

Pumping-Lemma und Varianten

Niklas Rieken

29. Januar 2018

Der in der Vorlesung (zu Unrecht, meiner Meinung nach) am meisten beworbene Weg um zu zeigen, dass eine Sprache nicht regulär ist, ist die Anwendung des Pumping-Lemmas, was für viele FoSAP-Hörer der Endgegner der Vorlesung ist, obwohl das dahinter stehende Argument recht einfach ist. Dieses Dokument soll beim Verständnis etwas helfen, indem wir das selbe Argument auf drei verschiedene Arten anwenden.

Wir betrachten jetzt die Sprache

$$L = \{uav \mid u, v \in \{a, b\}^* \text{ mit } |u| = |v|\}$$

und zeigen, dass diese nicht regulär ist auf drei verschiedene Arten: 1. Direktes Argument am Automaten, 2. Einfache Anwendung des Pumping-Lemmas, 3. Pumping-Lemma-Spiel.

Direkte Argumentation über Automaten

Angenommen L ist regulär, dann gibt es einen NFA $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$, der L erkennt. Angenommen \mathcal{A} habe n Zustände (d.h. $|Q| = n$). Wir betrachten das Wort $w = b^n ab^n$. Das Wort ist in der Sprache, also gibt es einen akzeptierenden Lauf $\varrho = (r_0, \dots, r_{2n+1})$ mit $r_0 = q_0$ und $r_{2n+1} =: q' \in F$. In dem Lauf gibt es insgesamt also $2n+2$ Zustände, der Automat hat jedoch nur n verschiedene Zustände. Daraus folgt mit *pigeon hole principle* (*Schubfachprinzip*), dass spätestens nach dem Zustand r_n ein Zustand doppelt in ϱ vorgekommen sein muss. Sei also $r_i = r_j =: q$ ($i < j \leq n$) ein Zustand, der zwei mal im Lauf vorkam. Wir wissen außerdem, dass die Zeichen, die zwischen der Zustandswiederholung gelesen wurden, alle aus dem Präfix b^n waren. Betrachten wir nun den Lauf

$$\varrho' = (r_0, \dots, r_i, r_{j+1}, \dots, r_{2n+1}).$$

Dieser entsteht, wenn wir die $j-i =: k > 0$ vielen bs , die zwischen der Wiederholung gelesen werden, weglassen, also $w' = b^i b^{n-j} ab^n = b^{n-k} ab^n$ lesen. Dieser Lauf existiert, denn $r_i = r_j$ und somit ist auch die Transition $(r_j, b, r_{j+1}) = (r_i, b, r_{j+1}) \in \Delta$. Alle anderen Transitionen sind ohnehin die gleichen wie in ϱ . Auch der Lauf ϱ' endet in $r_{2n+1} \in F$. Somit wird auch das Wort w' von \mathcal{A} akzeptiert, aber $w' \notin L$, da vor dem a weniger bs gelesen werden als nach dem a . D.h. \mathcal{A} erkennt eine andere Sprache als L . Da \mathcal{A} beliebig gewählt war, kann also kein solcher NFA für L existieren. Somit ist L nicht regulär. In Abbildung 1 ist eine Skizze zu diesem Argument.

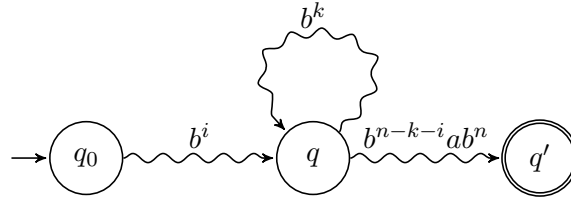


Abbildung 1: Zustandswiederholung in q . Es gibt also einen Lauf von q_0 nach q und einen Lauf von q nach p . Den Zwischenlauf von q nach q kann man also weglassen (oder auch beliebig oft wiederholen).

Anwendung des Pumping-Lemmas

Zunächst wiederholen wir das Lemma einmal im Wortlaut:

Lemma. *Sei L eine reguläre Sprache. Dann existiert ein $n \in \mathbb{N}_+$, sodass für alle $w \in L$ mit $|w| \geq n$ eine Zerlegung $w = xyz$ existiert für die gilt:*

- (i) $|xy| \leq n$,
- (ii) $y \neq \varepsilon$,
- (iii) $xy^iz \in L$ für alle $i \in \mathbb{N}$.

Um nun zu zeigen, dass L nicht regulär ist zeigen wir die Kontraposition des Lemmas, d.h.

Wenn für alle $n \in \mathbb{N}_+$ ein $w \in L$ mit $|w| \geq n$ existiert, sodass für mindestens eine Zerlegung $xyz = w$ mindestens eine der drei Bedingungen (i-iii) nicht erfüllt ist, dann ist L nicht-regulär.

Beachte: Das Pumping-Lemma kann nur dafür eingesetzt werden, um zu zeigen, dass eine Sprache nicht regulär ist! Es gibt auch nicht-reguläre Sprachen, die dem Pumping-Lemma genügen! Nun zum Beweis:

Angenommen L wäre regulär. Sei dann $n \in \mathbb{N}_+$ beliebig aber fest gewählt gemäß Pumping-Lemma. Betrachte das Wort $w = b^nab^n \in L$. Es gilt $|w| = 2n + 1 \geq n$. Sei nun $xyz = w$ eine beliebige Zerlegung, sodass (i+ii) erfüllt sind. Wir zeigen, dass dan (iii) nicht erfüllt sein kann. Es gilt, dass $y = b^k$ mit $k \leq n$ (wegen (i): $|y| \leq |xy| \leq n$, die ersten n Zeichen sind auch alles b s, sodass xy nur aus b s bestehen kann und somit auch y). Außerdem ist $k > 0$ (wegen (ii): $y \neq \varepsilon = b^0$). Wähle nun $i = 0$ und betrachte das Wort $w' = xy^0z = xz = b^{n-k}ab^n$. Nach Pumping-Lemma müsste dies zu L gehören, was jedoch offenbar nicht der Fall ist. Damit haben wir einen Widerspruch zum Pumping-Lemma und es folgt, dass L nicht regulär ist.

Bemerkung. Man sieht leicht wie analog die Argumentation über das Pumping-Lemma und über das direkte Argument am NFA verlaufen.

Bemerkung. Eine Wahl von $i \geq 2$ und das daraus resultierende "aufgepumpte" Wort $b^{n+(i-1)k}ab^n \notin L$ wäre natürlich ebenfalls möglich gewesen.

Das Pumping-Lemma-Spiel

Es ist auch möglich, das Pumping-Lemma als Spiel aufzufassen. Das Vorgehen ist dabei identisch zu Model-Checking-Spielen für First Order Logic (*Prädikatenlogik*), die man in der Vorlesung Mathematische Logik (normalerweise im 4. Semester) sieht. Wir beginnen mit den Spielern:

Alice versucht die Beweisführung der Nicht-Regularität zu sabotieren,
(*Verifiziererin, Defender*)

Bob möchte zeigen, dass eine Sprache nicht regulär ist.
(*Falsifizierer, Attacker*)

Die Spielregeln gehen wie folgt: Das Spielfeld ist die Sprache L .

- 1) Alice wählt eine Zahl $n \in \mathbb{N}_+$.
- 2) Bob wählt ein Wort $w \in L$ mit $|w| \geq n$.
- 3) Alice wählt eine Zerlegung $xyz = w$, mit $|xy| \leq n$ und $y \neq \varepsilon$.
- 4) Bob wählt eine Zahl $i \in \mathbb{N}$.

Alice gewinnt, wenn $xy^iz \in L$ oder Bob in Schritt 2 nicht ziehen kann (das wäre der Fall wenn jedes Wort aus L kürzer ist als das n , welches Alice vorgegeben hat – die Sprache wäre dann endlich und somit regulär).

Bob gewinnt, wenn $xy^iz \notin L$.

Bemerkung. Falls L regulär ist, so hat Alice eine Gewinnstrategie, d.h. sie kann ein n wählen, sodass sie für jedes Wort, mit dem Bob in Schritt 2 antworten könnte eine Zerlegung findet, die Bob in Schritt 4 nicht aus der Sprache rauspumpen kann.

Äquivalent dazu: Falls Bob eine Gewinnstrategie hat, so ist L nicht regulär.

Bemerkung. Beachtet, dass in Schritt 3 Alice(!) diejenige ist, die eine Zerlegung wählt. Wenn wir zeigen wollen, dass Bob eine Gewinnstrategie hat, heißt das, dass wir für alles(!) was Alice wählen könnte im 4. Schritt noch aus der Sprache herausgepumpt werden kann. Es wäre keine Gewinnstrategie, wenn Bob nur auf manche Züge von Alice eine passende Antwort findet.

Wir geben nun eine Gewinnstrategie für Bob auf der Sprache L von oben an.

- 1) Alice wählt eine Zahl n .
- 2) Bob wählt das Wort $w = b^n a b^n$ ($|w| \geq n$ ✓).
- 3) Alice wählt eine Zerlegung $xyz = w$ mit $|xy| \leq n$, $y \neq \varepsilon$ (d.h. $y = b^k$ für ein $0 < k \leq n$ mit der selben Begründung wie oben).
- 4) Bob wählt $i = 0$, unabhängig von Alice' Zerlegung.

Das Wort $xz = b^{n-k} a b^n$ gehört nicht zur Sprache L für egal welches $k > 0$ welches wir aus Alice' Zerlegung erhalten. Damit hat Bob eine Gewinnstrategie und somit ist L nicht regulär.