२५ मे २०१२
यशदा, पुणे

Training on-
Use of Unicode & ISM 6.0

Submitted by-

Pravin Kolhe

Executive Engineer,
MI Division No-2, Akola
Water Resource Departmen
Government of Maharashtra.

जा. क्र. १६९३/२०१२	दि. २८/०५/२०१२

प्रति,

अधीक्षक अभियंता,
अकोला पाटबंधारे मंडळ,
अकोला

विषय: Training on Unicode & ISM 6 या प्रशिक्षणाबाबतचा अहवाल सादर करणेबाबत

संदर्भ:-

१) महाराष्ट्र शासन, पाटबंधारे विभाग, परीपत्रक क्र टी.जी.ओ. १०००/(३७५/२०००)आ(प्रशि),
मंत्रालय, मुंबई, दिनांक २४ नोव्हेंबर २०००.

उपरोक्त विषयान्वये दि. २५ मे २०१२ रोजी, यशवंतराव चव्हाण विकास प्रशासन प्रबोधिनी, (यशदा), पुणे येथे Training on Unicode & ISM 6 या विषयावरील एकदिवसिय प्रशिक्षण वर्गास या विभागाचे कार्यकारी अभियंता श्री. प्र. शि. कोल्हे सहभागी झाले होते.

उपरोक्त संदर्भिय शासन परीपत्रकान्वये प्रशिक्षण बाबतचा सविस्तर अहवाल त्यांचे विभागप्रमुखांचे मार्फत शासनास सादर करण्याच्या सुचना निर्गमित केलेल्या आहेत. त्या अनुषंगाने सोबत प्रशिक्षण बाबतचा सविस्तर अहवाल आपल्या माहितीस्तव व पुढील कार्यवाहीस्तव सविनय सादर करत आहे.

प्र. शि. कोल्हे,
कार्यकारी अभियंता

सहपत्र:- उपरोक्त संदर्भिय अहवाल (२ प्रतींमध्ये)

Training Report on

Use of Unicode & ISM 6.0

25 May 2012

At

Yashawantrao Chavan Academy of Development Administration, Pune

Submitted by-

Er. Pravin Kolhe,
**Executive Engineer,
Minor Irrigation Division No.-2, Akola**

Executive Summary

महाराष्ट्र शासनाने कार्यालयाच्या संगणकीय कामकाजामधेय युनिकोड आज्ञावलीचा वापर करण्याबाबत दिनांक १० सप्टेंबर २००८ रोजी परिपत्रक निर्गमित केले होते. त्यानंतर दिनांक १६ डिसेंबर २०१० च्या शासन निर्णयान्वये सर्व शासकीय कार्यालयात तसेच सार्वजनिक उपक्रमांमध्ये मराठी युनिकोडचा वापर करण्याबाबत सविस्तर सुचना देण्यात आल्या होत्या.

शासकीय कामकाज हे मराठी (ISM 3.04) मध्ये होत असुन या कामकाजात वापरण्यात येणारी मराठी लिपी ही युनिकोड बेस नसल्याने मराठीतील पाठविण्यात येणारा डाटा हा शासनातील सर्व संगणकावर दिसेल, याबाबत खात्री नसल्याने माहिती व तंत्रज्ञान संचालनालयाने केंद्र शासित मे.सी-डॅक या कंपनीने विकसित केलेले Unicode ISM Version 6.0 हे महाराष्ट्र शासनातील सर्व कार्यालयांकरीता खरेदी केलेले आहे.

यापुर्वी वापरण्यात येणारे ISM 3.04 हे Microsoft Office 2003 (MS2003) किंवा Microsoft Office 2007 (MS2007) या आज्ञावलीवरील मराठीत केलेले कामकाज हे युनिकोड बेस नसल्याने सदर कामकाज युनिकोडमध्ये रुपांतरीत करण्याबाबत यशवंतराव चव्हाण विकास प्रशासन प्रबोधिनी, (यशदा), पुणे यांच्या मार्फत दिनांक २५ मे २०१२ रोजी प्रशिक्षण वर्ग आयोजित करण्यात आला होता.

या अहवालामध्ये सदर प्रशिक्षणादरम्यान विषयांकित प्रकरणी देण्यात आलेली माहिती सादर करण्यात येत आहे.

Acknowledgement

I take this opportunity to express my gratitude to those whose active help and support make this report possible in the present form.

First of all, I express my sincere gratitude to **Er. Rajendra Pawar**, Chief Engineer, Special Projects, Water Resources, Amravati for proposing me for the present training programme.

It is the endless guidance and constant encouragement of **Er. R.V. Jaltar**e, Superintending Engineer, Akola Irrigation Circle, Akola and I would like to express my heartfelt gratitude to him for allowing me to participate in the said workshop.

I also express my gratitude to the instructor and staff of Yashwantrao CHavan Academy of Development Administration, (YASHADA) Pune particularly, **Smt. Ujwala Udgonkar** from YASHADA & **Shri. Surya Pahauri**, from CDAC for explaining in depth about the topic.

Er. Pravin Kolhe,

Executive Engineer,

Minor Irrigation Division No.-2, Akola

Contents

Contents

1. Introduction to Unicode

1.1.   What is Unicode?

The Unicode Standard is the universal character encoding standard used for representation of text for computer processing. Unicode provides a consistent way of encoding multilingual plain text and brings ease in exchange of text file internationally.

Due to different non-standardized encoding styles it was very much difficult for computer users who deal with multilingual text -- business people, linguists, researchers, scientists, and others to exchange the files What is Unicode.

The Unicode Standard is the universal character encoding standard used for representation of text for computer processing. Unicode provides a consistent way of encoding multilingual plain text and brings ease in exchange of text file internationally.

Due to different non-standardized encoding styles it was very much difficult for computer users who deal with multilingual text -- business people, linguists, researchers, scientists, and others to exchange the files internationally. Now they are finding that the Unicode Standard greatly simplified their work. Mathematicians and technicians, who regularly use mathematical symbols and other technical characters, will also find the Unicode Standard valuable. UTF-8 and UTF-16 are the standard encodings for Unicode text in HTML documents, with UTF-8 as the preferred and most used encoding.

In simple words, Unicode provides a unique number for every character, no matter what the platform, no matter what the program, no matter what the language.

1.2. Why Unicode

In the past there was ISCII encoding was standardized for Indian language data storage. This was a 8 bit character encoding defined for majority of Indian languages. Various Indian language software are based on this encoding. There are proprietary font formats that are designed considering this as a base. Due the non uniform , rather proprietary nature of the fonts the user is dependant certain software for creation of the Indian language contents. Now with added support for Indian languages the dependency is no more in existence. Hence it is now very easy to work with Indian languages on computer.

Now with the advent of technology it has become mandatory to keep your data Unicode format. It becomes very easy to view this data across various operating systems like Windows, Linux, various devices like mobiles, PDAs, STBs. So contents created in Indian language Unicode is accessible to all the available devices which have Indian language support. Ex. You have created a word file and you wish publish this file on a website, it is possible to push on to web without any effort. This is because the latest browsers support Unicode. Now you are free from the worry of font download or presence of font on the machine. The browser will take care of this. IE 5.5 and above, Mozilla firefox 2.0 are Unicode enabled browsers so they

automatically detect the Unicode and display the contents in a proper manner.
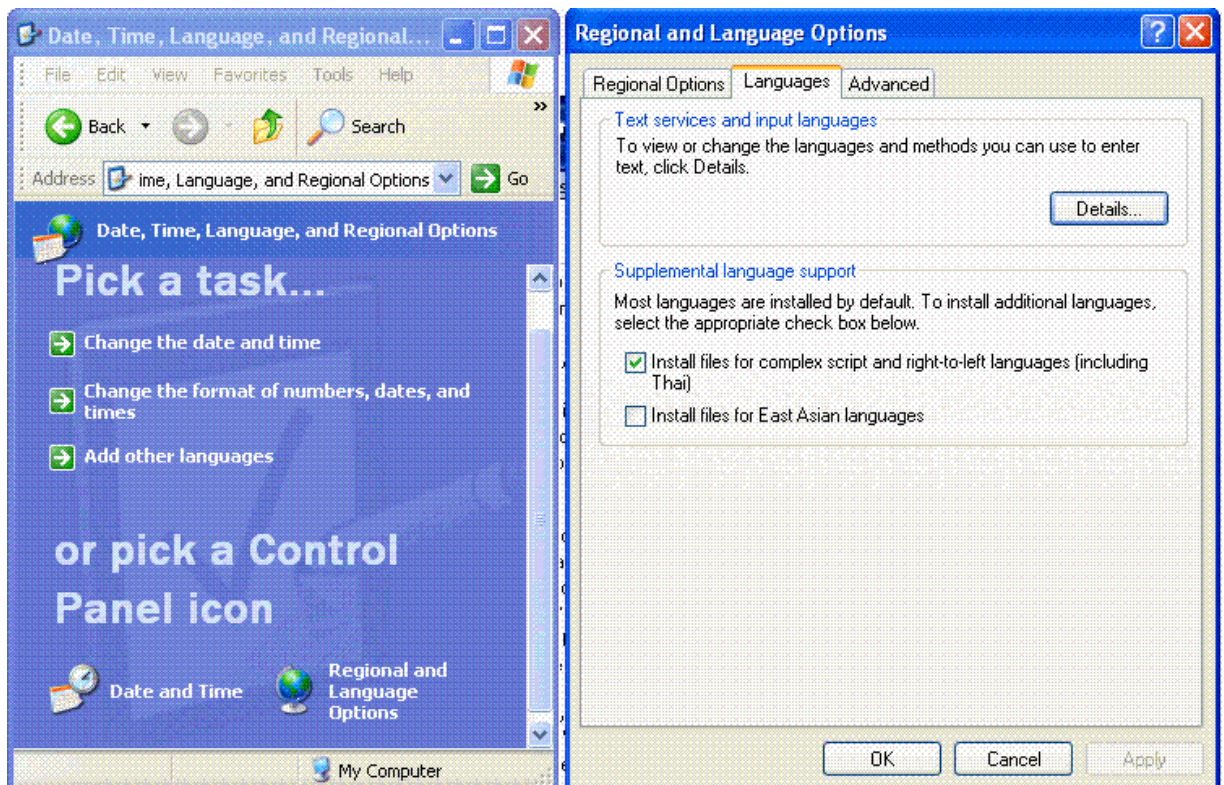
Unicode is-

• Universal Standard

• Independent of Operating system (Windows2000, XP and above)

• Supported by latest Microsoft Office and Open office suite

• Supported by popular browsers like IE, Mozilla, etc.

• Supported by web application development tools

• Supported by latest digital devices, mobile phones, STBs

1.3. How to enable Unicode in System For Windows XP

(Note: Windows XP CD may be required for the following operation):

• Click on "Start->Control Panel->(Date, Time, Language and Regional Options)->Regional and Language Options"

• Click on 'Languages' tab (next to 'Regional Options') at the top

• In 'Supplemental language support', see whether "Install files for Complex Script and right to left languages (including thai)" is kept ticked.

• If ticked, leave it as it is. If not ticked, tick it. It will start installing some required files and Windows XP CD may be required to complete the process.

You will have to restart the machine for the new setting to take effect.

1.4. How to enable Unicode in System For Windows 2000

(Note: Windows 2000 CD may be required for the following operation):

• Click on 'Start->Settings->Control Panel->Regional Options'

• You will see a list of languages under 'Language Settings for the system'. Scroll down the list. When you see 'Indic' see whether it is already ticked.

• If ticked, leave it as it is. If not ticked, tick it. It will start installing some required files and Windows 2000 CD will be required to complete the process.

• You might have to restart the machine for the new setting to take effect.

- For Windows Vista and Windows 7 the support for Indian languages is in-built so there is no setting required.

- The typical advantage of Unicode contents is universal acceptability. Then to migrate to Unicode one needs to convert the existing data to Unicode. There are converters available for conversion of Doc, txt, HTML files. So one can ensure the smooth migration from current data to Unicode.

So software package like ISM V6 can be suitable choice for migration from the older fonts to Unicode. It is bundled with all the converters, fonts and utilities required to work with Indian languages.

2. Introduction to ISM 6.0

2.1.　About ISM 6.0

ISM V6 is a software developed by the GIST group of Centre for Development of Advanced Computing, India. ISM is the acronym for Intelligent Script Manager. Using ISM is the easiest way to get started with Indian languages on your personal computer. It enables existing English applications to work for Indian Languages. User can almost immediately begin to type in Indian languages without having any previous experience in typing.

This software consists of various aesthetic Indian language fonts and tools that user require for working with Indian Languages on computers. This includes, a wide range of tools that assist the user in creating and editing Indian Language content, sorting Indian language data in Excel worksheets, various administrative and general dictionaries, transliteration and data migration tools. ISM can be also used for developing user own Indian Language Software.

ISM V6 has been developed to work with various Indian scripts. ISM enables users to create Indian language data in various Windows based applications. For viewing Indian language data the user needs to set the correct Indian language font in the application.

Hence viewing and editing Indian language data is possible only in applications with a provision to set different fonts. ISM facilitates the use of UNICODE as well as ISFOC fonts for display and printing, with Hardware and Software designed and developed primarily for English. Support for UNICODE has been provided for display (OT Fonts), typing, onscreen floating keyboard, MS-Office Macros, etc.
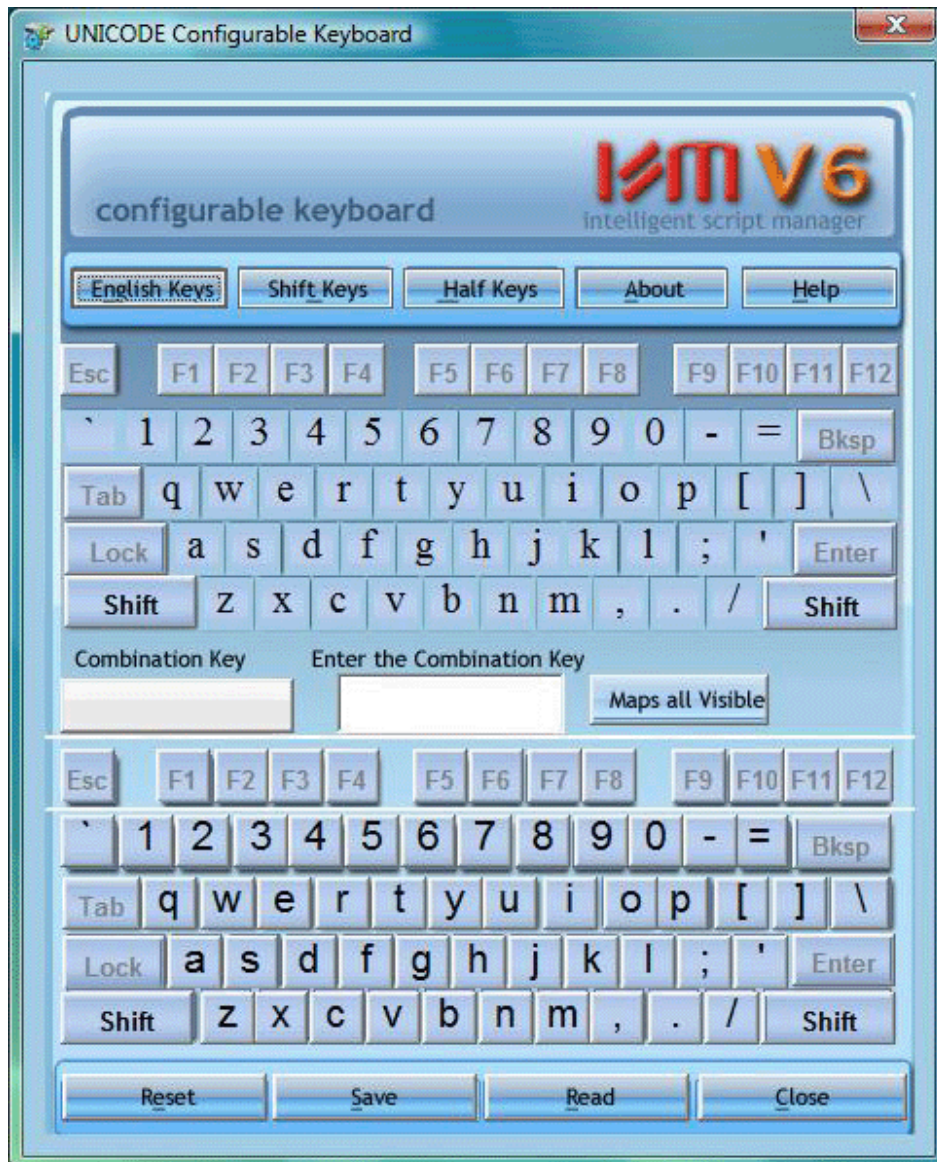
2.2. Salient Features of ISM V6

1) Supports UNICODE data and Open Type (OT) Fonts.

2) Enables creation of Indian language content using Windows based applications.

3) Enables creation of Indian language web content in applications like MS Front Page, Macromedia Dream Weaver.

4) Development of Indian Language applications using existing development platforms like Developer 2000, Visual Basic, Visual C++ etc.

5) Enabling support for Indian Languages in mailing, chatting and messaging applications like Outlook Express, MSN Messenger, ICQ Messenger, Yahoo Messenger etc.

6) On screen floating keyboard to help novice users in content creation and learning the keyboard.

7) Phonetic & Easy Phonetic typing support for users comfortable with the English keyboard.

8) Support for popular script specific keyboards layouts like Typewriter.

9) Facility to customize keyboard layout.

10) Insertion of frequently used Administrative / Banking words and phrases into documents with the use of official language dictionaries.

11) Insertion of date, time, and number in words into documents using Date & Time converter utility.

12) Facility to convert data from other data formats like ISCII, PC ISCII, 7 bit ISCII, Akruti, Akshar, Winkey, Indica, Shreelipi, are provided.

13) Word Macros for Spell Checkers, Find and Replace, Mail Merge, Official Language Dictionary, Synonym Dictionaries, Save as ISCII/UNICODE and Keyboard Shortcuts.

14) Excel Macros for sorting data.

15) Open Office Macros for Spell Checkers, Find and Replace and Mail Merge.

16) Easy to use with existing applications, without the learning curve of a new word processor.

17) A plethora of fonts, symbols & borders.

2.3. Working with ISM 6.0

2.4. Unicode Configuration Keyboard

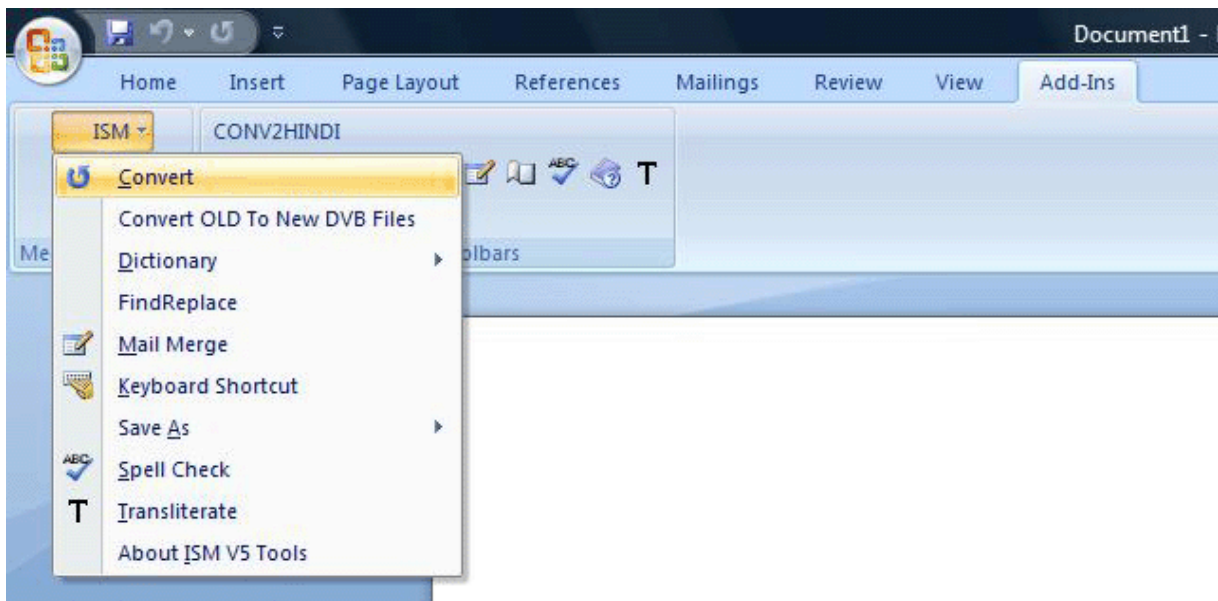## 3. Training on Unicode & ISM 6.0

3.1.   Training Schedule

Time	Topic	Speaker
09:00 to 09:30	Registration	Smt. Karuna Randhir
09:30 to 09:45	Introduction to Course	Smt. Ujwala Udgaonkar, Course Director
09:45 to 11:15	• Introduction to Unicode • Enable Unicode on Windows XP	Shri. Surya Pahauri, CDAC
11:15 to 11:30	Tea Break 1	
11:30 to 01:00	Use of ISM 6.0 & hands on	Shri. Surya Pahauri, CDAC
01:00 to 02:00	Lunch Break	
02:00 to 03:30	Conversion of ISM files to Unicode & Hands on	Shri. Surya Pahauri, CDAC
03:30 to 03:45	Tea Break 2	
03:45 to 04:30	• Using Unicode in mail • Google Transliterate • Hands on session	Shri. Surya Pahauri, CDAC
04:30 to 05:00	Feedback & Valedictory	Smt. Ujwala Udgaonkar, Course Director

4. Data Conversion from Various Form

4.1.   Conversion

1) Step 1:

Open the document you want to convert. Click on **Convert** in the **ISM** menu as shown below.
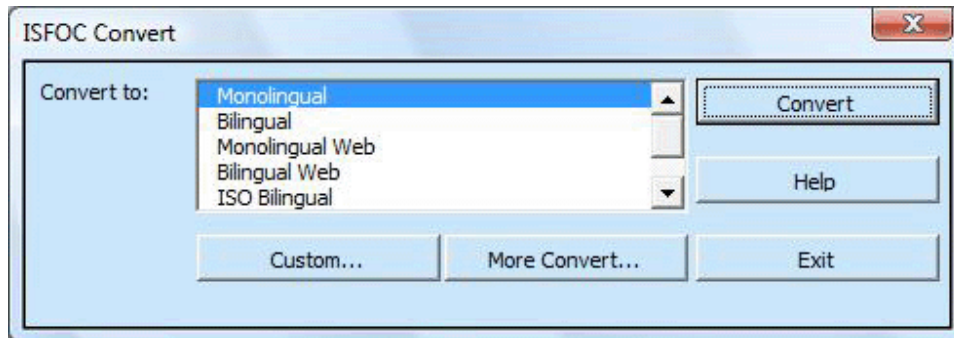

The Converter tool can also be accessed using the tool bar as shown below.

2) Step 2:

Click on **Custom** button on the Converter dialog to access the Custom Conversion options as shown in the screen shot of Step 3.
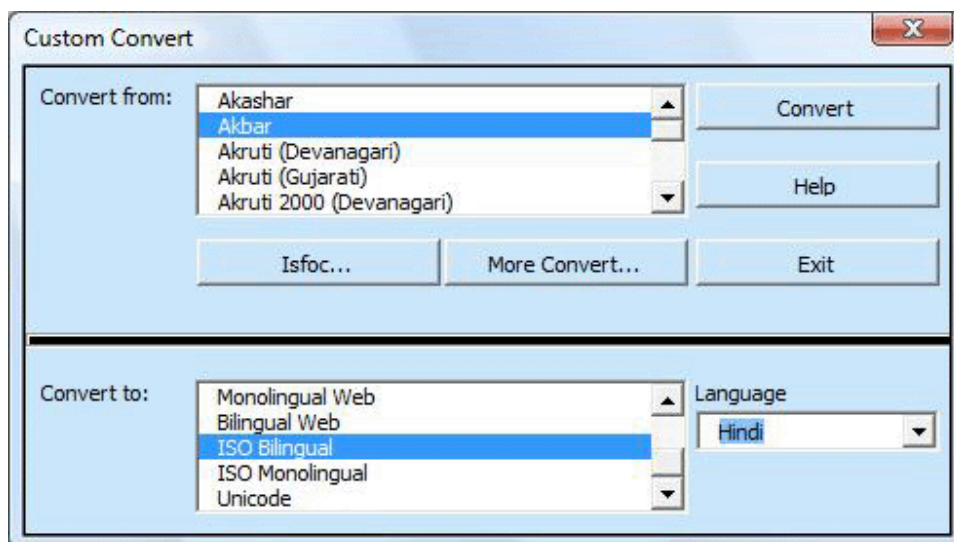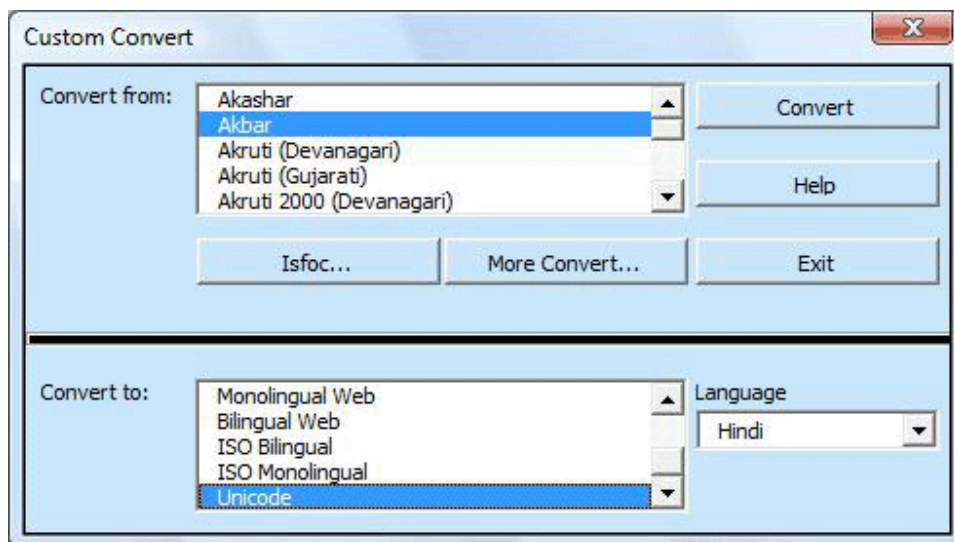
3) Step 3:

Specify the third party format of existing data in **Convert From**.
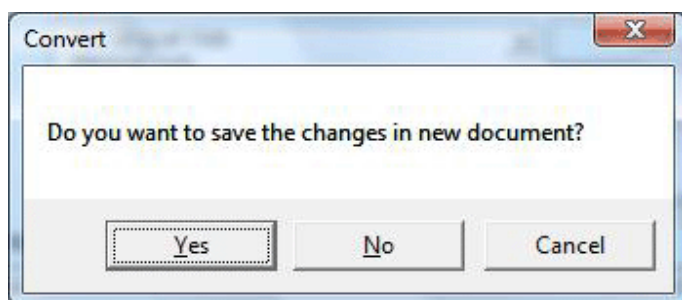
Specify the language in **Language** field.

Click the Convert **button.**

it also support third party to Unicode conversion…

The following message box will be displayed :

Click on **Yes** to save a copy of the original document before proceeding with conversion.

Click on **No** to start the conversion without making a copy.

Click on **Cancel** to stop the conversion.

The entire active document will gets converted by default. You can also convert part of a document by selecting the required portion before carrying out Step 1.

NOTE:

Conversion is done word by word. Thus a word with multiple font types is not a valid candidate for conversion.

Data in Windows Fonts such as Arial, Times New Roman etc, will get converted to garbage.

Data consisting of Indian language text interleaved with English will require the user to explicitly select the Indian language portions and do the conversion to prevent the Converter from converting the English data. If used otherwise the converter will proceed to convert the entire document, resulting in garbage converted output for the English data.

5. References

1) ISM 6.0 help file
2) http://www.cdac.in/html/gist/products/ismv6off.asp
3) www.cdac.in/gist
4) http://www.bis.org.in
5) "युनिकोडच्या सहाय्याने संगणकावर मराठी कसे वापरावे"-आशिष आल्मेडा व सुशांत देवळेकर