**Лабораторная работа 14 (4 часa)**

**Языки программирования**

**Разработка лексического распознавателя (IV часть)**

* Используйте материал лекций № 18.
* Используйте результаты лабораторных работ № 10-13.
* Создайте проект(VS201Х, C++, консольное приложение) с именем **LPLab14.**
* Ознакомьтесь со спецификацией языка SVV-2015.

|  |  |
| --- | --- |
| **Компоненты языка** | **Описание** |
| Символы | Windows-1251 |
| Символы-сепараторы | пробел - допускается везде кроме идентификаторов и ключевых слов;  **;**(точка с запятой) – разделитель инструкций;  **{}** – программный блок;  () – параметры;  ()-приоритетность операций. |
| Идентификаторы | только малые буквы, от 1 до 5 букв;  идентификатор не может совпадать с ключевыми словами;  максимальное количество идентификаторов 216 |
| Типы данных | **Integer** – целочисленные данные (четыре байта, диапазон от –231 до 231-1), автоматическая инициализация 0, LE;  **string** – строка, любые символы, (максимально 255 символов, первый байт – длина строки), автоматическая инициализация строкой длины 0 |
| Операции с данными | **+** – бинарный, суммирование, (integer, integer);  **+** – бинарный, конкатенация, (string, string);  **-** – бинарный, вычитание, (integer, integer);  **\*** – бинарный, умножение, (integer, integer);  **/** – бинарный, деление, (integer, integer) |
| Программные конструкции | Главная функция (точка входа):  main  {  return <integer-идентификатор> или <integer-литерал>;  }  Функции:  <тип данных> function<идентификатор> (  <тип данных> <идентификатор>, ...)  {  return <идентификатор>;  }  Параметры:  передаются по значению. |
| Литералы | числа в диапазоне от -231 до 231-1, интерпретируются как integer, могут быть только rvalue;  строки, символы, заключенные в '…' (одинарные кавычки), могут быть только rvalue |
| Выражения | арифметические с применением +, -, /, \*, ();  строковые с применением +,() |
| Инструкции | declare <тип данных> идентификатор; – объявление переменных;  declare <тип данных> function идентификатор (<тип данных> идентификатор,... (<тип данных> идентификатор ); – объявление внешних функций;  **=** присвоение значения;  print <идентификатор или литерал>; – вывод в стандартный поток вывода |
| Область видимости | сверху вниз (по принципу С++);  параметры – только внутри функции;  объявление внутри функции – видно только внутри функции |
| Память | Все переменные размещаются в стеке |
| Стандартная библиотека | interger strlen (string) – длина строки;  string substr (string, integer, integer) – подстрока; |
| Максимальное количество функций | 216 |
| Максимальное количество идентификаторов в функции | 216 |

* Текст программы на языке SVV-2015 для первоначального тестирования (контрольный пример):

*integer function fi(integer x, integer y)*

*{*

*declare integer z;*

*z= x\*(x+y);*

*return z;*

*};*

*string function fs (string a, string b)*

*{*

*declare string c;*

*declare string function substr(string a, integer p,*

*integer n);*

*c = substr(a, 1,3)+ b;*

*return c;*

*};*

*main*

*{*

*declare integer x;*

*declare integer y;*

*declare integer z;*

*declare string sa;*

*declare string sb;*

*declare string sc;*

*declare integer strlen(string p);*

*x = 1;*

*y = 5;*

*sa = '1234567890';*

*sb = '1234567890';*

*z = fi(x,y);*

*sc = fs(sa,sb);*

*print 'контрольный пример';*

*print z;*

*print sc;*

*print strlen(sc);*

*return 0;*

*};*

* Разработайте регулярные выражения, постройте графы переходов и соответствующие конечные автоматы для лексем из следующей таблицы (см. лабораторные 11-12).

|  |  |  |
| --- | --- | --- |
| Фраза SVV-2015 | лексема | примечание |
| integer  strinzg | t | ТИ: integer или  string,  значение по умолчанию: для integer – нуль,  для string – пустая строка |
| идентификатор | i | ТИ: строка идентификатора, усеченная до 5 символов.  Префикс: имя конструкции |
| литералы | l | integer или string,  значение. |
| function | f |  |
| declare | d |  |
| return | r |  |
| print | p |  |
| main | m |  |
| ; | ; |  |
| , | , |  |
| { | { |  |
| } | } |  |
| ( | ( |  |
| ) | ) |  |
| +  -  \*  / | v |  |

* Разработайте приложение, которое выполняет следующее:
* вводит текст программы из входного файла (заданного параметром   
  –in:);
* проверяет входные символы на допустимость, в соответствии с таблицей символов (лабораторная 10);
* удаляет лишние пробелы и добавляет сепаратор для вычисления номера строки для каждой лексемы;
* осуществляет лексический анализ всех фраз программы (лабораторная 12-13) и строит таблицу лексем (ТЛ) и таблицу идентификаторов (ТИ);
* формирует протокол работы: представляет содержимое таблицы лексем с разбивкой по строкам с нумерацией (во фрагменте ниже представлен примерный вид), а также диагностические сообщения (с применением таблицы ошибок) с указанием номера ошибочной строки и позиции в исходном тексте;

01 tfi(ti,ti)

02 {

03 dti;

04 i=i\*(i+i);

05 ri;

06 };

* при обработке ошибок использовать таблицу ошибок из лабораторной 10, добавив в нее необходимые коды ошибок. Все ошибки должны обрабатываться с помощью таблицы.

* В контрольном примере, должны вводиться строки, которые анализировались в лабораторной работе 12.
* Необходимо подготовить примеры, в которых демонстрируется обработка всех типов ошибок.
* При построении таблицы лексем используйте следующую спецификацию (h-файл LT.h)



* При построении таблицы идентификаторов используйте следующую спецификацию (h-файл IT.h)



