Fachhochschule Dortmund
University of Applied Sciences and Arts
we focus on students
Prof. Dr. Robert Preis

Kapitel 3

Kontextfreie Sprachen

3.4

Grenzen kontextfreier Sprachen

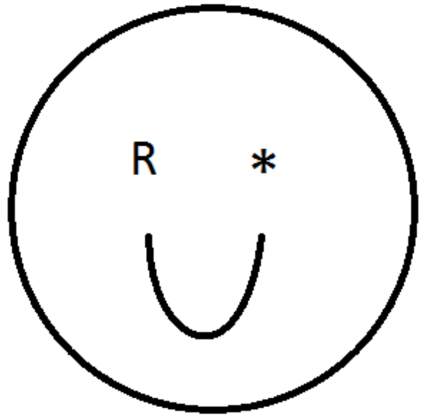
Prof. Dr. Robert Preis
Fachbereich Informatik
Fachhochschule Dortmund
Robert.Preis@fh-dortmund.de

Alle Materialien (Folien, Übungsblätter, etc.) dieser Veranstaltung sind urheberrechtlich geschützt und nur von Teilnehmern dieser Veranstaltung und im Rahmen dieser zu verwenden. Eine anderweitige Verwendung oder Verbreitung ist nicht gestattet.

Abgeschlossenheit kontextfreier Sprachen

Kontextfreie Sprachen sind abgeschlossen über

- $L_1 \cup L_2$ — Die Vereinigung zweier kf. Sprachen ist kf. .
- L^R — Die Spiegelung einer kf. Sprache ist kf. .
- $L_1 \circ L_2$ — Die Verkettung zweier kf. Sprachen ist kf. .
- L^* — Die Hülle einer kf. Sprache ist kf. .



Lars Aron Gehrke, 2017

Kontextfreie Sprachen sind **NICHT** abgeschlossen über

- $L_1 \cap L_2$ — Der Durchschnitt zweier kf. Sprachen muss nicht kf. sein:
  $L_1 = \{0^n1^n2^m \mid n,m \in \mathbf{N}\}$ und $L_2 = \{0^m1^n2^n \mid n,m \in \mathbf{N}\}$ sind kf.,
  $L_1 \cap L_2 = \{0^n1^n2^n \mid n \in \mathbf{N}\}$ ist aber nicht kf. (Beweis später).
- L^c — Das Komplement einer kf. Sprache muss nicht kf. sein:
  (dann wäre $L_1 \cap L_2 = (L_1^c \cup L_2^c)^c$ immer kf.)
- $L_1 \setminus L_2$ — Die Differenz zweier kf. Sprachen muss nicht kf. sein
  (dann wäre $L^c = \Sigma^* \setminus L$ immer kf.)

Das Pumping Lemma für kontextfreie Sprachen

Frage: *Was passiert,*

wenn ich bei einer Grammatik in CNF mit $|V|$ Variablen

ein Wort w abarbeite,

dessen Länge mindestens $|w| = 2^{|V|}$ ist?

Es gilt ein Pumping Lemma für Typ-2-Sprachen:

Für jede unendliche Typ-2-Sprache L lassen sich Wörter, wenn sie genügend lang sind, so zerlegen, dass es

zwei parallel zu vervielfachende „Pumpstellen“

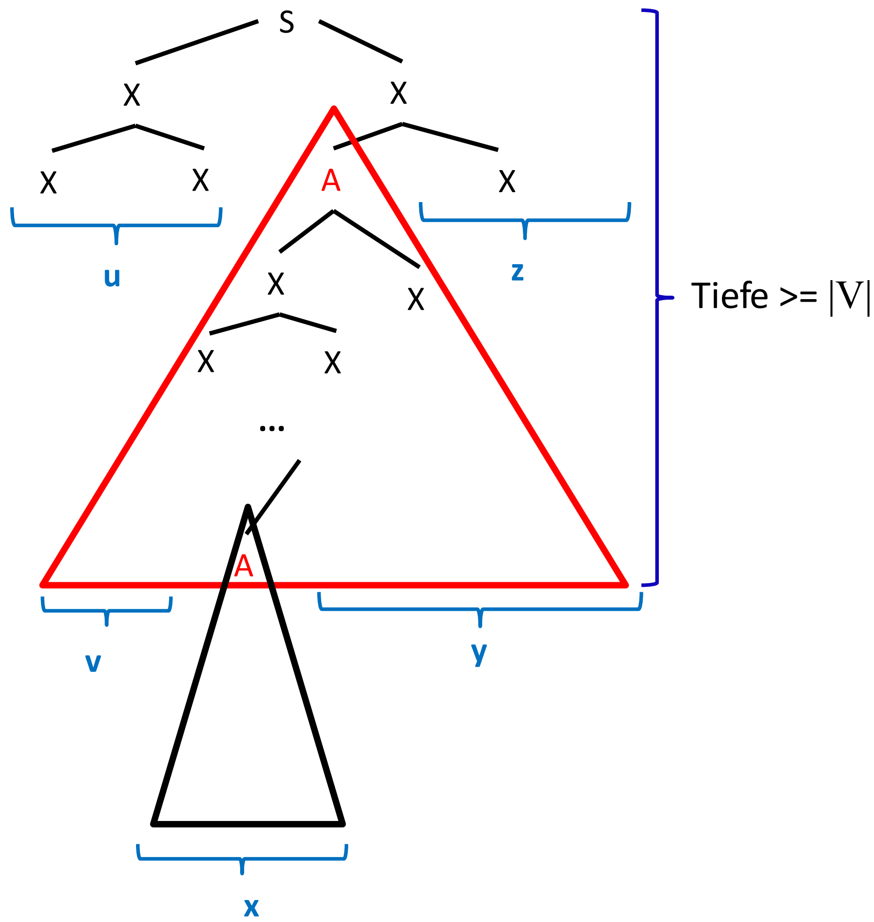
gibt.

Warum? Dazu kann man die CNF gut brauchen!

Ableitung eines langen Wortes w in einer Typ-2 Grammatik in CNF (Binärbaum)

Betrachten wir in einer Grammatik mit |V| Variablen den Ableitungsbaum eines Wortes w mit Länge $|w| >= 2^{|V|}$:

1. Der längste Ableitungsweg hat die Länge >= |V|.
2. Auf diesem Weg sind mindestens |V|+1 Variablen!
3. Also kommt mindestens eine Variable doppelt vor, z.B. A.
4. Das Wort w kann aufgeteilt werden in w=uvxyz.
   1. $vy \neq \varepsilon$
   2. $|vxy| \leq 2^{|V|}$
   3. für alle $k \in \mathbf{N}_0$ ist $uv^kxy^kz \in L$



Man kann v und y aufpumpen !

Das Pumping Lemma für kontextfreie Sprachen

Pumping Lemma:

Für jede kontextfreie Sprache L gibt es eine Pumpingzahl $n \in \mathbf{N}$, so dass jedes Wort $w \in L$ mit Länge $|w| \geq n$ zerlegt werden kann in $w = uvxyz$ mit den Eigenschaften

1. $vy \neq \varepsilon$
2. $|vxy| \leq n$
3. für alle $k \in \mathbf{N}_0$ ist $uv^kxy^kz \in L$.

Anwendung des Pumping Lemma

Behauptung: $L = \{0^m 1^m 2^m \mid m \in \mathbf{N}\}$ ist kontextfrei

- Dann gibt es eine (unbekannte, aber feste) Pumpingzahl $n \in \mathbf{N}$.
- Das Wort $w = 0^n 1^n 2^n$ muss in der Sprache sein.
- Laut Pumping Lemma:
  - Sei $u,v,x,y,z \in T^*$ eine beliebige Zerlegung von w in w = uvxyz mit
    1. $vy \neq \varepsilon$ und
    2. $|vxy| \leq n$ (vxy enthält keine Nullen oder keine Zweien)
  - Wir wählen i=0:
    - Es ist $uv^ixy^iz = uv^0xy^0z = uxz \in L$ (laut PL)
    - Aber $uxz \notin L$, weil entweder keine 0-en oder keine 2-en gelöscht werden
    - ... Das geht nicht, deshalb Widerspruch

Deshalb: **$L = \{0^m 1^m 2^m \mid m \in N\}$ ist NICHT kontextfrei**

Unentscheidbare Probleme für kontextfreie Sprachen

Frage 3: Sind die Sprachen von zwei kontextfreien Grammatiken identisch?

Frage 4: Kann man eine kontextfreie Grammatik minimieren?

Die folgenden Probleme sind

unentscheidbar (nicht berechenbar),

d.h. es kann keine Maschine oder Programm geben, die dieses Problem **für alle** möglichen Fälle beantwortet.

- *$L(G_1) = L(G_2)$* *für zwei kontextfreie Grammatiken G_1 und G_2?*

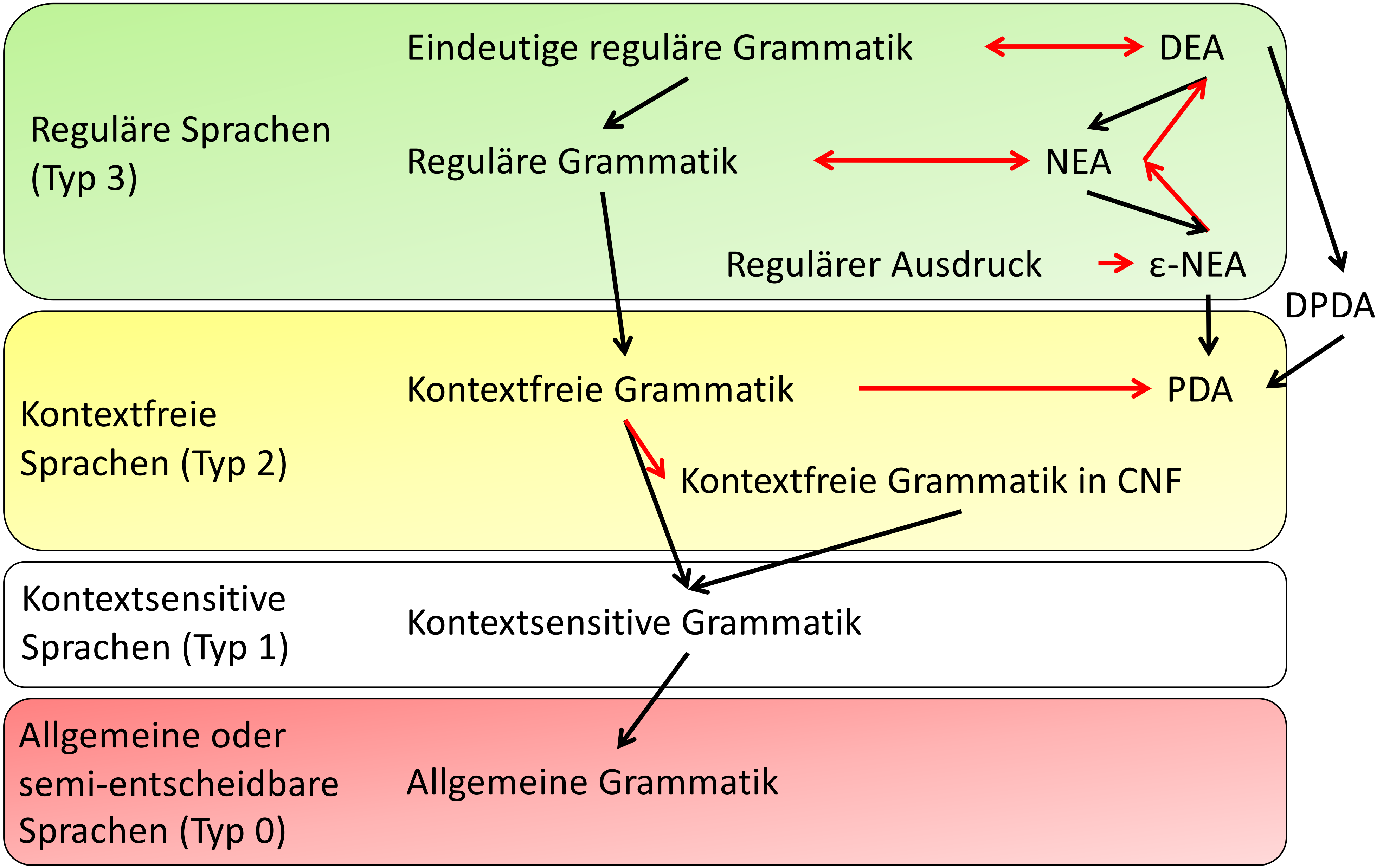
  …dann funktioniert eine Minimierung erst recht nicht!
- *$L(G_1) \subseteq L(G_2)$* *für zwei kontextfreie Grammatiken G_1 und G_2?*
- *$L(G_1) \cap L(G_2) = \emptyset$* *für zwei kontextfreie Grammatiken G_1 und G_2?*
- *$L(G)$ regulär* *für eine kontextfreie Grammatik G?*

Um das zu beweisen, brauchen wir die Theorie der Berechenbarkeit in der Vorlesung *Berechenbarkeit und Komplexitätstheorie,* die jeweils im SS im Master Informatik (Dozent: R.P.) angeboten wird.

Überblick Sprachklassen

	Typ-3	Typ-2	Typ-1	Typ-0
Frage 1: L={} ?	√	√		
Frage 2: $w \in L$?	√	√		
Frage 3: $L_1=L_2$?	√	ϟ		
Frage 4: minimal ?	√	ϟ		
$L_1 \cup L_2$	√	√		
$L_1 \cap L_2$	√	ϟ		
$L_1 \setminus L_2$	√	ϟ		
L^c	√	ϟ		
$L_1 \circ L_2$	√	√		
L^*	√	√		

Überblick der Transformationen



Zusammenfassung

- **Abschlußeigenschaften** kontextfreier Sprachen.
  - Abgeschlossenheit über $\cup$, ^{R}, $\circ$, $*$
  - Keine Abgeschlossenheit unter $\cap$, ^{c}, $\setminus$
- **Pumping Lemma**: In langen Wörtern von kontextfreien Sprachen gibt es immer ein oder zwei Wortteile in den ersten n Zeichen (n sprachabhängig), die man beliebig oft wiederholen kann, ohne die Sprache zu verlassen.
- Konsequenz: Es muss Sprachen geben, die nicht kontextfrei sind.
- Bezüglich kontextfreier Sprachen sind **wichtige Fragen unentscheidbar**.