INTRODUCTION

Nowadays, certainly every programmer is confronted with NullPointerExceptions in big Java Projects, whether it is for an enterprise or for private purposes. Not to mention even in small Java Projects they are also heavily present.

So what are those NullPointerExceptions? This thesis is going to attach importance to Java that is a concurrent, class-based, object-oriented programming language. We chose Java because NullPointerExceptions are more serious in this language than in others, e.g. Smalltalk. NullPointerException is a RuntimeException. In Java, an object reference can be assigned with a special null value. The exception is thrown when an application attempts to use an object reference that has the null value. (There are multiple ways this exception can be thrown, like: Calling an instance method on the object referred by a null reference; Accessing or modifying an instance field of the object referred by a null reference and so on.) In Java Projects developers always have to deal with a huge amount of references which means avoiding these NullPointerExceptions is as good as impossible.

On regular meetings among programmers they report what they have been doing and what they are planning to do for the next few weeks. But all too often it is stated that they are trying to fix bugs or have spent a lot of time fixing them. If there would be a way to minimize the time fixing exceptions and allow to work more efficiently, projects would progress much faster.

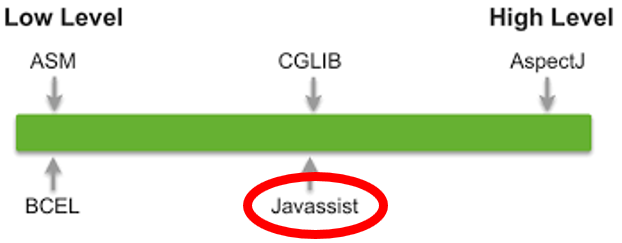
The main goal of the NullSpy application takes a step to that ideal vision. Anytime developers are facing a NullPointerException they don’t have to spent time on debugging finding where and why a reference was set to null. With NullSpy the exact location of the null assignment is shown next to the ordinary stack trace the Java virtual machine produces.

In this thesis it is explained how the goal mentioned above is achieved step by step, by using a class library Javassist (Java Programming Assistant) which allows us to deal with Java bytecode.

**Technical Background**

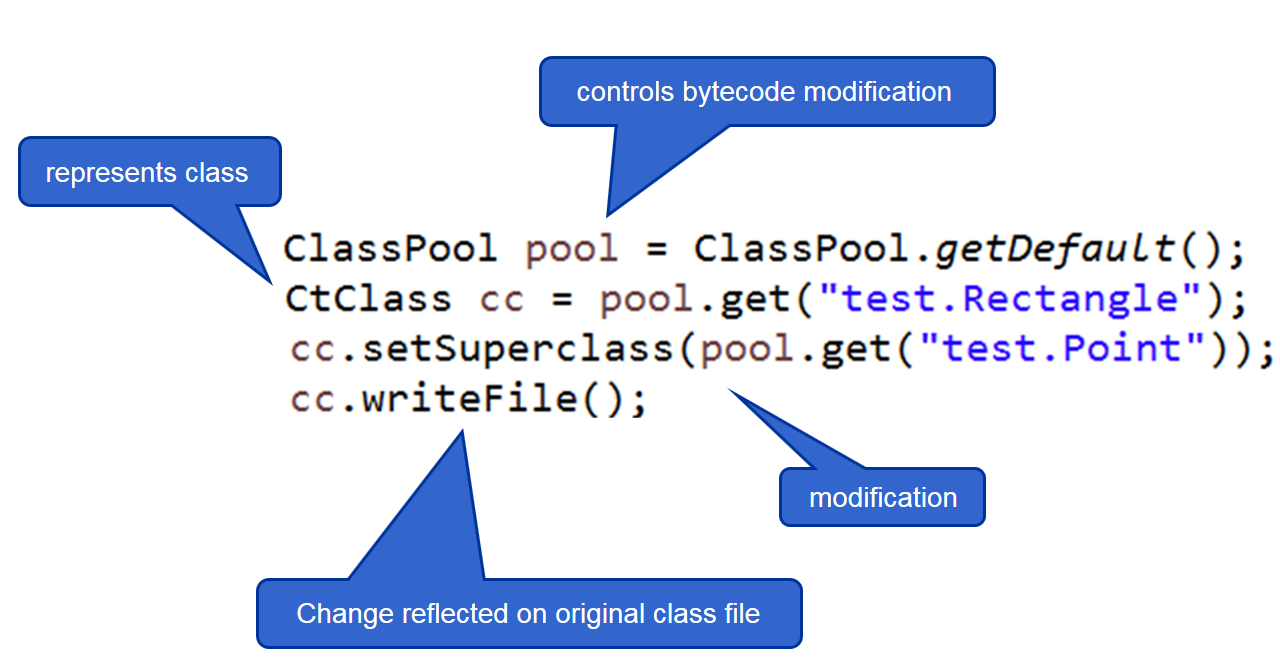
This chapter provides a short overview of works/technologies used in this project.

**Javassist**Javassist or Java Programming Assistant, a subproject of Jboss, is a class library which allows you to deal with Java bytecode. Since 1999 it is used as an engineering toolkit in a broad domain, and is still being extended by Shigeru Chiba. It enables developers to manipulate Java bytecode in a simplified way like defining a new class at runtime or modifying a class file when it is loaded by the JVM. All manipulations are performed at load-time through a provided class loader.

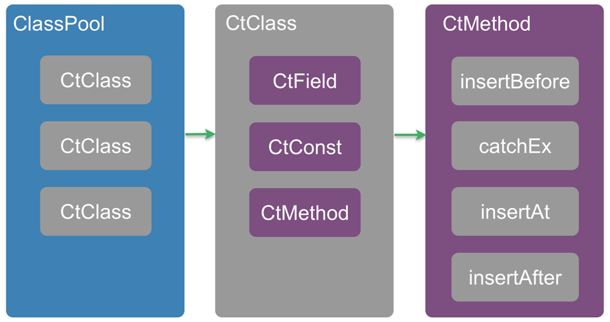


Unlike many other libraries Javassist offers two levels of API: source level and bytecode level. Using the source-level API, the user can edit a class file without any familiarity with the specifications of the Java bytecode. Only knowing the Java language is enough because the API is designed only with the vocabulary of Java. On this level the programmer just has to write normal source code and Javassist compiles it automatically. The bytecode level allows the user to modify classes directly in binary form like other editors, e.g. ASM.

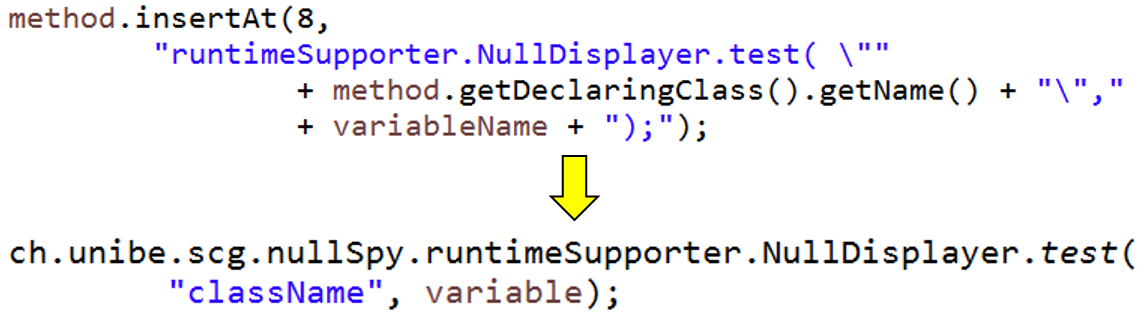
At this point, let us look at a small example to give you an idea how the bytecode manipulation works.



First a ClassPool object is obtained that controls bytecode modification with Javassist. With the ClassPool a class file can be read on demand for constructing a CtClass object. The class CtClass (compile-time class) is just an abstract representation of a class file which means all manipulations are performed on the CtClass. With the method invocation “get()” on ClassPoola reference to the class file “test.Rectangle” is obtained. In this example the superclass of “test.Rectangle” is just changed to “test.Point”. If the changes are done, the method call “writeFile()” on CtClass is necessary to make sure that the changes are reflected on the original class file.



This image gives you an understanding/overview how the main part of bytecode manipulation with Javassist is built up. The ClassPool is nothing else than a container of multiple CtClasses. As described before CtClass is just the abstract representation of a class file on which modifications are done. Like typical classes, it can hold several compile-time fields, constants or methods. While speaking about bytecode manipulation all the time, nothing but editing method is manly meant. It is possible to insert source code at the beginning of the method body, at the end or at a specific line. Next to these options a “catchBlock” can even be added.



**JHotDraw**To check whether the logic of the bytecode manipulation in this project NullSpy is working as desired, we had to perform them on a large working project. Thanks to Nevena Milojkovic and her experience with the combination Javassist and JHotDraw we as well decided to test NullSpy on the project JHotDraw.

It is an open-source Java GUI framework for technical and structured Graphics. Its original authors have been Erich Gamma and Thomas Eggenschwiler.

JAD  
JAva Decompiler is a decompiler for the Java programming language. A short explanation what a decompiler is: a computer program that takes an executable file as input, and attempts to create a high level, compatible source file that does the same thing. So it is used in software reverse engineering.

JAD is used in NullSpy since after running NullSpy on JHotDraw only modificated bytecodes are available. But to simplify the check if the modification by Javassist, e.g. inserting source code, has succeeded, a decompiler is needed.

**NullSpy**

As earlier explained in the introduction, this project is about providing the user with additional stack trace where the origin of a NullPointerException is actually rooted. Briefly worded, it shows the developer the exact location of where a method receiver, which causes the NPE, was assigned to null.

This is the main chapter of the thesis. Here we would like to give you a short insight of how we managed to successfully implement the core of the project NullSpy. Next to how it is built up, we will also let you know what challenges we were encountering during the implementation and about the limitations we planned for future work (ref chapter future work).

**High level overview/Rough Scheme**

The general approach of NullSpy is to statically analyze and add additional bytecode to a project. After reading the “Technical Background” section you should be more familiar with how bytecode manipulation with Javassist works.

What NullSpy first does is loading the project you want it to be able to track the null assignment if a NullPointerException is thrown. By loading the project to NullSpy, the compiled class files of the project are addressed only, which means the project itself does not have to be imported to the programming environment, e.g. Eclipse. Simultaneously at load time each class file is modified with help of Javassist; In what way will be discussed in the following section “Low level overview”.

Once the project modification is done it is stored in a destination folder that the user has chosen before. This means after the changes there will be another version of the project which can do additional stuff like tracking the null assignment. Because only the class files are accessed previously, the result which is stored in the destination folder is as expected only the modified bytecode.

The reason why we store the modified project in another directory than overwriting the existing one is because we would like to only involve Javassist library in the loading and editing part. This means the execution of the altered project does not include Javassist. This way the user does not have to download Javassist and integrate it to the project.

How do we check whether the instrumentation worked and the project really tracks the null assignment? The answer is wrapping the modified project into a jar file with which the modified project can be executed in the terminal or in Eclipse.

pic

**Low level overview**

NullSpy only handles the unhandled NullPointerException, meaning the main method of a project is wrapped with a \textit{''catchBlock''} where the NPE is treated. The idea in this \textit{''catchBlock''} is to load the data about the assignments and about the method receiver to compare them with each other. With the stacktrace the line number and the name of the Java class of where the NPE occured can be figured out. We look them up in the method receiver information and get some hold points with which we look up in the assignment information. If there is a match we provide the user with an additional stacktrace next to the usual one. Thanks to that link that points directly to the NPE root, debugging will be easier or it can even be skipped or avoided completely. To be able to provide this function data about the method receiver and variable assignment have to be collected initially.

So we first load the project which should be modified and then go through all class files of it. A NullPointerException can only be thrown if a method call was performed on a method receiver which is null. That means we have to gather information about the method receiver. Next to this also information about the variable assignments has to be collected to know the exact location of the null assignment. The main idea behind NullSpy actually is to get information comparing those together when a NullPointerException occurs and if there is a hit during the comparison the location of the null assignment can be obtained easily.

The implementation behind the NullSpy concept will be explained more fully.

Method Receiver Data Collection

Unfortunately, Javassist does not provide the function to directly get the method receiver. We got a suggestion to use AST (Abstract Syntax Tree footnote wiki) to get it but we decided to not go deeper into this and implement our own algorithm.

The algorithm contains following steps (abstract):  
1. getting pc-interval of method invocation  
2. storing all possible method receiver interval within the interval of step 1 into an ArrayList  
3. Getting the number of parameters, the method invocation takes  
4. Traversing back the ArrayList the amount of parameters obtained in step 3  
5. Result: method receiver interval  
6. Store variable name, type etc. into an external csv file

In step 1 we had big troubles getting the right interval of the method receiver because only by statically analyzing the bytecode it is unapparent where the method receiver is situated exactly. But more about the challenges you will learn more in the chapter [CHAPTER REF]

Statically analyzing bytecode for method receiver means looking for certain opcodes which matches all opcodes that matches with the regex “invoke.\*”. There are exactly five kinds of invocation opcode: invokedynamic, invokeinterface, invokespecial, invokestatic, invokevirtual. The bytecode instruction invokedynamic facilitates the dynamic-typed languages(footnote) through dynamic method invocation. In our case it can be ignored because NullSpy only supports the static-typed language(footnote) Java.

http://www.javaworld.com/article/2860079/learn-java/invokedynamic-101.html

In case of the \textit{invokestatic} instruction we do not have a method receiver. That is why NullSpy treats it extraordinary like ignoring it completely or wrap it as a possible method receiver when it is actually a parameter of a method invocation. In all other cases we normally use the algorithm to get the method receiver. The gathered data about method receivers are stored in an external csv file due to performance and overhead minimization.

Variable Data Collection

While going through the bytecode attention is paid to some opcodes (footnote). Right after each keyword that indicates a variable assignment we insert some bytecode. The inserted code represents a test method which tests whether the value of the assigned variable is null or not and store some information about it. Unlike in getting information about the method receiver in [SUBSECTION] the data about the variables are stored in a HashMap.

What kind of opcodes were NullSpy looking for? For instance or class/static variables the bytecode instruction “putfield” and “pustatic” were essential, for local variables the important opcodes were those which matches the regex “astore.\*”. Due to different types of variables and the limitation of Javassist, gathering information about them was performed differently. NullSpy first collects data of fields and then that of local variables. Again getting the necessary data about the variables we encountered many difficulties which will be discussed in the subsection [REFERENCE].

*Local variable*

Unfortunately, Javassist does not provide any support for gaining information about local variables that is why getting the needed data we had to understand how bytecode is constructed. At this point we would like to give you a small bytecode introduction.

pic

If a local variable is created, the value assigned to it is pushed onto the operand stack. With the bytecode instruction “.\*store.\*” the local variable is popped from the operand stack and stored into a local variable array slot. In which slot it is stored can be extracted from the instruction. Storing opcodes is composed of one, or in some cases two bytes. There are reserved machine commands for the first four local variables, index-linked from 0 to 3 and each of them contains one byte (astore\_0…3). If there is no slot number visible in the instruction, it indicates that the slot number is stored in the second byte from where it can be extracted. Next to storing the local variable loading it from the local variable is possible to, but only with the local variable slot number.

With this short introduction understanding the local variable table should be easier. Each method of a class file contains a local variable table (see figure x) with which many information can be read out of it, e.g. the lifespan of the local variable, what it is called, in which slot it is stored and what type it has.

Pic

We had to pay attention to be sure to get the right local variable. Every time when we bumped into the opcode “.\*store.\*” we could only get its slot and the pc where it is situated in the bytecode sequence. In the earlier paragraph the lifespan of the local variable was mentioned, the importance behind this is as soon as the lifespan of one ends, the slot can be used by the next instantiated local variable. This way, the local variable table could contain multiple entries with the same local variable slot.

Pic

After extracting the slot of the local variable we will get the first local variable table entry which contains that slot. If the pc of the local variable assignment is not included in the lifespan-pc-interval of the entry, the next entry with the same slot will be checked until both criteria (slot and pc) fits. Once those criteria are met we can be positive about having got the right local variable table entry to extract the information needed.

Next to the local variable table each methods of a class file also holds another attribute called line number attribute. This is just the mapping table from pc to source code line number. Since encountering the storing keyword the pc is available with it the line number can be easily obtained.

*Instance or class/static variables (Fields)*

Although Javassist does not support the access to local variable it provides a way for fields. Javassist allows modifying an expression in a method body with the class \textit{Javassist.expt.ExprEditor}. In our case we only want to extract some information about the fields instead of any modification, nevertheless this class can be used appropriately. What it does is to scan the bytecodes on instructions like ‘’getfield’’ or ‘’getstatic’’.

There are only two instructions that indicates an access to a field, but there are actually many different types. To get the meaning of different types see the following list:

(|…|: put value on operand stack for assigning to a field)  
1. this, |…|, putfield   
2. |…|, putstatic   
3. aload, |…|, putfield  
4. this, (getfield)+, |…|, putfield  
5. getstatic, (getfield)\*, |…|, putfield

* Code ThesisCodeExamples fields

Depending on the category the field belongs, different kind of information is stored. For fields 1-2 following information is needed and stored:

fieldID (which is field or localVar)  
fieldname  
fieldType  
fieldDeclaringClassName (in which class it is instantiated)  
isFieldStatic  
fieldLineNumber  
startPc (nonstatic: this; static: value loading bytecode for assigning to the field)  
storePc (putfield/putstatic)  
afterStorePc (right before this pc additional bytecode is inserted)  
classWhereFieldIsAccessed (in this case the same as fieldDeclaringClassName)  
behavior (method where it is accessed)  
indirectVariable (null - explained immediately)

In NullSpy we call these fields \textit{‘’directFields’’}. After reading \textit{‘’direct’’} you surely think of there must be something like \textit{‘’indirectFields’’}, and that is right. As \textit{indirectFields} the fields 3-4 are numbered among. Everything before the instruction \textit{putfield} is termed as \textit{‘’indirectVariable’’}. For those fields more data are needed to store for their identification. Nearly the same as above, except:

classWhereFieldIsAccessed (can be differend than fieldDeclaringClassName)  
indirectVariable:  
 indirectVariableName  
 indirectVariableType  
 indirectVariableDeclaringClassName  
 isIndirectVariableStatic  
 indirectVariableOpcode (to distinguisch if it is a localVariable or a field)

Bytecode adaptation

Each time encountering a variable assignment we first extract the needed data and directly after this we adapt the class file by adding extra bytecode right after the assignment to check whether the variable is null or not. If it is null, the information collected before is stored either into a \textit{‘’localVariableMap’’} or into a \textit{‘’fieldMap’’}.

The added bytecode represents a static method of a class named \textit{‘’VariableTester’’} which will be added to the modified project after class file modification is done. Depending on the kind of variable analyzed at the moment different bytecode is constructed, meaning different method with different parameter/variable data.

There were again some challenges we had to get over, like getting the wanted data after an instrumentation and entering the additional code in the right position of the bytecode sequence [REFERENCE].

Once we have gone through the bytecode of a Java class, the modified class files are stored in a destination directory as mentioned in [REFERENCE]. Next to the instrumentation supplementary supporter classes are added to the project. The most important ones are \textit{‘’VariableTester’’} which tests whether a variable is null or not and \textit{‘’NullDisplayer’’} which matches data and prints the location of a null assignment when a NullPointerException is thrown.

To do some information comparisons we first have to obtain some data about where a variable is actually set to null. To be able to get those data we load the class files with Javassist and then create CtClasses out of them, just as in chapter “Related Work” described. With the CtClass object we finally have access to look at and editing the bytecode of a loaded class file, which means that we for example adding code to the loaded Java class. That is exactly what we were doing.

For each class file we obtained and went through its bytecode and looked for opcodes like  
🡪 aastore  
astore  
putfield  
putstatic  
.

* BILD

All these opcodes represent some variable assignments. After bumping into one of those enumerated keywords the test method of the supporter class we added to the project is inserted right after the assignment. The idea behind this insertion is to collect some information about the variable which was just set to a value. We only store them if the value is the culprit null. So what kind of information are stored? The role of those information is to allow the identification of a variable and to get the exact location of its instantiation. That means the name, type, line number, the class name or method name and lot more. It depends on what kind of variable the examined one is, it could be a local variable or a class member, static or non-static one, a variable of a superclass etc.

**Challenges**

In this section we could like to give you an understanding of few difficulties occurred during the implementation of NullSpy.

*Method Receiver Data Collection*

Aforementioned in subsection \textit{Method Receiver Data Collection} \ref{subsec:methodReceiver} we were encountered with a persistent problem, namely getting the pc-interval of the method receiver when the interval covers multiple lines in source code. In many development environment the written code can be formatted automatically and as well manually. See following figures:

Pic 4

Figure shows a normal method invocation which is split into two lines in source code. This cannot be figured out by just looking at the bytecode therefore the line number attribute (REFERENCE) has to be consulted too. There is no method receiver in the figure because the method receiver would contain a method invocation itself (\textit{''Iconkit.instance()''}, what is not supported in NullSpy. Another method receivers NullSpy does not support are elements of collections due to complex structures that can be stored in the collections. But still it is a good example to show you how the bytecode and the line number attribute for this looks like. Applying the algorithm (REFERENCE) for gathering data about method receivers in this situation looks as follows:

1. pc-interval: 0-11 (cut everything off what has nothing to do with the method invocation)  
2. ArrayList: [0,[3,4],[7,8]]  
3. Number of parameters: 2  
4. Traversing back the amount of parameters: [7,8] (1) 🡪 [3,4] (2)  
5. Result: 0 (invokestatic) = possible method receiver  
6. Not storing it because NullSpy does not support this kind of method receiver

So how did we get the pc-interval? Just looking at bytecode is not enough, so we are looking at the line number attribute. Line number \textit{31} is listed twice, this indicates that the method invocation is split into multiple line in source code. The former declares the starting point of the interval and the latter the end of it. How to get the pc-interval will be presented shortly. But first take a look at more examples:

Pic 3

Special about this example is that the multiple line interval does not start with the smaller line number and end with the bigger one, apart from that it is alternated stored in the line number attribute.

Pic 2

Neither in bytecode nor in line number attribute we can extract the exact interval of method invocations. Generally, we cannot distinguish if the interval represents a method invocation or a loop. Another interesting thing is that in the figure multiple corresponding pcs are included.

Pic 5

By using the algorithm we only get the interval from pc 19-32, but actually pc 18 also belongs to the interval. That the missing pc is missing will be detected in step 4&5 when traversing back the ArrayList by the number of parameters. If it is not a static method invocation and there is not enough possible method receivers stored in that ArrayList to traverse back, the missing pc is added in retrospect.

* Pseudocode

This pseudocode will give you the main or a very rough idea how the invocation interval can be gained. If you are interested in knowing the exact way, please see the implementation of the class \textit{ch.unibe.scg.nullSpy.instrumentator.controller.methodInvocation.MultipleLineManager} and \textit{ch.unibe.scg.nullSpy.instrumentator.controller.methodInvocation.MethodInvocationAnalyzer}.

Field Data Collection

We previously mentioned that NullSpy first collects data about fields and we met some difficulties regarding this. By using the method \textit{loopBody} from class \textit{Javassist.expt.ExprEditor} [REFERENCE] for gathering information caused those troubles because it loops through the GIVEN \textit{methodInfo} which contains the bytecode sequence of a method. Even after finding a field assignment and adding some additional bytecode to it, the method still iterates through the unchanged \textit{methodInfo} it got as parameter without the additional bytecode. Getting the right startPc of a field assignment is not a straight forward process as it may seem.

After entered extra instructions the method \textit{loopBody} iterates onwards until it finds a key which indicates an access to a field. Normally, Javassist provides a method with which the starting pc of the field access can be find out easily, but in our case not due to bytecode alternation. This method returns the starting pc of the field as if there was no changes, however it actually has to return a bigger starting pc than it does. The starting pc is needed to distinguish in what category the field has to be assigned to (REFERENCE).

To find the right starting pc of the field assignment we compare the starting pc which a method of Javassist returns with the \textit{afterStorePc} of the previously found field assignment. Every time an assignment is found we store it as a reference for obtaining right data. Even for the storing pc of an assignment (\textit{''put.\* ''} the last found field is essential. If there is more interest how those pcs are obtained, please see the class \textit{ch.unibe.scg.nullSpy.instrumentator.controller.FieldAnalyzer}.

Bytecode adaptation

The reason why we decided to use Javassist for our NullSpy project is because the thought of only using the source-level API [REFERENCE] to implement everything. It would have been much easier to only operate at source-level instead of learning how to read bytecode or extract data from it or enter extra code into it, yet at the end we still had to do everything at bytecode-level.

One big problem encountered while inserting the test method between an assignment and a closing bracket \textit{''}''}. We tried to insert additional code as shown in figure [2.4.] by specifying the exact line number where it should be entered in source code. Unfortunately Javassist first checks the specified line whether it contains some code (only symbols excluded). If there is no code at that line it computes the next line that contains some and inserts the test method right before it. Please visualize a situation where for example a local variable is created/instantiated at the end of a \textit{if-body}. In this situation Javassist adds the extra code right before the next code line which is outside the existing scope of the local variable.

Pic bytecodeExample\_1

Pic bytecodeExample\_2

That is why changed the way to insert the test method at bytecode-level. Like this we first have to build up the bytecode sequence and then enter it before a specific pc. Please take a look at the class \textit{ch.unibe.scg.nullSpy.instrumentator.controller.ByteCodeAdapter} how the bytecode sequence is created.

------

There are of course many other small problems during the implementation of NullSpy but these mentioned are the most troublesome ones.

**General Problem/Challenge**

At the beginning of a project or even before starting it,

In this section I want to introduce some hurdles I had to overcome during the implementation of th

Get used to how Javassist works, tutorial, forgot things actually could take in, javassist problem not updating codeAttr after insertion after field assignments

Architect of project, first own project, without thinking started to code, naming, complex methods without any documentations, refactored many thimes, problem with concentrating on one task -> abschweifen, try to solve many things in parallel

**Limitations**

**Validation**

**Conclusion and Future work**