

Zpravodaj Československého sdružení uživatelů TeXu

Roman Plch

Matematika na čtečkách e-knih

Zpravodaj Československého sdružení uživatelů TeXu, Vol. 25 (2015), No. 1-2, 70–79

Persistent URL: <http://dml.cz/dmlcz/150228>

Terms of use:

© Československé sdružení uživatelů TeXu, 2015

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>

Mobilní zařízení se stala nedílnou součástí našeho každodenního života, a je proto přirozené, že ovlivňují i výukový proces. Tištěné výukové materiály jsou postupně nahrazovány digitálními a jejich využívání se stále častěji přesouvá z počítačů na tato mobilní zařízení (mobily, tablety, čtečky e-knih). Tento příspěvek popisuje možnosti zobrazování matematických textů na mobilních zařízeních a dále se věnuje konverzi matematického textu do podoby vhodné k zobrazení na čtečkách e-knih.

Formáty e-knih

Formát EPUB

EPUB (Electronic Publication, 2015) je standardizovaný otevřený souborový formát určený pro digitální knihy, vyvinutý organizací International Digital Publishing Forum (IDPF). Je založený na XML. Zjednodušeně se dá říci, že se jedná o sadu HTML stránek zabalenou s dalšími metadaty do ZIP archivu. Formát EPUB je pro čtenáře výhodný v tom, že umožňuje automaticky přizpůsobit text různým velikostem obrazovek zařízení (včetně chytrých telefonů, netbooků a čteček elektronických knih). EPUB také podporuje obsah s pevným rozvržením.

Poslední verze tohoto formátu (EPUB 3.01) už podporuje i jazyk MathML a umožňuje tak zobrazování matematických formulí přirozeným způsobem, tj. matematický text se zvětšuje/zmenšuje společně s ostatním textem, vzorečky se správně zarovnávají vzhledem k okolnímu textu a je možné v nich vyhledávat a kopírovat jejich obsah.

Bohužel podpora tohoto formátu je zatím omezená a mnohem častěji se používá starší verze EPUB 2, která pro zobrazování matematiky využívá obrázků, stejně jako všechny dále uváděné formáty.

Formát EPUB 2 podporují téměř všechny čtečky e-knih (významnou výjimkou je pouze Kindle od Amazonu).

Formáty MOBI, AZW, KF8

Čtečky Kindle od Amazonu nepodporují formát EPUB a používají vlastní formáty pro e-knihy. Na rozdíl od formátu EPUB se jedná o proprietární formáty. První čtečky Kindle používaly formát AZW, což je vlastně formát MOBI doplněný o podporu DRM restrikcí. Spolu se čtečkou Kindle Fire se objevuje formát KF8

(Kindle formát 8), známý též jako AZW3. Ten už podporuje i některé prvky HTML 5 a CSS3. Matematika může být ale uvedena pouze ve formě obrázků.

Pro převody z formátu EPUB do MOBI (AZW3) můžeme použít aplikaci KindleGen nebo níže uvedený program Calibre pro oboustranné převody.

Formát PDF

I tento základní souborový formát pro elektronické publikování má na čtečkách e-knih stále své místo, jeho „nevýhodou“ je ale fixní rozložení stránky, které neumožňuje přizpůsobovat velikost textu rozměrům čtečky.

Některé čtečky sice dnes už mají funkci „PDF Reflow“, která se pokusí původní strukturu dokumentu převést tak, aby se text dobře vešel na displej čtečky. Jak však uvidíme později, je funkce PDF Reflow prakticky nepoužitelná na texty s matematikou.

Kromě těchto základních formátů se můžeme setkat i s méně častými formáty, jako je eReader (PDB), FictionBook (FB2) a Microsoft Lit (LIT). V tomto příspěvku se ale budeme věnovat výhradně nejrozšířenějšímu formátu EPUB.

Typy čteček

Čtečky e-knih můžeme dělit do dvou základních kategorií – na softwarové a hardwarové čtečky.

Hardwarové čtečky

Jedná se o specializovaná elektronická zařízení na bázi elektronického inkoustu. Oproti zařízením s LCD displejem spotřebovává tato zobrazovací technologie mnohem méně energie a text je dobře čitelný i na intenzivním slunečním světle.

Mezi populární čtečky e-knih patří např. Kindle (Amazon), Nook (Barnes & Noble), Kobo (Rakuten), Cybook (Bookeen), PocketBook, SonyReader, JetBook (Ectaco) a Boox (Onyx). Podporu formátu EPUB3 ale uvádí pouze Kobo u své čtečky Kobo Glo HD (Kobo, 2015).

Softwarové čtečky

Druhou kategorií tvoří programy a aplikace umožňující čtení e-knih na počítači, tabletu a mobilním telefonu. Do výběru jsme zařadili některé programy, které jsou šířeny bezplatně, jsou multiplatformní a podporují více formátů e-knih (a především formát EPUB3). Kompletní přehled čteček s podporou EPUB3 najdeme např. na <http://docs.mathjax.org/en/latest/misc/epub.html>.

Calibre

(<http://calibre-ebook.com/>)

Volně šiřitelná multiplatformní aplikace Calibre je univerzální správce a konvertor elektronických knih. K dispozici jsou nástroje pro organizaci i prohlížení knih, editaci metadat i konverzi. Do osobní knihovny je možné přidávat knihy jednotlivě nebo v dávkách, je podporován široký rozsah nejrůznějších formátů.

Radium

(<http://idpf.org/forum/topic-2429>)

Rozšíření pro Google Chrome, umožňující čtení e-knih ve formátu EPUB 3.

Gitden Reader

(<http://gitden.com/>)

Jedna z nejlépe hodnocených čteček e-knih pro Android a IOS. Na základě našich zkušeností nejlépe zvládá zobrazování e-knih s matematickým textem ve formátu EPUB 3 na mobilních platformách.

Hardwarové čtečky s Androidem

Jedná se o zařízení s elektronickým inkoustem běžící pod operačním systémem Android. Umožňuje tedy instalaci Android aplikací a spojuje tak výhody hardwarových a softwarových čteček v jediném zařízení. Jako příklad můžeme uvést např. inkBOOK Onyx nebo eReading 4 Touch Light.

Konverze do EPUB

Po tomto stručném úvodu se budeme věnovat přípravě matematického textu pro čtečky e-knih. Vycházíme z předpokladu, že máme k dispozici zdrojový kód v \LaTeX u a chceme ho převést do formátu EPUB.

\TeX 4ebook

Jedná se o konvertor z \LaTeX u do formátu EPUB a MOBI. Ke konverzi využívá program `tex4ht` a skripty `Lua \TeX` . Instalační balíček a dokumentaci najdeme na <https://github.com/michal-h21/tex4ebook>, návod na instalaci najdeme např. v bakalářské práci (Antol, 2015).

Konverzi spustíme příkazem

```
tex4ebook [volby] jméno_souboru
```

Implicitní výstupní formát je EPUB, změnu na EPUB 3 (MOBI) provedeme parametrem `-f epub3 (mobi)`. Pro matematiku v MathML přidáme na konec příkazu ještě parametr `mathml`:

```
tex4ebook -f epub3 soubor.tex mathml
```

Při praktických testech konverze delšího matematického textu (Slovák, Panák a Bulant, 2013) jsme na instalaci T_EXLive 2014 narazili na chybové hlášení `Undefined control sequence`.

```
\pgfsys@svg@newline ->\Hnewline
```

Jedná se o známou chybu, kterou podle (T_EX – L^AT_EX Stack Exchange, 2014) vyřešíme úpravou v souboru `pgfsys-tex4ht.def`, kde řádek s textem `\def\pgfsys@svg@newline{\Hnewline}` nahradíme za

```
\def\pgfsys@svg@newline{^^J}.
```

Při výstupu do EPUB2 se matematika ukládá do obrázků formátu PNG, pomocí poměrně komplikovaného nastavení je možné docílit změny formátu na SVG. Potřebné nastavení je popsáno v (T_EX – L^AT_EX Stack Exchange, 2015).

Shrnutí, výhody a nevýhody

Konverze i složitějšího matematického textu s tabulkami sázenými pomocí balíčku `tabularx` proběhla v pořádku, zobrazení matematiky pomocí Calibre na počítači bylo taktéž bez chyb. Poznámky pod čarou se umísťují ihned za odstavcem textu, ve kterém je příkaz uveden. Do textu je možné vkládat obrázky ve formátech EPS, JPG a PNG.

Nevýhodou je poměrně komplikovaná instalace¹ a pro méně zkušené uživatele nutnost spouštět konverzi z příkazového řádku.

Pandoc

Pandoc je univerzální multiplatformní dokumentový konvertor podporující desítky různých formátů včetně EPUB3.

Instalace je v tomto případě jednoduchá, stačí ze stránek <http://pandoc.org/> stáhnout a spustit instalační soubor. K dispozici je i online verze (<http://pandoc.org/try/>), která má však omezené možnosti a nepodporuje výstup do EPUB.

Program se ovládá pouze z příkazového řádku (nemá grafické uživatelské rozhraní), jeho syntaxi si ukážeme na modelovém příkladu

```
pandoc soubor.tex -f latex -t epub -mathml -o soubor.epub
```

kde přepínač `f` zadává vstupní formát, přepínač `t` výstupní formát a přepínač `o` jméno výstupního souboru. Pro výstup ve formátu EPUB3 zadáme parametr `mathml`. Seznam všech možných parametrů získáme přepínačem `h`.

Shrnutí, výhody a nevýhody

Pandoc nezvládl konverzi celé řady matematických symbolů (balíček `amssymb`), dále nefunguje konverze matematických prostředí definovaných pomocí balíčku

¹T_EX4ebook se stal součástí aktuální instalace T_EXlive 2015, což situaci výrazně zjednodušuje.

`amsthm`. Po konverzi dále nefungují odkazy vytvářené pomocí příkazů `\label` a `\ref`. Ani konverze tabulek neproběhla bez chyb. Poznámky pod čarou jsou umístěny na konec knihy a v textu jsou umístěny odkazy. Do textu je možné vkládat obrázky ve formátech JPG a PNG.

Mezi výhody patří rychlá a snadná instalace a velké množství konverzních formátů. Pro složitější matematické texty není ale možné Pandoc doporučit.

$\text{T}_{\text{E}}\text{X}4\text{ht} + \text{Calibre}$

V tomto případě rozdělíme konverzi do dvou částí. Nejdříve pomocí programu $\text{T}_{\text{E}}\text{X}4\text{ht}$ (je součástí instalace $\text{T}_{\text{E}}\text{X}Live$) zkonvertujeme $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ ový zdroj do jazyka HTML. Získaný soubor poté pomocí Calibre převedeme do EPUB.

Syntaxe:

```
htllatex soubor.tex "xhtml, charset=utf-8,mathml" " -cunihtf -utf8"
```

Výstupem bude soubor `soubor.html` v kódování UTF8 s matematikou zapsanou v jazyce MathML.

Následně pomocí příkazu

```
ebook-convert soubor.html soubor.epub
```

získáme výsledný soubor ve formátu EPUB3. Pokud při konverzi vynecháme parametr `mathml`, bude matematika ve formě obrázků a výstup bude ve formátu EPUB2. Stejně tak při volbě jiného výstupního formátu (např. MOBI) bude matematika uložena ve formě obrázků. Tyto konverze je možné provádět i pomocí grafického uživatelského rozhraní Calibre.

Shrnutí, výhody a nevýhody

Výstup je srovnatelný s výsledky získanými pomocí $\text{T}_{\text{E}}\text{X}4\text{ebook}$. Poznámky pod čarou jsou umístěny na konci dokumentu, v textu jsou umístěny odkazy. Mnohem jednodušší je ale instalace, $\text{T}_{\text{E}}\text{X}4\text{ht}$ je součástí $\text{T}_{\text{E}}\text{X}Live$, stačí tedy nainstalovat pouze Calibre. Komplikací zůstává, že $\text{T}_{\text{E}}\text{X}4\text{ht}$ je při generování MathML velmi citlivý na čistotu zápisu matematiky ve zdrojovém textu dokumentu. Například zápis $\$M=\{x|x\$ je liché \$\}\$$ je korektní $\text{T}_{\text{E}}\text{X}$ ový zápis. $\text{T}_{\text{E}}\text{X}4\text{ht}$ ale v tomto případě nemá informaci o párování složených závorek. Je tedy nutné použít správnější zápis: $\$M=\{x|x \text{ \texttt{\textbackslash}text\{je liché\}\}}\$, viz Sojka a Růžička (2008).$

$\text{L}^{\text{A}}\text{T}_{\text{E}}\text{XML}$

$\text{L}^{\text{A}}\text{T}_{\text{E}}\text{XML}$ (<http://dmlf.nist.gov/LaTeXML/>) konvertuje zdrojový kód $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ u do formátů XML, HTML a EPUB. Primárním výstup je do XML, dalším zpracováním získáme formáty HTML a EPUB, matematické formule mohou být ve formě obrázků nebo MathML.

Vlastní použití je jednoduché, pomocí příkazu
`latexmlc soubor.tex --destination=soubor.epub`
dostaneme výsledný soubor ve formátu EPUB3.

Shrnutí, výhody a nevýhody

Konverze i složitějšího matematického textu proběhla v pořádku, text je na čtečkách na PC dobře čitelný (obr. 1), také konverze tabulek proběhla korektně. Poznámky pod čarou se zobrazují v samostatném okně (po najetí kurzorem myši na text poznámky). Do textu je možné vkládat obrázky ve formátu PDF, EPS, JPG a PNG.

Komplikací pro autory může být to, že pokud mají v dokumentu vlastní definice maker, musí tuto definici napsat i pro \LaTeX ML, aby věděl, jak tato makra expandovat. Dle našeho názoru se ale jedná o perspektivní a dobře použitelné řešení (pokud ovšem dojde k vylepšení softwaru pro čtení formátu EPUB3 na čtečkách e-knih a mobilních platformách).

$$(1 + ap)^{p^{\ell-1}} \equiv (1 + ap^{\ell-1})^p \pmod{p^{\ell+1}}.$$

Z binomické věty přitom plyne

$$(1 + ap^{\ell-1})^p = 1 + p \cdot a \cdot p^{\ell-1} + \sum_{k=2}^p \binom{p}{k} a^k p^{(\ell-1)k}$$

a vzhledem k tomu, že pro $1 < k < p$ platí

$p \mid \binom{p}{k}$, stačí ukázat $p^{\ell+1} \mid p^{1+(\ell-1)k}$, což je

ekvivalentní s $1 \leq (k-1)(\ell-1)$.

Rovněž pro $k = p$ dostáváme díky předpokladu

$\ell \geq 2$ vztah $p^{\ell+1} \mid p^{(\ell-1)p}$. ■

Lemma. Bud' p liché prvočíslo, $\ell \geq 2$ libovolné.

Pak pro libovolné $a \in \mathbb{Z}$ splňující $p \nmid a$ platí, že řád čísla $1 + ap$ modulo p^ℓ je roven $p^{\ell-1}$.

Důkaz. Podle předchozího lemmatu je

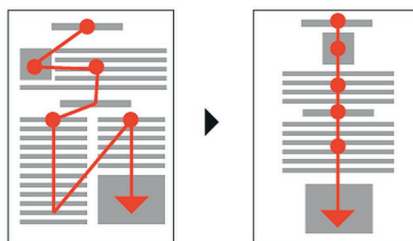
Obrázek 1: Export do EPUB3 zobrazený pomocí Calibre na PC

PDF formát na čtečkách e-knih

Formát PDF zachovává fixní rozložení stránky, proto se obecně nepovažuje za formát vhodný pro čtečky. Stále je však nejběžnější pro publikování elektronických dokumentů a při jeho použití máme jistotu, že i složitý matematický text bude vypadat přesně tak, jak jsme zamýšleli. V této části ukážeme, jak PDF soubory přizpůsobit k zobrazování na čtečkách e-knih.

PDF Reflow

Tato funkce se pokusí původní strukturu dokumentu převést tak, aby se text dobře vešel na displej čtečky, aby nikde nevyčníval (nepřetékal mimo displej) – aby se tedy dokument choval jako nativní elektronická kniha nebo webová stránka s průběžným zobrazením obsahu tak, jak jej nastavil uživatel, včetně velikosti písma a rozměrů stránky.



Obrázek 2: Schéma práce funkce PDF reflow, převzato (Bohdal, 2014)

Výhodou této funkce je, že umožňuje také určité zmenšení nebo zvětšení písma. Touto funkcí ale nejsou vybaveny všechny čtečky. Funguje jen na PDF se sazbou, nikoliv skenované soubory, které jsou vlastně jen obrázkem zabaleným do formátu PDF.

Čím jednodušší je dokument PDF, tím lépe funkce PDF Reflow pracuje. Pokud máte jen souvislý text s odstavci a kapitolami, např. běžný román, budete pravděpodobně s převodem vždy spokojeni. Pokud však máte odborný text s obrázky, tabulkami a vzorečky, výsledek může být katastrofální.

Změna rozměrů PDF souboru

K2pdfopt (<http://www.willus.com/k2pdfopt/>) je multiplatformní program, který optimalizuje soubory ve formátu PDF pro čtečky. Vstupním i výstupním souborem je PDF, přičemž volíme rozměry výsledného souboru. Nevýhodou je

nesprávné dělení slov a matematických výrazů na konci řádků a také velikost výsledného souboru, která několikanásobně převyšuje velikost původního souboru. Hodí se tedy pouze v případě, že máme k dispozici jen výsledný soubor a nikoliv zdrojový kód.

Podobně funguje celá řada dalších nástrojů, najít je můžeme např. na http://www.willus.com/k2pdfopt/pdf_conversion.shtml. Za vyzkoušení stojí i online nástroj <http://convert-kindle.com/>.

Balíček geometry, šablony

Mnohem výhodnější je ale upravit rozměry a vzhled výsledného PDF dokumentu ještě před překladem. K tomu můžeme využít např. balíček `geometry`, hodí se i balíček `extsizes`, který umožňuje měnit velikost základního fontu pro dokument.

Nejjednodušší možností je ale použít některou z šablon, dobré zkušenosti máme se šablonou eBook (<http://www.latextemplates.com/template/ebook>). Použitím šablony získáme výsledný soubor ve formátu PDF, který svými rozměry, nastavením okrajů, velikostí použitého fontu, číslováním stránek a celkovým vzhledem co nejlépe odpovídá zobrazení na dnešních čtečkách e-knih (obr. 3).

Závěr

Základním problémem zobrazování matematických textů na čtečkách e-knih je to, že hardwarové čtečky nepodporují formát EPUB 3. To vede k situaci, kdy se matematické knihy v tomto formátu prakticky nevydávají. Nedostatek poptávky pak způsobuje, že se výrobcům čteček nevyplatí investovat do podpory formátu EPUB 3. Pokud by se podařilo tento „začarovaný kruh“ nějakým způsobem vyřešit, tak je formát EPUB 3 určitě do budoucna perspektivním řešením.

Protože ani softwarové čtečky nejsou v současné době schopny korektně zobrazovat složitější matematický text, zůstává pro autory matematických textů jediné řešení – formátovat výsledný PDF dokument kromě verze pro tisk i ve verzi pro zobrazení na čtečkách e-knih, např. pomocí šablony eBook uvedené výše v článku.

Summary: Mathematics for readers of e-books

Mobile devices have become an integral part of our daily lives and that is why they have been affecting a learning process. Printed educational materials are being gradually replaced by a digital version. Displaying is increasingly shifting from desktop computers to the mobile devices (i.e. phones, tablets and e-book readers). This paper describes possibilities of displaying mathematical texts on the mobile devices and a possible conversion of mathematical texts for e-book readers is also analysed.

Důkaz. Plyne snadno z binomické věty s využitím matematické indukce vzhledem ℓ .

I. Pro $\ell = 2$ tvrzení zřejmě platí.

II. Necht' tvrzení platí pro ℓ , dokážeme jej i pro $\ell + 1$.

Tvrzení pro ℓ umocníme na p -tou, dostaneme

$$(1 + ap)^{p^{\ell-1}} \equiv (1 + ap^{\ell-1})^p \pmod{p^{\ell+1}}.$$

Z binomické věty přitom plyne

$$(1 + ap^{\ell-1})^p = 1 + p \cdot a \cdot p^{\ell-1} + \sum_{k=2}^p \binom{p}{k} a^k p^{(\ell-1)k}$$

a vzhledem k tomu, že pro $1 < k < p$ platí $p \mid \binom{p}{k}$, stačí ukázat $p^{\ell+1} \mid p^{1+(\ell-1)k}$, což je ekvivalentní s $1 \leq (k-1)(\ell-1)$. Rovněž pro $k = p$ dostáváme díky předpokladu $\ell \geq 2$ vztah $p^{\ell+1} \mid p^{(\ell-1)p}$. \square

Lemma. *Bud' p liché prvočíslo, $\ell \geq 2$ libovolné. Pak pro libovolné $a \in \mathbb{Z}$ splňující $p \nmid a$ platí, že řád čísla $1 + ap$ modulo p^ℓ je roven $p^{\ell-1}$.*

Obrázek 3: Ukázka použití šablony eBook

Literatura

- ANTOL, F. *Matematika na čtečkách e-knih* [on-line]. (Bakalářská práce.) Brno : Masarykova univerzita, 2015. [cit. 2015-07-17]. xii+29 s. Dostupné na: http://is.muni.cz/th/408834/prif_b/.
- BOHDAL, T. Jak číst PDF na ebook čtečkách [on-line]. In *Ebook Expert.cz*. 2014 [cit. 2015-07-29]. Dostupné na: <http://ebookexpert.cz/jak-cist-pdf-na-ebook-cteckach/>.
- EPUB [on-line]. In *International Digital Publishing Forum : The Trade and Standards Organization for the Digital Publishing Industry*. 2015 [cit. 2015-05-07]. Dostupné na: <http://idpf.org/epub>.
- Error using pgfsysdriver with T_EX4ht, only shows up with T_EXlive 2014, ok with

- TeXlive 2013 [on-line]. In *TeX – L^AT_EX Stack Exchange*. 2014 [cit. 2015-07-29].
Dostupné na: <http://tex.stackexchange.com/questions/185349/>.
- Kobo Device Comparison* [on-line]. 2015. [cit. 2015-07-30]. Dostupné na: <https://www.kobo.com/devices/compare#ereaders>.
- SOJKA, P., RŮŽIČKA, M. Parallel Electronic Publications. *Zpravodaj Československého sdružení uživatelů T_EXu*, 2008, vol. 18, issue 3, s. 116–129.
- SLOVÁK, J., PANÁK, M., BULANT, M. *Matematika drsně a svižně*. 1. vyd. Brno: Masarykova univerzita, 2013. 773 s. ISBN 978-80-210-6307-5.
- The best way for math in epub – font or SVG? (tex4ebook) [on-line]. In *TeX – L^AT_EX Stack Exchange*. 2015 [cit. 2015-07-29]. Dostupné na: <http://tex.stackexchange.com/questions/235423/the-best-way-for-math-in-epub-font-or-svg-tex4ebook>.

Roman Plch, plch@math.muni.cz