Lex Scanner Grading Criteria and DEMO

This assignment is worth 100 points (15% of the total score)

- Public test data：test1.java, test2.java, test3.java (15% each)
- Hidden test data：test4.java, test5.java (15% each)
- Symbol Table function implementation：10%
- Oral Q&A:：10%
- Readme.pdf：5%

Your Java Scanner must be able to handle the following cases.

(Same as the assignment specifications):

1. Tokens that will be sent to the Parser:：

    I. **Symbols**： 『,』、『:』、『;』、『()』、『[]』、『{}』

    II. **Operators**：『+ ++』、『- --』、『*』、『/ %』、『=』、『< <= >= > == !=』、『&& || !』

    III. **Reserved word**：boolean, break, byte, case, char, catch, class, const, continue, default, do, double, else, extends, false, final, finally, float, for, if, implements, int, long, main, new, print, private, protected, public, return, short, static, string, switch, this, true, try, void, while

    IV. **Identifier(ID)**：Java defines identifiers that can use any Unicode symbol, but conventionally they are based on the 26 uppercase and lowercase English letters, supplemented by numbers, underscores, and the dollar sign ($).

    peter, _db, a1 are valid IDs.

    1a , ^db, #db   are invalid IDs.

    V. **Integer**：A number that can be positive or negative.

VI. **Float**：Floating-point numbers can be positive or negative, and they also have decimal point representation and scientific notation.

For example:

1.0, 3.14, 7E-2, 12.25e+6, -7.5E+3, -7.5E+3f are valid IDs.

VII. **String**：String characters must be enclosed in double quotes.

For example:

Empty strings (""), whitespace (" "), "You'll see" represent the strings "You'll see" and "\"" are valid strings.

Single quotes ('aaa'), strings without escape characters ("""), "\", and strings with only one double quote ("ab") are invalid strings.

2. Tokens that will not be passed to the Parser:
   I. **Whitespace**：whitespace, tabs, newlines.
   II. **Comments**：Comments are divided into '/*' and '*/' that can span multiple lines, and '//' that cannot span multiple lines.

3. Recovery：
   I. Your Scanner can process all inputs without stopping at errors.

4. Symbol Table：
   I. Your Symbol Tables can store all of IDs and output all of IDs in the Scan result, and IDs with the same name cannot be printed repeatedly.
   II. II. 4 functions: create(), lookup(), insert(), dump().