Technische Universität Dresden

Formale Systeme

14. Vorlesung: Abschlusseigenschaften kontextfreier Sprachen

Markus Krötzsch

Professur für Wissensbasierte Systeme

TU Dresden, 1. Dezember 2025

Rückblick: Das Pumping Lemma

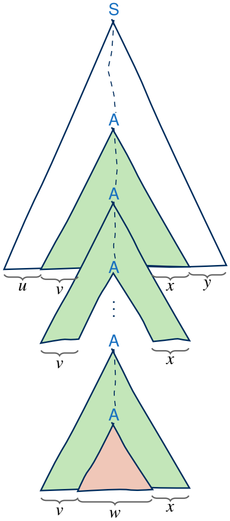
Pumpen für kontextfreie Sprachen

Satz (Pumping Lemma): Für jede kontextfreie Sprache **L**
gibt es eine Zahl $n \geq 0$, so dass gilt:
für jedes Wort $z \in$ **L** mit $|z| \geq n$
gibt es eine Zerlegung $z = uvwxy$ mit $|vx| \geq 1$ und $|vwx| \leq n$, s.d.:
für jede Zahl $k \geq 0$ gilt: $uv^k wx^k y \in$ **L**

Beispiel: Für die Sprache $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\}$ gilt der Satz. Wir wählen $n = 2$. Sei $z = \mathtt{a}^i\mathtt{b}^i$ mit $i \geq 1$ ein beliebiges Wort mit $|z| \geq 2$. Wir wählen die Zerlegung $u = \mathtt{a}^{i-1}$, $v = \mathtt{a}$, $w = \epsilon$, $x = \mathtt{b}$ und $y = \mathtt{b}^{i-1}$.
Dann ist $uv^k wx^k y = \mathtt{a}^{i-1}\mathtt{a}^k\mathtt{b}^k\mathtt{b}^{i-1} = \mathtt{a}^{i+k-1}\mathtt{b}^{i+k-1} \in$ **L** für alle $k \geq 0$.

Ableitungsbäume aufpumpen

Die Idee des Lemmas lässt sich gut am Ableitungsbaum darstellen:



Abgeleitetes Wort:
uv^kwx^ky

Beispiel

CNF-Grammatik für $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 1\}$:

$$\begin{aligned} \mathsf{S} &\to \mathsf{AB} \mid \mathsf{AC} \\ \mathsf{A} &\to \mathtt{a} \\ \mathsf{B} &\to \mathtt{b} \\ \mathsf{C} &\to \mathsf{SB} \end{aligned}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel

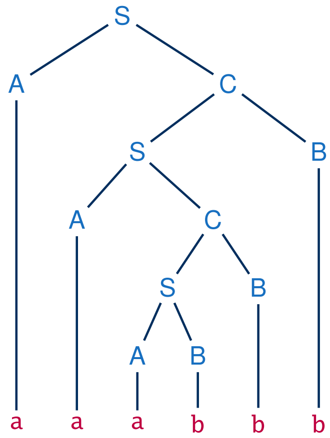
CNF-Grammatik für $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 1\}$:

$$\mathsf{S} \to \mathsf{AB} \mid \mathsf{AC}$$
$$\mathsf{A} \to \mathtt{a}$$
$$\mathsf{B} \to \mathtt{b}$$
$$\mathsf{C} \to \mathsf{SB}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für `aaabbb`



Beispiel

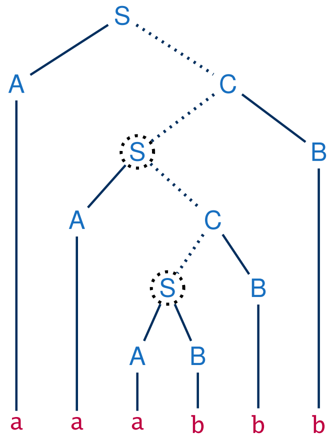
CNF-Grammatik für $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 1\}$:

$$\mathsf{S} \to \mathsf{AB} \mid \mathsf{AC}$$
$$\mathsf{A} \to \mathtt{a}$$
$$\mathsf{B} \to \mathtt{b}$$
$$\mathsf{C} \to \mathsf{SB}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für `aaabbb`



Beispiel

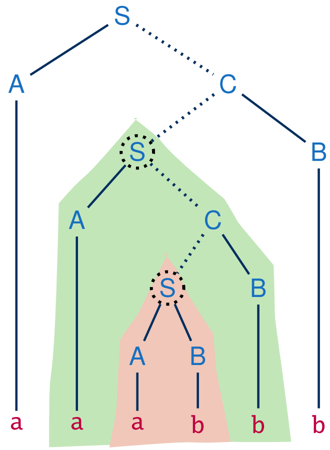
CNF-Grammatik für $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 1\}$:

$$S \to AB \mid AC$$
$$A \to \mathtt{a}$$
$$B \to \mathtt{b}$$
$$C \to SB$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für aaabbb



Beispiel

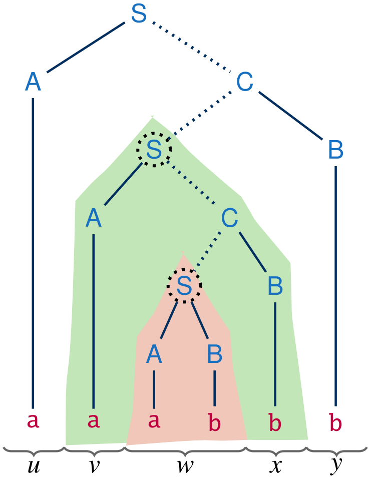
CNF-Grammatik für $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 1\}$:

$$\mathsf{S} \to \mathsf{AB} \mid \mathsf{AC}$$
$$\mathsf{A} \to \mathtt{a}$$
$$\mathsf{B} \to \mathtt{b}$$
$$\mathsf{C} \to \mathsf{SB}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für `aaabbb`



Beispiel (2)

CNF-Grammatik für $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 1\}$:

$$\mathsf{S} \to \mathsf{AB} \mid \mathsf{AC}$$

$$\mathsf{A} \to \mathtt{a}$$

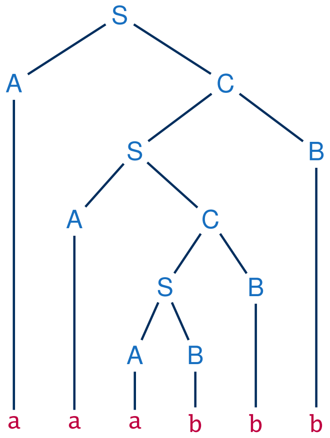
$$\mathsf{B} \to \mathtt{b}$$

$$\mathsf{C} \to \mathsf{SB}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für `aaabbb`



Beispiel (2)

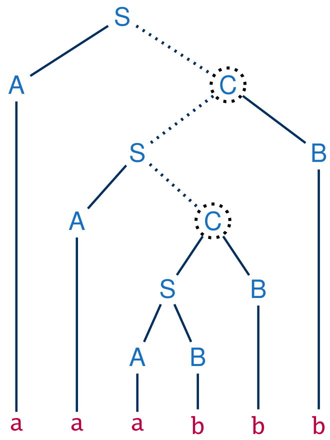
CNF-Grammatik für $\{a^i b^i \mid i \geq 1\}$:

$$\mathsf{S} \to \mathsf{AB} \mid \mathsf{AC}$$
$$\mathsf{A} \to \mathtt{a}$$
$$\mathsf{B} \to \mathtt{b}$$
$$\mathsf{C} \to \mathsf{SB}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für aaabbb



Beispiel (2)

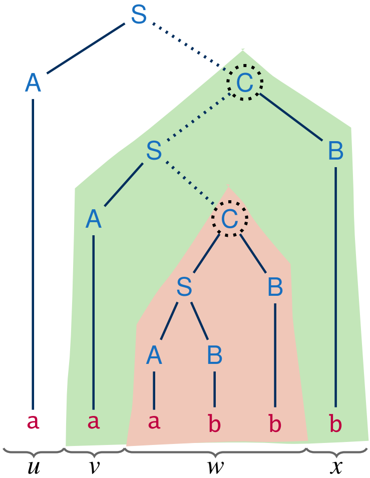
CNF-Grammatik für $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 1\}$:

$$\mathsf{S} \to \mathsf{AB} \mid \mathsf{AC}$$
$$\mathsf{A} \to \mathtt{a}$$
$$\mathsf{B} \to \mathtt{b}$$
$$\mathsf{C} \to \mathsf{SB}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für aaabbb



Beispiel (3)

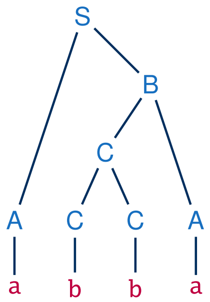
CNF-Grammatik für $\mathtt{ab^+a}$:

$$\begin{aligned} S &\to AB \\ A &\to \mathtt{a} \\ B &\to CA \\ C &\to CC \mid \mathtt{b} \end{aligned}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für $\mathtt{abba}$



Beispiel (3)

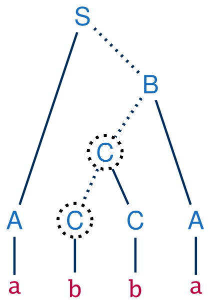
CNF-Grammatik für $\mathtt{ab}^{+}\mathtt{a}$:

$$S \to AB$$
$$A \to \mathtt{a}$$
$$B \to CA$$
$$C \to CC \mid \mathtt{b}$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für $\mathtt{abba}$



Beispiel (3)

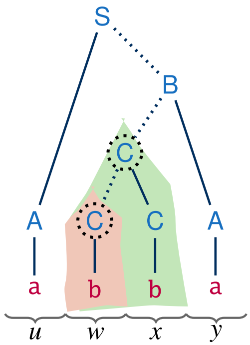
CNF-Grammatik für $\mathtt{ab^+a}$:

$$
\begin{aligned}
S &\to AB\\
A &\to \mathtt{a}\\
B &\to CA\\
C &\to CC \mid \mathtt{b}
\end{aligned}
$$

Konstante aus dem Beweis des Pumping-Lemma: $n = 2^4 = 16$

Aber man kann auch schon ab $n = 4$ pumpen

Beispiel: Ableitung für $\mathtt{abba}$



$v = \epsilon$

Abschlusseigenschaften kontextfreier Sprachen

Rückblick: Reguläre Sprachen

Satz: Wenn $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$ reguläre Sprachen sind, dann beschreiben auch die folgenden Ausdrücke reguläre Sprachen:

(1) $\mathbf{L_1} \cup \mathbf{L_2}$ (Abschluss unter Vereinigung)

(2) $\mathbf{L_1} \cap \mathbf{L_2}$ (Abschluss unter Schnitt)

(3) $\overline{\mathbf{L}}$ (Abschluss unter Komplement)

(4) $\mathbf{L_1} \circ \mathbf{L_2}$ (Abschluss unter Konkatenation)

(5) $\mathbf{L}^*$ (Abschluss unter Kleene-Stern)

Wie sieht es bei den kontextfreien Sprachen aus?

Abschluss für kontextfreie Sprachen?

Bei kontextfreien Sprachen ergibt sich ein anderes Bild:

Satz: Wenn $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$ kontextfreie Sprachen sind, dann beschreiben auch die folgenden Ausdrücke kontextfreie Sprachen:

(1) $\mathbf{L_1} \cup \mathbf{L_2}$ (Abschluss unter Vereinigung)

(2) $\mathbf{L_1} \circ \mathbf{L_2}$ (Abschluss unter Konkatenation)

(3) $\mathbf{L}^*$ (Abschluss unter Kleene-Stern)

Aber:

Satz: Es gibt kontextfreie Sprachen $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$, so dass die folgenden Ausdrücke keine kontextfreien Sprachen sind:

(1) $\mathbf{L_1} \cap \mathbf{L_2}$ (Nichtabschluss unter Schnitt)

(2) $\overline{\mathbf{L}}$ (Nichtabschluss unter Komplement)

Abschlusseigenschaften von Typ-2-Sprachen

Satz: Wenn $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$ kontextfreie Sprachen sind, dann beschreiben auch die folgenden Ausdrücke kontextfreie Sprachen:

(1) $\mathbf{L_1} \cup \mathbf{L_2}$ Abschluss unter Vereinigung

(2) $\mathbf{L_1} \circ \mathbf{L_2}$ (Abschluss unter Konkatenation)

(3) $\mathbf{L}^*$ (Abschluss unter Kleene-Stern)

Abschluss unter Vereinigung

Wir konstruieren eine Vereingungsgrammatik

Gegeben seien zwei formale Grammatiken $G_1 = \langle V_1, \Sigma, P_1, \mathsf{S}_1 \rangle$ und $G_2 = \langle V_2, \Sigma, P_2, \mathsf{S}_2 \rangle$ mit $V_1 \cap V_2 = \emptyset$ (o.B.d.A.).

Die Vereinigungsgrammatik $G_1 \uplus G_2$ ist gegeben durch

$$G_1 \uplus G_2 = \langle V_1 \cup V_2 \cup \{\mathsf{S}\}, \Sigma, P_1 \cup P_2 \cup \{\mathsf{S} \to \mathsf{S}_1 \mid \mathsf{S}_2\}, \mathsf{S} \rangle,$$

wobei S ein neues Startsymbol ist, das nicht in $V_1 \cup V_2$ vorkommt.

In Worten: Die neuen Ableitungsregeln $\mathsf{S} \to \mathsf{S}_1 \mid \mathsf{S}_2$ ermöglichen es, dass $G_1 \uplus G_2$ entweder Wörter aus G_1 oder aus G_2 generiert.

Abschluss unter Vereinigung

Wir konstruieren eine Vereingungsgrammatik

Gegeben seien zwei formale Grammatiken $G_1 = \langle V_1, \Sigma, P_1, \mathsf{S}_1\rangle$ und $G_2 = \langle V_2, \Sigma, P_2, \mathsf{S}_2\rangle$ mit $V_1 \cap V_2 = \emptyset$ (o.B.d.A.).

Die Vereinigungsgrammatik $G_1 \uplus G_2$ ist gegeben durch

$$G_1 \uplus G_2 = \langle V_1 \cup V_2 \cup \{\mathsf{S}\}, \Sigma, P_1 \cup P_2 \cup \{\mathsf{S} \to \mathsf{S}_1 \mid \mathsf{S}_2\}, \mathsf{S}\rangle,$$

wobei S ein neues Startsymbol ist, das nicht in $V_1 \cup V_2$ vorkommt.

In Worten: Die neuen Ableitungsregeln $\mathsf{S} \to \mathsf{S}_1 \mid \mathsf{S}_2$ ermöglichen es, dass $G_1 \uplus G_2$ entweder Wörter aus G_1 oder aus G_2 generiert.

Es ist daher leicht zu sehen:

Satz: $\mathbf{L}(G_1 \uplus G_2) = \mathbf{L}(G_1) \cup \mathbf{L}(G_2)$.

Vereinigungen in Typ 2?

Bisher haben wir nur erkannt:

Satz: $\mathbf{L}(G_1 \uplus G_2) = \mathbf{L}(G_1) \cup \mathbf{L}(G_2)$.

Vereinigungen in Typ 2?

Bisher haben wir nur erkannt:

Satz: $\mathbf{L}(G_1 \uplus G_2) = \mathbf{L}(G_1) \cup \mathbf{L}(G_2)$.

Für die Abschlusseigenschaft sollte noch mehr gelten:

Satz: Wenn G_1 und G_2 kontextfrei sind, dann ist auch $G_1 \uplus G_2$ kontextfrei.

Das ist leicht zu sehen, da die beiden zusätzlichen Regeln kontextfrei sind.

Vereinigungen in Typ 2?

Bisher haben wir nur erkannt:

Satz: $\mathbf{L}(G_1 \uplus G_2) = \mathbf{L}(G_1) \cup \mathbf{L}(G_2)$.

Für die Abschlusseigenschaft sollte noch mehr gelten:

Satz: Wenn G_1 und G_2 kontextfrei sind, dann ist auch $G_1 \uplus G_2$ kontextfrei.

Das ist leicht zu sehen, da die beiden zusätzlichen Regeln kontextfrei sind.

Daher gilt sogar noch ein stärkerer Satz:

Satz: Wenn G_1 und G_2 von Typ $i \in \{2, 1, 0\}$ sind, dann ist auch $G_1 \uplus G_2$ von Typ i.

$\leadsto$ Typ-0-Sprachen, kontextsensitive Sprachen und kontextfreie Sprachen sind unter Vereinigung abgeschlossen

(reguläre Sprachen auch, aber das haben wir anders gezeigt)

Abschlusseigenschaften von Typ-2-Sprachen

Satz: Wenn $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$ kontextfreie Sprachen sind, dann beschreiben auch die folgenden Ausdrücke kontextfreie Sprachen:

(1) $\mathbf{L_1} \cup \mathbf{L_2}$ (Abschluss unter Vereinigung)

(2) $\mathbf{L_1} \circ \mathbf{L_2}$ Abschluss unter Konkatenation

(3) $\mathbf{L}^*$ (Abschluss unter Kleene-Stern)

Konkatenation von Grammatiken

Wir erinnern uns: $\mathbf{L_1} \circ \mathbf{L_2} = \{w_1 w_2 \mid w_1 \in \mathbf{L_1}$ und $w_2 \in \mathbf{L_2}\}$

Es ist nicht schwer, eine passende Grammatik zu finden:

Gegeben seien zwei formale Grammatiken $G_1 = \langle V_1, \Sigma, P_1, \mathsf{S}_1\rangle$ und $G_2 = \langle V_2, \Sigma, P_2, \mathsf{S}_2\rangle$ mit $V_1 \cap V_2 = \emptyset$ (o.B.d.A.).

Die Grammatik $G_1 \circ G_2$ ist gegeben durch

$$G_1 \circ G_2 = \langle V_1 \cup V_2 \cup \{\mathsf{S}\}, \Sigma, P_1 \cup P_2 \cup \{\mathsf{S} \to \mathsf{S}_1\mathsf{S}_2\}, \mathsf{S}\rangle,$$

wobei S ein neues Startsymbol ist, das nicht in $V_1 \cup V_2$ vorkommt.

In Worten: Die neue Ableitungsregel $\mathsf{S} \to \mathsf{S}_1\mathsf{S}_2$ ermöglicht es, dass $G_1 \circ G_2$ Wörter aus G_1 gefolgt von Wörtern aus G_2 generiert.

Beispiel Konkatenation

Wir betrachten die Grammatiken

$$G_1 : \quad S_1 \to aS_1b \mid \epsilon$$
$$G_2 : \quad S_2 \to cS_2 \mid \epsilon$$

Beispiel Konkatenation

Wir betrachten die Grammatiken

$$\begin{aligned} G_1 &: \quad S_1 \to aS_1b \mid \epsilon \\ G_2 &: \quad S_2 \to cS_2 \mid \epsilon \end{aligned}$$

Damit ist $\mathbf{L}(G_1) = \{a^i b^i \mid i \geq 0\}$ und $\mathbf{L}(G_2) = \{c^k \mid k \geq 0\}$.

Beispiel Konkatenation

Wir betrachten die Grammatiken

$$G_1 : \quad \mathsf{S}_1 \to \mathtt{a}\mathsf{S}_1\mathtt{b} \mid \epsilon$$
$$G_2 : \quad \mathsf{S}_2 \to \mathtt{c}\mathsf{S}_2 \mid \epsilon$$

Damit ist $\mathbf{L}(G_1) = \{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\}$ und $\mathbf{L}(G_2) = \{\mathtt{c}^k \mid k \geq 0\}$.

Die Konkatenation der Grammatiken enthält die folgenden Regeln:

$$G_1 \circ G_2 : \qquad \mathsf{S} \to \mathsf{S}_1\mathsf{S}_2 \qquad \mathsf{S}_1 \to \mathtt{a}\mathsf{S}_1\mathtt{b} \mid \epsilon \qquad \mathsf{S}_2 \to \mathtt{c}\mathsf{S}_2 \mid \epsilon$$

Beispiel Konkatenation

Wir betrachten die Grammatiken

$$G_1: \quad \mathsf{S}_1 \to \mathtt{a}\mathsf{S}_1\mathtt{b} \mid \epsilon$$

$$G_2: \quad \mathsf{S}_2 \to \mathtt{c}\mathsf{S}_2 \mid \epsilon$$

Damit ist $\mathbf{L}(G_1) = \{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\}$ und $\mathbf{L}(G_2) = \{\mathtt{c}^k \mid k \geq 0\}$.

Die Konkatenation der Grammatiken enthält die folgenden Regeln:

$$G_1 \circ G_2: \qquad \mathsf{S} \to \mathsf{S}_1\mathsf{S}_2 \qquad \mathsf{S}_1 \to \mathtt{a}\mathsf{S}_1\mathtt{b} \mid \epsilon \qquad \mathsf{S}_2 \to \mathtt{c}\mathsf{S}_2 \mid \epsilon$$

Damit ergibt sich $\mathbf{L}(G_1 \circ G_2) = \{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^k \mid i \geq 0, k \geq 0\}$.

Korrektheit

Hypothese: $\mathbf{L}(G_1 \circ G_2) = \mathbf{L}(G_1) \circ \mathbf{L}(G_2)$

Korrektheit

Hypothese: $\mathbf{L}(G_1 \circ G_2) = \mathbf{L}(G_1) \circ \mathbf{L}(G_2)$

Leider stimmt das nicht:

Gegenbeispiel: G_1 sei die Grammatik mit der einen Regel $\mathsf{S}_1 \to \mathtt{a}$ und G_2 die Grammatik mit den Regeln $\mathsf{S}_2 \to \mathtt{b}$ und $\mathtt{a}\mathsf{S}_2 \to \mathtt{c}$.

Dann ist $\mathbf{L}(G_1) = \{\mathtt{a}\}$ und $\mathbf{L}(G_2) = \{\mathtt{b}\}$.
Trotzdem erlaubt $G_1 \circ G_2$ die Ableitung $\mathsf{S} \Rightarrow \mathsf{S}_1\mathsf{S}_2 \Rightarrow \mathtt{a}\mathsf{S}_2 \Rightarrow \mathtt{c}$.

Korrektheit

Hypothese: $\mathbf{L}(G_1 \circ G_2) = \mathbf{L}(G_1) \circ \mathbf{L}(G_2)$

Leider stimmt das nicht:

Gegenbeispiel: G_1 sei die Grammatik mit der einen Regel $\mathsf{S}_1 \to \mathtt{a}$ und G_2 die Grammatik mit den Regeln $\mathsf{S}_2 \to \mathtt{b}$ und $\mathtt{a}\mathsf{S}_2 \to \mathtt{c}$.

Dann ist $\mathbf{L}(G_1) = \{\mathtt{a}\}$ und $\mathbf{L}(G_2) = \{\mathtt{b}\}$.
Trotzdem erlaubt $G_1 \circ G_2$ die Ableitung $\mathsf{S} \Rightarrow \mathsf{S}_1\mathsf{S}_2 \Rightarrow \mathtt{a}\mathsf{S}_2 \Rightarrow \mathtt{c}$.

Korrekt ist dagegen:

Satz: Wenn G_1 und G_2 kontextfrei sind, dann ist $\mathbf{L}(G_1 \circ G_2) = \mathbf{L}(G_1) \circ \mathbf{L}(G_2)$.
Zudem ist $G_1 \circ G_2$ in diesem Fall kontextfrei.

Beweis: einfach (man betrachte den Ableitungsbaum).

Damit erhalten wir den gewünschten Abschluss.

Abschlusseigenschaften von Typ-2-Sprachen

Satz: Wenn $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$ kontextfreie Sprachen sind, dann beschreiben auch die folgenden Ausdrücke kontextfreie Sprachen:

(1) $\mathbf{L_1} \cup \mathbf{L_2}$ (Abschluss unter Vereinigung)

(2) $\mathbf{L_1} \circ \mathbf{L_2}$ (Abschluss unter Konkatenation)

(3) $\mathbf{L}^*$ Abschluss unter Kleene-Stern

Kleene-Stern für Grammatiken

Wir erinnern uns: $\mathbf{L}^* = \{w_1w_2\cdots w_i \mid i \geq 0, w_1, \ldots, w_i \in \mathbf{L}\} = \bigcup_{i\geq 0} \mathbf{L}^i$

Kleene-Stern für Grammatiken

Wir erinnern uns: $\mathbf{L}^* = \{w_1 w_2 \cdots w_i \mid i \geq 0, w_1, \ldots, w_i \in \mathbf{L}\} = \bigcup_{i\geq 0} \mathbf{L}^i$

Auch hier kann man leicht eine passende Grammatik finden:

Gegeben sei eine formale Grammatik $G = \langle V, \Sigma, P, \mathsf{S}\rangle$.

Die Grammatik G^* ist gegeben durch

$$G^* = \langle V \cup \{\mathsf{S}'\}, \Sigma, P \cup \{\mathsf{S}' \to \epsilon \mid \mathsf{SS}'\}, \mathsf{S}'\rangle,$$

wobei S' ein neues Startsymbol ist, das nicht in V vorkommt.

In Worten: Die neuen Ableitungsregeln $\mathsf{S}' \to \epsilon \mid \mathsf{SS}'$ ermöglichen es, dass G^* beliebig lange Ketten aus Wörtern aus G generiert.

Beispiel Kleene-Stern

Wir betrachten die Grammatik

$$G: \qquad \mathtt{S} \to \mathtt{SaSb} \mid \mathtt{SbSa} \mid \epsilon$$

Übung: Welche Sprache erzeugt diese Grammatik?

Beispiel Kleene-Stern

Wir betrachten die Grammatik

$$G: \qquad \mathtt{S} \to \mathtt{SaSb} \mid \mathtt{SbSa} \mid \epsilon$$

Übung: Welche Sprache erzeugt diese Grammatik?

Der Kleene-Abschluss dieser Grammatik ist

$$G^*: \qquad \mathtt{S'} \to \epsilon \mid \mathtt{SS'} \qquad\qquad \mathtt{S} \to \mathtt{SaSb} \mid \mathtt{SbSa} \mid \epsilon$$

Übung: Welche Sprache erzeugt diese Grammatik?

Korrektheit beim Kleene-Abschluss

Wie schon beim Konkatenation funktioniert diese Operation auf kontextfreien Grammatiken wie erwünscht:

Satz: Wenn G kontextfrei ist, dann ist $\mathbf{L}(G^*) = \mathbf{L}(G)^*$.
Zudem ist G^* in diesem Fall kontextfrei.

Beweis: einfach (man betrachte den Ableitungsbaum).

Korrektheit beim Kleene-Abschluss

Wie schon beim Konkatenation funktioniert diese Operation auf kontextfreien Grammatiken wie erwünscht:

Satz: Wenn G kontextfrei ist, dann ist $\mathbf{L}(G^*) = \mathbf{L}(G)^*$.
Zudem ist G^* in diesem Fall kontextfrei.

Beweis: einfach (man betrachte den Ableitungsbaum).

Für nicht kontextfreie Grammatiken gilt dieser Satz im Allgemeinen nicht:

Beispiel: Die (kontextsensitive) Grammatik G mit der einen Regel $\mathsf{SS} \to \mathtt{0}$ kann nichts hervorbringen.
Trotzdem erlaubt die Grammatik G^* die Ableitung eines Wortes:

$$\mathsf{S}' \Rightarrow \mathsf{SS}' \Rightarrow \mathsf{SSS}' \Rightarrow \mathsf{SS} \Rightarrow \mathtt{0}$$

Nichtabschlusseigenschaften

Satz: Es gibt kontextfreie Sprachen $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$, so dass die folgenden Ausdrücke keine kontextfreien Sprachen sind:

(1) $\mathbf{L_1} \cap \mathbf{L_2}$ Nichtabschluss unter Schnitt

(2) $\overline{\mathbf{L}}$ (Nichtabschluss unter Komplement)

Nichtabschluss unter $\cap$

Beweisansatz: Wir müssen kontextfreie Sprachen $\mathbf{L_1}$ und $\mathbf{L_2}$ finden, für die $\mathbf{L_1} \cap \mathbf{L_2}$ nicht kontextfrei ist.

Welche nichtkontextfreien Sprachen kennen wir?

Nichtabschluss unter $\cap$

Beweisansatz: Wir müssen kontextfreie Sprachen $\mathbf{L_1}$ und $\mathbf{L_2}$ finden, für die $\mathbf{L_1} \cap \mathbf{L_2}$ nicht kontextfrei ist.

Welche nichtkontextfreien Sprachen kennen wir?

$$\{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\}$$

Nichtabschluss unter $\cap$

Beweisansatz: Wir müssen kontextfreie Sprachen $\mathbf{L_1}$ und $\mathbf{L_2}$ finden, für die $\mathbf{L_1} \cap \mathbf{L_2}$ nicht kontextfrei ist.

Welche nichtkontextfreien Sprachen kennen wir?

$$\{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\}$$

Welche ähnlichen kontextfreien Sprachen kennen wir?

Nichtabschluss unter $\cap$

Beweisansatz: Wir müssen kontextfreie Sprachen $\mathbf{L_1}$ und $\mathbf{L_2}$ finden, für die $\mathbf{L_1} \cap \mathbf{L_2}$ nicht kontextfrei ist.

Welche nichtkontextfreien Sprachen kennen wir?

$$\{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\}$$

Welche ähnlichen kontextfreien Sprachen kennen wir?

$$\{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^k \mid i \geq 0, k \geq 0\}$$

Nichtabschluss unter $\cap$ (2)

Satz: Es gibt kontextfreie Sprachen $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$, so dass $\mathbf{L_1} \cap \mathbf{L_2}$ keine kontextfreie Sprache ist.

Beweis: Die folgenden Sprachen sind kontextfrei:

$$\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\} \qquad \text{(zuvor gezeigt)}$$

$$\{\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\} \qquad \text{(analog)}$$

$$\{\mathtt{a}^i \mid i \geq 0\} \qquad \text{(regulär)}$$

$$\{\mathtt{c}^i \mid i \geq 0\} \qquad \text{(regulär)}$$

Nichtabschluss unter ∩ (2)

Satz: Es gibt kontextfreie Sprachen **L**, $\mathbf{L_1}$ und $\mathbf{L_2}$, so dass $\mathbf{L_1} \cap \mathbf{L_2}$ keine kontextfreie Sprache ist.

Beweis: Die folgenden Sprachen sind kontextfrei:

$\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\}$	(zuvor gezeigt)
$\{\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\}$	(analog)
$\{\mathtt{a}^i \mid i \geq 0\}$	(regulär)
$\{\mathtt{c}^i \mid i \geq 0\}$	(regulär)

Dank Abschluss unter Konkatenation sind also auch kontextfrei:

$$\mathbf{L_1} = \{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\} \circ \{\mathtt{c}^i \mid i \geq 0\} = \{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^k \mid i \geq 0, k \geq 0\}$$
$$\mathbf{L_2} = \{\mathtt{a}^i \mid i \geq 0\} \circ \{\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\} = \{\mathtt{a}^i\mathtt{b}^k\mathtt{c}^k \mid i \geq 0, k \geq 0\}.$$

Nichtabschluss unter ∩ (2)

Satz: Es gibt kontextfreie Sprachen **L**, $\mathbf{L_1}$ und $\mathbf{L_2}$, so dass $\mathbf{L_1} \cap \mathbf{L_2}$ keine kontextfreie Sprache ist.

Beweis: Die folgenden Sprachen sind kontextfrei:

$$\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\} \qquad \text{(zuvor gezeigt)}$$
$$\{\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\} \qquad \text{(analog)}$$
$$\{\mathtt{a}^i \mid i \geq 0\} \qquad \text{(regulär)}$$
$$\{\mathtt{c}^i \mid i \geq 0\} \qquad \text{(regulär)}$$

Dank Abschluss unter Konkatenation sind also auch kontextfrei:

$$\mathbf{L_1} = \{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\} \circ \{\mathtt{c}^i \mid i \geq 0\} = \{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^k \mid i \geq 0, k \geq 0\}$$
$$\mathbf{L_2} = \{\mathtt{a}^i \mid i \geq 0\} \circ \{\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\} = \{\mathtt{a}^i\mathtt{b}^k\mathtt{c}^k \mid i \geq 0, k \geq 0\}.$$

Der Schnitt $\mathbf{L_1} \cap \mathbf{L_2}$ ist aber $\{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\}$ und also nicht kontextfrei (zuvor gezeigt). □

Nichtabschlusseigenschaften

Satz: Es gibt kontextfreie Sprachen $\mathbf{L}$, $\mathbf{L_1}$ und $\mathbf{L_2}$, so dass die folgenden Ausdrücke keine kontextfreien Sprachen sind:

(1) $\mathbf{L_1} \cap \mathbf{L_2}$ (Nichtabschluss unter Schnitt)

(2) $\overline{\mathbf{L}}$ Nichtabschluss unter Komplement

Nichtabschluss unter Komplement

Satz: Es gibt eine kontextfreie Sprache $\mathbf{L}$, so dass $\overline{\mathbf{L}}$ keine kontextfreie Sprache ist.

Nichtabschluss unter Komplement

Satz: Es gibt eine kontextfreie Sprache $\mathbf{L}$, so dass $\overline{\mathbf{L}}$ keine kontextfreie Sprache ist.

Beweis: Die Behauptung folgt unmittelbar aus dem bereits gezeigten

Nichtabschluss unter Komplement

Satz: Es gibt eine kontextfreie Sprache $\mathbf{L}$, so dass $\overline{\mathbf{L}}$ keine kontextfreie Sprache ist.

Beweis: Die Behauptung folgt unmittelbar aus dem bereits gezeigten:

- Angenommen, Typ-2-Sprachen wären unter Komplement abgeschlossen

Nichtabschluss unter Komplement

Satz: Es gibt eine kontextfreie Sprache $\mathbf{L}$, so dass $\overline{\mathbf{L}}$ keine kontextfreie Sprache ist.

Beweis: Die Behauptung folgt unmittelbar aus dem bereits gezeigten:

- Angenommen, Typ-2-Sprachen wären unter Komplement abgeschlossen
- Wir wissen bereits, dass Typ-2-Sprachen unter Vereinigung abgeschlossen sind

Nichtabschluss unter Komplement

Satz: Es gibt eine kontextfreie Sprache $\mathbf{L}$, so dass $\overline{\mathbf{L}}$ keine kontextfreie Sprache ist.

Beweis: Die Behauptung folgt unmittelbar aus dem bereits gezeigten:

- Angenommen, Typ-2-Sprachen wären unter Komplement abgeschlossen
- Wir wissen bereits, dass Typ-2-Sprachen unter Vereinigung abgeschlossen sind
- Laut den Gesetzen der Mengenlehre gilt:

$$\mathbf{L_1} \cap \mathbf{L_2} = \overline{\overline{\mathbf{L_1}} \cup \overline{\mathbf{L_2}}} \qquad \text{(De Morgan)}$$

Nichtabschluss unter Komplement

Satz: Es gibt eine kontextfreie Sprache $\mathbf{L}$, so dass $\overline{\mathbf{L}}$ keine kontextfreie Sprache ist.

Beweis: Die Behauptung folgt unmittelbar aus dem bereits gezeigten:

- Angenommen, Typ-2-Sprachen wären unter Komplement abgeschlossen
- Wir wissen bereits, dass Typ-2-Sprachen unter Vereinigung abgeschlossen sind
- Laut den Gesetzen der Mengenlehre gilt:

$$\mathbf{L_1} \cap \mathbf{L_2} = \overline{\overline{\mathbf{L_1}} \cup \overline{\mathbf{L_2}}} \qquad \text{(De Morgan)}$$

- Daraus folgt, dass Typ-2-Sprachen unter Schnitt abgeschlossen sind – Widerspruch □

Eine Warnung zum Nichtabschluss

Nichtabschluss unter Schnitt und Komplement bedeutet nicht, dass Schnitte bzw. Komplemente kontextfreier Sprachen grundsätzlich nicht kontextfrei sein können.

Eine Warnung zum Nichtabschluss

Nichtabschluss unter Schnitt und Komplement bedeutet nicht, dass Schnitte bzw. Komplemente kontextfreier Sprachen grundsätzlich nicht kontextfrei sein können.

Beispiel: Alle regulären Sprachen sind kontextfrei, aber ihre Schnitte und Komplemente sind weiterhin regulär, also auch kontextfrei.

Eine Warnung zum Nichtabschluss

Nichtabschluss unter Schnitt und Komplement bedeutet nicht, dass Schnitte bzw. Komplemente kontextfreier Sprachen grundsätzlich nicht kontextfrei sein können.

Beispiel: Alle regulären Sprachen sind kontextfrei, aber ihre Schnitte und Komplemente sind weiterhin regulär, also auch kontextfrei.

Beispiel: Das Komplement der nichtregulären Sprache $\mathbf{L} = \{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\}$ über dem Alphabet $\{\mathtt{a}, \mathtt{b}\}$ ist die Vereinigung der folgenden Sprachen:

$$\mathbf{L_1} = \{\mathtt{a}, \mathtt{b}\}^*\{\mathtt{ba}\}\{\mathtt{a}, \mathtt{b}\}^* \qquad \mathbf{L_2} = \{\mathtt{a}\}^+\mathbf{L} \qquad \mathbf{L_3} = \mathbf{L}\{\mathtt{b}\}^+$$

$\mathbf{L_1}$ ist regulär, also kontextfrei. $\mathbf{L_2}$ und $\mathbf{L_3}$ sind kontextfrei, da sie als Konkatenation zweier kontextfreier Sprachen entstehen. Die Vereinigung $\mathbf{L_1} \cup \mathbf{L_2} \cup \mathbf{L_3}$ ist also auch kontextfrei.

Ein nichtkontextfreies Komplement

Unser Widerspruchsbeweis zum Abschluss unter Komplement liefert uns noch kein konkretes Beispiel eines nichtkontextfreien Komplements.

Ein nichtkontextfreies Komplement

Unser Widerspruchsbeweis zum Abschluss unter Komplement liefert uns noch kein konkretes Beispiel eines nichtkontextfreien Komplements.

Wir können es aber aus den Beweisen ableiten:

- Seien $\mathbf{L_1} = \{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^k \mid i \geq 0, k \geq 0\}$ und $\mathbf{L_2} = \{\mathtt{a}^i\mathtt{b}^k\mathtt{c}^k \mid i \geq 0, k \geq 0\}$
- $\mathbf{L_1} \cap \mathbf{L_2} = \overline{\overline{\mathbf{L_1}} \cup \overline{\mathbf{L_2}}} = \{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\} = \mathbf{L}$ ist nicht kontextfrei

Ein nichtkontextfreies Komplement

Unser Widerspruchsbeweis zum Abschluss unter Komplement liefert uns noch kein konkretes Beispiel eines nichtkontextfreien Komplements.

Wir können es aber aus den Beweisen ableiten:

- Seien $\mathbf{L_1} = \{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^k \mid i \geq 0, k \geq 0\}$ und $\mathbf{L_2} = \{\mathtt{a}^i\mathtt{b}^k\mathtt{c}^k \mid i \geq 0, k \geq 0\}$
- $\mathbf{L_1} \cap \mathbf{L_2} = \overline{\overline{\mathbf{L_1}} \cup \overline{\mathbf{L_2}}} = \{\mathtt{a}^i\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\} = \mathbf{L}$ ist nicht kontextfrei
- $\overline{\mathbf{L_1}} \cup \overline{\mathbf{L_2}}$ ist die Vereinigung der folgenden Typ-2-Sprachen:
  - $\{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^* \circ \{\mathtt{ba}, \mathtt{ca}, \mathtt{cb}\} \circ \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^*$ (falsche Reihenfolge)
  - $\{\mathtt{a}\}^+ \circ \{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\} \circ \{\mathtt{c}\}^*$ (zu viele a für $\mathbf{L_1}$)
  - $\{\mathtt{a}^i\mathtt{b}^i \mid i \geq 0\} \circ \{\mathtt{b}\}^+ \circ \{\mathtt{c}\}^*$ (zu viele b für $\mathbf{L_1}$)
  - $\{\mathtt{a}\}^* \circ \{\mathtt{b}\}^+ \circ \{\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\}$ (zu viele b für $\mathbf{L_2}$)
  - $\{\mathtt{a}\}^* \circ \{\mathtt{b}^i\mathtt{c}^i \mid i \geq 0\} \circ \{\mathtt{c}\}^+$ (zu viele c für $\mathbf{L_2}$)
- $\overline{\mathbf{L_1}} \cup \overline{\mathbf{L_2}}$ ist daher kontextfrei, aber ihr Komplement nicht

Zusammenfassung und Ausblick

Kontextfreie Sprachen sind abgeschlossen unter Vereinigung, Konkatenation und Kleene-Stern

Kontextfreie Sprachen sind nicht abgeschlossen unter Komplement und Schnitt

Abschlüsse beruhen auf Grammatikoperationen, die man auf beliebige Grammatiken anwenden könnte, aber nur der Abschluss unter $\cup$ für Typ 0 und Typ 1 kann so gezeigt werden.

Offene Fragen:

- Haben kontextfreie Sprachen ein Berechnungsmodell?
- Welche Probleme auf kontextfreien Grammatiken kann man lösen?
- Was gibt es zu Typ 1 und Typ 0 zu sagen?