make use of created grammars.
3. Check the input from database.
4. Result will be spoken by the system.

The above scenario will work for the completely speech enabled application. Required features can be put in the application accordingly.

2.2 Sources of Data

Data will be provided by the user in the form of speech, and accordingly the response will be given to user. Suppose user says search e-learning, then relative content will be displayed to user.

Data will be also provided by the system itself when text to speech conversion need to do. Suppose the content about e-learning is displayed, if user says speak then the content will be presented in audio form. That is input is provided by the system and converted into audio form.

2.3 Research Design

This application will allow the user some facilities like

evaluation of learner by providing an e-test (including objective and subjective questions), user can search the content, go forward or backward etc. These all facilities will be speech enabled.

Suppose

1. User wants to study chapter 1 of asp.net, then he/she will say chapter 1 asp.net, then accordingly content will be displayed in form of text by converting speech into text. When user say speak this content will be converted into speech and will be presented in audio format.
2. User wants to check his/her performance then he/she can give the e-test. They just need to listen to the questions and give the answer accordingly. Answer will be checked and result will be given.
3. User wants to go from current page to back page, then just say go back to previous page and corresponding action will take place.

For speech recognition to work, it must know what phrases a user can speak, such as "A" or "B" or "C" or "D". A speech recognizer can listen for several hundred phrases at a time, and new phrases can be loaded and unloaded at any time. For a the computer to recognize a user's speech, the user must speak the full phrase exactly as it is written, without having words inserted or deleted. If the user says a phrase not on the list, the recognizer returns an "unrecognized" and the user has to rephrase his/her command. This way speech recognition works. So how this working will be done is going to show in the form of a flow chart.

Speech to text

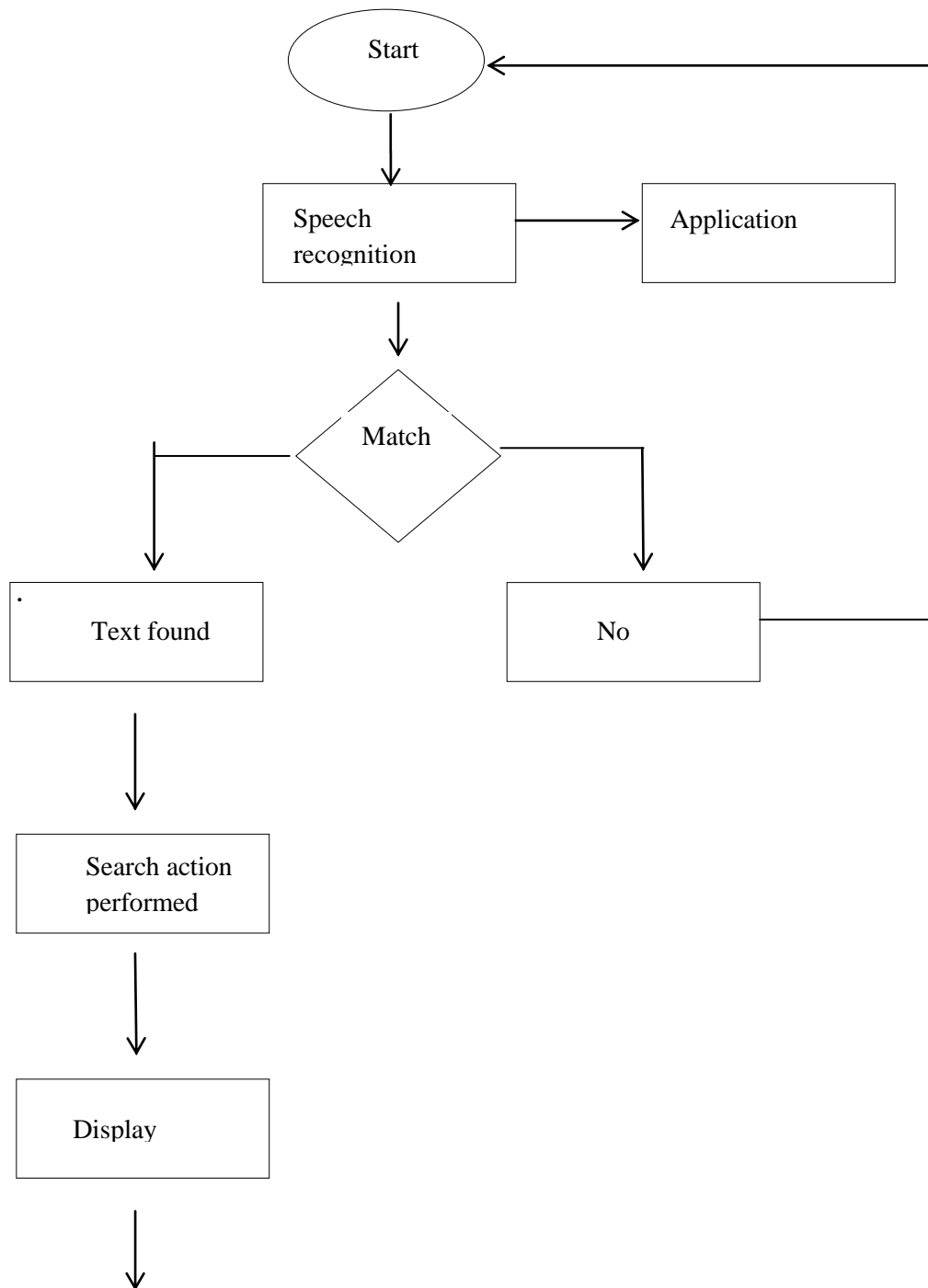


Fig2: Flow chart for speech recognition

The above flow chart shows the working of this application. When user will provide input then this input will be recognized by speech recognition engine and checked among the grammars which have been created. If input matches then the corresponding data will be collected by the system and

Text to speech

displayed to user. This collected data will be act as input to system if user want to listen the collected data. The data will be read by the system and user can make use of this data. How this text to speech conversion is done is shown in Fig 3 below.

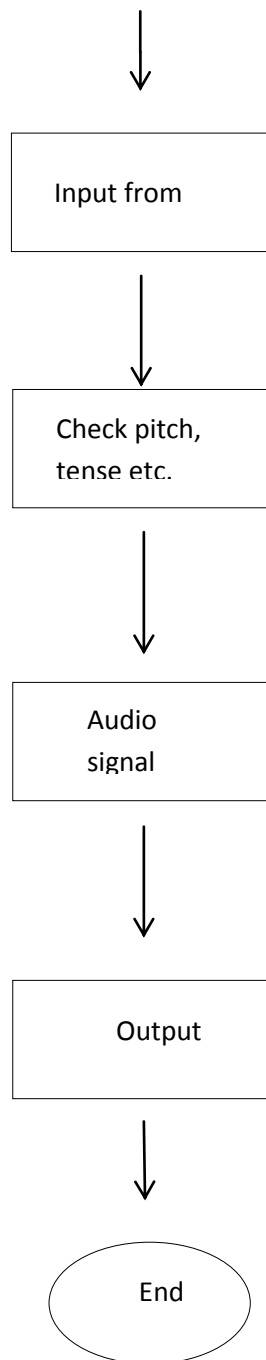


Fig3: Flow chart of speech synthesis

Algorithm

1. Input from user array of voice signals {s1,s2,s3....sn}
2. If (input) {
 signals sent to speech recognition engine
 }
 else {
 //ask to provide input or time out
 break;
 }
 3. check the speech grammars
 4. If (true)
 { //search the content }
 Else
 { // go to step1 }
 5. Display output
 6. automatic conversion of text to speech

7. Pitch and tense checking done
8. Audio signal generation
9. Voice result will be provided to user

2.4 Tools of Data Collection & Analysis

To implement this kind of system I will use the ASP.NET with C# which is provided by the Microsoft. Speech enabled system can be made very easily and efficiently. Microsoft provides a SDK (Software development kit) for this provision. This can be installed in the system and can be used.

Speech SDK can be used in either C#, C++, VB or any COM compliant language. Speech can be divided in to two paradigms. Text to speech conversion i.e Speech synthesis and other is speech recognition i.e speech to text conversion. The main thing which need to do is to map the user’s voice with the stored database and give the appropriate and correct response efficiently and fast. Speech recognition[1] can be of two types based on the grammar that the recognition is based on. (Grammar is in other words the list of possible recognition outputs that can be generated.) An application can limit the possible combination of the words spoken by choosing proper grammar in a command and control scenario. In this the accuracy of recognition is very high. In Dictation mode the recognition engine compares the input speech to the whole list of the dictionary words. The grammar choices are given manually here which needs to be matched that are a,b,c,d. By using Choices class in asp.net this can be done.

Before explaining steps required for speech recognition and synthesis lets discuss the reference required for this.

- Microsoft speech object library needs to use
- System.Speech class need to refer.

Few header files which are required to perform speech recognition and speech synthesis.

- Speechlib
- Speech.Recognition
- Speech.Synthesis

Steps of performing speech synthesis:

1. Create an object of spvoice class
2. Using this object you van set the rate and volume of system voice
3. Using speak method you can provide reference to system which it needs to speak

Steps of performing speech recognition

1. Create an object of SpeechRecognitionEngine class
2. To create your own choices to match with user input you can make use of Choices class
3. Grammar builder class can be used to append your created grammar
4. You can set the default input audio device
5. Load the grammars which you have created using load grammar method
6. Results of speech recognition can be stored in recognition result class
7. Now compare the results with the respective options

This is the general scenario which can be used to take user voice as input the system voice as output. Various methods are provided to manage the input and output voice.

E-test contains questions which are provided to the user in form of speech and input can be provided by the user in terms of speech. Now I am trying to do some work on the accent.

3. RESULTS ACHIEVED

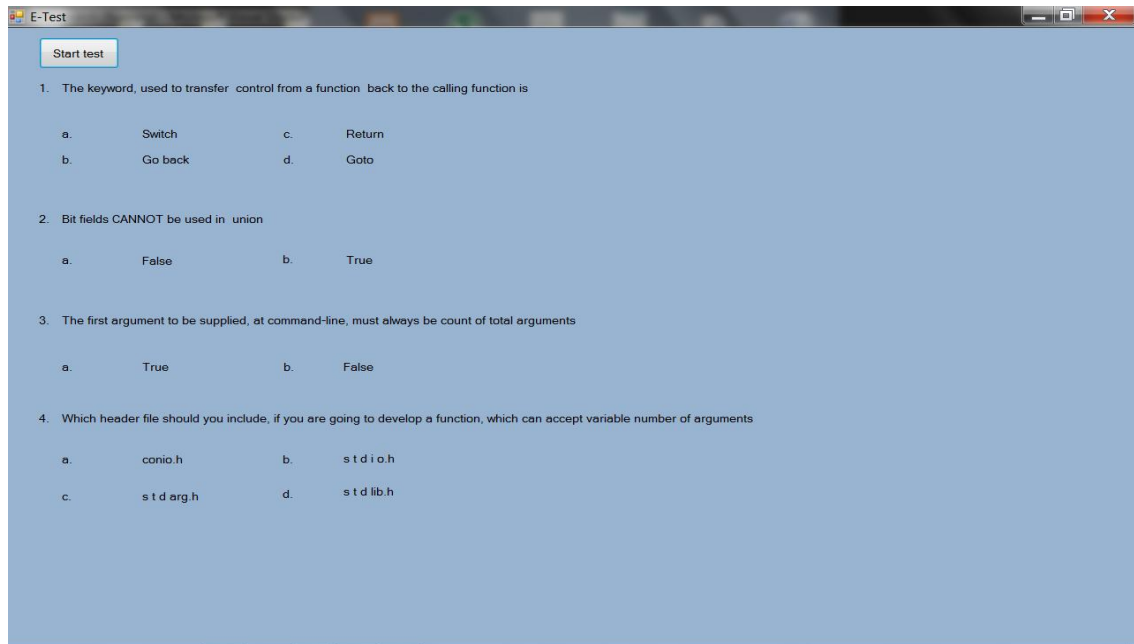


Fig4: Snapshot of the proposed work

1. When application will start then user needs to speak start test.
2. When user speaks start test then system will start speaking first question.
Eg: Firstly it will tell the user that you need to speak the options like "a","b","c","d"
Then it will speak the question like:

"The keyword used to transfer the control from a function back to the calling function is:"

Options are:

- A Switch
- B Go back
- C Return
- D Goto

Please speak the option

Then user needs to speak the option

Suppose user says

C

System will check the option said by user, if it is correct then system will speak "correct answer", if wrong then system will speak "wrong answer".

Then system will speak next question according to the questionair

All questions will be done in same manner

At last,

Suppose user has given three correct answers out of four then system will speak

"Scores are three out of four"

4. CONCLUSION

This application will be very helpful for common as well as disabled people. Because people are busy doing their work and cannot find extra time to learn something new. People who are physically disabled is very difficult for them to go for acquiring knowledge own their own. This application is the solution to this.

5. FUTURE WORK

In future I want to do is:

- evaluate the learner by means of subjective questions as well,
- User will give the input to the system which can be upto 4-5 words. This input will be checked against the database used for the questionnaire. On the basis of most appropriate matching results will be displayed.
- User provides input to system and system check it only with one particular voice accent, I want to implement that it should accept various accents of voice.

6. ACKNOWLEDGEMENT

I would like to express my deepest gratitude to **Mr. Sawal Tandon** (Dissertation Mentor) for his valuable guidance. It is only with his guidance that I could take up initiative of such a good topic of thesis. I am also very thankful to Lovely Professional University for giving me opportunity to propose and implement dissertation by the course of Dissertation.

I am gratified for the successful initiation of my thesis implementation I would like to convey thanks to all our friends who gave their full support and encouraged me for this thesis work.

7. REFERENCES

- [1] A. A. Azeta, C. K. Ayo, A. A. Atayero and N. A. Ikhu-Omoregbe (2009) “A Case-Based Reasoning Approach for Speech-Enabled e-Learning System” College of Science and Technology Covenant University, Ota, Nigeria.
- [2] Eryou Wei and Jun Yue (2008) “A New E-Learning System: Design and Implementation” LuDong University, Management College YanTa, China.
- [3] Fengming Yang and Wenjuan Wu “The Application of Virtual Reality in E-Learning” Computer Science and Technology Department East China Normal University (2010) IEEE
- [4] Hu Dali “Design and Implementation of E-Learning Performance Evaluation System ” Jiangxi University of Finance and Economics (2008) International Conference on Computer Science and Software Engineering
- [5] Luciana CARABANEANU(1), Romica TRANDAFIR(2), Ion MIERLUS-MAZILU(3) “Trends in e-learning” Department of Mathematics and Computer Science Technical University of Civil Engineering Bucharest B-ul Lacul TEI, Nr. 122-124, Sector 2, Bucharest, ROMANIA
- [6] Mari Ostendorfy Izhak Shafranz Rebecca Bates “Prosody models for conversational speech recognition”
- [7] Meysam Mohamad pour, Fardad Farokhi “An advanced method for speech recognition” World Academy of Science, Engineering and Technology 49 (2009)
- [8] Titus Felix FURTUNĂ Academy of Economic Studies, Bucharest “Dynamic Programming Algorithms in Speech Recognition” Revista Informatica Economica nr. 2(46)/(2008)
- [9] elearnspace
:http://www.elearnspace.org/resources/voicerecognition.htm
- [10] elearningeuropa
:http://www.elearningeuropa.info/en/category/freetags/speech-recognition
- [11] mapsofindia
:http://blog.mapsofindia.com/2010/12/16/current-trends-in-education-in-india-are-you-ready-to-acclimatize-and-succeed/
- [12] microsoft :http://msdn.microsoft.com/en-us/magazine/cc163663.aspx