# Algorithms and Data Structures (WIP)

Based on the online MSc in Computer Science by the University of York Notes taken by Nathan McKeown-Luckly (Discord Username: skeletman)

## Autumn 2023

# **Contents**

1	Software and the Java Programming language				
	1.1	Software and Code	3		
	1.2	The Java Programming Language	4		
	1.3	Command Line Interfaces (CLIs) and Integrated Development			
		Environments (IDEs)	6		
	1.4	User Interfaces	6		
	1.5	"Hello World" Program	6		
2	Vari	iables	7		
	2.1	Primitive Types	7		
	2.2	Declaring and assigning variables	7		
	2.3	Arithmetic Operators	8		
	2.4	Comparison Operators	9		
	2.5	Logical Operators	9		
3	Program design				
	3.1	Pseudocode	10		
4	Sele	ection and Iteration	11		
	4.1	If Statements	11		
	4.2	If-Else Statements	11		
	4.3	Switch Statements	11		
	4.4	For Loops	12		
	4.5	While loops	12		
	4.6	Do-While loops	13		
	4.7	Break statements	13		
	4.8	Continue statements	13		

5	Met	hods	14	
	5.1	Declaring, Defining and Calling Methods	14	
	5.2	Modifiers for Methods	15	
	5.3	Method Overloading	16	
6	Classes and Objects			
	6.1	Class Declaration	17	
	6.2	The Class Body	17	
	6.3	Instantiating an object	18	
	6.4	Accessing the Public Methods and Attributes of an Object or Class	18	
	6.5	The "this" keyword	18	
	6.6	Modifiers for Classes	19	
	6.7	Modifiers for Attributes	19	
7	Assignment and Equality			
	7.1	The Assignment Operator (=)	20	
	7.2	The new Keyword	21	
	7.3	The Equality Operator $(==)$	21	
8	Common Classes			
	8.1	Object	22	
	8.2	String	22	
	8.3	Wrappers	23	
	8.4	Scanner	23	
Inc	dex		25	

## 1 Software and the Java Programming language

#### 1.1 Software and Code

**Definition 1.1** (Computer Program). A *program* is a set of instructions to be given to a computer.

**Definition 1.2** (Software). *Software* means a program or set of programs.

**Definition 1.3** (Application Software). *Application software* or applications are software that is designed for the user to interact with the computer.

**Examples.** Word processor, image editor, computer game, spreadsheet.

**Definition 1.4** (System Software). *System software* is software that provides a platoform for other software to function.

**Examples.** Operating system, network software, game engine, search engine, compiler.

**Definition 1.5** (Embedded Software, Embedded System). *Embedded software* refers to software designed to control devices that are do not primarily function as computers. Such devices are known as *embedded systems*.

**Examples.** Microwave ovens, cars, modems, robots.

A computer can only understand instructions given to it in binary, we call such instructions *machine code*. However, this is not very convenient for humans, so we instead write in a human-readable *programming language*, for example *Java*. The code we write in these are translated into machine code by a piece of software called a *compiler*, to be executed by the computer.

**Definition 1.6** (Program code/Source Code). Code as it was written by a human, often in a programming language easily readable to humans, is called *program code* or *source code*.

**Definition 1.7** (Syntax). The rules that must be followed for a piece of code to be valid in a programming language is called *syntax*.

## 1.2 The Java Programming Language

Java is an *object oriented* programming language.

**Definition 1.8** (Object Oriented Programming). *Object oriented programming* or *OOP* is a programming paradigm based on the notion of objects. Objects may contain both data (fields, attributes or properties), and code (methods).

**Example.** Suppose we wish to simulate a physical scenario, where a cube has some force applied to it. In such a physics simulation, a cube object may have a field mass, and a method calculateAcceleration which takes an input of a force, and uses the mass field to calculate the acceleration that the cube would experience.

There are other programming paradigms: languages such as Haskell use the functional programming paradigm, which differs greatly from object oriented programming.

Compilers for many programming languages, such as C, compile program code directly into machine code. In figure 1 we see that this requires a different compiler for each type of machine for the program to be run on.

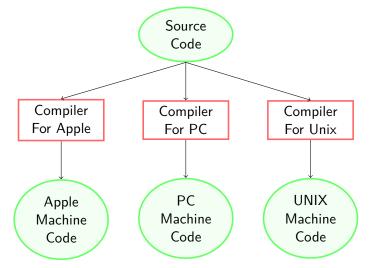


Figure 1: A typical pipeline for code compilation

However, Java is *platform-independent*: a compiled Java program can run on any type of machine. This is because Java compiles to *Java byte code*, which then can be run by the *Java virtual machine* or *JVM*, which is a piece of

software that runs on various types of machines to translate the Java byte code into machine code. There are versions of JVM for Windows, Mac, Unix, Linux, mobile phones, and even many embedded systems.

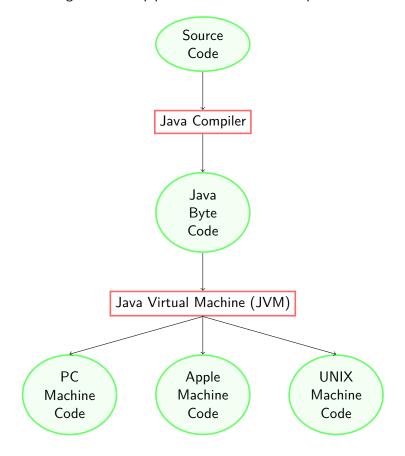


Figure 2: The pipeline for Java code compilation

**Definition 1.9** (Libraries, Packages). *Libraries* and *packages* are pre-compiled Java modules for use in your own programs.

The Java Runtime Environment or JRE is the JVM along with the Java libraries. The Java Development Kit or JDK consists of the JRE, the Java compiler, and some other tools for java development.

# 1.3 Command Line Interfaces (CLIs) and Integrated Development Environments (IDEs)

Henceforth we assume that the JDK is properly installed. If we wish to compile Java code using the command line, it should be saved as a .java file. Then we use javac to compile the program into a .class file, which contains Java byte code. Then to launch this with the JVM, we run the java command. For example, to compile a file called test.java, we would navigate to its directory, and enter javac test.java, and to launch the program we enter java test.

During development, instead of doing this every time we wish to test a program, programmers will often use an *Integrated Devevlopment Environment* or *IDE* to handle this. An IDE will often have many features designed for development, such as syntax highlighting, version management tools and debugging tools.

#### 1.4 User Interfaces

**Definition 1.10** (User Interfaces). A *user interface* or UI is the part of a program that prompts the user for input and provides the program output.

A console application is an application which uses a text based user interface. Alternatively, an application can use a graphical user interface or GUI, which allows for elements such a text boxes and buttons.

## 1.5 "Hello World" Program

Once a Java program is compiled, running the program will execute the main method in the class of the .java file that was compiled. In listing 1, the main method is executed and hence the words "Hello World" are printed to the console when the program is executed.

Listing 1: Hello World

```
public class HelloWorld
{
    public static void main(String[] args)
    {
        System.out.println("Hello World");
    }
}
```

For more on methods, see section 5. For more on classes, see section 6.

### 2 Variables

## 2.1 Primitive Types

**Definition 2.1** (Data type). A *data type* is a collection of data values, or the representation of these values on a machine.

In the Java programming language, the *primitive types* or *scalar types* are a collection of basic types which take up a fixed amount of memory. The primitive types along with the data which they store are displayed in figure 3.

Java type	Description of Data	Range of values
byte	Integers	-128 to $127$
short	Integers	-32,768 to $32,767$
int	Integers	-2,147,483,648 to
		2, 147, 483, 647
long	Integers	-9,223,372,036,854,775,808 to
		9,223,372,036,854,775,807
float	Real numbers	$\pm 1.4 \times 10^{-45} \text{ to } \pm 3.4 \times 10^{38}$
double	Real numbers	$\pm 4.9 \times 10^{-324} \text{ to } \pm 1.8 \times 10^{308}$
char	Characters	Unicode character set
boolean	True or false	True or false

Figure 3: Primitive types in Java

# 2.2 Declaring and assigning variables

**Definition 2.2** (Variable and Declaration). When we create a named location in memory to store a value of a given type, we are *declaring* a *variable*. They are called variables as the content may vary during the runtime of the program.

In Java, we declare variables by writing: dataType variableName. For example, if we wanted to declare an integer variable named score, we'd write

```
int score;
```

**Definition 2.3** (Assigning variables). When we give a variable a value, this is called *assigning* the variable a value.

In Java, we use the = symbol to assign variables, for example

```
score = 10;
```

We can declare and assign in the same line of code:

```
int score = 10;
```

Variables can be declared as *constant* by being preceded by the keyword final.

```
final double speedOfLight = 300000000;
```

## 2.3 Arithmetic Operators

Arithmetic operations take numeric values and output a numeric value of the same type. In Java, these can be done using the operators listed in figure 4.

Figure 4: Arithmetic operations in Java

Operation	Java operator
Addition	+
Subtraction	_
Multiplication	*
Division	/
Remainder or Modulus	%

**Definition 2.4** (Overloading). A method or operator is said to be *overloaded* if it has multiple definitions, which may depend on context such as the data type or number of inputs.

**Example.** The arithmetic operations are overloaded so we can use them on all of the primitive numeric types. For example / is defined for both int and double. In the int case, it returns a rounded down value, whereas in the double case it returns a precise (as possible given double precision) value: 3 / 2 will have value 1 of type int, but 3.0 / 2.0 will have value 1.5 of type double.

## 2.4 Comparison Operators

Comparison operations take numeric values and output a boolean value. In Java, these can be done using the operators listed in figure 5.

Figure 5: Comparison operations in Java

Operation	Java operator
Equal to	==
Not equal to	!=
Less than	<
Greater than	>
Less than or equal to	<=
Greater than or equal to	>=

**Note.** The == operator (and by extension !=) only compares the value for *primitive types*. We will see later in section 7.3 that for objects, == is unsuitable for comparing values.

## 2.5 Logical Operators

Logical operations take boolean values and output a boolean value. These can be done using the operators listed in figure 6

Figure 6: Logical operations in Java

Operation	Java operator
Logical NOT	!
Logical OR	11
Logical AND	&&

# 3 Program design

**Definition 3.1** (Program Design and Implementation). *Program design* is working out how to go about developing software. *Implementation* is the process of developing the software.

#### 3.1 Pseudocode

*Pseudocode* refers to code written with very loose syntax, not bound by that of any specific language.

Algorithm 1: An example of pseudocode to calculate a price after tax

- 1 display prompt for price;
- 2 price ← user input;
- 3 display prompt for tax percentage;
- 4 tax ← user input;
- **5 return** price \* (1 + tax/100);

**Algorithm 2:** An alternative way of writing the same algorithm as in algorithm  $\boldsymbol{1}$ 

- 1 DISPLAY prompt for price;
- 2 ENTER price;
- 3 DISPLAY prompt for tax percentage;
- 4 ENTER tax;
- **5** RETURN price \* (1 + tax/100);

# 4 Selection and Iteration

#### 4.1 If Statements

An *if statement* executes a piece of code conditional on a boolean condition. The Java syntax for an if statement is

```
if(condition)
{
    code-if-condition-true
}
```

#### 4.2 If-Else Statements

An *if-else statement* executes a piece of code conditional on a boolean condition, and if not executes another piece of code. The Java syntax for an if-else statement is

```
if(condition)
{
    code-if-condition-true
}
else
{
    code-if-condition-false
}
```

### 4.3 Switch Statements

A *switch statement* is a switch between many pieces of code based on the value of a varaible. In Java, the syntax for a switch statement is

```
switch(variable)
{
    case value1: code
        break;
    ...
    case valueN: code
        break;
    default: code
        break;
}
```

#### 4.4 For Loops

A *for loop* iterates a piece of code a fixed number of times, with a varying parameter. The Java syntax for a for loop is

```
for(initialisation; termination; increment)
{
    code
}
```

The code in the header behaves as follows

- initialisation runs once, at the beginning of the loop's execution
- termination should be a boolean expression: when this evaluates to false the loop will terminate.
- increment runs after each iteration of the loop

**Example.** A typical use of a for loop would be to add all the integers between 0 and 10. This would be written as follows

Listing 2: A typical use of a for loop

```
int total = 0;
for(int i = 0; i <= 10; i++)
{
    total += i;
}</pre>
```

In listing 2 we initialise a new variable int i = 0, and each iteration of the loop i++ is executed, incrementing i by 1. The loop terminates when i <= 10 no longer holds.

**Note.** When a variable is declared in the initialisation of a for loop, its scope is within that for loop.

## 4.5 While loops

A *while loop* iterates a piece of code while a boolean condition is true. The Java syntax for a for loop is

```
while(condition)
{
    code
}
```

## 4.6 Do-While loops

A *do-while loop* executes a piece of code once, then iterates it while a boolean condition is true. The Java syntax for a for loop is

```
do
{
    code
} while (condition)
```

#### 4.7 Break statements

While inside a loop, you can use a *break statement* to terminate iteration of the loop.

**Example.** If you wanted to terminate a for loop when its counter reaches 3, you could write the following

```
for(int i = 0; i < 100; i++){
   if (i == 3)
   {
      break;
   }
}</pre>
```

#### 4.8 Continue statements

While inside a loop, you can use a *continue statement* to skip to the next iteration of the loop.

**Example.** If you wanted to skip an iteration in a for loop when its counter reaches 3, you could write the following

```
for(int i = 0; i < 100; i++){
    if (i == 3)
    {
        continue;
    }
}</pre>
```

## 5 Methods

#### 5.1 Declaring, Defining and Calling Methods

Methods are defined using the following syntax:

```
modifiers return-type method-name(parameters)
{
    method-body
}
```

Convention is that a method name will have the first letter lowercase. The first line of the method declaration is called the *method header*.

**Example.** We could use a method to take an integer and return it multiplied by 2.

Listing 3: An example method

```
public static int multByTwo(int x)
{
    return x * 2;
}
```

In this example public and static are modifiers, int is the return type, multByTwo is the method name, int x is the parameter, and return x \* 2; is the method body.

Methods are called using the following syntax:

```
method-name(parameters)
```

**Example.** If we wanted to use the method in listing 3 to multiply 5 by 2 and assign it to variable num, we would write the following code.

```
int num = multByTwo(5);
```

Methods may affect the state of the program, so sometimes we don't need to put the output into a variable. We may use void as the return type for such methods. Regardless, any method can also be run without assigning the return value to a variable. The following example would run the method multByTwo, but not put the return value in a variable.

```
multByTwo(5);
```

#### 5.2 Modifiers for Methods

In a method header in Java, you can use *modifiers* to modify various behaviours of the method: one *access modifier* then potentially a combination of *non-access modifiers*. The access modifiers are as follows

- default (if you don't specify a modifier) The method is accessible only in the same package.
- public The method is accessible in all classes.
- private The method is only accessible in the declared class.
- protected The method is accessible in the same package and subclasses.

#### The non-access modifiers are

- final The method cannot be overridden or modified.
- static The method belongs to the class rather than objects of that class.
- abstract (can only be used in an abstract class) The method has no method body, which must be supplied by any subclass.
- transient Makes the method ignored when serialising the object containing it.
- synchronised The method can be accessed only by a single thread at a time.

## 5.3 Method Overloading

It is allowed in Java to give multiple methods the same name, as long as they accept a different set of parameters. This is called *overloading*. For example, the following is valid Java code.

```
static int max(int firstIn, int secondIn)
      if(firstIn > secondIn)
      {
          return firstIn;
      }
      else
          return secondIn;
10
      }
11
12
13 static int max(int firstIn, int secondIn, int thirdIn)
14 {
      return(max(firstIn, max(secondIn, thirdIn)))
15
16 }
```

**Definition 5.1** (Polymorphism). Methods and operators with the same name being allowed to perform different function is called *polymorphism*.

# 6 Classes and Objects

In definition 1.8, we defined object oriented programming, and said, objects may contain both data (fields, attributes or properties), and code (methods). The idea is that a *class* is a blueprint for an *object*.

#### 6.1 Class Declaration

The syntax for class declaration is as follows:

```
modifiers class class-name
{
    class-body
}
```

Convention is that a class name will have the first letter capitalised.

## 6.2 The Class Body

The class body may consist of method and attribute declarations. You may also provide a special method called a *constructor*, which has the same name as the class, and the return type is not written. For example, if we were to run a physics simulation of an object with mass under a gravitational force, a class PhysicsBox may be defined as follows

Listing 4: PhysicsBox

```
public class PhysicsBox
{
    private double mass;
    private static double gravity = 9.81;

public PhysicsBox(double massIn) {
        mass = massIn;
    }

public static void setGravity(double gravityIn) {
        gravity = gravityIn;
    }

public double getGravityForce() {
        return mass * gravityConst;
    }
}
```

Notice that we have declared the attributes as private, so the only way to access the attributes are through the class' methods. This kind of data protection is called *encapsulation*.

Also, we have *initialised* the gravity attribute, which sets the value at the start of the lifetime of the PhysicsBox class, or if it were not static, at the start of the lifetime of each PhysicsBox object. Java always initialises attributes to a given value, numerical primitives to 0, boolean to false, Object to null

## 6.3 Instantiating an object

Creating a new object is called *instantiation*. This is achieved in Java as follows

```
class-name object-name = new class-name(parameters)
```

For example, to create an object called box of the PhysicsBox class above, with mass 5.0, we would write

```
PhysicsBox box = new PhysicsBox(5.0)
```

# 6.4 Accessing the Public Methods and Attributes of an Object or Class

To access the public methods of an object we have instantiated, we use the *dot* operator.

```
object-name.method-name(parameters)
```

Similarly for attributes we use

object-name.attribute-name

For example, to get the force acting on our box

```
double force = box.getGravityForce();
```

If the static modifier has been used in the definition, instead we need to refer to the class name rather than the object name, for example, to set the gravity constant to 1, we would write

```
PhysicsBox.setGravity(1.0);
```

## 6.5 The "this" keyword

The this keyword refers to the object which owns a method or constructor. It is useful to distinguish between class attributes and parameters with the same name.

#### 6.6 Modifiers for Classes

When declaring an class, you can use *modifiers* to modify various behaviours of the class: one *access modifier* then potentially a combination of *non-access modifiers*. The access modifiers are as follows

- default (if you don't specify a modifier) The class is accessible only by classes in the same package.
- public The class is accessible in any other class.

The non-access modifiers are

- final The class cannot be inherited by other classes.
- abstract Objects of this class cannot be created.

#### 6.7 Modifiers for Attributes

When declaring an attribute, you can use *modifiers* to modify various behaviours of the attribute: one *access modifier* then potentially a combination of *non-access modifiers*. The access modifiers are as follows

- default (if you don't specify a modifier) The attribute is accessible only in the same package.
- public The attribute is accessible in all classes.
- private The attribute is only accessible in the declared class.
- protected The attribute is accessible in the same package and subclasses.

The non-access modifiers are

- final The attribute cannot be overridden or modified.
- static The attribute belongs to the class rather than objects of that class.
- transient Makes the attribute ignored when serialising the object containing it.
- synchronised The attribute can be accessed only by a single thread at a time.
- volatile The attribute is always read from main memory, not from a thread-local cache. In practice, this means any thread will be able to read the most recently updated value for this attribute.

# 7 Assignment and Equality

# 7.1 The Assignment Operator (=)

The = operator assigns a value to a variable. For example, the following assigns the value 10 to x.

```
int x = 10;
```

When dealing with objects instead of primitives, the value that is assigned is the memory address of the object on the right hand side.

Example. Suppose we have a class Rectangle

Listing 5: Rectangle

```
public class Rectangle
      private double length;
      private double width;
      public Rectangle(double length, double width)
           this.length = length;
           this.width = width;
10
11
      public setLength(double length)
12
13
           this.length = length;
14
15
16
      public setWidth(double width)
17
18
           this.width = width;
19
      }
20
22
      public double getArea()
23
           return length*width;
24
      }
25
26 }
```

The listing 6, line 1 creates a new Rectangle in memory with length 1.0 and width 1.0, and assign rect to its memory address.

In line 2, rect2 would be assigned the value of rect, i.e. the memory address of the rectangle.

Then in line 3, we use the setWidth method to set rect2.width to 5.0. As rect2 points to the same object as rect, this method sets rect.width to 5.0. In line 4, we set area to rect.getArea(). This returns 5.0, as rect.length is 1.0 and rect.width is 5.0.

Listing 6: A sample of code to interact with the Rectangle class

```
Rectangle rect = new Rectangle(1.0,1.0);
Rectangle rect2 = rect;
rect2.setWidth(5.0);
double area = rect.getArea();
```

**Definition 7.1** (Aliasing). *Aliasing* is when an object is assigned multiple names. In listing 6, we see rect2 is an *alias* of rect.

#### 7.2 The new Keyword

The *new keyword* creates a new object, at runtime it assigns an area in memory on the heap to store that object, and returns a reference to that object. It has one argument, being the constructor of the object to be initialised.

## 7.3 The Equality Operator (==)

The == operator will tell you if two non-null references refer to the same object in memory. This is called *reference equality*. In practice, this means that when comparing objects using ==, you are comparing their memory address, not their data. WARNING: It is a common mistake to use == to compare the data contained by an object. You should instead use a method which compares the data you require to be compared (often the overridden equals method). With the following declarations, we would have that a==b would return true, but a==c would return false, despite a and c containing the same data!

```
String a = new String("str");
String b = a;
String c = new String("str");
```

In contrast, for primitives, the == operator returns the *value equality*, i.e. true if and only if the variables contain the same value.

#### 8 Common Classes

#### 8.1 Object

Every class is a subclass of Object, so inherits its methods. Examples of such methods are

- public boolean equals(Object obj) Returns true if the current object is equal to obj. The equals method for the class Object this is the same as the == operator. WARNING: This means that two objects may be indistinguishable by the data and methods they hold, but are still distinguished by the equals method! As a result, for most classes the equals method is overriden.
- protected Object clone() Produces a clone of the current object, such that x.clone() == x will return false.
- public String toString() Returns a string representation of the object. Usually overridden to give more useful information.

## 8.2 String

A *string* is a sequence of characters. The String class represents *immutable* strings, that is, they cannot be modified by methods.

**Definition 8.1** (String Literal). A *string literal* is a piece of code which represents a fixed string.

In Java, string literals are represented as a string within quotation marks, for example "This is a string literal". String literals are implemented as instances of the class String.

Use of the same string literal twice will reference to the same String object in memory. For example, with the following declarations, a == b would return true.

```
String a = "str"
String b = "str"
```

The string class has many associated methods, such as:

- public int length() Returns the length of the string.
- public boolean equals (Object obj) Returns true if and only if obj is a String object and has the same value as the current string. Overrides the equals method for the Object class.

- public String substring(int beginIndex, int endIndex) -Returns the substring of this string beginning with the beginIndex<sup>th</sup> character, ending with the (endIndex-1)<sup>th</sup> character.
- public String toLowerCase() Returns the string in lowercase.
- public String toUpperCase() Returns the string in uppercase.

#### 8.3 Wrappers

For each primitive data type, there is a *wrapper class* which hold a single field of the corresponding primitive type.

Primitive Type	Wrapper Class
byte	Byte
short	Short
int	Integer
long	Long
float	Float
double	Double
boolean	Boolean
char	Character

Figure 7: The wrapper classes

Autoboxing allows us to use primitive values in place of wrapper objects, and unboxing allows vice versa.

#### 8.4 Scanner

The *Scanner class* allows text input to the program, with ability to parse primitive types, and use regular expressions to parse strings. Scanners can scan from user input to a console, from files, and from String objects.

This class comes in a package called util. We use the import keyword to import packages.

```
import java.util.Scanner;
```

or we can use an asterisk to import the whole package

```
import java.util.*;
```

Listing 7: Sample use of the Scanner class

```
Scanner sc = new Scanner(System.in);
int i = sc.nextInt();
long j = sc.nextLong();
String k = sc.next();
System.out.println(i + ", " + j + ", " + k);
sc.close();
```

In listing 7, 3 user inputs are taken from the console, and they are outputted separated by commas.

# Index

= operator, 20 == operator, 21 access modifier, 15, 19 alias, 21 Aliasing, 21	import keyword, 23 initialised, 18 instantiation, 18 Integrated Devevlopment Environment, 6
Application software, 3 assigning, 7 assignment operator, 20 Autoboxing, 23 break statement, 13 class, 17	Java, 3 Java byte code, 4 Java Development Kit, 5 Java Runtime Environment, 5 Java virtual machine, 4 JDK, 5 JRE, 5 JVM, 4
compiler, 3	3 V IVI, 4
console application, 6	Libraries, 5
constructor, 17 continue statement, 13	machine code, 3 method header, 14
data type, 7 declaring, 7 do-while loop, 13 dot operator, 18  Embedded software, 3 embedded systems, 3 encapsulation, 18 equality operator, 21	modifiers, 15, 19  new keyword, 21 non-access modifier, 15, 19  object, 17 object oriented, 4 Object oriented programming, 4 OOP, 4 overloaded, 8
for loop, 12	overloading, 16 packages, 5
graphical user interface, 6 GUI, 6	platform-independent, 4 polymorphism, 16
IDE, 6 if statement, 11 if-else statement, 11 immutable, 22 Implementation, 10	primitive types, 7 program, 3 program code, 3 Program design, 10 programming language, 3 Pseudocode, 10

reference equality, 21	text based user interface, 6 this keyword, 18
scalar types, 7 Scanner class, 23 Software, 3 source code, 3	UI, 6 unboxing, 23 user interface, 6
string, 22 string literal, 22 switch statement, 11	value equality, 21 variable, 7
syntax, 3 System software, 3	while loop, 12 wrapper class, 23