Belarusian−HOWTO

Alexander Mikhailian, mikhailian@altern.org

v.0.1.5, 25 February 2001

Short guide in setting up Belarusian language support in Linux console, X Window System, web−browsers, text editors, etc. Charsets used in the setup are either windows−1251 or iso−8859−5. Belarusian characters can also be found in koi8−ub, koi8−c, koi8−ru and, of course, Unicode. Although this HOWTO is Linux−specific, many advices are applicable to other UNIX−like systems.

Table of Contents

1. Introduction

1.1. History

This document was started in September 15, 1999 by Alexander Mikhailian

1.2. Comments

Comments on this HOWTO may be directed to the author mikhailian@altern.org.

1.3. New Versions

The newest version can always be found at the bellinux page.

1.4. Copyright

1.5. Acknowledgements and Thanks

Thanks to everyone who gave comments as I was writing this.

2. System−wide setup

2.1. Choosing Charset

As for now, the choice is mainly between windows−1251 and iso−8859−5. Windows−1251 provides compatibility with M$ Windows and it is is by far the most popular charset for nearly all slavic languages. On the other side, iso−8859−5 is better supported and easier to set up.

The present HOWTO will explain the use of both charsets in parallel. Unless otherwise stated in the text, Belarusian support package from the bellinux page contains all the files mentioned in the present HOWTO.

2.2. windows−1251 in the kernel

Windows−1251 support in the kernel is needed to visualize MS Windows filenames in cyrillic while the console and X Window system are localized in windows−1251. As latest FAT file systems store filenames in Unicode, we have to define the output charset of the Virtual File System layer and install the support for windows−1251 in the kernel.

There is a patch for 2.2.14 kernels that is included in the Belarusian support package. To apply the patch,

- ```
  cd /path_to_your_kernel_source_tree
  ```
- ```
  patch -p0 < patch_for_the_kernel.patch
  ```
- Recompile and reinstall the kernel modules. Refer to Kernel−HOWTO for the details.
- Add
  ```
  codepage=866,iocharset=microsoft-cp1251
  ```
  to your mount options in order to get the filenames in windows−1251.

  To test the setup, execute

  ```
  mount /dev/hda -t vfat -o codepage=866,iocharset=iso8859-5 /mnt/hda
  ```

as root. Of course, you should replace /dev/hda by the name of your vfat partitions containing names in cyrillic.

2.3. ISO−8859−5 in the kernel

There is built−in support for iso−8859−5 in Linux kernel. To enable iso−8859−5, make sure you compile this module in the nls section of the kernel setup. Normally, iso−8859−5 is compiled by default and can be loaded either automatically or by issuing

```
insmod nls_iso8859-5
```

as root.

Add

```
codepage=866,iocharset=iso8859−5
```

to your mount options in order to get the filenames in iso−8859−5.

2.4. Setting locale

- Belarusian locale is available in glibc 2.2 and later. If you do not have the belarusian locale installed in your system, you can compile it yourself. The source is available at the bellinux page. Issue
  ```
  localedef −f CP1251 −i be_BY be_BY.CP1251
  ```
  or
  ```
  localedef −f ISO−8899−5 −i be_BY be_BY.ISO−8899−5
  ```
  to compile the locale in windows−1251 encoding or iso8859−5 encoding.
- Check how it works by setting the locale
  ```
  set  LANG=be_BY.CP1251; export LANG
  ```
  or
  ```
  set  LANG=be_BY.ISO−8859−5; export LANG
  ```

and running a locale−aware program like `date` or `cal`.

2.5. Belarusian in console

2.5.1. Setting Belarusian with windows−1251

Some cyrillic console fonts are in fact russian fonts and lack many cyrillic characters. Fortunately, UniCyr fonts by Vadinm Zhitnikov have all the glyphs from cp866, cp1251, iso8859−5 and all printable symbols from koi8−r. Besides the belarusian keyboard maps which you can find at the the bellinux page , other files are fairly standard and are available in most linux distributions. To set up window−1251 in console,

- Load a UniCyr font
  ```
  consolechars −f UniCyr_8x16
  ```
- Load `by2_win.kmap` keymap.
  ```
  loadkeys by2_win
  ```
- Load Application−Charset Map `cp1251.acm`
  ```
  consolechars −m cp1251
  ```

2.5.2. Setting Belarusian with iso−8859−5

There are two ways to set up Belarusian with iso−8859−5

* Load iso−8859−5 font

  ```
  consolechars −f iso05.f16
  ```
* Load belarusian keymap `by2.kmap` or `by.kmap`

  ```
  loadkeys  by.kmap
  loadkeys by2.kmap
  ```

This method has one serious drawback − you will loose all pseudographic characters and, say, you Midnight Commander will look somewhat naked. The second method described below preserves all pseudographic characters:

* Load UniCyr font

  ```
  consolechars −f UniCyr_8x16
  ```
* Load `by2.kmap` keymap or `by.kmap`

  ```
  loadkeys  by.kmap
  loadkeys by2.kmap
  ```
* Load Application−Charset Map `iso05.acm`

  ```
  consolechars −m iso05
  ```

Beware that different Linux distributions have different console−related packages − `console−tools` or `kbd`. Abovementioned scripts are meant to work with `console−tools` which is by far more popular.

2.6. Belarusian in X Window System

2.6.1. ISO−8859−5 in X Window System through XKB

This is is quite easy to set up.

* Install ISO−8859−5 fonts for X Window and make sure they are before other fonts in the fontpath. Refer to Font−HOWTO for details on how to install fonts.
* Get the belarusian keyboard layout `by` from the the bellinux page and put it on the place of any other keyborad layout in `/usr/X11R6/lib/X11/xkb/symbols/`, e.g. instead of the belgian keyboard layout `/usr/X11R6/lib/X11/xkb/symbols/be`

  Put the following stings in your `/etc/X11/XF86COnfig`:

  ```
  XkbModel   "microsoft"
  XkbLayout  "be"
  XkbOptions "grp:caps_toggle"
  ```
* Don't forget to set up the system locale to be_BY.ISO8859−5 as described in Section 2.4. You will also want to make sure that iso−8859−5 fonts are before any other fonts in your font path.

2.6.2. Windows−1251 in X Window System through XKB

This works only with XFree 4.0.2 and higher.

- Install ISO−8859−5 fonts for X Window and make sure they are before other fonts in the fontpath.
- Get the belarusian keyboard layout `by` from the the bellinux page and put it on the place of any other keyborad layout in `/usr/X11R6/lib/X11/xkb/symbols/`, e.g. instead of the belgian keyboard layout `/usr/X11R6/lib/X11/xkb/symbols/be`
- Put the following strings in your `/etc/X11/XF86Config`:

  ```
  XkbModel   "microsoft"
  XkbLayout  "be"
  XkbOptions "grp:caps_toggle"
  ```
- Don't forget to set up the system locale to be_BY.CP1251 as described in Section 2.4.

To use windows−1251 with XFree 3.3.5 and 4.0, you have to apply a patch from Aleksey Novodvorsky that allows the use of windows−1251 with XKB. The original location is at ftp.logic.ru/pub/logic/linux/be−locale and it is also available from the bellinux page. Lucky users of Linux−Mandrake RE get a patched XFree86 out of box.

Belarusian keyboard layout is palnned for all XFree releases after 4.0.2. To make it work, you will have to add the following lines into `/etc/X11/XF86Config`:

```
XkbModel   "microsoft"
XkbLayout  "by"
XkbOptions "grp:caps_toggle"
```

2.6.3. Windows−1251 in X Window System through Xmodmap

- Uncomment

  ```
  XkbDisable
  ```

  line in your `/etc/X11/XF86COnfig`.
- Install windows−1251 fonts for X Windows and make sure thay are before other fonts in the path.
- Make and install xruskb package which can be downloaded from http://bellinux.sourceforge.net
- Replace .xmm files in your xruskb directory by those found at the bellinux page.
- Add the following lines in your .Xdefaults file

  ```
  xrus*modeButton1.labelString: BEL xrus*modeButton1.label: BEL
  ```
- Run

  ```
  xrus jcuken-cp1251
  ```

  or

  ```
  xrus jcuken-iso5
  ```

to start keyboard switcher.

---

3. Editing texts

3.1. Emacs

Emacs is able to display Belarusian characters out of box in many Linux distributions. However, the following conditions should satisfy:

- Cyrillic iso−8859−5 fonts should be installed.
- Emacs−mule package should be installed.

The user may want to change the default keyboard layout ("input method" in emacs slang) to jcuken. This feature is provided by `belarusian.el` along with some other goodies. Read the comments in the beginning of `belarusian.el` for details on the setup procedure.

After installing `belarusian.el`, you will be able to switch the input method, by typing **C−\−cyrillic−belarusian−RET**

For a more user−friendly setup, put the following lines in the .emacs file:

```
;; enables PC−like key mappings
(pc−bindings−mode)
;; enables PC−like selection
(pc−selection−mode)
;; highlight selection
(transient−mark−mode t)
;; use as many multilingual fonts as possible
(set−default−font "fontset−standard")
;; set up windows−1251 as default encoding everywhere
(codepage−setup 1251)
(set−terminal−coding−system 'cp1251)
(set−keyboard−coding−system nil)
(setq default−buffer−file−coding−system 'cp1251)
(set−selection−coding−system 'cp1251)
(prefer−coding−system 'cp1251)
(define−coding−system−alias 'windows−1251 'cp1251)
(define−coding−system−alias 'microsoft−1251 'cp1251)
(define−coding−system−alias 'microsoft−cp1251 'cp1251)
(define−coding−system−alias 'windows−cp1251 'cp1251)
(define−coding−system−alias 'koi8−u 'cyrillic−koi8)
(define−coding−system−alias 'KOI8−R 'cyrillic−koi8)
(modify−coding−system−alist 'file "\.*" 'cp1251)
(modify−coding−system−alist 'file "\\.el$" 'iso−2022−7bit)
```

This configuration is for windows−1251 setup.

3.2. Spell−checking

- If you do not have ispell 3.1.20, install it from any Linux distribution CD.
- Download the package belspell.tar.gz. Unpack it and run

```
buildhash belarusian.sml belarusian.aff belarusian.hash
```

- Copy belarusian.hash to </usr/lib/ispell or wherever your ispell hash tables are.
- in order to spell-check a file, type

  ```
  ispell -d belarusian yourfile.txt
  ```

Getting belarusian ispell dictionary working with emacs is a bit tricky.

- Copy ispell.el and ispell.elc into your $EMACSBIN/../lisp directory.
- Copy temporarily $EMACSBIN/../lisp/loaddefs.el to /usr/src/emacs-XX.X/lisp/loaddefs.el where XX.X stand for emacs version number.
- **M-x-update-file-autoloads-RET** and enter the path to the ispell.el.
- Copy back loaddefs.el
- Restart emacs

3.3. TeX

from Aleksey Novodvorsky:

You need TeX + babel + T2, e.g. teTeX >= 1.0. in order to get partial support for Belarusian,

- put babel.sty , belarusianb.ldf and belarusianb.sty into /usr/share/texmf/tex/generic/babel
- run texhash
- To use Belarusian in LaTeX: add the following lines in the preamble:

  ```
  \documentclass[belarusian]{article} \usepackage[cp1251]{inputenc}
  \usepackage{babel}
  ```

  or

  ```
  \documentclass[belarusian]{article} \usepackage[iso88595]{inputenc}
  \usepackage{babel}
  ```

---

4. Browsers

4.1. Netscape

It is often a problem to correctly visualize Belarusian−specific characters with Netscape. This is due to that Netscape 4.x looks for *any* koi8−r font and, if it finds one, they refuse to use other fonts for displaying Cyrillic pages.

The problem is solved by removing koi8−r fonts from everywhere − X Window System font path and font server path.Then install koi8−ru, iso−8859−5 or windows−1251 and run

```
xset fp rehash
```

or restart X server. You will probably have to delete `>preferences` and `preferences.js` from the Netscape home directory `.netscape` and even then you are not guaranteed to have Netscape work right. Netscape's handling of fonts has always been an obscure issue.

Links to koi8−ru, iso−8859−5 and windows−1251 cyrillic fonts for X Window System can be found at the bellinux page

---

4.2. Lynx

In order to view Belarusian sites, you have to set up you console to handle the encoding you need. See section Section 2.5 for more details.

Add the following lines to your `.linxrc` file.

```
character_set=Cyrillic (windows−1251) preferred_language=be
preferred_charset=windows−1251
```

or

```
character_set=Cyrillic (ISO−8859−5) preferred_language=be
preferred_charset=iso−8859−5
```

if you have set up iso−8859−5 in console.

If the page you are browsing does not have an explicit charset declaration, press **o** and set the document charset manually.

---

5. Mailers

5.1. Netscape Messenger

Netscape Messenger can not handle belarusian texts properly. The same applies to Mozilla.

5.2. Mutt

Mutt handles a multitude of charsets and encodings with ease. If your console has windows−1251 support, add the following lines to your `.muttrc`:

```
charset-hook "windows-1251" "cp1251" set charset="windows-1251"
```

Mutt automatically converts all incoming messages for the screen output. It can also convert outgoing messages basing itself on a versatile pattern matching mechanism, e.g. the following line in `.muttrc` will force mutt to convert all messages sent to be−locale@iatp.unibel.by to *koi8−r*.

```
send-hook '~t ^be-locale@iatp\.unibel\.by$' 'set
send_charset="koi8-r"'
```

6. Further support

A mailing list devoted entirely to Belarusian language support is available at be−locale@iatp.unibel.by. To subscribe, send a message to majordomo@iatp.unibel.by with the string "subscribe be−locale youremailaddress" in the body of the message

Another mailing list that treats mostly linguistic issues is available at movaznaustva@egroups.com. To subscribe to it, send a message with empty body to movaznaustva−subscribe@egroups.com.