

Übungen zur Vorlesung: Struktur und Implementierung von Programmiersprachen I

Blatt 3 (Syntaxanalyse (2), Scannergenerierung)

Hinweis: in beiden Aufgaben soll folgender Datentyp verwendet werden:

```
type dottedItem =  
  Dot (* Cursor *)  
  | Basic of char list (* Basismuster *)  
  | Conc of dottedItem * dottedItem (* Konkatenation *)  
  | Alt of dottedItem * dottedItem (* Alternative *)  
  | Star of dottedItem (* Sternoperation *)
```

Aufgabe 4 (Parsen regulärer Ausdrücke)

Gegeben ist eine Grammatik für reguläre Ausdrücke.

- Terminalzeichen (Token): $\{\text{BAS}, *, |, (,)\}$, wobei unter dem Token **BAS** alle Basismuster zusammengefasst werden.
- Nichtterminale: $\{A, A', C, C', S, S', B\}$; A steht für Alternativen, C für Konkatenationen, S für Sternoperationen und B für Basisfälle oder geklammerte Ausdrücke.
- Startsymbol: A
- Produktionen, linksfaktoriert und ohne Linksrekursion:

$$\begin{aligned} A &\rightarrow C A' \\ A' &\rightarrow | A \quad | \quad \varepsilon \\ C &\rightarrow S C' \\ C' &\rightarrow C \quad | \quad \varepsilon \\ S &\rightarrow B S' \\ S' &\rightarrow * \quad | \quad \varepsilon \\ B &\rightarrow \text{BAS} \quad | \quad (A) \end{aligned}$$

- (a) Berechnen Sie manuell die FIRST- und FOLLOW-Mengen für die Nichtterminale.

- (b) Schreiben Sie einen Recursive-Descend-Parser für Wörter der Sprache, die durch die Grammatik erzeugt wird. Die Eingabe soll ein String und die Ausgabe ein Element des Datentyps `dottedItem` (s.o.) sein. Der Konstruktor `Dot` ist für diese Aufgabe noch nicht relevant.
- (c) Stellen Sie zu der gegebenen Grammatik die Parsetabelle für einen Stackautomaten auf.

Aufgabe 5 (Funktionen auf Dotted Items)

Dotted Items sind Elemente des Typs `dottedItem` (s.o.), mit der Zusatzbedingung, dass jedes Item genau einen Dot enthält.

- (a) Schreiben Sie eine Funktion `isBasicItem`, die zu einem übergebenen Dotted Item den Wert `true` liefert, gdw. sich der Dot vor einem Basismuster befindet.
- (b) Schreiben Sie eine Funktion `epsmove`, die zu einem Dotted Item die Liste der Dotted Items liefert, die sich von diesem aus durch einen einzigen ε -Schritt erreichen lassen. Beachten Sie bitte, dass die Konkatenation assoziativ ist.