Package 'scopr'

September 3, 2017

Title Read Ethoscope Data

Date 2017-08-08

Version 0.0.0.9000

Description
Retreive behavioural data from the Ethocope platform (http://gilestrolab.github.io/ethoscope/).

Depends R (>= 3.00),
behavr

Imports data.table,
hms,
readr,
stringr,
RSQLite,
memoise

Suggests testthat

License GPL-3

Encoding UTF-8

LazyData true

URL https://github.com/rethomics/scopr

BugReports https://github.com/rethomics/scopr/issues

RoxygenNote 6.0.1

Roxygen list(markdown = TRUE)

R topics documented:

---

list_result_files   *List all available result files*

Description

This function discovers all ethoscope result files and put them in a [data.table::data.table](#). This is useful to figure out when and which experiments were performed.

Usage

```
list_result_files(result_dir, index_file = NULL)
```

Arguments

result_dir   the root directory where all data are saved

index_file   the name of an index_file, in `result_dir` (useful for loading remote data).

Value

a [data.table::data.table](#). each row is a single experimental file, and columns describe details such as: its `path`, start `date` and `time`, and the name and id of the ethoscope used.

See Also

- [query_ethoscopes](#) – to load the actual data

- [parse_query](#) – to generate a query

- [read_metadata](#) – to show the metadata of a specific experiment

Examples

```
todo
```

---

make_index_file	*Used by platform administrators to make an index file listing all experiments*

Description

This function builds an index which lists the available experimental file. It is useful to build indices when outing data on a remote server (e.g. FTP).

Usage

```
make_index_file(result_dir, index_file = "index.txt")
```

Arguments

result_dir   the root directory where all data are saved

index_file   the name of an index_file, in `result_dir` (useful for loading remote data).

Value

the path to the index file written (in `result_dir`))

---

no_region_id_query	*A simple toy query defining three experiments to query, with one condition (`test`) per experiment. Implicitly, 20 animals are in each experiment It serves as a simple example.*

Description

A simple toy query defining three experiments to query, with one condition (`test`) per experiment. Implicitly, 20 animals are in each experiment It serves as a simple example.

Usage

```
no_region_id_query
```

Format

An object of class `data.frame` with 3 rows and 4 columns.

Author(s)

Quentin Geissmann

---

parse_query *Build a query for loading ethoscope data*

Description

These functions make a query table, which is essential for subsequent loading of ethoscope data (via query_ethoscopes).

Usage

```
parse_query(x, result_dir = NULL, index_file = NULL)

parse_remote_query(x, remote_dir, result_dir, index_file = "index.txt",
  overwrite_local = FALSE, verbose = TRUE)
```

Arguments

x	object such as a data.frame, or the name of a file (see detail)
result_dir	the root directory where all data are saved
index_file	the name of an index_file, in result_dir (useful for loading remote data).
remote_dir	the url of the result directory on the data server
overwrite_local	
	whether to download all files. The default, FALSE, is to only fetch files if they are newer on the remote.
verbose	whether to print progress (a logical)

Details

These function will generate a query from three different types of inputs:

1. A vector of names of .db files (only for local queries with parse_query)

2. A data.frame (recomended) In this case, the function will try to match requested data with data available on result_dir. The provided data.table has typically one row per requested individial and the columns (not necessarily in this order):

   - machine_name – the name of the machine in which the individual was (e.g. "ETHOSCOPE_001")
   - date – the start date of the experiment formated as "YYYY-MM-DD"
   - region_id – the ROI in which the animal was. When *not provided, all regions are queried*.
   - time – the start time of the experiment formated as "HH:MM:SS". When *not provided, and multiple expriment for the same machine exist, only the last one is loaded*.
   - ??? – any number of arbitrary columns* to associate conditions/treatments/genotypes/... to the previous columns.

3. The name of a CSV file that contains a table as described in 2.

Value

a data.table::data.table. It contains the columns:

- path – the file to retreive the data from
- experiment_id – an autogenerated unique id for the experiment
- region_id – a unique identifier of each individual (region of interest) within an experiment
- Variables already present in x (if x was a data.frame).

References

- TODO – to understand the magical power of the query system!

See Also

- query_ethoscopes – to load the actual data
- list_result_files – to list available file

Examples

```
# a query with no region_id, all regions will be loaded with the same metadata
dir <- paste0(scopr_example_dir(), "/ethoscope_results/")
data(no_region_id_query)
query <- parse_query(no_region_id_query, dir)
print(query)
# a query with region_id, only stated regions will be loaded with specific metadata
data(region_id_query)
query <- parse_query(region_id_query, dir)
print(query)
```

---

query_ethoscopes   *Read data from ethoscope result files*

Description

This function is used to import behavioural data generated by the ethoscope platform. That is it loads multiple .db`` files into a singleR` behavr::behavr table.

Usage

```
query_ethoscopes(query, min_time = 0, max_time = Inf,
  reference_hour = NULL, verbose = TRUE, columns = NULL, cache = NULL,
  ncores = 1, FUN = NULL, ...)
```

Arguments

query	[data.table::data.table](#) obtained by [parse_query](#)
min_time	Exclude data before min_time (in seconds). This time is *relative to the start of the experiment*.
max_time	Exclude data after max_time (in seconds). It is also relative to the start of the experiment.
reference_hour	Hour, in the day, to use as ZT0 reference. When unspecified, time will be relative to the start of the experiment.
verbose	whether to print progress (a logical)
columns	optionnal vector of columns to be selected from the db file. Time (t) is always implicitly selected.
cache	the name of a local directory to cache results for faster subsequent data loading
ncores	number of cores to use for optionnal parallel processing
FUN	function (optional) to transform the data from each individual immediately after is has been loaded.
...	extra arguments to be passed to FUN

Details

todo

Value

a [behavr](#) table. The metadata contains all the query columns and an autogenerated id per animal. The data has the columns:

- id – autogenerated unique identifier, one per animal
- t – time
- Several variables recorded by ethoscopes (position, angle, width/height and others). Distance units (e.g. xy position, height/width) are expressed as a fraction of the width of the ROI they originate from.

References

- todo

See Also

- [behavr::behavr](#) – to understand the output format
- [parse_query](#) – to generate a query
- [read_metadata](#) – to show the metadata of a specific experiment
- [list_result_files](#) – to list available files

Examples

```
dir <- paste0(scopr_example_dir(), "/ethoscope_results/")
data(region_id_query)
query <- parse_query(region_id_query, dir)
print(query)

# default data loading
dt <- query_ethoscopes(query)
dt

# we use reference hour to set zt0 to 09:00 GMT
dt <- query_ethoscopes(query, reference_hour=9)
dt

#' # only load x and y positions
dt <- query_ethoscopes(query, columns=c("x", "y"), reference_hour=9)
dt
# apply function whilst loading the data
# todo use sleepr
dt <- query_ethoscopes(query, reference_hour=9, FUN=head)
dt
```

---

read_metadata   *Retrieve metadata from an ethoscope result file*

Description

This function is used to obtain metadata – such as *time and date* of the experiment, *acquisition device*, and *version of the software* – embedded in a result file (.db) generated by the ethoscope platform.

Usage

```
read_metadata(FILE)
```

Arguments

FILE   the name of the input file

Value

a list containing fields for metadata entries

See Also

- query_ethoscopes – to load the actual data
- list_result_files – to list available files

region_id_query	*A simple toy query experimental conditions for 15 animals in three experiments. the condition (*test*) is linked to the experiments, and another condition (*treatment*) is per animal. It serves as a simple example. #'*

Description

A simple toy query experimental conditions for 15 animals in three experiments. the condition (test) is linked to the experiments, and another condition (treatment) is per animal. It serves as a simple example. #'

Usage

```
region_id_query
```

Format

An object of class data.frame with 15 rows and 6 columns.

Author(s)

Quentin Geissmann

---

scopr_example	*Get path to scopr example*

Description

scopr comes with a sample DAM2 files in its inst/extdata directory. scopr_example allow make them easy to access.

Usage

```
scopr_example(path = NULL)

scopr_example_dir()
```

Arguments

path   Name of file. If NULL, the example files will be listed.

Author(s)

Hadley Wickham (modified from readr)

Examples

```
# list all files
scopr_example()
```

Index