Theorie der Informatik

10. Kontextfreie Sprachen I

Malte Helmert Gabriele Röger

Universität Basel

30. März 2015

Fahrplan für dieses Kapitel

- Wiederholung kontextfreie Grammatiken
- Keine Einschränkung der Ausdrucksstärke durch Anforderung an ε-Regeln
- Normalform für kontextfreie Grammatiken
- Pumping Lemma analog zur regulären Sprachen, um zu zeigen, dass Sprache nicht kontextfrei ist
- Abschlusseigenschaften
- Entscheidbarkeit

Kontextfreie Grammatiken

Wiederholung: Kontextfreie Grammatiken

Definition (Kontextfreie Grammatik)

Eine kontextfreie Grammatik ist ein 4-Tupel (Σ, V, P, S) mit

1. Σ endliches Terminalalphabet,
2. V endliche Menge von Variablen (mit $V \cap \Sigma = \emptyset$),
3. $P \subseteq (V \times (V \cup \Sigma)^+) \cup \{(S, \varepsilon)\}$ endliche Menge von Regeln,
4. falls $S \to \varepsilon \in P$, sind alle anderen Regeln in $V \times ((V \setminus \{S\}) \cup \Sigma)^+$.
5. $S \in V$ Startvariable.

Wiederholung: Kontextfreie Grammatiken

Definition (Kontextfreie Grammatik)

Eine kontextfreie Grammatik ist ein 4-Tupel (Σ, V, P, S) mit

1. Σ endliches Terminalalphabet,
2. V endliche Menge von Variablen (mit $V \cap \Sigma = \emptyset$),
3. $P \subseteq (V \times (V \cup \Sigma)^+) \cup \{(S, \varepsilon)\}$ endliche Menge von Regeln,
4. falls $S \to \varepsilon \in P$, sind alle anderen Regeln in $V \times ((V \setminus \{S\}) \cup \Sigma)^+$.
5. $S \in V$ Startvariable.

Regel $X \to \varepsilon$ ist nur erlaubt, wenn $X = S$ und S in keiner rechten Regelseite vorkommt.

Wiederholung: Kontextfreie Grammatiken

Definition (Kontextfreie Grammatik)

Eine kontextfreie Grammatik ist ein 4-Tupel (Σ, V, P, S) mit

1. Σ endliches Terminalalphabet,
2. V endliche Menge von Variablen (mit $V \cap \Sigma = \emptyset$),
3. $P \subseteq (V \times (V \cup \Sigma)^+) \cup \{(S, \varepsilon)\}$ endliche Menge von Regeln,
4. falls $S \to \varepsilon \in P$, sind alle anderen Regeln in $V \times ((V \setminus \{S\}) \cup \Sigma)^+$.
5. $S \in V$ Startvariable.

Regel $X \to \varepsilon$ ist nur erlaubt, wenn $X = S$ und S in keiner rechten Regelseite vorkommt.

Bei regulären Grammatiken konnte man diese Restriktion aufweichen. Wie ist das bei kontextfreien Grammatiken?

Epsilon-Regeln

Satz

Für jede Grammatik G mit Regeln $P \subseteq V \times (V \cup \Sigma)^*$ gibt es eine kontextfreie Grammatik G' mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Epsilon-Regeln

Satz

Für jede Grammatik G mit Regeln $P \subseteq V \times (V \cup \Sigma)^*$ gibt es eine kontextfreie Grammatik G' mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Beweis.

Sei $G = (\Sigma, V, P, S)$ Grammatik mit $P \subseteq V \times (V \cup \Sigma)^*$.

Sei $V_\varepsilon = \{A \in V \mid A \Rightarrow^* \varepsilon\}$. Wir finden diese Menge V_ε indem wir zunächst alle Variablen A mit Regel $A \to \varepsilon \in P$ aufsammeln und dann sukkzessive weitere Variablen B hinzufügen für die es eine Regel $B \to A_1 A_2 \ldots A_k \in P$ gibt, so dass die Variablen A_i bereits für alle $1 \leq i \leq k$ in der Menge sind. . . .

Epsilon-Regeln

Satz

Für jede Grammatik G mit Regeln $P \subseteq V \times (V \cup \Sigma)^*$ gibt es eine kontextfreie Grammatik G' mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Beweis (Fortsetzung).

Sei S die Startvariable von G. Sei P' die Regelmenge, die aus P hervorgeht, indem man alle Regeln der Form $A \to \varepsilon$ entfernt. Zudem fügt man für jede Regel der Form $B \to w$ mit $B \in V, w \in (V \cup \Sigma)^+$ neue Regeln hinzu, die alle Möglichkeiten Variablen in w auf ε abzuleiten vorwegnehmen: sei I_ε die Menge der Stellen, an denen w eine Variable $A \in V_\varepsilon$ enthält. Dann fügt man für jede nicht-leere Menge $I' \subseteq I_\varepsilon$ eine neue Regel $B \to w'$ hinzu, wobei w' aus w hervorgeht, indem die Variablen an allen Stellen in I' weggelassen werden. ...

Epsilon-Regeln

Satz

Für jede Grammatik G mit Regeln $P \subseteq V \times (V \cup \Sigma)^*$ gibt es eine kontextfreie Grammatik G' mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Beweis (Fortsetzung).

Damit ist $\mathcal{L}(G) \setminus \{\varepsilon\} = \mathcal{L}((\Sigma, V, P', S))$ und P' enthält keine Regel $A \to \varepsilon$. Ist das Startsymbol S von G nicht in V_ε, sind wir fertig.

□

Epsilon-Regeln

Satz

Für jede Grammatik G mit Regeln $P \subseteq V \times (V \cup \Sigma)^*$ gibt es eine kontextfreie Grammatik G' mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Beweis (Fortsetzung).

Damit ist $\mathcal{L}(G) \setminus \{\varepsilon\} = \mathcal{L}((\Sigma, V, P', S))$ und P' enthält keine Regel $A \to \varepsilon$. Ist das Startsymbol S von G nicht in V_ε, sind wir fertig.

Sonst sei S' neue Variable und P'' gehe aus P' hervor, indem man

1. alle Vorkommen von S in rechten Regelseiten durch S' ersetzt,
2. für jede Regel $S \to w$ die Regel $S' \to w$ hinzufügt, und
3. die Regel $S \to \varepsilon$ hinzufügt.

Dann ist $\mathcal{L}(G) = \mathcal{L}((\Sigma, V \cup \{S'\}, P'', S))$. □

Fragen

Fragen?

Chomsky-Normalform: Motivation

- Es gibt keine Normalform, die für kontextfreie Sprachen eindeutig ist.
- Chomsky-Normalform liefert aber interessante Eigenschaften für Ableitungen von Wörtern in einer Sprache.

Chomsky-Normalform: Definition

Definition (Chomsky-Normalform)

Eine kontextfreie Grammatik G heisst in Chomsky-Normalform (CNF), falls alle Regeln eine der folgenden drei Formen haben:

- $A \to BC$ mit A, B, C Variablen, oder
- $A \to a$ mit A Variable, a Terminalsymbol, oder
- $S \to \varepsilon$ mit S Startvariable.

Alternativ: Regelmenge $P \subseteq (V \times (VV \cup \Sigma)) \cup \{(S, \varepsilon)\}$

Chomsky-Normalform: Satz

Satz

Für jede kontextfreie Grammatik G gibt es eine kontextfreie Grammatik in Chomsky-Normalform mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Chomsky-Normalform: Satz

Satz

Für jede kontextfreie Grammatik G gibt es eine kontextfreie Grammatik in Chomsky-Normalform mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Beweis.

Das folgende Verfahren bringt die Regelmenge von G in CNF:

Schritt 1: Eliminiere Regeln der Form $A \to B$ mit A, B Variablen.

Gibt es Mengen von Variablen $\{B_1, \ldots, B_k\}$ mit Regeln $B_1 \to B_2, B_2 \to B_3, \ldots, B_{k-1} \to B_k, B_k \to B_1$, so ersetze diese Variablen durch eine neue Variable B.

Dann benenne alle Variablen so in $V = \{A_1, \ldots, A_n\}$ um, dass aus $A_i \to A_j \in P$ folgt, dass $i < j$. Für $k = n-1, \ldots, 1$: Eliminiere alle Regeln der Form $A_k \to A_{k'}$ mit $k' > k$ und füge für jede Regel $A_{k'} \to w$ mit $w \in (V \cup \Sigma)^+$ eine Regel $A_k \to w$ hinzu. ...

Chomsky-Normalform: Satz

Satz

Für jede kontextfreie Grammatik G gibt es eine kontextfreie Grammatik in Chomsky-Normalform mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Beweis (Fortsetzung).

Schritt 2: Eliminiere Regeln mit Terminalzeichen auf rechter Seite, die nicht die Form $A \to a$ haben.

Führe für jedes Terminalzeichen $a \in \Sigma$ eine neue Variable A_a und die Regel $A_a \to a$ ein.

Ersetze in allen Regeln, die nicht die Form $A \to a$ haben, alle Terminalzeichen durch die entsprechenden neu eingeführten Variablen. ...

Chomsky-Normalform: Satz

Satz

Für jede kontextfreie Grammatik G gibt es eine kontextfreie Grammatik in Chomsky-Normalform mit $\mathcal{L}(G) = \mathcal{L}(G')$.

Beweis (Fortsetzung).

Schritt 3: Eliminiere Regeln der Form $A \to B_1 B_2 \ldots B_k$ mit $k > 2$

Führe für jede Regel der Form $A \to B_1 B_2 \ldots B_k$ mit $k > 2$ neue Variablen $C_2, \ldots, C_{k-1}$ ein und ersetze die Regel durch

$$\begin{aligned} A &\to B_1 C_2 \\ C_2 &\to B_2 C_3 \\ &\vdots \\ C_{k-1} &\to B_{k-1} B_k \end{aligned}$$

$\square$

Chomsky-Normalform: Länge von Ableitungen

Beobachtung

Sei G eine Grammatik in Chomsky-Normalform und $w \in \mathcal{L}(G)$ ein nicht-leeres Wort, das von G erzeugt wird. Dann wird w in genau $2|w| - 1$ Ableitungsschritten erzeugt.

Beweis.

$\rightsquigarrow$ Übung $\square$

Ableitungsbaum: allgemein

Definition

Sei G eine kontextfreie Grammatik und $S \Rightarrow w_1 \Rightarrow w_2 \Rightarrow \ldots w_n$ eine Ableitung für ein nicht-leeres Wort $w_n \in \mathcal{L}(G)$. Der Ableitungsbaum T für diese Ableitung ist folgenderweise aufgebaut:

- Der Wurzel des Baumes ist die Startvariable S zugeordnet.
- Wird im i-ten Ableitungsschritt die Variable A durch Wort z ersetzt, hat der entsprechende A-Knoten $|z|$ Kinder, denen die einzelnen Zeichen von z (in der gleichen Reihenfolge) zugeordnet sind.

Den Blättern des resultierenden Baumes sind dann genau die Zeichen aus w_n zugeordnet.

Beispiel: Tafel

Chomsky-Normalform: Ableitungsbaum

Beobachtung

Sei G eine Grammatik in Chomsky-Normalform und $w \in \mathcal{L}(G)$ ein nicht-leeres Wort, das von G erzeugt wird. Jeder Ableitungsbaum von w ist – bis auf die Ableitung der Blätter – ein Binärbaum.

In einer Ableitung eines nicht-leeren Wortes können nur Formeln der Form $A \rightarrow BC$ und der Form $A \rightarrow a$ angewendet werden. Im ersten Fall bekommt der Knoten für A genau zwei Kinder und erfüllt damit die Binärbaumeigenschaft. Im zweiten Fall wird ein Terminalzeichen und damit ein Blatt abgeleitet.

Fragen

Fragen?

Pumping Lemma

Pumping Lemma für kontextfreie Sprachen

Mit dem Pumping Lemma aus dem vorherigen Kapitel konnten wir für Sprachen zeigen, das sie nicht regulär sind. Gibt es so ein Lemma auch für kontextfreie Sprachen?

Foto mit freundlicher Genehmigung von imagerymajestic / FreeDigitalPhotos.net

Pumping Lemma für kontextfreie Sprachen

Mit dem Pumping Lemma aus dem vorherigen Kapitel konnten wir für Sprachen zeigen, das sie nicht regulär sind. Gibt es so ein Lemma auch für kontextfreie Sprachen?

Ja!

Pumping Lemma

Satz (Pumping Lemma)

Sei L eine kontextfreie Sprache. Dann gibt es ein $n \in \mathbb{N}$ (eine Pumpingzahl für L), so dass sich alle Wörter $z \in L$ mit $|z| \geq n$ zerlegen lassen in $z = uvwxy$, so dass folgendes gilt:

1. $|vx| \geq 1$,
2. $|vwx| \leq n$, *und*
3. *für alle* $i = 0, 1, 2, \ldots$ *gilt:* $uv^iwx^iy \in L$.

Pumping Lemma

Satz (Pumping Lemma)

Sei L eine kontextfreie Sprache. Dann gibt es ein $n \in \mathbb{N}$ (eine Pumpingzahl für L), so dass sich alle Wörter $z \in L$ mit $|z| \geq n$ zerlegen lassen in $z = uvwxy$, so dass folgendes gilt:

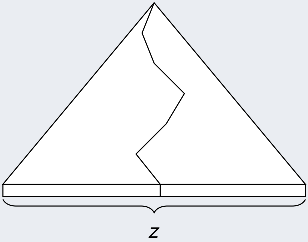
1. $|vx| \geq 1$,
2. $|vwx| \leq n$, *und*
3. *für alle* $i = 0, 1, 2, \ldots$ *gilt:* $uv^iwx^iy \in L$.

Beweis.

Sei $G = (\Sigma, V, P, S)$ Grammatik mit $\mathcal{L}(G) = L$ in Chomsky-Normalform.
Wir zeigen, dass $n = 2^{|V|}$ eine Pumpingzahl von L ist. Betrachte hierzu ein beliebiges Wort $z \in L$ mit $|z| \geq n$. ...

Pumping Lemma

Beweis (Fortsetzung).

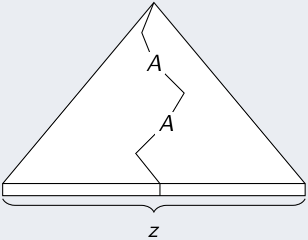


Da jeder Ableitungsbaum für eine Ableitung von z bis auf die Ableitung der Terminalsymbole ein Binärbaum ist und der Baum $\geq n$ Blätter hat, hat mindestens einen Pfad von der Wurzel zu einem Blatt Länge $\geq |V|$.

Auf solch einem Pfad befinden sich mindestens $|V| + 1$ Variablen, es muss also mindestens eine Variable doppelt vorkommen. ...

Pumping Lemma

Beweis (Fortsetzung).

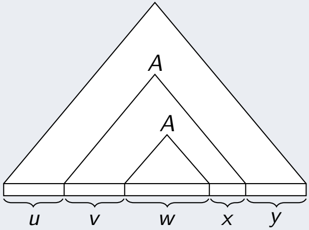


Wir wählen solch ein Doppelvorkommen, indem wir von unten nach oben einem Pfad der Länge $\geq |V|$ folgen bis eine Variable zum zweiten Mal vorkommt.

Die obere der beiden Variablen kann dann höchstens $|V| + 1$ Schritte von der Blattebene entfernt sein. ...

Pumping Lemma

Beweis (Fortsetzung).

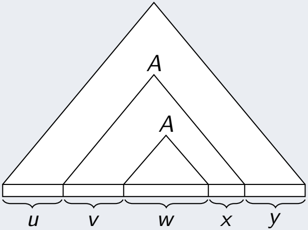


Die Teilbäume unter den Variablen definieren dann die Aufteilung $z = uvwxy$ von z in Teilwörter. Wir zeigen, dass die Aufteilung die geforderten drei Bedingungen erfüllt.

...

Pumping Lemma

Beweis (Fortsetzung).



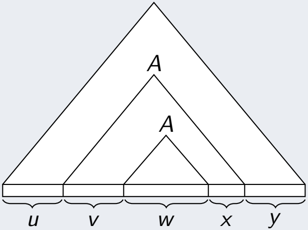
Die Teilbäume unter den Variablen definieren dann die Aufteilung $z = uvwxy$ von z in Teilwörter. Wir zeigen, dass die Aufteilung die geforderten drei Bedingungen erfüllt.

Da die obere Variable mit einer Regel der Form $A \to BC$ weiter abgeleitet wird und wegen der Restriktionen an die Verwendung von ε in G, ist v oder x (oder beide) nicht leer: Entweder ist v aus B abgeleitet oder x geht aus C hervor. Damit folgt die erste Bedingung $|vx| \geq 1$.

...

Pumping Lemma

Beweis (Fortsetzung).



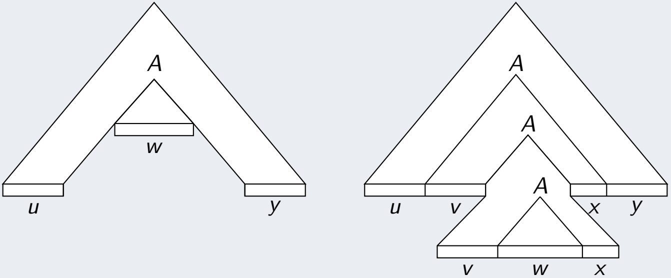
Die Teilbäume unter den Variablen definieren dann die Aufteilung $z = uvwxy$ von z in Teilwörter. Wir zeigen, dass die Aufteilung die geforderten drei Bedingungen erfüllt.

Da die obere Variable mit einer Regel der Form $A \to BC$ weiter abgeleitet wird und wegen der Restriktionen an die Verwendung von ε in G, ist v oder x (oder beide) nicht leer: Entweder ist v aus B abgeleitet oder x geht aus C hervor. Damit folgt die erste Bedingung $|vx| \geq 1$.

Da die obere Variable höchstens $|V| + 1$ Schritte von jedem Blatt entfernt ist, kann das abgeleitete Wort vwx höchstens Länge $2^{|V|}$ haben. Es gilt also auch die zweite Bedingung $|vwx| \leq n$. ...

Pumping Lemma

Beweis (Fortsetzung).



Die Ableitungsfolge kann aufgrund des Doppelvorkommens der Variable so modifiziert werden, dass man für jedes $i \geq 0$ das Wort uv^iwx^iy ableiten kann. Damit ist das Wort in der Sprache und auch die dritte Bedingung erfüllt. $\square$

Pumping Lemma: Beispiel

Beispiel

Die Sprache $L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ ist nicht kontextfrei.

Beweis.

Angenommen L ist kontextfrei. Dann sei p eine Pumpingzahl für L.

Das Wort $z = a^p b^p c^p$ ist in L und hat Länge $\geq p$. Sei $z = uvwxy$ eine Zerlegung mit den Eigenschaften aus dem PL.

Dann ist auch das Wort $z' = uv^2wx^2y$ in L. Da $|vwx| \leq p$ kann vwx nicht alle Symbole a, b und c enthalten und es bleibt die Anzahl mindestens eines Symbols unverändert. Wegen $|vx| \geq 1$ enthält z' aber für mindestens ein Symbol eine grössere Anzahl als das Wort z. Da daher in z' nicht mehr alle Symbole gleich oft vorkommen, ist z' nicht in L.

$\leadsto$ Widerspruch

$\leadsto$ L nicht kontextfrei. $\square$

Fragen

Fragen?

Abschlusseigenschaften

Abschlusseigenschaften der kontextfreien Sprachen

Satz

Die kontextfreien Sprachen sind abgeschlossen unter:

- *Vereinigung*
- *Produkt*
- *Stern*

Abschlusseigenschaften der kontextfreien Sprachen

Satz

Die kontextfreien Sprachen sind abgeschlossen unter:

- *Vereinigung*
- *Produkt*
- *Stern*

Beweis.

Abschluss unter Vereinigung:
Seien $G_1 = (\Sigma, V_1, P_1, S_1)$ und $G_2 = (\Sigma, V_2, P_2, S_2)$ kontextfreie Grammatiken mit $V_1 \cap V_2 = \emptyset$. Dann ist $G = (\Sigma, V_1 \cup V_2 \cup \{S\}, P_1 \cup P_2 \cup \{S \to S_1, S \to S_2\}, S)$ (wobei $S \notin (V_1 \cup V_2)$) eine kontextfreie Grammatik für die Vereinigungsmenge (evtl. Umformungen für ε-Regeln notwendig).

...

Abschlusseigenschaften der kontextfreien Sprachen

Satz

Die kontextfreien Sprachen sind abgeschlossen unter:

- *Vereinigung*
- *Produkt*
- *Stern*

Beweis (Fortsetzung).

Abschluss unter Produkt:
Seien $G_1 = (\Sigma, V_1, P_1, S_1)$ und $G_2 = (\Sigma, V_2, P_2, S_2)$ kontextfreie Grammatiken mit $V_1 \cap V_2 = \emptyset$. Dann ist $G = (\Sigma, V_1 \cup V_2 \cup \{S\}, P_1 \cup P_2 \cup \{S \to S_1 S_2\}, S)$ (wobei $S \notin V_1 \cup V_2$) eine kontextfreie Grammatik für das Produkt (evtl. Umformungen für ε-Regeln notwendig). ...

Abschlusseigenschaften der kontextfreien Sprachen

Satz

Die kontextfreien Sprachen sind abgeschlossen unter:

- *Vereinigung*
- *Produkt*
- *Stern*

Beweis (Fortsetzung).

Abschluss unter Sternoperation:
Sei $G_1 = (\Sigma, V_1, P_1, S_1)$ eine kontextfreie Grammatiken bei der o.B.d.A. S_1 auf keiner rechten Regelseite vorkommt. Dann ist $G = (\Sigma, V_1 \cup \{S\}, P, S)$ mit $S \notin (V_1 \cup V_2)$ und $P = (P_1 \cup \{S \to \varepsilon, S \to S_1, S_1 \to S_1 S_1\}) \setminus \{S_1 \to \varepsilon\}$ eine kontextfreie Grammatik für $\mathcal{L}(G_1)^*$. $\square$

Abschlusseigenschaften der kontextfreien Sprachen

Satz

Die kontextfreien Sprachen sind nicht abgeschlossen unter:

- *Schnitt*
- *Komplement*

Abschlusseigenschaften der kontextfreien Sprachen

Satz

Die kontextfreien Sprachen sind nicht abgeschlossen unter:

- *Schnitt*
- *Komplement*

Beweis.

Kein Abschluss unter Schnitt:
Die Sprachen $L_1 = \{a^i b^j c^j \mid i, j > 0\}$ und $L_2 = \{a^j b^i c^j \mid i, j > 0\}$ sind kontextfrei.
Zum Beispiel ist $G_1 = (\{a, b, c\}, \{S, A, B\}, P, S)$ mit $P = \{S \to AB, A \to a, A \to aA, B \to bc, B \to bBc\}$ eine kontextfreie Grammatik für L_1.
Wir haben vorhin aber gesehen, dass $L_1 \cap L_2 = \{a^i b^i c^i \mid i > 0\}$ nicht kontextfrei ist. ...

Abschlusseigenschaften der kontextfreien Sprachen

Satz

Die kontextfreien Sprachen sind nicht abgeschlossen unter:

- *Schnitt*
- *Komplement*

Beweis (Fortsetzung).

Kein Abschluss unter Komplement:
Wären die kontextfreien Sprachen unter Vereinigung und Komplement abgeschlossen, müssten sie wegen des Zusammenhangs

$$L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$$

ebenfalls unter Schnitt abgeschlossen sein. Wir haben aber bereits gezeigt, dass sie unter Vereinigung abgeschlossen, unter Schnitt jedoch nicht abgeschlossen sind. □

Fragen

Fragen?

Entscheidbarkeit

Entscheidbarkeit: Wortproblem

Satz (Wortproblem ist entscheidbar)

Das Wortproblem $\mathrm{P}_{\in}$ für kontextfreie Sprachen ist:

Gegeben: *Kontextfreie Grammatik G mit Terminalalphabet Σ und Wort $w \in \Sigma^*$*

Gefragt: *Ist $w \in \mathcal{L}(G)$?*

Das Wortproblem ist entscheidbar.

Beweis.

Falls $w = \varepsilon$, ist $w \in \mathcal{L}(G)$ gdw. $S \to \varepsilon$ mit Startvariable S Regel von G ist.
Da für alle anderen Regeln $w_l \to w_r$ von G gilt, dass $|w_l| \leq |w_r|$, können die Zwischenergebnisse im Laufe einer Ableitung eines nicht-leeren Wortes nicht kürzer werden. Man kann also systematisch alle (endlich vielen) Ableitungen von Wörtern bis zur Länge $|w|$ durchgehen und prüfen, ob sie das Wort w ableiten. □

Entscheidbarkeit: Leerheitsproblem

Satz (Leerheitsproblem ist entscheidbar)

Das Leerheitsproblem $\mathrm{P}_\emptyset$ *für kontextfreie Sprachen ist:*

Gegeben: *Kontextfreie Grammatik* G
Gefragt: *Ist* $\mathcal{L}(G) = \emptyset$*?*

Das Leerheitsproblem ist entscheidbar.

Beweis.

Man konstruiert zunächst eine Grammatik G' in CNF, die $\mathcal{L}(G)$ erzeugt. In G' bestimmt man alle Variablen, von denen Terminalwörter ableitbar sind: Man markiert zunächst alle Variablen A, für die es eine Regel der Form $A \to a$ gibt. Danach markiert man sukzessive alle Variablen A, für die es eine Regel $A \to BC$ gibt, so dass B und C bereits markiert sind.

$\mathcal{L}(G)$ ist leer gdw. die Startvariable nicht markiert ist. $\square$

Entscheidbarkeit: Endlichkeitsproblem

Satz (Endlichkeitsproblem ist entscheidbar)

Das Endlichkeitsproblem P_{∞} für kontextfreie Sprachen ist:

Gegeben: *Kontextfreie Grammatik G*

Gefragt: *Ist $|\mathcal{L}(G)| < \infty$?*

Das Endlichkeitsproblem ist entscheidbar.

Beweis.

Man konstruiert zunächst eine Grammatik G' in CNF, die $\mathcal{L}(G)$ erzeugt. Wir wissen, dass $n = 2^{|V'|}$ (wobei V' Variablenmenge von G' ist) eine Pumpingzahl für $\mathcal{L}(G)$ ist.

Wir zeigen, dass $|\mathcal{L}(G)| = \infty$ genau dann gilt, wenn es ein $w \in \mathcal{L}(G)$ gibt mit $n \leq |w| < 2n$:

$\Leftarrow$: Falls $w \in \mathcal{L}(G)$ mit $n \leq |w|$, dann gilt wg. PL $|\mathcal{L}(G)| = \infty$. . . .

Entscheidbarkeit: Endlichkeitsproblem

Satz (Endlichkeitsproblem ist entscheidbar)

Das Endlichkeitsproblem P_{∞} für kontextfreie Sprachen ist:

Gegeben: *Kontextfreie Grammatik G*

Gefragt: *Ist $|\mathcal{L}(G)| < \infty$?*

Das Endlichkeitsproblem ist entscheidbar.

Beweis (Fortsetzung).

$\Rightarrow$: Falls $\mathcal{L}(G) = \infty$, sei z ein kürzestes Wort in $\mathcal{L}(G)$ mit mindestens Länge $|z| \geq n$. Wenn $|z| \geq 2n$, dann kann z laut PL zerlegt werden, in $z = uvwxy$, so dass $z_0 = uwy \in \mathcal{L}(G)$, wobei $n \leq |z_0| < |z|$.

$\rightsquigarrow$ Widerspruch zur Minimalität der Länge von z. □

Unentscheidbarkeit: Äquivalenz- und Schnittproblem

Satz

Das Äquivalenzproblem und das Schnittproblem sind für kontextfreie Sprachen nicht entscheidbar.

Wir können das mit den uns zur Verfügung stehenden Mitteln noch nicht beweisen. Wir werden das jedoch im dritten Vorlesungsteil über Berechenbarkeitstheorie nachholen.

Fragen

Fragen?

Zusammenfassung

Zusammenfassung

- Für jede kontextfreie Sprache gibt es eine Grammatik in Chomsky-Normalform.
- Bei Grammatiken in Chomsky-Normalform sind Ableitungsbäume (fast) Binärbäume.
- Mit dem Pumping Lemma für kontextfreie Sprachen kann man zeigen, dass eine Sprache nicht kontextfrei ist.
- Die kontextfreien Sprachen sind unter Vereinigung, Produkt und Stern abgeschlossen.
- Die kontextfreien Sprachen sind unter Schnitt und Komplement nicht abgeschlossen.
- Das Wort-, das Leerheits- und das Endlichkeitsproblem sind für die kontextfreien Sprachen entscheidbar.
- Das Äquivalenz- und das Schnittproblem sind für die kontextfreien Sprachen nicht entscheidbar.