# Package 'tinycodet'

February 11, 2024

| 1 0010011 11, 202.   |
|--|
| Title Functions to Help in your Coding Etiquette   |
| Version 0.4.1  |
| <ul> <li>Description Adds some functions to help in your coding etiquette.     'tinycodet' primarily focuses on 4 aspects.     1) Safer decimal (in)equality testing, safer atomic conversions, and other functions for safer coding.     2) A new package import system,     that attempts to combine the benefits of using a package without attaching it,     with the benefits of attaching a package.     3) Extending the string manipulation capabilities of the 'stringi' R package.     4) Reducing repetitive code.     Besides linking to 'Rcpp', 'tinycodet' has only one other dependency, namely 'stringi'.</li> </ul> |
| License MIT + file LICENSE   |
| Encoding UTF-8   |
| LinkingTo Rcpp   |
| <b>Roxygen</b> list(markdown = TRUE)   |
| RoxygenNote 7.2.3  |
| Suggests knitr, rmarkdown, tinytest, roxygen2, collapse, kit   |
| <b>Depends</b> R (>= 4.1.0)  |
| Imports Rcpp (>= 1.0.11),<br>stringi (>= 1.7.12)   |
| <pre>URL https://github.com/tony-aw/tinycodet/, https://tony-aw.github.io/tinycodet/</pre>   |
| <pre>BugReports https://github.com/tony-aw/tinycodet/issues/ Language en-gb</pre>  |
| R topics documented:   |
| aaa0_tinycodet_help  |

2 aaa0\_tinycodet\_help

| Index |                       | 62       |
|-------|-----------------------|----------|
|       | x.import              | 59       |
|       | transform_if          | 57       |
|       | s_pattern             | 54       |
|       | subset_if             | 53       |
|       | str_subset_ops        | 52       |
|       | str_search            | 46       |
|       | str_arithmetic        | 44       |
|       | stri_locate_ith       | 39       |
|       | stri_join_mat         | 37       |
|       | strcut_loc            | 35       |
|       | source_selection      | 34       |
|       | report_inops          | 33       |
|       | pversion              | 31       |
|       | pkgs                  | 29       |
|       | matrix_ops            | 27       |
|       | logic ops             | 25       |
|       | lock                  | 23       |
|       | inplace               | 22       |
|       | import_Integration LL | 20       |
|       | import_inops.control  | 19       |
|       | import inops          | 17       |
|       | import_as             | 13<br>16 |
|       | decimal_truth         | 11       |
|       | atomic_conversions    | 9        |
|       | aaa5_tinycodet_misc   | 9        |
|       | aaa4_tinycodet_dry    | 8        |
|       |                       |          |

aaa0\_tinycodet\_help tinycodet: Functions to Help in your Coding Etiquette

# Description

Welcome to the 'tinycodet' introduction help page!

'tinycodet' adds some functions to help in your coding etiquette. It primarily focuses on 4 aspects:

- (1) Safer decimal (in)equality testing, safer atomic conversions, and other functions for safer coding;
- see tinycodet\_safer.
- (2) A new package import system, that attempts to combine the benefits of using a package without attaching it, with the benefits of attaching a package; see tinycodet\_import
- (3) Extending the string manipulation capabilities of the 'stringi' R-package; see tinycodet\_strings.

aaa1\_tinycodet\_safer 3

```
(4) Reducing repetitive code; see tinycodet_dry.
```

And some miscellaneous functionality; see tinycodet\_misc.

Please check the Change-log (see links below) regularly for updates (such as bug fixes).

'tinycodet' adheres to the tinyverse philosophy. Besides linking to 'Rcpp', 'tinycodet' only has one other dependency: 'stingi'. No other dependencies, thus avoiding "dependency hell". Most functions in this R-package are vectorized and optimised.

#### Author(s)

**Maintainer**: Tony Wilkes <tony\_a\_wilkes@outlook.com> (ORCID)

### References

The badges shown in the documentation of this R-package were made using the services of: https://shields.io/

#### See Also

Useful links:

- 'tinycodet' GitHub main page and Read-Me: https://github.com/tony-aw/tinycodet/
- 'tinycodet' package website: https://tony-aw.github.io/tinycodet/
- Report bugs at: https://github.com/tony-aw/tinycodet/issues/
- Changelog: https://github.com/tony-aw/tinycodet/blob/main/NEWS.md/or https://tony-aw.github.io/tinycodet/news/index.html
- The 'fastverse', which is related to the 'tinyverse': https://github.com/fastverse/fastverse/

# Description

To help make your code safer, the 'tinycodet' R-package introduces a few functions:

- Safer decimal (in)equality testing.
- Atomic type casting without stripping attributes.
- The lock\_TF function to set and lock T and F to TRUE and FALSE.
- The %<-c% operator to assign locked constants.

### See Also

tinycodet\_help

#### **Examples**

```
x <- c(0.3, 0.6, 0.7)
y <- c(0.1*3, 0.1*6, 0.1*7)
x == y # gives FALSE, but should be TRUE
x %d==% y # here it's done correctly</pre>
```

```
aaa2_tinycodet_import Overview of the 'tinycodet' Import System
```

# **Description**

The 'tinycodet' R-package introduces a new package import system.

One can **use** a package **without attaching** the package - for example by using the :: operator. Or, one can explicitly **attach** a package - for example by using the **library** function. The advantages and disadvantages of **using without attaching** a package versus **attaching** a package, at least those relevant here, are compactly presented in the following list:

```
(1) Prevent masking functions from other packages:
use without attach: Yes(advantage); attaching: No(disadvantage);
(2) Prevent masking core R functions:
use without attach: Yes(advantage); attaching: No(disadvantage);
(3) Clarify which function came from which package:
use without attach: Yes(advantage); attaching: No(disadvantage);
(4) Enable functions only in current/local environment instead of globally:
use without attach: Yes(advantage); attaching: No(disadvantage);
(5) Prevent namespace pollution:
use without attach: Yes(advantage); attaching: No(disadvantage);
(6) Minimise typing - especially for infix operators
(i.e. typing package:: `%op%` (x, y) instead of x %op% y is cumbersome):
use without attach: No(disadvantage); attaching: Yes(advantage);
(7) Use multiple related packages, without constantly switching between package prefixes
(i.e. doing packagename1::some_function1();
packagename2::some_function2();
packagename3::some_function3() is chaotic and cumbersome):
use without attach: No(disadvantage); attaching: Yes(advantage);
```

What 'tinycodet' attempts to do with its import system, is to somewhat find the best of both worlds. It does this by introducing the following functions:

- import\_as: Import a main package, and optionally its re-exports + its direct dependencies + its direct extensions, under a single alias. This essentially combines the attaching advantage of using multiple related packages (item 7 on the list), whilst keeping most advantages of using without attaching a package.
- import\_inops: Expose infix operators from a package or an alias object to the current environment. This gains the attaching advantage of less typing (item 6 on the list), whilst simultaneously avoiding the disadvantage of attaching functions from a package globally (item 4 on the list).
- import\_data: Directly return a data set from a package, to allow straight-forward assignment.

Furthermore, there are two miscellaneous import\_ - functions: import\_LL and import\_int.

The import system also includes general helper functions:

- The x.import functions: helper functions specifically for the 'tinycodet' import system.
- The pversion\_ functions: check mismatch between loaded package version and package version in library path.
- The pkgs functions: general helper functions regarding packages.

See the examples section below to get an idea of how the 'tinycodet' import system works in practice. More examples can be found on the website (https://tony-aw.github.io/tinycodet/)

### **Details**

### When to Use or Not to Use the 'tinycodet' Import System

The 'tinycodet' import system is helpful particularly for packages that have at least one of the following properties:

- The namespace of the package(s) conflicts with other packages.
- The namespace of the package(s) conflicts with core R, or with those of recommended R packages.
- The package(s) have function names that are generic enough, such that it is not obvious which function came from which package.

See examples below.

There is no necessity for using the 'tinycodet' import system with every single package. One can safely attach the 'stringi' package, for example, as 'stringi' uses a unique and immediately recognisable naming scheme (virtually all 'stringi' functions start with "stri\_"), and this naming scheme does not conflict with core R, nor with most other packages.

Of course, if one wishes to use a package (like 'stringi') **only** within a specific environment, like only inside a function, it becomes advantageous to still import the package using the 'tinycodet' import system. In that case the import\_LL function would be most applicable.

#### Some Additional Comments on the 'tinycodet' Import System

- (S3) Methods will automatically be registered.
- Pronouns, such as the .data and .env pronouns from the 'rlang' package, will work without any prefixes required.
- All functions imported by the import\_as, import\_inops, or import\_LL functions have a "package" attribute, so you will always know which function came from which package.

#### See Also

tinycodet\_help

```
all(c("dplyr", "powerjoin", "magrittr") %installed in% .libPaths())
\# NO packages are being attached in any of the following code
# import 'dplyr' + its re-exports + extension 'powerjoin', under alias "dpr.":
import_as(
  ~ dpr., "dplyr", re_exports = TRUE, extensions = "powerjoin"
# exposing infix operators from 'magrrittr' to current environment:
import_inops("magrittr")
# directly assigning dplyr's "starwars" dataset to object "d":
d <- import_data("dplyr", "starwars")</pre>
# See it in Action:
d %>% dpr.$filter(species == "Droid") %>%
  dpr.$select(name, dpr.$ends_with("color"))
male_penguins <- dpr.$tribble(</pre>
  ~name, ~species, ~island, ~flipper_length_mm, ~body_mass_g,
  "Giordan", "Gentoo", "Biscoe", 222L, 5250L, "Lynden", "Adelie", "Torgersen", 190L, 3900L,
  "Lynden",
  "Reiner", "Adelie",
                           "Dream",
                                                  185L,
                                                               3650L
female_penguins <- dpr.$tribble(</pre>
  ~name, ~species, ~island, ~flipper_length_mm, ~body_mass_g,
  "Alonda", "Gentoo", "Biscoe",
"Ola", "Adelie", "Dream",
            "Gentoo", "Biscoe", 211, 4500L,
                                                          3600L,
                                             190,
                                       215,
  "Mishayla", "Gentoo", "Biscoe",
                                                        4750L,
dpr.$check_specs()
dpr.$power_inner_join(
 male_penguins[c("species", "island")],
  female_penguins[c("species", "island")]
)
mypaste <- function(x, y) {</pre>
  import_LL("stringi", selection = "stri_c")
  stringi::stri_c(x, y)
mypaste("hello ", "world")
```

aaa3\_tinycodet\_strings

Overview of the 'tinycodet' Extension of 'stringi'

# **Description**

Virtually every programming language, even those primarily focused on mathematics, will at some point have to deal with strings. R's atomic classes boil down to some form of either numbers or characters. R's numerical functions are generally very fast. But R's native string functions are somewhat slow, do not have a unified naming scheme, and are not as comprehensive as R's impressive numerical functions. The primary R-package that fixes this is 'stringi'.

'stringi' is arguably the fastest and most comprehensive string manipulation package available at the time of writing. Many string related packages fully depend on 'stringi' (see its reverse-dependencies on CRAN).

As string manipulation is so important to programming languages, 'tinycodet' adds a little bit new functionality to 'stringi'.

'tinycodet' adds the following functions to extend 'stringi':

- Find  $i^{th}$  pattern occurrence (stri locate ith), or  $i^{th}$  text boundary (stri locate ith boundaries).
- Concatenate a character matrix row- or column-wise.

'tinycodet' adds the following operators, to complement the already existing 'stringi' operators:

- Infix operators for string arithmetic.
- Infix operators for string sub-setting, which get or remove the first and/or last n characters from strings.
- Infix operators for detecting patterns, and strfind()<- for locating/extracting/replacing found patterns.

And finally, 'tinycodet' adds the somewhat separate strcut\_-functions, to cut strings into pieces without removing the delimiters.

# Regarding Vector Recycling in the 'stringi'-based Functions

Generally speaking, vector recycling is supported as 'stringi' itself supports it also.

There are, however, a few exceptions.

First, matrix inputs (like in strcut\_loc and string sub-setting operators) will generally not be recycled.

Second, the i argument in stri\_locate\_ith does not support vector recycling.

Scalar recycling is virtually always supported.

8 aaa4\_tinycodet\_dry

#### References

Gagolewski M., **stringi**: Fast and portable character string processing in R, *Journal of Statistical Software* 103(2), 2022, 1–59, doi:10.18637/jss.v103.i02

### See Also

```
tinycodet_help, s_pattern
```

### **Examples**

```
# character vector:
x <- c("3rd 1st 2nd", "5th 4th 6th")
print(x)
# detect if there are digits:
x %s{}% "\\d"
# find second last digit:
loc <- stri_locate_ith(x, i = -2, regex = "\\d")</pre>
stringi::stri_sub(x, from = loc)
# cut x into matrix of individual words:
mat <- strcut_brk(x, "word")</pre>
# sort rows of matrix using the fast %row~% operator:
rank <- stringi::stri_rank(as.vector(mat)) |> matrix(ncol = ncol(mat))
sorted <- mat %row~% rank</pre>
sorted[is.na(sorted)] <- ""</pre>
# join elements of every row into a single character vector:
stri_c_mat(sorted, margin = 1, sep = " ")
```

aaa4\_tinycodet\_dry

Overview of the 'tinycodet' "Don't Repeat Yourself" Functionality

# Description

"Don't Repeat Yourself", sometimes abbreviated as "DRY", is the coding principle not to write unnecessarily repetitive code. To help in that effort, the 'tinycodet' R-package introduces a few functions:

- The transform\_if function
- The subset\_if operators and the in-place unreal modifier operator.
- The general in-place (mathematical) modification operator.

# See Also

```
tinycodet_help
```

aaa5\_tinycodet\_misc 9

#### **Examples**

```
object <- matrix(c(-9:8, NA, NA) , ncol=2)

# in base R:
ifelse( # repetitive, and gives unnecessary warning
    is.na(object > 0), -Inf,
    ifelse(
        object > 0, log(object), object^2
    )
)
mtcars$mpg[mtcars$cyl>6] <- (mtcars$mpg[mtcars$cyl>6])^2 # long

# with tinycodet:
object |> transform_if(\(\(x\) x > 0, log, \(x\) x^2, \(x\) -Inf) # compact & no warning
mtcars$mpg[mtcars$cyl > 6] %:=% \(x\) x^2 # short
```

aaa5\_tinycodet\_misc

Overview of the 'tinycodet' Miscellaneous Functionality

### **Description**

Some additional functions provided by the 'tinycodet' R-package:

- Infix logical operators for exclusive-or, not-and, not-in, number-type, and string-type.
- Infix operators for row- and column-wise re-ordering of matrices.
- Report infix operators present in the current environment, or a specified environment.
- source\_selection to source only selected objects.

### See Also

```
tinycodet_help()
```

atomic\_conversions

Atomic Type Casting Without Stripping Attributes

# Description

Atomic type casting in R is generally performed using the functions as.logical, as.integer, as.double, as.character.

Converting an object between atomic types using these functions strips the object of its attributes, including attributes such as names and dimensions.

The functions provided here by the 'tinycodet' package preserve all attributes - except the "class" attribute.

The functions are as follows:

10 atomic\_conversions

- as\_bool(): converts object to atomic type logical (TRUE, FALSE, NA).
- as\_int(): converts object to atomic type integer.
- as\_dbl(): converts object to atomic type double (AKA decimal numbers).
- as\_chr(): converts object to atomic type character.

Moreover, the function is\_wholenumber() is added, to safely test for whole numbers.

# Usage

```
as_bool(x, ...)
as_int(x, ...)
as_dbl(x, ...)
as_chr(x, ...)
is_wholenumber(x, tol = sqrt(.Machine$double.eps))
```

### **Arguments**

- x vector, matrix, array (or a similar object where all elements share the same type).
- ... further arguments passed to or from other methods.
- tol numeric, giving the tolerance.

#### Value

The converted object.

### See Also

tinycodet\_safer

```
x <- c(rep(0, 2), seq(0, 2.5, by=0.5)) |> matrix(ncol=2)
colnames(x) <- c("one", "two")
attr(x, "test") <- "test"
print(x)

# notice that in all following, attributes (except class) are conserved:
as_bool(x)
as_int(x)
as_dbl(x)
as_chr(x)

# is_wholenumber:
is_wholenumber(1:10 + c(0, 0.1))</pre>
```

decimal\_truth 11

decimal\_truth

Safer Decimal Number (In)Equality Testing Operators

### **Description**

The %d==%, %d!=% %d<%, %d>%, %d<=%, %d>=% (in)equality operators perform decimal (class "double") number truth testing.

They are virtually equivalent to the regular (in)equality operators,

```
==, !=, <, >, <=, >=,
```

except for 2 aspects:

1. The decimal number (in)equality operators assume that if the absolute difference between any 2 numbers x and y is smaller than the Machine tolerance, sqrt(.Machine\$double.eps), then x and y should be consider to be equal.

For example: 0.1\*7 == 0.7 returns FALSE, even though they are equal, due to the way decimal numbers are stored in programming languages like 'R' and 'Python'.

But 0.1\*7 %d==% 0.7 returns TRUE.

2. Only numeric input is allowed, so characters are not coerced to numbers.

I.e. 1 < "a" gives TRUE, whereas 1 %d<% "a" gives an error.

For character equality testing, see %s==% from the 'stringi' package.

Thus these operators provide safer decimal number (in)equality tests.

There are also the x %d{}% bnd and x %d!{}% 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 %d{}% bnd operator checks if x is within the closed interval with bounds defined by bnd.

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

### Usage

- x %d==% y
- x %d!=% y
- x %d<% y
- x %d>% y
- x %d<=% y
- x %d>=% y
- x %d{}% bnd
- x %d!{}% bnd

# **Arguments**

х, у numeric vectors, matrices, or arrays. 12 decimal\_truth

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) (or can be recycled to be 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.

#### Value

A logical vector with the same dimensions as x, indicating the result of the element by element comparison.

### See Also

tinycodet\_safer

```
x \leftarrow c(0.3, 0.6, 0.7)
y \leftarrow 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 %d==% y # here it's done correctly
x %d!=% y # correct
x %d<% y # correct
x %d>% y # correct
x %d<=% y # correct
x %d>=% y # correct
x < -c(0.3, 0.6, 0.7)
bnd <- cbind(x-0.1, x+0.1)
x %d{}% bnd
x %d!{}% bnd
# These operators work for integers also:
x <- 1L:5L
y <- 1L:5L
x %d==% y
x %d!=% y
x %d<% y
x %d>% y
x %d<=% y
x %d>=% y
x <- 1L:5L
y <- x+1
x %d==% y
x %d!=% y
x %d<% y
x %d>% y
x %d<=% y
x %d>=% y
```

import\_as 13

```
x <- 1L:5L
y <- x-1
x %d==% y
x %d!=% y
x %d<% y
x %d>% y
x %d<=% y
x %d>=% y
```

import\_as

Import R-package, its Re-exports, Dependencies, and/or Extensions, Under a Single Alias

### **Description**

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

### Usage

```
import_as(
  alias,
  main_package,
  re_exports = TRUE,
  dependencies = NULL,
  extensions = NULL,
  lib.loc = .libPaths(),
  import_order = c("dependencies", "main_package", "extensions")
)
```

# **Arguments**

alias

a syntactically valid non-hidden name giving the alias object where the package(s) are to be imported into.

This name can be given either as a single string (i.e. "alias."), or as a one-sided formula with a single term (i.e. ~alias.).

main\_package

a single string, giving the name of the main package to import under the given alias.

re\_exports

TRUE or FALSE.

- If re\_exports = TRUE the re-exports from the main\_package are added to the alias together with the main package. This is the default, as it is analogous to the behaviour of base R's :: operator.
- If re\_exports = FALSE, these re-exports are not added together with the main package. The user can still import the packages under the alias from which the re-exported functions came from, by specifying them in the dependencies argument.

14 import\_as

dependencies an optional character vector, giving the names of the dependencies of the main\_package

to be imported also under the alias.

Defaults to NULL, which means no dependencies are imported under the alias.

See pkg\_get\_deps to quickly get dependencies from a package.

extensions an optional character vector, giving the names of the extensions of the main\_package

to be imported also under the alias.

Defaults to NULL, which means no extensions are imported under the alias.

lib.loc character vector specifying library search path (the location of R library trees to

search through).

The lib.loc argument would usually be .libPaths().

See also loadNamespace.

import\_order the character vector

c("dependencies", "main\_package", "extensions"),

or some re-ordering of this character vector, giving the relative import order of

the groups of packages.

See Details section for more information.

#### **Details**

### **Expanded Definitions of Some Arguments**

- "Re-exports" are functions that are defined in the dependencies of the main\_package, but are re-exported in the namespace of the main\_package.
- "Dependencies" are here defined as any R-package appearing in the "Depends", "Imports", or "LinkingTo" fields of the Description file of the main\_package. So no recursive dependencies.
- "Extensions" are reverse-dependencies that actually extend the functionality of the main\_package. Programmatically, some package "E" is considered an extension of some "main\_package", if the following is TRUE:

"main\_package" %in% pkg\_get\_deps\_minimal("E")

# Why Aliasing Multiple Packages is Useful

To use an R-package with its extension packages or dependencies, whilst avoiding the disadvantages of attaching a package (see tinycodet\_import), one would traditionally use the :: operator like so:

```
main_package::some_function1()
dependency1::some_function2()
extension1::some_function3()
```

This becomes cumbersome as more packages are needed and/or as the package name(s) become longer.

The import\_as() function avoids this issue by allowing multiple **related** packages to be imported under a single alias, allowing one to code like this:

import\_as 15

```
import_as(
    ~ alias., "main_package",
    dependencies = "dependency1", extensions = "extension1",
    lib.loc = .libPaths()
)
alias.$some_function1()
alias.$some_function2()
alias.$some_function3()
```

Thus importing a package, or multiple directly related packages, under a single alias, which import\_as() provides, avoids the above issues. Importing a package under an alias is referred to as "aliasing" a package.

### **Alias Naming Recommendation**

To keep package alias object names easily distinguishable from other objects that can also be subset with the \$ operator, I recommend ending (not starting!) all alias names with a dot (.) or underscore (\_).

### Regarding import\_order

The order of the character vector given in the dependencies and extensions arguments 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.

```
The import_order argument defaults to the character vector c("dependencies", "main_package", "extensions"), which is the recommended setting.

This setting results in the following importing orders:
```

This setting results in the following importing order:

- 1. The dependencies, in the order specified by the dependencies argument.
- 2. The main\_package (see argument main\_package), including re-exports (if re\_exports = TRUE).
- 3. The extensions, in the order specified by the extensions argument.

#### **Other Details**

The import\_as() function does not support importing base/core R under an alias.

Packages that appear in the "Suggests" or "Enhances" fields of packages are not considered dependencies or extensions.

No more than 10 packages are allowed to be imported under a single alias.

### Value

A locked environment object, similar to the output of loadNamespace, with the name as specified in the alias argument, will be created.

This object, referred to as the "(package) alias object", will contain the exported functions from the specified package(s).

16 import\_data

The alias object will be placed in the current environment (like the global environment, or the environment within a function).

```
To use, for example, function "some_function()" from alias "alias.", use:
alias.$some_function()
To see the special attributes of this alias object, use attr.import.
To "unimport" the package alias object, simply remove it (i.e. rm(list = "alias.")).
```

### See Also

tinycodet\_import

# **Examples**

import\_data

Directly Return a Data-set From a Package

# **Description**

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

### Usage

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

# Arguments

package a single string, giving the name of the R-package.

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

lib.loc character vector specifying library search path (the location of R library trees to search through).

The lib.loc argument would usually be .libPaths().

See also loadNamespace.

import\_inops 17

### Value

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

#### See Also

```
tinycodet_import
```

### **Examples**

```
d <- import_data("datasets", "cars")
head(d)</pre>
```

import\_inops

(Un)Expose Infix Operators From Package Namespace in the Current Environment

# **Description**

import\_inops(expose = ...) exposes infix operators specified in a package or an alias object to the current environment (like the global environment or the environment within a function).

import\_inops (unexpose = ...) "unexposes" (i.e. removes) the infix operators specified in a package or an alias object from the current environment (like the global environment or the environment within a function).

Note that in this case only infix operators exposed by the 'tinycodet' import system will be removed from the current environment; "regular" (i.e. user-specified) infix operators will not be touched.

# Usage

```
import_inops(expose = NULL, unexpose = NULL, lib.loc = .libPaths(), ...)
```

### **Arguments**

expose, unexpose

either one of the following:

- an alias object as produced by the import\_as function.
- a string giving the package name.

lib.loc character vector specifying library search path (the location of R library trees to search through).

Only used when supplying a string to expose / unexpose, and ignored when supplying an alias object to expose / unexpose (the library is path already stored inside the alias object).

The lib.loc argument would usually be .libPaths().

See also loadNamespace.

.. additional arguments, only relevant if the expose argument is used.

See import\_inops.control.

18 import\_inops

#### **Details**

# Why Exposing Infix Operators Is Useful

To use a function from an R-package, while avoiding the disadvantages of attaching a package (see tinycodet\_import), one would traditionally use the :: operator like so:

```
packagename::function_name()
```

This is, however, cumbersome with infix operators, as it forces one to code like this:

```
packagename::`%op%`(x,y)
```

Exposing infix operators to the current environment, using the import\_inops() function, allows one to use infix operators without using cumbersome code, and without having to attach the infix operators globally.

#### **Other Details**

The import\_inops() function does not support overloading base/core R operators.

When using import\_inops() to remove infix operators from the current environment, it will use the attributes of those operators to determine if the infix operator came from the 'tinycodet' import system or not. Only infix operators exposed by the 'tinycodet' import system will be removed.

### Value

If using argument expose:

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

If using argument unexpose:

The infix operators specified in the given package or alias, exposed by import\_inops(), will be removed from the current environment (like the Global environment, or the environment within a function).

If such infix operators could not be found, this function simply returns NULL.

### See Also

```
tinycodet_import, import_inops.control(), report_inops()
```

```
import_inops(expose = "stringi") # expose infix operators from package
import_inops(unexpose = "stringi") # remove the exposed infix operators from environment
import_as(~ stri., "stringi")
import_inops(expose = stri.) # expose infix operators from alias
import_inops(unexpose = stri.) # unexposed infix operators from current environment

# additional arguments (only used when exposing, not unexposing):
import_inops(expose = "stringi", exclude = "%s==%")
import_inops(unexpose = "stringi")
```

import\_inops.control 19

```
import_inops(expose = "stringi", overwrite = FALSE)
import_inops(unexpose = "stringi")

import_as(~ stri., "stringi")
import_inops(expose = stri., include.only = "%s==%")
import_inops(unexpose = stri.)
import_inops(expose = stri., overwrite = FALSE)
import_inops(unexpose = stri.)
```

```
import_inops.control import_inops.control
```

### **Description**

Additional arguments to control exposing infix operators in the import\_inops function.

# Usage

```
import_inops.control(
  exclude = NULL,
  include.only = NULL,
  overwrite = TRUE,
  inherits = FALSE
)
```

### **Arguments**

exclude

a character vector, giving the infix operators NOT to expose to the current environment.

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

include.only

a character vector, giving the infix operators to expose to the current environment, and the rest of the operators will not be exposed.

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

overwrite

logical, indicating if it is allowed to overwrite existing infix operators.

- If TRUE (default), a warning is given when operators existing in the current environment are being overwritten, but the function continuous nonetheless.
- If FALSE, an error is produced when the to be exposed operators already exist in the current environment, and the function is halted.

inherits

logical; indicating whether enclosed environments, especially package namespaces, should also be taken into account (TRUE), or not (FALSE).

Defaults to FALSE.

See also exists.

20 import\_LL

#### **Details**

You cannot specify both the exclude and include only arguments. Only one or the other, or neither.

#### Value

This function is used internally in the import\_inops function.

#### See Also

```
import_inops(), tinycodet_import()
```

# **Examples**

```
# additional arguments (only used when exposing, not unexposing):
import_as(~ stri., "stringi")
import_inops(expose = stri., include.only = "%s==%")
import_inops(unexpose = stri.)
import_inops(expose = "stringi", exclude = "%s==%")
import_inops(unexpose = "stringi")
import_inops(expose = stri., overwrite = FALSE)
import_inops(unexpose = stri.)
import_inops(expose = "stringi", overwrite = FALSE)
import_inops(unexpose = "stringi")
```

import\_LL

Miscellaneous import\_ - Functions

# Description

The import\_LL() function places specific functions from a package in the current environment, and also locks (see lockBinding) the specified functions to prevent modification.

The primary use-case for this function is for exposing functions inside a local environment, like the environment within a function.

The import\_int() function directly returns an internal function from a package. It is similar to the ::: operator, but with 2 key differences:

- 1. import\_int() includes the lib.loc argument.
- 2. import\_int() only searches internal functions, not exported ones. This makes it clearer in your code that you're using an internal function, instead of making it ambiguous.

import\_LL 21

#### Usage

```
import_LL(package, selection, lib.loc = .libPaths())
import_int(form, lib.loc = .libPaths())
```

#### **Arguments**

package a single string, giving the name of the package to take functions from.

selection a character vector of function names (both regular functions and infix operators).

Internal functions or re-exported functions are not supported.

lib.loc character vector specifying library search path (the location of R library trees to

search through).

The lib.loc argument would usually be .libPaths().

See also loadNamespace.

form a two-sided formula, with one term on each side.

The term on the left hand side should give a single package name. The term on the right hand side should give a single internal function.

Example: package\_name ~ function\_name

#### **Details**

### Regarding the Locks in import\_LL()

The import\_as function returns a locked environment, just like loadNamespace, thus protecting the functions from accidental modification or re-assignment.

The import\_inops function returns infix operators, and though these are not locked, one needs to surround infix operators by back ticks to re-assign or modify them, which is unlikely to happen on accident.

The import\_LL() function, however, returns "loose" functions. And these functions (unless they are infix operators) do not have the protection due to a locked environment or due to the syntax.

Therefore, to ensure safety from (accidental) modification or re-assignment, the import\_LL() function locks these functions (see lockBinding). For consistency, infix operators exposed by import\_LL() are also locked.

# **Other Details**

The import\_LL() and import\_int() functions do not support importing functions from base/core R

#### Value

```
For import_LL():
```

The specified functions will be placed in the current environment (like the global environment, or the environment within a function), and locked.

To unexpose or overwrite the functions, simply remove them; i.e.:

```
rm(list=c("some_function1", "some_function2")).
```

```
For import_int():
```

The function itself is returned directly.

So one can assign the function directly to some variable, like so:

```
myfun <- import_int(...)</pre>
```

22 inplace

```
or use it directly without re-assignment like so: import_int(...)(...)
```

# See Also

tinycodet\_import

# **Examples**

```
# Using import_LL ====
import_LL(
  "stringi", "stri_sub"
# the stri_sub() function now cannot be modified, only used or removed, because it's locked:
bindingIsLocked("stri_sub", environment()) # TRUE
mypaste <- function(x, y) {</pre>
  import_LL("stringi", selection = "stri_c")
  stri_c(x, y)
mypaste("hello ", "world")
# Using internal function ====
# Through re-assignment:
fun <- import_int(tinycodet ~ .internal_paste, .libPaths())</pre>
fun("hello", "world")
# Or using directly:
import_int(
  tinycodet ~ .internal_paste, .libPaths()
)("hello", "world")
```

inplace

General In-place Modifier Operator

# **Description**

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

For example this:

```
mtcars$mpg[mtcars$cyl > 6] <- mtcars$mpg[mtcars$cyl>6]^2
```

Can now be re-written as:

```
mtcars$mpg\[mtcars$cyl > 6\] %:=% \(x) x^2
```

lock 23

#### **Usage**

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

### **Arguments**

x a variable.

f a (possibly anonymous) function to be applied in-place on x. The function must take one argument only.

### Value

This operator does not return any value:

It is an in-place modifier, and thus modifies the object directly.

#### See Also

```
tinycodet_dry
```

# **Examples**

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

lock

Lock T, Lock F, or Create Locked Constants

# **Description**

The lock\_TF() function locks the T and F values and sets them to TRUE and FALSE, respectively, to prevent the user from re-assigning them.

Removing the created T and F objects allows re-assignment again.

The X %<-c% A operator creates a constant X and assigns A to it.

Constants cannot be changed, only accessed or removed. So if you have a piece of code that requires some unchangeable constant, use this operator to create said constant.

Removing constant X also removes its binding lock. Thus to change a constant, simply remove it and re-create it.

# Usage

```
lock_TF(env)
X %<-c% A</pre>
```

24 lock

#### **Arguments**

env an optional environment to give, determining in which environment T and F

should be locked.

When not specified, the current environment (like the global environment, or the

environment within a function) is used.

X a syntactically valid unquoted name of the object to be created.

A any kind of object to be assigned to X.

#### **Details**

Note that following statement

```
x %<-c% 2+2
print(x)
```

returns

[1] 2

due to R's precedence rules. Therefore, in such cases, the right hand side of X %<-c% A need to be surrounded with brackets. I.e.:

```
x \% < -c\% (2 + 2)
```

Note that the lock\_TF() function and %s<-c% operator create constants through lockBinding. The constants are protected from modification by copy, but they are **not** protected from modification by reference (see for example collapse::setv).

# Value

For lock\_TF():

Two constants, namely T and F, set to TRUE and FALSE respectively, are created in the specified or else current environment, and locked. Removing the created T and F objects allows re-assignment again.

For X %<-c% A:

The object X containing A is created in the current environment, and this object cannot be changed. It can only be accessed or removed.

#### See Also

tinycodet\_safer

```
lock_TF() 
 X <<-c\% data.frame(x = 3, y = 2) # this data.frame cannot be changed. Only accessed or removed. 
 X[1, ,drop=FALSE]
```

logic\_ops 25

logic\_ops Additional 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 checks for every value of numeric vector n if it can be considered a number belonging to type numtype.

The s %=strtype% strtype operator checks for every value of character vector s if it can seen as a certain strtype.

# Usage

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

# **Arguments**

| x, y    | see Logic.   |
|---------|--|
| n       | a numeric vector.  |
| numtype | a single string giving the numeric type to be checked.<br>See Details section for supported types. |
| S       | a character vector.  |
| strtype | a single string giving the string type to be checked.<br>See Details section for supported types.  |

26 logic\_ops

#### **Details**

For argument numtype, 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);
- "I": Integers:
- "odd": odd integers;
- "even": even integers;
- "R": Real numbers;
- "unreal": infinity, NA, or NaN;

For argument strtype, 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 actual number 5.0.
- "special": checks if the string consists of only special characters.

### Value

A logical vector.

```
x <- c(TRUE, FALSE, TRUE, FALSE, NA, NaN, Inf, -Inf, TRUE, FALSE)
y <- c(FALSE, TRUE, TRUE, FALSE, rep(NA, 6))
outcome <- data.frame(
    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% "Br"]
n[n %=numtype% "I"]</pre>
```

matrix\_ops 27

```
n[n %=numtype% "odd"]
n[n %=numtype% "even"]
n[n %=numtype% "R"]
n[n %=numtype% "unreal"]

s <- c(" AbcZ123 ", " abc ", " 1.3 ", " !#$%^&*() ", " ", " NA ", " NaN ", " Inf ")
s[s %=strtype% "empty"]
s[s %=strtype% "unreal"]
s[s %=strtype% "numeric"]
s[s %=strtype% "special"]</pre>
```

matrix\_ops

Row- or Column-wise Re-ordering of Matrices

# **Description**

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

The x %row~% mat operator re-orders the elements of every row, each row ordered independently from the other rows, of matrix x, according to the ordering ranks given in matrix mat.

The x %col~% mat operator re-orders the elements of every column, each column ordered independently from the other columns, of matrix x, according to the ordering ranks given in matrix mat.

In short one can say:

- x %row~% x is a much faster version of apply(x, 1, sort)
- x %col~% x is a much faster version of apply(x, 2, sort)

# Usage

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

# **Arguments**

```
x a matrix mat a matrix with the
```

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

28 matrix\_ops

#### **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 is 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.

mat <- sample(seq_along(x)) |> matrix(ncol = ncol(x)))

then the code

x %row~% mat

will randomly shuffle the elements of every row of x, where the shuffling order in each row is independent from the shuffling order in the other rows.
```

Similarly,

x %col~% mat

will randomly shuffle the elements of every column of x, where the shuffling order in each column is independent from the shuffling order in the other columns.

Re-ordering/sorting every row/column of a matrix with these operators is generally faster than doing so through loops or apply-like functions.

Note that these operators strip all attributes except dimensions.

### Value

A modified matrix.

#### See Also

tinycodet\_misc

```
# numeric matrix ====

x <- matrix(sample(1:25), nrow = 5)
print(x)
x %row~% x # sort elements of every row independently
x %row~% -x # reverse-sort elements of every row independently
x %col~% x # sort elements of every column independently
x %col~% -x # reverse-sort elements of every column independently
x <- matrix(sample(1:25), nrow = 5)
print(x)
mat <- sample(seq_along(x)) |> matrix(ncol = ncol(x))
x %row~% mat # randomly shuffle every row independently
x %col~% mat # randomise shuffle every column independently
# character matrix ====
```

pkgs 29

```
x <- matrix(sample(letters, 25), nrow = 5)
print(x)
mat <- stringi::stri_rank(as.vector(x)) |> matrix(ncol=ncol(x))
x %row~% mat # sort elements of every row independently
x %row~% -mat # reverse-sort elements of every row independently
x %col~% mat # sort elements of every column independently
x %col~% -mat # reverse-sort elements of every column independently
x <- matrix(sample(letters, 25), nrow = 5)
print(x)
mat <- sample(seq_along(x)) |> matrix(ncol = ncol(x))
x %row~% mat # randomly shuffle every row independently
x %col~% mat # randomise shuffle every column independently
```

pkgs

Miscellaneous Package Related Functions

# **Description**

The pkgs %installed in% lib.loc operator checks if one or more packages (pkgs) exist in a library location (lib.loc), without loading the packages.

The syntax of this operator forces the user to make it syntactically explicit where to look for installed R-packages.

As pkgs %installed in% lib.loc does not even load a package, the user can safely use it without fearing any unwanted side-effects.

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

The pkg\_get\_deps\_minimal() function is the same as pkg\_get\_deps(), except with base, recom, rstudioapi, shared\_tidy all set to FALSE, and the default value for deps\_type is c("Depends", "Imports").

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

One handy use for this function is to, for example, globally attach all infix operators from a package using library, like so:

library(packagename, include.only = pkg\_lsf("packagename", type = "inops"))

# Usage

```
pkgs %installed in% lib.loc

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

30 pkgs

```
rstudioapi = TRUE,
    shared_tidy = TRUE
)

pkg_get_deps_minimal(
    package,
    lib.loc = .libPaths(),
    deps_type = c("Depends", "Imports")
)

pkg_lsf(package, type, lib.loc = .libPaths())
```

#### **Arguments**

pkgs a character vector with the package name(s).

lib.loc character vector specifying library search path (the location of R library trees to

search through).

The lib.loc argument would usually be .libPaths().

See also loadNamespace.

package a single string giving the package name.

deps\_type a character vector, giving the dependency types to be used.

The order of the character vector given in deps\_type affects the order of the

returned character vector; see Details sections.

base logical, indicating whether base/core R should be included (TRUE), or not in-

cluded (FALSE).

recom logical, indicating whether the pre-installed 'recommended' R-packages should

be included (TRUE), or not included (FALSE).

rstudioapi logical, indicating whether the 'rstudioapi' R-package should be included (TRUE),

or not included (FALSE).

shared\_tidy logical, indicating whether the shared dependencies of the 'tidyverse' should be

included (TRUE), or not included (FALSE).

**Details:** 

Some of the (often many) dependencies 'tidyverse' packages have are shared across the majority of the 'tidyverse'.

The "official" list of shared dependencies in the 'tidyverse' currently is the following:

'rlang', 'lifecycle', 'cli', 'glue', and 'withr'.

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**

For pkg\_get\_deps():

For each string in argument deps\_type, the package names in the corresponding field of the Description file are extracted, in the order as they appear in that field.

The order given in argument deps\_type also affects the order of the returned character vector:

For example, c("LinkingTo", "Depends", "Imports"),

means the package names are extracted from the fields in the following order:

pversion 31

```
1. "LinkingTo";
```

- 2. "Depends";
- 3. "Imports".

The unique (thus non-repeating) package names are then returned to the user.

### Value

For pkgs %installed in% lib.loc:

Returns a named logical vector, with the names giving the package names, and where the value TRUE indicates a package is installed, and the value FALSE indicates a package is not installed.

```
For pkg_get_deps():
```

A character vector of direct dependencies, without duplicates.

```
For pkg_lsf():
```

Returns a character vector of exported function names in the specified package.

# References

```
O'Brien J., elegantly extract R-package dependencies of a package not listed on CRAN. Stack Over-flow. (1 September 2023). https://stackoverflow.com/questions/30223957/elegantly-extract-r-package-d
```

### See Also

```
tinycodet_import
```

# **Examples**

```
"dplyr" %installed in% .libPaths()

pkg_get_deps_minimal("dplyr")
pkgs <- pkg_get_deps("dplyr")
pkgs %installed in% .libPaths()
pkg_lsf("dplyr", "all")</pre>
```

pversion

Check for Package Versions Mismatch

# **Description**

The pversion\_check4mismatch() function checks if there is any mismatch between the currently loaded packages and the packages in the specified library path.

The pversion\_report() function gives a table of all specified packages, with their loaded and installed versions, regardless if there is a mismatch or not.

32 pversion

#### Usage

```
pversion_check4mismatch(pkgs = NULL, lib.loc = .libPaths())
pversion_report(pkgs = NULL, lib.loc = .libPaths())
```

### **Arguments**

pkgs a character vector with the package name(s).

Packages that are not actually loaded will be ignored.

Base/core R will also be ignored.

If NULL, all loaded packages (see loadedNamespaces) excluding core/base R will

be checked.

lib.loc character vector specifying library search path (the location of R library trees to

search through).

The lib.loc argument would usually be .libPaths().

See also loadNamespace.

### Value

For pversion\_check4mismatch():

If no mismatch between loaded versions and those in 1ib.loc were found, returns NULL.

Otherwise it returns a data.frame, with the loaded version and library version of the specified packages.

For pversion\_report():

Returns a data.frame, with the loaded version and library version of the specified packages, as well as a logical column indicating whether the two versions are equal (TRUE), or not equal (FALSE).

# See Also

tinycodet\_import

```
"dplyr" %installed in% .libPaths()
import_as(~dpr., "dplyr")
pversion_check4mismatch()
pversion_report()
```

report\_inops 33

report\_inops

Report Infix Operators

# **Description**

The report\_inops() function returns a data.frame listing the infix operators defined in the current environment (like the global environment, or the environment within a function), or a user specified environment. It also reports from which packages the infix operators came from.

### Usage

```
report_inops(env)
```

# **Arguments**

env

an optional environment to give, where the function should look for infix oper-

ators.

When not specified, the current environment (like the global environment, or the

environment within a function) is used.

### Value

A data.frame. The first column gives the infix operator names. The second column gives the package the operator came from, or NA if it did not come from a package.

### See Also

```
tinycodet_misc()
```

```
report_inops()
`%paste%` <- function(x,y)paste0(x,y)
report_inops()
import_inops("stringi")
report_inops()</pre>
```

34 source\_selection

| source | selection |
|--------|-----------|

Source Specific Objects from Script

#### **Description**

The source\_selection() function is the same as base R's source function, except that it allows only placing the selected objects and functions into the current environment, instead of all objects.

The objects to be selected can be specified using any combination of the following:

- by supplying a character vector of exact object names to the select argument.
- by supplying a character vector of regex patterns to the regex argument.
- by supplying a character vector of fixed patterns to the fixed argument.

Note that the source\_selection() function does NOT suppress output (i.e. plots, prints, messages) from the sourced script file.

### Usage

```
source_selection(lst, select = NULL, regex = NULL, fixed = NULL)
```

### **Arguments**

1st a named list, giving the arguments to be passed to the source function.

The local argument should not be included in the list.

select a character vector, giving the exact names of the functions or objects appearing

in the script, to expose to the current environment.

regex a character vector of regex patterns (see about\_search\_regex).

These should give regular expressions that match to the names of the functions or objects appearing in the script, to expose to the current environment.

or objects appearing in the script, to expose to the current environment. For example, to expose the following methods to the current environment, mymethod.numeric() and mymethod.character() from generic mymethod(),

one could specify regex = "^mymethod".

about search: regex

fixed a character vector of fixed patterns (see about\_search\_fixed).

These should give fixed expressions that match to the names of the functions or objects appearing in the script, to expose to the current environment.

For example, to expose the following methods to the current environment, mymethod.numeric() and mymethod.character() from generic mymethod(),

one could specify fixed = "mymethod".

about search: fixed

# **Details**

One can specify which objects to expose using arguments select, regex, or fixed. The user can specify all 3 of them, but at least one of the 3 must be specified. It is not a problem if the specifications overlap.

streut\_loc 35

#### Value

Any specified objects will be placed in the current environment (like the Global environment, or the environment within a function).

#### See Also

```
tinycodet_misc, base::source()
```

# **Examples**

```
exprs <- expression({</pre>
helloworld = function()print("helloworld")
goodbyeworld <- function() print("goodbye world")</pre>
`%s+test%` <- function(x,y) stringi::`%s+%`(x,y)
`%s*test%` <- function(x,y) stringi::`%s*%`(x,y)
mymethod <- function(x) UseMethod("mymethod", x)</pre>
mymethod.numeric <- function(x)x * 2
mymethod.character \leftarrow function(x)chartr(x, old = "a-zA-Z", new = "A-Za-z")
})
source_selection(list(exprs=exprs), regex = "^mymethod")
mymethod(1)
mymethod("a")
temp.fun <- function(){</pre>
  source_selection(list(exprs=exprs), regex = "^mymethod", fixed = c("%", ":="))
  ls() \# list all objects residing within the function definition
}
temp.fun()
temp.fun <- function(){</pre>
  source_selection(list(exprs=exprs), select = c("helloworld", "goodbyeworld"))
  ls() # list all objects residing within the function definition
temp.fun()
```

strcut\_loc

Cut Strings

### **Description**

The strcut\_loc() function cuts every string in a character vector around a location range loc, such that every string is cut into the following parts:

- the sub-string **before** loc;
- the sub-string at loc itself;
- the sub-string after loc.

36 strcut\_loc

The location range loc would usually be matrix with 2 columns, giving the start and end points of some pattern match.

The strcut\_brk() function (a wrapper around stri\_split\_boundaries(..., tokens\_only = FALSE)) cuts every string into individual text breaks (like character, word, line, or sentence boundaries).

### Usage

```
strcut_loc(str, loc)
strcut_brk(str, type = "character", tolist = FALSE, n = -1L, ...)
```

# **Arguments**

str a string or character vector.

loc Either one of the following:

- the result from the stri\_locate\_ith function.
- a matrix of 2 integer columns, with nrow(loc)==length(str), giving the location range of the middle part.
- a vector of length 2, giving the location range of the middle part.

type either one of the following:

- a single string giving the break iterator type (i.e. "character", "line\_break", "sentence", "word", or a custom set of ICU break iteration rules).
- a list with break iteration options, like a list produced by stringi::stri\_opts\_brkiter.

about search: boundaries

tolist logical, indicating if strcut\_brk should return a list (TRUE), or a matrix (FALSE,

default).

n see stri\_split\_boundaries.

... additional arguments to be passed to stri\_split\_boundaries.

### **Details**

The main difference between the strcut\_ - functions and stri\_split / strsplit, is that the latter generally removes the delimiter patterns in a string when cutting, while the strcut\_-functions do not attempt to remove parts of the string by default, they only attempt to cut the strings into separate pieces. Moreover, the strcut\_ - functions return a matrix by default.

#### Value

For strcut\_loc():

A character matrix with length(str) rows and 3 columns, where for every row i it holds the following:

- the first column contains the sub-string **before** loc[i,], or NA if loc[i,] contains NA;
- the second column contains the sub\_string at loc[i,], or the uncut string if loc[i,] contains NA;

stri\_join\_mat 37

• the third and last column contains the sub-string after loc[i,], or NA if loc[i,] contains NA.

```
For strcut_brk(..., tolist = FALSE):
```

A character matrix with length(str) rows and a number of columns equal to the maximum number of pieces str was cut in.

Empty places are filled with NA.

```
For strcut_brk(..., tolist = TRUE):
```

A list with length(str) elements, where each element is a character vector containing the cut string.

## See Also

tinycodet\_strings

# **Examples**

```
x <- rep(paste0(1:10, collapse = ""), 10)
print(x)
loc <- stri_locate_ith(x, 1:10, fixed = as.character(1:10))</pre>
strcut_loc(x, loc)
strcut_loc(x, c(5,5))
strcut_loc(x, c(NA, NA))
strcut_loc(x, c(5, NA))
strcut_loc(x, c(NA, 5))
test <- "The\u00a0above-mentioned
                                        features are very useful. " %s+%
"Spam, spam, eggs, bacon, and spam. 123 456 789"
strcut_brk(test, "line")
strcut_brk(test, "word")
strcut_brk(test, "sentence")
strcut_brk(test)
strcut_brk(test, n = 1)
strcut_brk(test, "line", tolist = TRUE)
strcut_brk(test, "word", tolist = TRUE)
strcut_brk(test, "sentence", tolist = TRUE)
brk <- stringi::stri_opts_brkiter(</pre>
  type = "line"
strcut_brk(test, brk)
```

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.

38 stri\_join\_mat

#### **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 within 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 within 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.
```

#### Value

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

## See Also

tinycodet\_strings

# **Examples**

```
# Basic example
x \leftarrow matrix(letters[1:25], ncol = 5, byrow = TRUE)
print(x)
stri_join_mat(x, margin = 1)
x <- matrix(letters[1:25], ncol = 5, byrow = FALSE)</pre>
print(x)
stri_join_mat(x, margin = 2)
# sorting characters in strings ====
x <- c(paste(sample(letters), collapse = ""),</pre>
      paste(sample(letters), collapse = ""))
print(x)
mat <- strcut_brk(x)</pre>
rank <- stringi::stri_rank(as.vector(mat)) |> matrix(ncol=ncol(mat))
sorted <- mat %row~% rank</pre>
sorted[is.na(sorted)] <- ""</pre>
print(sorted)
```

```
stri_join_mat(sorted, margin = 1)
stri_join_mat(sorted, margin = 2)
# sorting words ====
x <- c("2nd 3rd 1st", "Goodbye everyone")
print(x)
mat <- strcut_brk(x, "word")</pre>
rank <- stringi::stri_rank(as.vector(mat)) |> matrix(ncol=ncol(mat))
sorted <- mat %row~% rank</pre>
sorted[is.na(sorted)] <- ""</pre>
stri_c_mat(sorted, margin = 1, sep = " ") # <- alias for stri_join_mat</pre>
stri_c_mat(sorted, margin = 2, sep = " ")
# randomly shuffling sentences ====
x <- c("Hello, who are you? Oh, really?! Cool!",
      "I don't care. But I really don't.")
print(x)
mat <- strcut_brk(x, "sentence")</pre>
rank <- sample(1:length(mat)) |> matrix(ncol = ncol(mat))
sorted <- mat %row~% rank</pre>
sorted[is.na(sorted)] <- ""</pre>
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 or Text Boundary

# **Description**

The stri\_locate\_ith() function locates the  $i^{th}$  occurrence of a pattern in each string of some character vector.

The stri\_locate\_ith\_boundaries() function locates the  $i^{th}$  text boundary (like character, word, line, or sentence boundaries).

# Usage

```
stri_locate_ith(str, i, ..., regex, fixed, coll, charclass)
stri_locate_ith_regex(str, pattern, i, ..., opts_regex = NULL)
stri_locate_ith_fixed(str, pattern, i, ..., opts_fixed = NULL)
stri_locate_ith_coll(str, pattern, i, ..., opts_collator = NULL)
```

```
stri_locate_ith_charclass(str, pattern, i, merge = TRUE, ...)
stri_locate_ith_boundaries(str, i, ..., opts_brkiter = NULL)
Arguments
```

str

a string or character vector.

i

an integer, or an integer vector of the same length as str.

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 occurrence of a pattern.
- stri\_locate\_ith(str, i = -1, ...) gives the position (range) of the last occurrence of a pattern.
- stri\_locate\_ith(str, i = 2, ...) gives the position (range) of the second occurrence 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 pattern occurrences n, the first (if i < -n) or last (if i > n) instance will be given.

For example: suppose a string has 3 instances of some pattern;

then if  $i \ge 3$  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 or stri\_locate\_all\_boundaries.

Do not supply the arguments omit\_no\_match or get\_length, as they are already specified internally. Supplying these arguments anyway will result in an error.

pattern, regex, fixed, coll, charclass

a character vector of search patterns, as in stri\_locate.

about search: regex about search: fixed about search: coll about search: charclass

opts\_regex, opts\_fixed, opts\_collator, opts\_brkiter

named list used to tune up the selected search engine's settings.

see stri\_opts\_regex, stri\_opts\_fixed, stri\_opts\_collator, and stri\_opts\_brkiter

NULL for the defaults. about search: regex about search: fixed about search: coll about search: charclass about search: boundaries

merge

logical, indicating if charclass locations should be merged or not.

#### **Details:**

For the charclass pattern type, the stri\_locate\_ith() function gives the start and end of **consecutive** characters by default, just like stri\_locate\_all. To give the start and end positions of single characters, much like stri\_locate\_first or stri\_locate\_last, set merge = FALSE.

#### **Details**

The 'stringi' functions only support operations on the first, last, or all occurrences of a pattern.

The stri\_locate\_ith() function allows locating the  $i^{th}$  occurrence of a pattern.

This allows for several workflows for operating on the  $i^{th}$  pattern occurrence.

See also the examples section.

## Extract $i^{th}$ Occurrence of a Pattern

For extracting the  $i^{th}$  pattern occurrence:

Locate the the  $i^{th}$  occurrence using stri\_locate\_ith(), and then extract it using, for example, stri\_sub.

# Replace/Transform $i^{th}$ Occurrence of a Pattern

For replacing/transforming the  $i^{th}$  pattern occurrence:

- 1. Locate the the  $i^{th}$  occurrence using stri\_locate\_ith().
- 2. Extract the occurrence using stri\_sub.
- 3. Transform or replace the extracted sub-strings.
- 4. Return the transformed/replaced sub-string back, using again stri\_sub.

# Capture Groups of $i^{th}$ Occurrence of a Pattern

The capture\_groups argument for regex is not supported within  $stri_locate_ith()$ . To capture the groups of the  $i^{th}$  occurrences:

- 1. Use stri\_locate\_ith() to locate the  $i^{th}$  occurrences without group capture.
- 2. Extract the occurrence using stri\_sub.
- 3. Get the matched group capture on the extracted occurrences using stri\_match.

## Value

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

If an empty string or empty pattern is supplied, a warning is given and a matrix with 0 rows is returned.

## Warning

The stri\_locate\_ith() function does not support long vectors, i.e. character vectors with more than 2^31 - 1 strings.

# See Also

tinycodet\_strings

# **Examples**

```
# practical example: transform regex pattern ====
# input character vector:
x \leftarrow c(paste0(letters[1:13], collapse = ""), paste0(letters[14:26], collapse = ""))
print(x)
# locate ith (second and second-last) vowel locations:
p \leftarrow rep("A|E|I|0|U", 2) # vowels
loc <- stri_locate_ith(x, c(2, -2), regex = p, case_insensitive = TRUE)
print(loc)
# extract ith vowels:
extr <- stringi::stri_sub(x, loc)</pre>
print(extr)
# transform & replace ith vowels with numbers:
repl <- chartr("aeiou", "12345", extr)</pre>
stringi::stri_sub(x, loc) <- repl</pre>
# result (notice ith vowels are now numbers):
# practical example: group-capture regex pattern ====
# input character:
# first group: c(breakfast=eggs, breakfast=bacon)
# second group: c(lunch=pizza, lunch=spaghetti)
x <- c('breakfast=eggs;lunch=pizza',</pre>
      'breakfast=bacon;lunch=spaghetti',
      'no food here') # no group here
print(x)
# locate ith=2nd group:
p <- '(\\w+)=(\\w+)'
loc <- stri_locate_ith(x, i = 2, regex = p)</pre>
print(loc)
# extract ith=2nd group:
extr <- stringi::stri_sub(x, loc)</pre>
print(extr)
# capture ith=2nd group:
stringi::stri_match(extr, regex = p)
# practical example: replace words using boundaries ====
# input character vector:
```

```
x <- c("good morning and good night",
"hello ladies and gentlemen")
print(x)
# report ith word locations:
loc <- stri_locate_ith_boundaries(x, c(-3, 3), type = "word")</pre>
print(loc)
# extract ith words:
extr <- stringi::stri_sub(x, from=loc)</pre>
print(extr)
# transform and replace words (notice ith words have inverted case):
tf <- chartr(extr, old = "a-zA-Z", new = "A-Za-z")</pre>
stringi::stri_sub(x, loc) <- tf</pre>
# result:
print(x)
# find pattern ====
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)
# simple pattern ====
x <- rep(paste0(1:10, collapse=""), 10)</pre>
out <- stri_locate_ith(x, 1:10, regex = as.character(1:10))</pre>
cbind(1:10, out)
x <- c(paste0(letters[1:13], collapse=""),</pre>
      paste0(letters[14:26], collapse = ""))
print(x)
p \leftarrow rep("a|e|i|o|u",2)
out <- stri_locate_ith(x, c(-1, 1), regex = p)</pre>
print(out)
substr(x, out[,1], out[,2])
# ignore case pattern ====
x <- c(paste0(letters[1:13], collapse = ""),</pre>
      paste0(letters[14:26], collapse = ""))
print(x)
p \leftarrow rep("A|E|I|0|U", 2)
out <- stri_locate_ith(x, c(1, -1), regex = p, case_insensitive = TRUE)
substr(x, out[,1], out[,2])
```

44 str\_arithmetic

```
# multi-character pattern ====
x <- c(paste0(letters[1:13], collapse = ""),
     paste0(letters[14:26], collapse = ""))
print(x)
# multi-character pattern:
p \leftarrow rep("AB", 2)
out <- stri_locate_ith(x, c(1, -1), regex = p, case_insensitive = TRUE)
print(out)
substr(x, out[,1], out[,2])
# Replacement transformation using stringi ====
x <- c("hello world", "goodbye world")</pre>
loc \leftarrow stri_locate_ith(x, c(1, -1), regex = "a|e|i|o|u")
extr <- stringi::stri_sub(x, from = loc)</pre>
repl <- chartr(extr, old = "a-zA-Z", new = "A-Za-z")
stringi::stri_sub_replace(x, loc, replacement = repl)
# Boundaries ====
test <- c(
paste0("The\u00a0above-mentioned features are very useful. ",
     "Spam, spam, eggs, bacon, and spam. 123 456 789"),
    "good morning, good evening, and good night"
    )
loc <- stri_locate_ith_boundaries(test, i = c(1, -1), type = "word")</pre>
stringi::stri_sub(test, from = loc)
```

str\_arithmetic

String Arithmetic Operators

## **Description**

String arithmetic operators.

The x %s+% y operator is exported from 'stringi', and concatenates character vectors x and y.

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

The x %s\*% n operator is exported from 'stringi', and duplicates each string in x n times, and concatenates the results.

str\_arithmetic 45

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

The x %s//% brk operator counts how often the text boundary specified in list brk occurs in each element of x.

The e1 %s\$% e2 operator is exported from 'stringi', and provides access to stri\_sprintf in the form of an infix operator.

The x %ss% p operator splits the strings in x by a delimiter character/pattern defined in p, and removes p in the process.

For cutting strings by text boundaries, or around a location, see strcut\_brk and strcut\_loc.

# Usage

```
x %s-% p
```

x %s/% p

x %s//% brk

x %ss% p

# **Arguments**

x a string or character vector.

p either a list with 'stringi' arguments (see s\_pattern), or else a character vector

with regular expressions.

about search: regex about search: fixed about search: coll about search: charclass

brk a list with break iteration options, like a list produced by stringi::stri\_opts\_brkiter.

about search: boundaries

#### Value

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

The %s/% and %s//% both return an integer vector of the same length as x.

The %s\$% operator returns a character vector.

The %ss% operator returns a list of the split strings - or, if simplify = TRUE / simplify = NA, returns a matrix of the split strings.

# See Also

tinycodet\_strings

# **Examples**

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

```
paste0(letters[14:26], collapse = ""))
print(x)
y <- c("a", "b")
p \leftarrow rep("a|e|i|o|u", 2) \# same as p \leftarrow list(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 %ss% p # split x around vowels, removing the vowels in the process
x %ss% s_regex(p, simplify = NA) # same as above, but in matrix form
test <- c(
paste0("The\u00a0above-mentioned
                                     features are very useful. ",
"Spam, spam, eggs, bacon, and spam. 123 456 789"),
"good morning, good evening, and good night"
)
test %s//% list(type = "character")
x <- c(paste0(letters[1:13], collapse = ""),</pre>
       paste0(letters[14:26], collapse = ""))
print(x)
y <- "a"
# pattern that ignores case:
p \leftarrow list(regex = rep("A|E|I|0|U", 2), case_insensitive = TRUE)
n <- c(2, 3)
x %s+% y # = paste0(x,y)
x %s-% p # remove all vowels from x
x %s/% p # count how often vowels appears in each string of vector x.
x <- c(paste(letters, collapse = ", "), paste(LETTERS, collapse = ", "))</pre>
print(x)
x %ss% ",
t(x %ss% s_fixed(", ", simplify = NA))
```

str\_search

'stringi' Pattern Search Operators

# Description

The  $x %s{}$  p operator checks for every string in character vector x if the pattern defined in p is present.

When supplying a list on the right hand side (see  $s_pattern$ ), one can optionally include the list element at = "start" or at = "end":

- Supplying at = "start" will check if the pattern appears at the start of a string (like stri\_startswith).
- Supplying at = "end" will check if the pattern appears at the end of a string (like stri\_endswith).

The  $x \%s!{}$  p operator is the same as  $x \%s{}$  p, except it checks for **absence** of the pattern, rather than presence.

For string (in)equality operators, see %s==% from the 'stringi' package.

strfind()<- locates, extracts, or replaces found patterns. It complements the other string-related operators, and uses the same s\_pattern API. It functions as follows:

- strfind() finds all pattern matches, and returns the extractions of the findings in a list, just like stri\_extract\_all.
- strfind(..., i = "all") finds all pattern matches, and reports the locations of the findings in a list, just like stri\_locate\_all.
- strfind(..., i = i), where i is an integer vector, locates the  $i^{th}$  occurrence of a pattern, and reports the locations in a matrix, just like stri\_locate\_ith.
- strfind() <- value finds pattern matches in variable x, replaces the pattern matches with the character vector specified in value, and assigns the transformed character vector back to x

This is similar to stri\_replace\_all, except the replacement is done in-place.

# Usage

```
x %s{}% p
x %s!{}% p
strfind(x, p, ..., i, rt)
strfind(x, p, ..., i, rt) <- value</pre>
```

# **Arguments**

x a string or character vector.

For strfind()<-, x must obviously be the variable containing the character vector/string, since strfind()<- performs assignment in-place.

p either a list with 'stringi' arguments (see s\_pattern), or else a character vector with regular expressions.

See also the Details section.

about search: regex about search: fixed about search: coll about search: charclass

. . . additional arguments to be specified.

i either one of the following can be given for i:

- $\bullet\,$  if i is not given or NULL, strfind() extracts all found pattern occurrences.
- if i = "all", strfind() locates all found pattern occurrences.

if i is an integer, strfind() locates the i<sup>th</sup> pattern occurrences.
 See the i argument in stri locate ith for details.

For strfind() <- value, i must not be specified.

use rt to specify the Replacement Type that strfind()<- should perform. Either one of the following can be given for rt:

- if rt is not given, NULL or "vec", strfind()<- performs regular, vector-ized replacement of **all** occurrences.
- if rt = "dict", strfind()<- performs dictionary replacement of **all** occurrences.
- if rt = "first", strfind()<- replaces only the first occurrences.
- if rt = "last", strfind()<- replaces only the last occurrences.

Note: rt = "first" and rt = "last" only exist for convenience; for more specific locational replacement, use stri\_locate\_ith or strfind(..., i) with numeric i (see the Examples section).

For strfind(), rt must not be specified.

value

rt

a character vector giving the replacement values.

## **Details**

## Right-hand Side List for the %s{}% and %s!{}% Operators

When supplying a list to the right-hand side of the  $s{}$  and s! operators, one can add the argument at.

If at = "start", the operators will check if the pattern is present/absent at the start of the string. If at = "end", the operators will check if the pattern is present/absent at the end of the string. Unlike stri\_startswith or stri\_endswith, regex is supported by the %s{}% and %s!{}% operators. See examples below.

# **Vectorized Replacement vs Dictionary Replacement**

- Vectorized replacement:
  - x, p, and value are of the same length (or recycled to become the same length). All occurrences of pattern p[j] in x[j] is replaced with value[j], for every j.
- Dictionary replacement:
  - p and value are of the same length, and their length is independent of the length of x. For every single string in x, all occurrences of pattern p[1] are replaced with value[1], all occurrences of pattern p[2] are replaced with value[2], etc.

Notice that for single replacement, i.e. rt = "first" or rt = "last", it makes no sense to distinguish between vectorized or dictionary replacement, since then only a single occurrence is being replaced per string.

See examples below.

#### Value

The x  $%s{}$  p and x  $%s!{}$  p operators return logical vectors.

strfind() returns a list with extractions of all found patterns.

```
strfind(..., i = "all") returns a list with all found pattern locations.
```

strfind(..., i = i), with i being an integer, returns an integer matrix with two columns, giving the start and end positions of the  $i^{th}$  matches, two NAs if no matches are found, and also two NAs if str is NA.

strfind() <- value returns nothing, but performs in-place replacement (but not by reference, technically speaking) of the found patterns in variable x.

## Warning

strfind()<- performs in-place replacement.</pre>

Therefore, the character vector or string to perform replacement on, must already exist as a variable. So take for example the following code:

```
strfind("hello", p = "e") <- "a" \# this obviously does not work y <- "hello" \# strfind(y, p = "e") <- "a" \# this works fine
```

In the above code, the first strfind()<- call does not work, because the string needs to exist as a variable.

#### See Also

tinycodet\_strings

# **Examples**

```
paste0(letters[14:26], collapse = ""))
x %s{}% "1"
x %s!{}% "1"
which(x %s{}% "1")
which(x %s!{}% "1")
x[x %s{}% "1"]
x[x %s!{}% "1"]
x[x %s{}]% "1"] <- "a"
x[x %s!{}% "1"] <- "a"
print(x)
# Example of %s{}% and %s!{}% with "at" argument ====
x <- c(paste0(letters, collapse = ""),</pre>
      paste0(rev(letters), collapse = ""), NA)
p <- s_fixed("abc", at = "start")</pre>
x %s{}% p
stringi::stri_startswith(x, fixed = "abc") # same as above
p <- s_fixed("xyz", at = "end")</pre>
x %s{}% p
stringi::stri_endswith(x, fixed = "xyz") # same as above
p <- s_fixed("cba", at = "end")</pre>
x %s{}% p
stringi::stri_endswith(x, fixed = "cba") # same as above
p <- s_fixed("zyx", at = "start")</pre>
x %s{}% p
stringi::stri_startswith(x, fixed = "zyx") # same as above
# Example of transforming ith occurrence ====
# new character vector:
x <- c(paste0(letters[1:13], collapse = ""),</pre>
      paste0(letters[14:26], collapse = ""))
print(x)
# report ith (second and second-last) vowel locations:
p <- s_regex( # vowels</pre>
 rep("A|E|I|0|U", 2),
 case\_insensitive = TRUE
loc <- strfind(x, p, i = c(2, -2))
print(loc)
# extract ith vowels:
extr <- stringi::stri_sub(x, from = loc)</pre>
print(extr)
```

```
# replace ith vowels with numbers:
repl <- chartr("aeiou", "12345", extr) # transformation</pre>
stringi::stri_sub(x, loc) <- repl</pre>
print(x)
# Example of strfind for regular vectorized replacement ====
x <- rep('The quick brown fox jumped over the lazy dog.', 3)
print(x)
p <- c('quick', 'brown', 'fox')</pre>
rp <- c('SLOW', 'BLACK', 'BEAR')</pre>
x %s{}% p
strfind(x, p)
strfind(x, p) \leftarrow rp
print(x)
# Example of strfind for dictionary replacement ====
x <- rep('The quick brown fox jumped over the lazy dog.', 3)</pre>
print(x)
p <- c('quick', 'brown', 'fox')</pre>
rp <- c('SLOW', 'BLACK', 'BEAR')</pre>
# thus dictionary is:
# guick => SLOW; brown => BLACK; fox => BEAR
strfind(x, p, rt = "dict") <- rp
print(x)
# Example of strfind for first and last replacement ====
x <- rep('The quick brown fox jumped over the lazy dog.', 3)</pre>
print(x)
p <- s_fixed("the", case_insensitive = TRUE)</pre>
rp <- "One"
strfind(x, p, rt = "first") <- rp</pre>
print(x)
x <- rep('The quick brown fox jumped over the lazy dog.', 3)</pre>
print(x)
p <- s_fixed("the", case_insensitive = TRUE)</pre>
rp <- "Some Other"</pre>
strfind(x, p, rt = "last") <- rp</pre>
print(x)
```

52 str\_subset\_ops

str\_subset\_ops

String Subsetting Operators

# **Description**

String subsetting operators.

The x %sget% ss operator gets a certain number of the first and last characters of every string in character vector x.

The x %strim% ss operator trims a certain number of the first and last characters of every string in character vector x.

## Usage

x %sget% ss

x %strim% ss

# **Arguments**

SS

x a character vector.

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 and non-missing integers, or be coerce-able to such integers. (thus negative integers, and missing values are not allowed; decimal numbers will be converted to integers).

The first element/column of ss gives the number of characters counting from

the left side to be extracted/removed from  $\boldsymbol{x}$ .

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

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

subset\_if 53

#### See Also

tinycodet\_strings

### **Examples**

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

 $subset_if$ 

Conditional Sub-setting and In-place Replacement of Unreal Values

# **Description**

The  $x \leq if \leq cond$  operator selects elements from vector/matrix/array x, for which the result of cond(x) returns TRUE.

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

The  $x \, \text{unreal} = \text{modifies}$  all unreal (NA, NaN, Inf, -Inf) values of  $x \, \text{with}$  replacement value repl.

```
Thus,
x %unreal =% repl,
is the same as,
x[is.na(x) | is.nan(x) | is.infinite(x)] <- repl
```

54 s\_pattern

#### Usage

```
x %[if]% cond
x %[!if]% cond
x %unreal =% repl
```

# **Arguments**

x a vector, matrix, or array.

cond a (possibly anonymous) function that returns a logical vector of the same

length/dimensions as x. For example: (x)x>0.

repl the replacement value.

## Value

For the x %[if]% cond and x %[!if]% cond operators:

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

For the x %unreal =% repl operator:

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

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

## See Also

tinycodet\_dry

# **Examples**

```
x <- c(-10:9, NA, NA)
object_with_very_long_name <- matrix(x, ncol=2)
print(object_with_very_long_name)
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>
```

s\_pattern

Pattern Specifications for String Related Operators

s\_pattern 55

# **Description**

The %s-%, %s/%, %ss% operators, as well as the string search operators (str\_search), perform pattern matching for some purpose, where the pattern is given on the right hand side.

When a character vector or string is given on the right hand side, this is interpreted as case-sensitive regex patterns from 'stringi'.

Instead of giving a string or character vector of regex patterns, one can also supply a list to specify exactly how the pattern should be interpreted. The list should use the exact same argument convention as 'stringi'.

For example:

```
list(regex = p, case_insensitive = FALSE, ...)
list(fixed = p, ...)
list(coll = p, ...)
list(charclass = p, ...)
```

All arguments in the list are simply passed to the appropriate functions in 'stringi'. For example:

```
x %s/% p
```

counts how often regular expression specified in character vector p occurs in x, whereas the following,

```
x %s/% list(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.

'tinycodet' adds some convenience functions based on the stri\_opts\_ - functions in 'stringi':

```
s_regex(p, ...) is equivalent to list(regex = p, ...)
s_fixed(p, ...) is equivalent to list(fixed = p, ...)
s_coll(p, ...) is equivalent to list(coll = p, ...)
s_chrcls(p, ...) is equivalent to list(charclass = p, ...)
```

With the ellipsis (...) being passed to the appropriate 'stringi'-functions when it matches their arguments.

'stringi' infix operators start with "%s", though they all have an alias starting with "%stri". In analogy to that, the above functions start with "s\_" rather than "stri\_", as they are all meant for operators only.

# Usage

```
s_regex(
  p,
  case_insensitive,
  comments,
```

56 s\_pattern

```
dotall,
      multiline,
      time_limit,
      stack_limit,
    )
    s_fixed(p, case_insensitive, overlap, ...)
    s_coll(
      р,
      locale,
      strength,
      alternate_shifted,
      french,
      uppercase_first,
      case_level,
      numeric,
      normalization,
    s_chrcls(p, ...)
Arguments
    р
                      a character vector giving the pattern to search for.
                      about search: regex
                      about search: fixed
                      about search: coll
                      about search: charclass
    case_insensitive
                      see stri_opts_regex and stri_opts_fixed.
    comments, dotall, multiline
                      see stri_opts_regex.
    time_limit, stack_limit
                      see stri_opts_regex.
                      additional arguments not part of the stri_opts - functions to be passed here.
                      For example: the at argument for the str_search operators.
    overlap
                      see stri_opts_fixed.
    locale, strength, alternate_shifted
                      see stri_opts_collator.
    french, normalization, numeric
                      see stri_opts_collator.
    uppercase_first, case_level
                      see stri_opts_collator.
```

# Value

A list with arguments to be passed to the appropriate operators.

transform\_if 57

#### See Also

tinycodet\_strings

## **Examples**

```
x <- c(paste0(letters[1:13], collapse = ""),</pre>
       paste0(letters[14:26], collapse = ""))
print(x)
p \leftarrow rep("a|e|i|o|u", 2) \# same as p \leftarrow list(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 = ""),</pre>
       paste0(letters[14:26], collapse = ""))
print(x)
x %s/% list(regex = rep("A|E|I|0|U", 2), case_insensitive = TRUE)
x %s/% s_{e}(rep("A|E|I|0|U", 2), case_insensitive = TRUE)
x <- c(paste0(letters[1:13], collapse = ""),</pre>
       paste0(letters[14:26], collapse = ""))
print(x)
p <- list(fixed = c("A", "A"), case_insensitive = TRUE)</pre>
x %s{}% p
x %s!{}% p
p <- s_fixed(c("A", "A"), case_insensitive = TRUE)</pre>
x %s{}% p
x %s!{}% p
x <- c(paste0(letters[1:13], collapse = ""),</pre>
      paste0(letters[14:26], collapse = ""), NA)
p \leftarrow s_fixed("abc", at = "start")
x %s{}% p
stringi::stri_startswith(x, fixed = "abc") # same as above
p \leftarrow s_fixed("xyz", at = "end")
x %s{}% p
stringi::stri_endswith(x, fixed = "xyz") # same as above
```

transform\_if

transform\_if: Conditional Sub-set Transformation of Atomic objects

# Description

The transform\_if() function transforms an object x, based on the logical result (TRUE, FALSE, NA) of condition function cond(x) or logical vector cond, such that:

- For every value where cond(x)==TRUE / cond==TRUE, function yes(x) is run or scalar yes is returned.
- For every value where cond(x)==FALSE / cond==FALSE, function no(x) is run or scalar no is returned.

58 transform\_if

• For every value where cond(x)==NA / cond==NA, function other(x) is run or scalar other is returned.

For a more ifelse-like function where yes, no, and other are vectors, see kit::iif.

# Usage

```
transform_if(x, cond, yes = function(x) x, no = function(x) x, other = NA)
```

# **Arguments**

x a vector, matrix, or array.

cond either an object of class logical with the same length as x,

or a (possibly anonymous) function that returns an object of class logical with

the same length as x. For example:  $\(x)x>0$ .

yes the (possibly anonymous) transformation function to use when function cond(x) == TRUE

/ logical cond==TRUE.

Alternatively, one can also supply an atomic scalar. If argument yes is not specified, it defaults to  $\(x)x$ .

no the (possibly anonymous) transformation function to use when function cond(x) = FALSE

/ logical cond==FALSE.

Alternatively, one can also supply an atomic scalar. If argument no is not specified, it defaults to  $\(x)x$ .

other the (possibly anonymous) transformation function to use when function cond(x)

/ logical cond returns NA.

Alternatively, one can also supply an atomic scalar. If argument other is not specified, it defaults to NA.

Note that function other(x) is run or scalar other is returned when function

cond(x) or logical cond is NA, not necessarily when x itself is NA.

#### **Details**

Be careful with coercion! For example the following code:

```
x <- c("a", "b")
transform_if(x, \(x) x == "a", as.numeric, as.logical)
returns:
```

[1] NA NA

due to the same character vector being given 2 incompatible classes.

## Value

The transformed vector, matrix, or array (attributes are conserved).

x.import 59

#### See Also

```
tinycodet_dry
```

### **Examples**

```
x <- c(-10:9, NA, NA)
object <- matrix(x, ncol = 2)
attr(object, "helloworld") <- "helloworld"
print(object)
y <- 0
z <- 1000

object |> transform_if(\(\nabla (x) \ x > y, \ log, \(\nabla (x) \ x^2, \ \(\nabla (x) - z) \)
object |> transform_if(object > y, \ log, \(\nabla (x) \ x^2, -z) \ # \ same \ as \ previous \ line
```

x.import

Helper Functions for the 'tinycodet' Package Import System

#### **Description**

The help.import() function finds the help file for functions or topics, including exposed functions/operators as well as functions in a package alias object.

The is.tinyimport() function checks if an alias object or an exposed function is of class tinyimport; i.e. if it is an object produced by the import\_as, import\_inops, or import\_LL function.

The attr.import() function gets one or all special attribute(s) from an alias object returned by import\_as.

# Usage

```
help.import(..., i, alias)
is.tinyimport(x)
attr.import(alias, which = NULL)
```

## **Arguments**

... further arguments to be passed to help.

i either one of the following:

- a function (use back-ticks when the function is an infix operator). Examples: myfun, `%operator%`, myalias.\$some\_function. If a function, the alias argument is ignored.
- a string giving the function name or topic (i.e. "myfun", "thistopic"). If a string, argument alias must be specified also.

alias the alias object as created by the import\_as function.

60 x.import

```
x the object/function to be tested.
```

which The attributes to list. If NULL, all attributes will be returned.

Possibilities: "pkgs", "conflicts", "args", and "ordered\_object\_names".

#### **Details**

```
For help.import(...):
```

Do not use the topic / package and i / alias argument sets together. It's either one set or the other.

For example:

```
import_as(~ mr., "magrittr")
import_inops(mr.)
help.import(i = mr.$add)
help.import(i = ~%>%^)
help.import(i = "add", alias = mr.)
help.import(topic = "%>%", package = "magrittr")
help.import("%>%", package = "magrittr") # same as previous line
```

#### Value

```
For help.import():
```

Opens the appropriate help page.

```
For is.tinyimport():
```

Returns TRUE if the function is produced by import\_as, import\_inops, or import\_LL, and returns FALSE if it is not.

```
For attr.import(alias, which = NULL):
```

All special attributes of the given alias object are returned as a list.

```
For attr.import(alias, which = "pkgs"):
```

Returns a list with 3 elements:

- packages\_order: a character vector of package names, giving the packages in the order they were imported in the alias object.
- main\_package: a string giving the name of the main package. Re-exported functions, if present, are taken together with the main package.
- re\_exports.pkgs: a character vector of package names, giving the packages from which the re-exported functions in the main package were taken.

```
For attr.import(alias, which = "conflicts"):
```

The order in which packages are imported in the alias object (see attribute pkgs\$packages\_order) matters: Functions from later named packages overwrite those from earlier named packages, in case of conflicts.

The "conflicts" attribute returns a data.frame showing exactly which functions overwrite functions from earlier named packages, and as such "win" the conflicts.

```
For attr.import(alias, which = "args"):
```

Returns a list of input arguments. These were the arguments supplied to import\_as when the alias

x.import 61

object in question was created.

For attr.import(alias, which = "ordered\_object\_names"):

Gives the names of the objects in the alias, in the order as they were imported.

For conflicting objects, the last imported ones are used for the ordering.

Note that if argument re\_exports is TRUE, re-exported functions are imported when the main package is imported, thus changing this order slightly.

## See Also

tinycodet\_import

# **Examples**

```
import_as(~ to., "tinycodet")
import_inops(to.)
`%s==%` <- stringi::`%s==%`

is.tinyimport(to.) # returns TRUE
is.tinyimport(`%:=%`) # returns TRUE
is.tinyimport(`%s==%`) # returns FALSE: not imported by tinycodet import system

attr.import(to., which = "conflicts")</pre>
```

# Index

| * join_mat                          | as.double,9  |
|-------------------------------------|--|
| stri_join_mat, 37                   | as.integer, 9  |
| ::, 4, 13, 14, 18                   | as.logical, 9  |
| :::, 20                             | as_bool (atomic_conversions), 9  |
| \$, <i>15</i>                       | as_chr (atomic_conversions), 9   |
| %:=%(inplace), 22                   | as_dbl (atomic_conversions), 9   |
| %<-c% (lock), 23                    | as_int (atomic_conversions), 9   |
| %=numtype% (logic_ops), 25          | Atomic type casting without stripping  |
| %=strtype% (logic_ops), 25          | attributes, 3  |
| %?=% (logic_ops), 25                | atomic_conversions, 9  |
| %[!if]% (subset_if), 53             | attaching, 4   |
| %[if]% (subset_if), 53              | attr.import, 16  |
| %col~% (matrix_ops), 27             | attr.import(x.import), 59  |
| %d!=% (decimal_truth), 11           | activity of the control of the contr |
| %d<=% (decimal_truth), 11           | base::source(), 35   |
| %d<% (decimal_truth), 11            | •  |
| %d==% (decimal_truth), 11           | Concatenate a character matrix row- or   |
| %d>=% (decimal_truth), 11           | column-wise, $7$   |
| %d>% (decimal_truth), 11            |  |
| %installed in% (pkgs), 29           | decimal_truth, 11  |
| %n&% (logic_ops), 25                | detecting patterns, 7  |
| %out% (logic_ops), 25               | outsta 10  |
| %row~% (matrix_ops), 27             | exists, 19   |
| %s-% (str_arithmetic), 44           | general in-place (mathematical)  |
| %s//%(str_arithmetic), 44           | modification operator, 8   |
| %s/%(str_arithmetic), 44            | modification operator, o   |
| %sget% (str_subset_ops), 52         | help, 59   |
| %ss% (str_arithmetic), 44           | help.import(x.import), 59  |
| %strim% (str_subset_ops), 52        | - F  |
| %unreal =% (subset_if), 53          | iif, 58  |
| %xor%(logic_ops), 25                | import_as, 5, 13, 17, 21, 59, 60   |
| %<-c%, 3                            | <pre>import_data, 5, 16</pre>  |
| %s-%, %s/%, %ss%, <i>55</i>         | import_inops, 5, 17, 19-21, 59, 60   |
| %s==%, 11, 47                       | <pre>import_inops(), 20</pre>  |
| , ,                                 | <pre>import_inops.control, 17, 19</pre>  |
| <pre>aaa0_tinycodet_help, 2</pre>   | <pre>import_inops.control(), 18</pre>  |
| aaa1_tinycodet_safer,3              | <pre>import_int, 5</pre>   |
| <pre>aaa2_tinycodet_import, 4</pre> | <pre>import_int (import_LL), 20</pre>  |
| aaa3_tinycodet_strings,7            | import_LL, 5, 20, 59, 60   |
| aaa4_tinycodet_dry,8                | Infix logical operators, 9   |
| <pre>aaa5_tinycodet_misc, 9</pre>   | Infix operators for row- and   |
| about_search_fixed, 34              | column-wise re-ordering of   |
| about_search_regex, 34              | matrices, 9  |
| as.character,9                      | inplace, 22  |
|                                     |  |

INDEX 63

| is.tinyimport(x.import), 59                      | stri_join, 38   |
|--|---|
| <pre>is_wholenumber(atomic_conversions), 9</pre> | stri_join_mat,37  |
|  | stri_locate,40  |
| library, 4                                       | stri_locate_all, <i>40</i> , <i>47</i>  |
| loadedNamespaces, 32                             | stri_locate_all_boundaries,40   |
| loadNamespace, 14–17, 21, 30, 32                 | stri_locate_first, 40   |
| lock, 23   | stri_locate_ith, 7, 36, 39, 47, 48  |
| lock_TF, 3                                       | <pre>stri_locate_ith_boundaries, 7</pre>  |
| lock_TF (lock), 23                               | stri_locate_ith_boundaries  |
| lockBinding, 20, 21, 24                          | <pre>(stri_locate_ith), 39</pre>  |
| Logic, 25  | stri_locate_ith_charclass   |
| logic_ops, 25                                    | (stri_locate_ith), 39   |
| 5 - 1 /  | stri_locate_ith_coll(stri_locate_ith,   |
| matrix_ops, 27                                   | 39  |
|  | stri_locate_ith_fixed   |
| pkg_get_deps, 14                                 | (stri_locate_ith), 39   |
| pkg_get_deps (pkgs), 29                          | stri_locate_ith_regex   |
| pkg_get_deps_minimal, 14                         | (stri_locate_ith), 39   |
| pkg_get_deps_minimal (pkgs), 29                  | stri_locate_last, 40  |
| pkg_lsf (pkgs), 29                               | stri_match, 41  |
| pkgs, 5, 29                                      | stri_opts_brkiter, 36, 40, 45   |
| pversion, <i>5</i> , 31                          | stri_opts_collator, 40, 56  |
| pversion_check4mismatch (pversion), 31           | stri_opts_fixed, 40, 56   |
| pversion_report (pversion), 31                   | stri_opts_regex, 40, 56   |
| F  | stri_paste_mat(stri_join_mat), 37   |
| Report infix operators present in the            | stri_replace_all, 47  |
| current environment, or a                        | stri_split, 36  |
| specified environment., 9                        | stri_split_boundaries, 36   |
| report_inops, 33                                 | stri_sprintf, 45  |
| report_inops(), 18                               | stri_startswith, 46, 48   |
|  | stri_sub, <i>41</i>   |
| s_chrcls(s_pattern), 54                          | string arithmetic, 7  |
| s_coll(s_pattern), 54                            | string sub-setting, 7   |
| s_fixed (s_pattern), 54                          | string sub-setting operators, 7   |
| s_pattern, 8, 45–47, 54                          | string sub-setting operators, 7 strsplit, 36                                    |
| s_regex (s_pattern), 54                          | subset_if, 53   |
| Safer decimal (in)equality testing, 3            |   |
| setv, 24   | <pre>subset_if operators and the in-place     unreal modifier operator, 8</pre> |
| source, <i>34</i>                                | unreal modifier operator, o   |
| source_selection, 9, 34                          | tinycodet(aaa0_tinycodet_help), 2   |
| str_arithmetic, 44                               | tinycodet-package   |
| str_search, 46, 55, 56                           | (aaa0_tinycodet_help), 2  |
| str_subset_ops, 52                               | tinycodet_dry, 3, 23, 54, 59  |
| strcutfunctions, 7                               | tinycodet_dry(aaa4_tinycodet_dry), 8  |
| strcut_brk, 45                                   | tinycodet_help, $3$ , $6$ , $8$   |
| strcut_brk (strcut_loc), 35                      | tinycodet_help(3,0,0) tinycodet_help(aaa0_tinycodet_help), 2                    |
| strcut_loc, 35, 45                               | tinycodet_help(), 9   |
| strfind (str_search), 46                         | tinycodet_inelp(), 9 tinycodet_import, 2, 14, 16–18, 22, 31, 32,                |
| strfind()<-, 7                                   | 61  |
| strfind()<-,/<br>strfind<- (str_search), 46      | tinycodet_import  |
| stri_c_mat (stri_join_mat), 37                   | (aaa2_tinycodet_import), 4  |
|  |   |
| stri_endswith, 46, 48                            | tinycodet_import(), 20  |
| stri_extract_all, 47                             | tinycodet_misc, $3$ , $28$ , $35$   |

INDEX