编译原理课程词法分析实验报告

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **专业班级** | | 计算机科学与技术02班 | **姓名** | | 徐海 | **学号** | 20174214 |
| **实验题目** | 词法分析程序的设计与实现 | | | | | | |
| **实验时间** | 10.1 | | | **实验地点** | A主410 | | |
| **实验成绩** |  | | | **实验性质** | **□验证性 □设计性 ■综合性** | | |
| 教师评价：  □算法/实验过程正确；□源程序/实验内容提交 □程序结构/实验步骤合理；  □实验结果正确； □语法、语义正确； □报告规范；  其他：  评价教师签名： | | | | | | | |
| 一、实验目的  了解编译器中词法分析环节的作用和意义  掌握词法分析基础  实现实验要求中C语言子集的词法分析器程序 | | | | | | | |
| 二、实验项目内容   * 实验内容   设计并实现给定C语言子集的词法分析程序，并实现功能要求:   * 必做项： * 可以识别出用C语言编写的源程序中的每个单词符号，并以记号的形式输出每个单词符号。（基本数据类型至少包含整数、布尔类型；至少包含赋值语句、变量申明语句、算数加/减法表达式、布尔表达式、循环语句、分支语句） * 可以识别并且跳过源程序中的空白。 * 检查源程序中存在的词法错误，并报告错误所在的位置。 * 可选项:   需在给定C语言子集的基础上进行扩充,实现下列要求：   * 识别带else的if语句 * 能识别8进制/16进制数。 * 可以统计源程序中的语句行数、各类单词的个数，以及字符总数，并输出统计结果 * 实验要求： * 自选程序设计语言（java，python，c/c++）作为实现语言，手工编写词法分析程序。 * 提交实验报告及源代码。实验报告需严格遵循学校文档规范，内容包含对应文法、词类编码表、词法分析测试用例。 * 给定C语言子集:   数据类型: **int**, 无符号整数, 取值范围0-9999  int a;  int a,b;  int a = 1;  算术运算符: **+**,**-**  a = b + 1;  a = b + c;  赋值运算符: **=**  a = 1;  关系运算符: **==** ,**>**,**<,!=,>=,<=**  a = (b==c);  a = (b>c);  a = (b<c);  逻辑运算符: **&&**, **||, !**  a = (b&&c);  a = (b||c);  a = (!b);  条件语句: **if**  if(a==b)  {  };  循环语句: **while**  while(a==b)  {  };  输入,输出: **get**,**put**  get(a);  put(a);  语句结束符: **;**  **加分项：**  条件语句 **if else**  if(a==b)  {  };  else  {  }; | | | | | | | |
| 三、实验过程或算法 | | | | | | | |
| 内容要点，供参考：  状态转换图不方便画， 用状态转与矩阵实现：    状态机类得实现    文件类的实现，用于实现记录文件字符的位置和 next prev方法    用于实现标识符分类 | | | | | | | |
| 四、实验测试结果  内容要点，供参考：  测试样例及结果展示：给出能涵盖全部要素的测试样例，包括正确源码样例及包含各种错误类型的样例   * 必选测试样例: * 简易计算器   //输入数据num1,num2,op，根据op确定操作进行运算，最后输出运算结果ans  int num1,num2,op,ans;  get(num1,num2,op);  if(op==0)  {  ans = num1 + num2;  };  if(op==1)  {  ans = num1 - num2;  };  if(op==2)  {  ans = num1 & num2;  };  if(op==3)  {  ans = num1 | num2;  };  put(ans);   * 跑马灯   //循环输入op，改变输出结果out，输入0则结束程序  int num0,num1,out,op;  num1 = 3333;  num2 = 6666;  num3 = 9999;  op = 1;  while(op>0)  {  if(op==1)  {  out = num1;  };  if(op==2)  {  out = num2;  };  if(op==2)  {  out = num3;  };  put(out);  get(op);  };  测试结果：    [  ['<postion object : 0, 4>', '保留字', 'int'],  ['<postion object : 0, 8>', '标识符', 'num1'],  ['<postion object : 0, 9>', '分隔符', ','],  ['<postion object : 0, 13>', '标识符', 'num2'],  ['<postion object : 0, 14>', '分隔符', ','],  ['<postion object : 0, 16>', '标识符', 'op'],  ['<postion object : 0, 17>', '分隔符', ','],  ['<postion object : 0, 20>', '标识符', 'ans'],  ['<postion object : 0, 21>', '分隔符', ';'],  ['<postion object : 1, 3>', '标识符', 'get'],  ['<postion object : 1, 4>', '运算符', '('],  ['<postion object : 1, 8>', '标识符', 'num1'],  ['<postion object : 1, 9>', '分隔符', ','],  ['<postion object : 1, 13>', '标识符', 'num2'],  ['<postion object : 1, 14>', '分隔符', ','],  ['<postion object : 1, 16>', '标识符', 'op'],  ['<postion object : 1, 17>', '运算符', ')'],  ['<postion object : 1, 18>', '分隔符', ';'],  ['<postion object : 2, 3>', '保留字', 'for'],  ['<postion object : 2, 4>', '运算符', '('],  ['<postion object : 2, 8>', '保留字', 'int'],  ['<postion object : 2, 9>', '标识符', 'i'],  ['<postion object : 2, 10>', '运算符', '='],  ['<postion object : 2, 11>', '常数', '0'],  ['<postion object : 2, 12>', '分隔符', ';'],  ['<postion object : 2, 13>', '标识符', 'i'],  ['<postion object : 2, 14>', '运算符', '<'],  ['<postion object : 2, 16>', '常数', '10'],  ['<postion object : 2, 17>', '分隔符', ';'],  ['<postion object : 2, 18>', '标识符', 'i'],  ['<postion object : 2, 19>', '运算符', '+'],  ['<postion object : 2, 20>', '运算符', '+'],  ['<postion object : 2, 21>', '分隔符', ';'],  ['<postion object : 2, 22>', '运算符', ')'],  ['<postion object : 2, 23>', '分隔符', '{'],  ['<postion object : 2, 24>', '分隔符', '}'],  ['<postion object : 3, 2>', '保留字', 'if'],  ['<postion object : 3, 3>', '运算符', '('],  ['<postion object : 3, 5>', '标识符', 'op'],  ['<postion object : 3, 7>', '运算符', '=='],  ['<postion object : 3, 8>', '常数', '0'],  ['<postion object : 3, 9>', '运算符', ')'],  ['<postion object : 4, 1>', '分隔符', '{'],  ['<postion object : 5, 8>', '标识符', 'ans'],  ['<postion object : 5, 9>', '运算符', '='],  ['<postion object : 5, 15>', '标识符', 'num1'],  ['<postion object : 5, 16>', '运算符', '+'],  ['<postion object : 5, 21>', '标识符', 'num2'],  ['<postion object : 5, 22>', '分隔符', ';'],  ['<postion object : 6, 1>', '分隔符', '}'],  ['<postion object : 6, 2>', '分隔符', ';'],  ['<postion object : 7, 2>', '保留字', 'if'],  ['<postion object : 7, 3>', '运算符', '('],  ['<postion object : 7, 5>', '标识符', 'op'],  ['<postion object : 7, 7>', '运算符', '=='],  ['<postion object : 7, 8>', '常数', '1'],  ['<postion object : 7, 9>', '运算符', ')'],  ['<postion object : 8, 1>', '分隔符', '{'],  ['<postion object : 9, 8>', '标识符', 'ans'],  ['<postion object : 9, 9>', '运算符', '='],  ['<postion object : 9, 15>', '标识符', 'num1'],  ['<postion object : 9, 16>', '运算符', '-'],  ['<postion object : 9, 21>', '标识符', 'num2'],  ['<postion object : 9, 22>', '分隔符', ';'],  ['<postion object : 10, 1>', '分隔符', '}'],  ['<postion object : 10, 2>', '分隔符', ';'],  ['<postion object : 11, 2>', '保留字', 'if'],  ['<postion object : 11, 3>', '运算符', '('],  ['<postion object : 11, 5>', '标识符', 'op'],  ['<postion object : 11, 7>', '运算符', '=='],  ['<postion object : 11, 8>', '常数', '2'],  ['<postion object : 11, 9>', '运算符', ')'],  ['<postion object : 12, 1>', '分隔符', '{'],  ['<postion object : 13, 8>', '标识符', 'ans'],  ['<postion object : 13, 9>', '运算符', '='],  ['<postion object : 13, 15>', '标识符', 'num1'],  ['<postion object : 13, 16>', '运算符', '&'],  ['<postion object : 13, 21>', '标识符', 'num2'],  ['<postion object : 13, 22>', '分隔符', ';'],  ['<postion object : 14, 1>', '分隔符', '}'],  ['<postion object : 14, 2>', '分隔符', ';'],  ['<postion object : 15, 2>', '保留字', 'if'],  ['<postion object : 15, 3>', '运算符', '('],  ['<postion object : 15, 5>', '标识符', 'op'],  ['<postion object : 15, 7>', '运算符', '=='],  ['<postion object : 15, 8>', '常数', '3'],  ['<postion object : 15, 9>', '运算符', ')'],  ['<postion object : 16, 1>', '分隔符', '{'],  ['<postion object : 17, 8>', '标识符', 'ans'],  ['<postion object : 17, 9>', '运算符', '='],  ['<postion object : 17, 15>', '标识符', 'num1'],  ['<postion object : 17, 16>', '标识符', '|'],  ['<postion object : 17, 21>', '标识符', 'num2'],  ['<postion object : 17, 22>', '分隔符', ';'],  ['<postion object : 18, 1>', '分隔符', '}'],  ['<postion object : 18, 2>', '分隔符', ';'],  ['<postion object : 19, 3>', '标识符', 'put'],  ['<postion object : 19, 4>', '运算符', '('],  ['<postion object : 19, 7>', '标识符', 'ans'],  ['<postion object : 19, 8>', '运算符', ')'],  ['<postion object : 19, 9>', '分隔符', ';'],  ['<postion object : 19, 11>', 'None', '']  ]  [  ['<postion object : 0, 4>', '保留字', 'int'],  ['<postion object : 0, 8>', '标识符', 'num0'],  ['<postion object : 0, 9>', '分隔符', ','],  ['<postion object : 0, 13>', '标识符', 'num1'],  ['<postion object : 0, 14>', '分隔符', ','],  ['<postion object : 0, 17>', '标识符', 'out'],  ['<postion object : 0, 18>', '分隔符', ','],  ['<postion object : 0, 20>', '标识符', 'op'],  ['<postion object : 0, 21>', '分隔符', ';'],  ['<postion object : 1, 5>', '标识符', 'num1'],  ['<postion object : 1, 6>', '运算符', '='],  ['<postion object : 1, 11>', '常数', '3333'],  ['<postion object : 1, 12>', '分隔符', ';'],  ['<postion object : 2, 5>', '标识符', 'num2'],  ['<postion object : 2, 6>', '运算符', '='],  ['<postion object : 2, 11>', '常数', '6666'],  ['<postion object : 2, 12>', '分隔符', ';'],  ['<postion object : 3, 5>', '标识符', 'num3'],  ['<postion object : 3, 6>', '运算符', '='],  ['<postion object : 3, 11>', '常数', '9999'],  ['<postion object : 3, 12>', '分隔符', ';'],  ['<postion object : 4, 3>', '标识符', 'op'],  ['<postion object : 4, 4>', '运算符', '='],  ['<postion object : 4, 6>', '常数', '1'],  ['<postion object : 4, 7>', '分隔符', ';'],  ['<postion object : 5, 5>', '保留字', 'while'],  ['<postion object : 5, 6>', '运算符', '('],  ['<postion object : 5, 8>', '标识符', 'op'],  ['<postion object : 5, 9>', '运算符', '>'],  ['<postion object : 5, 10>', '常数', '0'],  ['<postion object : 5, 11>', '运算符', ')'],  ['<postion object : 6, 1>', '分隔符', '{'],  ['<postion object : 7, 6>', '保留字', 'if'],  ['<postion object : 7, 7>', '运算符', '('],  ['<postion object : 7, 9>', '标识符', 'op'],  ['<postion object : 7, 11>', '运算符', '=='],  ['<postion object : 7, 12>', '常数', '1'],  ['<postion object : 7, 13>', '运算符', ')'],  ['<postion object : 8, 5>', '分隔符', '{'],  ['<postion object : 9, 13>', '标识符', 'out'],  ['<postion object : 9, 14>', '运算符', '='],  ['<postion object : 9, 19>', '标识符', 'num1'],  ['<postion object : 9, 20>', '分隔符', ';'],  ['<postion object : 10, 5>', '分隔符', '}'],  ['<postion object : 10, 6>', '分隔符', ';'],  ['<postion object : 11, 6>', '保留字', 'if'],  ['<postion object : 11, 7>', '运算符', '('],  ['<postion object : 11, 9>', '标识符', 'op'],  ['<postion object : 11, 11>', '运算符', '=='],  ['<postion object : 11, 12>', '常数', '2'],  ['<postion object : 11, 13>', '运算符', ')'],  ['<postion object : 12, 5>', '分隔符', '{'],  ['<postion object : 13, 13>', '标识符', 'out'],  ['<postion object : 13, 14>', '运算符', '='],  ['<postion object : 13, 19>', '标识符', 'num2'],  ['<postion object : 13, 20>', '分隔符', ';'],  ['<postion object : 14, 5>', '分隔符', '}'],  ['<postion object : 14, 6>', '分隔符', ';'],  ['<postion object : 15, 6>', '保留字', 'if'],  ['<postion object : 15, 7>', '运算符', '('],  ['<postion object : 15, 9>', '标识符', 'op'],  ['<postion object : 15, 11>', '运算符', '=='],  ['<postion object : 15, 12>', '常数', '2'],  ['<postion object : 15, 13>', '运算符', ')'],  ['<postion object : 16, 5>', '分隔符', '{'],  ['<postion object : 17, 13>', '标识符', 'out'],  ['<postion object : 17, 14>', '运算符', '='],  ['<postion object : 17, 19>', '标识符', 'num3'],  ['<postion object : 17, 20>', '分隔符', ';'],  ['<postion object : 18, 5>', '分隔符', '}'],  ['<postion object : 18, 6>', '分隔符', ';'],  ['<postion object : 19, 7>', '标识符', 'put'],  ['<postion object : 19, 8>', '运算符', '('],  ['<postion object : 19, 11>', '标识符', 'out'],  ['<postion object : 19, 12>', '运算符', ')'],  ['<postion object : 19, 13>', '分隔符', ';'],  ['<postion object : 20, 7>', '标识符', 'get'],  ['<postion object : 20, 8>', '运算符', '('],  ['<postion object : 20, 10>', '标识符', 'op'],  ['<postion object : 20, 11>', '运算符', ')'],  ['<postion object : 20, 12>', '分隔符', ';'],  ['<postion object : 21, 1>', '分隔符', '}'],  ['<postion object : 21, 2>', '分隔符', ';'],  ['<postion object : 21, 4>', 'None', '']  ]  [  ['<postion object : 0, 4>', '保留字', 'int'],  ['<postion object : 0, 8>', '标识符', 'num1'],  ['<postion object : 0, 9>', '分隔符', ','],  ['<postion object : 0, 13>', '标识符', 'num2'],  ['<postion object : 0, 14>', '分隔符', ','],  ['<postion object : 0, 16>', '标识符', 'op'],  ['<postion object : 0, 17>', '分隔符', ','],  ['<postion object : 0, 20>', '标识符', 'ans'],  ['<postion object : 0, 21>', '分隔符', ';'],  ['<postion object : 1, 3>', '标识符', 'get'],  ['<postion object : 1, 4>', '运算符', '('],  ['<postion object : 1, 8>', '标识符', 'num1'],  ['<postion object : 1, 9>', '分隔符', ','],  ['<postion object : 1, 13>', '标识符', 'num2'],  ['<postion object : 1, 14>', '分隔符', ','],  ['<postion object : 1, 16>', '标识符', 'op'],  ['<postion object : 1, 17>', '运算符', ')'],  ['<postion object : 1, 18>', '分隔符', ';'],  ['<postion object : 2, 2>', '保留字', 'if'],  ['<postion object : 2, 3>', '运算符', '('],  ['<postion object : 2, 5>', '标识符', 'op'],  ['<postion object : 2, 7>', '运算符', '=='],  ['<postion object : 2, 8>', '常数', '0'],  ['<postion object : 2, 9>', '运算符', ')'],  ['<postion object : 3, 1>', '分隔符', '{'],  ['<postion object : 4, 8>', '标识符', 'ans'],  ['<postion object : 4, 9>', '运算符', '='],  ['<postion object : 4, 15>', '标识符', 'num1'],  ['<postion object : 4, 16>', '运算符', '+'],  ['<postion object : 4, 21>', '标识符', 'num2'],  ['<postion object : 4, 22>', '分隔符', ';'],  ['<postion object : 5, 1>', '分隔符', '}'],  ['<postion object : 5, 2>', '分隔符', ';'],  ['<postion object : 6, 2>', '保留字', 'if'],  ['<postion object : 6, 3>', '运算符', '('],  ['<postion object : 6, 5>', '标识符', 'op'],  ['<postion object : 6, 7>', '运算符', '=='],  ['<postion object : 6, 8>', '常数', '1'],  ['<postion object : 6, 9>', '运算符', ')'],  ['<postion object : 7, 1>', '分隔符', '{'],  ['<postion object : 8, 8>', '标识符', 'ans'],  ['<postion object : 8, 9>', '运算符', '='],  ['<postion object : 8, 15>', '标识符', 'num1'],  ['<postion object : 8, 16>', '运算符', '-'],  ['<postion object : 8, 21>', '标识符', 'num2'],  ['<postion object : 8, 22>', '分隔符', ';'],  ['<postion object : 9, 1>', '分隔符', '}'],  ['<postion object : 9, 2>', '分隔符', ';'],  ['<postion object : 10, 2>', '保留字', 'if'],  ['<postion object : 10, 3>', '运算符', '('],  ['<postion object : 10, 5>', '标识符', 'op'],  ['<postion object : 10, 7>', '运算符', '=='],  ['<postion object : 10, 8>', '常数', '2'],  ['<postion object : 10, 9>', '运算符', ')'],  ['<postion object : 11, 1>', '分隔符', '{'],  ['<postion object : 12, 8>', '标识符', 'ans'],  ['<postion object : 12, 9>', '运算符', '='],  ['<postion object : 12, 15>', '标识符', 'num1'],  ['<postion object : 12, 16>', '运算符', '&'],  ['<postion object : 12, 21>', '标识符', 'num2'],  ['<postion object : 12, 22>', '分隔符', ';'],  ['<postion object : 13, 1>', '分隔符', '}'],  ['<postion object : 13, 2>', '分隔符', ';'],  ['<postion object : 14, 2>', '保留字', 'if'],  ['<postion object : 14, 3>', '运算符', '('],  ['<postion object : 14, 5>', '标识符', 'op'],  ['<postion object : 14, 7>', '运算符', '=='],  ['<postion object : 14, 8>', '常数', '3'],  ['<postion object : 14, 9>', '运算符', ')'],  ['<postion object : 15, 1>', '分隔符', '{'],  ['<postion object : 16, 8>', '标识符', 'ans'],  ['<postion object : 16, 9>', '运算符', '='],  ['<postion object : 16, 15>', '标识符', 'num1'],  ['<postion object : 16, 16>', '标识符', '|'],  ['<postion object : 16, 21>', '标识符', 'num2'],  ['<postion object : 16, 22>', '分隔符', ';'],  ['<postion object : 17, 1>', '分隔符', '}'],  ['<postion object : 17, 2>', '分隔符', ';'],  ['<postion object : 18, 3>', '标识符', 'put'],  ['<postion object : 18, 4>', '运算符', '('],  ['<postion object : 18, 7>', '标识符', 'ans'],  ['<postion object : 18, 8>', '运算符', ')'],  ['<postion object : 18, 9>', '分隔符', ';'],  ['<postion object : 18, 11>', 'None', '']  ]   * 自定义测试样例(包含自定义的词法错误类型):   自定义： | | | | | | | |
| 五、实验总结  内容要点，供参考：  1、实验过程中所遇到的问题及解决办法 | | | | | | | |