

Induktion über Wortlänge

Niklas Rieken
niklas@fosap.de

30. April 2018

In diesem Dokument zeigen wir, wie man beweist, dass ein gegebener DFA die richtige Sprache erkennt. Für die Automaten in dieser Vorlesung ist dies meist relativ offensichtlich, da sie wenige Zustände haben, aber die Beweistechnik sollte trotzdem an einfachen Beispielen geübt werden. Wir verwenden den Automaten aus Tutoriumsaufgabe 5 von Übungsblatt 2 aus dem Sommersemester 2018 im Fach Formale Systeme, Automaten, Prozesse an der RWTH Aachen.

Allgemeines zur vollständigen Induktion

Allgemein zeigt man mit der vollständigen Induktion Aussagen für *induktive Mengen*, meistens $\mathbb{N} := \{0, 1, 2, \dots\}$. Für eine Aussageform A zeigen wir, dass $A(n)$ für alle $n \in \mathbb{N}$ gilt. Klassische Beispiel für Aussagen, die man mit vollständiger Induktion zeigt sind die *Gauss'sche Summenformel* oder die *binomische Formel*. Wir verwenden die vollständige Induktion nun, um zu zeigen, dass ein DFA für jedes Wort *das richtige tut*, d.h. jedes Wort, was zur Sprache gehört, wird akzeptiert, alle anderen werden verworfen. In Anlehnung an die Notation oben könnten wir die Aussageform $A(w)$ für ein $w \in \Sigma^*$ interpretieren durch " w wird akzeptiert g.d.w. $w \in L$ ". Nun sind Wörter über einem Alphabet Σ keine natürlichen Zahlen. Wir können allerdings statt dem Wort selber die Länge des Wortes betrachten und erhalten ein analoges Schema wie z.B. für den Beweis der Summenformel. Dann müssen wir im Induktionsschritt (von Wortlänge n auf Wortlänge $n+1$) nur zeigen, dass die Transitionsfunktion für das angehängte Symbol das richtige tut, d.h. wenn für Wortlängen n (für alle Wörter $w = a_1 \dots a_n$) der Automat stets im richtigen Zustand endet, dann auch für das Wort wa_{n+1} für alle $a_{n+1} \in \Sigma$. Der Induktionsschritt involviert dann lediglich eine Fallunterscheidung über das zuletzt gelesene Symbol. Intuitiv werden die Beweise, wenn man den Zuständen des Automaten eine gewisse *Semantik* (eine Aufgabe) gibt, z.B. ist der Zustand q_b in Abbildung 1 ein Zustand der auf das Symbol a wartet, falls in den letzten beiden Symbolen kein a zu finden ist. Genauer es jetzt in der Aufgabe.

Aufgabe

Wir betrachten den Automaten in Abbildung 1. Zu zeigen ist nun, dass

$$L = \{w \in \{a, b\}^* \mid \text{das vorletzte Zeichen von } w \text{ ist } a\}$$

die Sprache dieses Automaten ist. Wir zeigen folgende Aussage, welche die Korrektheit des Automaten induziert (denn q_{aa}, q_{ab} sind genau die akzeptierenden Zustände von \mathcal{A}):

$S(n)$: Für jedes Wort $w = a_1 \dots a_n \in \{a, b\}^*$ der Länge n gilt für den Lauf $\varrho = (r_0, \dots, r_n)$ des Automaten \mathcal{A} aus Abbildung 1 auf w , dass

$$r_n = \begin{cases} q_\varepsilon, & w = \varepsilon \\ q_a, & a_n = a \text{ und } (a_{n-1} = b \text{ oder } n = 1) \\ q_b, & a_n = b \text{ und } (a_{n-1} = b \text{ oder } n = 1) \\ q_{aa}, & a_n = a_{n-1} = a \\ q_{ab}, & a_n = b \text{ und } a_{n-1} = a. \end{cases}$$

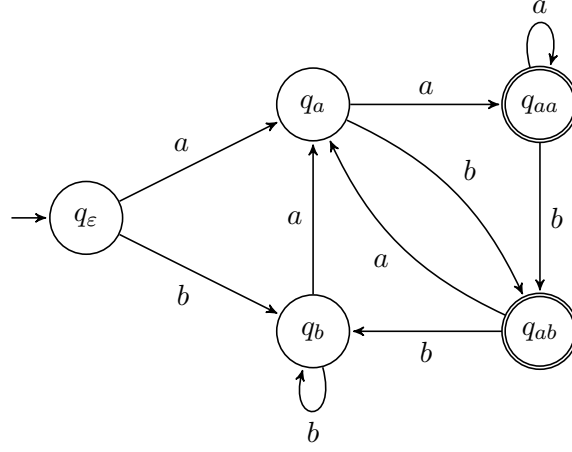


Abbildung 1: Automat \mathcal{A} , der die Sprache $L = \{w \in \{a, b\}^* \mid \text{das vorletzte Zeichen von } w \text{ ist } a\}$ erkennt.

Behauptung: $S(n)$ gilt für alle Wortlängen $n \in \mathbb{N}$. Dies beweisen wir vermöge einer vollständigen Induktion.

Induktionsverankerung: Sei $n = 0$, d.h. $w = \varepsilon$. Also hat \mathcal{A} den Lauf (q_ε) (nur Startzustand). Also ist $r_n = r_0 = q_\varepsilon$. Somit gilt $S(0)$. ✓

Induktionsschritt: Die Aussage $S(n)$ gelte für ein $n \in \mathbb{N}$ (Induktionshypothese (IH)). Wir folgern daraus, dass auch $S(n+1)$ gilt. Betrachte also das Wort $w = a_1 \dots a_n a_{n+1}$. Wir machen eine Fallunterscheidung zum Lauf auf dem Wort $a_1 \dots a_n$ (sic!).

1. $a_1 \dots a_n = \varepsilon$, d.h. $n+1 = 1$ und nach (IH) $r_n = q_\varepsilon$.
 - Wenn $a_{n+1} = a$, dann $r_{n+1} = \delta(q_\varepsilon, a) = q_a$.
 - Wenn $a_{n+1} = b$, dann $r_{n+1} = \delta(q_\varepsilon, b) = q_b$.
2. $a_n = a$, d.h. nach (IH) $r_n \in \{q_a, q_{aa}\}$. Wir können diese beiden Fälle zusammenfassen, da uns bei Betrachtung des Wortes $a_1 \dots a_{n+1}$ das Symbol a_{n-1} nicht mehr interessiert, denn dies ist nun das drittletzte Symbol und spielt keine Rolle mehr für die weitere Berechnung.
 - Wenn $a_{n+1} = a$, dann $r_{n+1} = \delta(q_a, a) = \delta(q_{aa}, a) = q_{aa}$.
 - Wenn $a_{n+1} = b$, dann $r_{n+1} = \delta(q_a, b) = \delta(q_{aa}, b) = q_{ab}$.
3. $a_n = b$, d.h. nach (IH) $r_n \in \{q_b, q_{ab}\}$. Wir können diese beiden Fälle wieder analog wie im zweiten Fall zusammenfassen.
 - Wenn $a_{n+1} = a$, dann $r_{n+1} = \delta(q_b, a) = \delta(q_{ab}, a) = q_a$.
 - Wenn $a_{n+1} = b$, dann $r_{n+1} = \delta(q_b, b) = \delta(q_{ab}, b) = q_b$.

Man sieht leicht, dass das Verhalten der Transitionsfunktion, wie in den sechs Bulletpoints beschrieben, das gewünschte Verhalten ist (überprüfe einfach mit den fünf Fällen für r_n zur Aussage $S(n)$ mit $n+1$ anstelle von n). Nach dem Prinzip der vollständigen Induktion folgt die Aussage für alle $n \in \mathbb{N}$. \square