Parsing, Sommersemester 2014: Hausaufgaben

Laura Kallmeyer, Magnus Roos

Aufgabe 1 (Grammatiken und die Chomsky-Hierarchie) *Betrachten Sie die folgenden drei Sprachen:*

- $L_1 = \{a^m b^n a^n b^m \mid m,n \geq 1\}$
- $L_2 = \{a^m b^n b^n b^n \mid m,n \geq 1\}$
- $L_3 = \{a^m b^n a^n b^n \mid m,n \geq 1\}$

Eine dieser Sprachen ist regulär, eine ist kontextfrei aber nicht regulär und eine ist nicht kontextfrei.

1. *Welche Sprache ist regulär? Geben Sie eine reguläre Grammatik zu der Sprache an.*
2. *Welche Sprache ist kontextfrei aber nicht regulär? Geben Sie eine kontextfreie Grammatik zu der Sprache an.*

Lösung:

1. L_2 ist regulär, eine zugehörige Grammatik ist $G_2 = \langle N,T,P,S \rangle$ mit $N = \{S,A\}$, $T = \{a,b\}$ und

$$\begin{array}{rcl} P = \{S & \rightarrow & aS \mid aA \\ A & \rightarrow & bbbA \mid bbb\}. \end{array}$$

2. L_1 ist kontextfrei aber nicht reguär, eine zugehörige Grammatik ist $G_1 = \langle N,T,P,S \rangle$ mit $N = \{S,A\}$, $T = \{a,b\}$ und

$$\begin{array}{rcl} P = \{S & \rightarrow & aSb \mid aAb \\ A & \rightarrow & bAa \mid ba\}. \end{array}$$

Folglich ist L_3 kontextsensitiv aber nicht kontextfrei.

Aufgabe 2 (Kontextfreie Grammatiken) *Betrachten Sie die Beispiel-Grammatik $G_{telescope}$ auf Folie 6 der Einführung (Sitzung vom 9.4.2014).*

1. *Bestimmen Sie die Menge der Nichtterminalsymbole N und die Menge der Terminalsymbole T dieser Grammatik.*
2. *Kann "John saw Mary with John" von dieser Grammatik erzeugt werden? Begründen Sie Ihre Antwort.*
3. *Zeigen Sie, dass es für jedes $n \geq 3$ ein Wort der Länge n gibt, welches von dieser Grammatik erzeugt wird.*

Lösung:

1. $N = \{S, NP, VP, D, N, PP, V\}$ und $T = \{man, girl, telescope, the, John, Mary, with, saw\}$.
2. Ja, z.B. durch die Ableitung S ⇒ NP VP ⇒ John VP ⇒ John VP PP ⇒ John V NP PP ⇒ John saw NP PP ⇒ John saw Mary PP ⇒ John saw Mary P NP ⇒ John saw Mary with NP ⇒ John saw Mary with John.

3. Ein Wort für $n = 3$ ist "John saw Mary".

   Für weitere Wörter ungerader Länge betrachte z.B. S $\Rightarrow$ NP VP $\stackrel{k}{\Rightarrow}$ NP VP PP^k $\Rightarrow$ NP V NP PP^k $\stackrel{k}{\Rightarrow}$ NP V NP (P NP)^k $\stackrel{2k+3}{\Rightarrow}$ John saw John (with John)^k. Dies ergibt für jedes $k \geq 0$ ein Wort der Länge $n = 2k + 3 = \{3,5,7,9,11,\ldots\}$.

   Ersetzt man nun am Wortanfang nicht NP $\Rightarrow$ John, sondern NP $\Rightarrow$ D N $\stackrel{2}{\Rightarrow}$ the girl, erhält man für $k \geq 0$ the girl saw John (with John)^k der Länge $n = 2k + 4 = \{4,6,8,10,12,..\}$.

   Insgesamt lässt sich für alle $n \geq 3$ ein Wort der Länge n mit dieser Grammatik generieren.

Aufgabe 3 (Knobelaufgabe) *Gegeben sei die Sprache $L = \{a^{2^n} \mid n \geq 0\}$ über dem Alphabet $X = \{a\}$. Es gilt also $L = \{a, a^2, a^4, a^8, a^{16}, a^{32}, \ldots\}$.*

1. *Geben Sie eine Grammatik mit (maximal) vier Produktionen an, die diese Sprache erzeugt.*
2. *Ist Ihre Grammatik kontextfrei? Begründen Sie Ihre Antwort.*

Lösung:

1. Betrachte $G = \langle N,T,P,S \rangle$ mit $N = \{A,B,S\}$, $T = \{a\}$ und

$$\begin{aligned} P = \{S &\rightarrow ASB \mid a \\ aB &\rightarrow Baa \\ AB &\rightarrow \varepsilon\}. \end{aligned}$$

2. G ist nicht kontextfrei, da z.B. die Regel $AB \rightarrow \varepsilon$ auf der linken Seite zwei Nichtterminale besitzt.

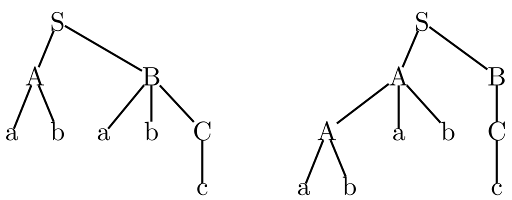
Aufgabe 4 (Parsbäume und Ambiguität) *Gegeben sei die Grammatik $G = \langle N,T,P,S \rangle$ mit $N = \{S,A,B,C\}$, $T = \{a,b,c\}$ und*

$$\begin{aligned} P = \{S &\rightarrow AB \\ A &\rightarrow Aab \mid ab \\ B &\rightarrow abC \mid C \\ C &\rightarrow cC \mid c\}. \end{aligned}$$

1. *Welche Sprache erzeugt die Grammatik G? Ordnen Sie die Sprache in die Chomsky-Hierarchie ein.*
2. *Zeigen Sie, dass die G ambig ist, indem Sie ein Wort finden zu dem Sie zwei verschiedene Parsbäume angeben.*
3. *Zeigen Sie, dass die Sprache $L(G)$ nicht inhärent ambig ist, indem Sie eine nicht ambige Grammatik G' mit $L(G') = L(G)$ angeben.*

Lösung:

1. Die Sprache ist $L = \{(ab)^m c^n \mid m,n \geq 1\}$, diese Sprache selbst ist sogar regulär.
2. Zwei Parsbäume für das Wort $w = ababc$:



3. Eine nicht ambige Grammatik für diese Sprache wäre z.B. $G' = \langle N,T,P,S \rangle$ mit $N = \{S,A,C\}$, $T = \{a,b,c\}$ und

$$\begin{aligned} P = \{S &\rightarrow AC \\ A &\rightarrow Aab|ab \\ C &\rightarrow cC|c\}. \end{aligned}$$

Aufgabe 5 (Normalformen; korrigiert) *Betrachten Sie die Grammatik* $G = \langle N,T,P,S \rangle$ *mit* $N = \{S,A,B,C\}$, $T = \{a,b,c\}$ *und*

$$\begin{aligned} P = \{S &\rightarrow AB \\ A &\rightarrow Aab|ab \\ B &\rightarrow abC|c \\ C &\rightarrow cC|c\}. \end{aligned}$$

1. *Überführen Sie G in eine äquivalente Grammatik* G_{CNF} *in CNF.*
2. *Überführen Sie G in eine äquivalente Grammatik* G_{GNF} *in GNF, ordnen Sie die Nichtterminasymbole dabei alphabetisch.*

Lösung:

1. G enthält keine ε-Produktionen und keine einfachen Regeln. Führe neue Nichtterminale mit zugehörigen Regeln ein und ersetze Terminalsymbole auf der rechten Seiten mit Länge > 1:

$$\begin{aligned} P = \{S &\rightarrow AB \\ A &\rightarrow AT_aT_b|T_aT_b \\ B &\rightarrow T_aT_bC|c \\ C &\rightarrow T_cC|c \\ T_a &\rightarrow a \\ T_b &\rightarrow b \\ T_c &\rightarrow c\}. \end{aligned}$$

   Ersetze nun noch die Nichtterminalen auf rechten Seiten mit Länge > 2 und erhalte die Grammatik $G_{CNF} = \langle N,T,P,S \rangle$ mit $N = \{S,A,B,C,T_a,T_b,T_c,D_1,D_2\}$, $T = \{a,b,c\}$ und

$$\begin{aligned} P = \{S &\rightarrow AB \\ A &\rightarrow AD_1|T_aT_b \\ D_1 &\rightarrow T_aT_b \\ B &\rightarrow T_aD_2|c \\ D_2 &\rightarrow T_bC \\ C &\rightarrow T_cC|c \\ T_a &\rightarrow a \\ T_b &\rightarrow b \\ T_c &\rightarrow c\}. \end{aligned}$$

2.

3. Für GNF eliminiere die Linksrekursion bei $A \to Aab|ab$ zu

$$\begin{array}{rcl} A & \to & abX|ab \\ X & \to & abX|ab \end{array}$$

mit einem neuen Nichtterminal X[1].

Bei der Regel $S \to AB$ werden nun die rechten Seiten von A eingesetzt: $S \to abXB|abB$.

Ersetze nun noch Regeln der Form $A \to \alpha a\beta$ mit $\alpha \neq \varepsilon$, dies trifft nur auf b zu, es genügt also ein neues Nichtterminal $Z \to b$. Es ergibt sich die Grammatik $G_{GNF} = \langle N,T,P,S\rangle$ mit $N = \{S_0,A_1,B_2,C_3,X_4,Z_5\}$, $T = \{a,b,c\}$ und

$$\begin{array}{rcl} P = \{S_0 & \to & aZ_5X_4B_2|aZ_5B_2 \\ A_1 & \to & aZ_5|aZ_5X_4 \\ B_2 & \to & aZ_5C_3|c \\ C_3 & \to & cC_3|c \\ X_4 & \to & aZ_5|aZ_5X_4 \\ Z_5 & \to & b\}, \end{array}$$

wobei die Nummerierung der Nichtterminalen mit angegeben ist.

Hinweis: Da A_1 auf keiner rechten Seite mehr auftaucht, können diese Regeln auch noch entfernt werden.

Aufgabe 6 (Kellerautomaten; korrigiert) *Gegeben sei die Sprache* $L = \{a^m b^n c d^n e^m \mid m,n \geq 0\}$. *Konstruieren Sie einen Kellerautomaten* $M = \langle Q,\Sigma,\Gamma,\delta,q_0,\#,F\rangle$ *mit* $L(M) = L$.

Lösung:

$Q = \{q_0,q_1,q_2,q_f\}$, $\Sigma = \{a,b,c,d,e\}$, $\Gamma = \{\#,A,B\}$, $F = \{q_f\}$ und δ wie folgt:

$$\begin{array}{lll} \delta(q_0,a,\varepsilon) = \{(q_0,A)\} & \delta(q_0,b,\varepsilon) = \{(q_1,B)\} & \delta(q_0,c,\varepsilon) = \{(q_2,\varepsilon)\} \\ \delta(q_1,b,\varepsilon) = \{(q_1,B)\} & & \delta(q_1,c,\varepsilon) = \{(q_2,\varepsilon)\} \\ \delta(q_2,d,B) = \{(q_2,\varepsilon)\} & \delta(q_2,e,A) = \{(q_2,\varepsilon)\} & \delta(q_2,\varepsilon,\#) = \{(q_f,\#)\} \end{array}$$

Aufgabe 7 (Unger-Parsing 1) *Gegeben sei die Grammatik* $G = \langle N,T,P,S\rangle$ *mit* $N = \{S,A\}$, $T = \{a,b\}$ *und*

$$\begin{array}{rcl} P = \{S & \to & aSa|A \\ A & \to & bA|b\}. \end{array}$$

1. *Bestimmen Sie mit Ungers Parser, ob das Wort* $w_1 = aba$ *zur Sprache* $L(G)$ *gehört.*
2. *Bestimmen Sie die Sprache* $L(G)$.

Lösung:

1. Es ergibt sich der folgende Trace

[1] Man könnte die Grammatik an dieser Stelle vereinfachen, da A und X die gleiche rechte Seite haben.

call	results	productions
unger($_1aba_3$,S)		
$S \to aSa$?		
unger($_1a_1$,a) $\to$ `true`	$\langle a, {}_1a_1, \texttt{true}\rangle$	
unger($_2b_2$,S)		
$S \to aSa$?		
first terminal does not match		
$S \to A$?		
unger($_2b_2$,A)		
$A \to bA$?		
no partition		
$A \to b$?		
unger($_2b_2$,b) $\to$ `true`	$\langle b, {}_2b_2, \texttt{true}\rangle$	
$\to$ `true`	$\langle A, {}_2b_2, \texttt{true}\rangle$	${}_2A_2 \to {}_2b_2$
unger($_3a_3$,a) $\to$ `true`	$\langle a, {}_3a_3, \texttt{true}\rangle$	${}_1S_3 \to {}_1a_1\, {}_2A_2\, {}_3a_3$
$S \to A$?		
unger($_1aba_3$,A)		
$A \to bA$?		
first terminal does not match		
$A \to b$?		
first terminal does not match		
$\to$ `false`	$\langle A, {}_1aba_3, \texttt{false}\rangle$	
$\to$ `true`	$\langle S, {}_1aba_3, \texttt{true}\rangle$	

2. $L(G) = \{a^n b^m a^n \mid m,n \geq 1\}$

Aufgabe 8 (Unger-Parsing 2) *Betrachten Sie die Grammatik* $G = \langle N,T,P,S\rangle$ *mit* $N = \{S,B,C\}$, $T = \{a,b,c\}$ *und*

$$\begin{aligned} P = \{S &\to aBC \\ B &\to bC \\ C &\to cS|c\}. \end{aligned}$$

Gegeben sei weiterhin der Trace aus Tabelle 1 vom Unger-Parser mit einigen der aus der Vorlesung bekannten Optimierungen (angegeben sind nur die calls*).*

Geben Sie die results *und die* productions *an, die sich hieraus ergeben.*

Lösung: Siehe Tabelle 2.

Aufgabe 9 (Top-Down-Parsing) *Gegeben sei die Grammatik* $G = \langle N,T,P,S\rangle$ *mit* $N = \{S,A,B\}$, $T = \{a,b\}$ *und*

$$\begin{aligned} P = \{S &\to AB|BA \\ A &\to a|AS \\ B &\to b|BS\}. \end{aligned}$$

1. *Finden Sie alle Parsbäume für das Eingabewort* $w = abab$.

2. *Geben Sie den Trace für den direktionalen Top-Down Parser an, wenn Sie Tiefensuche verwenden.*

Lösung:

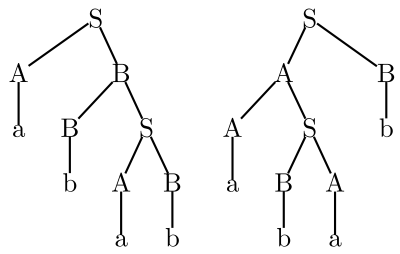
call	results	productions
unger($_1abcc_4$,S)		
$S \to aBC$?		
unger($_1ab_2, a$) $\to$ `false`		
unger($_3c_3, B$) $\to$ `false`		
unger($_4c_4, C$)		
$C \to cS$?		
no partition		
$C \to c$?		
unger($_4c_4, c$) $\to$ `true`		
$\to$ `true`		
unger($_1a_1, a$) $\to$ `true`		
unger($_2b_2, B$)		
$B \to bC$?		
no partition		
$\to$ `false`		
unger($_3cc_4, C$)		
$C \to cS$?		
unger($_3c_3$,c) $\to$ `true`		
unger($_4c_4$,S) $\to$ `false`		
$C \to c$?		
unger($_3cc_4$,c) $\to$ `false`		
$\to$ `false`		
unger($_1a_1, a$) $\to$ `true`		
unger($_2bc_3, B$)		
$B \to bC$?		
unger($_2b_2, b$) $\to$ `true`		
unger($_3c_3, C$)		
$C \to cS$?		
no partition		
$C \to c$?		
unger($_3c_3, c$) $\to$ `true`		
$\to$ `true`		
$\to$ `true`		
unger($_4c_4, C$) $\to$ `true`		
$\to$ `true`		

Table 1: Trace zu Aufgabe 8.

call	results	productions
unger($_1abcc_4$,S)		
$S \to aBC$?		
unger($_1ab_2, a$) $\to$ `false`	$\langle a, {}_1ab_2, false\rangle$	
unger($_3c_3, B$) $\to$ `false`	$\langle B, {}_3c_3, false\rangle$	
unger($_4c_4, C$)		
$C \to cS$?		
no partition		
$C \to c$?		
unger($_4c_4, c$) $\to$ `true`	$\langle c, {}_4c_4, true\rangle$	${}_4C_4 \to {}_4c_4$
$\to$ `true`	$\langle C, {}_4c_4, true\rangle$	
unger($_1a_1, a$) $\to$ `true`	$\langle a, {}_1a_1, true\rangle$	
unger($_2b_2, B$)		
$B \to bC$?		
no partition		
$\to$ `false`	$\langle B, {}_2b_2, false\rangle$	
unger($_3cc_4, C$)		
$C \to cS$?		
unger($_3c_3$,c) $\to$ `true`	$\langle c, {}_3c_3, true\rangle$	
unger($_4c_4$,S) $\to$ `false`	$\langle S, {}_4c_4, false\rangle$	
$C \to c$?		
unger($_3cc_4$,c) $\to$ `false`	$\langle c, {}_3cc_4, false\rangle$	
$\to$ `false`	$\langle C, {}_3cc_4, false\rangle$	
unger($_1a_1, a$) $\to$ `true`		
unger($_2bc_3, B$)		
$B \to bC$?		
unger($_2b_2, b$) $\to$ `true`	$\langle b, {}_2b_2, true\rangle$	
unger($_3c_3, C$)		
$C \to cS$?		
no partition		
$C \to c$?		
unger($_3c_3, c$) $\to$ `true`		${}_3C_3 \to {}_3c_3,\ {}_2B_3 \to {}_2b_2\,{}_3C_3$
$\to$ `true`	$\langle C, {}_3c_3, true\rangle$	
$\to$ `true`	$\langle B, {}_2bc_3, true\rangle$	
unger($_4c_4, C$) $\to$ `true`		${}_1S_4 \to {}_1a_1\,{}_2B_3\,{}_4C_4$
$\to$ `true`	$\langle S, {}_1abcc_4, true\rangle$	

Table 2: Lösung zu Aufgabe 8.

1. Die Parsbäume



2. Der Trace

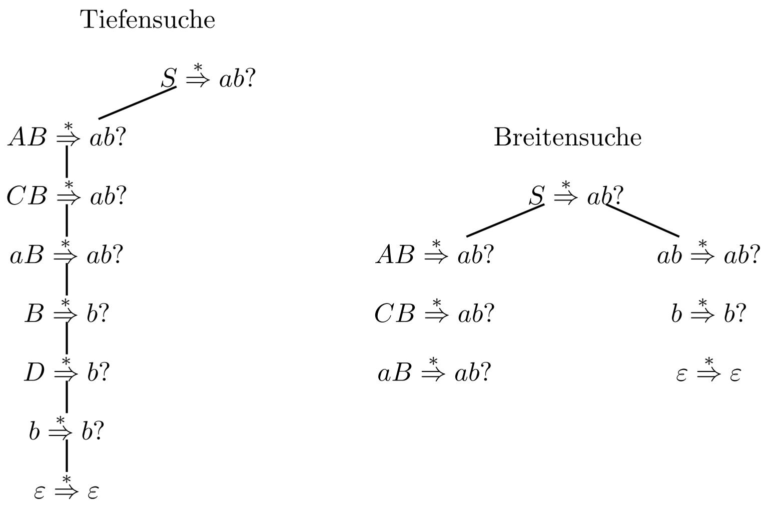
Resteingabe	Stack α	Analysestack
abab	S	
abab	AB	$\langle S{,}1\rangle$
abab	aB	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$
bab	B	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a
bab	b	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}1\rangle$
ab	ε	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}1\rangle$ b
bab	BS	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}2\rangle$
bab	bS	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}2\rangle$ $\langle B{,}1\rangle$
ab	S	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}2\rangle$ $\langle B{,}1\rangle$ b
ab	AB	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}2\rangle$ $\langle B{,}1\rangle$ b $\langle S{,}1\rangle$
ab	aB	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}2\rangle$ $\langle B{,}1\rangle$ b $\langle S{,}1\rangle$ $\langle A{,}1\rangle$
b	B	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}2\rangle$ $\langle B{,}1\rangle$ b $\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a
b	b	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}2\rangle$ $\langle B{,}1\rangle$ b $\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}1\rangle$
ε	ε	$\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}2\rangle$ $\langle B{,}1\rangle$ b $\langle S{,}1\rangle$ $\langle A{,}1\rangle$ a $\langle B{,}1\rangle$ b

Aufgabe 10 (Kontrollstrukturen) *Finden Sie eine Grammatik G und ein Eingabewort $w \in L(G)$, sodass beim direktionalen Top-Down Parsing die Breitensuche schneller die erste erfolgreiche Analyse findet, als die Tiefensuche. Geben Sie die entsprechenden Entscheidungsbäume an – der Entscheidungsbaum zur Breitensuche muss also weniger Knoten haben als der Entscheidungsbaum zur Tiefensuche.*

Lösung: Betrachte z.B. $G = \langle N{,}T{,}P{,}S\rangle$ mit $N = \{S{,}A{,}B{,}C{,}D\}$, $T = \{a{,}b\}$ und

$$\begin{array}{rcrcl} P & = & \{S & \to & AB|ab \\ & & A & \to & C \\ & & B & \to & D \\ & & C & \to & a \\ & & D & \to & b\}, \end{array}$$

sowie das Eingabewort $w = ab$. Die entsprechenden Entscheidungsbäume sehen wie folgt aus.



Aufgabe 11 (Unger-Parsing mit Deduktionsregeln) *Gegeben seien die Grammatik* $G = \langle N,T,P,S\rangle$ *mit* $N = \{S,A\}$, $T = \{a,b,c\}$ *und*

$$\begin{array}{rcrcl} P & = & \{S & \to & aSc|aA|ac \\ & & A & \to & cA|c\} \end{array}$$

und das Eingabewort $w = ac$. *Betrachen Sie Ungers Parser mit Deduktionsregeln.*

1. *Geben Sie die Items an, die der Parser für* w *generiert. Geben Sie zusätzlich die Antecedents an.*
2. *Ist* $w \in L(G)$*? Begründen Sie Ihre Antwort.*

Lösung:

1.

	Item	Operation	Antecedent
1	$[\bullet S, 0, 2]$	axiom	–
2	$[\bullet a,0,1]$	predict	1
3	$[\bullet A,1,2]$	predict	1
4	$[\bullet c,1,2]$	predict	1
5	$[a\bullet,0,1]$	scan	2
6	$[c\bullet,1,2]$	scan	4
7	$[A\bullet,1,2]$	complete	3,6
8	$[S\bullet, 0, 2]$	complete	1,5,6 oder 1,5,7

2. Da $[S\bullet, 0{,}2]$ ein Goal-Item und in der Chart ist, gilt $w \in L(G)$.

Aufgabe 12 (Top-Down-Parsing mit Deduktionsregeln) *Gegeben seien die Grammatik* $G = \langle N,T,P,S\rangle$ *mit* $N = \{S,A,B\}$, $T = \{a,b\}$ *und*

$$\begin{array}{rcrcl} P & = & \{S & \to & aB|bA \\ & & A & \to & a|aS|bAA \\ & & B & \to & b|bS|aBB\} \end{array}$$

und das Eingabewort $w = abba$. *Betrachen Sie die Deduktionsregeln für Top-Down Parsing.*

1. *Geben Sie die Items an, die der Parser für w generiert. Geben Sie zusätzlich die Antecedents an.*
2. *Ist $w \in L(G)$? Begründen Sie Ihre Antwort.*

Lösung:

1.

	Item	Operation	Antecedent
1	$[S, 0]$	axiom	–
2	$[aB,0]$	predict	1
3	$[bA,0]$	predict	1
4	$[B,1]$	scan	2
5	$[b,1]$	predict	4
6	$[bS,1]$	predict	4
7	$[aBB,1]$	predict	4
8	$[\varepsilon,2]$	scan	5
9	$[S,2]$	scan	6
10	$[aB,2]$	predict	9
11	$[bA,2]$	predict	9
12	$[A,3]$	scan	11
13	$[a,3]$	predict	12
14	$[aS,3]$	predict	12
15	$[bAA,3]$	predict	12
16	$[\varepsilon,4]$	scan	13
17	$[S,4]$	scan	14
18	$[aB,4]$	predict	17
19	$[bA,4]$	predict	17

2. Da $[\varepsilon,4]$ ein Goal-Item und in der Chart ist, gilt $w \in L(G)$.

Aufgabe 13 (Allgemeines CYK) *Gegeben seien die Grammatik $G = \langle N,T,P,S \rangle$ mit $N = \{S,A\}$, $T = \{a,b\}$ und*

$$
\begin{aligned}
P = \{S &\rightarrow Sb|A \\
A &\rightarrow aS|\varepsilon\}
\end{aligned}
$$

und das Eingabewort $w = abb$.

1. *Welche Chart erzeugt der allgemeine CYK-Recognizer bei Eingabe G und w?*
2. *Gilt $w \in L(G)$? Begründen Sie Ihre Antwort mit Hilfe der Chart.*

Lösung:

1. Chart:

$l \downarrow$				
3	S, A			
2	S, A	S		
1	a, S, A	b, S	b, S	
0	S, A	S, A	S, A	S, A
$i \rightarrow$	1	2	3	4

2. $w \in L(G)$, da wir ein S mit Anfangsposition 1 und Länge 3 gefunden haben.

Aufgabe 14 (CYK-Parsing) *Gegeben seien die Grammatik $G = \langle N,T,P,S \rangle$ mit $N = \{S\}$, $T = \{a\}$ und*

$$P = \{S \rightarrow SS|a\}$$

und das Eingabewort $w = aaaa$.

Verwenden Sie den CYK-Parser für kontextfreie Grammatiken in CNF, der mit Indizes annotierte Produktionen in der Chart ablegt. Wie sieht die Chart aus, die sich für G und w ergibt?

Lösung:

$l \downarrow$				
4	$S \rightarrow S_{1,1}S_{2,3}$ $S \rightarrow S_{1,2}S_{3,2}$ $S \rightarrow S_{1,3}S_{4,1}$			
3	$S \rightarrow S_{1,1}S_{2,2}$ $S \rightarrow S_{1,2}S_{3,1}$	$S \rightarrow S_{2,1}S_{3,2}$ $S \rightarrow S_{2,2}S_{4,1}$		
2	$S \rightarrow S_{1,1}S_{2,1}$	$S \rightarrow S_{2,1}S_{3,1}$	$S \rightarrow S_{3,1}S_{4,1}$	
1	$S \rightarrow a$	$S \rightarrow a$	$S \rightarrow a$	$S \rightarrow a$
$i \rightarrow$	1	2	3	4

Aufgabe 15 (Shift-Reduce-Parsing) *Gegeben sei die Grammatik $G = \langle N,T,P,S \rangle$ mit $N = \{S\}$, $T = \{a\}$ und*

$$P = \{S \rightarrow SS \,|\, a\}.$$

Geben Sie den Trace des Shift-Reduce-Parsers in der Form

#	Index	Stack	Resteingabe	Operation
1	0	ε	aaa	

für die Eingabe aaa an. Ist dieses Wort in $L(G)$ enthalten?

Lösung:

#	Index	Stack	Resteingabe	Operation
1	0	ε	aaa	
2	1	a	aa	shift 1 (a)
3	1	S	aa	reduce 2 ($S \to a$)
4	2	aa	a	shift 2 (a)
5	2	Sa	a	shift 3 (a)
6	2	aS	a	reduce 4 ($S \to a$)
7	3	aaa	ε	shift 4 (a)
8	2	SS	a	reduce 5 ($S \to a$)
9	3	Saa	ε	shift 5 (a)
10	3	aSa	ε	shift 6 (a)
11	3	aaS	ε	reduce 7 ($S \to a$)
12	2	S	a	reduce 8 ($S \to SS$)
13	3	SSa	ε	shift 8 (a)
14	3	SaS	ε	reduce 9 ($S \to a$)
15	3	aSS	ε	reduce 10 ($S \to a$)
16	3	Sa	ε	shift 12 (a)
17	3	SSS	ε	reduce 13 ($S \to a$)
18	3	aS	ε	reduce 15 ($S \to SS$)
19	3	SS	ε	reduce 16 ($S \to a$)
20	3	SS	ε	reduce 17 ($S \to SS$)
21	3	S	ε	reduce 19 ($S \to SS$)
22	3	S	ε	reduce 20 ($S \to SS$)

Da #21 und #22 Goal-Items sind, ist $aaa \in L(G)$.

Aufgabe 16 (LL(k)-Grammatiken) *Gegeben sei die Grammatik $G = \langle N,T,P,S\rangle$ mit $N = \{S,A,B\}$, $T = \{a,b\}$ und*

$$\begin{aligned} P = \{S &\to AB \\ A &\to aAa|\varepsilon \\ B &\to bBb|\varepsilon\}. \end{aligned}$$

1. *Bestimmen Sie die Sprache $L(G)$ und ordnen Sie die Sprache in die Chomsky-Hierarchie ein.*
2. *Bestimmen Sie die $First$- und $Follow$-Mengen.*
3. *Geben Sie die $LL(1)$-Parstabelle an.*
4. *Ist G eine $LL(1)$-Grammatik? Begründen Sie ihre Antwort.*

Lösung:

1. $L(G) = \{a^{2m}b^{2n} \mid m,n \geq 0\}$, diese Sprache ist regulär.
2. Die $First$- und $Follow$-Mengen der Nichtterminalsymbole:

$$\begin{aligned} First(S) = \{\varepsilon, a,b\} \quad &; \quad Follow(S) = \{\$\} \\ First(A) = \{\varepsilon, a\} \quad &; \quad Follow(A) = \{a,b,\$\} \\ First(B) = \{\varepsilon, b\} \quad &; \quad Follow(B) = \{b, \$\} \end{aligned}$$

   Die $First$-Mengen der rechten Seiten:

$$\begin{aligned} First(AB) &= \{a,b,\varepsilon\} \\ First(aAa) &= \{a\} \\ First(\varepsilon) &= \{\varepsilon\} \\ First(bBb) &= \{b\} \end{aligned}$$

3. Die $LL(1)$-Parstabelle

	S	A	B
a	$S \to AB$	$A \to aAa$ $A \to \varepsilon$	
b	$S \to AB$	$A \to \varepsilon$	$B \to bBb$ $B \to \varepsilon$
$\$$	$S \to AB$	$A \to \varepsilon$	$B \to \varepsilon$

4. Nein, denn $A \to \varepsilon$, $\varepsilon \in First(\varepsilon)$ und $Follow(A) \cap First(aAa) = \{a,b,\$\} \cap \{a\} = \{a\} \neq \emptyset$ – analog auch wegen $B \to \varepsilon$, $\varepsilon \in First(\varepsilon)$ und $Follow(B) \cap First(bBb) = \{b, \$\} \cap \{b\} = \{b\} \neq \emptyset$.

   Alternativ: Nein. denn es gibt in der $LL(1)$-Parstabelle Zellen mit mehr als einem Eintrag.

Aufgabe 17 (Left-Corner Parsing) *Betrachten Sie die Grammatik $G = \langle N, T, P, S\rangle$ mit $N = \{S\}$, $T = \{a,b\}$ und*

$$P = \{S \to aSb \mid aS \mid b\}.$$

Geben Sie den Trace, der sich beim Left-Corner-Parsing (nur Recognition) für die Eingabe abb ergibt, in der Form

#	Γ_{compl}	Γ_{td}	Γ_{lhs}	*Operation*
1	abb	S	ε	—
2	bb	$Sb\$S$	S	*Reduce von 1, $S \to aSb$*
⋮	⋮	⋮	⋮	⋮

an.

Lösung:

#	Γ_{compl}	Γ_{td}	Γ_{lhs}	Operation
1	abb	S	ε	—
2	bb	$Sb\$S$	S	Reduce von 1, $S \to aSb$
3	bb	$S\$S$	S	Reduce von 1, $S \to aS$
4	b	$\$Sb\$S$	SS	Reduce von 2, $S \to b$
5	b	$\$S\$S$	SS	Reduce von 3, $S \to b$
6	Sb	$Sb\$S$	S	Move von 4, S
7	Sb	$S\$S$	S	Move von 5, S
8	b	$b\$S$	S	Remove von 6, S
9	b	$\$S$	S	Remove von 7, S
10	ε	$\$S$	S	Remove von 8, b
11	ε	$\$b\$S$	SS	Reduce von 8, $S \to b$
12	Sb	S	ε	Move von 9, S
13	S	S	ε	Move von 10, S
14	S	$b\$S$	S	Move von 11, S
15	b	ε	ε	Remove von 12, S
16	ε	ε	ε	Remove von 13, $S \Longrightarrow$ GOAL
17	ε	$\$$	S	Reduce von 15, $S \to b$
18	S	ε	ε	Move von 17, S

Aufgabe 18 (Left-Corner Parsing) *Betrachten Sie die Deduktionsregeln für Left-Corner-Parsing.*

Erstellen Sie eine Regel

$$\varepsilon\text{-}Scan\text{:} \quad \frac{[antecedent]}{[consequent]} \quad side\ conditions,$$

die mit ε-Produktionen umgehen kann.

Lösung: Es ist eine zusätzliche Regel notwendig:

$$\varepsilon\text{-Scan:} \quad \frac{}{[A, i, 0]} \quad A \to \varepsilon \in P,\ 1 \leq i \leq n+1,$$

wobei n die Länge des Eingabewortes ist.

Aufgabe 19 (Earley Parsing) *Betrachten Sie die Grammatik $G = \langle N, T, P, S\rangle$ mit $N = \{S, A, B, C\}$, $T = \{a, b, c\}$ und*

$$\begin{array}{rcrcl} P & = & \{S & \to & A \\ & & A & \to & a \mid aBC \\ & & B & \to & b \mid bb \\ & & C & \to & c \mid \varepsilon\}. \end{array}$$

Gegeben ist der folgende Beginn der Tabulation zum Earley-Parsing für das Eingabewort $w = abb$.

3				
2	$A \to aB \bullet C$	$B \to b \bullet b$ $B \to b\bullet$		
1	$A \to a\bullet$ $A \to a \bullet BC$ $S \to A\bullet$	$B \to \bullet bb$ $B \to \bullet b$		
0	$S \to \bullet A$ $A \to \bullet a$ $A \to \bullet aBC$			
	0	*1*	*2*	*3*

Wie sehen zu diesem Zeitpunkt die Agenden zu den verschiedenen k aus?

Lösung: Die Agenda A_0 ist leer und die Agenda A_1 ist leer, da aktuell schon bei $k = 2$ gearbeitet wird. Entsprechend ist $A_2 = \{[B \to b \bullet b, 1, 2], [A \to aB \bullet C, 0, 2]\}$, da diese beiden Items schon in der Chart stehen aber noch nicht weiter verarbeitet wurden:

- $[B \to b \bullet b, 1, 2]$ würde ein $[B \to bb\bullet, 1, 3]$ zur Folge haben und
- $[A \to aB \bullet C, 0, 2]$ würde $[C \to \bullet c, 2, 2]$ und $[C \to \bullet, 2, 2]$ zur Folge haben.

Alle drei Items sind aber noch nicht vorhanden, darum können diese Items noch nicht aus der Agenda herausgenommen worden sein.

Aufgabe 20 (Earley Parsing) *Gegeben sei die Grammatik $G = \langle N, T, P, S\rangle$ mit $N = \{S, A, B, C\}$, $T = \{a, b\}$ und*

$$\begin{array}{rcrcl} P & = & \{S & \to & ABA \mid aCa \\ & & A & \to & a \mid aA \\ & & B & \to & bb \\ & & C & \to & bCb \mid \varepsilon\}. \end{array}$$

Geben Sie die Recognition-Chart vom Earley-Parser für das Eingabewort $w = abba$ an. Benutzen Sie für Predict und Complete folgende Regeln mit Lookahead:

Predict mit Lookahead: $$\frac{[A \to \alpha \bullet B\beta, i, j]}{[B \to \bullet\gamma, j, j]} \quad B \to \gamma \in P, w_{i+1} \in First(\gamma) \ \textit{or} \ \varepsilon \in First(\gamma)$$

Complete mit Lookahead: $$\frac{[A \to \alpha \bullet B\beta, i, j], [B \to \gamma\bullet, j, k]}{[A \to \alpha B \bullet \beta, i, k]} \quad w_{k+1} \in First(\beta) \ \textit{or} \ \varepsilon \in First(\beta)$$

Lösung:

4	$S \to ABA\bullet$ $S \to aCa\bullet$			$A \to a \bullet A$ $A \to a\bullet$	
3	$S \to AB \bullet A$ $S \to aC \bullet a$	$C \to bCb\bullet$ $B \to bb\bullet$	$C \to b \bullet Cb$	$A \to \bullet aA$ $A \to \bullet a$ $C \to \bullet$	
2		$C \to bC \bullet b$ $C \to b \bullet Cb$ $B \to b \bullet b$	$C \to \bullet bCb$ $C \to \bullet$		
1	$S \to A \bullet BA$ $A \to a\bullet$ $A \to a \bullet A$ $S \to a \bullet Ca$	$C \to \bullet bCb$ $B \to \bullet bb$ $C \to \bullet$			
0	$S \to \bullet aCa$ $S \to \bullet ABA$ $A \to \bullet a$ $A \to \bullet aA$				
	0	1	2	3	4

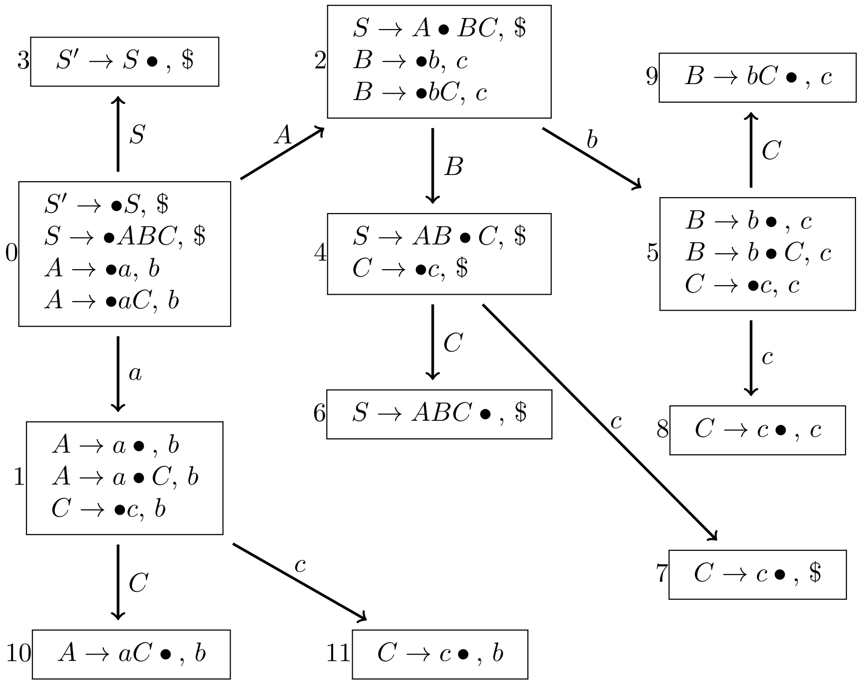
Aufgabe 21 (LR-Parsing) *Gegeben sei die Grammatik $G = \langle N, T, P, S\rangle$ mit $N = \{S,A,B,C\}$, $T = \{a,b,c\}$ und*

$$\begin{array}{rcl} P = \{S & \to & ABC \ (1) \\ A & \to & a \ (2) \mid aC \ (3) \\ B & \to & b \ (4) \mid bC \ (5) \\ C & \to & c \ (6)\}. \end{array}$$

1. *Konstruieren Sie die LR(1)-Parstabelle.*
2. *Ist G eine LR(1)-Grammatik? Begründen Sie Ihre Antwort.*
3. *Geben Sie den LALR-Automaten an.*
4. *Geben Sie den SLR-Automaten an.*

Lösung:

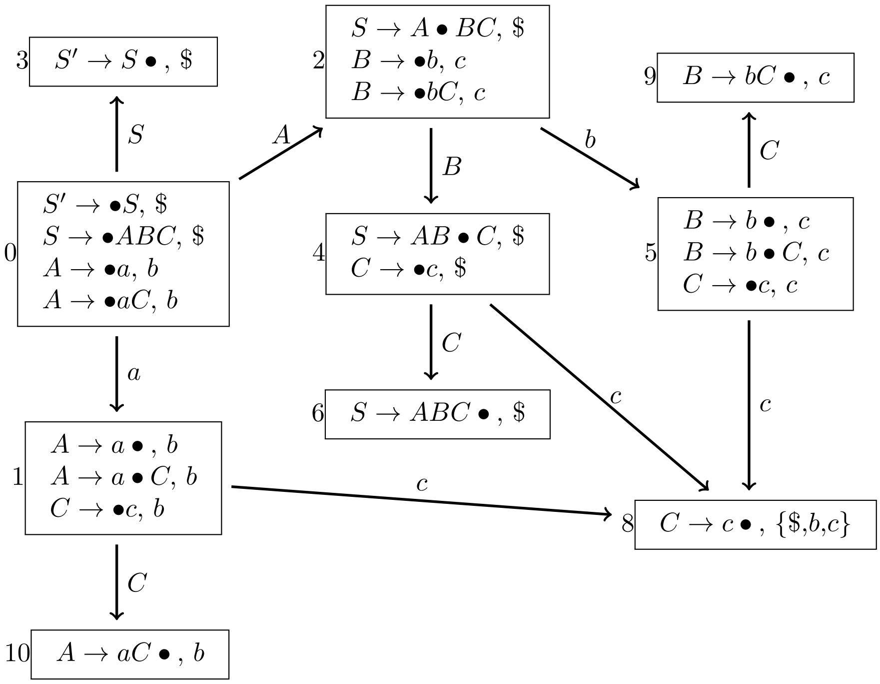
1. Zuerst der Automat:



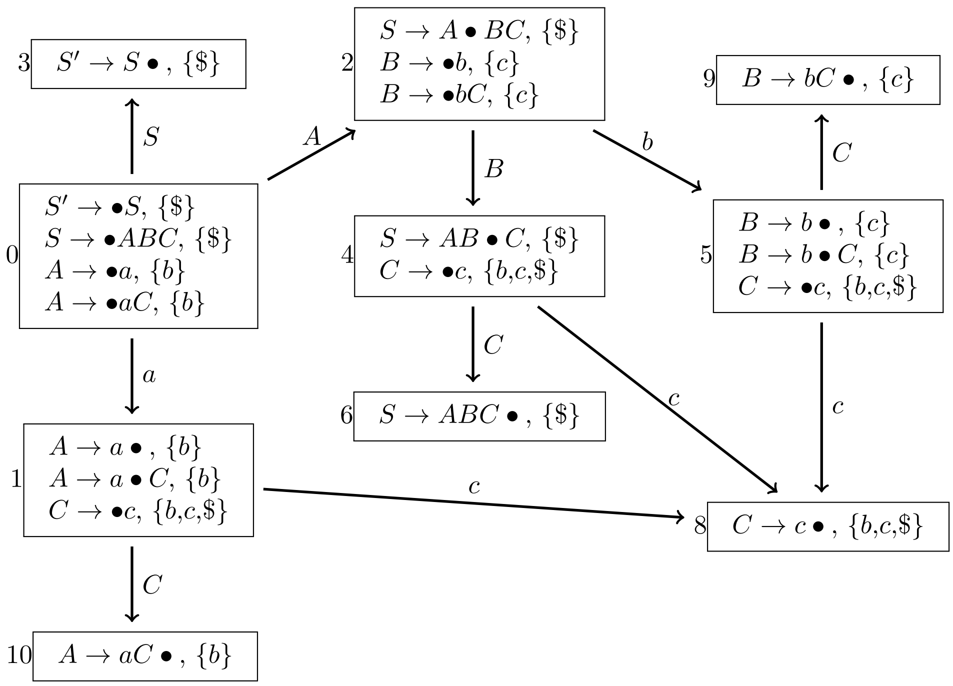
Dazu die LR(1)-Parstabelle

	action				goto			
	a	b	c	\$	A	B	C	S
0	s1				2			3
1		r2	s11				10	
2		s5				4		
3				acc				
4			s7				6	
5			s8, r4				9	
6				r1				
7				r6				
8			r6					
9			r5					
10		r3						
11		r6						

2. Die Grammatik ist keine LR(1)-Grammatik, da es bei Zustand 5 einen Shift-Reduce-Konflikt gibt.
3. Offensichtlich unterscheiden sich die Zustände 7, 8 und 11 nur durch ihren Lookahead, sie werden also zusammengefasst. Es ergibt sich folgender LALR-Automat:



4. Da $Follow(S') = \{\$\}$, $Follow(S) = \{\$\}$, $Follow(A) = \{b\}$, $Follow(B) = \{c\}$, $Follow(C) = \{\$,b,c\}$, ergibt sich der folgende SLR-Automat:



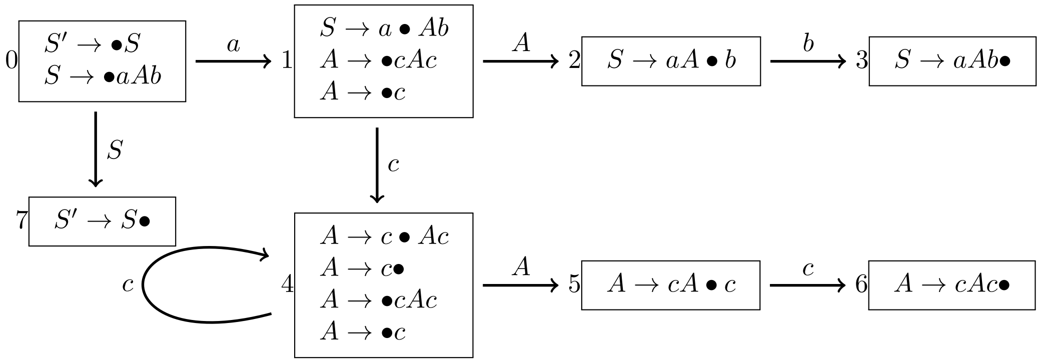
Aufgabe 22 (LR-Parsing) *Gegeben sei die Grammatik $G = \langle N, T, P, S\rangle$ mit $N = \{S,A\}$, $T = \{a,b,c\}$ und*

$$\begin{aligned} P = \{S &\to aAb \ (1) \\ A &\to cAc \ (2) \mid c \ (3)\}. \end{aligned}$$

1. *Konstruieren Sie die LR-Parstabelle*
2. *Zeigen Sie, dass G für kein k eine LR(k)-Grammatik ist.*

Lösung:

1. Zuerst der LR-Automat:



und nun die zugehörige Parstabelle:

	action				goto	
	a	b	c		A	S
0	s1					7
1			s4		2	
2		s3				
3				r1		
4			s4	r3	5	
5			s6			
6				r2		
7				acc		

Man erkennt an Zustand 4 schon, dass es sich nicht um eine LR(0)-Grammatik handeln kann.

2. Betrachte die beiden Ableitungen

$$S \Rightarrow_{rm}^{k+1} \underbrace{ac^k}_{\alpha}\underbrace{A}_{A}\underbrace{c^kb}_{w} \Rightarrow_{rm} ac^{2k+1}b = \underbrace{ac^k}_{\alpha}\underbrace{c}_{\beta}\underbrace{c^kb}_{w}$$
$$S \Rightarrow_{rm}^{k+2} \underbrace{ac^{k+1}}_{\gamma}\underbrace{A}_{B}\underbrace{c^{k+1}b}_{x} \Rightarrow_{rm} ac^{2k+3}b = \underbrace{ac^k}_{\alpha}\underbrace{c}_{\beta}\underbrace{c^{k+2}b}_{y}$$

und die Definition für LR(k)-Grammatiken: Es ist damit zwar $init_k(w) = c^k = init_k(y)$ für jedes k möglich, aber $\alpha = ac^k \neq ac^{k+1} = \gamma$. Folglich ist G für kein k eine LR(k)-Grammatik.

Aufgabe 23 (Tomita - korrigiert) *Gegeben sei die Grammatik $G = \langle N, T, P, S\rangle$ mit $N = \{S,A,B,X\}$, $T = \{a,b\}$ und*

$$\begin{aligned} P = \{S &\to ABA\ (1) \mid aXa\ (2) \\ X &\to bXb\ (3) \mid \varepsilon\ (4) \\ A &\to a\ (5) \mid aA\ (6) \\ B &\to bb\ (7)\}. \end{aligned}$$

Die zugehörige LR(1)-Parstabelle sei ebenfalls schon gegeben:

	action			*goto*			
	a	*b*	\$	*S*	*A*	*B*	*X*
0	*s1*			*4*	*5*		
1	*s8*,*r*4	*s2*,*r*5			*16*		*9*
2		*s3*,*r*4					*10*
3		*s3*,*r*4					*11*
4			*acc*				
5		*s13*				*6*	
6	*s14*				*7*		
7			*r*1				
8	*s8*	*r*5			*16*		
9	*s17*						
10		*s18*					
11		*s19*					
12	*r*7						
13		*s12*					
14	*s14*		*r*5		*15*		
15			*r*6				
16		*r*6					
17			*r*2				
18	*r*3						
19		*r*3					

Geben Sie den Trace für Tomitas Parser bei Eingabe $w = abba$ inkl. aller Zwischenschritte und Analysen an.

Lösung: 0 s1

0 — $\boxed{1}$ — 1 s2,r5 $\boxed{1}$: a

0 — $\boxed{1}$ — 1 s2
0 — $\boxed{2}$ — 5 s13 $\boxed{2}$: A($\boxed{1}$)

0 — $\boxed{1}$ — 1 — $\boxed{3}$ — 2 s3,r4 $\boxed{3}$: b
0 — $\boxed{2}$ — 5 — $\boxed{4}$ — 13 s12 $\boxed{4}$: b

2 — $\boxed{5}$ — 10 s18 $\boxed{5}$: X()
0 — $\boxed{1}$ — 1 — $\boxed{3}$ — 2 s3
0 — $\boxed{2}$ — 5 — $\boxed{4}$ — 13 s12

2 — $\boxed{5}$ — 10 — $\boxed{6}$ — 18 r3 $\boxed{6}$: b
0 — $\boxed{1}$ — 1 — $\boxed{3}$ — 2 — $\boxed{7}$ — 3 – $\boxed{7}$: b
0 — $\boxed{2}$ — 5 — $\boxed{4}$ — 13 — $\boxed{8}$ — 12 r7 $\boxed{8}$: b

0 — $\boxed{1}$ — 1 — $\boxed{9}$ — 9 s17 $\boxed{9}$: X($\boxed{3}$,$\boxed{5}$,$\boxed{6}$)
0 — $\boxed{2}$ — 5 — $\boxed{4}$ — 13 — $\boxed{8}$ — 12 r7 $\boxed{8}$: b

0 — $\boxed{1}$ — 1 — $\boxed{9}$ — 9 s17
0 — $\boxed{2}$ — 5 — $\boxed{10}$ — 6 s14 $\boxed{10}$: B($\boxed{4}$,$\boxed{8}$)

0 — $\boxed{1}$ — 1 — $\boxed{9}$ — 9 — $\boxed{11}$ — 17 r2 $\boxed{11}$: a
0 — $\boxed{2}$ — 5 — $\boxed{10}$ — 6 — $\boxed{12}$ — 14 r5 $\boxed{12}$: a

0 — $\boxed{13}$ — 4 acc $\boxed{13}$: S($\boxed{1}$,$\boxed{9}$,$\boxed{11}$)
$\boxed{2}$ — 5 — $\boxed{10}$ — 6 — $\boxed{12}$ — 14 r5

0 — $\boxed{13}$ — 4 acc
$\boxed{2}$ — 5 — $\boxed{10}$ — 6 — $\boxed{14}$ — 7 r1 $\boxed{14}$: A($\boxed{12}$)

0 — $\boxed{16}$ — 4 acc $\boxed{15}$: S($\boxed{2}$,$\boxed{10}$,$\boxed{14}$), $\boxed{16}$:[$\boxed{13}$,$\boxed{15}$]

Aufgabe 24 (PCFG) *Gegeben sei die Grammatik $G = \langle N,T,P,S,p\rangle$ mit $N = \{S,A,B\}$, $T = \{a,b\}$ und*

$$P = \{ \begin{array}{llll} 0{,}5 & S & \rightarrow & AS \\ 0{,}3 & S & \rightarrow & SB \\ 0{,}2 & S & \rightarrow & AB \\ 1 & A & \rightarrow & a \\ 1 & B & \rightarrow & b\,\}. \end{array}$$

Die Wahrscheinlichkeiten $p(N \rightarrow \gamma)$ sind jeweils vor den Produktionen angegeben.

1. *Bestimmen Sie die Chart der Inside-Wahrscheinlichkeiten für $w = aaabbb$.*
2. *Bestimmen Sie die Chart des probabilistischen CYK-Parsings für $w = aaabbb$.*

Lösung:

1. Chart der Inside-Wahrscheinlichkeiten:

6	(S, 0.027)	(S, 0.027)	(S, 0.018)			(B, 1)
5	(S, 0.045)	(S, 0.06)	(S, 0.06)		(B, 1)	
4	(S, 0.05)	(S, 0.1)	(S, 0.2)	(B, 1)		
3			(A, 1)			
2		(A, 1)				
1	(A, 1)					
	1	2	3	4	5	6

2. Der probabilistische CYK:

6	0.0045, $S \rightarrow AS$, 1					
5	0.015, $S \rightarrow AS$, 1	0.009, $S \rightarrow AS$, 1				
4	0.05, $S \rightarrow AS$, 1	0.03, $S \rightarrow AS$, 1	0.018, $S \rightarrow SB$, 3			
3		0.1, $S \rightarrow AS$, 1	0.06, $S \rightarrow SB$, 2			
2			0.2, $S \rightarrow AB$, 1			
1	1, $A \rightarrow a$	1, $A \rightarrow a$	1, $A \rightarrow a$	1, $B \rightarrow b$	1, $B \rightarrow b$	1, $B \rightarrow b$
	1	2	3	4	5	6

Hinweis: Bei einigen Feldern der Chart existieren mehrere mögliche Produktionen mit der gleichen Wahrscheinlichkeit.

Aufgabe 25 (Weighted Deductive Parsing) *Gegeben sei die Grammatik $G = \langle N,T,P,S,p\rangle$ mit $N = \{S,A,B\}$, $T = \{a,b\}$ und*

$$P = \{ 0{,}25 \quad S \rightarrow AS$$

$$\begin{array}{rlcl}
0{,}25 & S & \rightarrow & BS \\
0{,}5 & S & \rightarrow & c \\
0{,}875 & A & \rightarrow & a \\
0{,}125 & A & \rightarrow & b \\
0{,}125 & B & \rightarrow & a \\
0{,}875 & B & \rightarrow & b\,\}.
\end{array}$$

Die Wahrscheinlichkeiten $p(N \rightarrow \gamma)$ sind jeweils vor den Regeln angegeben.

1. *Bestimmen Sie mit dem deduktionsbasierten, probabilistischen CYK-Algorithmus die beste Ableitung für $w = abc$. Welche Items befinden sich in der Agenda, nachdem das (erste) Goal-Item in die Chart geschrieben wurde?*

2. *Bestimmen Sie die Chart des deduktionsbasierten, probabilistischen Left-Corner Parsers für $w = bac$.*

Benutzen Sie für die Gewichte den Logarithmus zur Basis 2 und runden Sie ggf. auf eine Nachkommastelle.

Lösung:

1. Die Chart:

3	$5{,}4 : \ [S \rightarrow AS,\ 1]$		
2		$3{,}2 : \ [S \rightarrow BS,\ 1]$	
1	$0{,}2 : \ [A \rightarrow a]$ $3 : \ [B \rightarrow a]$	$0{,}2 : \ [B \rightarrow b]$ $3 : \ [A \rightarrow b]$	$1 : \ [S \rightarrow c]$
	1	2	3

Die beste Ableitung ist folglich $S \Rightarrow AS \Rightarrow aS \Rightarrow aBS \Rightarrow abS \Rightarrow abc$. Die Agenda ist im Anschluss leer.

2. Nur die Chart:

3	$3{,}4 : \ S \rightarrow BS\bullet$ $5{,}4 : \ S$	$1{,}2 : \ S \rightarrow AS\bullet$ $3{,}2 : \ S$	$0 : \ c$ $0 : \ S \rightarrow c\bullet$ $1 : \ S$
2		$0 : \ a$ $0 : \ A \rightarrow a\bullet$ $0 : \ B \rightarrow a\bullet$ $0{,}2 : \ A$ $0{,}2 : \ S \rightarrow A \bullet S$ $3 : \ B$ $3 : \ S \rightarrow B \bullet S$	
1	$0 : \ b$ $0 : \ A \rightarrow b\bullet$ $0 : \ B \rightarrow b\bullet$ $0{,}2 : \ B$ $0{,}2 : \ S \rightarrow B \bullet S$ $3 : \ A$ $3 : \ S \rightarrow A \bullet S$		
	0	1	2

Aufgabe 26 (A^*-Parsing) *Betrachen Sie die Beispiel-Grammatik aus den A^*-Folien:* $G = \langle\{N,A\}, \{camping, car, nice, red, ugly, green, house, bike\}, P, N, p\rangle$ *mit*

$$
\begin{array}{rrrcl}
P = \{ & 0{,}1(1) & N & \to & NN \\
 & 0{,}2(0{,}7) & N & \to & AN \\
 & 0{,}1(1) & N & \to & red \\
 & 0{,}1(1) & N & \to & green \\
 & 0{,}1(1) & N & \to & car \\
 & 0{,}1(1) & N & \to & bike \\
 & 0{,}2(0{,}7) & N & \to & camping \\
 & 0{,}1(1) & N & \to & house \\
 & 0{,}3(0{,}5) & A & \to & nice \\
 & 0{,}25(0{,}6) & A & \to & ugly \\
 & 0{,}2(0{,}7) & A & \to & red \\
 & 0{,}25(0{,}6) & A & \to & green\,\}.
\end{array}
$$

Die Wahrscheinlichkeiten $p(N \to \gamma)$ sind jeweils vor den Regeln angegeben, in Klammern dahinter die jeweiligen Beträge der Logarithmen (gerundet).

Berechnen Sie das Gewicht der besten Analyse des Satzes "red ugly green house" mit dem CYK-Algorithmus bei Verwendung von SX-Schätzern.
Hinweis: die zugehörigen Deduktionsregeln sind auf Folie 18 zum A^*-Parsing angegeben.

Lösung:

4	$5: N \to AN, 1$	$4{,}8: N \to AN, 1$	$5{,}2: N \to AN, 1$	$5{,}1: N \to house$
3	$6: N \to NN, 1$	$5{,}2: N \to AN, 1$	$4{,}9: A \to green$ $5{,}1: N \to green$	
2		$4{,}4: A \to ugly$		
1	$4{,}5: A \to red$ $4{,}8: N \to red$			
(i,j)	0	1	2	3