Article Talk

Read Edit View history

Search

Unicode collation algorithm

From Wikipedia, the free encyclopedia
(Redirected from Unicode Collation Algorithm)

The **Unicode collation algorithm** (UCA) is an algorithm defined in Unicode Technical Report #10, which defines a customizable method to compare two strings. These comparisons can then be used to collate or sort text in any writing system and language that can be represented with Unicode.

Unicode Technical Report #10 also specifies the *Default Unicode Collation Element Table* (DUCET). This datafile specifies the default collation ordering. The DUCET is customizable for different languages. Some such customisations can be found in Common Locale Data Repository (CLDR).

An important open source implementation of UCA is included with the International Components for Unicode, ICU. ICU also supports tailoring and the collation tailorings from CLDR are included in ICU. You can see the effects of tailoring and a large number of language specific tailorings in the on-line **ICU Locale Explorer**.

See also   [edit]

- Collation
- ISO/IEC 14651
- European ordering rules (EOR)
- Common Locale Data Repository (CLDR)

External links and references   [edit]

- Unicode Collation Algorithm: Unicode Technical Standard #10
- Mimer SQL Unicode Collation Charts
- MySQL UCA-based Unicode Collation Charts

Tools   [edit]

- ICU Locale Explorer An online demonstration of the Unicode Collation Algorithm using International Components for Unicode
- msort A sort program that provides an unusual level of flexibility in defining collations and extracting keys.
- OpenRTL A library of functions using Unicode collation based on the Unicode collation algorithm. Also supports the customized Unicode collations for the locales defined by CLDR.

v · t · e	**Unicode**	[show]

This *algorithms* or *data structures*-related article is a *stub*. You can help Wikipedia by *expanding it*.

This *standards*- or *measurement*-related article is a *stub*. You can help Wikipedia by *expanding it*.

Categories: String collation algorithms | Unicode algorithms | Collation | Algorithms and data structures stubs | Computer science stubs | Standards and measurement stubs