資訊三乙 10727219 許志仲

**開發平台 :** window 10

**開發環境 :** Dev C++

**程式語言 :** C++

**選擇的組語 :** SIC

**程式設計**

**功能 :**

從指定的input檔中讀入指令，並依據SIC的規則將指令切成token，並將切好的token分類，選擇放入既有的table1~4或是新增到table5~7，最後依照分類結果印出資訊。

**流程 :**

程式首先會從目錄下讀入table1~4的資料，並存入vector中，接著讀入檔名，從input檔中將指令一行一行讀進來，將該行的每個token切開[1]並進行分類[2]，之後進行寫檔輸出，持續進行直到讀完input檔的所有資料。

[1] 切token的判斷 :

切token的方式參考了大一計算機概論project的作法，從該行指令的第一個字元開始檢查，先跳過所有的whitespace來到第一個”有效字元”，並根據此字元來判斷這個token會是甚麼類別，呼叫對應的function補完字串。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 第一個字元 | 字母(a, B, …) | 數字(1, 2, …) | 單引號 ‘ | 其他符號 |
| 類型 | 字串 | 數字 | ‘ ‘字串 | 符號 |
| 完整token | START | 9999 | ‘EOF’ | # |

特別注意 : 單引號(‘)之前的C, X 不輸出，僅作為table6或7的判斷。

[2] 分類的判斷 :

切好的token會送入最後的checktable()檢查，判斷要放進哪個table()，判斷順序大致上以特殊狀況優先檢查。

判斷順序 :

1. 是否是單個delimiter (table4)

2. 是否為頭尾帶單引號字串

3. 是否為table6,7(前面是”C”,”X”)

4. 是否為table1~3

5. 是否是純數字

6. 剩下都是table5

判斷完該放哪個table之後，最後只要計算index即可輸出答案。

|  |  |
| --- | --- |
| Table1~4 | Index為該token在table裡的索引值 |
| Table5~7 | 先計算token的hashvalue，檢查table[hashvalue]是否會碰撞(該位置已經有插入token)，沒碰撞就直接插入，有碰撞就用線性探測找到下一個可用的位置。  要注意的是如果table已經放滿100個token(full)，就不再放入，也不輸出。 |

使用的data structure :

vector<string> : 用來存放table1~4的內容

string table[100] : 大小為100的字串陣列，用來當作hash的空間

hash : 用來存放table5~7的內容。

**使用方式:**

將.cpp檔、input檔以及四個,.table檔放在同一個目錄之下。

執行程式

輸入input檔名(包括副檔名ex : "SIC\_input.txt")

輸出結果會印在螢幕上和寫檔，檔名為"ResultOutput.txt"