Markdown: Probably the most Easy-to-Use Markup Language

Tairan Li
Nanjing University
Nanjing, Jiangsu, China
tairanli@smail.nju.edu.cn

Abstract

Document editing is a basic need in work and academic occasions. For document editing, most people go for sophisticated softwares like Microsoft Word. However, these softwares have problems in terms of excessive file size and inconvenience with complicated formats. A type of programming language, called **markup language** (ML), can also serve the purpose of describing the content, format and structure of documents, while avoiding the aforementioned problems. This article, in particular, is going to dig into a quite favored markup language — **Markdown**, whose simplicity makes it stand out for most daily document-editing tasks. The simplicity of Markdown comes in the following aspects. First, "Easy to read and easy to write" is the core design principle of Markdown. Markdown's source file is highly readable as plain text. Meanwhile, it has clear, rich and extensible syntax that makes writing convenient. Also, Markdown's source file can be easily converted to equivalent HTML or exported as PDF, with softwares and tools facilitating its editing, rendering and conversion. It's no exaggerate to say that Markdown is building up a new ecology for document editing, with many programmers, writers and bloggers select it as the pivot of their workflow.

Keywords: Document Editing, Markup Language, Markdown

1 What is Markup Language

Document editing is an indispensable part in our daily life. Here by document I'm referring to any file that is intended to display a set of text, presumably in a certain format. When it comes to document editing, the first things that come into mind might be softwares like Microsoft Word. These are sophisticated softwares tailored to document editing and layout, providing a what-you-see-is-what-you-get (WYSIWYG) interface visualizing the end effect of documents, the reason for which we turn to them almost every day.

However, these softwares are not perfect in all cases. When the text is short and format isn't important, format consumes more space in the file than the text itself. In a complicated layout, clicking and dragging can mess up previously assembled contents. Also, these document files cannot be properly opened by text editors, as they are encoded in the binary form instead of plain text, making them more vulnerable to corruption.

An alternative to these document-processing softwares is the so-called markup language, which can effectively resolve the problems of traditional document files. **Markup language** [3] (ML) is a type of programming language that describes the content, format and structure of documents with plain text. Different from typical programming languages, they cannot execute and give computational results. The source file of markup languages is a combination of natural language and specialized tokens marking format and layout. Some well-known markup languages are listed as follows:

- **HTML** (HyperText Markup Language), the foundation of webpages
- **RTF** (Rich Text Format), a proprietary format developed by Microsoft
- **WikiText** [5] , powered by MediaWiki, employed by WikiPedia and similar online encyclopedias
- **LaTeX**, used for writing maths formulas and publishing essays, in which this article is written
- **Markdown**, which is the protagonist of this article that I will later introduce

Markup languages' source files are in plain text format, which reduces residual information, make them easier to maintain and more resilient to corruption. Document layout described by MLs are more stable and less sensitive to unwanted click-and-drag alterations. Also, setting special formats can be easier done with syntax than selecting text and searching for options in the crowded menu bar.

In the subsequent part of this article, I will introduce the markup language Markdown and discuss what makes it stand out as probably "the most easy-to-use markup language".

2 Markdown and Its Advantages

2.1 Markdown's Design Principle

Among all markup languages, **Markdown** [2], with file extension `.md`, might be the one most convenient to use and

easy to the eye. Markdown is categorized as a **lightweight markup language** [1] (LML), indicating its closeness to plain text, high source code readability and writing efficiency.

Readability is Markdown's key design principle. Markdown's creator, John Gruber, invented this markup language in 2004, with the aim of enabling people "to write using an easy-to-read and easy-to-write plain text format, optionally convert it to structurally valid XHTML (or HTML)" [9]. The syntax of Markdown is mainly inspired by pre-existing conventions for marking up plain text email and usenet posts, which include some early markup languages such as setex, Textile, reStructuredText and atx.

2.2 Syntax Simplicity and Extendability

Markdown achieves its "easy-to-read and easy-to-write" principle by providing a clear, rich and extensible syntax that blends in well with the surrounding context. Formats look obvious in its plain text form, so even if you are temporarily unable to render the document to see its output, you can still grab what it want to convey by peeking the source file. Compared to Markdown, the HTML tags are too complicated, RTF and LaTeX requires too many special commands preceding by backslash, making their source files harder to read, and even WikiText, which is already very close to plain text, still looks too heavily-marked with its use of multi-layer single quotes and brackets. Markdown is the simplification of HTML, reducing the most-used formats to human-readable text-based syntax while not losing HTML's functionality.

With syntax simplicity, Markdown can relieve you from the bother of formatting and layout of traditional document-processing softwares, allowing you to focus on the text itself and avoid distractions. Writing document with Markdown is almost no difference from writing plain text file except occasional markup for special formats like bold and italic. More complicated formats, like bullet lists, block quotes, links, images and code fragments, are also supported with highly intuitive syntax.

The original syntax of Markdown was published on Gruber's personal website, daringfireball.net [10], but it was only an informal document and contains unsettled ambiguities. Therefore, a later specification called **CommonMark** [7][8] was released to clear out ambiguities and set a standard. CommonMark's standardization effort makes the rendering effect of Markdown files more stable and predictable, hence rapidly gained in popularity among burgeoning websites and projects.

Table 1 is a showcase of the most frequently used Markdown syntax, based on the CommonMark standard. A sample document is attached in the appendix.

Most Markdown implementations today conform to the CommonMark standard, while some have extended it to power up more complicated features. Below are some examples:

Table 1. Frequently Used Markdown Syntax

Format	Syntax	Equv. HTML Tag
Italic	`*text*`	`<em>`
Bold	`**text**`	`<strong>`

Heading	`# Level 1`
	`## Level 2`
	`###### Level 6`
		`<h1>`,`<h2>`,
		$\cdots$,`<h6>`

Bullet List	`* Item 1`
	`* Item 2`
	`* Item 3`
		`<ul>`,`<li>`

Ordered List	`1. Item 1`
	`1. Item 2`
	`1. Item 3`
		`<ol>`,`<li>`

Inline Code	`` `code` ``	`<code>`
Hyperlink	`[text](url)`	`<a href="url">`
Block Quote	`> text`	`<blockquote>`

- GitHub has been providing its own variant of Markdown — **GitHub Flavored Markdown** [6] (GFM) based on CommonMark since 2017, adding new features like tables, strikethrough and to-do lists.
- **MultiMarkdown** [4][11] is an "upgraded" version of Markdown, with file extension .mmd, providing syntax for superscript, subscript, footnotes and LaTeXformulas, *etc.*, in addition to CommonMark.

2.3 Other Advantages

In Markdown's design principle, Gruber mentioned that Markdown can convert to structurally valid HTML file. In fact, more than that, with proper tools, softwares or extensions, Markdown is convertible to almost all popular document formats, such as .docx file, PDF, and e-book formats like ePub. Simple format exchange makes Markdown compatible with non-Markdown-centered workflow.

Rich softwares and extensions support is also what makes Markdown awesome. Among all softwares and extensions that facilitate the editing process of Markdown, there is one worthy of note, — the VSCode extension Markdown Preview Enhanced [12] (MPE). The Markdown standard this extension adopts is a fusion of GFM and MultiMarkdown. Beyond that, the extension supports customizable CSS stylesheet, importing outer files, and creating graphs with tools like Mermaid and Vega Lite. The most powerful feature the extension provides, however, is to influence the way your document is processed by editing a parser.js script. With this parser script, you can even define your own syntax apart from the Markdown standard and take advantage of Markdown's syntax extandability to the utmost.

References

[1] [n. d.]. WikiPedia - Lightweight Markup Language. https://en.wikipedia.org/wiki/Lightweight_markup_language

[2] [n. d.]. WikiPedia - Markdown. https://en.wikipedia.org/wiki/Markdown

[3] [n. d.]. WikiPedia - Markup Language. https://en.wikipedia.org/wiki/Markup_language

[4] [n. d.]. WikiPedia - MultiMarkdown. https://en.wikipedia.org/wiki/MultiMarkdown

[5] [n. d.]. WikiPedia - Wiki. https://en.wikipedia.org/wiki/Wiki

[6] 2019. GitHub Flavored Markdown Spec, Version 0.29-gfm. https://github.github.com/gfm/

[7] Jeff Atwood and John MacFarlane's group. 2014. CommonMark. https://commonmark.org/

[8] Jeff Atwood and John MacFarlane's group. 2021. CommonMark Spec, Version 0.30. https://spec.commonmark.org/0.30/

[9] John Gruber. 2004. Daring Fireball - Markdown. https://daringfireball.net/projects/markdown/

[10] John Gruber. 2004. Daring Fireball - Markdown: Syntax. https://daringfireball.net/projects/markdown/syntax

[11] Fletcher T. Penney. [n. d.]. MultiMarkdown v6 Syntax Guide. https://rawgit.com/fletcher/MultiMarkdown-6-Syntax-Guide/master/index.html

[12] Yiyi Wang. [n. d.]. Markdown Preview Enhanced document. https://shd101wyy.github.io/markdown-preview-enhanced/#/

Appendix

A Sample Markdown Document

Below is the source code of a sample Markdown document, from which you can clearly feel the simplicity of Markdown's syntax.

```
# A Brief Introduction to Markdown

**Markdown** is probably the most *easy-
    to-use* markup language. It largely
    simplifies the process of document-
    editing.

> John Gruber created Markdown in 2004.

## What Makes Markdown Stand out?

* Plain text format, easy to maintain
* Small file size, consumes little space
* Simplicity and source file readability
* Abundant softwares and extensions
    support

## How Can I Get Started?

1. Open a text editor (*e.g.* VSCode)
1. Create a file with extension `.md`
1. Enjoy writing!
```

```
Visit [Gruber's personal website](https
    ://daringfireball.net/projects/
    markdown/syntax) for the basic
    syntax of Markdown.
```