TI Übungsstunde 04

Marcel Schmid marcesch@student.ethz.ch

14.10.2020

---

1 Korrekturen

- Sehr gut gelöst, eine der wichtigsten Serien
- λ nicht als Eingabe vergessen
- Kl$[q]$ für alle Klassen, keine vergessen
- Kl[mühsamster Zustand] $= \Sigma^* \setminus \bigcup_{q \in Q, q \neq \text{mühsamster Zst}}$
- Nennt Zustände sinnvoll

$\Rightarrow$ am besten etwas wie q_i, nicht einfach eine Zahl

- ein einzlener Abfallstate/Sink verwenden, das erspart Schreibarbeit und ist eleganter
- $x^k = \underbrace{xxx\ldots x}_{k \text{ times}}$

$\Rightarrow$ ganz ähnlich wie bei Zahlen: $\{n^k \mid n \in \mathbb{N}\}$ sind auch nur Zahlen von der Form $n \cdot n \cdots n$

$\Rightarrow$ Verwechslung stammt von *Sprachen* wie bspw. $\{01, 11\}^k$.

- Bin(n) beginnt mit einer "1" (Ausnahme 0...), so wie unsere Darstellung von Zahlen nie mit 0 beginnt

  $\Rightarrow$ niemand schreibt 007 statt 7 im "richtigen" Leben
- K. Kompl. vs. Länge von Programmen vs. Binärdarstellung von Zahlen – nicht verwechseln!

2 Theorie/Repetition

2.1 Pumping Lemma / L3.4

Das Pumping Lemma ist eines der besten Tools, um die Nichtregularität einer Sprache zu zeigen. Die Beweise laufen immer nach dem gleichen Schema ab:

1. "Wir führen einen Widerspruchsbeweis und nehmen an, dass es einen EA A gibt, der ..."

$\Rightarrow$ deswegen gibt es ein $n_0 \in \mathbb{N}$ mit den Eigenschaften aus dem Lemma

2. Wir wählen ein Wort w in Abhängigkeit von n_0 (meistens so, dass w in der Sprache ist)
3. Zeige, dass $|w| \geq n_0$ ist
4. Zeige für *alle* möglichen Zerlegungen, dass (1), (2) und (3) nicht *gleichzeitig* erfüllt werden können.
5. Normalerweise so:
   (a) Finde *alle* Zerlegungen, welche (1) und (2) erfüllen
   (b) Zeige, dass für alle diese Zerlegungen (3) nicht gelten kann
   $\Rightarrow$ deswegen kann es sicherlich keine Zerlegung geben, die alle drei Bedingungen erfüllt
6. Widerspruch gefunden, deswegen Assumption dass es EA gibt falsch.

$\Rightarrow$ Tipp: Manchmal ist es einfach, wenn man ein "langes" Wort nimmt, z.B. $0^{2 \cdot n_0}$; die Länge muss nicht gleich sein wie n_0.

2.2 Kolmogorov-Methode / S3.1

- Recht "kompliziert", der Schlüssel ist es, die Präfixsprachen zu Verstehen:
$$L_x = \{y \in \Sigma^* \mid xy \in L\}$$
sprich *alle* "Ergänzungen"/Suffixe, so dass xy in L ist.
- In den Übungen/Anwendungen gibt das meistens unendlich viele solcher Sprachen mit je unendlich vielen Wörtern drin.
- Schauen wir uns beispielsweise die Sprache $L = \{0^n 1^n \mid n \in \mathbb{N}\}$ an. Definiere $L_0 = \{y \in \{0,1\}^* \mid 0y \in L\}$

⇒ Welche Wörter sind in L_0?
$$L_0 = \{1, 011, 00111, \dots, 0^i 1^{i+1}, \dots\}$$

- Sei $L_{00} = \{y \in \{0,1\}^* \mid 00y \in L\}$. Welche Sprachen sind da drin?
$$L_{00} = \{11, 0111, 001111, \dots, 0^i 1^{i+2}, \dots\}$$
- Wie sieht das dann mit L_{0^m} aus?

⇒ $L_{0^m} = \{1^m, 01^{m+1}, \dots, 0^i 1^{m+i}, \dots\}$

- Wir wissen wegen S3.1, dass für eine beliebige solcher Präfixsprachen L_x gilt, dass das j-te Wort x_j eine K.Kompl. $K(x_j) \leq \lceil \log(j+1) \rceil + c$ haben.
- Das heisst also bspw., dass für all unsere Sprachen $\{L_{0^i} \mid i \in \mathbb{N}\}$ gilt, dass das zweite Wort eine K-Kompl. $\leq \log(3) + c$ hat
- Wie sehen, für alle diese Sprahcen, die zweiten Wörter jeweils aus?

⇒ Alle sind von der Form $01^m + 1$ (wobei m die Länge des Präfixes ist)

- Damit haben wir aber unendlich viele Wörter gefunden, welche jeweils eine K. Komplexität kleiner als $\lceil \log 3 \rceil + c$ haben ⇒ Widerspruch.

3 Übungen

S19/11b) (Lemma 3.3.)

Sei $L = \{u\#v \mid u, v \in \{0,1\}^+ \text{ und Nummer}(u) = |v|\}$. Zeige anhand von L3.3, dass diese Sprache nicht regulär ist.

1. Wir gehen genau gleich wie letzte Woche vor: Angenommen, L sei regulär, dann gibt es einen EA $A = (Q, \Sigma, \delta, q_0, F)$ mit $L(A) = L$..
2. Wir wählen wieder viele Wörter, zum Beispiel die Wörter $10^i\#$ für $i \in \{1, \dots, |Q| + 1\}$.
3. Aus dem Pigeonhole Principle (da dies mehr Wörter sind als Zustände) folgt, dass es ein $i < j$ gibt so dass
$$\hat{\delta}(q_0, 10^i\#) = \hat{\delta}(q_0, 10^j\#)$$
4. Da L regulär ist (nach Annahme), gilt L3.3 und deswegen gilt auch für alle $z \in \Sigma$, dass
$$10^i\#z \in L \iff 10^j\#z \in L$$
5. Da das für alle möglichen $z \in \Sigma^*$ gilt, gilt das auch für $z = 1^{2^i}$. Das führ nämlich zu einem Widerspruch: Durch "#" ist der String $10^x\#z$ eindeutig "aufgeteilt" in zwei Teile $u, v \in \{0,1\}^+$. Wir haben aber:
   - $10^i\#z \in L$, da Nummer(10^i)$= 2^i$ ist und $|z| = |1^{2^i}| = 2^i$ gilt
   - $10^j\#z \notin L$, da Nummer(10^j)$= 2^j > 2^i$ und somit ungleich $|z|$
6. Somit haben wir einen Widerspruch zu unserer Annahme, dass L regulär sei.

Midterm 19, 2b) (Pumping Lemma)

Sei $L = \{w \in \{a,b\}^* \mid w = aub \text{ für } u \in \{a,b\}^* \text{ und } |w|_a \big| |w|_b\}$. Zeige mit L3.4, dass diese Sprache nicht regulär ist.

1. Widerspruchsbeweis: Wir nehmen an, L sei regulär. Dann gibt es einen EA $A = (Q, \Sigma, \delta, q_0, F$ mit $L(A) = L$. Weiter muss das L3.4 halten. Sei $n_0 \in \mathbb{N}$ wie aus dem Lemma.

2. Wir wählen uns daher ein "intelligentes", langes Wort, zum Beispiel

$$w = aub, \; u = a^{n_0} b^{n_0}$$

3. Aus dem Lemma wissen wir, dass es eine Zerteilung von w in 3 Teile gibt so dass $w = yxz$ gilt. Weiter gelten alle 3 Bedingungen an die Verteilung aus dem Lemma.

4. Wir schränken jetzt so quasi die Menge an mölichen Verteilungen ein, indem wir Schritt für Schritt nur solche Aufteilungen weiter betrachten, die auch die entsprechenden Bedingungen im Lemma erfüllen:

   (1) Aus (1) wissen wir, dass $|yx| \leq n_0$. In unserem Fall muss also $y = a^l$ und $x = a^m$ sein, wobei $l + m \leq n_0$.

   $\Rightarrow$ Wie sieht dann z aus? $z = a^n b^{n_0+1}$ wobei $l + m + n = n_0 + 1$

   (2) Aus (2) wissen wir, dass $m \geq 1$ ist.

   (3) Wir zeigen jetzt, dass keines der Wörter, die (1) und (2) erfüllen, auch (3) erfüllen *kann*, um zu einem Widerspruch zu gelangen: Wir wissen, dass $yxz = a^{n_0+1} b^{n_0+1} \in L$ ist, da $|w|_a = |w|_b$ und somit trivialerweise $|w|_a \, \big| \, |w|_b$.
   Aber: für $k = 2$ gilt $yx^k z = yx^2 z = a^{l+2m+n} b^{n_0+1} \notin L$. Warum? Da $|w|_a > |w|_b$ ist, da $m > 0$ und $l + m + n = n_0 + 1$ ist. Somit kann $|w|_a$ nicht $|w|_b$ teilen.

5. Da keine Aufteilung dieses Wortes alle drei Bedingungen *gleichzeitig* erfüllen kann, haben wir einen Widerspruch und unsere Annahme, dass L regulär sei, ist falsch.

Midterm 14, 4a)

Sei $L = \{0^{n!} \mid n \in \mathbb{N}\}$. Zeige anhand von Satz 3.1/Methode der K. Komplexität, dass diese Sprache nicht regulär ist.

1. Angenommen L ist regulär, dann gilt S3.1 für L.

2. Wir betrachten die Sprache

$$L_{0^{m!}} = \{y \in \{0\}^* \mid 0^{m!} y \in L\}$$

3. Wir betrachten uns die ersten paar Wörter aus $L_{0^{m!}}$:

$$L_{0^{m!}} = \{\lambda, 0^{m \cdot m!}, 0^{(m+2)! - m!}, \dots\}$$

4. Zu zeigen: $0^{m \cdot m!}$ ist immer das zweite Wort aus $L_{0^{m!}}$ (mit λ kriegen wir $0^{m!}$ und das nächste Wort in L ist wegen Monotonie von "!" $0^{(m+1)!}$, was genau das ist, was wir mit $0^{m \cdot m!}$ kriegen)

5. Nach Satz 3.1 ist deswegen für ein beliebiges m:

$$K(0^{m \cdot m!}) \leq \lceil \log(2+1) \rceil + c = c'$$

   für c konstant

6. $\Rightarrow$ Widerspruch! Es gibt unendlich viele Wörter von der Form $0^{m \cdot m!}$, aber nur endlich viele Wörter haben eine Kolmogorov-Komplexität von höchstens c'.

$\Rightarrow$ Unsere Annahme, dass L regulär ist, kann nicht stimmen.

4 Neue Serie

- Jede endliche Sprache ist regulär
- Am besten versucht ihr, über möglichst "schöne", klassische nicht-reguläre Sprachen zu argumentieren.
- Am besten auf triviale Cases (bspw. $\emptyset$) zurückführen