

## Übungsblatt 2: Übersetzer und sprachverarbeitende Werkzeuge (SS 2011)

Hand Out: 27. April 2011

Hand In: 4. Mai 2011

### Aufgabe 1 Reguläre Ausdrücke und deren Automaten

- Beschreiben Sie die in Java erlaubten Fließkommazahlen mit Hilfe eines regulären Ausdrucks. Suffixe, wie F und D, sowie binäre und hexadezimale Schreibweisen brauchen Sie nicht zu berücksichtigen. Verwenden Sie dazu die Java-Sprachspezifikation: <http://java.sun.com/docs/books/jls/>.
- Übersetzen Sie den regulären Ausdruck aus a) zu einem NFA.
- Übersetzen Sie den NFA zu einem DFA.

### Aufgabe 2 Scanner-Generierung mit JFlex

Hier werden wir die erste Vorversion für den Scanner des Compilers, den wir in diesem Semester erstellen werden, entwickeln. Dazu wird das in der Vorlesung erwähnte JFlex, <http://jflex.de>, verwendet.

- Erstellen Sie einen Scanner mit JFlex und eine Tokenrepräsentation, die beide für eine Java-Teilsprache geeignet sind. Folgende Schlüsselwörter sollen erkannt werden: **class**, **interface**, **extends**, **implements**, **void**, **int**, **boolean**, **return**, **if** und **while**. Folgende Sonderzeichen sollen Token sein: (, ), ,, {, }, ., +, -, \*, /, ;, == und >=. Außerdem sollen folgende Literale (Konstanten) als Token erkannt werden: ganzzahlige Zahlen mit Vorzeichen, **true** und **false**. Als Bezeichner sollen nur Wörter zugelassen sein, die mit einem Buchstaben oder Unterstrich beginnen und nur Buchstaben, Ziffern oder Unterstriche enthalten. Whitespace bestehend aus Leerzeichen, Tabulatoren und Newline werden nur überlesen.

Schreiben Sie ein kleines Hauptprogramm, welches eine Datei in der Kommandozeile nimmt und die Sequenz der erkannten Token ihres Scanners ausgibt.

- Erweitern Sie ihre Lösung aus a) um einfache Kommentare wie in Java, d.h. mit // und /\*. . \*/. Sie können annehmen, dass Kommentare nicht verschachtelt werden.
- Erweitern Sie ihre Lösung aus b) um die Möglichkeit die Position eines Tokens in der Eingabedatei abzufragen. Dazu müssen Sie ihre Tokenrepräsentation geeignet erweitern und beim Erzeugen eines Tokens die Zeilen- und Spaltennummer des Tokenanfangs festhalten. Sehen hierzu im Handbuch von JFlex nach wie sie an diese Information gelangen können.

Die Aufteilung in drei Schritte soll eine Hilfestellung sein, Sie können auch nur eine Gesamtlösung abgeben.