SPEECH ENABLED WEBBROWSER

PROJECT WORK DONE AT

KUMARAGURU COLLEGE OF TECHNOLOGY

PROJECT REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT

FOR THE AWARD OF

BACHELOR OF ENGINEERING DEGREE IN INFORMATION TECHNOLOGY

OF BHARATHIAR UNIVERSITY, COIMBATORE

SUBMITTED BY

R.ARUN KUMAR (0027S0065)

V.VIMAL DEV (0027S0122)

GUIDED BY

Ms. N. RAJATHI B.E., SENIOR LECTURER

Department of Information Technology

Estd-1984

KUMARAGURU COLLEGE OF TECHNOLOGY

COIMBATORE - 641 006

MARCH 2004

ISO 9001:2000 Certified

Estd-1984

ISO 9001:2000
Certified

DEPARTMENT OF INFORMATION TECHNOLOGY
KUMARAGURU COLLEGE OF TECHNOLOGY
COIMBATORE - 641 006

CERTIFICATE

This is to certify that the project entitled

"SPEECH ENABLED WEB BROWSER"

is the bonafide work of

R.ARUN KUMAR	(0027S0065)
V.VIMAL DEV	(0027S0122)

and submitted in partial fulfillment of the requirement for the award of Bachelor of Engineering Degree in Information Technology of Bharathiar university, Coimbatore

Head of the department
(Dr. S. THANGASAMY)

Guide
(Ms. N.RAJATHI)

Submitted for University Examination held on 26-3-2004

26/3
Internal Examiner

External Examiner

Acknowledgement

We are highly obliged and wish to express our deep sense of gratitude to the "**Department of Information Technology" Kumaraguru College of Technology,** for the encouragement and help in carrying out this project work and also making it as a success.

We wish to thank our respected principal **Dr.K.K. Padmanaban B.Sc(Engg)., M.Tech.,Ph.D.**, and Head of the Department **Dr.S.Thangasamy Ph.D.**, for their kind encouragement and for having provided us with all facilties.

We express our sincere thanks and indebtedness to our be-loved guide **Ms.N.Rajathi B.E.**, Senior Lecturer, Department of Information Technology for her valuable instructions and timely help throughout the work.

We are highly indebted to our course coordinator **Mrs.P.Devaki B.E.,M.S.**, for her valuable guidance in bringing out the project well. We are extremely grateful to our class advisor **Ms.P.Sudha B.E.**, for her support throughout the duration of the project.

We also express our thanks to all other faculty members and technical supporting staff of the Department of Computer Science and Engineering and the Department of Information Technology for their support and timely help.

SYNOPSIS

Speech Enabled Web browser is developed to recognize and convert speech into text and perform effective operations on the web browser according to the speech input.

Here the input is given as the speech to the system, which is received and converted into text format, by Microsoft speech SDK. Then the converted text is used to perform pattern matching with that of the database and perform effective operations on the browser.

This web browser is very effective since it restricts the use of keyboards and mouse.

This web browser is immensely useful for all users particularly for the handicapped users.

Already existing systems have too many constraints, which have been tried to overcome in this system. Speech Recognition is the area, which is of prime importance to researchers now, and in this system we have implemented research algorithms in this area to recognize the speech.

CONTENTS

1. Introduction

Speech Recognition is one area into which tremendous amount of research has gone in. We in this project have aimed at producing a Web browser by using these speech Recognition techniques , which would recognize and convert speech into text which could be manipulated on the computer to perform effective operations on the browser. This is an innovative effort to produce a web browser that is speech enabled and would produce better results than those which currently exist.

1.1. Existing system and limitations

The limitation of the existing systems is that input (i.e the speech) will vary from user to user which would bring the system to an ambiguity condition in recognizing the exact input. Also the input should be of a proper slang, of which the software had created. These constraints make the system all the more difficult for the user to use it.

1.2. Proposed system and advantages

Our software lets the user to give his own personalized inputs by providing each user a unique domain which he can use for his operations. Since every part of the browser is speech enabled, the necessary for keyboard and mouse are restricted .The main advantage of our software is to facilitate handicapped users. These limited constraints make the system all the more effective and opens new doors in this fixed field of speech recognition.

2. System Requirement Analysis

2.1. Product Definition

Speech Enabled Web browser extracts and recognizes speech and then converts them into a text for effective command processing purpose. Here the input is taken in the form of an audio, which is then converted to text format. This conversion is facilitated by Microsoft speech SDK .The converted text is then implemented for effective command operations in a browser.

3. Software Requirements Specification

3.1. Introduction

3.1.1. Purpose

The purpose of this system is to make the existing ordinary web browser, speech enabled. Each individual user will be personalized by providing an unique domain. The possible events in the browser are predefined in the database. The input is taken in the form of an audio format (speech) and when converted into text format it checks for the exact pattern in the database. If present then appropriate browser operations are performed. This project is aimed at easy manipulation of events of a web browser. If successful this project will be a time manageable product that would facilitate many users including handicapped users.

3.1.2. Scope

The Speech Enabled Web Browser, when completed will be useful for any user who would need to use his speech as input to manipulate or control the web browser activities only through speech. This would restrict the use of keyboard and mouse. To enable this, the project incorporates the following features.

An easy to use Graphical User Interface.

A robust system that could recognize and convert speech of various forms into required text format.

3.1.3. Definitions, Acronyms and Abbreviations

GUI	Graphical User Interface.
SDK	Software Development Kit.
SRS	Software Requirements Specification.
.WAV	wave files.
WWW	World Wide Web.
HTTP	Hyper Text Transfer Protocol.

3.2. General description

3.2.1. Product Perspective

The main feature of this project will be the provision of definite identity to users. Also the software to recognize and convert speech into text form. These soft wares will be capable of recognizing the speech of various forms that maybe influenced by the inherent speaking qualities of each user.

3.2.2. Product Functions

The function of this system is to receive the inputs relative to specific users with the help of microphone. With the help of speech SDK each word is converted into corresponding text. These texts are then compared with the texts that are predefined by the user in the database. If both the pattern matches then the specific activity of the web browser will be carried on.

3.2.3. User Characteristics

The user need not be familiar with any programming languages. He should be able to give the input in the form of speech legibly and should have a rough idea about the events present in the web browser.

3.2.4. General Constraints

The basic constraint is that the user should pronounce the exact format of the text to which it should be converted. Else he shall have his own user defined speech format for the required text.

The other important constraint is that the user should have known the events present in the browser and give event accordingly. Also the surrounding should be noise free to have an effective performance.

3.2.5. Assumptions and Dependencies

The user can competently read and follow the GUI's instructions.

The user enters the speech input, which in English language and are well defined. Wrong pronunciation of input may be processed incorrectly.

3.3. Functional Requirements

3.3.1. Functional Requirements-1

3.3.1.1. Introduction

This module gets the speech inputs from the user, converts it into text format with the help of speech SDK and performs effective operations relatively on the web browser.

3.3.1.2. Inputs

The inputs are received from the user in the form of speech via the microphone.

3.3.1.3. Processing

This received input is converted into text format with the help of speech SDK. This converted text is compared with that present in the database. If both the pattem matches then effective operations are performed.

3.3.1.4. Outputs

The relative web browser operations are performed according to the user input.

3.3.2. Functional Requirements-2

3.3.2.1. Introduction

This module gets the speech inputs from the user, converts it into text format with the help of speech SDK and performs effective operations relatively on scheduler and miscellaneous that includes audio player, video player, explorer etc.

3.3.2.2. Inputs

The inputs are received from the user in the form of speech via the microphone.

3.3.2.3. Processing

This received input is converted into text format with the help of speech SDK. This converted text is compared with that present in the database. If both the pattern matches then effective operations are performed.

3.3.2.4. Outputs

The relative operations in the scheduler, and other miscellaneous are performed according to the user input.

3.4. External Interface Requirements

3.4.1. User Interface

The user interface is a GUI that will enable the user to enter a image file as the input (location of the file). A button is provided for the user to prompt the system to convert the bitmap image in to text file.

3.4.2. S/W Interfaces

A database is used to store the information (events) needed for managing the browser for each user. The system accesses this information to match the pattern for any particular input. The software interface provides access to this database. A speech SDK is user for the conversion of speech to text.

3.5. Design Constraints

3.5.1. H/W Limitations

The minimum hardware requirements for this system are

Pentium III Processor
128 MB RAM
10 GB Hard Disk
104 keys keyboard
Microphone

3.5.2. Fault Tolerance and Reliability

The system developed will be able to process the speech input with exact pronunciation of that of corresponding text format. Improper pronunciation and inputs out of the database, will lead to improper functioning of the system.

4. System Design Specification

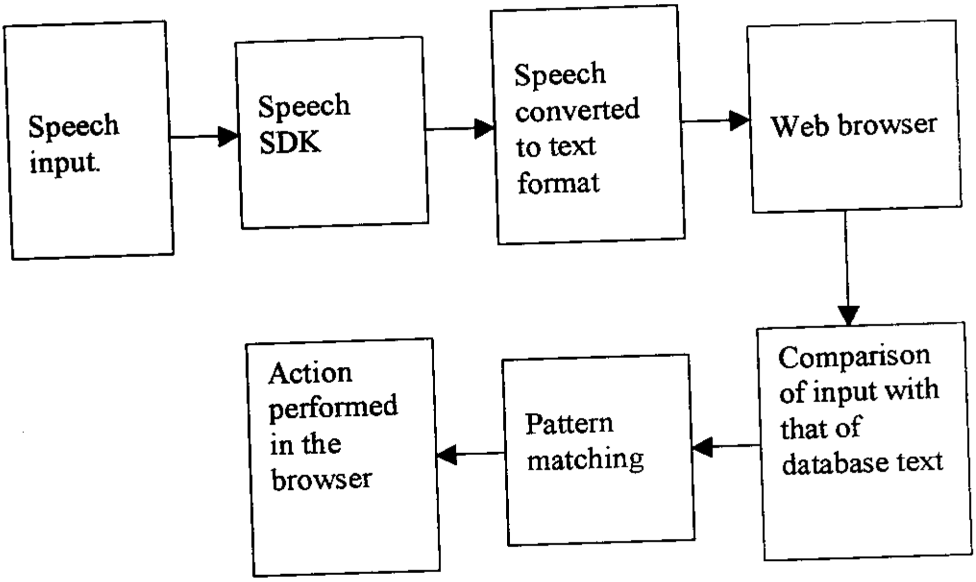
4.1. Overall Design:

Our system consists of a web browser which is created in such a way that it resembles the existing web browser consisting of all the functionalities in it. The

Specialty of this web browser is that it have a provision for each user to login under a personalized login. This browser has additional features such as personal history, favorites etc. The speech input via microphone is received and is converted into text format by the speech SDK. This input is compared with the database of that corresponding user. If both the pattern matches then effective action in the browser will take place.

In the next module, scheduler and other miscellaneous are created and made speech enabled. This module also uses the help of the speech SDK to convert speech into text to perform effective operations on the browser. The miscellaneous part includes audio player, video player, explorer, notepad.

4.2. Design Diagram:


Speech input.
Speech SDK
Speech converted to text format
Web browser
Comparison of input with that of database text
Pattern matching
Action performed in the browser

The user gives the speech input through microphone. This input should be of English with a specified format that supports our system.

Speech SDK is a software development kit that was developed by the Microsoft Company. The functionality of this kit is the speech to text conversion. The possible words in English will be present in this kit and also the possible pronunciations of these words are present here. These words are inputted by the manufactures in their own language and specific slang. This may be a bit uneasy for us to use. In order for us to use it effectively there is a provision which will enable us to train the SDK in such a way that it would recognize our words as well.

The Web browser is created in such a way that it resembles the conventional web browser that is in use. In addition to these, this browser has

a provision for each user to personalize his area. Entering into his area he can perform every operations on the browser personally. Also this browser has some extra features such as audio player , video player, explorer etc.

Pattern matching is an unique technique incorporated in this system. This process will compare the user inputs with that present in the database of corresponding user. If there is an exact match between the two, then respective operations are preformed relatively in the web browser.

5. Product Testing

Here, the various test strategies adopted in testing this system are outlined. The strategies include Unit Testing, Integrity Testing, and Output Testing.

5.1. Unit Testing

In this testing step, each module was found to be working satisfactory as per the expected output of the module. In the package development, each module was tested separately after it had been completed and checked with valid data. Unit testing exercises specific paths in the modules control structure to ensure complete coverage and maximum error detection.

In both the modules, tests are made in order to check all the menus, buttons, events etc works effectively. Also it is checked whether the speech entered by the user is received and recognized.

5.2. Integrity Testing

The individual modules are integrated to form the complete system. This system is then tested to find if the output of each module reaches the next module correctly. Various errors such as accessing of non-existing files were detected and corrected. A status field was included to let the user to know the current status of the system (which module of the system is being executed).

5.3. Output Testing

Output testing is the stage of implementation, which is aimed at ensuring that the system works accurately and efficiently before live operation commences. The input system, output documents were checked and required modifications have been made to suit the specifications. Then using the test data prepared, the whole system was tested and found to be a successful one.

Here various inputs are given to the system and the output is checked. The speech of different users may be different and the system was enhanced so that each user can configure his speech into the system before using it. This way the system was made more effective in recognizing the speech of different users.

6. Future Enhancements

This project could be in future extended to recognize user inputs in all forms and all formats, whatever may be the pronunciation. Further enhancements can be made such that the system becomes less sensitive to discrepancies in speech of various users. Also this system could be extended to make the whole computer system speech enabled with each and every component and software working through user speech.

7. Conclusion

This system developed has been able to recognize speech at a very good standard. The system's performance has been very impressive and further enhancements when considered, shows promise of an effective speech enabled web browser that could be of immense use to anyone who wishes to rely only upon their voice and not their hands . Also the system could be changed a bit to support different languages.

The techniques applied in the design of the programs provide a scope for expansion and implementation of changes, which may be required in future, all programs have been tested and have found to execute correctly. All the programs have been documented and can be easily understood.

8. References

8.1..Peter wright, “Beginning Visual Basics 6” ,Tata McGraw Hill, Second Edition,2001

8.2. Wayne S Freeze,”Expert Guide To Visual Basics 6”, Tata McGraw Hill Publishers and Distributors, Third Edition,2000

8.3. Mrod Stephens ,”Visual Basic Programming”, Dreamtech Press, First Edition.

WEBSITES

8.4. www.citeseer.com

8.5. www.microsoft.com/speech/SDK/

9. Appendix

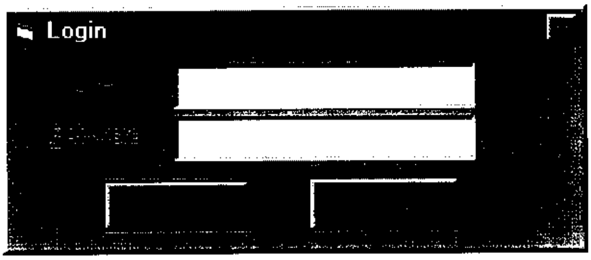
9.1. Source Code

```
Option Explicit
Public WithEvents RecoContext As SpSharedRecoContext
Public Sub read_text(str As String)
Dim a As Integer
If InStr(str, """") = 0 Then a = process_command(str)
End Sub

Private Function process_command(str As String)
Dim rs As rdoResultset
Set rs = cn.OpenResultset("select * from user_commands where
user_name='" & user_name & "' and module='main-1' and command='"
& str & "'", rdOpenDynamic, rdConcurRowVer)
        If rs.EOF = False Then
            Select Case rs("function_name")
                Case "change user"
                    sm_change_user_Click
                Case "configure users"
                    sm_configure_users_Click
                Case "change password"
                    sm_change_password_Click
                Case "logout"
                    sm_logout_Click
                Case "browser"
                            Click
        Case "scheduler"
          sm_scheduler_Click
        Case "commands"
          sm_commands_Click
        Case "play audio"
          sm_play_audio_Click
        Case "play video"
          sm_play_video_Click
        Case "windows explorer"
          sm_windows_explorer_Click
        Case "notepad"
          sm_notepad_Click
        Case "cascade"
          sm_cas_Click
        Case "tile horizontally"
          sm_tileh_Click
        Case "tile vertically"
          sm_tilev_Click
      End Select
    End If
End Function

Private Sub MDIForm_Activate()
  main = True
End Sub

Private Sub MDIForm_Deactivate()
  main = False
End Sub

Private Sub MDIForm_Load()
main = True
start_listen
End Sub
```

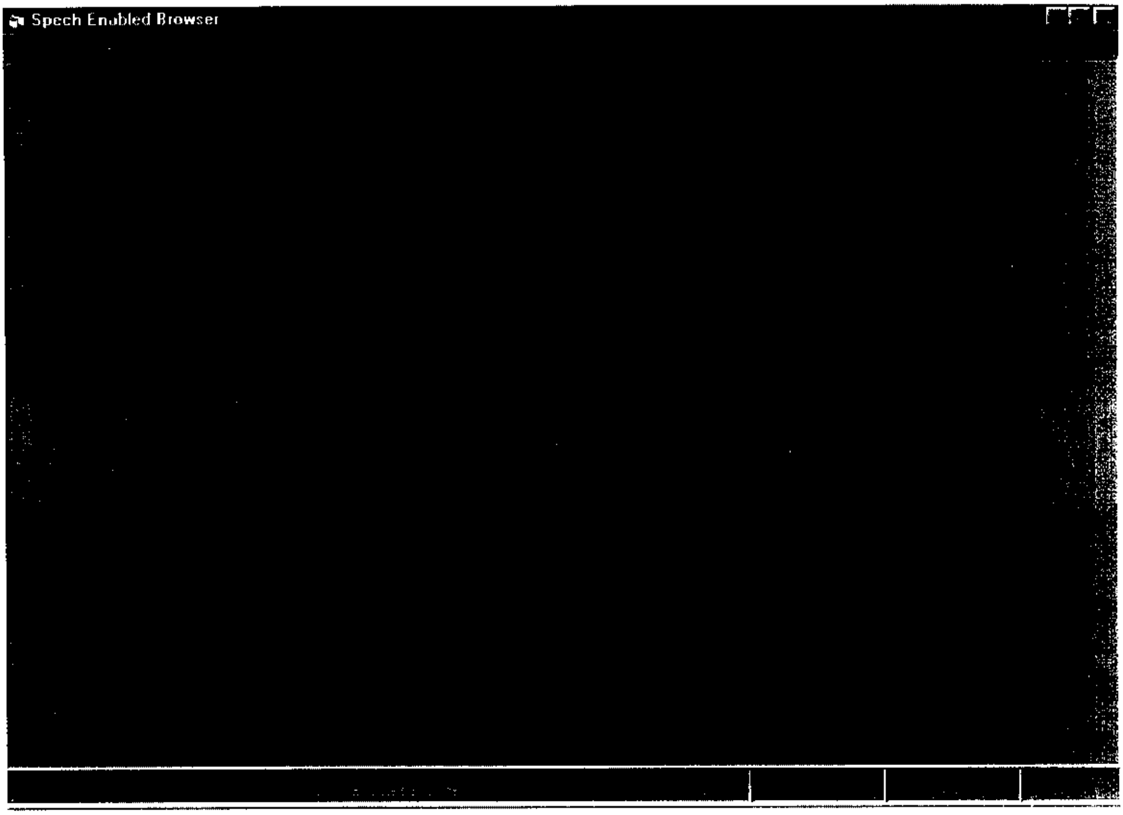
```
Private Sub RecoContext_Recognition(ByVal StreamNumber As Long,
ByVal StreamPosition As Variant, ByVal RecognitionType As
SpeechRecognitionType, ByVal Result As ISpeechRecoResult)
On Error Resume Next
    Dim strText As String
    Dim a As Form
    strText = Result.PhraseInfo.GetText
    StatusBar1.Panels(1).text = strText
    If main = False Then
       Set a = Me.ActiveForm
       a.read_text (strText)
    Else
       Me.read_text (strText)
    End If
 End Sub

 Private Sub MDIForm_Unload(Cancel As Integer)
 stop_listen
 End Sub

 Private Sub mm_browser_Click()
 frm_browser.Show
 frm_browser.SetFocus
 End Sub


 Private Sub sm_cas_Click()
    Me.Arrange (2)
 End Sub

 Private Sub sm_change_password_Click()
 frm_change_password.Show
```

```
End Sub

Private Sub sm_change_user_Click()
Unload Me
frm_Login.Show
End Sub

Private Sub sm_commands_Click()
   frm_user_commands.Show
   frm_user_commands.SetFocus
End Sub

Private Sub sm_configure_users_Click()
frm_configure_users.Show
frm_configure_users.SetFocus
End Sub

Private Sub sm_logout_Click()
If MsgBox("Are you Sure?", vbYesNo, "Alert") = vbYes Then
   Unload Me
End If
End Sub

Private Sub sm_notepad_Click()
frm_notepad.Show
frm_notepad.SetFocus
End Sub

Private Sub start_listen()
   If (RecoContext Is Nothing) Then
      Set RecoContext = New SpSharedRecoContext
      Set Grammar = RecoContext.CreateGrammar(1)
      Grammar.DictationLoad
   End If
   Grammar.DictationSetState SGDSActive
End Sub

Private Sub stop_listen()
   Grammar.DictationSetState SGDSInactive
End Sub



Private Sub sm_play_audio_Click()
   frm_audio.Show
   frm_audio.SetFocus
End Sub

Private Sub sm_play_video_Click()
   frm_video.Show
   frm_video.SetFocus
End Sub

Private Sub sm_scheduler_Click()
   frm_scheduler.Show
   frm_scheduler.SetFocus
End Sub



Private Sub sm_tileh_Click()
   Me.Arrange (1)
End Sub

Private Sub sm_tilev_Click()
   Me.Arrange (2)
End Sub

Private Sub sm_windows_explorer_Click()
   caller = 2
```

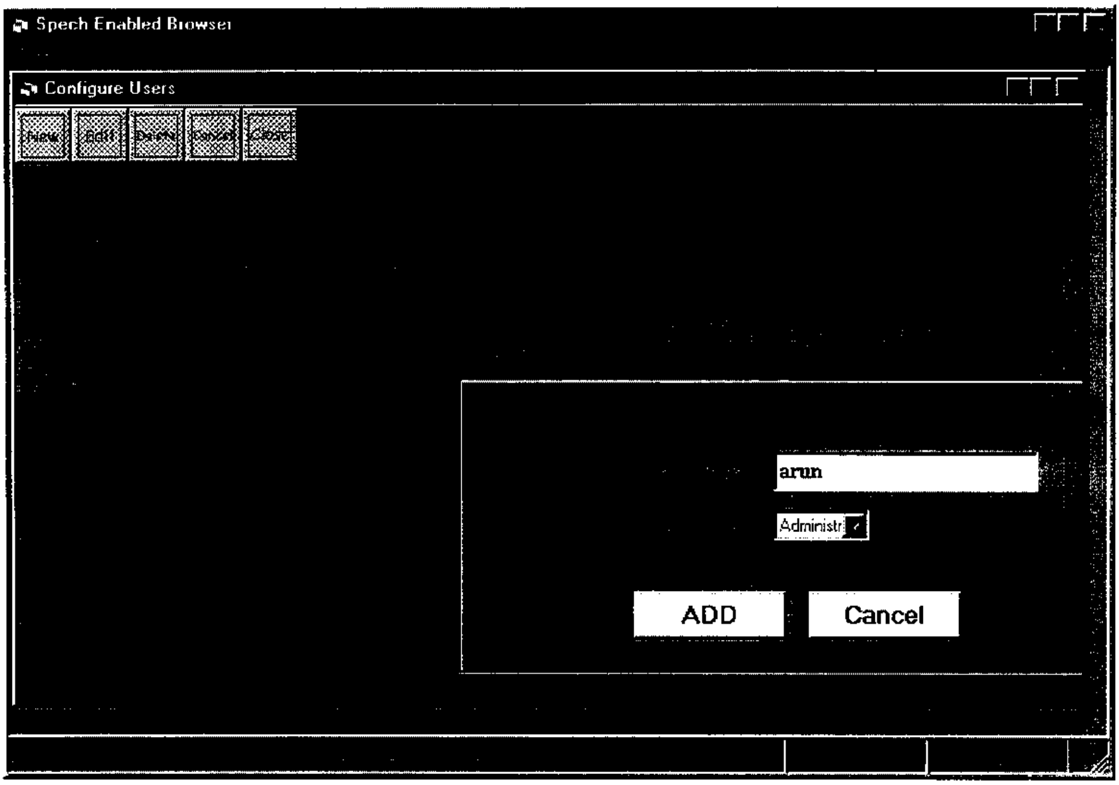
```
    frm_explorer.Show
    frm_explorer.SetFocus
End Sub

Private Sub Timer1_Timer()
    Dim rs As rdoResultset
'   MsgBox "select * from scheduler where user_name='" & user_name
& "' and schedule_date=#" & Format(Date, "dd-mmm-yyyy") & "# and
schedule_time=#" & Format(Time, "hh:mm") & "#"
    Set rs = cn.OpenResultset("select * from scheduler where
user_name='" & user_name & "' and schedule_date=#" & Format(Date,
"dd-mmm-yyyy") & "# and schedule_time=#" & Format(Time, "hh:mm")
& "#", rdOpenDynamic, rdConcurRowVer)
    If rs.EOF = False Then
        If rs("type") = "Message" Then
            frm_browser.Show
            frm_browser.SetFocus
            frm_browser.cboAddress = rs("url")
            frm_browser.cmd_go_Click
        Else
            MsgBox rs("message") & Chr(13) & Chr(10) & rs("URL")
        End If
    End If
```



Login

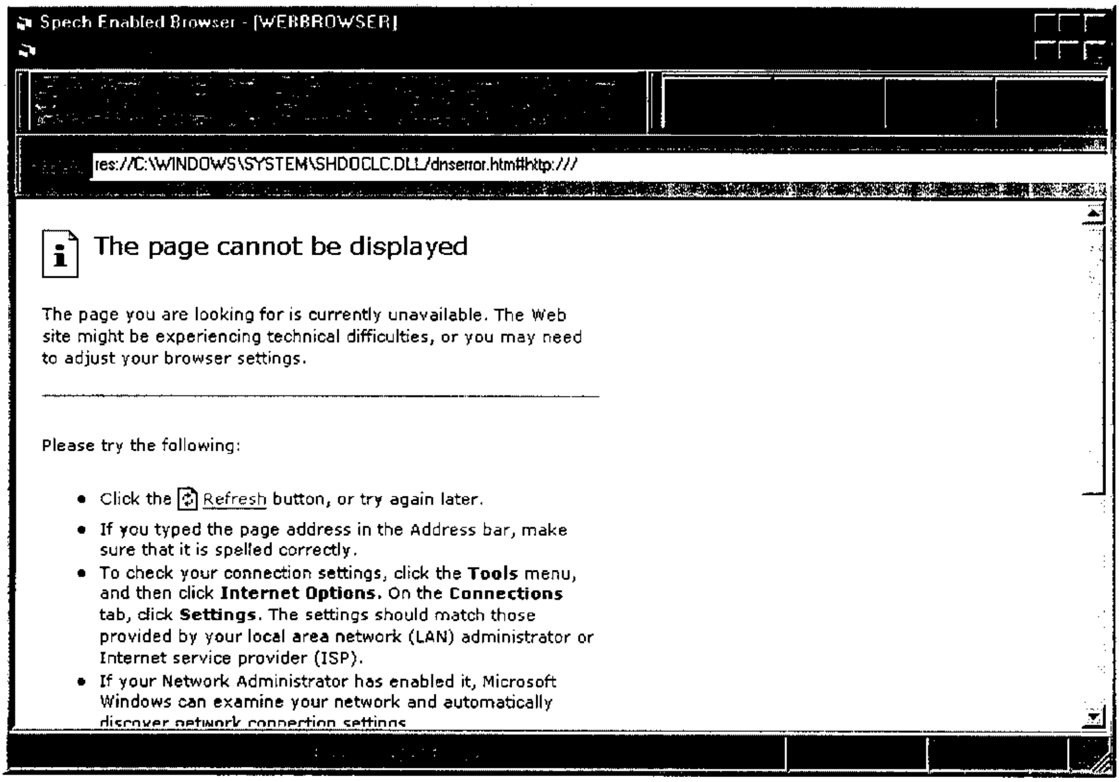


Spech Enabled Browser



arun

Administr



res://C:\WINDOWS\SYSTEM\SHDOCLC.DLL/dnserror.htm#http:///

The page cannot be displayed

The page you are looking for is currently unavailable. The Web site might be experiencing technical difficulties, or you may need to adjust your browser settings.

Please try the following:

- Click the Refresh button, or try again later.
- If you typed the page address in the Address bar, make sure that it is spelled correctly.
- To check your connection settings, click the **Tools** menu, and then click **Internet Options.** On the **Connections** tab, click **Settings**. The settings should match those provided by your local area network (LAN) administrator or Internet service provider (ISP).
- If your Network Administrator has enabled it, Microsoft Windows can examine your network and automatically discover network connection settings