Zadanie:

Rozważmy następującą gramatykę jednoznaczną dla n dwuargumentowych operatorów infiksowych:

$E_1 \rightarrow E_1\ \theta_1\ E_2 \mid E_2$
$E_2 \rightarrow E_2\ \theta_2\ E_3 \mid E_3$
...
$E_n \rightarrow E_n\ \theta_n\ E_{n+1} \mid E_{n+1}$
$E_{n+1} \rightarrow (E_1) \mid id$

a) Zbuduj zbiory sytuacji SLR dla tej gramatyki. Jak wiele zbiorów powstaje w zależności od n?
b) Zbuduj tablicę analizatora SLR.
c) Ilu kroków wymaga wyprowadzenie "$id\ \theta_i\ id\ \theta_j\ id$" ?

Rozwiązanie:

Ponumerujmy najpierw kolejne produkcje gramatyki. Dla potrzeb parsera SLR dopisujemy dodatkową produkcję początkową o numerze 0.

(0) $E' \rightarrow E_1$
(1.a) $E_1 \rightarrow E_1\ \theta_1\ E_2$
(1.b) $E_1 \rightarrow E_2$
(2.a) $E_2 \rightarrow E_2\ \theta_2\ E_3$
(2.b) $E_2 \rightarrow E_3$
....
(n.a) $E_n \rightarrow E_n\ \theta_n\ E_{n+1}$
(n.b) $E_n \rightarrow E_{n+1}$
(x) $E_{n+1} \rightarrow (E_1)$
(y) $E_{n+1} \rightarrow id$

Zbudujmy następnie kanoniczny system zbiorów SLR(1) sytuacji dopuszczalnych.

T_0 = { $[E' \rightarrow .\ E_1]$, $[E_1 \rightarrow .\ E_1\ \theta_1\ E_2]$, $[E_1 \rightarrow .\ E_2]$, $[E_2 \rightarrow .\ E_2\ \theta_2\ E_3]$, $[E_2 \rightarrow .\ E_3]$, ..., $[E_n \rightarrow .\ E_n\ \theta_n\ E_{n+1}]$, $[E_n \rightarrow .\ E_{n+1}]$, $[E_{n+1} \rightarrow .\ (\ E_1\ )]$, $[E_{n+1} \rightarrow .\ id]$ }
T_1 = GOTO(T_0, id) = { $[E_{n+1} \rightarrow id\ .]$ }
T_2 = GOTO(T_0, "(") = { $[E_{n+1} \rightarrow (\ .\ E_1\ )]$, $[E_1 \rightarrow .\ E_1\ \theta_1\ E_2]$, $[E_1 \rightarrow .\ E_2]$, $[E_2 \rightarrow .\ E_2\ \theta_2\ E_3]$, $[E_2 \rightarrow .\ E_3]$, ..., $[E_n \rightarrow .\ E_n\ \theta_n\ E_{n+1}]$, $[E_n \rightarrow .\ E_{n+1}]$, $[E_{n+1} \rightarrow .\ (\ E_1\ )]$, $[E_{n+1} \rightarrow .\ id]$ }
$T_{3.1}$ = GOTO(T_0, E_1) = { $[E' \rightarrow E_1\ .]$, $[E_1 \rightarrow E_1\ .\ \theta_1\ E_2]$ }
$T_{3.2}$ = GOTO(T_0, E_2) = { $[E_1 \rightarrow E_2\ .]$, $[E_2 \rightarrow E_2\ .\ \theta_2\ E_3]$ }
...
$T_{3.n}$ = GOTO(T_0, E_n) = { $[E_{n-1} \rightarrow E_n\ .]$, $[E_n \rightarrow E_n\ .\ \theta_n\ E_{n+1}]$ }
T_4 = GOTO(T_0, E_{n+1}) = { $[E_n \rightarrow E_{n+1}\ .]$ }

T_5 = GOTO(T_2, E_1) = { $[E_{n+1} \rightarrow (\ E_1\ .\ )]$, $[E_1 \rightarrow E_1\ .\ \theta_1\ E_2]$ }
GOTO(T_2, E_2) = $T_{3.2}$, GOTO(T_2, E_3) = $T_{3.3}$, ..., GOTO(T_2, E_n) = $T_{3.n}$, GOTO(T_2, E_{n+1}) = T_4
GOTO(T_2, "(") = T_2, GOTO(T_2, id) = T_1

$T_{6.1}$ = GOTO($T_{3.1}$, θ_1) = { $[E_1 \rightarrow E_1\ \theta_1\ .\ E_2]$, $[E_2 \rightarrow .\ E_2\ \theta_2\ E_3]$, $[E_2 \rightarrow .\ E_3]$, ..., $[E_n \rightarrow .\ E_n\ \theta_n\ E_{n+1}]$, $[E_n \rightarrow .\ E_{n+1}]$, $[E_{n+1} \rightarrow .\ (\ E_1\ )]$, $[E_{n+1} \rightarrow .\ id]$ }
$T_{6.2}$ = GOTO($T_{3.2}$, θ_2) = { $[E_2 \rightarrow E_2\ \theta_2\ .\ E_3]$, $[E_3 \rightarrow .\ E_3\ \theta_3\ E_4]$, $[E_3 \rightarrow .\ E_4]$, ..., $[E_n \rightarrow .\ E_n\ \theta_n\ E_{n+1}]$, $[E_n \rightarrow .\ E_{n+1}]$, $[E_{n+1} \rightarrow .\ (\ E_1\ )]$, $[E_{n+1} \rightarrow .\ id]$ }
...

$T_{6.n} = GOTO(T_{3.n}, \theta_n) = \{ [E_n \to E_n\ \theta_n\ .\ E_{n+1}], [E_{n+1} \to .\ (\ E_1\ )], [E_{n+1} \to .\ id] \}$

$T_7 = GOTO(T_5, \text{“)”}) = \{ [E_{n+1} \to (\ E_1\ )\ .] \}$
$GOTO(T_5, \theta_1) = T_{6.1}$

$T_{8.1} = GOTO(T_{6.1}, E_2) = \{ [E_1 \to E_1\ \theta_1\ E_2\ .], [E_2 \to E_2\ .\ \theta_2\ E_3] \}$
$GOTO(T_{6.1}, E_3) = T_{3.3}$, $GOTO(T_{6.1}, E_4) = T_{3.4}$, ..., $GOTO(T_{6.1}, E_n) = T_{3.n}$, $GOTO(T_{6.1}, E_{n+1}) = T_4$,
$GOTO(T_{6.1}, \text{“(“}) = T_2$, $GOTO(T_{6.1}, id) = T_1$

$T_{8.2} = GOTO(T_{6.2}, E_3) = \{ [E_2 \to E_2\ \theta_2\ E_3\ .], [E_3 \to E_3\ .\ \theta_3\ E_4] \}$
$GOTO(T_{6.2}, E_4) = T_{3.4}$, $GOTO(T_{6.2}, E_5) = T_{3.5}$, ..., $GOTO(T_{6.2}, E_n) = T_{3.n}$, $GOTO(T_{6.2}, E_{n+1}) = T_4$,
$GOTO(T_{6.2}, \text{“(“}) = T_2$, $GOTO(T_{6.2}, id) = T_1$

...

$T_{8.n} = GOTO(T_{6.n}, E_{n+1}) = \{ [E_n \to E_n\ \theta_n\ E_{n+1}\ .] \}$
$GOTO(T_{6.n}, \text{“(“}) = T_2$, $GOTO(T_{6.n}, id) = T_1$

$GOTO(T_{8.1}, \theta_2) = T_{6.2}$
$GOTO(T_{8.2}, \theta_3) = T_{6.3}$
...
$GOTO(T_{8.n-1}, \theta_n) = T_{6.n}$

To wszystkie stany jakie możemy utworzyć. Zatem odpowiedź na pytanie a) brzmi: zbiorów jest 3n+6 (T_0, T_1, T_2, T_4, T_5, T_7, oraz trzy serie $T_{3.1} – T_{3.n}$, $T_{6.1} – T_{6.n}$ i $T_{8.1} – T_{8.n}$).

Do zbudowania tablicy parsera będą nam jeszcze potrzebne zbiory FOLLOW nieterminali:

$FIRST_1(E') = \{ id, \text{“(“} \}$
$FIRST_1(E_1) = \{ id, \text{“(“} \}$
$FIRST_1(E_2) = \{ id, \text{“(“} \}$
...
$FIRST_1(E_n) = \{ id, \text{“(“} \}$
$FIRST_1(E_{n+1}) = \{ id, \text{“(“} \}$

$FOLLOW_1(E') = \{ \$ \}$
$FOLLOW_1(E_1) = \{ \$, \theta_1, \text{“)”} \}$
$FOLLOW_1(E_2) = \{ \$, \theta_1, \theta_2, \text{“)”} \}$
...
$FOLLOW_1(E_n) = \{ \$, \theta_1, \theta_2, ..., \theta_n, \text{“)”} \}$
$FOLLOW_1(E_{n+1}) = \{ \$, \theta_1, \theta_2, ..., \theta_n, \text{“)”} \}$

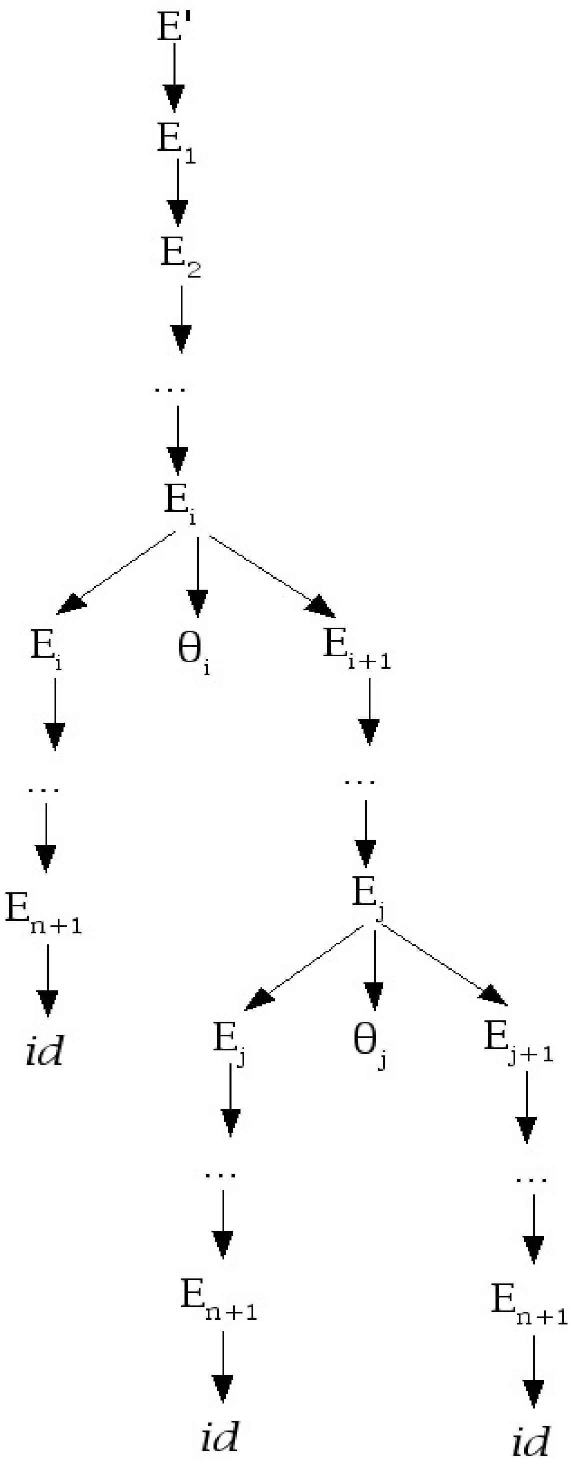
Teraz możemy zbudować tablicę; obydwa jej wymiary zależą od n, więc nie można narysować jej w całości. Dla ustalenia uwagi załóżmy, że n = 5, w przeciwnym wypadku tabela byłaby prawdopodobnie mało czytelna. Puste pola oznaczają błąd (err).

	θ_1	θ_2	θ_3	θ_4	θ_5	(	)	*id*	$
T_0						sh 2		sh 1	
T_1	red y	red y	red y	red y	red y		red y		red y
T_2						sh 2		sh 1	
$T_{3.1}$	sh 6.1								acc
$T_{3.2}$	red 1.b	sh 6.2					red 1.b		red 1.b
$T_{3.3}$	red 2.b	red 2.b	sh 6.3				red 2.b		red 2.b
$T_{3.4}$	red 3.b	red 3.b	red 3.b	sh 6.4			red 3.b		red 3.b
$T_{3.5}$	red 4.b	red 4.b	red 4.b	red 4.b	sh 6.5		red 4.b		red 4.b
T_4	red 5.b	red 5.b	red 5.b	red 5.b	red 5.b		red 5.b		red 5.b
T_5	sh 6.1						sh 7		
$T_{6.1}$						sh 2		sh 1	
$T_{6.2}$						sh 2		sh 1	
$T_{6.3}$						sh 2		sh 1	
$T_{6.4}$						sh 2		sh 1	
$T_{6.5}$						sh 2		sh 1	
T_7	red x	red x	red x	red x	red x		red x		red x
$T_{8.1}$	red 1.a	sh 6.2					red 1.a		red 1.a
$T_{8.2}$	red 2.a	red 2.a	sh 6.3				red 2.a		red 2.a
$T_{8.3}$	red 3.a	red 3.a	red 3.a	sh 6.4			red 3.a		red 3.a
$T_{8.4}$	red 4.a	red 4.a	red 4.a	red 4.a	sh 6.5		red 4.a		red 4.a
$T_{8.5}$	red 5.a	red 5.a	red 5.a	red 5.a	red 5.a		red 5.a		red 5.a

	E_1	E_2	E_3	E_4	E_5	E_6 (E_{n+1})
T_0	$T_{3.1}$	$T_{3.2}$	$T_{3.3}$	$T_{3.4}$	$T_{3.5}$	T_4
T_1						
T_2	T_5	$T_{3.2}$	$T_{3.3}$	$T_{3.4}$	$T_{3.5}$	T_4
$T_{3.1}$						
$T_{3.2}$						
$T_{3.3}$						
$T_{3.4}$						
$T_{3.5}$						
T_4						
T_5						
$T_{6.1}$		$T_{8.1}$	$T_{3.3}$	$T_{3.4}$	$T_{3.5}$	T_4
$T_{6.2}$			$T_{8.2}$	$T_{3.4}$	$T_{3.5}$	T_4
$T_{6.3}$				$T_{8.3}$	$T_{3.5}$	T_4
$T_{6.4}$					$T_{8.4}$	T_4
$T_{6.5}$						$T_{8.5}$
T_7						
$T_{8.1}$						
$T_{8.2}$						
$T_{8.3}$						
$T_{8.4}$						
$T_{8.5}$						

Ilu kroków wymaga wyprowadzenie: “*id* θ_i *id* θ_j *id*” ?

Oto drzewo rozbioru syntaktycznego danego wyrażenia: (zakładamy że i<j)



Do wyprowadzenia “*id* θ_i *id* θ_j *id*” z E' potrzebne są następujące kroki:

- i kroków od E' do E_i
- 1 krok: $E_i \Rightarrow E_i\ \theta_i\ E_{i+1}$
- n-i+2 kroków: $E_i\ \theta_i\ E_{i+1} \Rightarrow id\ \theta_i\ E_{i+1}$
- j-i-1 kroków: $id\ \theta_i\ E_{i+1} \Rightarrow id\ \theta_i\ E_j$
- 1 krok: $id\ \theta_i\ E_j \Rightarrow id\ \theta_i\ E_j\ \theta_j\ E_{j+1}$
- n-j+2 kroków: $id\ \theta_i\ E_j\ \theta_j\ E_{j+1} \Rightarrow id\ \theta_i\ id\ \theta_j\ E_{j+1}$
- n-j+1 kroków: $id\ \theta_i\ id\ \theta_j\ E_{j+1} \Rightarrow id\ \theta_i\ id\ \theta_j\ id$

Razem: 3n-i-j+6 (licząc razem z shiftami potrzebnymi na wczytanie wyrażenia – 3n-i-j+11).