Lecture 7: Characters and Strings

Jan 29 2015
char data type

- is a built-in (primitive) data type of Java
- is used to represent alpha-numerical information (characters) inside the computer
- uses the Unicode to encode characters from many different kinds of languages in the world
- uses 2 bytes of memory to store the Unicode value
- The Unicode includes the ASCII code which is used to encode English characters
Unicode

• Unicode provides a unique number for every character

• Example:
  – The number 65 of Unicode represents the (English) character A
  – The number 35 of Unicode represents the character #
  – The number 948 of Unicode represents the Greek character δ
  – The number 24373 of Unicode represents the Chinese character 張
The value 0 --- 127 of the Unicode is called the ASCII code.
The American Standard Code for Information Interchange (ASCII) code was designed to represent characters in the English alphabet.

http://www.ascii-code.com/
Character literals

• We write numeric literals as
  5
  8.23
  31.4e-1

• We can NOT write character literals
  A
  a
  8

Because they get confused with identifiers!
(for example, names of variables)
Character literals (continued)

• A character literal is written between quotes '...' 

Examples:
  – 'A' is the character literal A (Unicode code value 65)
  – 'a' is the character literal a (Unicode code value 97)
  – ‘8' is the character literal 1 (Unicode code value 56)
Important concept!!!

• A char variable is an integer typed variable.
• The expression 'A' is an integer (= 65) in a Java program !!!!
• Try running this program:
  public class Char01B {
      public static void main(String[] args) {
          char a;
          a = 65;  // 65 = Char code for the character 'A'
          System.out.println(a);  // Prints 'A' !!!
          int I;
          i = 'A' + 1;  // 'A' is just an integer !!!
          System.out.println(i);  // Prints 66 !!!
      }
  }
}
Defining character typed variables

• Syntax to define an character typed variable:
  ```
  char NameOfVariable;
  ```

• Notes:
  – The keyword char announces the variable definition clause
  – The NameOfVariable is an identifier which is the name of the variable.
  – The variable definition clause is must be ended with a semi-colon ";
  – A char typed variable can store the Unicode of one character
Defining character types

Example:

```java
public class Char01 {
    public static void main(String[] args) {
        char a;
        a = 'b';  // a = ASCII code of 'b' (98)
        System.out.println(a);  // Prints the character b !!!
    }
}
```
Printing characters vs integers

• Important difference in printing character variables and integer variables
• Although char type is a kind of integer type, it is treated differently when we want to "see" the value (e.g., printing)
• The print and println methods in System.out can tell the difference between:
  – char (which is an integer type !)
  – Other kinds of integer types.
• Different encoding methods used in printing:
  – When printing a char typed integer, the print and println methods will use the Unicode table to print the corresponding character.
  – When printing other kinds of integer types, the print and println methods will use the binary number encoding.
  – (I.e.: 00000001 is 1, 00000010 is 2, and so on)
**char** Summary

- A char type variable consists of 2 bytes
- It contains the Unicode code of some character
- A char type variable is an integer type variable that contains a positive number