Zpravodaj Československého sdružení uživatelů TeXu

Milan Šorm

Ligatura aneb začínáme s TeXem

Zpravodaj Československého sdružení uživatelů TeXu, Vol. 14 (2004), No. 3-4, 195–200

Persistent URL: http://dml.cz/dmlcz/149973

Terms of use:

© Československé sdružení uživatelů TeXu, 2004

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

Vytvořením určité alternativy by bylo možné dokázat, že systém TEX s jeho nadstavbami není jen hračkou pro vědce-matematiky.

Dalším stupněm zjednodušujícím využití standardních dokumentů je vytvoření webového rozhraní, kde jsou připravena pole pro vyplnění požadovaných textových segmentů, z nichž se automatizovaně sestaví, vysází a vrátí uživateli bezchybný dokument odpovídající normám a typografickým zásadám. Taková nadstavba je použitelná pro takovou třídu dokumentů, u nichž není počet textových segmentů příliš variabilní, ale takových dokumentů je ve zmíněné sestavě většina.

Závěr

Projekt zpřístupňující precizní sazbu systémem TEX včetně všech nadstaveb je ve své počáteční fázi. Možnosti, které jsou nyní k dispozici, lze shrnout do těchto bodů:

- Zpřístupnění instalace začínajícímu uživateli nebo zájemci, který chce vidět základní možnosti.
- Snížení objemu údržby učebnových instalací systému TEX.
- Využití připravených stylů pro tvorbu standardních dokumentů.
- Vytvoření další obalové aplikace využívající překladu jako jádra své činnosti a zpřístupňující standardní dokumenty přes webové rozhraní.

Z prvních ohlasů lze nabýt přesvědčení, že nabízené služby mohou být pro uživatele atraktivní a že se s minimálními nároky na údržbu může takový systém s úspěchem provozovat.

Odkazy

1. ČSN 01 6910 – Úprava písemností psaných strojem nebo upravených textovými editory. Praha: Normalizační institut, 1997.
2. Seznam formulářů ke stažení. Dokument HTML s odkazy na soubory formátu Microsoft DOC. Dostupné na `http://www.mendelu.cz/formulare`. 20. května 2004.

Ligatura aneb začínáme s TEXem

Milan Šorm

Z dlouholetých zkušeností s TEXem pozoruji, že nainstalovat nějakou TEXovou distribuci na systém Windows s fungující češtinou, editorem, zobrazováním a příp. připravenou nápovědou či ukázkovými styly je pro

běžného člověka velmi obtížné. Proto jsem se rozhodl pro takovéto začátečníky, pro diplomanty naší univerzity, a příp. pro další zájemce o pořádný sázecí program připravit jednoduchou distribuci pro naučení se základům TEXu. Distribuce se skládá z minimální části TEXLive pro tvorbu PDF (`pdfcslatex`), kontroly pravopisu (`ispell`), českého volně šiřitelného textového editoru vlastní provenience (určeného právě pro TEXování) a předpřipravených stylů. Příspěvek má za úkol tento projekt představit a příp. nalákat další zájemce ke spolupráci.

Klíčová slova: ligatura, sazba, TEX, začátečník.

Úvod

Jsem TEXovým nadšencem již osm let a poznal jsem během této doby řadu různých distribucí – emTEX, MikTEX, TeTEX, TEXLive apod. Každá z těchto instalací však obsahovala dvě zásadní nevýhody, se kterými se potýkala mimo mne i řada dalších lidí:

- obtížná instalace a
- nevyhovující editor.

První nevýhoda je dána architekturou TEXových distribucí, které se snaží vyhovět všem a obsahují desítky variant vlastního TEXu, stovky fontů a maker, které jsou stejně málo využívány. To vše ve složitých adresářových strukturách (které jsou sice jasné a přehledné, ale ovšem pouze pro ostříleného TEXového guru) a často rozmístěné v několika stromech po celém diskovém prostoru. Tyto distribuce navíc snahou o podporu maximálního množství platforem a jazyků končí u nepoužitelného gigantu, který neběží pořádně nikde a o funkčnosti češtiny si můžeme na první pokus instalace (babel) nechat jen znát.

Druhá nevýhoda je dána původním posláním editoru jako obecného nástroje pro všechny druhy pořizování textů – od programování přes sazbu, psaní poznámek až po prohlížení velkých souborů, na které nestačí běžné prohlížeče. Tyto editory nejsou obvykle specificky zaměřeny na TEX, ale v nejlepším případě na sazbu obecně.

Vedením diplomových a bakalářských prací jsem navíc na univerzitě nucen spolupracovat s řadou studentů nejen informatických studijních programů, kteří by rádi využili na naší univerzitě doporučovaný a vychvalovaný TEX, kteří však ale díky náročnosti instalace a komplikovanosti editoru nemají příliš šanci se s tímto nádherným systémem sžít.

V neposlední řadě také pozoruji šťastné majitele knih určených začátečníkům (jedna z nich vyšla i v našem sdružení uživatelů), kteří plni nadšení z pěkné knížky začínají s instalací TEXu, aby po několika hodinách či dnech tento systém zavrhli, protože se jim jej nepodařilo kvalitně nainstalovat a začít používat.

Všechny výše uvedené důvody mne vedly k založení nového projektu české miniTEXové distribuce pojmenované Ligatura. Převážná většina funkcí a nápadů realizovaných v této distribuci vychází z reálných potřeb studentů při psaní diplomových prací a potřeb mé ženy při tvorbě běžné korespondence (příspěvky na konference, dopisy, posudky) nebo potřeb mých (dopisy, fakturace, osvědčení, obchodní sestavy).

Vím, že předkládaný článek, projekt i celé téma je v našem sdružení velmi kontroverzní a obvykle vyvolá řadu otázek od Proč? Nač? A proč ne mou oblíbenou distribuci HCHKRDTNTEX? Projekt předkládaný k diskuzi širší TEXové obci si neklade za cíl být jedinou českou distribucí TEXu ani být profesionálním prostředím pro vydavatele a nakladatele. Jeho hlavní cíl je pomoci šíření TEXu a být odrazovým můstkem pro další poznávání TEXu u běžných uživatelů.

Nevymýšlet kolo

Základním principem, který jsem při tvorbě tohoto projektu přijal, je snaha „Nevymýšlet kolo.“. Celý projekt sám o sobě se však může zdát právě porušením uvedeného principu. Upravil bych tedy uvedenou větu do tvaru „Nevymýšlet kolo tam, kde to není nezbytně nutné.“. No a tento projekt minimálně pro mou rodinu a mé studenty nutný je.

Vzhledem k uvedenému principu tedy nebudeme tvořit distribuci na zelené louce pomocí `web2c`, ale vyjdeme z nějaké existující distribuce, kterou ještě důkladně probereme. Původně uvažovaná varianta očesání TEXLive se ukázala jako silně nereálná, protože se jedná o příliš rozsáhlou distribuci. Použil jsem proto značně menší distribuci fpTEX, která je obdobná, ale lze ji nainstalovat do minimálního prostoru v řádech desítek MB. Úkolem nebylo vytvoření nejmenší distribuce, ale malé a výkonné distribuce. Proto jsem ještě ručně nainstalovaný fpTEX probral a zrušil všechny nepotřebné formáty, pomocné programy, varianty a styly, které mi přišly jako redundantní.

Zásadní otázkou bylo, jaký formát zvolit jako jediný správný (možnost volby je věc pěkná a demokratická, která však později komplikuje instalaci a život uživatele) a jaké cílové formáty zvolit. Jakožto přesvědčený LATEXista jsem v první volbě měl jasno a vzhledem k obtížím se zobrazováním formátu PostScript (či dokonce DVI) jsem zvolil verzi produkující přímo formát PDF. Tedy jedinou podporovanou verzí TEXu v projektu Ligatura je varianta PDFLATEX. Dalším zásadním rozhodnutím bylo orientovat se jen na pořádné verze MS Windows, tedy na verze 2000 nebo XP.

Dále jsem v distribuci nahradil veškeré babelizované styly za originální styly a fonty CSTEXu, protože podle tvrzení TEXových odborníků okolo mne je to ta jediná správná česká distribuce s pěknými fonty a českými vzory pro dělení. Po odstranění všech redundancí jsem tak dosáhl 46 MB distribuce, na které budu dále ještě pracovat.

Tuto distribuci bylo nutné dále obohatit o vhodný editor zajišťující vyvolávání překladu, zobrazovač výstupů v PDF a vhodnou kontrolu pravopisu.

Dobrý editor je základ

Při své práci nejraději používám editor `vim`, který mi vyhovuje jak na programování, tak na přípravu textů. Ale např. pro psaní pošty mám raději editor `joe`. Z tohoto příkladu je vidět, že určení textového editoru je věc velmi záludná, protože každý uživatel má jiné představy a zvyklosti. Přesto ve světě MS Windows existuje řada nepsaných dohod, jak by měl textový editor pracovat. Pro práci s TEXem se řadě mých studentů i mně samotnému velmi osvědčil editor TextPad. Bohužel není zdarma a proto jeho užití by bylo velmi problematické. Při dlouhém hledání na Internetu se mi nepodařilo najít podobný editor, který by byl šířený zdarma.

Důsledkem předchozího zklamání z dlouhého hledání jsem se rozhodl jedno kolo znovu vymyslet v podobě naprogramování si v prostředí Borland Delphi nového editoru pro Windows, který bude přímo určen pro práci v Ligatuře. Jedná se tak vlastně o jedinou část projektu, kterou jsem byl nucen naprogramovat a vytvořit podle svých osobních představ.

Vyšel jsem z editoru TextPad a snažil se jej maximálně napodobit způsobem editace (syntax highlighting, číslování řádků, zalamování, doplňující vyhledávací funkce s regulárními výrazy apod.). Díky využití komponenty SynEdit, kterou jsem nalezl na Internetu zdarma k užití, byla tvorba editoru spíše hrou než náročnou prací. S postupem času jsem editor dále obohatil o možností označování pomocných značek vedle řádků (např. obrázek „je zde chyba, ještě se sem vrátím“) a především o funkce pro vkládání celých bloků textů na stisk klávesy či jejich kombinace (makra). Tuto funkci jsem pojmenoval rychlé vsuvky.

Editor jsem propojil pomocí horkých kláves s překladačem `pdfcslatex` (volá se BAT soubor umístěný v cestě, jeho úkolem je nastavit proměnné prostředí pro TEX a provést překlad). Díky dávkovému souboru není problém nahradit malý fpTEX později např. TEXLive a používat nadále editor z Ligatury. Na editor jsou napojeny i další externí programy – kontrola pravopisu, ovlnkování textu a zobrazování přeloženého výsledku.

Zatím ve verzi 0.05 (aktuální verze) obsahuje editor ještě drobné chybky a nepodporuje editaci více souborů najednou, ale jeho funkce postačily k napsání mé rigorózní i disertační práce a mé ženě k vytvoření práce diplomové. Současně i tento článek je tvořen v Ligatuře. Říká se, že pokud dokáže nástroj pracovat sám nad sebou, je konečně otestován. V případě Ligatury to platí dvojnásob, protože je u mne jako autora v každodenním používání.

Editor pomocí nástrojů ToolBar2000 získal vzhled moderní MS Office aplikace, abychom usnadnili uživatelům jejich přechod na tento editor.

Zobrazování výsledků

Zobrazování výsledků v Adobe Acrobat Readeru nese s sebou dva základní problémy. Tím prvním je nutnost zajistit instalaci celého prohlížeče na cílovém počítači a druhým problém nemožnost generovat nové PDF, dokud je staré v Adobe Acrobat Readeru otevřené.

Uvedené skutečnosti mne vedly k analýze dalších možností jak zobrazovat PDF. Při tomto výzkumu jsem u Adobe objevil komponentu `pdf.ocx`, což je ActiveX varianta PDF prohlížeče (Adobe Acrobatu Readeru). Tato komponenta může být integrována do okna Ligatury (tzn. mimo jednoho ActiveX prvku není nutné nic nového instalovat) a okno je částečně pod naší kontrolou. Je možné např. realizovat přegenerování PDF a jeho znovuzobrazení.

Uvedené ActiveX má i několik nevýhod:

- pokud je otevřen plný Adobe Acrobat Reader a následně uzavřen, jsou zničena i ActiveX okna (zjevně chyba),
- ActiveX rozhraní nezná funkci pro zjištění aktuální stránky (nejsme tedy schopni si ji zapamatovat a při znovuotevření stejného dokumentu není možné přejít ihned na příslušnou stránku),
- práce s ActiveX je natolik komplikovaná, že ne vždy se komponenty chovají podle očekávání (to klade zvýšené nároky na programování editoru).

Dosavadní zkušenosti z používání ActiveX komponenty prohlížeče PDF ukazují, že se jedná o plnohodnotnou náhradu Adobe Acrobat Readeru.

Kontrola pravopisu

V tomto okamžiku není propojení na kontrolu pravopisu realizováno. Předpokládám využití softwarových produktů `ispell` nebo `aspell`, přičemž vstupní text bude zasílán do uvedeného programu a na základě reakcí tohoto programu budou zobrazeny variantní správná slova pro každý nalezený gramatický jev. Druhou možností by mělo být trvalé kontrolování periodickým spouštěním tohoto programu na pozadí a podtrhování špatných výrazů červenou vlnovkou tak, jak to dělá program MS Word.

Vzhledem k častému výskytu překlepů a gramatických chyb v mých dokumentech bude nutné se na tuto funkci zapojenou do editoru soustředit co nejdříve.

Další plán rozvoje

Uvedený projekt je teprve v začátcích. Byl započat v prvních jarních měsících tohoto kalendářního roku a v tomto okamžiku umožňuje (s několika známými chybami, které se zatím nedaří opravit) plně editovat, překládat a zobrazovat poměrně rozsáhlé texty (chemická diplomka apod.).

Pro nejbližší dobu předpokládám dokončení stávajících funkcí uvnitř Ligatury a doplnění kontroly pravopisu. V tomto okamžiku při použití vhodného instalačního programu bude možné vyrobit jednoduchý instalační balíček. Průběžně jsou betaverze distribuce testovány širokým okruhem mých známých, kteří s TEXem pracují nebo začínají pracovat.

Další směry rozvoje projektu musí být stanoveny podle potřeby v budoucnosti. Po dosažení určitého základního stavu by bylo vhodné informovat o tomto projektu českou TEXovou veřejnost pro zajištění podpory této distribuce i jinými uživateli. Mezi tyto budoucí potřeby patří také webové stránky produktu, podpůrnou diskuzní skupinu apod.

Určitě bych v budoucnu nechtěl, aby se z programu stala „další“ existující distribuce, kterou by bylo nutné udržovat, ale která by nepomáhala základnímu cíli šířit TEX mezi nové uživatele.

Další zajímavá možnost rozvoje spočívá ve vytvoření desítek standardních českých stylů pro dopis, fakturu, zprávu apod., které by umožnily i běžnému uživateli začít produkovat typograficky kvalitní předpřipavené dokumenty (podobným stylem nabízí tyto dokumenty/šablony např. i program MS Word).

Závěr

Projekt Ligatura je zaměřen na vytvoření malé české TEXové distribuce zaměřené na produkci PDFCSLATEXových výstupů. Vychází z velmi zjednodušené distribuce fpTEXu a je dále obohacena vlastním jednoúčelovým editorem (určeným jen pro TEXování) a zobrazovačem PDF výstupů. Do budoucna se počítá s rozšířením editoru o funkce kontroly pravopisu a také přípravou TEXových stylů a dokumentů pro začínající uživatele. Jedním z konečných cílů je být na CD jako součást knihy o TEXu pro začátečníky. Projekt v současné betaverzi můžete nalézt na webové adrese `http://ligatura.milansorm.cz/`. Program bude po dokončení první produkční verze šířen nadále jako open-source (ale alespoň začátek bych rád podržel v pevných rukou a tak je program v betaverzích šířen jako freeware).

Rád bych na závěr poděkoval všem uživatelům a zejména mé ženě za dosavadní používání tohoto projektu při každodenním TEXování, kterým pomáhají postupně odstraňovat chyby, které brání nasazení produktu u uživatelů – začátečníků.

Uvítám také veškeré zájemce o podporu tohoto projektu ať již přiloženou rukou k vývoji (editor, distribuce), radou (co s chybami `pdf.ocx`) či testováním. Další diskuze o projektu je možná na mém e-mailu `sorm@pef.mendelu.cz`.