Supplementary Data Accompanying the Paper "Challenges of Annotation and Analysis in Computer-Assisted Language Comparison: A Case Study on Burmish Languages"

This is the supplementary material for the paper. It contains the webapplication that allows you to view the bipartite network, as well as the data and the Python code to create the newtork.

Web-Application

Just unpack the zip folder and open the file `index.html` in a web-browser.

Python Code

Just run the scripts (using Python3, make sure you have LingPy version 2.6 installed):

```
$ python3 c_bipartites.py
```

Will re-create the file `o_bipartite.gml`

```
$ python3 c_compare.py
```

Will print the comparison with STEDT which we mention in the paper to the screen and create additional files which are needed for the comparison (all prefixed with a `t` for "temporary").

Data

Data is given in the following files:

- concepts.tsv: the concepts linked to Concepticon
- languages.tsv: the languages linked to Glottolog
- d_bed.tsv: the BED data which was used for the study
- d_stedt.tsv: the original STEDT data for the Burmish languages which we extracted for this purpose

Data in CLDF (Forkel and List 2017)

Following the [specifications](#) of the [CLDF initiative](#), we provide the data in CLDF format as well. You find the data in the `cldf` folder. The script we used to convert the data in CLDF-format is `c_cldf.py`, and you can run it by writing:

```
$ python3 convert.py
```

References

- Robert Forkel, & Johann-Mattis List. (2017, July 27). glottobank/cldf: First release candidate for CLDF 1.0. Zenodo. http://doi.org/10.5281/zenodo.835502