

Vorlesungsskript
Einführung in die Theoretische
Informatik

Wintersemester 2016/17

Prof. Dr. Johannes Köbler
Humboldt-Universität zu Berlin
Lehrstuhl Komplexität und Kryptografie

15. Dezember 2016

Inhaltsverzeichnis

1	Einleitung	1
2	Reguläre Sprachen	2
2.1	Endliche Automaten	2
2.2	Nichtdeterministische endliche Automaten	5
2.3	Reguläre Ausdrücke	7
2.4	Relationalstrukturen	10
2.4.1	Ordnungs- und Äquivalenzrelationen	14
2.4.2	Abbildungen	17
2.4.3	Homo- und Isomorphismen	17
2.5	Minimierung von DFAs	19
2.6	Das Pumping-Lemma	23
2.7	Grammatiken	24
3	Kontextfreie Sprachen	27
3.1	Chomsky-Normalform	29
3.2	Das Pumping-Lemma für kontextfreie Sprachen	32
3.3	Der CYK-Algorithmus	33
3.4	Kellerautomaten	35
3.5	Deterministisch kontextfreie Sprachen	39
4	Kontextsensitive Sprachen	44
4.1	Kontextsensitive Grammatiken	44
4.2	Turingmaschinen	44

1 Einleitung

Rechenmaschinen spielen in der Informatik eine zentrale Rolle. In dieser Vorlesung beschäftigen wir uns mit mathematischen Modellen für Maschinentypen von unterschiedlicher Berechnungskraft. Unter anderem lernen wir das Rechenmodell der Turingmaschine (TM) kennen, mit dem sich alle anderen Rechenmodelle simulieren lassen. Ein weiteres wichtiges Thema der Vorlesung ist die Frage, welche Probleme algorithmisch lösbar sind und wo die Grenzen der Berechenbarkeit verlaufen.

Schließlich untersuchen wir die Komplexität von algorithmischen Problemen, indem wir den benötigten Rechenaufwand möglichst gut nach oben und unten abschätzen. Eine besondere Rolle spielen hierbei die NP-vollständigen Probleme, deren Komplexität bis heute offen ist.

Themen der Vorlesung

- Welche Rechenmodelle sind für bestimmte Aufgaben adäquat? (Automatentheorie)
- Welche Probleme sind lösbar? (Berechenbarkeitstheorie)
- Welcher Aufwand ist zur Lösung eines algorithmischen Problems nötig? (Komplexitätstheorie)

In den theoretisch orientierten Folgeveranstaltungen wird es dagegen um folgende Themen gehen.

Thema der Vorlesung Algorithmen und Datenstrukturen

- Wie lassen sich praktisch relevante Problemstellungen möglichst effizient lösen? (Algorithmik)

Thema der Vorlesung Logik in der Informatik

- Mathematische Grundlagen der Informatik, Beweise führen, Modellierung (Aussagenlogik, Prädikatenlogik)

Der Begriff *Algorithmus* geht auf den persischen Gelehrten **Muhammed Al Chwarizmi** (8./9. Jhd.) zurück. Der älteste bekannte nicht-triviale Algorithmus ist der nach *Euklid* benannte Algorithmus zur Berechnung des größten gemeinsamen Teilers zweier natürlicher Zahlen (300 v. Chr.). Von einem Algorithmus wird erwartet, dass er jede *Problemeingabe* nach endlich vielen Rechenschritten löst (etwa durch Produktion einer *Ausgabe*). Eine wichtige Rolle spielen Entscheidungsprobleme, bei denen jede Eingabe nur mit ja oder nein beantwortet wird. Problemeingaben können Zahlen, Formeln, Graphen etc. sein. Diese werden über einem *Eingabealphabet* Σ kodiert.

Definition 1.

- Ein **Alphabet** $\Sigma = \{a_1, \dots, a_m\}$ ist eine geordnete Menge von endlich vielen **Zeichen**.
- Eine Folge $x = x_1 \dots x_n$ von n Zeichen heißt **Wort** (der **Länge** $|x| = n$).
- Die Menge aller Wörter über Σ ist

$$\Sigma^* = \bigcup_{n \geq 0} \Sigma^n,$$

wobei $\Sigma^n = \{x_1 \dots x_n \mid n \geq 0 \text{ und } x_i \in \Sigma \text{ für } i = 1, \dots, n\}$ alle Wörter der Länge n enthält.

- Das (einzige) Wort der Länge $n = 0$ ist das **leere Wort**, welches wir mit ε bezeichnen.
- Jede Teilmenge $L \subseteq \Sigma^*$ heißt **Sprache** über dem Alphabet Σ .

Beispiel 2. Sei Σ ein Alphabet. Dann sind $\emptyset, \Sigma^*, \Sigma$ und $\{\varepsilon\}$ Sprachen über Σ . Die Sprache \emptyset enthält keine Wörter und heißt **leere Sprache**. Die Sprache Σ^* enthält dagegen alle Wörter über Σ , während die Sprache Σ alle Wörter über Σ der Länge 1 enthält. Die Sprache

$\{\varepsilon\}$ enthält nur das leere Wort, ist also einelementig. Einelementige Sprachen werden auch als **Singletonsprachen** bezeichnet.

Da Sprachen Mengen sind, können wir sie bzgl. Inklusion vergleichen. Zum Beispiel gilt

$$\emptyset \subseteq \{\varepsilon\} \subseteq \Sigma^*.$$

Wir können Sprachen auch vereinigen, schneiden und komplementieren. Seien A und B Sprachen über Σ . Dann ist

- $A \cap B = \{x \in \Sigma^* \mid x \in A, x \in B\}$ der **Schnitt** von A und B ,
- $A \cup B = \{x \in \Sigma^* \mid x \in A \vee x \in B\}$ die **Vereinigung** von A und B , und
- $\bar{A} = \{x \in \Sigma^* \mid x \notin A\}$ das **Komplement** von A .

Neben den Mengenoperationen gibt es auch spezielle Sprachoperationen.

Definition 3.

- Das **Produkt (Verkettung, Konkatenation)** der Sprachen A und B ist

$$AB = \{xy \mid x \in A, y \in B\}.$$

Ist $A = \{x\}$ eine Singletonsprache, so schreiben wir für $\{x\}B$ auch einfach xB .

- Die **n -fache Potenz A^n** einer Sprache A ist induktiv definiert durch

$$A^n = \begin{cases} \{\varepsilon\}, & n = 0, \\ A^{n-1}A, & n > 0. \end{cases}$$

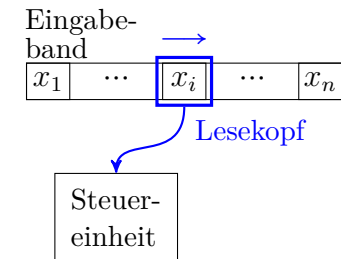
- Die **Sternhülle A^*** von A ist $A^* = \bigcup_{n \geq 0} A^n$.
- Die **Plushülle A^+** von A ist $A^+ = \bigcup_{n \geq 1} A^n = AA^*$.

2 Reguläre Sprachen

Wir betrachten zunächst Einschränkungen des TM-Modells, die vielfältige praktische Anwendungen haben, wie z.B. endliche Automaten (DFA, NFA), Kellerautomaten (PDA, DPDA) etc.

2.1 Endliche Automaten

Ein endlicher Automat führt bei einer Eingabe der Länge n nur n Rechenschritte aus. Um die gesamte Eingabe lesen zu können, muss der Automat also in jedem Schritt ein Zeichen der Eingabe verarbeiten.



Definition 4. Ein **endlicher Automat** (kurz: DFA; deterministic finite automaton) wird durch ein 5-Tupel $M = (Z, \Sigma, \delta, q_0, E)$ beschrieben, wobei

- $Z \neq \emptyset$ eine endliche Menge von **Zuständen**,
- Σ das **Eingabealphabet**,
- $\delta : Z \times \Sigma \rightarrow Z$ die **Überföhrungsfunktion**,
- $q_0 \in Z$ der **Startzustand** und
- $E \subseteq Z$ die Menge der **Endzustände** ist.

Die von M **akzeptierte** oder **erkannte Sprache** ist

$$L(M) = \left\{ x_1 \dots x_n \in \Sigma^* \mid \begin{array}{l} \text{es gibt } q_1, \dots, q_{n-1} \in Z, q_n \in E \text{ mit} \\ \delta(q_i, x_{i+1}) = q_{i+1} \text{ f\u00fcr } i = 0, \dots, n-1 \end{array} \right\}.$$

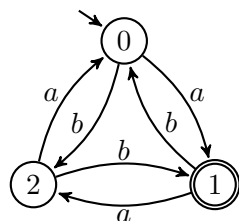
Eine Zustandsfolge q_0, q_1, \dots, q_n heißt **Rechnung** von $M(x_1 \dots x_n)$, falls $\delta(q_i, x_{i+1}) = q_{i+1}$ für $i = 0, \dots, n-1$ gilt. Sie heißt **akzeptierend**, falls $q_n \in E$ ist, und andernfalls verwerfend. Eine von einem DFA akzeptierte Sprache wird als **regulär** bezeichnet. Die zugehörige Sprachklasse ist

$$\text{REG} = \{L(M) \mid M \text{ ist ein DFA}\}.$$

Beispiel 5. Betrachte den DFA $M = (Z, \Sigma, \delta, 0, E)$ mit $Z = \{0, 1, 2\}$, $\Sigma = \{a, b\}$, $E = \{1\}$ und der Überföhrungsfunktion

δ	0	1	2
a	1	2	0
b	2	0	1

Graphische Darstellung:



Der Startzustand wird meist durch einen Pfeil und Endzustände werden durch einen doppelten Kreis gekennzeichnet.

Bei Eingabe $w_1 = aba$ führt M die akzeptierende Rechnung $0, 1, 0, 1$ durch, d.h. $w_1 \in L(M)$. Dagegen verwirft M das Wort $w_2 = abba$ (verwerfende Rechnung: $0, 1, 0, 2, 0$). \triangleleft

Bezeichne $\hat{\delta}(q, x)$ denjenigen Zustand, in dem sich M nach Lesen von x befindet, wenn M im Zustand q gestartet wird. Dann können wir die Funktion

$$\hat{\delta} : Z \times \Sigma^* \rightarrow Z$$

induktiv wie folgt definieren. Für $q \in Z$, $x \in \Sigma^*$ und $a \in \Sigma$ sei

$$\begin{aligned} \hat{\delta}(q, \varepsilon) &= q, \\ \hat{\delta}(q, xa) &= \delta(\hat{\delta}(q, x), a). \end{aligned}$$

Die von M erkannte Sprache lässt sich nun auch in der Form

$$L(M) = \{x \in \Sigma^* \mid \hat{\delta}(q_0, x) \in E\}$$

schreiben.

Behauptung 6. Der DFA M aus Beispiel 5 akzeptiert die Sprache

$$L(M) = \{x \in \Sigma^* \mid \#_a(x) - \#_b(x) \equiv_3 1\},$$

wobei $\#_a(x)$ die Anzahl der Vorkommen des Zeichens a in x bezeichnet und $i \equiv_m j$ (in Worten: i ist kongruent zu j modulo m) bedeutet, dass $i - j$ durch m teilbar ist.

Beweis. Da M nur den Endzustand 1 hat, ist $L(M) = \{x \in \Sigma^* \mid \hat{\delta}(0, x) = 1\}$, d.h. wir müssen folgende Äquivalenz zeigen:

$$\hat{\delta}(0, x) = 1 \Leftrightarrow \#_a(x) - \#_b(x) \equiv_3 1.$$

Hierzu reicht es, die Kongruenz

$$\hat{\delta}(0, x) \equiv_3 \#_a(x) - \#_b(x).$$

zu beweisen, wofür wir Induktion über die Länge n von x benutzen.

Induktionsanfang ($n = 0$): klar, da $\hat{\delta}(0, \varepsilon) = \#_a(\varepsilon) = \#_b(\varepsilon) = 0$ ist.

Induktionsschritt ($n \rightsquigarrow n + 1$): Sei $x = x_1 \dots x_{n+1}$ gegeben und sei $i = \hat{\delta}(0, x_1 \dots x_n)$. Nach IV gilt dann

$$i \equiv_3 \#_a(x_1 \dots x_n) - \#_b(x_1 \dots x_n).$$

Wegen $\delta(i, a) \equiv_3 i + 1$ und $\delta(i, b) \equiv_3 i - 1$ folgt daher

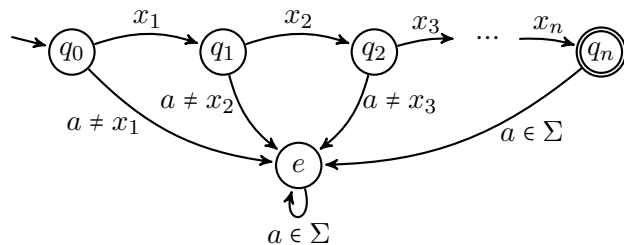
$$\begin{aligned} \delta(i, x_{n+1}) &\equiv_3 i + \#_a(x_{n+1}) - \#_b(x_{n+1}) \\ &\equiv_3 \#_a(x_1 \dots x_n) - \#_b(x_1 \dots x_n) + \#_a(x_{n+1}) - \#_b(x_{n+1}) \\ &= \#_a(x) - \#_b(x). \end{aligned}$$

und somit

$$\hat{\delta}(0, x) = \delta(\hat{\delta}(0, x_1 \dots x_n), x_{n+1}) = \delta(i, x_{n+1}) \equiv_3 \#_a(x) - \#_b(x).$$

Beobachtung 7. Alle Singletonsprachen sind regulär.

Beweis. Für jedes Wort $x = x_1 \dots x_n$ existiert ein DFA M_x mit $L(M_x) = \{x\}$:



Formal ist M_x also das Tupel $(Z, \Sigma, \delta, q_0, E)$ mit $Z = \{q_0, \dots, q_n, e\}$, $E = \{q_n\}$ und der Überföhrungsfunktion

$$\delta(q, a_j) = \begin{cases} q_{i+1}, & q = q_i \text{ für ein } i \text{ mit } 0 \leq i \leq n-1 \text{ und } a_j = x_{i+1} \\ e, & \text{sonst.} \end{cases}$$

Als nächstes betrachten wir Abschlusseigenschaften der Sprachklasse REG.

Definition 8. Ein **k-stelliger Sprachoperator** ist eine Abbildung op , die k Sprachen L_1, \dots, L_k auf eine Sprache $op(L_1, \dots, L_k)$ abbildet.

Beispiel 9. Der Schnittoperator \cap bildet zwei Sprachen L_1 und L_2 auf die Sprache $L_1 \cap L_2$ ab. \triangleleft

Definition 10. Eine Sprachklasse \mathcal{K} heißt unter op **abgeschlossen**, wenn gilt:

$$L_1, \dots, L_k \in \mathcal{K} \Rightarrow op(L_1, \dots, L_k) \in \mathcal{K}.$$

Der **Abschluss** von \mathcal{K} unter op ist die bzgl. Inklusion kleinste Sprachklasse \mathcal{K}' , die \mathcal{K} enthält und unter op abgeschlossen ist.

Beispiel 11. Der Abschluss der Singletonsprachen unter \cap besteht aus allen Singletonsprachen und der leeren Sprache.

Der Abschluss der Singletonsprachen unter \cup besteht aus allen nicht-leeren endlichen Sprachen. \triangleleft

Definition 12. Für eine Sprachklasse \mathcal{C} bezeichne $co\text{-}\mathcal{C}$ die Klasse $\{\bar{L} \mid L \in \mathcal{C}\}$ aller Komplemente von Sprachen in \mathcal{C} .

Es ist leicht zu sehen, dass \mathcal{C} genau dann unter Komplementbildung abgeschlossen ist, wenn $co\text{-}\mathcal{C} = \mathcal{C}$ ist.

Beobachtung 13. Mit $L_1, L_2 \in \text{REG}$ sind auch die Sprachen $\bar{L}_1 = \Sigma^* \setminus L_1$, $L_1 \cap L_2$ und $L_1 \cup L_2$ regulär.

Beweis. Sind $M_i = (Z_i, \Sigma, \delta_i, q_0, E_i)$, $i = 1, 2$, DFAs mit $L(M_i) = L_i$, so akzeptiert der DFA

$$\bar{M}_1 = (Z_1, \Sigma, \delta_1, q_0, Z_1 \setminus E_1)$$

das Komplement \bar{L}_1 von L_1 . Der Schnitt $L_1 \cap L_2$ von L_1 und L_2 wird dagegen von dem DFA

$$M = (Z_1 \times Z_2, \Sigma, \delta, (q_0, q_0), E_1 \times E_2)$$

mit

$$\delta((q, p), a) = (\delta_1(q, a), \delta_2(p, a))$$

akzeptiert (M wird auch **Kreuzproduktautomat** genannt). Wegen $L_1 \cup L_2 = \overline{\bar{L}_1 \cap \bar{L}_2}$ ist dann aber auch die Vereinigung von L_1 und L_2 regulär. (Wie sieht der zugehörige DFA aus?) \blacksquare

Aus Beobachtung 13 folgt, dass alle endlichen und alle co-endlichen Sprachen regulär sind. Da die in Beispiel 5 betrachtete Sprache weder endlich noch co-endlich ist, haben wir damit allerdings noch nicht alle regulären Sprachen erfasst.

Es stellt sich die Frage, ob REG neben den mengentheoretischen Operationen Schnitt, Vereinigung und Komplement unter weiteren Operationen wie etwa Produkt oder Sternhülle abgeschlossen ist. Im übernächsten Abschnitt werden wir sehen, dass die Klasse REG als der Abschluss der endlichen Sprachen unter Vereinigung, Produkt und Sternhülle charakterisierbar ist.

Beim Versuch, einen endlichen Automaten für das Produkt L_1L_2 zweier regulärer Sprachen zu konstruieren, stößt man auf die Schwierigkeit, den richtigen Zeitpunkt für den Übergang von (der Simulation von) M_1 zu M_2 zu finden. Unter Verwendung eines nichtdeterministischen Automaten lässt sich dieses Problem jedoch leicht beheben, da dieser den richtigen Zeitpunkt „erraten“ kann.

Im nächsten Abschnitt werden wir nachweisen, dass auch nichtdeterministische endliche Automaten nur reguläre Sprachen erkennen können.

2.2 Nichtdeterministische endliche Automaten

Definition 14. Ein **nichtdeterministischer endlicher Automat** (kurz: *NFA*; *nondeterministic finite automaton*) $N = (Z, \Sigma, \Delta, Q_0, E)$ ist ähnlich aufgebaut wie ein DFA, nur dass er mehrere Startzustände (zusammengefasst in der Menge $Q_0 \subseteq Z$) haben kann und seine Überföhrungsfunktion die Form

$$\Delta : Z \times \Sigma \rightarrow \mathcal{P}(Z)$$

hat. Hierbei bezeichnet $\mathcal{P}(Z)$ die **Potenzmenge** (also die Menge aller Teilmengen) von Z . Diese wird auch oft mit 2^Z bezeichnet. Die

von N akzeptierte Sprache ist

$$L(N) = \left\{ x_1 \dots x_n \in \Sigma^* \mid \begin{array}{l} \exists q_0 \in Q_0, q_1, \dots, q_{n-1} \in Z, q_n \in E: \\ q_{i+1} \in \Delta(q_i, x_{i+1}) \text{ f\u00fcr } i = 0, \dots, n-1 \end{array} \right\}.$$

Eine Zustandsfolge q_0, q_1, \dots, q_n hei\u00dft **Rechnung** von $N(x_1 \dots x_n)$, falls $q_{i+1} \in \Delta(q_i, x_{i+1})$ f\u00fcr $i = 0, \dots, n-1$ gilt.

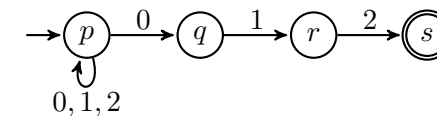
Ein NFA N kann bei einer Eingabe x also nicht nur eine, sondern mehrere verschiedene Rechnungen parallel ausf\u00fchren. Ein Wort x geh\u00f6rt genau dann zu $L(N)$, wenn $N(x)$ mindestens eine akzeptierende Rechnung hat.

Im Gegensatz zu einem DFA, dessen \u00cberf\u00f6hrungsfunktion auf der gesamten Menge $Z \times \Sigma$ definiert ist, kann ein NFA „stecken bleiben“. Das ist dann der Fall, wenn er in einen Zustand q gelangt, in dem das n\u00e4chste Eingabezeichen x_i wegen $\Delta(q, x_i) = \emptyset$ nicht gelesen werden kann.

Beispiel 15. Betrachte den NFA $N = (Z, \Sigma, \Delta, Q_0, E)$ mit Zustandsmenge $Z = \{p, q, r, s\}$, Eingabealphabet $\Sigma = \{0, 1, 2\}$, Start- und Endzustandsmenge $Q_0 = \{p\}$ und $E = \{s\}$ sowie der \u00cberf\u00f6hrungsfunktion

Δ	p	q	r	s
0	$\{p, q\}$	\emptyset	\emptyset	\emptyset
1	$\{p\}$	$\{r\}$	\emptyset	\emptyset
2	$\{p\}$	\emptyset	$\{s\}$	\emptyset

Graphische Darstellung:



Offensichtlich akzeptiert N die Sprache $L(N) = \{x012 \mid x \in \Sigma^*\}$ aller W\u00f6rter, die mit dem Suffix 012 enden. \triangleleft

Beobachtung 16. Sind $N_i = (Z_i, \Sigma, \Delta_i, Q_i, E_i)$ ($i = 1, 2$) NFAs, so werden auch die Sprachen $L(N_1)L(N_2)$ und $L(N_1)^*$ von einem NFA erkannt.

Beweis. Sei $L_i = L(N_i)$. Wir können $Z_1 \cap Z_2 = \emptyset$ annehmen. Dann akzeptiert der NFA

$$N = (Z_1 \cup Z_2, \Sigma, \Delta_3, Q_1, E)$$

mit

$$\Delta_3(p, a) = \begin{cases} \Delta_1(p, a), & p \in Z_1 \setminus E_1, \\ \Delta_1(p, a) \cup \bigcup_{q \in Q_2} \Delta_2(q, a), & p \in E_1, \\ \Delta_2(p, a), & \text{sonst} \end{cases}$$

und

$$E = \begin{cases} E_2, & Q_2 \cap E_2 = \emptyset \\ E_1 \cup E_2, & \text{sonst} \end{cases}$$

die Sprache $L_1 L_2$.

$L_1 L_2 \subseteq L(N)$: Seien $x = x_1 \dots x_k \in L_1$, $y = y_1 \dots y_l \in L_2$ und seien q_0, \dots, q_k und p_0, \dots, p_l akzeptierende Rechnungen von $N_1(x)$ und $N_2(y)$. Dann ist $q_0, \dots, q_k, p_1, \dots, p_l$ eine akz. Rechnung von $N(xy)$, da $q_0 \in Q_1$ und $p_l \in E_2$ ist, und

- im Fall $l \geq 1$ wegen $q_k \in E_1$, $p_0 \in Q_2$ und $p_1 \in \Delta_2(p_0, y_1)$ zudem $p_1 \in \Delta(q_k, y_1)$ und
- im Fall $l = 0$ wegen $q_k \in E_1$ und $p_l \in Q_2 \cap E_2$ zudem $q_k \in E$ ist.

$L(N) \subseteq L_1 L_2$: Sei $x = x_1 \dots x_n \in L(N)$ und sei q_0, \dots, q_n eine akz. Rechnung von $N(x)$. Dann gilt $q_0 \in Q_1$, $q_n \in E$, $q_0, \dots, q_i \in Z_1$ und $q_{i+1}, \dots, q_n \in Z_2$ für ein $i \leq n$. Wir zeigen, dass ein $q \in Q_2$ existiert, so dass q_0, \dots, q_i eine akz. Rechnung von $N_1(x_1 \dots x_i)$ und q, q_{i+1}, \dots, q_n eine akz. Rechnung von $N_2(x_{i+1} \dots x_n)$ ist.

- Im Fall $i < n$ impliziert der Übergang $q_{i+1} \in \Delta(q_i, x_{i+1})$, dass $q_i \in E_1$ und $q_{i+1} \in \Delta_2(q, x_{i+1})$ für ein $q \in Q_2$ ist.
- Im Fall $i = n$ ist $q_n \in E_1$ und $Q_2 \cap E_2 \neq \emptyset$ (d.h. $\varepsilon \in L_2$).

Ganz ähnlich lässt sich zeigen, dass der NFA

$$N^* = (Z_1 \cup \{q_{neu}\}, \Sigma, \Delta_4, Q_1 \cup \{q_{neu}\}, E_1 \cup \{q_{neu}\})$$

mit

$$\Delta_4(p, a) = \begin{cases} \Delta_1(p, a), & p \in Z_1 \setminus E_1, \\ \Delta_1(p, a) \cup \bigcup_{q \in Q_1} \Delta_1(q, a), & p \in E_1, \\ \emptyset, & \text{sonst} \end{cases}$$

die Sprache L_1^* akzeptiert. ■

Satz 17 (Rabin und Scott).

$\text{REG} = \{L(N) \mid N \text{ ist ein NFA}\}$.

Beweis. Die Inklusion von links nach rechts ist klar, da jeder DFA auch als NFA aufgefasst werden kann. Für die Gegenrichtung konstruieren wir zu einem NFA $N = (Z, \Sigma, \Delta, Q_0, E)$ einen DFA $M = (\mathcal{P}(Z), \Sigma, \delta, Q_0, E')$ mit $L(M) = L(N)$. Wir definieren die Überföhrungsfunktion $\delta : \mathcal{P}(Z) \times \Sigma \rightarrow \mathcal{P}(Z)$ von M mittels

$$\delta(Q, a) = \bigcup_{q \in Q} \Delta(q, a).$$

Die Menge $\delta(Q, a)$ enthält also alle Zustände, in die N gelangen kann, wenn N ausgehend von einem beliebigen Zustand $q \in Q$ das Zeichen a liest. Intuitiv bedeutet dies, dass der DFA M den NFA N simuliert, indem M in seinem aktuellen Zustand Q die Information speichert, in welchen Zuständen sich N momentan befinden könnte. Für die Erweiterung $\hat{\delta} : \mathcal{P}(Z) \times \Sigma^* \rightarrow \mathcal{P}(Z)$ von δ (siehe Seite 3) können wir nun folgende Behauptung zeigen.

Behauptung. $\hat{\delta}(Q_0, x)$ enthält alle Zustände, die N ausgehend von einem Startzustand nach Lesen von x erreichen kann.

Wir beweisen die Behauptung induktiv über die Länge n von x .

Induktionsanfang ($n = 0$): klar, da $\hat{\delta}(Q_0, \varepsilon) = Q_0$ ist.

Induktionsschritt ($n - 1 \rightsquigarrow n$): Sei $x = x_1 \dots x_n$ gegeben. Nach Induktionsvoraussetzung enthält

$$Q_{n-1} = \hat{\delta}(Q_0, x_1 \dots x_{n-1})$$

alle Zustände, die $N(x)$ in genau $n - 1$ Schritten erreichen kann.
Wegen

$$\hat{\delta}(Q_0, x) = \delta(Q_{n-1}, x_n) = \bigcup_{q \in Q_{n-1}} \Delta(q, x_n)$$

enthält dann aber $\hat{\delta}(Q_0, x)$ alle Zustände, die $N(x)$ in genau n Schritten erreichen kann.

Deklarieren wir nun diejenigen Teilmengen $Q \subseteq Z$, die mindestens einen Endzustand von N enthalten, als Endzustände des **Potenzmengenautomaten** M , d.h.

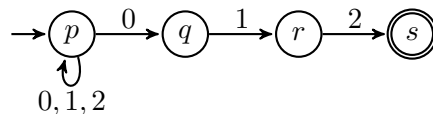
$$E' = \{Q \subseteq Z \mid Q \cap E \neq \emptyset\},$$

so folgt für alle Wörter $x \in \Sigma^*$:

- $x \in L(N) \Leftrightarrow N(x)$ kann in genau $|x|$ Schritten einen Endzustand erreichen
- $\Leftrightarrow \hat{\delta}(Q_0, x) \cap E \neq \emptyset$
- $\Leftrightarrow \hat{\delta}(Q_0, x) \in E'$
- $\Leftrightarrow x \in L(M)$.

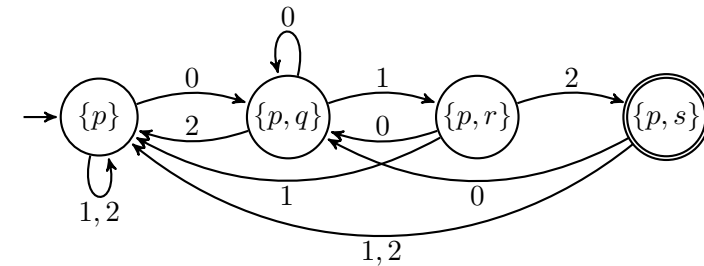
■

Beispiel 18. Für den NFA $N = (Z, \Sigma, \Delta, Q_0, E)$ aus Beispiel 15



ergibt die Konstruktion des vorigen Satzes den folgenden DFA M (nach Entfernen aller vom Startzustand $Q_0 = \{p\}$ aus nicht erreichbaren Zustände):

δ	0	1	2
$Q_0 = \{p\}$	$\{p, q\}$	$\{p\}$	$\{p\}$
$Q_1 = \{p, q\}$	$\{p, q\}$	$\{p, r\}$	$\{p\}$
$Q_2 = \{p, r\}$	$\{p, q\}$	$\{p\}$	$\{p, s\}$
$Q_3 = \{p, s\}$	$\{p, q\}$	$\{p\}$	$\{p\}$



◁

Im obigen Beispiel wurden für die Konstruktion des DFA M aus dem NFA N nur 4 der insgesamt $2^{|Z|} = 16$ Zustände benötigt, da die übrigen 12 Zustände in $\mathcal{P}(Z)$ nicht vom Startzustand $Q_0 = \{p\}$ aus erreichbar sind. Es gibt jedoch Beispiele, bei denen alle $2^{|Z|}$ Zustände in $\mathcal{P}(Z)$ für die Konstruktion des Potenzmengenautomaten benötigt werden (siehe Übungen).

Korollar 19. Die Klasse REG der regulären Sprachen ist unter folgenden Operationen abgeschlossen:

- Komplement,
- Produkt,
- Schnitt,
- Sternhülle.
- Vereinigung,

2.3 Reguläre Ausdrücke

Wir haben uns im letzten Abschnitt davon überzeugt, dass auch NFAs nur reguläre Sprachen erkennen können:

$$\text{REG} = \{L(M) \mid M \text{ ist ein DFA}\} = \{L(N) \mid N \text{ ist ein NFA}\}.$$

In diesem Abschnitt werden wir eine weitere Charakterisierung der regulären Sprachen kennenlernen:

REG ist die Klasse aller Sprachen, die sich mittels der Operationen Vereinigung, Schnitt, Komplement, Produkt und Sternhülle aus der leeren Menge und den Singletonsprachen bilden lassen.

Tatsächlich kann hierbei sogar auf die Schnitt- und Komplementbildung verzichtet werden.

Definition 20. Die Menge der **regulären Ausdrücke** γ (über einem Alphabet Σ) und die durch γ dargestellte Sprache $L(\gamma)$ sind induktiv wie folgt definiert. Die Symbole \emptyset , ϵ und a ($a \in \Sigma$) sind reguläre Ausdrücke, die

- die leere Sprache $L(\emptyset) = \emptyset$,
- die Sprache $L(\epsilon) = \{\epsilon\}$ und
- für jedes Zeichen $a \in \Sigma$ die Sprache $L(a) = \{a\}$

beschreiben. Sind α und β reguläre Ausdrücke, die die Sprachen $L(\alpha)$ und $L(\beta)$ beschreiben, so sind auch $\alpha\beta$, $(\alpha|\beta)$ und $(\alpha)^*$ reguläre Ausdrücke, die die Sprachen

- $L(\alpha\beta) = L(\alpha)L(\beta)$,
- $L(\alpha|\beta) = L(\alpha) \cup L(\beta)$ und
- $L((\alpha)^*) = L(\alpha)^*$

beschreiben.

Bemerkung 21.

- Um Klammern zu sparen, definieren wir folgende **Präzedenzordnung**: Der Sternoperator $*$ bindet stärker als der Produktoperator und dieser wiederum stärker als der Vereinigungsoperator. Für $((a|b(c)^*)|d)$ können wir also kurz $a|bc^*|d$ schreiben.
- Da der reguläre Ausdruck $\gamma\gamma^*$ die Sprache $L(\gamma)^+$ beschreibt, verwenden wir γ^+ als Abkürzung für den Ausdruck $\gamma\gamma^*$.

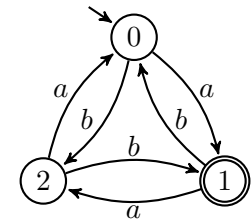
Beispiel 22. Die regulären Ausdrücke ϵ^* , \emptyset^* , $(0|1)^*00$ und $\epsilon 0|\emptyset 1^*$ beschreiben folgende Sprachen:

γ	ϵ^*	\emptyset^*	$(0 1)^*00$	$\epsilon 0 \emptyset 1^*$
$L(\gamma)$	$\{\epsilon\}^* = \{\epsilon\}$	$\emptyset^* = \{\epsilon\}$	$\{x00 \mid x \in \{0,1\}^*\}$	$\{0\}$

◁

Beispiel 23. Betrachte nebenstehenden DFA M . Um für die von M erkannte Sprache

$$L(M) = \{x \in \{a,b\}^* \mid \#_a(x) - \#_b(x) \equiv_3 1\}$$



einen regulären Ausdruck zu finden, betrachten wir zunächst die Sprache $L_{0,0}$ aller Wörter x , die den DFA M ausgehend vom Zustand 0 in den Zustand 0 überführen. Weiter sei $L_{0,0}^{\neq 0}$ die Sprache aller solchen Wörter $w \neq \epsilon$, die zwischendurch nicht den Zustand 0 besuchen. Dann setzt sich jedes $x \in L_{0,0}$ aus beliebig vielen Teilwörtern $w_1, \dots, w_k \in L_{0,0}^{\neq 0}$ zusammen, d.h. $L_{0,0} = (L_{0,0}^{\neq 0})^*$.

Jedes $w \in L_{0,0}^{\neq 0}$ beginnt entweder mit einem a (Übergang von 0 nach 1) oder mit einem b (Übergang von 0 nach 2). Im ersten Fall folgt eine beliebige Anzahl von Teilwörtern ab (Wechsel zwischen 1 und 2), an die sich entweder das Suffix aa (Rückkehr von 1 nach 0 über 2) oder das Suffix b (direkte Rückkehr von 1 nach 0) anschließt. Analog folgt im zweiten Fall eine beliebige Anzahl von Teilwörtern ba (Wechsel zwischen 2 und 1), an die sich entweder das Suffix a (direkte Rückkehr von 2 nach 0) oder das Suffix bb (Rückkehr von 2 nach 0 über 1) anschließt. Daher lässt sich $L_{0,0}^{\neq 0}$ durch den regulären Ausdruck

$$\gamma_{0,0}^{\neq 0} = a(ab)^*(aa|b) \mid b(ba)^*(a|bb) \mid \epsilon$$

beschreiben. Eine ähnliche Überlegung zeigt, dass sich die Sprache $L_{0,1}^{\neq 0}$ aller Wörter, die M ausgehend von 0 in den Zustand 1 überführen, ohne dass zwischendurch der Zustand 0 nochmals besucht

wird, durch den regulären Ausdruck $\gamma_{0,1}^{\neq 0} = (a|bb)(ab)^*$ beschreibbar ist. Somit erhalten wir für $L(M)$ den regulären Ausdruck

$$\gamma_{0,1} = (a(ab)^*(aa|b) | b(ba)^*(a|bb))^*(a|bb)(ab)^*.$$

◁

Satz 24. $\{L(\gamma) \mid \gamma \text{ ist ein regulärer Ausdruck}\} = \text{REG}.$

Beweis. Die Inklusion von rechts nach links ist klar, da die Basisausdrücke \emptyset , ϵ und a , $a \in \Sigma^*$, nur reguläre Sprachen beschreiben und die Sprachklasse REG unter Produkt, Vereinigung und Sternhülle abgeschlossen ist (siehe Beobachtungen 13 und 16).

Für die Gegenrichtung konstruieren wir zu einem DFA M einen regulären Ausdruck γ mit $L(\gamma) = L(M)$. Sei also $M = (Z, \Sigma, \delta, q_0, E)$ ein DFA, wobei wir annehmen können, dass $Z = \{1, \dots, m\}$ und $q_0 = 1$ ist. Dann lässt sich $L(M)$ als Vereinigung

$$L(M) = \bigcup_{q \in E} L_{1,q}$$

von Sprachen der Form

$$L_{p,q} = \{x \in \Sigma^* \mid \hat{\delta}(p, x) = q\}$$

darstellen. Folglich reicht es zu zeigen, dass die Sprachen $L_{p,q}$ durch reguläre Ausdrücke beschreibbar sind. Hierzu betrachten wir die Sprachen

$$L_{p,q}^r = \left\{ x_1 \dots x_n \in \Sigma^* \mid \begin{array}{l} \hat{\delta}(p, x_1 \dots x_n) = q \text{ und für} \\ i = 1, \dots, n-1 \text{ gilt } \hat{\delta}(p, x_1 \dots x_i) \leq r \end{array} \right\}.$$

Wegen $L_{p,q} = L_{p,q}^m$ reicht es, reguläre Ausdrücke $\gamma_{p,q}^r$ für die Sprachen $L_{p,q}^r$ anzugeben. Im Fall $r = 0$ enthält

$$L_{p,q}^0 = \begin{cases} \{a \in \Sigma \mid \delta(p, a) = q\} \cup \{\epsilon\}, & p = q, \\ \{a \in \Sigma \mid \delta(p, a) = q\}, & \text{sonst} \end{cases}$$

nur Buchstaben (und eventuell das leere Wort) und ist somit leicht durch einen regulären Ausdruck $\gamma_{p,q}^0$ beschreibbar. Wegen

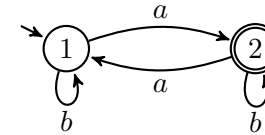
$$L_{p,q}^{r+1} = L_{p,q}^r \cup L_{p,r+1}^r (L_{r+1,r+1}^r)^* L_{r+1,q}^r$$

lassen sich aus den regulären Ausdrücken $\gamma_{p,q}^r$ für die Sprachen $L_{p,q}^r$ leicht reguläre Ausdrücke für die Sprachen $L_{p,q}^{r+1}$ gewinnen:

$$\gamma_{p,q}^{r+1} = \gamma_{p,q}^r \mid \gamma_{p,r+1}^r (\gamma_{r+1,r+1}^r)^* \gamma_{r+1,q}^r.$$

■

Beispiel 25. Betrachte den DFA



Da M insgesamt $m = 2$ Zustände und nur den Endzustand 2 besitzt, ist

$$L(M) = \bigcup_{q \in E} L_{1,q} = L_{1,2} = L_{1,2}^2 = L(\gamma_{1,2}^2).$$

Um $\gamma_{1,2}^2$ zu berechnen, benutzen wir die Rekursionsformel

$$\gamma_{p,q}^{r+1} = \gamma_{p,q}^r \mid \gamma_{p,r+1}^r (\gamma_{r+1,r+1}^r)^* \gamma_{r+1,q}^r$$

und erhalten

$$\begin{aligned} \gamma_{1,2}^2 &= \gamma_{1,2}^1 \mid \gamma_{1,2}^1 (\gamma_{2,2}^1)^* \gamma_{2,2}^1, \\ \gamma_{1,2}^1 &= \gamma_{1,2}^0 \mid \gamma_{1,1}^0 (\gamma_{1,1}^0)^* \gamma_{1,2}^0, \\ \gamma_{2,2}^1 &= \gamma_{2,2}^0 \mid \gamma_{2,1}^0 (\gamma_{1,1}^0)^* \gamma_{1,2}^0. \end{aligned}$$

Um den regulären Ausdruck $\gamma_{1,2}^2$ für $L(M)$ zu erhalten, genügt es also, die regulären Ausdrücke $\gamma_{1,1}^0, \gamma_{1,2}^0, \gamma_{2,1}^0, \gamma_{2,2}^0, \gamma_{1,2}^1$ und $\gamma_{2,2}^1$ zu berechnen:

r	p, q			
	1, 1	1, 2	2, 1	2, 2
0	ϵb	a	a	ϵb
1	-	$\underbrace{a (\epsilon b)(\epsilon b)^*a}_{b^*a}$	-	$\underbrace{(\epsilon b)a(\epsilon b)^*a}_{\epsilon b ab^*a}$
2	-	$\underbrace{b^*a b^*a(\epsilon b ab^*a)^*(\epsilon b ab^*a)}_{b^*a(b ab^*a)^*}$	-	-

◁

Korollar 26. Sei L eine Sprache. Dann sind folgende Aussagen äquivalent:

- L ist regulär (d.h. es gibt einen DFA M mit $L = L(M)$),
- es gibt einen NFA N mit $L = L(N)$,
- es gibt einen regulären Ausdruck γ mit $L = L(\gamma)$,
- L lässt sich mit den Operationen Vereinigung, Produkt und Sternhülle aus endlichen Sprachen gewinnen,
- L lässt sich mit den Operationen \cap, \cup , Komplement, Produkt und Sternhülle aus endlichen Sprachen gewinnen.

Wir werden bald noch eine weitere Charakterisierung von REG kennenlernen, nämlich durch reguläre Grammatiken. Zuvor befassen wir uns jedoch mit dem Problem, DFAs zu minimieren. Dabei spielen Relationen (insbesondere Äquivalenzrelationen) eine wichtige Rolle.

2.4 Relationalstrukturen

Sei A eine nichtleere Menge, R_i eine k_i -stellige Relation auf A , d.h. $R_i \subseteq A^{k_i}$ für $i = 1, \dots, n$. Dann heißt $(A; R_1, \dots, R_n)$ **Relationalstruktur**. Die Menge A heißt **Grundmenge**, **Trägermenge** oder **Individuenbereich** der Relationalstruktur.

Wir werden hier hauptsächlich den Fall $n = 1, k_1 = 2$, also (A, R) mit $R \subseteq A \times A$ betrachten. Man nennt dann R eine (**binäre**) **Relation** auf A . Oft wird für $(a, b) \in R$ auch die **Infix-Schreibweise** aRb benutzt.

Beispiel 27.

- (F, M) mit $F = \{f \mid f \text{ ist Fluss in Europa}\}$ und $M = \{(f, g) \in F \times F \mid f \text{ mündet in } g\}$.
- (U, B) mit $U = \{x \mid x \text{ ist Berliner}\}$ und $B = \{(x, y) \in U \times U \mid x \text{ ist Bruder von } y\}$.
- $(P(M), \subseteq)$, wobei $P(M)$ die Potenzmenge einer beliebigen Menge M und \subseteq die Inklusionsbeziehung auf den Teilmengen von M ist.
- (A, Id_A) , wobei $Id_A = \{(x, x) \mid x \in A\}$ die **Identität auf A** ist.
- (\mathbb{R}, \leq) .
- $(\mathbb{Z}, |)$, wobei $|$ die "teilt"-Relation bezeichnet (d.h. $a|b$, falls ein $c \in \mathbb{Z}$ mit $b = ac$ existiert). ◁

Da Relationen Mengen sind, sind auf ihnen die mengentheoretischen Operationen **Schnitt**, **Vereinigung**, **Komplement** und **Differenz** definiert. Seien R und S Relationen auf A , dann ist

$$\begin{aligned}
 R \cap S &= \{(x, y) \in A \times A \mid xRy \wedge xSy\}, \\
 R \cup S &= \{(x, y) \in A \times A \mid xRy \vee xSy\}, \\
 R - S &= \{(x, y) \in A \times A \mid xRy \wedge \neg xSy\}, \\
 \overline{R} &= (A \times A) - R.
 \end{aligned}$$

Sei allgemeiner $\mathcal{M} \subseteq \mathcal{P}(A \times A)$ eine beliebige Menge von Relationen auf A . Dann sind der **Schnitt über \mathcal{M}** und die **Vereinigung über \mathcal{M}** folgende Relationen:

$$\bigcap \mathcal{M} = \bigcap_{R \in \mathcal{M}} R = \{(x, y) \mid \forall R \in \mathcal{M} : xRy\},$$

$$\bigcup \mathcal{M} = \bigcup_{R \in \mathcal{M}} R = \{(x, y) \mid \exists R \in \mathcal{M} : xRy\}.$$

Die **transponierte (konverse) Relation** zu R ist

$$R^T = \{(y, x) \mid xRy\}.$$

R^T wird oft auch mit R^{-1} bezeichnet. Z.B. ist $(\mathbb{R}, \leq^T) = (\mathbb{R}, \geq)$.

Seien R und S Relationen auf A . Das **Produkt** oder die **Komposition** von R und S ist

$$R \circ S = \{(x, z) \in A \times A \mid \exists y \in A : xRy \wedge ySz\}.$$

Beispiel 28. Ist B die Relation "ist Bruder von", V "ist Vater von", M "ist Mutter von" und $E = V \cup M$ "ist Elternteil von", so ist $B \circ E$ die Onkel-Relation. ◁

Übliche Bezeichnungen für das Relationenprodukt sind auch $R;S$ und $R \cdot S$ oder einfach RS . Das n -fache Relationenprodukt $R \circ \dots \circ R$ von R wird mit R^n bezeichnet. Dabei ist $R^0 = Id$.

Vorsicht: Das n -fache Relationenprodukt R^n von R sollte nicht mit dem n -fachen kartesischen Produkt $R \times \dots \times R$ der Menge R verwechselt werden. Wir vereinbaren, dass R^n das n -fache Relationenprodukt bezeichnen soll, falls R eine Relation ist.

Eigenschaften von Relationen

Sei R eine Relation auf A . Dann heißt R

reflexiv,	falls $\forall x \in A : xRx$	(also $Id_A \subseteq R$)
irreflexiv,	falls $\forall x \in A : \neg xRx$	(also $Id_A \subseteq \overline{R}$)
symmetrisch,	falls $\forall x, y \in A : xRy \Rightarrow yRx$	(also $R \subseteq R^T$)
asymmetrisch,	falls $\forall x, y \in A : xRy \Rightarrow \neg yRx$	(also $R \subseteq \overline{R^T}$)
antisymmetrisch,	falls $\forall x, y \in A : xRy \wedge yRx \Rightarrow x = y$	(also $R \cap R^T \subseteq Id$)
konnex,	falls $\forall x, y \in A : xRy \vee yRx$	(also $A \times A \subseteq R \cup R^T$)
semikonnex,	falls $\forall x, y \in A : x \neq y \Rightarrow xRy \vee yRx$	(also $\overline{Id} \subseteq R \cup R^T$)
transitiv,	falls $\forall x, y, z \in A : xRy \wedge yRz \Rightarrow xRz$	(also $R^2 \subseteq R$)

gilt.

Die nachfolgende Tabelle gibt einen Überblick über die wichtigsten Relationalstrukturen.

	refl.	sym.	trans.	antisym.	asym.	konnex	semikon.
Äquivalenzrelation	✓	✓	✓				
(Halb-)Ordnung	✓		✓	✓			
Striktordnung			✓		✓		
lineare Ordnung			✓	✓		✓	
lin. Striktord.			✓		✓		✓
Quasiordnung	✓		✓				

In der Tabelle sind nur die definierenden Eigenschaften durch ein "✓" gekennzeichnet. Das schließt nicht aus, dass gleichzeitig auch noch weitere Eigenschaften vorliegen können.

Beispiel 29.

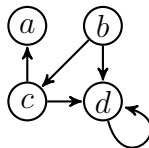
- Die Relation "ist Schwester von" ist zwar in einer reinen Damengesellschaft symmetrisch, i.a. jedoch weder symmetrisch noch asymmetrisch noch antisymmetrisch.

- Die Relation "ist Geschwister von" ist zwar symmetrisch, aber weder reflexiv noch transitiv und somit keine Äquivalenzrelation.
- $(\mathbb{R}, <)$ ist irreflexiv, asymmetrisch, transitiv und semikonnex und somit eine lineare Striktordnung.
- (\mathbb{R}, \leq) und $(P(M), \subseteq)$ sind reflexiv, antisymmetrisch und transitiv und somit Ordnungen.
- (\mathbb{R}, \leq) ist auch konnex und somit eine lineare Ordnung.
- $(P(M), \subseteq)$ ist zwar im Fall $\|M\| \leq 1$ konnex, aber im Fall $\|M\| \geq 2$ weder semikonnex noch konnex. \triangleleft

Graphische Darstellung von Relationen

Eine Relation R auf einer endlichen Menge A kann durch einen **gerichteten Graphen** (oder **Digraphen**) $G = (V, E)$ mit **Knotenmenge** $V = A$ und **Kantenmenge** $E = R$ veranschaulicht werden. Hierzu stellen wir jedes Element $x \in A$ als einen Knoten dar und verbinden jedes Knotenpaar $(x, y) \in R$ durch eine gerichtete Kante (Pfeil). Zwei durch eine Kante verbundene Knoten heißen **benachbart** oder **adjazent**.

Beispiel 30. Für die Relation (A, R) mit $A = \{a, b, c, d\}$ und $R = \{(b, c), (b, d), (c, a), (c, d), (d, d)\}$ erhalten wir folgende graphische Darstellung.



\triangleleft

Der **Ausgangsgrad** eines Knotens $x \in V$ ist $\deg^+(x) = \|R[x]\|$, wobei $R[x] = \{y \in V \mid xRy\}$ die Menge der **Nachfolger** von x ist. Entsprechend ist $\deg^-(x) = \|\{y \in V \mid yRx\}\|$ der **Eingangsgrad** von x und

$R^{-1}[x] = \{y \in V \mid yRx\}$ die Menge der **Vorgänger** von x . Falls R symmetrisch ist, werden die Pfeilspitzen meist weggelassen. In diesem Fall ist $d(x) = \deg^-(x) = \deg^+(x)$ der **Grad** von x und $R[x] = R^{-1}[x]$ heißt die **Nachbarschaft** von x . Ist R zudem irreflexiv, so ist G **schleifenfrei** und wir erhalten einen (**ungerichteten**) **Graphen**. Eine irreflexive und symmetrische Relation R wird meist als Menge der ungeordneten Paare $E = \{\{a, b\} \mid aRb\}$ notiert.

Darstellung durch Adjazenzmatrizen

Eine Relation R auf einer endlichen (geordneten) Menge $A = \{a_1, \dots, a_n\}$ lässt sich durch eine boolesche $n \times n$ -Matrix $M_R = (m_{ij})$ mit

$$m_{ij} := \begin{cases} 1, & a_i R a_j, \\ 0, & \text{sonst} \end{cases}$$

darstellen. Beispielsweise hat die Relation

$$R = \{(b, c), (b, d), (c, a), (c, d), (d, d)\}$$

auf der Menge $A = \{a, b, c, d\}$ die Matrixdarstellung

$$M_R = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Darstellung durch Adjazenzlisten

Eine weitere Möglichkeit besteht darin, eine endliche Relation R in Form einer Tabelle darzustellen, die jedem Element $x \in A$ seine Nachfolger in Form einer Liste zuordnet. Für obige Relation R erhalten wir folgende Listen:

$x:$	$R[x]$
$a:$	-
$b:$	c, d
$c:$	a, d
$d:$	d

Sind $M_R = (r_{ij})$ und $M_S = (s_{ij})$ boolesche $n \times n$ -Matrizen für R und S , so erhalten wir für $T = R \circ S$ die Matrix $M_T = (t_{ij})$ mit

$$t_{ij} = \bigvee_{k=1, \dots, n} (r_{ik} \wedge s_{kj})$$

Die Nachfolgermenge $T[x]$ von x bzgl. der Relation $T = R \circ S$ berechnet sich zu

$$T[x] = \bigcup \{S[y] \mid y \in R[x]\} = \bigcup_{y \in R[x]} S[y].$$

Beispiel 31. Betrachte die Relationen $R = \{(a, a), (a, c), (c, b), (c, d)\}$ und $S = \{(a, b), (d, a), (d, c)\}$ auf der Menge $A = \{a, b, c, d\}$.

Relation	R	S	$R \circ S$	$S \circ R$
Digraph				
Adjazenzmatrix	1 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0	0 1 0 0 0 0 0 0 0 0 0 0 1 0 1 0	0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1
Adjazenzliste	$a: a, c$ $b: -$ $c: b, d$ $d: -$	$a: b$ $b: -$ $c: -$ $d: a, c$	$a: b$ $b: -$ $c: a, c$ $d: -$	$a: -$ $b: -$ $c: -$ $d: a, b, c, d$

◁

Beobachtung: Das Beispiel zeigt, dass das Relationenprodukt nicht kommutativ ist, d.h. i.a. gilt nicht $R \circ S = S \circ R$.

Manchmal steht man vor der Aufgabe, eine gegebene Relation R durch eine möglichst kleine Modifikation in eine Relation R' mit vorgegebenen Eigenschaften zu überführen. Will man dabei alle in R enthaltenen Paare beibehalten, dann sollte R' aus R durch Hinzufügen möglichst weniger Paare hervorgehen.

Es lässt sich leicht nachprüfen, dass der Schnitt über eine Menge reflexiver (bzw. transitiver oder symmetrischer) Relationen wieder reflexiv (bzw. transitiv oder symmetrisch) ist. Folglich existiert zu jeder Relation R auf einer Menge A eine kleinste reflexive (bzw. transitive oder symmetrische) Relation R' , die R enthält.

Definition 32. Sei R eine Relation auf A .

- Die **reflexive Hülle** von R ist

$$h_{refl}(R) = \bigcap \{S \subseteq A \times A \mid S \text{ ist reflexiv und } R \subseteq S\}.$$

- Die **symmetrische Hülle** von R ist

$$h_{sym}(R) = \bigcap \{S \subseteq A \times A \mid S \text{ ist symmetrisch und } R \subseteq S\}.$$

- Die **transitive Hülle** von R ist

$$R^+ = \bigcap \{S \subseteq A \times A \mid S \text{ ist transitiv und } R \subseteq S\}.$$

- Die **reflexiv-transitive Hülle** von R ist

$$R^* = \bigcap \{S \subseteq A \times A \mid S \text{ ist reflexiv, transitiv und } R \subseteq S\}.$$

- Die **Äquivalenzhülle** von R ist

$$h_{\ddot{a}q}(R) = \bigcap \{S \mid S \text{ ist eine \ddot{A}quivalenzrelation auf } A \text{ und } R \subseteq S\}.$$

Satz 33. Sei R eine Relation auf A .

- (i) $h_{\text{refl}}(R) = R \cup Id_A$,
- (ii) $h_{\text{sym}}(R) = R \cup R^T$,
- (iii) $R^+ = \bigcup_{n \geq 1} R^n$,
- (iv) $R^* = \bigcup_{n \geq 0} R^n$,
- (v) $h_{\text{äq}}(R) = (R \cup R^T)^*$.

Beweis. Siehe Übungen. ■

Anschaulich besagt der vorhergehende Satz, dass ein Paar (a, b) genau dann in der reflexiv-transitiven Hülle R^* von R ist, wenn es ein $n \geq 0$ gibt mit $aR^n b$, d.h. es gibt Elemente $x_0, \dots, x_n \in A$ mit $x_0 = a$, $x_n = b$ und

$$x_0 R x_1 R x_2 \dots x_{n-1} R x_n.$$

In der Graphentheorie nennt man x_0, \dots, x_n einen **Weg** der Länge n von a nach b . Ein Digraph G heißt **zusammenhängend**, wenn es für je zwei Knoten a und b einen Weg von a nach b oder einen Weg von b nach a gibt. G heißt **stark zusammenhängend**, wenn es von jedem Knoten a einen Weg zu jedem Knoten b in G gibt.

2.4.1 Ordnungs- und Äquivalenzrelationen

Wir betrachten zunächst Äquivalenzrelationen, die durch die drei Eigenschaften reflexiv, symmetrisch und transitiv definiert sind.

Ist E eine Äquivalenzrelation, so nennt man die Nachbarschaft $E[x]$ die **von x repräsentierte Äquivalenzklasse** und bezeichnet sie mit $[x]_E$ oder einfach mit $[x]$. Eine Menge $S \subseteq A$ heißt **Repräsentantensystem**, falls sie genau ein Element aus jeder Äquivalenzklasse enthält.

Beispiel 34.

- Auf der Menge aller Geraden im \mathbb{R}^2 die Parallelität. Offenbar bilden alle Geraden mit derselben Richtung (oder Steigung)

jeweils eine Äquivalenzklasse. Daher wird ein Repräsentantensystem beispielsweise durch die Menge aller Ursprungsgeraden gebildet.

- Auf der Menge aller Menschen "im gleichen Jahr geboren wie". Hier bildet jeder Jahrgang eine Äquivalenzklasse.
- Auf \mathbb{Z} die Relation "gleicher Rest bei Division durch m ". Die zugehörigen Äquivalenzklassen sind

$$[r] = \{a \in \mathbb{Z} \mid a \equiv_m r\}, \quad r = 0, 1, \dots, m-1.$$

Ein Repräsentantensystem wird beispielsweise durch die Reste $0, 1, \dots, m-1$ gebildet. ◁

Die (bzgl. Inklusion) kleinste Äquivalenzrelation auf A ist die **Identität** Id_A , die größte die **Allrelation** $A \times A$. Die Äquivalenzklassen der Identität enthalten jeweils nur ein Element, d.h. $[x]_{Id_A} = \{x\}$ für alle $x \in A$, und die Allrelation erzeugt nur eine Äquivalenzklasse, nämlich $[x]_{A \times A} = A$ für jedes $x \in A$. Die Identität Id_A hat nur ein Repräsentantensystem, nämlich A . Dagegen kann jede Singletonmenge $\{x\}$ mit $x \in A$ als Repräsentantensystem für die Allrelation $A \times A$ fungieren.

Definition 35. Eine Familie $\{B_i \mid i \in I\}$ von nichtleeren Teilmengen $B_i \subseteq A$ heißt **Partition** der Menge A , falls gilt:

- a) die Mengen B_i **überdecken** A , d.h. $A = \bigcup_{i \in I} B_i$ und
- b) die Mengen B_i sind **paarweise disjunkt**, d.h. für je zwei verschiedene Mengen $B_i \neq B_j$ gilt $B_i \cap B_j = \emptyset$.

Wie der nächste Satz zeigt, bilden die Äquivalenzklassen einer Äquivalenzrelation E eine Partition $\{[x] \mid x \in A\}$ von A . Diese Partition wird auch **Quotienten-** oder **Faktormenge** genannt und mit A/E bezeichnet. Die Anzahl der Äquivalenzklassen von E wird auch als der **Index** von E bezeichnet.

Für zwei Äquivalenzrelationen $E \subseteq E'$ sind auch die Äquivalenzklassen $[x]_E$ von E in den Klassen $[x]_{E'}$ von E' enthalten. Folglich ist

jede Äquivalenzklasse von E' die Vereinigung von (evtl. mehreren) Äquivalenzklassen von E . E bewirkt also eine **feinere** Partitionierung als E' . Demnach ist die Identität die **feinste** und die Allrelation die **größte** Äquivalenzrelation.

Satz 36. Sei E eine Relation auf A . Dann sind folgende Aussagen äquivalent.

- (i) E ist eine Äquivalenzrelation auf A .
- (ii) Es gibt eine Partition $\{B_i \mid i \in I\}$ von A mit

$$xEy \Leftrightarrow \exists i \in I : x, y \in B_i.$$

Beweis.

(i) \Rightarrow (ii) Sei E eine Äquivalenzrelation auf A . Wir zeigen, dass dann $\{E[x] \mid x \in A\}$ eine Partition von A mit der gewünschten Zusatzeigenschaft bildet:

Da E reflexiv ist, gilt xEx und somit $x \in E[x]$, d.h. $A = \bigcup_{x \in A} E[x]$.

Ist $E[x] \cap E[y] \neq \emptyset$ und $u \in E[x] \cap E[y]$, so folgt $E[x] = E[y]$:

$$z \in E[x] \Leftrightarrow xEz \stackrel{xEu}{\Leftrightarrow} uEz \stackrel{yEu}{\Leftrightarrow} yEz \Leftrightarrow z \in E[y]$$

Zudem gilt

$$\begin{aligned} \exists z \in A : x, y \in E[z] &\Leftrightarrow \exists z : z \in E[x] \cap E[y] \\ &\Leftrightarrow E[x] = E[y] \stackrel{y \in E[y]}{\Leftrightarrow} xEy \end{aligned}$$

(ii) \Rightarrow (i) Existiert umgekehrt eine Partition $\{B_i \mid i \in I\}$ von A mit $xEy \Leftrightarrow \exists i \in I : x, y \in B_i$, so ist E

- reflexiv, da zu jedem $x \in A$ eine Menge B_i mit $x \in B_i$ existiert,
- symmetrisch, da aus $x, y \in B_i$ auch $y, x \in B_i$ folgt, und

- transitiv, da aus $x, y \in B_i$ und $y, z \in B_j$ wegen $y \in B_i \cap B_j$ die Gleichheit $B_i = B_j$ und somit $x, z \in B_i$ folgt. ■

Als nächstes betrachten wir Ordnungsrelationen, die durch die drei Eigenschaften reflexiv, antisymmetrisch und transitiv definiert sind.

Beispiel 37.

- $(\mathcal{P}(M), \subseteq)$, (\mathbb{Z}, \leq) , (\mathbb{R}, \leq) und $(\mathbb{N}, |)$ sind Ordnungen. $(\mathbb{Z}, |)$ ist keine Ordnung, aber eine Quasiordnung.
- Für jede Menge M ist die relationale Struktur $(\mathcal{P}(M); \subseteq)$ eine Ordnung. Diese ist nur im Fall $\|M\| \leq 1$ linear.
- Ist R eine Relation auf A und $B \subseteq A$, so ist $R_B = R \cap (B \times B)$ die Einschränkung von R auf B .
- Einschränkungen von (linearen) Ordnungen sind ebenfalls (lineare) Ordnungen.
- Beispielsweise ist (\mathbb{Q}, \leq) die Einschränkung von (\mathbb{R}, \leq) auf \mathbb{Q} und $(\mathbb{N}, |)$ die Einschränkung von $(\mathbb{Z}, |)$ auf \mathbb{N} . ◁

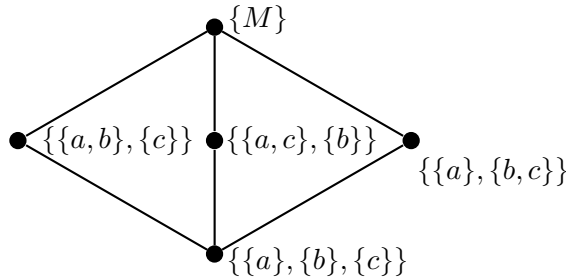
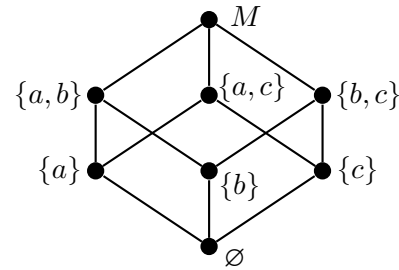
Ordnungen lassen sich sehr anschaulich durch Hasse-Diagramme darstellen. Sei \leq eine Ordnung auf A und sei $<$ die Relation $\leq \cap \overline{Id}_A$. Um die Ordnung \leq in einem **Hasse-Diagramm** darzustellen, wird nur der Graph der Relation

$$< = < \setminus <^2, \text{ d.h. } x < y \Leftrightarrow x < y \wedge \neg \exists z : x < z < y$$

gezeichnet. Für $x < y$ sagt man auch, y ist **oberer Nachbar** von x . Weiterhin wird im Fall $x < y$ der Knoten y oberhalb vom Knoten x gezeichnet, so dass auf Pfeilspitzen verzichtet werden kann.

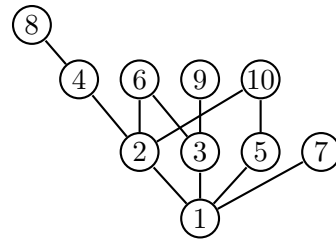
Beispiel 38.

Das Hasse-Diagramm rechts zeigt die Inklusionsrelation auf der Potenzmenge $\mathcal{P}(M)$ von $M = \{a, b, c\}$.



Das Hasse-Diagramm links zeigt die feiner-Relation auf der Menge aller Partitionen von $M = \{a, b, c\}$.

Schränken wir die "teilt"-Relation auf die Menge $\{1, 2, \dots, 10\}$ ein, so erhalten wir nebenstehendes Hasse-Diagramm.



Definition 39. Sei \leq eine Ordnung auf A und sei b ein Element in einer Teilmenge $B \subseteq A$.

- b heißt **kleinstes Element** oder **Minimum** von B (kurz $b = \min B$), falls gilt:

$$\forall b' \in B : b \leq b'.$$

- b heißt **größtes Element** oder **Maximum** von B (kurz $b = \max B$), falls gilt:

$$\forall b' \in B : b' \leq b.$$

- b heißt **minimal** in B , falls es in B kein kleineres Element gibt:

$$\forall b' \in B : b' \leq b \Rightarrow b' = b.$$

- b heißt **maximal** in B , falls es in B kein größeres Element gibt:

$$\forall b' \in B : b \leq b' \Rightarrow b = b'.$$

Bemerkung 40. Da Ordnungen antisymmetrisch sind, kann es in jeder Teilmenge B höchstens ein kleinstes und höchstens ein größtes Element geben. Die Anzahl der minimalen und maximalen Elemente in B kann dagegen beliebig groß sein.

Definition 41. Sei \leq eine Ordnung auf A und sei $B \subseteq A$.

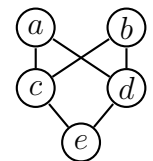
- Jedes Element $u \in A$ mit $u \leq b$ für alle $b \in B$ heißt **untere** und jedes $o \in A$ mit $b \leq o$ für alle $b \in B$ heißt **obere Schranke** von B .
- B heißt **nach oben beschränkt**, wenn B eine obere Schranke hat, und **nach unten beschränkt**, wenn B eine untere Schranke hat.
- B heißt **beschränkt**, wenn B nach oben und nach unten beschränkt ist.
- Besitzt B eine größte untere Schranke i , d.h. besitzt die Menge U aller unteren Schranken von B ein größtes Element i , so heißt i das **Infimum** von B (kurz $i = \inf B$):

$$(\forall b \in B : b \geq i) \wedge [\forall u \in A : (\forall b \in B : b \geq u) \Rightarrow u \leq i].$$

- Besitzt B eine kleinste obere Schranke s , d.h. besitzt die Menge O aller oberen Schranken von B ein kleinstes Element s , so heißt s das **Supremum** von B ($s = \sup B$):

$$(\forall b \in B : b \leq s) \wedge [\forall o \in A : (\forall b \in B : b \leq o) \Rightarrow s \leq o]$$

Beispiel 42. Betrachte nebenstehende Ordnung. Die folgende Tabelle zeigt für verschiedene Teilmengen $B \subseteq \{a, b, c, d, e\}$ alle minimalen und maximalen Elemente, alle unteren und oberen Schranken sowie Minimum, Maximum, Infimum und Supremum von B (falls existent).



B	minimal	maximal	min	max	untere Schranken	obere Schranken	inf	sup
$\{a, b\}$	a, b	a, b	-	-	c, d, e	-	-	-
$\{c, d\}$	c, d	c, d	-	-	e	a, b	e	-
$\{a, b, c\}$	c	a, b	c	-	c, e	-	c	-
$\{a, b, c, e\}$	e	a, b	e	-	e	-	e	-
$\{a, c, d, e\}$	e	a	e	a	e	a	e	a

◁

Bemerkung 43.

- Es kann nicht mehr als ein Supremum und ein Infimum geben.
- Auch in linearen Ordnungen muss nicht jede beschränkte Teilmenge ein Supremum oder Infimum besitzen. So hat in der linear geordneten Menge (\mathbb{Q}, \leq) die Teilmenge

$$\{x \in \mathbb{Q} \mid x^2 \leq 2\} = \{x \in \mathbb{Q} \mid x^2 < 2\}$$

weder ein Supremum noch ein Infimum.

- Dagegen hat in (\mathbb{R}, \leq) jede beschränkte Teilmenge ein Supremum und ein Infimum (aber eventuell kein Maximum oder Minimum).

2.4.2 Abbildungen

Definition 44. Sei R eine binäre Relation auf einer Menge M .

- R heißt **rechtseindeutig**, falls für alle $x, y, z \in M$ gilt:

$$xRy \wedge xRz \Rightarrow y = z.$$

- R heißt **linkseindeutig**, falls für alle $x, y, z \in M$ gilt:

$$xRz \wedge yRz \Rightarrow x = y.$$

- Der **Nachbereich** $N(R)$ und der **Vorbereich** $V(R)$ von R sind

$$N(R) = \bigcup_{x \in M} R[x] \quad \text{und} \quad V(R) = \bigcup_{x \in M} R^T[x].$$

- Eine rechtseindeutige Relation R mit $V(R) = A$ und $N(R) \subseteq B$ heißt **Abbildung** oder **Funktion von A nach B** (kurz $R: A \rightarrow B$).

Bemerkung 45.

- Wie üblich werden wir Abbildungen meist mit kleinen Buchstaben f, g, h, \dots bezeichnen und für $(x, y) \in f$ nicht xfy sondern $f(x) = y$ oder $f: x \mapsto y$ schreiben.
- Ist $f: A \rightarrow B$ eine Abbildung, so wird der Vorbereich $V(f) = A$ der **Definitionsbereich** und die Menge B der **Wertebereich** oder **Wertevorrat** von f genannt.
- Der Nachbereich $N(f)$ wird als **Bild** von f bezeichnet.

Definition 46.

- Im Fall $N(f) = B$ heißt f **surjektiv**.
- Ist f linkseindeutig, so heißt f **injektiv**. In diesem Fall impliziert $f(x) = f(y)$ die Gleichheit $x = y$.
- Eine injektive und surjektive Abbildung heißt **bijektiv**.
- Ist f injektiv, so ist auch $f^{-1}: N(f) \rightarrow A$ eine Abbildung, die als die zu f inverse Abbildung bezeichnet wird.

Man beachte, dass der Definitionsbereich $V(f^{-1}) = N(f)$ von f^{-1} nur dann gleich B ist, wenn f auch surjektiv, also eine Bijektion ist.

2.4.3 Homo- und Isomorphismen

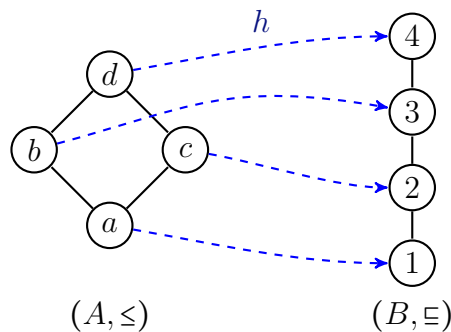
Definition 47. Seien (A_1, R_1) und (A_2, R_2) Relationalstrukturen.

- Eine Abbildung $h : A_1 \rightarrow A_2$ heißt **Homomorphismus**, falls für alle $a, b \in A_1$ gilt:

$$aR_1b \Rightarrow h(a)R_2h(b).$$

- Sind (A_1, R_1) und (A_2, R_2) Ordnungen, so spricht man von **Ordnungshomomorphismen** oder einfach von **monotonen Abbildungen**.
- Injektive Ordnungshomomorphismen werden auch **streng monotone** Abbildungen genannt.

Beispiel 48. Folgende Abbildung $h : A_1 \rightarrow A_2$ ist ein bijektiver Ordnungshomomorphismus.



Obwohl h ein bijektiver Homomorphismus ist, ist die Umkehrung h^{-1} kein Homomorphismus, da h^{-1} nicht monoton ist. Es gilt nämlich

$$2 \subseteq 3, \text{ aber } h^{-1}(2) = b \not\subseteq c = h^{-1}(3).$$

◁

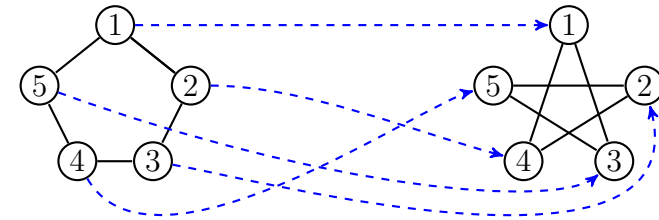
Definition 49. Ein bijektiver Homomorphismus $h : A_1 \rightarrow A_2$, bei dem auch h^{-1} ein Homomorphismus ist, d.h. es gilt

$$\forall a, b \in A_1 : aR_1b \Leftrightarrow h(a)R_2h(b).$$

heißt **Isomorphismus**. In diesem Fall heißen die Strukturen (A_1, R_1) und (A_2, R_2) **isomorph** (kurz: $(A_1, R_1) \cong (A_2, R_2)$).

Beispiel 50.

- Die beiden folgenden Graphen G und G' sind isomorph. Zwei Isomorphismen sind beispielsweise h_1 und h_2 .

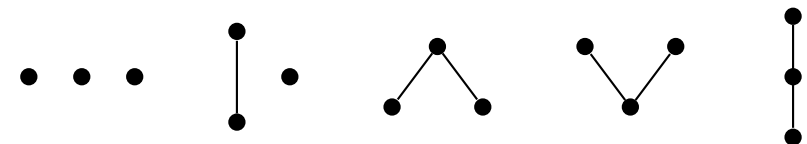


$G = (V, E)$	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px 10px;"></td> <td style="padding: 2px 10px;">1</td> <td style="padding: 2px 10px;">2</td> <td style="padding: 2px 10px;">3</td> <td style="padding: 2px 10px;">4</td> <td style="padding: 2px 10px;">5</td> </tr> </table>		1	2	3	4	5	$G' = (V, E')$
	1	2	3	4	5			
	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px 10px;">$h_1(v)$</td> <td style="padding: 2px 10px;">1</td> <td style="padding: 2px 10px;">3</td> <td style="padding: 2px 10px;">5</td> <td style="padding: 2px 10px;">2</td> <td style="padding: 2px 10px;">4</td> </tr> </table>	$h_1(v)$	1	3	5	2	4	
$h_1(v)$	1	3	5	2	4			
	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px 10px;">$h_2(v)$</td> <td style="padding: 2px 10px;">1</td> <td style="padding: 2px 10px;">4</td> <td style="padding: 2px 10px;">2</td> <td style="padding: 2px 10px;">5</td> <td style="padding: 2px 10px;">3</td> </tr> </table>	$h_2(v)$	1	4	2	5	3	
$h_2(v)$	1	4	2	5	3			

- Während auf der Knotenmenge $V = [3]$ insgesamt $2^3 = 8$ verschiedene Graphen existieren, gibt es auf dieser Menge nur 4 verschiedene nichtisomorphe Graphen:



- Die Abbildung $h : \mathbb{R} \rightarrow \mathbb{R}^+$ mit $h(x) = e^x$ ist ein Ordnungsisomorphismus zwischen (\mathbb{R}, \leq) und (\mathbb{R}^+, \leq) .
- Es existieren genau 5 nichtisomorphe Ordnungen mit 3 Elementen:



Anders ausgedrückt: Die Klasse aller dreielementigen Ordnungen zerfällt unter der Äquivalenzrelation \cong in fünf Äquivalenzklassen, die durch obige fünf Hasse-Diagramme repräsentiert werden.

- Für $n \in \mathbb{N}$ sei $T_n = \{k \in \mathbb{N} \mid k \text{ teilt } n\}$ die Menge aller Teiler von n und $P_n = \{p \in T_n \mid p \text{ ist prim}\}$ die Menge aller Primteiler von n . Dann ist die Abbildung

$$h : k \mapsto P_k$$

ein (surjektiver) Ordnungshomomorphismus von $(T_n, |)$ auf $(\mathcal{P}(P_n), \subseteq)$. h ist sogar ein Isomorphismus, falls n quadratfrei ist (d.h. es gibt kein $k \geq 2$, so dass k^2 die Zahl n teilt).

◁

2.5 Minimierung von DFAs

Wie können wir feststellen, ob ein DFA $M = (Z, \Sigma, \delta, q_0, E)$ unnötige Zustände enthält? Zunächst einmal können alle Zustände entfernt werden, die nicht vom Startzustand aus erreichbar sind. Im folgenden gehen wir daher davon aus, dass M keine unerreichbaren Zustände enthält.

Offensichtlich können zwei Zustände q und p zu einem Zustand verschmolzen werden (kurz: $q \sim_M p$), wenn M von q und von p ausgehend jeweils dieselben Wörter akzeptiert. Bezeichnen wir den DFA $(Z, \Sigma, \delta, q, E)$ mit M_q , so sind q und p genau dann verschmelzbar, wenn $L(M_q) = L(M_p)$ ist. Offensichtlich ist \sim_M eine Äquivalenzrelation.

Fassen wir alle mit einem Zustand z verschmelzbaren Zustände in dem neuen Zustand

$$[z]_{\sim_M} = \{z' \in Z \mid L(M_{z'}) = L(M_z)\}$$

zusammen (wofür wir auch kurz $[z]$ oder \tilde{z} schreiben) und ersetzen wir Z und E durch $\tilde{Z} = \{\tilde{z} \mid z \in Z\}$ und $\tilde{E} = \{\tilde{z} \mid z \in E\}$, so erhalten wir den DFA $M' = (\tilde{Z}, \Sigma, \delta', \tilde{q}_0, \tilde{E})$ mit

$$\delta'(\tilde{q}, a) = \widetilde{\delta(q, a)}.$$

Hierbei bezeichnet \tilde{Q} für eine Teilmenge $Q \subseteq Z$ die Menge $\{\tilde{q} \mid q \in Q\}$ aller Äquivalenzklassen \tilde{q} , die mindestens ein Element $q \in Q$ enthalten. Der nächste Satz zeigt, dass M' tatsächlich der gesuchte Minimalautomat ist.

Satz 51. Sei $M = (Z, \Sigma, \delta, q_0, E)$ ein DFA, der nur Zustände enthält, die vom Startzustand q_0 aus erreichbar sind. Dann ist $M' = (\tilde{Z}, \Sigma, \delta', \tilde{q}_0, \tilde{E})$ mit

$$\delta'(\tilde{q}, a) = \widetilde{\delta(q, a)}$$

ein DFA für $L(M)$ mit einer minimalen Anzahl von Zuständen.

Beweis. Wir zeigen zuerst, dass δ' wohldefiniert ist, also der Wert von $\delta'(\tilde{q}, a)$ nicht von der Wahl des Repräsentanten q abhängt. Hierzu zeigen wir, dass im Fall $p \sim_M q$ auch $\delta(q, a)$ und $\delta(p, a)$ äquivalent sind:

$$\begin{aligned} L(M_q) = L(M_p) &\Rightarrow \forall x \in \Sigma^* : x \in L(M_q) \leftrightarrow x \in L(M_p) \\ &\Rightarrow \forall x \in \Sigma^* : ax \in L(M_q) \leftrightarrow ax \in L(M_p) \\ &\Rightarrow \forall x \in \Sigma^* : x \in L(M_{\delta(q,a)}) \leftrightarrow x \in L(M_{\delta(p,a)}) \\ &\Rightarrow L(M_{\delta(q,a)}) = L(M_{\delta(p,a)}). \end{aligned}$$

Als nächstes zeigen wir, dass $L(M') = L(M)$ ist. Sei $x = x_1 \dots x_n$ eine Eingabe und seien

$$q_i = \hat{\delta}(q_0, x_1 \dots x_i), \quad i = 0, \dots, n$$

die von M bei Eingabe x durchlaufenen Zustände. Wegen

$$\delta'(\tilde{q}_{i-1}, x_i) = \widetilde{\delta(q_{i-1}, x_i)} = \tilde{q}_i$$

durchläuft M' dann die Zustände

$$\tilde{q}_0, \tilde{q}_1, \dots, \tilde{q}_n.$$

Da aber q_n genau dann zu E gehört, wenn $\tilde{q}_n \in \tilde{E}$ ist, folgt $L(M') = L(M)$ (man beachte, dass \tilde{q}_n entweder nur Endzustände oder nur Nicht-Endzustände enthält, vgl. Beobachtung 52).

Es bleibt zu zeigen, dass M' eine minimale Anzahl $\|\tilde{Z}\|$ von Zuständen hat. Dies ist sicher dann der Fall, wenn bereits M minimal ist. Es reicht also zu zeigen, dass die Anzahl $k = \|\tilde{Z}\| = \|\{L(M_z) \mid z \in Z\}\|$ der Zustände von M' nicht von M , sondern nur von $L = L(M)$ abhängt. Für $x \in \Sigma^*$ sei

$$L_x = \{y \in \Sigma^* \mid xy \in L\}.$$

Dann gilt $\{L_x \mid x \in \Sigma^*\} \subseteq \{L(M_z) \mid z \in Z\}$, da $L_x = L(M_{\delta(q_0, x)})$ ist. Die umgekehrte Inklusion gilt ebenfalls, da nach Voraussetzung jeder Zustand $q \in Z$ über ein $x \in \Sigma^*$ erreichbar ist. Also hängt $k = \|\{L(M_z) \mid z \in Z\}\| = \|\{L_x \mid x \in \Sigma^*\}\|$ nur von L ab. ■

Eine interessante Folgerung aus obigem Beweis ist, dass für eine reguläre Sprache $L \subseteq \Sigma^*$ die Menge $\{L_x \mid x \in \Sigma^*\}$ nur endlich viele verschiedene Sprachen enthält, und somit die durch

$$x \sim_L y \Leftrightarrow L_x = L_y$$

auf Σ^* definierte Äquivalenzrelation \sim_L endlichen Index hat. Die Relation \sim_L wird als *Nerode-Relation* von L bezeichnet.

Für die algorithmische Konstruktion von M' aus M ist es notwendig herauszufinden, ob zwei Zustände p und q von M äquivalent sind oder nicht.

Bezeichne $A \triangle B = (A \setminus B) \cup (B \setminus A)$ die *symmetrische Differenz* von zwei Mengen A und B . Dann ist die Inäquivalenz $p \not\sim_M q$ zweier Zustände p und q gleichbedeutend mit $L(M_p) \triangle L(M_q) \neq \emptyset$. Wir nennen ein Wort $x \in L(M_p) \triangle L(M_q)$ einen *Unterscheider* zwischen p und q . Für $i \geq 0$ sei D_i die Menge aller Paare $\{p, q\}$, die einen Unterscheider x der Länge $|x| \leq i$ haben. Dann enthält $D = \bigcup_{i \geq 0} D_i$ alle inäquivalenten Paare $\{p, q\}$, d.h. es gilt $p \sim_M q \Leftrightarrow \{p, q\} \notin D$.

Beobachtung 52.

- Das leere Wort ε unterscheidet Endzustände und Nichtendzustände, d.h.

$$D_0 = \{\{p, q\} \subseteq Z \mid p \in E, q \notin E\}.$$

- Zudem haben p und q genau dann einen Unterscheider ax der Länge $i + 1$, wenn x die beiden Zustände $\delta(p, a)$ und $\delta(q, a)$ unterscheidet, d.h.

$$D_{i+1} = D_i \cup \left\{ \{p, q\} \subseteq Z \mid \exists a \in \Sigma : \{\delta(p, a), \delta(q, a)\} \in D_i \right\}.$$

Da es nur endlich viele Zustandspaare gibt, muss es ein k geben mit $D = D_k$. Offensichtlich gilt

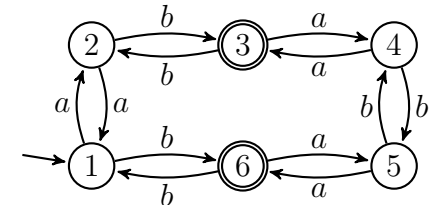
$$D = D_k \Leftrightarrow D_{k+1} = D_k.$$

Der folgende Algorithmus berechnet für einen beliebigen DFA M den zugehörigen Minimal-DFA M' .

Algorithmus min-DFA(M)

-
- 1 **Input:** DFA $M = (Z, \Sigma, \delta, q_0, E)$
 - 2 entferne alle nicht erreichbaren Zustände
 - 3 $D' := \{\{z, z'\} \mid z \in E, z' \notin E\}$
 - 4 **repeat**
 - 5 $D := D'$
 - 6 $D' := D \cup \{\{p, q\} \mid \exists a \in \Sigma : \{\delta(p, a), \delta(q, a)\} \in D\}$
 - 7 **until** $D' = D$
 - 8 **Output:** $M' = (\tilde{Z}, \Sigma, \delta', \tilde{q}_0, \tilde{E})$, wobei für jeden Zustand $z \in Z$ gilt: $\tilde{z} = \{z' \in Z \mid \{z, z'\} \notin D\}$
-

Beispiel 53. Betrachte den DFA M :



Dann enthält D_0 die Paare

$$\{1, 3\}, \{1, 6\}, \{2, 3\}, \{2, 6\}, \{3, 4\}, \{3, 5\}, \{4, 6\}, \{5, 6\}.$$

Die Paare in D_0 sind in der folgenden Matrix durch den Unterscheider ε markiert.

2					
3	ε	ε			
4	a	a	ε		
5	a	a	ε		
6	ε	ε		ε	ε
	1	2	3	4	5

Wegen

$\{p, q\}$	$\{1, 4\}$	$\{1, 5\}$	$\{2, 4\}$	$\{2, 5\}$
$\{\delta(q, a), \delta(p, a)\}$	$\{2, 3\}$	$\{2, 6\}$	$\{1, 3\}$	$\{1, 6\}$

enthält D_1 die zusätzlichen Paare $\{1, 4\}, \{1, 5\}, \{2, 4\}, \{2, 5\}$ (in obiger Matrix durch den Unterscheider a markiert). Da nun jedoch keines der verbliebenen Paare $\{1, 2\}, \{3, 6\}, \{4, 5\}$ wegen

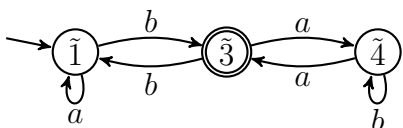
$\{p, q\}$	$\{1, 2\}$	$\{3, 6\}$	$\{4, 5\}$
$\{\delta(p, a), \delta(q, a)\}$	$\{1, 2\}$	$\{4, 5\}$	$\{3, 6\}$
$\{\delta(p, b), \delta(q, b)\}$	$\{3, 6\}$	$\{1, 2\}$	$\{4, 5\}$

zu D_2 hinzugefügt werden kann, gilt $D_2 = D_1$ und somit $D = D_1$.

Aus den unmarkierten Paaren $\{1, 2\}, \{3, 6\}$ und $\{4, 5\}$ erhalten wir die Äquivalenzklassen

$$\tilde{1} = \{1, 2\}, \quad \tilde{3} = \{3, 6\} \quad \text{und} \quad \tilde{4} = \{4, 5\},$$

die auf folgenden Minimal-DFA M' führen:



◁

Es ist auch möglich, einen Minimalautomaten M_L direkt aus einer regulären Sprache L zu gewinnen (also ohne einen DFA M für L zu kennen). Da wegen

$$\begin{aligned} \widehat{\delta}(q_0, x) = \widehat{\delta}(q_0, y) &\Leftrightarrow \widehat{\delta}(q_0, x) \sim_M \widehat{\delta}(q_0, y) \\ &\Leftrightarrow L(M_{\widehat{\delta}(q_0, x)}) = L(M_{\widehat{\delta}(q_0, y)}) \Leftrightarrow L_x = L_y \end{aligned}$$

zwei Eingaben x und y den DFA M' genau dann in denselben Zustand $\widehat{\delta}(q_0, x) = \widehat{\delta}(q_0, y)$ überführen, wenn $L_x = L_y$ ist, können wir den von M' bei Eingabe x erreichten Zustand $\widehat{\delta}(q_0, x)$ auch mit der Sprache L_x bezeichnen. Dies führt auf den zu M' isomorphen (also bis auf die Benennung der Zustände mit M' identischen) DFA $M_L = (Z_L, \Sigma, \delta_L, L_\varepsilon, E_L)$ mit

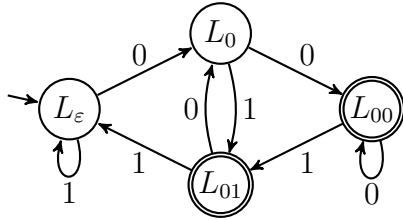
$$\begin{aligned} Z_L &= \{L_x \mid x \in \Sigma^*\}, \\ E_L &= \{L_x \mid x \in L\} \text{ und} \\ \delta_L(L_x, a) &= L_{xa}. \end{aligned}$$

Notwendig und hinreichend für die Existenz von M_L ist, dass die Nerode-Relation \sim_L von L endlichen Index hat, also die Menge $\{L_x \mid x \in \Sigma^*\}$ endlich ist.

Beispiel 54. Für $L = \{x_1 \dots x_n \in \{0, 1\}^* \mid n \geq 2 \text{ und } x_{n-1} = 0\}$ ist

$$L_x = \begin{cases} L, & x \in \{\varepsilon, 1\} \text{ oder } x \text{ endet mit } 11, \\ L \cup \{0, 1\}, & x = 0 \text{ oder } x \text{ endet mit } 10, \\ L \cup \{\varepsilon, 0, 1\}, & x \text{ endet mit } 00, \\ L \cup \{\varepsilon\}, & x \text{ endet mit } 01. \end{cases}$$

Somit erhalten wir den folgenden Minimalautomaten M_L .



◁

Im Fall, dass M bereits ein Minimalautomat ist, sind alle Zustände von M' von der Form $\tilde{q} = \{q\}$, so dass M isomorph zu M' und damit auch isomorph zu M_L ist. Dies zeigt, dass alle Minimalautomaten für eine Sprache L isomorph sind.

Satz 55 (Myhill und Nerode).

1. Sei L regulär und sei $\text{index}(\sim_L)$ der Index von \sim_L . Dann gibt es für L bis auf Isomorphie genau einen Minimal-DFA. Dieser hat $\text{index}(\sim_L)$ Zustände.
2. $\text{REG} = \{L \mid \sim_L \text{ hat endlichen Index}\}$.

Beispiel 56. Sei $L = \{a^i b^i \mid i \geq 0\}$. Wegen $b^i \in L_{a^i} \Delta L_{a^j}$ für $i \neq j$ hat \sim_L unendlichen Index, d.h. L ist nicht regulär. ◁

Die Zustände von M_L können anstelle von L_x auch mit den Äquivalenzklassen \tilde{x} von \sim_L benannt werden. Der resultierende Minimal-DFA $M_{\sim_L} = (Z, \Sigma, \delta, \tilde{\varepsilon}, E)$ mit

$$\begin{aligned} Z &= \{\tilde{x} \mid x \in \Sigma^*\}, \\ E &= \{\tilde{x} \mid x \in L\} \text{ und} \\ \delta(\tilde{x}, a) &= \tilde{x}a \end{aligned}$$

wird auch als **Äquivalenzklassenautomat** bezeichnet.

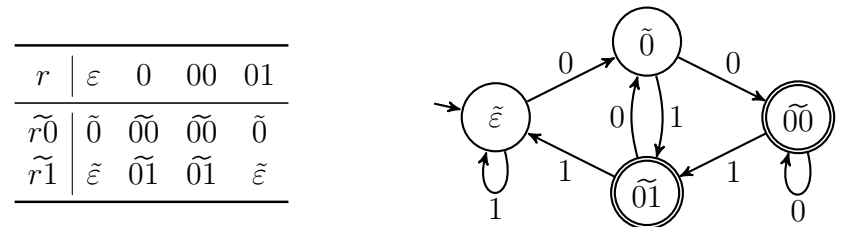
Die Konstruktion von M_{\sim_L} ist meist einfacher als die von M_L , da die Bestimmung der Sprachen L_x entfällt. Um die Überföhrungsfunktion von M_{\sim_L} aufzustellen, reicht es, ausgehend von $r_1 = \varepsilon$ eine Folge

r_1, \dots, r_k von paarweise bzgl. \sim_L inäquivalenten Wörtern zu bestimmen, so dass zu jedem Wort $r_i a$, $a \in \Sigma$, ein r_j mit $r_i a \sim_L r_j$ existiert. In diesem Fall ist $\delta(\tilde{r}_i, a) = \tilde{r}_i a = \tilde{r}_j$.

Beispiel 57. Für die Sprache $L = \{x_1 \dots x_n \in \{0, 1\}^* \mid x_{n-1} = 0\}$ lässt sich M_{\sim_L} wie folgt konstruieren:

1. Wir beginnen mit $r_1 = \varepsilon$.
2. Da $r_1 0 = 0 \not\sim_L \varepsilon$ ist, erhalten wir $r_2 = 0$ und setzen $\delta(\tilde{\varepsilon}, 0) = \tilde{0}$.
3. Da $r_1 1 = 1 \sim_L \varepsilon$ ist, setzen wir $\delta(\tilde{\varepsilon}, 1) = \tilde{\varepsilon}$.
4. Da $r_2 0 = 00 \not\sim_L r_i$ für $i = 1, 2$ ist, erhalten wir $r_3 = 00$ und setzen $\delta(\tilde{0}, 0) = \tilde{00}$.
5. Da $r_2 1 = 01 \not\sim_L r_i$ für $i = 1, 2, 3$ ist, erhalten wir $r_4 = 01$ und setzen $\delta(\tilde{0}, 1) = \tilde{01}$.
6. Da zudem $r_3 0 = 000 \sim_L 00$, $r_3 1 = 001 \sim_L 01$, $r_4 0 = 010 \sim_L 0$ und $r_4 1 = 011 \sim_L \varepsilon$ gilt, setzen wir $\delta(\tilde{00}, 0) = \tilde{00}$, $\delta(\tilde{00}, 1) = \tilde{01}$, $\delta(\tilde{01}, 0) = \tilde{0}$ und $\delta(\tilde{01}, 1) = \tilde{\varepsilon}$.

Wir erhalten also folgenden Minimal-DFA M_{\sim_L} :



◁

Wir fassen nochmals die wichtigsten Ergebnisse zusammen.

Korollar 58. Für jede Sprache L sind folgende Aussagen äquivalent:

- L ist regulär (d.h. es gibt einen DFA M mit $L = L(M)$),
- es gibt einen NFA N mit $L = L(N)$,
- es gibt einen regulären Ausdruck γ mit $L = L(\gamma)$,
- die Nerode-Relation \sim_L von L hat endlichen Index.

Wir werden im nächsten Abschnitt noch eine weitere Methode kennenlernen, mit der man beweisen kann, dass eine Sprache nicht regulär ist, nämlich das Pumping-Lemma.

2.6 Das Pumping-Lemma

Wie kann man von einer Sprache nachweisen, dass sie nicht regulär ist? Eine Möglichkeit besteht darin, die Kontraposition folgender Aussage anzuwenden.

Satz 59 (Pumping-Lemma für reguläre Sprachen).

Zu jeder regulären Sprache L gibt es eine Zahl $l \geq 0$, so dass sich alle Wörter $x \in L$ mit $|x| \geq l$ in $x = uvw$ zerlegen lassen mit

1. $v \neq \varepsilon$,
2. $|uv| \leq l$ und
3. $uv^i w \in L$ für alle $i \geq 0$.

Falls eine Zahl $l \geq 0$ mit diesen Eigenschaften existiert, wird das kleinste solche l die **Pumping-Zahl** von L genannt.

Beispiel 60. Die Sprache

$$L = \{x \in \{a, b\}^* \mid \#_a(x) - \#_b(x) \equiv_3 1\}$$

hat die Pumping-Zahl $l = 3$. Sei nämlich $x \in L$ beliebig mit $|x| \geq 3$. Dann lässt sich innerhalb des Präfixes von x der Länge drei ein nichtleeres Teilwort v finden, das gepumpt werden kann:

1. Fall: x hat das Präfix ab (oder ba).

Zerlege $x = uvw$ mit $u = \varepsilon$ und $v = ab$ (bzw. $v = ba$).

2. Fall: x hat das Präfix aab (oder bba).

Zerlege $x = uvw$ mit $u = a$ (bzw. $u = b$) und $v = ab$ (bzw. $v = ba$).

3. Fall: x hat das Präfix aaa (oder bbb).

Zerlege $x = uvw$ mit $u = \varepsilon$ und $v = aaa$ (bzw. $v = bbb$). \triangleleft

Beispiel 61. Eine endliche Sprache L hat die Pumping-Zahl

$$l = \begin{cases} 0, & L = \emptyset, \\ \max\{|x| + 1 \mid x \in L\}, & \text{sonst.} \end{cases}$$

Tatsächlich lässt sich jedes Wort $x \in L$ der Länge $|x| \geq l$ „pumpen“ (da solche Wörter gar nicht existieren), weshalb die Pumping-Zahl höchstens l ist. Zudem gibt es im Fall $l > 0$ ein Wort $x \in L$ der Länge $|x| = l - 1$, das sich nicht „pumpen“ lässt, weshalb die Pumping-Zahl nicht kleiner als l sein kann. \triangleleft

Wollen wir mit Hilfe des Pumping-Lemmas von einer Sprache L zeigen, dass sie nicht regulär ist, so genügt es, für jede Zahl l ein Wort $x \in L$ der Länge $|x| \geq l$ anzugeben, so dass für jede Zerlegung von x in drei Teilwörter u, v, w mindestens eine der drei in Satz 59 aufgeführten Eigenschaften verletzt ist.

Beispiel 62. Die Sprache

$$L = \{a^j b^j \mid j \geq 0\}$$

ist nicht regulär, da sich für jede Zahl $l \geq 0$ das Wort $x = a^l b^l$ der Länge $|x| = 2l \geq l$ in der Sprache L befindet, welches offensichtlich nicht in Teilwörter u, v, w mit $v \neq \varepsilon$ und $uv^2 w \in L$ zerlegbar ist. \triangleleft

Beispiel 63. Die Sprache

$$L = \{a^{n^2} \mid n \geq 0\}$$

ist ebenfalls nicht regulär. Andernfalls müsste es nämlich eine Zahl l geben, so dass jede Quadratzahl $n^2 \geq l$ als Summe von natürlichen Zahlen $u + v + w$ darstellbar ist mit der Eigenschaft, dass $v \geq 1$ und $u + v \leq l$ ist, und für jedes $i \geq 0$ auch $u + iv + w$ eine Quadratzahl ist. Insbesondere müsste also $u + 2v + w = n^2 + v$ eine Quadratzahl sein, was wegen

$$n^2 < n^2 + v \leq n^2 + l < n^2 + 2l + 1 = (n + 1)^2$$

ausgeschlossen ist. \triangleleft

Beispiel 64. Auch die Sprache

$$L = \{a^p \mid p \text{ prim}\}$$

ist nicht regulär, da sich sonst jede Primzahl p einer bestimmten Mindestgröße l als Summe von natürlichen Zahlen $u + v + w$ darstellen ließe, so dass $v \geq 1$ und für alle $i \geq 0$ auch $u + iv + w = p + (i - 1)v$ prim ist. Dies ist jedoch für $i = p + 1$ wegen

$$p + (p + 1 - 1)v = p(1 + v)$$

nicht der Fall. ◁

Bemerkung 65. Mit Hilfe des Pumping-Lemmas kann nicht für jede Sprache $L \notin \text{REG}$ gezeigt werden, dass L nicht regulär ist, da seine Umkehrung falsch ist. So hat beispielsweise die Sprache

$$L = \{a^i b^j c^k \mid i = 0 \text{ oder } j = k\}$$

die Pumping-Zahl 1 (d.h. jedes Wort $x \in L$ mit Ausnahme von ϵ kann „gepumpt“ werden). Dennoch ist L nicht regulär (siehe Übungen).

2.7 Grammatiken

Eine beliebte Methode, Sprachen zu beschreiben, sind Grammatiken. Implizit haben wir diese Methode bei der Definition der regulären Ausdrücke bereits benutzt.

Beispiel 66. Die Sprache RA aller regulären Ausdrücke über einem Alphabet $\Sigma = \{a_1, \dots, a_k\}$ lässt sich aus dem Symbol R durch wiederholte Anwendung folgender Regeln erzeugen:

$$\begin{array}{ll} R \rightarrow \emptyset, & R \rightarrow RR, \\ R \rightarrow \epsilon, & R \rightarrow (R|R), \\ R \rightarrow a_i, i = 1, \dots, k, & R \rightarrow (R)^*. \end{array} \quad \triangleleft$$

Definition 67. Eine **Grammatik** ist ein 4-Tupel $G = (V, \Sigma, P, S)$, wobei

- V eine endliche Menge von **Variablen** (auch **Nichtterminalsymbole** genannt),
- Σ das **Terminalalphabet**,
- $P \subseteq (V \cup \Sigma)^+ \times (V \cup \Sigma)^*$ eine endliche Menge von **Regeln** (oder **Produktionen**) und
- $S \in V$ die **Startvariable** ist.

Für $(u, v) \in P$ schreiben wir auch kurz $u \rightarrow_G v$ bzw. $u \rightarrow v$, wenn die benutzte Grammatik aus dem Kontext ersichtlich ist.

Definition 68. Seien $\alpha, \beta \in (V \cup \Sigma)^*$.

- a) Wir sagen, β ist aus α in einem Schritt ableitbar (kurz: $\alpha \Rightarrow_G \beta$), falls eine Regel $u \rightarrow_G v$ und Wörter $l, r \in (V \cup \Sigma)^*$ existieren mit

$$\alpha = lur \text{ und } \beta = lvr.$$

Hierfür schreiben wir auch $\underline{lur} \Rightarrow_G \underline{lvr}$. (Man beachte, dass durch Unterstreichen von u in α sowohl die benutzte Regel als auch die Stelle in α , an der u durch v ersetzt wird, eindeutig erkennbar sind.)

- b) Eine Folge $\sigma = (l_0, u_0, r_0), \dots, (l_m, u_m, r_m)$ von Tripeln (l_i, u_i, r_i) heißt **Ableitung** von β aus α , falls gilt:

- $l_0 u_0 r_0 = \alpha$, $l_m = l_m u_m r_m = \beta$ und
- $l_i u_i r_i \Rightarrow l_{i+1} u_{i+1} r_{i+1}$ für $i = 0, \dots, m - 1$.

Die **Länge** von σ ist m und wir notieren σ auch in der Form

$$l_0 \underline{u_0} r_0 \Rightarrow l_1 \underline{u_1} r_1 \Rightarrow \dots \Rightarrow l_{m-1} \underline{u_{m-1}} r_{m-1} \Rightarrow l_m u_m r_m.$$

- c) Die durch G erzeugte Sprache ist

$$L(G) = \{x \in \Sigma^* \mid S \Rightarrow_G^* x\}.$$

d) Ein Wort $\alpha \in (V \cup \Sigma)^*$ mit $S \Rightarrow_G^* \alpha$ heißt **Satzform** von G .

Zur Erinnerung: Die Relation \Rightarrow^* bezeichnet die reflexive, transitive Hülle der Relation \Rightarrow , d.h. $\alpha \Rightarrow^* \beta$ bedeutet, dass es ein $n \geq 0$ gibt mit $\alpha \Rightarrow^n \beta$. Hierzu sagen wir auch, β ist aus α (in n Schritten) **ableitbar**. Die Relation \Rightarrow^n bezeichnet das n -fache Produkt der Relation \Rightarrow , d.h. es gilt $\alpha \Rightarrow^n \beta$, falls Wörter $\alpha_0, \dots, \alpha_n$ existieren mit

- $\alpha_0 = \alpha$, $\alpha_n = \beta$ und
- $\alpha_i \Rightarrow \alpha_{i+1}$ für $i = 0, \dots, n-1$.

Beispiel 69. Wir betrachten nochmals die Grammatik $G = (\{R\}, \Sigma \cup \{\emptyset, \epsilon, (,), *, | \}, P, R)$, die die Menge der regulären Ausdrücke über dem Alphabet Σ erzeugt, wobei P die oben angegebenen Regeln enthält. Ist $\Sigma = \{0, 1\}$, so lässt sich der reguläre Ausdruck $(01)^*(\epsilon|\emptyset)$ beispielsweise wie folgt ableiten:

$$\begin{aligned} R &\Rightarrow \underline{R}R \Rightarrow (\underline{R})^*R \Rightarrow (RR)^*R \Rightarrow (\underline{R}R)^*(R|R) \\ &\Rightarrow (0\underline{R})^*(R|R) \Rightarrow (01)^*(\underline{R}|R) \Rightarrow (01)^*(\epsilon|\underline{R}) \Rightarrow (01)^*(\epsilon|\emptyset) \quad \triangleleft \end{aligned}$$

Man unterscheidet vier verschiedene Typen von Grammatiken.

Definition 70. Sei $G = (V, \Sigma, P, S)$ eine Grammatik.

1. G heißt vom **Typ 3** oder **regulär**, falls für alle Regeln $u \rightarrow v$ gilt: $u \in V$ und $v \in \Sigma V \cup \Sigma \cup \{\epsilon\}$.
2. G heißt vom **Typ 2** oder **kontextfrei**, falls für alle Regeln $u \rightarrow v$ gilt: $u \in V$.
3. G heißt vom **Typ 1** oder **kontextsensitiv**, falls für alle Regeln $u \rightarrow v$ gilt: $|v| \geq |u|$ (mit Ausnahme der ϵ -Sonderregel, siehe unten).
4. Jede Grammatik ist automatisch vom **Typ 0**.

ϵ -Sonderregel: In einer kontextsensitiven Grammatik $G = (V, \Sigma, P, S)$ kann auch die verkürzende Regel $S \rightarrow \epsilon$ benutzt werden. Aber nur, wenn das Startsymbol S nicht auf der rechten Seite einer Regel in P vorkommt.

Die Sprechweisen „vom Typ i “ bzw. „regulär“, „kontextfrei“ und „kontextsensitiv“ werden auch auf die durch solche Grammatiken erzeugte Sprachen angewandt. (Der folgende Satz rechtfertigt dies für die regulären Sprachen, die wir bereits mit Hilfe von DFAs definiert haben.) Die zugehörigen neuen Sprachklassen sind

$$\text{CFL} = \{L(G) \mid G \text{ ist eine kontextfreie Grammatik}\},$$

(context free languages) und

$$\text{CSL} = \{L(G) \mid G \text{ ist eine kontextsensitive Grammatik}\}$$

(context sensitive languages). Da die Klasse der Typ 0 Sprachen mit der Klasse der rekursiv aufzählbaren (recursively enumerable) Sprachen übereinstimmt, bezeichnen wir diese Sprachklasse mit

$$\text{RE} = \{L(G) \mid G \text{ ist eine Grammatik}\}.$$

Die Sprachklassen

$$\text{REG} \subset \text{CFL} \subset \text{CSL} \subset \text{RE}$$

bilden eine Hierarchie (d.h. alle Inklusionen sind echt), die so genannte **Chomsky-Hierarchie**.

Als nächstes zeigen wir, dass sich mit regulären Grammatiken gerade die regulären Sprachen erzeugen lassen. Hierbei erweist sich folgende Beobachtung als nützlich.

Lemma 71. Zu jeder regulären Grammatik $G = (V, \Sigma, P, S)$ gibt es eine äquivalente reguläre Grammatik G' , die keine Produktionen der Form $A \rightarrow a$ hat.

Beweis. Betrachte die Grammatik $G' = (V', \Sigma, P', S)$ mit

$$\begin{aligned} V' &= V \cup \{X_{\text{neu}}\}, \\ P' &= \{A \rightarrow aX_{\text{neu}} \mid A \rightarrow_G a\} \cup \{X_{\text{neu}} \rightarrow \epsilon\} \cup P \setminus (V \times \Sigma). \end{aligned}$$

Es ist leicht zu sehen, dass G' die gleiche Sprache wie G erzeugt. ■

Satz 72. $REG = \{L(G) \mid G \text{ ist eine reguläre Grammatik}\}.$

Beweis. Sei $L \in REG$ und sei $M = (Z, \Sigma, \delta, q_0, E)$ ein DFA mit $L(M) = L$. Wir konstruieren eine reguläre Grammatik $G = (V, \Sigma, P, S)$ mit $L(G) = L$. Setzen wir

$$\begin{aligned} V &= Z, \\ S &= q_0 \text{ und} \\ P &= \{q \rightarrow ap \mid \delta(q, a) = p\} \cup \{q \rightarrow \varepsilon \mid q \in E\}, \end{aligned}$$

so gilt für alle Wörter $x = x_1 \dots x_n \in \Sigma^*$:

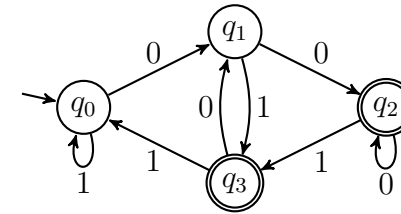
$$\begin{aligned} x \in L(M) &\Leftrightarrow \exists q_1, \dots, q_{n-1} \in Z \exists q_n \in E : \\ &\delta(q_{i-1}, x_i) = q_i \text{ für } i = 1, \dots, n \\ &\Leftrightarrow \exists q_1, \dots, q_n \in V : \\ &q_{i-1} \rightarrow_G x_i q_i \text{ für } i = 1, \dots, n \text{ und } q_n \rightarrow_G \varepsilon \\ &\Leftrightarrow \exists q_1, \dots, q_n \in V : \\ &q_0 \Rightarrow_G^i x_1 \dots x_i q_i \text{ für } i = 1, \dots, n \text{ und } q_n \rightarrow_G \varepsilon \\ &\Leftrightarrow x \in L(G) \end{aligned}$$

Für die entgegengesetzte Inklusion sei nun $G = (V, \Sigma, P, S)$ eine reguläre Grammatik, die keine Produktionen der Form $A \rightarrow a$ enthält. Dann können wir die gerade beschriebene Konstruktion einer Grammatik aus einem DFA „umdrehen“, um ausgehend von G einen NFA $M = (Z, \Sigma, \delta, \{S\}, E)$ mit

$$\begin{aligned} Z &= V, \\ E &= \{A \mid A \rightarrow_G \varepsilon\} \text{ und} \\ \delta(A, a) &= \{B \mid A \rightarrow_G aB\} \end{aligned}$$

zu erhalten. Genau wie oben folgt nun $L(M) = L(G)$. ■

Beispiel 73. Der DFA



führt auf die Grammatik $(\{q_0, q_1, q_2, q_3\}, \{0, 1\}, P, q_0)$ mit

$$\begin{aligned} P : \quad &q_0 \rightarrow 1q_0, 0q_1, \\ &q_1 \rightarrow 0q_2, 1q_3, \\ &q_2 \rightarrow 0q_2, 1q_3, \varepsilon, \\ &q_3 \rightarrow 0q_1, 1q_0, \varepsilon. \end{aligned}$$

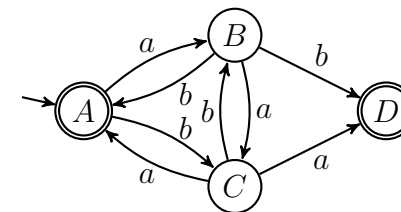
Umgekehrt führt die Grammatik $G = (\{A, B, C\}, \{a, b\}, P, A)$ mit

$$\begin{aligned} P : \quad &A \rightarrow aB, bC, \varepsilon, \\ &B \rightarrow aC, bA, b, \\ &C \rightarrow aA, bB, a \end{aligned}$$

über die Grammatik $G' = (\{A, B, C, D\}, \{a, b\}, P', A)$ mit

$$\begin{aligned} P' : \quad &A \rightarrow aB, bC, \varepsilon, \\ &B \rightarrow aC, bA, bD, \\ &C \rightarrow aA, bB, aD, \\ &D \rightarrow \varepsilon \end{aligned}$$

auf den NFA



3 Kontextfreie Sprachen

Wie wir gesehen haben, ist die Sprache $L = \{a^n b^n \mid n \geq 0\}$ nicht regulär. Es ist aber leicht, eine kontextfreie Grammatik für L zu finden:

$$G = (\{S\}, \{a, b\}, \{S \rightarrow aSb, S \rightarrow \varepsilon\}, S).$$

Damit ist klar, dass die Klasse der regulären Sprachen echt in der Klasse der kontextfreien Sprachen enthalten ist. Als nächstes wollen wir zeigen, dass die Klasse der kontextfreien Sprachen wiederum echt in der Klasse der kontextsensitiven Sprachen enthalten ist:

$$\text{REG} \subsetneq \text{CFL} \subsetneq \text{CSL}.$$

Kontextfreie Grammatiken sind dadurch charakterisiert, dass sie nur Regeln der Form $A \rightarrow \alpha$ haben. Dies lässt die Verwendung von beliebigen ε -Regeln der Form $A \rightarrow \varepsilon$ zu. Eine kontextsensitive Grammatik darf dagegen höchstens die ε -Regel $S \rightarrow \varepsilon$ haben. Voraussetzung hierfür ist, dass S das Startsymbol ist und dieses nicht auf der rechten Seite einer Regel vorkommt. Daher sind nicht alle kontextfreien Grammatiken kontextsensitiv. Beispielsweise ist die Grammatik $G = (\{S\}, \{a, b\}, \{S \rightarrow aSb, S \rightarrow \varepsilon\}, S)$ nicht kontextsensitiv, da sie die Regel $S \rightarrow \varepsilon$ enthält, obwohl S auf der rechten Seite der Regel $S \rightarrow aSb$ vorkommt.

Es lässt sich jedoch zu jeder kontextfreien Grammatik eine äquivalente kontextfreie Grammatik G' konstruieren, die auch kontextsensitiv ist. Hierzu zeigen wir zuerst, dass sich zu jeder kontextfreien Grammatik G , in der nicht das leere Wort ableitbar ist, eine äquivalente kontextfreie Grammatik G' ohne ε -Regeln konstruieren lässt.

Satz 74. *Zu jeder kontextfreien Grammatik G gibt es eine kontextfreie Grammatik G' ohne ε -Produktionen mit $L(G') = L(G) \setminus \{\varepsilon\}$.*

Beweis. Zuerst sammeln wir mit folgendem Algorithmus alle Variablen A , aus denen das leere Wort ableitbar ist. Diese werden auch als ε -ableitbar bezeichnet.

```

1  $E' := \{A \in V \mid A \rightarrow \varepsilon\}$ 
2 repeat
3    $E := E'$ 
4    $E' := E \cup \{A \in V \mid \exists B_1, \dots, B_k \in E : A \rightarrow B_1 \dots B_k\}$ 
5 until  $E = E'$ 

```

Nun konstruieren wir $G' = (V, \Sigma, P', S)$ wie folgt:

Nehme zu P' alle Regeln $A \rightarrow \alpha'$ mit $\alpha' \neq \varepsilon$ hinzu, für die P eine Regel $A \rightarrow \alpha$ enthält, so dass α' aus α durch Entfernen von beliebig vielen Variablen $A \in E$ hervorgeht.

■

Beispiel 75. *Betrachte die Grammatik $G = (V, \Sigma, P, S)$ mit $V = \{S, T, U, X, Y, Z\}$, $\Sigma = \{a, b, c\}$ und den Regeln*

$$P: \quad S \rightarrow aY, bX, Z; \quad Y \rightarrow bS, aYY; \quad T \rightarrow U; \\ X \rightarrow aS, bXX; \quad Z \rightarrow \varepsilon, S, T, cZ; \quad U \rightarrow abc.$$

Bei der Berechnung von $E = \{A \in V \mid A \Rightarrow^ \varepsilon\}$ ergeben sich der Reihe nach folgende Belegungen für die Mengenvariablen E und E' :*

$$\begin{array}{c|cc} \hline E' & \{Z\} & \{Z, S\} \\ E & \{Z, S\} & \{Z, S\} \\ \hline \end{array}$$

Um nun die Regelmenge P' zu bilden, entfernen wir aus P die einzige ε -Regel $Z \rightarrow \varepsilon$ und fügen die Regeln $X \rightarrow a$ (wegen $X \rightarrow aS$), $Y \rightarrow b$ (wegen $Y \rightarrow bS$) und $Z \rightarrow c$ (wegen $Z \rightarrow cZ$) hinzu:

$$P': \quad S \rightarrow aY, bX, Z; \quad Y \rightarrow b, bS, aYY; \quad T \rightarrow U; \\ X \rightarrow a, aS, bXX; \quad Z \rightarrow c, S, T, cZ; \quad U \rightarrow abc.$$

◁

Als direkte Anwendung des obigen Satzes können wir die Inklusion der Klasse der Typ 2 Sprachen in der Klasse der Typ 1 Sprachen zeigen.

Korollar 76. $\text{REG} \not\subseteq \text{CFL} \subseteq \text{CSL} \subseteq \text{RE}$.

Beweis. Die Inklusionen $\text{REG} \subseteq \text{CFL}$ und $\text{CSL} \subseteq \text{RE}$ sind klar. Wegen $\{a^n b^n | n \geq 0\} \in \text{CFL} - \text{REG}$ ist die Inklusion $\text{REG} \subseteq \text{CFL}$ auch echt. Also ist nur noch die Inklusion $\text{CFL} \subseteq \text{CSL}$ zu zeigen. Nach obigem Satz ex. zu $L \in \text{CFL}$ eine kontextfreie Grammatik $G = (V, \Sigma, P, S)$ ohne ε -Produktionen mit $L(G) = L \setminus \{\varepsilon\}$. Da G dann auch kontextsensitiv ist, folgt hieraus im Fall $\varepsilon \notin L$ unmittelbar $L(G) = L \in \text{CSL}$. Im Fall $\varepsilon \in L$ erzeugt die kontextsensitive Grammatik

$$G' = (V \cup \{S'\}, \Sigma, P \cup \{S' \rightarrow S, \varepsilon\}, S')$$

die Sprache $L(G') = L$, d.h. $L \in \text{CSL}$. ■

Als nächstes zeigen wir folgende Abschlusseigenschaften der kontextfreien Sprachen.

Satz 77. Die Klasse CFL ist abgeschlossen unter Vereinigung, Produkt und Sternhülle.

Beweis. Seien $G_i = (V_i, \Sigma, P_i, S_i)$, $i = 1, 2$, kontextfreie Grammatiken für die Sprachen $L(G_i) = L_i$ mit $V_1 \cap V_2 = \emptyset$ und sei S eine neue Variable. Dann erzeugt die kontextfreie Grammatik

$$G_3 = (V_1 \cup V_2 \cup \{S\}, \Sigma, P_1 \cup P_2 \cup \{S \rightarrow S_1, S_2\}, S)$$

die Vereinigung $L(G_3) = L_1 \cup L_2$. Die Grammatik

$$G_4 = (V_1 \cup V_2 \cup \{S\}, \Sigma, P_1 \cup P_2 \cup \{S \rightarrow S_1 S_2\}, S)$$

erzeugt das Produkt $L(G_4) = L_1 L_2$ und die Sternhülle $(L_1)^*$ wird von der Grammatik

$$G_5 = (V_1 \cup \{S\}, \Sigma, P_1 \cup \{S \rightarrow S_1 S, \varepsilon\}, S)$$

erzeugt. ■

Offen bleibt zunächst, ob die kontextfreien Sprachen auch unter Schnitt und Komplement abgeschlossen sind. Um dies zu verneinen, müssen wir für bestimmte Sprachen nachweisen, dass sie nicht kontextfrei sind. Dies gelingt mit einem Pumping-Lemma für kontextfreie Sprachen.

Satz (Pumping-Lemma für kontextfreie Sprachen).

Zu jeder kontextfreien Sprache L gibt es eine Zahl l , so dass sich alle Wörter $z \in L$ mit $|z| \geq l$ in $z = uvwxy$ zerlegen lassen mit

1. $vx \neq \varepsilon$,
2. $|vwx| \leq l$ und
3. $uv^iwx^iy \in L$ für alle $i \geq 0$.

Für den Beweis benötigen wir Grammatiken in Chomsky-Normalform, die wir im nächsten Abschnitt behandeln werden.

Beispiel 78. Betrachte die Sprache $L = \{a^n b^n | n \geq 0\}$. Dann lässt sich jedes Wort $z = a^n b^n$ mit $|z| \geq 2$ pumpen: Zerlege $z = uvwxy$ mit $u = a^{n-1}$, $v = a$, $w = \varepsilon$, $x = b$ und $y = b^{n-1}$. ◁

Beispiel 79. Die Sprache $\{a^n b^n c^n | n \geq 0\}$ ist nicht kontextfrei. Für eine vorgegebene Zahl $l \geq 0$ hat nämlich $z = a^l b^l c^l$ die Länge $|z| = 3l \geq l$. Dieses Wort lässt sich aber nicht pumpen, da für jede Zerlegung $z = uvwxy$ mit $vx \neq \varepsilon$ und $|vwx| \leq l$ das Wort $z' = uv^2wx^2y$ nicht zu L gehört:

- Wegen $vx \neq \varepsilon$ ist $|z| < |z'|$.
- Wegen $|vwx| \leq l$ kann in vx nicht jedes der drei Zeichen a, b, c vorkommen.
- Kommt aber in vx beispielsweise kein a vor, so ist

$$\#_a(z') = \#_a(z) = l = |z|/3 < |z'|/3,$$

also kann z' nicht zu L gehören. ◁

Die Chomsky-Normalform ist auch Grundlage für einen effizienten Algorithmus zur Lösung des Wortproblems für kontextfreie Grammatiken, das wie folgt definiert ist.

Wortproblem für kontextfreie Grammatiken:

Gegeben: Eine kontextfreie Grammatik G und ein Wort x .

Gefragt: Ist $x \in L(G)$?

Satz. Das Wortproblem für kontextfreie Grammatiken ist effizient entscheidbar.

3.1 Chomsky-Normalform

Definition 80. Eine Grammatik (V, Σ, P, S) ist in **Chomsky-Normalform (CNF)**, falls $P \subseteq V \times (V^2 \cup \Sigma)$ ist, also alle Regeln die Form $A \rightarrow BC$ oder $A \rightarrow a$ haben.

Um eine kontextfreie Grammatik in Chomsky-Normalform zu bringen, müssen wir neben den ε -Regeln $A \rightarrow \varepsilon$ auch sämtliche Variablenumbenennungen $A \rightarrow B$ loswerden.

Definition 81. Regeln der Form $A \rightarrow B$ heißen **Variablenumbenennungen**.

Satz 82. Zu jeder kontextfreien Grammatik G ex. eine kontextfreie Grammatik G' ohne Variablenumbenennungen mit $L(G') = L(G)$.

Beweis. Zuerst entfernen wir sukzessive alle Zyklen

$$A_1 \rightarrow A_2 \rightarrow \dots \rightarrow A_k \rightarrow A_1,$$

indem wir diese Regeln aus P entfernen und alle übrigen Vorkommen der Variablen A_2, \dots, A_k durch A_1 ersetzen. Falls sich unter den entfernten Variablen A_2, \dots, A_k die Startvariable S befindet, sei A_1 die neue Startvariable.

Nun entfernen wir sukzessive die restlichen Variablenumbenennungen, indem wir

- eine Regel $A \rightarrow B$ wählen, so dass in P keine Variablenumbenennung $B \rightarrow C$ mit B auf der rechten Seite existiert,
- diese Regel $A \rightarrow B$ aus P entfernen und
- für jede Regel $B \rightarrow \alpha$ in P die Regel $A \rightarrow \alpha$ zu P hinzunehmen. ■

Beispiel 83. Ausgehend von den Produktionen

$$\begin{aligned} P: S &\rightarrow aY, bX, Z; & Y &\rightarrow b, bS, aYY; & T &\rightarrow U; \\ X &\rightarrow a, aS, bXX; & Z &\rightarrow c, S, T, cZ; & U &\rightarrow abc \end{aligned}$$

entfernen wir den Zyklus $S \rightarrow Z \rightarrow S$, indem wir die Regeln $S \rightarrow Z$ und $Z \rightarrow S$ entfernen und dafür die Produktionen $S \rightarrow c, T, cS$ (wegen $Z \rightarrow c, T, cZ$) hinzunehmen:

$$\begin{aligned} S &\rightarrow aY, bX, c, T, cS; & Y &\rightarrow b, bS, aYY; & T &\rightarrow U; \\ X &\rightarrow a, aS, bXX; & U &\rightarrow abc. \end{aligned}$$

Nun entfernen wir die Regel $T \rightarrow U$ und fügen die Regel $T \rightarrow abc$ (wegen $U \rightarrow abc$) hinzu:

$$\begin{aligned} S &\rightarrow aY, bX, c, T, cS; & Y &\rightarrow b, bS, aYY; & T &\rightarrow abc; \\ X &\rightarrow a, aS, bXX; & U &\rightarrow abc. \end{aligned}$$

Als nächstes entfernen wir dann auch die Regel $S \rightarrow T$ und fügen die Regel $S \rightarrow abc$ (wegen $T \rightarrow abc$) hinzu:

$$\begin{aligned} S &\rightarrow abc, aY, bX, c, cS; & Y &\rightarrow b, bS, aYY; & T &\rightarrow abc; \\ X &\rightarrow a, aS, bXX; & U &\rightarrow abc. \end{aligned}$$

Da T und U nun nirgends mehr auf der rechten Seite vorkommen, können wir die Regeln $T \rightarrow abc$ und $U \rightarrow abc$ weglassen:

$$S \rightarrow abc, aY, bX, c, cS; Y \rightarrow b, bS, aYY; X \rightarrow a, aS, bXX. \quad \triangleleft$$

Nach diesen Vorarbeiten ist es nun leicht, eine gegebene kontextfreie Grammatik in Chomsky-Normalform umzuwandeln.

Satz 84. *Zu jeder kontextfreien Sprache $L \in \text{CFL}$ gibt es eine CNF-Grammatik G' mit $L(G') = L \setminus \{\varepsilon\}$.*

Beweis. Aufgrund der beiden vorigen Sätze hat $L \setminus \{\varepsilon\}$ eine kontextfreie Grammatik $G = (V, \Sigma, P, S)$ ohne ε -Produktionen und ohne Variablenumbenennungen. Wir transformieren G wie folgt in eine CNF-Grammatik.

- Füge für jedes Terminalsymbol $a \in \Sigma$ eine neue Variable X_a zu V und eine neue Regel $X_a \rightarrow a$ zu P hinzu.
- Ersetze alle Vorkommen von a durch X_a , außer wenn a alleine auf der rechten Seite einer Regel steht.
- Ersetze jede Regel $A \rightarrow B_1 \dots B_k$, $k \geq 3$, durch die $k - 1$ Regeln

$$A \rightarrow B_1 A_1, A_1 \rightarrow B_2 A_2, \dots, A_{k-3} \rightarrow B_{k-2} A_{k-2}, A_{k-2} \rightarrow B_{k-1} B_k,$$

wobei A_1, \dots, A_{k-2} neue Variablen sind. ■

Beispiel 85. *In der Produktionsmenge*

$$P: S \rightarrow abc, aY, bX, c, cS; X \rightarrow a, aS, bXX; Y \rightarrow b, bS, aYY$$

ersetzen wir die Terminalsymbole a , b und c durch die Variablen A , B und C (außer wenn sie alleine auf der rechten Seite einer Regel vorkommen) und fügen die Regeln $A \rightarrow a$, $B \rightarrow b$, $C \rightarrow c$ hinzu:

$$S \rightarrow c, ABC, AY, BX, CS; X \rightarrow a, AS, BXX;$$

$$Y \rightarrow b, BS, AYY; A \rightarrow a; B \rightarrow b; C \rightarrow c.$$

Ersetze nun die Regeln $S \rightarrow ABC$, $X \rightarrow BXX$ und $Y \rightarrow AYY$ durch die Regeln $S \rightarrow AS'$, $S' \rightarrow BC$, $X \rightarrow BX'$, $X' \rightarrow XX'$ und $Y \rightarrow AY'$, $Y' \rightarrow YY'$:

$$S \rightarrow c, AS', AY, BX, CS; S' \rightarrow BC;$$

$$X \rightarrow a, AS, BX'; X' \rightarrow XX'; Y \rightarrow b, BS, AY'; Y' \rightarrow YY';$$

$$A \rightarrow a; B \rightarrow b; C \rightarrow c. \quad \triangleleft$$

Eine interessante Frage ist, ob in einer kontextfreien Grammatik G jedes Wort $x \in L(G)$ "eindeutig" ableitbar ist. Es ist klar, dass in diesem Kontext Ableitungen, die sich nur in der Reihenfolge der Regelanwendungen unterscheiden, nicht als verschieden betrachtet werden sollten. Dies erreichen wir dadurch, dass wir die Reihenfolge der Regelanwendungen festlegen.

Definition 86. *Sei $G = (V, \Sigma, P, S)$ eine kontextfreie Grammatik.*

a) *Eine Ableitung*

$$\alpha_0 = l_0 \underline{A_0} r_0 \Rightarrow l_1 \underline{A_1} r_1 \Rightarrow \dots \Rightarrow l_{m-1} \underline{A_{m-1}} r_{m-1} \Rightarrow \alpha_m.$$

heißt **Linksableitung** von α (kurz $\alpha_0 \Rightarrow_L^* \alpha_m$), falls in jedem Ableitungsschritt die am weitesten links stehende Variable ersetzt wird, d.h. es gilt $l_i \in \Sigma^*$ für $i = 0, \dots, m - 1$.

b) **Rechtsableitungen** $\alpha_0 \Rightarrow_R^* \alpha_m$ sind analog definiert.

c) G heißt **mehrdeutig**, wenn es ein Wort $x \in L(G)$ gibt, das zwei verschiedene Linksableitungen $S \Rightarrow_L^* x$ hat. Andernfalls heißt G **eindeutig**.

Offenbar gelten für alle Wörter $x \in \Sigma^*$ folgende Äquivalenzen:

$$x \in L(G) \Leftrightarrow S \Rightarrow^* x \Leftrightarrow S \Rightarrow_L^* x \Leftrightarrow S \Rightarrow_R^* x.$$

Beispiel 87. *Wir betrachten die Grammatik $G = (\{S\}, \{a, b\}, \{S \rightarrow aSbS, \varepsilon\}, S)$. Offenbar hat das Wort $aabb$ in G acht verschiedene Ableitungen, die sich allerdings nur in der Reihenfolge der Regelan-*

wendungen unterscheiden:

$$\begin{aligned} \underline{S} &\Rightarrow a\underline{S}bS \Rightarrow aa\underline{S}bSbS \Rightarrow aa\underline{S}bbS \Rightarrow aabb\underline{S} \Rightarrow aabb \\ \underline{S} &\Rightarrow a\underline{S}bS \Rightarrow aa\underline{S}b\underline{S}bS \Rightarrow aa\underline{S}bb\underline{S} \Rightarrow aa\underline{S}bb \Rightarrow aabb \\ \underline{S} &\Rightarrow a\underline{S}bS \Rightarrow aa\underline{S}bSbS \Rightarrow aab\underline{S}bS \Rightarrow aabb\underline{S} \Rightarrow aabb \\ \underline{S} &\Rightarrow a\underline{S}bS \Rightarrow aa\underline{S}bSbS \Rightarrow aab\underline{S}b\underline{S} \Rightarrow aab\underline{S}b \Rightarrow aabb \\ \underline{S} &\Rightarrow a\underline{S}bS \Rightarrow aa\underline{S}bSb\underline{S} \Rightarrow aa\underline{S}bSb \Rightarrow aab\underline{S}b \Rightarrow aabb \\ \underline{S} &\Rightarrow a\underline{S}bS \Rightarrow aa\underline{S}bSb\underline{S} \Rightarrow aa\underline{S}b\underline{S}b \Rightarrow aa\underline{S}bb \Rightarrow aabb \\ \underline{S} &\Rightarrow a\underline{S}b\underline{S} \Rightarrow a\underline{S}b \Rightarrow aa\underline{S}bSb \Rightarrow aab\underline{S}b \Rightarrow aabb \\ \underline{S} &\Rightarrow a\underline{S}b\underline{S} \Rightarrow a\underline{S}b \Rightarrow aa\underline{S}b\underline{S}b \Rightarrow aa\underline{S}bb \Rightarrow aabb. \end{aligned}$$

Darunter sind genau eine Links- und genau eine Rechtsableitung:

$$\underline{S} \Rightarrow_L a\underline{S}bS \Rightarrow_L aa\underline{S}bSbS \Rightarrow_L aab\underline{S}bS \Rightarrow_L aabb\underline{S} \Rightarrow_L aabb$$

und

$$\underline{S} \Rightarrow_R a\underline{S}b\underline{S} \Rightarrow_R a\underline{S}b \Rightarrow_R aa\underline{S}bSb \Rightarrow_R aa\underline{S}bb \Rightarrow_R aabb.$$

Die Grammatik G ist eindeutig. Dies liegt daran, dass in jeder Satzform $\alpha S \beta$ von G das Suffix β entweder leer ist oder mit einem b beginnt. Daher muss jede Linksableitung eines Wortes $x \in L(G)$ die am weitesten links stehende Variable der aktuellen Satzform $\alpha S \beta$ genau dann nach $aSbS$ expandieren, falls das Präfix α in x von einem a gefolgt wird.

Dagegen ist die Grammatik $G' = (\{S\}, \{a, b\}, \{S \rightarrow aSbS, ab, \varepsilon\}, S)$ mehrdeutig, da das Wort $x = ab$ zwei Linksableitungen hat:

$$\underline{S} \Rightarrow_L ab \text{ und } \underline{S} \Rightarrow_L a\underline{S}bS \Rightarrow_L ab\underline{S} \Rightarrow_L ab. \quad \triangleleft$$

Ableitungen in einer kontextfreien Grammatik lassen sich graphisch sehr gut durch einen Syntaxbaum (auch **Ableitungsbaum** genannt, engl. *parse tree*) veranschaulichen.

Definition 88. Sei $G = (V, E)$ ein Digraph.

- Ein **v_0 - v_k -Weg** in G ist eine Folge von Knoten v_0, \dots, v_k mit $(v_i, v_{i+1}) \in E$ für $i = 0, \dots, k-1$. Seine **Länge** ist k .
- Ein Weg heißt **einfach** oder **Pfad**, falls alle seine Knoten paarweise verschieden sind.
- Ein u - v -Weg der Länge ≥ 1 mit $u = v$ heißt **Zyklus**.
- G heißt **azyklisch**, wenn es in G keinen Zyklus gibt.
- G heißt **gerichteter Wald**, wenn G azyklisch ist und jeder Knoten $v \in V$ Eingangsgrad $\deg^-(v) \leq 1$ hat.
- Ein Knoten $u \in V$ vom Ausgangsgrad $\deg^+(u) = 0$ heißt **Blatt**.
- Ein Knoten $w \in V$ heißt **Wurzel** von G , falls alle Knoten $v \in V$ von w aus erreichbar sind (d.h. es gibt einen w - v -Weg in G).
- Ein **gerichteter Wald**, der eine Wurzel hat, heißt **gerichteter Baum**.
- Da die Kantenrichtungen durch die Wahl der Wurzel eindeutig bestimmt sind, kann auf ihre Angabe verzichtet werden. Man spricht dann auch von einem **Wurzelbaum**.

Definition 89. Wir ordnen einer Ableitung

$$\underline{A_0} \Rightarrow l_1 \underline{A_1} r_1 \Rightarrow \dots \Rightarrow l_{m-1} \underline{A_{m-1}} r_{m-1} \Rightarrow \alpha_m.$$

den Syntaxbaum T_m zu, wobei die Bäume T_0, \dots, T_m induktiv wie folgt definiert sind:

- T_0 besteht aus einem einzigen Knoten, der mit A_0 markiert ist.
- Wird im $(i+1)$ -ten Ableitungsschritt die Regel $A_i \rightarrow v_1 \dots v_k$ mit $v_j \in \Sigma \cup V$ für $j = 1, \dots, k$ angewandt, so entsteht T_{i+1} aus T_i , indem wir das Blatt A_i in T_i durch folgenden Unterbaum ersetzen:

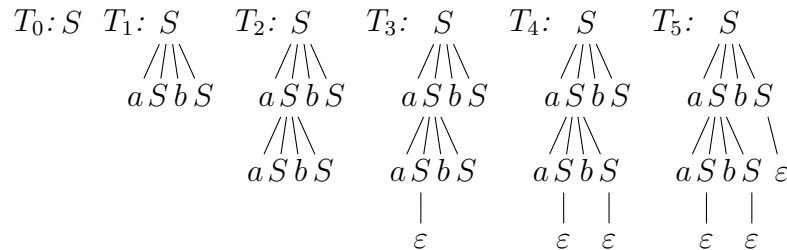
$$k > 0: \begin{array}{c} A_i \\ / \quad \backslash \\ v_1 \quad \dots \quad v_k \end{array} \quad k = 0: \begin{array}{c} A_i \\ | \\ \varepsilon \end{array}$$

- Hierbei stellen wir uns die Kanten von oben nach unten gerichtet und die Kinder $v_1 \dots v_k$ von links nach rechts geordnet vor.

Beispiel 90. Betrachte die Grammatik $G = (\{S\}, \{a, b\}, \{S \rightarrow aSbS, \varepsilon\}, S)$ und die Ableitung

$$\underline{S} \Rightarrow a\underline{S}bS \Rightarrow aaSb\underline{S}bS \Rightarrow aa\underline{S}bbS \Rightarrow aabb\underline{S} \Rightarrow aabb.$$

Die zugehörigen Syntaxbäume sind dann



Die Satzform α_i ergibt sich aus T_i , indem wir die Blätter von T_i von links nach rechts zu einem Wort zusammensetzen. \triangleleft

Bemerkung 91.

- Aus einem Syntaxbaum ist die zugehörige Linksableitung eindeutig rekonstruierbar. Daher führen unterschiedliche Linksableitungen auch auf unterschiedliche Syntaxbäume. Linksableitungen und Syntaxbäume entsprechen sich also eineindeutig. Ebenso Rechtsableitungen und Syntaxbäume.
- Ist T Syntaxbaum einer CNF-Grammatik, so hat jeder Knoten in T höchstens zwei Kinder (d.h. T ist ein Binärbaum).

3.2 Das Pumping-Lemma für kontextfreie Sprachen

In diesem Abschnitt beweisen wir das Pumping-Lemma für kontextfreie Sprachen. Dabei nutzen wir die Tatsache aus, dass die Syntaxbäume einer CNF-Grammatik Binäräume sind.

Definition 92. Die **Tiefe** eines Baumes mit Wurzel w ist die maximale Pfadlänge von w zu einem Blatt.

Lemma 93. Ein Binärbaum B der Tiefe k hat höchstens 2^k Blätter.

Beweis. Wir führen den Beweis durch Induktion über k .

$k = 0$: Ein Baum der Tiefe 0 kann nur einen Knoten haben.

$k \rightsquigarrow k + 1$: Sei B ein Binärbaum der Tiefe $k + 1$. Dann hängen an B 's Wurzel maximal zwei Teilbäume. Da deren Tiefe $\leq k$ ist, haben sie nach IV höchstens 2^k Blätter. Also hat $B \leq 2^{k+1}$ Blätter. \blacksquare

Korollar 94. Ein Binärbaum B mit mehr als 2^{k-1} Blättern hat mindestens Tiefe k .

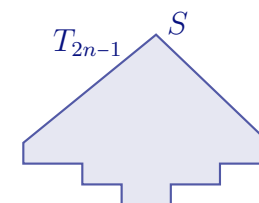
Beweis. Würde B mehr als 2^{k-1} Blätter und eine Tiefe $\leq k - 1$ besitzen, so würde dies im Widerspruch zu Lemma 93 stehen. \blacksquare

Satz 95 (Pumping-Lemma für kontextfreie Sprachen).

Zu jeder kontextfreien Sprache L gibt es eine Zahl l , so dass sich alle Wörter $z \in L$ mit $|z| \geq l$ in $z = uvwx$ zerlegen lassen mit

1. $vx \neq \varepsilon$,
2. $|vwx| \leq l$ und
3. $uv^iwx^i \in L$ für alle $i \geq 0$.

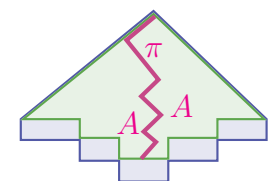
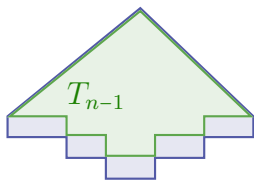
Beweis. Sei $G = (V, \Sigma, P, S)$ eine CNF-Grammatik für $L \setminus \{\varepsilon\}$. Dann gibt es in G für jedes Wort $z = z_1 \dots z_n \in L$ mit $n \geq 1$, eine Ableitung



$$S = \alpha_0 \Rightarrow \alpha_1 \dots \Rightarrow \alpha_m = z.$$

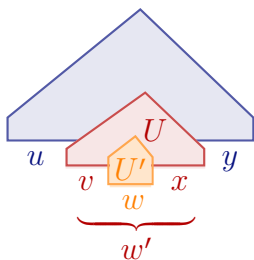
Da G in CNF ist, werden hierbei $n - 1$ Regeln der Form $A \rightarrow BC$ und n Regeln der Form $A \rightarrow a$ angewandt, d.h. $m = 2n - 1$ und z hat den Syntaxbaum T_{2n-1} . Wir können annehmen,

dass zuerst alle Regeln der Form $A \rightarrow BC$ und danach die Regeln der Form $A \rightarrow a$ zur Anwendung kommen. Dann besteht die Satzform α_{n-1} aus n Variablen und der Syntaxbaum T_{n-1} hat ebenfalls n Blätter. Setzen wir $l = 2^k$, wobei $k = \|V\|$ ist, so hat T_{n-1} im Fall $n \geq l$ mindestens

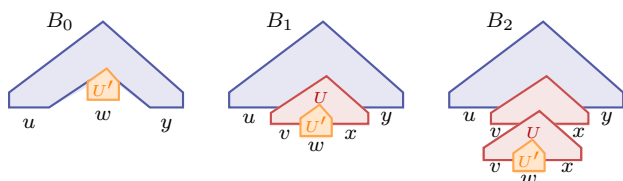


$l = 2^k > 2^{k-1}$ Blätter und daher mindestens die Tiefe k . Sei π ein von der Wurzel ausgehender Pfad maximaler Länge in T_{n-1} . Dann hat π die Länge $\geq k$ und unter den letzten $k + 1$ Knoten von π müssen zwei mit derselben Variablen A markiert sein.

Seien U und U' die von diesen Knoten ausgehenden Unterbäume des vollständigen Syntaxbaums T_{2n-1} . Nun zerlegen wir z wie folgt. w' ist das Teilwort von $z = uw'y$, das von U erzeugt wird und w ist das Teilwort von $w' = vwx$, das von U' erzeugt wird. Jetzt bleibt nur noch zu zeigen, dass diese Zerlegung die geforderten 3 Eigenschaften erfüllt.



- Da U mehr Blätter hat als U' , ist $vx \neq \varepsilon$ (Bedingung 1).
- Da der Baum $U^* = U \cap T_{n-1}$ die Tiefe $\leq k$ hat (andernfalls wäre π nicht maximal), hat U^* höchstens $2^k = l$ Blätter. Da U^* genau $|vwx|$ Blätter hat, folgt $|vwx| \leq l$ (Bedingung 2).
- Für den Nachweis von Bedingung 3 lassen sich schließlich Syntaxbäume B^i für die Wörter w^iwx^iy , $i \geq 0$, wie folgt konstruieren:



B^0 entsteht also aus $B^1 = T_{2n-1}$, indem wir U durch U' ersetzen, und B^{i+1} entsteht aus B^i , indem wir U' durch U ersetzen. ■

Satz 96. Die Klasse CFL ist nicht abgeschlossen unter Schnitt und Komplement.

Beweis. Die beiden Sprachen

$$L_1 = \{a^n b^m c^m \mid n, m \geq 0\} \quad \text{und} \quad L_2 = \{a^n b^n c^m \mid n, m \geq 0\}$$

sind kontextfrei. Nicht jedoch $L_1 \cap L_2 = \{a^n b^n c^n \mid n \geq 0\}$. Also ist CFL nicht unter Schnitt abgeschlossen.

Da CFL zwar unter Vereinigung aber nicht unter Schnitt abgeschlossen ist, kann CFL wegen de Morgan nicht unter Komplementbildung abgeschlossen sein. ■

3.3 Der CYK-Algorithmus

In diesem Abschnitt stellen wir den bereits angekündigten effizienten Algorithmus zur Lösung des Wortproblems für kontextfreie Grammatiken vor.

Wortproblem für kontextfreie Grammatiken:

Gegeben: Eine kontextfreie Grammatik G und ein Wort x .
Gefragt: Ist $x \in L(G)$?

Wir lösen das Wortproblem, indem wir G zunächst in Chomsky-Normalform bringen und dann den nach seinen Autoren Cocke, Younger und Kasami benannten CYK-Algorithmus anwenden, welcher auf dem Prinzip der Dynamischen Programmierung beruht.

Satz 97. Das Wortproblem für kontextfreie Grammatiken ist effizient entscheidbar.

Beweis. Seien eine Grammatik $G = (V, \Sigma, P, S)$ und ein Wort $x = x_1 \dots x_n$ gegeben. Falls $x = \varepsilon$ ist, können wir effizient prüfen, ob $S \Rightarrow^* \varepsilon$

gilt. Andernfalls transformieren wir G in eine CNF-Grammatik G' für die Sprache $L(G) \setminus \{\varepsilon\}$. Chomsky-Normalform. Es lässt sich leicht verifizieren, dass die nötigen Umformungsschritte effizient ausführbar sind. Nun setzen wir den CYK-Algorithmus auf das Paar (G', x) an, der die Zugehörigkeit von x zu $L(G')$ wie folgt entscheidet.

Bestimme für $l = 1, \dots, n$ und $k = 1, \dots, n - l + 1$ die Menge

$$V_{l,k}(x) = \{A \in V \mid A \Rightarrow^* x_k \dots x_{k+l-1}\}$$

aller Variablen, aus denen das an Position k beginnende Teilwort $x_k \dots x_{k+l-1}$ von x der Länge l ableitbar ist. Dann gilt offensichtlich

$$x \in L(G') \Leftrightarrow S \in V_{n,1}(x).$$

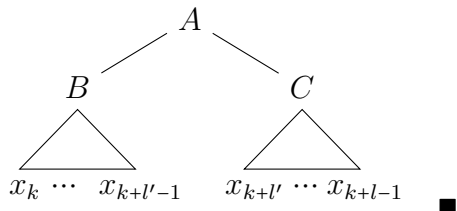
Für $l = 1$ ist

$$V_{1,k}(x) = \{A \in V \mid A \rightarrow x_k\}$$

und für $l = 2, \dots, n$ ist

$$V_{l,k}(x) = \{A \in V \mid \exists l' < l \exists B \in V_{l',k}(x) \exists C \in V_{l-l',k+l'}(x) : A \rightarrow BC\}.$$

Eine Variable A gehört also genau dann zu $V_{l,k}(x)$, $l \geq 2$, falls eine Zahl $l' \in \{1, \dots, l - 1\}$ und eine Regel $A \rightarrow BC$ existieren, so dass $B \in V_{l',k}(x)$ und $C \in V_{l-l',k+l'}(x)$ sind.



Algorithmus CYK(G, x)

```

1 Input: CNF-Grammatik  $G = (V, \Sigma, P, S)$  und ein Wort  $x = x_1 \dots x_n$ 
2 for  $k := 1$  to  $n$  do
3    $V_{1,k} := \{A \in V \mid A \rightarrow x_k \in P\}$ 
4 for  $l := 2$  to  $n$  do
5   for  $k := 1$  to  $n - l + 1$  do
6      $V_{l,k} := \emptyset$ 

```

```

7   for  $l' := 1$  to  $l - 1$  do
8     for all  $A \rightarrow BC \in P$  do
9       if  $B \in V_{l',k}$  and  $C \in V_{l-l',k+l'}$  then
10         $V_{l,k} := V_{l,k} \cup \{A\}$ 
11 if  $S \in V_{n,1}$  then accept else reject

```

Der CYK-Algorithmus lässt sich leicht dahingehend modifizieren, dass er im Fall $x \in L(G)$ auch einen Syntaxbaum T von x ausgibt. Hierzu genügt es, zu jeder Variablen A in $V_{l,k}$ den Wert von l' und die Regel $A \rightarrow BC$ zu speichern, die zur Aufnahme von A in $V_{l,k}$ geführt haben. Im Fall $S \in V_{n,1}(x)$ lässt sich dann mithilfe dieser Information leicht ein Syntaxbaum T von x konstruieren.

Beispiel 98. Betrachte die CNF-Grammatik mit den Produktionen

$$S \rightarrow AS', AY, BX, CS, c; \quad S' \rightarrow BC; \quad X \rightarrow AS, BX', a; \quad X' \rightarrow XX;$$

$$Y \rightarrow BS, AY', b; \quad Y' \rightarrow YY; \quad A \rightarrow a; \quad B \rightarrow b; \quad C \rightarrow c.$$

Dann erhalten wir für das Wort $x = abb$ folgende Mengen $V_{l,k}$:

$x_k:$	a	b	b
$l:1$	$\{X, A\}$	$\{Y, B\}$	$\{Y, B\}$
2	$\{S\}$	$\{Y'\}$	
3	$\{Y\}$		

Wegen $S \notin V_{3,1}(abb)$ ist $x \notin L(G)$.

Dagegen gehört das Wort $y = aababb$ wegen $S \in V_{6,1}(aababb)$ zu $L(G)$:

	a	a	b	a	b	b
	$\{X, A\}$	$\{X, A\}$	$\{Y, B\}$	$\{X, A\}$	$\{Y, B\}$	$\{Y, B\}$
	$\{X'\}$	$\{S\}$	$\{S\}$	$\{S\}$	$\{Y'\}$	
	$\{X\}$	$\{X\}$	$\{Y\}$	$\{Y\}$		
	$\{X'\}$	$\{S\}$	$\{Y'\}$			
	$\{X\}$	$\{Y\}$				
	$\{S\}$					

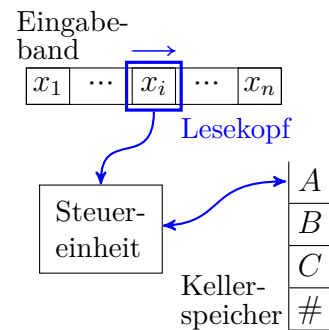


3.4 Kellerautomaten

Wie müssen wir das Maschinenmodell des DFA erweitern, damit die Sprache $L = \{a^n b^n \mid n \geq 0\}$ und alle anderen kontextfreien Sprachen erkannt werden können? Dass ein DFA die Sprache $L = \{a^n b^n \mid n \geq 0\}$ nicht erkennen kann, liegt an seinem beschränkten Speichervermögen, das zwar von L aber nicht von der Eingabe abhängen darf.

Um L erkennen zu können, reicht bereits ein so genannter Kellerspeicher (Stapel, engl. *stack*, *pushdown memory*) aus. Dieser erlaubt nur den Zugriff auf die höchste belegte Speicheradresse. Ein Kellerautomat

- verfügt über einen Kellerspeicher,
- kann ε -Übergänge machen,
- liest in jedem Schritt das aktuelle Eingabezeichen und das oberste Kellersymbol,
- kann das oberste Kellersymbol entfernen (durch eine **pop-Operation**) und
- durch beliebig viele Symbole ersetzen (durch eine **push-Operation**).



Für eine Menge M bezeichne $\mathcal{P}_e(M)$ die Menge aller endlichen Teilmengen von M , d.h.

$$\mathcal{P}_e(M) = \{A \subseteq M \mid A \text{ ist endlich}\}.$$

Definition 99. Ein **Kellerautomat** (kurz: PDA; pushdown automaton) wird durch ein 6-Tupel $M = (Z, \Sigma, \Gamma, \delta, q_0, \#)$ beschrieben, wobei

- $Z \neq \emptyset$ eine endliche Menge von **Zuständen**,
- Σ das **Eingabealphabet**,

- Γ das **Kelleralphabet**,
- $\delta : Z \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \rightarrow \mathcal{P}_e(Z \times \Gamma^*)$ die **Überföhrungsfunktion**,
- $q_0 \in Z$ der **Startzustand** und
- $\# \in \Gamma$ das **Kelleranfangszeichen** ist.

Wenn q der momentane Zustand, A das oberste Kellerzeichen und $u \in \Sigma$ das nächste Eingabezeichen (bzw. $u = \varepsilon$) ist, so kann M im Fall $(p, B_1 \dots B_k) \in \delta(q, u, A)$

- in den Zustand p wechseln,
- den Lesekopf auf dem Eingabeband um $|u|$ Positionen vorrücken und
- das Zeichen A im Keller durch die Zeichenfolge $B_1 \dots B_k$ ersetzen.

Hierfür sagen wir auch, M führt die **Anweisung** $quA \rightarrow pB_1 \dots B_k$ aus. Da im Fall $u = \varepsilon$ kein Eingabezeichen gelesen wird, spricht man auch von einem **spontanen** Übergang (oder **ε -Übergang**). Eine **Konfiguration** wird durch ein Tripel

$$K = (q, x_i \dots x_n, A_1 \dots A_l) \in Z \times \Sigma^* \times \Gamma^*$$

beschrieben und besagt, dass

- q der momentane Zustand,
- $x_i \dots x_n$ der ungelesene Rest der Eingabe und
- $A_1 \dots A_l$ der aktuelle Kellerinhalt ist (A_1 steht oben).

Eine Anweisung $quA_1 \rightarrow pB_1 \dots B_k$ (mit $u \in \{\varepsilon, x_i\}$) überföhrt die Konfiguration K in die **Folgekonfiguration**

$$K' = (p, x_j \dots x_n, B_1 \dots B_k A_2 \dots A_l) \text{ mit } j = i + |u|.$$

Hierfür schreiben wir auch kurz $K \vdash K'$. Eine **Rechnung** von M bei Eingabe x ist eine Folge von Konfigurationen $K_0, K_1, K_2 \dots$ mit $K_0 = (q_0, x, \#)$ und $K_0 \vdash K_1 \vdash K_2 \dots$. K_0 heißt **Startkonfiguration**

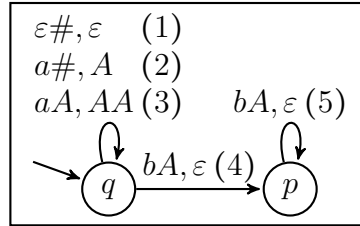
von M bei Eingabe x . Die reflexive, transitive Hülle von \vdash bezeichnen wir wie üblich mit \vdash^* . Die von M **akzeptierte** oder **erkannte Sprache** ist

$$L(M) = \{x \in \Sigma^* \mid \exists p \in Z : (q_0, x, \#) \vdash^* (p, \varepsilon, \varepsilon)\}.$$

Ein Wort x wird also genau dann von M akzeptiert, wenn es eine Rechnung gibt, bei der M das gesamte Eingabewort bis zum Ende liest und den Keller leert. Man beachte, dass bei leerem Keller kein weiterer Übergang mehr möglich ist.

Beispiel 100. Sei $M = (Z, \Sigma, \Gamma, \delta, q, \#)$ ein PDA mit $Z = \{q, p\}$, $\Sigma = \{a, b\}$, $\Gamma = \{A, \#\}$ und den Anweisungen

$$\begin{aligned} \delta : q\varepsilon\# \rightarrow q & \quad (1) & qa\# \rightarrow qA & \quad (2) \\ qaA \rightarrow qAA & \quad (3) & qbA \rightarrow p & \quad (4) \\ pbA \rightarrow p & \quad (5) \end{aligned}$$



Dann akzeptiert M die Eingabe $aabb$:

$$(q, aabb, \#) \vdash_{(2)} (q, abb, A) \vdash_{(3)} (q, bb, AA) \vdash_{(4)} (p, b, A) \vdash_{(5)} (p, \varepsilon, \varepsilon).$$

Allgemeiner akzeptiert M das Wort $x = a^n b^n$ mit folgender Rechnung:

$$n = 0: (q, \varepsilon, \#) \vdash_{(1)} (p, \varepsilon, \varepsilon).$$

$$\begin{aligned} n \geq 1: (q, a^n b^n, \#) & \vdash_{(2)} (q, a^{n-1} b^n, A) \vdash_{(3)}^{n-1} (q, b^n, A^n) \\ & \vdash_{(4)} (p, b^{n-1}, A^{n-1}) \vdash_{(5)}^{n-1} (p, \varepsilon, \varepsilon). \end{aligned}$$

Dies zeigt $\{a^n b^n \mid n \geq 0\} \subseteq L(M)$. Als nächstes zeigen wir, dass jede von M akzeptierte Eingabe $x = x_1 \dots x_n$ die Form $x = a^m b^m$ hat.

Ausgehend von der Startkonfiguration $(q, x, \#)$ sind nur die Anweisungen (1) oder (2) ausführbar. Falls M Anweisung (1) wählt, wird der Keller geleert. Daher kann M in diesem Fall nur das leere Wort $x = \varepsilon = a^0 b^0$ akzeptieren.

Falls die akzeptierende Rechnung mit Anweisung (2) beginnt, muss $x_1 = a$ sein. Danach ist nur Anweisung (3) ausführbar, bis M das erste b liest:

$$\begin{aligned} (q, x_1 \dots x_n, \#) & \vdash_{(2)} (q, x_2 \dots x_n, A) \vdash_{(3)}^{m-1} (q, x_{m+1} \dots x_n, A^m) \\ & \vdash_{(4)} (p, x_{m+2} \dots x_n, A^{m-1}) \end{aligned}$$

mit $x_1 = x_2 = \dots = x_m = a$ und $x_{m+1} = b$. Damit M den Keller leeren kann, müssen jetzt noch genau $m - 1$ b 's kommen, weshalb x auch in diesem Fall die Form $a^m b^m$ hat. \triangleleft

Als nächstes zeigen wir, dass PDAs genau die kontextfreien Sprachen erkennen.

Satz 101. $CFL = \{L(M) \mid M \text{ ist ein PDA}\}.$

Beweis. Wir zeigen zuerst die Inklusion von links nach rechts.

Idee: Konstruiere zu einer kontextfreien Grammatik $G = (V, \Sigma, P, S)$ einen PDA $M = (\{q\}, \Sigma, \Gamma, \delta, q_0, S)$ mit $\Gamma = V \cup \Sigma$, so dass gilt:

$$S \Rightarrow_L^* x_1 \dots x_n \text{ gdw. } (q, x_1 \dots x_n, S) \vdash^* (q, \varepsilon, \varepsilon).$$

Hierzu fügen wir für jede Regel $A \rightarrow_G \alpha$ in P die Anweisung $q\varepsilon A \rightarrow q\alpha$ und für jedes $a \in \Sigma$ die Anweisung $qaa \rightarrow q\varepsilon$ zu δ hinzu.

M berechnet also nichtdeterministisch eine Linksableitung für die Eingabe x . Da M hierbei den Syntaxbaum von oben nach unten aufbaut, wird M als *Top-Down Parser* bezeichnet. Nun ist leicht zu sehen, dass sogar folgende Äquivalenz gilt:

$$S \Rightarrow_L^l x_1 \dots x_n \text{ gdw. } (q, x_1 \dots x_n, S) \vdash^{l+n} (q, \varepsilon, \varepsilon).$$

Daher folgt

$$x \in L(G) \Leftrightarrow S \Rightarrow_L^* x \Leftrightarrow (q, x, S) \vdash^* (q, \varepsilon, \varepsilon) \Leftrightarrow x \in L(M).$$

Als nächstes zeigen wir die Inklusion von rechts nach links.

Idee: Konstruiere zu einem PDA $M = (Z, \Sigma, \Gamma, \delta, q_0, \#)$ eine kontextfreie Grammatik $G = (V, \Sigma, P, S)$ mit Variablen X_{pAq} , $A \in \Gamma$, $p, q \in Z$, so dass folgende Äquivalenz gilt:

$$X_{pAq} \Rightarrow^* x \text{ gdw. } (p, x, A) \vdash^* (q, \varepsilon, \varepsilon). \quad (*)$$

Ein Wort x soll also genau dann in G aus X_{pAq} ableitbar sein, wenn M ausgehend vom Zustand p bei Lesen von x in den Zustand q gelangen kann und dabei das Zeichen A aus dem Keller entfernt. Um dies zu erreichen, fügen wir für jede Anweisung $puA \rightarrow p_1A_1 \dots A_k$, $k \geq 0$, die folgenden $\|Z\|^k$ Regeln zu P hinzu:

Für jede Zustandsfolge p_2, \dots, p_{k+1} : $X_{pAp_{k+1}} \rightarrow uX_{p_1A_1p_2} \dots X_{p_kA_kp_{k+1}}$.

Um damit alle Wörter $x \in L(M)$ aus S ableiten zu können, benötigen wir jetzt nur noch für jeden Zustand $q \in Z$ die Regel $S \rightarrow X_{q_0\#q}$. Die Variablenmenge von G ist also

$$V = \{S\} \cup \{X_{pAq} \mid p, q \in Z, A \in \Gamma\}$$

und P enthält neben den Regeln $S \rightarrow X_{q_0\#q}$, $q \in Z$, für jede Anweisung $puA \rightarrow p_1A_1 \dots A_k$, $k \geq 0$, von M und jede Zustandsfolge p_2, \dots, p_{k+1} die Regel $X_{pAp_{k+1}} \rightarrow uX_{p_1A_1p_2} \dots X_{p_kA_kp_{k+1}}$.

Unter der Voraussetzung, dass die Äquivalenz $(*)$ gilt, lässt sich nun leicht die Korrektheit von G zeigen. Es gilt

$$\begin{aligned} x \in L(M) &\Leftrightarrow (q_0, x, \#) \vdash^* (q, \varepsilon, \varepsilon) \text{ für ein } q \in Z \\ &\Leftrightarrow S \Rightarrow X_{q_0\#q} \Rightarrow^* x \text{ für ein } q \in Z \\ &\Leftrightarrow x \in L(G). \end{aligned}$$

Wir müssen also nur noch die Gültigkeit von $(*)$ zeigen. Hierzu zeigen wir durch Induktion über m für alle $p, q \in Z$, $A \in \Gamma$ und $x \in \Sigma^*$ folgende stärkere Behauptung:

$$X_{pAq} \Rightarrow^m x \text{ gdw. } (p, x, A) \vdash^m (q, \varepsilon, \varepsilon) \quad (**)$$

$m = 0$: Da sowohl $X_{pAq} \Rightarrow^0 x$ als auch $(p, x, A) \vdash^0 (q, \varepsilon, \varepsilon)$ falsch sind, ist die Äquivalenz $(**)$ für $m = 0$ erfüllt.

$m \rightsquigarrow m + 1$: Wir zeigen zuerst die Implikation von links nach rechts. Sei x aus X_{pAq} in $m + 1$ Schritten ableitbar und sei α die im ersten Schritt abgeleitete Satzform:

$$X_{pAq} \Rightarrow \alpha \Rightarrow^m x.$$

Wegen $X_{pAq} \rightarrow_G \alpha$ gibt es eine Anweisung $puA \rightarrow p_1A_1 \dots A_k$, $k \geq 0$, und Zustände $p_2, \dots, p_{k+1} \in Z$ mit

$$\alpha = u X_{p_1A_1p_2} \dots X_{p_kA_kp_{k+1}},$$

wobei $p_{k+1} = q$ ist. Wegen $\alpha \Rightarrow^m x$ ex. eine Zerlegung $x = uu_1 \dots u_k$ und Zahlen $m_i \geq 1$ mit $m_1 + \dots + m_k = m$ und

$$X_{p_iA_i p_{i+1}} \Rightarrow^{m_i} u_i \quad (i = 1, \dots, k).$$

Nach IV gibt es somit Rechnungen

$$(p_i, u_i, A_i) \vdash^{m_i} (p_{i+1}, \varepsilon, \varepsilon), \quad i = 1, \dots, k,$$

aus denen sich die gesuchte Rechnung der Länge $m + 1$ zusammensetzen lässt:

$$\begin{aligned} (p, uu_1 \dots u_k, A) \vdash & (p_1, u_1 \dots u_k, A_1 \dots A_k) \\ & \vdash^{m_1} (p_2, u_2 \dots u_k, A_2 \dots A_k) \\ & \vdots \\ & \vdash^{m_{k-1}} (p_k, u_k, A_k) \\ & \vdash^{m_k} (p_{k+1}, \varepsilon, \varepsilon). \end{aligned}$$

Sei nun umgekehrt eine Rechnung der Länge $m + 1$ gegeben und sei $puA \rightarrow p_1A_1 \dots A_k$ die erste in dieser Rechnung zur Ausführung kommende Anweisung (d.h. $x = ux'$):

$$(p, x, A) \vdash (p_1, x', A_1 \dots A_k) \vdash^m (q, \varepsilon, \varepsilon)$$

Zudem sei p_i für $i = 2, \dots, k + 1$ der erste Zustand in dieser Rechnung, in den M mit dem Kellerinhalt $A_i \dots A_k$ gelangt (d.h. $p_{k+1} = q$). Dann enthält P die Regel $X_{pAp_{k+1}} \rightarrow uX_{p_1A_1p_2} \dots X_{p_kA_kp_{k+1}}$. Weiter sei u_i für $i = 1, \dots, k$ das Teilwort von x' , das M zwischen den Besuchen von p_i und p_{i+1} liest.

Dann gibt es Zahlen $m_i \geq 1$ mit $m_1 + \dots + m_k = m$ und

$$(p_i, u_i, A_i) \vdash^{m_i} (p_{i+1}, \varepsilon, \varepsilon)$$

für $i = 1, \dots, k$. Nach IV gibt es daher Ableitungen

$$X_{p_iA_i p_{i+1}} \Rightarrow^{m_i} u_i, \quad i = 1, \dots, k,$$

die wir zu der gesuchten Ableitung zusammensetzen können:

$$\begin{aligned} X_{pAp_{k+1}} &\Rightarrow uX_{p_1A_1p_2} \dots X_{p_{k-1}A_{k-1}p_k} X_{p_kA_kp_{k+1}} \\ &\Rightarrow^{m_1} uu_1 \dots X_{p_{k-1}A_{k-1}p_k} X_{p_kA_kp_{k+1}} \\ &\vdots \\ &\Rightarrow^{m_{k-1}} uu_1 \dots u_{k-1} X_{p_kA_kp_{k+1}} \\ &\Rightarrow^{m_k} uu_1 \dots u_k = x. \end{aligned}$$

■

Beispiel 102. Sei $G = (\{S\}, \{a, b\}, P, S)$ mit

$$P: S \rightarrow aSbS, \quad (1) \quad S \rightarrow a. \quad (2)$$

Der zugehörige PDA besitzt dann die Anweisungen

$$\begin{aligned} \delta: \quad qaa \rightarrow q\varepsilon, \quad (0) \quad qbb \rightarrow q\varepsilon, \quad (0') \\ q\varepsilon S \rightarrow qaSbS, \quad (1') \quad q\varepsilon S \rightarrow qa. \quad (2') \end{aligned}$$

Der Linksableitung

$$\underline{S} \xRightarrow{(1)} a\underline{S}bS \xRightarrow{(2)} aab\underline{S} \xRightarrow{(2)} aaba$$

in G entspricht beispielsweise die akzeptierende Rechnung

$$\begin{aligned} (q, aaba, S) &\vdash_{(1')} (q, aaba, aSbS) \vdash_{(0)} (q, aba, SbS) \\ &\vdash_{(2')} (q, aba, abS) \vdash_{(0)} (q, ba, bS) \\ &\vdash_{(0')} (q, a, S) \vdash_{(2')} (q, a, a) \vdash_{(0)} (q, \varepsilon, \varepsilon) \end{aligned}$$

von M und umgekehrt. ◁

Beispiel 103. Sei M der PDA $(\{p, q\}, \{a, b\}, \{A, \#\}, \delta, p, \#)$ mit

$$\begin{aligned} \delta: p\varepsilon\# \rightarrow q\varepsilon, \quad (1) \quad paA \rightarrow pAA, \quad (3) \quad qbA \rightarrow q\varepsilon. \quad (5) \\ pa\# \rightarrow pA, \quad (2) \quad pbA \rightarrow q\varepsilon, \quad (4) \end{aligned}$$

Dann erhalten wir die Grammatik $G = (V, \Sigma, P, S)$ mit der Variablenmenge

$$V = \{S, X_{p\#p}, X_{p\#q}, X_{q\#p}, X_{q\#q}, X_{pAp}, X_{pAq}, X_{qAp}, X_{qAq}\}.$$

Die Regelmengemenge P enthält neben den beiden Startregeln

$$S \rightarrow X_{p\#p}, X_{p\#q} \quad (0, 0')$$

die folgenden Produktionen:

Anweisung	k	p_2, \dots, p_{k+1}	zugehörige Regel
$puA \rightarrow p_1A_1 \dots A_k$			$X_{pAp_{k+1}} \rightarrow uX_{p_1A_1p_2} \dots X_{p_{k-1}A_{k-1}p_k}$
$p\varepsilon\# \rightarrow q\varepsilon$ (1)	0	-	$X_{p\#q} \rightarrow \varepsilon$ (1')
$pa\# \rightarrow pA$ (2)	1	p	$X_{p\#p} \rightarrow aX_{pAp}$ (2')
		q	$X_{p\#q} \rightarrow aX_{pAq}$ (2'')
$paA \rightarrow pAA$ (3)	2	p, p	$X_{pAp} \rightarrow aX_{pAp}X_{pAp}$ (3')
		p, q	$X_{pAq} \rightarrow aX_{pAp}X_{pAq}$ (3'')
		q, p	$X_{pAp} \rightarrow aX_{pAq}X_{qAp}$ (3''')
		q, q	$X_{pAq} \rightarrow aX_{pAq}X_{qAq}$ (3''')
$pbA \rightarrow q\varepsilon$ (4)	0	-	$X_{pAq} \rightarrow b$ (4')
$qbA \rightarrow q\varepsilon$ (5)	0	-	$X_{qAq} \rightarrow b$ (5')

Der akzeptierenden Rechnung

$$(p, aabb, \#) \stackrel{(2)}{\vdash} (p, abb, A) \stackrel{(3)}{\vdash} (p, bb, AA) \stackrel{(4)}{\vdash} (q, b, A) \stackrel{(5)}{\vdash} (q, \varepsilon, \varepsilon)$$

von M entspricht dann die Ableitung

$$S \stackrel{(0')}{\Rightarrow} X_{p\#q} \stackrel{(2'')}{\Rightarrow} aX_{pAq} \stackrel{(3''')}{\Rightarrow} aaX_{pAq}X_{qAq} \stackrel{(4')}{\Rightarrow} aabX_{qAq} \stackrel{(5')}{\Rightarrow} aabb$$

in G und umgekehrt. ◁

3.5 Deterministisch kontextfreie Sprachen

Von besonderem Interesse sind kontextfreie Sprachen, die von einem deterministischen Kellerautomaten erkannt werden können.

Definition 104. Ein Kellerautomat heißt **deterministisch**, falls \vdash eine rechtseindeutige Relation ist:

$$K \vdash K_1 \wedge K \vdash K_2 \Rightarrow K_1 = K_2.$$

Äquivalent hierzu ist, dass die Überföhrungsfunktion δ für alle $(q, a, A) \in Z \times \Sigma \times \Gamma$ folgende Bedingung erfüllt:

$$\|\delta(q, a, A)\| + \|\delta(q, \varepsilon, A)\| \leq 1. \quad (*)$$

Hat nämlich eine Konfiguration $K = (q, x_1 \dots x_n, A_1 \dots A_l)$ zwei Folgekonfigurationen $K_1 \neq K_2$, so gibt es 3 Fälle:

- In K_1 und K_2 wurde x_i gelesen. Dann ist $\|\delta(q, x_i, A_1)\| \geq 2$.
- Weder in K_1 noch in K_2 wurde x_i gelesen. In diesem Fall ist $\|\delta(q, \varepsilon, A_1)\| \geq 2$.
- In K_1 wurde x_i gelesen, aber nicht in K_2 . In diesem Fall ist $\|\delta(q, x_i, A_1)\| \geq 1$ und $\|\delta(q, \varepsilon, A_1)\| \geq 1$.

D.h. in allen 3 Fällen ist $(*)$ verletzt. Gilt umgekehrt $(*)$ nicht, so lassen sich für $K = (q, a, A)$ leicht zwei Folgekonfigurationen angeben.

Beispiel 105. Der PDA $M = (\{q_0, q_1, q_2\}, \{a, b, c\}, \{A, B, \#\}, \delta, q_0, \#)$ mit der Überföhrungsfunktion

$$\begin{array}{llll} \delta: q_0a\# \rightarrow q_0A\# & q_0b\# \rightarrow q_0B\# & q_0aA \rightarrow q_0AA & q_0bA \rightarrow q_0BA \\ q_0aB \rightarrow q_0AB & q_0bB \rightarrow q_0BB & q_0cA \rightarrow q_1A & q_0cB \rightarrow q_1B \\ q_1aA \rightarrow q_1 & q_1bB \rightarrow q_1 & q_1\varepsilon\# \rightarrow q_2 & \end{array}$$

erkennt die Sprache $L(M) = \{x_cx^R \mid x \in \{a, b\}^+\}$. Um auf einen Blick erkennen zu können, ob M deterministisch ist, empfiehlt es sich, δ in Form einer Tabelle darzustellen:

δ	$q_0, \#$	q_0, A	q_0, B	$q_1, \#$	q_1, A	q_1, B	$q_2, \#$	q_2, A	q_2, B
ε	-	-	-	q_2	-	-	-	-	-
a	$q_0A\#$	q_0AA	q_0AB	-	q_1	-	-	-	-
b	$q_0B\#$	q_0BA	q_0BB	-	-	q_1	-	-	-
c	-	q_1A	q_1B	-	-	-	-	-	-

Man beachte, dass jedes Tabellenfeld höchstens eine Anweisung enthält und jede Spalte, die einen ε -Eintrag in der ersten Zeile hat, sonst keine weiteren Einträge enthält. Daher ist für alle $(q, a, A) \in Z \times \Sigma \times \Gamma$ die Bedingung

$$\|\delta(q, a, A)\| + \|\delta(q, \varepsilon, A)\| \leq 1$$

erfüllt. ◁

Verlangen wir von einem deterministischen Kellerautomaten, dass er seine Eingabe durch Leeren des Kellers akzeptiert, so können nicht alle regulären Sprachen von deterministischen Kellerautomaten erkannt werden. Um beispielsweise die Sprache $L = \{a, aa\}$ zu erkennen, muss der Keller von M nach Lesen von a geleert werden. Daher ist es M nicht mehr möglich, die Eingabe aa zu akzeptieren. Deterministische Kellerautomaten können also durch Leeren des Kellers nur

präfixfreie Sprachen L akzeptieren (d.h. kein Wort $x \in L$ ist Präfix eines anderen Wortes in L). Wir können das Problem aber lösen, indem wir deterministischen Kellerautomaten erlauben, ihre Eingabe durch Erreichen eines Endzustands zu akzeptieren.

Definition 106.

- Ein **Kellerautomat mit Endzuständen** wird durch ein 7-Tupel $M = (Z, \Sigma, \Gamma, \delta, q_0, \#, E)$ beschrieben. Dabei sind die Komponenten $Z, \Sigma, \Gamma, \delta, q_0, \#$ dieselben wie bei einem PDA und zusätzlich ist $E \subseteq Z$ eine Menge von **Endzuständen**.
- Die von M **akzeptierte** oder **erkannte Sprache** ist

$$L(M) = \{x \in \Sigma^* \mid \exists p \in E, \alpha \in \Gamma^* : (q_0, x, \#) \vdash^* (p, \varepsilon, \alpha)\}.$$

- M ist ein **deterministischer Kellerautomat mit Endzuständen** (kurz: **DPDA**), falls M zusätzlich für alle $(q, a, A) \in Z \times \Sigma \times \Gamma$ folgende Bedingung erfüllt:

$$\|\delta(q, a, A)\| + \|\delta(q, \varepsilon, A)\| \leq 1.$$

- Die Klasse der deterministisch kontextfreien Sprachen ist definiert durch

$$\text{DCFL} = \{L(M) \mid M \text{ ist ein DPDA}\}.$$

Als nächstes zeigen wir, dass DCFL unter Komplementbildung abgeschlossen ist. Versuchen wir, die End- und Nichtendzustände eines DPDA M einfach zu vertauschen, um einen DPDA \bar{M} für $\overline{L(M)}$ zu erhalten, so ergeben sich folgende Schwierigkeiten:

1. Falls M eine Eingabe x nicht zu Ende liest, wird x weder von M noch von \bar{M} akzeptiert.
2. Falls M nach dem Lesen von x noch ε -Übergänge ausführt und dabei End- und Nichtendzustände besucht, wird x von M und von \bar{M} akzeptiert.

Der nächste Satz zeigt, wie sich Problem 1 beheben lässt.

Satz 107. *Jede Sprache $L \in \text{DCFL}$ wird von einem DPDA M' erkannt, der alle Eingaben zu Ende liest.*

Beweis. Sei $M = (Z, \Sigma, \Gamma, \delta, q_0, \#, E)$ ein DPDA mit $L(M) = L$. Falls M eine Eingabe $x = x_1 \dots x_n$ nicht zu Ende liest, muss einer der folgenden drei Gründe vorliegen:

1. M gerät in eine Konfiguration $(q, x_i \dots x_n, \varepsilon)$, $i \leq n$, mit leerem Keller.
2. M gerät in eine Konfiguration $(q, x_i \dots x_n, A\gamma)$, $i \leq n$, in der wegen $\delta(q, x_i, A) = \delta(q, \varepsilon, A) = \emptyset$ keine Anweisung ausführbar ist.
3. M gerät in eine Konfiguration $(q, x_i \dots x_n, A\gamma)$, $i \leq n$, so dass M ausgehend von der Konfiguration (q, ε, A) eine unendliche Folge von ε -Anweisungen ausführt.

Die erste Ursache schließen wir aus, indem wir ein neues Zeichen \square auf dem Kellerboden platzieren:

$$(a) \quad s\varepsilon\# \rightarrow q_0\#\square \quad (\text{dabei sei } s \text{ der neue Startzustand}).$$

Die zweite Ursache schließen wir durch Hinzunahme eines Fehlerzustands r sowie folgender Anweisungen aus (hierbei ist $\Gamma' = \Gamma \cup \{\square\}$):

$$(b) \quad qaA \rightarrow rA, \quad \text{für alle } (q, a, A) \in Z \times \Sigma \times \Gamma' \text{ mit } A = \square \text{ oder } \delta(q, a, A) = \delta(q, \varepsilon, A) = \emptyset,$$

$$(c) \quad raA \rightarrow rA, \quad \text{für alle } a \in \Sigma \text{ und } A \in \Gamma'.$$

Als nächstes verhindern wir die Ausführung einer unendlichen Folge von ε -Übergängen. Dabei unterscheiden wir die beiden Fälle, ob M hierbei auch Endzustände besucht oder nicht. Falls ja, sehen wir einen

Umweg über den neuen Endzustand t vor.

(d) $q\epsilon A \rightarrow rA$, für alle $q \in Z$ und $A \in \Gamma$, so dass M ausgehend von der Konfiguration (q, ϵ, A) unendlich viele ϵ -Übergänge ausführt ohne dabei einen Endzustand zu besuchen.

(e) $q\epsilon A \rightarrow tA$
 $t\epsilon A \rightarrow rA$, für alle $q \in Z$ und $A \in \Gamma$, so dass M ausgehend von der Konfiguration (q, ϵ, A) unendlich viele ϵ -Übergänge ausführt und dabei auch Endzustände besucht.

Schließlich übernehmen wir von M die folgenden Anweisungen:

(f) alle Anweisungen aus δ , soweit sie nicht durch Anweisungen vom Typ (d) oder (e) überschrieben wurden.

Zusammenfassend transformieren wir M in den DPDA

$$M' = (Z \cup \{r, s, t\}, \Sigma, \Gamma', \delta', s, \#, E \cup \{t\})$$

mit $\Gamma' = \Gamma \cup \{\square\}$, wobei δ' die unter (a) bis (f) genannten Anweisungen enthält. ■

Beispiel 108. Wenden wir diese Konstruktion auf den DPDA

$$M = (\{q_0, q_1, q_2\}, \{a, b, c\}, \{A, B, \#\}, \delta, q_0, \#, \{q_2\})$$

mit der Überföhrungsfunktion

δ	$q_0, \#$	q_0, A	q_0, B	$q_1, \#$	q_1, A	q_1, B	$q_2, \#$	q_2, A	q_2, B
ϵ	-	-	-	q_2	-	-	$q_2\#$	-	-
a	$q_0A\#$	q_0AA	q_0AB	-	q_1	-	-	-	-
b	$q_0B\#$	q_0BA	q_0BB	-	-	q_1	-	-	-
c	-	q_1A	q_1B	-	-	-	-	-	-

an, so erhalten wir den DPDA

$$M' = (\{q_0, q_1, q_2, r, s, t\}, \{a, b, c\}, \{A, B, \#, \square\}, \delta', s, \#, \{q_2, t\})$$

mit folgender Überföhrungsfunktion δ' :

δ'	$q_0, \#$	q_0, A	q_0, B	q_0, \square	$q_1, \#$	q_1, A	q_1, B	q_1, \square	$q_2, \#$	q_2, A	q_2, B	q_2, \square
ϵ	-	-	-	-	q_2	-	-	-	$t\#$	-	-	-
a	$q_0A\#$	q_0AA	q_0AB	$r\square$	-	q_1	rB	$r\square$	-	rA	rB	$r\square$
b	$q_0B\#$	q_0BA	q_0BB	$r\square$	-	rA	q_1	$r\square$	-	rA	rB	$r\square$
c	$r\#$	q_1A	q_1B	$r\square$	-	rA	rB	$r\square$	-	rA	rB	$r\square$
Typ	(f, b)	(f)	(f)	(b)	(f)	(f, b)	(f, b)	(b)	(e)	(b)	(b)	(b)
	$s, \#$	s, A	s, B	s, \square	$r, \#$	r, A	r, B	r, \square	$t, \#$	t, A	t, B	t, \square
ϵ	$q_0\#\square$	-	-	-	-	-	-	-	$r\#$	-	-	-
a	-	-	-	-	$r\#$	rA	rB	$r\square$	-	-	-	-
b	-	-	-	-	$r\#$	rA	rB	$r\square$	-	-	-	-
c	-	-	-	-	$r\#$	rA	rB	$r\square$	-	-	-	-
Typ	(a)				(c)	(c)	(c)	(c)	(e)			

Satz 109. Die Klasse DCFL ist unter Komplement abgeschlossen, d.h. es gilt $DCFL = co\text{-}DCFL$.

Beweis. Sei $M = (Z, \Sigma, \Gamma, \delta, q_0, \#, E)$ ein DPDA, der alle Eingaben zu Ende liest, und sei $L(M) = L$. Wir konstruieren einen DPDA \bar{M} für \bar{L} .

Die Idee dabei ist, dass sich \bar{M} in seinem Zustand (q, i) neben dem aktuellen Zustand q von M in der Komponente i merkt, ob M nach Lesen des letzten Zeichens (bzw. seit Rechnungsbeginn) einen Endzustand besucht hat ($i = 1$) oder nicht ($i = 2$). Möchte M das nächste Zeichen lesen und befindet sich \bar{M} im Zustand $(q, 2)$, so macht \bar{M} noch einen Umweg über den Endzustand $(q, 3)$.

Konkret erhalten wir $\overline{M} = (Z \times \{1, 2, 3\}, \Sigma, \Gamma, \delta', s, \#, Z \times \{3\})$ mit

$$s = \begin{cases} (q_0, 1), & q_0 \notin E, \\ (q_0, 2), & \text{sonst,} \end{cases}$$

indem wir zu δ' für jede Anweisung $q \in A \rightarrow_M p \gamma$ die beiden Anweisungen

$$\begin{aligned} (q, 1) \varepsilon A &\rightarrow (p, i) \gamma \\ (q, 2) \varepsilon A &\rightarrow (p, 2) \gamma \end{aligned} \quad \text{mit} \quad i = \begin{cases} 1, & p \in E, \\ 2, & p \notin E, \end{cases}$$

sowie für jede Anweisung $qaA \rightarrow_M p \gamma$ die drei Anweisungen

$$\begin{aligned} (q, 1) aA &\rightarrow (p, i) \gamma \\ (q, 2) \varepsilon A &\rightarrow (q, 3) A \\ (q, 3) aA &\rightarrow (p, i) \gamma \end{aligned} \quad \text{mit} \quad i = \begin{cases} 1, & p \in E, \\ 2, & p \notin E \end{cases}$$

hinzufügen. ■

Eine nützliche Eigenschaft von \overline{M} ist, dass \overline{M} in einem Endzustand keine ε -Übergänge macht.

Beispiel 110. Angenommen, ein DPDA $M = (Z, \Sigma, \Gamma, \delta, q_0, \#, E)$ führt bei der Eingabe $x = a$ folgende Rechnung aus:

$$(q_0, a, \#) \vdash (q_1, \varepsilon, \gamma_1) \vdash (q_2, \varepsilon, \gamma_2).$$

Dann würde \overline{M} im Fall $E = \{q_0, q_2\}$ (d.h. $x \in L(M)$) die Rechnung

$$((q_0, 1), a, \#) \vdash ((q_1, 2), \varepsilon, \gamma_1) \vdash ((q_2, 1), \varepsilon, \gamma_2)$$

ausführen. Da $(q_1, 2), (q_2, 1) \notin Z \times \{3\}$ sind, verwirft also \overline{M} das Wort a . Dagegen würde \overline{M} im Fall $E = \{q_0\}$ (d.h. $x \notin L(M)$) die Rechnung

$$((q_0, 1), a, \#) \vdash ((q_1, 2), \varepsilon, \gamma_1) \vdash ((q_2, 2), \varepsilon, \gamma_2) \vdash ((q_2, 3), \varepsilon, \gamma_2)$$

ausführen. Da $(q_2, 3) \in Z \times \{3\}$ ein Endzustand von \overline{M} ist, würde \overline{M} nun also das Wort a akzeptieren. ◁

Satz 111. Die Klasse DCFL ist nicht abgeschlossen unter Schnitt, Vereinigung, Produkt und Sternhülle.

Beweis. Die beiden Sprachen

$$L_1 = \{a^n b^m c^m \mid n, m \geq 0\} \quad \text{und} \quad L_2 = \{a^n b^n c^m \mid n, m \geq 0\}$$

sind deterministisch kontextfrei (siehe Übungen). Da der Schnitt $L_1 \cap L_2 = \{a^n b^n c^n \mid n \geq 0\}$ nicht kontextfrei ist, liegt er auch nicht in DCFL, also ist DCFL nicht unter Schnitt abgeschlossen.

Da DCFL unter Komplementbildung abgeschlossen ist, kann DCFL wegen de Morgan dann auch nicht unter Vereinigung abgeschlossen sein. Beispielsweise sind folgende Sprachen deterministisch kontextfrei:

$$L_3 = \{a^i b^j c^k \mid i \neq j \wedge i, j, k \geq 1\} \quad \text{und} \quad L_4 = \{a^i b^j c^k \mid j \neq k \wedge i, j, k \geq 1\}.$$

Ihre Vereinigung $L_3 \cup L_4 = \{a^i b^j c^k \mid (i \neq j \vee j \neq k) \wedge i, j, k \geq 1\}$ gehört aber nicht zu DCFL, d.h. $L_3 \cup L_4 \in \text{CFL} \setminus \text{DCFL}$. DCFL ist nämlich unter Schnitt mit regulären Sprachen abgeschlossen (siehe Übungen). Daher wäre mit $L_3 \cup L_4$ auch die Sprache

$$(\overline{L_3 \cup L_4}) \cap L(a^+ b^+ c^+) = \{a^n b^n c^n \mid n \geq 1\}$$

(deterministisch) kontextfrei.

Als nächstes zeigen wir, dass DCFL nicht unter Produktbildung abgeschlossen ist. Wir wissen bereits, dass $L = L_3 \cup L_4 \notin \text{DCFL}$ ist. Dann ist auch die Sprache

$$0L = 0L_3 \cup 0L_4 \notin \text{DCFL},$$

da sich ein DPDA $M = (Z, \Sigma, \Gamma, \delta, q_0, \#, E)$ für $0L$ leicht zu einem DPDA für L umbauen ließe. Sei nämlich (p, ε, γ) die Konfiguration,

die M nach Lesen der Eingabe 0 erreicht. Dann erkennt der DP-DA $M' = (Z \cup \{s\}, \Sigma, \Gamma, \delta', s, \#, E)$ die Sprache L , wobei δ' wie folgt definiert ist:

$$\delta'(q, u, A) = \begin{cases} (p, \gamma), & (q, u, A) = (s, \varepsilon, \#), \\ \delta(q, u, A), & (q, u, A) \in Z \times (\Sigma \cup \{\varepsilon\}) \times \Gamma. \end{cases}$$

Es ist leicht zu sehen, dass die beiden Sprachen $\{\varepsilon, 0\}$ und $L_5 = L_3 \cup 0L_4$ in DCFL sind (siehe Übungen). Ihr Produkt $\{\varepsilon, 0\}L_5 = L_5 \cup 0L_5 = L_3 \cup 0L_4 \cup 0L_3 \cup 00L_4$ gehört aber nicht zu DCFL. Da DCFL unter Schnitt mit regulären Sprachen abgeschlossen ist (siehe Übungen), wäre andernfalls auch

$$\{\varepsilon, 0\}L_5 \cap L(0a^*b^*c^*) = 0L_3 \cup 0L_4$$

in DCFL, was wir bereits ausgeschlossen haben. ■

Dass DCFL auch nicht unter Sternhüllenbildung abgeschlossen ist, lässt sich ganz ähnlich zeigen (siehe Übungen). Wir fassen die bewiesenen Abschlusseigenschaften der Klassen REG, DCFL und CFL in folgender Tabelle zusammen:

	Vereinigung	Schnitt	Komplement	Produkt	Sternhülle
REG	ja	ja	ja	ja	ja
DCFL	nein	nein	ja	nein	nein
CFL	ja	nein	nein	ja	ja

Die Klasse der deterministisch kontextfreien Sprachen lässt sich auch mit Hilfe von speziellen kontextfreien Grammatiken charakterisieren, den so genannten $LR(k)$ -Grammatiken.

Der erste Buchstabe L steht für die Leserichtung bei der Syntaxanalyse, d.h. das Eingabewort x wird von links (nach rechts) gelesen. Der zweite Buchstabe R bedeutet, dass bei der Syntaxanalyse eine

Rechtsableitung entsteht. Schließlich gibt der Parameter k an, wieviele Zeichen man über das aktuelle Eingabezeichen hinauslesen muss, damit der nächste Schritt eindeutig feststeht (k wird auch als *Lookahead* bezeichnet).

Durch $LR(0)$ -Grammatiken lassen sich nur die präfixfreien Sprachen in DCFL erzeugen. Dagegen erzeugen die $LR(k)$ -Grammatiken für jedes $k \geq 1$ genau die Sprachen in DCFL.

Daneben gibt es noch $LL(k)$ -Grammatiken, die für wachsendes k immer mehr deterministisch kontextfreie Sprachen erzeugen.

4 Kontextsensitive Sprachen

In diesem Kapitel führen wir das Maschinenmodell des linear beschränkten Automaten (LBA) ein und zeigen, dass LBAs genau die kontextsensitiven Sprachen erkennen. Die Klasse CSL ist unter Komplementbildung abgeschlossen. Es ist jedoch offen, ob die Klasse DCSL der von einem deterministischen LBA erkannten Sprachen eine echte Teilklasse von CSL ist (diese Frage ist als *LBA-Problem* bekannt).

4.1 Kontextsensitive Grammatiken

Zur Erinnerung: Eine Grammatik $G = (V, \Sigma, P, S)$ heißt **kontextsensitiv**, falls für alle Regeln $\alpha \rightarrow \beta$ gilt: $|\beta| \geq |\alpha|$. Als einzige Ausnahme hiervon ist die Regel $S \rightarrow \varepsilon$ erlaubt. Allerdings nur dann, wenn das Startsymbol S nicht auf der rechten Seite einer Regel vorkommt.

Das nächste Beispiel zeigt, dass die Sprache $L = \{a^n b^n c^n \mid n \geq 0\}$ von einer kontextsensitiven Grammatik erzeugt wird. Da L nicht kontextfrei ist, ist also die Klasse CFL echt in der Klasse CSL enthalten.

Beispiel 112. Betrachte die kontextsensitive Grammatik $G = (V, \Sigma, P, S)$ mit $V = \{S, B\}$, $\Sigma = \{a, b, c\}$ und den Regeln

$$P: S \rightarrow aSBc, abc \quad (1, 2) \quad cB \rightarrow Bc \quad (3) \quad bB \rightarrow bb \quad (4)$$

In G läßt sich beispielsweise das Wort $w = aabbcc$ ableiten:

$$\underline{S} \xrightarrow{(1)} a\underline{S}Bc \xrightarrow{(2)} aabc\underline{B}c \xrightarrow{(3)} aab\underline{B}cc \xrightarrow{(4)} aabbcc$$

Allgemein gilt für alle $n \geq 1$:

$$S \xrightarrow{(1)} a^{n-1} S (Bc)^{n-1} \xrightarrow{(2)} a^n b c (Bc)^{n-1} \xrightarrow{(3)} a^n b B^{n-1} c^n \xrightarrow{(4)} a^n b^n c^n$$

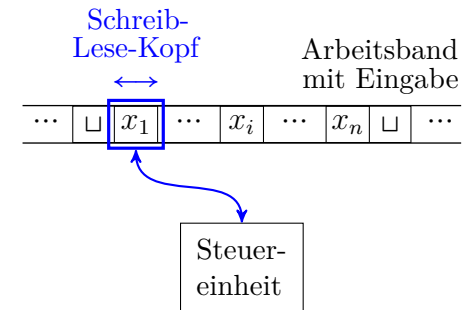
Also gilt $a^n b^n c^n \in L(G)$ für alle $n \geq 1$. Umgekehrt folgt durch Induktion über die Ableitungslänge m , dass jede Satzform u mit $S \Rightarrow^m \alpha$ die folgenden Bedingungen erfüllt:

- $\#_a(\alpha) = \#_b(\alpha) + \#_B(\alpha) = \#_c(\alpha)$,
- links von S und links von einem a kommen nur a 's vor,
- links von einem b kommen nur a 's oder b 's vor.

Daraus ergibt sich, dass in G nur Wörter der Form $w = a^n b^n c^n$ ableitbar sind. \triangleleft

4.2 Turingmaschinen

Um ein geeignetes Maschinenmodell für die kontextsensitiven Sprachen zu finden, führen wir zunächst das Rechenmodell der nichtdeterministischen Turingmaschine (NTM) ein. Eine NTM erhält ihre Eingabe auf einem nach links und rechts unbegrenzten Band. Während ihrer Rechnung kann sie den Schreib-Lese-Kopf auf dem Band in beide Richtungen bewegen und dabei die besuchten Bandfelder lesen sowie gelesenen Zeichen gegebenenfalls überschreiben.



Es gibt mehrere Arten von Turingmaschinen (u.a. mit einseitig unendlichem Band oder mit mehreren Schreib-Lese-Köpfen auf dem Band). Wir verwenden folgende Variante der Mehrband-Turingmaschine.

Definition 113. Sei $k \geq 1$.

- a) Eine **nichtdeterministische k -Band-Turingmaschine** (kurz **k -NTM** oder einfach **NTM**) wird durch ein 6-Tupel $M = (Z, \Sigma, \Gamma, \delta, q_0, E)$ beschrieben, wobei

- Z eine endliche Menge von Zuständen,
- Σ das Eingabealphabet (wobei $\sqcup \notin \Sigma$),
- Γ das Arbeitsalphabet (wobei $\Sigma \cup \{\sqcup\} \subseteq \Gamma$),
- $\delta: Z \times \Gamma^k \rightarrow \mathcal{P}(Z \times \Gamma^k \times \{L, R, N\}^k)$ die Überföhrungsfunktion,
- q_0 der Startzustand und
- $E \subseteq Z$ die Menge der Endzustände ist.

b) Eine k -NTM M heißt **deterministisch** (kurz: M ist eine **k -DTM** oder einfach **DTM**), falls für alle $(q, a_1, \dots, a_k) \in Z \times \Gamma^k$ die Ungleichung $\|\delta(q, a_1, \dots, a_k)\| \leq 1$ gilt.

Für $(q', a'_1, \dots, a'_k, D_1, \dots, D_k) \in \delta(q, a_1, \dots, a_k)$ schreiben wir auch

$$(q, a_1, \dots, a_k) \rightarrow (q', a'_1, \dots, a'_k, D_1, \dots, D_k).$$

Eine solche Anweisung ist ausführbar, falls

- q der aktuelle Zustand von M ist und
- sich für $i = 1, \dots, k$ der Lesekopf des i -ten Bandes auf einem mit a_i beschrifteten Feld befindet.

Ihre Ausführung bewirkt, dass M

- vom Zustand q in den Zustand q' übergeht,
- auf Band i das Symbol a_i durch a'_i ersetzt und
- den Kopf gemäß D_i bewegt (L: ein Feld nach links, R: ein Feld nach rechts, N: keine Bewegung).

Definition 114. Sei $M = (Z, \Sigma, \Gamma, \delta, q_0, E)$ eine k -NTM.

a) Eine **Konfiguration** von M ist ein $(3k + 1)$ -Tupel

$$K = (q, u_1, a_1, v_1, \dots, u_k, a_k, v_k) \in Z \times (\Gamma^* \times \Gamma \times \Gamma^*)^k$$

und besagt, dass

- q der momentane Zustand ist und

- das i -te Band mit $\dots \sqcup u_i a_i v_i \sqcup \dots$ beschriftet ist, wobei sich der Kopf auf dem Zeichen a_i befindet.

Im Fall $k = 1$ schreiben wir für eine Konfiguration (q, u, a, v) auch kurz $uqav$.

b) Die **Startkonfiguration** von M bei Eingabe $x = x_1 \dots x_n \in \Sigma^*$ ist

$$K_x = \begin{cases} (q_0, \varepsilon, x_1, x_2 \dots x_n, \varepsilon, \sqcup, \varepsilon, \dots, \varepsilon, \sqcup, \varepsilon), & x \neq \varepsilon, \\ (q_0, \varepsilon, \sqcup, \varepsilon, \dots, \varepsilon, \sqcup, \varepsilon), & x = \varepsilon. \end{cases}$$

c) Eine Konfiguration $K' = (q, u'_1, a'_1, v'_1, \dots, u'_k, a'_k, v'_k)$ heißt **Folgekonfiguration** von $K = (p, u_1, a_1, v_1, \dots, u_k, a_k, v_k)$ (kurz $K \vdash K'$), falls eine Anweisung

$$(q, a_1, \dots, a_k) \rightarrow (q', b_1, \dots, b_k, D_1, \dots, D_k)$$

existiert, so dass für $i = 1, \dots, k$ gilt:

im Fall $D_i = N$:	$D_i = R$:	$D_i = L$:
$K: \overline{u_i \boxed{a_i} v_i}$	$K: \overline{u_i \boxed{a_i} v_i}$	$K: \overline{u_i \boxed{a_i} v_i}$
$K': \overline{u_i \boxed{b_i} v_i}$	$K': \overline{u_i b_i \boxed{a'_i} v'_i}$	$K': \overline{u'_i \boxed{a'_i} b_i v_i}$
$u'_i = u_i,$ $a'_i = b_i$ und $v'_i = v_i.$	$u'_i = u_i b_i$ und $a'_i v'_i = \begin{cases} v_i, & v_i \neq \varepsilon, \\ \sqcup, & \text{sonst.} \end{cases}$	$u'_i a'_i = \begin{cases} u_i, & u_i \neq \varepsilon, \\ \sqcup, & \text{sonst} \end{cases}$ und $v'_i = b_i v_i.$

Man beachte, dass sich die Länge der Bandinschrift $u_i a_i v_i$ beim Übergang von K zu K' genau dann um 1 erhöht, wenn in K' zum ersten Mal ein neues Feld auf dem i -ten Band besucht wird. Andernfalls bleibt die Länge von $u_i a_i v_i$ unverändert. Die Länge von $u_i a_i v_i$ entspricht also genau der Anzahl der auf dem i -ten Band besuchten Felder (inkl. Eingabezeichen im Fall $i = 1$).

d) Eine **Rechnung** von M bei Eingabe x ist eine Folge von Konfigurationen $K_0, K_1, K_2 \dots$ mit $K_0 = K_x$ und $K_0 \vdash K_1 \vdash K_2 \dots$.

e) Die von M **akzeptierte** oder **erkannte Sprache** ist

$$L(M) = \{x \in \Sigma^* \mid \exists K \in E \times (\Gamma^* \times \Gamma \times \Gamma^*)^k : K_x \vdash^* K\}.$$

M akzeptiert also eine Eingabe x (hierfür sagen wir kurz $M(x)$ akzeptiert), falls es eine Rechnung $K_x = K_0 \vdash K_1 \vdash K_2 \dots \vdash K_l$ von $M(x)$ gibt, bei der ein Endzustand erreicht wird.

Beispiel 115. Betrachte die 1-DTM $M = (Z, \Sigma, \Gamma, \delta, q_0, E)$ mit $Z = \{q_0, \dots, q_4\}$, $\Sigma = \{a, b\}$, $\Gamma = \Sigma \cup \{A, B, \sqcup\}$, $E = \{q_4\}$, wobei δ folgende Anweisungen enthält:

$q_0a \rightarrow q_1AR$ (1) Anfang der Schleife: Ersetze das erste a durch A .

$q_1a \rightarrow q_1aR$ (2) Bewege den Kopf nach rechts bis zum ersten b
 $q_1B \rightarrow q_1BR$ (3) und ersetze dies durch ein B (falls kein b mehr
 $q_1b \rightarrow q_2BL$ (4) vorhanden ist, dann halte ohne zu akzeptieren).

$q_2a \rightarrow q_2aL$ (5) Bewege den Kopf zurück nach links bis ein A
 $q_2B \rightarrow q_2BL$ (6) kommt, gehe wieder ein Feld nach rechts und wie-
 $q_2A \rightarrow q_0AR$ (7) derhole die Schleife.

$q_0B \rightarrow q_3BR$ (8) Falls kein a am Anfang der Schleife, dann teste,
 $q_3B \rightarrow q_3BR$ (9) ob noch ein b vorhanden ist. Wenn ja, dann halte
 $q_3\sqcup \rightarrow q_4\sqcup N$ (10) ohne zu akzeptieren. Andernfalls akzeptiere.

Dann führt M bei Eingabe $aabb$ folgende Rechnung aus:

$$\begin{array}{llll} q_0aabb \vdash Aq_1abb & \vdash Aaq_1bb & \vdash Aq_2aBb & \\ (1) & (2) & (4) & \\ \vdash q_2AaBb & \vdash Aq_0aBb & \vdash AAq_1Bb & \\ (5) & (7) & (1) & \\ \vdash AABq_1b & \vdash AAq_2BB & \vdash Aq_2ABB & \\ (3) & (4) & (6) & \\ \vdash AAq_0BB & \vdash AABq_3B & \vdash AABBq_3\sqcup & \vdash AABBq_4\sqcup \\ (7) & (8) & (9) & (10) \end{array}$$

Ähnlich läßt sich $a^n b^n \in L(M)$ für ein beliebiges $n \geq 1$ zeigen. Andererseits führt die Eingabe abb auf die Rechnung

$$q_0abb \vdash Aq_1bb \vdash q_2ABb \vdash Aq_0Bb \vdash ABq_3b,$$

(1) (4) (7) (8)

die nicht weiter fortsetzbar ist. Da M deterministisch ist, kann $M(abb)$ auch nicht durch eine andere Rechnung den Endzustand q_4 erreichen. D.h. abb gehört nicht zu $L(M)$. Tatsächlich läßt sich durch Betrachtung der übrigen Fälle ($x = a^n b^m$, $n > m$, $x = a^n b^m a^k$, $m, k \geq 1$, etc.) zeigen, dass M nur Eingaben der Form $a^n b^n$ akzeptiert, und somit $L(M) = \{a^n b^n \mid n \geq 1\}$ ist. \triangleleft