Technische Universität Braunschweig

Theoretische Informatik 2

Arne Schmidt

Unentscheidbare Sprachen

Nicht-trivial, nicht-monoton

Eigenschaften der von Turing-Maschinen erkannten Sprachen

Semi-entscheidbare Sprachen

PCP

HP

Totality

Accept

(Many-One-) Reduktionen

Kapitel 6 – Unentscheidbare Probleme kontextfreier Sprachen

Unterschied kontextfrei und kontextsensitiv

Betrachte folgende Grammatik $G = (\{S, B, C\}, \{a, b, c\}, P, S)$ mit folgenden Produktionsregeln.

1. $S \rightarrow aSBC$
2. $S \rightarrow aBC$
3. $CB \rightarrow BC$
4. $aB \rightarrow ab$
5. $bB \rightarrow bb$
6. $bC \rightarrow bc$
7. $cC \rightarrow cc$

$\{a^n b^n c^n \mid n \geq 1\}$

Kontextsensitiv

1. $S \rightarrow aSBC$
2. $S \rightarrow aBC$
3. $B \rightarrow b$
4. $B \rightarrow Bb$
5. $C \rightarrow Bc$
6. $C \rightarrow c$

$\{a^n (b^+ c)^n \mid n \geq 1\}$

Kontextfrei

Entscheidbare Probleme

Wortproblem für kontextfreie Sprachen (WORD-CFL)
Gegeben: Kontextfreie Grammatik G, Wort w.
Frage: Gilt $w \in \mathcal{L}(G)$?

CYK-Algorithmus

Leerheit kontextfreie Sprachen (EMPTINESS-CFL)
Gegeben: Kontextfreie Grammatik G.
Frage: Gilt $\mathcal{L}(G) = \emptyset$?

Unendlichkeit kontextfreie Sprachen
Gegeben: Kontextfreie Grammatik G.
Frage: Gilt $\mathcal{L}(G)$ unendlich?

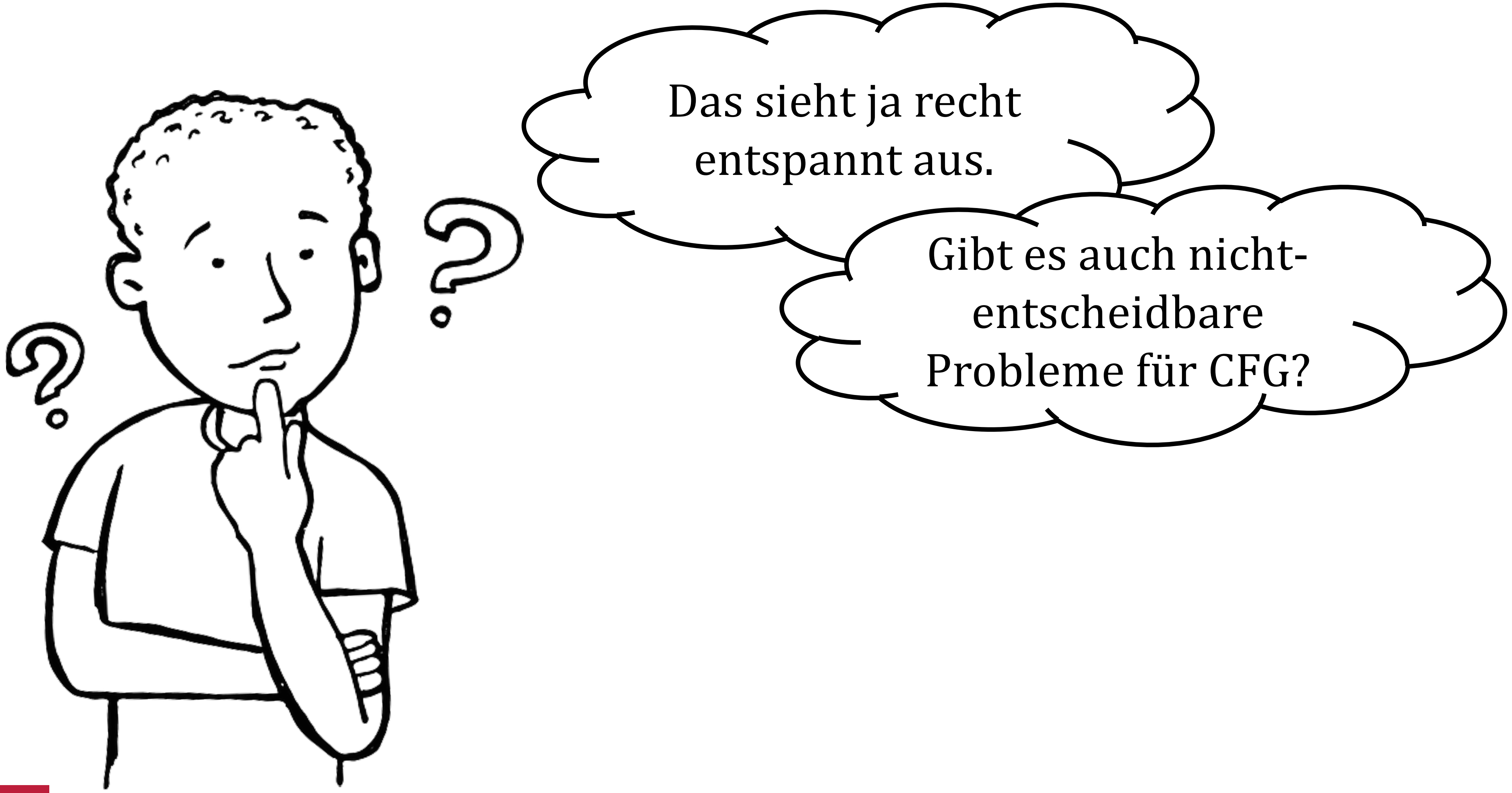
Reguläre Inklusion kontextfreier Sprachen

Gegeben: Kontextfreie Grammatik G, NFA A.
Frage: Gilt $\mathcal{L}(G) \subseteq \mathcal{L}(A)$?

NFA: Nicht-det. Endlicher Automat

Dieses Problem ist entscheidbar!

- Konstruiere NFA $\bar{A}$ mit $\mathcal{L}(\bar{A}) = \Sigma^* \setminus \mathcal{L}(A)$, sowie Grammatik G' mit $\mathcal{L}(G') = \mathcal{L}(G) \cap \mathcal{L}(\bar{A})$
- G' ist kontextfrei (CFG ist unter reg. Schnitt abgeschlossen)
- $\mathcal{L}(G) \subseteq \mathcal{L}(A) \Leftrightarrow \mathcal{L}(G') = \emptyset$
- $\mathcal{L}(G') = \emptyset$ ist entscheidbar!
- $\Rightarrow$ Problem entscheidbar.


Das sieht ja recht entspannt aus.
Gibt es auch nicht-entscheidbare Probleme für CFG?

Schnitt Probleme

INTERSECTION-EMPTINESS-CFL
Gegeben: Kontextfreie Grammatiken G_1, G_2.
Frage: Gilt $\mathcal{L}(G_1) \cap \mathcal{L}(G_2) = \emptyset$?

Unendlichkeit des Schnitts CFLs
Gegeben: Kontextfreie Grammatiken G_1, G_2.
Frage: Gilt $\mathcal{L}(G_1) \cap \mathcal{L}(G_2)$ unendlich?

Kontextfreiheit des Schnitts CFLs
Gegeben: Kontextfreie Grammatiken G_1, G_2.
Frage: Ist $\mathcal{L}(G_1) \cap \mathcal{L}(G_2)$ kontextfrei?

Unentscheidbare Probleme

Theorem 6.1
Alle nicht-trivialen Eigenschaften des Schnitts kontextfreier Sprachen sind unentscheidbar.
Insbesondere:
a) Schnittleerheit kontextfreier Sprachen (INTERSECTION-EMPTINESS-CFL)
b) Unendlichkeit des Schnitts kontextfreier Sprachen
c) Kontextfreiheit des Schnitts kontextfreier Sprachen

Beweisskizze:
Reduziere 0-1-PCP auf die drei Probleme.

Danach argumentieren wir, warum die Aussage für alle nicht-trivialen Eigenschaften gelten muss

0-1-PCP ≤ INTERSECTION-EMPTINESS-CFL

Sei $K = (x_1, y_1), \dots, (x_k, y_k)$ die PCP-Instanz mit $x_i, y_i \in \{0,1\}^*$ für alle i.
Betrachte folgende zwei Grammatiken G_1, G_2 mit Terminalalphabet $\{0,1,\#, z_1, \dots, z_k\}$

Produktionsregeln G_1

$$
\begin{aligned}
S &\to A\#B\ , \\
A &\to z_1 A x_1 \mid \dots \mid z_k A x_k\ , \\
A &\to z_1 x_1 \mid \dots \mid z_k x_k\ , \\
B &\to y_1^{\text{reverse}} B z_1 \mid \dots \mid y_k^{\text{reverse}} B z_k \\
B &\to y_1^{\text{reverse}} z_1 \mid \dots \mid y_k^{\text{reverse}} z_k\ .
\end{aligned}
$$

Produktionsregeln G_2

$$
\begin{aligned}
S &\to z_1 S z_1 \mid \dots \mid z_k S z_k \mid T \\
T &\to 0T0 \mid 1T1 \mid \#\ .
\end{aligned}
$$

$$\mathcal{L}(G_1) = \{z_{i_n} \dots z_{i_1} x_{i_1} \dots x_{i_n} \# y_{j_m}^{\text{reverse}} \dots y_{j_1}^{\text{reverse}} z_{j_m} \dots z_{j_1} \mid n, m \geq 1, \forall s,t\colon i_s, j_t \in \{1, \dots, k\}\}$$

$$\mathcal{L}(G_2) = \{u\ v\ \# v^{\text{reverse}} u^{\text{reverse}} \mid v \in \{0,1\}^*, u \in \{z_1, \dots, z_k\}^*\}$$

0-1-PCP $\leq$ INTERSECTION-EMPTINESS-CFL

Sei $K = (x_1, y_1), \dots, (x_k, y_k)$ die PCP-Instanz mit $x_i, y_i \in \{0,1\}^*$ für alle i.
Betrachte folgende zwei Grammatiken G_1, G_2 mit Terminalalphabet $\{0,1,\#, z_1, \dots, z_k\}$

$$\mathcal{L}(G_1) = \left\{ z_{i_n} \dots z_{i_1} x_{i_1} \dots x_{i_n} \# y_{j_m}^{\text{reverse}} \dots y_{j_1}^{\text{reverse}} z_{j_m} \dots z_{j_1} \;\middle|\; n, m \geq 1, \forall s, t\colon i_s, j_t \in \{1, \dots, k\} \right\}$$

$$\mathcal{L}(G_2) = \{ u\ v\ \# v^{\text{reverse}} u^{\text{reverse}} \mid v \in \{0,1\}^*, u \in \{z_1, \dots, z_k\}^* \}$$

Man erkennt, dass $\mathcal{L}(G_1) \cap \mathcal{L}(G_2)$ genau dann nicht-leer ist, wenn K eine nicht-leere Lösung $\ell_1, \dots, \ell_n$ besitzt. Dann existiert nämlich das Wort:

$$w = z_{\ell_n} \dots z_{\ell_1} x_{\ell_1} \dots x_{\ell_n} \# y_{\ell_n}^{\text{reverse}} \dots y_{\ell_1}^{\text{reverse}} z_{\ell_n} \dots z_{\ell_1}$$

Damit ist Teil a) gezeigt.

0-1-PCP $\leq$ Unendlichkeit des Schnitts

Zunächst:
Gibt es eine Lösung $\ell \coloneqq \ell_1, \dots, \ell_n$ für die PCP-Instanz, dann sind $\{\, \ell^k \mid k \geq 1 \,\}$ Lösungen für die PCP-Instanz.

$\Rightarrow \mathcal{L}(G_1) \cap \mathcal{L}(G_2)$ ist unendlich mit G_1, G_2 aus vorheriger Reduktion.

Andersherum: Gibt es keine Lösung, dann ist der Schnitt leer!

Damit ist Teil b) gezeigt.

0-1-PCP $\leq$ Kontextfreiheit des Schnitts

Zunächst:
Gibt es keine Lösung, dann ist der Schnitt leer.
Das ist eine kontextfreie Sprache!

Anderherum:
Gibt es eine Lösung für die PCP Instanz, dann kann das Pumping-Lemma für kontextfreie Sprachen genutzt werden, um zu zeigen, dass $\mathcal{L}(G_1) \cap \mathcal{L}(G_2)$ nicht kontextfrei ist.

Damit ist Teil c) gezeigt.

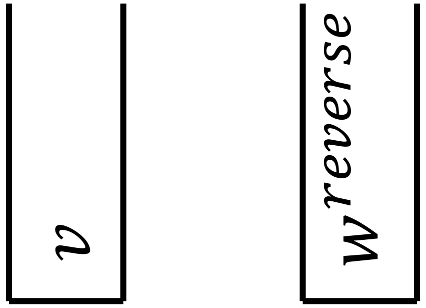
Schnitt von zwei CFL

Betrachte Sprachen $\mathcal{L}_1, \mathcal{L}_2$, welche von Pushdown-Automaten P_1 bzw. P_2 erkannt werden.
Der Schnitt $\mathcal{L}_1 \cap \mathcal{L}_2$ entspricht dem Produkt $P_1 \times P_2$ der beiden Automaten.

Dies entspricht einem Multipushdown-Automaten! Er kann in diesem Fall zwei Stacks unabhängig nutzen.

> Multipushdown-Automaten sind gleichmächtig wie Turing-Maschinen.

Konfiguration $v\ q\ w$ $\longrightarrow$



Stack 1 Stack 2

Für eine TM M existieren also zwei Sprachen $\mathcal{L}(P_1), \mathcal{L}(P_2)$ mit $\mathcal{L}(P_1) \cap \mathcal{L}(P_2) = \mathcal{L}(M)$.

Satz von Rice liefert zu zeigende Behauptung

Determinismus vs. Nicht-Determinismus

Macht es einen Unterschied, wenn wir uns auf CFL restriktieren, die von det. PDA akzeptiert werden müssen?

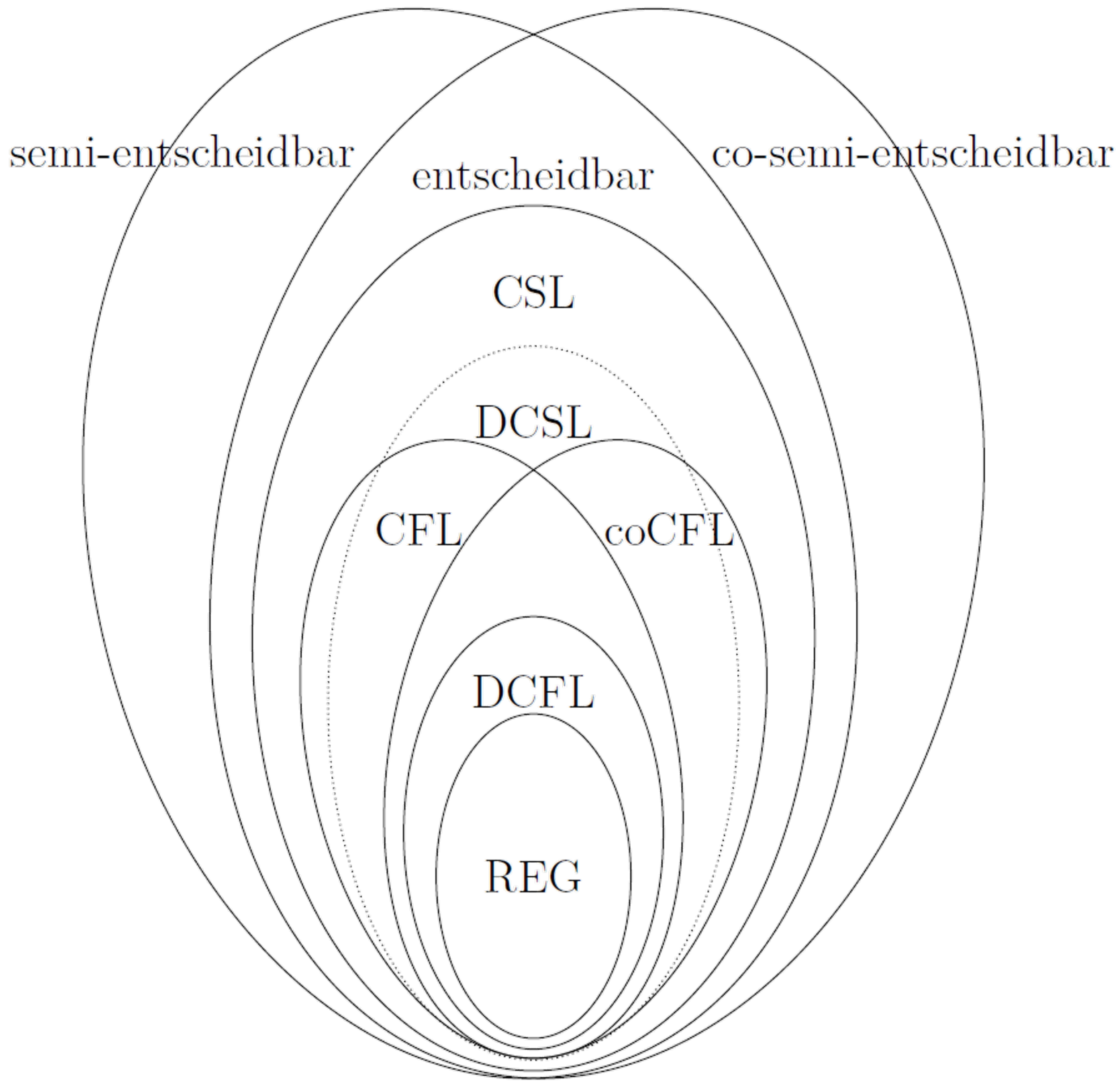
Sei im Folgenden DCFL $:= \{\, \mathcal{L}(P) \mid P \text{ det. PDA} \,\}$

DCFL

Es gilt: REG ⊂ DCFL ⊂ CFL.

Bekanntermaßen ist
$\{ a^n b^n \mid n \in \mathbb{N} \}$ eine nicht-reguläre Sprache,
kann aber mit einem det. PDA erkannt werden.

Allerdings kann
$\{ ww^{reverse} \mid w \in \{0,1\}^* \}$ nicht durch einen det. PDA erkannt werden, gehört aber zu CFL.



Abgeschlossenheit

DCFL und CFL unterscheiden sich bzgl. Abgeschlossenheit:

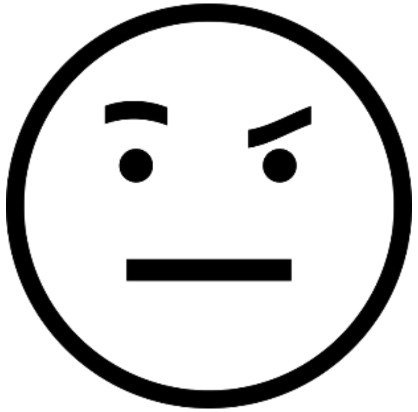
	DCFL	CFL
Komplement	Ja	Nein
Vereinigung	Nein	Ja
Schnitt	Nein	Nein

Unentscheidbare Probleme für DCFL

$$\mathcal{L}(G_1) = \left\{ z_{i_n} \ldots z_{i_1} x_{i_1} \ldots x_{i_n} \# y_{j_m}^{\text{reverse}} \ldots y_{j_1}^{\text{reverse}} z_{j_m} \ldots z_{j_1} \;\middle|\; n, m \geq 1, \forall s, t\colon i_s, j_t \in \{1, \ldots, k\} \right\}$$

$$\mathcal{L}(G_2) = \{ u\ v\ \# v^{\text{reverse}} u^{\text{reverse}} \mid v \in \{0,1\}^*, u \in \{z_1, \ldots, z_k\}^* \}$$

Diese beiden Sprachen können von det. PDAs erkannt werden!



Wir haben doch gerade festgehalten, dass gerade Palindrome nicht von det. PDAs erkannt werden können!

Hier haben wir allerdings ein Trennzeichen!

Die in Theorem 6.1 genannten Eigenschaften sind auch für det. kontextfreie Sprachen unentscheidbar.

Intersection DCFL

INTERSECTION-EMPTINESS-DCFL
Gegeben: Det. PDAs P_1, P_2.
Frage: Gilt $\mathcal{L}(P_1) \cap \mathcal{L}(P_2) = \emptyset$?

Korollar 6.3
INTERSECTION-EMPTINESS-DCFL ist unentscheidbar.

Weitere unentscheidbare Probleme

INCLUSION-CFL
Gegeben: Kontextfreie Grammatiken G_1, G_2.
Frage: Gilt $\mathcal{L}(G_1) \subseteq \mathcal{L}(G_2)$?

EQUIVALENCE-CFL
Gegeben: Kontextfreie Grammatiken G_1, G_2.
Frage: Gilt $\mathcal{L}(G_1) = \mathcal{L}(G_2)$?

UNIVERSALITY-CFL
Gegeben: Kontextfreie Grammatik G mit Terminalalphabet Σ.
Frage: Gilt $\mathcal{L}(G) = \Sigma^*$?

Weitere unentscheidbare Probleme

Theorem 6.4
Die folgenden Probleme sind unentscheidbar.
a) INCLUSION-CFL
b) EQUIVALENCE-CFL
c) UNIVERSALITY-CFL

Aus c) folgt direkt das Korollar:

Korollar
Gegeben sei eine Kontextfreie Grammatik G und ein NFA A. Dann sind die folgenden Probleme sind unentscheidbar.
a) Gilt $\mathcal{L}(A) \subseteq \mathcal{L}(G)$?
b) Gilt $\mathcal{L}(A) = \mathcal{L}(G)$?

INTERSECTION-EMPTINESS-DCFL $\leq$ INCLUSION-CFL

Wir haben gegeben zwei det. PDAs P_1, P_2. Dann gilt:

$$\mathcal{L}(P_1) \cap \mathcal{L}(P_2) = \emptyset \Leftrightarrow \mathcal{L}(P_1) \subseteq \overline{\mathcal{L}(P_2)}$$

DCFL ist unter Komplement abgeschlossen, d.h. wir können einen PDA $\overline{P_2}$ mit $\mathcal{L}(\overline{P_2}) = \overline{\mathcal{L}(P_2)}$ konstruieren.

Aus TheoInf 1: Aus jedem PDA kann eine CFG abgeleitet werden.

→ Konstruiere Grammatiken $G_1, \overline{G_2}$ mit $\mathcal{L}(G_1) = \mathcal{L}(P_1)$ und $\mathcal{L}(\overline{G_2}) = \mathcal{L}(\overline{P_2})$

→ Die Funktion $f(P_1, P_2) = (G_1, \overline{G_2})$ ist die gesuchte Reduktion.

Korollar

INCLUSION-DCFL (also mit Beschränkung auf det. Kontextfreie Sprachen) ist auch unentscheidbar.

INCLUSION-CFL $\leq$ EQUIVALENCE-CFL

Betrachte zwei Grammatiken G_1, G_2. Dann gilt:

$$\mathcal{L}(G_1) \subseteq \mathcal{L}(G_2) \Leftrightarrow \mathcal{L}(G_1) \cup \mathcal{L}(G_2) = \mathcal{L}(G_2)$$

CFL ist unter Vereinigung abgeschlossen, d.h. betrachte Grammatik G_3 mit

$$\mathcal{L}(G_3) = \mathcal{L}(G_1) \cup \mathcal{L}(G_2)$$

Dann ist

$$\mathcal{L}(G_1) \subseteq \mathcal{L}(G_2) \Leftrightarrow \mathcal{L}(G_3) = \mathcal{L}(G_2)$$

Die Funktion $g(G_1, G_2) = (G_3, G_2)$ ist die gesuchte Reduktion.

INTERSECTION-EMPTINESS-DCFL $\leq$ UNIVERSALITY-CFL

Wir haben gegeben zwei det. PDAs P_1, P_2. Dann gilt:

$$\mathcal{L}(P_1) \cap \mathcal{L}(P_2) = \emptyset \Leftrightarrow \overline{\mathcal{L}(P_1)} \cup \overline{\mathcal{L}(P_2)} = \Sigma^*$$

DCFL ist unter Komplement abgeschlossen, d.h. wir können PDAs $\overline{P_1}, \overline{P_2}$ mit $\mathcal{L}(\overline{P_1}) = \overline{\mathcal{L}(P_1)}$ und $\mathcal{L}(\overline{P_2}) = \overline{\mathcal{L}(P_2)}$ konstruieren.

Aus TheoInf 1: Aus jedem PDA kann eine CFG abgeleitet werden.

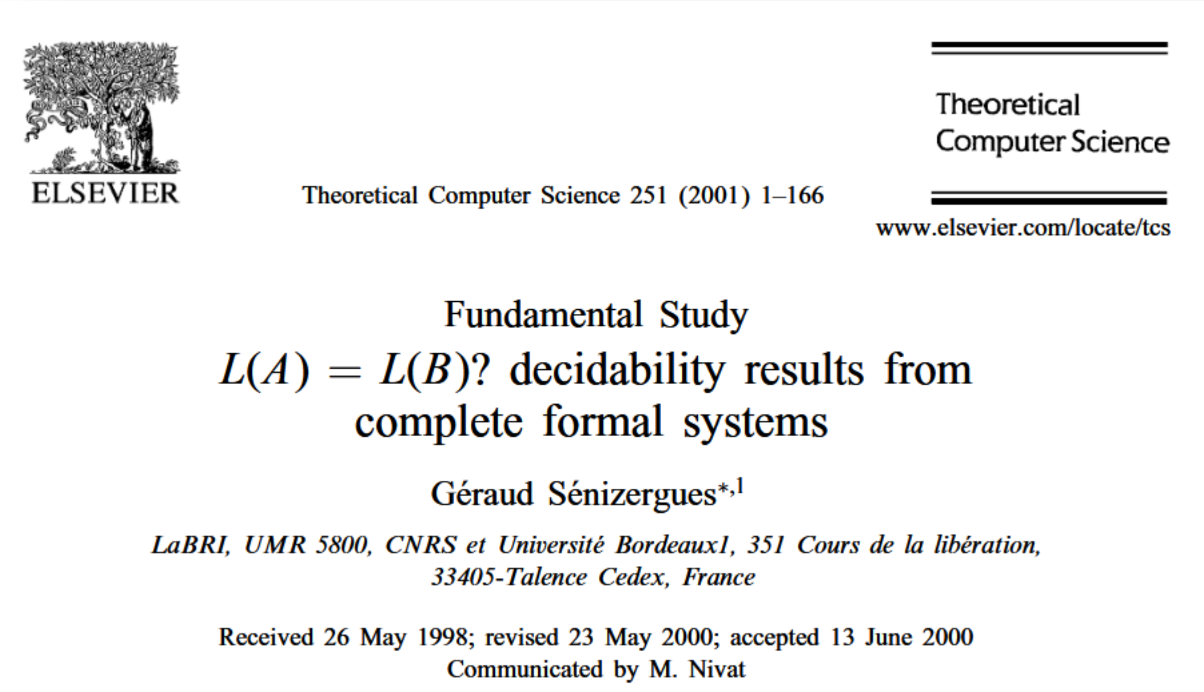
→ Konstruiere Grammatiken $\overline{G_1}, \overline{G_2}$ mit $\mathcal{L}(\overline{G_1}) = \mathcal{L}(\overline{P_1})$ und $\mathcal{L}(\overline{G_2}) = \mathcal{L}(\overline{P_2})$

→ Konstruiere Grammatik $G = \overline{G_1} \cup \overline{G_2}$

→ Die Funktion $f(P_1, P_2) = G$ ist die gesuchte Reduktion.

Bemerkung 6.5

Die deterministischen Varianten EQUIVALENCE-DCFL, sowie UNIVERSALITY-DCFL sind entscheidbar.


ELSEVIER

Theoretical Computer Science 251 (2001) 1–166

Theoretical Computer Science

www.elsevier.com/locate/tcs

Fundamental Study

$L(A) = L(B)$? decidability results from complete formal systems

Géraud Sénizergues*,[1]

LaBRI, UMR 5800, CNRS et Université Bordeaux1, 351 Cours de la libération, 33405-Talence Cedex, France

Received 26 May 1998; revised 23 May 2000; accepted 13 June 2000

Communicated by M. Nivat

Géraud Sénizergues

Noch mehr unentscheidbare Probleme

<u>REGULARITY-CFL</u>
Gegeben: Kontextfreie Grammatik G.
Frage: Ist $\mathcal{L}(G)$ regulär?

<u>CONTEXTFREE-COMPLEMENT-CFL</u>
Gegeben: Kontextfreie Grammatik G.
Frage: Ist $\overline{\mathcal{L}(G)}$ kontextfrei?

<u>DCFL-MEMBERSHIP-CFL</u>
Gegeben: Kontextfreie Grammatik G.
Frage: Ist $\mathcal{L}(G)$ deterministisch-kontextfrei?

Folgen

Theorem 6.6
Die folgenden Probleme sind unentscheidbar.

a) REGULARITY-CFL
b) CONTEXTFREE-COMPLEMENT-CFL
c) DCFL-MEMBERSHIP-CFL

Konsequenz aus Th. 6.6 c):
Es gibt keinen Algorithmus, der einen gegebenen PDA und einen dPDA konstruiert, falls existent.

Nächstes Kapitel

Es gibt viele unentscheidbare Probleme.

Was können wir über **entscheidbare** Probleme aussagen?

Wie effizient können wir diese lösen?

Det. vs Nicht-Det.

Speicher vs. Laufzeit

1-Band vs. m-Band