

Zpravodaj Československého sdružení uživatelů TeXu

Vít Novotný

Markdown 2.8.1: Směle k trůnu odlehčeného značkování v TeXu

Zpravodaj Československého sdružení uživatelů TeXu, Vol. 30 (2020), No. 1-2, 48–56

Persistent URL: <http://dml.cz/dmlcz/150268>

Terms of use:

© Československé sdružení uživatelů TeXu, 2020

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ*:
The Czech Digital Mathematics Library <http://dml.cz>

Markdown 2.8.1: Směle k trůnu odlehčeného značkování v $\text{T}_{\text{E}}\text{X}$ u

VÍT NOVOTNÝ

Markdown je odlehčený značkovací jazyk, který je určený pro přípravu strukturně jednoduchých dokumentů. Existující systémy pro sazbu dokumentů v jazyce Markdown využívají často $\text{T}_{\text{E}}\text{X}$, ale zachází s ním jako s černou skříňkou, čímž znemožňují použití standardních $\text{T}_{\text{E}}\text{X}$ ových balíků a nástrojů. Oproti tomu balík Markdown umožňuje sazbu dokumentů v jazyce Markdown přímo z $\text{T}_{\text{E}}\text{X}$ u místo toho, aby se snažil nahradit $\text{T}_{\text{E}}\text{X}$ omezenějším systémem.

Od svého vydání v roce 2016 obdržel balík Markdown množství významných aktualizací, které přidaly nové funkce a zlepšily uživatelskou přístupnost balíku. V tomto článku si balík Markdown krátce představíme a podíváme se na novinky v jeho funkcionalitě a dokumentaci.

Klíčová slova: Markdown, Lua, plain $\text{T}_{\text{E}}\text{X}$, $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$, $\text{C}\text{O}\text{N}\text{T}_{\text{E}}\text{X}\text{T}$, Pandoc

Úvod

Za hlavní přednosti $\text{T}_{\text{E}}\text{X}$ u lze považovat jeho programovatelnost a typografické funkce a až v druhé řadě jeho syntax. S výjimkou matematiky jsou odlehčené značkovací jazyky jako Markdown [1] snáze (nau)čitelné a umožňují export do desítek výstupních formátů.

Existující nástroje pro sazbu dokumentů v jazyce Markdown, jakými jsou např. Pandoc a MultiMarkdown, mají několik významných nevýhod [2, 3]: zachází s $\text{T}_{\text{E}}\text{X}$ em jako s černou skříňkou, mají nekonzistentní podporu pro používání $\text{T}_{\text{E}}\text{X}$ ových maker uvnitř markdownových dokumentů, znesnadňují údržbu existujících markdownových dokumentů a nejsou dostupné ve webových službách postavených na $\text{T}_{\text{E}}\text{X}$ u, jako je např. webový editor Overleaf. Balík Markdown řeší veškeré zmíněné nevýhody.

V předchozích článcích [2, 3] jsem představil balík Markdown [4] ve verzi 2.5.3, která trpěla několika vlastními problémy: Balík nefungoval při uvedení volby `-output-directory` na příkazové řádce $\text{T}_{\text{E}}\text{X}$ u a zaplňoval souborový systém konvertovanými markdownovými dokumenty, které musel uživatel ručně mazat. Dokumentace balíku byla vyčerpávající, ale pro běžného uživatele příliš technická.

V tomto článku představím verzi 2.8.1 balíku Markdown, která zmíněné problémy předchozích verzí řeší a zároveň přináší nové funkce, jako je porcování obsahu, značkování tabulek a konverze markdownových dokumentů z příkazové řádky.

Nové funkce

Mezi verzemi 2.5.3 a 2.8.1 balíku Markdown byla jedna významná záplatová verze, 2.5.4, a tři významné malé verze, 2.6.0, 2.7.0 a 2.8.0. Verze 2.5.4 přidala podporu pro volbu `-output-directory` \TeX u. Verze 2.6.0 představila rozhraní pro konverzi markdownových dokumentů z příkazové řádky a přidala podporu pro \LaTeX ový balík `doc` [5]. Verze 2.7.0 umožnila porcování obsahu a představila nový uživatelský manuál. Verze 2.8.0 rozšířila porcování obsahu a umožnila značkovat tabulky pomocí syntaktického rozšíření jazyka Markdown.

V této sekci představím jednotlivé nové funkce balíku Markdown. Ačkoliv balík podporuje i formáty plain \TeX a \CONTEXT , veškeré ukázky jsou pro přístupnost ve formátu \LaTeX .

Změna pracovního adresáře

\TeX je možné spustit s volbou `-output-directory`, která z pohledu \TeX ového dokumentu změni pracovní adresář. Toto umožňuje uživateli přesměrovat dočasné soubory \TeX u do samostatného adresáře. Veškeré externí programy spuštěné pomocí mechanismu `\write18` jsou ale spuštěné v původním pracovním adresáři: Například externí interpret jazyka Lua spouštěný balíkem Markdown pracuje s dočasnými soubory \TeX u, a pokud je nenalezne v pracovním adresáři, selže.

Pro vyřešení tohoto problému představila verze 2.5.4 balíku Markdown volbu `outputDir`, která informuje interpret jazyka Lua o tom, v jakém adresáři se nachází pomocné soubory \TeX u. Pro vyzkoušení si vytvořte textový dokument `dokument.tex` s následujícím obsahem:

```
1 \documentclass{article}
2 \usepackage[outputDir=/dev/shm]{markdown}
3 \begin{document}
4 \begin{markdown}
5 Nadpis první úrovně
6 =====
7 Nadpis druhé úrovně
8 -----
9 Lorem ipsum dolor sit amet, consectetur adipiscing elit.
10 Vestibulum ac commodo. Toto je pouhý odstavec textu.
11 \end{markdown}
12 \end{document}
```

Nakonec spusťte následující příkaz pro vysázení markdownového dokumentu s jednou sekcí, jednou podsekcí a jedním odstavcem (vizte obrázek 1):

```
13 pdflatex -output-directory /dev/shm -shell-escape dokument.tex
```

Dočasné soubory budou přesměrovány do adresáře `/dev/shm`, což je RAM disk dostupný v aktuálním Linuxovém jádře.

Nadpis první úrovně

Nadpis druhé úrovně

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum ac commodo. Toto je pouhý odstavec textu.

Obrázek 1: Markdown dokument s jednou sekcí, podsekcí a odstavcem

Markdown nám umožňuje část textu *zdůraznit*.

Pomocí dvou hvězdiček nebo podtržíték značíme **silný důraz**.

Obrázek 2: Markdown dokument s jedním odstavcem, který obsahuje zdůrazněný a silně zdůrazněný text

Rozhraní pro příkazovou řádku

Balík Markdown předává markdownové dokumenty externímu interpretu jazyka Lua. Intepret jazyka Lua převede markdownové dokumenty do \TeX u a předá je zpět balíku Markdown, který je vysází, vizte obrázek 3. Výhodou tohoto postupu je automatizovanost. Zároveň však tento postup skýtá i množství nevýhod: Převedené markdownové dokumenty jsou ukládány na souborový systém, kde zabírají stále rostoucí množství místa. Pokud použitá varianta \TeX u neobsahuje interpret jazyka Lua, vyžaduje balík Markdown přístup k příkazové řádce systému, což je bezpečnostní riziko. Postup je komplexní a z pohledu uživatele obtížně laditelný.

Řešením výše uvedených problémů je rozdělení postupu na dva kroky. V prvním kroku uživatel převede markdownové dokumenty do \TeX u a v druhém kroku uživatel převedené markdownové dokumenty vysází, vizte obrázek 4. Před prvním krokem může uživatel odstranit všechny dříve převedené markdownové dokumenty. Před druhým krokem může uživatel libovolně upravovat převedené markdownové dokumenty. Během druhého kroku nepotřebuje balík Markdown přístup k příkazové řádce systému. Oddělení jednotlivých kroků postupu usnadňuje ladění.

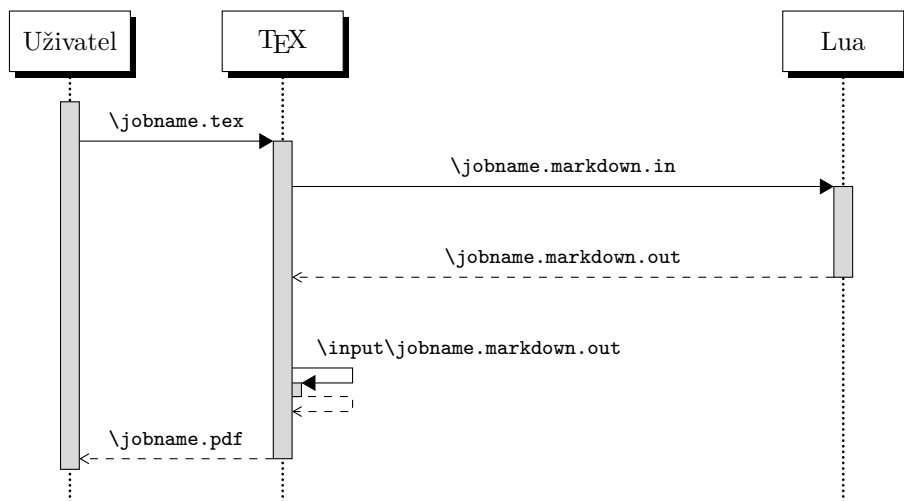
Verze 2.6.0 balíku Markdown proto zavedla rozhraní pro příkazovou řádku ve formě samostatného programu v jazyce Lua. Pro vyzkoušení si vytvořte textový dokument `priklad.md` s následujícím obsahem:

```
14 Markdown nám umožňuje část textu *zdůraznit*.  
15
```

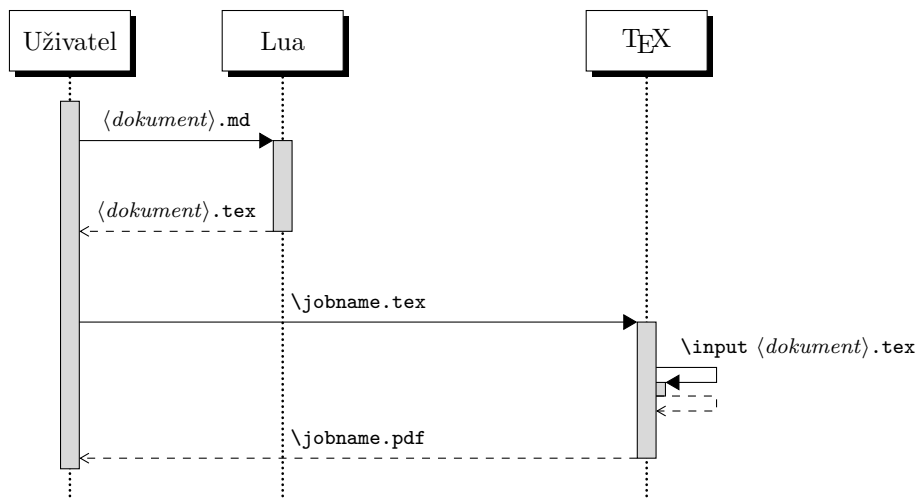
```
16 Pomocí dvou hvězdiček nebo podtržíték značíme **silný důraz**.  
Dále spusťte následující příkaz pro zjištění umístění programu, který zpřístupňuje  
příkazovou řádku balíku Markdown:
```

```
16 kpsewhich markdown-cli.lua
```

Možným umístěním v distribuci \TeX Live 2019 může být například `/usr/local/texlive/2019/texmf-dist/scripts/markdown/markdown-cli.lua`. Posléze



Obrázek 3: Sekvenční diagram sázení dokumentu v jazyce Markdown pomocí TeXového rozhraní balíku Markdown



Obrázek 4: Sekvenční diagram sázení dokumentu v jazyce Markdown pomocí rozhraní balíku Markdown pro příkazovou řádku

spusťte následující příkaz pro převod markdownového dokumentu `priklad.md` na \TeX ový dokument `priklad.tex`:

```
17 texlua /umístění/markdown-cli.lua -- priklad.md priklad.tex
```

Dále vytvořte textový dokument `dokument.tex` s následujícím obsahem:

```
18 \documentclass{article}
19 \usepackage{markdown}
20 \begin{document}
21 \input priklad
22 \end{document}
```

Nakonec spusťte následující příkaz pro vysázení markdownového dokumentu s jedním odstavcem, který obsahuje zdůrazněný a silně zdůrazněný text (vizte obrázek 2):

```
23 pdflatex dokument.tex
```

Dokumentace \LaTeX ových balíků

\LaTeX ový balík `doc` umožňuje dokumentaci jiných balíků vepsáním \LaTeX ového dokumentu s dokumentací do komentářů ve zdrojovém textu balíku. Tento způsob vychází z Knuthova literárního programování [6] a je oblíbený mezi autory \LaTeX ových balíků, protože udržuje dokumentaci poblíž dokumentovaného kódu.

Použití odlehčeného značkování v dokumentaci může zvýšit čitelnost zdrojového kódu a učinit ho tím snáze přístupným vývojářské komunitě. Za tímto účelem zavedla verze 2.6.0 balíku Markdown volbu `stripPercentSigns`, která informuje interpret jazyka Lua, aby ignoroval znak uvozující komentáře. Pro vyzkoušení si vytvořte textový dokument `dokument.dtx` s následujícím obsahem:

```
24 % \iffalse
25 \documentclass{ltxdoc}
26 \usepackage[T1]{fontenc}
27 \usepackage[stripPercentSigns, hashEnumerators]{markdown}
28 \begin{document}
29 \DocInput{dokument.dtx}
30 \end{document}
31 % \fi
32 % \begin{markdown}
33 % * Bylo smažno, lepě svihlí tlové
34 % * se batoumali v dálnici,
35 % + chrudošní byli borolové,
36 % - na mamné krsy žárnící.
37 %
38 % 1. „Ó synu, střež se Žvahlava,
39 % 2. má zuby, drápy přeostré;
40 % #. střež se i Ptáka Neklava,
41 % #. zuřmící Bodostre!"
42 % \end{markdown}
```

- | | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> • Bylo smažno, lepě svihlí tlové • se batoumali v dálnici, • chrudošní byli borolové, • na mamné krsy žárnící. | <ol style="list-style-type: none"> 1. „Ó synu, střež se Žvahlava, 2. má zuby, drápy přeostré; 3. střež se i Ptáka Neklava, 4. zuřmící Bodostre!“ |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Obrázek 5: Markdown dokument s jedním odrážkovým seznamem a jedním číslovaným seznamem. Text je převzatý z Alenčina dobrodružství v říši divů a za zrcadlem v překladu Jaroslava Čísaře.

Nakonec spusťte následující příkaz pro vysázení markdownového dokumentu s jedním odrážkovým seznamem a jedním číslovaným seznamem (vizte obrázek 5):

```
43 pdflatex -shell-escape dokument.dtx
```

Porcování obsahu

Při sazbě dokumentů je často užitečné, z důvodů jak z časových tak autorských, tyto dokumenty rozdělit nebo vhodně přeskládat jejich obsah: Přednáškové slajdy pro celý semestr lze rozdělit na samostatné dokumenty pro jednotlivé týdny. Z desítek minut strávených sazbou monografie se stanou vteřiny, pokud se omezíme na vybrané části. Neuspořádané poznámky můžeme při sazbě přeskládat do soudržného dokumentu.

Pro usnadnění formátování markdownových dokumentů existuje syntaktické rozšíření jazyka Markdown, které umožňuje přidělovat prvkům jazyka HTML atributy. Díky HTML atributům se můžeme jednoznačně odkazovat na jednotlivé sekce markdownových dokumentů. Verze 2.7.0 balíku Markdown podporuje zmíněné syntaktické rozšíření pomocí volby `headerAttributes`. Pomocí voleb `slice` a `shiftHeadings` můžeme následně vysázet pouze malou část markdownového dokumentu a měnit úroveň nadpisů. Pro vyzkoušení si vytvořte textový dokument `dokument.tex` s následujícím obsahem:

```
44 \documentclass{article}
45 \usepackage[headerAttributes]{markdown}
46 \begin{filecontents*}{recept.md}
47 # Palačinky
48 Moučník z mouky, mléka a vajec. Podávejte s marmeládou.
49 ## Třetí krok {#krok3}
50 Rychlým pohybem katapultujeme palačinku na strop. Hola hop!
51 ### První krok {#krok1}
52 Přísady rozmixujeme.
53 ## Druhý krok {#krok2}
54 Výslednou směs nalijeme na rozpálenou pánev.
55 \end{filecontents*}
```

```

56 \begin{document}
57 \markdownInput[slice=~ ^krok3]{recept.md}
58 \markdownInput[slice=krok1, shiftHeadings=-1]{recept.md}
59 \markdownInput[slice=krok2]{recept.md}
60 \markdownInput[slice=krok3 ^krok1]{recept.md}
61 \end{document}

```

Nakonec spusťte následující příkaz pro vysazení markdownového dokumentu s jednou sekci a třemi podsekcemi ve správném pořadí (vizte obrázek 6):

```

62 pdflatex -shell-escape dokument.tex

```

Značkování tabulek

Jedním z nejsložitějších prvků odlehčeného značkovacího jazyka bývají zpravidla tabulky. Jazyk Markdown tabulky nepodporuje, ale existuje množství syntaktických rozšíření, které podporu pro značkování tabulek přidávají.

Od verze 2.8.0 podporuje balík Markdown syntaktické rozšíření pro značkování tabulek pomocí volby `pipeTables` a syntaktické rozšíření pro značkování popisků tabulek pomocí volby `tableCaptions`. Pro vyzkoušení si vytvořte textový dokument `dokument.tex` s následujícím obsahem:

```

63 \documentclass{article}
64 \usepackage[pipeTables, tableCaptions]{markdown}
65 \begin{filecontents*}{tabulka.md}
66 | Příjmení | Jméno | Funkce | Přítomnost |
67 |-----:|:-----:|:-----:|:-----:|
68 | Němec | Karel | náčelník | zde |
69 | Šofr | Vojtěch | lékárník | zde |
70 | Fryštenský | Varel | potah | zde |
71 | Poustka | Václav | pomocný učitel | chybí |
72 : Docházka bývalé polární výpravy

```

Palačinky

Moučník z mouky, mléka a vajec. Podávejte s marmeládou.

První krok

Prísady rozmixujeme.

Druhý krok

Výslednou směs nalijeme na rozpálenou pánev.

Třetí krok

Rychlým pohybem katapultujeme palačinku na strop. Hola hop!

Obrázek 6: Markdown dokument s jednou sekci a třemi podsekcemi

Příjmení	Jméno	Funkce	Přítomnost
Němec	Karel	náčelník	zde
Šofr	Vojtěch	lékárník	zde
Fryštenský	Varel	potah	zde
Poustka	Václav	pomocný učitel	chybí

Tabulka 1: Docházka bývalé polární výpravy

Obrázek 7: Markdown dokument s jednou tabulkou a jejím popiskem

```

73 \end{filecontents*}
74 \begin{document}
75 \markdownInput{tabulka.md}
76 \end{document}

```

Nakonec spusťte následující příkaz pro vysázení markdownového dokumentu s tabulkou a jejím popiskem (vizte obrázek 7):

```

77 pdflatex -shell-escape dokument.tex

```

Uživatelský manuál

Tufte [7] tvrdí, že systémy pro přípravu prezentačních slajdů se nedílně pojí s kognitivním stylem, který brání efektivní komunikaci. Stejně tak literární programování v podobě, v jaké ho aplikují autoři \TeX ových balíků, často produkuje špatně čitelnou dokumentaci, která je příliš pevně svázána se strukturou dokumentovaného kódu.¹ Některé \TeX ové balíky obsahují samostatný uživatelský manuál, který je připravený nezávisle na kódu dokumentovaného balíku, což obvykle zvyšuje čitelnost, ale za cenu horší udržitelnosti dokumentace.

Před verzí 2.5.6 obsahoval balík Markdown pouze technickou dokumentaci [8], která byla připravena literárním programováním a která byla určena pro vývojáře a pokročilé uživatele. Od verze 2.5.6 obsahoval balík Markdown také uživatelský manuál [9], který se zaměřoval na běžné uživatele. Stejně jako technická dokumentace byl ale i tento uživatelský manuál přímočaře vytvořen literárním programováním, díky čemuž byl špatně strukturovaný a obtížně čitelný. Od verze 2.7.0 je proto uživatelský manuál vytvořen ze tří textů, které popisují uživatelské rozhraní, volby balíku a prvky jazyka Markdown. Výsledkem je (doufám) čitelný uživatelský manuál, který je ale stále připraven pomocí literárního programování.

¹Knuthovo literární programování [6] umožňuje rekurzivně definovat program po malých samostatně dokumentovaných částech. \LaTeX ové balíky `doc` a `docstrip` umožňují pouze generovat soubor se zdrojovým textem programu v takovém sledu, v jakém je program dokumentován.

Odkazy

1. GRUBER, John. *Daring Fireball: Markdown* [online]. 2013 [cit. 2019-03-31]. Dostupné z: <https://daringfireball.net/projects/markdown>.
2. NOVOTNÝ, Vít. Sazba textu označovaného v jazyce Markdown uvnitř T_EXových dokumentů. *Zpravodaj Československého sdružení uživatelů T_EXu* [online]. 2016, roč. 26, č. 1–4, s. 78–217 [cit. 2019-03-31]. ISSN 1213-8185. Dostupné z DOI: 10.5300/2016-1-4/78.
3. NOVOTNÝ, Vít. Using Markdown Inside T_EX Documents. *TUGboat* [online]. 2017, roč. 38, č. 2, s. 214–217 [cit. 2019-03-31]. ISSN 0896-3207. Dostupné z: <https://tug.org/TUGboat/tb38-2/tb119novotny.pdf>.
4. NOVOTNÝ, Vít. *A package for converting and rendering markdown documents inside T_EX* [online]. 2019 [cit. 2019-11-26]. Dostupné z: <https://ctan.org/pkg/markdown>.
5. MITTELBACH, Frank. *Doc – Format L^AT_EX documentation* [online]. 2018 [cit. 2019-03-31]. Dostupné z: <https://ctan.org/pkg/doc>.
6. KNUTH, Donald Ervin. Literate programming. *The Computer Journal*. 1984, roč. 27, č. 2, s. 97–111. Dostupné z DOI: 10.1093/comjnl/27.2.97.
7. TUFTE, Edward Rolf. *The Cognitive Style of PowerPoint*. 2. vyd. Graphics Press Cheshire, CT, 2006. ISBN 0-961-39216-9.
8. NOVOTNÝ, Vít. *A Markdown Interpreter for T_EX* [online]. 2019 [cit. 2019-11-26]. Dostupné z: <http://mirrors.ctan.org/macros/generic/markdown/markdown.pdf>. Verze 2.8.1.
9. NOVOTNÝ, Vít. *Markdown Package User Manual* [online]. 2019 [cit. 2019-11-26]. Dostupné z: <http://mirrors.ctan.org/macros/generic/markdown/markdown.html>. Verze 2.8.1.

Summary: Markdown 2.8.1: Boldly Unto the Throne of Lightweight Markup in T_EX

Markdown is a lightweight markup language that makes it easy to write structurally simple documents. Existing tools for rendering markdown documents to PDF treat T_EX as a black box. In contrast, the Markdown package provides support for styling and typesetting markdown documents in T_EX, extending a T_EXie’s toolbox rather than forcing them to replace T_EX with a more limited tool.

Since its release in 2016, the package has received several important updates improving the functionality and user experience. In this article, I will reintroduce the package, and describe its new functionality and documentation.

Keywords: Markdown, Lua, plain T_EX, L^AT_EX, CONT_EX_T, Pandoc

Vít Novotný, witiko@mail.muni.cz