

Stefan Sokołowski

JĘZYKI FORMALNE I METODY KOMPILACJI

wykład 8

Inst. Informatyki Stosowanej, ANS
Elbląg, 2024/2025

Wykład 8, str. 1

Rozbiór wstępujący

Rozbiór zstępujący — budujemy drzewo rozbioru od korzenia do liści:



drzewo

„Z prądem” działania
produkcji gramatyki.

Rozbiór wstępujący — budujemy drzewo rozbioru od liści do korzenia



drzewo

„Pod prąd” działania
produkcji gramatyki.

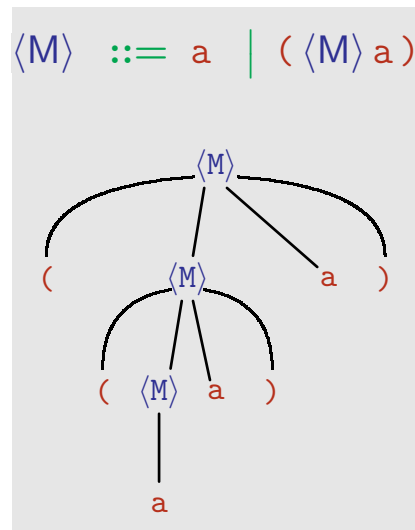
Rozbiór wstępny — gramatyki z pierwszeństwem

— systematyczny sposób przypasowywania prawych stron produkcji do terminali na wejściu.

W ciągu symboli wejściowych znaleźć *podstawę redukcji* — podciąg wejścia zgodny **znak ze znakiem** z *prawą stroną* produkcji.

Dokonać *redukcji* — zastąpić *podstawę* nie-terminalem z *lewej strony* tej produkcji.

Te kroki powtarzać, aż powstanie całe drzewo.



W jaki sposób znajdujemy podstawę redukcji?

Rozbiór wstępny — gramatyki z pierwszeństwem

W jaki sposób znajdujemy podstawę redukcji?

Relacje wpisane w tablicę związane są z kolejnością redukowania poszczególnych symboli:

- jeśli $x \doteq y$ to symbole x i y będą się redukować razem (jeśli stoją koło siebie i jeden należy do podstawy, to drugi też);
- jeśli $x < y$ to symbol x powinien poczekać, aż y zostanie zredukowany (jeśli stoją koło siebie i y należy do podstawy, to x nie należy do podstawy);
- jeśli $x > y$ to symbol y powinien poczekać, aż x zostanie zredukowany (jeśli stoją koło siebie i x należy do podstawy, to y nie należy do podstawy);
- jeśli nie ma między nimi żadnej relacji, to nie powinny stać obok siebie (jeśli stoją koło siebie, to jest to błąd).

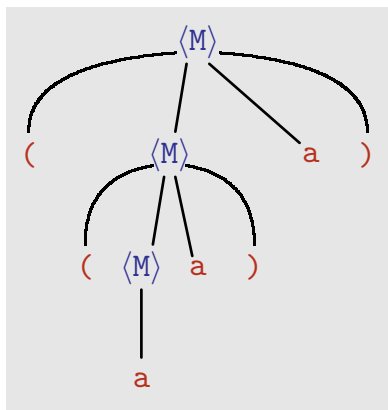
$\langle M \rangle ::= a \mid (\langle M \rangle a)$

Tablica pierwszeństwa redukcji:

	(	$\langle M \rangle$	a	)
(	<	$\doteq$	<	
$\langle M \rangle$			$\doteq$	
a			>	$\doteq$
)			>	

Rozbiór wstępny — gramatyki z pierwszeństwem

W jaki sposób znajdujemy podstawę redukcji?



$\langle M \rangle ::= a \mid (\langle M \rangle a)$

Tablica pierwszeństwa redukcji:

	(	$\langle M \rangle$	a	)
(	<	$\dot{=}$	<	
$\langle M \rangle$			$\dot{=}$	
a			>	$\dot{=}$
)			>	

- Jak skonstruować tablicę pierwszeństw?
- Dla jakich gramatyk da się to zrobić?

Konstrukcja tablic pierwszeństwa

Niech $G = \langle \Sigma, N, P, S \rangle$ będzie gramatyką bezkontekstową.

Definicja:

Niech $A \in N$.

$$fst A \stackrel{\text{def}}{=} \{x \in \Sigma \cup N \mid \text{istnieje w } P \text{ jakaś produkcja } A \rightarrow xv\}$$

$$fst^+ A \stackrel{\text{def}}{=} fst A \cup \bigcup \{fst^+ B \mid B \in N \wedge B \in fst A\}$$

$$lst A \stackrel{\text{def}}{=} \{x \in \Sigma \cup N \mid \text{istnieje w } P \text{ jakaś produkcja } A \rightarrow wx\}$$

$$lst^+ A \stackrel{\text{def}}{=} lst A \cup \bigcup \{lst^+ B \mid B \in N \wedge B \in lst A\}$$

Przykład:

$$\langle M \rangle ::= a \mid (\langle M \rangle a) \quad \text{—} \quad fst \langle M \rangle = \{a, ($$

Przykład:

$$\left. \begin{array}{l} \langle M \rangle ::= a \mid \langle Q \rangle a \\ \langle Q \rangle ::= (\langle M \rangle \end{array} \right\} \quad \text{—} \quad \begin{cases} fst \langle M \rangle = \{a, \langle Q \rangle\} \\ fst^+ \langle M \rangle = \{a, \langle Q \rangle\} \cup fst^+ \langle Q \rangle \\ = \{a, \langle Q \rangle, (\end{cases}$$

Konstrukcja tablic pierwszeństwa

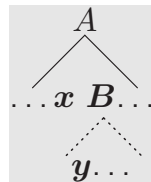
Niech $G = \langle \Sigma, N, P, S \rangle$ będzie gramatyką bezkontekstową.

Definicja:

Relacje $\doteq$, $\triangleleft$, $\triangleright$ w $\Sigma \cup N$ określone są następująco:

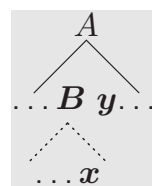
$x \doteq y \iff$ w P istnieje produkcja $A \rightarrow wxyv$

$x \triangleleft y \iff$ w P istnieje produkcja $A \rightarrow wxByv$
taka że $y \in \text{fst}^+ B$



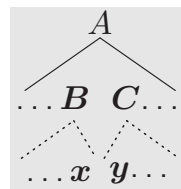
$x \triangleright y \iff$

w P istnieje produkcja $A \rightarrow wByv$
taka że $x \in \text{lst}^+ B$



lub

w P istnieje produkcja $A \rightarrow wBCv$
taka że $x \in \text{lst}^+ B$ i $y \in \text{fst}^+ C$



Konstrukcja tablic pierwszeństwa

Przykład:

$\langle M \rangle ::= a \mid (\langle M \rangle a)$

$x \doteq y \iff$ istnieje produkcja $A \rightarrow wxyv$

$(\doteq \langle M \rangle \quad \langle M \rangle \doteq a \quad a \doteq)$

$x \triangleleft y \iff$ istnieje produkcja $A \rightarrow wxByv$ taka że $y \in \text{fst}^+ B$

$(\triangleleft a \quad (\triangleleft (\quad (\text{bo } \text{fst}^+ \langle M \rangle = \{a, (\})$

$x \triangleright y \iff$

istnieje produkcja $A \rightarrow wByv$ taka że $x \in \text{lst}^+ B$

$a \triangleright a \quad) \triangleright a \quad (\text{bo } \text{lst}^+ \langle M \rangle = \{a,) \}$

lub istnieje produkcja $A \rightarrow wBCv$ taka że $x \in \text{lst}^+ B$ i $y \in \text{fst}^+ C$

Konstrukcja tablic pierwszeństwa

	(	$\langle M \rangle$	a	)
(	$\leq$	$\doteq$	$\leq$	
$\langle M \rangle$			$\doteq$	
a			$\geq$	$\doteq$
)			$\geq$	

Uwaga:

- nie jest prawdą, że zawsze $x \doteq x$
(np. nie zachodzi $a \doteq a$);
- nie jest prawdą, że zawsze z $x \doteq y$ wynika $y \doteq x$
(np. zachodzi $a \doteq)$ ale nie zachodzi $) \doteq a$);
- nie jest prawdą, że zawsze z $x \leq y$ wynika $y \geq x$
(np. zachodzi $(\leq a$ ale nie zachodzi $a \geq ($);
- nie jest prawdą, że zawsze z $x \leq y$ i $y \doteq z$ wynika $x \leq z$
(np. zachodzi $(\leq a$ i $a \doteq)$ ale nie zachodzi $(\leq)$).

Zakres stosowalności gramatyk z pierwszeństwem

Przykład:

$\langle W \rangle ::= \langle S \rangle \mid \langle W \rangle + \langle S \rangle$
 $\langle S \rangle ::= \ell \mid (\langle W \rangle)$

Ponieważ $($ i $\langle W \rangle$ sąsiadują w tej samej produkcji, więc $(\doteq \langle W \rangle$.
Ponieważ $\langle W \rangle \in \text{fst}^+ \langle W \rangle$, więc $(\leq \langle W \rangle$.

Dla tej gramatyki nie da się więc skonstruować tablicy pierwszeństwa.

Przykład:

$\langle A \rangle ::= a \mid \langle B \rangle a$
 $\langle B \rangle ::= a$

Nie wiadomo, do jakiego nieterminalu zredukować a .

Dla tej gramatyki nie da się zastosować analizy z pierwszeństwem.

Definicja:

Gramatyka bezkontekstowa $G = \langle \Sigma, N, P, S \rangle$ jest *gramatyką z pierwszeństwem* jeśli

- dla każdych dwóch symboli $x, y \in \Sigma \cup N$, zachodzi najwyżej jeden ze związków $x \doteq y$, $x \leq y$, $x \geq y$ (może nie zachodzić żaden),
- nie istnieją dwie różne produkcje o tej samej prawej stronie, ani produkcja o pustej prawej stronie: $\langle A \rangle \rightarrow \lambda$.

Poprawianie na gramatykę z pierwszeństwem

Przykład:

$\langle W \rangle ::= \langle S \rangle \mid \langle W \rangle + \langle S \rangle$
 $\langle S \rangle ::= \ell \mid (\langle W \rangle)$

	$\langle W \rangle$	$\langle S \rangle$	$+$	ℓ	$($	$)$
$\langle W \rangle$			$\doteq$			$\doteq$
$\langle S \rangle$			$\triangleright$			$\triangleright$
$+$		$\doteq$		$\triangleleft$	$\triangleleft$	
ℓ			$\triangleright$			$\triangleright$
$($	$\doteq / \triangleleft$	$\triangleleft$		$\triangleleft$	$\triangleleft$	
$)$			$\triangleright$			$\triangleright$

$\langle W \rangle ::= \langle V \rangle$
 $\langle V \rangle ::= \langle S \rangle \mid \langle V \rangle + \langle S \rangle$
 $\langle S \rangle ::= \ell \mid (\langle W \rangle)$

	$\langle W \rangle$	$\langle S \rangle$	$+$	ℓ	$($	$)$	$\langle V \rangle$
$\langle W \rangle$						$\doteq$	
$\langle S \rangle$			$\triangleright$			$\triangleright$	
$+$		$\doteq$		$\triangleleft$	$\triangleleft$		
ℓ			$\triangleright$			$\triangleright$	
$($	$\doteq$	$\triangleleft$		$\triangleleft$	$\triangleleft$		$\triangleleft$
$)$			$\triangleright$			$\triangleright$	
$\langle V \rangle$			$\doteq$			$\triangleright$	

Użycie tablicy pierwszeństwa do rozbioru

$\langle W \rangle ::= \langle V \rangle$
 $\langle V \rangle ::= \langle S \rangle \mid \langle V \rangle + \langle S \rangle$
 $\langle S \rangle ::= \ell \mid (\langle W \rangle)$

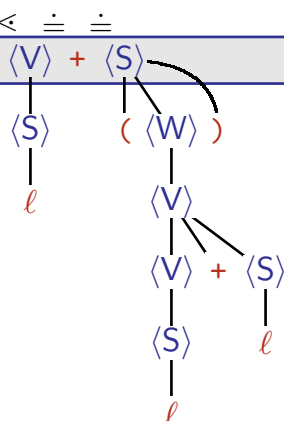
	$\langle W \rangle$	$\langle S \rangle$	$+$	ℓ	$($	$)$	$\langle V \rangle$	$\square$
$\langle W \rangle$						$\doteq$		$\triangleright$
$\langle S \rangle$			$\triangleright$			$\triangleright$		$\triangleright$
$+$		$\doteq$		$\triangleleft$	$\triangleleft$			$\triangleright$
ℓ			$\triangleright$			$\triangleright$		$\triangleright$
$($	$\doteq$	$\triangleleft$		$\triangleleft$	$\triangleleft$		$\triangleleft$	$\triangleright$
$)$			$\triangleright$			$\triangleright$		$\triangleright$
$\langle V \rangle$			$\doteq$			$\triangleright$		$\triangleright$
$\square$	$\triangleleft$	$\triangleleft$	$\triangleleft$	$\triangleleft$	$\triangleleft$	$\triangleleft$	$\triangleleft$	

STOS: $\square \langle V \rangle + \langle S \rangle$



WEJŚCIE:

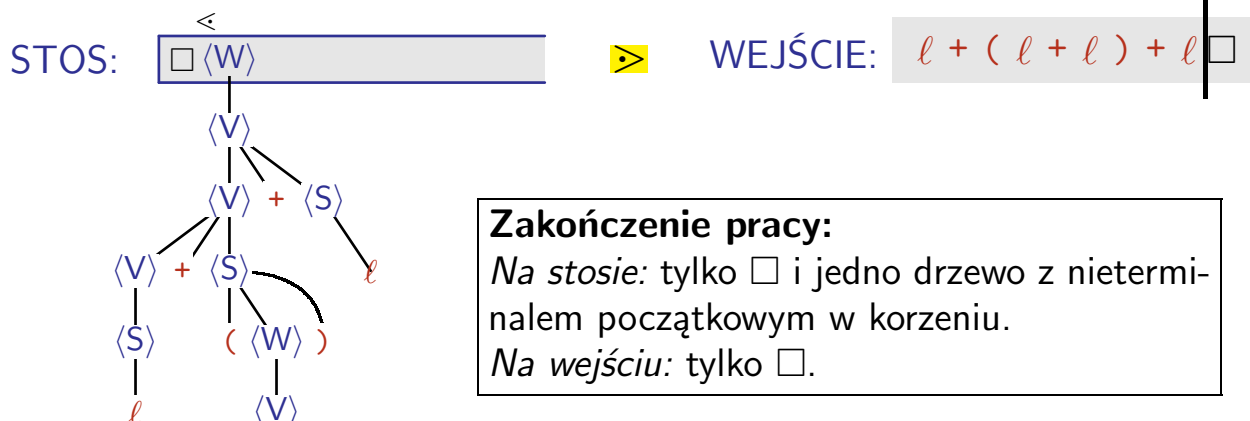
$\ell + (\ell + \ell) + \ell \square$



Użycie tablicy pierwszeństwa do rozbioru

$\langle W \rangle ::= \langle V \rangle$
 $\langle V \rangle ::= \langle S \rangle \mid \langle V \rangle + \langle S \rangle$
 $\langle S \rangle ::= \ell \mid (\langle W \rangle)$

	$\langle W \rangle$	$\langle S \rangle$	$+$	ℓ	$($	$)$	$\langle V \rangle$	$\square$
$\langle W \rangle$							$\doteq$	$\triangleright$
$\langle S \rangle$			$\triangleright$				$\triangleright$	$\triangleright$
$+$		$\doteq$		$\triangleleft$	$\triangleleft$			$\triangleright$
ℓ			$\triangleright$				$\triangleright$	$\triangleright$
$($	$\doteq$	$\triangleleft$		$\triangleleft$	$\triangleleft$		$\triangleleft$	$\triangleright$
$)$			$\triangleright$				$\triangleright$	$\triangleright$
$\langle V \rangle$			$\doteq$				$\triangleright$	$\triangleright$
$\square$		$\triangleleft$	$\triangleleft$	$\triangleleft$	$\triangleleft$	$\triangleleft$	$\triangleleft$	



Użycie tablicy pierwszeństwa do rozbioru

```

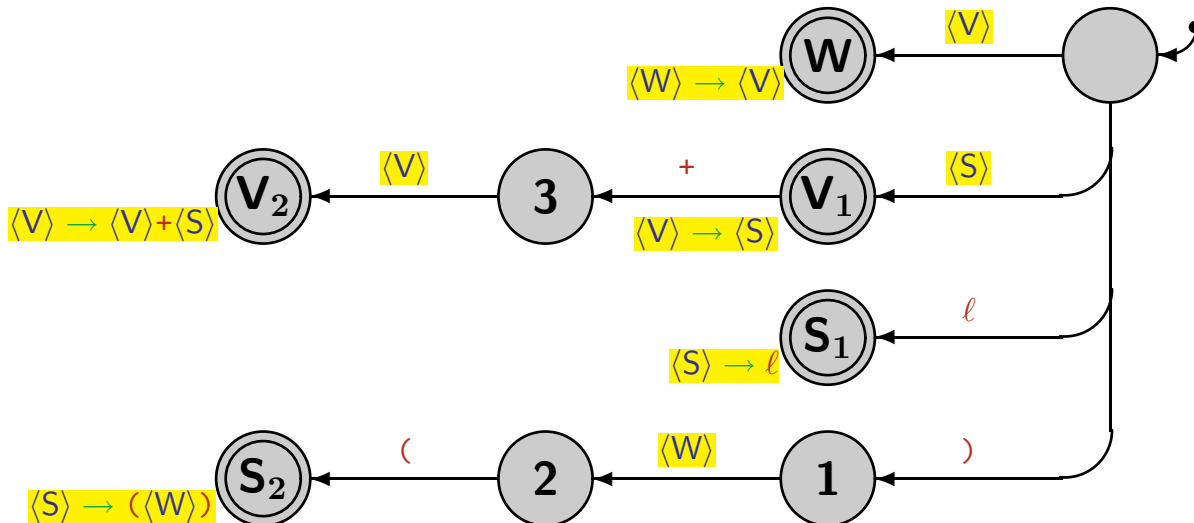
inicjalizacje ();
do {
    switch (pierwszenstwo [szczyt_stosu()][leks]) {
        case nieokr:
            blad ("Te symbole nie moga wystepowac obok siebie: ",
                znak_z_symb (szczyt_stosu()), znak_z_symb (leks));
            break;
        case mniejsze:
        case rowne:
            drz = drzewo_z_terminalu (leks); na_stos (drz);
            nast_leks();
            break;
        case wieksze: redukcja(); break;
    }
}
while (wys_stosu != 2 || szczyt_stosu() != wyr || leks != pusty);
drz = ze_stosu (); drukuj_drzewo (drz);

```

Wykonanie redukcji

$\langle W \rangle ::= \langle V \rangle$
 $\langle V \rangle ::= \langle S \rangle \mid \langle V \rangle + \langle S \rangle$
 $\langle S \rangle ::= \ell \mid (\langle W \rangle)$

- zdejmujemy ze stosu symbole między $\leq$ and $\geq$;
- wyszukujemy właściwą produkcję za pomocą automatu skończonego zbudowanego wg gramatyki;
- wkładamy na stos nieterminal z lewej strony tej produkcji.



Użycie tablicy pierwszeństwa do rozbioru

```

void redukcja (void) {
    Stan st = stan_pocz;  drzewo drz;  int i;
    drz.czyterm = FALSE;
    drz.u.s.il_syn = 0;
    do {
        for (i=drz.u.s.il_syn; i>0; i--)
            drz.u.s.syn[i] = drz.u.s.syn[i-1];
        (drz.u.s.il_syn)++;
        drz.u.s.syn[0] = (drzewo*)malloc(sizeof(drzewo));
        *(drz.u.s.syn[0]) = ze_stosu();
        st = automat [st][symb_drzewa(*(drz.u.s.syn[0]))];
    }
    while (pierwszenstwo [szczyt_stosu()]
           [symb_drzewa(*(drz.u.s.syn[0]))] == rowne);
    drz.u.s.ntm = nterm_ze_stanu (st);
    na_stos (drz);
}

```

Gramatyki z pierwszeństwem — podsumowanie

- Bezpośrednia *wstępująca* konstrukcja drzewa wyvodu z użyciem stosu.
- **Zalety:**
 - prostota;
 - o wiele szybsze działanie i mniejsza zajętość pamięci niż w rozbiorach zstępujących;
 - możliwość dobrej sygnalizacji błędów.
- **Wady:**
 - konieczność poprawienia gramatyki przez dodanie nowych nieterminali i produkcji — czasem bardzo wielu.