

Code Clone Refactoring in C# with Lambda Expressions

Takuto Kawamoto

Yoshiki Higo*

tk-kawam@ist.osaka-u.ac.jp

higo@ist.osaka-u.ac.jp

the University of Osaka

Suita, Osaka, Japan

Abstract

“Extract Method” refactoring is a technique for consolidating code clones. Parameterization approaches are used to extract a single method from multiple code clones that contain differences. This approach parameterizes expressions and behaviors within a method. In particular, behavior parameterization has been extensively studied in Java programs, but little research has been conducted on other programming languages.

Lambda expressions can be used to parameterize behaviors, but the specifications of each programming language significantly affect the applicability of this technique. Therefore, the optimal “Extract Method” approach may vary depending on the programming language.

In this study, we propose a C#-specific technique that uses lambda expressions to analyze and consolidate code clones. We evaluated our proposed method by applying it to code clones detected by the NiCad clone detector and measuring how many of them could be successfully consolidated.

In total, 2,217 clone pairs from 22 projects were included in our evaluation. For the clone pairs determined to be refactorable, we also attempted refactoring actually. The proposed approach determined that 35.0% of all clone pairs were suitable for refactoring. Among these, 28.9% were successfully refactored.

Keywords

Refactoring, Code clone, Lambda expression

ACM Reference Format:

Takuto Kawamoto and Yoshiki Higo. 2025. Code Clone Refactoring in C# with Lambda Expressions. In *2025 The 7th World Symposium on Software Engineering (WSSE 2025), October 24–26, 2025, Okayama, Japan*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3779657.3779670>

1 INTRODUCTION

Code clones are similar code fragments within the source code. Pairs of such code fragments are called clone pairs [5]. Refactoring code clones reduces the size of the code base, helps prevent inconsistent modifications to code clones [15], and lowers maintenance costs [8].

“Extract Method” is a refactoring technique that consolidates code clones into a single method [2]. Since clones often contain minor

differences, this technique parameterizes the differing portions to unify the similar fragments [14].

When differences occur at the statement level, expression parameterization may be insufficient to extract them. This is because expressions are smaller units than statements. One approach to dealing with statement-level differences is to move the differing statements outside the consolidated code fragment. However, this reordering can alter program behavior due to side effects, requiring careful analysis to ensure correctness [9]. Previous studies have investigated the feasibility of the Extract Method refactoring through statement reordering and expression parameterization, using data flow analysis [9].

Some code clones cannot be consolidated simply by reordering statements. Alternatively, refactoring approaches that avoid statement reordering may instead parameterize the statements themselves. Parameterization of statements is called behavior parameterization.

Parameterized expressions are evaluated when the extracted method is invoked, which may alter the original timing of execution [10]. In contrast, parameterized behavior does not change the execution order. Lambda expressions are used to parameterize behaviors. Language-specific features significantly affect how behavior can be parameterized using lambda expressions. Existing research [10] has examined behavior parameterization using lambda expressions in Java. However, it has not addressed its applicability to other programming languages.

This study proposes techniques for behavior parameterization using lambda expressions in C#. Using our proposed method, we analyzed 2,217 clone pairs collected from 22 projects on GitHub to assess their refactorability and performed refactoring on those deemed suitable.

The contributions of this study are as follows:

- We provide useful techniques for Extract Method refactoring based on behavior parameterization.
- We demonstrate the effectiveness of behavior parameterization in Extract Method refactoring. Our results suggest that it broadens the range of code clones that are refactorable.



This work is licensed under a Creative Commons Attribution 4.0 International License.

WSSE 2025, Okayama, Japan

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1577-8/2025/10

<https://doi.org/10.1145/3779657.3779670>

This paper is organized as follows: Section 2 introduces background and preparatory concepts for Extract Method refactoring. Section 3 presents the proposed approach. Section 4 describes the evaluation and discusses the results. Section 5 discusses the results. Section 6 considers threats to the validity of the study. Section 7 summarizes the conclusions.

2 PREPARATION

2.1 Code clone

Code clones are similar code fragments within the source code. Pairs of code fragments that are similar to each other are called clone pairs [5]. Inconsistent changes between code clones can result in missed updates, potentially leading to bugs [16]. Based on their degree of similarity, code clones are classified into four types [14].

- Type-1 The code fragments are entirely identical except for the changes that may exist in whitespace and comments.
- Type-2 The structure of the code fragments is the same, but the identifier names, types, whitespace, and comments may differ.
- Type-3 In addition to changes in identifier names, types, whitespace, and comments, some statements may have been modified or deleted, and others may have been newly added.
- Type-4 The code fragments differ syntactically, but are semantically equivalent.

2.2 Refactoring code clones

The “Extract Method” refactoring is used to consolidate code clones into a single method. Parameters can be introduced to apply the “Extract Method” refactoring to Type-1 and Type-2 code clones [13]. The parameters are evaluated when the extracted method is called, which may differ from the timing in the original source code. If the parameterized expressions include method calls or object instantiations, which may have side effects, the difference in evaluation order may affect the program’s behavior [10].

Prior studies have found that most syntactically similar code clones are Type-2 or Type-3, while Type-1 clones are relatively rare [9]. Therefore, refactoring techniques that can consolidate Type-2 and Type-3 code clones play an important role.

2.3 “Extract Method” with behavior parameterization

Some code clones cannot be consolidated into a single method, even with expression parameterization. Figure 1 shows such an example. In these code fragments, the highlighted parts differ, which classifies them as Type-3 code clones.

These code fragments cannot be consolidated into a single method with expression parameterization. They cannot be parameterized because the differing parts are statements rather than expressions. A combined behavior and expression parameterization approach has been proposed [10]. Function objects, such as lambda expressions, defer evaluation until runtime, thereby preserving program behavior.

In some programming languages, lambda expressions can capture variables and functions from the scope in which the code clone appears. For example, lambda expressions can be used for behavior parameterization in Java, C#, and Python. In JavaScript, arrow functions serve the same purpose. For simplicity, we collectively refer to such function objects, including arrow functions, as lambda expressions throughout this study.

Figure 2 shows an example of method refactoring using behavior parameterization, based on the code in Figure 1. The common parts between the code fragments are highlighted in orange. Furthermore, the differing parts are highlighted in distinct colors. In both code fragments, the original code has been rewritten to call the newly

```

1 StreamReader reader = new StreamReader("Path");
2 while (reader.Peek() != -1) {
3     string line = reader.ReadLine();
4     Console.WriteLine(line);
5 }
6 reader.Close();

1 StreamReader reader = new StreamReader("OtherPath");
2 while (reader.Peek() != -1) {
3     string line = reader.ReadLine();
4     for (int i = 0; i < line.Length; i++) {
5         Console.WriteLine(line[i]);
6     }
7 }
8 reader.Close();

```

Figure 1: Examples difficult to extract via expression parameterization

```

1 Subroutine("Path", (line) => {
2     Console.WriteLine(line);
3 });
4
5 Subroutine("OtherPath", (line) => {
6     for (int i = 0; i < line.Length; i++) {
7         Console.WriteLine(line[i]);
8     }
9 });

```

Figure 2: Example of refactoring Figure 1

generated method `Subroutine`, where the second argument is a lambda expression representing the differing parts of the original source code. The definition of `Subroutine` is shown in Figure 3. The definition contains the common parts of both code fragments and invokes to the lambda expressions passed as arguments to replace the differing parts.

The lambda expression may be executed multiple times during a single run of the code fragment, outputting a string to the console each time. This behavior cannot be achieved with simple expression parameterization.

Previous research [10] proposed a technique for “Extract Method” refactoring using lambda expressions in Java. However, this technique cannot be directly applied to code clones in other programming languages. This is because the proposed technique is specific to the Java language. For example, this technique accounts for checked exceptions and final variables [12], both of which are not present in C#. Instead, C# includes features such as properties, which require careful consideration when performing expression parameterization. According to a recent review [11], apart from the technique proposed for Java, no other refactoring techniques using lambda expressions have been proposed. Therefore, this study proposes a technique for “Extract Method” refactoring using lambda expressions in C#. Additionally, we examine the applicability of the proposed technique to other programming languages.

For simplicity, we refer to “Extract Method” refactoring using behavior parameterization simply as “Extract Method” refactoring, and to the extracted method as a subroutine.

```

1 delegate void LambdaType(string v);
2 void Subroutine(string v1, LambdaType lambda1) {
3     StreamReader reader = new StreamReader(v1);
4     while (reader.Peek() != -1) {
5         string line = reader.ReadLine();
6         lambda1.Invoke(line);
7     }
8     reader.Close();
9 }
```

Figure 3: Method consolidated with Figure 2

3 OUR APPROACH

3.1 Overview

Our approach takes the source code and a Type-1, Type-2, or Type-3 clone pair as input, and applies “Extract Method” refactoring to produce the refactored source code. The code fragments in the clone pair are limited to sequences of one or more consecutive statements.

Figure 4 provides an overview of our approach. The approach identifies the portions of the given clone pair that should be parameterized using lambda expressions and then generates a subroutine. Specifically, the subroutine generation process is divided into five steps:

- (1) Statement mapping,
- (2) Variable mapping,
- (3) Exit block checking,
- (4) Argument determination, and
- (5) Source code generation.

In the remainder of this section, we explain each step in detail.

3.2 Statement mapping

First, this approach maps the statements in the input clone pair. It scans the code fragments from the beginning and identifies matching statement pairs, while ignoring differences in identifiers.

Through this mapping, the statements in each fragment are classified as either matching or non-matching with those in the other fragment. During code generation, non-matching statements are parameterized using lambda expressions, while matching ones are directly included in the subroutine.

We apply the Longest Common Subsequence (LCS) algorithm to perform this mapping. In our approach, the LCS algorithm is used to obtain the longest common subsequence of statements, rather than substrings.

Typically, the LCS algorithm is applied to strings to find the longest common subsequence of characters. It also requires a comparison function for the sequence elements. When applied to strings, these elements are characters, and thus a simple character-level comparison suffices. However, in our case, the elements are program statements, which require a comparison function capable of evaluating their similarity. The following describes the statement comparison function used in our approach.

3.2.1 Statement comparator. The LCS algorithm requires a comparison function to determine equivalence. We consider statements equivalent if they are so similar that they do not require behavior parameterization. Our approach uses a comparator that considers statements to match if they differ only in whitespace, type names, or identifiers.

Ultimately, statement comparison is performed using a simple string comparison. To allow for specific differences, the statements to be compared are normalized. In other words, the comparison function normalizes both statements and then simply compares the normalized versions.

Whitespace and comments can easily be ignored in the abstract syntax tree obtained from the source code. Furthermore, for identifier normalization, our approach masks identifier names. For example, constants and variables are masked because they can often be parameterized.

There are constraints on expressions that can be parameterized, so identifier normalization must be handled carefully. Expressions in the source code are further categorized into finer-grained types, and some grammar rules restrict which types can be used in specific contexts. For example, expressions used in case labels of `switch` statements must be constant expressions whose values are determined at compile time. Since the parameters are clearly not constant expressions, the identifier names appearing within the `switch` case labels must not be masked. We conducted a detailed analysis of the grammar rules of C# and restricted identifier parameterization to those that meet the following criteria:

- (1) Member identifiers,
- (2) Type identifiers,
- (3) Method and constructor identifiers,
- (4) Identifiers in case labels of `switch` statements, and
- (5) Identifiers used in pattern matching.

3.2.2 Nested statements matching. For statements containing multiple inner statements—such as `for` or `while` loops—simply determining whether they match is not sufficient. We refer to the statements that contain inner statements as nested statements.

Partial behavior parameterization may be desirable within a `for` loop. For example, most of the inner statements of a `for` loop may match, while some inner statements may differ.

For nested statements, the comparison is first performed by ignoring their inner statements. The LCS algorithm is then applied to the inner statement sequences of the nested statements that are determined to match. This process of applying the LCS algorithm is repeated recursively according to the nesting depth.

The way inner statements are ignored varies depending on the type of statement. For example, in the case of an `if` statement, only the conditional expression is compared. For a `for` statement, the statement comparator ignores inner statements and considers only the loop variable declaration, the loop continuation condition, and the update statement.

3.3 Variable mapping

Second, in expression parameterization, we map corresponding expressions within matching statements. These expressions may include local variables, members, fields, and literals. However, in this study, we exclude expressions that contain operators.

From the matching statements identified in the statement mapping step, we extract parameterizable expressions that occur at the same positions in both statements. Our approach then assigns parameters to these expression pairs. Identical expression pairs share the same parameter, while distinct pairs receive different parameters.

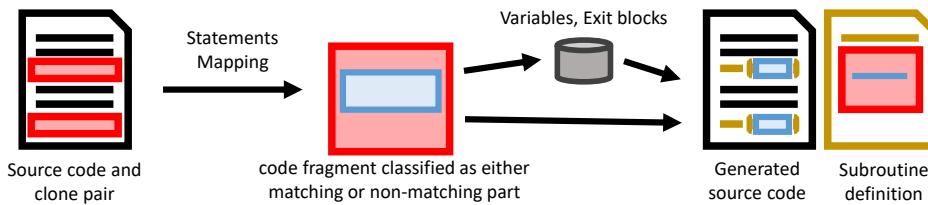


Figure 4: Overview of our approach

```

1 List<int> list = [1, 2, 3];
2 list.Add(5);
3 list.Insert(2, 6);
4 list.RemoveAt(0);
5 list.Add(10);

1 HashSet<int> set = [4, 3, 6];
2 set.Add(5);
3 set.Overlaps([1, 2, 3]);
4 set.Add(10);
  
```

Figure 5: Code clones requiring generics

```

1 List<int> list = [1, 2, 3];
2 Subroutine<List<int>>(list, (c) => {
3     c.Insert(2, 6);
4     c.RemoveAt(0);
5 });

1 HashSet<int> set = [4, 3, 6];
2 Subroutine<HashSet<int>>(set, (c) => {
3     c.Overlaps([1, 2, 3]);
4 });
  
```

Figure 6: Example of refactoring Figure 5

```

1 delegate void LambdaType<T>(T v);
2 void Subroutine<T>(T collection
, LambdaType<T> lambda) where T : ICollection<int> {
3     collection.Add(5);
4     lambda.Invoke(collection);
5     collection.Add(10);
6 }
  
```

Figure 7: Method consolidated with Figure 6

```

1 int total = 0;
2 StreamReader reader = new StreamReader(filePath);
3 string line;
4 while ((line = reader.ReadLine()) != null) {
5     if (int.TryParse(line, out int number)) {
6         if (number == 0) break;
7         total += number;
8     }
9 }
  
```

```

1 int max = 0;
2 StreamReader reader = new StreamReader(filePath);
3 string line;
4 while ((line = reader.ReadLine()) != null) {
5     if (int.TryParse(line, out int number)) {
6         if (number < 0) break;
7         if (max < number) max = number;
8     }
9 }
  
```

Figure 8: Code clones containing jump statements

3.4 Exit block checking

Handling statements that alter control flow, such as break statements, poses a significant challenge for the Extract Method refactoring. In a Control Flow Graph, which represents program execution paths as a directed graph, the final node to be executed is called an

exit block. This study defines an exit block as either the last executable statement in a code fragment or the end of the code fragment if control can reach it. In C#, such last executable statements include break, continue, return, and goto statements. Different exit blocks lead to distinct subsequent behaviors after exiting the code

```

1 int total = 0;
2 Subroutine((number) => {
3     if (number == 0) return true;
4     total += number;
5     return false;
6 });

```

```

1 int max = 0;
2 Subroutine((number) => {
3     if (number < 0) return true;
4     if (max < number) max = number;
5     return false;
6 });

```

Figure 9: Example of refactoring Figure 8

```

1 delegate bool LambdaType(int number);
2 void Subroutine(string path, LambdaType lambda) {
3     StreamReader reader = new StreamReader(filePath);
4     string line;
5     while ((line = reader.ReadLine()) != null) {
6         if (int.TryParse(line, out int number)) {
7             if (lambda.Invoke(number)) break;
8         }
9     }
10 }

```

Figure 10: Method consolidated with Figure 9

fragment. Crucially, a code fragment with a single exit block cannot emulate the behavior of one with multiple exit blocks.

A function call statement has only a single exit block, whereas a code fragment composed of several statements may contain multiple exit blocks. The Extract Method refactoring replaces a code fragment with a function call. A code fragment with multiple exit blocks cannot be replaced by a simple function call. Therefore, we emulate the behavior of multiple exit blocks by carefully designing the subroutine's return value and its handling at the call site.

For simplicity, `break`, `continue`, and `return` statements are considered the end of code fragments in this study. We excluded the `goto` statement from our analysis. Since `goto` statements can jump to arbitrary locations and make control flow difficult to track, we omitted them. A more detailed discussion on the impact of including `goto` is discussed later. This exclusion is justified because the use of `goto` is widely considered poor practice that reduces code readability [4].

Specifically, the subroutine's return value indicates the type of exit block to the call site. The call site of the subroutine inserts an `if` statement after the function call to execute the appropriate behavior for each exit block, thereby emulating the original control flow containing multiple exit blocks. To ensure that only necessary exit blocks are emulated, it is essential to check the exit blocks included in the code fragments of the clone pair and in parameterized behavior.

Figure 8 illustrates an example requiring the emulation of exit block behavior. The matching code segments are highlighted in orange, while the behaviors parameterized by a lambda expression are highlighted in blue and green. This parameterized behavior contains a `break` statement, resulting in multiple exit blocks. Figure 9 shows the source code after refactoring, and Figure 10 presents the definition of the extracted subroutine. Within the parameterized statement, the `break` statement is replaced by a `return` statement

that returns a Boolean value. This return value informs the call site which behavior to emulate.

The extracted method is similar to the common higher-order function `reduce`. However, unlike `reduce`, this method can terminate the loop prematurely. Our approach provides developers with insights into refactoring using higher-order functions.

This approach can be a bit tricky to apply in some programming languages. One notable programming language is Java. Java supports labeled `continue` and `break` statements, which can result in multiple possible target locations for control transfer. To accurately emulate the behavior of `continue` and `break` statements, it is necessary to take labels into account. JavaScript also has similar characteristics. A similar problem occurs with the `goto` statement, which, like labeled statements, enables transfers to multiple destinations. This is why our research does not consider the `goto` statement.

3.5 Argument determination

Arguments must be appropriately passed to both subroutines and lambda expressions used for behavior parameterization.

If a local variable declared in a subroutine is accessed within a parameterized lambda expression, it must be passed as an argument from the subroutine to the lambda expression. Additionally, the subroutine must receive the lambda expression to be invoked as an argument.

When a lambda expression is passed as an argument for behavior parameterization, its type must also be explicitly specified. In C#, `delegate` types, which are used to reference functions, must be declared in advance.

3.6 Source code generation

The subroutine is generated using one of the code fragments in the clone pair as a template. In our approach, the static class defining the subroutine is placed in the global namespace. To define this static class, a new source file is added to the project.

In addition to the static class, this source file defines the `delegate` type for lambda expressions and an enumeration representing exit block types. Both types are declared with the `internal` access modifier. The `internal` modifier allows access only within the same project. Compared to the `public` modifier, `internal` supports more refactoring scenarios because it avoids the need to expose refactored components to external projects.

3.7 Related characteristics of C#

Our approach relies significantly on the specifications of the programming language. This section outlines the C# language features most pertinent to our approach.

3.7.1 Parameter modifier. In C#, parameters are generally passed by value. In this case, a function cannot reassign a value to the caller's variable through a parameter. However, by adding the `ref` modifier to a parameter, the function can pass it by reference, allowing it to modify the caller's variable.

The `ref` modifier can only be applied to variables that have already been initialized. If a variable is uninitialized, the `out` modifier can be used instead of `ref`, enabling the function to initialize the variable

within its body. Our approach uses appropriate parameter modifiers to create subroutines that reassign parameter values without significantly changing the original code structure.

3.7.2 Restriction of parameter modifier. C# supports asynchronous methods and iterator methods, also known as generators in other languages. However, parameter modifiers cannot be used in these methods.¹ Therefore, this study excludes from refactoring any code clones that appear in contexts where parameter modifiers cannot be used.

3.7.3 Property. A property is a class or structure member consisting of a getter, a setter, or both. Since properties are accessed the same way as fields or variables, it is impossible to distinguish between property access and field or variable access based solely on static syntactic analysis. However, unlike fields and other members, properties cannot be passed by reference to functions using parameter modifiers. Nevertheless, to achieve pass-by-reference behavior, getter and setter functions must be passed as delegates.

3.8 Integration with code clone detector

The proposed approach takes a clone pair as input. The suitability of the clone pair for Extract Method refactoring is not verified beforehand. The proposed approach must determine whether the given clone pair is suitable for refactoring, either before or during the process. During the process, the approach may terminate with an error. Such an error can be interpreted as an indication that the clone pair is unsuitable for refactoring.

The reasons why a clone pair may be deemed unsuitable are as follows:

- (1) Cross-project code clones,
- (2) No matching statements,
- (3) Variable type mismatch,
- (4) The required `ref` or `out` modifiers cannot be added,
- (5) Code clones containing method calls,
- (6) Code clones containing `yield`, `async`, `await`, and
- (7) Source code analysis failed.

(1) indicates that no matching statements could be found during statement mapping. (2) indicates that the parameter types could not be determined during variable mapping. (3) indicates that `ref` or `out` modifiers are required in variable mapping but cannot be applied. Sometimes, `ref` and `out` modifiers cannot be used inside lambda expressions. (4) indicates that the code clone contains an instance method call. Since the subroutine is defined in a separate class, the instance methods from the original class cannot be invoked directly. (5) indicates that the code clones are located within iterator methods or asynchronous methods. This study excludes code clones within iterator or asynchronous methods from refactoring targets. (6) represents other abnormal cases, such as broken projects that could not be correctly analyzed.

In addition to performing refactoring, the proposed approach may report one of the errors listed in (1) through (6). In particular, detecting errors (1) through (5) indicates that the clone pair is either out of scope or not suitable for refactoring.

¹As of the latest C# version released in 2025, parameter modifiers are allowed in asynchronous and iterator methods under certain conditions.

4 EVALUATION

4.1 Procedure

To evaluate the accuracy and applicability of the proposed subroutine generator, we applied it to clone pairs extracted from real-world software projects. We assessed the correctness of the generated subroutines by comparing project test results before and after applying the subroutine generator to the input clone pairs. If the test results were identical, we determined that the project's behavior was preserved after subroutine generation.

The input code clones for the subroutine generator were obtained from the NiCad code clone detector [3], which detects Type-1, Type-2, and Type-3 code clones. We collected 22 C# projects from GitHub that include tests and evaluated 2,217 clone pairs detected by NiCad.

However, this simple comparison is not sufficient. If the refactored portions of code are not executed by any test, the behavior may appear unchanged even when the refactoring is incorrect. To address this, we intentionally applied incorrect refactoring and compared the resulting test outcomes. If the test results before and after applying the incorrect refactoring differed, we determined that the affected code was exercised by the tests. From this point onward, we refer to the subroutine generator that applies incorrect refactoring as the incorrect subroutine generator. The only difference between the subroutine generator and the incorrect subroutine generator is the function body of the generated subroutines. Specifically, the incorrect subroutine generator generates subroutines that throw exceptions. Both generators output the same source code, except for the function body. The evaluation procedure using the incorrect subroutine generator is described below. The authors prepared three versions of the project:

- (1) original project,
- (2) project after applying the subroutine generator, and
- (3) project after applying the incorrect subroutine generator.

We ran tests for each version and compared their test results. If the test results of the original project differed from those of the project with the incorrect subroutine generator, we assumed that the refactored code had been run by the tests. If those results matched the ones from the project with the correct subroutine generator, we concluded that the subroutine generator correctly preserved the program's behavior. To eliminate non-deterministic test results, we ran each test four times and confirmed no variations in the test outcomes.

4.2 Results

Table 1 shows the classification of code clones produced by our proposed approach. Items (1) to (7) correspond to the error categories described in Section 3.8. 35.0% of the total, 775 to be exact, clone pairs were determined to be refactorable, and were actually refactored. They were then evaluated using the procedure described earlier.

The evaluation results for each clone pair are shown in Table 2. The clone pairs were classified into three categories based on build success and test results. "Refactored successfully" refers to clone pairs for which the refactored source code built successfully. The correctness of the refactoring was also verified through testing. "Refactored, not tested" refers to clone pairs where the refactored source code built successfully. However, there were no tests covering the refactored portions, so the correctness of the refactoring could not be confirmed. "Failed to refactor" refers to clone pairs for which refactoring failed.

Among the refactorable clone pairs, 32.1% had refactored source code that both built successfully and passed the tests. In addition, 61.0% of the refactorable clone pairs built successfully.

5 DISCUSSION

5.1 Why do code clone statements not match?

Table 1 shows that in 38.0% of the clone pairs, no matching statements were found. Since clone pairs consist of similar groups of statements, the absence of matching statements may seem surprising. NiCad detects Type-2 code clones, which are identified by ignoring differences in identifier names. NiCad disregards identifier name differences regardless of the type of identifier. In contrast, our approach selectively ignores identifier name differences based on the type of identifier.

In practice, some clone pairs included switch statements with differing case labels. Such differences can explain the absence of matching statements.

Additionally, we observed cases where matches inside nested constructs were missed. In this study, for statements were treated as one of the constructs that form nested structures. A for statement was considered a match only when its loop variable definition and continuation condition were identical. If the code fragment consisted solely of for statements and there were slight differences in their continuation conditions, no match was detected.

An example where refactoring was not possible due to minor differences in the continuation conditions of for statements is shown in Figure 11. In this example, all statements inside the for loops match except for differences in identifier names. However, the loop variable definitions and continuation conditions differ significantly, making it difficult to consolidate the code fragments into a single method through expression parameterization. Although only the highlighted line containing the for loop condition differs, the rest of the source code is nearly identical. Nonetheless, the entire code fragment cannot be directly extracted as a single method. While extracting the entire code fragment as a method is not straightforward, it would be possible to extract only the inner part of the for loop.

Table 1: Classification Results

Category	Code clones
Refactorable code clones	775
(1) No matching statements	842
(2) Variable type mismatch	46
(3) ref, out modifiers cannot be added	100
(4) Code clones containing method calls	119
(5) Code clones containing yield, async, await	293
(6) Source code analysis failed	42
Total	2,217

Table 2: Evaluation Results

Category	Code clones
Refactored successfully	249
Refactored, not tested	224
Failed to refactor	302
Total	775

```

1 for (int i = 0; i < str.Length; i++) {
2     char c = str [i];
3     if (c == '\n') {
4         if (idx == -1)
5             idx = i;
6         idx = i;
7     }
8     else if (c == '\u000d' || c == '\t')
9         continue;
10    else
11        break;
12 }

1 for (int i = str.Length - 1; i >= 0; i--) {
2     char c = str [i];
3     if (c == '\n') {
4         if (index == -1)
5             index = i;
6         index = i;
7     }
8     else if (c == '\u000d' || c == '\t')
9         continue;
10    else
11        break;
12 }

```

Figure 11: Example of code clone with nested syntax

As a possible countermeasure, we could consider applying the Extract Method refactoring within the nested structure itself. This is applicable when the surrounding nested constructs do not match. However, attempting to refactor every possible combination of nested structures within a code fragment results in exponential growth in computational complexity. This complexity increases with the number of nested constructs.

Another possible improvement is to implement more flexible expression parameterization to absorb larger differences in loop variable declarations and continuation conditions.

5.2 Why does refactoring fail?

Table 2 shows that refactoring failed in 39.0% of the refactorable code clones. One factor contributing to refactoring failure is the use of private classes. The class that defines the subroutine is external to the original classes that contain the code clones. Therefore, it cannot access private classes used in the original source code. The proposed approach does not take into account the accessibility of classes used within the code fragments.

To solve this problem, it is effective to check the accessibility of the used classes and members. Another solution is to consider changing the location where the subroutine is defined. Defining the subroutine in the class that contains the code clone can resolve accessibility issues and avoid problems related to instance method calls. However,

Table 3: Evaluation Results with Behavior Parameterization

Category	Code clones
Refactored successfully	175
Refactored, not tested	34
Total	209

this approach cannot be used when the code fragments in the clone pair belong to different classes.

Additionally, relying on interfaces can sometimes resolve accessibility problems. If there is a public interface with a private implementation, using the interface instead of the implementation can avoid accessibility issues. This measure is effective if using the interface sufficiently replicates the original behavior. In our variable mapping, our approach selected the most specific common class that closely matches the actual variable types. To address accessibility issues, the parameter type should be selected as the most specific common class that is publicly accessible and closest to the actual variable types.

5.3 How effective is behavior parameterization?

The proposed approach performs behavior parameterization only when the clone pair contains statements that differ. We investigated how many clone pairs were refactored using behavior parameterization. Table 3 shows how many clone pairs were refactored using behavior parameterization. Among the refactorable clone pairs, 27.0% were successfully refactored using behavior parameterization. This result demonstrates that behavior parameterization is effective in a substantial number of cases.

6 THREATS TO VALIDITY

There are two primary threats to the validity of this study.

Internal validity may be threatened by the evaluation method. To verify the correctness of refactoring, we compared test results. The quality of the tests affects the reliability of the results. Since test results are not always stable, there is a risk of coincidental matches or mismatches. To eliminate unstable tests, we ran the tests multiple times. Additionally, to avoid false negatives from tests that do not cover the refactored portions, we introduced an incorrect subroutine generator. We then evaluated whether the tests could detect the incorrect behavior.

External validity may be threatened by the quality of the evaluation data. The results of the evaluation are affected by the quality of the code clone pairs used for testing. To reduce bias in clone pair quality, we conducted evaluations across 22 different projects. Furthermore, the quality of the clone pairs also depends on the code clone detector used. In this study, we used NiCad [3]. Other code clone detectors capable of detecting clones in C# include CCFinder [6] and Simian [1]. Future studies are encouraged to consider using different code clone detectors or conducting validation on different projects.

7 CONCLUSION

In this study, we proposed and implemented an Extract Method refactoring approach for C#, which utilizes behavior parameterization. Furthermore, we evaluated our approach using projects that included test cases. We applied the proposed refactoring and analysis approach to 2,217 clone pairs detected from 22 C# projects. We then verified the correctness of the refactoring for the clone pairs that were successfully refactored.

The proposed approach determined that 35.0% of all clone pairs were suitable for refactoring.

Furthermore, 61.0% of the refactorable projects built successfully. For 32.1% of these projects, the provided tests passed, confirming the behavioral correctness. In the remaining 28.9% of cases, behavioral

correctness could not be confirmed because the test suite did not cover the refactored source code.

In this study, we used the detection results from NiCad [3]. Future work could investigate how the results differ when using other code clone detectors. Variable names in both the subroutine and the caller were generated mechanically using sequential numbering. Improving readability through variable name refactoring [7] is a promising direction for future work.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 25K03102, 24H00692, 23K24823.

References

- [1] 2021. Simian. <https://simian.quandarypeak.com/>. Accessed: 5.25.2025.
- [2] N. Yoshida, S. Numata, E. Choiz, and K. Inoue. 2019. Proactive Clone Recommendation System for Extract Method Refactoring. In *2019 IEEE/ACM 3rd International Workshop on Refactoring (IWoR)*. IEEE, Montreal, QC, Canada, 67–70. doi:10.1109/IWoR.2019.00020
- [3] J. R. Cordy, and C. K. Roy. 2011. The NiCad Clone Detector. In *2011 IEEE 19th International Conference on Program Comprehension*. IEEE, Kingston, ON, Canada, 219–220. doi:10.1109/ICPC.2011.26
- [4] E. W. Dijkstra. 1972. *Chapter I: Notes on structured programming*. Academic Press Ltd., GBR, 1–82.
- [5] L. Moura, M. Sant'Anna, I.D. Baxter, A. Yahin and L. Bier. 1998. Clone detection using abstract syntax trees. In *Proceedings. International Conference on Software Maintenance (Cat. No. 98CB36272)*. IEEE, Bethesda, MD, USA, 368–377. doi:10.1109/ICSM.1998.738528
- [6] T. Kamiya, S. Kusumoto and K. Inoue. 2002. CCFinder: A Multilingual Token-Based Code Clone Detection System for Large Scale Source Code. *IEEE Transaction on Software Engineering* vol. 28, no. 7, pp. 654–670 (July 2002). (July 2002).
- [7] H. Liu, Y. Wang, Z. Wei, Y. Xu, J. Wang, H. Li, and R. Ji. 2023. RefBERT: A Two-Stage Pre-trained Framework for Automatic Rename Refactoring. In *Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis* (Seattle, WA, USA) (ISSTA 2023). Association for Computing Machinery, New York, NY, USA, 740–752. doi:10.1145/3597926.3598092
- [8] A. Lozano, and M. Wermelinger. 2008. Assessing the effect of clones on changeability. In *2008 IEEE International Conference on Software Maintenance*. IEEE, Beijing, China, 227–236. doi:10.1109/ICSM.2008.4658071
- [9] N. Tsantalis, D. Mazinanian, and G. P. Krishnan. 2015. Assessing the Refactorability of Software Clones. *IEEE Transactions on Software Engineering* 41, 11 (Nov 2015), 1055–1090. doi:10.1109/TSE.2015.2448531
- [10] N. Tsantalis, D. Mazinanian, and S. Rostami. 2017. Clone Refactoring with Lambda Expressions. In *2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE)*. IEEE, Buenos Aires, Argentina, 60–70. doi:10.1109/ICSE.2017.14
- [11] E. A. AlOmar, M. W. Mkaouer, and A. Ouni. 2024. Behind the Intent of Extract Method Refactoring: A Systematic Literature Review. *IEEE Transactions on Software Engineering* 50, 4 (2024), 668–694. doi:10.1109/TSE.2023.3345800
- [12] Oracle. 2014. Java Documentation, Lambda Expressions. <https://docs.oracle.com/javase/tutorial/java/javaOO/lambdaexpressions.html>. Accessed: 2.19.2025.
- [13] T. Robert, and G. Jeff. 2012. Increasing clone maintenance support by unifying clone detection and refactoring activities. *Information and Software Technology* 54, 12 (2012), 1297–1307. doi:10.1016/j.infsof.2012.06.011 Special Section on Software Reliability and Security.
- [14] C. Roy, and J. Cordy. 2007. A Survey on Software Clone Detection Research. *School of Computing TR 2007-541* 541, 115 (01 2007), 3–7.
- [15] M. Mondal, C. K. Roy, and K. A. Schneider. 2014. Automatic Identification of Important Clones for Refactoring and Tracking. In *2014 IEEE 14th International Working Conference on Source Code Analysis and Manipulation*. IEEE, Victoria, BC, Canada, 11–20. doi:10.1109/SCAM.2014.11
- [16] M. Mondal, C. K. Roy, and K. A. Schneider. 2020. A Fine-Grained Analysis on the Inconsistent Changes in Code Clones. In *2020 IEEE International Conference on Software Maintenance and Evolution (ICSME)*. IEEE, Adelaide, SA, Australia, 220–231. doi:10.1109/ICSME46990.2020.00030