C CSV Parser Documentation

# Overview

The included files provide functions to parse CSV (or other character separated) files and/or strings into an accessible structure in the C programming language. The underlying parsed structures make use of doubly linked lists and are allocated on the heap of the running program.

The reason I wrote this CSV parser for C is its ease of access. While there are several other CSV parsers available, I could not find anyone that was as convenient as calling a function on a string or filename parameter and the parsed CSV structure being returned. Most of the parsers I found were more complicated than what I needed.

This documentation is intended to provide information on how to use the provided CSV parser.

# Basic Tutorial

Once the header file and the implementation file are included in your project, CSV files and strings can be parsed using the following functions:

|  |
| --- |
| struct csv\_table \* parse\_string\_to\_csv\_table(char \* string, char delim, char quot\_char, int strip\_spaces, int discard\_empty\_cells);  struct csv\_table \* open\_and\_parse\_file\_to\_csv\_table(char \* filename, char delim, char quot\_char, int strip\_spaces, int discard\_empty\_cells); |

The functions take the string to parse (string) or the address of the file to read and parse (filename). The other function parameters control the parser behaviour:

|  |  |
| --- | --- |
| **Parameter** | **Description** |
| delim | This is the character used to separate values. For CSVs it is , but it can be set to any symbol. |
| quot\_char | Character used for making quotes, e.g. " or ' |
| strip\_spaces | Boolean flag. If true (1), parsed CSV values will be stripped of leading and trailing spaces. |
| discard\_empty\_cells | Boolean flag. If true (1), any value that is an empty string ("") will not be added to the parsed structure. |

*Note: If the delimiter is a space character* (' ')*,* strip\_spaces *and* discard\_empty\_cells *will be overridden to* 0 *(false) and* 1 *(true) respectively.*

The return value is a pointer to the structure used to store the parsed CSV values, a struct csv\_table.

struct csv\_table is the highest level of the parsed CSV structure. It is made up of a doubly linked list of struct csv\_row, which are in turn made up of a doubly linked list of struct csv\_cell. The struct csv\_cell contains the appropriate parsed string value in its str field.

Diagram

Description automatically generated

Figure 1 - Parsed CSV Structure Layout

The structure is designed to mirror the table layout in the unparsed CSV, and provides functions to get a row/cell at a specific row and column index (see <some link here>).

The parsed CSV structure is allocated on the heap and therefore must be freed manually. The header files include free functions for each of the CSV structures to make sure all allocated memory is freed (see <some link here>).

# In-depth Tutorial

## Creating CSV Structures

## Freeing CSV Structures

## Getting Nested CSV Structures

## Combining CSV Structures

## Separating and Deleting CSV Structures

## Comparing CSV Structures

## Checking for Strings in CSV Structures

## Iterating Through Rows/Cells