



Sudoku s vepsanými kandži: integrace čínských glyfů s grafikou na úrovni METAPOSTu

Denis Roegel

► To cite this version:

Denis Roegel. Sudoku s vepsanými kandži: integrace čínských glyfů s grafikou na úrovni METAPOSTu. Zpravodaj (Československého sdružení uživatelů TeXu, ISSN 1211-6661), 2010, 20 (3), pp.220-226. inria-00548912

HAL Id: inria-00548912

<https://inria.hal.science/inria-00548912>

Submitted on 27 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sudoku s vepsanými kandži: integrace čínských glyfů s grafikou na úrovni METAPOSTu

DENIS ROEGEL

Abstrakt

Článek představuje metodu, jak lze technicky za pomoci METAPOSTu vykreslit sudoku s čínskými glyfy (kandži). Makra nejsou rozsáhlá a jsou v úplné podobě součástí tohoto článku.

Na vykreslení sudoku jen s arabskými číslicemi můžete použít balíček `sudoku`; na generování, vykreslení a řešení sudoku nahlédněte, prosím, na balíček `sudoku-bundle`. Oba balíčky jsou dostupné z CTAN.ORG.

Sudoku, česky též magický čtverec, je logická hra, kterou vymyslel Howard Garns v roce 1979. Své obliby se dočkala především v Japonsku, ale prakticky se hraje po celém světě. Cílem hry je doplnit chybějící čísla 1 až 9 v předem dané předvyplněné tabulce. K dříve vyplněným číslům je potřeba dopsat další čísla tak, aby platilo, že v každé řadě, v každém sloupci a v každém z devíti čtverců byla použita vždy všechna čísla jedna až devět. Pořadí čísel není důležité. Čísla se nesmí opakovat v žádném sloupci, řádku nebo ve vnitřních malých čtvercích o devíti polích. Kdo by si to rád zkusil, nechť navštíví např. <http://profuvsvet.ic.cz/data/sudoku.html> nebo <http://sudokuonline.cz/>.

Přehledně o typech sudoku, jejich algoritmech a matematickém pozadí, viz např. Wikipedie [1]. Za zajímavost stojí, že první mistrovství světa v sudoku v roce 2006 vyhrála Češka Jana Tylová z Ústí nad Labem. Turnaj v sudoku můžete absolvovat např. na mezinárodním festivalu Czech Open. Ustavující Guinnessův rekord drží z 23. února 2008 Pavel Jaselský z Vojtovce na Slovensku.

Klíčová slova: METAPOST, hra sudoku, balíček CJK, balíček `latexmp`, kódování UTF-8, čínština, čínské glyfy, kandži.

doi: 10.5300/2010-3/220

1. Úvodem

Není tomu zas tak dávno, co jsem potřeboval ke svému povídání o čínském kalendáři využít výborný balíček Wernera Lemberga CJK. K této příležitosti jsem chtěl připravit obrázky v METAPOSTu s čínským značením. Práce to byla téměř bezproblémová.

This article is a translation of the article “Kanji-Sudokus: Integrating Chinese and Graphics”, which first appeared in *TUGboat*, Volume 29 (2), pp. 317–319, 2008. Reprinted with permission. Translation and Czech abstract by Pavel Stríž. The author took the opportunity of this translation in order to make a few minor changes for the sake of clarity, on the suggestion of the translator.

O té bezproblémovosti to ve skutečnosti není zcela pravda, pravdou však je, že v současnosti je balíček CJK stále lépe a lépe zabudováván do verzí T_EX Live. Sazba dokumentů s čínštinou, japonštinou a korejštinou (ČJK) se stává v rychle se měnícím T_EXovém světě téměř nadpozemskou záležitostí. Zdaleka tomu tak nebylo v polovině roku 2007. Stále je potřeba nainstalovat různé, řekněme, linuxové balíčky a i poté hrozí, že nám drobná, ale nenahraditelná záplata chybí. Například se podívejme na mou verzi Ubuntu z poloviny roku 2008. T_EX Live nebyl v té době tak kompletní jako dnes, chyběla mu korejská písma a byl jsem nahraný. Počítám však s tím, že to je již zapomenutá bolavá minulost a tato konkrétní situace již byla dávno vyřešená.

Abych shrnul své postřehy k polovině roku 2008, tak za pomoci T_EX Live 2007, možná s několika doplňkovými linuxovými balíčky, a zároveň s poslední verzí editoru Emacs jste naprosto soběstační k přípravě prvotřídních dokumentů s ČJK. Sazba ČJK se stala ještě snazší záležitostí, protože téměř vše můžeme sepisovat v kódování UTF-8, bez potřeby využívat konverzí vstupního souboru, např. pomocí maker editoru Emacs. Tento postup se používal ke generování souboru `.cjk`, který mohl být následně L^AT_EXem zpracován. Teď je celý proces ještě rychlejší; soubor, který připravíte, je zároveň tím, který se zpracovává.

Pro tvorbu obrázků přes METAPOST bylo toto zpracování ČJK připraveno taktéž příjemněji. Ještě donedávna, když jsme chtěli zahrnout čínštinu do METAPOSTu, tak jsme museli vygenerovat soubor `.cjk`, ten však bohužel nemohl být přímo METAPOSTem zpracován. Do souboru `.cjk` se muselo drobně zasáhnout, protože makra Emacsu formát METAPOSTu nebrala v úvahu.

Situace se mohla řešit na úrovni maker Emacsu, nu, ale ve skutečnosti, když už je tento druh konverzí do `.cjk` souboru téměř kompletně zapomenutou historií, tak i tento druh problému je bezpředmětný. Tedy má rada je nejen vstoupit do světa Linuxu, T_EX Live v poslední verzi, ale také psát dokumenty s ČJK v kódování UTF-8. Šlape to jak švýcarské hodinky!

2. Motivační ukázka

Osvětlím vám technické pozadí integrace čínštiny a METAPOSTu na následující drobné ukázce. Nakreslím si hrací pole s mřížkou, v polích však ne s arabskými číslovkami, ale s čínskými glyfy. Číslovky jsou následující 一 (1), 二 (2), 三 (3), 四 (4), 五 (5), 六 (6), 七 (7), 八 (8) a 九 (9).

Prázdná mřížka (ve všech proměnných `sol` nastavit nuly a v makru `fillgrid` přepnout `false` na `true`) a typická ukázka zadání ze hry sudoku vypadá jako následující příklad (tento konkrétní problém je převzat z Wikipedie [1]).

2.1. Vykreslení mřížky

Zadání sudoku může být nakresleno v METAPOSTu takto:

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

```

beginfig(1);
  string sol[];
  drawgrid(1.5pt,.5pt);
  % První řádek je ve spodní části mřížky.
  % Poslední řádek je naopak v její horní části.
  sol1="000080079"; sol2="000419005";
  sol3="060000280"; sol4="700020006";
  sol5="400803001"; sol6="800060003";
  sol7="098000060"; sol8="600195000";
  sol9="530070000";
  fillgrid(sol)(false);
endfig;

```

Makro `drawgrid` je přímočaře naprogramované a vykreslí potřebné horizontální a vertikální linky (s parametrem `u` nastaveným například na 1 cm):

```

def drawgrid(expr tha,thb)=
  pickup pencircle scaled thb;
  for i=0 upto 9:
    draw (i*u,0)--(i*u,9u);
    draw (0,i*u)--(9u,i*u);
  endfor;
  pickup pencircle scaled tha;
  for i=0 upto 3:
    draw (3i*u,0)--(3i*u,9u);
    draw (0,3i*u)--(9u,3i*u);
  endfor;
enddef;

```

2.2. Vyplnění polí

Předtím, než se nám podaří vypsát hodnotu do políčka, si musíme zajistit posun do pozice (i, j) , kde i je pořadové číslo sloupce a j je hodnota řádku. Počáteční hodnota je v obou směrech nastavena na jedničku a počátek tvoří levá spodní buňka. Následuje příprava makra, které si načte argumenty i , j a popisek umístí do středu příslušné buňky. V našem konkrétním případě je popisek zvětšen na 200 %, s tím, že aktuální změna závisí na velikosti hracího pole a velikosti písma.

```
def pos(expr i,j,l)=  
  label(1 scaled 2,((i-.5)*u,(j-.5)*u));  
enddef;
```

2.3. Vysázení hodnot

Abychom zajistili vysázení, řekněme, hodnoty tři na pozici (2,9), můžeme zapsat:
`pos(2,9,btex 3 etex);`

My samozřejmě budeme chtít situaci zevšeobecnit a vysázet jednotlivé hodnoty z textového řetězce `sol`. Touto úvahou mohou některé nástroje a balíčky nabídnout řešení, stejně tak i potenciální problém. Abychom této situaci s přibráním problémů jiných nástrojů předešli, rozšířme naši úvahu na:

```
pos(2,9,TEX(s));
```

kde `s` je textový řetězec k načtení makrem `TEX`. Zmíněné makro je načteno balíčkem `TEX`:

```
input TEX;
```

Vypadá to hezky, ale bohužel tento postup není příliš vhodný, důvod je ten, že `TeX` je celkově volán jedenaosmdesátkrát. A vedle toho nemáme stále zajištění podporu čínských glyfů, jakmile dojde na její volání. Hledejme ještě chvíli postup, který by byl více použitelný a lépe nám vyhovoval.

To, co jsme usilovně hledali, je balíček `METAPOSTu` nazvaný `latexmp`. Dříve zmíněný popisek na pozici (2,9) zapíšeme takto:

```
pos(2,9,texttext(s));
```

Jednou z výhod balíčku `latexmp` je, že potřebuje jen dvě spuštění `LATEXu` místo toho, aby si jej volal pro každý popisek zvlášť. Balíček navíc umožňuje snadné načtení dalších `LATEXových` balíčků, především těch pro čínštinu. Dostáváme se k tomu, jak by mohly vypadat první řádky zdrojového kódu v `METAPOSTu`:

```
input latexmp;  
setupLaTeXMP(class="article",  
  packages="CJKutf8",  
  preamble=(  
    "\let\N\newcommand"
```

```

&"\N\0{\}\N\1{一}\N\2{二}\N\3{三}"
&"\N\4{四}\N\5{五}\N\6{六}\N\7{七}"
&"\N\8{八}\N\9{九}"
&"\AtBeginDocument{"
&  "\begin{CJK}{UTF8}{bsmi}}"
&"\AtEndDocument{\end{CJK}}");

```

Funguje to tak, že si hlavička načte balíček CJKutf8, který potřebujeme na práci se vstupními soubory formátu UTF-8. Následuje definování příkazů \0 (prázdné pole), \1 (čínská číslovka 1), \2 (čínská číslovka 2), postupně až do příkazu \9.

Pro tyto číslovky zařídíme načtení nezbytného prostředí pro ČJK hned ze začátku těla dokumentu takto:

```

\AtBeginDocument{
  \begin{CJK}{UTF8}{bsmi}}

```

Otevřené prostředí L^AT_EXu uzavřeme obdobně:

```

\AtEndDocument{\end{CJK}}

```

2.4. Složení dílčích částí do fungujícího celku

V této chvíli máme připravenou definici jednotlivých buněk, můžeme si vykreslit hrací pole v podobě mřížky a máme připravená makra na čínské číslovky. Co nám zbývá? Ještě bychom potřebovali dvě maličkosti.

Za prvé bychom si přáli, abychom mohli vykreslit zadání versus řešení. Pro naše účely je zadání jen pole hodnot buněk, kdy některé jsou rovny nule. Hodnoty nula budou vypadat jako prázdné buňky. Když už jsme u tohoto přepínače, tak bychom si přáli, aby se nenulové hodnoty zobrazily buď v podobě arabských číslic nebo právě číslic čínských.

Naše zavedená konvence na hodnoty buněk je vskutku jednoduchá, neboť použijeme 4, chceme-li výstup v podobě arabských číslic, nebo \4, chceme-li číslice čínské. Je to jedna z možností. Pak už si jen dáваме pozor, máme-li před číslovkou zpětné lomítko, či nikoliv. Zvláštní pozornost věnujme nule, protože nula u arabských číslic bude prázdné pole, ovšem u čínských číslic může a nemusí být pole vyplněno glyfem pro nulu. Záleží na stylu.

Přepínač je naprogramován pod makrem **zerospace**. Makro si bere znak **s** a nahradí jej mezerou, jen pokud se jedná o nulu a režim není nastaven na čínské číslovky.

```

def zerospace(expr chinese,s)=
  if not chinese and (s="0"): " " else: s fi
enddef;

```

Dostáváme se do finále, neboť mřížka je vyplněna makrem `fillgrid`. Nechtě je prvním parametrem název pole s textovým řetězcem a druhý nechtě je přepínač mezi číslovkami čínskými a arabskými. Příkaz `substring` je použit na oddělení jednotlivých znaků.

```
def fillgrid(text grid)(expr chinese)=
  for i=1 upto 9: for j=1 upto 9:
    pos(j,i,texttext(if chinese: "\" & fi
      zerospace(chinese,
        substring(j-1,j) of grid[i])));
  endfor; endfor;
enddef;
```

Výsledky naší práce následují. První obrázek je zadání sudoku (levý obrázek) a druhý je jeho jednoznačné řešení (pravý obrázek), vše se zapnutým režimem pro čínštinu.

五	三			七				
六			一	九	五			
	九	八					六	
八				六				三
四			八		三			一
七				二				六
	六					二	八	
			四	一	九			五
				八			七	九

五	三	四	六	七	八	九	一	二
六	七	二	一	九	五	三	四	八
一	九	八	三	四	二	五	六	七
八	五	九	七	六	一	四	二	三
四	二	六	八	五	三	七	九	一
七	一	三	九	二	四	八	五	六
九	六	一	五	三	七	二	八	四
二	八	七	四	一	九	六	三	五
三	四	五	二	八	六	一	七	九

3. Závěrem

Příklad nám ukazuje, jak je v současnosti snadné včlenit čínštinu do META-POSTu. Zbývá nám již jediné, a to napojit na představená makra dostatečně obecného řešitele sudoku. To nechtě zůstane čtenářům jako otevřený a v článku neřešený problém.

Seznam literatury

- [1] Wikipedia, the free encyclopedia. [Wikipedie, otevřená encyklopedie.]
[online cit. 15. 1. 2010] Typy sudoku, o algoritmech a matematickém pozadí.
Dostupné z webových stránek na URL:
V angličtině: <http://en.wikipedia.org/wiki/Sudoku>
V angličtině: http://en.wikipedia.org/wiki/Algorithmics_of_sudoku
V angličtině: http://en.wikipedia.org/wiki/Mathematics_of_Sudoku
V češtině: <http://cs.wikipedia.org/wiki/Sudoku>

Summary:

Kanji-Sudokus: Integrating Chinese and Graphics

The article illustrates the integration of Chinese and METAPOST on a small example. The author draws a Sudoku grid, not with Hindu-Arabic numerals, but with Chinese numerals. This article demonstrates how straightforward the integration of Chinese and METAPOST has become.

The source codes are included as small parts in the article commented in detail. You may find the original English version of the article in *TUGboat*, see <http://tug.org/TUGboat/Articles/tb29-2/tb92roegel.pdf>.

Keywords: METAPOST, the game of sudoku, CJK package, latexmp package, UTF-8, Chinese glyphs, kanji.

*Denis Roegel, roegel@loria.fr
<http://www.loria.fr/~roegel>
LORIA – Campus Scientifique, BP 239
F-54506 Vandœuvre-lès-Nancy Cedex, France*