

Package ‘TableToLongForm’

January 10, 2019

Type Package

Title Automatically Convert Hierarchical for-Human Tables to
Machine-Readable LongForm Dataframes

Version 1.3.2

Date 2019-01-07

Description

A wrapper to a set of algorithms designed to recognise positional cues present in hierarchical for-human Tables (which would normally be interpreted visually by the human brain) to decompose, then reconstruct the data into machine-readable LongForm Dataframes.

License GPL-3

URL <https://www.stat.auckland.ac.nz/~joh024/Research/TableToLongForm/>

NeedsCompilation no

Author Jimmy Oh [aut, cre]

Maintainer Jimmy Oh <joh024@aucklanduni.ac.nz>

Repository CRAN

Date/Publication 2019-01-10 08:00:03 UTC

R topics documented:

TableToLongForm-package	2
print.plist	2
TableToLongForm	3
TCData	4
TTLFaliasAdd	5
TTLFaliasList	6
Index	7

TableToLongForm-package

Convert a Table to a LongForm data.frame

Description

TableToLongForm automatically converts hierarchical Tables intended for a human reader into a simple LongForm Dataframe that is machine readable.

Details

Package: TableToLongForm
 Type: Package
 Version: 1.3.1
 Date: 2014-08-01
 License: GPL-3

Call TableToLongForm() on a Table to automatically convert it to a LongForm data.frame.

Examples of Tables that can be converted are found in data(TCData).

For more details on what TableToLongForm does and what sorts of Tables it can convert, refer to the website: <https://www.stat.auckland.ac.nz/~joh024/Research/TableToLongForm/>

Available help: help(TableToLongForm) help(TCData)

Author(s)

Jimmy Oh

Maintainer: Jimmy Oh <joh024@aucklanduni.ac.nz>

print.plist

Print Method for plist Objects

Description

A print method for class `plist`, which are nested lists with a numeric vector at the lowest level, used as `print.default` is rather inefficient (and much uglier) when displaying such nested lists.

Arguments

`x` a `plist` object.

`...` potential further arguments (required by generic), unused by this method

Details

plist objects are created as part of diagnostic output for TableToLongForm. For more information, refer to the website: <https://www.stat.auckland.ac.nz/~joh024/Research/TableToLongForm/>

TableToLongForm	<i>Convert a Table to a LongForm data.frame</i>
-----------------	---

Description

TableToLongForm automatically converts hierarchical Tables intended for a human reader into a simple LongForm Dataframe that is machine readable.

Use this function to run TableToLongForm on the specified matrix Table. All other arguments are optional.

Once the conversion is complete, the user is recommended to check the result for correctness and to consider tidying up the variable names.

Usage

```
TableToLongForm(Table, IdentResult = NULL,
                 IdentPrimary = "compound",
                 IdentAuxiliary = "sequence",
                 ParePreRow = NULL,
                 ParePreCol = c("mismatch", "misalign", "multirow"),
                 fulloutput = FALSE,
                 diagnostics = FALSE, diagnostics.trim = TRUE)
```

Arguments

Table	the Table to convert, given as a character matrix. Also accepts a data.frame, which is coerced to a matrix with a warning.
IdentResult	an optional list specifying the locations of the various elements of the Table. By default this is automatically generated but it can be specified manually where the automatic detection fails.
IdentPrimary	The Primary Ident algorithm, of which one is chosen. See details.
IdentAuxiliary	Auxiliary Ident algorithms, of which any combination, in any order, can be chosen. They are called after the Primary algorithm, to refine the IdentResult. See details.
ParePreRow	Pre-requisite algorithms that tidy up the Row Labels for correct operation of the Main Parentage algorithm. Any combination of these algorithms, in any order, can be chosen. See details.
ParePreCol	Pre-requisite algorithms that tidy up the Column Labels for correct operation of the Main Parentage algorithm. Any combination of these algorithms, in any order, can be chosen. See details.

fulloutput	if TRUE, returns a list containing additional information primarily useful for diagnostic purposes. Otherwise, and by default, the function only returns the converted data.frame object.
diagnostics	a character vector specifying the name of the file diagnostic output will be written to. Can also be TRUE, in which case the file name will be the name of the object specified in Table.
diagnostics.trim	a logical indicating whether the diagnostics output should be trimmed. A good idea to keep TRUE (default) as trimmed output is generally more useful.

Details

For more details on TableToLongForm refer to the website: <https://www.stat.auckland.ac.nz/~joh024/Research/TableToLongForm/>

Specifically, the 'Technical Report' gives a rounded introduction to TableToLongForm, including a short user manual, some examples and a complete gallery of recognised patterns.

'Working with Modules' gives an introduction to creating new modules/algorithms for TableToLongForm, to extend its capabilities.

Finally, the Literate Document has the complete documentation of the source code for TableToLongForm.

Value

The converted Table as a data.frame object.

Examples

```
## load Toy Examples data
data(TCData)

## Convert ToyExComplete
TableToLongForm(TCData$ToyExComplete)
```

TCData

Example hierarchical Tables

Description

A list containing a number of example Tables that can be converted to LongForm dataframes by TableToLongForm. Each Tables is stored as a character matrix.

These datasets are generally not immediately useful as data, as they must first be converted (e.g. by using TableToLongForm).

If the user prefers to have these data in their Global Environment rather than nested inside a single list, they can use `attach(TCData)`.

Usage

```
data(TCData)
```

Format

list containing character matrices of varying size.

Source

Department of Internal Affairs (New Zealand) (2012) New Zealand Qualifications Authority (2012)
 Statistics New Zealand (2013)

Examples

```
data(TCData)

## list all Tables
names(TCData)

## One such Toy Example Table
TCData$ToyExByEmptyBelow
```

TTLFaliasAdd

Register a new Module to TableToLongForm

Description

TableToLongForm is partially modular and can be extended in some ways with external modules. Registration of these modules with this function is necessary.

Arguments

Type	e.g. IdentPrimary
Fname	the name of the Function/Algorithm
Falias	the alias for the Function/Algorithm, which is used for the call to TableToLongForm
Author	(optional) name of the author of the algorithm
Description	(optional) a short description of the purpose of the algorithm

Details

For more details on modules, refer to the “Working with Modules” document on the website:
<https://www.stat.auckland.ac.nz/~joh024/Research/TableToLongForm/>

`TTLFaliasList`*List registered Modules for TableToLongForm*

Description

TableToLongForm is partially modular and can be extended in some ways with external modules. This function is used to list currently registered modules.

Details

For more details on modules, refer to the “Working with Modules” document on the website: <https://www.stat.auckland.ac.nz/~joh024/Research/TableToLongForm/>

Index

*Topic **datasets**

TCData, [4](#)

*Topic **package**

TableToLongForm-package, [2](#)

print.plist, [2](#)

TableToLongForm, [3](#)

TableToLongForm-package, [2](#)

TCData, [4](#)

TTLFaliasAdd, [5](#)

TTLFaliasList, [6](#)