

Zpravodaj Československého sdružení uživatelů TeXu

Zdeněk Wagner

Babylón v TeX Live 2007

Zpravodaj Československého sdružení uživatelů TeXu, Vol. 17 (2007), No. 1, 21–23

Persistent URL: <http://dml.cz/dmlcz/150024>

Terms of use:

© Československé sdružení uživatelů TeXu, 2007

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>

V příloze tohoto čísla Zpravodaje najdete DVD T_EX Live 2007. V této verzi ještě není obsažen nový modul pro češtinu a slovenštinu, o němž jste si mohli přečíst v článku Petra Tesaříka. Musíte si jej nainstalovat sami, a přitom zjistíte, že to podle návodu jaksi nejde. T_EX Live 2007 totiž má jednu novinku, již je X_ƎT_EX. A kvůli němu došlo k drobné změně v `language.dat`. Instalaci nové češtiny a slovenštiny je tedy nutno provést nepatrně jinak.

Návod Petra Tesaříka může na méně zkušené uživatele působit poněkud komplikovaně. Návod je totiž psán obecně pro jakoukoliv distribuci T_EXu, takže začátečník si s tím nemusí poradit. A pokud chcete češtinu a slovenštinu instalovat do více počítačů, například do stolního počítače doma, do počítače v zaměstnání a do notebooku, nějaký automatický skript by se hodil. Takový skript byl vytvořen a funguje nejen v T_EX Live, ale i v teT_EXu (včetně verze pro OS/2 resp. eComStation) a v MiKT_EXu. Instalační skript zajistí též funkčnost češtiny a slovenštiny v X_ƎL^AT_EXu. Odkaz na instalační balíček najdete na www.cstug.cz v rubrice software.

Po instalaci češtiny a slovenštiny do Babylónu je nutno znovu vygenerovat formáty. Pokud čtete tento článek ještě před instalací T_EX Live, máte jistou výhodu. T_EX Live totiž při instalaci formáty nevytváří, ale pouze nakopíruje datové soubory a připraví konfiguraci. Formát se vygeneruje automaticky při jeho prvním použití. Pokud nainstalujete češtinu a slovenštinu (a případně hindštinu) ihned po instalaci T_EX Live, nemusíte se o generování formátů starat.

Po instalaci T_EX Live si nastavíme PATH tak, jak je to instalačním skriptem navrženo. Pak rozbalíme instalační balíček češtiny, slovenštiny a hindštiny. Instalační skript není kumulativní, je nutno všechny jazyky nainstalovat najednou. Pokud nainstalujete jen češtinu a slovenštinu budete chtít doinstalovat později, budete muset nejprve celý balíček odinstalovat a potom oba jazyky nainstalovat znovu, což může vést za jistých okolností k problémům. Podrobnější popis je v manuálu dodávaném s balíčkem. Pokud instalujeme do čerstvě nainstalované distribuce T_EX Live, nebudeme generovat formáty.

Poté se podíváme do konfigurace formátů. Uživatelé nových linuxových distribucí, kteří mají v `locales` nastaveno kódování UTF-8, pravděpodobně budou chtít aktivovat `encTEX` [1], aby i české písmeno bylo pro T_EX obyčejným znakem kategorie 11 a nemuselo se číst s využitím aktivních znaků po částech. Při použití `encTEX`u totiž i česká a slovenská písmena s diakritikou fungují správně v konstrukcích s `\futurelet` a v podmínkách `\if`. `EncTEX` zapneme v (pdf)L^AT_EXu parametrem `-enc`, ale nesmíme změnit defaultní kódování, to musí zůstat `cp227`.

V opačném případě se nenačtou bezchybně vzory dělení pro další jazyky. Po spuštění programu TEXCONFIG a volbě položky „Formats“ se otevře v editoru soubor `fmtutil.cnf`. Konfigurace (pdf)LaTeXu se zapnutým `encTeXem` vypadá např. takto:

```
latex pdfetex language.dat -enc -default-translate-file=cp227.tcx *latex.ini
pdflatex pdfetex language.dat -enc -default-translate-file=cp227.tcx *pdflatex.ini
```

Podobně můžete aktivovat `encTeX` pro \mathcal{C} plain. Postupujte podle komentářů v souboru `fmtutil.cnf`. Tím je instalace hotova a můžete začít psát dokumenty.

Se zapnutým `encTeXem` lze volit vstupní kódování jak pomocí TCX tabulek, tak novými primitivami `encTeXu`. Ty využijeme zejména v případě, kdy je vstupní soubor zapsán v kódování UTF-8. Použijete-li fonty v kódování IL2 (\mathcal{C} -fonty), pak český dokument bude začínat řádky:

```
\input csenc-u
\documentclass{article}
\usepackage[IL2]{fontenc}
\usepackage[czech]{babel}
```

Soubor `csenc-u.tex` je součástí `encTeXu` a najdete jej v `TeX Live`. Chcete-li použít fonty v kódování T1, změní se jen dva řádky:

```
\input csenc-u-t1
\documentclass{article}
\usepackage[T1]{fontenc}
\usepackage[czech]{babel}
```

Soubor `csenc-u-t1.tex` včetně instalačního návodu najdete na WWW serveru autora.

Poznámka: pro funkčnost hindštiny v Babylónu potřebujete nejnovější verzi balíčku DEVNAG z CVS [2].

Literatura

1. Petr Olšák: Nový `encTeX` – kódování UTF-8 v `TeXu`. Zpravodaj Československého sdružení uživatelů `TeXu`, **13** (2), 98–106 (2003).
2. Devanāgarī for `TeX`. <http://devnag.sarovar.org/>

Summary: Babel in `TeX Live 2007`

With inclusion of `XYTeX` into `TeX Live` the structure of `language.dat` has been changed slightly. Due to this fact the new Czech and Slovak module, which is

not yet distributed with official Babel, cannot be installed smoothly. The article introduces an installation package of the new module not only for T_EX Live but also for other famous T_EX distributions. Functionality of X_YL^AT_EX is also preserved.

Abecední řazení a sestavování rejstříků výhradně pomocí makrojazyka T_EXu

PETR BŘEZINA

Ve světě T_EXu se pro abecední řazení rejstříků obvykle užívá externích programů, jako je např. `makeindex`. T_EX ovšem poskytuje silný makrojazyk, který umožňuje vytvořit mechanismus schopný kvalitně provádět řazení podle abecedy. Hlavní výhoda využití takového mechanismu pro tvorbu rejstříků by spočívala v tom, že dokument by bylo možno zpracovat včetně rejstříku na každém počítači, kde je nainstalován T_EX, nezávisle na dalším programovém vybavení (externí programy nemusí být všude k dispozici, mohou být závislé na použitém operačním systému atd.). Přesto se v praxi této možnosti většinou nevyužívá. V jedenácté kapitole knihy *Typografický systém T_EX* [3] autor vyjmenovává několik důvodů: „Algoritmus by byl velmi těžkopádný, zatěžoval by paměť T_EXu, zdržoval by formátování a v neposlední řadě úpravy algoritmu pro třídění podle národních abeced by byly velmi obtížné.“ Tato nedůvěra ve schopnosti T_EXu se pro mě stala výzvou, abych se pokusil v makrojazyku T_EXu napsat prakticky využitelný program na abecední řazení.

Při tvorbě makra `sort`, jak jsem nakonec nazval výsledek svého úsilí, jsem se inspiroval programem `csr` Petra Olšáka [cf. 2], odkud jsem převzal myšlenku čtyřprůchodového abecedního řazení a myšlenku třídící tabulky. Posléze jsem k tomuto makru udělal soubor obslužných maker na tvorbu rejstříku; při návrhu vzhledu rejstříku jsem našel inspiraci ve článku *An indexing facility for T_EX* [4]. Celý makrobaliček, který dostal název `index`, si čtenář může stáhnout z mých internetových stránek na adrese <http://www.volny.cz/petr-brezina/>.

Poté co jsem dokončil makrobaliček `index`, jeví se mi výše citované argumenty jako nepřesvědčivé. Pokusím se proto v krátkosti zhodnotit výhody a nevýhody svého řešení v konkurenci s externími programy. Jedna významná přednost makrobaličku `index` zazněla již v prvním odstavci tohoto článku; na tomto místě ji už opakovat nebudu a přejdu k dalším vlastnostem. Řadicí algoritmus lze pomocí třídící tabulky snadno přizpůsobit pro různé jazyky; navíc je celé makro `sort` pouze T_EXovský kód, takže si ho šikovný T_EXista může vždycky přizpůsobit.