

2. Übungsblatt – Abgabe: 07.11.2021, 23:55 Uhr

Aufgabe 2.1 Endliche Automaten und Morphologie

- (a) Das Adjektiv-Endungs-Diagramm auf der Vorlesungsfolie akzeptiert das Wort *st*. Die Superlativendung *kleinst* kommt aber für sich allein gar nicht vor. Korrigieren Sie den NEA mit möglichst wenig Aufwand so, dass *st* für sich genommen nicht mehr akzeptiert wird (alle übrigen Endungen dagegen wie bisher).
- (b) Entwerfen Sie einen möglichst einfachen NEA, der die einfachen flektierten Formen des Verbs *reden* spezifiziert (Präsens und Präteritum, ohne Partizipialendungen!).

Aufgabe 2.2 Endliche Automaten und Syntax

- (a) In der Vorlesung haben wir einen NEA betrachtet, der einige zulässige Wortartketten für Nominalausdrücke spezifiziert. Der NEA spezifiziert Nominalausdrücke mit Artikel, Gattungssubstantiv, und einem potentiellen pränominalen adjektivischen Attribut, möglicherweise mit Gradpartikeln (GPRT). Erweitern bzw. modifizieren Sie den Automaten so, dass es auch Nominalausdrücken mit post-nominalen Präpositionalausdrücken umfasst. Also z.B.:

ART NN (*das Auto*)

ART ADJA ADJA NN (*das neue schnelle Auto*)

ART ADJA NN APPR ART NN (*das grüne Auto auf dem Parkplatz*)

ART NN APPR ART NN APPR ART ADJA NN (*das Auto auf dem Parkplatz bei dem neuen Institutsgebäude*)

Aufgabe 2.3 Formale Schreibweise von Automaten

Gegeben sei der Automat $A = \langle K, \Sigma, \Delta, s, F \rangle$ mit

$$K = \{1, 2, 3, 4\}$$

$$\Sigma = \{l, a, !\}$$

$$s = 1$$

$$F = \{1, 4\}$$

$$\Delta = \{\langle 1, l, 2 \rangle, \langle 2, a, 3 \rangle, \langle 3, l, 2 \rangle, \langle 3, !, 4 \rangle\}$$

- (a) Malen Sie das Zustandsdiagramm für den Automaten
- (b) Beschreiben Sie informell in Worten die Sprache, die der Automat akzeptiert. Wie lang ist das kürzeste, wie lang das längste akzeptierte Wort?
- (c) Geben Sie drei verschiedene Möglichkeiten an, A so zu verändern, dass auch Wörter erkannt werden, die kein Ausrufezeichen am Ende haben, (aber ansonsten genauso wie in a) angegeben aufgebaut sind). Zulässige Veränderungen dabei sind: Das Einfügen einer neuen Transition in Δ , sowie das Hinzufügen eines Zustandes zu F .

(d) Ändern sie A so, dass Wörter mit beliebig vielen a's (aber mehr als null) statt einem a hinter dem l (z.B. lalaa!, laalaaa!) erkannt werden. Ändern Sie dabei nur Δ und lassen Sie den Automaten ansonsten unverändert.

(e) In dieser Aufgabe geht es um reguläre Ausdrücke (Englisch: *regular expression*, oder einfach *Regex*). Wenn Sie das Konzept der regulären Ausdrücke nicht kennen, finden Sie zahlreiche Einführungen dazu in Text- und Videoform - einige Beispiele sind unter der Aufgabenbeschreibung angegeben.

In dieser Aufgabe sollen Sie aber erst einmal spielerisch mit regulären Ausdrücken umgehen, indem Sie 'Regex Crosswords' lösen:

<https://regexcrossword.com/>

Auf der verlinkten Webseite finden Sie verschiedene Kreuzworträtsel, bei denen aus sich überschneidenden regulären Ausdrücken die korrekten Strings abgelesen werden müssen. Unter 'Tutorial' finden sich die ersten Kreuzworträtsel, mittels derer man an einzelne Symbole herangeführt wird, die in regulären Ausdrücken benutzt werden. Nehmen Sie sich eine halbe Stunde Zeit, so viele Rätsel wie möglich zu lösen. Bei Interesse können Sie sich natürlich noch mehr Zeit nehmen - es macht nämlich sogar Spaß! Schreiben Sie in Ihrer Lösung für diesen Teil einfach, wie weit Sie gekommen sind.

Geben Sie außerdem nun einen regulären Ausdruck an, der alle nichtleeren Strings beschreibt, die der ursprüngliche Automat in Aufgabe 2.3 a) erkennt.

Hinweise: In c) und d) sollen Sie keine Zustandsdiagramme zeichnen, sondern die neuen Automaten formal angeben.

Einen 'Quick Start' und ein Tutorial zu regulären Ausdrücken finden Sie zum Beispiel auf regular-expressions.info:

<https://www.regular-expressions.info/tutorial.html>

<https://www.regular-expressions.info/quickstart.html>

Es lassen sich auch zahlreiche 'Cheat Sheets' finden, zum Beispiel hier:

<https://cheatography.com/davechild/cheat-sheets/regular-expressions/>

Die wichtigsten Symbole sind zusätzlich hier noch einmal aufgeführt:

.	irgendein beliebiges Zeichen
.*	kein oder beliebig viele Zeichen
[a-e]	a, b, c, d, e
[^a-e]	alle Zeichen außer a, b, c, d, e
(maus hund)	Zeichenfolge maus oder hund
(ab)*	kein oder beliebig viele ab
(ab)+	mindestens ein oder beliebig viele ab
(ab)?	kein oder ein ab
^	Anfang einer Zeile
\$	Ende einer Zeile

Abgabe via Moodle. Bei Fragen posten Sie im MS Teams Fragenchannel.