# VOICE NAVIGATION SYSTEM

Project Report

Table of Contents

# Project Schedule

The project is of one month duration and will be carried out in four phases, devoting one week for each phase. Following is a detailed time schedule for the project:

## Phase I (Starting Week 1 : May 1,2002)

- Introduction to MS Speech Application Programming Interface (SAPI) SDK
- Basic VB Programs containing SAPI Components
- Routine Tasks on VB to enhance programming skills
- Concepts on working of Text to Speech (TTS) and Speech Recognition (SR)
- Starting with Project Ideas - Rough outlines, skeletons and System Features
- Emphasis on VB Components necessary for features to be included in VNS
- Start with preparing Software Requirements and Specifications

## Phase II (Week 2 : May 8,2002)

- Have a full-featured SRS document to Start - Design, Coding, Testing Cycle
- A Framework of VNS having scope of "Functional Modifications"
- Misc SDK Issues
  -Installing .vce files on systems
  -Installing SR Engine files on systems
  -Other necessary details
- Continue the Design, Coding, Testing Cycle till a Full Featured VNS is Running

## Phase III (Week 3 : May 15,2002)

- Exclude the Extra features, and let remain the Required VNS Goal Components
- Start system testing for First Final Release
- Consider additional issues (Application Features)
- Look for porting VNS to VC++
- First Final Release

## Phase IV (Week 4 : May 22,2002)

- Further testing for Final Release
- Define future scope of the project
- Provide documentation in the project
- Final Project Application
- Project Report

# Speech Recognition
# Based Computer Navigation System

## Introduction:

The main objective of speech recognition based computer navigation is to provide a voice command interface with GUI to the user through which user can navigate the system.

The goals of this software are:

- Provides a voice command interface to run all applications present in the system.
- Provide speech response to all text files.
- Editing option, which edits the existing commands to any new command.
- Interface to create new CFG file.
- Interface for updating the existing CFG files.
- Exploring & opening files using voice commands.
- Interface for controlling the attributes of text-to-speech engine.

## Information description:

In order to navigate the system the user has to use either keyboard or mouse. But in case if the user is busy with his hands/eyes or if the user is in a repetitive stress or injuries in such cases it is very difficult to navigate the system through keyboard or mouse.

This software provides a voice command interface for navigating the system. Using this software the user can run any applications or can open any files using voice commands. This software also provides the flexibility to user for creating their own voice commands & editing the existing commands.

### Project Components:

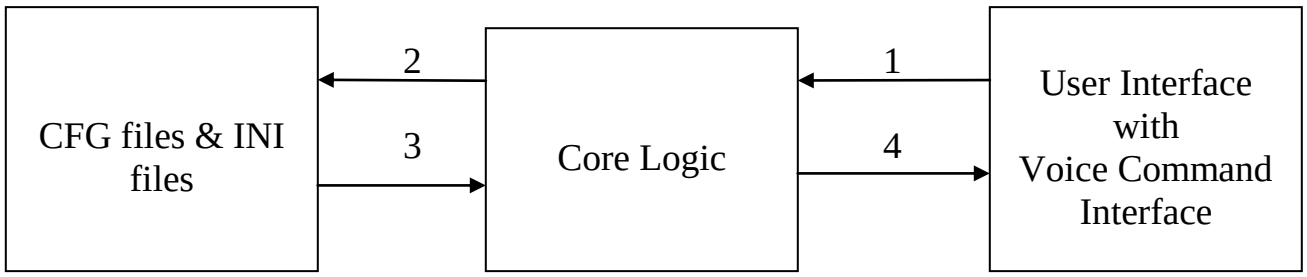



**Figure:** Module diagram of VNS

**Layer1:**

Layer 1 is the top layer, provides voice command & graphical user interfaces to user.

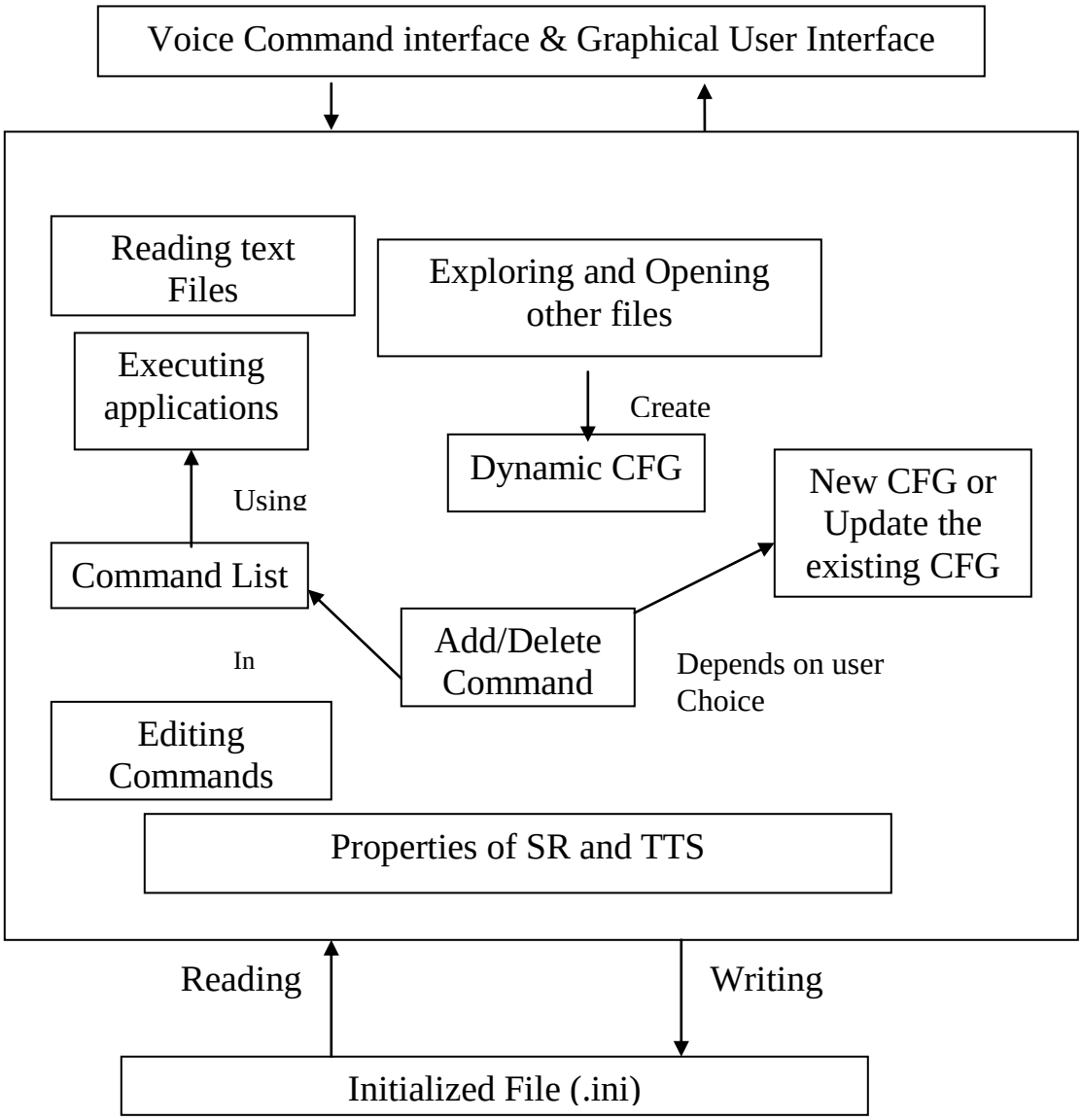
**Layer2:**

Layer 2 is the core logic module, which recognizes the voice commands, parsed it, performs computation and gives the appropriate result.

**Layer3:**

Layer 3 is a collection of CFG and INI files, which are used by the core logic.

## Module Diagram:



## Program Flow:

Speech recognition based computer navigation system runs in background and in foreground, while running in foreground it will serve some useful functionality and utilities.

1. Application starts with welcome with optional show/hide. The welcome screen is loaded only once during the start of the application.
2. When main application is runs in the background, it is visible in the system tray and it listens to previously selected grammar.
3. View command list; list all the commands with the respective paths from the selected grammar. This software listens to the previously selected grammar until a new grammar is selected it also provides the option of editing the commands in the grammar file.
4. User commands is a utility, which facilitates the user to create the own grammar file and selecting the grammar instead of working with the predefined grammar. It also deletes the selected grammar from the grammar file.
5. Plot four game gives the demo about how voice commands can be used in the game to make it more interactive and realistic. This game uses a very simple grammar. When game is loaded grammar is set to game CFG file and it will remain until game is closed.
6. System navigator is another utility, which uses dynamic CFG. Dynamic grammar is stored in the string. System navigator enables the user to navigate the system very easily through voice commands.

## User Interface and Voice Command Interface:

- Accepts the Voice Commands from the user.
- Runs the appropriate application.
- Reads the text file.
- Provides the interface to modify the properties of TTS.
- Provides the interface for creating and updating the CFG files.
- Provide interface to add/delete voice commands

## Hardware requirements:

- Sound card
- Micro head phone
- Pentium II/III processor
- 32/64 MB RAM

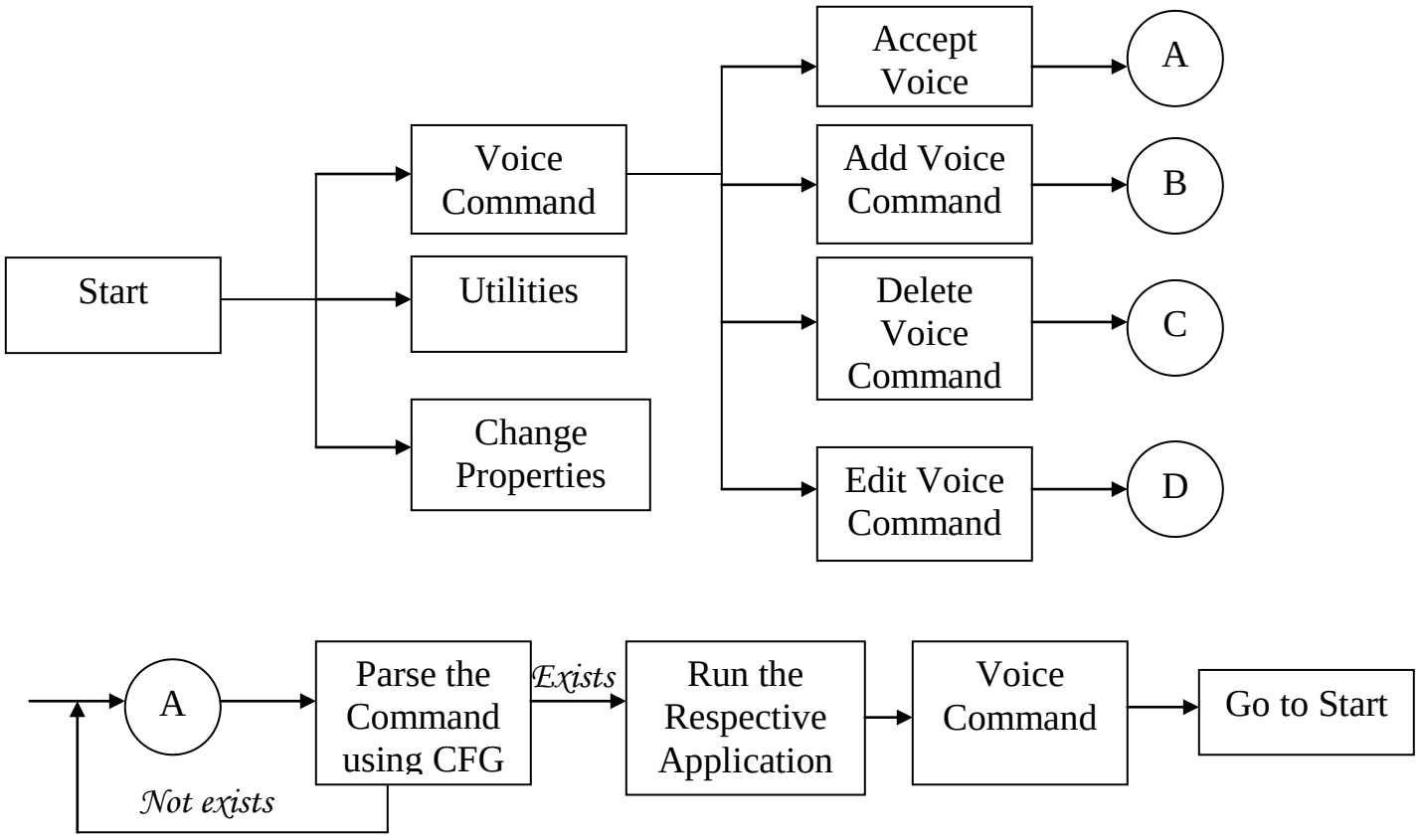
## Software requirements:

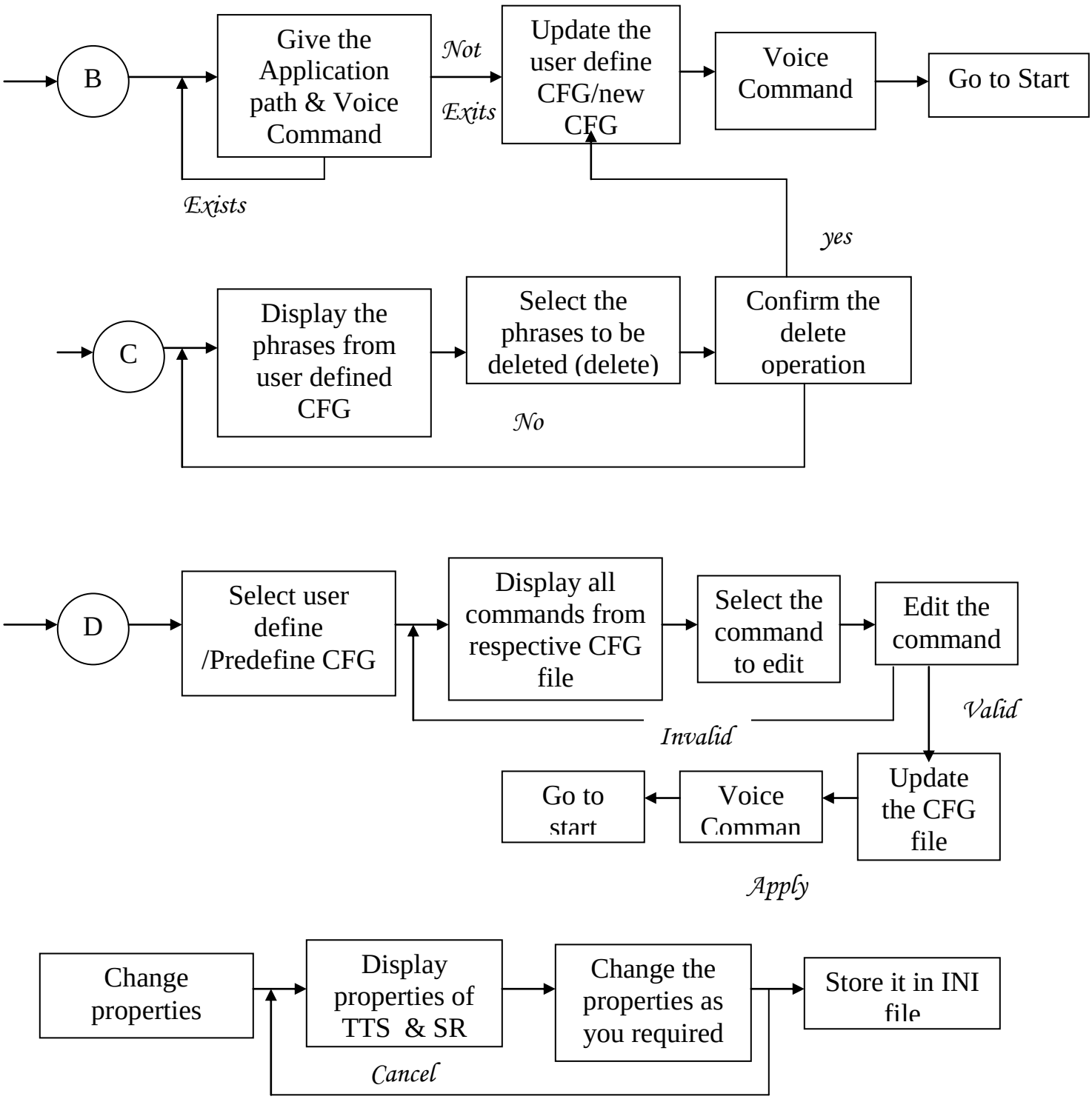
- Visual Basic version 6
- Microsoft Speech SDK version 4
- Windows NT, windows95/98 operating system
- Speech Recognition and Text-to-Speech engines

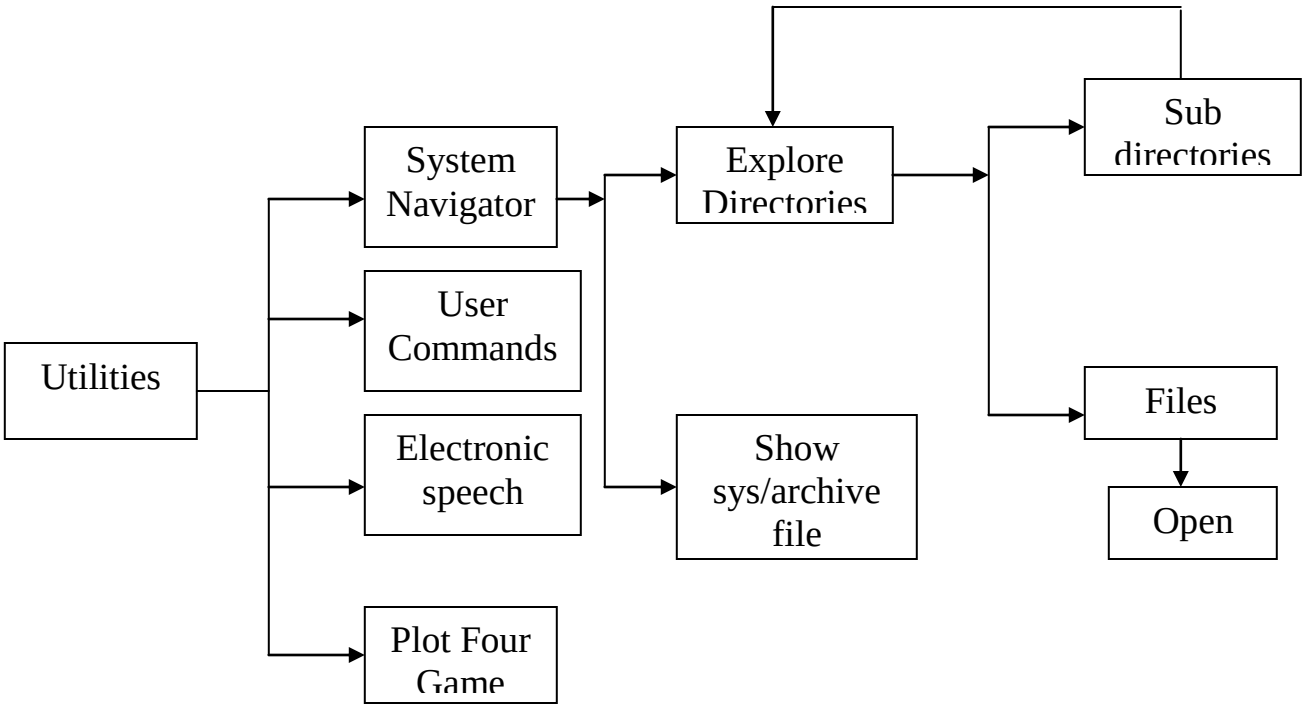
# Functional description:

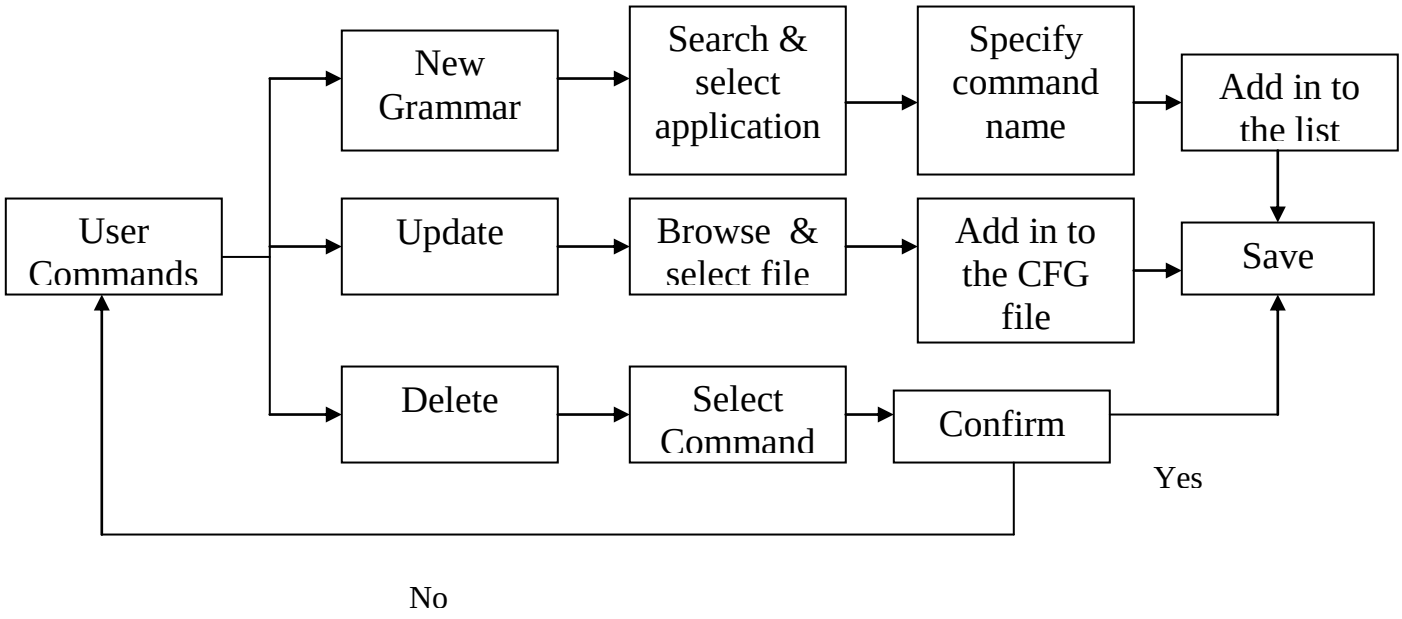
The application starts with the welcome form which contains options for running the application in background, show application & view command list, which displays a command list. Depending on users selection, respective tasks are performed.

**Block Diagram:**




B
Give the Application path & Voice Command
Not
Exits
Update the user define CFG/new CFG
Voice Command
Go to Start
Exists
yes
C
Display the phrases from user defined CFG
Select the phrases to be deleted (delete)
Confirm the delete operation
No
D
Select user define /Predefine CFG
Display all commands from respective CFG file
Select the command to edit
Edit the command
Valid
Invalid
Update the CFG file
Voice Comman
Go to start
Apply
Change properties
Display properties of TTS & SR
Change the properties as you required
Store it in INI file
Cancel




**Graphical User Interface**

**Menus**

Main form contains various menu options. Those are as follows:

**File menu:**

It contains the menu items as show application & run in system tray, which runs the application in foreground & background respectively.

**View menu:**

Menu items of view menu are command list & Options. Command list displays all the commands from the selected CFG file. It also includes editing of existing commands & replacing the edited commands with the default commands.

Option menu item is used to set the properties and attributes of text-to-speech engine. It set the random speech response and all speech responses depending on the users need.

**Utilities:**

Sub-Menu items are:

System Navigator:

It explores & opens the file & subdirectories present in the system by voice commands. Whenever a directory is explored will create a dynamic CFG and grammar will be selected from this dynamic CFG. Speaking the file names will cause the files to open with their associated editors. It also searches files depending on the file specification provided by the user. It is capable of displaying all hidden files, archive files & system files.

**User commands:**

This utility displays the commands along with the associated files from the CFG file selected from the combo box. Provides the flexibility for creating new CFG file by saying new grammar or clicking on the new grammar button, which invokes a form for searching .exe files, file are selected by clicking on the list containing file path. Clicking on the list gives an input box asking the command for the selected file. After clicking ok button will add all commands along with their file path in the user command form. Clicking on save button will create a new CFG file. CFG files can be updated by clicking update button or speaking update which invokes a add form where the user has to specify command and file name after clicking ok will add the new command to the existing file.

Commands can be deleted from the existing CFG file by clicking the delete button or speaking delete, which delete the selected command from the CFG file.

**Electronic Speech:**

It is the utility to provide speech response to all the text files. The user can open the text file & can hear the speech response to the opened file by clicking on the speak button. Speech response can be stopped at any time by clicking on the stop button. There are options for setting the properties and attributes of the text-to-speech engine.

**Plot Four Game:**
Plot four game gives the demo, how voice commands can be attached in the game to make the game very interactive and realistic. Player has to speak the column numbers to insert coin into the column.

**Help menu:**
Provides help about how to use speech recognition based computer navigation software.

## Bibliography and references:

Mastering Visual Basic 6