

# The xe-inputenc override package for XeTeX

Will Robertson

2008/05/15      v0.1

## I Introduction

When new users migrate from pdfTeX to XeTeX, some are often confused by the fact that the inputenc package is no longer needed; it does seem somewhat intuitive that the correct incantation should now be `\usepackage[utf8]{inputenc}`. Unfortunately, this is not the case due to differences between the way the 8-bit TeX and pdfTeX engines parse unicode files in comparison to XeTeX.

It seems useful, then, to provide an ‘override package’ for inputenc that interprets the wishes of the user without adding the code to break XeTeX’s attempted reading of the input file.

In this first hasty release of the package, any package options given to the override xe-inputenc are passed directly to XeTeX’s low-level command `\XeTeXinputencoding`. This command sets the encoding for subsequent reading of the current file. All other files are still interpreted as UTF8 else problems arise when reading .aux files (etc.), which are always *written* in the UTF8 encoding.

The best solution, possibly, would be to declare `\XeTeXinputencoding` explicitly to “utf8” for all XeTeX-written files, and use `\XeTeXdefaultencoding` to read all other input files in the encoding specified by the options to this package.

An exhaustive list of input encodings supported by XeTeX is shown here: <http://www.iana.org/assignments/character-sets>. Luckily, most

or all of the encoding names supported by the traditional `inputenc` share their names here. TODO: provide a complete and correct mapping from `inputenc`'s encoding options to XeTeX's.

## File I

# Implementation

This is the package implementation.

## Change History

VO.I

General: Initial version.

2

Do the right thing and indicate we've overridden L<sup>A</sup>T<sub>E</sub>X's `inputenc`:

```
1 \typeout{%
2   Loading a XeTeX-specific package to
3   override LaTeX's inputenc package.}
```

Just to be safe:

```
4 \RequirePackage{ifxetex}
5 \RequireXeTeX
```

Pass the package option to XeTeX's input encoding primitives:

```
6 \DeclareOption*{%
7   \PackageWarning{xe-inputenc}{^^J\space\space
8     XeTeX input encoding set to '\CurrentOption' for the rest of^^J%
9     \space\space file '\jobname'.^^J\space\space
10    All subsequent files to be read will be interpreted as UTF8.^^J
11    This message occurred}%
12 \XeTeXinputencoding "\CurrentOption"}
```

It might be better not to execute anything, but for now just do utf8 by default:

```
13 \ExecuteOptions{utf8}
14 \ProcessOptions
```

And that's it!

# Index

Numbers written in *italic* refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in *roman* refer to the code lines where the entry is used.

C		\ProcessOptions . . . . . I4	
\CurrentOption . . . . .	8, <i>12</i>	R	
D		\RequirePackage . . . . . 4	
\DeclareOption . . . . .	6	\RequireXeTeX . . . . . 5	
E		S	
\ExecuteOptions . . . . .	<i>13</i>	\space . . . . .	7, 9
J		T	
\jobname . . . . .	9	\typeout . . . . .	I
P		X	
\PackageWarning . . . . .	7	\XeTeXinputencoding . . . . .	<i>12</i>