CS321 Languages and Compiler Design I, Fall 2014 10/7/14

Prof. Jingke Li (FAB120-06, lij@pdx.edu), Tue 10:00-11:50 @SRTC 155, Labs: Tue 14:00-15:50 & Wed 12:00-13:50 @FAB 88-10

The miniJava Lexical Rules

The miniJava language's token definitions follow Java's lexical rules in most cases; but miniJava differs from Java in some places (for simplification purpose).

- miniJava is case sensitive — upper and lower-case letters are *not* considered equivalent.
- The following are miniJava's *reserved* words — they must be written in the exact form as given:

  ```
  class extends static public void int boolean new if else while return
  main true false String System out println
  ```

  Note that the words in the second row are not reserved in Java. They are made reserved in miniJava to simplify compatibility with Java. (For example, you can use `System.out.println` to print in miniJava even though miniJava does not support packages.)
- *Identifiers* are strings of letters and digits starting with a letter, *excluding* the reserved words. There is no length limit.
- *Integer* literals contain only digits; their values must be in the range 0 to $2^{31} - 1$. Note that an integer literal's value is always non-negative. To get a negative integer value, an unary minus operator can be used.
- *String* literals begin and end with a double quote (`"`) and contain any sequence of ASCII characters, except double quotes (`"`), carridge returns (`\r`), and newlines (`\n`). A string can be of arbitrary length, including zero.
- *Comments* can be in two forms: a single-line comment starts with `//` and ends with a (invisible) newline character (`\n`); multi-line comments are enclosed in the pair `/*`, `*/`; they cannot be nested. All ASCII characters are legal in a comment.
- The following are miniJava's *operators* and remaining *delimiters*:

  ```
  operator  = "+"|"-"|"*"|"/"|"&&"|"||"|"!"|"=="|"!="|"<"|"<="|">"|">="
  delimiter = "="|";"|","|"."|"("|")"|"["|"]"|"{"|"}"
  ```

That is all. There are no other lexical entities in miniJava. Many Java lexical features are not supported in miniJava: for example, there are no floating point literals or hexadecimal numbers.