

O bibliografickém systému BibTeX

Vítězslav Švejdar

12. února 2005

1 Úvod

BibTeX je subsystém systému L^AT_EX, který umožňuje automatizovat vytváření seznamů literatury při psaní odborných publikací. Tento text je stručnou informací o jeho použití obecně a o jeho použití v počítačové síti Filozofické fakulty. Uvádím zde také návod na použití mých dodatků k BibTeXu a náměty k vytváření univerzálních bibliografických databází, nezávislých na tom, zda cílová publikace je česky nebo anglicky.

T_EX je jak známo založen na názoru, že při vytváření dokumentů vysoké typografické kvality mají být úvahy o vizuální podobě dokumentu odděleny od úvah o jeho logické struktuře. To znamená, že autor dokumentu má o každé části dokumentu určit, je-li odstavcem, matematickou větou, nadpisem kapitoly nebo něčím jiným. Zvlášť má pak (týž autor nebo někdo, kdo rozumí typografii) určit, jak velké mají být vertikální mezery mezi odstavci a jak má být odsazen jejich první řádek, jakým písmem mají být vysázena znění matematických vět, jak velkým a jak tučným písmem se mají sázet nadpisy kapitol a že v nich je nebo není povoleno dělení slov.

BibTeX umožňuje oddělit úvahy o logické struktuře a o vizuální podobě i při psaní seznamů literatury. Jeho vstupem je *bibliografická databáze* (nebo více takových databází najednou), a dále *bibliografický stylový soubor*. Bibliografická databáze je textový soubor, ve kterém jsou údaje o knihách, článcích a jiných bibliografických položkách zapsány v pevně definovaném formátu. Stylový soubor určuje, jakým písmem a v jakém pořadí mají být v seznamu literatury finálního dokumentu zapsány údaje o autorovi, časopisu, čísle, roku vydání atd. Určuje také, jak mají být vizuálně odlišeny například knihy od časopiseckých článků.

2 Spolupráce BibTeXu a L^AT_EXu

Máte-li bibliografickou databázi, stačí vědět, jak se *jmenují* záznamy v ní uvedené, a můžete se na ně při psaní svého zdrojového textu odkazovat pomocí příkazu `\cite`:

```
V~\cite{papa:comp94} je vyvrácena domněnka  
uvedená v~\cite{ladn:mode77}.
```

Pak stačí do místa textu, kde má \LaTeX vytvořit seznam literatury, napsat příkazy například

```
\bibliographystyle{abbrv}
\bibliography{log,compsci}
```

Tyto příkazy předpokládají, že ve své \TeX ovské instalaci máte stylový soubor `abbrv.bst` a dvě bibliografické databáze `log.bib` a `compsci.bib` obsahující záznamy (knihy, články, ...) pojmenované `papa:comp94` a `ladn:mode77`. $\text{Bib}\TeX$ tedy ze sjednocení databází jmenovaných mezi složenými závorkami v argumentu příkazu `\bibliography` vybere ty záznamy, které se ve vašem dokumentu vyskytují jako argumenty příkazu `\cite`, sestaví z nich seznam literatury a dá mu úpravu podle pokynů uvedených ve stylovém souboru jmenovaném mezi složenými závorkami v argumentu příkazu `\bibliographystyle`. Seznam literatury sestavený příkazem `\bibliography` má pak tvar jako je například v *tomto* dokumentu vidět na str. 8, výsledek užití příkazů `\cite` je vidět například v poslední větě tohoto odstavce. Chcete-li, aby do seznamu literatury byla zařazena nějaká položka, přestože v textu není explicitně citována, užíjte příkaz například `\nocite{shoe:math67}`. Tento příkaz negeneruje žádný text v místě, ve kterém je použit. Někdy se může hodit i příkaz `\nocite{*}`, který do seznamu literatury zařadí *všechny* dostupné položky. Některé odborné časopisy poskytují svým autorům vlastní stylové soubory, které definují různé varianty příkazu `\cite`. Argumentem příkazů `\cite` i `\nocite` může být i výčet, tj. několik jmen záznamů oddělených čárkami. Podrobná dokumentace k $\text{Bib}\TeX$ u je v knihách [Lam94] a [GMS94].

Požadavky na zařazení určitých položek do seznamu literatury našeho dokumentu, které \LaTeX u sdělíme použitím příkazu `\cite`, \LaTeX uloží do pomocného souboru, který má stejné jméno jako zdrojový text a typ `aux`. $\text{Bib}\TeX$ je implementován jako samostatný program, který zpracuje soubor typu `aux` a vytvoří z něj soubor typu `bbl`. Soubor `bbl` je pomocí příkazu `\include` v rámci expanze makra `\bibliography` vložen do našeho dokumentu, což může nastat nejdříve při následujícím zpracování dokumentu \TeX ovským překladačem. Pak nastane situace podobná, jako když nově použijeme příkaz `\label`: všechny odkazy jsou správně až po dalších dvou překladech. Přesnější představa o spolupráci \LaTeX u a $\text{Bib}\TeX$ u pravděpodobně není nutná, je však dobré vědět, že vyžádáte-li si pomocí příkazu `\cite` nebo `\nocite` bibliografickou položku, která dosud požadována nebyla, jsou nutné *čtyři* kroky, než je odkaz v pořádku a překlad negeneruje varovnou hlášku: \LaTeX , $\text{Bib}\TeX$, \LaTeX a opět \LaTeX .

3 Integrace $\text{Bib}\TeX$ u do MikTeXu

Nadále předpokládám, že pracujete s instalací MikTeXu, kterou jste si v počítačové síti Filozofické fakulty zpřístupnil(a) pomocí instalačního programu `MikTxCSetup.exe` nebo podle návodu `MikTxCCustom.html`, nebo která byla na

nějakém jiném počítači pořízena například podle návodu `MikInst24.html`. V tom případě máte v záložce Roots programu MikTeX Options (nadále MO) nastaveno několik \TeX ovských adresářů (stromů), z nichž první, *soukromý konfigurační adresář*, se jmenuje `texmflocal` (nejspíš `h:\MikTeX\texmflocal`). Kromě něj máte ještě globální konfigurační adresář pojmenovaný rovněž `texmflocal` (tj. `z:\MikTeX\texmflocal`), sekundární konfigurační adresář `texmfvar` a instalační adresář `texmf`. Soukromý konfigurační adresář máte označen jako Local TEXMF directory, čímž je řečeno, že tento adresář hraje také úlohu adresáře pro generování fontů. Což teď není důležité. Pracujete-li s domácí instalací, máte pravděpodobně jen jeden konfigurační adresář `texmflocal` a jako adresář pro generování fontů je určen adresář `texmfvar`. Vyhledávací cesta je nastavena na adresář `texmf\miktex\bin`.

Máte-li ve svém pokusném adresáři ukázkový soubor `demo.tex`, ukázkovou bibliografickou databázi `texinfo.bib` a pomocný soubor `demo.aux` vzniklý překladem souboru `demo.tex`,¹ můžete fungování BIB \TeX u ověřit tak, že v příkazovém řádku si pokusný adresář pomocí příkazu `cd` nastavíte jako pracovní a pak zadáte příkaz `bibtex demo`. Tím je z adresáře `texmf\miktex\bin` spuštěn program `bibtex.exe`, zpracován soubor `demo.aux` a vytvořen pomocný soubor `demo.bbl`. Po dalších dvou překladech souboru `demo.tex` překladačem `cslatex` se už neobjeví žádné varovné hlášky a výsledný dokument bude vybaven kýženým seznamem literatury. Při zpracování souboru `demo.aux` byly požadované bibliografické databáze (tj. soubor `texinfo.bib`) hledány v pracovním adresáři a pak v adresářích `bibtex\bib` jednotlivých \TeX ovských stromů. Bibliografický stylový soubor `abbrv.bst` byl hledán v pracovním adresáři a pak v adresářích `bibtex\bst` jednotlivých \TeX ovských stromů, což se podařilo až v posledním, `texmf\bibtex\bst`.

Z uvedeného vyplývá, že pokud si pořídíte specializované bibliografické stylové soubory nebo bibliografické databáze určené pro jediný dokument (přesněji řečeno pro dokumenty umístěné v jediném adresáři), můžete je prostě okopírovat do příslušného (pracovního) adresáře a o víc se nestarat. Stylové soubory a databáze, které mají být přístupné bez ohledu na umístění zdrojového souboru, ve kterém mají být použity, můžete umístit do adresáře `bibtex\bst` resp. `bibtex\bib` svého soukromého konfiguračního adresáře. Nesmíte ovšem zapomenout, že o každé změně v soukromém konfiguračním adresáři (o přidání či odebrání souboru) musíte svou instalaci zpravit provedením funkce `refresh fndb`, čili stisknutím tlačítka Refresh na záložce General programu MO. Kromě souborů, které do své instalace takto přidáte, můžete samozřejmě využívat také stylové soubory a databáze, které naleznete v centrální části instalace.

Podrobný návod k psaní bibliografických databází je v již zmíněných knihách [Lam94] a [GMS94]. Určitý návod k vytváření bibliografických stylových souborů je v druhé z nich. Vytváření vlastních stylových souborů ale nedoporučuji.

¹Tj. zadáním příkazu například `cslatex demo`; první takový překlad musel hlásit nedefinované odkazy.

Jednak to není snadné, protože stylový soubor je vlastně program ve specifickém a nepříliš názorném programovacím jazyce, jednak v adresáři `texmf\bibtex\bst` jsou desítky stylových souborů, ze kterých si lze vybírat. Naproti tomu vytváření vlastních bibliografických databází se vyhnout nelze. Avšak není to tak obtížné, prohlédnete-li si formát databází v adresářích `bibtex\bib`, snadno napíšete databázi vlastní.

Kromě programu `bibtex.exe` je v instalaci MikTeXu ještě alternativní program `bibtex8.exe`. To je tzv. “osmibitový BIB_TE_X,” který přijímá i znaky z horní poloviny tabulky ascií a je těmto znakům ochoten přiřadit správné (předem řečené v pomocném souboru) abecední pořadí. Protože `bibtex8` je použitelný při psaní dokumentů jak v angličtině, tak v češtině, můžete jej použít pro všechny účely místo programu `bibtex`. V profilech editoru TxC vytvořených programem MikTxCSetup.exe je volání programu `bibtex8` již přednastaveno v profilu LaTeX ⇒ DVI, BBL. Potřebujete-li takový profil vytvořit ručně, postupujte následovně. V menu Build / Define Output Profiles zvolte profil určený pro spouštění BIB_TE_Xu, nebo takový profil vytvořte okopírováním profilu LaTeX ⇒ DVI. Odzaškrtněte checkbox Do not use BibTeX in this profile. V editačním okénku Path to BibTeX executable, v kterém je pravděpodobně vyplněno `bibtex` (nebo obsírněji `bibtex.exe` případně i s cestou), vyplňte `bibtex8`. V okénku Command line arguments to pass to BibTeX vyplňte `-B -c cp1250cz.csf "%bm"`. Parametrem `-B` je kapacita programu nastavena na “big,” což je možná na úkor rychlosti, ale pochybuji, že lze pozorovat nějaký rozdíl. Parametrem `-c cp1250cz.csf` je řečeno, že české abecední pořadí písmen s diakritikou určuje soubor `cp1250cz.csf`. Ten jsem pro kódování Win1250 pořídil převedením podobných souborů pro cizí jazyky a v instalaci na Filozofické fakultě je umístěn v centrálním konfiguračním adresáři. V domácí instalaci si jej do konfiguračního adresáře můžete umístit, a to do `bibtex\csf` (a nezapomeňte na refresh fndb). Kdybyste se souborem `csf` chtěl(a) experimentovat (třeba kvůli nějakému jinému kódování), je dobré vědět, že MikTeX *nepočítá* s jeho umístěním do pracovního adresáře, čili *musíte* jej umístit do konfiguračního adresáře své instalace. Pokud v bibliografických databázích nehodláte používat písmena s diakritikou (nebo hodláte diakritiku psát pomocí T_EXovských příkazů, o čemž ještě bude řeč), můžete vystačit s programem `bibtex.exe`. V tom případě v okénkách určujících příkazový řádek vyplňte v prvním `bibtex` a v druhém `"%bm"` (to tam pravděpodobně bylo vyplněno již od začátku). Je dobré vědět, že programy `bibtex8` a `bibtex` mají odlišné názory na otázku, zda publikace od určitého autora uvádět dřív nebo později než publikace téhož autora napsané se spoluautory. Jiné rozdíly jsem mezi nimi nenašel.

4 Univerzalizované bibliografické stylové soubory

Mezi stylovými soubory v `texmf\bibtex\bst` jsou čtyři *standardní stylové soubory*. `Plain.bst` vypíše seznam literatury v pokud možno nejběžnějším formátu s polož-

kami označenými čísly v hranatých závorkách. `Abbrv.bst` se zachová podobně, místo křestních jmen autorů ale napíše zkratky. Zkrátí také jméno měsíce (např. “Dec.” místo “December”), je-li uvedeno (v poli `month`). `Unsr.bst` funguje jako `plain`, ale zachová pořadí položek v bibliografické databázi, záznamy netřídí. A konečně `alpha.bst` zachází s křestními jmény také jako `plain`, ale do hranatých závorek místo čísla položky uvede zkratku, kterou automaticky sestaví ze jmen autorů a roku vydání. To je použito v *tomto* poučném textu.

Z těchto čtyř souborů jsem utvořil (a umístil do globálního konfiguračního stromu) jejich *univerzalizované verze*, jejichž názvy začínají písmenem `u`, tedy `uplain.bst`, `uabbrv.bst`, `uunstr.bst` a `ualpha.bst`. Univerzalizovanou verzí míním, že slova jako je `volume` nejsou v souboru pevně “zadrátována,” jednoduchým parametrem lze určit, zda se ve výsledném dokumentu má objevit “volume” nebo “svazek”. Univerzalizovaná verze je tedy použitelná pro anglické i české publikace.

Předpokládejme, že ve svém česky psaném dokumentu, řekněme disertaci, citujete záznamy z databáze `compsci.bib`, a dále záznamy z druhotné databáze `mythesis.bib`. Plně Vám vyhovují zkrácená křestní jména, ostatně u některých autorů plná křestní jména ani neznáte. V tom případě na ono místo dokumentu, kde má být umístěn seznam literatury, napíšete

```
\bibliographystyle{uabbrv}
\bibliography{czx,compsci,mythesis}
```

a je to. Ještě je třeba vysvětlit smysl souboru `czx.bib`. Bibliografická databáze může kromě záznamů obsahovat také definice textových řetězců (neboli zkratok). Soubor `czx.bib` je zvláštní bibliografická databáze, která obsahuje nula záznamů a pouze ony definice zkratok. Tento soubor je zodpovědný za to, abyste místo “A. Tarski, A. Mostowski, and R. M. Robinson, editors, *Interpretability and Undecidability*, volume 60 of *Mundus Demonstrandus*” měli v disertaci vysázeno “A. Tarski, A. Mostowski a R. M. Robinson, editoři, *Interpretability and Undecidability*, svazek 60 řady *Mundus Demonstrandus*”. Pokud jste při pořízení svých databází dodrželi doporučení uvedená dále v oddílu 5, také “ch” bude na svém místě: článek od Churcha bude zařazen až za knihou od Hájka. Klíčových slov jako je “and” a “editors” jsem ve stylových souborech našel patnáct.² Písmeno `x` v názvu souboru `czx` má připomínat slovo *prefix*. Prefixem je míněn soubor, který ničím “materiálním” nepřispívá, avšak musí být jmenován před ostatními.

Rozhodnete-li se později, že budete publikovat anglickou verzi disertace, stačí z argumentu příkazu `\bibliography` odstranit soubor `czx`. Klíčová slova pak budou anglicky a Church se přesune před Hájka a dokonce před Cohena. V anglické publikaci se ovšem nemusíte omezovat na univerzalizované stylové soubory

²Cvičení: našli jste *pět* rozdílů, kterými se text v uvozovkách určený do české publikace liší od předchozího textu pro anglickou publikaci?

začínající na `u`, můžete použít kterýkoliv z ostatních, které najdete v instalaci. Některé bibliografické soubory vyžadují, aby byl do zdrojového textu pomocí příkazu `\usepackage` importován také nějaký L^AT_EXovský stylový soubor. Avšak netýká se to žádného z těch, o kterých je zde řeč.

Kromě čtyř standardních jsou v konfiguračním adresáři připraveny ještě další dva univerzalizované stylové soubory `uacm.bst` a `uanabbrv.bst`. `Uacm.bst` sází jména autorů pomocí kapitálek (malých velkých písmen). `Uanabbrv.bst` připouští v záznamech kromě standardních polí jako je `author` a `title` ještě pole `annotate` pro psaní anotací.

5 Formát bibliografické databáze

je podrobně popsán v příručkách [Lam94] a [GMS94]. V nich si lze také přečíst například o formátu víceslovních příjmení typu “Lopez Fernandez” a o jiných problémech. V tomto oddílu uvádím jen několik poznámek a nápady, které jsem nad rámec popisu v obou příručkách implementoval pro pořizování databází nezávislých na jazyce. Před psaním vlastních bibliografických databází se také můžete inspirovat prohlédnutím databází, které najdete v instalaci, zejména upozorňuji na databázi `log.bib`. Ta je zamýšlena jako databáze publikací, u kterých je naděje, že budou citovány v pracích vznikajících na katedře logiky. Stále se rozrůstá a obsahuje nyní asi 200 záznamů. K jejímu vytvoření přispěl Michal Peliš. Obsah databáze `log.bib` je vidět v souboru `log.pdf`. Logikové mohou databázi `log.bib` přímo použít, tj. na její záznamy se odkazovat bez jakéhokoliv kopírování či přepisování jejího obsahu. Pro tento účel může pomoci, že jména záznamů obsažených v databázi `log.bib`, tj. ta jména, která můžete použít jako argument příkazu `\cite`, jsou v souboru `log.pdf` zvýrazněna v rámečcích. Kdybyste chtěl(a) použít databázi `log.bib`, avšak nikoliv některý z mých univerzalizovaných stylových souborů, musíte použít také prefixový soubor `lx.bib`, tj. napsat například `\bibliography{lx,mythesis,log}`. Zbývající typy jsou určeny i nelogikům.

U všech bibliografických souborů, u kterých připadá v úvahu přenesení pod nějakou jinou instalaci, doporučuji nevyužívat onu osmibitovou vlastnost programu `bibtex8` a diakritiku psát pomocí T_EXovských příkazů. Než se spoléhat na to, že například v redakci zahraničního časopisu budou něco vědět o kódování češtiny, je asi lepší riskovat, že háčky a čárky ve jménech českých autorů budou nepatrně jiné, než kdyby byly použity kanonické české fonty. Takže jmenují-li se autoři F. Müller a G. Řehák, pište

```
author = "F. M{\u}ller and G. {\v R}eh{\`a}k",
```

Přitom respektujte to, že *zobecněný akcentovaný znak* je pro BIB_TE_X definován jako text, který začíná levou složenou závorkou bezprostředně následovanou zpětným lomítkem a který končí pravou složenou závorkou na příslušné úrovni. Nepište tedy `\v{R}eh\`ak` ani `{\v Reh\`ak}`, přestože z hlediska TeXu by to

také bylo dobře. Poctivě musím přiznat jednu nevýhodu psaní akcentů pomocí \TeX ovských příkazů. Zobecněný akcentovaný znak $\{\backslash v e\}$ (což je totéž jako $\{\backslash v{e}\}$) považuje $\text{BIB}\TeX$ za znak umístěný v abecedě mezi e a f. To znamená, že publikaci od Běliny $\text{BIB}\TeX$ při užití mnou doporučeného způsobu psaní umístí až za publikaci od Beránka, ačkoliv podle českých zvykostí by to mělo být naopak (protože ‘ě’ a ‘e’ mají stejnou váhu a ‘l’ předchází ‘r’). Jsem ale přesvědčen, že jen v málokterých dokumentech se tato nevýhoda projeví.

Místo všech jen trochu běžných textů užívejte *zkratky*. Pište např.

```
journal = js1,
```

tím se vyhnete úvahám, zda psát "J. Symb. Logic", "J.~Symb. Logic", nebo "Journ. Symb. Logic", a také budete mít zaručeno jednotné vysázení všech výskytů. Takovéto zkratky užívejte pro jména časopisů, pro jména měsíců (jan, feb, atd.), známé vydavatele (AW je "Addison-Wesley", NH je "North-Holland") a pro často se vyskytující texty (jako je "Unpublished" nebo "Lecture Notes"). Existující zkratky lze zjistit vyhledáním slova `MACRO` v příslušném stylovém souboru a vyhledáním slova `@String` v bibliografické databázi. Další zkratky můžete definovat ve vlastních databázích či prefixových souborech. Zkratky a konstanty (explicitně uvedené texty) můžete spojovat dohromady pomocí znaku #:

```
note = "Po z{\`a}niku nakladatelstv{\`i} Bibliopolis
      znovu vydal " # AW,
```

Přítom mezera za slovem vydal je významná (odděluje "vydal" od "Addison"), mezery okolo znaku # významné nejsou. Malá a velká písmena ve jménech zkratek také významná nejsou.

Program `bibtex8` má sice vědomost o abecedním pořadí českých písmen s diakritikou, neví však, že "ch" je v české abecedě až za "h". Při použití univerzalizovaných souborů můžete tuto potíž vyřešit použitím předdefinovaných zkratek. Všechna "ch" (přesněji řečeno ta "ch", která jsou na asi tak třech prvních pozicích v příjmeních autorů, jinde to sotva má vliv na abecední pořadí) pište pomocí předdefinovaných zkratek `ch` pro velké "Ch" a `lch` pro malé "ch":

```
author = "Alonzo "#ch#"urch and Detlef S"#lch#"langer",
```

Tento způsob psaní zaručuje správné abecední pořadí citací v anglické i v české publikaci. Za zkratkou `ch` se skrývá určité \TeX ovské makro, a jiné \TeX ovské makro v případě užití prefixového souboru `czx.bib`. Obě makra vyhovují definici zobecněného akcentovaného znaku. Druhé z těchto maker se jeví jako akcent nad `IO`. Tím je ošizen třídící algoritmus $\text{BIB}\TeX$ u. Při zpracování \LaTeX em ale obě makra expandují na týž text "Ch". Zkratka `lch` se chová analogicky.

Při použití univerzalizovaných stylových souborů můžete do svých databázových položek zapracovat i další anglicko-české alternativní texty, a to pomocí konstrukce `\ifenglish .. \else .. \fi{}`. Například:

```
edition = "\ifenglish Eleventh\else Jeden{\'a}ct{\'e}\fi{ }",
```

Přitom je nutné nezapomenout na závěrečné `\fi{ }`. Všimněte si také, že před `\else` a `\fi{ }` nejsou mezery a že oba texty začínají velkým písmenem. Místo příkazu `\ifenglish` můžete použít také kratší variantu `\ifen`.

Nakonec pro jistotu dodávám, i když to nepatří k problematice zkratk a česko-anglických alternativ, že má-li publikace více autorů, autoři se spojují pomocí `and` a poslední jméno může znít `others`. Takže například z pole

```
author = "J. L. Balc{\'a}zar and J. D{\'i}az and J. Gabarr{\'o}",
```

vznikne v anglické publikaci “J. L. Balcázar, J. Díaz, and J. Gabarró”, v české “J. L. Balcázar, J. Díaz a J. Gabarró”. Z pole

```
author = "W. Hodges and others",
```

vznikne v obou případech “W. Hodges et. al.”.

Literatura

- [Chi93] *The Chicago Manual of Style*, čtrnácté vydání, 1993.
- [GMS94] Michel Goossens, Frank Mittelbach a Alexander Samarin. *The L^AT_EX Companion*. Addison-Wesley, Reading, Massachusetts, 1994.
- [Knu86] Donald E. Knuth. *The T_EXbook*. Addison-Wesley, 1986.
- [Lam94] Leslie Lamport. *L^AT_EX: A Document Preparation System*. Addison-Wesley, 1994.
- [vL92] Mary-Claire van Leunen. *A Handbook for Scholars*. Oxford University Press, New York, Oxford, opravené vydání, 1992.