

1. Formale Sprachen

1.2 Grammatiken formaler Sprachen

Die Regeln zur Bildung korrekter Wörter einer Sprache kann man in einer natürlichen Sprache formulieren. Da dies jedoch wieder Mehrdeutigkeiten mit sich bringen kann, ist es notwendig, eindeutige **Produktionsregeln** für die Wörter einer Sprache zu formulieren.

1.2 Grammatiken formaler Sprachen

Beispiel

Die Wörter einer Sprache sollen korrekte Klassenbezeichnungen an einem Gymnasium mit höchstens sieben Klassen in einem Jahrgang sein.

Korrekte Wörter sind also 5a, 7f, 10g, Q11, Q12.

Das Alphabet dieser Sprache soll sein

$A = \{5, 6, 7, 8, 9, 10, 11, 12, a, b, c, d, e, f, g, Q\}$.

1.2 Grammatiken formaler Sprachen

Die Produktionsregeln kann man mit Hilfe **syntaktischer Variablen** festlegen:

R1: <Stufe> → <Unter- oder Mittelstufe>

R2: <Stufe> → <Oberstufe>

R3: <Unter- oder Mittelstufe> → <ZahlKlasse><BezeichnungKlasse>

R4: <ZahlKlasse> → '5'

R5: <ZahlKlasse> → '6'

R6: <ZahlKlasse> → '7'

R7: <ZahlKlasse> → '8'

R8: <ZahlKlasse> → '9'

R9: <ZahlKlasse> → '10'

R10: <BezeichnungKlasse> → 'a'

R11: <BezeichnungKlasse> → 'b'

R12: <BezeichnungKlasse> → 'c'

R13: <BezeichnungKlasse> → 'd'

R14: <BezeichnungKlasse> → > 'e'

R15: <BezeichnungKlasse> → 'f'

R16: <BezeichnungKlasse> → > 'g'

R17: <Oberstufe> → 'Q'<ZahlOberstufe>

R18: <ZahlOberstufe> → '11'

R19: <ZahlOberstufe> → '12'

1.2 Grammatiken formaler Sprachen

Man kann die Regeln auch folgendermaßen zusammenfassen:

Z1: <Stufe> → <Unter- oder Mittelstufe> | <Oberstufe>

Z2: <Unter- oder Mittelstufe> →
<ZahlKlasse><BezeichnungKlasse>

Z3: <ZahlKlasse> → '5' | '6' | '7' | '8' | '9' | '10'

Z4: <BezeichnungKlasse> → 'a' | 'b' | 'c' | 'd' | 'e' | 'f' | 'g' |

Z5: <Oberstufe> → 'Q'<ZahlOberstufe>

Z6: <ZahlOberstufe> → '11' | '12'

Diese Schreibweise der Produktionsregeln heißt

Backus-Naur-Form (BNF)

1.2 Grammatiken formaler Sprachen

Das Wort 'Q12' wird folgendermaßen gebildet:

<Stufe> \rightarrow (R2, bzw. Z1) **<Oberstufe>** \rightarrow (R7, bzw. Z5)
'Q'<ZahlOberstufe> \rightarrow (R19 bzw. Z6) **'Q12'**

Man bezeichnet diese Vorgehensweise als

Ableitung des Wortes.

Dazu ist es notwendig, genau eine der syntaktischen Variablen als **Startvariable** festzulegen.

Zeichen aus dem Alphabet nennt man auch **Terminale**,
syntaktische Variablen nennt man **Nichtterminale**.

1.2 Grammatiken formaler Sprachen

Formale Sprachen können beliebig viele Wörter enthalten, die auch unbegrenzte Wortlängen haben können.

Die Anzahl der Produktionsregeln muss jedoch endlich sein.

Zur Bildung von Wörtern mit unbegrenzter Wortlänge eignet sich **Rekursion**.

Beispiel:

Sprache aller nichtnegativen ganzen Zahlen $\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\}$

Wörter der Sprache (führende Nullen erlaubt) :

0, 1, 2, ..., 10, 11, 12, ..., 84, 084, ...

Alphabet $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

Produktionsregeln :

R1: $\langle \text{zahl} \rangle \rightarrow \langle \text{ziffer} \rangle \mid \langle \text{zahl} \rangle \langle \text{ziffer} \rangle$

R2: $\langle \text{ziffer} \rangle \rightarrow '0' \mid '1' \mid '2' \mid '3' \mid \dots \mid '9'$

Startvariable: $\langle \text{zahl} \rangle$

Ableitung der Zahl (des Wortes) '084':

$\langle \text{zahl} \rangle \rightarrow \langle \text{zahl} \rangle \langle \text{ziffer} \rangle \rightarrow \langle \text{zahl} \rangle '4' \rightarrow \langle \text{zahl} \rangle \langle \text{ziffer} \rangle '4' \rightarrow \langle \text{zahl} \rangle '84' \rightarrow$
 $\rightarrow \langle \text{ziffer} \rangle '84' \rightarrow '084'$

1.2 Grammatiken formaler Sprachen

Das Alphabet ist identisch mit der Menge der Terminale.

Ein Alphabet kann auch ein Leerzeichen als Terminal enthalten.

Eine Sprache kann auch das sogenannte **leere Wort ϵ** enthalten.

Es hat die Wortlänge 0.

Im Gegensatz dazu hat ein Wort, das aus dem Leerzeichen besteht, die Wortlänge 1.

1.2 Grammatiken formaler Sprachen

Beispiel für die Verwendung von ε :

Nichtnegative ganze Zahlen (ohne führende Nullen)

Alphabet $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

Wörter: 0, 1, 2, ..., 9, 10, 11, 12, ..., 804, ...

Produktionsregeln

R1: $\langle \text{zahl} \rangle \rightarrow \langle 0\text{ziffer} \rangle \mid \langle \text{ziffer} \rangle \langle \text{rest} \rangle$

R2: $\langle \text{rest} \rangle \rightarrow \langle 0\text{ziffer} \rangle \langle \text{rest} \rangle \mid \varepsilon$

R3: $\langle \text{ziffer} \rangle \rightarrow '1' \mid '2' \mid '3' \mid \dots \mid '9'$

R4: $\langle 0\text{ziffer} \rangle \rightarrow '0' \mid \langle \text{ziffer} \rangle$

Startvariable: $\langle \text{zahl} \rangle$

Ableitung der Zahl (des Wortes) '804':

$\langle \text{zahl} \rangle \rightarrow \langle \text{ziffer} \rangle \langle \text{rest} \rangle \rightarrow '8' \langle \text{rest} \rangle \rightarrow '8' \langle 0\text{ziffer} \rangle \langle \text{rest} \rangle \rightarrow '80' \langle \text{rest} \rangle \rightarrow$
 $\rightarrow '80' \langle 0\text{ziffer} \rangle \langle \text{rest} \rangle \rightarrow '804' \langle \text{rest} \rangle \rightarrow '804' \varepsilon \rightarrow '804'$

1.2 Grammatiken formaler Sprachen

Die Menge A^* eines Alphabets bezeichnet alle möglichen Wörter, die sich (ohne Existenz bestimmter Produktionsregeln) mit den Elementen aus A bilden lassen.

Somit kann man sagen, dass eine **formale Sprache eine Teilmenge von A^*** ist.

1.2 Grammatiken formaler Sprachen

Das Alphabet A einer Sprache, die Menge V der syntaktischen Variablen, die Startvariable S aus V und die Menge P der Produktionsregeln bilden zusammen die **Grammatik einer Sprache**. $G = (A, V, S, P)$
 A und V dürfen dabei kein gemeinsames Element haben.

1.2 Grammatiken formaler Sprachen

Übung 1

In einer Stadt gibt es zwei U-Bahn- und drei S-Bahnlinien:

U1, U2, S1, S2, S3

Bestimme die Grammatik einer formalen Sprache, welche diese Wörter erzeugt.

Übung 2

Uebung2_Grammatik_BNF.pdf

Aufgaben

2, 3 (ohne Verwendung von ε), 6, 9, 10, 11

1.2 Grammatiken formaler Sprachen

Die Ableitung eines Wortes kann man auch mit Hilfe eines **Ableitungsbaumes** darstellen.

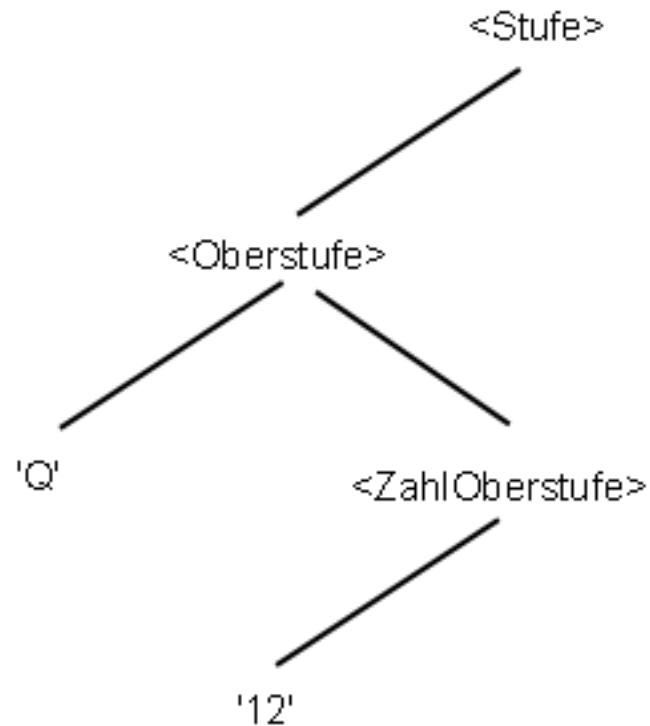
Die Wurzel ist die Startvariable. Die Nachfolger sind jeweils die folgenden Terminale oder Nichtterminale.

Die Blätter des Baumes sind schließlich die Terminale des abgeleiteten Wortes.

1.2 Grammatiken formaler Sprachen

Beispiel

Ableitung von Q12 :



1.2 Grammatiken formaler Sprachen

Andere Notationsformen:

Die erweiterte Backus-Naur-Form (EBNF)

Die BNF wird durch einige Abkürzungsregeln ergänzt:

1.2 Grammatiken formaler Sprachen

Andere Notationsformen:

Die erweiterte Backus-Naur-Form (EBNF)

Der Produktionspfeil \rightarrow wird durch ein Gleichheitszeichen = ersetzt.

Nichtterminale schreibt man ohne spitze Klammern,

Terminale schreibt man mit Hochkomma. *

Jede Regel endet mit einem Strichpunkt.

* In der BNF kann das Hochkomma bei den Terminalen auch weggelassen werden

1.2 Grammatiken formaler Sprachen

Der Produktionspfeil \rightarrow wird durch ein Gleichheitszeichen = ersetzt.

Nichtterminale schreibt man ohne spitze Klammern,

Terminale schreibt man mit Hochkomma.

Jede Regel endet mit einem Strichpunkt.

Beispiel

Datumsangaben :

datum = tag '.' monat '.' jahr ;

1.2 Grammatiken formaler Sprachen

Optionale Teile eines Worts werden in eckige Klammern [] gesetzt.

Beispiel

jahr = [ziffer ziffer] ziffer ziffer ;

heißt, dass man 2019 oder 19 für das Jahr schreiben kann.

In BNF :

<jahr> --> <ziffer><ziffer> | <ziffer><ziffer><ziffer><ziffer>

1.2 Grammatiken formaler Sprachen

Wiederholungen werden in geschweifte Klammern { } gesetzt.

Beispiel

zahl = ziffer {ziffer} ;

Die Wiederholung beinhaltet auch den Fall, dass 0 mal wiederholt wird.

In BNF:

$\langle \text{zahl} \rangle \rightarrow \langle \text{ziffer} \rangle \mid \langle \text{ziffer} \rangle \langle \text{zahl} \rangle$

1.2 Grammatiken formaler Sprachen

Gruppierungen werden in runde Klammern () gesetzt.

Beispiel

wort = ('a' | 'b') rest ;

produziert alle Wörter, die entweder mit 'a' oder 'b' anfangen.

In BNF:

<wort> --> 'a' <rest> | 'b' <rest>

1.2 Grammatiken formaler Sprachen

Übung 3

Buch, S. 28 - 29, Aufgaben 2, 5 und 6

Übung 4

Uebung2_Grammatik_BNF.pdf

Aufgabe 10 mit EBNF

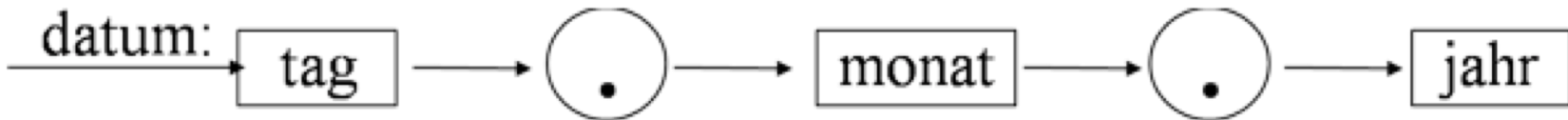
1.2 Grammatiken formaler Sprachen

Andere Notationsformen:

Syntaxdiagramme

Syntaxdiagramme helfen, die Produktionsregeln zu veranschaulichen

Beispiel



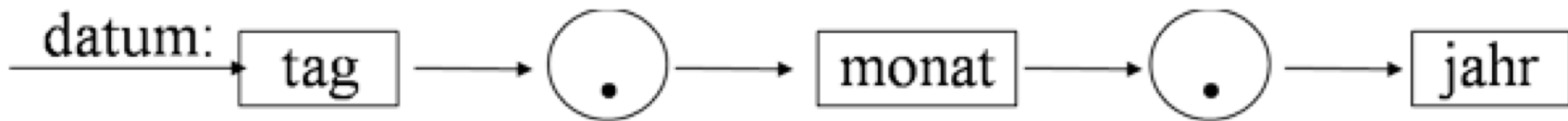
1.2 Grammatiken formaler Sprachen

Andere Notationsformen:

Syntaxdiagramme

Syntaxdiagramme helfen, die Produktionsregeln zu veranschaulichen

Beispiel



Hinter der Variable, die verarbeitet werden soll, setzt man einen Doppelpunkt.

Nichtterminale schreibt man in Rechtecke, Terminale in Kreise.

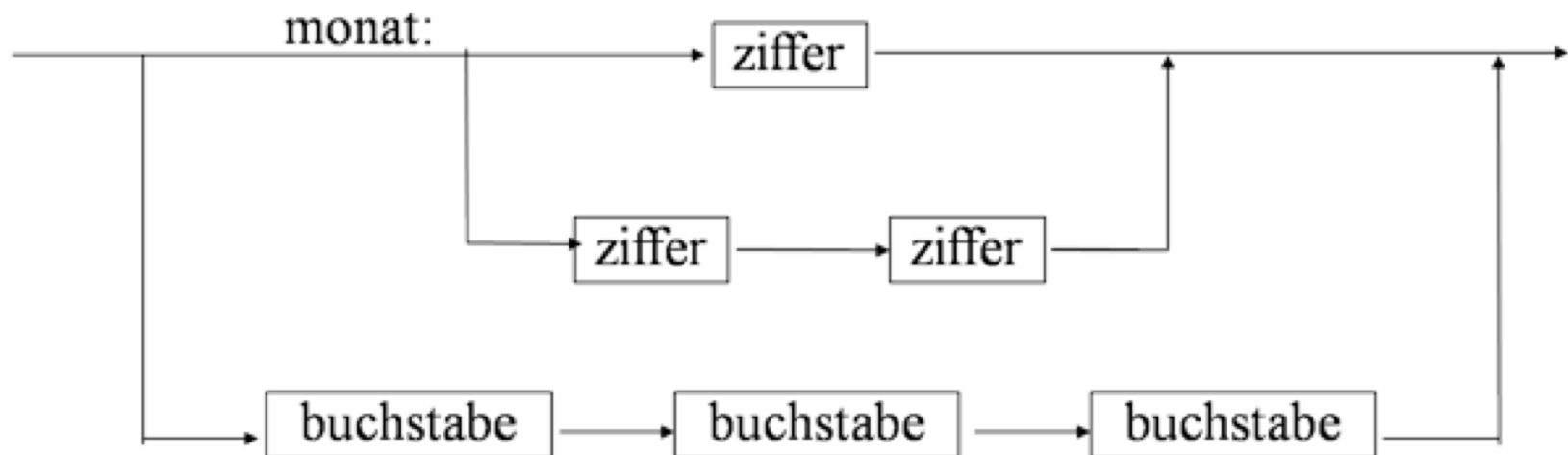
1.2 Grammatiken formaler Sprachen

Andere Notationsformen:

Syntaxdiagramme

Beispiel

Auswahl



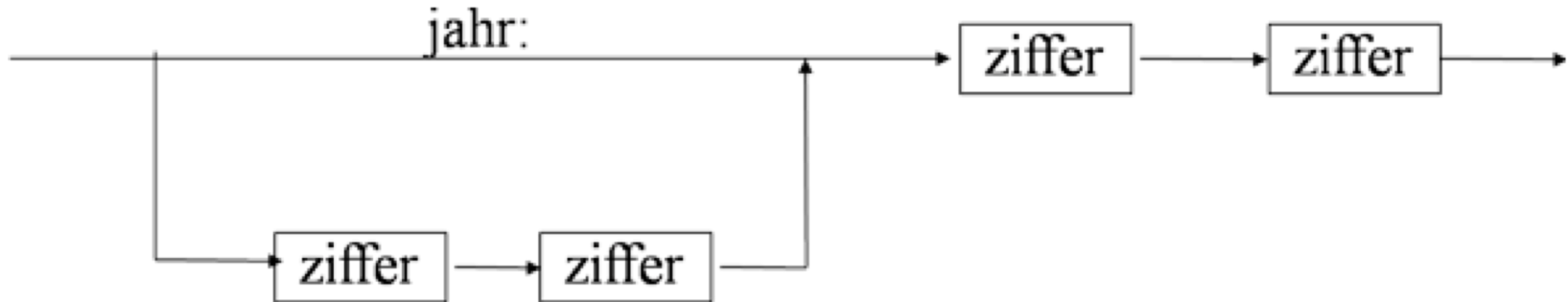
1.2 Grammatiken formaler Sprachen

Andere Notationsformen:

Syntaxdiagramme

Beispiel

Optionale Teile



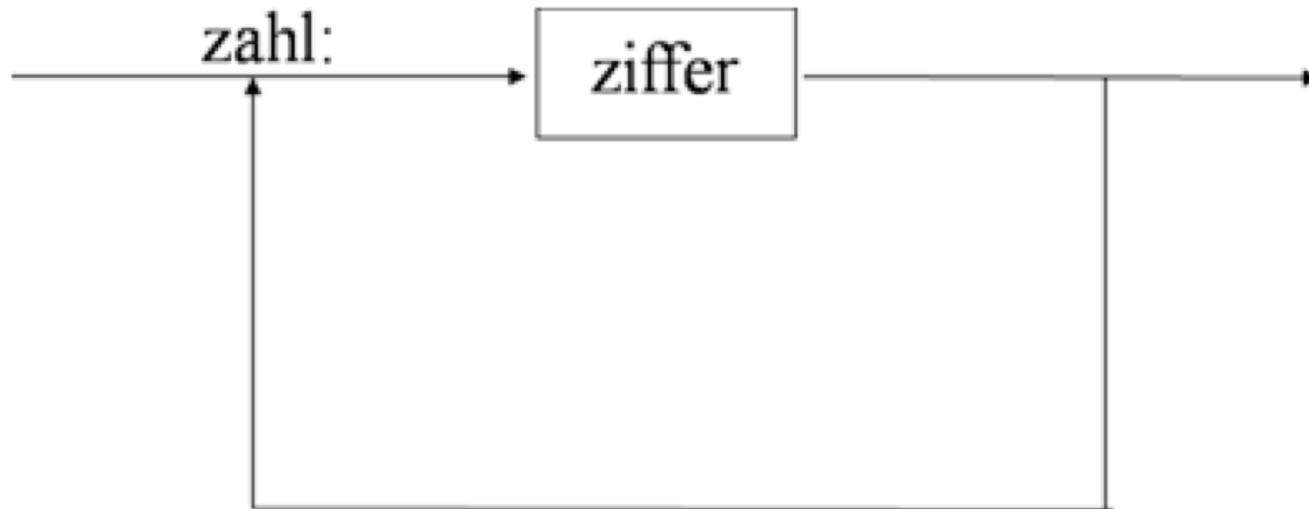
1.2 Grammatiken formaler Sprachen

Andere Notationsformen:

Syntaxdiagramme

Beispiel

Wiederholungen



1.2 Grammatiken formaler Sprachen

Übung 5

Buch, S. 29, Aufgabe 7