

# Move CLDR collation files from UCA to CLDR

Author: Markus Scherer

Date: 2012-oct-11

## Proposal

Move the CLDR root collation files from the UCA release folder to the CLDR repository.

Specifically:

1. Beginning with a future UCA release, omit the files `CollationAuxiliary.html` and `CollationAuxiliary.zip` from the `http://www.unicode.org/Public/UCA/<version>/` folder.
2. Merge the contents of `CollationAuxiliary.html` into LDML (UTS #35), with appropriate modifications.
3. Modify the references between UCA (UTS #10) and the merged LDML (UTS #35) as appropriate.
4. Move `CollationAuxiliary.zip` (or a subset of its contents) into the CLDR repository, for release as part of CLDR. Details to be discussed in the CLDR TC. Strawman:
  - a. Add `allkeys_CLDR.txt`, `CollationTest_CLDR_NON_IGNOREABLE.txt`, `CollationTest_CLDR_SHIFTED.txt`, `FractionalUCA.txt` and `UCA_Rules.txt` into an appropriate location in the CLDR repository. (Possibly with modified file names.)
  - b. The other files are redundant (representing the same data except with fewer or no comments, or using different syntax) and should be omitted.
5. Add a note to the UCA `ReadMe.txt` about the new location of the CLDR root collation files.
6. Adjust the UCA beta/release process and its documentation appropriately.
7. Adjust the CLDR beta/release process and its documentation appropriately.

## Rationale

Starting with UCA 6.0, the UCA release has included the CLDR root collation files. CLDR root is a minor tailoring of the DUCET, and the CLDR root data files are normally generated at the same time and from the same tools as the DUCET `CollationTest` files.

However, the CLDR files are unnecessary for the specification of the DUCET, the release timing is wrong for the CLDR files, and it is unclear where to track bugs and tasks for the CLDR root collation that do not affect UCA+DUCET. It would be much cleaner to `publish/release/version/track` the CLDR root collation files as part of CLDR.

With the proposed move, all of the files would still be generated and tested during the UCA development and beta period as before.

## List of UCA 6.2 files

<http://www.unicode.org/Public/UCA/6.2.0/>

- allkeys.txt
- **CollationAuxiliary.html** (see below)
- **CollationAuxiliary.zip**
- CollationTest.html
- CollationTest.zip (see below)
- decomp.txt
- ReadMe.txt

<http://www.unicode.org/Public/UCA/6.2.0/CollationAuxiliary.zip>:

- allkeys\_CLDR.txt
- CollationTest\_CLDR\_NON\_IGNOREABLE\_SHORT.txt
- CollationTest\_CLDR\_NON\_IGNOREABLE.txt
- CollationTest\_CLDR\_SHIFTED\_SHORT.txt
- CollationTest\_CLDR\_SHIFTED.txt
- FractionalUCA\_SHORT.txt
- FractionalUCA\_summary.txt
- FractionalUCA.txt
- UCA\_Rules\_NoCE.txt
- UCA\_Rules\_NoCE.xml
- UCA\_Rules\_SHORT.txt
- UCA\_Rules\_SHORT.xml
- UCA\_Rules.txt
- UCA\_Rules.xml

<http://www.unicode.org/Public/UCA/6.2.0/CollationTest.zip>:

- CollationTest\_NON\_IGNOREABLE\_SHORT.txt
- CollationTest\_NON\_IGNOREABLE.txt
- CollationTest\_SHIFTED\_SHORT.txt
- CollationTest\_SHIFTED.txt