:mod:`unicodedata` --- Unicode Database

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 1`); *backlink*

Unknown interpreted text role "mod".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 4`)

Unknown directive type "module".

```
.. module:: unicodedata
   :synopsis: Access the Unicode Database.
```

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 7`)

Unknown directive type "moduleauthor".

```
.. moduleauthor:: Marc-André Lemburg <mal@lemburg.com>
```

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 8`)

Unknown directive type "sectionauthor".

```
.. sectionauthor:: Marc-André Lemburg <mal@lemburg.com>
```

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 9`)

Unknown directive type "sectionauthor".

```
.. sectionauthor:: Martin v. Löwis <martin@v.loewis.de>
```

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 11`)

Unknown directive type "index".

```
.. index::
   single: Unicode
   single: character
   pair: Unicode; database
```

This module provides access to the Unicode Character Database (UCD) which defines character properties for all Unicode characters. The data contained in this database is compiled from the UCD version 14.0.0.

The module uses the same names and symbols as defined by Unicode Standard Annex #44, "Unicode Character Database". It defines the following functions:

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 29`)

Unknown directive type "function".

```
.. function:: lookup(name)

   Look up character by name.  If a character with the given name is found, return
   the corresponding character.  If not found, :exc:`KeyError` is raised.

   .. versionchanged:: 3.3
      Support for name aliases [#]_ and named sequences [#]_ has been added.
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 38)

Unknown directive type "function".

```
.. function:: name(chr[, default])

   Returns the name assigned to the character *chr* as a string. If no
   name is defined, *default* is returned, or, if not given, :exc:`ValueError` is
   raised.
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 45)

Unknown directive type "function".

```
.. function:: decimal(chr[, default])

   Returns the decimal value assigned to the character *chr* as integer.
   If no such value is defined, *default* is returned, or, if not given,
   :exc:`ValueError` is raised.
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 52)

Unknown directive type "function".

```
.. function:: digit(chr[, default])

   Returns the digit value assigned to the character *chr* as integer.
   If no such value is defined, *default* is returned, or, if not given,
   :exc:`ValueError` is raised.
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 59)

Unknown directive type "function".

```
.. function:: numeric(chr[, default])

   Returns the numeric value assigned to the character *chr* as float.
   If no such value is defined, *default* is returned, or, if not given,
   :exc:`ValueError` is raised.
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 66)

Unknown directive type "function".

```
.. function:: category(chr)

   Returns the general category assigned to the character *chr* as
   string.
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\cpython-main\Doc\library\[cpython-main][Doc][library]unicodedata.rst, line 72)

Unknown directive type "function".

```
.. function:: bidirectional(chr)

   Returns the bidirectional class assigned to the character *chr* as
   string. If no such value is defined, an empty string is returned.
```

Unknown directive type "function".

```
.. function:: combining(chr)

   Returns the canonical combining class assigned to the character *chr*
   as integer. Returns ``0`` if no combining class is defined.
```

Unknown directive type "function".

```
.. function:: east_asian_width(chr)

   Returns the east asian width assigned to the character *chr* as
   string.
```

Unknown directive type "function".

```
.. function:: mirrored(chr)

   Returns the mirrored property assigned to the character *chr* as
   integer. Returns ``1`` if the character has been identified as a "mirrored"
   character in bidirectional text, ``0`` otherwise.
```

Unknown directive type "function".

```
.. function:: decomposition(chr)

   Returns the character decomposition mapping assigned to the character
   *chr* as string. An empty string is returned in case no such mapping is
   defined.
```

Unknown directive type "function".

```
.. function:: normalize(form, unistr)

   Return the normal form *form* for the Unicode string *unistr*. Valid values for
   *form* are 'NFC', 'NFKC', 'NFD', and 'NFKD'.

   The Unicode standard defines various normalization forms of a Unicode string,
   based on the definition of canonical equivalence and compatibility equivalence.
   In Unicode, several characters can be expressed in various way. For example, the
   character U+00C7 (LATIN CAPITAL LETTER C WITH CEDILLA) can also be expressed as
   the sequence U+0043 (LATIN CAPITAL LETTER C) U+0327 (COMBINING CEDILLA).

   For each character, there are two normal forms: normal form C and normal form D.
   Normal form D (NFD) is also known as canonical decomposition, and translates
   each character into its decomposed form. Normal form C (NFC) first applies a
   canonical decomposition, then composes pre-combined characters again.

   In addition to these two forms, there are two additional normal forms based on
   compatibility equivalence. In Unicode, certain characters are supported which
```

normally would be unified with other characters. For example, U+2160 (ROMAN
NUMERAL ONE) is really the same thing as U+0049 (LATIN CAPITAL LETTER I).
However, it is supported in Unicode for compatibility with existing character
sets (e.g. gb2312).

The normal form KD (NFKD) will apply the compatibility decomposition, i.e.
replace all compatibility characters with their equivalents. The normal form KC
(NFKC) first applies the compatibility decomposition, followed by the canonical
composition.

Even if two unicode strings are normalized and look the same to
a human reader, if one has combining characters and the other
doesn't, they may not compare equal.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-`
`main\Doc\library\[cpython-main][Doc][library]unicodedata.rst`, line 136)

Unknown directive type "function".

```
.. function:: is_normalized(form, unistr)

   Return whether the Unicode string *unistr* is in the normal form *form*. Valid
   values for *form* are 'NFC', 'NFKC', 'NFD', and 'NFKD'.

   .. versionadded:: 3.8
```

In addition, the module exposes the following constant:

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-`
`main\Doc\library\[cpython-main][Doc][library]unicodedata.rst`, line 146)

Unknown directive type "data".

```
.. data:: unidata_version

   The version of the Unicode database used in this module.
```

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\cpython-`
`main\Doc\library\[cpython-main][Doc][library]unicodedata.rst`, line 151)

Unknown directive type "data".

```
.. data:: ucd_3_2_0

   This is an object that has the same methods as the entire module, but uses the
   Unicode database version 3.2 instead, for applications that require this
   specific version of the Unicode database (such as IDNA).
```

Examples:

```
>>> import unicodedata
>>> unicodedata.lookup('LEFT CURLY BRACKET')
'{'
>>> unicodedata.name('/')
'SOLIDUS'
>>> unicodedata.decimal('9')
9
>>> unicodedata.decimal('a')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ValueError: not a decimal
>>> unicodedata.category('A')  # 'L'etter, 'u'ppercase
'Lu'
>>> unicodedata.bidirectional('\u0660') # 'A'rabic, 'N'umber
'AN'
```

Footnotes

[1]   https://www.unicode.org/Public/14.0.0/ucd/NameAliases.txt

[2]   https://www.unicode.org/Public/14.0.0/ucd/NamedSequences.txt