- 1. lex 版本: flex 2.6.4
- 2. 作業平台: ubuntu 22.04
- 3. 執行方式: ./a.out < testfile.pas

4.

(1) 保留字是 case-sensitive,可以建一個這個表:

```
[Aa]
Ь
         [Bb]
c
         [Cc]
d
         [Dd]
e
         [Ee]
f
         [Ff]
g
h
         [Gg]
         [Hh]
i
         [Ii]
j
         [Jj]
k
         [Kk]
ι
         [LL]
M
         [Mm]
n
         [Nn]
0
         [0o]
P
         [Pp]
q
         [Qq]
         [Rr]
s
         [Ss]
t
         [Tt]
u
         [Uu]
         [Vv]
         [Ww]
X
         [XX]
         [Yy]
         [Zz]
```

往後遇到字母都可以用{a}等方式同時表示大小寫。

- (2) 識別字最長到 15 字元,可以簡單用 if-else 處理。
- (3) 字串 30 過長的部分也可以用 if-else 處理,另外像: 'You''ll see' 可以像 這樣寫個小段程式來去掉多餘的單引號:

```
int i;
for(i=1;i<=yyleng-2;i++){
         if(yytext[i]=='\''&&yytext[i+1]=='\''&&i+1!=yyleng-1){
               printf("%c", yytext[i]);
                i++;
                continue;
         }
         printf("%c", yytext[i]);
}</pre>
```

(4) (*ab*)**)我還是沒辦法很完美的解決,如果(*ab*)**)出現之前有一個合法的註解,就會沒辦法正確辨認。目前我的 comment 和 wrong comment 的 regular expression 是這樣的:

wrong_string (\'([^'\r\n;]|\")+)|{letter}+\'
comment \/\/[^\r\n]*|"{"[^}]*"}"|"(*"([^*]|[*]+[^*)])*[*]+")"

5. 遇到的問題:

我覺得最難的部分是讓 scanner 除了辨認正確的 token,還能辨認錯誤的 token,因為不合法 token 的可能性很多。我覺得這要對 pascal 有一定的熟悉,且思考周到才有辦法處理大部分錯誤。

第二個讓我產生問題的東西是 regular expression,延續第一個問題,由於有太多可能性要考慮,且又要將想法轉換成 regular expression 來表示,這個部分造成了我最多問題與停頓。Regular expression 需要不少時間熟悉,才能用的比較顯活。

第三個我覺得難的部分是 lex,lex 有許多功能,因此我覺得跟 regular expression 一樣,熟悉其功能及原理需要一些時間,在寫作業時,有時候也會因為不熟悉 lex 的規則而產生不必要的錯誤。

6. 執行結果:

```
----Scanner start----
          1, starts at char. position: 1, program "res
1, starts at char. position: 9, test "identifier'
1, starts at char. position: 13,; "punctuation"
Line:
                                                                             "reserved word"
Line:
                                                                  "punctuation symbol"
Line:
                                                                  "reserved word'
           2, starts at char. position: 1, var
Line:
                                                                  "identifier'
          3, starts at char. position: 3, i
Line:
          3, starts at char. position: 5, : '3, starts at char. position: 7, integer
                                                                  "punctuation symbol"
Line:
                                                                     "reserved word"
Line:
           3, starts at char. position: 14, ;
                                                                  "punctuation symbol"
Line:
          4, starts at char. position: 14, begin
5, starts at char. position: 3, read
5, starts at char. position: 7, (
5, starts at char. position: 8, i
5, starts at char. position: 9, )
                                                                  "reserved word'
Line:
Line:
                                                                  "reserved word"
                                                                  "punctuation symbol"
Line:
                                                                  "identifier"
Line:
                                                                  "punctuation symbol"
          5, starts at char. position: 10, ;
                                                                  "punctuation symbol"
Line:
          6, starts at char. position: 1, end 6, starts at char. position: 4,;
Line:
                                                                  "reserved word'
                                                                  "punctuation symbol"
Line:
Identifier Table:
test
-----Scanner end-----
-----Scanner start-----
Line: 1, starts at char. position: 1, program "res
Line: 1, starts at char. position: 9, test "identifier"
                                                                            "reserved word"
                                                                  "punctuation symbol"
Line: 1, starts at char. position: 13, ;
Line: 2, starts at char. position: 1, var "reserved word"
Line: 3, starts at char. position: 3, 3i "wrong identifi
Line: 3, starts at char. position: 6,: "punctuation sy
Line: 3, starts at char. position: 8, string "reserved word"
                                                                  "wrong identifier"
                                                                  "punctuation symbol"
          3, starts at char. position: 14, ;
                                                                  "punctuation symbol"
Line:
          4, starts at char. position: 1, begin
5, starts at char. position: 3, 3i
5, starts at char. position: 6, :=
5, starts at char. position: 9, 'ab
                                                                  "reserved word'
line:
                                                                  "wrong identifier"
Line:
                                                                  "operator"
line:
                                                                  "wrong character string"
                                                                  "punctuation symbol'
          5, starts at char. position: 12, ;
Line:
          6, starts at char. position: 1, end 6, starts at char. position: 4,;
                                                                  "reserved word'
Line:
                                                                  "punctuation symbol"
Line:
Identifier Table:
test
-----Scanner end-----
-----Scanner start-----
Line: 1, starts at char. position: 1,
(* comment 1
  comment 2 *)
 "comment"
Line: 3, starts at char. position: 1, program
Line: 3, starts at char. position: 9, test "
Line: 3, starts at char. position: 13 : "
                                                                            "reserved word"
                                                                  "identifier
          3, starts at char. position: 13, ;
                                                                  "punctuation symbol"
Line:
        4, starts at char. position: 1, var
                                                                  "reserved word"
Line:
         5, starts at char. position: 3, i
5, starts at char. position: 5, :
5, starts at char. position: 7, integer
                                                                  "identifier'
Line:
                                                                  "punctuation symbol"
Line:
                                                                             "reserved word"
Line:
           5, starts at char. position: 14, ;
                                                                  "punctuation symbol"
line:
          "reserved word"
Line:
                                                                  "reserved word"
Line:
                                                                  "punctuation symbol"
Line:
                                                                  "identifier"
Line:
Line:
                                                                  "punctuation symbol"
                                                                  "punctuation symbol"
Line:
          7, starts at char. position: 10, ;
          8, starts at char. position: 1, end 8, starts at char. position: 4,;
                                                                  "reserved word
Line:
                                                                  "punctuation symbol"
Line:
Identifier Table:
test
 ·----Scanner end-----
```

```
----Scanner start-
          1, starts at char. position: 1, program "re
1, starts at char. position: 9, test "identifier
Line:
                                                                            "reserved word"
Line:
                                                                 "punctuation symbol"
           1, starts at char. position: 13, ;
Line:
Line:
          2, starts at char. position:
                                                  1, var
                                                                  reserved word
                                                                 "identifier
Line:
          3, starts at char. position:
                                                                  punctuation symbol"
Line:
          3, starts at char. position:
          3, starts at char. position: 7, float "reserved word"
Line:
Line:
          3, starts at char. position: 12, ;
                                                                  "punctuation symbol"
                                                                 "reserved word
"identifier"
          4, starts at char. position: 1, begin 5, starts at char. position: 3, f
Line:
Line:
                                                                 "operator"
-6 "float"
Line:
          5, starts at char. position: 5, :=
Line:
          5, starts at char. position: 8, 12.25e+6
                                                                "punctuation symbol"
"reserved word"
Line:
          5, starts at char. position: 16, ;
          6, starts at char. position: 1, end 6, starts at char. position: 4,;
Line:
                                                                 "punctuation symbol"
Line:
Identifier Table:
test
 -----Scanner end-----
 ----Scanner start-----
Line: 1, starts at char. position: 1,
(* a**b) *)
"comment"
Line: 2, starts at char. position: 1, program "re
Line: 2, starts at char. position: 9, test "identifier
                                                                            "reserved word"
          2, starts at char. position: 13, ;
                                                                 "punctuation symbol"
Line:
          3, starts at char. position: 1, var
Line:
                                                                  reserved word
                                                                 "identifier
Line:
          4, starts at char. position:
          4, starts at char. position: 5, :
4, starts at char. position: 7, integer
                                                                 "punctuation symbol"
line:
                                                                 "reserved word"
"punctuation symbol"
"identifier"
Line:
          4, starts at char. position: 14, ;
Line:
          5, starts at char. position: 3, _s
Line:
          5, starts at char. position: 5, , 5, starts at char. position: 7, _s2
                                                                 "punctuation symbol"
Line:
                                                                 "identifier
Line:
          5, starts at char. position: 10, , 5, starts at char. position: 12, _s3
                                                                "punctuation symbol"
"identifier"
Line:
Line:
          5, starts at char. position: 15, , 5, starts at char. position: 17, _s4
                                                                 "punctuation symbol"
Line:
                                                                 "identifier
Line:
          5, starts at char. position: 20, , 5, starts at char. position: 22, _s5
                                                                 "punctuation symbol"
"identifier"
Line:
Line:
          5, starts at char. position: 26, :
                                                                  "punctuation symbol"
Line:
          5, starts at char. position: 28, string
                                                                 "reserved word
Line:
          5, starts at char. position: 34,;
6, starts at char. position: 1, begin
7, starts at char. position: 3, i
7, starts at char. position: 5,:=
                                                                  "punctuation symbol"
Line:
Line:
                                                                 "reserved word
                                                                 "identifier
Line:
                                                                 "operator
Line:
          7, starts at char. position: 8, -100
7, starts at char. position: 12, ;
                                                                 "integer
Line:
                                                                  punctuation symbol"
Line:
          8, starts at char. position: 3, s
8, starts at char. position: 6, :=
8, starts at char. position: 9, 'db lab'
                                                                  "identifier
Line:
Line:
                                                                 "operator
                                                                            "character string"
Line:
          8, starts at char. position: 17, ;
                                                                  "punctuation symbol
Line:
          9, starts at char. position: 3, s2 "iden
9, starts at char. position: 7, := "oper
9, starts at char. position: 10, 'You'll see'
                                                                  "identifier
Line:
                                                                 operator
Line:
                                                                         "character string"
Line:
                                                                  punctuation symbol
Line:
          9, starts at char. position: 23, ;
         10, starts at char. position: 3, _s3
10, starts at char. position: 7, :=
10, starts at char. position: 10, ''
                                                                 "identifier
Line:
                                                                 "operator
Line:
         10, starts at char. position: 10,
                                                                 "character string'
Line:
         10, starts at char. position: 12, ;
                                                                  'punctuation symbol"
Line:
         11, starts at char. position: 3, _s4
11, starts at char. position: 7, :=
11, starts at char. position: 10, '''
Line:
                                                                 "identifier
                                                                 "operator
Line:
                                                                 "character string"
Line:
         11, starts at char. position: 14, ;
                                                                  punctuation symbol"
Line:
         12, starts at char. position: 3, _s5
12, starts at char. position: 7, :=
                                                                 "identifier
Line:
Line:
                                                                 "operator
         12, starts at char. position: 10,
Line:
                                                                 "character string"
         12, starts at char. position: 13, ;
                                                                 "punctuation symbol"
Line:
         13, starts at char. position: 1, end
13, starts at char. position: 4,;
 Line:
                                                                  reserved word
Line:
                                                                 "punctuation symbol"
Identifier Table:
test
_s2
_s3
_s4
  -----Scanner end-----
```

```
----Scanner start----
        1, starts at char. position: 1, ProGram
1, starts at char. position: 9, test "
1, starts at char. position: 13,; "
Line:
                                                            "reserved word"
                                                   "identifier'
Line:
                                                   "punctuation symbol"
Line:
        2, starts at char. position: 1, var
3, starts at char. position: 3, #db
3, starts at char. position: 7, :
3, starts at char. position: 9, float
Line:
                                                   "reserved word"
Line:
                                                   "wrong identifier"
Line:
                                                   "punctuation symbol"
                                                   "reserved word"
Line:
        3, starts at char. position: 14, ;
                                                   "punctuation symbol"
Line:
                                                   "identifier"
        4, starts at char. position: 3, _f2
Line:
        4, starts at char. position: 7, :
                                                   "punctuation symbol"
Line:
                                                   "reserved word"
        4, starts at char. position: 9, float
Line:
                                                   "punctuation symbol"
        4, starts at char. position: 14, ;
Line:
                                                   "reserved word"
        5, starts at char. position: 1, begin
Line:
                                                   "wrong identifier"
Line:
        6, starts at char. position: 3, #db
                                                   "operator"
Line:
        6, starts at char. position: 7, :=
                                                   "wrong float"
Line:
        6, starts at char. position: 10, .1
                                                   "punctuation symbol"
        6, starts at char. position: 12, ;
Line:
        7, starts at char. position: 3, _f2
                                                   "identifier"
Line:
                                                   "operator
        7, starts at char. position: 7, :=
Line:
        7, starts at char. position: 10, 12.100 "float"
Line:
        7, starts at char. position: 16, ;
                                                   "punctuation symbol"
Line:
        8, starts at char. position: 1, end
                                                   "reserved word"
Line:
                                                  "punctuation symbol"
Line:
        8, starts at char. position: 4, ;
Identifier Table:
test
 f2
 -----Scanner end-----
-----Scanner start-----
Line: 1, starts at char. position: 1,
(* This line is a comment. *)
"comment"
Line: 2, starts at char. position: 1, program
                                                           "reserved word"
        2, starts at char. position: 9, test
                                                   "identifier"
Line:
                                                   "punctuation symbol"
Line:
       2, starts at char. position: 13, ;
       3, starts at char. position: 1, var
                                                   "reserved word"
Line:
        4, starts at char. position: 3, i
                                                   "identifier"
Line:
        4, starts at char. position: 5, :
                                                   "punctuation symbol"
Line:
        4, starts at char. position: 7, integer
                                                      "reserved word"
Line:
                                                   "punctuation symbol"
        4, starts at char. position: 14, ;
        5, starts at char. position: 1, begin
                                                 "reserved word"
Line:
                                                   "identifier'
        6, starts at char. position: 3, i
Line:
        6, starts at char. position: 5, :=
                                                   "operator
Line:
                                                   "integer"
        6, starts at char. position: 8, 1
Line:
        6, starts at char. position: 9, +
                                                   "operator"
Line:
                                                   "integer"
        6, starts at char. position: 10, 2
Line:
        6, starts at char. position: 11, ;
                                                   "punctuation symbol"
Line:
                                                   "reserved word"
        7, starts at char. position: 1, end
Line:
        7, starts at char. position: 4, ;
Line:
                                                   "punctuation symbol"
Identifier Table:
test
-----Scanner end-----
```