

Grundlagen der Programmierung

Compiler B: Grammatiken

Prof. Dr. Manfred Schmidt-Schauß

Sommersemester 2018

Beschreibung der Sprache L_S aller Programme einer Programmiersprache:

- Finde kontextfreie Grammatik (CFG) G mit eindeutigen Syntaxbäumen, so dass $L(G)$ Obermenge von L_S
- Weitere Einschränkung von $L(G)$ (durch Bedingungen) ergibt L_S .

4-Tupel $G = (N, T, P, \sigma)$ mit

- 1 N : endliche Menge von *Hilfszeichen (Nonterminals)*
- 2 T endliche Menge von *Terminalzeichen (Terminals)*,
wobei $N \cap T = \emptyset$.
- 3 $P \subseteq N \times (N \cup T)^*$ endliche Menge von *Regeln (Produktionen)*
- 4 $\sigma \in N$ ist *Startzeichen*

Regeln schreibt man auch $A \rightarrow w$ oder $A ::= w$.

Erzeugung:	Starte	mit σ .
(Herleitung)	Iteriere	Wende eine Regel an
	Ende	Wort aus Nichtterminalen

Erzeugte Sprache $L(G)$ = erreichbare Worte über T .
= $\{w \in T^* \mid \sigma \xrightarrow{*}_G w\}$

Die AZ-Grammatik

A ::= **A + Z** | **A - Z** | **Z**

Z ::= 0 | ... | 9

Die AA-Grammatik:

A ::= **A + A**

A ::= **A - A**

A ::= 0 | ... | 9

Beide Grammatiken erzeugen dieselbe Sprache.

Diese besteht aus sehr einfachen arithmetische Ausdrücken:

$z_1 \text{ op}_1 z_2 \dots \text{op}_n z_{n+1}$

mit $z_i \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 0\}$

und $\text{op}_i \in \{+, -\}$

z.B. $9 + 2 - 5 + 3$

Beispiel: $x := 1$; $y := 2$; IF $x < y$ THEN $z := 5$ ELSE $z := 100$

Grammatik:

Programm ::= skip | **Befehl ; Programm**

fett: Nichtterminal

blau: Terminal-Texte

andere Meta-Symbole der Grammatik

Beispiel: $x := 1$; $y := 2$; IF $x < y$ THEN $z := 5$ ELSE $z := 100$

Grammatik:

Programm	::=	skip Befehl ; Programm
Befehl	::=	Zuweisung Verzweigung

fett: **Nichtterminal**

blau: **Terminal-Texte**

andere **Meta-Symbole der Grammatik**

Beispiel: $x := 1$; $y := 2$; IF $x < y$ THEN $z := 5$ ELSE $z := 100$

Grammatik:

Programm	::=	skip Befehl ; Programm
Befehl	::=	Zuweisung Verzweigung
Zuweisung	::=	Variable := Ausdruck
Verzweigung	::=	IF Bedingung THEN Befehl ELSE Befehl

fett: Nichtterminal

blau: Terminal-Texte

andere Meta-Symbole der Grammatik

Beispiel: $x := 1$; $y := 2$; IF $x < y$ THEN $z := 5$ ELSE $z := 100$

Grammatik:

Programm	::=	skip Befehl ; Programm
Befehl	::=	Zuweisung Verzweigung
Zuweisung	::=	Variable := Ausdruck
Verzweigung	::=	IF Bedingung THEN Befehl ELSE Befehl
Bedingung	::=	Ausdruck Vergleichsop Ausdruck
Ausdruck	::=	Variable Zahl (- Ausdruck) (Ausdruck Operator Ausdruck)
...		...

fett: Nichtterminal

blau: Terminal-Texte

andere Meta-Symbole der Grammatik

Linksherleitung: Ersetze stets das linkeste Nichtterminal

Rechtsherleitung: Ersetze stets das rechteste Nichtterminal

Die **AA-Grammatik**:

A ::= **A+A**

A ::= **A - A**

A ::= 0 | ... | 9

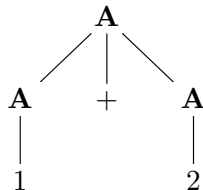
Beispiel „1+2“

Linksherleitung:

A → **A+A** → 1+**A** → 1+2

Rechtsherleitung:

A → **A+A** → **A**+2 → 1+2



Zur kontextfreien Grammatik $G = (N, T, P, \sigma)$

Kann man Herleitungen mittels

Herleitungsbäumen, bzw. Parse-Bäumen darstellen.

- pro Herleitungsbaum gibt es
genau eine Rechtsherleitung
und genau eine Linksherleitung
- Die Bedeutung eines Programms wird festgelegt
anhand des Herleitungsbaumes

Definition

Eine kontextfreie Grammatik G heißt *eindeutig*, wenn für alle $w \in L(G)$ *genau ein* Herleitungsbaum existiert. Andernfalls heißt die Grammatik *mehrdeutig*.

Wenn G eindeutig ist,
Dann gibt es für *jedes* Wort in $L(G)$
genau einen Herleitungsbaum
und *genau eine Rechts-Herleitung*
und *genau eine Links-Herleitung*

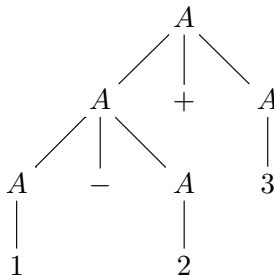
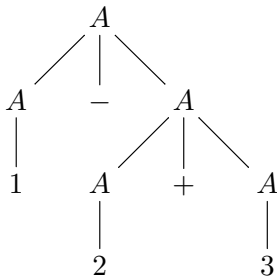
Beispiel zu AA-Grammatik

A ::= **A+A**

A ::= **A-A**

A ::= 0 | ... | 9

Zwei verschiedene Herleitungsbäume für “1 - 2 + 3” sind:



Die AZ-Grammatik ist eindeutig.

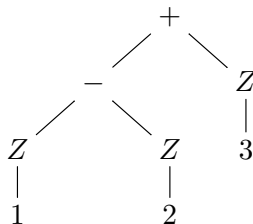
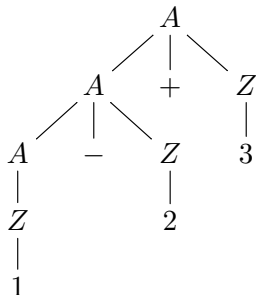
A ::= **A**+**Z**

A ::= **A**-**Z**

A ::= **Z**

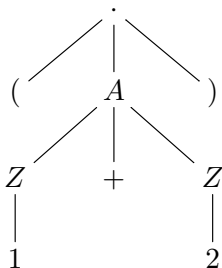
Z ::= 0 | ... | 9

Der einzige (d.h. eindeutige) AZ-Herleitungsbaum für $1 - 2 + 3$,
und ein Syntaxbaum dazu ist:

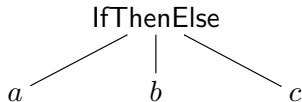
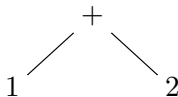


Beispiele:

Herleitungsbaum



Syntaxbaum



Die Struktur eines Ausdrucks/Programms wird erkannt
durch den Parsebaum / Syntaxbaum;
danach die Semantik anhand des Syntaxbaums

Mehrdeutige Grammatik entspricht i.a.
mehrdeutiger (d.h. undefinierter) Semantik

Reparatur: Erzeuge äquivalente eindeutige Grammatik
Beachte Grammatiken, die äquivalent bzgl $L(.)$ sind,
könnten **verschiedene operationale Semantik** für den gleichen
String festlegen