Inhaltsverzeichnis

Modul:Unicode data/Doku

Dies ist die Dokumentationsseite für Modul:Unicode data

Inhaltsverzeichnis

1 Usage 2
2 Functions 2
3 Data modules 2
4 Copyright 3

Usage

This module provides functions that access information on Unicode code points. The information is retrieved from data modules generated from the Unicode Character Database, or derived by rules given in the Unicode Specification. It and its submodules were copied from English Wiktionary and then modified; see there for more information.

Functions

Vorlage:Code

Receives a code point (number) and returns its name or label; for example, Vorlage:Code returns Vorlage:Code.

For example, Vorlage:Tnull → <reserved-0061>

Vorlage:Code

Template-invokable functions that allow access to the functions starting with `lookup` and `is`. Replace the first underscore in the function name with a pipe. For most of the functions, add the code point in hexadecimal base as the next parameter, but for `is_Latin`, `is_rtl`, and `is_valid_pagename`, add text. HTML character references in the text are decoded by the module into code points.

For example, Vorlage:Tnull → **Lua-Fehler in Modul:Unicode_data, Zeile 293: attempt to index local 'data_module' (a boolean value)**.

Data modules

The data used by functions in this module is found in submodules. Some are generated by AWK scripts shown at User:Kephir/Unicode on English Wiktionary, others by Lua scripts on the `/make` subpages of the submodules.

- Module:Unicode data/aliases: the formal name aliases for characters (from NameAliases.txt)
- Module:Unicode data/blocks: the list of Unicode blocks (from Blocks.txt)
- Module:Unicode data/category: data mapping characters to their General Category (from DerivedGeneralCategory.txt)

- Module:Unicode data/control: data for identifying characters that belong to the General Categories of Separator and Other (from DerivedGeneralCategory.txt)
- Module:Unicode data/combining: data mapping characters to their Combining Classes (from DerivedCombiningClass.txt)
- Module:Unicode data/Hangul: data used to generate the names of Hangul syllables (from Jamo.txt)
- Module:Unicode data/scripts: data mapping characters to their Unicode script properties (from Scripts.txt).

The name data modules (Module:Unicode data/names/*xxx*) were compiled from UnicodeData.txt. Each one contains, at maximum, code points U+xxx000 to U+xxxFFF. **Lua-Fehler in mw.title.lua, Zeile 206: too many expensive function calls**

Copyright

The Unicode database is released by Unicode Inc. under the following terms:

> Copyright © 1991-2018 Unicode, Inc. All rights reserved. Distributed under the Terms of Use in https://www.unicode.org/copyright.html.
>
> Permission is hereby granted, free of charge, to any person obtaining a copy of the Unicode data files and any associated documentation (the "Data Files") or Unicode software and any associated documentation (the "Software") to deal in the Data Files or Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, and/or sell copies of the Data Files or Software, and to permit persons to whom the Data Files or Software are furnished to do so, provided that either (a) this copyright and permission notice appear with all copies of the Data Files or Software, or (b) this copyright and permission notice appear in associated Documentation.
>
> THE DATA FILES AND SOFTWARE ARE PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OF THIRD PARTY RIGHTS. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR HOLDERS INCLUDED IN THIS NOTICE BE LIABLE FOR ANY CLAIM, OR ANY SPECIAL INDIRECT OR CONSEQUENTIAL DAMAGES, OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THE DATA FILES OR SOFTWARE.
>
> Except as contained in this notice, the name of a copyright holder shall not be used in advertising or otherwise to promote the sale, use or other dealings in these Data Files or Software without prior written authorization of the copyright holder.