Vorlesungsskript

Einführung in die Theoretische Informatik

Wintersemester 2013/14

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

6. November 2013

Inhaltsverzeichnis

Inhaltsverzeichnis

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.

a) Ein ***Alphabet*** *$\Sigma = \{a_1, \dots, a_m\}$ ist eine geordnete Menge von endlich vielen* ***Zeichen****.*

b) Eine Folge $x = x_1 \dots x_n$ von n Zeichen heißt ***Wort*** *(der* ***Länge*** *n).*

c) 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.

d) Das (einzige) Wort der Länge $n = 0$ ist das ***leere Wort****, welches wir mit ε bezeichnen.*

e) Jede Teilmenge $L \subseteq \Sigma^*$ heißt ***Sprache*** *über dem Alphabet Σ.*

Beispiel 2. *Sei Σ ein Alphabet. Dann sind $\varnothing, \Sigma^*, \Sigma$ und $\{\varepsilon\}$ Sprachen über Σ. Die Sprache $\varnothing$ 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 **Singleton-Sprachen** bezeichnet.*

Da Sprachen Mengen sind, können wir sie bzgl. Inklusion vergleichen. Zum Beispiel gilt
$$\varnothing \subseteq \{\varepsilon\} \subseteq \Sigma^*.$$
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
- $\overline{A} = \{x \in \Sigma^* \mid x \notin A\}$ das ***Komplement*** von A.

Neben den Mengenoperationen gibt es auch spezielle Sprachoperationen.

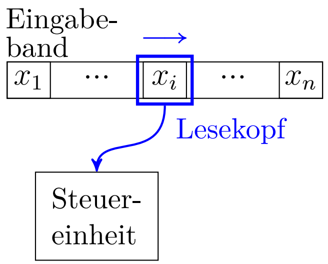
Definition 3.

- *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*

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

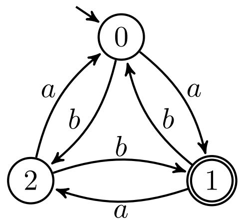
*Die von M **akzeptierte** oder **erkannte Sprache** ist*
$$L(M) = \left\{ x_1 \dots x_n \in \Sigma^* \;\middle|\; \begin{array}{l} \text{es gibt } q_1, \dots, q_{n-1} \in Z, q_n \in E \text{ mit} \\ \delta(q_i, x_{i+1}) = q_{i+1} \text{ für } i = 0, \dots, n-1 \end{array} \right\}.$$

$q_0, q_1, \ldots, q_n$ *heißt* ***Rechnung*** *von* $M(x_1 \ldots x_n)$, *falls* $\delta(q_i, x_{i+1}) = q_{i+1}$ *für* $i = 0, \ldots, n-1$ *gilt. Sie heißt* ***akzeptierend***, *falls* $q_n \in E$ *ist.*

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

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. $\triangleleft$

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

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

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* $j \equiv_m k$ *bedeutet, dass* $j - k$ *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 $x = x_1 \ldots x_{n+1}$ gegeben und sei $i = \hat{\delta}(0, x_1 \ldots x_n)$. Nach IV gilt dann

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

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

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

und somit

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

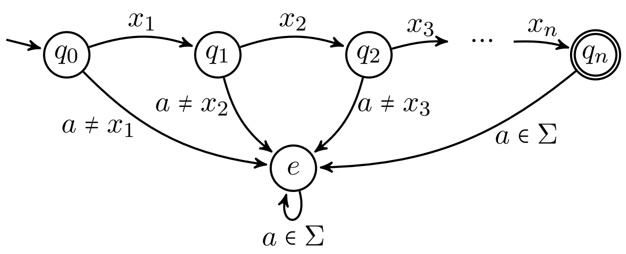
$\blacksquare$

Eine von einem DFA akzeptierte Sprache wird als **regulär** bezeichnet. Die zugehörige Sprachklasse ist

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

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 Überführungsfunktion

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

■

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.* ◁

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 nichtleeren endlichen Sprachen. ◁

Definition 12. *Für eine Sprachklasse $\mathcal{C}$ bezeichne co-$\mathcal{C}$ 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 $L_1, L_2 \in \mathsf{REG}$ sind auch die Sprachen $\overline{L_1} = \Sigma^* \setminus L_1$, $L_1 \cap L_2$ 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

$$\overline{M_1} = (Z_1, \Sigma, \delta_1, q_0, Z_1 \setminus E_1)$$

das Komplement $\overline{L_1}$ 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 $L_1 \cup L_2 = \overline{(\overline{L_1} \cap \overline{L_2})}$ ist dann aber auch die Vereinigung von L_1 und L_2 regulär. (Wie sieht der zugehörige DFA aus?) ■

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 Überführungsfunktion die Form*

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

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*

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

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

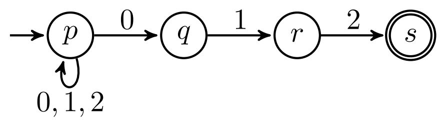
Ein NFA N kann bei einer Eingabe x also nicht nur eine, sondern mehrere verschiedene Rechnungen parallel ausführen. Ein Wort x gehört genau dann zu $L(N)$, wenn $N(x)$ mindestens eine akzeptierende Rechnung hat.

Im Gegensatz zu einem DFA, dessen Überführungsfunktion 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 nächste Eingabezeichen x_i wegen $\Delta(q, x_i) = \varnothing$ 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 Überführungsfunktion*

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

Graphische Darstellung:



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

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 $Z_1 \cap Z_2 = \varnothing$ 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

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

die Sprache $L_1 L_2$.

Beweis von $L_1L_2 \subseteq L(N)$: Seien $x = x_1 \cdots x_k \in L_1, y = y_1 \cdots y_l \in L_2$ und seien $q_0, \ldots, q_k$ und $p_0, \ldots, p_l$ akzeptierende Rechnungen von $N_1(x)$ und $N_2(y)$. Dann gilt $q_0 \in Q_1$, $q_k \in E_1$ und $p_0 \in Q_2$, $p_l \in E_2$.

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

Also ist $q_0, \ldots, q_k, p_1, \ldots, p_l$ eine akzeptierende Rechnung von $N(xy)$.

Beweis von $L(N) \subseteq L_1L_2$: Sei $x = x_1 \cdots x_n \in L(N)$ und sei $q_0, \ldots, q_n$ eine akz. Rechnung von $N(x)$. Dann gilt $q_0 \in Q_1$, $q_n \in E$, $q_0, \ldots, q_i \in Z_1$ und $q_{i+1}, \ldots, q_n \in Z_2$ für ein $i \in \{0, \ldots, n\}$.

- Im Fall $i = n$ ist $q_n \in E_1$ (d.h. $x \in L_1$) und $Q_2 \cap E_2 \neq \varnothing$ (d.h. $\varepsilon \in L_2$).
- 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.

Also ist $q_0, \ldots, q_i$ eine akz. Rechnung von $N_1(x_1 \cdots x_i)$ und $q, q_{i+1}, \ldots, q_n$ eine akz. Rechnung von $N_2(x_{i+1} \cdots x_n)$, d.h. $x \in L_1L_2$.

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

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

die Sprache L_1^* akzeptiert. ■

Satz 17 (Rabin und Scott)**.**
$\mathsf{REG} = \{L(N) \mid N \textit{ ist ein NFA}\}$.

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 Überführungsfunktion $\delta : \mathcal{P}(Z) \times \Sigma \to \mathcal{P}(Z)$ 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 $\hat{\delta} : \mathcal{P}(Z) \times \Sigma^* \to \mathcal{P}(Z)$ 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 $x = x_1 \ldots x_n$ gegeben. Nach Induktionsvoraussetzung enthält

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

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.

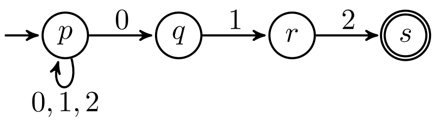
$$E' = \{Q \subseteq Z \mid Q \cap E \neq \varnothing\},$$

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

$$\begin{aligned} x \in L(N) &\Leftrightarrow N(x) \text{ kann in genau } |x| \text{ Schritten einen Endzustand erreichen} \\ &\Leftrightarrow \hat{\delta}(Q_0, x) \cap E \neq \varnothing \\ &\Leftrightarrow \hat{\delta}(Q_0, x) \in E' \\ &\Leftrightarrow x \in L(M). \end{aligned}$$

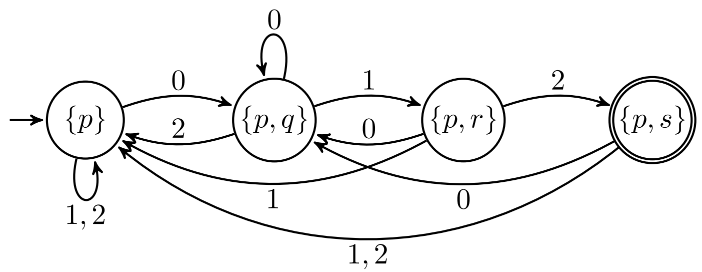
■

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,*
- *Schnitt,*
- *Vereinigung,*
- *Produkt,*
- *Sternhülle.*

2.3 Reguläre Ausdrücke

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

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

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

> 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* $\varnothing$, ϵ *und* a $(a \in \Sigma)$ *sind reguläre Ausdrücke, die*

- *die leere Sprache* $L(\varnothing) = \varnothing$,
- *die Sprache* $L(\epsilon) = \{\varepsilon\}$ *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.

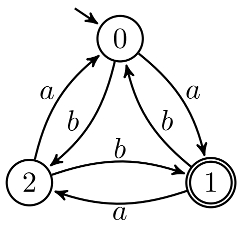
- *Um Klammern zu sparen, definieren wir folgende* **Präzedenzordnung***: Der Sternoperator* * *bindet stärker als der Produktoperator und dieser wiederum stärker als der Vereinigungsoperator. Für* $((a|b(c)^*)|d)$ *können wir also kurz* $a|bc^*|d$ *schreiben.*
- *Da der reguläre Ausdruck* $\gamma\gamma^*$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

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

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

◁

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* $w \neq \varepsilon$, *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*

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

beschreiben. Eine ähnliche Überlegung zeigt, dass sich die 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

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

◁

Satz 24. $\{L(\gamma) \mid \gamma$ *ist ein regulärer Ausdruck*$\} \subseteq \mathsf{REG}$.

Beweis. Die Inklusion von rechts nach links ist klar, da die Basisausdrücke $\varnothing$, 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 $Z = \{1, \ldots, m\}$ 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

$$L^r_{p,q} = \left\{ x_1 \ldots x_n \in \Sigma^* \;\middle|\; \begin{array}{c} \hat{\delta}(p, x_1 \ldots x_n) = q \text{ und für} \\ i = 1, \ldots, n-1 \text{ gilt } \hat{\delta}(p, x_1 \ldots x_i) \le r \end{array} \right\}.$$

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

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

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

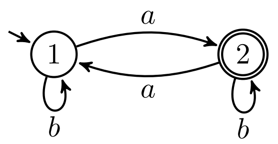
$$L^{r+1}_{p,q} = L^r_{p,q} \cup L^r_{p,r+1}(L^r_{r+1,r+1})^* L^r_{r+1,q}$$

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

$$\gamma^{r+1}_{p,q} = \gamma^r_{p,q} | \gamma^r_{p,r+1}(\gamma^r_{r+1,r+1})^* \gamma^r_{r+1,q}.$$

■

Beispiel 25. *Betrachte den DFA*



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

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

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

$$\gamma^{r+1}_{p,q} = \gamma^r_{p,q} | \gamma^r_{p,r+1}(\gamma^r_{r+1,r+1})^* \gamma^r_{r+1,q}$$

und erhalten

$$\begin{aligned} \gamma^2_{1,2} &= \gamma^1_{1,2} | \gamma^1_{1,2}(\gamma^1_{2,2})^* \gamma^1_{2,2}, \\ \gamma^1_{1,2} &= \gamma^0_{1,2} | \gamma^0_{1,1}(\gamma^0_{1,1})^* \gamma^0_{1,2}, \\ \gamma^1_{2,2} &= \gamma^0_{2,2} | \gamma^0_{2,1}(\gamma^0_{1,1})^* \gamma^0_{1,2}. \end{aligned}$$

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

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

◁

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

- *L ist regulär,*
- *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 $i = 1, \ldots, n$. Dann heißt $(A; R_1, \ldots, R_n)$ ***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 $F = \{f \mid f$ ist Fluss in Europa$\}$ und*
$$M = \{(f, g) \in F \times F \mid f \text{ mündet in } g\}.$$
- *(U, B) mit $U = \{x \mid x$ ist Berliner$\}$ und*
$$B = \{(x, y) \in U \times U \mid x \text{ ist Bruder von } y\}.$$
- *$(\mathsf{P}(M), \subseteq)$, wobei $\mathsf{P}(M)$ 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:

$$\begin{aligned} \bigcap \mathcal{M} &= \bigcap_{R \in \mathcal{M}} R = \{(x, y) \mid \forall R \in \mathcal{M} : xRy\}, \\ \bigcup \mathcal{M} &= \bigcup_{R \in \mathcal{M}} R = \{(x, y) \mid \exists R \in \mathcal{M} : xRy\}. \end{aligned}$$

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 $R \circ \cdots \circ R$ 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 $R \times \cdots \times R$ 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.

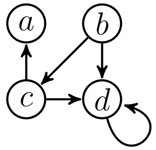
- *Die Relation "ist Schwester von" ist zwar in einer reinen Damengesellschaft symmetrisch, i.a. jedoch weder symmetrisch noch asymmetrisch noch antisymmetrisch.*
- *Die Relation "ist Geschwister von" ist zwar symmetrisch, aber weder reflexiv noch transitiv und somit keine Äquivalenzrelation.*
- $(\mathbb{R},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

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.*



◁

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***.

Darstellung durch eine Adjazenzmatrix

Eine Relation R auf einer endlichen (geordneten) Menge $A = \{a_1, \ldots, a_n\}$ 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 eine Adjazenzliste

Eine weitere Möglichkeit besteht darin, eine endliche Relation R in Form einer Tabelle darzustellen, die jedem Element $x \in A$ seine Nachfolgermenge $R[x]$ in Form einer Liste zuordnet:

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

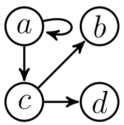
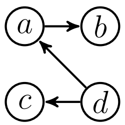
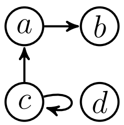
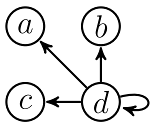
Sind $M_R = (r_{ij})$ und $M_S = (s_{ij})$ boolesche $n \times n$-Matrizen für R und S, so erhalten wir für $T = R \circ S$ die Matrix $M_T = (t_{ij})$ mit

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

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				
Adjazenz-matrix	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
Adjazenz-liste	$a : a, c$ $b :$ - $c : b, d$ $d :$ -	$a : b$ $b :$ - $c :$ - $d : a, c$	$a : b$ $b :$ - $c : a, c$ $d :$ -	$a :$ - $b :$ - $c :$ - $d : a, b, c, d$

◁

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

Als nächstes zeigen wir, dass die Menge $\mathcal{R} = \mathsf{P}(A \times A)$ aller binären Relationen auf A mit dem Relationenprodukt $\circ$ als binärer Operation ein ***Monoid***) (also eine Halbgruppe mit neutralem Element) bildet.

Satz 32. *Seien Q, R, S Relationen auf A. Dann gilt*

(i) $(Q \circ R) \circ S = Q \circ (R \circ S)$, *d.h. $\circ$ ist assoziativ,*

(ii) $Id \circ R = R \circ Id = R$, *d.h. Id ist neutrales Element.*

Beweis.

(i) Es gilt:

$$\begin{aligned}
x\,(Q \circ R) \circ S\,y &\Leftrightarrow \exists u \in A : x\,(Q \circ R)\,u \wedge u\,S\,y \\
&\Leftrightarrow \exists u \in A : (\exists v \in A : x\,Q\,v\,R\,u) \wedge u\,S\,y \\
&\Leftrightarrow \exists u, v \in A : x\,Q\,v\,R\,u\,S\,y \\
&\Leftrightarrow \exists v \in A : x\,Q\,v \wedge (\exists u \in A : v\,R\,u \wedge u\,S\,y) \\
&\Leftrightarrow \exists v \in A : x\,Q\,v\,(R \circ S)\,y \\
&\Leftrightarrow x\,Q \circ (R \circ S)\,y
\end{aligned}$$

(ii) Wegen $x\,Id \circ R\,y \Leftrightarrow \exists z : x = z \wedge z\,R\,y \Leftrightarrow x\,R\,y$ folgt $Id \circ R = R$. Die Gleichheit $R \circ Id = R$ folgt analog.

■

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 33. *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*

$$h_{äq}(R) = \bigcap \{S \mid S \text{ ist eine Äquivalenzrelation auf } A \text{ und } R \subseteq S\}.$$

Satz 34. *Sei R eine Relation auf A.*

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

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^nb, d.h. es gibt Elemente $x_0, \ldots, x_n \in A$ mit $x_0 = a$, $x_n = b$ und

$$x_0Rx_1Rx_2 \ldots x_{n-1}Rx_n.$$

In der Graphentheorie nennt man $x_0, \ldots, x_n$ einen ***Weg*** der Länge n von a nach b.

2.4.1 Ordnungs- und Äquivalenzrelationen

Wir betrachten zunächst ***Ordnungsrelationen***, die durch die drei Eigenschaften reflexiv, antisymmetrisch und transitiv definiert sind.

Beispiel 35.

- $(\mathsf{P}(M), \subseteq)$, $(\mathbb{Z}, \leq)$, $(\mathbb{R}, \leq)$ *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 $\mathbb{N}$.* ◁

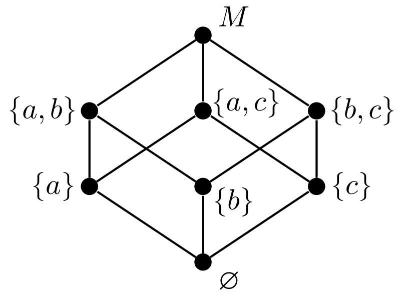
Ordnungen lassen sich sehr anschaulich durch Hasse-Diagramme darstellen. Sei $\leq$ eine Ordnung auf A und sei $<$ die Relation $\leq \cap \overline{Id_A}$. Um die Ordnung $\leq$ in einem ***Hasse-Diagramm*** darzustellen, wird nur der Graph der Relation

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

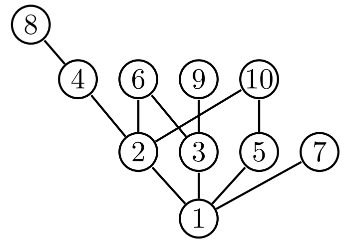
gezeichnet. Für $x \lessdot y$ sagt man auch, y ist ***oberer Nachbar*** von x. Weiterhin wird im Fall $x \lessdot y$ der Knoten y oberhalb vom Knoten x gezeichnet, so dass auf Pfeilspitzen verzichtet werden kann.

Beispiel 36.

Die Inklusionsrelation auf der Potenzmenge $\mathcal{P}(M)$ von $M = \{a,b,c\}$ lässt sich durch nebenstehendes Hasse-Diagramm darstellen.



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



◁

Definition 37. *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 38. *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 39. *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 $i = \inf B$):*

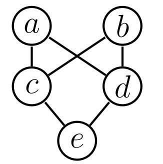
$$(\forall b \in B : b \geq i) \wedge [\forall u \in A : (\forall b \in B : b \geq u) \Rightarrow u \leq i].$$

- *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 ($s = \sup B$):*

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

Bemerkung 40. *B kann nicht mehr als ein Supremum und ein Infimum haben.*

Beispiel 41. *Betrachte nebenstehende Ordnung auf der Menge $A = \{a, b, c, d, e\}$. Die folgende Tabelle zeigt für verschiedene Teilmengen $B \subseteq A$ alle minimalen und maximalen Elemente in B Minimum und Maximum, alle unteren und oberen Schranken, sowie 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

$\triangleleft$

Bemerkung 42.

- *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*

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

weder ein Supremum noch ein Infimum.

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

Als nächstes betrachten wir 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 43.

- *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*

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

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

Definition 44. *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 $B_i \cap B_j = \varnothing$.*

Die Äquivalenzklassen einer Äquivalenzrelation E bilden eine Partition $\{[x] \mid x \in A\}$ von A (siehe Satz 45). 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. Wie der nächste Satz zeigt, beschreiben Äquivalenzrelationen auf A und Partitionen von A denselben Sachverhalt.

Satz 45. *Sei E eine Relation auf A. Dann sind folgende Aussagen äquivalent.*

(i) E ist eine Äquivalenzrelation auf A.

(ii) Für alle $x, y \in A$ gilt

$$xEy \Leftrightarrow E[x] = E[y] \quad (*)$$

(iii) 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. Da E transitiv ist, impliziert xEy die Inklusion $E[y] \subseteq E[x]$:

$$z \in E[y] \Rightarrow yEz \Rightarrow xEz \Rightarrow z \in E[x].$$

Da E symmetrisch ist, folgt aus xEy aber auch $E[x] \subseteq E[y]$.

Umgekehrt folgt aus $E[x] = E[y]$ wegen der Reflexivität von E, dass $y \in E[y] = E[x]$ enthalten ist, und somit xEy. Dies zeigt, dass E die Äquivalenz $(*)$ erfüllt.

$(ii) \Rightarrow (iii)$ Wir zeigen, dass die Äquivalenzklassen $E[x]$, $x \in A$, die Menge A partitionieren, falls E die Bedingung $(*)$ erfüllt.

Wegen $E[x] = E[x]$ folgt xEx und somit $x \in E[x]$. Folglich überdecken die Mengen $E[x]$ die Menge A.

Ist $E[x] \cap E[y] \neq \varnothing$ und z ein Element in $E[x] \cap E[y]$, so gilt xEz und yEz und daher folgt $E[x] = E[z] = E[y]$.

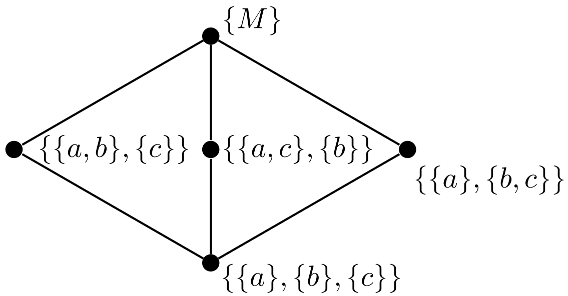
$(iii) \Rightarrow (i)$ Existiert schließlich eine Partition $\{B_i \mid i \in I\}$ von A mit $xEy \Leftrightarrow \exists i \in I : x, y \in B_i$, so ist E reflexiv, da zu jedem $x \in A$ eine Menge B_i mit $x \in B_i$ existiert. Zudem ist E symmetrisch, da aus $x, y \in B_i$ auch $y, x \in B_i$ folgt. Und E ist 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.

■

Die 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. $A/Id_A = \{\{x\} \mid x \in A\}$, und die Allrelation erzeugt nur eine Äquivalenzklasse, nämlich $A/(A \times A) = \{A\}$.

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 ***gröbste*** Äquivalenzrelation.

Die feiner-Relation auf der Menge aller Partitionen von $M = \{a, b, c\}$ hat das folgende Hasse-Diagramm:



2.4.2 Abbildungen

Definition 46. *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*

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

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

Bemerkung 47.

- *Wie üblich werden wir Abbildungen meist mit kleinen Buchstaben $f, g, h, \ldots$ bezeichnen und für $(x, y) \in f$ nicht xfy sondern $f(x) = y$ oder $f : x \mapsto y$ schreiben.*
- *Ist $f : A \to B$ 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 48.

- *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 $f^{-1} : N(f) \to A$ 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

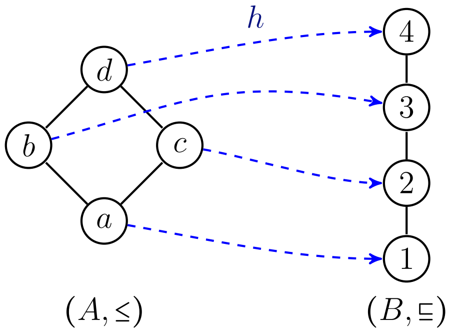
Definition 49. *Seien (A_1, R_1) und (A_2, R_2) Relationalstrukturen.*

- *Eine Abbildung $h : A_1 \to A_2$ 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 50. *Folgende Abbildung $h : A_1 \to A_2$ 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

$$2 \sqsubseteq 3, \;\; aber\; h^{-1}(2) = b \nleq c = h^{-1}(3).$$

$\triangleleft$

Definition 51. *Ein bijektiver Homomorphismus $h : A_1 \to A_2$, bei dem auch h^{-1} ein Homomorphismus ist, d.h. es gilt*

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

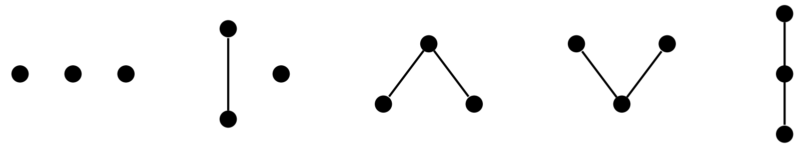
heißt **Isomorphismus***. In diesem Fall heißen die Strukturen (A_1, R_1) und (A_2, R_2)* **isomorph** *(kurz: $(A_1, R_1) \cong (A_2, R_2)$).*

Beispiel 52.

- *Die Abbildung $h : \mathbb{R} \to \mathbb{R}^+$ mit*

  $$h : x \mapsto e^x$$

  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*

  $$T_n = \{k \in \mathbb{N} \mid k \text{ teilt } n\}$$

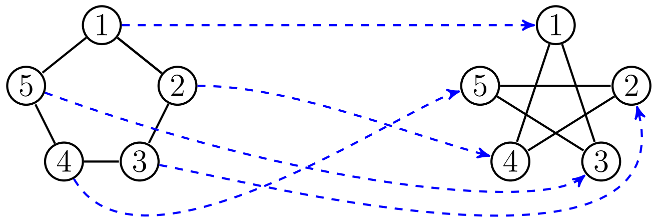
  die Menge aller Teiler von n und

  $$P_n = \{p \in T_n \mid p \text{ ist prim}\}$$

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



$G = (V, E)$ $\qquad$ $G' = (V, E')$

v	1 2 3 4 5
$h_1(v)$	1 3 5 2 4
$h_2(v)$	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:*

◁

Bemerkung 53. *Auf der Knotenmenge $V = \{1, \ldots, n\}$ existieren genau $2^{\binom{n}{2}}$ verschiedene Graphen. Sei $a(n)$ die Anzahl aller nichtisomorphen Graphen auf V. Da jede Isomorphieklasse mindestens einen und höchstens $n!$ verschiedene Graphen enthält, ist $2^{\binom{n}{2}}/n! \le a(n) \le 2^{\binom{n}{2}}$.*

Tatsächlich ist $a(n)$ **asymptotisch gleich** *$u(n) = 2^{\binom{n}{2}}/n!$ (in Zeichen: $a(n) \sim u(n)$), d.h.*

$$\lim_{n \to \infty} a(n)/u(n) = 1.$$

Also gibt es auf $V = \{1, \ldots, n\}$ nicht wesentlich mehr als $u(n)$ nichtisomorphe Graphen.

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: $q \sim p$), 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.

Fassen wir alle zu einem Zustand z äquivalenten Zustände in dem neuen Zustand

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

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 54. *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 $p \sim q$ 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

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

die von M beim Abarbeiten von x durchlaufenen Zustände. Wegen

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

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 55).

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 $L_x = L(M_{\hat{\delta}(q_0,x)})$ 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

$$x\, R_L\, y \Leftrightarrow L_x = L_y$$

auf Σ^* definierte Äquivalenzrelation R_L endlichen Index hat.

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 $A \Delta B = (A \setminus B) \cup (B \setminus A)$ die *symmetrische Differenz* von zwei Mengen A und B. Dann ist die Inäquivalenz $p \nsim q$ zweier Zustände p und q gleichbedeutend mit $L(M_p) \Delta L(M_q) \neq \emptyset$. Wir nennen ein Wort $x \in L(M_p) \Delta L(M_q)$ einen *Unterscheider* zwischen p und q.

Beobachtung 55.

- *Endzustände $p \in E$ sind nicht mit Zuständen $q \in Z \setminus E$ äquivalent (da sie durch ε unterschieden werden).*
- *Wenn $\delta(p,a)$ und $\delta(q,a)$ inäquivalent sind, dann auch p und q (da jeder Unterscheider x von $\delta(p,a)$ und $\delta(q,a)$ einen Unterscheider ax von p und q liefert).*

Wenn also D nur Paare von inäquivalenten Zuständen enthält, dann trifft dies auch auf die Menge

$$D' = \{\{p,q\} \mid \exists a \in \Sigma : \{\delta(p,a), \delta(q,a)\} \in D\}$$

zu. Wir können somit ausgehend von der Menge

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

eine Folge von Mengen

$$D_0 \subseteq D_1 \subseteq \dots \subseteq \{\{z,z'\} \subseteq Z \mid z \neq z'\}$$

mittels der Vorschrift

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

berechnen, indem wir zu D_i alle Paare $\{p,q\}$ hinzufügen, für die eines der Paare $\{\delta(p,a),\delta(q,a)\}$, $a \in \Sigma$, bereits zu D_i gehört. Da Z endlich ist, muss es ein j mit $D_{j+1} = D_j$ geben. In diesem Fall gilt (siehe Übungen):

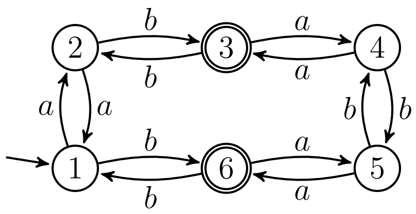
$$p \not\sim q \Leftrightarrow \{p,q\} \in D_j.$$

Folglich kann M' durch Verschmelzen aller Zustände p,q mit $\{p,q\} \notin D_j$ gebildet werden. Der folgende Algorithmus berechnet für einen beliebigen DFA M den zugehörigen Minimal-DFA M'.

Algorithmus `min-DFA`(M)

Input: DFA $M = (Z,\Sigma,\delta,q_0,E)$
  entferne alle nicht erreichbaren Zustaende
  $D' := \{\{z,z'\} \mid z \in E, z' \notin E\}$
  repeat
    $D := D'$
    $D' := D \cup \{\{p,q\} \mid \exists a \in \Sigma : \{\delta(p,a),\delta(q,a)\} \in D\}$
  until $D' = D$
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 56. *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

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

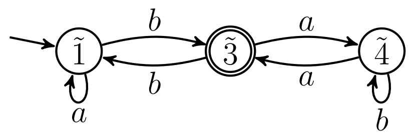
enthält D_1 zusätzlich die Paare $\{1,4\}$, $\{1,5\}$, $\{2,4\}$, $\{2,5\}$ (in obiger Matrix durch den Unterscheider a markiert). Da die verbliebenen Paare $\{1,2\}$, $\{3,6\}$, $\{4,5\}$ wegen

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

nicht zu D_1 hinzugefügt werden können, ist $D_2 = D_1$. Aus den unmarkierten Paaren $\{1,2\}$, $\{3,6\}$ und $\{4,5\}$ erhalten wir die Äquivalenzklassen

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

die auf folgenden Minimal-DFA M' führen:



◁