Package 'rematch'

August 30, 2023

Title Match Regular Expressions with a Nicer 'API'

Version 2.0.0

Author Gabor Csardi

Maintainer Gabor Csardi <csardi.gabor@gmail.com>

Description A small wrapper on 'regexpr' to extract the matches and captured groups from the match of a regular expression to a character vector.

License MIT + file LICENSE

URL https://github.com/gaborcsardi/rematch

BugReports https://github.com/gaborcsardi/rematch/issues

RoxygenNote 5.0.1.9000

Suggests covr, testthat

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2023-08-30 16:50:02 UTC

R topics documented:

rematch 1
re_match 2
re_match_all 3

Index **4**

`rematch` *Match Regular Expressions with a Nicer 'API'*

Description

A small wrapper on 'regexpr' to extract the matches and captured groups from the match of a regular expression to a character vector. See `re_match`.

re_match *Match a regular expression to a character vector*

Description

This function is a small wrapper on the regexpr base R function, to provide an API that is easier to use.

Usage

```
re_match(pattern, text, ...)
```

Arguments

pattern	Regular expression, defaults to be a PCRE expression. See regex for more about regular expressions.
text	Character vector.
...	Additional arguments to pass to regexpr.

Details

Currently only the first occurence of the pattern is used.

Value

A character matrix of the matched (sub)strings. The first column is always the full match. This column is named .match. The result of the columns are capture groups, with appropriate column names, if the groups are named.

Examples

```
dates <- c("2016-04-20", "1977-08-08", "not a date", "2016",
  "76-03-02", "2012-06-30", "2015-01-21 19:58")
isodate <- "([0-9]{4})-([0-1][0-9])-([0-3][0-9])"
re_match(text = dates, pattern = isodate)

# The same with named groups
isodaten <- "(?<year>[0-9]{4})-(?<month>[0-1][0-9])-(?<day>[0-3][0-9])"
re_match(text = dates, pattern = isodaten)
```

re_match_all *Extract all matches of a regular expression*

Description

This function is a thin wrapper on the gregexpr base R function, to provide an API that is easier to use. It is similar to re_match, but extracts all matches, including potentially named capture groups.

Usage

```
re_match_all(pattern, text, ...)
```

Arguments

pattern	Regular expression, defaults to be a PCRE expression. See regex for more about regular expressions.
text	Character vector.
...	Additional arguments to pass to regexpr.

Value

A list of character matrices. Each list element contains the matches of one string in the input character vector. Each matrix has a .match column that contains the matching part of the string. Additional columns are added for capture groups. For named capture groups, the columns are named.

Index

gregexpr, *3*

re_match, *1*, 2, *3*
re_match_all, 3
regex, *2, 3*
regexpr, *2, 3*
rematch, 1
rematch-package *(*rematch*)*, 1