

Mathematische Grundlagen der Computerlinguistik

Organisatorisches

Dozentin: Wiebke Petersen

0. Foliensatz

Das 1. Semester

Zentrale Grundlagenveranstaltungen:

- Informatik
- Logik
- math. Grundlagen
- Grundkurs Linguistik

Tipps:

- stellen Sie sich den Prüfungen
- überprüfen Sie am Ende des 1. Semesters Ihre Studienfachwahl:
 - wenn Ihnen die Prüfungen schwergefallen sind, Sie sich aber für das Fach begeistern können, machen Sie weiter
 - wenn Sie bereits am Ende des 1. Semesters Ihr Studium nur als Pflicht betrachten, erwägen Sie einen Fachwechsel
 - wenn Sie Informationswiss. und Sprachtech. studieren, müssen Sie sich für beide Fachbereiche begeistern können
- beteiligen Sie sich aktiv an Ihrem Studiengang (Mitarbeit in Veranstaltungen, Fachschaft, Institutsfeiern, Fachvorträge, ...)
- blicken Sie über den Tellerrand (Veranstaltungen anderer Fachbereiche)

Spezielles Hilfsangebot für Sie

was: Sprechstunden für Fragen zur Computerlinguistik (fachliche und organisatorische), Hausaufgabenhilfe und Klausurvorbereitung

wann: **Oliver:** Mo., 12:00 - 13:30 Uhr in 23.21.00.91 (oliver.hanraths@uni-duesseldorf.de)

Patrick: Di., 12:30 Uhr in 23.21.04.87 (patrick.hommers@uni-duesseldorf.de)

durch wen: Studierende höherer Fachsemester

für wen: alle Studierende der Studiengänge „BA Informationswissenschaft und Sprachtechnologie“ und „BA Linguistik mit Schwerpunkt Computerlinguistik“

Semesterplan (math. Grundlagen)

- 1 Mengen und Mengenoperationen
- 2 Relationen und Funktionen
- 3 formale Sprachen
- 4 Ordnungsrelationen
- 5 Graphen und Bäume
- 6 Beweise
- 7 Kombinatorik
- 8 Wahrscheinlichkeitstheorie
- 9 (Statistik)

Kurshomepage: http://user.phil-fak.uni-duesseldorf.de/~petersen/WiSe1112_mathGrundl/WiSe1112_mathGrundl_Petersen.html

Voraussetzung für Beteiligungsnachweis (BN)

- Wöchentliche Hausaufgaben
 - Anmelden: <https://app.phil-fak.uni-duesseldorf.de/moodle/login/signup.php>
 - Einschreiben: Mathematische Grundlagen der Computerlinguistik
 - Passwort: mathgrundl
 - Mindestpunktzahl: 50%
 - Abgabe: bis Do. 12:00 Uhr
- Bei Problemen: Email an oliver.hanraths@uni-duesseldorf.de

Mathematische Grundlagen der Computerlinguistik

Mengen und Mengenoperationen

Dozentin: Wiebke Petersen

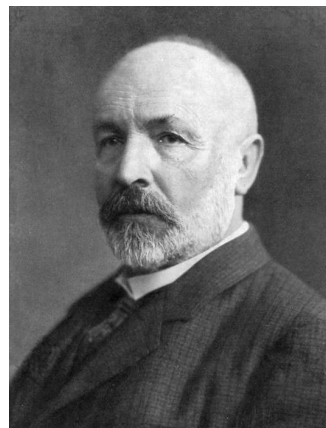
1. Foliensatz

Mengen

Georg Cantor (1845-1918)

„Eine Menge ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die ‚Elemente‘ genannt werden) zu einem Ganzen.“

- Mengen werden über ihre Elemente bestimmt.
- Elemente von Mengen können selber Mengen sein.
- Mengen können endlich oder unendlich sein.



Notation und Terminologie

- Variablen für Mengen: $A, B, C, \dots, M, N, \dots$
- Variablen für Elemente: a, b, c, \dots, x, y, z
- Ist m ein Element von M so schreibt man $m \in M$.
- Ist m kein Element von M so schreibt man $m \notin M$.
- Zwei Mengen A und B sind genau dann **identisch** oder **gleich**, wenn jedes Element von A auch Element von B ist und wenn jedes Element von B auch Element von A ist.
- Es gibt genau eine Menge, die keine Elemente enthält, die **leere Menge** (Symbol: \emptyset).
- Mengen mit genau einem Element werden **Einermengen** (*singleton*) genannt.
- $\mathbb{N} = \{1, 2, 3, \dots\}$ ist die Menge der natürlichen Zahlen
- $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ ist die Menge der natürlichen Zahlen mit 0
- $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ ist die Menge der ganzen Zahlen
- \mathbb{Q} ist die Menge der rationalen Zahlen (alle ‚Bruchzahlen‘).
- \mathbb{R} ist die Menge der reellen Zahlen (alle ‚Kommazahlen‘).

Bertrand Russell (1872-1970)

Russels Antinomie (1901)

- Sei M die Menge aller Mengen, die sich nicht selbst als Element enthalten.
- Gilt $M \in M$ oder $M \notin M$?

Ausweg: ‚Theorie der Typen‘ (Principia Mathematica, Russel & Whitehead 1910-13)

Mengen werden stufenweise aufgebaut und sind immer von einem höheren Typ als ihre Elemente.



Grellings Paradoxie

Ein Adjektiv heie

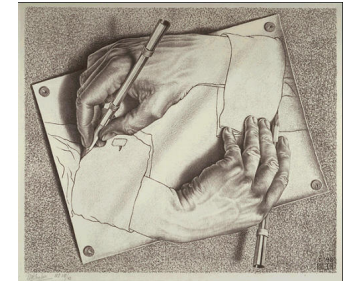
autologisch, wenn es sich selbst beschreibt (Bsp.: dreisilbig, haplogistisch, kurz, xenonymisch, adjektivisch, verbal, vokalenthaltend, exquisit, ...)

heterologisch, wenn es sich nicht selbst beschreibt (Bsp.: zweisilbig, essbar, grn, ...)

Ist ‚heterologisch‘ heterologisch? (nach D.R. Hofstadter: *Gdel, Escher, Bach*)

In diesem Kurs werden Mengen so beschrieben, dass keine Paradoxien auftreten.

Paradoxien der Selbstbezglichkeit



zeichnende Hnde von M.C. Escher

Mengenbeschreibungen

Explizite Mengendarstellung

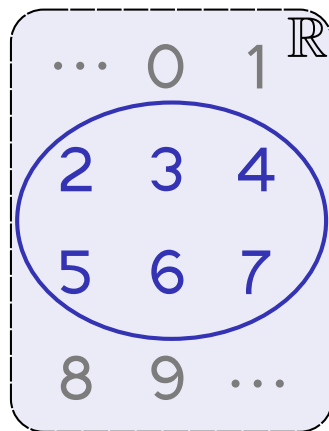
$\{a_1, a_2, \dots, a_n\}$ ist die Menge, die genau die Elemente a_1, a_2, \dots, a_n enthlt.

Beispiel:
 $\{2, 3, 4, 5, 6, 7\}$

Implizite Mengendarstellung

$\{x|A\}$ ist die Menge, die genau die Objekte x enthlt, auf die die Aussage A zutrifft.

Beispiel:
 $\{x \in \mathbb{R} | x \in \mathbb{N} \text{ und } 1 < x \text{ und } x < 8\}$,



Hinweise zur expliziten Mengendarstellung

- Beschreibung durch Aufzhlung oder -listung
- nur fr endliche Mengen mglich
- Die Klammern $\{$ und $\}$ heien **Mengenklammern** oder geschweifte Klammern.
- Reihenfolge der Elemente spielt keine Rolle: $\{a, b, c\} = \{c, a, b\}$
- Elemente knnen in der Klammernotation mehrfach auftreten:
 $\{a, b, c\} = \{a, b, a, b, a, b, c\}$

Hinweise zur impliziten Mengendarstellung

Beschreibung mittels charakteristischer Eigenschaft

- $\{ \text{Element} \in \text{Grundbereich} \mid \text{Eigenschaft von Element} \}$
- $\{x \in G \mid E(x)\}$ („Menge aller x in G mit der Eigenschaft E “)
- Bsp.: $\{x \in \mathbb{N} \mid x \text{ ist eine gerade Zahl}\}$
- Wenn der Grundbereich aus dem Kontext bekannt ist oder sich aus der Eigenschaft ergibt, kann er weggelassen werden.
- Bsp.: $\{x \mid x \text{ ist eine Primzahl}\}$

Hinweise zur impliziten Mengendarstellung

Beschreibung mittels rekursiver Definition

Beispiel: Menge der Nachkommen von Georg Cantor

- 1 **Festlegung endlich vieler Startelemente:**
Die Kinder von Cantor sind Nachkommen von Cantor
 - 2 **Konstruktionsvorschrift für zusätzliche Elemente:**
Wenn x ein Nachkomme von Cantor ist, dann ist jedes Kind von x ein Nachkomme von Cantor.
 - 3 **Einschränkung:**
Nichts sonst ist ein Nachkomme von Cantor.
- Was ist, wenn Cantor keine Kinder hatte?
 - Lässt sich so auch die Menge der Nachkommen von Aristoteles definieren? oder die von Merlin?

Teilmengen

Eine Menge N ist eine **Teilmenge** der Menge M (in Zeichen: $N \subseteq M$) genau dann, wenn alle Elemente von N auch Elemente von M sind.

- Wenn $x \in N$, dann $x \in M$
- Wenn $y \in M$, dann muss $y \in N$ nicht unbedingt gelten, es kann aber gelten.

Eine Menge N ist eine **echte Teilmenge** der Menge M (in Zeichen: $N \subset M$) genau dann, wenn N eine Teilmenge von M ist und wenn M und N ungleich sind.

- $N \subseteq M$ und $N \neq M$
- Es gibt ein $y \in M$ mit $y \notin N$.

Wenn $N \subseteq M$, dann ist M eine **Übermenge** von N (in Zeichen: $M \supseteq N$).

Wenn $M \supseteq N$ und $M \neq N$ dann ist M eine **echte Übermenge** von N (in Zeichen: $M \supset N$).

Teilmengen

$x \in M$: x ist ein **Element** der Menge M

- $2 \in \{1, 2, 3\}$
- $2 \notin \{1, 3, 5\}$
- $\{3\} \in \{M \mid M \text{ ist eine Einermenge}\}$
- $\{3\} \notin \{3\}$

$N \subseteq M$: Die Menge N ist eine **Teilmenge** der Menge M

- $\{2, 3\} \subseteq \{1, 2, 3, 4\}$
- $\{2, 3\} \subseteq \{2, 3\}$
- $\emptyset \subseteq \{1, 2, 3, 4\}$
(Die leere Menge ist eine Teilmenge jeder Menge!)
- $\{3\} \not\subseteq \{M \mid M \text{ ist eine Einermenge}\}$

$N \subset M$: Die Menge N ist eine **echte Teilmenge** der Menge M

- $\{1\} \subset \{1, 2\}$
- $\{1, 2\} \not\subset \{1, 2\}$

Vorsicht: Die Element-von- und die Teilmengenrelation müssen streng unterschieden werden!

Mächtigkeit von Mengen

Zwei Mengen M und N haben dieselbe **Mächtigkeit** oder heißen **gleichmächtig** (in Zeichen: $|M| = |N|$), wenn es eine eindeutige Zuordnung der Elemente von M auf N gibt (d.h., die Zuordnung ordnet jedem Element aus M genau ein Element aus N und jedem Element aus N genau ein Element aus M zu.)

endliche Mengen

Die **Mächtigkeit** einer endlichen Menge (in Zeichen: $|M|$) ist die Anzahl ihrer Elemente.

Beispiele:

- $|\emptyset| = 0$
- $|\{1, 2\}| = 2$
- $|\{\{1, 2\}\}| = 1$

Vorsicht: nicht alle unendlichen Mengen sind gleichmächtig!

Mengenoperationen (unäre Potenzmengenoperation)

Mengenoperationen sind Abbildungen, die einer oder mehreren Mengen eindeutig eine Menge zuordnen. Einstellige Operationen werden auch **unäre** und zweistellige auch **binäre** Operationen genannt.

Die Potenzmengenoperation ist eine unäre Operation, die jeder Menge ihre Potenzmenge zuordnet.

Die **Potenzmenge** einer Menge M ist die Menge aller möglichen Teilmengen von M , also $\mathcal{P}\mathcal{O}\mathcal{T}(M) = \{N | N \subseteq M\}$.

$$\mathcal{P}\mathcal{O}\mathcal{T}(\{1, 2, 3\}) = \left\{ \begin{array}{l} \{ \quad \quad \quad \}, \\ \{ 1 \quad \quad \quad \}, \\ \{ \quad 2 \quad \quad \quad \}, \\ \{ \quad \quad 3 \quad \quad \quad \}, \\ \{ 1, 2 \quad \quad \quad \}, \\ \{ 1, \quad \quad 3 \quad \quad \quad \}, \\ \{ \quad 2, \quad 3 \quad \quad \quad \}, \\ \{ 1, 2, 3 \quad \quad \quad \} \end{array} \right\}$$

Mächtigkeit der Potenzmenge

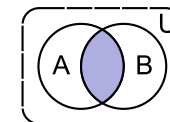
Für endliche Mengen gilt: ist M eine n -elementige Menge, so ist $|\mathcal{P}\mathcal{O}\mathcal{T}(M)| = 2^n$.

| 1 | 2 | 3 | ... | n |
|------------|------------|------------|-----|---|
| - | - | - | ... | - |
| + | - | - | ... | - |
| - | + | - | ... | - |
| - | - | + | ... | - |
| ⋮ | | | | ⋮ |
| - | - | - | ... | + |
| + | + | - | ... | - |
| + | - | + | ... | - |
| ⋮ | | | | ⋮ |
| + | + | + | ... | + |
| $2 \times$ | $2 \times$ | $2 \times$ | ... | 2 |

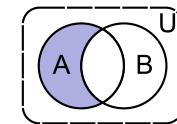
2^n Möglichkeiten

Mengenoperationen (binäre Operationen)

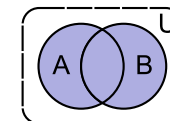
Schnitt: $A \cap B$
 „A geschnitten mit B“
 $A \cap B = \{x | x \in A \text{ und } x \in B\}$



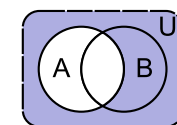
Differenz: $A \setminus B$
 „A ohne B“
 $A \setminus B = \{x | x \in A \text{ und } x \notin B\}$



Vereinigung: $A \cup B$
 „A vereinigt mit B“
 $A \cup B = \{x | x \in A \text{ oder } x \in B\}$



Komplement (in U): $C_U(A)$
 „Komplement von A in U“
 $C_U(A) = U \setminus A$



Wenn U feststeht, schreibt man auch \bar{A}

Mengenoperationen

Beispiele

$A = \{1, 2, 3, 4\}$, $B = \{3, 4, 5\}$, $U = \{1, 2, 3, 4, 5, 6, 7\}$,

- $A \cup B = \{1, 2, 3, 4, 5\}$, $A \cap B = \{3, 4\}$
- $A \setminus B = \{1, 2\}$, $\bar{A} = \{5, 6, 7\}$

Notation

- Zwei Mengen A und B mit leerem Schnitt heißen **disjunkt** ($A \cap B = \emptyset$).
- Wenn A eine Menge von Mengen ist, schreiben wir $\cup A$ für die Vereinigung aller Elemente von A (Bsp.: $\cup\{B, C, D\} = B \cup C \cup D$)
- Wenn A eine Menge von Mengen ist, schreiben wir $\cap A$ für den Schnitt aller Elemente von A (Bsp.: $\cap\{B, C, D\} = B \cap C \cap D$)
- Häufig werden auch Indizes und Indexmengen zur Notation verwendet. Bsp.: Sei $A_i = \{x \in \mathbb{N}_0 \mid x \leq i\}$, dann

$$\bigcup_{3 \leq i \leq 5} A_i = \{0, 1, 2, 3, 4, 5\} \text{ und } \bigcap_{3 \leq i \leq 5} A_i = \{0, 1, 2, 3\}$$

Eigenschaften der Mengeoperationen (Schnitt und Vereinigung)

Kommutativgesetz:

$$A \cap B = B \cap A$$
$$A \cup B = B \cup A$$

Assoziativgesetz:

$$(A \cap B) \cap C = A \cap (B \cap C)$$
$$(A \cup B) \cup C = A \cup (B \cup C)$$

Distributivgesetz:

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

Idempotenzgesetz:

$$A \cap A = A$$
$$A \cup A = A$$

\emptyset ist **neutrales Element** der Vereinigung: $A \cup \emptyset = \emptyset \cup A = A$

Gibt es auch ein neutrales Element des Schnitts?

Gesetze der Komplementoperation

de Morgan:

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$
$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

weitere Gesetze:

$$\overline{\bar{A}} = A$$
$$\bar{A} \cap A = \emptyset$$

Mathematische Grundlagen der Computerlinguistik

Relationen und Funktionen

Dozentin: Wiebke Petersen

2. Foliensatz

n -Tupel und Cartesisches Produkt

Mengen sind ungeordnet, häufig werden jedoch geordnete Listen benötigt:

n -Tupel

Ein n -Tupel ist eine Liste mit $n \geq 1$ Elementen. Im Gegensatz zu Mengen ist die Reihenfolge festgelegt und jedes Element kann beliebig oft vorkommen.

Beispiel: $\langle 2, 3, 1 \rangle$, $\langle b, e, e, s, i, i, p, l \rangle$

2-Tupel werden auch (**geordnete**) **Paare** genannt.

Cartesisches Produkt

Das **Cartesische Produkt** (oder Kreuzprodukt) von n Mengen $M_1 \dots M_n$ ist die Menge aller n -Tupel deren i -tes Element aus M_i stammt.

$$M_1 \times \dots \times M_n := \{ \langle x_1, \dots, x_n \rangle \mid x_i \in M_i \text{ für } i = 1, \dots, n \}$$

Statt $M \times M \times \dots \times M$ schreibt man auch M^n , wenn M genau n -mal auftritt.

Beispiel

$$M_1 = \{a, b, c\}, M_2 = \{a, d\}$$

$$M_1 \times M_2 = \{ \langle a, a \rangle, \langle a, d \rangle, \langle b, a \rangle, \langle b, d \rangle, \langle c, a \rangle, \langle c, d \rangle \}$$

$$M_1 \times \emptyset = \emptyset$$

Relationen

Definition

Eine **Teilmenge** des Cartesischen Produktes von n Mengen $R \subseteq M_1 \times \dots \times M_n$ heißt **n -stellige Relation**.

Eine Relation R ist also eine Menge von n -Tupeln.

Hinweis: Relationen werden **extensional** definiert. Es ist unerheblich, wie die Relation charakterisiert (oder benannt) wird. Wichtig ist allein, welche Objekte zueinander in der Relation stehen.

Für Relationen werden häufig die Buchstaben R, S, T verwendet.

Beispiele

- Schwester von
- Mutter von
- weibliches Elternteil von
- bilden ein Quartet
- Teilmenge von

binäre Relationen

- binäre Relationen sind Mengen geordneter Paare
- wenn a in der Relation R zu b steht, dann schreibt man
 - $\langle a, b \rangle \in R$ oder
 - aRb oder
 - $R(a, b)$ oder
 - Rab
- Wenn $R \subseteq A \times B$, dann sagt man, dass R eine Relation zwischen A und B ist.
- Wenn $R \subseteq A \times A$, dann sagt man, dass R eine Relation auf A ist.

inverse und komplementäre Relation

inverse Relation

Die **inverse Relation** zu einer binären Relation $R \subseteq A \times B$ ist die Relation

$$R^{-1} = \{ \langle b, a \rangle \in B \times A \mid \langle a, b \rangle \in R \}.$$

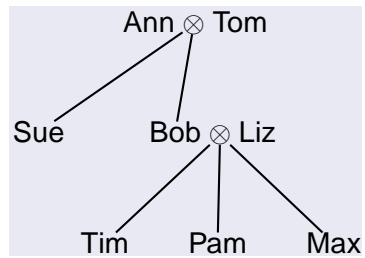
komplementäre Relation

Die **komplementäre Relation** zu einer binären Relation $R \subseteq A \times B$ zwischen A und B ist die Relation

$$R' = A \times B \setminus R.$$

Beispiel: Verwandtschaftsterme

Beispielfamilie

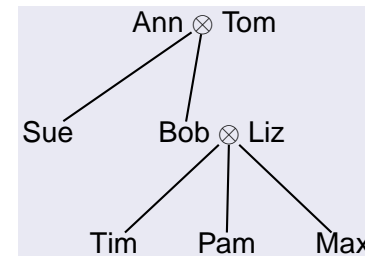


„hat als Sohn“

| | | |
|-----|------------------|-----|
| Ann | R_{son} | Bob |
| Tom | R_{son} | Bob |
| Bob | R_{son} | Max |
| Bob | R_{son} | Tim |
| Liz | R_{son} | Max |
| Liz | R_{son} | Tim |

Beispiel: Verwandtschaftsterme

Beispielfamilie



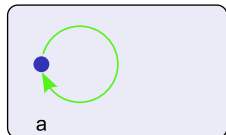
„hat als Mutter“

| | | |
|-----|---------------------|-----|
| Sue | R_{mother} | Ann |
| Bob | R_{mother} | Ann |
| Tim | R_{mother} | Liz |
| Pam | R_{mother} | Liz |
| Max | R_{mother} | Liz |

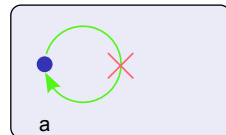
Eigenschaften binärer Relationen

Sei $R \subseteq A \times A$ eine binäre Relation auf A .

R ist **reflexiv** g.d.w. für alle $a \in A$ gilt, dass aRa .



R ist **irreflexiv** g.d.w. für kein $a \in A$ gilt, dass aRa .

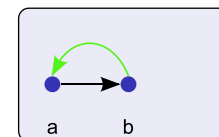


- Die Relation ‚hat am selben Tag Geburtstag‘ auf der Menge der Menschen ist reflexiv.
- Die Relation ‚ist Mutter von‘ auf der Menge der Menschen ist irreflexiv.
- Die Relation ‚kann die Quersumme des Geburtstags von berechnen‘ auf der Menge der Menschen ist weder reflexiv noch irreflexiv.
- Welche Bedingungen erfüllen die Beispielrelationen an der Tafel?

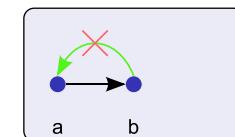
Eigenschaften binärer Relationen

Sei $R \subseteq A \times A$ eine binäre Relation auf A .

R ist **symmetrisch** g.d.w. für alle $a, b \in A$ mit aRb auch bRa gilt.



R ist **asymmetrisch** g.d.w. für $a, b \in A$ niemals sowohl aRb als auch bRa gilt.



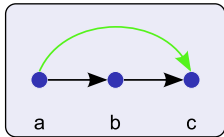
R ist **antisymmetrisch** g.d.w. für alle $a, b \in A$ aus aRb und bRa folgt, dass $a = b$.

- Die Relation ‚ist verheiratet mit‘ ist symmetrisch.
- Die Relation ‚ist größer als‘ ist asymmetrisch.
- Die Relation ‚ist Teilmenge von‘ ist antisymmetrisch.
- Welche Bedingungen erfüllen die Beispielrelationen an der Tafel?

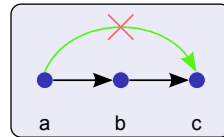
Eigenschaften binärer Relationen

Sei $R \subseteq A \times A$ eine binäre Relation auf A .

R ist **transitiv** g.d.w. für alle $a, b, c \in A$ aus aRb und bRc immer aRc folgt.



R ist **intransitiv** g.d.w. für alle $a, b, c \in A$ mit aRb und bRc niemals aRc gilt.



- Die Relation ‚ist Vorfahr von‘ ist transitiv.
- Die Relation ‚steht genau eine Treppenstufe höher als‘ ist intransitiv.
- Die Relation ‚kennt‘ ist weder transitiv noch intransitiv.
- Welche Bedingungen erfüllen die Beispielrelationen an der Tafel?

Definitions- und Wertebereich einer Relation

Wenn $R \subseteq A \times B$ eine binäre Relation ist, dann heißt

$$\text{dom}(R) = \{a \in A \mid \text{es gibt ein } b \in B \text{ mit } (a, b) \in R\}$$

der **Definitionsbereich (domain)** von R .

Die Menge

$$\text{rng}(R) = \{b \in B \mid \text{es gibt ein } a \in A \text{ mit } (a, b) \in R\}$$

heißt der **Wertebereich (range)** von R .

Beispiel:

$$A = \{a, b, c, d\}, B = \{1, 2, 3, 4, 5\}, R = \{(b, 1), (b, 2), (c, 3)\}$$

$$\text{dom}(R) = \{b, c\}, \text{rng}(R) = \{1, 2, 3\}$$

Äquivalenzrelation

Äquivalenzrelation

Eine Relation $R \subseteq A \times A$ ist eine **Äquivalenzrelation** auf A , g.d.w. R reflexiv, symmetrisch und transitiv ist.

Wenn R eine Äquivalenzrelation ist und aRb gilt, dann sagt man, dass a äquivalent ist zu b bezüglich R .

Für Äquivalenzrelationen verwendet man häufig das Symbol \sim .

Beispiele:

- Gleichheit
- ist im selben Semester wie
- hat gleich viele Elemente wie
- hat die selbe Farbe wie
- Welche der Beispielrelationen an der Tafel sind Äquivalenzrelationen?

Äquivalenzrelation

Äquivalenzklasse

Sei R eine Äquivalenzrelation auf A . Die **Äquivalenzklasse** eines Elements $a \in A$ ist die Menge aller zu a äquivalenten Elemente von A , also

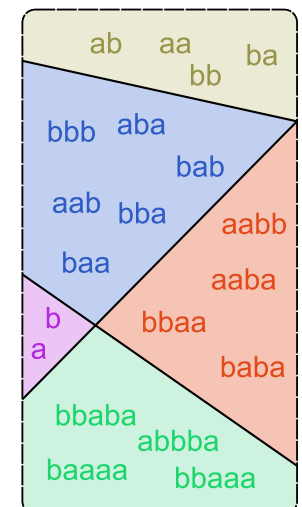
$$[a]_R = \{b \in A \mid aRb\}.$$

Die Menge

$$A/R = \{[a]_R \mid a \in A\}$$

aller Äquivalenzklassen von Elementen aus A bezüglich R heißt **Quotient** von A bezüglich R .

Hinweis: Äquivalenzklassen können per Definition nicht leer sein.



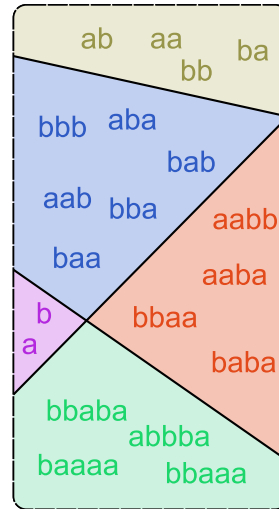
Äquivalenzrelation

Sei R eine Äquivalenzrelation auf A .
Dann gilt:

- Zwei Äquivalenzklassen von R sind entweder disjunkt oder identisch: für alle $a, b \in A$ gilt entweder $[a]_R \cap [b]_R = \emptyset$ oder $[a]_R = [b]_R$.
- Die Äquivalenzklassen von R decken ganz A ab: $\bigcup A/R = A$.

Eine Menge $P \subseteq \mathcal{P}\mathcal{O}\mathcal{T}(A)$ ist eine **Partition** (oder disjunkte Zerlegung) von A , g.d.w. $\bigcup P = A$ und für alle $X, Y \in P$ mit $X \neq Y$ gilt $X \cap Y = \emptyset$.

Folglich bildet der Quotient einer Äquivalenzrelation eine Partition der Grundmenge.



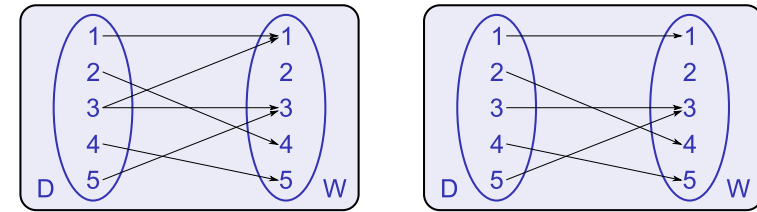
Funktionen

Definition

Eine Relation $R \subseteq D \times W$ ist eine **Funktion** (oder **Abbildung**), wenn sie **jedem** Element aus D **genau ein** Element aus W zuordnet.
Funktionen müssen also die Bedingungen der Existenz und Eindeutigkeit erfüllen:

Existenz: Für **jedes** $x \in D$ gibt es ein $y \in W$ mit $\langle x, y \rangle \in R$.

Eindeutigkeit: Wenn $\langle x, y \rangle \in R$ und $\langle x, z \rangle \in R$, dann $y = z$.



Eine Relation, für die die Eindeutigkeitsbedingung (aber nicht unbedingt die Existenzbedingung) gilt, heißt **partielle Funktion**.

Notation und Terminologie

- Für Funktionen verwendet man häufig die Buchstaben f, g, h, F, G, H .
- Wenn $f \subseteq A \times B$ eine Funktion ist, dann sagt man, dass f eine Funktion von A nach B ist, und schreibt $f : A \rightarrow B$.
- Wenn $\langle a, b \rangle \in f$, dann sagt man, dass die Funktion f dem Element a den Wert b zuweist, und schreibt $f(a) = b$ oder $f : a \mapsto b$.
- Elemente des Definitionsbereiches heißen **Argumente** und Elemente des Wertebereiches heißen **Werte** einer Funktion.
- Wenn $C \subset A$ und $f : A \rightarrow B$, dann bezeichnet $f|_C : C \rightarrow B$ die **Einschränkung** der Funktion f auf C . Für alle $c \in C$ gilt $f|_C(c) = f(c)$.
- Im Kontext von partiellen Funktionen werden Funktionen, die die Existenzbedingung erfüllen, häufig **totale Funktionen** genannt.

Beispiele

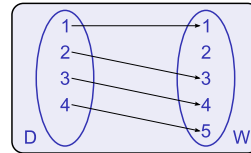
Sei $A = \{a, b, c, d\}$ $B = \{1, 2, 3, 4, 5\}$

- Die Relation $R \subseteq A \times B$ mit $R = \{(b, 1), (b, 2), (c, 3)\}$ ist keine partielle Funktion.
- Die Relation $R \subseteq A \times B$ mit $R = \{(b, 1), (c, 3), (d, 1)\}$ ist eine partielle aber keine totale Funktion.
- Die Relation $R \subseteq A \times B$ mit $R = \{(a, 2), (b, 1), (c, 3), (d, 1)\}$ ist eine totale und folglich auch eine partielle Funktion.

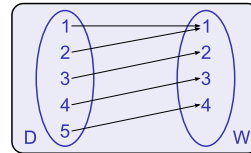
Funktionseigenschaften

Sei $f: D \rightarrow W$ eine Funktion.

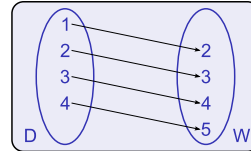
f ist **injektiv** (Engl.: one-to-one), wenn keine zwei verschiedenen Elemente des Definitionsbereiches denselben Wert zugewiesen bekommen. Wenn also für alle $x, y \in D$ gilt:
 $f(x) = f(y)$ g.d.w. $x = y$.



f ist **surjektiv** (Engl.: onto), wenn jedes Element von W mindestens einem Element von D als Wert zugewiesen wird. Wenn es also für jedes $y \in W$ ein $x \in D$ gibt, für das $f(x) = y$ gilt.



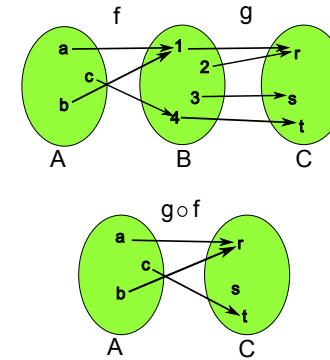
f ist **bijektiv**, wenn f **injektiv** **und** **surjektiv** ist. Merke: Eine Funktion f ist bijektiv, g.d.w. f^{-1} eine Funktion ist.



Komposition von Funktionen

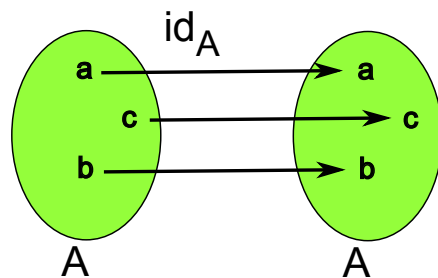
Seien $f: A \rightarrow B$ und $g: B \rightarrow C$ zwei Funktionen. Die Funktion $g \circ f: A \rightarrow C$ mit $g \circ f = \{(x, z) \in A \times C \mid \text{es gibt ein } y \in B \text{ mit } (x, y) \in f \text{ und } (y, z) \in g\}$ ist die **Komposition** (oder **Verkettung**) von f und g .

Es gilt $(g \circ f)(x) = f(g(x))$. Die Funktion $g \circ f$ weist einem Element $x \in A$ das Element aus C zu, das man erhält, wenn man zunächst g auf x anwendet und auf das Ergebnis noch f anwendet.



Identitätsfunktion

Die Funktion $id_A: A \rightarrow A$ mit $f = \{(a, a) \in A \times A\}$ (oder $f(a) = a$ für alle $a \in A$) heißt die **Identität(sfunktion)** auf A .



mehrstellige Funktionen

- Der Definitionsbereich einer Funktion kann selbst eine Relation sein.
- Eine Funktion $A_1 \times A_2 \times \dots \times A_n \rightarrow B$ heißt n -stellige Funktion.
- Beispiel: Die Addition der natürlichen Zahlen $+: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{N}_0$ kann als zweistellige Funktion aufgefasst werden.
- Zweistellige Operationen bilden zweistellige Funktionen (Bsp.: Schnitt, Vereinigung, ...).
- n -stellige Funktionen sind $n+1$ -stellige Relationen (Bsp: Mutter)

Exkurs: algebraische Struktur

- Die Mathematik ist die Wissenschaft der Strukturen
- Strukturen sind Paare bestehend aus einer Grundmenge und einer endlichen Menge von Operationen auf der Grundmenge. (Operationen sind Abbildungen/Funktionen von M^n in M).
- Wir kennen bereits die Struktur $(\mathcal{POT}(M), \{\cup, \cap, -, \setminus\})$
- Oder auch $(\mathbb{N}_0, \{+\})$

Sei $(M, \{\otimes\})$ eine Struktur, dann

- erfüllt sie das **Kommutativgesetz**, wenn $a \otimes b = b \otimes a$ für alle $a, b \in M$.
- erfüllt sie das **Assoziativgesetz**, wenn $(a \otimes b) \otimes c = a \otimes (b \otimes c)$ für alle $a, b, c \in M$.
- e ist das **neutrale Element** der Struktur, wenn $a \otimes e = e \otimes a = a$ für alle $a \in M$.

Mathematische Grundlagen der Computerlinguistik

formale Sprachen

Dozentin: Wiebke Petersen

3. Foliensatz

Alphabete und Wörter

Definition

- **Alphabet Σ** : endliche Menge von **Symbolen / Zeichen**.
- **Wort**: eine endliche Kette/Folge $x_1 \dots x_n$ von Symbolen/Zeichen eines Alphabets (mit $n \geq 0$). Das Wort, das aus null Zeichen besteht heißt **leeres Wort** und wird mit ε bezeichnet.
- Die Menge aller Wörter über einem Alphabet Σ bezeichnen wir mit Σ^* .
 $\Sigma^+ = \Sigma^* \setminus \{\varepsilon\}$ ist die Menge der nichtleeren Wörter.
- **Länge** eines Wortes $|w|$: Gesamtzahl der Zeichen eines Wortes w ($|abbaca| = 6$, $|\varepsilon| = 0$)

Leersymbol, leeres Wort und leere Menge

Vorsicht Verwechslungsgefahr!

Das **Leersymbol** \sqcup ist ein **Zeichen** des Alphabets, also ist ein Wort, das nur aus dem Leersymbol besteht, ein Wort der Länge 1.

Das **leere Wort** ε ist ein **Wort** der Länge 0.

Die **leere Menge** \emptyset ist eine **Menge**.

Übung: Alphabete und Wörter

Sei $\Sigma = \{a, b, c\}$ ein Alphabet:

- Geben Sie zwei Wörter der Länge 4 über Σ an.
- Welche der folgenden Ausdrücke sind Wörter über Σ und welche Länge haben sie?:
'aa', 'caab', 'da'
- Was ist der Unterschied zwischen Σ^* , Σ^+ und Σ ?
- Wieviele Elemente haben Σ , Σ^* und Σ^+ ?

Operationen auf Wörtern

Verkettung / Konkatenation

Die **Konkatenation / Verkettung** zweier Wörter $u = a_1 a_2 \dots a_n$ und $v = b_1 b_2 \dots b_m$ mit $n, m \geq 0$ ist

$$u \circ v = a_1 \dots a_n b_1 \dots b_m$$

Häufig schreiben wir uv statt $u \circ v$.

$$w \circ \varepsilon = \varepsilon \circ w = w \quad \text{Neutrales Element}$$

$$u \circ (v \circ w) = (u \circ v) \circ w \quad \text{Assoziativität}$$

Ist die Konkatenationsoperation kommutativ?

Symbolpolitik der Mathematik

Vorsicht:

- Obwohl die Symbole für die Komposition von Funktionen und die Konkatenation von Wörtern übereinstimmen, handelt es sich um unterschiedliche Operationen!
- In der Mathematik finden sie häufig mehrdeutige Symbole, deren Bedeutung sich aus dem jeweiligen Kontext ergibt.
- Sie müssen sich also bei dem Symbol \circ immer fragen, ob es zwischen Funktionen oder Wörtern steht (wir werden auch noch eine Operation auf Mengen kennenlernen, die mit demselben Symbol bezeichnet wird).
- Bedenken Sie, dass die Alternative die Verwendung einer unbegrenzten Zahl verschiedener Symbole wäre, da es theoretisch unendlich viele Operationen gibt. Jedes dieser Symbole müsste in Zeichensätzen vorgehalten werden, was unmöglich ist, da Alphabete endlich sein müssen. Stellen Sie sich außerdem vor, ich würde an der Tafel versuchen eine Vielzahl von sehr ähnlichen Symbolen zu verwenden (Beispiel: Kreis mit dickem Punkt in der Mitte, Kreis mit kleinem Punkt, Kreis ohne Punkt, Kreis mit zwei Umrandungen, ...), Sie würden das nicht lesen wollen!

Operationen auf Wörtern

Exponenten

- w^n : w wird n -mal mit sich selbst verkettet.
- $w^0 = \varepsilon$: w wird '0-mal' mit sich selbst verkettet.

Umkehrung

- Die **Umkehrung** eines Wortes w wird mit w^R bezeichnet.
 $(abcd)^R = dcba$.
- Ein Wort w , für das $w = w^R$ gilt, heißt **Palindrom**.

(madam, reliefpfeiler, otto, anna, regallager ...)

Übung: Operationen auf Wörtern

Seien $w = aabc$ und $v = bcc$ Wörter, ermitteln Sie:

- $w \circ v$
- $((w^R \circ v)^R)^2$
- $w \circ (v^R \circ w^3)^0$

Formale Sprache

Definition

Eine **formale Sprache** L ist eine Menge von Wörtern über einem Alphabet Σ , also $L \subseteq \Sigma^*$.

Beispiele:

- Sprache L_{rom} der gültigen römischen Zahldarstellungen über dem Alphabet $\Sigma_{rom} = \{\mathbf{I}, \mathbf{V}, \mathbf{X}, \mathbf{L}, \mathbf{C}, \mathbf{D}, \mathbf{M}\}$.
- Sprache L_{Mors} der Buchstaben des lateinischen Alphabets dargestellt im Morsecode. $L_{Mors} = \{ \cdot - , - \cdot \cdot , \dots , - - \cdot \}$
- Sprache L_{pal} der Palindrome im deutschen Duden $L_{pal} = \{\text{Madam, reliefpfeiler, } \dots\}$
- Leere Menge
- Menge der Wörter der Länge 13 über dem Alphabet $\{a, b, c\}$
- Sprache der syntaktisch wohlgeformten Java-Programme
- Deutsch?

Operationen auf Sprachen

Seien $L \subseteq \Sigma^*$ und $K \subseteq \Sigma^*$ zwei Sprachen über dem Alphabet Σ , dann entstehen durch die Verknüpfung mit Mengenoperatoren neue Sprachen über Σ :

$$K \cup L, K \cap L, K \setminus L$$

Die Verkettung von Wörtern kann ausgedehnt werden auf die Verkettung von Sprachen:

$$K \circ L := \{v \circ w \in \Sigma^* \mid v \in K, w \in L\}$$

Beispiel: Sei $K = \{abb, a\}$ und $L = \{bbb, ab\}$

- $K \circ L = \{abbbbb, abbab, abbb, aab\}$ und $L \circ K = \{bbbabb, bbba, ababb, aba\}$
- $K \circ \emptyset = \emptyset$
- $K \circ \{\epsilon\} = K$
- $K^2 = K \circ K = \{abbabb, abba, aabb, aa\}$

Potenzen von Sprachen, Iteration, Kleene-Stern

Die n -te Potenz einer Sprache L ist die n -fache Verkettung von L mit sich selbst:

$$L^n = \underbrace{L \circ L \circ L \dots \circ L}_{n\text{-mal}}$$

Induktive Definition:

$$L^0 = \{\epsilon\}, L^{n+1} = L^n \circ L$$

Die Iteration (Kleene-Stern) von L ist

$$L^* := \bigcup_{n \geq 0} L^n$$

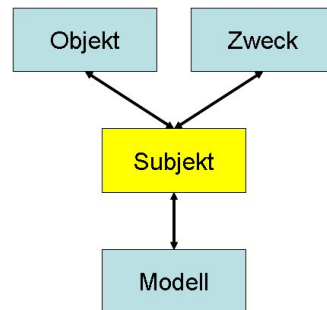
Für jede beliebige Sprache L gilt: $\epsilon \in L^*$

Modell

- künstlich geschaffen
- materiell oder immateriell
- vereinfachtes Abbild
- zweckgerichtet
- Abstraktion
- Repräsentation
- Modellierungsannahmen

Modellierung

Ein **Subjekt** entwirft zu einem **Original** ein **Modell** zu einem bestimmten **Zweck**.



Modellierung natürlicher Sprachen

Formale Sprachen

Formale Sprachen sind Mengen von **Wörtern** (entspricht in natürlichen Sprachen den **Sätzen**), die ihrerseits aus **Zeichen/Symbolen** (in natürlichen Sprachen **Wörtern**) aufgebaut sind. Was in der Menge ist, ist ein "grammatisch korrektes Wort", alles andere nicht.

Für "strukturierte" formale Sprachen lassen sich endliche Mengen von Regeln/Grammatiken angeben, die diese beschreiben.

Sprachmodell

Formale Sprachen dienen als Modell für natürliche Sprachen.

Wir gehen davon aus, daß alle natürlichen Sprachen durch endlich viele Regeln beschreibbar sind, da wir sie ansonsten nicht sprechen / verstehen könnten.

Welche Modellannahmen werden hier implizit gemacht?

Sprachbeschreibung durch Aufzählung aller Wörter

- Peter says that Mary has fallen off the tree.
- Oskar says that Peter says that Mary has fallen off the tree.
- Lisa says that Oskar says that Peter says that Mary has fallen off the tree.
- ...

Scheitert bei unendlichen Sprachen.

Aufzählungen erfassen keine Generalisierungen.

Sprachbeschreibung durch Angabe einer Grammatik

Grammatik

- Eine formale Grammatik ist ein generativer Mechanismus zur Erzeugung von Zeichenketten.
- Grammatiken sind endliche Regelsysteme.
- Die Menge aller Ketten, die von einer Grammatik generiert werden, bilden die von der Grammatik beschriebene formale Sprache.

| | | | | | | | | |
|----|---|--------|----|---|-----------|----|---|-----------|
| S | → | NP VP | VP | → | V | VP | → | VP and VP |
| NP | → | D N | NP | → | NP and NP | D | → | the |
| N | → | cat | N | → | dog | V | → | sleeps |
| V | → | dreams | | | | | | |

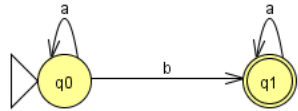
Generiert: the cat sleeps, the dog sleeps, the cat sleeps and dreams, ...

aber auch: the cat and the dog sleeps and dreams, ...

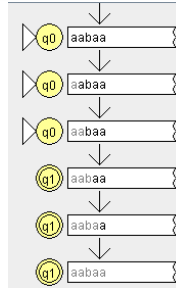
Sprachbeschreibung durch Automaten

Automaten

- Ein Automat ist eine abstrakte Maschine, die bestimmte Zeichenketten akzeptiert.
- Die Menge aller Ketten, die von einem Automaten akzeptiert werden, bilden die von dem Automaten beschriebene formale Sprache.



akzeptiert die Sprache $\{a\}^* \circ \{b\} \circ \{a\}^*$



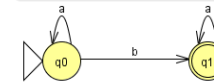
einfachstes Automatenmodell: endliche Automaten

Definition

Ein **endlicher Automat** ist ein 5-Tupel $(Q, \Sigma, \Delta, q_0, F)$ bestehend aus:

- 1 Q : Alphabet der **Zustände**
- 2 Σ : **Eingabealphabet** (Q und Σ müssen disjunkt sein)
- 3 Δ : **Übergangsrelation** ($\Delta \subseteq Q \times \Sigma \times Q$)
- 4 q_0 : **Startzustand** ($q_0 \in Q$)
- 5 F : Menge der **Endzustände** $F \subseteq Q$.

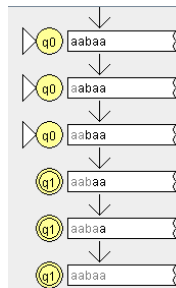
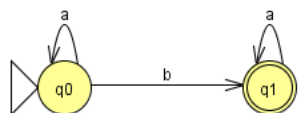
Der Automat heißt **deterministisch**, wenn die Übergangsrelation Δ eine (partielle) Funktion ist ($\Delta : Q \times \Sigma \rightarrow Q$).



endliche Automaten: Akzeptanz von Wörtern

Ein endlicher Automat **akzeptiert** ein Wort w , wenn es möglich ist

- beginnend im Startzustand
- das Wort Symbol für Symbol abzuarbeiten, indem man den Zustand gemäß der Übergangsrelation wechselt
- bis das Wort vollständig abgearbeitet ist,
- und wenn man sich am Ende in einem Endzustand befindet.



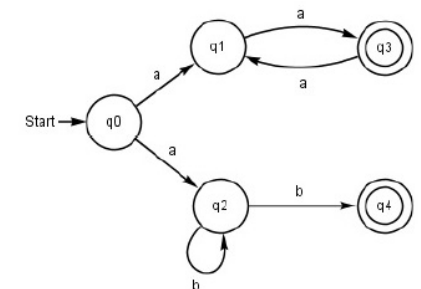
Beispiel: endlicher Automat

als 5-Tupel:

$(Q, \Sigma, \Delta, q_0, F)$ mit

- $Q = \{q_0, q_1, q_2, q_3, q_4\}$
- $\Sigma = \{a, b\}$
- $\Delta = \{(q_0, a, q_1), (q_0, a, q_2), (q_1, a, q_3), (q_3, a, q_1), (q_2, b, q_2), (q_2, b, q_4)\}$
- $F = \{q_3, q_4\}$

als Übergangnetz:



Dieser Automat ist **nicht deterministisch**

(am Übergangnetz ablesbar an identisch beschrifteten Kanten, die von demselben Knoten ausgehen)

Beispiel: endlicher Automat

als 5-Tupel:

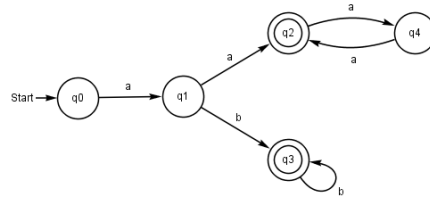
$(Q, \Sigma, \Delta, q_0, F)$ mit

- $Q = \{q_0, q_1, q_2, q_3, q_4\}$
- $\Sigma = \{a, b\}$
- $\Delta = \{(q_0, a, q_1), (q_1, a, q_2), (q_1, b, q_3), (q_3, b, q_3), (q_2, a, q_4), (q_4, a, q_2)\}$
- $F = \{q_2, q_3\}$

Dieser Automat ist **deterministisch** und akzeptiert dieselbe Sprache wie der Automat der vorangegangenen Folie, nämlich $\{a\} \circ ((\{a\} \circ (\{a\} \circ \{a\})^*) \cup (\{b\} \circ \{b\}^*))$.

Dies ist die Sprache aller Wörter über dem Alphabet $\{a, b\}$, die aus einem a gefolgt von einer beliebigen, nichtleeren Kette von b 's oder aus einer nichtleeren Kette von a 's gerader Länge bestehen.

als Übergangsnetz:



Endliche Automaten: Terminologie

- Zwei Automaten, die dieselbe Sprache akzeptieren, heißen **äquivalent** (Beispiel: die Automaten der letzten beiden Folien sind äquivalent)
- Satz: Zu jedem nichtdeterministischen endlichen Automaten gibt es einen äquivalenten deterministischen endlichen Automaten.
- Übergangsrelationen werden häufig als **Übergangstabellen** dargestellt. Beispiel: $\Delta = \{(q_0, a, q_1), (q_1, a, q_2), (q_1, b, q_3), (q_3, b, q_3), (q_2, a, q_4), (q_4, a, q_2)\}$

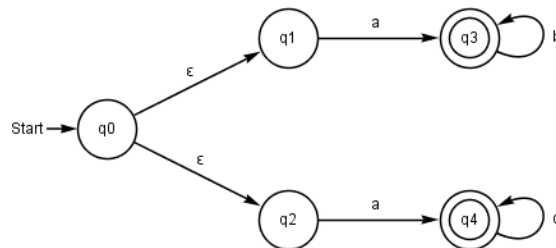
| | a | b |
|----|-------|----|
| q0 | q1 | |
| q1 | q2 q3 | |
| q2 | q4 | |
| q3 | | q3 |
| q4 | q2 | |

- Ist die Übergangsrelation eines endlichen Automaten eine totale Funktion (steht also in jeder Zelle der Übergangstabelle genau ein Element), so ist der Automat ein **endlicher Automat mit vollständiger Übergangsfunktion**

Sind endliche Automaten mit vollständiger Übergangsfunktion immer deterministisch?

Das Programm **Exorciser** bietet sehr gute Übungsmöglichkeiten für die Arbeit mit endlichen Automaten (Website)

endliche Automaten mit ϵ -Übergängen



Zu jedem endlichen Automaten mit ϵ -Übergängen gibt es einen äquivalenten endlichen Automaten ohne ϵ -Übergänge.

Übung

Erstellen Sie endliche Automaten, die die folgenden Sprachen über dem Alphabet $\{a, b\}$ akzeptieren:

- 1 die Sprache aller Wörter, die nicht länger als 3 sind.
- 2 die Sprache aller Wörter, die mit 'ab' beginnen.
- 3 die Sprache aller Wörter, in denen die Kette 'aa' vorkommt.
- 4 die Sprache aller Wörter, die ungleich der Kette 'abb' sind.
- 5 die Sprache aller Wörter, die auf die Kette 'aa' enden.
- 6 die Sprache aller Wörter, in denen eine gerade Zahl von a 's vorkommt.
- 7 die Sprache aller Wörter, in denen mindestens zwei a 's vorkommen.

reguläre Sprache

Gegeben ein Alphabet Σ .

- \emptyset ist eine reguläre Sprache über dem Alphabet Σ .
- $\{e\}$ ist eine reguläre Sprache über dem Alphabet Σ .
- Für jedes $a \in \Sigma$ ist $\{a\}$ eine reguläre Sprache über dem Alphabet Σ .
- Wenn A und B reguläre Sprachen sind, dann ist auch $A \cup B$ eine reguläre Sprache über dem Alphabet Σ .
- Wenn A und B reguläre Sprachen sind, dann ist auch $A \circ B$ eine reguläre Sprache über dem Alphabet Σ .
- Wenn A eine reguläre Sprachen ist, dann ist auch A^* eine reguläre Sprache über dem Alphabet Σ .
- Nichts sonst ist eine reguläre Sprache über dem Alphabet Σ .

Satz von Kleene



(Stephen C. Kleene, 1909 - 1994)

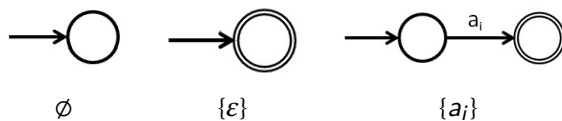
Jede Sprache, die von einem endlichen Automaten akzeptiert wird, ist regulär und jede reguläre Sprache wird von einem endlichen Automaten akzeptiert.

Endliche Automaten akzeptieren reguläre Sprachen

Theorem (Kleene)

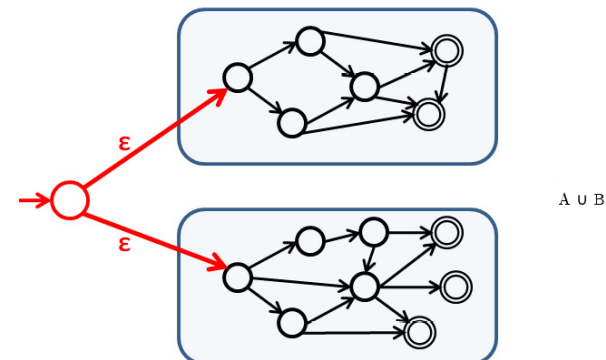
Jede Sprache, die von einem endlichen Automaten akzeptiert wird ist regulär und jede reguläre Sprache wird von einem endlichen Automaten akzeptiert.

Beweisidee (eine Richtung): Zu jeder regulären Sprache gibt es einen endlichen Automaten, der diese akzeptiert:



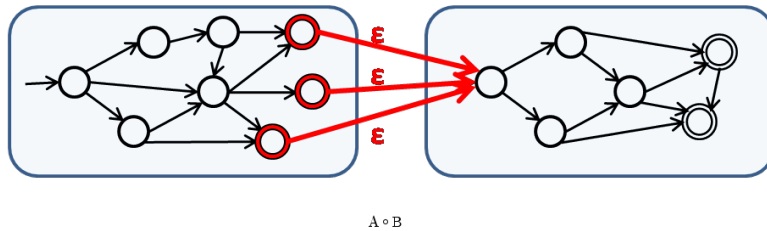
Beweis des Satzes von Kleene (Fortsetzung)

Wenn A und B zwei reguläre Sprachen sind, die von den Automaten A_A und A_B akzeptiert werden, dann wird die reguläre Sprache $A \cup B$ von dem folgenden Automaten akzeptiert:



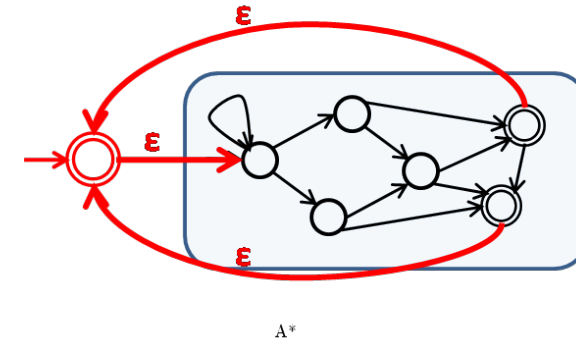
Beweis des Satzes von Kleene (Fortsetzung)

Wenn A und B zwei reguläre Sprachen sind, die von den Automaten A_A und A_B akzeptiert werden, dann wird die reguläre Sprache $A \circ B$ von dem folgenden Automaten akzeptiert:



Beweis des Satzes von Kleene (Fortsetzung)

Wenn A eine reguläre Sprache ist, die von dem Automaten A_A akzeptiert wird, dann wird die reguläre Sprache A^* von dem folgenden Automaten akzeptiert:



Formale Grammatik

Definition

Eine **formale Grammatik** ist ein 4-Tupel $G = (N, T, S, P)$ aus

- einem Alphabet von Terminalsymbolen T (häufig auch Σ)
- einem Alphabet von Nichtterminalsymbolen N mit $N \cap T = \emptyset$
- einem Startsymbol $S \in N$
- einer Menge von Regeln/Produktionen $P \subseteq \{(\alpha, \beta) \mid \alpha, \beta \in (N \cup T)^* \text{ und } \alpha \notin T^*\}$.

Für eine Regel (α, β) schreiben wir auch $\alpha \rightarrow \beta$.

Formale Grammatiken werden auch **Typ0-** oder **allgemeine Regelgrammatiken** genannt.

$$\begin{array}{l} S \rightarrow NP VP \quad VP \rightarrow V \quad NP \rightarrow D N \\ D \rightarrow the \quad N \rightarrow cat \quad V \rightarrow sleeps \end{array}$$

Generiert: the cat sleeps

Terminologie

$G = (\{S, NP, VP, N, V, D, EN\}, \{the, cat, peter, chases\}, S, P)$

$$P = \left\{ \begin{array}{lll} S \rightarrow NP VP & VP \rightarrow V NP & NP \rightarrow D N \\ NP \rightarrow EN & D \rightarrow the & N \rightarrow cat \\ EN \rightarrow peter & V \rightarrow chases & \end{array} \right\}$$

“NP VP” ist **in einem Schritt ableitbar** aus S

“the cat chases peter” ist **ableitbar** aus S :

$$\begin{array}{lll} S \rightarrow NP VP & \rightarrow NP V NP & \rightarrow NP V EN \\ \rightarrow NP V peter & \rightarrow NP chases peter & \rightarrow D N chases peter \\ \rightarrow D cat chases peter & \rightarrow the cat chases peter & \end{array}$$

Die Menge aller aus dem Startsymbol S ableitbarer Wörter (= Ketten aus Terminalsymbolen) ist die von der Grammatik G **erzeugte Sprache** $L(G)$.

$$L(G) = \left\{ \begin{array}{ll} the\ cat\ chases\ peter & peter\ chases\ the\ cat \\ peter\ chases\ peter & the\ cat\ chases\ the\ cat \end{array} \right\}$$

Hinweis: für Terminalsymbole verwendet man in der Regel Klein- und für Nichtterminalsymbole Großbuchstaben.

kontextfreie Grammatiken

Eine formale Grammatik in der jede linke Regelseite aus genau einem Nichtterminalsymbol besteht heißt **kontextfrei**.

Beispiel:

$G = (\{S, NP, VP, N, V, D, EN\}, \{the, cat, peter, chases\}, S, P)$

$$P = \left\{ \begin{array}{l} S \rightarrow NP VP \quad VP \rightarrow V NP \quad NP \rightarrow D N \\ NP \rightarrow EN \quad D \rightarrow the \quad N \rightarrow cat \\ EN \rightarrow peter \quad V \rightarrow chases \end{array} \right\}$$

Chomskyhierarchie

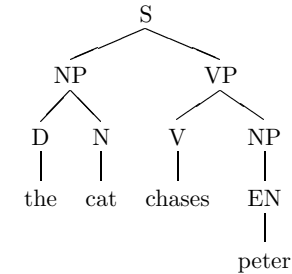
- Wenn man die Form der Regeln einschränkt erhält man Teilmengen der Menge aller durch eine Grammatik erzeugten Sprachen.
- Die Chomskyhierarchie ist eine Hierarchie über die Regelbedingungen (den verschiedenen Sprachklassen entsprechen Einschränkungen über die rechten und linken Regelseiten).
- Die Chomskyhierarchie reflektiert eine spezielle Form der Komplexität, andere Kriterien sind denkbar und führen zu anderen Hierarchien.
- Die Sprachklassen der Chomskyhierarchie sind in der Informatik intensiv untersucht worden (Berechnungskomplexität, effektive Parser).
- Für Linguisten ist die Chomsky Hierarchie besonders interessant, da sie die Form der Regeln zentral stellt, und somit Aussagen über Grammatikformalismen zuläßt.

Linksableitung (kontextfreie Grammatiken)

Gegeben eine kontextfreie Grammatik G . Eine Ableitung bei der stets das am weitesten links stehende nichtterminale Symbol ersetzt wird, heißt

Linksableitung

$S \rightarrow NP VP \quad \rightarrow D N VP \quad \rightarrow the N VP$
 $\rightarrow the cat VP \quad \rightarrow the cat V NP \quad \rightarrow the cat chases NP$
 $\rightarrow the cat chases EN \quad \rightarrow the cat chases peter$



Zu jeder Linksableitung gibt es genau einen **Ableitungsbaum** und zu jedem Ableitungsbaum gibt es genau eine Linksableitung.

Noam Chomsky



Noam Chomsky

(* 7.12.1928, Philadelphia)

Noam Chomsky, *Three Models for the Description of Language*, IRE Transactions on Information Theory (1956).

| Sprache | Automat | Grammatik | Erkennung | Abhängigkeit |
|------------------------|-------------------------------|---------------------------------------------------------------------------|--------------------|------------------|
| rekursiv aufzählbar | Turing Maschine | unbeschränkt $Baa \rightarrow \epsilon$ | unentscheidbar | beliebig |
| kontext- sensitiv | linear gebunden | kontext- sensitiv $\gamma A \delta \rightarrow \gamma \beta \delta$ | NP-vollständig | überkreuzt |
| kontext- frei | Kellerautomat (Stapel) | kontextfrei $C \rightarrow bABa$ | polynomiell | eingebettet |
| regulär | endlicher Automat | regulär $A \rightarrow bA$ | linear | strikt lokal |

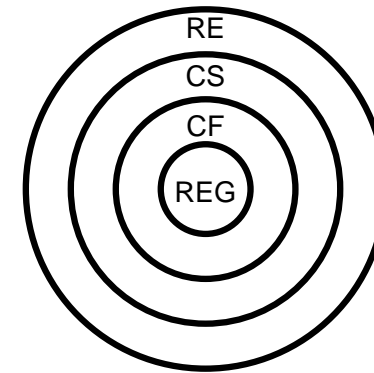
Mathematische Grundlagen der Computerlinguistik

Ordnungsrelationen

Dozentin: Wiebke Petersen

4. Foliensatz

$$\text{REG} \subset \text{CF} \subset \text{CS} \subset \text{RE}$$



REG: reguläre Sprachen, CF: kontextfreie Sprachen, CS: kontextsensitive Sprachen, RE: rekursiv-aufzählbare Sprachen

starke / schwache Ordnungen

Eine **Ordnung** R einer Menge A ist eine binäre Relation $R \subseteq A \times A$.
Man unterscheidet zwischen **starken** und **schwachen** Ordnungen:

Eine binäre Relation ist eine **schwache Ordnung**, gdw. sie

- transitiv,
- reflexiv und
- anti-symmetrisch

ist.

Eine binäre Relation ist eine **starke Ordnung**, gdw. sie

- transitiv,
- irreflexiv und
- asymmetrisch

ist.

Starke Ordnungen werden auch **strikte** Ordnungen genannt.

korrespondierende Ordnungen

Eine schwache Ordnung $R \subseteq A \times A$ und eine starke Ordnung S **korrespondieren** zueinander gdw.

$$R = S \cup id_A$$

Beispiele: Sei $A = \{a, b, c, d\}$

- $R_1 = \{\langle a, b \rangle, \langle a, c \rangle, \langle a, d \rangle, \langle b, c \rangle, \langle a, a \rangle, \langle b, b \rangle, \langle c, c \rangle, \langle d, d \rangle\}$
- $R_2 = \{\langle b, a \rangle, \langle c, b \rangle, \langle c, a \rangle, \langle b, b \rangle, \langle a, a \rangle, \langle c, c \rangle, \langle d, d \rangle\}$
- $R_3 = \{\langle d, c \rangle, \langle d, b \rangle, \langle d, a \rangle, \langle c, b \rangle, \langle c, a \rangle, \langle b, a \rangle, \langle a, a \rangle, \langle b, b \rangle, \langle c, c \rangle, \langle d, d \rangle\}$

korrespondierende starke Ordnungen:

- $S_1 = \{\langle a, b \rangle, \langle a, c \rangle, \langle a, d \rangle, \langle b, c \rangle\}$
- $S_2 = \{\langle b, a \rangle, \langle c, b \rangle, \langle c, a \rangle\}$
- $S_3 = \{\langle d, c \rangle, \langle d, b \rangle, \langle d, a \rangle, \langle c, b \rangle, \langle c, a \rangle, \langle b, a \rangle\}$

geordnete Mengen

Eine **geordnete Menge** ist ein Paar (M, R) , bestehend aus einer Menge M und einer Ordnung R von M .

Beispiele:

- $(\mathcal{POT}(M), \subseteq)$ ist eine schwach geordnete Menge.
 $(\mathcal{POT}(M), \subset)$ ist die korrespondierende stark geordnete Menge.
- (\mathbb{N}, \leq) ist eine schwach geordnete Menge.
 $(\mathbb{N}, <)$ ist die korrespondierende stark geordnete Menge.

Terminologie

Sei (M, R) eine (stark oder schwach) geordnete Menge.

- a ist ein **Vorgänger** von b gdw. $R(a, b)$.
- a ist ein **Nachfolger** von b gdw. $R(b, a)$.
- a ist ein **unmittelbarer Vorgänger** (oder **unterer Nachbar**) von b gdw.
 - $a \neq b$,
 - $R(a, b)$, und
 - es gibt kein $c \in M$ mit $c \notin \{a, b\}$ so dass $R(a, c)$ und $R(c, b)$.
- a ist ein **unmittelbarer Nachfolger** (oder **oberer Nachbar**) von b gdw. b ein unmittelbarer Vorgänger von a ist.

Wenn a ein unmittelbarer Vorgänger von b ist, dann schreibt man häufig $a < b$.

Hassediagramm

Konstruktion

Eine endliche geordnete Mengen (M, R) kann durch ein **Hassediagramm** veranschaulicht werden; dieses erhält man, indem man für jedes Element von M einen Punkt zeichnet und zwar so, daß a unterhalb von b liegt, wenn $a \neq b$ und $(a, b) \in R$.

Zwei Punkte a und b werden mit einer Linie verbunden, wenn $a < b$.

Übung: Zeichnen sie die folgenden Hasse-Diagramme

Hasse-Diagramm von $(\{a, b, c\}, R_2)$ mit

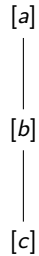
$$R_2 = \{\langle b, a \rangle, \langle c, b \rangle, \langle c, a \rangle, \langle b, b \rangle, \langle a, a \rangle, \langle c, c \rangle\}$$

Hasse-Diagramm von

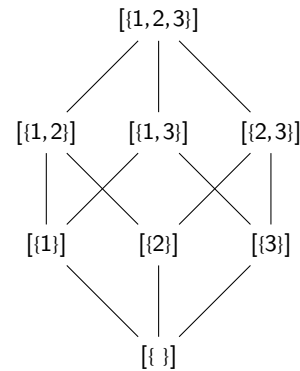
$$(\mathcal{POT}(\{1, 2, 3\}), \subseteq)$$

Beispiele

Hasse-Diagramm von $(\{a, b, c\}, R_2)$ mit
 $R_2 = \{\langle b, a \rangle, \langle c, b \rangle, \langle c, a \rangle, \langle b, b \rangle, \langle a, a \rangle, \langle c, c \rangle\}$



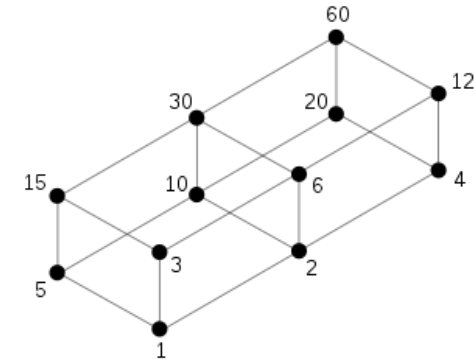
Hasse-Diagramm von
 $(\mathcal{P}\mathcal{O}\mathcal{T}(\{1, 2, 3\}), \subseteq)$



Hasse-Diagramme: Beispiel Teilbarkeit

Sei $M = \{x \in \mathbb{N} \mid 60 \text{ ist durch } x \text{ ohne Rest teilbar}\}$, und
 $R = \{\langle x, y \rangle \in M \times M \mid y \text{ ist durch } x \text{ ohne Rest teilbar}\}$.

Hassediagramm der geordneten Menge (M, R) :



Übung

Zeichnen sie ein Hasse-Diagramm zur geordneten Menge
 $M = (\{\{1, 2, 3, 4, 5\}, \{1, 2, 3, 5\}, \{1, 3, 4\}, \{2, 4, 5\}, \{1, 2, 3\}, \{1, 3\}, \{2, 4\}, \{1, 5\}, \{1\}, \{3\}, \{4\}, \{5\}, \emptyset\}, \subseteq)$.

totale/partielle Ordnung

Eine binäre Ordnungsrelation ist eine **totale** Ordnung, gdw. sie **konnex** ist.

Eine binäre Relation $R \subseteq M \times M$ ist **konnex** (bzw. **linear**) gdw. für alle $x, y \in M$ mit $x \neq y$ gilt: $\langle x, y \rangle \in R$ oder $\langle y, x \rangle \in R$ (oder beides).

- Das Hassediagramm einer total geordneten, endlichen Menge bildet eine Linie. Kein Element hat mehr als einen oberen oder unteren Nachbarn.
- Totale Ordnungen werden auch **lineare** Ordnungen genannt.
- In Abgrenzung zu totalen Ordnungen werden allgemeine Ordnungen auch **partielle** Ordnungen (oder **Halbordnungen**) genannt. Im Englischen spricht man von **'poset'** (partially ordered set).

minimale und maximale Elemente

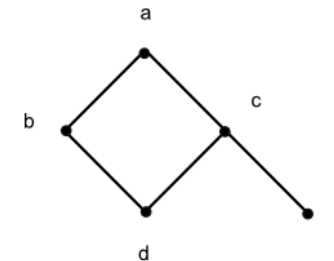
Sei $R \subseteq A \times A$ eine Ordnung (stark oder schwach).

- Ein Element $x \in A$ ist **minimal** gdw. es kein $y \neq x$ gibt, das Vorgänger von x ist.
- Ein Element $x \in A$ ist **maximal** gdw. es kein $y \neq x$ gibt, das Nachfolger von x ist.
- $x \in A$ ist das **Minimum** von A , wenn x Vorgänger jedes anderen Elements von A ist (für alle $y \in A$ mit $x \neq y$ gilt $x < y$).
- $x \in A$ ist das **Maximum** von A , wenn x Nachfolger jedes anderen Elements von A ist (für alle $y \in A$ mit $x \neq y$ gilt $y < x$).

Hinweise:

- eine total geordnete Menge kann höchstens ein minimales und höchstens ein maximales Element haben.
- eine partiell geordnete Menge kann beliebig viele minimale und maximale Elemente aber höchstens ein Minimum und höchstens ein Maximum haben.

Beispiel



- a ist das einzige maximale Element und somit das Maximum der geordneten Menge.
- d und e sind die minimalen Elemente der geordneten Menge.
- die geordnete Menge hat kein Minimum,

Vergleichbarkeit / Kette / Antikette

Sei (M, R) eine geordnete Menge und seien a und b Elemente von M . a und b heißen **vergleichbar**, falls aRb oder bRa ; sonst **unvergleichbar**. Eine Teilmenge K von M heißt **Kette**, g.d.w. für beliebige $a, b \in K$ gilt, daß sie vergleichbar sind. Eine Teilmenge A von M heißt **Antikette**, g.d.w. für beliebige $a, b \in A$ gilt, daß sie unvergleichbar sind.

Satz von Dilworth

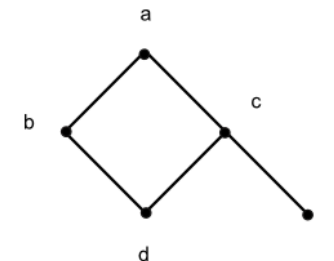
Für eine geordnete endliche Menge (M, R) gilt: Die maximale Anzahl von Elementen in einer Antikette von (M, R) ist gleich der kleinsten Anzahl von Ketten von (M, R) , die man für eine Partition von M benötigt.

Höhe / Breite

Die **Höhe** einer endlichen geordneten Menge (M, R) ist gleich der maximalen Anzahl von Elementen einer Kette von (M, R) .

Die **Breite** einer endlichen geordneten Menge (M, R) ist gleich der maximalen Anzahl von Elementen einer Antikette von (M, R) .

Beispiel



- Die Elemente a und b sind vergleichbar.
- d und e sind unvergleichbar.
- $\{a, b, d\}$ ist eine Kette der geordneten Menge.
- $\{b, c\}$ ist Antikette der geordneten Menge.
- Die geordnete Menge hat die Höhe 3 und die Breite 2.
- Die Ketten $\{a, b, d\}$ und $\{c, e\}$ bilden eine minimale Partition in Ketten der geordneten Menge.

Intervall / Ideal / Filter

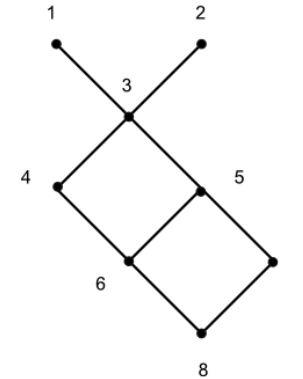
Sei (M, \leq) eine geordnete Menge:

Intervall: $[a, b] := \{x \in M \mid a \leq x \leq b\}$

Hauptideal: $(b) := \{x \in M \mid x \leq b\}$

Hauptfilter: $[a) := \{x \in M \mid a \leq x\}$

Beispiel



- $[6, 1] = \{1, 3, 4, 5, 6\}$ (Intervall von 6 bis 1)
- $(4) = \{4, 6, 8\}$ (Hauptideal von 4)
- $[6) = \{1, 2, 3, 4, 5, 6\}$ (Hauptfilter von 6).

Ordnungserhaltende/monotone Abbildungen

Definition

Seien (M, \leq) und (M', \leq') zwei geordnete Mengen. Eine Abbildung (Funktion) $f : M \rightarrow M'$ heißt **ordnungserhaltend** oder **monoton**, wenn für alle $x, y \in M$ gilt:

$$\text{wenn } x \leq y, \text{ dann } f(x) \leq' f(y)$$

Beispiele:

- $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ mit $f(x) = 2x$ ist eine monotone Abbildung von (\mathbb{N}_0, \leq) nach (\mathbb{N}_0, \leq) .
- $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ mit $f(x) = x^2$ ist eine monotone Abbildung von (\mathbb{N}_0, \leq) nach (\mathbb{N}_0, \leq) .
- $f : \mathbb{Z} \rightarrow \mathbb{Z}$ mit $f(x) = x^2$ ist keine monotone Abbildung von (\mathbb{Z}, \leq) nach (\mathbb{Z}, \leq) .
- Sei M eine endliche Menge. $f : \mathcal{P}(\mathcal{T}(M)) \rightarrow \mathbb{N}_0$ mit $f(A) = |A|$ ist eine ordnungserhaltende Abbildung von $(\mathcal{P}(\mathcal{T}(M)), \subseteq)$ nach (\mathbb{N}_0, \leq) .

Ordnungseinbettung

Definition

Eine monotone Funktion heißt **Ordnungseinbettung**, wenn sie injektiv ist, und **Ordnungsisomorphismus**, wenn sie bijektiv ist.

Ein Ordnungsisomorphismus von (M, R) in sich selbst wird auch **Ordnungsautomorphismus** genannt.

Beispiele:

- $f : \mathbb{Z} \rightarrow \mathbb{Z}$ mit $f(x) = -x$ ist ein Ordnungsisomorphismus von (\mathbb{Z}, \leq) nach (\mathbb{Z}, \geq) .
- $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = \frac{x}{2}$ ist ein Ordnungsautomorphismus auf (\mathbb{R}, \leq) .

Quasiordnung

Der Begriff der Quasiordnung ist schwächer als der der Ordnung:

Definition

Eine binäre Relation $R \subseteq M \times M$ ist eine **Quasiordnung** (oder **Präordnung**), wenn R

- reflexiv und
- transitiv ist.

Beispiel:

- Die Ordnung \leq_{abs} , die die ganzen Zahlen nach ihrem Betrag ordnet ist eine Quasiordnung aber keine Ordnung (beachte, dass $-3 \leq_{abs} 3$ und $3 \leq_{abs} -3$ aber $-3 \neq 3$).

Zusammenfassung: Ordnungen

schwache Ordnungen

| | transitiv | reflexiv | anti-symmetrisch | linear/total |
|-------------------|-----------|----------|------------------|--------------|
| Quasiordnung | ★ | ★ | | |
| partielle Ordnung | ★ | ★ | ★ | |
| totale Ordnung | ★ | ★ | ★ | ★ |

Bemerkung: (Schwache) lineare Ordnungsrelationen werden häufig mit \leq , bzw. partielle Ordnungsrelationen mit \subseteq bezeichnet, auch wenn es sich bei der gegebenen Ordnung weder um eine numerische Größenordnung noch um die Mengeninklusion handelt.

Zusammenfassung: Ordnungen

strikte Ordnungen

| | transitiv | irreflexiv | asymmetrisch | linear/total |
|---------------------------|-----------|------------|--------------|--------------|
| strikte partielle Ordnung | ★ | ★ | ★ | |
| strikte totale Ordnung | ★ | ★ | ★ | ★ |

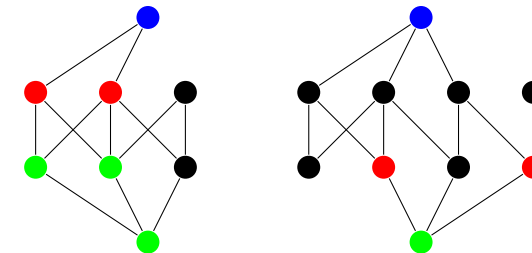
Bemerkung: Strikte Ordnungsrelationen werden häufig mit $<$, bzw. mit \subset bezeichnet.

Man könnte strikte Ordnungen äquivalent auch als transitive, irreflexive und antisymmetrische Relationen definieren, da eine Relation, die irreflexiv und antisymmetrisch ist, immer asymmetrisch ist.

obere / untere Schranke

Sei (M, \leq) eine (partiell) geordnete Menge und K eine Teilmenge von M . Ein Element x von M ist

- eine **obere Schranke** von K , g.d.w. für alle $y \in K : y \leq x$;
- eine **untere Schranke** von K , g.d.w. für alle $y \in K : x \leq y$.



Die Abbildungen zeigen die Hasse diagramme zweier geordneter Mengen. Die rot markierten Elementen haben die blau markierten Elemente als obere und die grün markierten als untere Schranken.

kleinste obere / größte untere Schranke

x heißt **kleinste obere Schranke** oder **Supremum** von K in M , wenn x eine obere Schranke von K ist und für jede obere Schranke $y \in M$ von K mit $x \neq y$ die Ungleichung $x \leq y$ gilt. Wir schreiben $\sup K$ oder $\bigvee K$ für das Supremum von K (lese \vee als 'join').

x heißt **größte untere Schranke** oder **Infimum** von K in M , wenn x eine untere Schranke von K ist und für jede untere Schranke $y \in M$ von K mit $x \neq y$ die Ungleichung $y \leq x$ gilt. Wir schreiben $\inf K$ oder $\bigwedge K$ für das Infimum von K (lese \wedge als 'meet').

Wir schreiben $x \vee y$ statt $\bigvee\{x, y\}$ und $x \wedge y$ statt $\bigwedge\{x, y\}$.

Die Beispiele der vorangegangenen Folie zeigen, daß es geordnete Mengen M gibt, für die nicht jede Teilmenge $K \subseteq M$ ein Supremum oder Infimum hat.

Das Infimum ist also das Maximum aller unteren Schranken und das Supremum ist das Minimum aller oberen Schranken.

Beispiele

- Für die linear geordnete Menge (\mathbb{R}, \leq) gilt: $\sup[1, 4] = 4$ und $\inf[1, 4] = 1$.
- Für die partiell geordnete Menge $(\mathcal{POT}(M), \subseteq)$ mit $M = \{1, 2, 3, 4\}$ ist das Supremum von $K = \{\{1, 2\}, \{2, 4\}, \{1\}\}$ die Vereinigung aller Elemente von K , also $\sup K = \{1, 2, 4\}$. Das Infimum von K ist der Durchschnitt aller Elemente von K , also $\inf K = \emptyset$.

Verbände

Verband: ordnungstheoretische Definition

Eine geordnete Menge (V, \leq) ist ein **Verband**, g.d.w. zu je zwei Elementen x und y aus V auch das Supremum von x und y und das Infimum von x und y Elemente von V sind.

vollständiger Verband

Ein Verband (V, \leq) ist ein **vollständiger Verband**, falls für alle $K \subseteq V$ gilt, daß $\sup K \in V$ und $\inf K \in V$.

Jeder vollständige Verband hat ein größtes Element $\sup V$, das **Einselement** (1_V) genannt, und ein kleinstes Element $\inf V$, das **Nullelement** (0_V) genannt.

Die oberen Nachbarn des Nullelements nennt man die **Atome** und die unteren Nachbarn des Einselements die **Koatome** des Verbands.

Bemerkungen

- Jeder endliche Verband ist vollständig.
- Da $\inf \emptyset = 1_V$ und $\sup \emptyset = 0_V$ gilt, gibt es keinen vollständigen Verband mit leerer Menge V .

- $(\mathcal{POT}(M), \subseteq)$ ist ein vollständiger Verband, \vee entspricht \cup und \wedge entspricht \cap .
- $([2, 5], \leq)$ ist ein vollständiger Verband.
- (\mathbb{R}, \leq) ist ein Verband, aber nicht vollständig.
- $(\{\{1, 2\}, \{2, 4\}, \{1\}\}, \subseteq)$ ist kein Verband.

Algebren (algebraische Strukturen)

Eine **Algebra** \mathbf{A} ist eine Menge A zusammen mit einer oder mehreren n -stelligen **Operationen** (Verknüpfungen) f_i .

In diesem Kurs beschränken wir uns auf Algebren mit ein oder zwei binären Operationen.

Die Operationen einer Algebra müssen die folgenden Axiome erfüllen:

Abgeschlossenheit: A ist unter der Operation \otimes abgeschlossen, d.h. für beliebige $a, b \in A$ gibt es ein Element $c \in A$, sodass $a \otimes b = c$.

Eindeutigkeit: Wenn $a = a'$ und $b = b'$, dann gilt $a \otimes b = a' \otimes b'$.

An was erinnern Sie die beiden Axiome?

Alternative Definition: Eine Algebra \mathbf{A} ist eine Menge A zusammen mit einer oder mehreren n -stelligen Funktionen $f_i: A^n \rightarrow A$.

Eigenschaften von Operationen

Assoziativgesetz

Eine Operation \otimes auf A ist **assoziativ**, g.d.w. für alle $a, b, c \in A$ gilt:

$$a \otimes (b \otimes c) = (a \otimes b) \otimes c$$

Kommutativgesetz

Eine Operation \otimes auf A ist **kommutativ**, g.d.w. für alle $a, b \in A$ gilt:

$$a \otimes b = b \otimes a$$

Idempotenzgesetz

Eine Operation \otimes auf A ist **idempotent**, g.d.w. für alle $a \in A$ gilt:

$$a \otimes a = a$$

Distributivgesetz

Für zwei Operationen \oplus und \otimes auf A ist **distributiert** \oplus über \otimes , g.d.w. für alle $a, b, c \in A$ gilt:

$$a \oplus (b \otimes c) = (a \oplus b) \otimes (a \oplus c)$$

neutrale und inverse Elemente

neutrales Element

Gegeben eine Operation \oplus auf A . Ein Element $e \in A$ ist das **neutrale Element** von \oplus , g.d.w. für alle a in A gilt:

$$e \oplus a = a \oplus e = a.$$

inverses Element

Gegeben eine Operation \oplus auf A mit neutralem Element e . Ein Element $a^{-1} \in A$ ist das **inverse Element** eines Elements $a \in A$, g.d.w.:

$$a^{-1} \oplus a = a \oplus a^{-1} = e$$

Bsp: Drehungen eines gleichseitigen Dreiecks $((\Delta_D, \circ))$

Grundmenge: $\{0, \curvearrowright, \curvearrowleft\}$

(id: 0°-Drehung; \curvearrowright : 120°-Drehung nach rechts; \curvearrowleft : 120°-Drehung nach links)

Operation: \circ Hintereinanderausführung der Drehungen.

| \circ | id | \curvearrowright | \curvearrowleft |
|--------------------|--------------------|--------------------|--------------------|
| id | id | \curvearrowright | \curvearrowleft |
| \curvearrowright | \curvearrowright | \curvearrowright | id |
| \curvearrowleft | \curvearrowleft | id | \curvearrowright |

- neutrales Element: id
- inverse Elemente: $\text{id}^{-1} = \text{id}$; $\curvearrowright^{-1} = \curvearrowleft$; $\curvearrowleft^{-1} = \curvearrowright$
- Eigenschaften von \circ : assoziativ, kommutativ

Bsp: Drehungen und horizontale Spiegelung eines gleichseitigen Dreiecks

Grundmenge: $\{0, \curvearrowright, \curvearrowleft, \leftrightarrow\}$

(id: 0°-Drehung; \curvearrowright : 120°-Drehung nach rechts; \curvearrowleft : 120°-Drehung nach links, \leftrightarrow : horizontale Spiegelung)

Operation: \circ Hintereinanderausführung der Drehungen und Spiegelungen.

Diese Struktur bildet keine Algebra, da u.a. $\leftrightarrow \circ \curvearrowright$ kein Element der Grundmenge ist (Verletzung der Abgeschlossenheit).

Wenn man alle drei Spiegelungen entlang aller drei Spiegelachsen hinzunimmt, erhält man wieder eine Algebra.

Beispiel: Restklassen modulo 3 $((\mathbb{N} \bmod 3, \oplus_3))$

Grundmenge: $\{[0], [1], [2]\}$

Operation: \oplus_3 : Summe modulo 3

| \oplus_3 | [0] | [1] | [2] |
|------------|-----|-----|-----|
| [0] | [0] | [1] | [2] |
| [1] | [1] | [2] | [0] |
| [2] | [2] | [0] | [1] |

- neutrales Element: [0]
- inverse Elemente: $[0]^{-1} = [0]$; $[1]^{-1} = [2]$; $[2]^{-1} = [1]$
- Eigenschaften von \oplus_3 : assoziativ, kommutativ

- $(\mathbb{N}_0, +)$
- $(\mathbb{Z}, +, \cdot)$
- $(\mathbb{R}, +, \cdot)$
- $(\mathcal{POT}(M), \cap, \cup)$
- (Σ^*, \circ)

Morphismus

Ein **Morphismus** ($\varphi: \mathbf{A} \rightarrow \mathbf{B}$) von einer Algebra \mathbf{A} in eine Algebra \mathbf{B} ist eine Abbildung, die zum einen eine Funktion von der Menge der ersten in die Menge der zweiten Algebra definiert ($F: A \rightarrow B$), und zum anderen die Operationen der ersten Algebra auf die zweite Algebra projiziert (hierzu müssen beide Algebren gleichviele Operationen gleicher Stelligkeit haben).

Homomorphismus

Gegeben zwei Algebren $\mathbf{A} = (A, \oplus, \otimes)$ und $\mathbf{B} = (B, \star, \circ)$. Ein Morphismus $\varphi: \mathbf{A} \rightarrow \mathbf{B}$ ist ein **Homomorphismus**, g.d.w. für alle x, y in A gilt:

$$\begin{aligned}\varphi(x) \star \varphi(y) &= \varphi(x \oplus y) \text{ und} \\ \varphi(x) \circ \varphi(y) &= \varphi(x \otimes y)\end{aligned}$$

Isomorphismus

Gegeben zwei Algebren $\mathbf{A} = (A, \oplus, \otimes)$ und $\mathbf{B} = (B, \star, \circ)$. Ein Morphismus $\varphi: \mathbf{A} \rightarrow \mathbf{B}$ ist ein **Isomorphismus**, g.d.w. $\varphi: A \rightarrow B$ bijektiv ist und wenn für alle x, y in A gilt:

$$\begin{aligned}\varphi(x) \star \varphi(y) &= \varphi(x \oplus y) \text{ und} \\ \varphi(x) \circ \varphi(y) &= \varphi(x \otimes y)\end{aligned}$$

Zwei Algebren sind **isomorph**, wenn es einen Isomorphismus zwischen ihnen gibt.

Automorphismus

Ein Automorphismus einer Algebra \mathbf{A} ist ein Isomorphismus $\varphi: \mathbf{A} \rightarrow \mathbf{A}$.

Beispiele

- $\varphi: (\mathbb{N}_0, +) \rightarrow (\mathbb{N} \bmod 3, \oplus_3)$ mit $\varphi(n) = n \bmod 3$ ist ein Homomorphismus, aber kein Isomorphismus
- $\varphi: (\mathbb{N}_0, +) \rightarrow (\{a\}^*, \circ)$ mit $\varphi(n) = a^n$ ist ein Isomorphismus.
- $\varphi: (\Delta_D, \circ) \rightarrow (\mathbb{N} \bmod 3, \oplus_3)$ mit $\varphi(\text{id}) = [0]$, $\varphi(\curvearrowright) = [1]$, $\varphi(\curvearrowleft) = [0]$ ist ein Isomorphismus.

Semigruppe, Monoid, Gruppe

Semigruppe

Eine **Semigruppe** (Halbgruppe) $\mathbf{G} = (G, \otimes)$ ist eine Algebra, bestehend aus einer Menge G und einer binären Operation \otimes , die folgende Bedingungen erfüllt:

G1 \otimes ist assoziativ

Monoid

Ein **Monoid** $\mathbf{G} = (G, \otimes)$ ist eine Algebra mit:

G1 \otimes ist assoziativ

G2 G enthält ein neutrales Element

Gruppe

Eine **Gruppe** $\mathbf{G} = (G, \otimes)$ ist eine Algebra mit:

G1 \otimes ist assoziativ

G2 G enthält ein neutrales Element

G3 jedes Element aus G hat ein inverses Element in G .

Beispiele

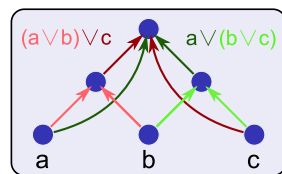
- $(\mathbb{N}, +)$ ist eine Semigruppe
- $(\mathbb{N}_0, +)$ ist ein Monoid
- $(\mathbb{Z}, +)$ ist eine Gruppe
- $(\mathcal{POT}(M), \cup)$ ist ein Monoid
- (Σ^*, \circ) ist ein Monoid
- $(\mathbb{N} \text{ mod } 3, \oplus_3)$ ist eine Gruppe
- (Δ_D, \circ) ist eine Gruppe

Verbände

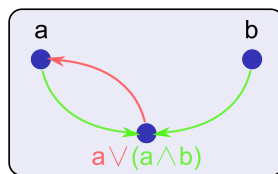
Verband: algebraische Definition

Ein **Verband** $\mathbf{V} = (V, \vee, \wedge)$ ist eine Algebra, bestehend aus einer Menge V und zwei binären Operationen \vee und \wedge , die folgende Bedingungen erfüllen:

- Kommutativgesetz: $a \vee b = b \vee a$ und $a \wedge b = b \wedge a$
- Assoziativgesetz: $a \vee (b \vee c) = (a \vee b) \vee c$ und $a \wedge (b \wedge c) = (a \wedge b) \wedge c$
- Idempotenzgesetz: $a \vee a = a$ und $a \wedge a = a$
- Absorptionsgesetz: $a \vee (a \wedge b) = a$ und $a \wedge (a \vee b) = a$



$$a \vee (b \vee c) = (a \vee b) \vee c$$



$$a \vee (a \wedge b) = a$$

Verbände

Zusammenhang algebraischer und ordnungstheoretischer Verband

- $\mathbf{V} = (V, \leq)$ sei ein (ordnungstheoretisch definierter) Verband. Setze $a \wedge b = \inf\{a, b\}$ und $a \vee b = \sup\{a, b\}$. Dann ist $\mathbf{V} = (V, \vee, \wedge)$ ein (algebraisch definierter) Verband.
- $\mathbf{V} = (V, \vee, \wedge)$ sei ein (algebraisch definierter) Verband. Setze $a \leq b$ g.d.w. $a \wedge b = a$. Dann ist $\mathbf{V} = (V, \leq)$ ein (ordnungstheoretisch definierter) Verband.

Beispiel: $(\mathcal{POT}(M), \subseteq)$ und $(\mathcal{POT}(M), \cup, \cap)$

Mathematische Grundlagen der Computerlinguistik

Bäume

Dozentin: Wiebke Petersen

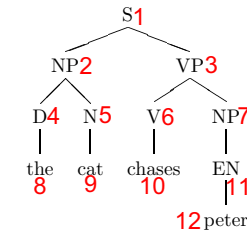
6. Foliensatz
(basierend auf Folien von Gerhard Jäger)

Bäume

Baumdiagramme

Ein Baumdiagramm eines Satzes stellt drei Arten von Information dar:

- die Konstituenten-Struktur des Satzes,
- die grammatische Kategorie jeder Konstituente, sowie
- die lineare Anordnung der Konstituenten.



Bäume

Konventionen

- Ein Baum besteht aus **Knoten**, die durch
- **Kanten** verbunden werden.
- Kanten sind implizit von oben nach unten **gerichtet** (ähnlich zu Hasse-Diagrammen, wo die implizite Richtung aber von unten nach oben ist.)
- Jeder Knoten ist mit einem **Etikett** (engl. **label**) versehen.

Bäume

Dominanz

- Ein Knoten x **dominiert** Knoten y wenn es eine zusammenhängende (möglicherweise leere) Sequenz von abwärts gerichteten Ästen gibt, die mit x beginnt und mit y endet.
- Für einen Baum T bildet

$$D_T = \{(x,y) | x \text{ dominiert } y \text{ in } T\}$$

die zugehörige **Dominanz-Relation**

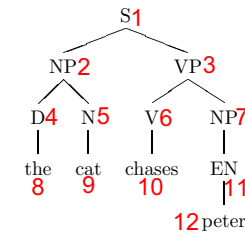
- D_T ist eine schwache Ordnung, also reflexiv, transitiv und anti-symmetrisch.

Bäume

Konventionen

- Wenn x nach D_T der unmittelbare Vorgänger von y ist, dann **dominiert** x y **unmittelbar**.
- Der unmittelbare Vorgänger von x bzgl. D_T heißt der **Mutterknoten** von x .
- Die unmittelbaren Nachfolger von x heißen **Tochterknoten** von x .
- Wenn zwei Knoten nicht identisch sind, aber den selben Mutterknoten haben, heißen sie **Schwesterknoten**.
- Jeder Baum hat endlich viele Knoten.
- Jeder Baum hat ein Infimum bezüglich der Ordnung D_T . Das Infimum heißt **Wurzel** oder **Wurzelknoten** des Baums. Vorsicht: Die Baumdiagramme sind auf den Kopf gestellte Hasse-Diagramme (die Wurzel ist der oberste Knoten des Baumdiagramms, also der Knoten, der als einziges keinen Mutterknoten hat)
- Die maximalen Elemente eines Baumes heißen **Blätter** (Blätter stehen in einem Baumdiagramm ganz unten. Blätter sind diejenigen Knoten, die keine Töchter haben).

Beispiel



- Knoten 2 dominiert Knoten 8 ($\langle 2, 8 \rangle \in D_T$)
- Knoten 2 dominiert Knoten 5 unmittelbar
- Knoten 2 dominiert Knoten 2
- Knoten 2 ist der Mutterknoten von Knoten 5
- Knoten 4 und Knoten 5 sind Schwesterknoten
- Knoten 1 ist der Wurzelknoten des Baums
- Knoten 10 ist ein Blatt des Baums

Bäume

Präzedenz

- Baum-Diagramme beinhalten (anders als Hasse-Diagramm) Informationen über die lineare Abfolge der Knoten.
- Knoten x **geht** Knoten y **voran** (engl. x *precedes* y) g.d.w. x links von y steht und keiner der beiden Knoten den anderen dominiert.
- Für einen Baum T bildet

$$P_T = \{\langle x, y \rangle \mid x \text{ geht } y \text{ voran}\}$$

die zugehörige **Präzedenz-Relation**.

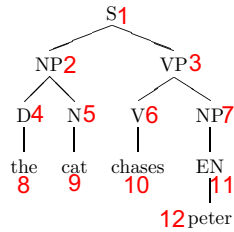
- P_T ist eine starke Ordnung, also irreflexiv, transitiv und asymmetrisch.

Bäume

Exklusivität

In einem Baum T stehen die Knoten x und y in der Präzedenz-Relation (also $P_t(x, y)$ oder $P_t(y, x)$) g.d.w. sie nicht in der Dominanz-Relation stehen (also weder $D_T(x, y)$ noch $D_T(y, x)$).

Beispiel



- Knoten 7 und Knoten 1 stehen in der Dominanz-Relation
- Knoten 7 und Knoten 2 stehen in der Präzedenz-Relation
- Knoten 7 und Knoten 9 stehen in der Präzedenz-Relation
- Knoten 7 und Knoten 12 stehen in der Dominanz-Relation
- Knoten 7 und Knoten 10 stehen in der Präzedenz-Relation

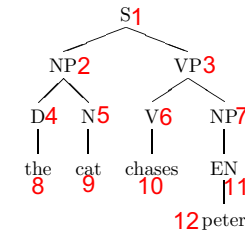
Bäume

Nicht-Überkreuzung

Wenn in einem Baum der Knoten x dem Knoten y vorangeht, dann geht jeder Knoten x' , der von x dominiert wird, jedem Knoten y' voran, der von y dominiert wird.

Diese Bedingung schließt aus, dass

- ein Knoten mehrere Mutterknoten hat, oder dass
- sich Äste überkreuzen.

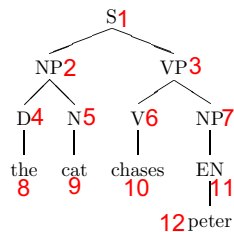


Bäume

Etikettierung

Für jeden Baum T gibt es eine Etikettierungs-Funktion L_T , die jedem Knoten ein Etikett zuweist.

- L_T muss nicht injektiv sein (mehrere Knoten können das selbe Etikett tragen).
- Bei Ableitungsbäumen werden Blätter (auch **Terminal-Knoten** genannt) auf Terminalsymbole abgebildet und alle anderen Knoten auf Nichtterminal-Symbole.



Bäume

Mit Hilfe dieser Eigenschaften von Bäumen können *Theoreme* bewiesen werden, also Sachverhalte, die für alle Bäume gelten. Zum Beispiel

Theorem 1

Wenn x und y Schwesterknoten sind, dann gilt entweder $P_T(x, y)$ oder $P_T(y, x)$.

Theorem 2

Die Menge der Blätter eines Baumes sind durch P_T total geordnet.

- Bäume repräsentieren die relevanten Aspekte einer Ableitung
- Zusammenhang zwischen Ableitung und Baum am einfachsten, wenn alle Regeln der Grammatik die Form

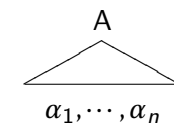
$$A \rightarrow \alpha$$

haben (mit $A \in V_N$ und $\alpha \in (V_T \cup V_N)^*$)

Definition 3

Eine Grammatik $G = (N, T, S, P)$, bei der alle Regeln als linke Seite genau ein Nichtterminal-Symbol haben, **generiert** einen Baum B genau dann wenn

- die Wurzel von B mit S etikettiert ist,
- die Blätter entweder mit Terminalsymbolen oder mit ϵ etikettiert sind, sowie
- es für jeden Teilbaum

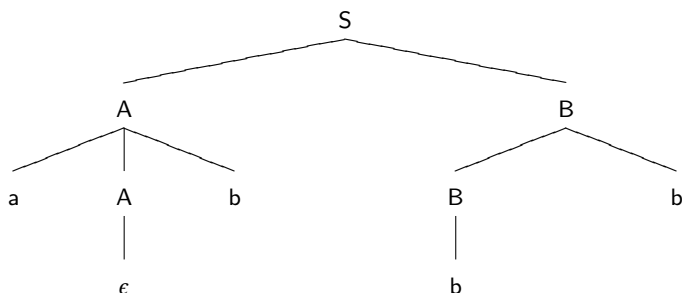


$A \rightarrow \alpha_1, \dots, \alpha_n$ in P gibt.

Beispiel-Grammatik

$$G = (\{S, A, B\}, \{a, b\}, S, P) \quad P = \left\{ \begin{array}{ll} S \rightarrow AB & B \rightarrow Bb \\ A \rightarrow aAb & B \rightarrow b \\ A \rightarrow \epsilon & \end{array} \right\}$$

Diese Grammatik generiert z.B. folgende Baum:



Frage: Welche Sprache wird durch diese Grammatik generiert?

Mathematische Grundlagen der Computerlinguistik

Kombinatorik

Dozentin: Wiebke Petersen

7. Foliensatz

Kombinatorik

- Thema der Kombinatorik ist die Bestimmung der Anzahl möglicher Anordnungen oder Auswahlen.
- Typische kombinatorische Aufgaben sind Urnenaufgaben: Wieviele Möglichkeiten gibt es k Kugeln aus einer Urne mit n Kugeln zu ziehen?
- Hierbei unterscheidet man
 - ob die gezogenen Kugeln wieder zurückgelegt werden oder nicht, und
 - ob die Reihenfolge, in der die Kugeln gezogen werden, beachtet wird oder nicht.



kombinatorische Grundaufgaben: Beispiele

| | ohne Zurücklegen | mit Zurücklegen |
|--------------------------------|-----------------------|-----------------|
| mit Beachtung der Reihenfolge | 3er-Wette (Rennsport) | Toto |
| ohne Beachtung der Reihenfolge | Lotto / Skat | Eisbecher |

Ziehen mit Zurücklegen mit Beachtung der Reihenfolge

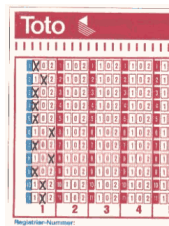
Beispiel: Tippen der Ergebnisse von 11 Fußballspielen (1: Sieg Heimmannschaft, 2: Sieg Gastmannschaft, 0: unentschieden).

$$3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 3^{11} = 177147$$

Es gibt n^k

Möglichkeiten k Objekte aus einer Menge von n Objekten mit Beachtung ihrer Reihenfolge und mit Zurücklegen auszuwählen.

Beispiel: Toto (11er-Wette)



Ziehen ohne Zurücklegen mit Beachtung der Reihenfolge

Spezialfall: alle Kugeln werden gezogen ($n = k$)

Permutationen

n Objekte lassen sich auf $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$ verschiedene Arten in einer Reihe anordnen.

Der Ausdruck $n!$ wird ‚ n Fakultät‘ gelesen.

Als **Permutation** bezeichnet man eine bijektive Abbildung einer endlichen Menge auf sich selbst.

Zu einer n -elementigen Menge gibt es $n!$ Permutationen.

Permutationen sind ein Spezialfall ($k = n$) des ‚Ziehens ohne Zurücklegen mit Beachtung der Reihenfolge‘

Lineare Anordnungsmöglichkeiten für 3 verschiedenfarbige Kugeln:

$$3! = 1 \cdot 2 \cdot 3 = 6$$



Ziehen ohne Zurücklegen mit Beachtung der Reihenfolge

Beispiel: Tippen der ersten 3 Plätze bei einem Pferderennen, wenn 10 Pferde starten.

$$10 \cdot 9 \cdot 8 = \frac{10!}{(10-3)!} = 720$$

Es gibt

$$\frac{n!}{(n-k)!}$$

Möglichkeiten k Objekte aus einer Menge von n Objekten mit Beachtung ihrer Reihenfolge und ohne Zurücklegen auszuwählen.

**Beispiel: 3er-Wette
Pferderennsport**



Ziehen ohne Zurücklegen ohne Beachtung der Reihenfolge

Beispiel: Lottospiel (6 aus 49)

$$\frac{49!}{(49-6)! \cdot 6!} = \frac{49!}{(49-6)! \cdot 6!} = \binom{49}{6} = 13983816$$

Beispiel: Skathände (10 aus 32)

$$\frac{32!}{(32-10)! \cdot 10!} = \frac{32!}{(32-10)! \cdot 10!} = \binom{32}{10} = 64512240$$

Es gibt

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

Möglichkeiten k Objekte aus einer Menge von n Objekten ohne Beachtung ihrer Reihenfolge und ohne Zurücklegen auszuwählen.

Beispiel: Lotto



Beispiel: Skat



Herleitung: ohne Reihenfolge, ohne Zurücklegen

Es gibt

$$\frac{n!}{(n-k)!}$$

Möglichkeiten k Objekte aus einer Menge von n Objekten ohne Zurücklegen aber **mit** Beachtung der Reihenfolge auszuwählen.

Jede k -Auswahl ohne Wiederholungen lässt sich auf $k!$ Arten anordnen.

Folglich gibt es

$$\frac{n!}{(n-k)!} : k! = \frac{n!}{(n-k)!} \cdot \frac{1}{k!} = \frac{n!}{k! \cdot (n-k)!}$$

verschiedene ungeordnete k -Auswahlen aus einer n -Menge ohne Wiederholungen.

Die Zahlen $\frac{n!}{k! \cdot (n-k)!}$ sind die **Binomialkoeffizienten** und werden oft mit $\binom{n}{k}$ bezeichnet (in Worten „ n über k “).

Es gibt

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

Möglichkeiten k Objekte aus einer Menge von n Objekten ohne Beachtung ihrer Reihenfolge und ohne Zurücklegen auszuwählen.

Ziehen mit Zurücklegen ohne Beachtung der Reihenfolge

Beispiel: Eisbecher mit 3 Kugeln aus 10 Eissorten zusammenstellen.

$$\binom{10+3-1}{3} = 220$$

Es gibt

$$\binom{n+k-1}{k}$$

Möglichkeiten k Objekte aus einer Menge von n Objekten ohne Beachtung ihrer Reihenfolge und mit Zurücklegen auszuwählen.

Beispiel: Eisbecher



Herleitung: ohne Reihenfolge, mit Zurücklegen

Beispiel: gemischte Eisbecher mit 3 Kugeln aus 5 Eissorten

| Schoko | Nuß | Orange | Erdbeer | Banane | Becher | Kodierung |
|--------|-----|--------|---------|--------|--------|-----------|
| | •• | | | • | ••• | •• • |
| | | • | • | • | ••• | • • • |
| • | | | •• | | ••• | • •• |
| | | | ••• | | ••• | •• |

Die Kodierung der Eisbecher ist so gewählt, dass sich das Problem der Wahl von k Eiskugeln aus n Eissorten auf das Problem der linearen Anordnung von k ununterscheidbaren Kugeln und $n-1$ ununterscheidbaren Strichen reduziert. Dieses Problem lässt sich als Auswahl von k Positionen (die Kugelpositionen) aus $k+n-1$ Positionen auffassen.

Hierfür gibt es $\binom{k+n-1}{k}$ Möglichkeiten

kombinatorische Grundaufgaben: Zusammenfassung

Anzahl der k -Auswahlen aus einer n -er-Menge:

| | ohne Wiederholungen | mit Wiederholungen |
|--------------------------------|---------------------------------------------|--------------------|
| mit Beachtung der Reihenfolge | $\binom{n}{k} \cdot k! = \frac{n!}{(n-k)!}$ | n^k |
| ohne Beachtung der Reihenfolge | $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$ | $\binom{n+k-1}{k}$ |

Hinweise:

- Bearbeiten Sie bitte das Modul Kombinatorik (Link)
- Berechnung von Binomialkoeffizienten (Link)

Mathematische Grundlagen der Computerlinguistik

Wahrscheinlichkeit

Dozentin: Wiebke Petersen

8. Foliensatz

Motivation

- In vielen Bereichen der CL kommt Wahrscheinlichkeitstheorie zur Anwendung, da es oft unmöglich ist, mit rein symbolischen Ansätzen ein vollständiges Bild aller möglichen Strukturen einschließlich Präferenzen bei Ambiguitäten zu gewinnen.
- Wir haben es meist mit einer endlichen oder abzählbar unendlichen Menge von sogenannten Ergebnissen zu tun, deren Wahrscheinlichkeit irgendwie abgeschätzt werden muss.

Bsp.:

- Wahrscheinlichkeit dafür, dass VP \rightarrow VP PP verwendet wird, vorausgesetzt, man möchte eine VP generieren.
- Wahrscheinlichkeit dafür, dass *chair* eine Nomen ist.

ideales Zufallsexperiment (Modell)

Anforderungen an ein ideales Zufallsexperiment:

- Das Experiment wird unter genau festgelegten Versuchsbedingungen durchgeführt.
- Die Menge der möglichen Ergebnisse ist vor der Durchführung des Experiments bekannt.
- Das Experiment kann zumindest prinzipiell beliebig oft unter gleichen Bedingungen wiederholt werden.

Ergebnisraum

Die Menge der möglichen Ergebnisse eines idealen Zufallsexperiments bildet den **Ergebnisraum** und wird mit Ω ('Omega') bezeichnet.

Ω wird auch der **Stichprobenraum** genannt.

Ist der Ergebnisraum nicht leer und abzählbar, dann heißt er **diskret**.

Zufallsexperiment und Ereignisse

Wir unterscheiden einzelne Ergebnisse und Ereignisse, die Mengen von Ergebnissen sind.

- Ein **Ereignis** bildet eine Teilmenge von Ω .
- \emptyset ist das **unmögliche** Ereignis.
- Ω ist das **sichere** Ereignis.
- Zwei Ereignisse E_1 und E_2 heißen **unvereinbar**, wenn $E_1 \cap E_2 = \emptyset$.
- Die Einermengen $\{e\}$ ($e \in \Omega$) heißen **Elementarereignisse**.
- Das Komplement eines Ereignisses E , also \bar{E} , heißt **Gegenereignis** zu E .

Beispiel Zufallsexperiment: Würfeln mit einem Würfel

- $\Omega = \{\square, \square, \square, \square, \square, \square\}$
- Der Wurf einer 3 ist das Elementarereignis $\{\square\}$ des Zufallsexperiments.
- $\{\square, \square, \square\}$ ist das Ereignis 'Wurf einer geraden Augenzahl'
- Das Gegenereignis von 'Wurf einer geraden Augenzahl' ist 'Wurf einer ungeraden Augenzahl'

Beispiel Zufallsexperiment: Augensumme bei zweimaligem Würfeln

| | | |
|----------|-------------------------|------------------------------------------------------------------------------------------------------|
| Summe 2 | entspricht dem Ereignis | $\{\square\square\}$ |
| Summe 3 | entspricht dem Ereignis | $\{\square\square, \square\square\}$ |
| Summe 4 | entspricht dem Ereignis | $\{\square\square, \square\square, \square\square\}$ |
| Summe 5 | entspricht dem Ereignis | $\{\square\square, \square\square, \square\square, \square\square\}$ |
| Summe 6 | entspricht dem Ereignis | $\{\square\square, \square\square, \square\square, \square\square, \square\square\}$ |
| Summe 7 | entspricht dem Ereignis | $\{\square\square, \square\square, \square\square, \square\square, \square\square, \square\square\}$ |
| Summe 8 | entspricht dem Ereignis | $\{\square\square, \square\square, \square\square, \square\square, \square\square\}$ |
| Summe 9 | entspricht dem Ereignis | $\{\square\square, \square\square, \square\square, \square\square\}$ |
| Summe 10 | entspricht dem Ereignis | $\{\square\square, \square\square, \square\square\}$ |
| Summe 11 | entspricht dem Ereignis | $\{\square\square, \square\square\}$ |
| Summe 12 | entspricht dem Ereignis | $\{\square\square\}$ |

Diskrete Wahrscheinlichkeitsräume

Ein **diskreter Wahrscheinlichkeitsraum** ist ein Paar (Ω, P) , bestehend aus

1. einer nicht leeren, abzählbaren Menge Ω von **Ergebnissen** (diskreter Ergebnisraum) und
2. einem **Wahrscheinlichkeitsmaß** $P: \mathcal{POT}(\Omega) \rightarrow \mathbb{R}$, so dass
 1. $P(A) \geq 0$ für alle $A \in \mathcal{POT}(\Omega)$;
 2. $P(\Omega) = 1$;
 3. für paarweise disjunkte Mengen $A_n \in \mathcal{POT}(\Omega)$, $n \in \mathbb{N}$ gilt

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n)$$

Diskrete Wahrscheinlichkeitsräume

Es ergeben sich folgende Eigenschaften für Wahrscheinlichkeitsmaße:

- 1 $P(\emptyset) = 0$
- 2 Für Ereignisse A, B mit $A \cap B = \emptyset$ gilt $P(A \cup B) = P(A) + P(B)$.
- 3 $P(A) + P(\bar{A}) = 1$ für alle $A \subseteq \Omega$ (Tertium non datur)
- 4 Impliziert Ereignis A das Ereignis B (d.h. $A \subseteq B$), dann gilt $P(B \setminus A) = P(B) - P(A)$.
- 5 Kein Ereignis kann eine Wahrscheinlichkeit über 1 haben.

Diskrete Wahrscheinlichkeitsräume

Bsp.: $\Omega = \{\text{is-noun, has-plural-s, is-adjective, is-verb}\}$.

Frage: Kann die Funktion f mit

$$\begin{aligned} f(\text{is-noun}) &= 0.45 \\ f(\text{has-plural-s}) &= 0.2 \\ f(\text{is-adjective}) &= 0.25 \\ f(\text{is-verb}) &= 0.3 \end{aligned}$$

zu einem Wahrscheinlichkeitsmaß $f: \mathcal{P}(\Omega) \rightarrow \mathbb{R}$ ergänzt werden?

Nein, da dann $f(\Omega) = 0.45 + 0.2 + 0.25 + 0.3 = 1.2 > 1$ wäre.

Diskrete Wahrscheinlichkeitsräume

Besser: $\Omega = \{\text{is-noun-with-plural-s, is-noun-without-plural-s, is-adjective, is-verb}\}$.

$$\begin{aligned} f(\text{is-noun-with-plural-s}) &= 0.09 \\ f(\text{is-noun-without-plural-s}) &= 0.36 \\ f(\text{is-adjective}) &= 0.25 \\ f(\text{is-verb}) &= 0.3 \end{aligned}$$

Laplace-Räume

Laplace-Räume sind diskrete Wahrscheinlichkeitsräume, in denen alle Ergebnisse gleich wahrscheinlich sind.

Bsp.: Würfelexperiment. $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Jedes Ergebnis hat die Wahrscheinlichkeit $\frac{1}{|\Omega|} = \frac{1}{6}$.

In Laplace-Räumen gilt also

$$P(A) = \frac{|A|}{|\Omega|}$$

Beispiel Laplace-Raum: zweimaliges Würfeln und Augensumme

| Augensumme | Ereignis | Wahrscheinlichkeit |
|------------|--------------------------------------------|--------------------|
| 2 | {(1,1)} | $\frac{1}{36}$ |
| 3 | {(1,2), (2,1)} | $\frac{2}{36}$ |
| 4 | {(1,3), (2,2), (3,1)} | $\frac{3}{36}$ |
| 5 | {(1,4), (2,3), (3,2), (4,1)} | $\frac{4}{36}$ |
| 6 | {(1,5), (2,4), (3,3), (4,2), (5,1)} | $\frac{5}{36}$ |
| 7 | {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)} | $\frac{6}{36}$ |
| 8 | {(2,6), (3,5), (4,4), (5,3), (6,2)} | $\frac{5}{36}$ |
| 9 | {(3,6), (4,5), (5,4), (6,3)} | $\frac{4}{36}$ |
| 10 | {(4,6), (5,5), (6,4)} | $\frac{3}{36}$ |
| 11 | {(5,6), (6,5)} | $\frac{2}{36}$ |
| 12 | {(6,6)} | $\frac{1}{36}$ |

Beispiel Laplace-Raum: Geburtstagsproblem

Bsp.: Wahrscheinlichkeit dafür, dass in einer Gruppe von 30 Personen mindestens zwei am gleichen Tag Geburtstag haben.

Vereinfachung: Wir ignorieren Schaltjahre und saisonale Variationen.

D.h., Wahrscheinlichkeit dafür, an einem bestimmten Tag Geburtstag zu haben, ist $\frac{1}{365}$.

Wahrscheinlichkeitsraum:

- $\Omega = \{1, \dots, 365\}^{30}$, also alle Folgen von 30 Zahlen aus $\{1, \dots, 365\}$.
- $|\Omega| = 365^{30}$. Alle Folgen sind gleichwahrscheinlich (Laplace-Raum).

Hinweis

Für die Modellierung als Laplace-Raum ist es unerlässlich, die Geburtstagsverteilung als Urnenproblem mit Beachtung der Reihenfolge zu betrachten.

Würde die Reihenfolge vernachlässigt und Ω als die Menge aller ungeordneten Kombinationen möglicher Geburtstagsverteilungen betrachtet (also $|\Omega| = \binom{n+k-1}{k}$), so wären die Ereignisse in dem Ereignisraum nicht gleichwahrscheinlich.

Beispiel:

Die Wahrscheinlichkeit für das Ereignis, dass alle am 1. Januar Geburtstag haben ist $(\frac{1}{365})^{30}$ während die Wahrscheinlichkeit dafür, dass die Geburtstage genau auf die ersten 30 Tage des Jahres fallen $(\frac{1}{365})^{30} * 30!$ ist.

Beispiel Laplace-Raum: Geburtstagsproblem

Ziel: Wahrscheinlichkeit dafür, dass eine Folge eintritt, in der sich mindestens ein Element wiederholt.

Einfacher: Wahrscheinlichkeitsermittlung über das Komplement.

Wieviel Folgen gibt es, in denen sich kein Element wiederholt?

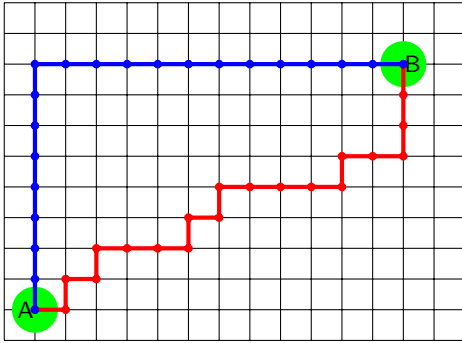
$$365 \times 364 \times \dots \times (365 - 29) = \frac{365!}{(365 - 30)!}$$

⇒ Wahrscheinlichkeit dafür, dass zwei am gleichen Tag Geburtstag haben ist

$$1 - \frac{365!}{365^{30}(365 - 30)!} \approx 1 - 0.29 = 0.71$$

Beispiel Laplace-Raum: Wege im Gitter

Wege von A nach B . Regel: Schritt nach oben oder Schritt nach rechts.
Beide Richtungen gleichwahrscheinlich.



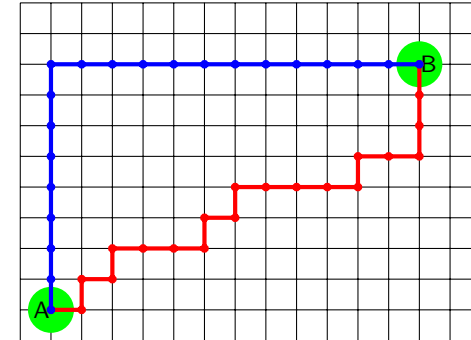
Wieviele Wege gibt es von A nach B ?

Jeder erfolgreiche Weg besteht aus 20 Schritten (8 nach oben, 12 nach rechts).

Zahl der möglichen Wege: $|A \rightarrow^{20} B| = \binom{20}{8} = 125970$

Beispiel Laplace-Raum: Wege im Gitter

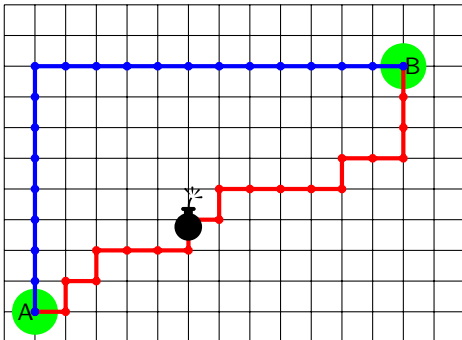
Wege von A nach B . Regel: Schritt nach oben oder Schritt nach rechts.
Beide Richtungen gleichwahrscheinlich.



Wie hoch ist die Wahrscheinlichkeit für das Ereignis nach 20 Schritten bei Punkt B anzukommen?

$$P(A \rightarrow^{20} B) = \frac{|A \rightarrow^{20} B|}{\Omega} = \frac{\binom{20}{8}}{2^{20}} = \frac{125970}{1048576} \approx 0,12$$

Beispiel Laplace-Raum: Gitter mit Bombe

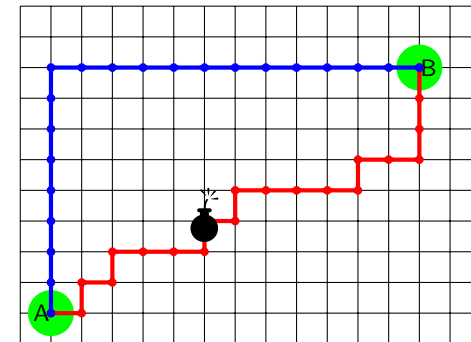


Wie hoch ist die Wahrscheinlichkeit für das Ereignis auf einem Weg von A nach B nicht auf die Bombe zu stoßen? Also die Wahrscheinlichkeit unter allen Wegen von A nach B einen zu wählen, der nicht auf die Bombe stößt.

Wegen von A nach B , die die Bombe treffen: $\binom{8}{3} \cdot \binom{12}{5} = 56 \cdot 792 = 44352$

Wahrscheinlichkeit Weg von A nach B ohne Bombe: $\frac{125970 - 44352}{125970} \approx 0,65$

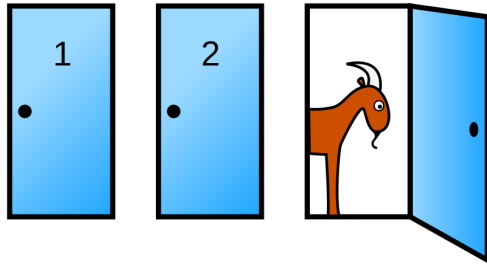
Beispiel Laplace-Raum: Gitter mit Bombe



Wie hoch ist die Wahrscheinlichkeit für das Ereignis nach 20 Schritten sowohl bei Punkt B anzukommen, als auch auf dem Weg nicht auf die Bombe zu treffen?

$$\frac{\binom{20}{8} - \binom{8}{3} \cdot \binom{12}{5}}{2^{20}} = \frac{125970 - 44352}{1048576} \approx 0,078$$

Ziegenproblem



Situation: 3 verschlossene Türen, hinter einer der Türen befindet sich ein Gewinn, hinter zwei Türen befinden sich Nieten (Ziegen).

- 1 der Kandidat wählt eine Tür
- 2 der Moderator öffnet von den verbleibenden beiden Türen eine Ziegentür
- 3 der Kandidat darf die Tür wechseln

Frage: Lohnt sich ein Wechsel?

Bedingte Wahrscheinlichkeiten

Bsp.:

- Wahrscheinlichkeit für eine Produktion $VP \rightarrow V NP$ für die Generierung einer VP, gegeben, dass es sich um das Verb *kisses* (bzw. *sleeps*) handelt.
- Wahrscheinlichkeit dafür, dass *chairs* ein Nomen ist, gegeben die Tatsache, dass das vorangehende Wort ein Artikel ist.
- Wahrscheinlichkeit dafür, dass *chairs* ein Nomen ist, gegeben die Tatsache, dass das nachfolgende Wort ein Artikel ist.

Bedingte Wahrscheinlichkeiten

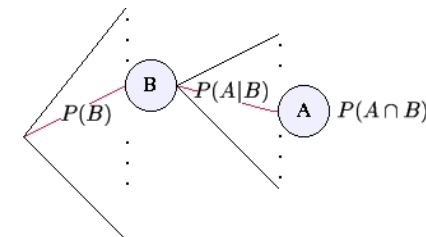
In einem diskreten Wahrscheinlichkeitsraum $\langle \Omega, P \rangle$, gegeben ein Ereignis $A \subseteq \Omega$ mit $P(A) > 0$, ist durch

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

das durch A bedingte Wahrscheinlichkeitsmaß $P(\cdot|A) : \Omega \rightarrow \mathbb{R}$ auf $\mathcal{P}(\Omega)$ definiert.

$\langle \mathcal{P}(\Omega), P(\cdot|A) \rangle$ ist ein diskreter Wahrscheinlichkeitsraum.

bedingte Wahrscheinlichkeiten: Produktregel



Produktregel

$$P(A \cap B) = P(B) \cdot P(A|B)$$

Unabhängigkeit von Ereignissen

Zwei Ereignisse A und B heißen **unabhängig**, falls $P(A \cap B) = P(A) \cdot P(B)$. Das heißt $P(A|B) = P(A)$.

Bsp. Würfelexperiment.

- Die Ereignisse, dass (A) eine gerade Zahl gewürfelt wird und (B) eine Zahl ≤ 2 , sind unabhängig:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{2\})}{P(\{1,2\})} = 0.5 = P(A)$$

- Die Ereignisse A wie oben und B , dass genau die 2 gewürfelt wird, sind nicht unabhängig:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{2\})}{P(\{2\})} = 1 \neq P(A)$$

Die Formel von Bayes



Thomas Bayes (1701-1761)

Die Formel von Bayes

Ziel: $P(A|B)$ berechnen auf der Grundlage von $P(B|A)$, $P(A)$ und $P(B)$.

Laut Definition gilt

$$P(A \cap B) = P(A|B) \cdot P(B) \text{ und } P(B \cap A) = P(B|A) \cdot P(A)$$

Daraus ergibt sich

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Die Formel von Bayes

Man kann das Theorem von Bayes noch verallgemeinern: Angenommen, es gibt eine endliche oder abzählbar unendliche Folge $(A_i)_{i \in \mathbb{N}}$ von paarweise disjunkten Ereignissen mit $A_i \subseteq \Omega$ und $P(A_i) > 0$ für alle $i \in \mathbb{N}$, die eine Zerlegung von Ω bilden, dann gilt für jedes Ereignis $B \subseteq \Omega$: $(B \cap A_i)_{i \in \mathbb{N}}$ bildet eine disjunkte Zerlegung von B , und daher

$$P(B) = \sum_{i \in \mathbb{N}} P(B|A_i)P(A_i)$$

Spezialfall: Zerlegung in A und \bar{A} :

$$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$$

$$P(B) = \frac{P(B \cap A)}{P(A)}P(A) + \frac{P(B \cap \bar{A})}{P(\bar{A})}P(\bar{A})$$

$$P(B) = P(B \cap A) + P(B \cap \bar{A})$$

Die Formel von Bayes

Aus

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

und

$$P(B) = \sum_{i \in \mathbb{N}} P(B|A_i)P(A_i)$$

ergibt sich dann für die Folge $(A_i)_{i \in \mathbb{N}}$ und das Ereignis B wie oben die verallgemeinerte Formel von Bayes:

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i \in \mathbb{N}} P(B|A_i)P(A_i)}$$

Die Formel von Bayes

Aus

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

und

$$P(B) = \sum_{i \in \mathbb{N}} P(B|A_i)P(A_i)$$

ergibt sich dann für die Folge $(A_i)_{i \in \mathbb{N}}$ und das Ereignis B wie oben die verallgemeinerte Formel von Bayes:

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i \in \mathbb{N}} P(B|A_i)P(A_i)}$$

Die Formel von Bayes

Bsp.: Angenommen, wir interessieren uns für eine relativ seltene Konstruktion, z.B. *Parasitic Gaps*, die ungefähr alle 100.000 Sätze einmal vorkommt.¹ Joe Linguist hat einen Pattern-Matching Algorithmus zur Erkennung von Parasitic Gaps implementiert, der, falls ein Satz ein Parasitic Gap enthält, dies mit einer Wahrscheinlichkeit von 0.95 auch erkennt. Enthält ein Satz kein Parasitic Gap, liefert der Algorithmus mit einer Wahrscheinlichkeit von 0.005 das falsche Ergebnis, dass ein Parasitic Gap in dem Satz vorhanden ist.

Frage: Angenommen, der Test meldet ein Parasitic Gap in einem Satz. Wie wahrscheinlich ist es, dass es sich wirklich um eines handelt?

¹Z.B. *which book did she review* _ without reading _?

Die Formel von Bayes

Sei G das Ereignis eines parasitic gaps, T das eines positiven Tests. Wir kennen die Werte $P(G) = \frac{1}{100.000} = 0,00001$, $P(\bar{G}) = 1 - P(G) = 0,99999$, $P(T|G) = 0,95$ und $P(T|\bar{G}) = 0,005$.

Wir wollen $P(G|T)$ berechnen.

$$P(G|T) = \frac{P(G \cap T)}{P(T)} = \frac{P(T|G) \cdot P(G)}{P(T)}$$

$P(T)$ lässt sich über $P(T|G)$ und $P(T|\bar{G})$ berechnen:

$$P(T) = P(G \cap T) + P(T \cap \bar{G}) = P(T|G) \cdot P(G) + P(T|\bar{G}) \cdot P(\bar{G})$$

Also erhalten wir

$$\begin{aligned} P(G|T) &= \frac{P(T|G)P(G)}{P(T|G)P(G) + P(T|\bar{G})P(\bar{G})} \\ &= \frac{0.95 \times 0.00001}{0.95 \times 0.00001 + 0.005 \times 0.99999} \\ &\approx 0.002 \end{aligned}$$

Gefangenenparadoxon

Aus drei zum Tode verurteilten Gefangene (Anton, Bernd und Clemens) wird einer zur Begnadigung ausgewählt. Anton erfährt, dass Bernd hingerichtet wird. Anton erzählt dies Clemens weiter.

- Anton: entweder Clemens wird begnadigt oder er selbst, so dass seine Überlebenschance von $1/3$ auf $1/2$ gestiegen sei.
- Clemens: Überlebenschance von $1/3$ auf $2/3$ gestiegen

Wer von beiden Gefangenen schätzt seine Chancen korrekt ein?

Berkeley 1973 (Simpsonsche Paradoxon)

- Annahmequote Universität
für Männer: 44,3% für Frauen: 34,6%
- liegt hier eine Benachteiligung der Frauen vor?

Nicht notwendig, siehe: [Statistikmodul Mathe-Prisma \(Link\)](#)

Hinweis

Arbeiten Sie bitte das Mathe-Prisma Modul zur bedingten Wahrscheinlichkeit durch ([Link](#))