



Chapter 6: Software Construction for Robustness

6.3 Assertions and Defensive Programming 断言与防御式编程

May 8, 2020

Outline

- Recall: Designing an ADT
- Assertions
 - What to Assert and What not to?
 - Guidelines for Using Assertions
- Defensive Programming
- Techniques for defensive programming
- Checklist for Defensive Programming
- The SpotBugs tool
- Summary

6-2节:使用错误处理和exception提高 correctness和robustness,但主要还是面向 robustness。

本节:第2种技术——断言、防御式编程,侧重于correctness。

Reading

- MIT 6.031: 09
- CMU 17-214: Oct 1
- 代码大全: 第8章





1 Recall: Designing an ADT

First Defense: Make Bugs Impossible

- The best defense against bugs is to make them impossible by design. 最好的防御就是不要引入bug ☺
 - Static checking: eliminates many bugs by catching them at compile time.
 - Dynamic checking: Java makes array overflow bugs impossible by catching them dynamically. If you try to use an index outside the bounds of an array or a List, then Java automatically produces an error. --- unchecked exception / runtime error
 - Immutability: An immutable type is a type whose values can never change once they have been created.
 - Immutable values: by final, which can be assigned once but never reassigned.
 - Immutable references: by final, which makes the reference unreassignable, but the object that the reference points to may be mutable or immutable.

Second Defense: Localize Bugs

- If we can't prevent bugs, we can **try to localize them to a small part of the program**, so that we don't have to look too hard to find the cause of a bug. 如果无法避免,尝试着将bug限制在最小的范围内
 - When localized to a single method or small module, bugs may be found simply by studying the program text. 限定在一个方法内部,不扩散

- **Fail fast**: the earlier a problem is observed (the closer to its cause), the easier it is to fix. 尽快失败,就容易发现、越早修复

```
/**
  * @param x requires x >= 0
  * @return approximation to square root of x
  */
public double sqrt(double x) {
    ...
    Since the bad call
    however, the manner
}
```

Pre-condition 如果违反,该方法 可以做任何事

Since the bad call indicates a bug in the caller, however, the most useful behavior would point out the bug as early as possible. 应该尽可能早的指出client的bug

Second Defense: Localize Bugs

- **Assertions** 断言: When the precondition is not satisfied, this code terminates the program by throwing an **AssertionError** exception. The effects of the caller's bug are prevented from propagating. **Fail** fast,避免扩散
- Checking preconditions is an example of defensive programming 检查前置条件是防御式编程的一种典型形式
 - Real programs are rarely bug-free.
 - Defensive programming offers a way to mitigate the effects of bugs even if you don't know where they are.



2 Assertions





(1) What and Why Assertions?

What is assertion?

- An assertion is code that's used during development that allows a program to check itself as it runs, i.e., to test your assumptions about your program logic (such as pre-conditions, post-conditions, and invariants). 断言: 在开发阶段的代码中嵌入,检验某些"假设"是否成立。若成立,表明程序运行正常,否则表明存在错误。
 - When an assertion is true, that means everything is operating as expected.
 - When it's false, that means it has detected an unexpected error in the code.

An example

What is assertion?

- Each assertion contains a boolean expression that you believe will be true when the program executes.
 - If it is not true, the JVM will throw an AssertionError.
 - This error signals you that you have an invalid assumption that needs to be fixed. 出现AssertionError,意味着内部某些假设被违反了
 - The assertion confirms your assumptions about the behavior of your program, increasing your confidence that the program is free of errors. 增强程序员对代码质量的信心:对代码所做的假设都保持正确
- Assertion is much better than using if-else statements, as it serves as proper documentation on your assumptions, and it does not carry performance liability in the production environment. 断言即是对代码中程序员所做假设的文档化,也不会影响运行时性能(在实际使用时,assertion都会被disabled)

What is assertion?

- An assertion usually takes two arguments
 - A boolean expression that describes the assumption supposed to be true
 - A message to display if it isn't.
- The Java language has a keyword assert with two forms:
 - assert condition;
 - assert condition : message;
 - Both statements evaluate the condition and throw an AssertionError if the boolean expression evaluates to false.
 - In the second statement, the expression is passed to the constructor of the AssertionError object and turned into a message string. The description is printed in an error message when the assertion fails, so it can be used to provide additional details to the programmer about the cause of the failure.

所构造的message在发生错误时显示给用户,便于快速发现错误所在

Assertion in Java

- To assert that x is non-negative, you can simply use the statement assert $x \ge 0$;
- Or pass the actual value of x into the AssertionError object, so that it gets displayed later:

assert
$$x >= 0$$
: "x is" + x;

• If x == -1, then this assertion fails with the error message

This information is often enough to get started in finding the bug.

Assertion Example

```
public class AssertionSwitchTest {
   public static void main(String[] args) {
      // assumed either '+', '-', '*', '/' only
      char operator = '%';
      int operand1 = 5, operand2 = 6, result = 0;
      switch (operator) {
         case '+': result = operand1 + operand2; break;
         case '-': result = operand1 - operand2; break;
         case '*': result = operand1 * operand2; break;
         case '/': result = operand1 / operand2; break;
         default: assert false : "Unknown operator: " + operator;
      System.out.println(operand1 + " " + operator + " "
              + operand2 + " = " + result);
```





(2) What to Assert and What not to?

When use assertions?

- Assertion can be used for verifying:
 - Internal Invariants 内部不变量: Assert that a value is within a certain constraint, e.g., assert x > 0.
 - **Rep Invariants** 表示不变量: Assert that an object's state is within a constraint. What must be true about each instance of a class before or after the execution of a method? Class invariants are typically verified via private boolean method, e.g., checkRep().
 - Control-Flow Invariants 控制流不变量: Assert that a certain location will not be reached. For example, the default clause of a switch-case statement.
 - Pre-conditions of methods 方法的前置条件: What must be true when a method is invoked? Typically expressed in terms of the method's arguments or the states of its objects.
 - Post-conditions of methods 方法的后置条件: What must be true after a method completes successfully?

What to Assert?

- Pre-condition: method argument requirements
- Post-condition: Method return value requirements
 - This kind of assertion is sometimes called a *self check* .

```
public double sqrt(double x) {
    assert x >= 0;
    double r;
    ... // compute result r
    assert Math.abs(r*r - x) < .0001;
    return r;
}</pre>
```

What to Assert?

Control-flow: covering all

 If a conditional statement or switch does not cover all the possible cases, it is good practice to use an assertion to block the illegal cases.

```
void foo() {
    for (...) {
        if (...)
            return;
    }
    assert false;
    // Execution should never reach this point!
    switch (vowel) {
        case 'a':
        case 'e':
        case 'i':
        case 'i':
        case 'u': return "A";
        default: assert false;
}
```

But don't use the assert statement here, because it can be turned off. Instead, throw an exception in the illegal cases, so that the check will always happen:

What to Assert: more scenarios

- The value of an input-only variable is not changed by a method
- A pointer is non-NULL
- An array or other container passed into a method can contain at least X number of data elements
- A table has been initialized to contain real values
- A container is empty (or full) when a method begins executing (or when it finishes)
- The results from a highly optimized, complicated method match the results from a slower but clearly written routine

Example

```
/**
 * Solves quadratic equation ax^2 + bx + c = 0.
 * @param a quadratic coefficient, requires a != 0
* @param b linear coefficient
* @param c constant term
* @return a list of the real roots of the equation
*/
public static List<Double> quadraticRoots(final int a, final int b, final int c) {
   List<Double> roots = new ArrayList<Double>();
   // A
    ... // compute roots
   // B
   return roots;
```

What statements would be reasonable to write at position A?

And at position B?

```
assert a != 0;
assert b != 0;
assert c != 0;
assert roots.size() >= 0;
assert roots.size() <= 2;</li>
for (double x : roots) { assert
Math.abs(a*x*x + b*x + c) < 0.0001; }</li>
```

When to use assertions?

- Normally, you don't want users to see assertion messages in production code; assertions are primarily for use during development and maintenance. 断言主要用于开发阶段,避免引入 和帮助发现bug
- Assertions are normally compiled into the code at development time and compiled out of the code for production. 实际运行阶段, 不再使用断言
- During production, they are compiled out of the code so that the assertions don't degrade system performance. 避免降低性能
- When should you write runtime assertions? 使用断言的主要目的是为了在开发阶段调试程序、尽快避免错误
 - As you write the code, not after the fact. When you're writing the code, you have the invariants in mind.
 - If you postpone writing assertions, you're less likely to do it, and you're liable to omit some important invariants.

Avoid putting executable code in assertions

- Since assertions may be disabled, the correctness of your program should never depend on whether or not the assertion expressions are executed.
- In particular, asserted expressions should not have side-effects.
 - For example, if you want to assert that an element removed from a list was actually found in the list, don't write it like this:
- If assertions are disabled, the entire expression is skipped, and x is never removed from the list. Write it like this instead:

```
// don't do this:
assert list.remove(x);
boolean found = list.remove(x);
assert found;
```

Don't Assert External Conditions

- Never use assertions to test conditions that are external to your program. 程序之外的事,不受你控制,不要乱断言
 - Such as the existence of **files**, the availability of the **network**, or the correctness of **input** typed by a human user. 文件/网络/用户输入等
 - Assertions test the internal state of your program to ensure that it is within the bounds of its specification. 断言只是检查程序的内部状态是否符合规约
 - When an assertion fails, it indicates that the program has run off the rails in some sense, into a state in which it was not designed to function properly. Assertion failures therefore indicate bugs. 断言一旦false,程序就停止执行
 - External failures are not bugs, and there is no change you can make to your program in advance that will prevent them from happening. 你的代码无法保证不出现此类外部错误 (recall section 7-2)
 - External failures should be handled using exceptions instead. Avoid trivial assertions, just as you would avoid uninformative comments. 外部错误要使用Exception机制去处理

Turn on/off Assert in different phases

- Many assertion mechanisms are designed so that assertions are executed only during testing and debugging, and turned off when the program is released to users.
 - Assertions are a great tool for keeping your code safe from bugs,
 but Java has them off by default! Java缺省关闭断言,要记得打开(-ea)
- The advantage of this approach is that you can write very expensive assertions that would otherwise seriously degrade the performance of your program. 断言非常影响运行时的性能
 - For example, a procedure that searches an array using binary search has a requirement that the array be sorted.
 - Asserting this requirement requires scanning through the entire array, however, turning an operation that should run in logarithmic time into one that takes linear time.
 - You should be willing to pay this cost during testing, since it makes debugging much easier, but not after the program is released to users.

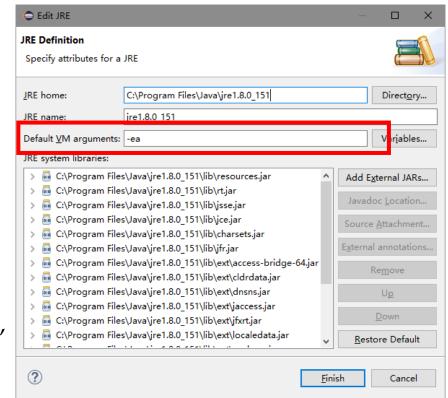
Enable & Disable assertions in Java

Enable assertions

- Running the program with the -enableassertions or -ea option:
 - java -enableassertions MyApp
- The option -ea... turns on assertions in all classes of the default package.
 - java -ea:MyClass MyApp

Disable assertions

- Running the program with the-disableassertions or -da option:
 - java -ea:... -da:MyClass MyApp
- By default, assertions are disabled.
- Enable assertions in Eclipse:
 - In preferences, go to Java → Installed JREs . Click "Java SE 8", click "Edit...", and in the "Default VM arguments" box enter: -ea







(3) Guidelines for Using Assertions

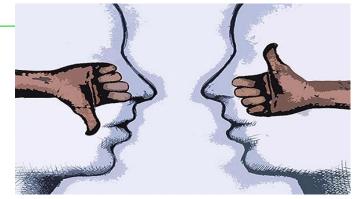
Assertion vs. Exception?

- Assertions generally cover correctness issues of program.
 - If an assertion is fired for an anomalous condition, the corrective action is not merely to handle an error gracefully the corrective action is to change the program's source code, recompile, and release a new version of the software. 断言→Correctness
- Exceptions generally cover robustness issues of program.
 - If error handling code is used to address an anomalous condition, the error handling will enable the program to respond to the error gracefully.
 错误/异常处理→Robustness
- Assertions are especially useful in large, complicated programs and in high reliability programs.
 - They enable programmers to more quickly flush out mismatched interface assumptions, errors that creep in when code is modified, and so on.

Assertion vs. Exception?

- Use error handling code (exception) for conditions you expect to occur 使用异常来处理你"预料到可以发生"的不正常情况
 - Error handling code checks for off-nominal circumstances that might not occur very often, but that have been anticipated by the programmer who wrote the code and that need to be handled by the production code.
- Use assertions for conditions that should never occur 使用断言处理 "绝不应该发生"的情况
 - Assertions check for bugs in the code.

 Another viewpoint: Do not use assertions for argument checking in public methods. 在其他一些开发者眼里,不应该针对参数的 合法性使用断言。



Reasons:

- Argument checking is typically part of the
 published specifications (or contract) of a method, and these specifications
 must be obeyed whether assertions are enabled or disabled. 不管是否-ea,
 spec中的pre-/post-conditions都能够被保证
- Another problem with using assertions for argument checking is that erroneous arguments should result in an appropriate runtime exception (such as IllegalArgumentException, IndexOutOfBoundsException, or NullPointerException). An assertion failure will not throw an appropriate exception. 即使spec被违反,也不应通过assert直接fail,而是应抛出具体的runtime异常

Use Assertions for pre-/post- conditions

- If the variables latitude, longitude, and elevation were coming from an external source, invalid values should be checked and handled by error handling code rather than assertions. 如果参数来 自于外部(不受自己控制),使用异常处理
- If the variables are coming from a trusted, internal source, however, and the routine's design is based on the assumption that these values will be within their valid ranges, then assertions are appropriate. 如果来自于自己所写的其他代码,可以使用断言来帮助发现错误(例如postcondition就需要)

```
float latitude;
float longitude;
float elevation:
//Preconditions
assert latitude>=-90 && latitude<=90;</pre>
assert longitude>=0 && longitude<360;</pre>
assert elevation>=-500 && elevation<=7500
//Postconditions
assert Velocity>=0 && Velocity <= 600
return Velocity;
```

```
/**
  * Sets the refresh rate.
   @param rate refresh rate, in frames per second.
  * @throws IllegalArgumentException if rate <= 0 or
  * rate > MAX_REFRESH_RATE.
*/
public void setRefreshRate(int rate) {
  // Enforce specified precondition in public method
  if (rate <= 0 || rate > MAX_REFRESH_RATE)
   throw new IllegalArgumentException("Illegal rate: " + rate);
  setRefreshInterval(1000/rate);
}
```

You can use an assertion to test a nonpublic method's precondition that you believe will be true no matter what a client does with the class.

```
/**
* Sets the refresh interval (which must correspond to a legal rate).
  @param interval refresh interval in milliseconds.
*/
private void setRefreshInterval(int interval) {
 // Confirm adherence to precondition in nonpublic method
 assert interval > 0
         && interval <= 1000/MAX_REFRESH_RATE : interval;</pre>
 ... // Set the refresh interval
```

 You can test postcondition with assertions in both public and nonpublic methods.

```
/**
  * Returns a BigInteger whose value is (this-1 mod m).
  *
  * @param m the modulus.
  * @return this-1 mod m.
  * @throws ArithmeticException m <= 0, or this BigInteger
  * has no multiplicative inverse mod m (that is, this BigInteger
  * is not relatively prime to m).
  */
public BigInteger modInverse(BigInteger m) {
  if (m.signum <= 0)</pre>
    throw new ArithmeticException("Modulus not positive: " + m);
  ... // Do the computation
  assert this.multiply(result).mod(m).equals(ONE) : this;
  return result;
```

Combine assert & exception handling for robustness

- Both assertions and exception handling code might be used to address the same error. 断言和异常处理都可以处理同样的错误
 - In the source code for Microsoft Word, for example, conditions that should always be true are asserted, but such errors are also handled by error-handling code in case the assertion fails.
 - For extremely large, complex, long-lived applications like Word, assertions are valuable because they help to flush out as many development-time errors as possible.
- But the application is so complex (million of lines of code) and has gone through so many generations of modification that it isn't realistic to assume that every conceivable error will be detected and corrected before the software ships, and so errors must be handled in the production version of the system as well.

开发阶段用断言尽可能消除bugs 在发行版本里用异常处理机制处理漏掉的错误





3 Defensive Programming

What is defensive programming?

- Defensive programming is a form of defensive design intended to ensure the continuing function of a piece of software under unforeseen circumstances.
 - Defensive programming practices are often used where high availability, safety or security is needed.

The idea can be viewed as reducing or eliminating the prospect of

Murphy's Law having effect.

"Your code should fail as early as possible!"



Murphy's Law

IF ANYTHING CAN GO WRONG

ITWILL

9

MURPHY'S LAW

What is defensive programming?

- The idea is based on defensive driving
 - You adopt the mind-set that you're never sure what the other drivers are going to do.
 - That way, you make sure that if they do something dangerous you won't be hurt.
 - You take responsibility for protecting yourself even when it might be the other driver's fault.
 - 眼观六路,耳听八方,一旦 其他车辆有对你产生危险的 症状,马上采取防御式行动







4 Techniques for defensive programming

Techniques for defensive programming

- Protecting programs from invalid inputs
- Assertions
- Exceptions
- Specific error handling techniques
- Barricade
- Debugging aids
- The best form of defensive coding is not inserting errors in the first place.
- You can use defensive programming in combination with the other techniques.

(1) Protecting Programs From Invalid Inputs

- "Garbage in, garbage out"
 - That expression is essentially software development's version of caveat emptor: let the user beware. 货物出门概不退换
- For production software, garbage in, garbage out isn't good enough.
- A good program never puts out garbage, regardless of what it takes in.
 - "Garbage in, nothing out"
 - "Garbage in, error message out"
 - "No garbage allowed in"



"Garbage in, garbage out" is the mark of a sloppy, non-secure program.

Protecting Programs From Invalid Inputs

- Check the values of all data from external sources 对来自外部的数据源要仔细检查,例如:文件、网络数据、用户输入等
 - When getting data from a file, a user, the network, or some other external interface, check to be sure that the data falls within the allowable range.

Examples:

- Make sure that numeric values are within tolerances and that strings are short enough to handle.
- If a string is intended to represent a restricted range of values (such as a financial transaction ID or something similar), be sure that the string is valid for its intended purpose; otherwise reject it.
- If you're working on a secure application, be especially leery of data that might attack your system: attempted buffer overflows, injected SQL commands, injected html or XML code, integer overflows, and so on.

Protecting Programs From Invalid Inputs

- Check the values of all routine input parameters 对每个函数的输入 参数合法性要做仔细检查,并决定如何处理非法输入
 - Checking the values of routine input parameters is essentially the same as checking data that comes from an external source, except that the data comes from another routine instead of from an external interface.
- Decide how to handle bad inputs
 - Once you've detected an invalid parameter, what do you do with it?
 - Depending on the situation, you might choose any of a dozen different approaches, which are described in detail later in this chapter.

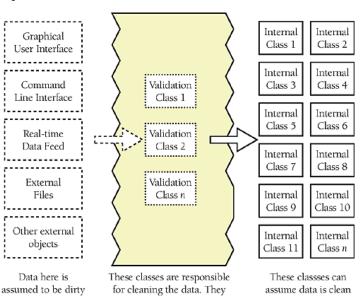
(2) Barricade 设置路障

and untrusted.

Barricades are a damage-containment strategy.

- The reason is similar to that for having isolated compartments in the hull of a ship and firewalls in a building.
- One way to barricade for defensive programming purposes is to designate certain interfaces as boundaries to "safe" areas.
- Check data crossing the boundaries of a safe area for validity and respond sensibly if the data isn't valid.

and trusted.



make up the barricade.

Defining some parts of the software that work with dirty data and some that work with clean can be an effective way to relieve the majority of the code of the responsibility for checking for bad data.

Barricade

- The class's public methods assume the data is unsafe, and they are responsible for checking the data and sanitizing it. 类的public方法接收到的外部数据都应被认为是dirty的,需要处理干净再传递到private方法——隔离舱
 - Once the data has been accepted by the class's public methods, the class's private methods can assume the data is safe.
- Another way is as an operating-room technique. 操作间技术
 - Data is sterilized before it's allowed to enter the operating room. Anything that's in the operating room is assumed to be safe.
 - The key design decision is deciding what to put in the operating room, what to keep out, and where to put the doors—which routines are considered to be inside the safety zone, which are outside, and which sanitize the data.
 - The easiest way to do this is usually by sanitizing external data as it arrives, but data often needs to be sanitized at more than one level, so multiple levels of sterilization are sometimes required.

Relationship between Barricades and Assertions

- The use of barricades makes the distinction between assertions and error handling clean cut.
 - Routines that are outside the barricade should use error handling because it isn't safe to make any assumptions about the data.
 - Routines inside the barricade should use assertions, because the data passed to them is supposed to be sanitized before it's passed across the barricade. If one of the routines inside the barricade detects bad data, that's an error in the program rather than an error in the data.
- "隔离舱"外部的函数应使用异常处理, "隔离舱"内的函数应使用断言。
 - The use of barricades also illustrates the value of deciding at the architectural level how to handle errors.
 - Deciding which code is inside and which is outside the barricade is an architecture-level decision.
- Proxy设计模式? ——隔离



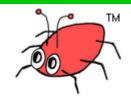


5 Checklist for Defensive Programming



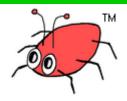
6 The SpotBugs tool

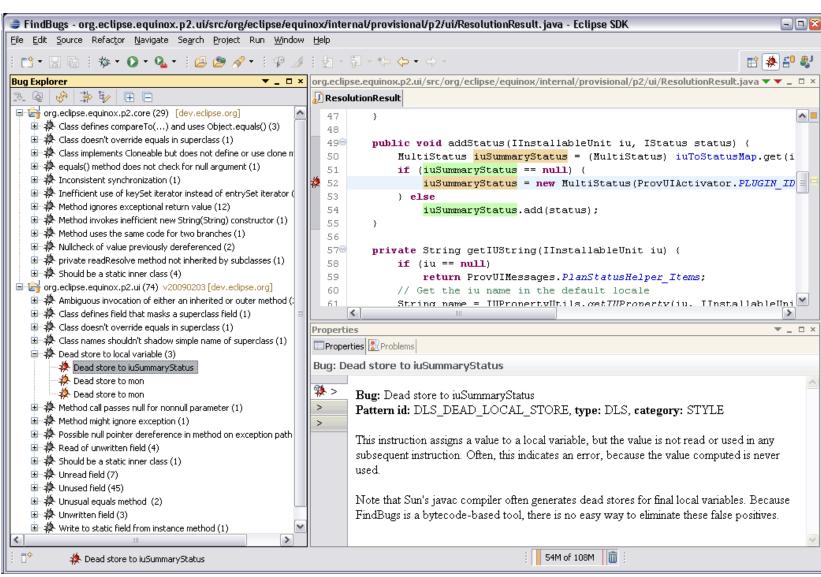
早期版本: FindBugs

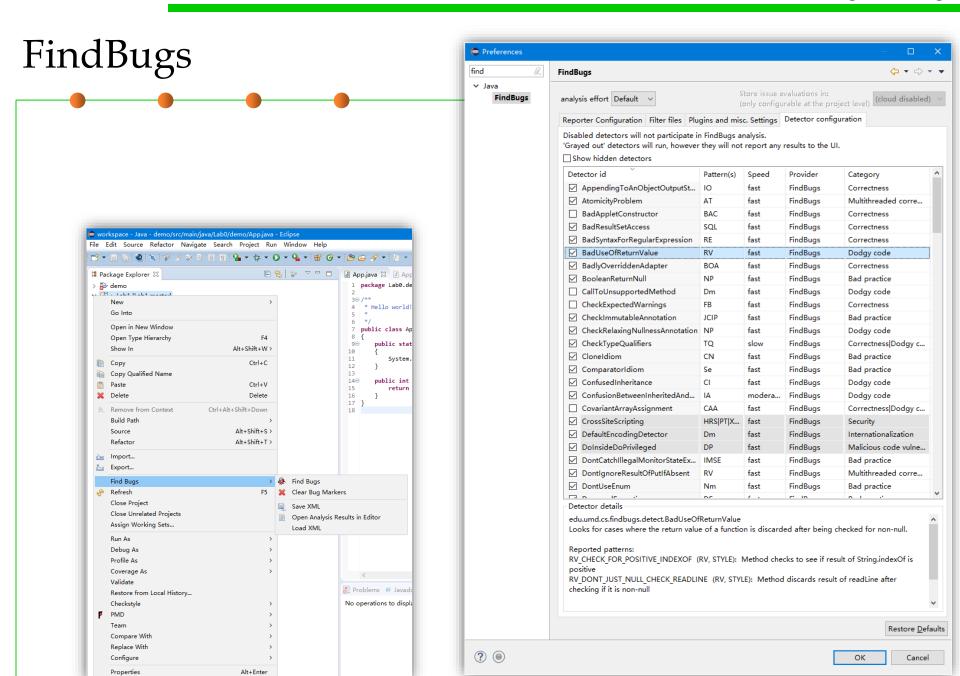


- **FindBugs** is a program which uses **static analysis** to look for bugs in Java code. **Java**静态代码分析工具
 - It operates on Java bytecode.
- Potential errors are classified in four ranks: scariest, scary, troubling and of concern. This is a hint to the developer about their possible impact or severity.
 - Bug list can be found in http://findbugs.sourceforge.net/bugDescriptions.html
- It is distributed as a stand-alone GUI application, but also plug-ins available for Eclipse, Gradle, Maven, and Jenkins.
 - http://findbugs.sourceforge.net/
 - http://findbugs.cs.umd.edu/eclipse/

FindBugs







SpotBugs



- SpotBugs is a program which uses static analysis to look for bugs in Java code.
 - It is the spiritual successor of FindBugs, carrying on from the point where it left off with support of its community.
- It checks for more than 400 bug patterns
 - https://spotbugs.readthedocs.io/en/latest/bugDescriptions.html
- It can be used standalone and through several integrations, including: Ant, Maven, Gradle, and Eclipse
- 开源Git仓库:
 - https://github.com/spotbugs/spotbugs

Examples of SpotBugs patterns

NP: equals() method does not check for null argument (NP_EQUALS_SHOULD_HANDLE_NULL_ARGUMENT)

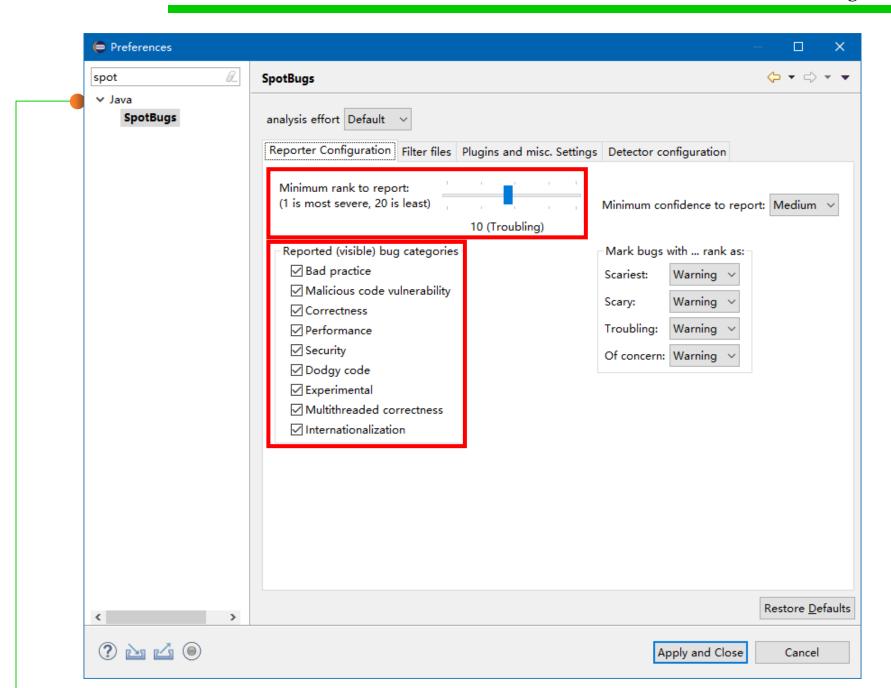
This implementation of equals(Object) violates the contract defined by java.lang.Object.equals() because it does not check for null being passed as the argument. All equals() methods should return false if passed a null value.

HE: Class defines equals() and uses Object.hashCode() (HE_EQUALS_USE_HASHCODE)

This class overrides <code>equals(Object)</code>, but does not override <code>hashCode()</code>, and inherits the implementation of <code>hashCode()</code> from <code>java.lang.Object</code> (which returns the identity hash code, an arbitrary value assigned to the object by the VM). Therefore, the class is very likely to violate the invariant that equal objects must have equal hashcodes.

If you don't think instances of this class will ever be inserted into a HashMap/HashTable, the recommended hashcode implementation to use is:

```
public int hashCode() {
   assert false : "hashCode not designed";
   return 42; // any arbitrary constant will do
}
```





Summary

Summary

- Recall: Designing an ADT
- Assertions
 - What to Assert and What not to?
 - Guidelines for Using Assertions
- Defensive Programming
- Techniques for defensive programming
- Checklist for Defensive Programming
- The SpotBugs tool



The end

May 8, 2020