GTI Übungsblatt 4
Tutor: Marko Schmellenkamp
ID: MS1
Übung: Mi 16-18

Max Springenberg, 177792

4.1 Pumping Lemma

4.1.1 $L_1 = \{cucv | u, v \in \{a, b\}^*, |u| = |v|\}$

Insbesondere gilt für alle Wörter aus L_2, dass sie genau 2 c enthalten.

Es gelte:

$w = xyz$, mit $w \in L_1$ beliebig
$|w| = 2n$, mit $|cu| = |cv| = n$
$(i) y \neq \epsilon$
$(ii) |xv| \leq n$

Aus $(i), (ii)$ folgt, dass y mindestens ein Zeichen von cu enthält.

wähle $k = 0$:

1.Fall y enthält nur das erste c:
$xy^0z = xz \overset{\#_c(wz) \neq 2}{\notin} L_1$

2. Fall mindestens ein Zeichen von u in y:
$xy^0z = xz \overset{|u|<|v|}{\notin} L_1$

Da dies für eine beliebiges n gilt, ist L_1 nicht regulär.

4.1.2 $L_2 = \{ua^i | u \in \{a, b, c\}^*, i \in \mathbb{N}_0, \#_a(u) = i\}$

Es gelte:

$w = xyz$, mit $w = ua^n, u = a^n$
$|w| = 2n$
$(i) y \neq \epsilon$
$(ii) |xv| \leq n$

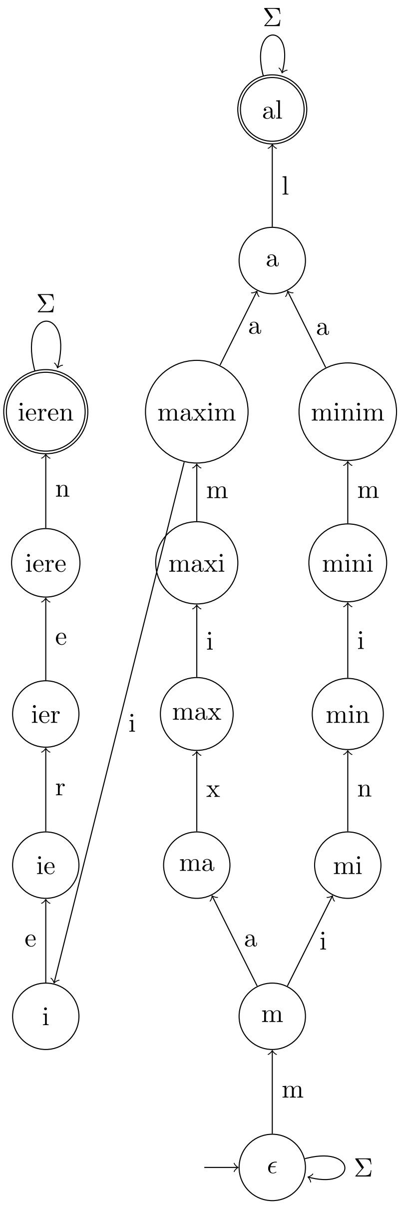
Aus $(i), (ii)$ folgt für y, dass y nur Zeichen von u enthält.

wähle $k = 0$:
$xy^0z = xz$
Da nun xz eine ungerade Anzahl an a enthält, gibt es auch keine Aufteilung von $xz = u'a^i$, sodass gelten würde $\#_a(u') = a^i$.
Daraus folgt: $wz \notin L_2$

Da dies für beliebige n gilt, ist L_2 nicht regulär.

4.2

4.2.1



Bei einem NFA für Zeichenkettensuche ist wichtig, dass der Automat genau dann akzeptiert wenn eines der Zu suchenden Wörter als Teilwort enthalten ist.
Ferner soll für alle anderen Wörter über dem gegebenem Alphabet nicht akzeptiert wird.

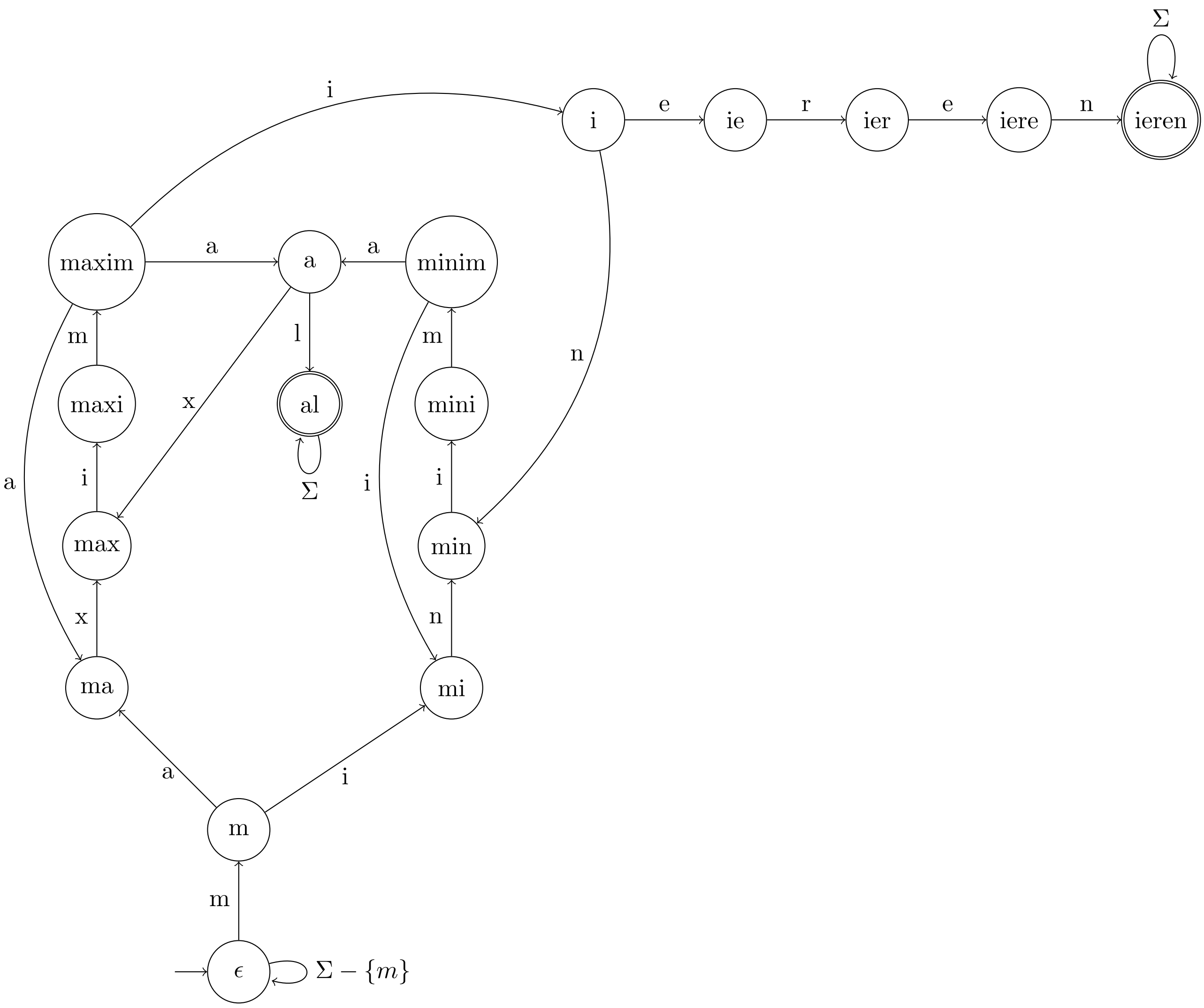
Das Verwerfen für Wörter über dem Alphabet, die nicht eines der gesuchten Teilwörter enthalten ist durch die Transition $\epsilon \overset{\sigma\in\Sigma}{\to} \epsilon$ umgesetzt, da der Zustand ϵ nicht akzeptierend ist und für alle

Wörter die keinen akzeptierenden Lauf haben die Transition für alle Zeichen angewandt werden kann.

Das Akzeptieren erfolgt über das Sequenzzielle durchlaufen aller Zeichen für jedes gesuchte Wort.
Dabei werden insbesondere die Präfixe $maxim, minim$ durch die Läufe $\delta^*(\epsilon, minim) = minim, \delta^*(\epsilon, maxim) = maxim$ realisierten, von denen aus die gesuchten Wörter $maximal, maximieren, minimal$ über die Läufe $\delta^*(minim, al) = al, \delta^*(maxim, al) = al, \delta^*(maxim, ieren) = ieren$, mit den akzeptierenden Zuständen $al, ieren$, erreicht werden können.
Da das einmalige einlesen eines Teilwortes, das einem gesuchten Wort entspricht bereits reicht, wird für alle weiteren eingelesenen Zeichen in den akzeptierenden Zuständen verweilt.

4.2.2

Konvention:
Alle fehlenden Transitionen führen in den Startzustand ϵ und werden für die Übersichtlichkeit nicht gezeichnet.



Der Aufbau des DFA's erfolgt wie auch bei 4.2.1. Jedoch wurden noch transitionen , die für den Determinismus notwendendig sind ergänzt.
Zunächst wurden die Transitionen der akzeptierenden Zustände, für alle Zeichen, auf sich selbst beibehalten, da weiterhin nur ein gesuchtes Wort als Teilwort enthalten sein muss.
Die Transition des Startzustandes auf sich selbst wird um die Zeichen, deren einlesen auch noch in einen anderen Zustand führt veringert.
Damit sind Start und Akzeptierenden Zustände bereits deterministisch.

Für die verbleibenden Zustände müssen noch Transitionen ergänzt werden, die nach Lesen eines Zeichens zu einem Präfix führen, der durch einen Zustand im Automaten repräsentiert wird.
Wurde zuvor ein m eingelesen, so kann wieder mit den Präfixen ma und mi angefangen werden. Daher die Transitionen $maxim \overset{a}{\rightarrow} ma, minim \overset{a}{\rightarrow} ma, minim \overset{i}{\rightarrow} mi$.
Für den Zustand $maxim$ wird keine Transition in mi angefügt, da mi nicht akzeptierend ist und ein lauf zu $ieren$ noch möglich ist. Dafür existiert die Transition $i \overset{n}{\rightarrow} min$, die nachträglich zum Präfix

min wechselt, da wir uns an der Stelle sicher sein können, dass kein anderer Präfix mehr möglich ist. Zudem gibt es noch die Möglichkeit, dass nach dem einlesen von ma wieder ein Präfix max entsteht. Daher die Transition $a \xrightarrow{x} max$.
Alle Verbliebenen Transitionen je Zustand, führen nach Konvention in den Startzustand.

Damit existieren für jeden Zustand Transitionen für alle Zeichen aus dem Alphabet.

4.3

4.3.1

Eine informelle Beschreibung der Sprache $L(G)$ kann wie folgt geführt werden.

Die Sprache $L(G)$ umfasst nur Wörter deren Zeichen von einem a und einem a oder einem b und einem b oder einem c und einem a oder einem c und einem b eingeklammert werden, Bestandteil der Klammerung oder c's, die inmitten aller Klammerungen stehen sind.

Diese informelle Beschreibung umfasst die Regeln:

D:
Durch die mittigen c's und dem möglichen Fall das alle Zeichen Teil einer Klammerung sind und kein Zeichen innerhalb der innersten Klammerung steht.

B:
Durch die Klammerung von a und a, sowie b und b und der Bechreibung zu D

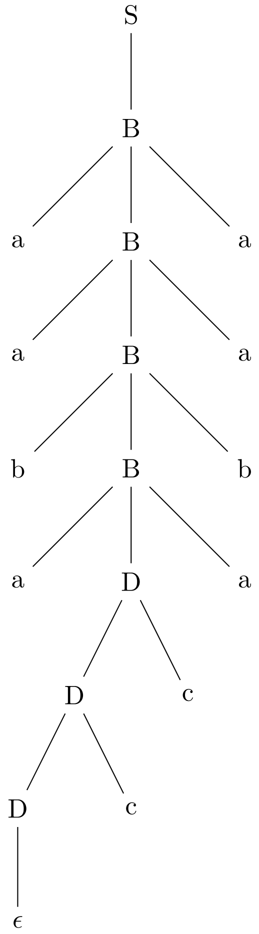
C:
Durch die Klammerung von c und a, sowie c und b und der Bechreibung zu D

S:
Durch die Beschreibung von B und C.

Damit ist die Grammatik G informell beschrieben worden.

4.3.2

Nach Aufgabenstellung soll ein Ableitungsbaum ohne Begründung angegeben werden. Ein möglicher Ableitungsbaum ist:



mit dem Blätterstring $aabaccabaa = w$

4.3.3

Annahme: G ist eindeutig.
Dann existiert nach der Eigenschaft von eindeutigen Grammatiken kein Wort $w \in L(G)$, mit zwei verschiedenen Ableitungsbäumen.

Wir betrachten das Wort $w = \epsilon$, mit den Ableitungsbäumen:

(i)

$$\begin{array}{c} \text{S} \\ | \\ \text{B} \\ | \\ \text{D} \\ | \\ \epsilon \end{array}$$

(ii)

$$\begin{array}{c} \text{S} \\ | \\ \text{C} \\ | \\ \text{D} \\ | \\ \epsilon \end{array}$$

↯ Da die Ableitungsbäume verschieden sind, ist die Eigenschaft für die Eindeutigkeit von Grammatiken verletzt.

Damit ist die Grammatik G mehrdeutig.