Idee zu einem Parser:

- Der Parser verwaltet ein zuverlässiges Präfix $\alpha = X_1 \ldots X_m$ auf dem Keller und benutzt $LR(G)$, um Reduktionsstellen zu entdecken.
- Er kann mit einer Regel $A \to \gamma$ reduzieren, falls $[A \to \gamma \bullet]$ für α gültig ist :-)
- Damit der Automat nicht immer wieder neu über den Kellerinhalt laufen muss, kellern wir anstelle der X_i jeweils die Zustände !!!

Achtung:

Dieser Parser ist nur dann deterministisch, wenn jeder Endzustand des kanonischen $LR(0)$-Automaten keine Konflikte enthält ...

... im Beispiel:

$$q_1 = \{[S' \to E \bullet], \\ [E \to E \bullet + T]\}$$

$$q_2 = \{[E \to T \bullet], \\ [T \to T \bullet * F]\}$$

$$q_9 = \{[E \to E + T \bullet], \\ [T \to T \bullet * F]\}$$

$$q_3 = \{[T \to F \bullet]\}$$

$$q_{10} = \{[T \to T * F \bullet]\}$$

$$q_4 = \{[F \to \mathsf{int} \bullet]\}$$

$$q_{11} = \{[F \to (\, E\,) \bullet]\}$$

Die Endzustände q_1, q_2, q_9 enthalten mehr als ein Item :-(

Aber wir haben ja auch noch nicht Vorausschau eingesetzt :-)

Die Konstruktion des $LR(0)$-Parsers:

Zustände: $Q \cup \{f\}$ (f neu :-)

Anfangszustand: q_0

Endzustand: f

Übergänge:

Shift: $(p, a, p\,q)$ falls $q = \delta(p, a) \neq \emptyset$

Reduce: $(p\,q_1 \ldots q_m, \epsilon, p\,q)$ falls $[A \to X_1 \ldots X_m \bullet] \in q_m$, $q = \delta(p, A)$

Finish: $(q_0\,p, \epsilon, f)$ falls $[S' \to S \bullet] \in p$

wobei $LR(G) = (Q, T, \delta, q_0, F)$.

Zur Korrektheit:

Man zeigt:

Die akzeptierenden Berechnungen des $LR(0)$-Parsers stehen in eins-zu-eins Beziehung zu denen des Shift-Reduce-Parsers $M_G^{(1)}$.

Wir folgern:

$\Longrightarrow$ Die akzeptierte Sprache ist genau $\mathcal{L}(G)$:-)

$\Longrightarrow$ Die Folge der Reduktionen einer akzeptierenden Berechnung für ein Wort $w \in T$ liefert eine reverse Rechts-Ableitung von G für w :-)

Leider ist der $LR(0)$-Parser i.a. nicht-deterministisch :-(

Wir identifizieren zwei Gründe:

Reduce-Reduce-Konflikt:

$$[A \to \gamma \bullet] \,,\quad [A' \to \gamma' \bullet] \quad \in \quad q \quad \text{mit} \quad A \neq A' \vee \gamma \neq \gamma'$$

Shift-Reduce-Konflikt:

$$[A \to \gamma \bullet] \,,\quad [A' \to \alpha \bullet a\, \beta] \quad \in \quad q \quad \text{mit} \quad a \in T$$

für einen Zustand $q \in Q$.

Solche Zustände nennen wir ungeeignet.

Idee:

Benutze k-Vorausschau, um Konflikte zu lösen.

Wir definieren:

Die reduzierte kontextfreie Grammatik G heißt $LR(k)$-Grammatik, falls für $\mathsf{First}_k(w) = \mathsf{First}_k(x)$ aus:

$$\left.\begin{array}{llllll} S & \rightarrow_R^* & \alpha\, A\, w & \rightarrow & \alpha\, \beta\, w \\ S & \rightarrow_R^* & \alpha'\, A'\, w' & \rightarrow & \alpha\, \beta\, x \end{array}\right\} \quad \text{folgt:} \quad \alpha = \alpha' \;\wedge\; A = A' \;\wedge\; w' = x$$

Beispiele:

(1) $S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

... ist nicht $LL(k)$ für jedes k — aber $LR(0)$:

Sei $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Dann ist $\alpha\,\underline{\beta}$ von einer der Formen:

$$\underline{A}\ ,\ \underline{B}\ ,\ a^n\,\underline{a\,A\,b}\ ,\ a^n\,\underline{a\,B\,b\,b}\ ,\ a^n\,\underline{0}\ ,\ a^n\,\underline{1} \qquad (n \geq 0)$$

Beispiele:

(1) $S \to A \mid B \qquad A \to a\,A\,b \mid 0 \qquad B \to a\,B\,b\,b \mid 1$

... ist nicht $LL(k)$ für jedes k — aber $LR(0)$:

Sei $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Dann ist $\alpha\,\underline{\beta}$ von einer der Formen:

$$\underline{A}\,,\ \underline{B}\,,\ a^n\,\underline{a\,A\,b}\,,\ a^n\,\underline{a\,B\,b\,b}\,,\ a^n\,\underline{0}\,,\ a^n\,\underline{1} \qquad (n \geq 0)$$

(2) $S \to a\,A\,c \qquad A \to A\,b\,b \mid b$

... ist ebenfalls $LR(0)$:

Sei $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$. Dann ist $\alpha\,\underline{\beta}$ von einer der Formen:

$$a\,\underline{b}\,,\ a\,\underline{A\,b\,b}\,,\ \underline{a\,A\,c}$$

(3) $S \to a\,A\,c \qquad A \to b\,b\,A \;\mid\; b$ ... ist nicht $LR(0)$, aber $LR(1)$:

Für $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$ mit $\{y\} = \text{First}_k(w)$ ist $\alpha\,\underline{\beta}\,y$ von einer der Formen:

$$a\,b^{2n}\,\underline{b}\,c\ ,\ a\,b^{2n}\,\underline{b\,b\,A}\,c\ ,\ \underline{a\,A\,c}$$

(3) $S \to a\,A\,c \qquad A \to b\,b\,A \mid b$ ... ist nicht $LR(0)$, aber $LR(1)$:

Für $S \to_R^* \alpha\,X\,w \to \alpha\,\beta\,w$ mit $\{y\} = \text{First}_k(w)$ ist $\alpha\,\underline{\beta}\,y$ von einer der Formen:

$$a\,b^{2n}\,\underline{b}\,c\;,\;a\,b^{2n}\,\underline{b\,b\,A}\,c\;,\;\underline{a\,A\,c}$$

(4) $S \to a\,A\,c \qquad A \to b\,A\,b \mid b$ ... ist nicht $LR(k)$ für jedes $k \geq 0$:

Betrachte einfach die Rechtsableitungen:

$$S \to_R^* a\,b^n\,A\,b^n\,c \to a\,b^n\,\underline{b}\,b^n\,c$$

In der Tat gilt:

Satz:

Die reduzierte Grammatik G ist genau dann $LR(0)$ wenn der kanonische $LR(0)$-Automat $LR(G)$ keine ungeeigneten Zustände enthält.

In der Tat gilt:

Satz:

Die reduzierte Grammatik G ist genau dann $LR(0)$ wenn der kanonische $LR(0)$-Automat $LR(G)$ keine ungeeigneten Zustände enthält.

Beweis:

Enthalte G einen ungeeigneten Zustand q.

Fall 1: $[A \to \gamma\bullet],\ [A' \to \gamma'\bullet] \in q$ mit $A \to \gamma \neq A' \to \gamma'$

Dann gibt es ein zuverlässiges Präfix $\alpha\,\gamma = \alpha'\,\gamma'$ mit

$$S \to_R^* \alpha\, A\, w \to \alpha\,\gamma\, w \quad \wedge \quad S \to_R^* \alpha'\, A'\, x \to \alpha'\,\gamma'\, x$$

$\Longrightarrow$ G ist nicht $LR(0)$:-)

Fall 2: $[A \to \gamma\bullet]\,,\ [A' \to \beta \bullet a\, \beta'] \ \in\ q$

Dann gibt es ein zuverlässiges Präfix $\alpha\,\gamma = \alpha'\,\beta$ mit

$$S \to_R^* \alpha\, A\, w \to \alpha\,\gamma\, w \quad \wedge \quad S \to_R^* \alpha'\, A'\, x \to \alpha'\,\beta\, a\, \beta'\, x$$

Ist $\beta' \in T^*$, dann ist G nicht $LR(0)$:-)

Andernfalls $\beta' \to_R^* v_1\, X\, v_2 \to v_1\, u\, v_2$. Damit erhalten wir:

$$S \to_R^* \alpha'\,\beta\, a\, v_1\, X\, v_2\, x \to \alpha'\,\beta\, a\, v_1\, u\, v_2\, x$$

$$\Longrightarrow \qquad G \quad \text{ist nicht } LR(0) \quad \text{:-)}$$

Fall 2: $[A \to \gamma\bullet]\,,\ [A' \to \beta \bullet a\,\beta'] \ \in\ q$

Dann gibt es ein zuverlässiges Präfix $\alpha\,\gamma = \alpha'\,\beta$ mit

$$S \to_R^* \alpha\,A\,w \to \alpha\,\gamma\,w \quad \wedge \quad S \to_R^* \alpha'\,A'\,x \to \alpha'\,\beta\,a\,\beta'\,x$$

Ist $\beta' \in T^*$, dann ist G nicht $LR(0)$:-)

Andernfalls $\beta' \to_R^* v_1\,X\,v_2 \to v_1\,u\,v_2$. Damit erhalten wir:

$$S \to_R^* \alpha'\,\beta\,a\,v_1\,X\,v_2\,x \to \alpha'\,\beta\,a\,v_1\,u\,v_2\,x$$

$\Longrightarrow$ G ist nicht $LR(0)$:-)

Enthalte $LR(G)$ keine ungeeigneten Zustände. Betrachte:

$$S \to_R^* \alpha\,A\,w \to \alpha\,\gamma\,w \qquad\qquad S \to_R^* \alpha'\,A'\,w' \to \alpha'\,\gamma'\,x$$

Sei $\delta(q_0, \alpha\,\gamma) = q$. Insbesondere ist $[A \to \gamma\bullet] \in q$.

Annahme: $(\alpha, A, w') \neq (\alpha', A', x)$.

Fall 1: $w' = x$. Dann muss q $[A' \to \gamma'\bullet]$ enthalten :-)

Fall 2: $w' \neq x$. Weitere Fallunterscheidung :-))

Sei $k > 0$.

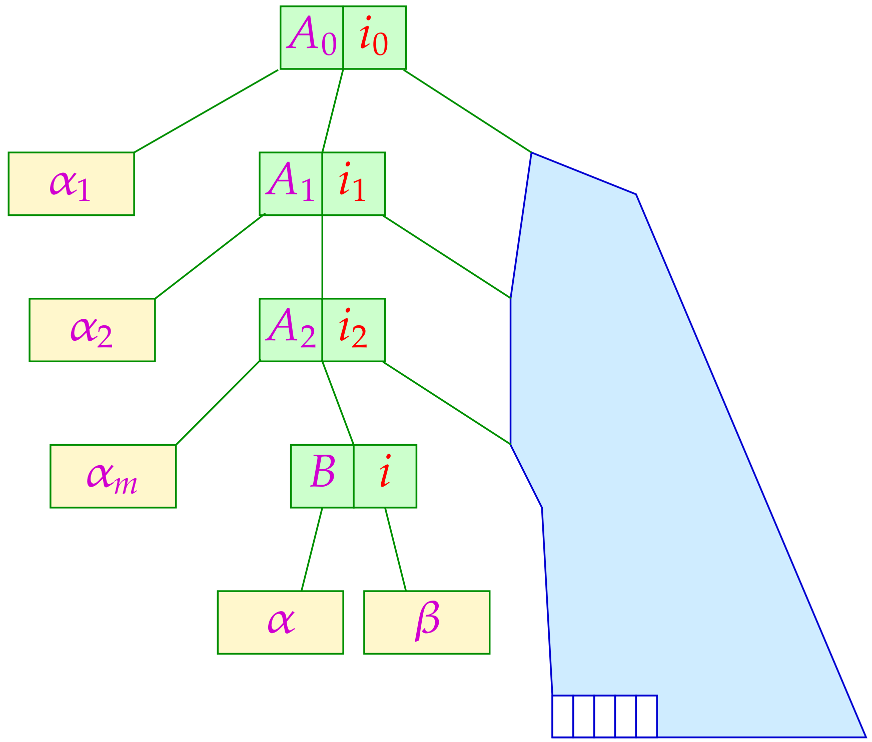
Idee: Wir statten Items mit k-Vorausschau aus :-)

Ein $LR(k)$-Item ist dann ein Paar:

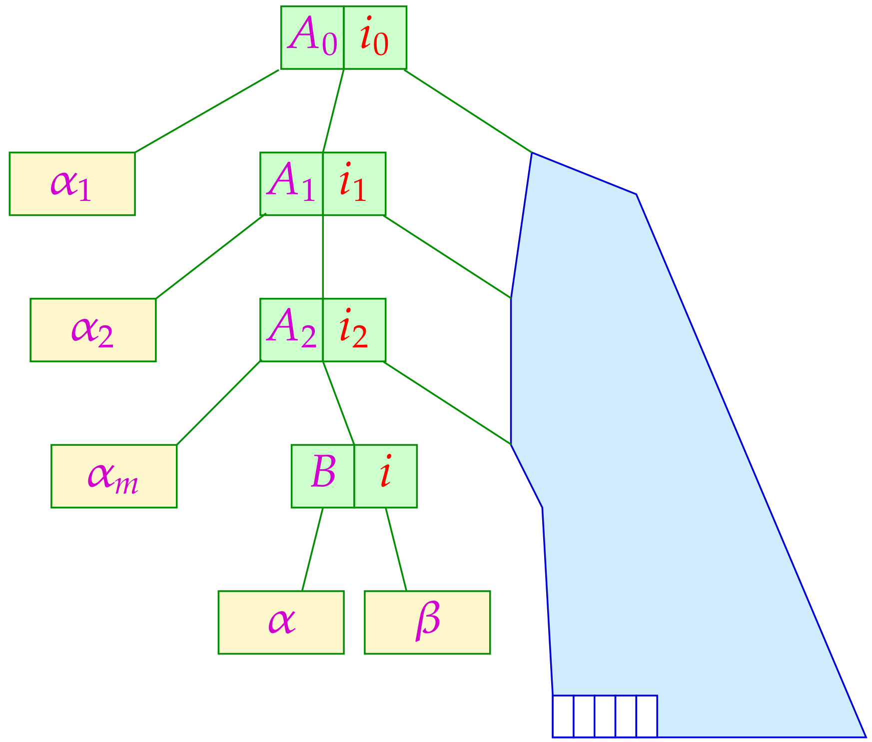
$$[B \to \alpha \bullet \beta,\, x]\,, \quad x \in \mathsf{Follow}_k(B)$$

Dieses Item ist gültig für $\gamma\,\alpha$ falls:

$$S \to^*_R \gamma\, B\, w \qquad \text{mit} \qquad \{x\} = \mathsf{First}_k(w)$$



... wobei $\quad \alpha_1 \ldots \alpha_m = \gamma$



... wobei $\alpha_1 \ldots \alpha_m = \gamma$

Die Menge der gültigen $LR(k)$-Items für zuverlässige Präfixe berechnen wir wieder mithilfe eines endlichen Automaten :-)

Der Automat $c(G,k)$:

Zustände: $LR(k)$-Items :-)

Anfangszustand: $[S' \to \bullet\, S,\ \epsilon]$

Endzustände: $\{[B \to \gamma\bullet,\ x] \mid B \to \gamma \,\in\, P, x \,\in\, \text{Follow}_k(B)\}$

Übergänge:

(1) $([A \to \alpha \bullet X\,\beta,\ x], X, [A \to \alpha\, X \bullet \beta,\ x]), \quad X \,\in\, (N \cup T)$

(2) $([A \to \alpha \bullet B\,\beta,\ x], \epsilon,\ [B \to \bullet\,\gamma,\ x'])$,

$$A \to \alpha\, B\, \beta\ ,\quad B \to \gamma \,\in\, P, x' \,\in\, \text{First}_k(\beta) \odot \{x\};$$

Der Automat $c(G,k)$:

Zustände: $LR(k)$-Items :-)

Anfangszustand: $[S' \to \bullet\, S,\ \epsilon]$

Endzustände: $\{[B \to \gamma\bullet,\ x] \mid B \to \gamma\ \in\ P, x\ \in\ \mathsf{Follow}_k(B)\}$

Übergänge:

(1) $([A \to \alpha \bullet X\,\beta,\ x], X, [A \to \alpha\, X \bullet \beta,\ x]),\quad X\ \in\ (N \cup T)$

(2) $([A \to \alpha \bullet B\,\beta,\ x], \epsilon,\ [B \to\ \bullet \gamma,\ x']),$

$$A \to \alpha\, B\,\beta\ ,\quad B \to \gamma\ \in\ P, x'\ \in\ \mathsf{First}_k(\beta) \odot \{x\};$$

Dieser Automat arbeitet wie $c(G)$ — verwaltet aber zusätzlich ein k-Präfix aus dem Follow_k der linken Seiten.

Den kanonischen $LR(k)$-Automaten $LR(G,k)$ erhält man aus $c(G,k)$, indem man nach jedem Übergang beliebig viele ϵ liest und dann den Automaten deterministisch macht ...

Den kanonischen $LR(k)$-Automaten $LR(G,k)$ erhält man aus $c(G,k)$, indem man nach jedem Übergang beliebig viele ϵ liest und dann den Automaten deterministisch macht ...

Man kann ihn aber auch direkt aus der Grammatik konstruieren werden :-)

Wie bei $LR(0)$ benötigt man eine Hilfsfunktion:

$$\begin{aligned}\delta^*_\epsilon(q) = q \cup \{[B \to \bullet\, \gamma, x] \mid \quad & \exists\, [A \to \alpha \bullet B'\, \beta', x'] \in q\,, \\ & \beta \in (N \cup T)^* : \quad B' \to^* B\, \beta\} \wedge \\ & x \in \mathsf{First}_k(\beta\, \beta') \odot \{x'\}\}\end{aligned}$$

Den kanonischen $LR(k)$-Automaten $LR(G, k)$ erhält man aus $c(G, k)$, indem man nach jedem Übergang beliebig viele ϵ liest und dann den Automaten deterministisch macht ...

Man kann ihn aber auch direkt aus der Grammatik konstruieren werden :-)

Wie bei $LR(0)$ benötigt man eine Hilfsfunktion:

$$\begin{aligned}\delta_\epsilon^*(q) = q \cup \{[B \to \bullet\, \gamma, x] \mid\ & \exists\, [A \to \alpha \bullet B'\, \beta', x'] \in q,\\ & \beta \in (N \cup T)^* : \quad B' \to^* B\, \beta\} \wedge\\ & x \in \mathsf{First}_k(\beta\, \beta') \odot \{x'\}\}\end{aligned}$$

Dann definiert man:

Zustände: Mengen von $LR(k)$-Items;

Anfangszustand: $\delta_\epsilon^*\, \{[S' \to \bullet\, S, \epsilon]\}$

Endzustände: $\{q \mid \exists\, A \to \alpha \in P : \ [A \to \alpha \bullet, x] \in q\}$

Übergänge: $\delta(q, X) = \delta_\epsilon^*\, \{[A \to \alpha\, X \bullet\, \beta, x] \mid [A \to \alpha \bullet X\, \beta, x] \in q\}$

Im Beispiel:

$$\begin{array}{lll}
q_0 & = & \{[S' \to \bullet\, E],\\
& & [E \to \bullet\, E + T],\\
& & [E \to \bullet\, T],\\
& & [T \to \bullet\, T * F],\\
& & [T \to \bullet\, F],\\
& & [F \to \bullet\, (\, E\,)],\\
& & [F \to \bullet\, \mathsf{int}]\}
\end{array}$$

$$\begin{array}{lllll}
q_1 & = & \delta(q_0, E) & = & \{[S' \to E\, \bullet],\\
& & & & [E \to E\, \bullet + T]\}
\end{array}$$

$$\begin{array}{lllll}
q_2 & = & \delta(q_0, T) & = & \{[E \to T\, \bullet],\\
& & & & [T \to T\, \bullet * F]\}
\end{array}$$

$$\begin{array}{lllll}
q_3 & = & \delta(q_0, F) & = & \{[T \to F\, \bullet]\}
\end{array}$$

$$\begin{array}{lllll}
q_4 & = & \delta(q_0, \mathsf{int}) & & \{[F \to \mathsf{int}\, \bullet]\}
\end{array}$$

$$\begin{array}{lllll}
q_5 & = & \delta(q_0, (\,) & = & \{[F \to (\, \bullet\, E\,)],\\
& & & & [E \to \bullet\, E + T],\\
& & & & [E \to \bullet\, T],\\
& & & & [T \to \bullet\, T * F],\\
& & & & [T \to \bullet\, F],\\
& & & & [F \to \bullet\, (\, E\,)],\\
& & & & [F \to \bullet\, \mathsf{int}]\}
\end{array}$$

Im Beispiel:

$$\begin{array}{lllll}
q_0 & = & & & \{[S' \to \bullet\, E,\ \{\epsilon\}],\\
& & & & [E \to \bullet\, E + T,\ \{\epsilon, +\}],\\
& & & & [E \to \bullet\, T,\ \{\epsilon, +\}],\\
& & & & [T \to \bullet\, T * F,\ \{\epsilon, +, *\}],\\
& & & & [T \to \bullet\, F,\ \{\epsilon, +, *\}],\\
& & & & [F \to \bullet\, (\, E\,),\ \{\epsilon, +, *\}],\\
& & & & [F \to \bullet\, \mathsf{int},\ \{\epsilon, +, *\}]\}
\end{array}$$

$$\begin{array}{lllll}
q_1 & = & \delta(q_0, E) & = & \{[S' \to E\, \bullet\quad],\\
& & & & [E \to E\, \bullet + T\quad]\}
\end{array}$$

$$\begin{array}{lllll}
q_2 & = & \delta(q_0, T) & = & \{[E \to T\, \bullet\quad],\\
& & & & [T \to T\, \bullet * F\quad]\}
\end{array}$$

$$\begin{array}{lllll}
q_3 & = & \delta(q_0, F) & = & \{[T \to F\, \bullet\quad]\}
\end{array}$$

$$\begin{array}{lllll}
q_4 & = & \delta(q_0, \mathsf{int}) & & \{[F \to \mathsf{int}\, \bullet\quad]\}
\end{array}$$

$$\begin{array}{lllll}
q_5 & = & \delta(q_0, (\,) & = & \{[F \to (\, \bullet\, E\,)\quad],\\
& & & & [E \to \bullet\, E + T\quad],\\
& & & & [E \to \bullet\, T\quad],\\
& & & & [T \to \bullet\, T * F\quad],\\
& & & & [T \to \bullet\, F\quad],\\
& & & & [F \to \bullet\, (\, E\,)\quad],\\
& & & & [F \to \bullet\, \mathsf{int}\quad]\}
\end{array}$$

Im Beispiel:

$$
\begin{array}{lllll}
q_0 & = & & & \{[S' \to \bullet\, E,\ \{\epsilon\}], \\
 & & & & [E \to \bullet\, E + T,\ \{\epsilon, +\}], \\
 & & & & [E \to \bullet\, T,\ \{\epsilon, +\}], \\
 & & & & [T \to \bullet\, T * F,\ \{\epsilon, +, *\}], \\
 & & & & [T \to \bullet\, F,\ \{\epsilon, +, *\}], \\
 & & & & [F \to \bullet\, (\, E\,),\ \{\epsilon, +, *\}], \\
 & & & & [F \to \bullet\, \mathsf{int},\ \{\epsilon, +, *\}]\}
\end{array}
$$

$$
\begin{array}{lllll}
q_1 & = & \delta(q_0, E) & = & \{[S' \to E \bullet,\ \{\epsilon\}], \\
 & & & & [E \to E \bullet + T,\ \{\epsilon, +\}]\}
\end{array}
$$

$$
\begin{array}{lllll}
q_2 & = & \delta(q_0, T) & = & \{[E \to T \bullet,\ \{\epsilon, +\}], \\
 & & & & [T \to T \bullet * F,\ \{\epsilon, +, *\}]\}
\end{array}
$$

$$
\begin{array}{lllll}
q_3 & = & \delta(q_0, F) & = & \{[T \to F \bullet,\ \{\epsilon, +, *\}]\}
\end{array}
$$

$$
\begin{array}{lllll}
q_4 & = & \delta(q_0, \mathsf{int}) & & \{[F \to \mathsf{int} \bullet,\ \{\epsilon, +, *\}]\}
\end{array}
$$

$$
\begin{array}{lllll}
q_5 & = & \delta(q_0, (\,) & = & \{[F \to (\, \bullet\, E\,) \qquad], \\
 & & & & [E \to \bullet\, E + T \qquad], \\
 & & & & [E \to \bullet\, T \qquad], \\
 & & & & [T \to \bullet\, T * F \qquad], \\
 & & & & [T \to \bullet\, F \qquad], \\
 & & & & [F \to \bullet\, (\, E\,) \qquad], \\
 & & & & [F \to \bullet\, \mathsf{int} \qquad]\}
\end{array}
$$

Im Beispiel:

$$
\begin{array}{lllll}
q_0 & = & & & \{[S' \to \bullet\, E,\ \{\epsilon\}], \\
& & & & [E \to \bullet\, E + T,\ \{\epsilon, +\}], \\
& & & & [E \to \bullet\, T,\ \{\epsilon, +\}], \\
& & & & [T \to \bullet\, T * F,\ \{\epsilon, +, *\}], \\
& & & & [T \to \bullet\, F,\ \{\epsilon, +, *\}], \\
& & & & [F \to \bullet\, (\, E\,),\ \{\epsilon, +, *\}], \\
& & & & [F \to \bullet\, \mathsf{int},\ \{\epsilon, +, *\}]\} \\
\\
q_1 & = & \delta(q_0, E) & = & \{[S' \to E \bullet,\ \{\epsilon\}], \\
& & & & [E \to E \bullet + T,\ \{\epsilon, +\}]\} \\
\\
q_2 & = & \delta(q_0, T) & = & \{[E \to T \bullet,\ \{\epsilon, +\}], \\
& & & & [T \to T \bullet * F,\ \{\epsilon, +, *\}]\}
\end{array}
$$

$$
\begin{array}{lllll}
q_3 & = & \delta(q_0, F) & = & \{[T \to F \bullet,\ \{\epsilon, +, *\}]\} \\
\\
q_4 & = & \delta(q_0, \mathsf{int}) & & \{[F \to \mathsf{int} \bullet,\ \{\epsilon, +, *\}]\} \\
\\
q_5 & = & \delta(q_0, (\,) & = & \{[F \to (\, \bullet\, E\,),\ \{\epsilon, +, *\}], \\
& & & & [E \to \bullet\, E + T,\ \{\,),+\}], \\
& & & & [E \to \bullet\, T,\ \{\,),+\}], \\
& & & & [T \to \bullet\, T * F,\ \{\,),+,*\}], \\
& & & & [T \to \bullet\, F,\ \{\,),+,*\}], \\
& & & & [F \to \bullet\, (\, E\,),\ \{\,),+,*\}], \\
& & & & [F \to \bullet\, \mathsf{int},\ \{\,),+,*\}]\}
\end{array}
$$

$$
\begin{aligned}
q_5' &= \delta(q_5, (\,) = \{[F \rightarrow (\,\bullet E\,)], \\
&\quad [E \rightarrow \bullet E + T], \\
&\quad [E \rightarrow \bullet T], \\
&\quad [T \rightarrow \bullet T * F], \\
&\quad [T \rightarrow \bullet F], \\
&\quad [F \rightarrow \bullet (\,E\,)], \\
&\quad [F \rightarrow \bullet\, \mathsf{int}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_6 &= \delta(q_1, +) = \{[E \rightarrow E + \bullet T], \\
&\quad [T \rightarrow \bullet T * F], \\
&\quad [T \rightarrow \bullet F], \\
&\quad [F \rightarrow \bullet (\,E\,)], \\
&\quad [F \rightarrow \bullet\, \mathsf{int}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_7 &= \delta(q_2, *) = \{[T \rightarrow T * \bullet F], \\
&\quad [F \rightarrow \bullet (\,E\,)], \\
&\quad [F \rightarrow \bullet\, \mathsf{int}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_8 &= \delta(q_5, E) = \{[F \rightarrow (\,E \bullet)]\} \\
&\quad [E \rightarrow E \bullet + T]\}
\end{aligned}
$$

$$
\begin{aligned}
q_9 &= \delta(q_6, T) = \{[E \rightarrow E + T \bullet], \\
&\quad [T \rightarrow T \bullet * F]\}
\end{aligned}
$$

$$
q_{10} = \delta(q_7, F) = \{[T \rightarrow T * F \bullet]\}
$$

$$
q_{11} = \delta(q_8,)\,) = \{[F \rightarrow (\,E\,) \bullet]\}
$$

$$
\begin{aligned}
q_5' &= \delta(q_5, (\,) &= \{&[F \rightarrow (\bullet E\,), \{\,),+,*\}],\\
&&&[E \rightarrow \bullet E + T, \{\,),+\}],\\
&&&[E \rightarrow \bullet T, \{\,),+\}],\\
&&&[T \rightarrow \bullet T * F, \{\,),+,*\}],\\
&&&[T \rightarrow \bullet F, \{\,),+,*\}],\\
&&&[F \rightarrow \bullet (\,E\,), \{\,),+,*\}],\\
&&&[F \rightarrow \bullet \mathsf{int}, \{\,),+,*\}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_6 &= \delta(q_1, +) &= \{&[E \rightarrow E + \bullet T \quad],\\
&&&[T \rightarrow \bullet T * F \quad],\\
&&&[T \rightarrow \bullet F \quad],\\
&&&[F \rightarrow \bullet (\,E\,) \quad],\\
&&&[F \rightarrow \bullet \mathsf{int} \quad]\}
\end{aligned}
$$

$$
\begin{aligned}
q_7 &= \delta(q_2, *) &= \{&[T \rightarrow T * \bullet F \quad],\\
&&&[F \rightarrow \bullet (\,E\,) \quad],\\
&&&[F \rightarrow \bullet \mathsf{int} \quad]\}
\end{aligned}
$$

$$
\begin{aligned}
q_8 &= \delta(q_5, E) &= \{&[F \rightarrow (\,E \bullet) \quad]\}\\
&&&[E \rightarrow E \bullet + T \quad]\}
\end{aligned}
$$

$$
\begin{aligned}
q_9 &= \delta(q_6, T) &= \{&[E \rightarrow E + T \bullet \quad],\\
&&&[T \rightarrow T \bullet * F \quad]\}
\end{aligned}
$$

$$
q_{10} = \delta(q_7, F) = \{[T \rightarrow T * F \bullet \quad]\}
$$

$$
q_{11} = \delta(q_8,)\,) = \{[F \rightarrow (\,E\,) \bullet \quad]\}
$$

$$
\begin{aligned}
q_5' &= \delta(q_5, (\,) &&= \{[F \rightarrow (\, \bullet E\,),\ \{\,)\,, +, *\}],\\
&&&\quad [E \rightarrow \bullet E + T,\ \{\,)\,, +\}],\\
&&&\quad [E \rightarrow \bullet T,\ \{\,)\,, +\}],\\
&&&\quad [T \rightarrow \bullet T * F,\ \{\,)\,, +, *\}],\\
&&&\quad [T \rightarrow \bullet F,\ \{\,)\,, +, *\}],\\
&&&\quad [F \rightarrow \bullet (\, E\,),\ \{\,)\,, +, *\}],\\
&&&\quad [F \rightarrow \bullet \mathsf{int},\ \{\,)\,, +, *\}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_6 &= \delta(q_1, +) &&= \{[E \rightarrow E + \bullet T,\ \{\epsilon, +\}],\\
&&&\quad [T \rightarrow \bullet T * F,\ \{\epsilon, +, *\}],\\
&&&\quad [T \rightarrow \bullet F,\ \{\epsilon, +, *\}],\\
&&&\quad [F \rightarrow \bullet (\, E\,),\ \{\epsilon, +, *\}],\\
&&&\quad [F \rightarrow \bullet \mathsf{int},\ \{\epsilon, +, *\}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_7 &= \delta(q_2, *) &&= \{[T \rightarrow T * \bullet F \quad],\\
&&&\quad [F \rightarrow \bullet (\, E\,) \quad],\\
&&&\quad [F \rightarrow \bullet \mathsf{int} \quad]\}
\end{aligned}
$$

$$
\begin{aligned}
q_8 &= \delta(q_5, E) &&= \{[F \rightarrow (\, E \bullet\,) \quad]\}\\
&&&\quad [E \rightarrow E \bullet + T \quad]\}
\end{aligned}
$$

$$
\begin{aligned}
q_9 &= \delta(q_6, T) &&= \{[E \rightarrow E + T \bullet \quad],\\
&&&\quad [T \rightarrow T \bullet * F \quad]\}
\end{aligned}
$$

$$
q_{10} = \delta(q_7, F) = \{[T \rightarrow T * F \bullet \quad]\}
$$

$$
q_{11} = \delta(q_8, \,)\,) = \{[F \rightarrow (\, E\,) \bullet \quad]\}
$$

$$
\begin{aligned}
q_5' &= \delta(q_5, (\,) &&= \{[F \to (\bullet E), \{), +, *\}],\\
&&& [E \to \bullet E + T, \{), +\}],\\
&&& [E \to \bullet T, \{), +\}],\\
&&& [T \to \bullet T * F, \{), +, *\}],\\
&&& [T \to \bullet F, \{), +, *\}],\\
&&& [F \to \bullet (E), \{), +, *\}],\\
&&& [F \to \bullet \mathsf{int}, \{), +, *\}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_6 &= \delta(q_1, +) &&= \{[E \to E + \bullet T, \{\epsilon, +\}],\\
&&& [T \to \bullet T * F, \{\epsilon, +, *\}],\\
&&& [T \to \bullet F, \{\epsilon, +, *\}],\\
&&& [F \to \bullet (E), \{\epsilon, +, *\}],\\
&&& [F \to \bullet \mathsf{int}, \{\epsilon, +, *\}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_7 &= \delta(q_2, *) &&= \{[T \to T * \bullet F, \{\epsilon, +, *\}],\\
&&& [F \to \bullet (E), \{\epsilon, +, *\}],\\
&&& [F \to \bullet \mathsf{int}, \{\epsilon, +, *\}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_8 &= \delta(q_5, E) &&= \{[F \to (E \bullet), \{\epsilon, +, *\}]\}\\
&&& [E \to E \bullet + T, \{), +\}]\}
\end{aligned}
$$

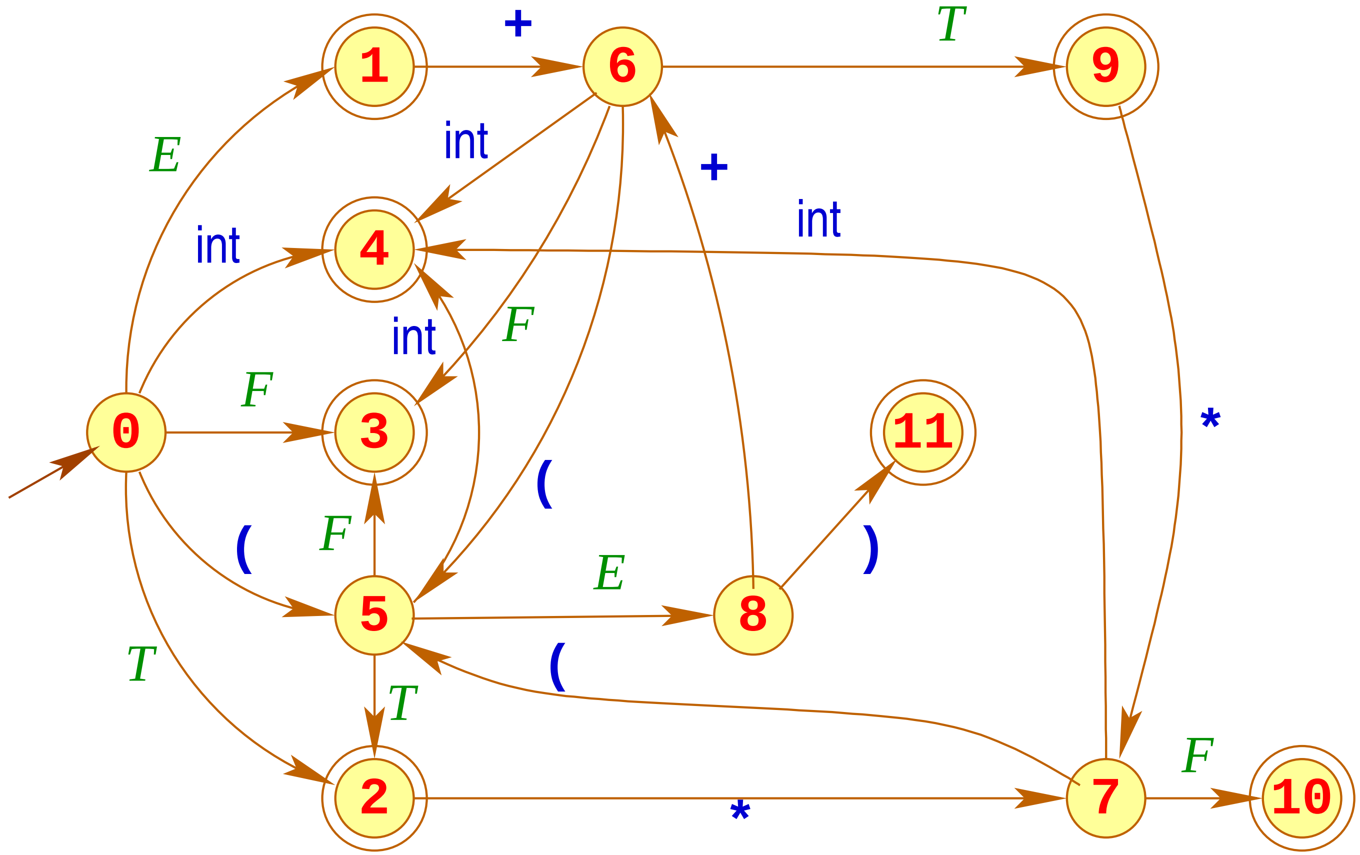
$$
\begin{aligned}
q_9 &= \delta(q_6, T) &&= \{[E \to E + T \bullet, \{\epsilon, +\}],\\
&&& [T \to T \bullet * F, \{\epsilon, +, *\}]\}
\end{aligned}
$$

$$
q_{10} = \delta(q_7, F) = \{[T \to T * F \bullet, \{\epsilon, +, *\}]\}
$$

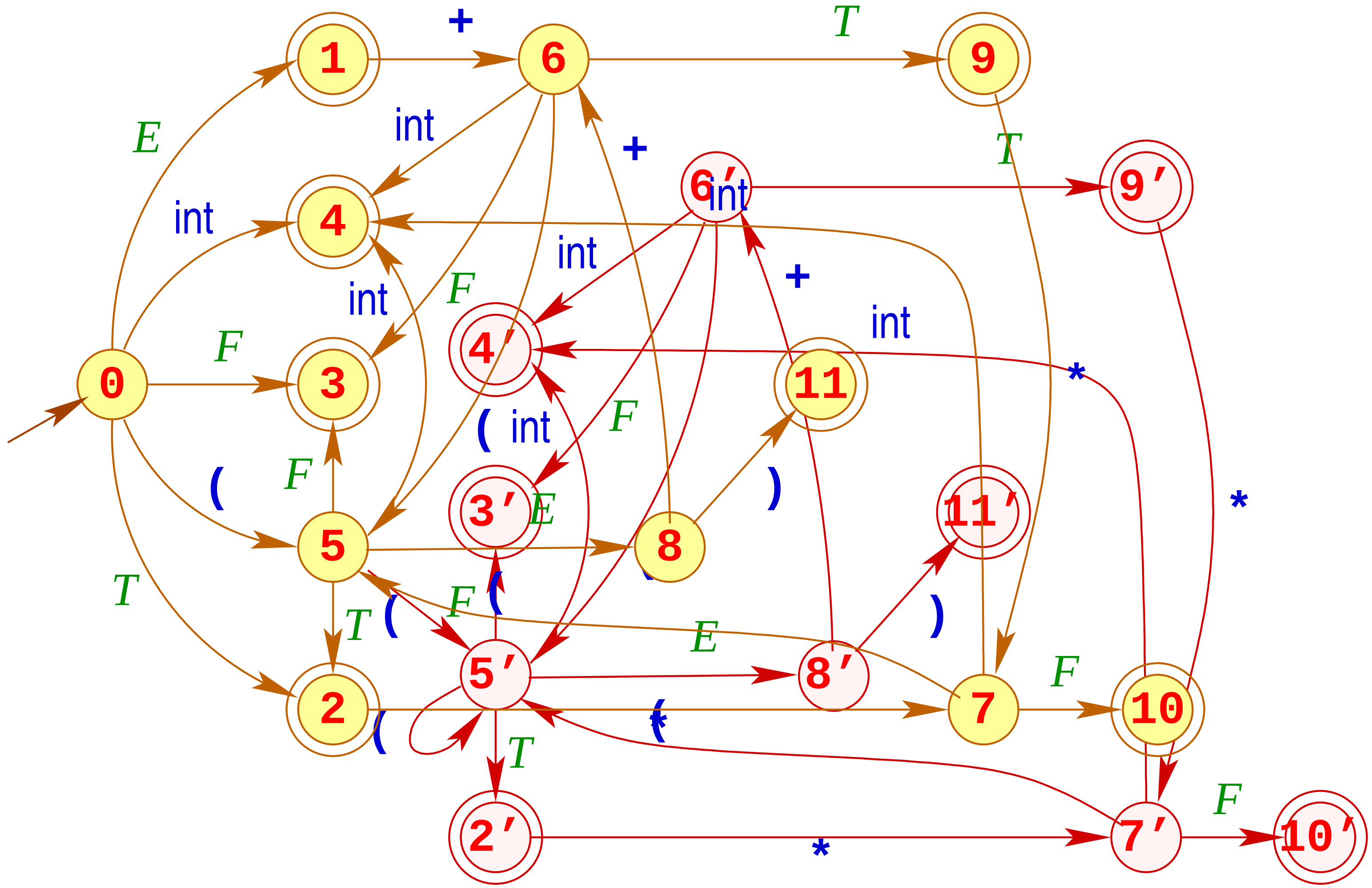
$$
q_{11} = \delta(q_8,)\,) = \{[F \to (E) \bullet, \{\epsilon, +, *\}]\}
$$

$$
\begin{aligned}
q_2' &= \delta(q_5', T) = \{[E \to T\bullet, \{\,)\,, +\}], \\
&\qquad [T \to T \bullet * F, \{\,)\,, +, *\}]\} \\[1em]
q_3' &= \delta(q_5', F) = \{[F \to F\bullet, \{\,)\,, +, *\}]\} \\[1em]
q_4' &= \delta(q_5', \mathsf{int}) = \{[F \to \mathsf{int}\bullet, \{\,)\,, +, *\}]\} \\[1em]
q_6' &= \delta(q_8, +) = \{[E \to E + \bullet\, T, \{\,)\,, +\}], \\
&\qquad [T \to \bullet\, T * F, \{\,)\,, +, *\}], \\
&\qquad [T \to \bullet\, F, \{\,)\,, +, *\}], \\
&\qquad [F \to \bullet\, (\,E\,), \{\,)\,, +, *\}], \\
&\qquad [F \to \bullet\, \mathsf{int}, \{\,)\,, +, *\}]\}
\end{aligned}
$$

$$
\begin{aligned}
q_7' &= \delta(q_9, *) = \{[T \to T * \bullet\, F, \{\,)\,, +, *\}], \\
&\qquad [F \to \bullet\, (\,E\,), \{\,)\,, +, *\}], \\
&\qquad [F \to \bullet\, \mathsf{int}, \{\,)\,, +, *\}]\} \\[1em]
q_8' &= \delta(q_5', E) = \{[F \to (\,E \bullet\,), \{\,)\,, +, *\}]\} \\
&\qquad [E \to E \bullet + T, \{\,)\,, +\}]\} \\[1em]
q_9' &= \delta(q_6', T) = \{[E \to E + T\bullet, \{\,)\,, +\}], \\
&\qquad [T \to T \bullet * F, \{\,)\,, +, *\}] \\[1em]
q_{10}' &= \delta(q_7', F) = \{[T \to T * F\bullet, \{\,)\,, +, *\}]\} \\[1em]
q_{11}' &= \delta(q_8', \,)\,) = \{[F \to (\,E\,)\bullet, \{\,)\,, +, *\}]\}
\end{aligned}
$$



0
1
2
3
4
5
6
7
8
9
10
11
E
T
F
int
(
+
*
)


0
1
2
3
4
5
6
7
8
9
10
11
2'
3'
4'
5'
6'
7'
8'
9'
10'
11'
E
T
F
int
(
)
+
*

Diskussion:

- Im Beispiel hat sich die Anzahl der Zustände fast verdoppelt :-)

  Es kann noch schlimmer kommen :-(

- Die Konflikte in den Zuständen q_1, q_2, q_9 sind nun aufgelöst ...

  Z.B. haben wir für:

$$q_9 = \{[E \to E + T\bullet,\ \{\epsilon, +\}],\\ [T \to T \bullet * F,\ \{\epsilon, +, *\}]\}$$

mit:

$$\{\epsilon, +\} \cap (\mathsf{First}_1(* F) \odot \{\epsilon, +, *\}) = \{\epsilon, +\} \cap \{*\} = \emptyset$$

Allgemein: Wir identifizieren zwei Konflikte:

Reduce-Reduce-Konflikt:

$$[A \to \gamma\,\bullet,\, x]\,,\;\; [A' \to \gamma'\,\bullet,\, x] \;\in\; q \quad \text{mit} \quad A \neq A' \vee \gamma \neq \gamma'$$

Shift-Reduce-Konflikt:

$$[A \to \gamma\,\bullet,\, x]\,,\;\; [A' \to \alpha\,\bullet\, a\,\beta,\, y] \;\in\; q \quad \text{mit} \quad a \in T \quad \text{und}$$

$$x \;\in\; \{a\} \odot \mathsf{First}_k(\beta) \odot \{y\}\,.$$

für einen Zustand $q \in Q$.

Solche Zustände nennen wir jetzt $LR(k)$-ungeeignet :-)

Satz

Eine reduzierte kontextfreie Grammatik G ist genau dann $LR(k)$ wenn der kanonische $LR(k)$-Automat $LR(G, k)$ keine $LR(k)$-ungeeigneten Zustände besitzt.

Satz

Eine reduzierte kontextfreie Grammatik G ist genau dann $LR(k)$ wenn der kanonische $LR(k)$-Automat $LR(G, k)$ keine $LR(k)$-ungeeigneten Zustände besitzt.

Diskussion:

- Unser Beispiel ist offenbar $LR(1)$:-)
- Im Allgemeinen hat der kanonische $LR(k)$-Automat sehr viel mehr Zustände als $LR(G) = LR(G, 0)$:-(
- Man betrachtet darum i.a. Teilklassen von $LR(k)$-Grammatiken, bei denen man nur $LR(G)$ benutzt ...

Satz

Eine reduzierte kontextfreie Grammatik G ist genau dann $LR(k)$ wenn der kanonische $LR(k)$-Automat $LR(G, k)$ keine $LR(k)$-ungeeigneten Zustände besitzt.

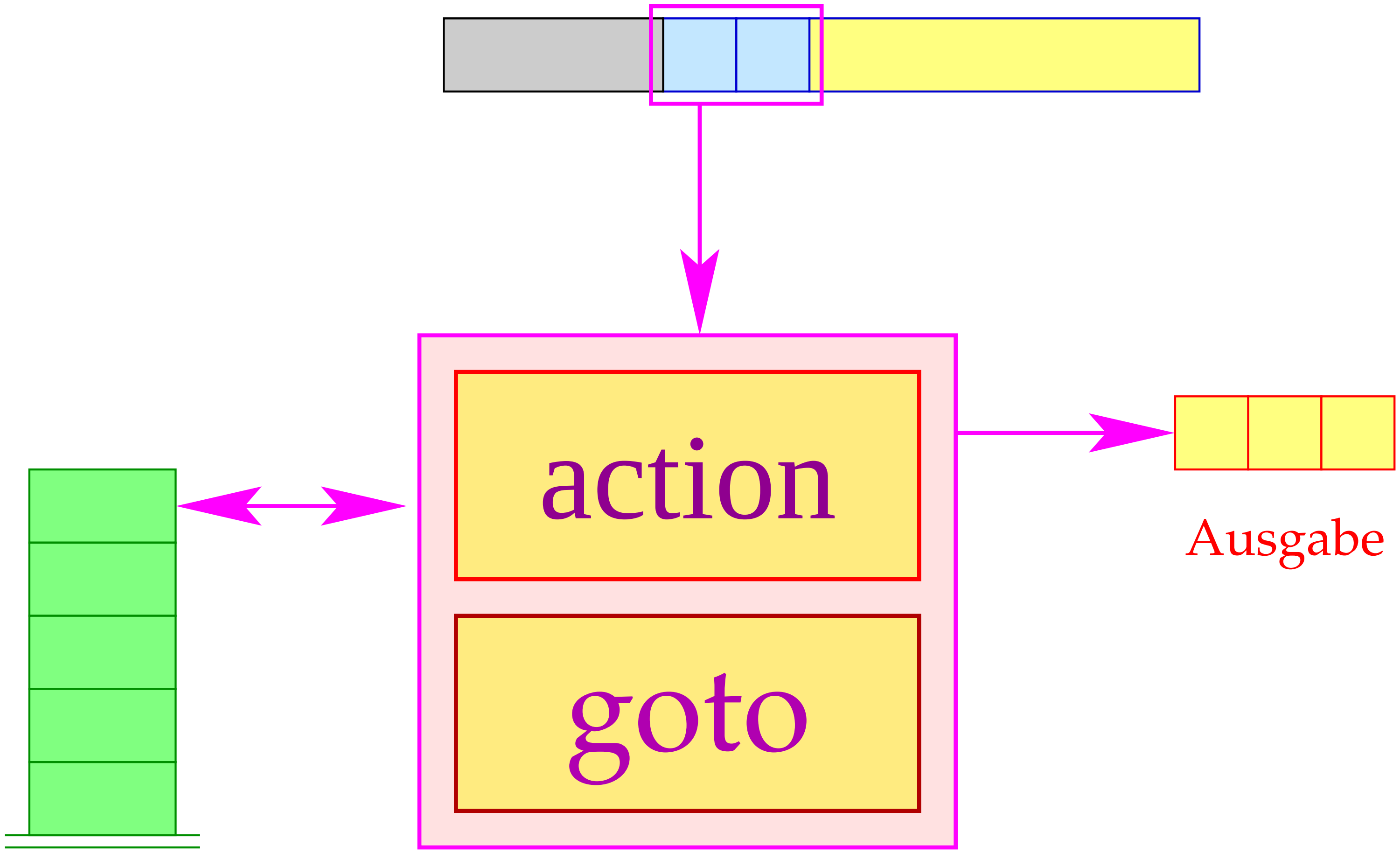
Diskussion:

- Unser Beispiel ist offenbar $LR(1)$:-)
- Im Allgemeinen hat der kanonische $LR(k)$-Automat sehr viel mehr Zustände als $LR(G) = LR(G, 0)$:-(
- Man betrachtet darum i.a. Teilklassen von $LR(k)$-Grammatiken, bei denen man nur $LR(G)$ benutzt ...
- Zur Konflikt-Auflösung ordnet man den Items in den Zuständen Vorausschau-Mengen zu:

  (1) Die Zuordnung ist unabhängig vom Zustand $\Longrightarrow$ Simple $LR(k)$

  (2) Die Zuordnung hängt vom Zustand ab $\Longrightarrow$ $LALR(k)$

Der $LR(k)$-Parser:



Erläuterung:

- Die goto-Tabelle kodiert die Zustandsübergänge:

$$\text{goto}[q, X] = \delta(q, X) \quad \in \quad Q$$

- Die action-Tabelle beschreibt für jeden Zustand q und möglichen Look-ahead w die erforderliche Aktion.

Diese sind:

shift	//	Shift-Operation
reduce $(A \to \gamma)$	//	Reduktion mit Ausgabe
error	//	Fehler

... im Beispiel:

$$
\begin{array}{lcl}
E & \to & E + T^{\,0} \quad | \quad T^{\,1} \\
T & \to & T * F^{\,0} \quad | \quad F^{\,1} \\
F & \to & (\, E\,)^{\,0} \quad | \quad \mathsf{int}^{\,1}
\end{array}
$$

action	ϵ	int	(	)	$+$	$*$
q_1	$S', 0$					s
q_2	$E, 1$					s
q'_2				$E, 1$		s
q_3	$T, 1$				$T, 1$	$T, 1$
q'_3				$T, 1$	$T, 1$	$T, 1$
q_4	$F, 1$				$F, 1$	$F, 1$
q'_4				$F, 1$	$F, 1$	$F, 1$
q_9	$E, 0$				$E, 0$	s
q'_9				$E, 0$	$E, 0$	s
q_{10}	$T, 0$				$T, 0$	$T, 0$
q'_{10}				$T, 0$	$T, 0$	$T, 0$
q_{11}	$F, 0$				$F, 0$	$F, 0$
q'_{11}				$F, 0$	$F, 0$	$F, 0$

2.7 Spezielle Bottom-up-Verfahren mit $LR(G)$

Idee 1: Benutze Follow_k-Mengen zur Konflikt-Lösung ...

Reduce-Reduce-Konflikt:

Falls für $[A \to \gamma \bullet]$, $[A' \to \gamma' \bullet] \in q$ mit $A \neq A' \vee \gamma \neq \gamma'$,

$$\text{Follow}_k(A) \cap \text{Follow}_k(A') \neq \emptyset$$

Shift-Reduce-Konflikt:

Falls für $[A \to \gamma \bullet]$, $[A' \to \alpha \bullet a\, \beta] \in q$ mit $a \in T$,

$$\text{Follow}_k(A) \cap (\{a\} \odot \text{First}_k(\beta) \odot \text{Follow}_k(A')) \neq \emptyset$$

für einen Zustand $q \in Q$.

Dann nennen wir den Zustand q $SLR(k)$-ungeeignet :-)

Die reduzierte Grammatik G nennen wir $SLR(k)$ (simple $LR(k)$:-), falls der kanonische $LR(0)$-Automat $LR(G)$ keine $SLR(k)$-ungeeigneten Zustände enthält :-)

Die reduzierte Grammatik G nennen wir $SLR(k)$ (simple $LR(k)$:-), falls der kanonische $LR(0)$-Automat $LR(G)$ keine $SLR(k)$-ungeeigneten Zustände enthält :-)

... im Beispiel:

Bei unserer Beispiel-Grammatik treten Konflikte möglicherweise in den Zuständen q_1, q_2, q_9 auf:

$$\begin{array}{lll} q_1 & = & \{[S' \to E \bullet], \\ & & \;[E \to E \bullet + T]\} \end{array} \qquad \begin{array}{lll} \mathsf{Follow}_1(S') \cap \{+\} \odot \{\ldots\} & = & \{\epsilon\} \cap \{+\} \\ & = & \emptyset \end{array}$$

Die reduzierte Grammatik G nennen wir $SLR(k)$ (simple $LR(k)$:-), falls der kanonische $LR(0)$-Automat $LR(G)$ keine $SLR(k)$-ungeeigneten Zustände enthält :-)

... im Beispiel:

Bei unserer Beispiel-Grammatik treten Konflikte möglicherweise in den Zuständen q_1, q_2, q_9 auf:

$$
\begin{array}{lll}
q_1 & = & \{[S' \to E\bullet], \\
 & & \;[E \to E \bullet + T]\}
\end{array}
\qquad
\begin{array}{lll}
\mathsf{Follow}_1(S') \cap \{+\} \odot \{\ldots\} & = & \{\epsilon\} \cap \{+\} \\
 & = & \emptyset
\end{array}
$$

$$
\begin{array}{lll}
q_2 & = & \{[E \to T\bullet], \\
 & & \;[T \to T \bullet * F]\}
\end{array}
\qquad
\begin{array}{lll}
\mathsf{Follow}_1(E) \cap \{*\} \odot \{\ldots\} & = & \{\epsilon, +,)\} \cap \{*\} \\
 & = & \emptyset
\end{array}
$$

$$
\begin{array}{lll}
q_9 & = & \{[E \to E + T\bullet], \\
 & & \;[T \to T \bullet * F]\}
\end{array}
\qquad
\begin{array}{lll}
\mathsf{Follow}_1(E) \cap \{*\} \odot \{\ldots\} & = & \{\epsilon, +,)\} \cap \{*\} \\
 & = & \emptyset
\end{array}
$$