Lab 1 Lexer

Sorin Macaluso

February 10, 2024

1 Test Cases

These are the test cases that I used for the program these are taken from Lexer without spaces and Gabriel Arnell Hall of Fame project. I have added changes to them to have unrecognized tokens and key words, as well a numbers in strings, pretty much anything that is not a lower case letter or space (since lowercase letters and space's are the only things allowed inside quotes).

I attempted to make it so that the tokens would match. Could not get all of them without ruining the document.

1.1 Test Case 1

```
1  /*Long Test Case - Everything Except Boolean Declaration */
2  {
3      /* Int Declaration */
4      int a
5      int b
6      a = 0
7      b=0
8      /* While Loop */
9      while (a != 3) {
10         print(a)
11         while (b != 3) {
12             print(b)
13             b = 1 + b
14             if (b == 2) {
15                 /* Print Statement */
16                 print({"there is no spoon" /* This will do nothing */ )
17             }
18         }
19         b = 0
20         a = 1+a
21     }
22 }
23 $
24
25 Output:
26 Begin_Block [ { ] Found at line 2 : 1
27 TYPE [ int ] Found at line 4 : 11
28 ID [ a ] Found at line 4 : 13
29 TYPE [ int ] Found at line 5 : 11
30 ID [ b ] Found at line 5 : 13
```

```
31  ID [ a ] Found at line 6 : 9
32  AssignmentStatement [ = ] Found at line 6 : 11
33  DIGIT [ 0 ] Found at line 6 : 13
34  ID [ b ] Found at line 7 : 9
35  AssignmentStatement [ = ] Found at line 7 : 10
36  DIGIT [ 0 ] Found at line 7 : 11
37  While_Statment [ while ] Found at line 9 : 13
38  Open_Expression [ ( ] Found at line 9 : 15
39  ID [ a ] Found at line 9 : 16
40  Not_Equal [ != ] Found at line 9 : 19
41  DIGIT [ 3 ] Found at line 9 : 21
42  Close_Expression [ ) ] Found at line 9 : 22
43  Begin_Block [ { ] Found at line 9 : 24
44  Print_Statment [ print ] Found at line 10 : 17
45  Open_Expression [ ( ] Found at line 10 : 18
46  ID [ a ] Found at line 10 : 19
47  Close_Expression [ ) ] Found at line 10 : 20
48  While_Statment [ while ] Found at line 11 : 17
49  Open_Expression [ ( ] Found at line 11 : 19
50  ID [ b ] Found at line 11 : 20
51  Not_Equal [ != ] Found at line 11 : 23
52  DIGIT [ 3 ] Found at line 11 : 25
53  Close_Expression [ ) ] Found at line 11 : 26
54  Begin_Block [ { ] Found at line 11 : 28
55  Print_Statment [ print ] Found at line 12 : 21
56  Open_Expression [ ( ] Found at line 12 : 22
57  ID [ b ] Found at line 12 : 23
58  Close_Expression [ ) ] Found at line 12 : 24
59  ID [ b ] Found at line 13 : 17
60  AssignmentStatement [ = ] Found at line 13 : 19
61  DIGIT [ 1 ] Found at line 13 : 21
62  InTop [ + ] Found at line 13 : 23
63  ID [ b ] Found at line 13 : 25
64  If_Statment [ if ] Found at line 14 : 18
65  Open_Expression [ ( ] Found at line 14 : 20
66  ID [ b ] Found at line 14 : 21
67  Equal [ == ] Found at line 14 : 24
68  DIGIT [ 2 ] Found at line 14 : 26
69  Close_Expression [ ) ] Found at line 14 : 27
70  Begin_Block [ { ] Found at line 14 : 29
71  Print_Statment [ print ] Found at line 16 : 25
72  Open_Expression [ ( ] Found at line 16 : 26
73  QUOTE [ " ] Found at line 16 : 27
74  CHAR [ t ] Found at line 16 : 28
75  CHAR [ h ] Found at line 16 : 29
76  CHAR [ e ] Found at line 16 : 30
77  CHAR [ r ] Found at line 16 : 31
78  CHAR [ e ] Found at line 16 : 32
79  SPACE [   ] Found at line 16 : 33
80  CHAR [ i ] Found at line 16 : 34
81  CHAR [ s ] Found at line 16 : 35
82  SPACE [   ] Found at line 16 : 36
83  CHAR [ n ] Found at line 16 : 37
84  CHAR [ o ] Found at line 16 : 38
```

```
85   SPACE [   ] Found at line 16 : 39
86   CHAR [ s ] Found at line 16 : 40
87   CHAR [ p ] Found at line 16 : 41
88   CHAR [ o ] Found at line 16 : 42
89   CHAR [ o ] Found at line 16 : 43
90   CHAR [ n ] Found at line 16 : 44
91   QUOTE [ " ] Found at line 16 : 45
92   Close_Expression [ ) ] Found at line 16 : 74
93   End_Block [ } ] Found at line 17 : 17
94   End_Block [ } ] Found at line 18 : 13
95   ID [ b ] Found at line 19 : 13
96   AssignmentStatement [ = ] Found at line 19 : 15
97   DIGIT [ 0 ] Found at line 19 : 17
98   ID [ a ] Found at line 20 : 13
99   AssignmentStatement [ = ] Found at line 20 : 15
100  DIGIT [ 1 ] Found at line 20 : 17
101  InTop [ + ] Found at line 20 : 18
102  ID [ a ] Found at line 20 : 19
103  End_Block [ } ] Found at line 21 : 9
104  End_Block [ } ] Found at line 22 : 5
105  END_OF_PROGRAM [ $ ] Found at line 23 : 1
106  Number of Errors is 0 :(
107  End of program 1
```

1.2 Test Case 2

```
1  /*LongTestCase-EverythingExceptBooleanDeclaration*/{/*IntDeclaration*/intaintba=0b=0/
2  *WhileLoop*/while(a!=3){print(a)while(b!=3){print(b)b=1+bif(b==2)
3  {/*PrintStatement*/print"there is no 1spoon"/*Thiswilldonothing*/)}}b=-0a=1+a}}$
4
5  Output:
6  Begin_Block [ { ] Found at line 25 : 52
7  TYPE [ int ] Found at line 25 : 73
8  ID [ a ] Found at line 25 : 74
9  TYPE [ int ] Found at line 25 : 77
10 ID [ b ] Found at line 25 : 78
11 ID [ a ] Found at line 25 : 79
12 AssignmentStatement [ = ] Found at line 25 : 80
13 DIGIT [ 0 ] Found at line 25 : 81
14 ID [ b ] Found at line 25 : 82
15 AssignmentStatement [ = ] Found at line 25 : 83
16 DIGIT [ 0 ] Found at line 25 : 84
17 While_Statment [ while ] Found at line 25 : 102
18 Open_Expression [ ( ] Found at line 25 : 103
19 ID [ a ] Found at line 25 : 104
20 Not_Equal [ != ] Found at line 25 : 106
21 DIGIT [ 3 ] Found at line 25 : 107
22 Close_Expression [ ) ] Found at line 25 : 108
23 Begin_Block [ { ] Found at line 25 : 109
24 Print_Statment [ print ] Found at line 25 : 114
25 Open_Expression [ ( ] Found at line 25 : 115
26 ID [ a ] Found at line 25 : 116
27 Close_Expression [ ) ] Found at line 25 : 117
```

```
28  While_Statment [ while ] Found at line 25 : 122
29  Open_Expression [ ( ] Found at line 25 : 123
30  ID [ b ] Found at line 25 : 124
31  Not_Equal [ != ] Found at line 25 : 126
32  DIGIT [ 3 ] Found at line 25 : 127
33  Close_Expression [ ) ] Found at line 25 : 128
34  Begin_Block [ { ] Found at line 25 : 129
35  Print_Statment [ print ] Found at line 25 : 134
36  Open_Expression [ ( ] Found at line 25 : 135
37  ID [ b ] Found at line 25 : 136
38  Close_Expression [ ) ] Found at line 25 : 137
39  ID [ b ] Found at line 25 : 138
40  AssignmentStatement [ = ] Found at line 25 : 139
41  DIGIT [ 1 ] Found at line 25 : 140
42  InTop [ + ] Found at line 25 : 141
43  ID [ b ] Found at line 25 : 142
44  If_Statment [ if ] Found at line 25 : 144
45  Open_Expression [ ( ] Found at line 25 : 145
46  ID [ b ] Found at line 25 : 146
47  Equal [ == ] Found at line 25 : 148
48  DIGIT [ 2 ] Found at line 25 : 149
49  Close_Expression [ ) ] Found at line 25 : 150
50  Begin_Block [ { ] Found at line 25 : 151
51  Print_Statment [ print ] Found at line 25 : 174
52  Open_Expression [ ( ] Found at line 25 : 175
53  QUOTE [ " ] Found at line 25 : 176
54  CHAR [ t ] Found at line 25 : 177
55  CHAR [ h ] Found at line 25 : 178
56  CHAR [ e ] Found at line 25 : 179
57  CHAR [ r ] Found at line 25 : 180
58  CHAR [ e ] Found at line 25 : 181
59  SPACE [   ] Found at line 25 : 182
60  CHAR [ i ] Found at line 25 : 183
61  CHAR [ s ] Found at line 25 : 184
62  SPACE [   ] Found at line 25 : 185
63  CHAR [ n ] Found at line 25 : 186
64  CHAR [ o ] Found at line 25 : 187
65  SPACE [   ] Found at line 25 : 188
66  Error: int not allowed in string [ 1 ] Found at line 25 : 189
67  CHAR [ s ] Found at line 25 : 190
68  CHAR [ p ] Found at line 25 : 191
69  CHAR [ o ] Found at line 25 : 192
70  CHAR [ o ] Found at line 25 : 193
71  CHAR [ n ] Found at line 25 : 194
72  QUOTE [ " ] Found at line 25 : 195
73  Close_Expression [ ) ] Found at line 25 : 217
74  End_Block [ } ] Found at line 25 : 218
75  End_Block [ } ] Found at line 25 : 219
76  ID [ b ] Found at line 25 : 220
77  AssignmentStatement [ = ] Found at line 25 : 221
78  Error: non recognized symbol [ - ] Found at line 25 : 222
79  DIGIT [ 0 ] Found at line 25 : 223
80  ID [ a ] Found at line 25 : 224
81  AssignmentStatement [ = ] Found at line 25 : 225
```

```
82  DIGIT [ 1 ] Found at line 25 : 226
83  InTop [ + ] Found at line 25 : 227
84  ID [ a ] Found at line 25 : 228
85  End_Block [ } ] Found at line 25 : 229
86  End_Block [ } ] Found at line 25 : 230
87  END_OF_PROGRAM [ $ ] Found at line 25 : 231
88  Number of Errors is 2 :(
89  Lexer failed :(
90  End of program 2
```

1.3 Test Case 3

```
1  {/* Comment!!!*/ () print=whileif"teststring" intstringbooleanfalse true
2  ==!=+ a 0123456789}$
3
4  Output:
5  Begin_Block [ { ] Found at line 27 : 1
6  Open_Expression [ ( ] Found at line 27 : 18
7  Close_Expression [ ) ] Found at line 27 : 19
8  Print_Statment [ print ] Found at line 27 : 25
9  AssignmentStatement [ = ] Found at line 27 : 26
10  While_Statment [ while ] Found at line 27 : 31
11  If_Statment [ if ] Found at line 27 : 33
12  QUOTE [ " ] Found at line 27 : 34
13  CHAR [ t ] Found at line 27 : 35
14  CHAR [ e ] Found at line 27 : 36
15  CHAR [ s ] Found at line 27 : 37
16  CHAR [ t ] Found at line 27 : 38
17  CHAR [ s ] Found at line 27 : 44
18  CHAR [ t ] Found at line 27 : 44
19  CHAR [ r ] Found at line 27 : 44
20  CHAR [ i ] Found at line 27 : 44
21  CHAR [ n ] Found at line 27 : 44
22  CHAR [ g ] Found at line 27 : 44
23  QUOTE [ " ] Found at line 27 : 45
24  TYPE [ int ] Found at line 27 : 49
25  TYPE [ string ] Found at line 27 : 55
26  TYPE [ boolean ] Found at line 27 : 62
27  Boolean_Value [ false ] Found at line 27 : 67
28  Boolean_Value [ true ] Found at line 27 : 72
29  Equal [ == ] Found at line 27 : 75
30  Not_Equal [ != ] Found at line 27 : 77
31  InTop [ + ] Found at line 27 : 78
32  ID [ a ] Found at line 27 : 80
33  DIGIT [ 0 ] Found at line 27 : 91
34  DIGIT [ 1 ] Found at line 27 : 91
35  DIGIT [ 2 ] Found at line 27 : 91
36  DIGIT [ 3 ] Found at line 27 : 91
37  DIGIT [ 4 ] Found at line 27 : 91
38  DIGIT [ 5 ] Found at line 27 : 91
39  DIGIT [ 6 ] Found at line 27 : 91
40  DIGIT [ 7 ] Found at line 27 : 91
41  DIGIT [ 8 ] Found at line 27 : 91
```

```
42  DIGIT [ 9 ] Found at line 27 : 91
43  End_Block [ } ] Found at line 27 : 92
44  END_OF_PROGRAM [ $ ] Found at line 27 : 93
45  Number of Errors is 0 :(
46  End of program 3
47  Number of Errors is 0 :(
48  End of program 3
```

1.4 Test Case 4

```
1   /*  Boolean If Statements output: if statement worked*/
2   {
3   if true {
4       print("if statement worked")
5   }
6   if false{
7       print("if statement failed")
8   }
9   }$
10
11  Output:
12  Begin_Block [ { ] Found at line 30 : 1
13  If_Statment [ if ] Found at line 31 : 2
14  Boolean_Value [ true ] Found at line 31 : 7
15  Begin_Block [ { ] Found at line 31 : 9
16  Print_Statment [ print ] Found at line 32 : 9
17  Open_Expression [ ( ] Found at line 32 : 10
18  QUOTE [ " ] Found at line 32 : 11
19  CHAR [ i ] Found at line 32 : 13
20  CHAR [ f ] Found at line 32 : 13
21  SPACE [   ] Found at line 32 : 14
22  CHAR [ s ] Found at line 32 : 15
23  CHAR [ t ] Found at line 32 : 16
24  CHAR [ a ] Found at line 32 : 17
25  CHAR [ t ] Found at line 32 : 18
26  CHAR [ e ] Found at line 32 : 19
27  CHAR [ m ] Found at line 32 : 20
28  CHAR [ e ] Found at line 32 : 21
29  CHAR [ n ] Found at line 32 : 22
30  CHAR [ t ] Found at line 32 : 23
31  SPACE [   ] Found at line 32 : 24
32  CHAR [ w ] Found at line 32 : 25
33  CHAR [ o ] Found at line 32 : 26
34  CHAR [ r ] Found at line 32 : 27
35  CHAR [ k ] Found at line 32 : 28
36  CHAR [ e ] Found at line 32 : 29
37  CHAR [ d ] Found at line 32 : 30
38  QUOTE [ " ] Found at line 32 : 31
39  Close_Expression [ ) ] Found at line 32 : 32
40  End_Block [ } ] Found at line 33 : 1
41  If_Statment [ if ] Found at line 34 : 2
42  Boolean_Value [ false ] Found at line 34 : 8
43  Begin_Block [ { ] Found at line 34 : 9
```

```
44  Print_Statment [ print ] Found at line 35 : 9
45  Open_Expression [ ( ] Found at line 35 : 10
46  QUOTE [ " ] Found at line 35 : 11
47  CHAR [ i ] Found at line 35 : 13
48  CHAR [ f ] Found at line 35 : 13
49  SPACE [    ] Found at line 35 : 14
50  CHAR [ s ] Found at line 35 : 15
51  CHAR [ t ] Found at line 35 : 16
52  CHAR [ a ] Found at line 35 : 17
53  CHAR [ t ] Found at line 35 : 18
54  CHAR [ e ] Found at line 35 : 19
55  CHAR [ m ] Found at line 35 : 20
56  CHAR [ e ] Found at line 35 : 21
57  CHAR [ n ] Found at line 35 : 22
58  CHAR [ t ] Found at line 35 : 23
59  SPACE [    ] Found at line 35 : 24
60  CHAR [ f ] Found at line 35 : 25
61  CHAR [ a ] Found at line 35 : 26
62  CHAR [ i ] Found at line 35 : 27
63  CHAR [ l ] Found at line 35 : 28
64  CHAR [ e ] Found at line 35 : 29
65  CHAR [ d ] Found at line 35 : 30
66  QUOTE [ " ] Found at line 35 : 31
67  Close_Expression [ ) ] Found at line 35 : 32
68  End_Block [ } ] Found at line 36 : 1
69  End_Block [ } ] Found at line 37 : 1
70  END_OF_PROGRAM [ $ ] Found at line 37 : 2
71  Number of Errors is 0 :(
72  End of program 4
```

1.5   Test Case 5

```
1
2   /*This is suppose to crash at the top since there is a space before the first line */
3    {
4         int a
5         a = 3
6         int b
7         b = 4
8         a = b
9         print(a)
10        if (a == b) {
11            print(a)
12        }
13       }$
14
15  Output:
16  Begin_Block [ { ] Found at line 40 : 2
17  TYPE [ int ] Found at line 41 : 10
18  ID [ a ] Found at line 41 : 12
19  ID [ a ] Found at line 42 : 8
20  AssignmentStatement [ = ] Found at line 42 : 10
21  DIGIT [ 3 ] Found at line 42 : 12
```

```
22  TYPE [ int ] Found at line 43 : 10
23  ID [ b ] Found at line 43 : 12
24  ID [ b ] Found at line 44 : 8
25  AssignmentStatement [ = ] Found at line 44 : 10
26  DIGIT [ 4 ] Found at line 44 : 12
27  ID [ a ] Found at line 45 : 8
28  AssignmentStatement [ = ] Found at line 45 : 10
29  ID [ b ] Found at line 45 : 12
30  Print_Statment [ print ] Found at line 46 : 12
31  Open_Expression [ ( ] Found at line 46 : 13
32  ID [ a ] Found at line 46 : 14
33  Close_Expression [ ) ] Found at line 46 : 15
34  If_Statment [ if ] Found at line 47 : 9
35  Open_Expression [ ( ] Found at line 47 : 11
36  ID [ a ] Found at line 47 : 12
37  Equal [ == ] Found at line 47 : 15
38  ID [ b ] Found at line 47 : 17
39  Close_Expression [ ) ] Found at line 47 : 18
40  Begin_Block [ { ] Found at line 47 : 20
41  Print_Statment [ print ] Found at line 48 : 15
42  Open_Expression [ ( ] Found at line 48 : 16
43  ID [ a ] Found at line 48 : 17
44  Close_Expression [ ) ] Found at line 48 : 18
45  End_Block [ } ] Found at line 49 : 8
46  End_Block [ } ] Found at line 50 : 6
47  END_OF_PROGRAM [ $ ] Found at line 50 : 7
48  Number of Errors is 0 :(
49  End of program 5
```

1.6 Test Case 6

```
1   {
2           int a
3           boolean b
4           string c
5         a = 9
6          b = true
7           {
8             print(a)
9             print(b)
10            b = false
11            c = "hello world"
12            int b
13            b = 0
14            {
15              print(c)
16              a = 1 + 2 + a
17              {
18                print(b)
19              }
20            }
21            b = a
22            print(b)
```

```
23            }
24            print(b)
25
26
27
28
29    }$
30
31    Begin_Block [ { ] Found at line 52 : 1
32    TYPE [ int ] Found at line 53 : 11
33    ID [ a ] Found at line 53 : 13
34    TYPE [ boolean ] Found at line 54 : 15
35    ID [ b ] Found at line 54 : 17
36    TYPE [ string ] Found at line 55 : 14
37    ID [ c ] Found at line 55 : 16
38    ID [ a ] Found at line 56 : 8
39    AssignmentStatement [ = ] Found at line 56 : 10
40    DIGIT [ 9 ] Found at line 56 : 12
41    ID [ b ] Found at line 57 : 9
42    AssignmentStatement [ = ] Found at line 57 : 11
43    Boolean_Value [ true ] Found at line 57 : 16
44    Begin_Block [ { ] Found at line 58 : 9
45    Print_Statment [ print ] Found at line 59 : 16
46    Open_Expression [ ( ] Found at line 59 : 17
47    ID [ a ] Found at line 59 : 18
48    Close_Expression [ ) ] Found at line 59 : 19
49    Print_Statment [ print ] Found at line 60 : 16
50    Open_Expression [ ( ] Found at line 60 : 17
51    ID [ b ] Found at line 60 : 18
52    Close_Expression [ ) ] Found at line 60 : 19
53    ID [ b ] Found at line 61 : 12
54    AssignmentStatement [ = ] Found at line 61 : 14
55    Boolean_Value [ false ] Found at line 61 : 20
56    ID [ c ] Found at line 62 : 12
57    AssignmentStatement [ = ] Found at line 62 : 14
58    QUOTE [ " ] Found at line 62 : 16
59    CHAR [ h ] Found at line 62 : 17
60    CHAR [ e ] Found at line 62 : 18
61    CHAR [ l ] Found at line 62 : 19
62    CHAR [ l ] Found at line 62 : 20
63    CHAR [ o ] Found at line 62 : 21
64    SPACE [   ] Found at line 62 : 22
65    CHAR [ w ] Found at line 62 : 23
66    CHAR [ o ] Found at line 62 : 24
67    CHAR [ r ] Found at line 62 : 25
68    CHAR [ l ] Found at line 62 : 26
69    CHAR [ d ] Found at line 62 : 27
70    QUOTE [ " ] Found at line 62 : 28
71    TYPE [ int ] Found at line 63 : 14
72    ID [ b ] Found at line 63 : 16
73    ID [ b ] Found at line 64 : 12
74    AssignmentStatement [ = ] Found at line 64 : 14
75    DIGIT [ 0 ] Found at line 64 : 16
76    Begin_Block [ { ] Found at line 65 : 12
```

```
77   Print_Statment [ print ] Found at line 66 : 19
78   Open_Expression [ ( ] Found at line 66 : 20
79   ID [ c ] Found at line 66 : 21
80   Close_Expression [ ) ] Found at line 66 : 22
81   ID [ a ] Found at line 67 : 15
82   AssignmentStatement [ = ] Found at line 67 : 17
83   DIGIT [ 1 ] Found at line 67 : 19
84   InTop [ + ] Found at line 67 : 21
85   DIGIT [ 2 ] Found at line 67 : 23
86   InTop [ + ] Found at line 67 : 25
87   ID [ a ] Found at line 67 : 27
88   Begin_Block [ { ] Found at line 68 : 15
89   Print_Statment [ print ] Found at line 69 : 22
90   Open_Expression [ ( ] Found at line 69 : 23
91   ID [ b ] Found at line 69 : 24
92   Close_Expression [ ) ] Found at line 69 : 25
93   End_Block [ } ] Found at line 70 : 15
94   End_Block [ } ] Found at line 71 : 12
95   ID [ b ] Found at line 72 : 12
96   AssignmentStatement [ = ] Found at line 72 : 14
97   ID [ a ] Found at line 72 : 16
98   Print_Statment [ print ] Found at line 73 : 16
99   Open_Expression [ ( ] Found at line 73 : 17
100  ID [ b ] Found at line 73 : 18
101  Close_Expression [ ) ] Found at line 73 : 19
102  End_Block [ } ] Found at line 74 : 9
103  Print_Statment [ print ] Found at line 75 : 13
104  Open_Expression [ ( ] Found at line 75 : 14
105  ID [ b ] Found at line 75 : 15
106  Close_Expression [ ) ] Found at line 75 : 16
107  End_Block [ } ] Found at line 80 : 1
108  END_OF_PROGRAM [ $ ] Found at line 80 : 2
109  Number of Errors is 0 :(
110  End of program 6
```

2 References

Java Regular Expression
Enum
Switch cases (Was more of an idea)
Double Quote For Regular Expressions
Groups in Regular Expressions
More Group stuff
Index for Regular Expressions

I used these all to help me understand how to use Regular Expression and get the data that I need in order to make the tokens. Enum and Cases where not used but more of an ideas that I was thing about using but never did end up using them