The Data Frame Input/Output Reference Manual

Common Lisp library for reading data from text files (eg CSV), version 1.3

Steve Nunez <steve@symbolics.tech>
Tamas Papp <tkpapp@gmail.com>

Table of Contents

1 Systems

The main system appears first, followed by any subsystem dependency.

1.1 dfio

Maintainer
Steve Nunez <steve@symbolics.tech>

Author Tamas Papp <tkpapp@gmail.com>

License MS-PL

Description
Common Lisp library for reading data from text files (eg CSV).

Version 1.3

Dependencies

- alexandria
- anaphora
- cl-csv
- data-frame
- let-plus

Source [dfio.asd], page 3, (file)

Directory s:/src/dfio/

Components

- [decimal.lisp], page 3, (file)
- [string-table.lisp], page 3, (file)
- [data-column.lisp], page 4, (file)
- [dfio.lisp], page 4, (file)

2 Files

Files are sorted by type and then listed depth-first from the systems components trees.

2.1 Lisp

2.1.1 dfio.asd

Location dfio.asd

Systems [dfio], page 1, (system)

2.1.2 dfio/decimal.lisp

Parent [dfio], page 1, (system)

Location decimal.lisp

Packages [dfio.decimal], page 5,

Exported Definitions

- [parse-rational], page 8, (function)
- [parse-rational-error], page 10, (condition)
- [parse-real], page 8, (function)

Internal Definitions

- [+exponent-chars+], page 12, (special variable)
- [gobble-positive-integer], page 12, (function)
- [gobble-sign], page 12, (function)

2.1.3 dfio/string-table.lisp

Dependency
[decimal.lisp], page 3, (file)

Parent [dfio], page 1, (system)

Location string-table.lisp

Packages [dfio.string-table], page 5,

Exported Definitions

- [string-table], page 9, (function)
- [string-table], page 10, (structure)
- [string-table-add], page 9, (function)
- [string-table-count], page 9, (function)
- [string-table-duplicate], page 10, (condition)
- [string-table-intern], page 9, (function)
- [string-table-lookup], page 9, (function)
- [string-table-not-found], page 10, (condition)
- [string-table-strings], page 9, (function)

Internal Definitions

- [copy-string-table], page 12, (function)
- [string-table-get], page 12, (function)
- [(setf string-table-get)], page 13, (function)

- [string-table-p], page 13, (function)
- [string-table-table], page 13, (function)
- [(setf string-table-table)], page 13, (function)

2.1.4 dfio/data-column.lisp

Dependency
[string-table.lisp], page 3, (file)

Parent [dfio], page 1, (system)

Location data-column.lisp

Packages [dfio.data-column], page 6,

Exported Definitions

- [data-column], page 7, (function)
- [data-column], page 11, (class)
- [data-column-add], page 7, (function)
- [data-column-counts], page 7, (function)
- [data-column-vector], page 7, (function)

Internal Definitions
[non-negative-integer], page 13, (type)

2.1.5 dfio/dfio.lisp

Dependency
[data-column.lisp], page 4, (file)

Parent [dfio], page 1, (system)

Location dfio.lisp

Packages [dfio], page 6,

Exported Definitions

- [csv-to-data-frame], page 7, (function)
- [data-frame-to-csv], page 7, (function)
- [string-to-keyword], page 9, (function)

Internal Definitions

- [2d-array-to-list], page 12, (function)
- [csv-to-data-columns], page 12, (function)

3 Packages

Packages are listed by definition order.

3.1 dfio.decimal

Source [decimal.lisp], page 3, (file)

Use List

- let-plus
- anaphora
- common-lisp

Used By List
[dfio.data-column], page 6,

Exported Definitions

- [parse-rational], page 8, (function)
- [parse-rational-error], page 10, (condition)
- [parse-real], page 8, (function)

Internal Definitions

- [+exponent-chars+], page 12, (special variable)
- [gobble-positive-integer], page 12, (function)
- [gobble-sign], page 12, (function)

3.2 dfio.string-table

Source [string-table.lisp], page 3, (file)

Use List

- let-plus
- anaphora
- alexandria
- common-lisp

Used By List
[dfio.data-column], page 6,

Exported Definitions

- [string-table], page 9, (function)
- [string-table], page 10, (structure)
- [string-table-add], page 9, (function)
- [string-table-count], page 9, (function)
- [string-table-duplicate], page 10, (condition)
- [string-table-intern], page 9, (function)
- [string-table-lookup], page 9, (function)
- [string-table-not-found], page 10, (condition)
- [string-table-strings], page 9, (function)

Internal Definitions

- [copy-string-table], page 12, (function)

- [string-table-get], page 12, (function)
- [(setf string-table-get)], page 13, (function)
- [string-table-p], page 13, (function)
- [string-table-table], page 13, (function)
- [(setf string-table-table)], page 13, (function)

3.3 dfio.data-column

Source [data-column.lisp], page 4, (file)

Use List

- let-plus
- [dfio.string-table], page 5,
- [dfio.decimal], page 5,
- anaphora
- common-lisp

Used By List

- vglt
- [dfio], page 6,

Exported Definitions

- [data-column], page 7, (function)
- [data-column], page 11, (class)
- [data-column-add], page 7, (function)
- [data-column-counts], page 7, (function)
- [data-column-vector], page 7, (function)

Internal Definitions

[non-negative-integer], page 13, (type)

3.4 dfio

Source [dfio.lisp], page 4, (file)

Use List

- [dfio.data-column], page 6,
- let-plus
- cl-csv
- anaphora
- alexandria
- common-lisp

Used By List

vglt

Exported Definitions

- [csv-to-data-frame], page 7, (function)
- [data-frame-to-csv], page 7, (function)
- [string-to-keyword], page 9, (function)

Internal Definitions

- [2d-array-to-list], page 12, (function)
- [csv-to-data-columns], page 12, (function)

4 Definitions

Definitions are sorted by export status, category, package, and then by lexicographic order.

4.1 Exported definitions

4.1.1 Functions

`csv-to-data-frame` *STREAM-OR-STRING* **&key** *SKIP-FIRST-ROW? COLUMN-KEYS-OR-FUNCTION* [Function]

Read a CSV file (or stream, or string) into a DATA-FRAME, which is returned.

When SKIP-FIRST-ROW?, the first row is read separately and COLUMN-KEYS-OR-FUNCTION is used to form column keys.

When COLUMN-KEYS-OR-FUNCTION is a sequence, it is used for column keys, regardless of the value of SKIP-FIRST-ROW?.

Package [`dfio`], page 6,

Source [`dfio.lisp`], page 4, (file)

`data-column` **&key** *MAP-ALIST DEFAULT-FLOAT-FORMAT* [Function]

Package [`dfio.data-column`], page 6,

Source [`data-column.lisp`], page 4, (file)

`data-column-add` *DATA-COLUMN STRING* [Function]

Package [`dfio.data-column`], page 6,

Source [`data-column.lisp`], page 4, (file)

`data-column-counts` *DATA-COLUMN* [Function]

Return the counts.

Package [`dfio.data-column`], page 6,

Source [`data-column.lisp`], page 4, (file)

`data-column-vector` *DATA-COLUMN* [Function]

Return the collected elements as a vector.

Package [`dfio.data-column`], page 6,

Source [`data-column.lisp`], page 4, (file)

`data-frame-to-csv` *DF* **&key** *STREAM ADD-FIRST-ROW* (*SEPARATOR* ***SEPARATOR***) (*QUOTE* ***QUOTE***) (*ESCAPE* ***QUOTE-ESCAPE***) (*NEWLINE* ***WRITE-NEWLINE***) (*ALWAYS-QUOTE* ***ALWAYS-QUOTE***) [Function]

Write a data-frame to a stream.

Keywords:
stream: stream to write to. Default: nil.
nil - writes the rows to a string and returns it
an open stream
a pathname (overwrites if the file exists)
quote: quoting character. Defaults to *quote*

escape: escaping character. Defaults to *quote-escape*
newline: newline character. Defaults to *write-newline*
always-quote: Defaults to *always-quote*
add-first-row: Add column names as the first

Notes:
The :newline keyword requires a sequence, so use :newline '(#newline) or use cl-interpol

Package [`dfio`], page 6,

Source [`dfio.lisp`], page 4, (file)

`parse-rational` *STRING* **&key** *START END EXPONENT-CHARS* [Function]

Parse a decimal rational in (subseq string start end) of the form [sign][whole][.[fraction]][exponent] where

sign ::= `+` | `-` | empty
whole ::= digit*
fraction ::= digit*
exponent ::= exponent-char[sign]digit`+`

with the restriction that WHOLE and FRACTION cannot be empty at the same time. EXPONENT-CHAR is a string and contains the valid exponent chars.

Whitespace is NOT trimmed, and leads to an error. In case of a parsing failure, PARSE-RATIONAL-ERROR is used.

Return (values NUMBER DECIMAL-DOT? EXPONENT-CHAR). NUMBER is a RATIONAL, DECIMAL-DOT? is T when a decimal dot is present, otherwise NIL, EXPONENT-CHAR contains the exponent character, NIL if not present.

Numbers of the form .112 and 112. are valid syntax, representing 0.112 and 112.0, respectively.

Examples:

(parse-rational "7") => (values 7 NIL NIL)
(parse-rational "7.") => (values 7 T NIL)
(parse-rational "0.7") => (values 7/10 T NIL)
(parse-rational ".7") => (values 7/10 T NIL)
(parse-rational "7.e2") => (values 700 T #e)
(parse-rational ".7d1") => (values 7 T #d)

Package [`dfio.decimal`], page 5,

Source [`decimal.lisp`], page 3, (file)

`parse-real` *STRING* **&key** *START END S-FLOAT F-FLOAT D-FLOAT L-FLOAT E-FLOAT* [Function]

Wrapper for PARSE-RATIONAL, converting non-integers to floats. The float type is determined by the -float arguments for each exponent character. Integers are not converted to floats. Return a single value, type of (or integer float).

See PARSE-RATIONAL for accepted formats, errors, etc.

Package [`dfio.decimal`], page 5,

Source [`decimal.lisp`], page 3, (file)

`string-table` **&key** (*TABLE* **TABLE**) [Function]

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

`string-table-add` *STRING-TABLE STRING* **&optional** *VALUE* [Function]

Add STRING mapped to VALUE to STRING-TABLE, raising STRING-TABLE-DUPLICATE if STRING is already in the table. Return VALUE.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

`string-table-count` *STRING-TABLE* [Function]

Number of distinct strings in the table.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

`string-table-intern` *STRING-TABLE STRING* **&optional** *NEW-VALUE* [Function]

If STRING is already in STRING-TABLE, return its value, otherwise add it and return NEW-VALUE. When used with the default argument for NEW-VALUE, EQUAL strings are always mapped to values that are EQ.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

`string-table-lookup` *STRING-TABLE STRING* [Function]

Return the value corresponding to STRING in STRING-TABLE, or raise the STRING-TABLE-NOT-FOUND error.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

`string-table-strings` *STRING-TABLE* [Function]

List of strings in STRING-TABLE.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

`string-to-keyword` *STRING* [Function]

Map string to a keyword.

This is the default for constructing column keys for CSV files.

The current implementation replaces #. and #space with a #-, and upcases all other characters.

Package [`dfio`], page 6,

Source [`dfio.lisp`], page 4, (file)

4.1.2 Conditions

`parse-rational-error` () [Condition]

Error used by parse-rational and parse-real.

Package [`dfio.decimal`], page 5,

Source [`decimal.lisp`], page 3, (file)

Direct superclasses
`error` (condition)

Direct slots

`string` [Slot]

Initform `(quote :string)`

`message` [Slot]

Initform `(quote :message)`

`string-table-duplicate` () [Condition]

String is already in the table.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

Direct superclasses
`error` (condition)

`string-table-not-found` () [Condition]

String not found in table.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

Direct superclasses
`error` (condition)

4.1.3 Structures

`string-table` () [Structure]

A table of distinct strings, optionally mapping each one to a value.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

Direct superclasses
`structure-object` (structure)

Direct methods
`print-object` (method)

Direct slots

`table` [Slot]

Type `hash-table`

Initform `(make-hash-table :test (function equalp))`

Readers [`string-table-table`], page 13, (function)

Writers [`(setf string-table-table)`], page 13, (function)

4.1.4 Classes

`data-column` () [Class]

Package [`dfio.data-column`], page 6,

Source [`data-column.lisp`], page 4, (file)

Direct superclasses
`standard-object` (class)

Direct slots

`reverse-elements` [Slot]

Type `list`

`default-float-format` [Slot]

Type `symbol`

Initargs `:default-float-format`

`float-count` [Slot]

Type `dfio.data-column::non-negative-integer`

Initform `0`

`integer-count` [Slot]

Type `dfio.data-column::non-negative-integer`

Initform `0`

`integer-min` [Slot]

Type `integer`

Initform `0`

`integer-max` [Slot]

Type `integer`

Initform `0`

`map-count` [Slot]

Type `dfio.data-column::non-negative-integer`

Initform `0`

`map-table` [Slot]

Type `dfio.string-table:string-table`

Initargs `:map-table`

`string-count` [Slot]

Type `dfio.data-column::non-negative-integer`

Initform `0`

`string-table` [Slot]

Type `dfio.string-table:string-table`

Initform `(dfio.string-table:string-table)`

4.2 Internal definitions

4.2.1 Special variables

`+exponent-chars+` [Special Variable]

Default exponent characters.

Package [`dfio.decimal`], page 5,

Source [`decimal.lisp`], page 3, (file)

4.2.2 Functions

`2d-array-to-list` *ARRAY* [Function]

Helper for CSV writing.

Package [`dfio`], page 6,

Source [`dfio.lisp`], page 4, (file)

`copy-string-table` *INSTANCE* [Function]

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

`csv-to-data-columns` *STREAM-OR-STRING SKIP-FIRST-ROW?* [Function]

Read a CSV file (or stream, or string), accumulate the values in DATA-COLUMNs, return a list of these. Rows are checked to have the same number of elements.

When SKIP-FIRST-ROW?, the first row is read separately and returned as the second value (list of strings), otherwise it is considered data like all other rows.

Package [`dfio`], page 6,

Source [`dfio.lisp`], page 4, (file)

`gobble-positive-integer` *STRING START END* [Function]

If (SUBSEQ STRING START END) starts with a nonnegative integer (ie a sequence of digits 0-9), return the integer and position at which it ends as two values.

Otherwise, return NIL and 0.

START < END has to hold, END cannot be NIL. Consequences are undefined when START >= END.

Package [`dfio.decimal`], page 5,

Source [`decimal.lisp`], page 3, (file)

`gobble-sign` *STRING START* [Function]

Return (values SIGNUM INDEX), where SIGNUM is -1 or 1 depending on whether (CHAR STRING START) was a sign, and INDEX is the index of the subsequent character (START or START+1).

Package [`dfio.decimal`], page 5,

Source [`decimal.lisp`], page 3, (file)

`string-table-get` *STRING-TABLE STRING* [Function]

Synonym for GETHASH, used internally.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

Writer [`(setf string-table-get)`], page 13, (function)

`(setf string-table-get)` *VALUE STRING-TABLE STRING* [Function]

Synonym for (SETF GETHASH), used internally, checks that STRING is a string.

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

Reader [`string-table-get`], page 12, (function)

`string-table-p` *OBJECT* [Function]

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

`string-table-table` *INSTANCE* [Function]
`(setf string-table-table)` *VALUE INSTANCE* [Function]

Package [`dfio.string-table`], page 5,

Source [`string-table.lisp`], page 3, (file)

4.2.3 Types

`non-negative-integer ()` [Type]

Package [`dfio.data-column`], page 6,

Source [`data-column.lisp`], page 4, (file)

Appendix A Indexes

A.1 Concepts

A.2 Functions

A.3 Variables

A.4 Data types