

Wintersemester 2024/25

Übungen zur Vorlesung Automaten und Formale Sprachen

Blatt 2

Termin: 14. Oktober 2024

Reguläre Ausdrücke

Aufgabe 1

Gegen Sie jeweils die erzeugten Sprachen der regulären Ausdrücke an.

Beispiel: $L((a \mid b) \cdot c^+) = \{ac, bc, acc, bcc, accc, bccc, \dots\}$

1. $r \mid r^*$
2. $(r \cdot t)^*$
3. $r^* \cdot t^*$
4. $(t \cdot r)^*$
5. $(r \mid t)^*$
6. $r^* \mid t^*$
7. $(0^* \cdot 1)^* \cdot 0^*$
8. $(1 \mid 0)^*$

Geben Sie ein Gegenbeispiel an, falls die Äquivalenz nicht gelten sollte. Ansonsten zeigen Sie, dass die erzeugten Sprachen der beiden regulären Ausdrücke gleich sind.

Hinweis zu den folgenden Aufgaben

Verwenden Sie zur Lösung der Aufgaben 2 – 4 die von Java bereit gestellte Klasse `Pattern` aus dem Package `java.util.regex`. Die Klasse `Pattern` stellt u.a. die `matches(r,s)` Methode zur Verfügung, mit der überprüft werden kann, ob eine Zeichenfolge `s` in der von einem regulären Ausdruck `r` erzeugten Sprache enthalten ist, d.h. ob $s \in L(r)$ gilt.

Zur Definition eines regulären Ausdrucks wird u.a. folgende Notation angeboten:

1. das Zeichen selbst, beispielsweise `x` für das Zeichen ‘`x`’
2. eine der Ziffern `0, 1, 2, \dots, 9`: `\d`
3. ein Zeichen aus `a – z` sowie `A – Z`: `\w`
4. genau ein beliebiges Zeichen: `.`
5. Gruppierung: `(...)`
6. `X \mid Y`: Alternative: entweder `X` oder `Y`
7. `[abcde]`: Auswahl, entspricht `(a \mid b \mid c \mid d \mid e)`

8. X^* : beliebig viele X , auch kein Auftreten von X
9. X^+ : mindestens ein Auftreten von X
10. $X^?$: optionales Auftreten von X

Um die besondere Bedeutung von Zeichen (etwa '+' oder '.') aufzuheben, werden zwei Backslashes benötigt wie z.B. in " $(\\+|-)?4\\.2$ ".

Folgendes Listing zeigt Beispiele für reguläre Ausdrücke:

```
String file_pattern = "Blatt*.pdf";
System.out.println(Pattern.matches(file_pattern, "Blatt1.pdf")); // true
System.out.println(Pattern.matches(file_pattern, "Blatt11.pdf")); // true
System.out.println(Pattern.matches(file_pattern, "Blatt.pdf")); // true
System.out.println(Pattern.matches(file_pattern, "blatt1.pdf")); // false

String int_pattern = "0|[123456789][0123456789]*";
System.out.println(Pattern.matches(int_pattern, "0")); // true
System.out.println(Pattern.matches(int_pattern, "007")); // false
System.out.println(Pattern.matches(int_pattern, "5")); // true
System.out.println(Pattern.matches(int_pattern, "4009")); // true
System.out.println(Pattern.matches(int_pattern, "")); // false
```

Aufgabe 2

Erstellen Sie einen regulären Ausdruck für die Darstellung von Gleitkommazahlen. Eine Gleitkommazahl besteht aus einer ganzen Zahl mit einem optionalen Vorzeichen, die einen Dezimal- und einen Exponentialanteil haben kann (beide oder einer von beiden kann fehlen). Der Dezimalanteil beginnt mit einem Punkt, dem eine beliebige nicht leere Ziffernfolge folgen muss. Der Exponentialanteil beginnt mit dem Symbol E oder e , dem eine ganze Zahl folgen muss.

Aufgabe 3

Betrachten Sie folgendes Format für die Angabe von Zeitpunkten:

yyyy-MM-ddTHH:mm:ss.SSS

Konkretes Beispiel: 2019-04-01T14:00:30.469. Erstellen Sie einen regulären Ausdruck, der die Darstellung solcher Angaben möglichst genau wiedergibt.

Aufgabe 4

Geben Sie einen regulären Ausdruck an, der den strukturellen Aufbau deutscher KFZ-Kennzeichen formal spezifiziert: ein-, zwei- oder dreibuchstabige Kürzel, gefolgt von einer ein- oder zweielementigen Buchstabenkombination, gefolgt von einer höchstens vierstelligen Zahl ungleich Null und ohne führende Nullen. Zwischen den drei genannten Teilen befindet sich jeweils ein Blank (leerer Zwischenraum). Optional kann entweder das Zeichen H oder E folgen.

Hinweise zur Abgabe:

- Der Aufgabeordner ist Blatt 2.
- In einem pdf-Dokument mit dem Namen Blatt2_<Nachname1>_<Nachname2>.pdf hinterlegen Sie Ihre Lösung von Aufgabe 1.

- Legen Sie ein Java-Programm mit den von Ihnen erstellten regulären Ausdrücken für die Aufgaben 2 – 4 ebenfalls in dem Ordner **Blatt 2** ab. Testen Sie Ihre Lösungen insbesondere mit den vorgegebenen Testfällen (siehe `Blatt_2.RegexTest.java`). Erstellen Sie jeweils einen Screen-Shot mit dem Testprotokoll und nehmen Sie diese in das pdf-Dokument auf.
- Legen Sie die Dokumente bis spätestens **Freitag, 18. September**, in dem Abgabep Ordner **Blatt 2** ab.