

Funkcionální projekt **SIMPLIFY-BKG**

(Funkcionální a logické programování **2016/2017**)

Marek Kidoň
ikidon@fit.vutbr.cz

29. ledna 2017

1 Úvod

Tento dokument popisuje požadavky na projekt **SIMPLIFY-BKG** v předmětu FLP v akademickém roce **2016/2017**. Obecné pokyny pro všechna zadání jsou sepsány ve zvláštním dokumentu, který si nezapomeňte důkladně pročíst.

2 Zadání

Vytvořte program provádějící odstranění zbytečných symbolů z bezkontextových gramatik (dále BKG). Postupujte podle algoritmu 4.3 z opory předmětu TIN, který pracuje ve dvou krocích.

2.1 Rozhraní programu

Program se bude jmenovat **simplify-bkg** a bude jej možné spustit takto:

```
simplify-bkg [volby] [vstup]
```

kde

- **vstup** je jméno vstupního souboru (pokud není specifikováno, program čte standardní vstup) obsahujícího BKG ve formátu popsaném v sekci 3.
- **volby** jsou parametry ovlivňující chování programu, standardně tyto:
 - i dojde pouze k vypsání načtené a do Vaší vnitřní reprezentace převedené BKG na stdout ve formátu popsaném v sekci 4. Nevypisujte jen načtený řetězec, tato volba ověřuje, že Váš program dokáže gramatiku naparsovat a znovu vypsát.
 - 1 dojde k vypsání BKG \overline{G} (po prvním kroku algoritmu 4.3) na stdout ve formátu popsaném v sekci 4.
 - 2 dojde k vypsání BKG G' (po druhém kroku algoritmu 4.3) na stdout ve formátu popsaném v sekci 4.

3 Formát vstupu

BKG $G = (N, \Sigma, P, S)$ na vstupu odpovídá standardní definici BKG (viz opora předmětu TIN). Pro zjednodušení je ale abeceda N podmnožinou množiny velkých písmen $[A-Z]$, a abeceda Σ je podmnožinou množiny malých písmen $[a-z]$. Vstupní textová reprezentace BKG o N pravidlech má formát:

```
<seznam všech nonterminálů>\n
<seznam všech terminálů>\n
<počáteční nonterminál>\n
<pravidlo 1>\n
...
<pravidlo N>\n
```

Seznam o M symbolech má tvar `symbol_1,symbol_2,...,symbol_M`. Pravidlo $A \rightarrow \alpha$ má tvar `A->alpha`. Epsilon je reprezentováno znakem `#`. Epsilon pravidlo má tedy tvar `A->#`.

Například, reprezentace BKG z příkladu 4.1 opory předmětu TIN, s přidaným pravidlem $S \rightarrow \epsilon$, je následující:

```
S,A,B
a,b,c,d
S
S->#
S->AB
A->aAb
A->ab
B->cBd
B->cd
```

4 Formát výstupu

Textová reprezentace BKG na výstupu má stejný formát jako BKG na vstupu ze sekce 3. V žádném případě ale neměňte označení stavů nebo dokonce sémantiku pravidel. Jen odstraňte zbytečné symboly a pravidla. Ponechané symboly a pravidla musí zůstat úplně stejná, jako byla zadána.