

Perso-Arabic Input Methods

And Making More Emacs Applications BIDI Aware

Mohsen Banan — محسن بنان

Email: emacs@mohsen.1.banan.byname.net

<http://www.by-star.net/PLPC/180063>

November 02, 2021

Contours Of This Presentation

A quick overview of:

- 1 Perso-Arabic Scripts
- 2 Persian Input Methods
 - farsi-isiri-9147 — farsi-transliterate-banan
- 3 Making More Emacs Applications BIDI (Bidirectional) Aware
 - Problems and Challenges For Emacs Developers
- 4 Persian Blee (By* Libre-Halaal Emacs Environment)
 - Towards a complete convivial Persian digital environment
 - Direction statements and preview of coming attractions
- 5 Pointers and References

Contours Of This Presentation

And Intended Audience

| **Perso-Arabic Writer** | **Non**

Emacs User:

Emacs Developer:

Considering To Convert To Emacs:

Persian Blee User:

Shaping And Bidirectionality

Left-To-Right:
Right-To-Left:
Bidirectional:
Top-To-Bottom:

| **Shaped Alphabet** | **Un-Shaped Alphabet** | **M**

Emacs: A Truly Multilingual Capable

Editor And Environment

Emacs Fully Supports Perso-Arabic Scripts:

- Unicode: Since 1990s
- Many Quail based input methods: Since 1990s
- BIDI (Bidirectional): Since 2012 (v24)
- Shaping: Since 2000s. With Harfbuzz since v27
- Various Persian Input Methods: Since 2012 (v24)

Significance Of Emacs Support For Perso-Arabic

- By Perso-Arabic script we mean Arabic script with various extensions used for writing Arabic languages, Persian families of languages and several other languages.
- Languages that use Perso-Arabic script as their writing system include: Arabic, Farsi, Dari, Urdu, Afghan, Pashto, Kurdish, Balochi, Lurish, Kashmiri.
- Perso-Arabic is the *second* most widely used writing system in the world by the *number of countries*.
- It is the *third* by the *number of users*, after the Latin and Chinese scripts.

About Emacs Input Methods

Input Methods allow you to enter characters that are not supported by your keyboard. With Quail maps we can map ASCII key strings to multilingual characters. So, we can input any text from an ASCII keyboard.

Main facilities are:

- `M-x set-input-method`: Select input method — `C-x C-m C-\`
- `M-x describe-input-method`: `C-h C-\` or `C-h I`
- `M-x toggle-input-method`: `C-\`

Emacs Persian Input Methods

Emacs comes with two built-in Persian input methods:

[farsi-isiri-9147](#): A Persian keyboard based on the Islamic Republic of Iran's ISIRI-9147 specification.

This is the traditional one-to-one mapping of keys on a computer keyboard to Persian letters.

[farsi-transliterate-banan](#): An intuitive transliteration keyboard for Farsi.

This is a more powerful method which converts sequences of characters into one letter. For example “kh” becomes خ.

Emacs Built In Documentation

Emacs is a self-documenting editor. Input methods' keyboard layouts can be displayed with `describe-input-method` and BIDI is well documented in Emacs Manual.

- `farsi-isiri-9147::` (`describe-input-method 'farsi-isiri-9147`)
- `farsi-transliterate-banan::` (`describe-input-method 'farsi-transliterate-banan`)
- Emacs Manual: Bidirectional Editing::
https://www.gnu.org/software/emacs/manual/html_node/emacs/Bidirectional-Editing.html
- Emacs Manual: Bidirectional Display::

Pointers To Code

The code for the quail keyboard mappings are at:

- Part of Emacs Installation At:

`/usr/share/emacs/27.1/lisp/leim/quail/persian.el.gz`

- As a Git Repo At:

<https://github.com/bx-blee/persian-input-method>

Complete Documentation

Complete Documentation For Persian Input Methods Is PLPC-120036:

<http://mohsen.1.banan.byname.net/PLPC/120036>

Persian Input Methods

For Emacs And More Broadly Speaking

شیوه‌های درج به فارسی

Various Related Information Is Also Available At PersoArabic.org:

<http://www.persoarabic.org>

Ramification Of BIDI And Perso-Arabic On Apps

BIDI And Perso-Arabic related glitches (or more than glitches) in various Emacs Apps:

- Gnus
 - For perso-arabic, Gnus columns don't line up in the Summary mode.
 - For perso-arabic, Subject and from fields should respect direction in Summary mode.
- bbdb
 - Paragraph and field directionality is not respected.
- calendar and calfw
 - For calendar, print persian and print islamic, can now produce perso-arabic letters.
(Starting point in place)
 - For calfw, perso-arabic entries don't line up and don't respect direction.
- AUCTeX, XeLaTeX

Making More Emacs Applications

BIDI And Perso-Arabic Aware

- BIDI Aware Topics

- Consider use of explicit specificatin of directioanlity at buffer, paragraph and field level.

For example, bbdb can easily be fixed to allow for right-to-left fields.

- Sometimes Combine left-to-right and right-to-left to convey more information.

For example, In calendar applications, Christian and Islamic dates can face each other on the same line.

- Use explicit html direction specification in email and other communications.

While emacs detects paragraph directionality properly, Firefox and Chrome don't.

- Perso-Arabic Topics

- Shaping makes fixed width character counting impractical. Count pixels not letters.

Group columns don't line up in the Summary mode. Subject and from fields should

Emacs Native Markup Language - ENML:

A Model For Apps Development

Let's put support for BIDI in an apps dev framework that all Emacs apps can use.

- Let's create "Emacs Native Markup Language – ENML" as a lispish super-set of html5.
- Let's mimic the web apps development frameworks in Emacs.
- Let's consider ENML as the primary Native Emacs Mode.
- Let's use ENML in all basic emacs buffers. Help buffers, doc-strings, etc.
- Let's merge ENML and org-mode.
- Let's make ENML BIDI aware.
- Let's transition all emacs apps to use ENML

About ByStar And BISOS

<http://www.by-star.net>

The Libre-Halaal By* (ByStar) Digital Ecosystem
For Preservation Of The Individual's Autonomy and Privacy
A Moral Alternative To The Proprietary American Digital Ecosystem

<http://www.by-star.net/PLPC/180054>

About Blee And Persian Blee

Blee: ByStar Libre-Halaal Emacs Environment is full integration of:

- Emacs +
- Lots Of Emacs Apps
- Lots Of Blee Apps
- Lots Of Debian Facilities
- Lots Of ByStar Services
- Lots Of BISOS (ByStar Internet Services OS) Facilities

Persian Blee: Blee for Persian Speakers