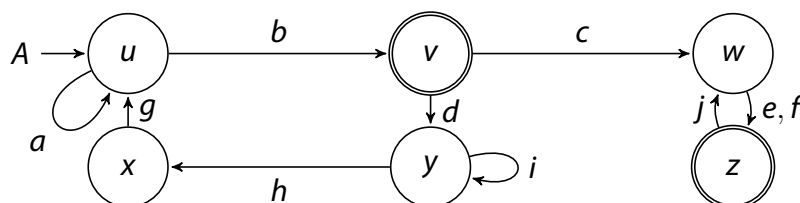


Berechnung eines Regulären Ausdrucks



Das Gleichungssystem lautet:

$$\mathcal{L}(A) = U = aU \cup bV$$

$$V = \varepsilon \cup cW \cup dY$$

$$W = (euf)Z$$

$$X = gU$$

$$Y = hX \cup iY$$

$$Z = \varepsilon \cup jW$$

Die Variablen sollen für Mengen von Wörtern stehen, mit denen der Automat vom jeweiligen Zustand in einen akzeptierenden Zustand gelangen kann.

Daher ist das leere Wort bei den akzeptierenden Zuständen immer enthalten. Ansonsten betrachte man für jeden Zustand nur die ausgehenden Kanten.

Je nach Strategie entsteht ein anderer Ausdruck als Lösung.

Insbesondere muss aber jeder Kreis im Transitionsgraphen, der nicht aus kleineren Kreisen besteht, mittels Ardens Lemma transformiert werden.

Solange man sich nicht wiederholt, benötigt man höchstens sooft Arden's Lemma, wie es einfache Kreise im Graphen gibt.

Dieser Automat hat vier Kreise: $u \rightarrow u$, $y \rightarrow y$, $u \rightarrow v \rightarrow y \rightarrow x \rightarrow u$ und $w \rightarrow z \rightarrow w$.

Allerdings können überschneidende Kreise gleichzeitig vereinfacht werden.

Für A reichen hier drei Anwendungen.

Eine Strategie kann sein, zuerst die Kreise ohne u abzuarbeiten.

$$Y = i^* hX$$

Arden's Lemma

$$Z = \varepsilon \cup j(euf)Z$$

Einfügen von W

$$= (j(euf))^*$$

Arden's Lemma

$$U = aU \cup b(\varepsilon \cup cW \cup dY)$$

Einfügen von V

$$= aU \cup b(\varepsilon \cup c(euf)Z \cup dY)$$

Einfügen von W

$$= aU \cup b(\varepsilon \cup c(euf)(j(euf))^* \cup dY)$$

Einfügen von Z

$$= aU \cup b(\varepsilon \cup c(euf)(j(euf))^* \cup di^* hX)$$

Einfügen von Y

$$= aU \cup b(\varepsilon \cup c(euf)(j(euf))^* \cup di^* hgU)$$

Einfügen von X

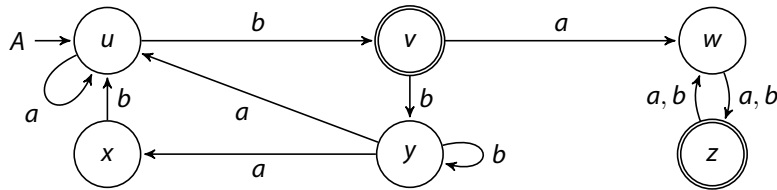
$$= aU \cup b \cup bc(euf)(j(euf))^* \cup bdi^* hgU$$

Distributivität

$$= (a \cup bdi^* hg)^* (b \cup bc(euf)(j(euf))^*)$$

Arden's Lemma

Potenzmengenkonstruktion nach Rabin & Scott



Es reichen die erreichbaren Zustände. Dazu starte mit $\{q_0\}$.

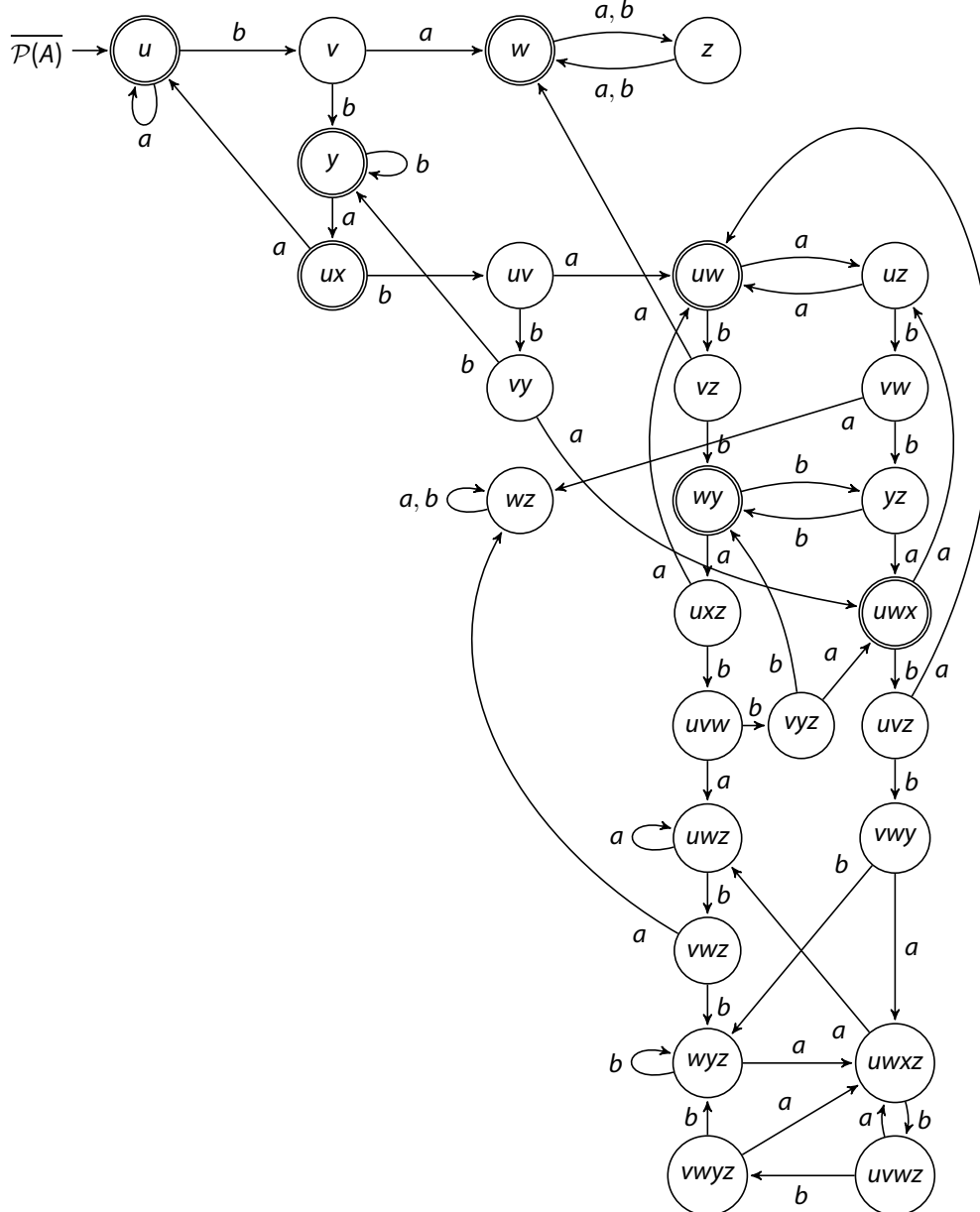
Für jedes Symbol in Σ , bestimme die Folgezustandsmenge und füge eine Kante hinzu.

Falls es nun keine Zustände gibt, denen ausgehende Kanten fehlen, ist man fertig.

Sonst, wiederhole den Prozess mit einem solchen Zustand.

(Akzeptierend sind die Zustandsmengen, die einen akzeptierenden Zustand enthalten.)

Tipp: Bevorzugt zuerst kleinere Teilmengen vor Größeren.



Komplementierung eines DFA: Akzeptierend sind nun die Zustände, die keinen akzeptierenden Zustand enthalten. $Q_F'' := \{ Q' \subseteq Q \mid \text{nicht } \exists q \in Q' : q \in Q_F \}$

Produktautomat

Theorem

Seien Σ ein endliches Alphabet und $L, M \subseteq \Sigma^*$ regulär. Der Durchschnitt $L \cap M$ ist regulär.

Klar mit deMorgan'schen Regeln und der bereits bekannten Abgeschlossenheit unter Komplement: $L \cap M = \overline{\overline{L} \cup \overline{M}}$. □

Dennoch geht es praktischer. Seien $A = \langle Q_A, q_{0A}, \rightarrow_A, Q_{FA} \rangle$ und $B = \langle Q_B, q_{0B}, \rightarrow_B, Q_{FB} \rangle$ endliche Automaten mit $\mathcal{L}(A) = L$ und $\mathcal{L}(B) = M$. Der Produktautomat $A \cap B = \langle Q_A \times Q_B, \langle q_{0A}, q_{0B} \rangle, \rightarrow', Q_{FA} \times Q_{FB} \rangle$ enthalte genau dann eine Transition $\langle p_A, p_B \rangle \xrightarrow{s} \langle q_A, q_B \rangle$, wenn gleichzeitig $p_A \xrightarrow{s}_A q_A$ und $p_B \xrightarrow{s}_B q_B$ existieren. Man sagt, $A \cap B$ simuliert gleichzeitig A und B und akzeptiert genau dann, wenn beide Simulationen akzeptieren.

Zeige $\mathcal{L}(A \cap B) = \mathcal{L}(A) \cap \mathcal{L}(B)$.

Sei $w = w_1 w_2 \dots w_n \in \Sigma^*$. Wir zeigen beide Teilmengen-Beziehungen gleichzeitig.

$$\begin{aligned} w \in \mathcal{L}(A \cap B) &\text{ gdw. } \langle q_{0A}, q_{0B} \rangle \xrightarrow{w_1} \langle q_{1A}, q_{1B} \rangle \dots \xrightarrow{w_n} \langle q_{nA}, q_{nB} \rangle \in Q_{FA} \times Q_{FB} \\ &\text{ gdw. } q_{0A} \xrightarrow{w_1}_A q_{1A} \dots \xrightarrow{w_n}_A q_{nA} \in Q_{FA} \text{ und } q_{0B} \xrightarrow{w_1}_B q_{1B} \dots \xrightarrow{w_n}_B q_{nB} \in Q_{FB} \\ &\text{ gdw. } w \in \mathcal{L}(A) \cap \mathcal{L}(B) \end{aligned}$$

□

Shuffle

Definition

Seien Σ ein Alphabet und $v, w \in \Sigma^*$. Der Shuffle $v \sqcup w$ ist die Menge der Worte, die durch eine beliebige Abwechslung der Buchstaben von v und w hervorgehen. Dabei soll die Reihenfolge der Buchstaben aus demselben Wort erhalten bleiben. Der Shuffle zweier Sprachen $L \subseteq \Sigma^*$ und $M \subseteq \Sigma^*$ ist die Sprache

$$L \sqcup M := \begin{cases} \{v_1 \cdot w_1\} & \text{falls } L = \{v_1\} \text{ und } M = \{w_1\} \text{ und } (v_1 = \varepsilon \text{ oder } w_1 = \varepsilon) \\ a \cdot (x \sqcup w_2) \cup b \cdot (v_2 \sqcup y) & \text{falls } L = \{v_2\} \text{ und } M = \{w_2\} \text{ und } v_2 = a \cdot x \text{ und } w_2 = b \cdot y \\ \bigcup_{v_3 \in L, w_3 \in M} \{v_3\} \sqcup \{w_3\} & \text{sonst} \end{cases}$$

Beispiel

$$aab \sqcup ba = \{\underline{a} \underline{a} \underline{b} \underline{b} \underline{a}, \underline{a} \underline{a} \underline{b} \underline{b} \underline{a}, \underline{a} \underline{a} \underline{b} \underline{a} \underline{b}, \underline{a} \underline{b} \underline{a} \underline{b} \underline{a}, \underline{a} \underline{b} \underline{a} \underline{a} \underline{b}, \underline{a} \underline{b} \underline{a} \underline{a} \underline{b}, \underline{b} \underline{a} \underline{a} \underline{b} \underline{a}, \underline{b} \underline{a} \underline{a} \underline{a} \underline{b}, \underline{b} \underline{a} \underline{a} \underline{a} \underline{b}\}$$

Beispiele für reguläre Sprachen sind

$$\begin{aligned} (aa)^* \sqcup b^* &= b^* (ab^* ab^*)^* & (ab)^* \sqcup (ba)^* &= (ab \cup ba)^* \\ a^* \sqcup b^* &= \{a, b\}^* & a \sqcup a^* \cdot b \cdot a^* &= a^* \cdot (ab \cup ba) \cdot a^* \end{aligned}$$

Proposition

Falls L und M regulär sind, dann ist es auch $L \sqcup M$.

Erster Ansatz: Nutze Charakterisierung als Automaten.

Seien $A = \langle Q_A, q_{0A}, \rightarrow_A, Q_{FA} \rangle$ und $B = \langle Q_B, q_{0B}, \rightarrow_B, Q_{FB} \rangle$ NFAs auf dem Alphabet Σ . Betrachte den NFA $A \sqcup B := \langle Q_A \times Q_B, \langle q_{0A}, q_{0B} \rangle, \rightarrow, Q_{FA} \times Q_{FB} \rangle$ mit

$$\langle p_A, p_B \rangle \xrightarrow{s} \langle q_A, q_B \rangle \quad \text{gdw.} \quad \left(p_A \xrightarrow{s}_A p_A \wedge p_B = q_B \right) \vee \left(p_A = q_A \wedge p_B \xrightarrow{s}_B q_B \right)$$

Nun gilt es, die Korrektheit dieser Konstruktion zu zeigen.

Beweisskizze

Zeige $\mathcal{L}(A \sqcup B) = \mathcal{L}(A) \sqcup \mathcal{L}(B)$ mit Antisymmetrie.

Sei $x = x_1 \dots x_k \in \mathcal{L}(A \sqcup B)$.

Dann gibt es einen akzeptierenden Lauf $q_0 \xrightarrow{x_1} \langle a_1, b_1 \rangle \xrightarrow{x_2} \dots \xrightarrow{x_k} \langle a_k, b_k \rangle \in Q_F$.

Dabei simuliert gemäß der obigen Konstruktion jede dieser Transitionen entweder eine A -Transition oder eine B -Transition, während der simulierte Zustand des jeweils anderen Automaten unverändert bleibt.

Die Aneinanderreihung der A -Transitionen liefert nun ein Subwort $y \in \Sigma^*$ mit einen akzeptierenden Lauf $q_{0A} \xrightarrow{y}_A^* a_k \in Q_{FA}$, der aufgrund der Konstruktion von $A \sqcup B$ in A existieren musste.

Analog erhalte über den B -Transitionen das andere Subwort $z \in \Sigma^*$ und einen akzeptierenden Lauf $q_{0B} \xrightarrow{z}_B^* b_k \in Q_{FB}$.

Es gilt $x \in \{y\} \sqcup \{z\} \subseteq \mathcal{L}(A) \sqcup \mathcal{L}(B)$.

Da dies für alle x gilt, folgt $\mathcal{L}(A \sqcup B) \subseteq \mathcal{L}(A) \sqcup \mathcal{L}(B)$.

Sei $x = x_1 \dots x_k \in \mathcal{L}(A) \sqcup \mathcal{L}(B)$.

Dann gibt es $y = y_1 \dots y_n \in \mathcal{L}(A)$ und $z = z_1 \dots z_m \in \mathcal{L}(B)$ mit $x \in \{y\} \sqcup \{z\}$ und $n + m = k$.

Genauer betrachte eine beliebige Buchstabenzuweisung $f : \{1, \dots, k\} \rightarrow \{0, 1\}$ mit $y = \langle x_i \rangle_{f(i)=0}$ und $z = \langle x_i \rangle_{f(i)=1}$, falls mehrere existieren.

Dann gibt es akzeptierenden Läufe $q_{0A}=a_0 \xrightarrow{y_1}_A \dots \xrightarrow{y_n}_A a_n \in Q_{FA}$ und $q_{0B}=b_0 \xrightarrow{z_1}_B \dots \xrightarrow{z_m}_B b_m \in Q_{FB}$.

In der gleichen Anordnung wie im Wort, lässt sich nun der folgende Lauf konstruieren:

$$q_0 = \langle a'_0, b'_0 \rangle \xrightarrow{x_1} \langle a'_1, b'_1 \rangle \xrightarrow{x_2} \dots \xrightarrow{x_k} \langle a'_k, b'_k \rangle$$

$$\langle a'_{i+1}, b'_{i+1} \rangle := \begin{cases} \langle a_{j+1}, b'_i \rangle & \text{falls } a'_i = a_j \text{ und } f(i+1) = 0 \\ \langle a'_i, b_{j+1} \rangle & \text{falls } b'_i = b_j \text{ und } f(i+1) = 1 \end{cases}$$

Per Konstruktion existieren alle erforderlichen Transitionen tatsächlich und es gilt $\langle a'_k, b'_k \rangle = \langle a_n, b_m \rangle \in Q_F$. Also wird x von $A \sqcup B$ akzeptiert.

Da dies für alle x gilt, folgt $\mathcal{L}(A) \sqcup \mathcal{L}(B) \subseteq \mathcal{L}(A \sqcup B)$.

Alternativ: Nutze Abgeschlossenheit unter homomorphen Bildern und Urbildern.

Idee: Erweitere kurzfristig das Alphabet zu $\Gamma = \Sigma \times \{0, 1\}$. Markiere für jeden Buchstaben, zu dem Wort welcher Sprache es gehören soll.

Betrachte Homomorphismen $\pi_\Sigma, \sigma_0, \sigma_1 : \Gamma^* \rightarrow \Sigma^*$ mit

$$\pi_\Sigma(\langle a, m \rangle) := a \quad \sigma_0(\langle a, m \rangle) := \begin{cases} a & \text{falls } m = 0 \\ \varepsilon & \text{falls } m = 1 \end{cases} \quad \sigma_1(\langle a, m \rangle) := \begin{cases} a & \text{falls } m = 1 \\ \varepsilon & \text{falls } m = 0 \end{cases}.$$

Zeige induktiv $\forall x, y, z \in \Sigma^*, x \in \{y\} \sqcup \{z\} : \exists x' \in \Gamma^* : \pi_\Sigma(x') = x \wedge \sigma_0(x') = y \wedge \sigma_1(x') = z$.

Anfang: Für $x = \varepsilon$ gilt $y = z = \varepsilon$. Wähle $x' = \varepsilon$.

Voraussetzung: Aussage gelte bis Länge $|w| = n$.

Schluss: Sei $a.x \in \{y\} \sqcup \{z\}$.

Falls $z = \varepsilon$, wähle $x' := \langle y_1, 0 \rangle \dots \langle y_n, 0 \rangle$. Falls $y = \varepsilon$, wähle $x' := \langle z_1, 1 \rangle \dots \langle z_m, 1 \rangle$.

Anderenfalls seien $y = b.y'$ und $z = c.z'$. Es gilt $a.x \in b(y' \sqcup z) \cup c(y \sqcup z')$.

Falls $a.x \in b(y' \sqcup z)$, dann gibt es ein x'' mit $\pi_\Sigma(x'') = x, \sigma_0(x'') = y'$ und $\sigma_1(x'') = z$. Wähle $x' := \langle b, 0 \rangle x''$.

Anderenfalls gibt es ein x''' mit $\pi_\Sigma(x''') = x, \sigma_0(x''') = y$ und $\sigma_1(x''') = z'$. Wähle $x' := \langle c, 1 \rangle x'''$.

Zeige induktiv $\forall x' \in \Gamma^* : \pi_\Sigma(x') \in \{\sigma_0(x')\} \sqcup \{\sigma_1(x')\}$.

Anfang: Falls $w' = \varepsilon$, dann $w = \sigma_0(w') = \sigma_1(w') = \varepsilon \in \{\varepsilon\} \sqcup \{\varepsilon\}$.

Voraussetzung: Aussage gelte bis Länge $|w'| = n$.

Schluss: Sei $x' := \langle a_0, m_0 \rangle \langle a_1, m_1 \rangle \dots \langle a_n, m_n \rangle$.

Falls $\sigma_1(x') = \varepsilon$, dann gelten $\pi_\Sigma(x') = \sigma_0(x')$ und $\{\sigma_0(x')\} \sqcup \{\sigma_1(x')\} = \{\sigma_0(x')\}$. Fall $\sigma_1(x') = \varepsilon$ ist analog.

Anderenfalls seien $\sigma_0(x') = b.y, \sigma_1(x') = c.z$. Betrachte $x'' := \langle a_1, m_1 \rangle \dots \langle a_n, m_n \rangle$.

Falls $m_0 = 0$, folgen $a_0 = b, \sigma_0(x'') = y$ und $\sigma_1(x'') = \sigma_1(x')$.

Es folgt zunächst $\pi_\Sigma(x'') \in \{y\} \sqcup \{c.z\}$ nach Voraussetzung.

Damit gilt schließlich $\pi_\Sigma(x') = a_0.\pi_\Sigma(x'') \in b.(\{y\} \sqcup \{c.z\}) \subseteq \{\sigma_0(w')\} \sqcup \{\sigma_1(w')\}$.

Fall $m_0 = 1$ ist analog.

Zeige $L \sqcup M = \pi_\Sigma(\sigma_0^{-1}(L) \cap \sigma_1^{-1}(M))$ mittels Antisymmetrie.

Sei $w \in L \sqcup M$.

Dann gibt es Wörter $u \in L$ und $v \in M$ mit $w \in \{u\} \sqcup \{v\}$.

Weiterhin gibt es ein $w' \in \Gamma^*$ mit $\pi_\Sigma(w') = w, \sigma_0(w') = u$ und $\sigma_1(w') = v$ wie oben induktiv gezeigt.

Erhalte $w' \in \sigma_0^{-1}(u) \subseteq \sigma_0^{-1}(L)$ und $w' \in \sigma_1^{-1}(v) \subseteq \sigma_1^{-1}(M)$.

Damit liegt $w \in \pi_\Sigma(\sigma_0^{-1}(L) \cap \sigma_1^{-1}(M))$.

Da dies für alle w gilt, ist $L \sqcup M \subseteq \pi_\Sigma(\sigma_0^{-1}(L) \cap \sigma_1^{-1}(M))$.

Sei nun $w \in \pi_\Sigma(\sigma_0^{-1}(L) \cap \sigma_1^{-1}(M))$.

Dann gibt es ein $w' \in \sigma_0^{-1}(L) \cap \sigma_1^{-1}(M)$ mit $\pi_\Sigma(w') = w$.

Es folgt $w \in \{\sigma_0(w')\} \sqcup \{\sigma_1(w')\}$ wie oben induktiv gezeigt.

Weiterhin gelten $\sigma_0(w') \in L$ und $\sigma_1(w') \in M$, da $w \in \sigma_0^{-1}(L)$ und $w \in \sigma_1^{-1}(M)$.

Zusammen erhält man $w \in L \sqcup M$.

Da dies für alle w gilt, ist $\pi_\Sigma(\sigma_0^{-1}(L) \cap \sigma_1^{-1}(M)) \subseteq L \sqcup M$. □