Java – Coding Guidelines

**Revision History**

|  |  |  |  |
| --- | --- | --- | --- |
| Rev. | Changes Made | Author | Date |
| 0.1 draft | Initial draft | Mihir Shelat | 28th June 2015 |
| 0.2 draft | Have added sections on security, performance, logging practices and code analyzers | Mihir Shelat | 3rd July 2015 |
|  |  |  |  |
|  |  |  |  |

Table of Contents

[2 Introduction 4](#_Toc423702978)

[2.1 Purpose i.e. why do we need coding guidelines? 4](#_Toc423702979)

[2.2 Objective & Scope 5](#_Toc423702980)

[2.3 References 5](#_Toc423702981)

[3 Coding Standards 5](#_Toc423702982)

[3.1 Clean code 6](#_Toc423702983)

[3.1.1 Names – And opportunity to reveal intent 6](#_Toc423702984)

[6](#_Toc423702985)

[3.1.2 Meaningful Names 6](#_Toc423702986)

[3.1.3 Avoid Noise 7](#_Toc423702987)

[3.1.4 Code Structuring 8](#_Toc423702988)

[3.1.5 Some good practices 10](#_Toc423702989)

[4 Design Principles 13](#_Toc423702990)

[4.1 GRASP - General Responsibility Assignment Software Patterns 13](#_Toc423702991)

[4.1.1 Information Expert 14](#_Toc423702992)

[4.1.2 Creator 14](#_Toc423702993)

[4.1.3 Low Coupling 15](#_Toc423702994)

[4.1.4 High Cohesion 16](#_Toc423702995)

[4.2 SOLID Principles 18](#_Toc423702996)

[4.2.1 Single responsibility principle 18](#_Toc423702997)

[4.2.2 Open closed principle 19](#_Toc423702998)

[20](#_Toc423702999)

[4.2.3 Liskov Substitution Principle 20](#_Toc423703000)

[4.2.4 Interface Segregation Principle 22](#_Toc423703001)

[4.2.5 Dependency Injection or Inversion Principle 23](#_Toc423703002)

[4.3 Other Principles 25](#_Toc423703003)

[4.3.1 Programming for Interface not implementation 25](#_Toc423703004)

[4.3.2 Favor Composition over Inheritance 25](#_Toc423703005)

[4.3.3 Encapsulate what varies 25](#_Toc423703006)

[5 Best practices 26](#_Toc423703007)

[5.1 Java coding 26](#_Toc423703008)

[5.1.1 Avoid creating unnecessary objects and always prefer to do Lazy Initialization 26](#_Toc423703009)

[5.1.2 Never make an instance fields of class public 26](#_Toc423703010)

[5.1.3 Always try to minimize Mutability of a class 27](#_Toc423703011)

[5.1.4 Try to prefer Interfaces instead of Abstract classes 27](#_Toc423703012)

[5.1.5 Always try to limit the scope of Local variable 27](#_Toc423703013)

[5.1.6 Wherever possible try to use Primitive types instead of Wrapper classes 28](#_Toc423703014)

[5.1.7 Use Strings with utmost care 28](#_Toc423703015)

[5.1.8 Defensive copies are savior 28](#_Toc423703016)

[5.1.9 Never let exception come out of finally block & Never throw “Exception” 29](#_Toc423703017)

[5.1.10 Exception handling – Best practices 29](#_Toc423703018)

[5.1.11 Other good practices 30](#_Toc423703019)

[ Avoid using implementation types (i.e., HashSet). Instead use an interface (i.e. Set). 30](#_Toc423703020)

[ Avoid Reassigning Parameters 30](#_Toc423703021)

[ Final Field Could Be Static 30](#_Toc423703022)

[ Position Literal First In Comparison 30](#_Toc423703023)

[ Close resources 31](#_Toc423703024)

[6 Analyzers 32](#_Toc423703025)

[6.1 Checkstyle 32](#_Toc423703026)

[6.2 PMD 33](#_Toc423703027)

[6.3 Findbugs 33](#_Toc423703028)

[7 Effective Logging practices 34](#_Toc423703029)

[7.1 Use the appropriate tools for the job 34](#_Toc423703030)

[7.2 Logging Levels 35](#_Toc423703031)

[7.3 Know what you are logging 35](#_Toc423703032)

[7.4 Be concise and descriptive 36](#_Toc423703033)

[7.5 Conversion pattern 36](#_Toc423703034)

[7.6 Method arguments & return value 37](#_Toc423703035)

[7.7 Integrating with external systems 37](#_Toc423703036)

[7.8 Log exceptions properly 37](#_Toc423703037)

[7.9 Don’t log secure information 37](#_Toc423703038)

[8 Java Performance 38](#_Toc423703039)

[8.1 Avoid producing a lot of garbage 38](#_Toc423703040)

[8.2 Recycle and/ or cache objects 38](#_Toc423703041)

[8.3 Variables 38](#_Toc423703042)

[8.4 Avoid explicitly calling the Garbage Collector 39](#_Toc423703043)

[8.5 Avoid synchronization 39](#_Toc423703044)

[8.6 Use StringBuffer /StringBuilder 39](#_Toc423703045)

[8.7 Explicitly close resources 39](#_Toc423703046)

[8.8 Limit number of threads 40](#_Toc423703047)

[8.9 Avoid excessive writing to the Java console 40](#_Toc423703048)

[8.10 Data Types 40](#_Toc423703049)

[8.11 When possible, declare methods as final 41](#_Toc423703050)

[9 Java Security – Best practices 42](#_Toc423703051)

[9.1 Sanitize inputs 42](#_Toc423703052)

[9.2 Limit access 42](#_Toc423703053)

[9.3 Make everything final 42](#_Toc423703054)

[9.4 Secure your data 42](#_Toc423703055)

[9.5 Don't always trust third-party code 42](#_Toc423703056)

[9.6 Make your classes non-serializable 43](#_Toc423703057)

# Introduction

## Purpose i.e. why do we need coding guidelines?

We have all seen situations where a new project starts whose objective is to develop from scratch an application in a leading-edge technology. Everything goes very fast; first, second, third release and then all of a sudden, the team's productivity & velocity starts to decrease. Fourth release is postponed for the third time, fixing something breaks something else.

Investing some efforts delivering good code quality code helps maintaining the code easily. Badly written code is difficult to maintain and bug-fixing, enhancement becomes error-prone and time consuming. Thus managing the source code quality is all about optimizing ROI – Return on Investment

As per Itay Maman - *A well-written program is a program where the cost of implementing a feature is constant throughout the program's lifetime*

In other words, a badly written program is a program where the cost of implementing a feature grows throughout time. Everyone must take care that code is not getting worse in subsequent releases. Some attention should be given to coverage of the code, following best practices and adhering to well-known standards.

Code which is not maintained rots & results in:

* Rigidity - hard to change because every change affects too many other parts of the system. A small changes causes the entire system to rebuild.
* Fragility - When you make a change, unexpected parts of the system break. In other words, changes to one module causes other unrelated modules to misbehave. Imagine a car system in which changing the radio station affects windows.
* Immobility - hard to reuse in another application because it cannot be disengaged from the current application caused by couplings and dependencies between different modules.

Whereas clean code is simple and direct. Clean code makes designers intent visible, which is full of abstractions and straight forward lines of controls.

## Objective & Scope

The objective is to define some guidelines & best practices to develop a clean code & maintain code quality throughout the lifecycle of project i.e. right from the beginning.

This document covers the following aspects of Java coding:

* Coding Standards
* Coding Principles
* Best practices
* Checklist
* Code Analyzers
* Logging practices

## References

* <https://en.wikipedia.org/wiki/GRASP_(object-oriented_design)>
* <https://en.wikipedia.org/wiki/SOLID_(object-oriented_design)>
* <https://en.wikipedia.org/wiki/Code_smell>

# Coding Standards

Coding conventions are a set of guidelines for a specific programming language that recommend programming style, practices and methods for each aspect of a piece program written in this language. These conventions usually cover file organization, indentation, comments, declarations, statements, white space, naming conventions etc. These are guidelines for software structural quality & hence programmers are highly recommended to follow these guidelines to help improve the readability of their source code and make software maintenance easier. Reducing the cost of software maintenance is the most often cited reason for following coding conventions.

Both Oracle (i.e. Sun) & Google have come up with Java coding standards which are very similar & recommend similar coding styles.

Here is the Sun specific coding convention:

<http://www.oracle.com/technetwork/java/codeconventions-150003.pdf>

But considering Sun specific standards were last modified in year 2000, the following one which Google has suggested in 2014 is latest and relevant.

<https://google-styleguide.googlecode.com/svn/trunk/javaguide.html#s4.4-column-limit>

## Clean code

### Names – And opportunity to reveal intent

**public** List<**int**[]> getThem() {

List<**int**[]> list1 = **new** ArrayList<**int**[]>();

**for** (**int**[] x : theList)

**if** (x[0] == 4)

list1.add(x);

**return** list1;

}

Tell me what does it do?

### 

**public** List<Cell> getFlaggedCells() {

List<Cell> flaggedCells = **new** ArrayList<Cell>();

**for** (Cell cell : gameBoard)

**if** (cell.isFlagged())

flaggedCells.add(cell);

**return** flaggedCells;

}

And now?

What is the difference?

You express functionality which anyone understands in given context. Nothing should be left to imagination.

**public** **static** **void** copyChars(**char** a1[], **char** a2[]) {

**for** (**int** i = 0; i < a1.length; i++) {

a2[i] = a1[i];

}

}

**public** **static** **void** copyCharacters(**char** source[], **char** destination[]) {

**for** (**int** currentIndex = 0; currentIndex < source.length; currentIndex++) {

destination[currentIndex] = source[currentIndex];

}

}

Avoid mental mapping…

Don’t put reader to comprehend in their head what a1, a2 and i signify.

### Meaningful Names

This is incomplete information, Demands more...

* **getCustomer()**
* **retrieveCustomer()**
* **fetchCustomer()**

Which one to use?

Chose one & stick to it

**Int days;**

**Int daysSinceCreation;**

**Int daysSinceLastModification ;**

**Int durationInDays;**

Add context to make the name meaningful…

### Avoid Noise

The type is list so the name need not say it again ….

**List<Client> clientList;**

**List<Client> clients;**

**List<Client> regularClients;**

**List<Client> newClients;**

### Code Structuring

Make it read like a story…. Like a thriller movie, it should start with abstractions which become clearer as story unfolds until it gives you the Aha moment in the climax… Basically you need to refactor your code when you smell deeper problem such as:

* Duplicate code
* Long methods
* Big classes
* Big switch statements
* Long navigations (e.g. a.b().c().d())
* Too much checking for null objects
* Un-encapsulated fields / expressions.

#### Avoid Mixing Abstractions….

**public** **void** generateAggregateReportFor(

**final** List<StoryTestResults> storyResults,

**final** List<FeatureResults> featureResults) **throws** IOException {

LOGGER.info("Generating summary report for user stories to" + getOutputDirectory());

copyResourcesToOutputDirectory();

// create story report

Map<String, Object> storyContext = **new** HashMap<String, Object>();

storyContext.put("stories", storyResults);

storyContext.put("storyContext", "All stories");

addFormattersToContext(storyContext);

writeReportToOutputDirectory("stories.html",

mergeTemplate(STORIES\_TEMPLATE\_PATH).usingContext(storyContext));

// create feature report

Map<String, Object> featureContext = **new** HashMap<String, Object>();

addFormattersToContext(featureContext);

featureContext.put("features", featureResults);

writeReportToOutputDirectory(

"features.html",

mergeTemplate(FEATURES\_TEMPLATE\_PATH).usingContext(featureContext));

// generate home page for story and feature

generateReportHomePage(storyResults, featureResults);

// update history for feature

getTestHistory().updateData(featureResults);

generateHistoryReport();

}

Mixing abstractions …hard to understand… too many things going on

**private** **void** generateAggregateReportFor(

**final** List<StoryTestResults> storyResults,

**final** List<FeatureResults> featureResults) **throws** IOException {

copyResourcesToOutputDirectory();

generateStoriesReportFor(storyResults);

generateFeatureReportFor(featureResults);

generateReportHomePage(storyResults, featureResults);

updateHistoryFor(featureResults);

generateHistoryReport();

}

* Use extract method and move method techniques to create consistent abstractions…
* If you require more than few minutes to understand the intent of the method/class then, it probably needs improvement…

#### Encapsulate Complex Expressions…

**for** (TestStep currentStep : testSteps) {

**if** (!currentStep.isAGroup() && currentStep.getScreenshots() != **null**) {

**for** (RecordedScreenshot screenshot : currentStep.getScreenshots()) {

screenshots.add(**new** Screenshot(screenshot.getScreenshot().getName(),

currentStep.getDescription(), widthOf(screenshot.getScreenshot()),currentStep.getException()));

}

}

}

Explain yourself in code…

**if (employee.isEligibleForFullBenefits())**

**// Check to see if the employee is eligible for fullbenefits**

**if ((employee.flags & HOURLY\_FLAG) &&(employee.age > 65))**

When there is more complex conditional expressions (e.g. && / || etc) then abstract it for better readability…

**public boolean needsScreenshots() {**

**return (!isAGroup() && getScreenshotsAndHtmlSources() != null)**

**}**

**for (TestStep currentStep : testSteps) {**

**if (currentStep.needsScreenshots()) {**

What does this expression mean?

### Some good practices

#### Tell, Don’t Ask

* A method m of an object O may only invoke the methods of the following kinds of objects:
  + itself
  + m's parameters
  + Any objects created/instantiated within m
  + O's direct component objects
  + A global variable, accessible by O, in the scope of m
* Only talk to your immediate friends, don’t talk to strangers…
* In object oriented world, law can be stated simply as "use only one dot". That is, the code a.b.Method() breaks the law where a.Method() does not.

**walkDog(Dog dog){**

**dog.getLegs().move();**

**}**

**walkDog(Dog dog){**

**dog.walk();**

**}**

#### Don’t return or pass NULL

Null is a Billion dollar mistake – Avoid at all costs

Null was invented by Charles Antony Richard Hoare and he calls it his “billion dollar mistake”, it’s because failure to check have caused Null Pointers and in turn have causes huge losses commercially - <http://en.wikipedia.org/wiki/Tony_Hoare>

Before Java 8, people used to implement Null Object Pattern:

<https://en.wikipedia.org/wiki/Null_Object_pattern>

<http://www.tutorialspoint.com/design_pattern/null_object_pattern.htm>

But now Java 8 has alternative idiom Optional<?> to handle situations of failure/null returns.

<http://examples.javacodegeeks.com/core-java/util/optional/java-8-optional-example/>

//return Empty List if not found

**public** List<Employee> getEmployeesOnLeave(Date leaveDate) {

List<Employee> employeeOnLeave = selectEmployeeOnLeave(leaveDate);

If(employeeOnLeave==**null**) Collections.emptyList();

}

//return Null if not found

**public** List<Employee> getEmployeesOnLeave(Date leaveDate) {

**return** selectEmployeeOnLeave(leaveDate);

}

#### It’s good to be DRY (Don’t Repeat Yourself)

As name suggest DRY (don’t repeat yourself) means don’t write duplicate code, instead use abstraction to abstract common things in one place. If you use a hardcoded value more than one time consider making it public final constant, if you have block of code in more than two place consider making it a separate method.

**int** countItems(List<Item> items) {

**if**(items==**null** || items.isEmpty())

**return** -1;

**else** **if**(items!=**null** && !items.isEmpty())

....

}

Item selectItem(List<Item> items) {

**if**(items==**null** || items.isEmpty())

**return** Item.Blank;

**else** **if**(items!=**null** && !items.isEmpty())

....

}

**private** <T> isNotNullOrEmpty( **final** List<T> listUnderTest) {

**return** listUnderTest!=**null** && !listUnderTest.isEmpty();

}

**int** countItems(List<Item> items) {

**if**(isNotNullOrEmpty(items))

....

}

If you used common code to validate OrderID and SSN it doesn’t mean they are same or they will remain same in future. By using common code for two different functionality or thing you closely couple them forever and when your OrderID changes its format, your SSN validation code will break. So be aware of such coupling and just don’t combine anything which uses similar code but are not related.

#### Avoid Deep Nesting

**if** (cond1) {

Logger.debug("cond1 is true");

**if** (cond2) {

Logger.debug("cond1 and cond2 are both true");

**if** (cond3) {

Logger.debug("cond1, cond2 and cond3 are all true");

}

}

}

**if** (cond1) {

Logger.debug("Executing when cond1 is true");

}

**if** (cond1&& cond2) {

Logger.debug("Executing when cond1 and cond2 are both true");

}

**if** (cond1 && cond2 && cond3) {

Logger.debug("Executing when all cond1, cond2 & cond3 are true");

}

#### Move Method

**public** **class** Student {

**public** **boolean** isEnrolledFor(Course course) {

course.getEnrolledStudents().contains(**this**);

}

}

**public** **class** Course {

**private** List<Student> enrolledStudents;

**public** List<Student> getEnrolledStudents() {

**return** **this**.students();

}

}

**public** **class** Student{…}

**public** **class** Course{

**private** List<Student> enrolledStudents;

**public** List<Student> getEnrolledStudents(){

**return** **this**.enrolledStudents();

}

**public** **boolean** isEnrolledFor(Student student){

enrolledStudents.contains(student);

}

}

If a method on one class uses (or is used by) another class more than the class on which It’s defined, move it to the other class…

# Design Principles

While designing the application, some of the most obvious object oriented design questions we come across are:

* Which class has what responsibility?
* What will be the associations?
* Which object will create/contain which object?
* How to have minimum dependencies?
* I need to maximize cohesion and reduce coupling
* I need to increase the reuse etc. …

Many of us knows various design principles but have not put enough attention in following them while designing a solution. One should always strive for highly cohesive and loosely couple solution, code or design and looking open source code from Apache and Sun are good examples of Java design principles or how design principles should be used in Java coding. The following two are widely used principles to create a system that is easy to maintain and extend over time:

* **GRASP** i.e. General Responsibility Assignment Software Patterns (or Principles), consists of guidelines for assigning responsibility to classes and objects in object-oriented design.
* **SOLID** i.e. (Single responsibility, Open-closed, Liskov substitution, Interface segregation and Dependency inversion)

While these patterns/principles answer some software problem, and in almost every case these problems are common to almost every software development project. Please note:

* These are principles, not rules, laws or truth which must be followed under all conditions – not a panacea to all problems.
* These are advices, just like age old “An apple a day keeps doctor away” makes you healthy but does not actually keep the doctor away.
* So, at times, in pragmatic situation it may make perfect sense to overrule them but, you will know when you have and why …

## GRASP - General Responsibility Assignment Software Patterns

The different patterns and principles used in GRASP are: Controller, Creator, Indirection, Information Expert, High Cohesion, Low Coupling, Polymorphism, Protected Variations, and Pure Fabrication. All these patterns answer some software problem, and in almost every case these problems are common to almost every software development project.

* Here responsibility can be:
* Accomplished by a single object
* A group of object collaboratively accomplish a responsibility.
* GRASP helps us in deciding which responsibility should be assigned to which object/class.
* Identify the objects and responsibilities from the problem domain, and also identify how objects interact with each other.
* Define blue print for those objects – i.e. class with methods implementing those responsibilities.

### Information Expert

* **Problem:** How to assign responsibility a class?
* **Solution:** Assign a responsibility to the class that has the information necessary to fulfil the responsibility. The class which has all the information needed to perform operations, or in some cases they collaborate with others to fulfill their responsibilities. These responsibilities include methods, computed fields, and so on.

Cards

Game

Deck

Players

Has

Played by

Contains

**Question:** Which is class should have the logic of shuffling cards?

* Assume we need to get all the videos of a VideoStore

|  |
| --- |
| VideoStore |
| getAllVideos() |

|  |
| --- |
| Video |
| title |

* Since VideoStore knows about all the videos, we can assign this responsibility of giving all the videos can be assigned to VideoStore class.
* VideoStore is the information expert.

### Creator

* **Problem:** Who creates an Object? Or who should create a new instance of some class?
* **Solution:** Determine which class should create instances of a class based on the relationship between potential creator classes and the class to be instantiated

Assign class B responsibility of creating instance of class A if:

* + B aggregates A objects
  + B contains A objects
  + B records instances of A objects
  + B closely uses A objects
  + B has the initializing data for creating A objects

Where there is a choice, prefer B aggregates or contains A objects

Cards

Game

Deck

Players

Has

Played by

Contains

Deck contains Cards so, Deck should be creator of Card objects.

|  |
| --- |
| VideoStore |
| getAllVideos() |

|  |
| --- |
| Video |
| title |

* + Consider VideoStore and Video in that store.
  + VideoStore has an aggregation association with Video. I.e, VideoStore is the container and the Video is the contained object.
  + So, we can instantiate video object in VideoStore class

### Low Coupling

Coupling is a measure of how strongly one element is connected to, has knowledge of, or relies on other elements.

**Problem:** How to support low dependency, low change impact, increased reuse?

**Solution:** Assign responsibilities such that:

* lower dependency between the classes
* change in one class having lower impact on other classes
* higher reuse potential



Here class Rent knows about both VideoStore and Video objects. Rent is depending on both the classes.



VideoStore and Video class are coupled, and Rent is coupled with VideoStore. Thus providing low coupling.

### High Cohesion

High Cohesion is an evaluative pattern that attempts to keep objects appropriately focused, manageable and understandable. High cohesion is generally used in support of Low Coupling. High cohesion means that the responsibilities of a given element are strongly related and highly focused. Alternatively, low cohesion is a situation in which a given element has too many unrelated responsibilities. Elements with low cohesion often suffer from being hard to comprehend, hard to reuse, hard to maintain and averse to change

**Problem:** How to keep classes focused and manageable?

**Solution:** Assign responsibility so that cohesion remains high. Benefits of it are:

* Easily understandable and maintainable.
* Code reuse
* Low coupling

Here is the example of LOW Cohesion where all Student + database specific responsibilities are in the same class:



For High Cohesion, move database specific responsibilities to a DB class as show below.



## SOLID Principles

It's an acronym introduced by Michael Feathers for the "first five principles" of object-oriented programming and design named by Robert Martin. The principles (Single responsibility, Open-closed, Liskov substitution, Interface segregation and Dependency inversion), when applied together, intend to make it more likely that a programmer will create a system that is easy to maintain and extend over time.

### Single responsibility principle

The Single Responsibility Principle (SRP) states that there should never be more than one reason for a class to change. **One class should have one and only one responsibility**. This means that every class, or similar structure, in your code should have only one job to do.

Everything in the class should be related to that single purpose, i.e. be cohesive. It does not mean that your classes should only contain one method or property.

There can be a lot of members as long as they relate to the single responsibility. It may be that when the one reason to change occurs, multiple members of the class may need modification. It may also be that multiple classes will require updates.

**class** Employee {

**public** Money calculatePay() {...}

**public** **void** save() {...}

**public** String reportHours() {...}

}

Here we have pay 1) calculation logic with 2) database logic and 3) reporting logic all mixed up within one class. If you have multiple responsibilities combined into one class, it might be difficult to change one part without breaking others. Mixing responsibilities also makes the class harder to understand and harder to test, decreasing cohesion. The easiest way to fix this is to split the class into three different classes, with each having only one responsibility: database access, calculating pay and reporting, all separated.

So basically gather together the things that change for the same reasons. Separate those things that change for different reasons.

### Open closed principle

The Open-Closed Principle (OCP) states that classes (and even methods) should be open for extension but closed for modification. “Open to extension” means that you should design your classes so that new functionality can be added as new requirements are generated. “Closed for modification” means that once you have developed a class you should never modify it, except to correct bugs.

These two parts of the principle appear to be contradictory. However, if you correctly structure your classes and their dependencies, you can add functionality without editing existing source code.

Generally you achieve this by referring to abstractions for dependencies, such as interfaces or abstract classes, rather than using concrete classes. Functionality can be added by creating new classes that implement the interfaces.

Applying OCP to your projects limits the need to change source code once it has been written, tested and debugged. This reduces the risk of introducing new bugs to existing code, leading to more robust software.

Another side effect of the use of interfaces for dependencies is reduced coupling and increased flexibility.

**class** GraphicEditor {

**public** **void** drawShape(Shape s) {

**if** (s.m\_type == 1)

drawRectangle(s);

**else** **if** (s.m\_type == 2)

drawCircle(s);

}

**public** **void** drawCircle(Circle r) {....}

**public** **void** drawRectangle(Rectangle r) {....}

}

**class** Shape {

**int** m\_type;

}

**class** Rectangle **extends** Shape {

Rectangle() {

**super**.m\_type = 1;

}

}

**class** Circle **extends** Shape {

Circle() {

**super**.m\_type = 2;

}

}

What happens when you add a new shape?

* Impossible to add a new Shape without modifying GraphEditor
* Important to understand GraphEditor to add a new Shape
* Tight coupling between GraphEditor and Shape
* Difficult to test a specific Shape without involving GraphEditor

**if** (cond1) {

System.***out***.println("Executing when cond1 is true");

}

**if** (cond1&& cond2) {

System.***out***.println("Executing when cond1 and cond2 are both true");

}

**if** (cond1 && cond2 && cond3) {

System.***out***.println("Executing when all cond1, cond2 & cond3 are true");

}

### 

**class** GraphicEditor {

**public** **void** drawShape(Shape s) {

s.draw();

}

}

**class** Shape {

**abstract** **void** draw();

}

**class** Rectangle **extends** Shape {

**public** **void** draw() {

// draw the rectangle

}

}

A good example of this principle is a framework like struts or spring, you will see that you cannot change their core logic and request processing, But you can modify the desired application flow just by extending some classes and plugin them in configuration files.

### Liskov Substitution Principle

All the time we design a program module and we create some class hierarchies. Then we extend some classes creating some derived classes.

We must make sure that the new derived classes just extend without replacing the functionality of old classes. Otherwise the new classes can produce undesired effects when they are used in existing program modules.

Likov's Substitution Principle states that if a program module is using a Base class, then the reference to the Base class can be replaced with a Derived class without affecting the functionality of the program module.

All subclasses must, therefore, operate in the same manner as their base classes. The specific functionality of the subclass may be different but must conform to the expected behavior of the base class. To be a true behavioral subtype, the subclass must not only implement the base class’s methods and properties, but also conform to its implied behavior.

**public** **class** Product {

**public** string Name { get; set; }

**public** string Author { get; set; }

}

**public** **class** Book : Product {}

**public** **class** Movie : Product {}

If someone had a Product object (which was actually a Movie) and asked for the Author, what should it do (a Movie doesn’t have an Author)? – Must have seen UnsupportedOperationException, it’s a LSP violation. This principle is just an extension of the Open Close Principle and it means that we must make sure that new derived classes are extending the base classes without changing their behavior.

Here is one classic example violating Liskov’s Substitution principle. In the example 2 classes are used: Rectangle and Square. Let's assume that the Rectangle object is used somewhere in the application. We extend the application and add the Square class. The square class is returned by a factory pattern, based on some conditions and we don't know the exact what type of object will be returned. But we know it's a Rectangle. We get the rectangle object, set the width to 5 and height to 10 and get the area. For a rectangle with width 5 and height 10 the area should be 50. Instead the result will be 100

// Violation of Likov's Substitution Principle

**class** Rectangle {

**protected** **int** m\_width;

**protected** **int** m\_height;

**public** **void** setWidth(**int** width) {

m\_width = width;

}

**public** **void** setHeight(**int** height) {

m\_height = height;

}

**public** **int** getArea() {

**return** m\_width \* m\_height;

}

}

**class** Square **extends** Rectangle {

**public** **void** setWidth(**int** width) {

m\_width = width;

m\_height = width;

}

**public** **void** setHeight(**int** height) {

m\_width = height;

m\_height = height;

}

}

**class** LspTest {

**private** **static** Rectangle getNewRectangle() {

// it can be an object returned by some factory…

**return** **new** Square();

}

**public** **static** **void** main(String args[]) {

Rectangle r = LspTest.*getNewRectangle*();

r.setWidth(5);

r.setHeight(10);

// user knows that r it's a rectangle.

// It assumes that he's able to set the width and height as for the base class

Logger.debug(r.getArea());

// now he's surprised to see that the area is 100 instead of 50.

}

}

### Interface Segregation Principle

It is applicable to interfaces as single responsibility principle holds to classes. It says: "Clients should not be forced to implement unnecessary methods which they will not use". Instead of one fat interface many small interfaces are preferred based on groups of methods, each one serving one submodule.

When we have non-cohesive interfaces, the ISP guides us to create multiple, smaller, cohesive interfaces. Smaller interfaces are easier to implement, improving flexibility and the possibility of reuse. As fewer classes share these interfaces, the number of changes that are required in response to an interface modification is lowered, which increases robustness.

interface Worker {

void work();

void eat();

}

ManWorker implements Worker {

void work() {…};

void eat() {30 min break;};

}

RobotWorker implements Worker {

void work() {…};

void eat() {

//Not Applicable for a RobotWorker

};

}

Bad example (polluted interface)

**interface** Workable {

**public** **void** work();

}

**interface** Feedable{

**public** **void** eat();

}

ManWorker **implements** Workable, Feedable {

**void** work() {...};

**void** eat() {...};

}

RobotWorker **implements** Workable {

**void** work() {...};

}

Solution: Split into two interfaces

Wherever you find that client of the interface doesn’t use/need all methods then create a specific interface and, the generic interface should extend the specific interface for backwards compatibility (if required).

### Dependency Injection or Inversion Principle

Don't ask for dependency it will be provided to you by framework. This has been very well implemented in Spring framework, beauty of this design principle is that any class which is injected by DI framework is easy to test with mock object and easier to maintain because object creation code is centralized in framework and client code is not littered with that.

The Dependency Inversion Principle (DIP) states that high-level modules should not depend upon low-level modules; they should depend on abstractions. Secondly, abstractions should not depend upon details; details should depend upon abstractions. The idea is that we isolate our class behind a boundary formed by the abstractions it depends on. If all the details behind those abstractions change, then our class is still safe. This helps keep coupling low and makes our design easier to change. DIP also allows us to test things in isolation, details like database are plugins to our system.

Below is an example which violates the Dependency Inversion Principle. We have the manager class which is a high level class, and the low level class called Worker. We need to add a new module to our application to model the changes in the company structure determined by the employment of new specialized workers. We created a new class SuperWorker for this.

Let's assume the Manager class is quite complex, containing very complex logic. And now we have to change it in order to introduce the new SuperWorker. Let's see the disadvantages:

* We have to change the Manager class (remember it is a complex one and this will involve time and effort to make the changes).
* Some of the current functionality from the manager class might be affected.
* Unit testing should be redone.

**class** Worker {

**public** **void** work() {

// ....working

}

}

**class** Manager {

Worker worker;

**public** **void** setWorker(Worker w) {

worker = w;

}

**public** **void** manage() {

worker.work();

}

}

**class** SuperWorker {

**public** **void** work() {

// .... working much more

}

}

Bad example of DI principle

All those problems could take a lot of time to be solved and they might induce new errors in the old functionality. The situation would be different if the application had been designed following the Dependency Inversion Principle. It means we design the manager class, an IWorker interface and the Worker class implementing the IWorker interface. When we need to add the SuperWorker class all we have to do is implement the IWorker interface for it. No additional changes in the existing classes.

**interface** IWorker {

**public** **void** work();

}

**class** Worker **implements** IWorker{

**public** **void** work() {

// ....working

}

}

**class** SuperWorker **implements** IWorker{

**public** **void** work() {

//.... working much more

}

}

**class** Manager {

IWorker worker;

**public** **void** setWorker(IWorker w) {

worker = w;

}

**public** **void** manage() {

worker.work();

}

}

As in this new design a new abstraction layer is added through the IWorker Interface, the problems from the bad code are solved:

* Manager class doesn't require changes when adding SuperWorkers.
* Minimized risk to affect old functionality present in Manager Class since we don't change it.
* No need to redo the unit testing for Manager Class.

|  |  |
| --- | --- |
| Principle | Tips |
| Single Responsibility Principle | A class should have one, and only one, reason to change. |
| Open Closed Principle | You should be able to extend a classes behavior, without modifying it |
| Liskov Substitution Principle | Derived classes must be substitutable for their base classes |
| Interface Segregation Principle | Make fine grained interfaces that are client specific |
| Dependency Inversion Principle | Depend on abstractions, not on concretions |

## Other Principles

### Programming for Interface not implementation

Always program for interface and not for implementation this will lead to flexible code which can work with any new implementation of interface. So use interface type on variables, return types of method or argument type of methods in Java.

### Favor Composition over Inheritance

Always favor composition over inheritance if possible. Some of you may argue this but I found that Composition is lot more flexible than Inheritance. Composition allows to change behavior of a class at runtime by setting property during runtime and by using Interfaces to compose a class we use polymorphism which provides flexibility of to replace with better implementation any time

### Encapsulate what varies

Only one thing is constant in software field and that is “Change”, so encapsulate the code you expect or suspect to be changed in future. Benefit of this OOPS Design principle is that It’s easy to test and maintain proper encapsulated code. If you are coding in Java then follow principle of making variable and methods private by default and increasing access step by step e.g. from private to protected and not public. Several of design pattern in Java uses Encapsulation, Factory design pattern is one example of Encapsulation which encapsulate object creation code and provides flexibility to introduce new product later with no impact on existing code.

# Best practices

## Java coding

### Avoid creating unnecessary objects and always prefer to do Lazy Initialization

Object creation in Java is one of the most expensive operation in terms of memory utilization and performance impact. It is thus advisable to create or initialize an object only when it is required in the code

**public class Countries {**

**private List countries;**

**public List getCountries() {**

**//initialize only when required**

**if(null == countries) {**

**countries = new ArrayList();**

**}**

**return countries;**

**}**

**}**

### Never make an instance fields of class public

Making a class field public can cause lot of issues in a program. For instance you may have a class called MyCalender. This class contains an array of String weekdays. You may have assume that this array will always contain 7 names of weekdays. But as this array is public, it may be accessed by anyone. Someone by mistake also may change the value and insert a bug!

Best approach as many of you already know is to always make the field private and add a getter method to access the elements.

**private String[] weekdays = {"Sun", "Mon", "Tue", "Thu", "Fri", "Sat", "Sun"};**

**public String[] getWeekdays() {**

**return weekdays;**

**}**

But writing getter method does not exactly solve our problem. The array is still accessible. Best way to make it unmodifiable is to return a clone of array instead of array itself. Thus the getter method will be changed to.

**public String[] getWeekdays() {**

**return weekdays.clone();**

**}**

### Always try to minimize Mutability of a class

Making a class immutable is to make it unmodifiable. The information the class preserve will stay as it is throughout the lifetime of the class. Immutable classes are simple, they are easy to manage. They are thread safe. They makes great building blocks for other objects.

However creating immutable objects can hit performance of an app. So always choose wisely if you want your class to be immutable or not. Always try to make a small class with less fields immutable.

To make a class immutable you can define its all constructors private and then create a public static method to initialize and object and return it.

**public class Employee {**

**private String firstName;**

**private String lastName;**

**//private default constructor**

**private Employee(String firstName, String lastName) {**

**this.firstName = firstName;**

**this.lastName = lastName;**

**}**

**public static Employee valueOf (String firstName, String lastName) {**

**return new Employee(firstName, lastName);**

**}**

**}**

### Try to prefer Interfaces instead of Abstract classes

First you cannot inherit multiple classes in Java but you can definitely implements multiple interfaces. It’s very easy to change the implementation of an existing class and add implementation of one more interface rather than changing full hierarchy of class.

Again if you are 100% sure what methods an interface will have, then only start coding that interface. As it is very difficult to add a new method in an existing interface without breaking the code that has already implemented it. On contrary a new method can be easily added in Abstract class without breaking existing functionality.

### Always try to limit the scope of Local variable

Local variables are great. But sometimes we may insert some bugs due to copy paste of old code. Minimizing the scope of a local variable makes code more readable, less error prone and also improves the maintainability of the code.

Thus, declare a variable only when needed just before its use.

Always initialize a local variable upon its declaration. If not possible at least make the local instance assigned null value.

### Wherever possible try to use Primitive types instead of Wrapper classes

Wrapper classes are great. But at same time they are slow. Primitive types are just values, whereas Wrapper classes are stores information about complete class.

### Use Strings with utmost care

Always carefully use Strings in your code. A simple concatenation of strings can reduce performance of program. For example if we concatenate strings using + operator in for loop then every time + is used, it creates a new String object. This will affect both memory usage and performance time.

Also whenever you want to instantiate a String object, never use its constructor but always instantiate it directly. For example:

**//slow instantiation**

**String slow = new String("Yet another string object");**

**//fast instantiation**

**String fast = "Yet another string object";**

### Defensive copies are savior

Defensive copies are the clone objects created to avoid mutation of an object. For example in below code we have defined a Student class which has a private field birth date that is initialized when the object is constructed.

**public class Student {**

**private Date birthDate;**

**public Student(birthDate) {**

**this.birthDate = birthDate;**

**}**

**public Date getBirthDate() {**

**return this.birthDate;**

**}**

**}**

**//Now we may have some other code that uses the Student object**

**public static void main(String []arg) {**

**Date birthDate = new Date();**

**Student student = new Student(birthDate);**

**birthDate.setYear(2019);**

**logger.debug(student.getBirthDate());**

**}**

In above code we just created a Student object with some default birthdate. But then we changed the value of year of the birthdate. Thus when we print the birth date, its year was changed to 2019!

To avoid such cases, we can use Defensive copies mechanism. Change the constructor of Student class to following which ensure we have another copy of birthdate that we use in Student class

**public Student(birthDate) {**

**this.birthDate = new Date(birthDate);**

**}**

### Never let exception come out of finally block & Never throw “Exception”

Finally blocks should never have code that throws exception. Always make sure finally clause does not throw exception. If you have some code in finally block that does throw exception, then log the exception properly and never let it come out.

Also never throw java.lang.Exception directly. It defeats the purpose of using checked Exceptions. Also there is no useful information getting conveyed in caller method.

### Exception handling – Best practices

* Never swallow the exception in catch block
* Declare the specific checked exceptions that your method can throw
* Do not catch the Exception class rather catch specific sub classes
* Never catch Throwable class
* Always correctly wrap the exceptions in custom exceptions so that stack trace is not lost
* Either log the exception or throw it but never do the both
* Never throw any exception from finally block
* Always catch only those exceptions that you can actually handle
* Don't use printStackTrace() statement or similar methods
* Use finally blocks instead of catch blocks if you are not going to handle exception
* Remember "Throw early catch late" principle
* Always clean up after handling the exception
* Throw only relevant exception from a method
* Never use exceptions for flow control in your program
* Validate user input to catch adverse conditions very early in request processing
* Always include all information about an exception in single log message
* Pass all relevant information to exceptions to make them informative as much as possible
* Always terminate the thread which it is interrupted
* Use template methods for repeated try-catch
* Document all exceptions in your application in javadoc

### Other good practices

### Avoid using implementation types (i.e., HashSet). Instead use an interface (i.e. Set).

**public** **class** Bar {

**private** ArrayList list = **new** ArrayList(); // Use List instead

**public** HashSet getFoo() { // Use Set instead

**return** **new** HashSet();

}

}

**public** **class** Bar1 {

**private** List list = **new** ArrayList(); // Use List instead

**public** Set getFoo() { // Use Set instead

**return** **new** HashSet();

}

}

### Avoid Reassigning Parameters

Reassigning values to parameters is a questionable practice. Use a temporary local variable instead or use only when required.

**private void foo(String bar) {**

**bar = "something else";**

**}**

### Final Field Could Be Static

If a final field is assigned to a compile-time constant, it could be made static, thus saving overhead in each object at runtime.

**public static final int BAR = 42;**

### Position Literal First In Comparison

Position literals first in String comparisons - that way if the String is null you won't get a NullPointerException, it'll just return false.

**boolean bar(String x) {**

**return 2.equals("x");**

**}**

### Close resources

Ensure that resources (like Connection, Statement and Streams) are always closed after use.

**public** **void** foo() {

Connection conn = pool.getConnection();

**try** {

// do stuff

} **catch** (SQLException ex) {

// handle exception

} **finally** {

conn.close();

}

}

# Analyzers

For code review, you need to gather several programmers at regular times to review a fresh code or re-review a code after recommended changes have been applied to it. This is time consuming + incur cost.

Static code analyzers can be embedded in IDE like eclipse and can also run as a part of build tools like Ant, Maven. Static analyzers can be used to:

* Detecting errors in source code
* Recommendations on code formatting such as number of indents in various constructs, unused imports, use of spaces/tabs and so on

There are quite a few static analyzers for Java such as:

* Checkstyle – Besides some static code analysis, it can be used to show violations of a configured coding standard.
* FindBugs – An open-source static bytecode analyzer for Java.
* PMD – A static ruleset based Java source code analyzer that identifies potential problems.
* SonarQube – is an open source platform for Continuous Inspection of code quality.

## Checkstyle

Checkstyle is a static code analysis tool used in software development for checking if Java source code complies with coding rules. The performed checks mainly limit themselves to the presentation and don't analyze content, and do not confirm the correctness or completeness of the program. Checkstyle can examine:

* Javadoc comments for classes, attributes and methods
* Naming conventions of attributes and methods
* Limit of the number of function parameters, line lengths
* Presence of mandatory headers
* The use of packets imports, of classes, of scope modifiers and of instructions blocks
* The spaces between some characters
* The good practices of class construction
* Duplicated code sections

For further reading & installing checkstyle plugin in eclipse:

* <http://checkstyle.sourceforge.net/>
* <http://eclipse-cs.sf.net/update>

## PMD

PMD is a static rule-set based Java source code analyzer which finds common programming flaws like unused variables, empty catch blocks, unnecessary object creation, and so forth. It supports Java, JavaScript, PLSQL, Apache Velocity, XML, XSL. Additionally it includes CPD, the copy-paste-detector. CPD finds duplicated code in Java and other languages. It identifies potential problems like:

* Possible bugs—Empty try/catch/finally/switch blocks.
* Dead code—Unused local variables, parameters and private methods.
* Empty if/while statements.
* Overcomplicated expressions—Unnecessary if statements, for loops that could be while loops.
* Suboptimal code—Wasteful String/StringBuffer usage.
* Classes with high Cyclomatic Complexity measurements.
* Duplicate code—Copied/pasted code can mean copied/pasted bugs, and decreases maintainability.

For further reading & installing PMD plugin in eclipse:

* http://pmd.sourceforge.net/
* http://sourceforge.net/projects/pmd/files/pmd-eclipse/update-site

## Findbugs

Findbugs is an open source project for a static analysis of the Java bytecode to identify potential software bugs, such as null pointer references, infinite recursive loops, bad uses of the Java libraries, deadlocks, etc. Findbugs provides early feedback about potential errors in the code. This helps the developer to access these problems early in the development phase.

Findbugs categories bug patterns in:

* Correctness - the code is doing something wrong, you should look at it
  + Illegal format string
  + Null value is guaranteed to be dereferenced
* Bad practice - the code violates good practice
  + Dodgy Code
    - Useless control flow
    - Redundant null check of value known to be null
  + Concurrency
  + Performance
    - Method concatenates strings using + in a loop
    - Method invokes inefficient Boolean constructor
  + Miscellaneous
    - Confusing method names
    - Method may fail to close stream
    - Comparison of String parameter using == or !=
* Security defect
  + Finalizer should be protected, not public
  + Field isn't final and can't be protected from malicious code
  + Hardcoded constant database password
  + A prepared statement is generated from a variable String

For further reading & installing Findbugs plugin in eclipse:

* http://findbugs.sourceforge.net/
* https://findbugs.cs.umd.edu/eclipse
* http://findbugs.sourceforge.net/bugDescriptions.html

# Effective Logging practices

This is pretty basic java logging question and everybody argue that if we Java System.out.println() for printing messages then whey we use logging. Everybody who starts java starts with System.out.println() for printing message in java console. But this is not at all powerful as compared to advanced Java logging API like log4j and java.util.logging. Many of you already know this, but logging is just a way of monitoring your application, and when used right this can provide you with information of what is happening during you’re application’s runtime process.

You can log errors and actions your applications encounter, which is pretty critical unless you never write any bugs ☺. Logging other things can be very informative e.g. the start and end of a batch process, the changes made by administrators, amount of processed files, duration within a certain method call and so on. It’s kind of a double-edged sword. You want the right kinds of information available to you, but don’t want to store tons of old logs that don’t mean anything to you a year later. You want to try to write concise informative messages that are self-explanatory and able to be dumped soon after, as needed.

People have recommended few basic principles of proper application logging such as:

## Use the appropriate tools for the job

There are quite a few logging tools available in Java such as Log4j, SLF4j, Commons logging etc. In my opinion, **SLF4J** is the best logging API available, mostly because of a great pattern substitution support:

*log.debug("Found {} records matching filter: '{}'", records, filter);*

In **Log4j** you would have to use:

*log.debug("Found " + records + " records matching filter: '" + filter + "'");*

This is not only longer and less readable, but also inefficient because of extensive use of string concatenation. SLF4J adds a nice {} substitution feature. Also, because string concatenation is avoided and toString() is not called if the logging statement is filtered, there is no need for isDebugEnabled() anymore. BTW, have you noticed single quotes around filter string parameter?

SLF4J is just a façade. As an implementation I would recommend the Logback framework, already advertised, instead of the well-established Log4J. It has many interesting features and, in contrary to Log4J, is actively developed.

## Logging Levels

The primary purpose of a logging level is to help us filter useful information out of the noise. To avoid using the wrong level and thus reducing the usefulness of log messages, developers must be given clear guidelines before they start coding. Somehow programmers never pay attention to logging levels, simply logging everything on the same level, typically INFO or DEBUG. Why? Logging frameworks have two major benefits over System.out, i.e. categories and levels. Both allow you to selectively filter logging statements permanently or only for diagnostics time. The most common logging levels could be:

* **ERROR**: An error message indicates a serious problem in the system. The problem is usually non-recoverable and requires manual intervention i.e. something terribly wrong had happened, that must be investigated immediately. No system can tolerate items logged on this level. Example: NPE, database unavailable, mission critical use case cannot be continued.
* **WARN**: A warning message indicates a potential problem in the system. Here the process might be continued, but take extra caution. For example, if the message category is related to security, a warning message should be produced if a dictionary attack is detected.
* **INFO**: These messages contain some contextual information to help trace execution. Reading these messages, one should be able to understand INFO messages and quickly find out what the application is doing. Other definition of INFO message: each action that changes the state of the application significantly (database update, external system request
* **DEBUG**: Messages in this level contain extensive contextual information. They are mostly useful to track the flow or to diagnose the problem.
* **TRACE**: Very detailed information, intended only for development. You might keep trace messages for a short period of time after deployment on production environment, but treat these log statements as temporary, that should or might be turned-off eventually. The distinction between DEBUG and TRACE is the most difficult, but if you put logging statement and remove it after the feature has been developed and tested, it should probably be on TRACE level.

Note: The list above is just a suggestion, you can create your own set of instructions to follow, but it is important to have some.

## Know what you are logging

Every time you issue a logging statement, take a moment and have a look at what exactly will land in your log file. Read your logs afterwards and spot malformed sentences. First of all, avoid NPEs like this:

*log.debug("Processing request with id: {}", request.getId());*

Are you absolutely sure that request is not null here?

Another pitfall is logging collections. If you fetched collection of domain objects from the database using Hibernate and carelessly log them like here:

*log.debug("Returning users: {}", users);*

SLF4J will call toString() only when the statement is actually printed, which is quite nice. But if it does… Out of memory error, N+1 select problem, thread starvation (logging is synchronous!), lazy initialization exception, logs storage filled completely – each of these might occur.

It is a much better idea to log, for example, only ids of domain objects (or even only size of the collection).

## Be concise and descriptive

Each logging statement should contain both data and description. Show the data being processed and show its meaning. Show what the program is actually doing. Good logs can serve as a great documentation of the application code itself

Instead of having multiple logs statements:

*log.debug("Message processed");*

*log.debug(message.getJMSMessageID());*

Use the one given below:

*log.debug("Message with id '{}' processed", message.getJMSMessageID());*

## Conversion pattern

Logging pattern is a wonderful tool that transparently adds a meaningful context to every logging statement you make. But you must consider very carefully which information to include in your pattern. For example, logging date when your logs roll every hour is pointless as the date is already included in the log file name. On the contrary, without logging the thread name you would be unable to track any process using logs when two threads work concurrently – the logs will overlap. This might be fine in single-threaded applications but not otherwise.

The performant & informative logging pattern should include (of course except the logged message itself): current time (without date, milliseconds precision), logging level, name of the thread, simple logger name (not fully qualified) and the message. In Logback it is something like:

*%d{HH:mm:ss.SSS} %-5level [%thread][%logger{0}] %m%n*

You should never include file name, class name and line number, although it’s very tempting as logging class name, method name and/or line number has a serious performance impact. Like the one given below may affect application performance:

*[%d{DATE}] [%-5p] %-5c{1}.%M() - %m%n*

## Method arguments & return value

When you find a bug during development, you typically run a debugger trying to track down the potential cause. Now imagine for a while that you can’t use a debugger. For example, because the bug manifested itself on a customer environment few days ago and everything you have is logs. Would you be able to find anything in them?

If you follow the simple rule of logging each method input and output (arguments and return values), you don’t even need a debugger any more. Of course, you must be reasonable but every method that: accesses external system (including database), blocks, waits, etc. should be considered. Simply follow this pattern:

*public String printDocument(Document doc, Mode mode) {*

*log.debug("Entering printDocument(doc={}, mode={})", doc, mode);*

*String id = //Lengthy printing operation*

*log.debug("Leaving printDocument(): {}", id);*

*return id;*

*}*

Because you are logging both the beginning and the end of method invocation, you can manually discover inefficient code and even detect possible causes of deadlocks and starvation – simply by looking after “entering” without corresponding “leaving”.

## Integrating with external systems

This is the special case of the previous tip: if you communicate with an external system, consider logging every piece of data that comes out from your application and gets in. Period. Integration is a tough job and diagnosing problems between two applications (think two different vendors, environments, technology stacks and teams) is particularly hard.

## Log exceptions properly

Logging exceptions is one of the most important roles of logging at all, but many programmers tend to treat logging as a way to handle the exception.

*log.error("Error reading configuration file", e);*

## Don’t log secure information

There is nothing that prevents us from logging a user's username and password in plain text. We must also protect other sensitive information, such as credit card information, social security number or account information. Using a security-specific logger for sensitive messages can help reduce risk. You can configure this logger with a special appender to store messages in an encrypted format or in a secured location.

# Java Performance

One of the most common Java fallacies is; it is slow because compiled Java programs run on the Java Virtual Machine rather than directly on the computer's processor like C and C++ programs do; however, in newer Java versions the execution performance has been optimized significantly mainly thanks to the introduction of just-in-time compilation. Java 7 and 8 introduce a number of new features and tools that make it even easier to get the best possible performance from a Java application. However good the performance of Java is, a badly written program, not taking care of some best practices definitely degrades the performance of your system. Here are some guidelines which may help you write better code which may prevent performance problems.

## Avoid producing a lot of garbage

The Java memory management is quite simple. The memory can be allocated in a program and when that memory is no longer referenced, it is garbage collected. You don’t have to care about these aspects. Theoretically, the garbage collector is executed only when the system load is low, and it is not supposed to affect an application. In real life a heavy used application can become very slow and the run of the garbage collector has an impact.

Always try to avoid generating garbage, because this will reduce the need to run a garbage collection. Just think about three things:

* Reuse existing objects
* Avoid creating unnecessary objects
* Create object pools (Use Object Pool only when object creation is very expensive like threads or data connection

## Recycle and/ or cache objects

The cost of object creation and garbage collection is quite high. If existing objects can be reused, then savings in memory and runtime performance can be realized. Look for opportunities in writing code where existing objects can be reused and recycled instead of creating new ones. Cache objects in frequently used methods so that they will persist between calls, but make sure that the cached object’s state is set properly before subsequent use. For example, if you cached a date object, then prior to using it again, ensure that it is set to the proper date. Some objects are not as straightforward as a date object, so use care.

## Variables

In contrast to a compiled language such as C++, application performance in Java is noticeably affected by what types of variables are accessed and how they are accessed. For example, while stack variables are directly addressable (and may even be placed in registers), instance variables typically require an extra level of indirection to be accessed.

This implies the potential value of data location shifting, changing the storage location of data based on the access patterns. For example, a data-intensive operation would benefit from first copying instance variables into stack variables, operating on the stack variables, and, finally, copying the stack variables back to the permanent instance variables.

This technique is particularly useful when a method variable is accessed repeatedly within a loop. For example, the common loop construct:

*for (int i = 0; ++ i <= limit; )*

Can be improved by 25 percent by rewriting it as:

*for (int i = limit; -- i >= 0; )*

to reduce the number of accesses to the limit variable.

## Avoid explicitly calling the Garbage Collector

Invoking garbage collection when responsiveness is expected, (e.g. when processing a mouse button event) can slow the program down at a time when the user is expecting fast processing. In most circumstances, invoking System.gc() explicitly will not be needed. Invoking System.gc() will not run the Garbage Collector at once, it just says that the Garbage Collector will run next time the system has time to run the Garbage Collector.

## Avoid synchronization

The JVM uses class locks while resolving class references. Class references are resolved on an as used basis (versus resolving at class load time) so it is not easy to predict when the JVM will want a class lock. So, do not synchronize on a class.

An alternate approach would be to create a special purpose lock object instead. For example:

*private static final Object lock = new Object();*

However, synchronization activity can consume significant system resources and affect Java performance. This may be true even if there is only one thread being executed, due to constant checking of lock availability. As a result, synchronization should be reduced as much as possible.

## Use StringBuffer /StringBuilder

Since strings are immutable, any change to a string will create at least one more string object. This degrades performance and unnecessarily creates objects that will eventually need to be garbage collected. StringBuffers, however, are modifiable and can be used to avoid creating temporary String objects.

StringBuilder are more advance API written on the top of StringBuffer can also be used for the same purpose.

## Explicitly close resources

When using the FileInputStream, do not rely on the object finalizer to close the file for you. Explicitly close the file when you are done. The problem with allowing the finalizer to close the file is there is a potentially long delay before the finalizers are actually run. This will keep the operating system file handle in use until the time the finalizers do run potentially creating a situation where the pool of operating system file handles become exhausted. Explicitly closing the file will cause the file handle to be released immediately.

## Limit number of threads

Every java thread created requires dedicated memory for its’ native stack frame. On many systems, the size of this native stack frame is controlled by the -ss Java command line parameter and is the same for every Java thread crea ted. The default stack size on some platforms is as much as 32 kilobytes. For an application that has 20 threads this represents 32KB\*20 or 640KB. Limiting the number of threads in the application will help reduce the system memory requirements. It may also improve performance by giving more CPU time to threads doing real work.

## Avoid excessive writing to the Java console

Writing information to the java console takes time and resources, and should not be done unless necessary. Although helpful in debugging, writing to the console usually involves a great deal of string manipulations, text formatting, and output. These are typically slow operations. Even when the console is not displayed, performance can be adversely affected.

## Data Types

Primitive types are faster than classes encapsulating types. Avoid the costs of object creation and manipulation by using primitive types for variables when prudent. Memory can be reduced and variable access times can be improved.

In the following example, the second declaration is smaller and quicker:

*Currency {*

*public double amount;*

*}*

*double currency\_amount;*

Unlike C++, casting in Java is not done at compile time. Since there is a cost at run-time, avoid unnecessary recasting of variables.

Use int instead of long when possible on 32-bit systems.

long is 64-bit while int is 32-bit data type. 32-bit operations are executed faster than 64-bit on 32-bit systems. Example 1 took about half of the time of Example 2. It is worthwhile noting that Example 2 will run faster on systems with 64-bit addressing.

Use static final when creating constants. When data is invariant, declare it as static and final. By reducing the number of times variables need to be initialized and giving better optimization information to the JVM, performance can be improved.

## When possible, declare methods as final

Declare methods as final whenever possible. Final methods can be handled better by the JVM, leading to improved performance. In this example, the second method will execute faster than the first one.

*void doThing(){*

*for (int i=0; i<100; ++i){*

*dosomething;*

*}*

*}*

*final void doThingfinal(){*

*for (int i=0; i<100; ++i){*

*dosomething;*

*}*

*}*

# Java Security – Best practices

Writing security-conscious Java code can help you avoid security surprises. There are some best practices emerged based on the experiences of various programmers but certainly they won't provide any guarantee that your code is completely secure. It is easy to write insecure code that follows these best practices. By following these practices, however, you will minimize or eliminate certain kinds of security attacks that you might not have thought of.

## Sanitize inputs

SQL injection and remote file inclusion are two of the most common and most dangerous security vulnerabilities around. Basically sanitize user inputs.

## Limit access

Every class, method, and variable that is not private provides a potential entry point for an attacker. By default, everything should be private. Make something non private only with good reason, and document that reason.

## Make everything final

By default, everything should be final. Be aware that non final classes and methods could be extended/overridden by some non-anticipated subclasses. This may lead to unstable code or may even open up some possibility to an attacker to misuse the application by extending some classes. If a class or method isn't final, an attacker could try to extend it in a dangerous and unforeseen way. So make something non final only if there is a good reason, and document that reason.

Do not use non-final public static variables - While final static variables are treated as constants (already at compile time), values of non-final public static variables can be changed from everywhere during runtime, making it impossible to guarantee a consistent state of such variables.

## Secure your data

It's not acceptable these days to handle sensitive data without encrypting it in transit and at rest. That includes user names; passwords; personally identifiable information; and data covered by state, federal, and international regulations.

You might be tempted to store secrets such as cryptographic keys in the code for your application or library. Secrets stored in this way are completely accessible to anybody who runs your code. There is nothing to stop a malicious programmer or virtual machine from looking inside your code and learning its secrets.

Code obfuscation is another way of storing a secret in your code; in the case of obfuscation the secret is simply the algorithm used by your code. There's not much harm in using an obfuscator, but you shouldn't believe it will provide strong protection.

## Don't always trust third-party code

Often, third-party components are poorly managed and rife with exploitable vulnerabilities that may have gone unnoticed so suggestion is to at the least, review recent guidance on ensuring the reliability of third-party software

## Make your classes non-serializable

Serialization is dangerous because it allows adversaries to get their hands on the internal state of your objects. An adversary can serialize one of your objects into a byte array that can be read. This allows the adversary to inspect the full internal state of your object, including any fields you marked private, and including the internal state of any objects you reference.

You can read more about Java security at:

* <http://www.oracle.com/technetwork/java/seccodeguide-139067.html>
* <http://www.secologic.org/downloads/java/051207_Draft_EUROSEC_Whitepaper_Secure_Java_Programming.pdf>