Bi-Lingual Audio Assistance Supported Screen Reading Software for the People with Visual Impairments

Md.Ruhul Amin
Computer Science &
Engineering,
Shahjalal University of Science
& Technology,
Sylhet,Bangladesh
shajib.sust@gmail.com

Bijan Paul
Computer Science &
Engineering,
Shahjalal University of Science
& Technology,
Sylhet,Bangladesh
bijancse@gmail.com

Faisal Ahamed Khan
Computer Science &
Engineering,
Shahjalal University of Science
& Technology,
Sylhet,Bangladesh
faisal.cse06@gmail.com

Abstract— Accessibility of computer for visually impaired people is a Challenging task. This paper presents software named Mongol Dip, which provides audio-based interactive interfaces to help people with visual impairments to explore and navigate windows operating system. In this software we have used text to speech technology (TTS) to echo every operation done by the user. Mongol Dip provides the easiest interface that helps the visually impaired people to access computer. Visually impaired people cannot read out Bangla text document via computer because existing software they use to operate computer do not support Bangla text. Bi-lingual (English & Bangla Language) support is the main significant part of our software. Hence the main goal of this software is to assist the people with visual impairments to operate computers and connecting with the digital world using minimum operations and maximum usability using bi-lingual TTS support.

Keywords— Mongol Dip, TTS, Visual impairments, Subachan

I. INTRODUCTION

According to the statistics of World Health Organization (WHO) and the International Agency for the Prevention of Blindness (IAPB) [1] in World Sight Day-2011 Survey – the number of visually impaired people in our country is 7, 50,000. If these visually impaired people can use computers like other normal human being then it will open a new opportunity for their education and self improvement. To fulfill this objective, we visited the IER department of the University of Dhaka, Bangladesh. There we met some people (students and teachers) who were visually impaired. They were operating computers using JAWS [2], a screen reading software in English language only. But during a discussion they pointed out that screen reading software with bi-lingual TTS support would let them use the computers more easily and enable them to participate in traditional education system. So we decided to develop Mongol Dip to help them to use computers efficiently and also give bilingual support so that they can read Bangla text.

To develop Mongol Dip, we studied the ways a visually impaired people can follow to use computer easily. Then we

decided to provide couple of very important functionalities in our software which are mostly used by a normal person, such as audio assisted accessibility to any file and folder from the hard disk drive using customized browser, working with documents (using MS-Word, MS-Excel, MS-PowerPoint, Notepad, PDF etc), listening music (using Media Player), installing or removing any kind of software, and sending or receiving email using customized client. To provide the functionalities successfully, we have developed our own customized browser using which user can locate programs and files easily and can work with them.

The remaining sections of this paper are organized as follows: section 2 describes related works; section 3 describes system requirement. Section 4 describes system architecture of Mongol Dip, section 5 shows performance measure & feedbacks of Mongol Dip software, and Section 6 presents comparison of Mongol Dip with other software. We conclude in Section 7 and at the end references are added.

II. RELATED WORKS

There is several software such as JAWS [2], NVDA [3], SuperNova [4] for visually impaired people to operate computer. Most of the software mainly uses screen reading technique and none of them support any Bangla text reading. Among them JAWS is most popular which is used mainly for working with documents. In general, a screen reader [5] software application reads aloud information displayed on a computer monitor screen. It reads aloud text within a document, information within dialog boxes and error messages. It also reads aloud menu selections, text with the graphical icons on the desktop. On the other hand, Mongol Dip can access everything within a computer using a custom built browser which is absent in other screen reading application. We have used Microsoft Windows Accessibility API like other application to detect user key-stroke event and echo every operation done by user. We achieved significant difference with other software in terms of computer accessibility for the visually impaired people by providing the support for reading and writing in both the English and Bangla text. We also provide a customize browser with short-cut key features using which anyone can operate computer easily and this feature can be considered unique comparative to other software.

III. SYSTEM REQUIREMENTS

This software is designed to operate Windows 2007 and Windows XP.

A. Software Requirement

- o Java platform (Run Time Environment)
- o .Net framework 3.5
- Microsoft office 2007
- Subachan Bangla TTS Software

B. API & PACKAGES

To handle various types of applications for customized input we have used system32.dll, user32.dll, Kernel32 etc. For every operation, Mongol Dip echoes the corresponding audio instruction to assist user to complete the desired task. In case of reading MS office application we have used MS office object library and for pdf we have used PDFBox library to get text from these files. For sending/receiving mail we have built an application by which a user can send/receive mail without going to Internet browser. For sending mail we have used SMTP [6] protocol and for getting mail IMAP [7] protocol. Users can also use internet by using a normal internet browser and during browsing they can get sound output of the text at the position of their pointer. Different types of packages, APIs and libraries we have used in Mongol Dip are described below:

- Microsoft TTS (text-to-speech): Microsoft TTS [8] helps the operating system to play back printed text as spoken words. Microsoft TTS engine recognizes the text and using a synthesized voice to speak the written text.
- 2) Bangla TTS (Subachan): Subachan [9] is a Bangla TTS software developed in SUST to read Bangla text as spoken Bangla word. We integrated Subachan with Mongol Dip software. Thus Mongol Dip gives Bilingual (English and Bangla Language) audio assistance support.
- 3) *MS office object library:* This library is used to handle the Microsoft office properly [10].
- 4) *Windows media player package:* It is from Microsoft.Net to handle the windows media player.
- 5) Global hook: We have used this module for getting low level keyboard and mouse I/O and monitoring keyboard and mouse operations from everywhere in the system [11].

- 6) System dll files: System dll files [12] are used to get menu text, window title, button text etc from different types of window, as well as to get system resolution, screen image, global keyboard and mouse hook.
- 7) *SMTP & IMAP protocol:* For sending mail we have used SMTP [6] protocol and for getting mail IMAP [7] protocol.
- 8) *Microsoft Accessibility API:* To access windows all the windows functionality from Mongol Dip, we have used Microsoft Accessibility API [13].

IV. SYSTEM- ARCHITECTURE

System-Architecture is the basic structure of the system that defines the essential core design features. By using flow chart, here we show Mongol Dip System-Architecture.

A. System Architecture of Mongol Dip Browser In the flowchart, at the end of every operation users always return to the Mongol Dip Browser.

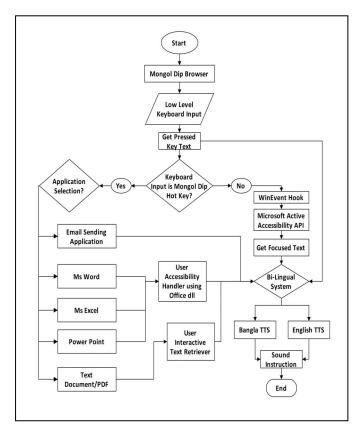


Figure 1. Overall flow-chart of our system

B. System Architecture of Sending/Receiving E-mail

To send or receive mail using our software, a user has to run the application using hot key Ctrl+3. At first, users need to login. Then by selecting send button they can send mail, can read mail by selecting access mail button, and close button for exiting E-mail sending window application.

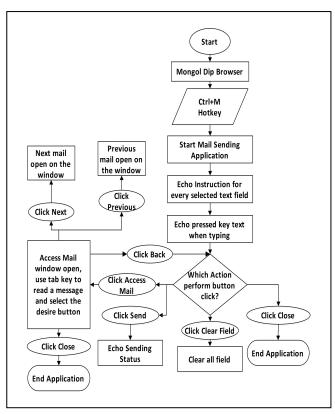


Figure 2. Flow-chart of sending/receiving mail

C. System Architecture of Microsoft Office Operation

To launch MS word new document users have to press hotkey Ctrl + W. For editing a document they need to know the key combinations for working on the document. The various operations for working with ms document are shown in the following two flowcharts.

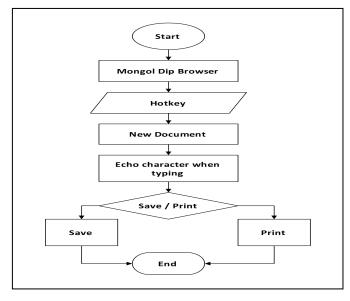


Figure 3. Flow-chart for MS Open operation

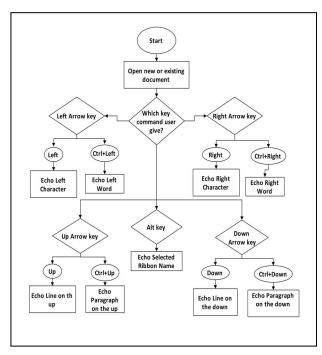


Figure 4. Flow-chart for MS Edit operation

D. Bi-lingual System Pseudo Code

(6) Exit

```
Here 'Text' contained Bangla English mixed text.
BiLanguageSystem(Text):
(1) Initialize a character array name charArray and string array name
stringArray
(2) Initialize c=0, englishText="", and banglaText=""
(3) Text= Text+' '+"Null"
(4) stringArray=Text.split(' ')
(5) for every word in stringArray repeat:
  (i) if(word!=""):
     (i.1) if(word!="Null"):
          (a) charArray= convertToCharacterArray(word)
          (b) if(charArray[0]>=0 and charArray[0]<=127):
                  (b.1) englishText=englishText+word+' '
                   (b.2) if(banglaText!=""):
                         (b.2.i) speak(banglaText)
                                   (b.2.ii) while(True):
                              if(banglaText speaking is complete):
                     (b.3) c=1
          (c)else if(unicode(charArray[0])>=2432 and
               (unicode(charArray[0])<=2559)):
               (c.1) banglaText=banglaText+ word +' '
              (c.2) if(englishText!=""):
                       (c.2.i) speak(englishText)
                       (c.2.ii) while(True):
                              if(englishText speaking is complete):
                (c.3) c=2
     (i.2) else:
               (a) if (c==1 and englishText!=""):
                        speak(englishText)
               (b) if (c==2 and banglaText!=""):
                        speak(banglaText)
```

Note:

- (1) If we find Bangla word at first add this in banglaText variable until we will find English word.
- (2) If we find English word at first add this in englishText variable until we will find Bangla word.
- (3) When find English word then speak previous banglaText.
- (4) When find Bangla word then speak previous englishText.
- (5) If word == "Null" then speak remaining banglaText or englishText.

E. Integrate Microsoft Accessibility API with Mongol Dip Pseudo Code

When Mongol Dip Start SetWinEventHook for call back every focus event

- (a) Initialize IntPtr name handler and object name ChildId
- (b) #region APIs:

[DllImport("oleacc.dll")]

public static extern uint WindowFromAccessibleObject(IAccessible
pacc, ref IntPtr phwnd);

[DllImport("oleacc.dll")]

private static extern IntPtr AccessibleObjectFromEvent(IntPtr hwnd, uint dwObjectID, uint dwChildID, out IAccessible ppace, [MarshalAs(UnmanagedType.Struct)] out object pvarChild); #endregion

(c) SetWinEventHook(

Initialise system event
IntPtr.Zero,
CallBackFunction,
(uint)0,
(uint)0,
WINEVENT_OUTOFCONTEXT

- (d) CallBackFunction(IntPtr hWinEventHook, uint eventType, IntPtr hWnd, uint idObject, uint idChild, uint dwEventThread, uint dwmsEventTime)
- (1) handler=AccessibleObjectFromEvent(hWnd, idObject, idChild, out iAccessible, out ChildId)
 - (2) WindowFromAccessibleObject(iAccessible, ref handler)
 - (3) if (iAccessible != null && ChildId != null):

Speak(iAccessible.get accName(ChildId).toString())

(e) Exit

Note:

- (1) Set WinEvent hook, which call back the function for every windows event
- (2) In call back function do the following:
 - (a) Call AccessibleObjectFromEvent to retrieve an IAccessible object
 - (b) Call WindowFromAccessibleObject
- (c) Retrieve focus text using get_accName method of IAccessible object
- (3) Exit.

V. PERFORMANCE MEASURE & FEEDBACK

Using computer with minimum number of operations will be a great achievement for the visually impaired people. So our target is to support them by providing maximum usability with minimal operations. We also want to ensure that one will be able to access the computer using this software without any prior knowledge. For this reason we are controlling applications by hot key combinations. They can start using the software by pressing "F1" that provide audio directions for using different key combinations to launch different programs. Again for each application same key combination (F1) provides the audio directions for different operations attached to that application. When we went to the IER dept. after developing demo version of our software, the visually impaired student appreciated this idea very much. To get real user feedbacks we went to Bangladesh Visually Impaired People's Society (BVIPS) [14] and Green Disabled Foundation (GDF) [15]. Bangladesh Visually Impaired People's Society (BVIPS) is a self-help voluntary organization of the visually impaired people. In Mongol Dip software development phase, we elicited requirement via meeting with visually impaired people of BVIPS several times. They appreciated us for the Bi-lingual features very much.

In the following sections we are providing the details of using some important applications through Mongol Dip along with the user's feedback that we have collected from the teachers and students of IER Dept. of Dhaka University, Bangladesh Visually Impaired People's Society (BVIPS) [14] and Green Disabled Foundation [15]. The number of users who provided their valuable opinion for this software is as follows:

Name	Attendance(Person)
IER Dept, Dhaka University	30
BVIPS	40
GDF	20

A. Hotkey executed Programs

All the application that can be used through Mongol Dip can be launched and executed by using hotkeys. The process of how these applications are launched and executed is shown in the following simple figure.

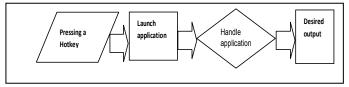


Figure 5. Way to handle hotkey executed application

B. MS Office Support

Our software can process MS word 97-2007, MS excel, MS PowerPoint, Notepad, WordPad and many other text editors. For example: One can launch Ms Word, MS Excel, MS PowerPoint or Text editor by pressing "Ctrl+W" or "Ctrl+E" or "Ctrl+P" or "Ctrl+T" respectively, where one can type anything in Bengali/English. Mongol Dip makes echo of each character at the time of typing and also reads out the complete word after the completion of typing a single word to help composing.

A person can have audio assistance for each character, or for a complete word, or for a complete line, or for a complete paragraph, or from the start to end of a document using different key combinations. For example if I want to have audio output of a line above my current position then I have to press "Ctrl+ Up" arrow.

Below we are giving a table showing the users feedback in case of using MS-Office.

TABLE I. FEEDBACK FROM IER DEPT. & BVIPS ABOUT MS OFFICE

Operation	Scope	Feedback	Satisfaction
			(%)
Launching	Press Ctrl + W(Word)/	Useful feature	Above 70%
new	E(Excel)/P(PowerPoint)/		
document	T(Text Editor) Key		
Typing,	Same as JAWS with	Need to learn	50%
editing	diff. hotkey	the combination	
Audio	We used Microsoft TTS	Little bit speedy,	Suggested
output	and Bengali TTS	and Bengali	us to
-		voice is not	change the
		clear	voice

C. Sending and receiving mail

Users will be able to run the program for sending mail and reading all unread mail using "Ctrl+M". Firstly, one has to login using the login form, then for sending mail s/he has to fill up the field To, Subject, Attachment, Body then press send key. By pressing receive button, one can get his all unread messages info (such as: sender name, delivery date, no of attachment and the message). When one goes to the desired field using tab key our software provide audio output of the field contents. Suppose when user will go to the Body field of a compose mail our software will echo, "Type your message body here". Below we are providing the feedback in case of using customized mailing application of Mongol Dip.

TABLE II. FEEDBACK FROM IER DEPT. & BVIPS ABOUT MAILING

	Scope	Feedback	Satisfaction
Operation			(%)
Launching	Hotkey Ctrl + M	Very user	Satisfied (90%)
Ū		friendly	, ,
Selecting	Just to use tab key	Also appreciated	70%
new, inbox	-		
Audio	Same as MS-Word	Same as MS-	Same as MS-
output		Word	Word

1) Comparisons between requires time to send mail from this software and send mail using browser

Case Study	Net Speed	Visual Impaired User	Normal user
1	15KB/S	10 minute	3 minute
2	15KB/S	8 minute	4 minute
3	15KB/S	5 minute	2.5 minute

2) Comparisons between require time to read mail message using this software and read mail message using browser

Case Study	Net Speed	Visual Impaired User	Normal user
1	15KB/S	5 minute	2 minute
2	15KB/S	3 minute	2 minute
3	15KB/S	2 minute	3 minute

D. Media files:

By accessing a folder using the customized browser of Mongol Dip, when a user selects a file s/he hears the name of the file and its artist. When one presses an enter key then our software launch the Windows Media Player and start to play. One can use "n" key for the next media file, "p" for the previous file, "<space>" for pause and resume, "+" and "–" for increasing and decreasing volume, and left and right arrow for backward and forward.

TABLE III. FEEDBACK FROM IER DEPT. & BVIPS ABOUT HANDLING MEDIA FILE

Operation	Scope	Feedback	Satisfaction (%)
Launching	Just to select media file and press Enter key	Good	90%
Controlling	Some shortcut key for handling media files	Easy to handle	70%

Mongol Dip provides a customize browser that helps visually impaired people to operate computer easily. Because of using software like JAWS, NVDA for a long time it might be a bit difficult for visually impaired people to maintain Mongol Dip browser. But in the long run it would be beneficial for its extended capability. This is because we have added various short cut key for accessing the windows files and program via Mongol Dip Browser.

VI. COMPARISON

Operating windows is easier for visually enabled users, but not for visually impaired. So we tried to build a system which is easier to use in their perspective. JAWS[2] is a screen reading software that is used by visually disabled people. But it only gives audio output of every operation and the texts wherever user's mouse pointer resides. It has no customized options which can help the visually disabled people to use Windows. In a sense it is similar to the Microsoft narrator program. From this perspective, Mongol Dip has some very attractive and innovative features which are very important for visually impaired people. In the following table we have shown comparisons between JAWS and Mongol Dip.

TABLE IV. COMPARISON BETWEEN JAWS AND MONGOL DIP

SL.	Criteria	JAWS	Mongol Dip
1	Language	Only for English	Bi-Lingual for Bengali & English
2	Mail	Sending & receiving is complicated, need to go to web browser	No need to go to web browser
3	Text editor	Composing is tough	Easier than JAWS
4	Launching application	No hotkeys to launch	Hotkey for each application
5	Media player	Not so much easier	Easier than JAWS
6	Customize Browser	No Customize browser	Provide customize browser to access computer easily.

VII. CONCLUSION

The main strength of this software is that it fills an essential need of visually impaired population. Generally people have five senses using which they interact with the outer environment. In these days of information technology everyone wants to be updated with latest technology and services. But the visually impaired people did not have much scope to get the benefits of technology revolution. Mongol Dip will give visually impaired population a great opportunity to connect with and contribute to the digital world. Mongol Dip helps the visually impaired people to use Windows operating system to complete official tasks easily like a normal human being. We hope and believe that Mongol Dip will make the visually impaired people technology enabled manpower of our nation. The Mongol Dip screenshots and operation details are given in annex1 and annex2.

ACKNOWLEDGMENT

"Mongol Dip: Light for the People with Visual Impairment" is a HCI type project. This project is developed for the visual impaired people of Bangladesh and the project is sponsored by "Ministry of Science and Information Technology, Bangladesh. The first version of this software has been released on 6th October, 2012 in Dhaka. Bangladesh Visually Impaired People's Society (BVIPS) members always co-operated with us for the development of this software.

REFERENCES

- Statistics of World Health Organization (WHO) and the International Agency for the Prevention of Blindness (IAPB) in World Sight Day-2011.
 - http://www.almustafatrust.org/content/Downloads/forms/eyecamp-2012-web.pdf
- [2] Jaws, "Screen Reading Software", http://www.freedomscientific.com/products/fs/jaws-product-page.asp
- [3] NVDA, "Screen Reading Software",
- http://www.nvda-project.org/wiki/About
- [4] SuperNova, "Screen Reading Software", http://www.enablemart.com/Microsoft/SuperNova-Screen-Reader
- [5] Screen Reader http://www.abilityhub.com/vision/index.htm
- [6] SMTP protocol http://en.wikipedia.org/wiki/Simple Mail Transfer Protocol
- [7] IMAP protoco http://en.wikipedia.org/wiki/Internet_Message_Access_Protocol
- [8] Diana Huggins, "Configure Text-to-Speech in Microsoft Windows XP", Link:
- http://i.t.com.com/i/tr/downloads/home/hdi_configure_text_speech.pdf

 [9] Abu Naser, Devjoyti Aich, Md. Ruhul Amin, "Bengali Text to Speech Synthesis Software", Developed for the fulfilment of undergraduate thesis, Dept. of CSE, Shahjalal University of Science and Technology, ICECE 2010.

 http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5700757&abstra
- ctAccess=no&userType=inst
 Programming Office Applications Using Visual C#,
- http://msdn.microsoft.com/en-us/library/aa537157%28office.11%29.aspx
 [11] Global hooking, A Simple C# Global Low Level Keyboard Hook, Link:
 http://www.codeproject.com/KB/cs/CSLLKeyboardHook.aspx
- [12] System dlls, Wiki for .NET Developers, Link: http://pinvoke.net/
- [13] http://en.wikipedia.org/wiki/Microsoft_Active_Accessibility
- [14] http://www.sightsavers.org/in_depth/quality_and_learning/organisationa l_monitoring/[SIM_Card_case_studies]/13980.html
- [15] http://gdfbangladesh.org/2010/03/welcome-to-the-green-disabled-foundation/

ANNEX1: SCREENSHOTS OF MONGOL DIP



Figure 6. Screen Shot of Mongol Dip Browse

By using the customize browser[Figure 6] user can locate files easily and can work with them. Because of Microsoft Windows Accessibility API user can access anything from computer.



Figure 7. Screen shot of sending mail window

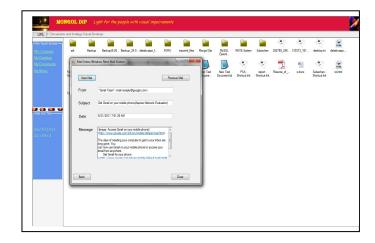


Figure 8. Screen shot of receiving mail window



Figure 9. Subachan Bangla Spell Checker

ANNEX2: OPERATION DETAILS OF MONGOL DIP

A. Mongol Dip Browser

Events	Short-cut Keys
To know any keyboard key name	Press the desire key
Instruction of how to operate Mongol Dip Browser	F1
To Select File/Folder from current link	Press Right/left /up/down key
To Open/Execute/Enter current selected File/Folder	Press Enter
To Browse link such as My computer, My document, Desktop, Music	Press Tab
To Terminate Any Opening Application	Press Altr + F4
To Hear Today's Date & Time	Press Insert + F12
To Shut-Down Computer	Press Insert + F4
To Open New Microsoft Word Document	Press Ctrl + W
To Open New Text Editor Document	Press Ctrl + T
To Open New Excel Document	Press Ctrl + E
To Open New PowerPoint Document	Press Ctrl + P
To Open E-mail Sending & Receiving Window	Press Ctrl + M
Copy File/Folder	Press Ctrl + C
Paste File/Folder	Press Ctrl+ V
Edit File Folder Name From Mongol Dip	Press F2
Browser Delete File/Folder	D. D.L.
_ 57575 _ 57575	Press Delete
To Increase Speaker Volume	Press Ctrl + D9 or Press +
To Decrease Speaker Volume	Press Ctrl + D8 or Press
To Increase Speech Speed	Press Ctrl + D7 or Press *
To Decrease Speech Speed	Press Ctrl + D6 or Press
To Add a new Folder	Press Ctrl + F
To on-off Speaker Sound	Press Space key

B. Microsoft Word Document

Events	Short-cut Keys
Instruction of how to operate Word Document	Press Ctrl + Numpad 0 or
	Ctrl + D7
To Listen One Character From Left	Press Left Arrow
To Listen One Character From Right	Press Right Arrow
To Listen One Word From left	Press Ctrl + Right Arrow
To Listen One Word From Right	Press Ctrl + Left Arrow

To Listen One Up Paragraph	Press Ctrl + Up Arrow
To Listen One Down Paragraph	Press Ctrl + Down Arrow
To Listen Current Page Number, Current Text	Press F2
Font Size, Current Text Font Name Instruction	
To Detect Table Cell Position	Press Tab/Up.
Line Number of Document	Press Up or Down Key
To create new office Document	Press Ctrl + n / N

C. Microsoft PowerPoint Document

Events	Short-cut Keys	
To Listen One Character From Left in Text Edit Mode	Press Left Arrow	
To Listen One Character From Right in Text Edit Mode	Press Right Arrow	
To Listen One Word From Left in Text Edit Mode	Press Ctrl + Left Arrow	
To Listen One Word From Right in Text Edit Mode	Press Ctrl + Right Arrow	
To Listen Paragraph From Up to Down	Press Ctrl + Down Arrow	
To Listen Paragraph From Down to Up	Press Ctrl + Up Arrow	
To go next slide	Press Down Arrow	
To go previous slide	Press Up Arrow	
To select Text area/Shape from slide document	Press tab	
Edit operation when make slide	Press F2 or Enter	
To deselect anything	Press Escape key	
Slide show start and stop	Press F5 key	
Go prior slide when slideshow running	Press PageUp / BackSpace key	
Go next slide when slideshow running	Press PageDown / Space key	
To read text of slide when slideshow running	Press Up / Down / Right / Left Arrow key	
To create new presentation	Press Ctrl + n / N	
To create new slide	Press Ctrl + m / M	
To Save Currently Working PowerPoint document	Press Ctrl + S	
To Save Current working PowerPoint document & close	Press Altr + F4	
To go top most slide	Press Ctrl + Home	
To go last slide	Press Ctrl + End	

D. Microsoft Excel Document

Events	Short-cut Keys
Detect current selected cell position & it's text	Using Up, Down, Right,
_	Left, Tab keys
To select current selected all rows	Press Shift + Space key
To select current selected all columns	Press Ctrl + Space key
Sheet Instruction	Press Ctrl + Shift +
	PageUp / PageDown key
To Know Current Cell Top Header	Press Ctrl + E key
Current cell 1 st to 4 th row text	Press Ctrl + D6 key
Current cell 1 st to 4 th column text	Press Ctrl + D7 key
Today's Date Entry in current cell	Press Ctrl+ Semicolon key
Today's Time Entry in current cell	Press Ctrl+ Shift +
	Semicolon key
To update cell text	Press F2 edit mode active
	then press right or left key
	to read and type to write
To know Current Cell Position & Text again	Press Ctrl+ M key