

CSE307: Software Engineering

Code Smell Detection using SonarLint and CheckStyle

Submitted By:

Sub-Section: A2

Group Members:

1. Rasman Mubtasim Swargo (ID: 1705051)
2. Apurba Saha (ID: 1705056)

Department: CSE

Level 3 Term 1

Date of submission: 28 July, 2021

Overview

Code Smells are issues with the style of the code that makes code harder to read, understand, and maintain. They are super common. Code smells are usually not bugs. They are not technically incorrect and do not prevent the program from functioning. Instead, they indicate weaknesses in design that may slow down development or increase the risk of bugs or failures in the future.

In this assignment, we are checking some of our past codes for bad code smell. For C++ codes, we are using SonarLint and for Java codes, we are using CheckStyle and built-in IntelliJ tool.

All the codes can be found in the following github repository

<https://github.com/diponsaha007/Code-Smell>

Note: As all the codes lack comments, function and class documentation, consistent naming, we will not mention these everytime.

CPP Codes

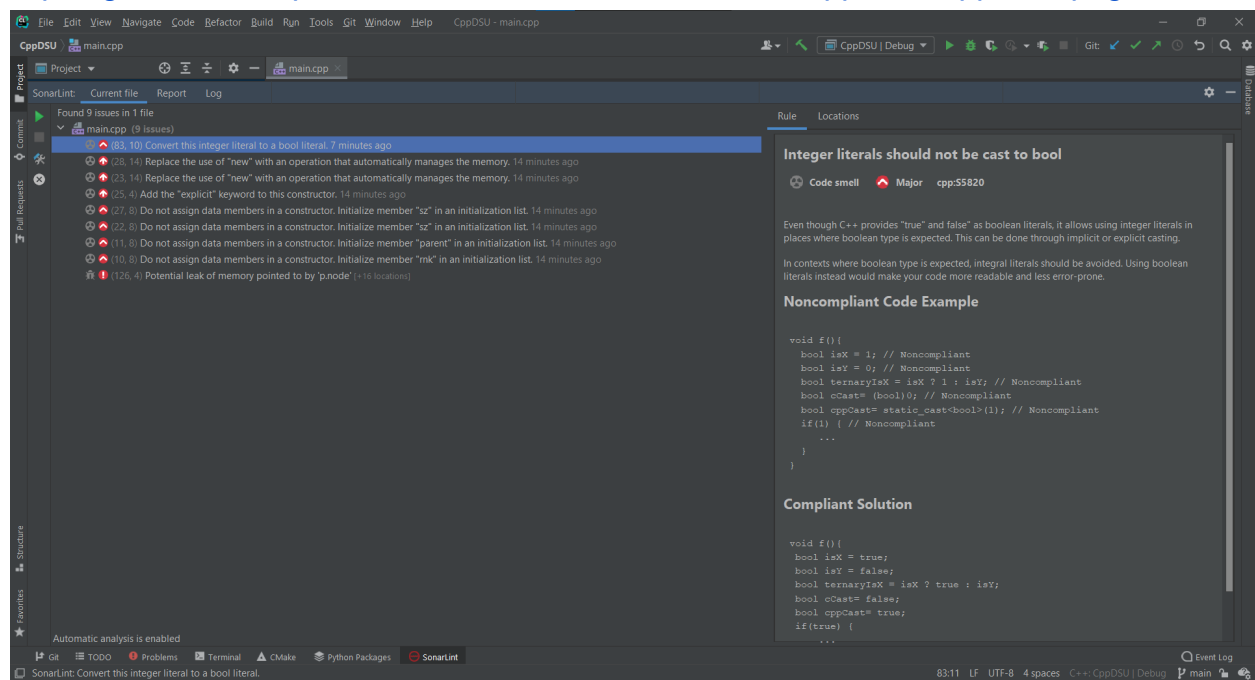
Code 1:

<https://github.com/diponsaha007/Code-Smell/blob/main/CppDSU/main.cpp>

This code implements a simple Disjoint Set Union (DSU) class in C++. When we analyzed the code with SonarLint the following issues shown in the image were raised.

Link to image for better quality:

<https://github.com/diponsaha007/Code-Smell/blob/main/CppDSU/CppDSU.png>



Overview of the bad smells detected:

- **Dynamically allocated memory is not released:** At line 23 and 28, we allocate memory with new operator for node* but there is no deallocation of memory creating potential memory leak. Dynamically allocated memory should be released.

```
node= new Node[sz];
```

- **Memory is managed manually:** Data structures such as vector or list can be used for automatic memory management.

- **"explicit" is not used on single parameter constructors and conversion operators:** At line 25, there is no explicit keyword before single parameter constructor. When the expected parameter is a class with a single-argument constructor, the compiler will implicitly pass the method argument to that constructor to implicitly create an object of the correct type for the method invocation. Alternatively, if the wrong type has a conversion operator to the correct type, the operator will be called to create an object of the needed type. But using implicit conversions makes the execution flow difficult to understand. Readers may not notice that a conversion occurs, and if they do notice, it will raise a lot of questions: Is the source type able to convert to the destination type? Is the destination type able to construct an instance from the source? Is it both? And if so, which method is called by the compiler? Moreover, implicit promotions can lead to unexpected behavior, so they should be prevented by using the explicit keyword on single-argument constructors and conversion operators. Doing so will prevent the compiler from performing implicit conversions.

```
Disjoint_Set(int n)
{
    sz=n;
    node= new Node[sz];
}
```

- **Class members are not initialized in an initialization list:** Member data should be initialized in-class or in a constructor initialization list. At line 10,11,22 and 27 data members are assigned in a constructor. Member data should be initialized in-class or in a constructor initialization list.

```
Node()
{
    rnk=-1;
    parent=-1;
}
```

- **Integer literal is used instead of Bool:** At line 83, integer literal is used inside 'while' condition which should be bool. Even though C++ provides "true" and "false" as boolean literals, it allows using integer literals in places where boolean type is expected. Where boolean type is expected, integral literals should be avoided. Using boolean literals instead would make the code more readable and less error-prone

```
while(1){..}
```

- **Magic Number:** Magic numbers are numeric representations that should be replaced with a symbolic representation. At line 22, we have

```
sz = 100;
```

Here 100 is the default size when no size is passed in the constructor. However, using numerical constants makes it harder for the reader to understand what the constant signifies. So we could do,

```
#define DEFAULT_SIZE 100  
sz = DEFAULT_SIZE;
```

This makes the code more readable.

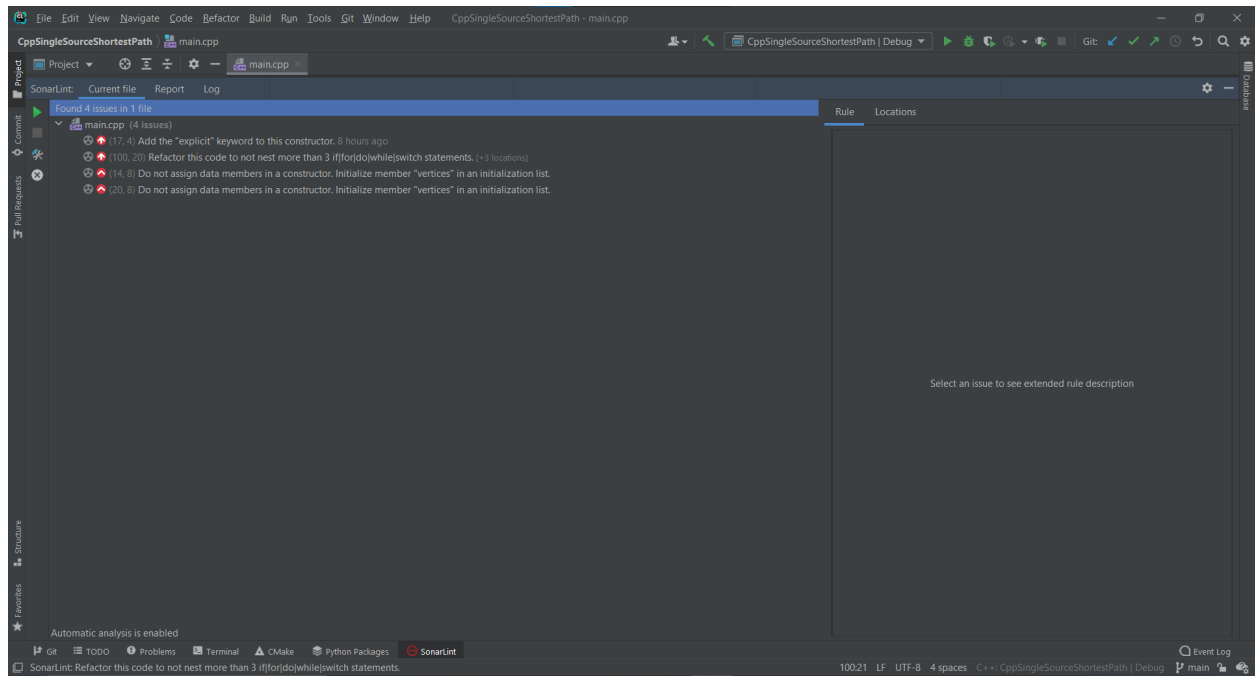
Code 2:

<https://github.com/diponsaha007/Code-Smell/blob/main/CppSingleSourceShortestPath/main.cpp>

This code implements a graph class which contains two single source shortest path algorithms (Dijkstra and Bellman-Ford) in C++. When we analyzed the code with SonarLint the following issues shown in the image were raised.

Link to image for better quality:

<https://github.com/diponsaha007/Code-Smell/blob/main/CppSingleSourceShortestPath/CppSSSP.png>



Overview of the bad smells detected:

- **"explicit" is not used on single parameter constructors and conversion operators:** At line 17, there is no explicit keyword before single parameter constructor.

```
Graph(int n)
{
    adj.resize(n);
    vertices = n;
}
```

- **Class members are not initialized in an initialization list:** At line 14 and 20 data members are assigned in a constructor. Member data should be initialized in-class or in a constructor initialization list.

```
Graph()
{
    vertices = 0;
}
```

- **More than 3 nested if | for | while statements:** At line 94 - 107, we have 4 nested control flow statements. Nested if, for, do, while, switch and try statements is a key ingredient for making what's known as "Spaghetti code".

```

for (int i = 1; i <= vertices - 1; i++)
{
    for (int a = 0; a < vertices; a++)
    {
        for (pair<int, int> b : adj[a])
        {
            if (dist[b.first] > dist[a] + b.second)
            {
                dist[b.first] = dist[a] + b.second;
                parent[b.first] = a;
            }
        }
    }
}

```

Such code is hard to read, refactor and therefore maintain. Though this might be an overkill here as the algorithm requires 4 nested control flow statements, however we can make it 3. The 2 nested loops at line 96 and 98 basically traverse each edge of the graph. So we can use a define for that like below:

```

#define FOR_EACH_EDGE(a,b,vertices,adj) for (int a = 0; a < vertices; a++) for (pair<int, int> b : adj[a])

```

Then the nested control flow statements can be nested as,

```

for (int i = 1; i <= vertices - 1; i++) //level 1
    FOR_EACH_EDGE(a,b,vertices,adj) //level 2
        if (dist[b.first] > dist[a] + b.second) //level 3
        {
            dist[b.first] = dist[a] + b.second;
            parent[b.first] = a;
        }

```

Again, this might be an overkill from the context of the bellman ford algorithm.

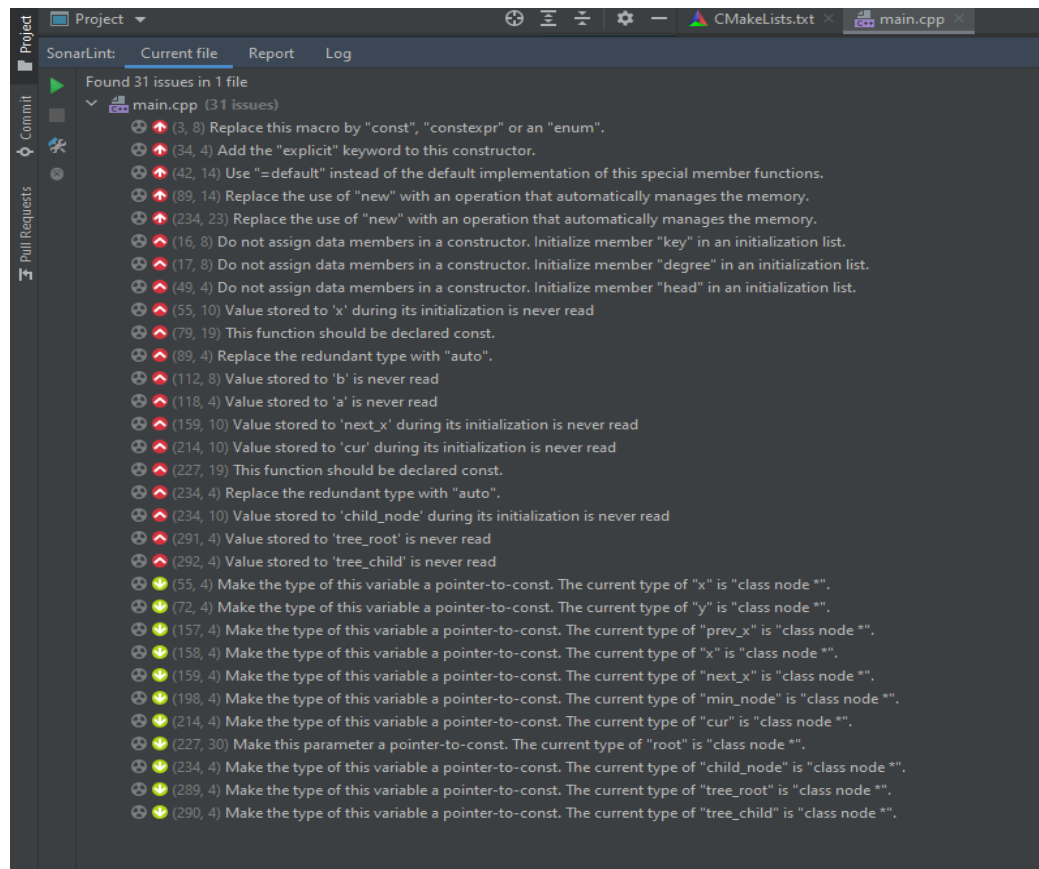
Code 3:

<https://github.com/diponsaha007/Code-Smell/blob/main/CPPBinomialHeap/main.cpp>

This code implements a Binomial Heap in C++. When we analyzed the code with SonarLint the following issues shown in the image were raised.

Link to image for better quality:

<https://github.com/diponsaha007/Code-Smell/blob/main/CPPBinomialHeap/CPPBinHeap.png>



Overview of the bad smells detected:

- **Macros should not be used to define constants:**

At line 3 we used macro to define "inf". A macro is a textual replacement, which means that it's not respecting the type system, it's not respecting scoping rules... There is no reason not to use a constant instead.

```
#define inf 99999
```


- **"explicit" is not used on single parameter constructors and conversion operators:**

At line 34, there is no explicit keyword before single parameter constructor.

```
BinomialHeap(node* x);
```

- **Special member function should not be defined unless a non standard behavior is required:**

At line 42, the default constructor was defined. But all special member functions (default constructor, copy and move constructors, copy and move assignment operators, destructor) can be automatically generated by the compiler if we don't prevent it (for many classes, it is good practice to organize your code so that you can use these default versions).

```
BinomialHeap::BinomialHeap()  
{  
    head = NULL;  
}
```

- **Memory should not be managed manually**

At line 89, 234.

If we manage memory manually, it's our responsibility to delete all memory created with new, and to make sure it's deleted once and only once. Ensuring this is done is error-prone, especially when your function can have early exit points.

```
node* x = new node(k);
```

```
node* child_node = new node();
```

- **Member data should be initialized in-class or in a constructor initialization list**

At lines 16, 17, 49.

There are three ways to initialize a non-static data member in a class:

- With an in-class initializer
- In the initialization list of a constructor
- In the constructor body

We should use those methods in that order of preference. When applicable, in-class initializers are best, because they apply automatically to all constructors of the class (except for default copy/move constructors and constructors where an explicit initialization for this member is provided). But they can only be used for initialization with constant values.

If our member value depends on a parameter, we can initialize it in the constructor's initialization list. If the initialization is complex, you can define a function to compute the value, and use this function in the initializer list.

```
node(int val = -1, int dgr = 0)
{
    key = val;
    degree = dgr;
    child = NULL;
    sibling = NULL;
    parent = NULL;
}
```

- **Unused assignments should be removed**

At line 55, 12, 18, 59, 214, 234, 291, 292.

A dead store happens when a local variable is assigned a value that is not read by any subsequent instruction.

```
node* x = head;
```

- **"auto" should be used to avoid repetition of types**

At line 89, 234.

When used as a type specifier in a declaration, `auto` allows the compiler to deduce the type of a variable based on the type of the initialization expression.

```
node* x = new node(k);
```

```
node* child_node = new node();
```

- **Member functions that don't mutate their objects should be declared "const"**

At line 79, 227.

No member function can be invoked on a const-qualified object unless the member function is declared "const".

Qualifying member functions that don't mutate their object with the "const" qualifier makes your interface easier to understand; you can deduce without diving into implementation if a member function is going to mutate its object

```
void BinomialHeap::binomial_link(node*y, node*z)
{
    y->parent = z;
    y->sibling = z->child;
    z->child = y;
    z->degree = z->degree + 1;
}
```

- **Pointer and reference local variables should be "const" if the corresponding object is not modified**

At lines 55, 72, 157, 158, 159, 198, 214, 227, 234, 289, 290.

This rule leads to greater precision in the definition of local variables by making the developer's intention about modifying the variable explicit.

```
node* x = head;
```

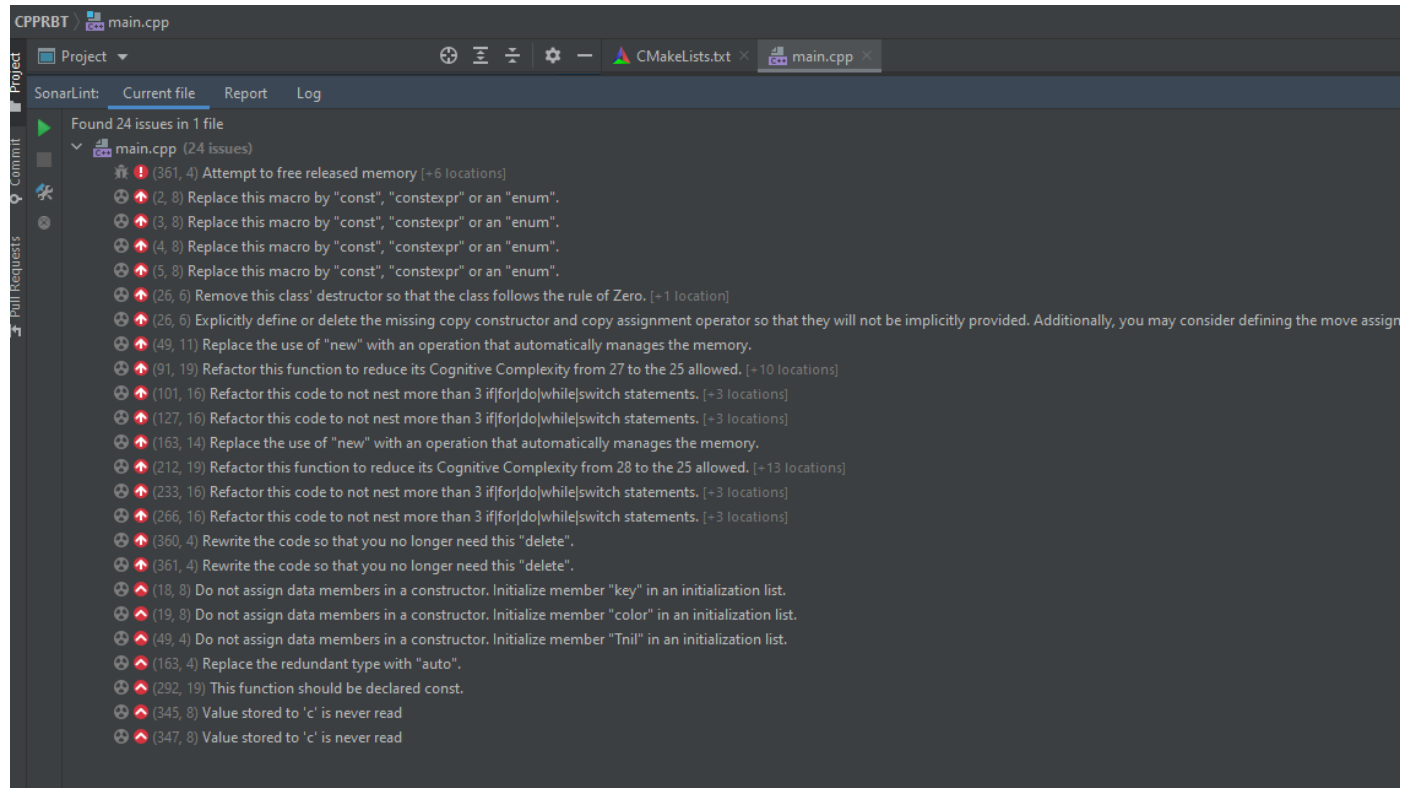
Code 4:

<https://github.com/diponsaha007/Code-Smell/blob/main/CPPRBT/main.cpp>

This code implements a Red Black Tree in C++. When we analyzed the code with SonarLint the following issues shown in the image were raised.

Link to image for better quality:

<https://github.com/diponsaha007/Code-Smell/blob/main/CPPRBT/CPPRBT.png>



Overview of the bad smells detected:

- **Macros should not be used to define constants**

At lines 2, 3, 4, 5.

```
#define RED 1
#define BLACK 2
#define DBLACK 3
#define nil -1
```

- **Memory locations should not be released more than once**

At line 361.

Using `free(...)` or `delete` releases the reservation on a memory location, making it immediately available for another purpose. So releasing the same memory location twice can lead to corrupting the program's memory.

```
delete Tnil;
```

- **The "Rule-of-Zero" should be followed**

At line 26.

Classes that avoid directly handling resources don't need to define any of the special member functions required to properly handle resources: destructor, copy constructor, move constructor, copy-assignment operator, move-assignment operator. That's because the versions of those functions provided by the compiler do the right thing automatically, which is especially useful because writing these functions correctly is typically tricky and error-prone.

Omitting all of these functions from a class is known as the Rule of Zero because no special function should be defined.

```
class RedBlackTree
{
    node *root;
    node *Tnil;
    void left_rotate(node *x);
    void right_rotate(node *y);
    void insert_fixup(node *z);
    void preorder(node *ptr);
    node* tree_minimum(node *x);
    void rb_transplant(node* u, node* v);
    void delete_fixup(node* x);
    node* tree_search(node*x, int k);
public:
    RedBlackTree();
    void rbinser(int k);
    void search_value(int k);
    void rbdelete(int k);
    ~RedBlackTree();
};
```

- **When the "Rule-of-Zero" is not applicable, the "Rule-of-Five" should be followed**

At line 26.

In C++, we should not directly manipulate resources (a database transaction, a network connection, a mutex lock), but encapsulate them in RAII wrapper classes that will allow us to manipulate them safely. When defining one of those wrapper classes, we cannot rely on the compiler-generated special member functions to manage the class' resources for you. We must define those functions yourself to make sure the class' resources are properly copied, moved, and destroyed.

```
class RedBlackTree
{
    node *root;
    node *Tnil;
    void left_rotate(node *x);
    void right_rotate(node *y);
    void insert_fixup(node *z);
    void preorder(node *ptr);
    node* tree_minimum(node *x);
    void rb_transplant(node* u, node* v);
    void delete_fixup(node* x);
    node* tree_search(node*x, int k);
public:
    RedBlackTree();
    void rbinser(int k);
    void search_value(int k);
    void rbdelete(int k);
    ~RedBlackTree();
};
```

- **Memory should not be managed manually**

At lines 49, 163, 360, 361.

```
Tnil = new node(-1, BLACK);
```

- **The cognitive Complexity of functions should not be too high**

At line 91, 212.

Cognitive Complexity is a measure of how hard the control flow of a function is to understand. Functions with high Cognitive Complexity will be difficult to maintain.

```
void RedBlackTree::insert_fixup(node *z)
{
    if(z->parent!=NULL)
    {
        while (z->parent->color == RED)
        {
            if (z->parent == z->parent->parent->left)
            {

                node *y = z->parent->parent->right;
                if (y->color==RED)
                {
                    z->parent->color = BLACK;
                    y->color = BLACK;
                    z->parent->parent->color=RED;
                    z=z->parent->parent;
                }

                else
                {
                    if(z==z->parent->right)
                    {
                        z = z->parent;
                        left_rotate(z);
                    }
                    z->parent->color = BLACK;
                    z->parent->parent->color = RED;
                    right_rotate(z->parent->parent);
                }
            }

            else
            {

                node *y = z->parent->parent->left;
                if (y->color==RED)
                {
                    z->parent->color = BLACK;
                    y->color = BLACK;
                    z->parent->parent->color=RED;
                    z=z->parent->parent;
```

```

    }
    else
    {
        if(z==z->parent->left)
        {
            z = z->parent;
            right_rotate(z);

        }
        z->parent->color = BLACK;

        z->parent->parent->color = RED;

        left_rotate(z->parent->parent);

    }
}
}

}
root->color=BLACK;
}

```

- **Control flow statements "if", "for", "while", "switch" and "try" should not be nested too deeply**

At lines 101, 127, 133, 166.

Nested if, for, do, while, switch and try statements is a key ingredient for making what's known as "Spaghetti code".

Such code is hard to read, refactor and therefore maintain.

```

void RedBlackTree::insert_fixup(node *z)
{
    if(z->parent!=NULL)
    {
        while (z->parent->color == RED)
        {
            if (z->parent == z->parent->parent->left)
            {

```



```

        if (y->color==RED)
        {
        }

        else
        {
            if(z==z->parent->right)
            {

            }

        }

    }

    else
    {
        if (y->color==RED)
        {
        }
        else
        {
            if(z==z->parent->left)
            {
            }

        }

    }

}

}

}

```

- **Member data should be initialized in-class or in a constructor initialization list**

At lines 18, 19, 49.

```
node(int val = -1, int clr = RED)
{
    key = val;
    color = clr;
    left = NULL;
    right = NULL;
    parent = NULL;
}
```

- **"auto" should be used to avoid repetition of types**
At line 163.

```
node *z = new node(k);
```

- **Member functions that don't mutate their objects should be declared "const"**
At line 292.

```
void RedBlackTree::search_value(int k)
{
    if(tree_search(root, k)==Tnil) cout<<"False"<<endl;
    else cout<<"True"<<endl;
}
```

- **Unused assignments should be removed**
At lines 345, 347.

```
if(ptr->color==BLACK)
    c='b';
```

Java Codes:

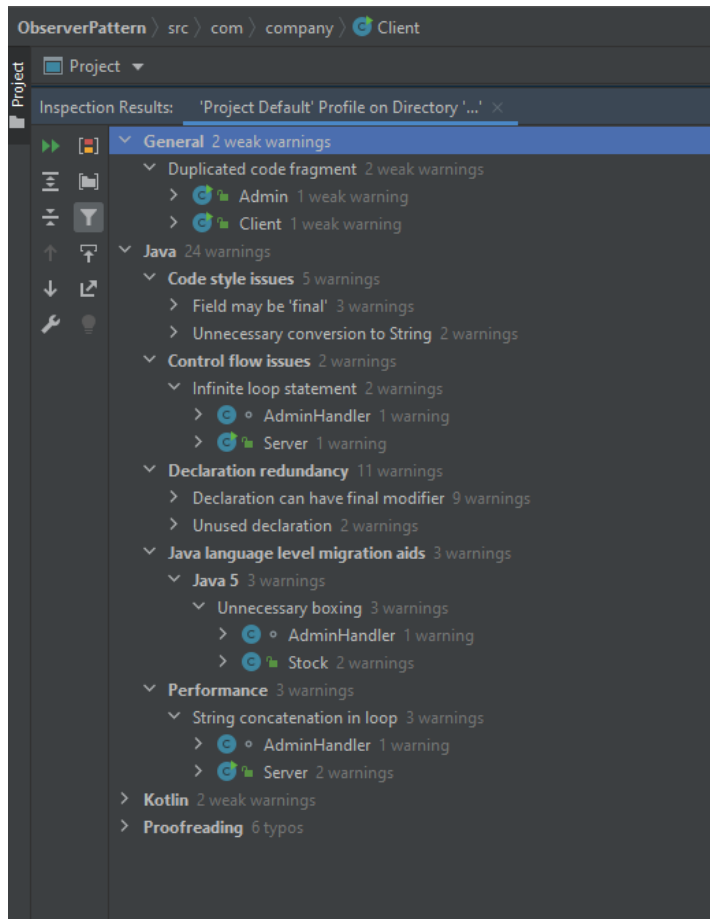
Code 1:

<https://github.com/diponsaha007/Code-Smell/tree/main/JavaObserverPattern/src>

This code basically implements a demo Observer Pattern in Java. When we analyzed the code with the default analyzer the following issues shown in the image arose.

Link to image for better quality:

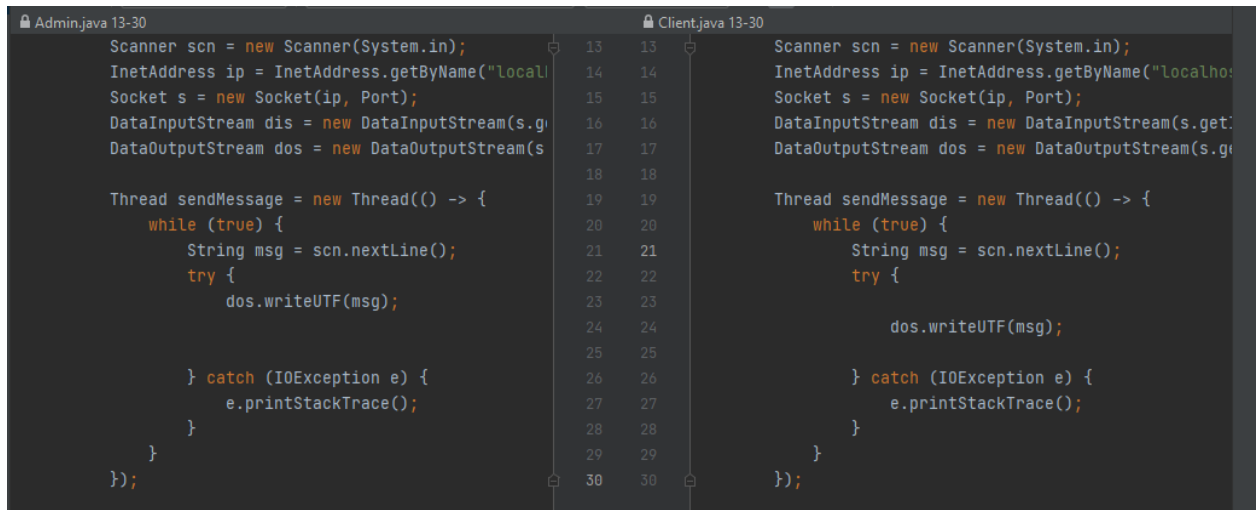
<https://github.com/diponsaha007/Code-Smell/blob/main/JavaObserverPattern/JAVAObserverPattern.png>



Overview of the bad smells detected:

- **Duplicate code fragment**

Duplicated blocks of code from the selected scope: the same file, same module, dependent modules, or the entire project.



- **Field may be 'final'**

Any fields which may safely be made final. A static field may be final if it is initialized in its declaration or in one static class initializer, but not both. A non-static field may be final if it is initialized in its declaration or in one non-static class initializer or in all constructors.

```
private int id;
```

- **Unnecessary conversion to string**

Any calls to static methods like `String.valueOf()` or `Integer.toString()` used in string concatenations and as arguments to the `print()` and `println()` methods of `java.io.PrintWriter` and `java.io.PrintStream`, the `append()` method of `java.lang.StringBuilder` and `java.lang.StringBuffer` or the `trace()`, `debug()`, `info()`, `warn()` and `error()` methods of `org.slf4j.Logger`. In these cases the conversion to string will be handled by the underlying library methods and an explicit call to `String.valueOf()` is not needed.

```
String UpdateMessage = "You subscribed stock " + StockName + ". Its current price: " + Float.toString(StockPrice) + ", number of available stocks: " + Integer.toString(StockCount);
```

- **Infinite loop statement**

for, while, or do statements which can only exit by throwing an exception. While such statements may be correct, they are often a symptom of coding errors.

```
while (true){...}
```

- **Declaration can have final modifiers**

Fields, methods or classes, found in the specified inspection scope, that may have a final modifier added to their declarations.

```
Socket s;
```

- **Unused declaration**

Classes, methods or fields in the specified inspection scope that are not used or not reachable from entry points. It also reports parameters that are not used by their methods and all method implementations/overrides and local variables that are declared but not used. Some unused members might not be reported during in-editor highlighting. Due to performance reasons, a non-private member is checked only when its name rarely occurs in the project.

```
Socket s;
```

- **Unnecessary boxing**

Explicit boxing, i.e. wrapping of primitive values in objects. Explicit manual boxing is unnecessary under Java 5 and newer, and can be safely removed.

```
StockCount = Integer.valueOf(st.nextToken());  
StockPrice = Float.valueOf(st.nextToken());
```

- **String concatenation in loop**

String concatenation in loops. As every String concatenation copies the whole String, usually it is preferable to replace it with explicit calls to `StringBuilder.append()` or `StringBuffer.append()`.

```
for (StockInterface si: Server.StockList)  
{  
    updated+=si.getName()+"\t\t\t\t"+ si.getCount() +"\t\t\t\t"+ si.getPrice()  
    +"\n";  
}
```

Code 2:

<https://github.com/diponsaha007/Code-Smell/tree/main/JavaMessenger/src/com/jetbrains>

This code implements a simple messenger using socket programming in Java where users can send messages among themselves. When we analyzed the code with CheckStyle and the default analyzer the following issues shown in the image arose.

Link to image for better quality:

Default analyzer:

<https://github.com/diponsaha007/Code-Smell/blob/main/JavaMessenger/Smells%20with%20intellij.png>

CheckStyle:

<https://github.com/diponsaha007/Code-Smell/blob/main/JavaMessenger/Smells%20with%20checkStyle.png>

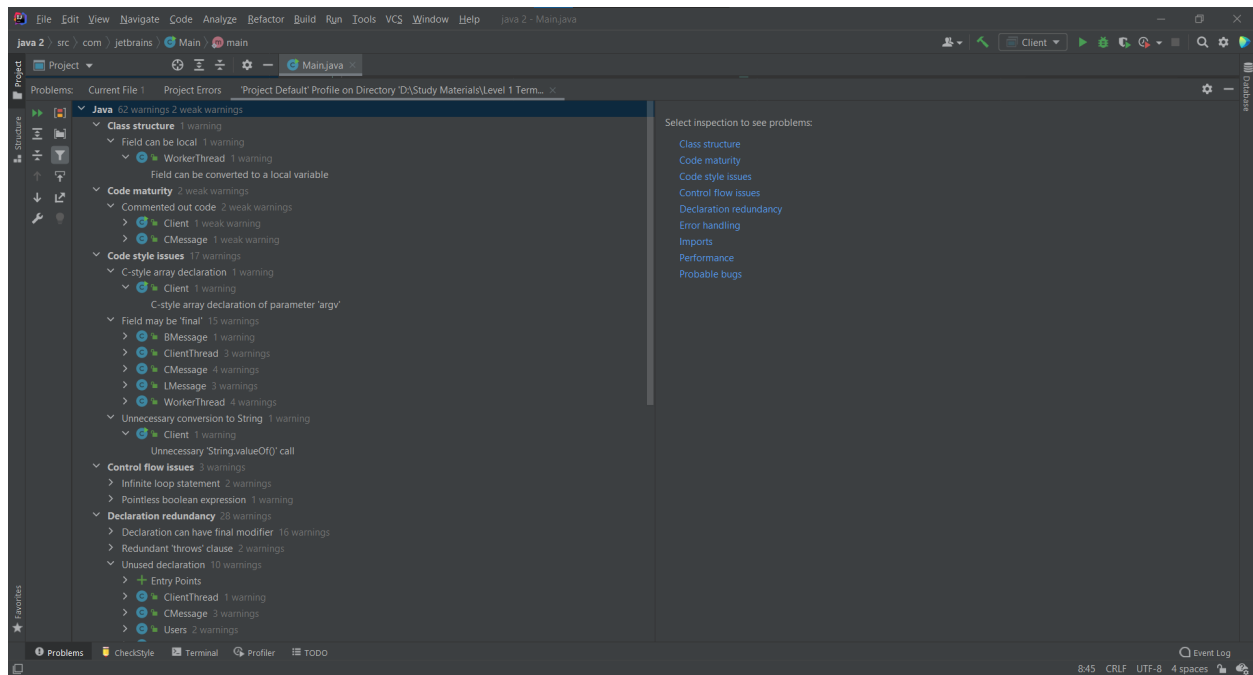


Fig (a) : Smells with default intellij analyzer

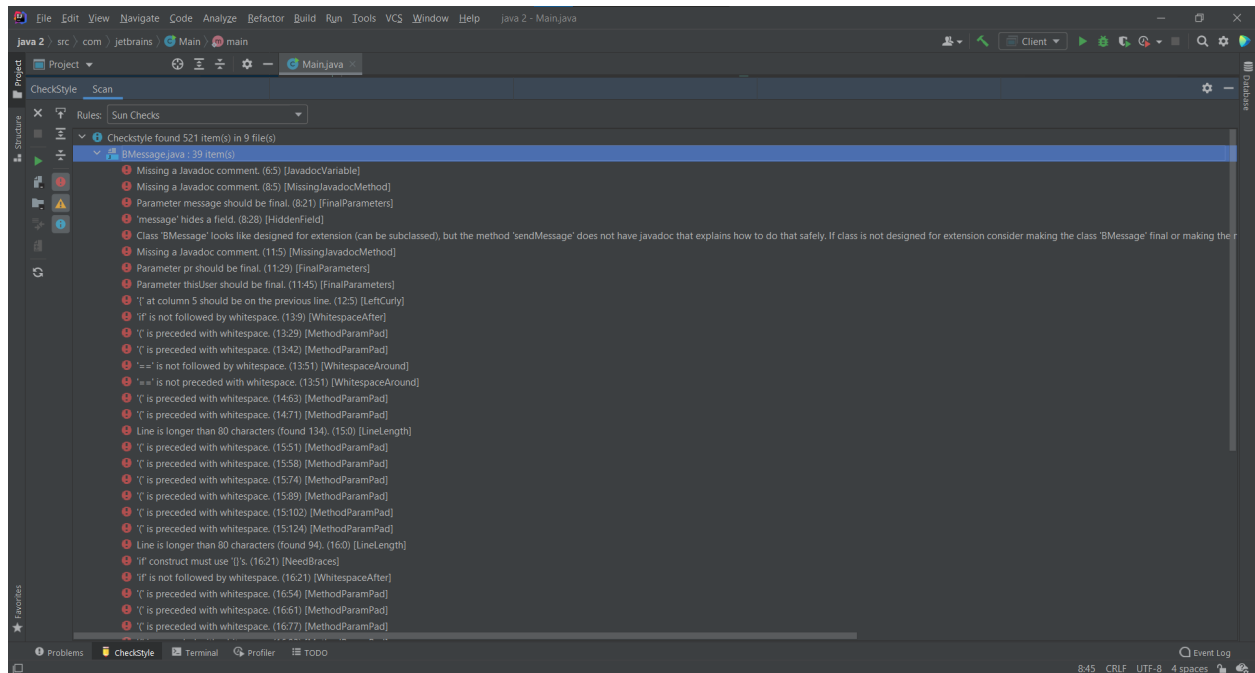


Fig (b) : Smells with CheckStyle-IDEA

Overview of the bad smells detected:

- **Field can be local** : The 'id' variable in the 'WorkerThread' class can be made local in the constructor.

```
public class WorkerThread implements Runnable {
    private int id ;
    public WorkerThread(Socket s, int id) throws Exception {
        this.id = id;
    }
}
```

- **C-style array declaration** : Though C-style and Java-style arrays are semantically identical. The "int array[]" syntax was only added to help C programmers get used to java. "int[] array" is much preferable, and less confusing.

```
String argv[]
```


- **Field may be final** : There are many fields in the code that can be made final. A final field cannot have its value changed. A final field must have an initial value assigned to it, and once set, the value cannot be changed again. Making a field final can make the code less error-prone.

```
public class BMessage {
    private String message;
}
```

```
public class ClientThread implements Runnable {
    private PrintWriter outToServer;
    private BufferedReader inFromServer;
    private Socket clientSocket;
}
```

- **Infinite loop statements** : There are infinite loops in “Client” and “Main” class which can’t be completed without throwing an exception.
- **Redundant ‘throws’ clause** : The ‘sendFile’ function inside the ‘CMessage’ class and in the ‘WorkerThread’ constructor the declared exception is never thrown. So redundant ‘throws’ should be removed.

```
public void sendFile(Users thisUser,String string) throws Exception{...}
```

```
public WorkerThread(Socket s, int id)throws Exception{...}
```

- **Unused declaration** : There are some unused declarations in the code that are never used. They should be removed.

```
String string = inFromServer.readLine ();
```

```
private String filename;
```

- **Empty 'catch' block** : There are some empty 'catch' blocks inside the 'run' method of 'ClientThread' and 'WorkerThread' class. The 'catch' parameter should be renamed to 'ignored' in case of nothing to be done.

```
catch (Exception e) {}
```

- **Unused imports** : Unused imports should be removed from the code.

```
import java.awt.*;
import java.net.ServerSocket;
import java.nio.file.Files;
```

- **String Concatenation inside loop** : Inside the 'action' method of 'SMessage' class 'String s' is concatenated inside a loop. Variable 's' should be converted into 'StringBuilder' as java strings are immutable.

```
s= s+ WorkerThread.getWorkerThreads ().get (i).getThisUser ().getUsername ()+ " ";
```

- **NullPointerException** : Method 'read' , 'write' and 'close' inside 'ClientThread' class may produce NullPointerException. So it must be handled with a 'try catch' block.

```
while ((count = in.read(bytes)) > 0) {
    out.write(bytes, 0, count);
}

out.close();
```

- **Result of method call ignored** : Result of 'BufferedInputStream.read()' is ignored. This is a bad coding style.

```
bis.read (contents, 0, size);
```

- **Long lines** : There are some lines longer than 80 characters. Long lines are hard to follow. So writing long lines should be avoided or these lines should be splitted.