# Package 'tinyoperators'

June 27, 2023

| June 27, 2023   |
|---|
| Title Infix operators and some functions to help in proper coding etiquette   |
| Version 0.0.9   |
| Description The 'tinyoperators' R-package adds some infix operators, and a few functions.  It primarily focuses on 4 things.  (1) Float truth testing.  (2) Reducing repetitive code.  (3) Extending the string manipulation capabilities of the 'stringi' R package.  (4) A new package and module import system, that combines the benefits of aliasing a package with the benefits of attaching a package. The 'tinyoperators' R-package has only one dependency, namely 'stringi', though it does allows multi-threading of some of the string-related functions (when appropriate) via the suggested 'stringfish' R-package. |
| License MIT + file LICENSE  |
| Encoding UTF-8  |
| <b>Roxygen</b> list(markdown = TRUE)  |
| RoxygenNote 7.2.3   |
| Suggests rlang, knitr, rmarkdown, stringfish (>= 0.15.7), tinytest  |
| VignetteBuilder knitr   |
| <b>Depends</b> R (>= 4.1.0)   |
| Imports stringi (>= 1.7.12)   |
| URL https://github.com/tony-aw/tinyoperators  |
| BugReports https://github.com/tony-aw/tinyoperators/issues  |
| R topics documented:  |
| float_truth_testing   |

2 float\_truth\_testing

float\_truth\_testing Safer float (in)equality operators

## **Description**

**Index** 

The %f==%, %f!=% %f<%, %f>%, %f<=%, %f>=% (in)equality operator perform float truth testing. They are virtually equivalent to the regular (in)equality operators,

except for one aspect. The float truth testing operators assume that if the absolute difference between x and y is smaller than the Machine tolerance, sqrt(.Machine\$double.eps), then x and y ought to be consider to be equal.

Thus these provide safer float truth testing.

For example: 0.1\*7 == 0.7 returns FALSE, even though they are equal, due to the way floating numbers are stored in programming languages like R. But 0.1\*7 %f==% 0.7 returns TRUE.

There are also the  $x \%f{}\%$  bnd and  $x \%f{}\%$  bnd operators, where bnd is a vector of length 2, or a 2-column matrix (nrow(bnd)==length(x) or nrow(bnd)==1).

The x %f{}% bnd operator checks if x is within the closed interval with bounds defined by bnd.

The  $x \%f!{}\%$  bnd operator checks if x is outside the closed interval with bounds defined by bnd.

## Usage

x %f==% y

x %f!=% y

x %f<% y

x %f>% y

x %f<=% y

x %f>=% y

x %f{}% bnd

float\_truth\_testing 3

```
x %f!{}% bnd
```

#### **Arguments**

x, y numeric vectors, matrices, or arrays, though these operators were specifically

designed for floats (class "double").

bnd either a vector of length 2, or a matrix with 2 columns and 1 row, or else a matrix

with 2 columns where nrow(bnd) == length(x).

The first element/column of bnd gives the lower bound of the closed interval; The second element/column of bnd gives the upper bound of the closed interval;

```
x \leftarrow c(0.3, 0.6, 0.7)
y <- c(0.1*3, 0.1*6, 0.1*7)
print(x); print(y)
x == y \# gives FALSE, but should be TRUE
x!= y \# gives TRUE, should be FALSE
x > y \# not wrong
x < y \# gives TRUE, should be FALSE
x %f==% y # here it's done correctly
x %f!=% y # correct
x %f<% y # correct
x %f>% y # correct
x \%f \le y \# correct
x %f>=% y # correct
x < -c(0.3, 0.6, 0.7)
bnd <- matrix(c(0.29, 0.59, 0.69, 0.31, 0.61, 0.71), ncol=2)
x %f{}% bnd
x %f!{}% bnd
# These operators still work for non-float numerics also:
x <- 1:5
y <- 1:5
x %f==% y
x %f!=% y
x %f<% y
x %f>% y
x %f<=% y
x %f>=% y
x <- 1:5
y <- x+1
x %f==% y
x %f!=% y
x %f<% y
x %f>% y
x %f<=% y
x %f>=% y
x <- 1:5
y <- x-1
x %f==% y
```

4 import

```
x %f!=% y
x %f<% y
x %f>% y
x %f<=% y
x %f>=% y
```

import

Additional package import management

#### **Description**

These functions implement a new package import system, that attempts to combine the benefits of aliasing a package with the benefits of attaching a package.

```
import_as:
```

The import\_as() function imports the namespace of an R package, and optionally also its dependencies, enhances, and extensions, under the same alias. The specified alias will be placed in the current environment (like the global environment, or the environment within a function).

```
import_inops:
```

The import\_inops() function exposes the infix operators of the specified packages to the current environment (like the global environment, or the environment within a function).

To ensure the user can still verify which operator function came from which package, a "package" attribute is added to each exposed operator.

Naturally, the namespaces of the operators remain intact.

```
import_data:
```

The import\_data() function gets a specified data set from a package.

Unlike utils::data(), the import\_data() function returns the data set directly, and allows assigning the data set like so:

```
mydata <- import_data(...).</pre>
```

### import\_lsf:

The import\_lsf(package, ...) function gets a list of exported functions/operators from a package.

# Usage

```
import_as(
   alias,
   main_package,
   depends = FALSE,
   enhances = NULL,
   extends = NULL,
   loadorder = c("depends", "main_package", "enhances", "extends"),
   lib.loc = .libPaths()
)
import_inops(pkgs, lib.loc = .libPaths(), exclude, include.only)
```

import 5

```
import_data(dataname, package, lib.loc = .libPaths())
import_lsf(package, type, lib.loc = .libPaths())
```

#### **Arguments**

alias a variable name (unquoted), giving the (not yet existing) object where the pack-

age(s) are to be assigned to.

Syntactically invalid names are not allowed for the alias name.

main\_package a single string, giving the name of the main package to load under the given

alias.

depends either logical, or a character vector.

If FALSE (default), no dependencies are loaded under the alias.

If TRUE, ALL dependencies of the main\_package are loaded under the alias, but **excluding** core/base R packages, and also **excluding** pre-installed "recommended" R packages. See also pkgs\_get\_deps

If a character vector, then it is taken as the dependencies of the package to be loaded also under the alias.

NOTE (1): "Dependencies" here are defined as any package appearing in the "Depends", "Imports", or "LinkingTo" sections of the Description file of the main\_package.

NOTE (2): If depends is a character vector: The order of the character vector matters! If multiple packages share objects with the same name, the package named last will overwrite the earlier named packages.

enhances

a character vector, giving the names of the packages enhanced by the main\_package to be loaded also under the alias.

NOTE(1): Enhances are defined as packages appearing in the "Enhances" section of the Description file of the main\_package.

NOTE (2): The order of the character vector matters! If multiple packages share objects with the same name, the objects of the package named last will overwrite those of the earlier named packages.

extends

a character vector, giving the names of the extensions / reverse-dependencies of the main\_package to be loaded also under the alias. Defaults to NULL, which means no extensions are loaded.

NOTE (1): "Extensions" here are defined as reverse-depends or reverse-imports. It does not matter if these are CRAN or non-CRAN packages.

NOTE (2): The order of the character vector matters! If multiple packages share objects with the same name, the objects of the package named last will overwrite those of the earlier named packages.

loadorder the character vector

c("depends", "main\_package", "enhances", "extends"),

or some re-ordering of this character vector, giving the relative load order of the groups of packages.

By default this is the character vector

c("depends", "main\_package", "enhances", "extends").

lib.loc

character vector specifying library search path (the location of R library trees to search through). This is usually .libPaths(). See also loadNamespace.

6 import

pkgs a single string, or character vector, with the package name(s).

NOTE (1): The order of the character vector matters! If multiple packages share objects with the same name, the objects of the package named last will overwrite

those of the earlier named packages.

NOTE (2): The import\_inops function performs a basic check that the packages are mostly (reverse) dependencies of each other. If not, it will give an error.

exclude a character vector, giving the infix operators NOT to expose to the current envi-

ronment.

This can be handy to prevent overwriting any (user defined) infix operators already present in the current environment.

NOTE: You cannot specify both the exclude and include.only arguments.

Only one or the other, or neither.

include.only a character vector, giving the infix operators to expose to the current environ-

ment, and the rest of the operators will not be exposed.

This can be handy to prevent overwriting any (user defined) infix operators al-

ready present in the current environment.

NOTE: You cannot specify both the exclude and include.only arguments.

Only one or the other, or neither.

dataname a single string, giving the name of the data set.

package the quoted package name.

type The type of functions to list. Possibilities:

"inops" or "operators": Only infix operators.

"regfuns": Only regular functions (thus excluding infix operators). "all": All functions, both regular functions and infix operators.

## **Details**

Regarding "For a more detailed description of the import system introduced by the tinyoperators R package, please refer to the Read Me file on the GitHub main page:

https://github.com/tony-aw/tinyoperators

#### Value

For import\_as:

The variable named in the alias argument will be created (if it did not already exist), in the current environment (like the global environment, or the environment within a function). The alias object will contain the (merged) package environment.

For import\_inops():

The infix operators from the specified packages will be placed in the current environment (like the Global environment, or the environment within a function).

For import\_data():

Returns the data directly. Thus, one can assign the data like so: mydata <- import\_data(...).

For import\_lsf():

Returns a character vector of function and/or operator names.

inplace 7

## **Examples**

```
## Not run:
depends <- unlist(tools::package_dependencies("devtools"))
pkgs <- c(depends, "devtools")
import_as(devt, "devtools", depends = TRUE) # this creates the devt object
import_inops(pkgs)
d <- import_data("chicago", "gamair")
head(d)
## End(Not run)</pre>
```

inplace

Generalized in-place (mathematical) modifier

# Description

Generalized in-place (mathematical) modifier.

The x %:=% f operator allows performing in-place modification of some object x with a function f.

```
For example this:
object[object > 0] %:=% \(x) x + 1

Is the same as:
object[object > 0] <- object[object > 0] + 1
```

This function-based method is used instead of the more traditional in-place mathematical modification like += to prevent precedence issues (functions come before mathematical arithmetic in R).

## Usage

```
x %:=% f
```

## **Arguments**

x an object, with properties such that function f can be use on it.

For example, when function f is mathematical, x should be a number or numeric (or 'number-like') vector, matrix, or array.

f a function to be applied in-place on x.

## Value

This operator does not return any value: it is an in-place modifiers, and thus modifies x directly.

## **Examples**

```
set.seed(1)
object <- matrix(rpois(10, 10), ncol=2)
print(object)
object %:=% \(x) x+3 # same as object <- object + 3
print(object)</pre>
```

```
inplace_str_arithmetic
```

In place modifying string arithmetic

# Description

In-place modifier versions of string arithmetic:

```
x %s+ =% y is the same as x <- x %s+% y
x %s- =% p is the same as x <- x %s-% p
x %s* =% n is the same as x <- x %s*% n
x %s/ =% p is the same as x <- x %s/% p</pre>
```

See also the documentation on string arithmetic: string arithmetic.

# Usage

```
x %s+ =% y
x %s- =% p
x %s* =% n
x %s/ =% p
```

# **Arguments**

```
x, y, p, n see string arithmetic and s_pattern.
```

## Value

These operators do not return any value: they are in-place modifiers, and thus modify x directly.

inplace\_str\_arithmetic 9

```
y <- "a"
p \leftarrow "a|e|i|o|u"
n < -c(2, 3)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
x %s+ =% y # same as x <- x %s+% y
print(x)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
x \%s- =\% p \# same as x <- x \%s-\% p
print(x)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
x %s* =% n # same as <math>x <- x %s\\*% n
print(x)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
x \%s/ = % p # same as <math>x < - x \%s/\% p
print(x)
y <- "a"
# pattern with ignore.case=TRUE:
p <- s_pattern(regex = "A|E|I|0|U", ignore.case=TRUE)</pre>
n < -c(3, 2)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
x \%s + = % y # same as <math>x < - x \%s + % y
print(x)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
x %s - = % p # same as x < - x %s - % p
print(x)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
x %s* =% n # same as <math>x <- x %s\*% n
print(x)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
x \%s/ = % p # same as <math>x < - x \%s/\% p
print(x)
```

inplace\_str\_subset

inplace\_str\_subset

In place modifying string subsetting

## **Description**

In-place modifier versions of string subsetting:

```
x %sget =% ss is the same as x <- x %sget% ss
x %strim =% ss is the same as x <- x %strim% ss</pre>
```

See also the documentation on string subsetting (string subset). Note that there is no in-place modifier versions of %ss%.

## Usage

```
x %sget =% ss
x %strim =% ss
```

## **Arguments**

```
x, ss see string subset.
```

#### Value

These operators do not return any value: they are in-place modifiers, and thus modify x directly.

```
ss <- c(2,2)
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))
print(x)</pre>
```

logic\_ops 11

```
x %sget =% ss # same as x <- x %sget% ss
print(x)

x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))
print(x)
x %strim =% ss # same as x <- x %strim% ss
print(x)</pre>
```

logic\_ops

Logic operators

## **Description**

Additional logic operators:

The x %xor% y operator is the "exclusive-or" operator, the same as xor(x, y).

The x n& operator is the "not-and" operator, the same as (!x) & (!y).

The x %out% y operator is the same as !x %in% y.

The x %?=% y operator checks if x and y are **both** unreal or unknown (i.e. NA, NaN, Inf, -Inf).

The n %=numtype% numtype operator is a vectorized operator that checks for every value of numeric vector n if it can be considered a number belonging to type numtype. See arguments for details.

The s %=strtype% strtype operator is a vectorized operator that checks for every value of character vector s

if it can seen as a certain strtype. See arguments for details.

The s %sgrep% p operator is a vectorized operator that checks for every value of character vector s if it has pattern p.

## Usage

```
x %xor% y
x %n&% y
x %out% y
x %?=% y
s %sgrep% p
n %=numtype% numtype
s %=strtype% strtype
```

12 logic\_ops

#### **Arguments**

x, y see Logic.

s a character vector.

p the result from s\_pattern, or else a character vector of the same length as s with

regular expressions.

n a numeric vector.

numtype a single string giving the type if numeric to be checked. The following options are supported:

- "~0": zero, or else a number whose absolute value is smaller than the Machine tolerance (sqrt(.Machine\$double.eps)).
- "B": binary numbers (exactly 0 or exactly 1);
- "prop": proportions numbers between 0 and 1 (exactly 0 or 1 is also allowed);
- "N": Natural numbers (non-negative integers including zero);
- "I": Integers;
- "odd": odd integers;
- "even": even integers;
- "R": Real numbers;
- "unreal": infinity, NA, or NaN;

strtype

a single string giving the type of string to be checked. The following options are supported:

- "empty": checks if the string only consists of empty spaces.
- "unreal": checks if the string is NA, or if it has literal string "NA", "NaN" or "Inf", regardless if it has leading or trailing spaces.
- "numeric": checks if the string can be converted to a number, disregarding leading and trailing spaces. I.e. the string "5.0" can be converted to the the actual number 5.0.
- "special": checks if the string consists of only special characters.

```
x \leftarrow c(TRUE, FALSE, TRUE, FALSE, NA, NaN, Inf, -Inf, TRUE, FALSE) y \leftarrow c(FALSE, TRUE, TRUE, FALSE, rep(NA, 6)) outcome <- data.frame(
```

matrix\_ops 13

```
x=x, y=y,
  "x %xor% y"=x %xor% y, "x %n&% y" = x %n&% y, "x %?=% y" = x %?=% y,
  check.names = FALSE
print(outcome)
1:3 %out% 1:10
1:10 %out% 1:3
n <- c(0:5, 0:-5, 0.1, -0.1, 0, 1, Inf, -Inf, NA, NaN)
1e-20 %=numtype% "~0"
n[n %=numtype% "B"]
n[n %=numtype% "prop"]
n[n %=numtype% "N"]
n[n %=numtype% "I"]
n[n %=numtype% "odd"]
n[n %=numtype% "even"]
n[n %=numtype% "R"]
n[n %=numtype% "unreal"]
s <- c(" AbcZ123 ", " abc ", " 1.3 ", " !#$%^&*() ", " ", " NA ", " NA ", " Inf ")\\
s[s %=strtype% "empty"]
s[s %=strtype% "unreal"]
s[s %=strtype% "numeric"]
s[s %=strtype% "special"]
s <- c("Hello world", "Goodbye world")</pre>
p <- s_pattern(regex = c("Hello", "Hello"))</pre>
s %sgrep% p
```

matrix\_ops

Infix operators for row- and column-wise re-ordering of matrices

# Description

Infix operators for custom row- and column-wise re-ordering of matrices

The  $x \%row^{\infty}$  mat operator re-orders the elements of every row of matrix x according to the ordering ranks given in matrix mat.

The  $x \col^{\infty}$  mat operator re-orders the elements of every column of matrix x according to the ordering ranks given in matrix mat.

# Usage

```
x %row~% mat
x %col~% mat
```

14 matrix\_ops

#### **Arguments**

x a matrix
mat a matrix with the same dimensions as x, giving the ordering ranks of every element of matrix x.

#### **Details**

If matrix x is a numeric matrix, and one wants to sort the elements of every row or column numerically, x %row~% x or x %col~% x would suffice, respectively.

If matrix x is not numeric, sorting the elements using x %row~% x and x %col~% x are still possible, but probably not the best option. In the non-numeric case, providing a matrix of ordering ranks for mat would be faster and give more accurate ordering. See the examples section.

If mat is a matrix of non-repeating random integers (i.e. sample(1:length(x)), replace=FALSE) sample(1:length(x)) |> matrix(ncol=ncol(x))), x %row~% mat will randomly shuffle the elements of every row, where the shuffling order of every row is independent of the other rows. Similarly, x %col~% mat will randomly shuffle the elements of every column, where the shuffling order of every column is independent of the other columns.

These operators internally only use vectorized operations (no loops or apply-like functions), and are faster than re-ordering matrices using loops or apply-like functions.

#### Value

A modified matrix.

```
# numeric matrix ====
x <- matrix(sample(1:25), nrow=5)</pre>
print(x)
x %row~% x # sort elements of every row
x %row~% -x # reverse-sort elements of every row
x %col~% x \# sort elements of every column
x %col~% -x # reverse-sort elements of every column
x <- matrix(sample(1:25), nrow=5)</pre>
mat <- sample(1:length(x)) |> matrix(ncol=ncol(x)) # matrix of non-repeating random integers
x %row~% mat # randomly shuffle every row independently
x %col~% mat # randomize shuffle every column independently
# character matrix ====
x <- matrix(sample(letters, 25), nrow=5)</pre>
print(x)
mat <- stringi::stri_rank(as.vector(x)) |> matrix(ncol=ncol(x))
x %row~% mat # sort elements of every row
x %row~% -mat # reverse-sort elements of every row
x \ \%col^{-}\% mat # sort elements of every column
```

pkgs\_get\_deps 15

```
x %col~% -mat # reverse-sort elements of every column

x <- matrix(sample(letters, 25), nrow=5)
print(x)
mat <- sample(1:length(x)) |> matrix(ncol=ncol(x)) # matrix of non-repeating random integers
x %row~% mat # randomly shuffle every row independently
x %col~% mat # randomize shuffle every column independently
```

pkgs\_get\_deps

Miscellaneous package functions

## **Description**

The pkgs %installed in% lib.loc operator checks if one or more package(s) pkgs exist(s) in library location lib.loc.

Now you no longer have to attach a package with require() simply to check if it exists.

Moreover, this operator makes it syntactically explicit in your code where you are looking for your R package(s).

The pkgs\_get\_deps() function gets the dependencies of a package from the Description file. It works on non-CRAN packages also.

#### Usage

```
pkgs_get_deps(
  package,
  lib.loc = .libPaths(),
  deps_type = c("Depends", "Imports", "LinkingTo"),
  base = FALSE,
  recom = FALSE
)

pkgs %installed in% lib.loc
```

## **Arguments**

| package   | a single string giving the package name   |  |
|-----------|---|--|
| lib.loc   | character vector specifying library search path (the location of R library trees to search through). This is usually .1ibPaths(). See also loadNamespace.   |  |
| deps_type | a character vector, giving the dependency types to be used.  Defaults to c("Depends", "Imports", "LinkingTo").  |  |
| base      | logical, indicating whether base/core R packages should be included (TRUE), or not included (FALSE; the default).   |  |
| recom     | logical, indicating whether the pre-installed "recommended" R packages should be included (TRUE), or not included (FALSE; the default). Note that only the recommended R packages actually installed in your system are taken into consideration. |  |
| pkgs      | a single string, or character vector, with the package name(s).   |  |

16 source\_module

#### Value

For pkgs %installed in% lib.loc: Returns a logical vector, where TRUE indicates a package is installed, and FALSE indicates a package is not installed.

For pkgs\_get\_deps(): A character vector of dependencies.

#### References

https://stackoverflow.com/questions/30223957/elegantly-extract-r-package-dependencies-of-a-package-not-listed-on-cran

## **Examples**

```
## Not run:
pkgs <- c(unlist(tools::package_dependencies("devtools")), "devtools")
pkgs %installed in% .libPaths()
## End(Not run)</pre>
```

source\_module

Additional module import management

# Description

The alias %@source% list(file=...) operator imports all objects from a source-able script file under an alias.

The source\_inops() function exposes the infix operators defined in a source-able script file to the current environment (like the global environment, or the environment within a function).

Note that the alias %@source% list(file=...) operator and the source\_inops() function do NOT suppress output (i.e. plots, prints, messages) from the sourced module file.

#### Usage

```
alias %@source% lst
source_inops(...)
```

## **Arguments**

| alias | a variable name (unquoted), giving the (not yet existing) object where the sourced objects from the module are to be assigned to.  Syntactically invalid names are not allowed for the alias name. |
|-------|--|
| lst   | a named list, giving the arguments to be passed to the source function. For example: alias %@source% list(file="mydir/myscript.R")   |
|       | arguments to be passed to the source function, such as the file argument.  |

stri\_join\_mat 17

#### Value

For the alias %@source% list(file=...) operator:

The variable named as the alias will be created (if it did not already exist) in the current environment, and will contain all objects from the sourced script.

For source\_inops():

The infix operators from the specified module will be placed in the current environment (like the Global environment, or the environment within a function).

## **Examples**

```
## Not run:
alias %@source% list(file="mydir/mymodule.R")
source_inops(file="mydir/mymodule.R")
## End(Not run)
```

stri\_join\_mat

Concatenate Character Matrix Row-wise or Column-wise

## **Description**

The stri\_join\_mat() function (and their aliases stri\_c\_mat and stri\_paste\_mat) perform rowwise (margin=1; the default) or column-wise (margin=2) joining of a matrix of strings, thereby transforming a matrix of strings into a vector of strings.

## Usage

```
stri_join_mat(mat, margin = 1, sep = "", collapse = NULL)
stri_c_mat(mat, margin = 1, sep = "", collapse = NULL)
stri_paste_mat(mat, margin = 1, sep = "", collapse = NULL)
```

## **Arguments**

mat a matrix of strings

margin the margin over which the strings must be joined.

If margin=1, the elements in each row of matrix mat are joined into a single

string. Thus if the matrix has 10 rows, it returns a vector of 10 strings.

If margin=2, the elements in each column of matrix mat are joined into a single string. Thus if the matrix has 10 columns, it returns a vector of 10 strings.

sep, collapse as in stri\_join.

18 stri\_locate\_ith

#### **Details**

The examples section show the uses of the stri\_join\_mat() function.

#### Value

The stri\_join\_mat() function, and its aliases, return a vector of strings.

#### **Examples**

```
# re-ordering characters in strings ====
x <- c("Hello world", "Goodbye world")</pre>
print(x)
mat <- stringi::stri_split_boundaries(x, simplify = TRUE, type="character")</pre>
rank <- stringi::stri_rank(as.vector(mat)) |> matrix(ncol=ncol(mat))
sorted <- mat %row~% rank
print(sorted)
stri_join_mat(sorted, margin=1)
stri_join_mat(sorted, margin=2)
# re-ordering words ====
x <- c("Hello everyone", "Goodbye everyone")</pre>
print(x)
mat <- stringi::stri_split_boundaries(x, simplify = TRUE, type="word")</pre>
rank <- stringi::stri_rank(as.vector(mat)) |> matrix(ncol=ncol(mat))
sorted <- mat %row~% rank
print(sorted)
stri_c_mat(sorted, margin=1) # <- alias for stri_join_mat</pre>
stri_c_mat(sorted, margin=2)
# re-ordering sentences ====
x <- c("Hello, who are you? Oh, really?! Cool!", "I don't care. But I really don't.")
mat <- stringi::stri_split_boundaries(x, simplify = TRUE, type="sentence")</pre>
rank <- stringi::stri_rank(as.vector(mat)) |> matrix(ncol=ncol(mat))
sorted <- mat %row~% rank
print(sorted)
stri_paste_mat(sorted, margin=1) # <- another alias for stri_join_mat</pre>
stri_paste_mat(sorted, margin=2)
```

stri\_locate\_ith

Locate i^th Pattern Occurrence

## **Description**

The  $stri_locate_ith$  function locates the  $i^{th}$  occurrence of a pattern in each string of some character vector.

stri\_locate\_ith

## Usage

```
stri_locate_ith(str, i, ..., regex, fixed, coll, charclass)
```

## **Arguments**

```
str
                   a string or character vector.
                   a number, or a numeric vector of the same length as str. This gives the i^{th}
i
                   instance to be replaced.
                   Positive numbers are counting from the left. Negative numbers are counting
                   from the right. I.e.:
                   stri_locate_ith(str, i=1, ...) gives the position (range) of the first occur-
                   rence of a pattern.
                   stri_locate_ith(str, i=-1, ...) gives the position (range) of the last oc-
                   currence of a pattern.
                   stri_locate_ith(str, i=2, ...) gives the position (range) of the second oc-
                   currence of a pattern.
                   stri_locate_ith(str, i=-2, ...) gives the position (range) of the second-
                   last occurrence of a pattern.
                   If abs(i) is larger than the number of instances, the first (if i < 0) or last (if i > 0)
                   0) instance will be given.
                   For example: suppose a string has 3 instances of some pattern;
                   then if i \ge 4 the third instance will be located,
                   and if i \le -3 the first instance will be located.
                   more arguments to be supplied to stri_locate and stri_count.
regex, fixed, coll, charclass
                   a character vector of search patterns, as in stri_locate.
```

#### Value

The stri\_locate\_ith() function returns an integer matrix with two columns, giving the start and end positions of the  $i^{th}$  matches, and two NAs if no matches are found.

```
# simple pattern ====

x <- rep(paste0(1:10, collapse=""), 10)
print(x)
out <- stri_locate_ith(x, 1:10, regex = as.character(1:10))
cbind(1:10, out)

x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))
print(x)
p <- rep("a|e|i|o|u",2)
out <- stri_locate_ith(x, c(-1, 1), regex=p)
print(out)
substr(x, out[,1], out[,2])</pre>
```

20 str\_arithmetic

```
# ignore case pattern ====
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
p <- rep("A|E|I|0|U", 2)</pre>
out <- stri_locate_ith(x, c(1, -1), regex=p, case_insensitive=TRUE)</pre>
substr(x, out[,1], out[,2])
# multi-character pattern ====
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
# multi-character pattern:
p \leftarrow rep("AB", 2)
out <- stri_locate_ith(x, c(1, -1), regex=p, case_insensitive=TRUE)</pre>
print(out)
substr(x, out[,1], out[,2])
# Replacement transformation using stringi ====
x <- c("hello world", "goodbye world")</pre>
loc <- stri_locate_ith(x, c(1, -1), regex="a|e|i|o|u")</pre>
extr <- stringi::stri_sub(x, from=loc)</pre>
repl <- chartr(extr, old = "a-zA-Z", new = "A-Za-z")</pre>
stringi::stri_sub_replace(x, loc, replacement=repl)
```

str\_arithmetic

String arithmetic

#### **Description**

String arithmetic operators.

The x %s+%y operator is equivalent to stringi::stri\_c(x,y).

The x %s-% p operator removes character/pattern defined in p from x.

The x %s \*% n operator repeats every element of x for n times, and glues them together.

The x %s/% p operator counts how often regular expression or character pattern p occurs in each element of x.

str\_arithmetic 21

## Usage

```
x %s+% yx %s-% px %s*% nx %s/% p
```

## **Arguments**

| Χ | a string or character vector.   |
|---|---|
| у | a string, or a character vector of the same length as x.  |
| р | the result from s_pattern, or else a character vector of the same length as x with regular expressions. |
| n | a number, or a numeric vector of the same length as x.  |

#### **Details**

These operators and functions serve as a way to provide straight-forward string arithmetic, missing from base R.

#### Value

The %s+%, %s-%, and %s\*% operators return a character vector of the same length as x. The %s/% returns a integer vector of the same length as x.

```
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))</pre>
print(x)
y <- c("a", "b")
p \leftarrow rep("a|e|i|o|u", 2) \# same as p \leftarrow s_pattern(regex=rep("a|e|i|o|u", 2))
n < -c(3, 2)
x %s+% y # =paste0(x,y)
x %s-% p # remove all vowels from x
x %s*% n
x %s/% p \# count how often vowels appear in each string of vector x.
x \leftarrow c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))
print(x)
y <- "a"
# pattern that ignores case:
p \leftarrow s_pattern(regex=rep("A|E|I|0|U", 2), ignore.case=TRUE)
n <- c(2, 3)
x %s+% y # =paste0(x,y)
```

22 str\_subset\_ops

```
x %s-% p # remove all vowels from x x %s+% n x %s/% p # count how often vowels appears in each string of vector x.
```

str\_subset\_ops

String subsetting operators

# Description

String subsetting operators.

The x %ss% s operator allows indexing a single string as-if it is an iterable object.

The x %sget% ss operator gives a certain number of the first and last characters of x.

The x %strim% ss operator removes a certain number of the first and last characters of x.

# Usage

- x %ss% s
- x %sget% ss
- x %strim% ss

# Arguments

- x a string or character vector.
- s a numeric vector giving the subset indices.
- a vector of length 2, or a matrix with 2 columns with nrow(ss)==length(x). The object ss should consist entirely of non-negative integers (thus 0, 1, 2, etc. are valid, but -1, -2, -3 etc are not valid). The first element/column of ss gives the number of characters counting from the left side to be extracted/removed from x. The second element/column of ss gives the number of characters counting

from the right side to be extracted/removed from x.

# Details

These operators serve as a way to provide straight-forward string sub-setting.

#### Value

The %ss% operator always returns a vector or matrix, where each element is a single character.

#### **Examples**

```
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))
print(x)
ss <- c(2,3)
x %sget% ss

x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))
print(x)
ss <- c(1,0)
x %sget% ss

x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))
print(x)
ss <- c(2,3)
x %strim% ss

x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse=""))
print(x)
ss <- c(1,0)
x %strim% ss</pre>
```

substr\_repl

Substr - functions

## **Description**

Fully vectorized sub-string functions.

These functions extract, replace, add-in, transform, or re-arrange, the  $i^{th}$  pattern occurrence or position range.

The  $substr\_repl(x, rp, ...)$  function replaces a position (range) with string rp.

The  $substr_chartr(x, old, new, ...)$  function transforms the sub-string at a position (range) using chartr(old, new).

The  $substr_addin(x, addition, side, ...)$  function adds the additional string addition at the side (specified by argument side) of a position.

The  $substr_extract(x, type, ...)$  function extracts the string at, before, or after some position.

The substr\_arrange(x, arr, ...) function sorts (alphabetically or reverse-alphabetically) or reverses the sub-string at a position (range).

# Usage

```
substr_repl(x, rp, ..., loc = NULL, start = NULL, end = NULL, fish = FALSE)
substr_chartr(
  х,
  old = "a-zA-Z",
  new = ^{\prime\prime}A-Za-z^{\prime\prime},
  . . . ,
  loc = NULL,
  start = NULL,
  end = NULL,
  fish = FALSE
substr_addin(
  х,
  addition,
  side = "after",
  ...,
  loc = NULL,
  at = NULL,
  fish = FALSE
)
substr_extract(
  Х,
  type = "at",
  ...,
  loc = NULL,
  start = NULL,
  end = NULL,
  fish = FALSE
)
substr_arrange(
  Х,
  arr = "incr",
  loc = NULL,
  start = NULL,
  end = NULL,
  opts_collator = NULL,
  fish = FALSE
```

## **Arguments**

a string or character vector.

| rp   | a string, or a character vector of the same length as x, giving the replacing strings.   |  |
|--|--|--|
| • • •  | only applicable if fish=TRUE; other arguments to be passed to the stringfish functions.  |  |
| loc  | The result from the <a href="stri_locate_ith">stri_locate_ith</a> function.  NOTE: you cannot fill in both loc and start, end, or both loc and at. Choose one or the other.  |  |
| start, end   | integers, or integer vectors of the same length as x, giving the start and end position of the range to be modified.   |  |
| fish   | although tinyoperators has no dependencies other than stringi, it does allow<br>the internal functions to use the multi-threadable stringfish functions. To do<br>so, set fish=TRUE; this requires stringfish to be installed. |  |
| old, new   | see chartr. Defaults to old="a-zA-Z", new="A-Za-z", which means upper case characters will be transformed to lower case characters, and vice-versa.  |  |
| addition   | a string, or a character vector of the same length as x, giving the string(s) to add-in.   |  |
| side   | which side of the position to add in the string. Either "before" or "after".   |  |
| at   | an integer, or integer vector of the same length as $x$ , giving the position after or before which the string is to be added.   |  |
| type   | a single string, giving the part of the string to extract. 3 options available:  |  |
| • type = "at": extracts the string part at the position range; |  |  |
|  | • type = "before": extracts the string part before the position range;   |  |
|  | • type = "after": extracts the string part after the position range.   |  |
| arr  | a single string, giving how the sub-string should be arranged. 3 options available:  |  |
|  | • arr = "incr": sort the sub-string alphabetically.  |  |

- arr = "decr": sort the sub-string reverse alphabetically.
- arr = "rev": reverse the sub-string.
- arr = "rand": randomly shuffles the sub-string.

opts\_collator as in stri\_rank. Only used when arr = "incr" or arr = "decr".

## **Details**

These functions serve as a way to provide straight-forward sub-string modification and/or extraction.

All substr\_ functions internally only use fully vectorized R functions (no loops or apply-like functions).

# Value

A modified character vector. If no match is found in a certain string of character vector x, the unmodified string is returned. The exception is for the substr\_extract() function: in this function, non-matches return NA.

```
# numerical substr ====
x \leftarrow rep("12345678910", 2)
start=c(1, 2); end=c(3,4)
substr_extract(x, start=start, end=end)
substr_extract(x, type="before", start=start, end=end)
substr_extract(x, type="after", start=start, end=end)
substr_repl(x, c("??", "!!"), start=start, end=end)
substr_chartr(x, start=start, end=end)
substr_addin(x, c("", "~"), "after", at=end)
substr_addin(x, c(" ", "~"), "before", at=start)
substr_arrange(x, start=start, end=end)
substr_arrange(x, "decr", start=start, end=end)
substr_arrange(x, "rev", start=start, end=end)
substr_arrange(x, "rand", start=start, end=end)
start=10; end=11
substr_extract(x, start=start, end=end)
substr\_extract(x, type="before", start=start, end=end)
substr_extract(x, type="after", start=start, end=end)
substr_repl(x, "??", start=start, end=end)
substr_chartr(x, start=start, end=end)
substr_addin(x, " ", "after", at=end)
substr_addin(x, " ", "before", at=start)
start=5; end=6
substr_extract(x, start=start, end=end)
substr_extract(x, type="before", start=start, end=end)
substr_extract(x, type="after", start=start, end=end)
substr_repl(x, "??", start=start, end=end)
substr_chartr(x, start=start, end=end)
substr_addin(x, " ", "after", at=end)
substr_addin(x, " ", "before", at=start)
# simple pattern ====
x <- c("goodGOODGoodgOOd", "goodGOODGoodgOOd", paste0(letters[1:13], collapse=""))</pre>
print(x)
loc <- stri_locate_ith(</pre>
 # locate second-last occurrence of "good" of each string in x:
 x, -2, regex="good", case_insensitive=TRUE
substr_extract(x, loc=loc) # extract second-last "good"
substr_repl(x, c("??", "!!", " "), loc=loc) # replace second-last "good"
substr_chartr(x, loc=loc) # switch upper/lower case of second-last "good"
substr\_addin(x,\ c("\ ",\ "^{-}",\ "\ "),\ "after",\ loc=loc)\ \#\ add\ white\ space\ after\ second-last\ "good"
substr_addin(x, c("", "~", ""), "before", loc=loc) \# add white space before second-last "good" 
substr_arrange(x, loc=loc) # sort second-last "good"
substr_arrange(x, "decr", loc=loc) # reverse-sort second-last "good"
substr_arrange(x, "rev", loc=loc) # reverse second-last "good"
substr_arrange(x, "rand", loc=loc) # randomly shuffles "good"
```

s\_pattern 27

s\_pattern

Pattern attribute assignment

## **Description**

The %s-% and %s/% operators, their in-place equivalents, as well as the %sgrep% operator, all perform pattern matching for some purpose. By default the pattern matching is interpreted as case-sensitive regex patterns from stringi.

The s\_pattern function allows the user to specify exactly how the pattern should be interpreted. To use more refined pattern definition, simply replace the right-hand-side expression p in the relevant operators with a call from the s\_pattern() function.

The s\_pattern() function uses the exact same argument convention as stringi. For example:

```
• s_pattern(regex=p, case_insensitive=FALSE, ...)
```

- s\_pattern(fixed=p, ...)
- s\_pattern(coll=p, ...)
- s\_pattern(charclass=p, ...)

All arguments in s\_pattern() are simply passed to the appropriate functions in stringi. For example:

x %s/% p counts how often regular expression p occurs in x,

whereas x %s/% s\_pattern(fixed=p, case\_insensitive=TRUE) will do the same, except it uses fixed (i.e. literal) expression, and it does not distinguish between upper case and lower case characters

For consistency with base R, one can also fill in ignore.case=TRUE or ignore\_case=TRUE instead of case\_insensitive=TRUE, and s\_pattern will still understand that.

## Usage

```
s_pattern(...)
```

## **Arguments**

pass stringi arguments here. I.e. regex=p, coll=p, charclass=p, case\_insensitive=FALSE, etc. See the documentation in the stringi R package.

#### **Details**

The s\_pattern() function only works in combination with the functions and operators in this package. It does not affect functions from base R or functions from other packages.

#### Value

The s\_pattern(...) call returns a list with arguments that will be passed to the appropriate functions in stringi.

28 tinyoperators\_help

#### **Examples**

```
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse="")) print(x) p <- rep("a|e|i|o|u", 2) # same as p <- s_pattern(regex=rep("a|e|i|o|u", 2)) x %s/% p # count how often vowels appear in each string of vector x.   
x <- c(paste0(letters[1:13], collapse=""), paste0(letters[14:26], collapse="")) print(x) p <- s_pattern(regex=rep("A|E|I|O|U", 2), case_insensitive=TRUE) x %s/% p # count how often vowels appear in each string of vector x.
```

tinyoperators\_help

The tinyoperators help page

### **Description**

Welcome to the tinyoperators help page!

The tinyoperators R-package adds some infix operators, and a few functions. It primarily focuses on 4 things:

- (1) Float truth testing.
- (2) Reducing repetitive code.
- (3) Extending the string manipulation capabilities of the stringi R package.
- (4) A new package and module import system, that combines the benefits of aliasing a package with the benefits of attaching a package.

The tinyoperators R-package has only one dependency, namely stringi, though it does allows multi-threading of some of the string-related functions (when appropriate) via the suggested stringfish R-package.

The tinyoperators R package adds the following functionality:

- Infix logical operators for exclusive-or, not-and, not-in, number-type, and string-type.
- Safer (in)equality operators for floating numbers.
- Infix operators for string arithmetic.
- Infix operators for string sub-setting.
- Several operators for the "Don't Repeat Yourself" coding principle (DRY). This includes the generalized in-place (mathematical) modification operator, infix operators for In-place modifying string arithmetic, and infix operators for In-place modifying string sub-setting.

transform\_if 29

- Infix operators for row- and column-wise re-ordering of matrices.
- The tinyoperators package adds additional stringi functions, namely stri\_locate\_ith and stri\_join\_mat (and aliases). These functions use the same naming and argument convention as the rest of the stringi functions, thus keeping your code consistent.
- The fully vectorized sub-string functions, that extract, replace, add-in, transform, or re-arrange, the  $i^{th}$  pattern occurrence or location.
- The s\_pattern helper function for string operators.
- New package import functions, and new module sourcing functions.
- Most stringi pattern expressions options are available for the string-pattern-related functions, when appropriate.
- This R package has only one dependency: stringi. No other dependencies, as to avoid "dependency hell".
- Although this package has no other dependencies, it allows multi-threading of the sub-string functions through the stringfish R package.

Please also have a look at the Read-Me file on the GitHub main page before using this package: https://github.com/tony-aw/tinyoperators

#### Usage

tinyoperators\_help()

transform if

The transform\_if function and the subset\_if operators

## **Description**

Consider the following code:

ifelse(cond(x), f(x), g(x))

Here a conditional subset of the object x is transformed, where the condition is using a function referring to x itself. Consequently, reference to x is written four times!

The tinyoperators package therefore adds the transform\_if() function which will tiny this up.

The tinyoperators package also adds 2 "subset\_if" operators:

The x %[if]% cond operator selects elements from vector/matrix/array x, for which the result of cond(x) returns TRUE.

The x %[!if]% cond operator selects elements from vector/matrix/array x, for which the result of cond(x) returns FALSE.

The tinyoperators package also adds the x %unreal =% repl operator: x %unreal = % repl is the same as x[is.na(x)|is.na(x)|is.infinite(x)] <- repl

30 transform\_if

#### Usage

```
transform_if(x, cond, trans_T = function(x) x, trans_F = function(x) x)
x %[if]% cond
x %[!if]% cond
x %unreal =% repl
```

#### **Arguments**

x a vector, matrix, or array.

cond a function that returns a binary logic (TRUE, FALSE) vector of the same length/dimensions

as x (for example: is.na).

trans\_T the transformation function to use when cond(x) == TRUE.

For example: log.

If this is not specified,  $trans_T$  defaults to function(x)x.

trans\_F the transformation function to use when cond(x) == FALSE.

For example: log.

If this is not specified, trans\_F defaults to function(x)x.

repl the replacement value.

#### **Details**

The transform\_if(x, cond, trans) function does not rely on any explicit or implicit loops, nor any third-party functions.

#### Value

```
For transform_if():
```

Similar to ifelse. However, unlike ifelse(), the transformations are evaluated as trans\_T(x[cond(x)]) and trans\_F(x[!cond(x)]), ensuring no unnecessary warnings or errors occur.

The subset\_if - operators all return a vector with the selected elements.

The x %unreal =% repl operator does not return any value:

It is an in-place modifiers, and thus modifies x directly. The object x is modified such that all NA, NaN and Inf elements are replaced with repl.

```
object_with_very_long_name <- matrix(-10:9, ncol=2)
print(object_with_very_long_name)
object_with_very_long_name |> transform_if(\(x)x>0, log)
object_with_very_long_name |> transform_if(\(x)x>0, log, \(x)x^2)
object_with_very_long_name %[if]% \(x)x %in% 1:10
object_with_very_long_name %[!if]% \(x)x %in% 1:10

x <- c(1:9, NA, NaN, Inf)
print(x)
x %unreal =% 0 # same as x[is.na(x)|is.nan(x)|is.infinite(x)] <- 0
print(x)</pre>
```

# Index

| * join_mat                                  | import, 4   |
|---|---|
| stri_join_mat, 17                           | <pre>import_as(import), 4</pre>                   |
| %:=% (inplace), 7                           | <pre>import_data(import), 4</pre>                 |
| %=numtype% (logic_ops), 11                  | <pre>import_inops (import), 4</pre>               |
| %=strtype%(logic_ops), 11                   | <pre>import_lsf(import), 4</pre>                  |
| %?=% (logic_ops), 11                        | In-place modifying string arithmetic,             |
| %[!if]%(transform_if), 29                   | 28  |
| %[if]%(transform_if), 29                    | <pre>In-place modifying string sub-setting,</pre> |
| %col~% (matrix_ops), 13                     | 28  |
| %f!=%(float_truth_testing), 2               | Infix logical operators, 28                       |
| %f<=% (float_truth_testing), 2              | Infix operators for row- and                      |
| %f<%(float_truth_testing), 2                | column-wise re-ordering of                        |
| %f==% (float_truth_testing), 2              | matrices, 29                                      |
| %f>=% (float_truth_testing), 2              | inplace, 7  |
| %f>% (float_truth_testing), 2               | <pre>inplace_str_arithmetic, 8</pre>              |
| %installed in% (pkgs_get_deps), 15          | inplace_str_subset, 9                             |
| %n&% (logic_ops), 11                        | [   |
| %out% (logic_ops), 11                       | loadNamespace, 5, 15                              |
| %row~% (matrix_ops), 13                     | Logic, <i>11</i>                                  |
| %s* =% (inplace_str_arithmetic), 8          | logic_ops, 11                                     |
| %s*% (str_arithmetic), 20                   |   |
| %s+ =% (inplace_str_arithmetic), 8          | matrix_ops, 13                                    |
| %s+% (str_arithmetic), 20                   | new module sourcing functions, 29                 |
| %s- =% (inplace_str_arithmetic), 8          | New package import functions, 29                  |
| %s-% (str_arithmetic), 20                   | New package import runctions, 27                  |
| %s/ =% (inplace_str_arithmetic), 8          | pkgs_get_deps, 5, 15                              |
| %s/% (str_arithmetic), 20                   |   |
| %sget =%(inplace_str_subset),9              | s_pattern, 8, 11, 21, 27, 29                      |
| %sget% (str_subset_ops), 22                 | Safer (in)equality operators for                  |
| %sgrep%(logic_ops), 11                      | floating numbers, $28$                            |
| %ss% (str_subset_ops), 22                   | source, <i>16</i>                                 |
| <pre>%strim =%(inplace_str_subset), 9</pre> | <pre>source_inops (source_module), 16</pre>       |
| %strim% (str_subset_ops), 22                | source_module, 16                                 |
| %unreal =% (transform_if), 29               | str_arithmetic, 20                                |
| %xor% (logic_ops), 11                       | str_subset_ops, 22                                |
|   | stri_c_mat(stri_join_mat), 17                     |
| chartr, 25                                  | stri_count, 19                                    |
|   | stri_join, <i>17</i>                              |
| <pre>float_truth_testing, 2</pre>           | stri_join_mat, 17, 29                             |
|   | stri_locate, <i>19</i>                            |
| generalized in-place (mathematical)         | stri_locate_ith, 18, 25, 29                       |
| modification operator, $28$                 | <pre>stri_paste_mat(stri_join_mat), 17</pre>      |
|   | stri_rank, 25                                     |
| ifelse, <i>30</i>                           | string arithmetic, $8,28$                         |

32 INDEX

```
string sub-setting, 28
string subset, 10
sub-string functions, 29
substr_addin(substr_repl), 23
substr_arrange(substr_repl), 23
substr_chartr(substr_repl), 23
substr_extract(substr_repl), 23
substr_repl, 23
tinyoperators_help, 28
transform_if, 29
```