

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Костюк Ю. Л.

**ТЕОРИЯ АВТОМАТОВ И
ФОРМАЛЬНЫХ ЯЗЫКОВ**

Лекция 4

ОБРАТНАЯ ПОЛЬСКАЯ СТРОКА, КАК ПРОМЕЖУТОЧНАЯ ФОРМА ПРОГРАММЫ

Обратная польская строка (ОПС) является бесскобочной записью арифметических выражений. В ОПС вначале пишутся операнды, а затем – знак операции. При этом операнды могут быть простыми или сложными, т.е. результатами других операций.

Формула	ОПС
---------	-----

$a + b$	$a\ b\ +$
---------	-----------

$a*(c + d)$	$a\ c\ d\ +\ *$
-------------	-----------------

$(x + y)*(a*x - b*y)$	$x\ y\ +\ a\ x\ *\ b\ y\ *\ -\ *$
-----------------------	-----------------------------------

$(a + b)*(c - d)$	$a\ b\ +\ c\ d\ -\ *$
-------------------	-----------------------

Порядок исполнения действий в ОПС полностью определяется местом операций, количество операндов в операции определяется знаком операции.

При переходе от формулы к ОПС порядок простых операндов ***не изменяется!***

Вычисление ОПС выполняется интерпретатором

В интерпретаторе используется магазин, в который записываются значения операндов и результаты вычислений.

ОПС просматривается слева направо.

Если очередной элемент в ОПС – операнд, то его значение записывается в магазин.

Если очередной элемент в ОПС – операция, то из магазина считываются операнды для этой операции, после чего операция выполняется, а ее результат записывается в магазин.

Пример вычисления ОПС $a\ b + c\ d - *$

- 1) в магазин заносится a ;
- 2) в магазин заносится b ;
- 3) из магазина извлекается b , затем a , выполняется сложение, сумма заносится в магазин;
- 4) в магазин заносится c ;
- 5) в магазин заносится d ;
- 6) из магазина извлекается d , затем c , выполняется вычитание $c - d$, разность заносится в магазин;
- 7) из магазина извлекаются два числа (сумма и разность), выполняется умножение, результат заносится в магазин.

Генерация ОПС для арифметических выражений

Генерация ОПС выполняется одновременно с действиями LL(1)-анализатора, для этого наряду с таблицей анализатора необходима таблица генератора, в которой записываются семантические действия по генерации элементов ОПС. При этом каждой правой части правил порождения соответствует последовательность семантических действий, количество которых ***в точности совпадает*** с длиной правой части.

В процессе генерации ОПС используется ***дополнительный магазин***, занесение в него и извлечение семантических действий выполняется синхронно с магазином LL(1)-анализатора.

Семантические действия выполняются при их извлечении из дополнительного магазина. Они обозначены символами:

□ – пустое действие;

a – запись в ОПС операнда из входной цепочки, распознанного лексическим анализатором (переменной или константы);

+

*

Пример. LL(1)-анализатор грамматики простых арифметических выражений, совмещенный с генератором ОПС

	+	*	()	a	\perp
S			$(S)VU$ □□□□□		aVU $a□□$	
U	$+ TU$ □□+	λ	λ	λ	λ	λ
T			$(S)V$ □□□□		aV $a□$	
V	λ	$* FV$ □□*	λ	λ	λ	λ
F			(S) □□□		a a	

Пример. Анализ цепочки $a + b * c \perp$ и генерация ОПС

№ шага	Вход	Магазин	Правило	Генератор	ОПС
1	$a + b * c \perp$	$S \perp$	$S \rightarrow aVU$	$\square\square$	
2	$a + b * c \perp$	$aVU \perp$		$a\square\square\square$	a
3	$+ b * c \perp$	$VU \perp$	$V \rightarrow \lambda$	$\square\square\square$	a
4	$+ b * c \perp$	$U \perp$	$U \rightarrow + TU$	$\square\square$	a
5	$+ b * c \perp$	$+ TU \perp$		$\square\square+\square$	a
6	$b * c \perp$	$TU \perp$	$T \rightarrow aV$	$\square+\square$	a
7	$b * c \perp$	$aVU \perp$		$a\square+\square$	a b
8	$* c \perp$	$VU \perp$	$V \rightarrow * FV$	$\square+\square$	a b
9	$* c \perp$	$* FVU \perp$		$\square\square*+\square$	a b
10	$c \perp$	$FVU \perp$	$F \rightarrow a$	$\square*+\square$	a b
11	$c \perp$	$aVU \perp$		$a*+\square$	a b c
12	\perp	$VU \perp$	$V \rightarrow \lambda$	$*+\square$	a b c *
13	\perp	$U \perp$	$U \rightarrow \lambda$	$+\square$	a b c * +
14	\perp	\perp		\square	a b c * +

Грамматика с дополнительными операциями

Операция вычитания в грамматике аналогична операции сложения, а операция деления – операции умножения:

$$S \rightarrow S - T$$

$$T \rightarrow T / F$$

Унарные операции (+ и –) требуют дополнительных нетерминалов и порождающих правил:

$$F \rightarrow +G \mid -GZ$$

$$G \rightarrow (S) \mid a$$

$$Z \rightarrow \lambda$$

Операция унарный плюс в ОПС не отображается, так как её не требуется вычислять.

Операция унарный минус должна генерироваться в ОПС *после того*, как будет сгенерирована ОПС для операнда, поэтому в грамматике потребовалось использовать особый нетерминал Z , который *всегда* порождает только λ .

Обозначение в ОПС операции унарный минус должно отличаться от обозначения операции вычитания!

Грамматика с дополнительными операциями

После ее преобразования к нестрогой нормальной форме Грейбах получим:

$$S \rightarrow (S)VU \mid aVU \mid +GVU \mid -GVU$$

$$U \rightarrow + TU \mid - TU \mid \lambda$$

$$T \rightarrow (S)V \mid aV \mid +GV \mid -GV$$

$$V \rightarrow * FV \mid / FV \mid \lambda$$

$$F \rightarrow (S) \mid a \mid +G \mid -GZ$$

$$G \rightarrow (S) \mid a$$

$$Z \rightarrow \lambda$$

Пример. LL(1)-анализатор и генератор ОПС для грамматики простых арифметических выражений с дополнительными операциями: вычитание ($-$) , деление ($/$), унарный минус ($-'$).

	$+$	$-$	$*$	$/$	$($	$)$	a	\perp
S	$+GVU$ $\square\square\square\square$	$-GVU$ $\square\square-'\square$			$(S)VU$ $\square\square\square\square\square$		aVU $a\square\square$	
U	$+ TU$ $\square\square+$	$- TU$ $\square\square-$	λ	λ	λ	λ	λ	λ
T	$+GV$ $\square\square\square$	$-GV$ $\square\square-'$			$(S)V$ $\square\square\square\square$		aV $a\square$	
V	λ	λ	$* FV$ $\square\square*$	$/ FV$ $\square\square/$	λ	λ	λ	λ
F	$+G$ $\square\square$	$-GZ$ $\square\square-'$			(S) $\square\square\square$		a a	
G					(S) $\square\square\square$		a a	
Z	λ	λ	λ	λ	λ	λ	λ	λ

Грамматика присваиваний и арифметических выражений с индексами

Здесь A – начальный нетерминал, знаки $+$, $-$, $*$, $/$, $:=$, запятая, скобки круглые, скобки квадратные, переменная a , константа k – терминалы:

$$A \rightarrow aH := SZ$$

$$S \rightarrow S + T \mid S - T \mid T$$

$$T \rightarrow T * F \mid T / F \mid F$$

$$F \rightarrow (S) \mid + G \mid - GZ \mid aH \mid k$$

$$G \rightarrow (S) \mid aH \mid k$$

$$H \rightarrow [S] \mid [S, S] \mid \lambda$$

$$Z \rightarrow \lambda$$

Грамматика присваиваний и арифметических выражений с индексами

После устранения левой рекурсии, преобразования к нестрогой нормальной форме Грейбах и факторизации:

$$A \rightarrow aH := SZ$$

$$S \rightarrow (S)VU \mid aHVU \mid kVU \mid +GVU \mid -GVU$$

$$U \rightarrow + TU \mid - TU \mid \lambda$$

$$T \rightarrow (S)V \mid aHV \mid kV \mid +GV \mid -GV$$

$$V \rightarrow * FV \mid / FV \mid \lambda$$

$$F \rightarrow (S) \mid aH \mid k \mid +G \mid -GZ$$

$$G \rightarrow (S) \mid aH \mid k$$

$$H \rightarrow [SK \mid \lambda$$

$$K \rightarrow] \mid ,S]$$

$$Z \rightarrow \lambda$$

ДАЛЕЕ: LL(1)-анализатор и генератор ОПС

	+	−	*	/	()	<i>a</i>	<i>k</i>	[]	,	:=	⊥
<i>A</i>							<i>aH</i> := <i>SZ</i> <i>a</i> □□□:=						
<i>S</i>	+ <i>GVU</i> □□□□	− <i>GVU</i> □□−'□			(<i>S</i>) <i>VU</i> □□□□□		<i>aHVU</i> <i>a</i> □□□	<i>kVU</i> <i>k</i> □□					
<i>U</i>	+ <i>TU</i> □□+	− <i>TU</i> □□−	λ	λ	λ	λ	λ	λ	λ	λ	λ	λ	λ
<i>T</i>	+ <i>GV</i> □□□	− <i>GV</i> □□−'			(<i>S</i>) <i>V</i> □□□□		<i>aHV</i> <i>a</i> □□	<i>kV</i> <i>k</i> □					
<i>V</i>	λ	λ	* <i>FV</i> □□*	/ <i>FV</i> □□/	λ	λ	λ	λ	λ	λ	λ	λ	λ
<i>F</i>	+ <i>G</i> □□	− <i>GZ</i> □□−'			(<i>S</i>) □□□		<i>aH</i> <i>a</i> □	<i>k</i> <i>k</i>					
<i>G</i>					(<i>S</i>) □□□		<i>aH</i> <i>a</i> □	<i>k</i> <i>k</i>					
<i>H</i>	λ	λ	λ	λ	λ	λ	λ	λ	[<i>SK</i> □□□	λ	λ	λ	
<i>K</i>] i	, <i>S</i>] □□ i2		
<i>Z</i>	λ	λ	λ	λ	λ	λ	λ	λ	λ	λ	λ	λ	λ

Построенный LL(1)-анализатор – генератор ОПС в процессе разбора входной цепочки символов записывает в ОПС:

Операнды.

a - ссылки на переменные (помещенные в таблицы скалярных переменных и массивов).

k - ссылки на константы (помещенные в таблицу констант).

Операции.

Бинарные: + (сложение), – (вычитание), * (умножение), / (деление), := (присваивание),
i (индексирование в одномерном массиве).

Унарная операция: –' (унарный минус).

Тернарная операция: i2 (индексирование в двумерном массиве).

Каждый элемент ОПС состоит из двух частей:

- 1) тип элемента;
- 2) если элемент – операнд, то ссылка на соответствующую таблицу.

Исполнение (интерпретация) сформированной ОПС:

В интерпретаторе используется магазин, каждый элемент в магазине состоит из двух частей: вида содержимого и собственно содержимого.

Варианты видов содержимого в магазине:

- 1) ссылка на переменную,
- 2) ссылка на константу,
- 3) значение, как результат некоторой операции,
- 4) ссылка на массив (его описание, называемое паспортом массива),
- 5) ссылка на элемент массива.

Интерпретатор поочередно просматривает элементы в ОПС, и если встретился операнд, то его вид и ссылка на него записывается в магазин. Если встретилось обозначение операции, то из магазина считываются операнды для этой операции (их количество определяется этой операцией), после чего операция выполняется, а ее результат для таких операций, как арифметические, записывается в магазин.

Арифметические операции должны выполняться над числовыми значениями операндов, если же встречается операнд – ссылка, то по этой ссылке вначале прочитывается значение, и лишь затем выполняется сама операция.

Для операции присваивания первый операнд должен быть ссылкой, второй операнд используется, как числовое значение, которое записывается в соответствующую таблицу переменных по ссылке. *При этом в магазин ничего не записывается.*

Вычисление операции индексирования i для одномерного массива.

Первый операнд – ссылка на таблицу переменных, где находится описание (паспорт) массива. В паспорте есть ссылка на расположение начального элемента массива с индексом 0, количество элементов, длина одного элемента.

Второй операнд – значение индекса.

Результат операции – ссылка на индексируемый элемент массива согласно формуле:

$$M + d * k,$$

где M – ссылка на элемент массива с индексом [0],
 d – длина элемента массива, k – значение индекса.

Вычисленная ссылка на индексируемый элемент записывается в магазин.

Вычисление операции индексирования i2 для двумерного массива.

Первый операнд – ссылка на паспорт массива.

Второй операнд – значение индекса по первому измерению.

Третий операнд – значение индекса по второму измерению.

Результат операции – ссылка на индексируемый элемент массива согласно формуле:

$$M + d * (k * m + j),$$

где M – ссылка на элемент массива с индексами $[0,0]$,

d – длина элемента массива,

m – количество элементов в массиве по второму измерению,

k, j – значения индексов.

Вычисленная ссылка на индексируемый элемент записывается в магазин.

ПРИМЕР. На вход поступает цепочка:

$$M[k] := a * L[k, j + d] \perp$$

LL(1)-анализатор – генератор сформирует следующую ОПС:

$$M \ k \ i \ a \ L \ k \ j \ d \ + \ i2 \ * \ :=$$

Здесь k, a, j, d – простые переменные; M, L – ссылки на массивы (точнее, на паспорта массивов). Операции обозначены символами: $i, i2, +, *, :=$.

Вычисление ОПС:

Шаг	Магазин	Операция
1	M, k	i
2	$M[k], a, L, k, j, d$	$+$
3	$M[k], a, L, k, j+d$	$i2$
4	$M[k], a, L[k, j+d]$	$*$
5	$M[k], a * L[k, j+d]$	$:=$
6		