

Лабораторная работа 2. Лексический анализ программ.

Разработка лексического анализатора подмножества языка программирования, определённого в лабораторной работе 1. Программа анализа определяет лексические правила и выполняет перевод потока символов программ лабораторной работы 1 в поток лексем (токенов).

На вход программы подается текстовый файл, содержащий строки текста программы.

Например, строка присваивания переменной значения арифметического выражения в виде

ПЕРЕМЕННАЯ = ВЫРАЖЕНИЕ.

Выражение может включать:

Знаки сложения и умножения («+» и «*»);

Круглые скобки («(» и «)»);

Константы (например, 5; 3.8; 1e+18, 8.41E-10);

Имена переменных.

Имя переменной – это последовательность букв и цифр, начинающаяся с буквы.

Разбор выражения $COST = (PRICE + TAX) * 0.98$.

Проанализируем выражение:

$COST$, $PRICE$ и TAX – лексемы-идентификаторы;

0.98 – лексема-константа; $=$, $+$, $*$ – просто лексемы.

Пусть все константы и идентификаторы можно отображать в лексемы типа <идентификатор> (<ИД>). Тогда выходом лексического анализатора будет последовательность лексем $\langle ИД_1 \rangle = (\langle ИД_2 \rangle + \langle ИД_3 \rangle) * \langle ИД_4 \rangle$.

Вторая часть компоненты лексемы (указатель, т.е. номер лексемы в таблице имен) – показана в виде индексов. Символы «=», «+» и «*» трактуются как лексемы, тип которых представляется ими самими. Они не имеют связанных с ними данных и, следовательно, не имеют указателей.

Примеры выделения лексем (токенов).

VARIABLES TABLE	
j	variable of "int" type
d	variable of "double" type
a	variable of "char" type

Рис. 1. Таблица переменных

KEY WORDS TABLE	
const	key word "const"
unsigned	key word "unsigned" is variables type
if	key word "if" is the condition statement
char	key word "char" is variables type
return	return statement
int	key word "int" is variables type
double	key word "double" is variables type

Рис. 2. Таблица ключевых слов

OPERATORS TABLE	
-	is arithmetic operator
+=	is arithmetic operator
=	is arithmetic operator
+	is arithmetic operator
==	is comparison operator/relational operator
%	is arithmetic operator
*	is arithmetic operator

Рис. 3. Таблица операторов

CONSTANTS TABLE	
i	const of "int" type
0	int constant
7	int constant
8	int constant
5	int constant
6	int constant

Рис. 4. Таблица констант

Нахождение и локализация лексических ошибок программы.

Не менее 4-х ошибок. Локализация - № строки, № позиции в строке

Примеры:

```
input.c:4:5: warning: implicit declaration of function 'iff'
    iff(i==j)
    ^
```

```
input.c:9:9: error: expected expression before ';' token
    a+++;
    ^
```