
Python Myanmar Documentation

Release 1.9.0

Thura Hlaing

Nov 24, 2020

Contents:

1	Installation	3
1.1	Installation	3
1.2	Syllabification	4
1.3	Encodings	4
1.4	Transliteration	5
1.5	Phonenumbers	5
1.6	Myanmar NRC	6
1.7	Credits	6
2	Indices and tables	7
	Python Module Index	9
	Index	11

This package contains a variety of python modules for Myanmar text proccesing, such as *syllabification*, *romanization*, *encoding conversion*, *nrc validation* etc. Only python3 is currently supported at the moment.

CHAPTER 1

Installation

The package is distributed on PyPI and can be installed with pip:

```
pip install python-myanmar
```

For more information, please read the full documentation [here](#).

1.1 Installation

1.1.1 Stable release

To install Python Myanmar, run this command in your terminal:

```
$ pip install python-myanmar
```

This is the preferred method to install Python Myanmar, as it will always install the most recent stable release.

If you don't have pip installed, this [Python installation guide](#) can guide you through the process.

1.1.2 From sources

The sources for Python Myanmar can be downloaded from the [Github repo](#).

You can either clone the public repository:

```
$ git clone git://github.com/trhura/python-myanmar
```

Or download the [tarball](#):

```
$ curl -OL https://github.com/trhura/python-myanmar/tarball/master
```

Once you have a copy of the source, you can install it with:

```
$ python setup.py install
```

1.2 Syllabification

Morphological and phonetic syllable break for Burmese text. Syllable break with zawgyi text will not be accurate / reliable. You probably should convert it to unicode before processing.

`myanmar.language.MorphoSyllableBreak` (*text, encoding*)

Return an iterable of morphological / visual syllables in text.

```
>>> from myanmar.encodings import UnicodeEncoding
>>> slb = list(MorphoSyllableBreak("", UnicodeEncoding()))
>>> list(s['syllable'] for s in slb)
['', '', '']
>>> slb[2]
{'syllable': '', 'consonant': '', 'eVowel': '', 'visarga': ''}
```

`myanmar.language.PhonemicSyllableBreak` (*text, encoding*)

Return an iterable of phonemic syllables in text.

```
>>> from myanmar.encodings import UnicodeEncoding
>>> slb = list(PhonemicSyllableBreak("", UnicodeEncoding()))
>>> list(s['syllable'] for s in slb)
['', '', '']
>>> slb[0]
{'syllable': '', 'consonant': '', 'iVowel': '', 'visarga': ''}
```

1.3 Encodings

Convert text in various Myanmar encodings. It currently supports **wininnwa**, **zawgyi**, **unicode**. Performance-wise, it is not as good as other regex-based converters.

`myanmar.converter.convert` (*text, fromenc, toenc*)

Convert text in *fromenc* encoding to *toenc* encoding.

```
>>> convert('', 'unicode', 'zawgyi')
''
>>> convert('', 'zawgyi', 'unicode')
''
>>> convert('&[ef;', 'wininnwa', 'unicode')
''
```

`myanmar.converter.get_supported_encodings` ()

Get a list of encodings supported by `converter` module.

```
>>> get_supported_encodings()
['unicode', 'zawgyi', 'wininnwa']
```


1.4 Transliteration

Transliterate Burmese text with latin characters. Currently, romanization based on BGN_PCGN, MLCTS, IPA systems are available.

`myanmar.romanizer.romanize(string, system)`

Transliterate Burmese text with latin letters.

```
>>> romanize("", IPA)
'kpjùtà'
>>> romanize("", MLC)
'padezaraja'
>>> romanize("", BGN_PCGN)
'bin-gala-aw'
```

1.5 Phonenumbers

Validation and normalization for Myanmar phonenumbers. Based on `mm_phonenumber` module from Melomap.

class `myanmar.phonenumber.Operator`

An enumeration.

`myanmar.phonenumber.get_landline_operator(phonenumbers)`

Get operator type for a given landline number.

```
>>> get_landline_operator('+95674601234')
'MyanmarAPN'
>>> get_landline_operator('9524261234')
'MyanmarSpeedNet'
>>> get_landline_operator('14681234')
'VoIPMyanmarGroup'
```

`myanmar.phonenumber.get_phone_operator(phonenumbers)`

Get operator type for a given phonenumbers.

```
>>> get_phone_operator('+959262624625')
<Operator.Mpt: 'MPT'>
>>> get_phone_operator('09970000234')
<Operator.Ooredoo: 'Ooredoo'>
>>> get_phone_operator('123456789')
<Operator.Unknown: 'Unknown'>
```

`myanmar.phonenumber.is_valid_phonenumber(phonenumbers)`

Checks whether a given phonenumbers is a valid Myanmar number or not.

```
>>> is_valid_phonenumber('09420028187')
True
>>> is_valid_phonenumber('+959420028187')
True
>>> is_valid_phonenumber(9420028187)
False
>>> is_valid_phonenumber(94200281870)
False
```

`myanmar.phonenumber.normalize_phonenumber(phonenumbers)`

Normalize a given phonenumbers into 959xxx number format.

```
>>> normalize_phonenumber('09420028187')
959420028187
>>> normalize_phonenumber('+959420028187')
959420028187
>>> normalize_phonenumber('420028187')
959420028187
```

1.6 Myanmar NRC

Validation and normalization for Myanmar NRC number.

`myanmar.nrc.is_valid_nrc(nrc)`

Check whether the given string is valid Myanmar national registration ID or not

```
>>> is_valid_nrc('12/LMN (N) 144144')
True
>>> is_valid_nrc('5/PMN (N) 123456')
False
```

`myanmar.nrc.normalize_nrc(nrc)`

Check the given string is valid myanmar nrc or not and normalize the string to simplest form if the string is valid

```
>>> normalize_nrc('9/pmn(n) 123456')
'9 pamana n 123456'
>>> normalize_nrc('1/bkn(n) 123456')
'1 bakana n 123456'
```

1.7 Credits

1.7.1 Development Lead

- Thura Hlaing <trhura at gmail.com>

1.7.2 Contributors

- Set Kyar Wa Lar <setkyar16 at gmail.com>
- Aye Chan Mon <polestar.mon20 at gmail.com>
- Soe Zayar <soezayar019@gmail.com>
- Kyaw Naing Win <kyawnaingwinknw@gmail.com>

CHAPTER 2

Indices and tables

- `genindex`
- `modindex`
- `search`

m

- `myanmar.converter`, 4
- `myanmar.language`, 4
- `myanmar.nrc`, 6
- `myanmar.phonenumber`, 5
- `myanmar.romanizer`, 5

C

`convert()` (in module *myanmar.converter*), 4

G

`get_landline_operator()` (in module *myanmar.phonenumber*), 5

`get_phone_operator()` (in module *myanmar.phonenumber*), 5

`get_supported_encodings()` (in module *myanmar.converter*), 4

I

`is_valid_nrc()` (in module *myanmar.nrc*), 6

`is_valid_phonenumber()` (in module *myanmar.phonenumber*), 5

M

`MorphoSyllableBreak()` (in module *myanmar.language*), 4

`myanmar.converter` (module), 4

`myanmar.language` (module), 4

`myanmar.nrc` (module), 6

`myanmar.phonenumber` (module), 5

`myanmar.romanizer` (module), 5

N

`normalize_nrc()` (in module *myanmar.nrc*), 6

`normalize_phonenumber()` (in module *myanmar.phonenumber*), 5

O

`Operator` (class in *myanmar.phonenumber*), 5

P

`PhonemicSyllableBreak()` (in module *myanmar.language*), 4

R

`romanize()` (in module *myanmar.romanizer*), 5