

Kiril — Additional Cyrillic encodings

kiril version 1.05

Georgi N. Boshnakov

Copyright © 2006-2011 Georgi N. Boshnakov

Permission is granted to make and distribute verbatim copies of this manual provided the copyright notice and this permission notice are preserved on all copies.

Permission is granted to copy and distribute modified versions of this manual under the conditions for verbatim copying, provided that the entire resulting derived work is distributed under the terms of a permission notice identical to this one.

1 Introduction

Package ‘kiril’ provides input methods for the Bulgarian and Russian languages for use on top of UK or US layouts of the background keyboard.

For the Bulgarian language, a modernised keyboard layout based on A. Zinoviev’s proposal¹ is implemented. In practice, there is no difficulty in switching to this layout (from the standard BDS) since most of the keys are the same and the few that have changed actually make life easier.

For the Russian language, a standard layout is given (todo: give exact details).

The input encodings are prefixed with ‘`bulgarian-bg`’ (Bulgarian) and ‘`russian-ru`’ (Russian).

The ‘background’ layout is the layout used for Latin text. Currently the supported layouts are US English and British English. The special variants for the UK background layout ensure that all Cyrillic characters appear where expected when using the British keyboard. Also, on the US layout two of the keys generate the same codes and as a consequence the Cyrillic variants for the US background are slightly different from the British.

For users of LaTeX I give slightly changed layouts. The key holding ‘\’ and ‘|’ on the UK or US keyboards is left as is. This is because of the special use of backslash in TeX documents. The remaining changes are optional. In the TeX layouts the Capital cyrillic letters i-grave, soft-sign and eri, can be obtained by pressing twice in succession the key for the corresponding lowercase letter. This can be requested for the non-TeX version as well by setting a variable ‘`FIXME: not implemented yet`’.

I wrote this package for personal use in order to be able to enter conveniently Bulgarian text, often in conjunction with LaTeX. The standard keyboard layout for the Bulgarian language is known as BDS and is available as input method ‘`bulgarian-bds`’ in Emacs. I am working on Windows machines using normally the British keyboard. When switching to the BDS input method some of the letters annoyingly are not where they should be. Moreover, the BDS layout was designed for typewriters and contains features which are hardly needed with computers. For example, it includes several Latin letters for typing Roman numerals. Once I figured out how to set up an encoding, introducing additional encodings and modifications to the layouts did not involve much additional work.

I have tested this package on Windows only but, to my knowledge, there is nothing platform specific in it.

¹ Anton Zinoviev, Extended Bulgarian keyboard layouts, CoRR, abs/0905.0484, 2009, <http://arxiv.org/abs/0905.0484>

2 Installation

This package can be installed through the packaging system distributed with Emacs 24 and available as add-on for older versions of Emacs.

The package currently resides in my package archive at <http://www.maths.manchester.ac.uk/~gb/emacs/packages/>. To make the packages provided there known to the package system, add the above site to the list of package archives. This can be done, for example, by putting the following command in your `.emacs` file (and restarting Emacs).

```
(setq package-archives
      '(("gnu" . "http://elpa.gnu.org/packages/")
        ("marmalade" . "http://marmalade-repo.org/packages/")
        ("georgi" . "http://www.maths.manchester.ac.uk/~gb/emacs/packages/")))
```

Once this is done, you can run the command `'list-packages'` and choose package `'Kiril'`. The encodings provided by it should become available immediately after installation. The Info version of this documentation should also appear at the top level of the Info help.

If you have special settings, you may need to put the following line in your `' .emacs'` file.

```
(require 'kiril)
```

and restart Emacs.

The old method of installation will work as well - download the package from <http://www.maths.manchester.ac.uk/~gb/emacs/packages/> (it is a tar file) extract `'kiril.el'` and place it where Emacs looks for lisp files, e.g. in subdirectory `'site-lisp'` of the Emacs installation directory, put the above `'require'` command in your `' .emacs'` file and restart Emacs.

Here are some commands that you may wish to set in your `' .emacs'` file to activate some of the features.

3 Usage

Package *kiril* adds several input encodings to Emacs. They can be used as any other Emacs encodings. You can select any of them using the usual Emacs key bindings, commands or menu items. The quickest method is probably the `C-\` key (`toggle-input-method`). Alternatively, toggle or select the input method from the ‘Options->Mule’ menu.

Description of an encoding can be obtained with `C-h I` (`describe-input-method`) or the equivalent menu item in the ‘Help’ menu.

The initial part of the names of most input methods in Emacs indicate the designated language. Thus, input methods for Bulgarian language start with ‘`bulgarian-`’, for Russian with ‘`russian-`’, etc.

Input methods (encodings) introduced by package *Kiril* start with ‘`bulgarian-bg`’ and ‘`russian-ru`’, respectively. The next part of an encodings’ name designates the background layout, as follows.

`-for-us-` for underlying US layout

`-for-uk-` for underlying UK layout

Make sure that you choose the encoding corresponding to your background layout.

There are also modified variants of the above encodings suitable for work with \TeX files. Their names contain the string ‘`LaTeX`’.

3.1 Input methods

Here is the full list (at the time of this writing) of the input methods (encodings) provided by *kiril*. The documentation string of each input method shows a table of its layout.

If you encounter complications with the letters lowercase and uppercase Cyrillic *i*-grave, See [Section 3.4 \[Cyrillic i with grave accent\]](#), page 4.

‘`bulgarian-bg-for-uk-layout`’
Bulgarian layout, UK background

‘`bulgarian-bg-for-us-layout`’
Bulgarian layout, US background

‘`bulgarian-bg`’
equivalent to ‘`bulgarian-bg-for-uk-layout`’ if ‘`kiril-for-uk-only`’ is bound; equivalent to ‘`bulgarian-bg-for-us-layout`’ if ‘`kiril-for-us-only`’ is bound; undefined in all other cases.

‘`russian-ru-for-uk-layout`’
Russian layout, UK background

‘`russian-ru-for-us-layout`’
Russian layout, US background

‘`bulgarian-bgLaTeX-for-uk-layout`’
Bulgarian layout, LaTeX variant, UK background

‘`bulgarian-bgLaTeX-for-us-layout`’
Bulgarian layout, LaTeX variant, US background

3.2 Shorter names for input methods

If you prefer shorter names, put the following line in your `.emacs` file.

```
(setq kiril-for-uk-only t)
```

Then *kiril* will set up the encoding ‘`bulgarian-bg`’ to be equivalent to the UK layout, ‘`bulgarian-bg-for-uk-layout`’.

Similarly, if you set

```
(setq kiril-for-us-only t)
```

then the encoding ‘`bulgarian-bg`’ will be equivalent to the US layout, ‘`bulgarian-bg-for-us-layout`’.

The longer names for the input encodings are always available.

3.3 Opening and closing quotes

Computer typesetting makes it possible to use proper opening and closing quotes. There are slots for them in the encodings provided here. The neutral double quotes are also available at their usual keyboard place.

3.4 Cyrillic i with grave accent

The material in this section was written some years ago when Unicode support by operating systems and editors was not complete. The accented letters should work fine these days. Use combining accents only if the standard option does not work.

The grave accent is the only one needed in the Bulgarian language. It is used to show the position of the stress but is rarely needed. The only accented letter that occurs not so rarely is *Cyrillic-i-grave* (the Cyrillic *i* with grave accent) when it stands on its own (meaning ‘her’). Putting the accent helps the reader since the non-accented *Cyrillic-i* is the conjunction ‘and’ but omitting the accent is acceptable and, in fact, *Cyrillic-i-grave* was not available on typewriters. Some typists used to replace it with *Cyrillic-i-short* which many consider ugly.

Note that the standard Bulgarian keyboard layouts do not have *Cyrillic-i-grave* but the encodings in this package do.

Users of the \TeX typesetting system have always been able to put accents on letter using standard commands.

Unicode also provides a way to put accents on letters using ‘combining accents’. More recently, a slot for *Cyrillic-i-grave* was introduced in Unicode. Both, combining accents and the dedicated letter are supported by browsers and ‘pdf’ readers but the support by editors varies. The dedicated slots for *Cyrillic-i-grave* (lowercase and uppercase) may show as hollow boxes if the system fonts have not been updated.

Combining accents may appear as the letter followed by the accent (this was the case Emacs 22 and in MS word some years ago, I do not know what is the situation now). Emacs 23 handles combined accents more or less correctly, albeit not very pretty but I have observed some funny effects when moving point over letters accented via combining accents.

The Bulgarian encodings provided by package *kiril* use the dedicated Unicode slots for *Cyrillic-i-grave* but the package also provides commands to switch to the combining accents and, for the \TeX modes, to the corresponding \TeX commands.

3.4.1 Old discussion partly in Bulgarian

`\' , \ Auctex. , , . \n` (space , ,). `auctex \ .`
`\u045D \u040D` -. , , . ().
`\u0438\u0300` .. (,) 23. 22 "" (MS Word).
 "" utf8x .

3.4.2 Functions reference for Cyrillic-i-grave

`kiril-cyri-grave-via-combine-accent` [Variable]

If bound, use Unicode combining accents for ‘Cyrillic-i-grave’ letters, not the dedicated Unicode slots.

`kiril-cyri-grave-unicode` [Command]

Insert cyri-grave using the appropriate Unicode slot.

`kiril-cyrI-grave-unicode` [Command]

Insert cyrI-grave using the appropriate Unicode slot.

Unicode combination accent for cyri and cyrI.

`kiril-cyri-grave-combine` [Command]

Insert cyri-grave as Unicode combined accent.

`kiril-cyrI-grave-combine` [Command]

Insert cyrI-grave as Unicode combined accent.

`kiril-toggle-cyrIi-unicode ()` [Command]

Toggle the cyrillic letter cyrillic-i-grave, lower- and uppercase, between its Unicode slot and the Unicode combining accent. The need for this should be diminishing. However the glyphs for this letter are not always supported yet, so the generic combining accent may be a solution in some cases. It seems that this letter is absent from the Emacs fonts. At the time of writing (Sep 2009) the support for the combining accent is not ideal — on the screen the accent goes over the following letter and appears only when moving the point backwards over it (the letter is ok, if saved, other program show it).

Index

(Index is nonexistent)

Table of Contents

1	Introduction	1
2	Installation	2
3	Usage	3
3.1	Input methods.....	3
3.2	Shorter names for input methods.....	4
3.3	Opening and closing quotes.....	4
3.4	Cyrillic i with grave accent.....	4
3.4.1	Old discussion partly in Bulgarian.....	5
3.4.2	Functions reference for Cyrillic-i-grave.....	5
	Index	6