

Informatika pro moderní fyziky (7) Regulární výrazy. Tvorba textových dokumentů.

František HAVLŮJ

e-mail: haf@ujv.cz

ÚJV Řež

oddělení Reaktorové fyziky a podpory palivového cyklu

akademický rok 2016/2017

16. listopadu 2016

- 1 Co jsme se naučili minule
- 2 Najdi to, nevím co: regex
- 3 Výroba dokumentu v praxi
- 4 Na šablony chytře

Obsah

- 1 Co jsme se naučili minule
- 2 Najdi to, nevím co: regex
- 3 Výroba dokumentu v praxi
- 4 Na šablony chytře

- technika šablonování a její použití pro tvorbu vstupních souborů pro výpočetní programy
- generování většího množství grafů s využitím šablon (nebo bez nich?)
- vygenerování PDF pomocí LaTeXu

Obsah

- 1 Co jsme se naučili minule
- 2 Najdi to, nevím co: regex**
- 3 Výroba dokumentu v praxi
- 4 Na šablony chytře

Úkoly

- najít všechna celá čísla
- najít všechny emailové adresy
- vlnka: nedělitelné mezery v textu před k, s, v, z
- nahradit desetinné tečky čárkami

Jak na to: regulární výrazy (regexy)

- můžu definovat chytrou 'masku'
- nejjednodušší: `/text/` - text
- skupiny písmen: `/1[abc]/` - 1a, 1b ...
- speciální skupiny: `/da/` - 1a, 2a, 3a ...
- opakování:
- hračka a vysvětlení: rubular.com

Regexy v Ruby

- `String#match` vrací `MatchData` objekt, `m[0]` je ten samotný řetězec, `m[i]` jsou skupiny
- `String#gsub` nahrazuje
- `String#scan` vrací pole všech výskytů

Obsah

- 1 Co jsme se naučili minule
- 2 Najdi to, nevím co: regex
- 3 Výroba dokumentu v praxi**
- 4 Na šablony chytře

Úkol na dnešek

- pro jeden blok JE mám provozní data - v určitých dnech hodnotu koncentrace kyseliny borité a axiálního offsetu - pro několik kampaní (blíže neurčený počet)
- chci vyrobit přehledové PDF, které bude hezky prezentovat grafy obou veličin pro každou kampaň a k tomu i tabulky
- data pro jednotlivé kampaně mám v CSV souborech, každý má tři sloupce (datum, cB, AO)

Rozbor

- načíst tabulky a vykreslit grafy umíme (a taky už jsme to udělali minule)
- převést tabulky v CSV na tabulky v LaTeXu se záhy naučíme
- vložit obrázek do latexu taky umíme
- předem neznámý počet souborů nás netrápí
(`Dir["*.csv"]`)

Jak na tabulky

- tabulky budou dost rozsáhlé a montovat je přímo nějak do latexových vstupů je asi spíš nepraktické, naštěstí to jde i jinak
- naštěstí má LaTeX příkaz `\input`, kterým můžeme prostě vložit do dokumentu nějaký externí soubor
- takže si nejdřív přichystáme soubory s tabulkami a pak se na ně budeme už jenom odkazovat

Jak na tabulky v LaTeXu (1)

Základem tabulky je prostředí `tabular` s definicí počtu a zarovnání sloupců:

```
\begin{tabular}{lrr}  
...  
\end{tabular}
```

Jak na tabulky v LaTeXu (2)

Uvnitř tabulky se sloupce oddělují ampersandem a řádky dvojítm backslashem:

```
\begin{tabular}{lrr}  
  Data 1 & a & 1.0 \\  
  Data 2 & b & 2.0 \\  
  Data 3 & c & 3.0 \\  
\end{tabular}
```

Jak na tabulky v LaTeXu (3)

Přidání mřížky je nesnadné, leč proveditelné a vlastně docela dobře vymyšlené - přidáváme jednotlivé čáry po sloupcích a řádcích:

```
\begin{tabular}{|l|r|r|}  
  \hline  
  Data 1 & a & 1.0 \\  
  \hline  
  Data 2 & b & 2.0 \\  
  \hline  
  Data 3 & c & 3.0 \\  
  \hline  
\end{tabular}
```

Úkol na teď: výroba tabulek

- vyrobit z CSV souboru (tři sloupce) dvě LaTeX tabulky (po dvou sloupcích)
- postarat se, aby byly hezké
- chytré je vyrobit tabulku třeba o šesti sloupcích (jakože tři dvousloupce), pak už se to na stránku v klidu vejde

Obsah

- 1 Co jsme se naučili minule
- 2 Najdi to, nevím co: regex
- 3 Výroba dokumentu v praxi
- 4 Na šablony chytře**

Úskalí šablon

- snadno umíme nahradit jeden řetězec druhým
- trochu méně pohodlné pro větší bloky textu
- navíc by se hodila nějaká logika (cyklus) přímo v šabloně
- naštěstí jsou na to postupy

ERb (Embedded Ruby)

- lepší šablona - “aktivní text”
- používá se například ve webových aplikacích
- hodí se ale i na generování latexových dokumentů, resp. všude, kde nám nesejde na whitespace
- poměrně jednoduchá syntax, zvládne skoro všechno (viz předmět MAA3)

Základní syntaxe ERb (1)

Jakýkoli Ruby příkaz, přiřazení, výpočet ...

```
<% a = b + 5 %>  
<% list = ary * ", " %>
```

Základní syntaxe ERb (2)

Pokud chci něco vložit, stačí přidat rovnítko

```
<%= a %>
```

```
<%= ary[1] %>
```

```
<%= b + 5 %>
```

Základní syntaxe ERb (3)

Radost je možnost použít bloky a tedy i iterátory apod.
v propojení s vkládaným textem:

```
<% (1..5).each do |i| %>  
Number <%= i %>  
<% end %>  
<% ary.each do |x| %>  
Array contains <%= x %>  
<% end %>
```

ERb – shrnutí

- dobrý sluha, ale špatný pán
- můžu s tím vyrobit hromadu užitečných věcí na malém prostoru
- daň je velké riziko zamotaného kódu a nízké přehlednosti (struktura naprosto není patrná na první pohled, proto je namístě ji držet maximálně jednoduchou)

Důležité upozornění

- oddělení modelu a view
- přestože lze provádět zpracování dat a výpočty přímo v ERb, je to nejvíc nejhorší nápad
- je chytré si všechno připravit v modelu (tj. v Ruby skriptu, kterým data chystáme)
- a kód ve view (tj. v ERb šabloně) omezit na naprosté minimum

Jak ze šablony udělat výsledek

Příklad překladač ERb

```
require 'erb_compiler'  
  
erb(template, filename, {:x => 1, :y => 2})
```

Příklad – kreslení grafů z minula

template.gp

```
set terminal png
set output "plot_<%=n%>.png"
plot "data_<%=n%>.csv"
```

```
(1..10).each do |i|
  erb("template.gp", "plot_#{i}.gp", {:n => i})
end
```

Takže v latexu třeba

```
\subsection{Koncentrace kyseliny borité}  
  
<% files.each do |f| %>  
  
\subsubsection{Kampaň <%= f.split('_').last %>}  
\begin{center}  
\includegraphics[width=0.8\textwidth]{<%= f %>_bc.eps}  
\end{center}  
<% end %>
```

A teď už to jenom dejte dohromady...

- 1 připravit si základní kostru dokumentu v latexu
- 2 převést na šablonu: mít seznam souborů, správně generovat kapitoly
- 3 vyrobit grafy
- 4 vložit grafy do šablony
- 5 vyrobit tabulky
- 6 vložit tabulky do šablony
- 7 A JE TO!

Co jsme se naučili minule
Najdi to, nevím co: regex
Výroba dokumentu v praxi
Na šablony chytře

A to je vše, přátelé!

