# Čeština na počítačích unix12, unix21 a pc064

(Jan Kučera)

20. července 1998

## 1 Úvodní poznámky

Implementace češtiny je poměrně rozsáhlý problém, který má řadu aspektů souvisejících zejména s nejednotností jejího kódování. Podrobnější poučení o celé problematice nalezne zájemce na adrese http://www.cestina.cz/.

Z důvodů praktické potřeby je čeština na všech počítačích implementována pouze v X-Windows. Pokud by někdo potřeboval i češtinu v režimu konzole nechť mi laskavě dá vědět.

## 2 Kódování

Na všech počítačích je použito standardního unixového kódování podle normy ISO-8859-2. Technicky je samozřejmě možné vnutit systému i kódování win-1250, nevidím však důvod, proč této aroganci ustupovat. Budoucnost představuje *Unicode*, který mohu již nyní bez větších technických problémů instalovat. Problém je zatím nedostatek volných (free) fontů v tomto kódování.

## 3 Fonty

K disposici jsou proporcionální i neproporcionální fonty, bitmapové fonty i (v dohledné době budou) Type-1 fonty. Většina neproporcionálních fontů s "aliasem" typu "WxH" (např. 7x13, 10x20, . . . ), kde W je šířka a H je výška písmene v počtu pixliček, jsou české fonty. Český X-Terminál pak voláme např. příkazem

```
xterm -fn 10x20 -sb & .
```

Pro pořádek podotýkám, že možnost používat českých fontů záleží na tom, jaké fonty má k disposici X-server *na tom počítači, u kterého fyzicky pracujeme* a ne na tom jaké fonty jsou k disposici na počítači, kde běží aplikace. V případě potřeby mohu z některého z unixu (nejspíše unix12) příslušné adresáře exportovat přes NFS. Jiná možnost je zkrátka potřebné fonty si zkopírovat a instalovat (podrobnější instrukce uvedu dodatečně).

# 4 Klávesnice

X-windows v instalované verzi podporují české “locales” a rovněž mrtvé klávesy. Bohužel existují momentálně alespoň čtyři důvody k tomu, abych českých locales ani mrtvých kláves nepoužíval:

1. Ve standardní distribuci Debianu je chyba (a to i v té nejnovější), která použití čekých locales a českých klávesnic s mrtvými klávesami ztěžuje (jinými slovy, jsem líný se s tím babrat. . . a ani to na dálku moc dobře nejde).

2. Většina programů (z angloamerické oblasti) výše zmíněný standard zatím stejně nevzala na vědomí a chová se patologicky.

3. Čeští debianisté zatím standard českých klávesnic stejně nevytvořili.

4. Emulátor Windows — “Wabi” zmíněné nové standarty nepodporuje.

Vedle toho existují užitečné programy, které výše zmíněný standard podporují (např. Staroffice).

S ohledem na výše zmíněné skutečnosti jsem implementoval vlastní (velmi nestandardní) českou klávesnici bez mrtvých kláves, pouze s jednoduchým přemapováním pomocí “xmodmap”.

## 4.1 Spuštění klávesnice

Klávesnice se spouští příkazem “czech” (univerzální), resp. “czech.pc”, je-li naše klávesnice fyzicky připojena k linuxovému PC, resp. “czech.sgi”, je-li naše klávesnice fyzicky připojena k SGI. Modifikátory jsou na této klávesnici **pravý** *ALT* a *SHIFT*. Fungování klávesnice popisuje následující tabulka:

| Označení klávesy | Klávesa | Klávesa + SHIFT | Klávesa + ALT | Klávesa + SHIFT + ALT |
|---|---|---|---|---|
| 1 | ú | Ú | 1 | ! |
| 2 | ě | Ě | 2 | @ |
| 3 | š | Š | 3 | # |
| 4 | č | Č | 4 | $ |
| 5 | ř | Ř | 5 | % |
| 6 | ž | Ž | 6 | ˆ |
| 7 | ý | Ý | 7 | & |
| 8 | á | Á | 8 | * |
| 9 | í | Í | 9 | ( |
| 0 | é | É | 0 | ) |
| - | ů | Ů | - | _ |
| = | ó | Ó | = | + |
| e | e | E | ë | Ë |
| t | t | T | ť | Ť |
| u | u | U | ü | Ü |
| o | o | O | ö | Ö |
| a | a | A | ä | Ä |
| s | s | S | ß | ß |
| d | d | D | ď | Ď |
| n | n | N | ň | Ň |

ostatní znaky se zůstávají nezměněny jako na původní americké klávesnici.

## 4.2 Zrušení klávesnice

K odstranění české klávesnice (resp. k uvedení klávesnice do “amerického” stavu) slouží příkaz “unczech”, resp. “unczech.pc”, “unczech.sgi”. Po provedení této operace je klávesnice popsána následující tabulkou:

| Označení klávesy | Klávesa | Klávesa + SHIFT | Klávesa + ALT | Klávesa + SHIFT + ALT |
|---|---|---|---|---|
| 1 | 1 | ! | ú | Ú |
| 2 | 2 | @ | ě | Ě |
| 3 | 3 | # | š | Š |
| 4 | 4 | $ | č | Č |
| 5 | 5 | % | ř | Ř |
| 6 | 6 | ^ | ž | Ž |
| 7 | 7 | & | ý | Ý |
| 8 | 8 | * | á | Á |
| 9 | 9 | ( | í | Í |
| 0 | 0 | ) | é | É |
| - | - | _ | ů | Ů |
| = | = | + | ó | Ó |
| e | e | E | ë | Ë |
| t | t | T | ť | Ť |
| u | u | U | ü | Ü |
| o | o | O | ö | Ö |
| a | a | A | ä | Ä |
| s | s | S | ß | ß |
| d | d | D | ď | Ď |
| n | n | N | ň | Ň |

Znovu podotýkám, že toto řešení je pouze dočasné a patrně se v brzké době změní. Na druhé straně se mi tato klávesnice dobře osvědčila na sunech a na sgi v době, kdy jsem neměl přístup superuživatele a kdy ještě X-windows nepodporovaly mrtvé klávesy.

Perspektivně k této klávesnici vyrobím klikací ikonku, která umožní automatické přepínání.

# 5 Čeština a Ispell

nainstaloval jsem databázi češtiny pro program Ispell pro kontrolu pravopisu. Jediný rozdíl oproti použití s angličtinou (default) je příznak *-d czech*. Podpora funguje v emacsu a v LyXu, prý také v editoru joe. Zatím mám ale spíše negativní zkušenosti. Toto téma ještě rozvedu.

# 6 Čeština v Emacsu

uživatele editoru Emacs potěší (ale patrně nepřekvapí), že češtinu plně podporuje editor Emacs (včetně plnohodnotné vlastní české klávesnice, českého vyhledávání apod.)

# 7 TEX/LATEX

Český (i slovenský) TEX i LATEX2e jsou nainstalovány. Vyvolávají se příkazem *cstex*, resp. *cslatex*. Jedná se o verzi pánů Šnajdra, Wágnera a Zlatušky, vycházející z původní instalace *CS*TEXu, která se opírala o typograficky kvalitní *csfonty* pana Olšáka v kódování ISO-8859-2. Výhodou je jednak zpětná kompatibilita s rozšířeným *CS*TEXem pro DOS, dále vnitřní kompatibilita se standardním unixovým kódováním češtiny a dále vysoká typografická kvalita odpovídající českým zvyklostem (umístění akcentů, odsazování,...) Vedle češtiny podporuje instalovaná verze ještě angličtinu. Nevýhodou je nekompatibilita se světovým sandardem *Babel*, umožňujícím podporu více jazyků v jednom dokumentu (ovšem za cenu vážných prohřešků proti českým typografickým zvyklostem). Tato mezinárodní verze TEXu je spustitelná pomocí *tex*, resp. *latex*. Bohužel nepodporuje češtinu, jelikož obě verze TEXu podporující češtinu nemohou *při pužití standardních csfontů* v jedné instalaci koexistovat. Na druhé straně můžeme využít schopnosti nového *cslatexu* používat jiného kódování na vstupu a na výstupu (viz níže).

Za zmínku stojí dále pomocný program *vlna*, který zavede spojí vlnkami (tilde) jednoslabičné předložky s následujícími slovy (v místě mezery nelze takové spojení zlomit).

Následující informace se týkají pouze dokumentů vytvořených v *CS*LATEXu 2e.

## 7.1 Příklad: Použití csfontů

Chceme-li s *csfonty* dosáhnout co nejlepšího typografického výsledku, bude hlavička našeho souboru obsahovat následující řádek:

```
\usepackage{lczech}
```

Přitom dojde i k využití dalších českých maker, např. místo *chapter* se vypíše jako “Kapitola”, apod.

(Namísto latexovského makra *lczech* jsme mohli stejně dobře použít obecného texovského makra *czech*, jež v latexovém souboru stejně načítá makro *lczech*. Slovenští uživatelé mohou namísto *lczech* použít *lslovak*.

Správného načtení souboru a použití českých fontů (i když podle mých zkušeností v o něco menší typografické kvalitě, což se týká i českého dělení; rovněž českých maker nebude využito) dosáhneme kombinací následujících dvou příkazů:

```
\usepackage[latin2]{inputenc}
\usepackage[IL2]{fontenc}
```

První příkaz říká, že výstupní kódování standardních fontů (což jsou u *cslatexu* právě csfonty) je *IL2* (tj. standardní unixové kódování češtiny ISO-Latin-2, ISO-8859-2). Druhý příkaz sděluje, že soubor je právě v tomto kódování.

## 7.2 Příklad: Použití Postskriptových fontů

Sučástí standardní distribuce TeTEXu je podpora standardních postskriptových fontů ve standardním kódování *T1* (Cork). Toho můžeme s výhodou využít zejména tehdy, neobsahuje-li náš text matematické vzorce. Chceme-li např. vysázet text v Timesech (v nichž je vysázena i postskriptová verze tohoto dokumentu), bude v záhlaví následující kombinace příkazů:

```
\usepackage[latin2]{inputenc}
\usepackage[T1]{fontenc}
\usepackage{times}
```

Takto vytvořené postskriptové soubory neobsahují bitmapové fonty, tj. jsou výrazně kratší a optimálně využívají kvality výstupního zařízení. Příkazem *ps2pdf* lze navíc vytvořit ještě kratší soubor ve formátu pdf, který mohou snadno číst a tisknout pomocí programu Acrobat-Reader uživatelé praktický všech počítačových platforem (včetně Windows všeho druhu a počítačů Apple-Macintosh).

Chceme-li přitom využít českých LATEXových maker a českého dělení, bude naše hlavička ještě o něco komplikovanější. Např. pro sazbu v řezu "Utopia" obsahuje hlavička řádky:

```
\usepackage[latin2]{inputenc}
\usepackage[IL2,T1]{fontenc}
\usepackage{utopia}
\usepackage{lczech}
```

Pro správné fungování je přitom důležité uvedené *pořadí* zavádění packages. Podrobnější informace o podporovaných postscriptových fontech viz soubory **.sty* v adresáři

*/usr/lib/texmf/tex/latex/psnfss*

# 8 LyX

LyX je velmi užitečný program, který se tváří jako textový procesor (je podobný MS-Wordu), ale používá k sazbě svých dokumentů LATEX. Umožňuje i sazbu vzorečků pomocí obrazovky a myši, ale zrovna tak může uživatel psát latexovské výrazy, jež se mu přímo před očima mění ve formulky, krásně a pohodlně se v něm sázejí tabulky, umí RevTEX apod. Zatím neumožňuje načítání textů v LATEXu (tj., chceme-li s dokumentem dlouhodobě pracovat, musíme ukládat a upravovat jeho LyXovou podobu), ale na příslušném konverzním programu se intenzivně pracuje. Nejnovější verze 0.12 už nezavádí do TEXovského souboru žádná makra (s výjimkou svého loga) a hodí se tedy např. pro psaní dokumentů, určených k další úpravě. Mám s tímto programem velmi

pozitivní zkušenosti. Instalovaná verze *plně* podporuje češtinu, a to včetně *vlastní instalované klávesnice* (je nastavena jako druhá v pořadí a lze ji přepnout v options). Kontrola českého pravopisu je rovněž podporována.

Navíc jsem "předinstaloval" soubor *defaults.lyx*, který obsahuje mimo jiné nastavení implicitního kódování na *latin2* a volání maker *czech* (resp. *lczech* (např. pro správné české dělení; viz předchozí povídání o LaTeXu),

Jestliže z nějakého důvodu tohoto balíku nechceme používat, musíme v menu "Layout" v položce "Latex preamble" vymazat řádku:

```
\usepackage{czech}
```

# 9 Tisk českých dokumentů v Postscriptu

Některé programy (např. Netscape) generují postskriptové dokumenty mlčky odkazující na (neexistující) české Type-1 fonty. Pošleme-li tyto dokumenty přímo na postskriptovou tiskárnu, vytiskne se nám něco zmateného (místo očekávaných fontů v kódování ISO-8859-2 budou použity fonty v ISO-8859-1).

Chceme-li, aby se příslušný soubor vytiskl správněji, musíme použít tiskovou frontu "-Pczps". Příslušný filtr vytvoří virtuální "uživatelské" (Type-3) fonty, které už budou ve správném kódování a umístí je na začátek tisku. Chceme-li takto upravit postscriptový soubor, použijeme přímo filtru

"/usr/local/bin/filter_czps".