Package ‘cnum’

April 13, 2020

Type Package
Title Chinese Numerals Processing
Version 0.1.1
Maintainer Elgar Teo <elgarteo@connect.hku.hk>

URL https://github.com/elgarteo/cnum/

BugReports https://github.com/elgarteo/cnum/issues

Description Chinese numerals processing in R, such as conversion between
Chinese numerals and Arabic numerals as well as detection and extraction of
Chinese numerals in character objects and string. This package supports
the casual scale naming system and the respective SI prefix systems used
in mainland China and Taiwan:
```
```

License MIT + file LICENSE
Encoding UTF-8
Depends R(>= 2.10)
Imports stringr
Suggests magrittr
LazyData true
RoxygenNote 7.1.0

NeedsCompilation no
Author Elgar Teo [aut, cre]
Repository CRAN

Date/Publication 2020-04-13 08:00:02 UTC
R topics documented:

- c2num ......................................................... 2
- default_cnum_lang ........................................ 4
- tools ....................................................... 5

Index 8

---

c2num  Chinese Numerals Conversion

Description

Functions to convert between Chinese and Arabic numerals.

Usage

c2num(
  x,
  lang = default_cnum_lang(),
  mode = "casual",
  financial = FALSE,
  literal = FALSE
)

um2c(
  x,
  lang = default_cnum_lang(),
  mode = "casual",
  financial = FALSE,
  literal = FALSE,
  single = FALSE
)

Arguments

- x  the Arabic/Chinese numerals to be converted, or a vector of them. The absolute value must not be greater than 1e+18.
- lang the language of the Chinese numerals. "tc" for Traditional Chinese. "sc" for Simplified Chinese. The default is "tc", but this can be changed by setting options(cnum.lang = "sc").
- mode the scale naming system to be enforced. See the ‘Details’ section for the list of supported modes.
- financial logical: should the financial numerals be used (daxie shuzi)?
- literal logical: should the numerals be converted literally? (e.g. 721 to be converted to "qi er yi" instead of "qibai ershiyi" and vice versa)
- single logical: should the return result with one scale character only? (e.g. 1.5e+08 as "yi dian wuyi" instead of "yiyi wuqianwan"
Value

c2num returns a numeric vector.
num2c returns a character vector.

Functions

- c2num: Convert Chinese Numerals to Arabic Numerals.
- num2c: Convert Arabic Numerals to Chinese Numerals.

Details

The following scale naming systems are supported:

- "casual": the casual naming system used outside of mainland China, i.e. 1e+09 is referred to as "yi zhao".
- "casualPRC": the casual naming system used in mainland China, i.e. 1e+9 is referred to as "yi wanyi".
- "SIprefix": the SI prefix system used in Taiwan as stipulated in the document *Names, Definitions and Symbols of the Legal Units of Measurement and the Decimal Multiples and Submultiples*.
- "SIprefixPRC": the SI prefix system used in mainland China as stipulated in the document *China Statutory Measurement Units*.
- "SIprefixPRClong": a variant of "SIprefixPRC" with long prefixes, e.g. 1e+09 is referred to as "jika" instead of "ji".

Warnings

The modes "casual" and "casualPRC" implements a "myriad scale" with an interval of 1e+04 for large numbers, i.e. "yi" is 10,000 times of "wan", which is different from some of the interval systems used in ancient Chinese writings.

The precision of converting large numbers and long decimals is limited in base R. You might be able to improve the accuracy by running options(digits = 22).

References


See Also

Functions for detection and extraction
Examples

c2num("hello")

num2c(721)
num2c(-6)
num2c(3.14)
num2c(721, literal = TRUE)
num2c(1.45e12, financial = TRUE)
num2c(6.85e12, lang = "sc", mode = "casualPRC")
num2c(1.5e9, mode = "SIprefix", single = TRUE)

Description

Function to check the default language for cnum functions.

Usage

default_cnum_lang()

Details

This package supports Traditional Chinese and Simplified Chinese. The language can be specified with the lang parameter in every function, with "tc" for Traditional Chinese and "sc" for Simplified Chinese. The default is "tc", but this can be changed by setting options(cnum.lang = "sc").

Value

The default language for cnum functions.

See Also

- Functions for conversion
- Functions for detection and extraction

Examples

# Set the default language to Simplified Chinese
options(cnum.lang = "sc")
default_cnum_lang()
Chinese Numerals Detection and Extraction

Description
Functions to detect and extract Chinese numerals in character object and string.

Usage

```r
is_cnum(
  x,
  lang = default_cnum_lang(),
  mode = "casual",
  financial = FALSE,
  literal = FALSE,
  strict = FALSE,
  ...
)

has_cnum(
  x,
  lang = default_cnum_lang(),
  mode = "casual",
  financial = FALSE,
  ...
)

extract_cnum(
  x,
  lang = default_cnum_lang(),
  mode = "casual",
  financial = FALSE,
  ...
)
```

Arguments

- **x** the character object or string to be tested or to extract from.
- **lang** the language of the Chinese numerals. "tc" for Traditional Chinese. "sc" for Simplified Chinese. The default is "tc", but this can be changed by setting `options(cnum.lang = "sc")`.
- **mode** the scale naming system to be enforced. See the ‘Details’ section for the list of supported modes.
- **financial** logical: should the financial numerals be used (`daxie shuzi`)?
- **literal** logical: should the numerals be converted literally? (e.g. 721 to be converted to "qi er yi" instead of "qibai ershiyi" and vice versa)
strict logical: Should the Chinese numerals format be strictly enforced? A casual test only checks if x contains Chinese numerals characters. A strict test checks if x is valid Chinese numerals. (e.g. "yì bāi yì" will pass the casual test and fail the strict test)

... optional arguments to be passed to grepl (for is_cnum and has_cnum) or str_extract_all (for extract_cnum). Disregarded when strict = TRUE.

Value

is_cnum returns a logical vector indicating is Chinese numerals or not for each element of x).

has_cnum returns a logical vector indicating contains Chinese numerals or not for each element of x).

extract_cnum returns a list of character vectors containing the extracted Chinese numerals.

Functions

- **is_cnum**: Test if character object is Chinese numerals
- **has_cnum**: Test if string contains Chinese numerals
- **extract_cnum**: Extract Chinese numerals from string

Details

The following scale naming systems are supported:

- "casual": the casual naming system used outside of mainland China, i.e. 1e+09 is referred to as "yī zhào".
- "casualPRC": the casual naming system used in mainland China, i.e. 1e+9 is referred to as "yī wàn yì".
- "SIprefix": the SI prefix system used in Taiwan as stipulated in the document Names, Definitions and Symbols of the Legal Units of Measurement and the Decimal Multiples and Submultiples.
- "SIprefixPRC": the SI prefix system used in mainland China as stipulated in the document China Statutory Measurement Units.
- "SIprefixPRClong": a variant of "SIprefixPRC" with long prefixes, e.g. 1e+09 is referred to as "jīkā" instead of "jī".

References


See Also

Functions for conversion
Examples

is_cnum("hello")

has_cnum("hello")

extract_cnum("hello")
Index

c2num, 2
conversion (c2num), 2

default_cnum_lang, 4

extract_cnum (tools), 5

Functions for conversion, 4, 6
Functions for detection and extraction, 3, 4

grepl, 6

has_cnum (tools), 5

is_cnum (tools), 5

num2c (c2num), 2

options, 2, 4, 5

str_extract_all, 6

tools, 5