A Study and Toolkit for Asynchronous Programming in C#

Semih Okur, David L. Hartveld, Danny Dig and Arie van Deursen

Report TUD-SERG-2013-016
A Study and Toolkit for Asynchronous Programming in C#
Semih Okur1, David L. Hartveld2, Danny Dig3, Arie van Deursen2
1University of Illinois  2Delft University of Technology  3Oregon State University
okur2@illinois.edu d.l.hartveld@student.tudelft.nl digd@eecs.oregonstate.edu arie.vandeursen@tudelft.nl

ABSTRACT
Asynchronous programming is in demand today, because responsiveness is increasingly important on all modern devices. Yet, we know little about how developers use asynchronous programming in practice. Without such knowledge, developers, researchers, language and library designers, and tool vendors can make wrong assumptions.

We present the first study that analyzes the usage of asynchronous programming in a large experiment. We analyzed 1378 open source Windows Phone (WP) apps, comprising 12M SLOC, produced by 3376 developers. Using this data, we answer 2 research questions about use and misuse of asynchronous constructs. Inspired by these findings, we developed (i) Asyncifier, an automated refactoring tool that converts callback-based asynchronous code to the new async/await; (ii) Corrector, a tool that finds and corrects common misuses of async/await. Our empirical evaluation shows that these tools are (i) applicable and (ii) efficient. Developers accepted 313 patches generated by our tools.

1. INTRODUCTION
User interfaces are usually designed around the use of a single user interface (UI) event thread [16, 17, 24, 25]: every operation that modifies UI state is executed as an event on that thread. The UI “freezes” when it cannot respond to input, or when it cannot be redrawn. It is recommended that long-running CPU-bound or blocking I/O operations execute asynchronously so that the application (app) continues to respond to UI events.

Asynchronous programming is in demand today because responsiveness is increasingly important on all modern devices: desktop, mobile, or web apps. Therefore, major programming languages have APIs that support non-blocking, asynchronous operations (e.g., to access the web, or for file operations). While these APIs make asynchronous programming possible, they do not make it easy.

Asynchronous APIs rely on callbacks. However, callbacks invert the control flow, are awkward, and obfuscate the intent of the original synchronous code [8].

Recently, major languages (F# [38], C# and Visual Basic [8] and Scala [7]) introduced async constructs that resemble the straightforward coding style of traditional synchronous code. Thus, they recognize asynchronous programming as a first-class citizen.

Yet, we know little about how developers use asynchronous programming and specifically the new async constructs in practice. Without such knowledge, other developers cannot educate themselves about the state of the practice, language and library designers are unaware of any misuse, researchers make wrong assumptions, and tool vendors do not provide the tools that developers really need. This knowledge is also important as a guide to designers of other major languages (e.g., Java) planning to support similar constructs. Hence, asynchronous programming deserves first-class citizenship in empirical research and tool support, too.

We present the first study that analyzes the usage of asynchronous libraries and new language constructs, async/await in a large experiment. We analyzed 1378 open source Windows Phone (WP) apps, comprising 12M SLOC, produced by 3376 developers. While all our empirical analysis and tools directly apply to any platform app written in C# (e.g., desktop, console, web, tablet), in this paper we focus on the Windows Phone platform.

We focus on WP apps because we expect to find many exemplars of asynchronous programming, given that responsiveness is critical. Mobile apps can easily be unresponsive because mobile devices have limited resources and have high latency (excessive network accesses). With the immediacy of touch-based UIs, even small hiccups in responsiveness are more obvious and jarring than when using a mouse or keyboard. Some sluggishness might motivate the user to uninstall the app, and possibly submit negative comments in the app store [37]. Moreover, mobile apps are becoming increasingly more important. According to Gartner, by 2016 more than 300 billion apps will be downloaded annually [18].

The goal of this paper is twofold. First, we obtain a deep understanding of the problems around asynchronous programming. Second, we present a toolkit (2 tools) to address exactly these problems. To this end, we investigate 1378 WP apps through tools and by hand, focusing on the following research questions:

RQ1: How do developers use asynchronous programming?
RQ2: To what extent do developers misuse async/await?

We found that developers heavily use callback-based asyn-
chronous idioms. However, Microsoft no longer officially recommends these asynchronous idioms [30] and has started to replace them with new idioms in new libraries (e.g., WinRT). Developers need to refactor callback-based idioms to new idioms that can take advantage of the async/await language constructs. The changes that the refactoring requires are non-trivial, though. For instance, developers have to inspect deep call graphs. Furthermore, they need to be extra careful to preserve the exception-handling. Thus, we implemented the refactoring as an automated tool, Asyncifier.

We also found that nearly half of WPS apps have started to use the 9-month-old async/await keywords. However, developers misuse async/await in various ways. We define misuse as anti-patterns, which hurt performance and might cause serious problems like deadlocks. For instance, we found that 14% of methods that use (the expensive) async/await do this unnecessarily. 10% of developers do not follow an important good practice [22]. 1 out of 5 apps miss opportunities in async methods to increase asynchronicity, and developers (almost) always unnecessarily capture context, hurting performance. Thus, we implemented a transformation tool, Corrector, that finds and corrects the misused async/await.

This paper makes the following contributions:

Empirical Study: To the best of our knowledge, this is the first large-scale empirical study to answer questions about asynchronous programming and new language constructs, async/await, that will be available soon in other major programming languages. We present implications of our findings from the perspective of four main audiences: developers, language and library designers, researchers, and tool vendors.

Toolkit: We implemented the analysis and transformation algorithms to address the challenges (Asyncifier and Corrector).

Evaluation: We evaluated our tools by using the code corpus and applied the tools hundreds of times. We show that our tools are highly applicable and efficient. Developers find our transformations useful. Using Asyncifier, we applied and reported refactorings in 10 apps. 9 replied and accepted each one of our 28 refactorings. Using Corrector, we found and reported misuses in 19 apps. 18 replied and accepted each of our 285 patches.

Outreach: Because developers learn new language constructs through both positive and negative examples, we designed a website, http://LearnAsync.NET/, to show hundreds of such usages of asynchronous idioms and async/await keywords.

2. BACKGROUND

When a button click event handler executes a synchronous long-running CPU-bound or blocking I/O operation, the user interface will freeze because the UI event thread cannot respond to events. Code listing 1 shows an example of such an event handler, method Button_Click. It uses the GetFromUrl method to download the contents of a URL, and place it in a text box. Because GetFromUrl is waiting for the network operation to complete, the UI event thread is blocked, and the UI is unresponsive.

Keeping UIs responsive thus means keeping the UI event thread free of those long-running or blocking operations. If these operations are executed asynchronously in the background, the foreground UI event thread does not have to busy-wait for completion of the operations. That frees the UI event thread to respond to user input, or redraw the UI: the user will experience the UI to be responsive.

CPU-bound operations can be executed asynchronously by (i) explicitly creating threads, or (ii) by reusing a thread from the thread pool.

I/O operations are more complicated to offload asynchronously. The naive approach would be to just start another thread to run the synchronous operation asynchronously, using the same mechanics as used for CPU-bound code. However, that would still block the new thread, which consumes significant resources, hurting scalability.

The solution is to use asynchronous APIs provided by the platform. The .NET framework mainly provides two models for asynchronous programming: (1) the Asynchronous Programming Model (APM), that uses callbacks, and (2) the Task Asynchronous Pattern (TAP), that uses Tasks, which are similar to the concept of futures found in many other languages such as Java, Scala or Python.

2.1 Asynchronous Programming Model

APM, the Asynchronous Programming Model, was part of the first version of the .NET framework, and has been in existence for 10 years. APM asynchronous operations are started with a Begin method invocation. The result is obtained with an End method invocation. In Code listing 2, BeginGetResponse is such a Begin method, and EndGetResponse is an End method. BeginGetResponse is used to initiate an asynchronous HTTP GET request. The .NET framework starts the I/O operation in the background (in this case, sending the request to the remote web server). Control is returned to the calling method, which can then continue to do something else. When the server responds, the .NET framework will “call back” to the application to notify that the response is ready. EndGetResponse is then used in the callback code to retrieve the actual result of the operation. See Figure 1 for an illustration of this flow of events.

The APM Begin method has two pattern-related parameters. The first parameter is the callback delegate (which is a managed, type-safe equivalent of a function pointer). It can be defined as either a method reference, or a lambda expression. The second parameter allows the developer to pass any single object reference to the callback, and is called state.

The .NET framework will execute the callback delegate on the thread pool once the asynchronous background operation completes. The EndGetResponse method is then used in the callback to obtain the result of the operation, the actual WebResponse.

Note a subtle difference between the synchronous, sequential example in Code listing 1 and the asynchronous, APM-

---

Code 1  Synchronous example

```csharp
1 void Button_Click(...) {
2     string contents = GetFromUrl(url);
3     MessageBox.Text = contents;
4     string GetFromUrl(string url) { 
5         WebRequest request = WebRequest.Create(url);
6         WebResponse response = request.GetResponse();
7         Stream stream = response.GetResponseStream();
8         return stream.ReadAsString();
9     }
10 }
```

---

The .NET framework will execute the callback delegate on the thread pool once the asynchronous background operation completes. The EndGetResponse method is then used in the callback to obtain the result of the operation, the actual WebResponse.

Note a subtle difference between the synchronous, sequential example in Code listing 1 and the asynchronous, APM-

---

Code 1  Synchronous example

```csharp
1 void Button_Click(...) {
2     string contents = GetFromUrl(url);
3     MessageBox.Text = contents;
4     string GetFromUrl(string url) { 
5         WebRequest request = WebRequest.Create(url);
6         WebResponse response = request.GetResponse();
7         Stream stream = response.GetResponseStream();
8         return stream.ReadAsString();
9     }
10 }
```
2.2 Task-based Asynchronous Pattern

TAP, the Task-based Asynchronous Pattern, provides for a slightly different approach. TAP methods have the same base operation name as APM methods, without ‘Begin’ or ‘End’ prefixes, and instead have an ‘Async’ suffix. The API consists of methods that start the background operation and return a Task object. The Task represents the operation in progress, and its future result.

The Task can be (1) queried for the status of the operation, (2) synchronized upon to wait for the result of the operation, or (3) set up with a continuation that resumes in the background when the task completes (similar to the callbacks in the APM model).

2.3 Drawbacks of APM and plain TAP

Using APM and plain TAP directly has two main drawbacks. First, the code that must be executed after the asynchronous operation is finished, must be passed explicitly to the Begin method invocation. For APM, even more scaffolding is required: The End method must be called, and that usually requires the explicit passing and casting of an ‘async state’ object instance - see Code listing 2, lines 9-10. Second, even though the Begin method might be called from the UI event thread, the callback code is executed on a thread pool thread. To update the UI after completion of the asynchronous operation from the thread pool thread, an event must be sent to the UI event thread explicitly - see Code listing 2, line 13-15.

2.4 Pause’n’play with async & await

To solve this problem, the async and await keywords have been introduced. When a method has the async keyword modifier in its signature, the await keyword can be used to define pausing points. When a Task is awaited in an await expression, the current method is paused and control is returned to the caller. When the await’ed Task’s background operation is completed, the method is resumed from right after the await expression. Code listing 3 shows the TAP- & async/await-based equivalent of Code listing 2, and Figure 2 illustrates its flow of execution.

The code following the await expression can be considered a continuation of the method, exactly like the callback that needs to be supplied explicitly when using APM or plain TAP. Methods that have the async modifier will thus run synchronously up to the first await expression (and if it does not have any, it will complete synchronously). Merely adding the async modifier does not magically make a method be asynchronously executed in the background.

2.5 Where is the code executing?

There is one important difference between async/await continuations, and APM or plain TAP callback continuations: APM and plain TAP always execute the callback on a thread pool thread. The programmer needs to explicitly schedule a UI event to interface with the UI, as shown in Code listing 2 and Figure 1.

In async/await continuations, the await keyword, by default, captures information about the thread in which it is executed. This captured context is used to schedule execution of the rest of the method in the same context as when the asynchronous operation was called. For example, if the
The `await` keyword is encountered in the UI event thread, it will capture that fact. Once the background operation is completed, the continuation of the rest of the method is scheduled back onto the UI event thread. This behavior allows the developer to write asynchronous code in a sequential manner. See Code listing 3 for an example.

Comparing the code examples in listings 1 and 3 will show that the responsive version based on TAP & `async/await` only slightly differs from the sequential version. It is readable in a similar fashion, and even the UI update (setting the contents of the text box) is back to its original place.

By default, `await` expressions capture the current context. However, it is not always needed to make the expensive context switch back to the original context. To forestall a context switch, an `await`ed Task can be set to ignore capturing the current context by using `ConfigureAwait(false)`. In Code listing 3, in `GetFromUrlAsync`, none of the statements following the `await` expressions require access to the UI. Hence, the `await`ed Task is set with `ConfigureAwait(false)`. In `Button_Click`, the statement following `await GetFromUrlAsync()` does need to update the UI. So that `await` expression should capture the original context, and the task should not be set up with `ConfigureAwait(false).

3. RESEARCH QUESTIONS

We are interested in assessing the usage of state of the art asynchronous programming in real world WP apps.

3.1 Methodology

Corpus of Data: We chose Microsoft’s CodePlex [11] and GitHub [19] as sources of the code corpus of WP apps. According to a recent study [27], most C# apps reside in these two repositories. We developed WPCollector to create our code corpus. It is available online [10] for reuse by other researchers.

We used WPCollector to download all recently updated WP apps which have a WP-related signature in their project files. It ignores (1) apps without commits since 2012, and (2) apps with less than 500 non-comment, non-blank lines of code (SLOC). The latter “toy apps” are not representative of production code.

WPCollector makes as many projects compilable as possible (e.g., by resolving-installing dependencies), because the Roslyn APIs that we rely on (see Analysis Infrastructure) require compilable source code.

WPCollector successfully downloaded and prepared 1378 apps, comprising 12M SLOC, produced by 1028 developers, which we all used in our analysis, without sampling.

WP7, released in October 2010, is targeted by 1115 apps. WP8, released in October 2012, is targeted by 342 apps. 86 apps target both platforms, because WP8 apps cannot run on WP7 devices.

Analysis Infrastructure: We developed AsyncAnalyzer to perform the static analysis of asynchronous programming construct usage. We used Microsoft’s recently released Roslyn [31] SDK, which provides an API for syntactic and semantic program analysis, AST transformations, and editor services in Visual Studio. Because the publicly available version of Roslyn is incomplete and does not support the `async/await` keywords yet, we used an internal build obtained from Microsoft.

We executed AsyncAnalyzer over each app in our corpus.

---

<table>
<thead>
<tr>
<th></th>
<th>WP7</th>
<th>WP8</th>
</tr>
</thead>
<tbody>
<tr>
<td>I/O APM</td>
<td>1028</td>
<td>217</td>
</tr>
<tr>
<td>I/O TAP</td>
<td>123</td>
<td>269</td>
</tr>
<tr>
<td>New Thread</td>
<td>183</td>
<td>28</td>
</tr>
<tr>
<td>BG Worker</td>
<td>149</td>
<td>11</td>
</tr>
<tr>
<td>ThreadPool</td>
<td>386</td>
<td>52</td>
</tr>
<tr>
<td>New Task</td>
<td>51</td>
<td>182</td>
</tr>
</tbody>
</table>

Table 1: Usage of asynchronous idiom. The three columns per platform show the total number of idiom instances, the total number of apps with instances of the idiom, and the percentage of apps with instances of the idiom.

For each of these apps, it inspects the version from the main development branch as of August 1st, 2013. We developed a specific analysis to answer each research question.

3.2 How do developers use asynchronous programming?

**Asynchronous APIs:** We detected all APM and TAP methods that are used in our code corpus as shown in Table 1. Because in WP7 apps, TAP methods are only accessible via additional libraries, Table 1 tabulates the usage statistics for WP7 and WP8 apps separately. The data shows that APM is more popular than TAP for both WP7 and WP8.

We also manually inspected all APM and TAP methods used and categorized them based on the type of I/O operations: network (1012), file system (310), database (145), user interaction (102) and other I/O (e.g., speech recognition) (68). We found that asynchronous operations are most commonly used for network operations.

There are two ways to offload CPU-bound operations to another thread: by creating a new thread, or by reusing threads from the thread pool. Based on C# books and references [1], we distinguish 3 different approaches developers use to access the thread pool: (1) the BackgroundWorker class, (2) accessing the ThreadPool directly, and (3) creating Tasks. Table 1 tabulates the usage statistics of all these approaches. Because `Task` is only available in WP7 apps by using additional libraries, the table shows separate statistics for WP7 and WP8 apps. The data shows that `Task` is used significantly more in WP8 apps, most likely because of availability in the core platform.

**Language Constructs:** `async/await` have become accessible for WP development in last quarter of 2012. While they are available by default in WP8, WP7 apps have to reference Microsoft.Bcl.Async library to use them.

We found that 45% (157) of WP8 apps use `async/await` keywords. While nearly half of all WP8 apps have started to use `async/await`, only 10 WP7 apps use them.

**Callback-based APM is the most widely used idiom. While nearly half of all WP8 apps have started to use `async/await`, only 10 WP7 apps use them.**
3.3 Do developers misuse async/await?

Because async/await are relatively new language constructs, we have also investigated how developers misuse these constructs. We define misuse as anti-patterns which hurt performance and might cause serious problems like deadlocks. We detected the following typical misuse idiom.

3.3.1 Fire & Forget methods

799 of 2382 async/await methods are “fire&forget”, which return void. Unless a method is only called as a UI event handler, it must be callable. Otherwise, it is a code smell because it complicates control flow and makes error detection & correction difficult. Exceptions in fire&forget methods cannot be caught in the calling method, causing termination of the app. Instead, they should return Task which does not force the method to return anything; but it enables easier error-handling, composability, and testability.

However, we found that only 339 out of these 799 async void methods are event handlers. It means that 19% of all async methods (400 out of 2383) are not following this important practice [22].

One in five async methods violate the principle that an async method should be awaitable unless it is the top level event handler.

3.3.2 Unnecessary async/await methods

Consider the example from “Cimbalino Windows Phone Toolkit” [3]:

```csharp
public async Task<Stream> OpenFileForReadAsync(...) {
    return await Storage.OpenStreamForReadAsync(path);
}
```

The OpenStream method is a TAP call, which is awaited in the OpenFile method. However, there is no need to await it. Because there is no statement after the await expression except for the return, the method is paused without reason: the Task that is returned by Storage.OpenStream can be immediately returned to the caller. The snippet below behaves exactly the same as the one above:

```csharp
public Task<Stream> OpenFileForReadAsync(...) {
    return Storage.OpenStreamForReadAsync(path);
}
```

It is important to detect this kind of misuse. Adding the async modifier comes at a price: the compiler generates some code in every async method.

We discovered that in 26% of the 167 apps, 324 out of all 2383 methods unnecessarily use async/await. There is no need to use async/await in 14% of all async methods.

3.3.3 Long-running operations under async methods

We also noticed that developers use some potentially long running operations under async methods even though there are corresponding asynchronous versions of these methods in .NET or third-party libraries. Consider the following example from indulged-flickr [15], a Flickr:

```csharp
public async void GetPhotoStreamAsync(...) {
    var response = await DispatchRequest(...);
    using (StreamReader reader = new StreamReader(...)) {
        string jsonString = reader.ReadToEnd();
    }
}
```

The developer might use await ReadToEndAsync() instead of the synchronous ReadToEnd call, especially if the stream is expected to be large.

In the example below from iRacerMotionControl [23], the situation is more severe.

```csharp
private async void BT2Arduino_Send(string WhatToSend) {
    ... 
    await BTSock.OutputStream.WriteAsync(datab);
    txtBTStatus.Text = "sent";
    System.Threading.Thread.Sleep(5000);
    ...
}
```

The UI event thread calls BT2Arduino_Send, which blocks the UI thread by busy-waiting for 5 seconds. Instead of using the blocking Thread.Sleep method, the developer should use the non-blocking Task.Delay(5000) method call to preserve similar timing behavior, and await it to prevent the UI to freeze for 5 seconds.

We found 115 instances of potentially long-running operations in 22% of the 167 apps that use async/await. 1 out of 5 apps miss opportunities in at least one async method to increase asynchronicity.

3.3.4 Unnecessarily capturing context

async/await introduce new risks if the context is captured without specifying ConfigureAwait(false). For example, consider the following example from asdnes [2]:

```csharp
    void GetMessage(byte[] response) {
        ... 
        ReceiveAsync(response).Wait();
        ... 
    }
    async Task<bool> ReceiveAsync(byte[] message) {
        ... 
        return await tcs.Task;
    }
```

If GetMessage is called from the UI event thread, the thread will wait for completion of ReceiveAsync because of the Wait call. When the await completes in ReceiveAsync, it attempts to execute the remainder of the method within the captured context, which is the UI event thread. However, the UI event thread is already blocked, waiting for the completion of ReceiveAsync. Therefore, a deadlock occurs.

To prevent the deadlock, the developer needs to set up the await expression to use ConfigureAwait(false). Instead of attempting to resume the ReceiveAsync method on the UI event thread, it now resumes on the thread pool, and the blocking wait in GetMessage does not cause a deadlock any more. In the example above, although ConfigureAwait(false) is a solution, we fixed it by removing await because it was also an instance of unnecessary async/await use. The developer of the app accepted our fix as a patch.

We found 5 different cases for this type of deadlock which can happen if the caller method executes on UI event thread.

Capturing the context can also cause another problem: it hurts performance. As asynchronous GUI applications grow larger, there can be many small parts of async methods all using the UI event thread as their context. This can cause sluggishness as responsiveness suffers from thousands of paper cuts. It also enables a small amount of parallelism: some asynchronous code can run in parallel with the UI event thread instead of constantly badgering it with bits of work to do.

To mitigate these problems, developers should await the Task with ConfigureAwait(false) whenever they can. If the statements after the await expression do not update the
UI, `ConfigureAwait(false)` must be set. Detecting this misuse is important because using `ConfigureAwait(false)` might prevent future bugs like deadlocks and improve the performance.

1786 out of 2383 async methods do not update GUI elements in their call graph after `await` expressions. We found that `ConfigureAwait(false)` is used in only 16 out of these 1786 async methods in `await` expressions. All 1770 other async methods should have used `ConfigureAwait(false)`. 99% of the time, developers did not use `ConfigureAwait(false)` where this was needed.

```
<table>
<thead>
<tr>
<th>Misuse</th>
<th>#</th>
<th>Method%</th>
<th>App%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)Fire&amp;Forget</td>
<td>460</td>
<td>19%</td>
<td>74%</td>
</tr>
<tr>
<td>(2)Unnec. Async</td>
<td>324</td>
<td>14%</td>
<td>22%</td>
</tr>
<tr>
<td>(3)Potential LongRunning</td>
<td>115</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>(4)Unnec. Context</td>
<td>1770</td>
<td>74%</td>
<td>86%</td>
</tr>
</tbody>
</table>
```

### 4. TOOLKIT

Based on our findings, we developed a two-fold approach to support the developer: (1) **Asyncifier**, a refactoring tool to upgrade legacy callback-based APM code to take advantage of `async/await` construct (see section 4.1) and (2) Corrector, a tool for detecting and fixing misuses of `async/await` in code (see Section 4.2).

Asyncifier helps the developer in two ways: (1) the code is upgraded without errors, retaining original behavior, and (2) it shows how to correctly use the `async/await` keywords in production code. If the developer manually introduces `async/await`, Corrector will help in finding and removing misuses.

#### 4.1 Refactoring APM to async & await

##### 4.1.1 Challenges

There are three main challenges that make it hard to execute the refactoring quick and flawlessly by hand. First, the developer needs to understand if the APM instance is a candidate for refactoring based on the preconditions in Section 4.1.2. Second, he must transform the code while retaining the original behavior of the code - both functionally and in terms of scheduling. This is non-trivial, especially in the presence of (1) exception handling, and (2) APM `End` methods that are placed deeper in the call graph.

##### Exception handling

The refactoring from APM to `async/await` should retain the functional behavior of the original program, both in the normal case and under exceptional circumstances. In 52% of all APM instances, `try-catch` blocks are in place to handle those exceptions. The `try-catch` blocks surround the `End` method invocation, which throws an exception if the background operation results in an exceptional circumstance. These catch blocks can contain business logic: for example, a network error sometimes needs to be reported to the user (“Please check the data or WiFi connection”). Code listing 4 shows such an example.

The naive approach to introducing `async/await` is to replace the `Begin` method invocation with an invocation to the corresponding TAP method, and await the result immediately. However, the `await` expression is the site that can throw the exception when the background operation failed. Thus, the exception would be thrown at a different site, and this can drastically change behavior. By introducing the `await` expression as replacement of the `End` method call at the exact same place, existing exception handling will work exactly as it did before. This is not a non-trivial insight for developers, because online examples of `async/await` only show the refactoring for extremely simple cases, where this is not a concern.

### Hidden End methods

The developer needs to take even more care when the `End` method is not immediately called in the `callback` lambda expression, but is ‘hidden’ deeper down the call chain. In that case, the `Task` instance must be passed down to where the `End` method invocation was to retain exceptional behavior. This requires an inter-procedural analysis of the code: each of the methods, through which the `IAsyncResult` parameter, with `Task<TResult>` parameter (with `T` being the return type of the TAP method, (2) replace the return type with `async Task<T>` or void with `async void` or `async Task`, and (3) introduce `ConfigureAwait(false) at each `await` expression. As shown in the results of the empirical study, when its presence is critical to retain UI responsiveness, developers almost never use `ConfigureAwait(false)` where it should be used. Code listing 5 shows such an example.

#### 4.1.2 Algorithm precondition

An invocation of a `Begin` method is a candidate for refactoring to `async/await` based constructs, if it adheres to the following preconditions and restrictions:

```csharp
void Button_Click(...) {
    WebRequest request = WebRequest.Create(url);
    request.BeginGetResponse(Callback, request);
}

void Callback(IAsyncResult ar) {
    WebRequest request = (WebRequest)ar.AsyncState;
    try {
        var response = request.EndGetResponse(ar);
        // Do something with successful response
    } catch (WebException e) {
        // Error handling
    }
}
```
Based on encountered cases in the analyzed code corpus, we have identified and (partially) implemented several such refactorings in Asyncifier. Examples are (1) identification of unused state arguments which can be replaced with null (solves violations of P5), and (2) rewriting of some callback argument expressions (solves violations of P3).

### 4.1.3 Refactoring APM instances

Asyncifier detects all Begin method invocations that fulfill the preconditions. It takes the following steps to refactor the APM instance to async/await-based constructs.

#### Traveling the call graph from APM Begin to End

First, Asyncifier explores the call graph of the body of the callback lambda expression to find the invocation path to the End invocation. It does a depth-first search of the call graph, by looking up the symbols of any non-virtual method that is encountered. There are two possible scenarios: the End method invocation (1) is placed directly in the lambda expression, or (2) it is found on the call graph of the lambda body in another method’s body. Code listing 6 is an example of the first case.

In the second case, Asyncifier identifies three different methods which are on the call graph path: (1) the initiating method, the method containing the Begin method invocation, (2) the result-obtaining method, the method containing the End method invocation, and (3) intermediate methods, the remaining methods on the path. Code listing 7 is an example of the second case. This example is used in the description of the following steps.

#### Rewriting the initiating method

In both cases, the initiating method needs to be rewritten. Asyncifier adds the async modifier to the signature of the initiating method. It changes the return value is to either Task instead of void, or Task<T> for any other return type T.

```csharp
void GetFromUrl(string url) {
    // Do something with response.
}
```

7
Rewriting the result-obtaining method

`Asyncifier` updates the signature of the result-obtaining method as follows: (1) it adds the `async` modifier, (2) it replaces return type `void` with `Task`, or any other `T` with `Task<T>`, and (3) it replaces the `AsyncResult` parameter with `Task<R>`, with `R` the return type of the `End` method.

```csharp
void Callback(Task<R> asyncResult, WebRequest request) {
    var response = asyncResult.ConfigureAwait(false);
    var content = response.GetResponseStream().ReadAsStringAsync();
    // Do something with response.
}
```

Rewriting intermediate methods

Intermediate methods must be rewritten if the `End` method is not invoked in the callback lambda expression body. `Asyncifier` recursively refactors every method recursively, applying the same steps as for the result-obtaining method. Additionally, at the call site of each method, the reference to the (removed) `result` parameter is replaced with a reference to the (newly introduced) `task` parameter.

### 4.1.4 Retaining original behavior

It is crucial that the refactored code has the same behavior in terms of scheduling as the original code. With both the `Begin` method and the `TAP` method, the asynchronous operation is started. In the `APM` case, the callback is only executed once the background operation is completed. With `async/await`, the same `happens-before` relationship exists between the `await` expression and the statements that follow the `await` of the `Task` returned by the `TAP` method. Because the statements in callbacks are placed after the `await` expression that pauses execution until completion of the background operation, this timing behavior is preserved.

4.1.5 Implementation limitations

The set of candidates is restricted by tool limitations related to re-use of `Begin` or `End` methods. First, there should not be other call graph paths leading from `Begin` method call to the target `End` method, which means so much as that the specific `End` method invocation must not be shared between multiple `Begin` invocations. Second, recursion in the callback through another `Begin` call that references the same callback again is not allowed (essentially, this is also sharing of an `End` method call). Third, `Asyncifier` does not support multiple `End` method invocations that correspond to a single `Begin` method invocation, for example through the use of branching. However, this case is very rare.

### 4.2 Corrector

We implemented another tool, `Corrector`, that detects and corrects common misuses that we explained in RQ4. `Corrector` gets the project file as an input and automatically corrects the misuses if it finds any without user intervention. Although this batch mode works to fix present misuses, it does not prevent users to make mistakes. Hence, `Corrector` also supports Quick Fix mode for Visual Studio. This mode shows a small icon close to the location of the misuse and offers a transformation to fix the problem, similar to the one in Eclipse.

1. **Fire & Forget methods**: There is no fix that can be automated for this misuse. If `fire & forget` method is converted to `async Task` method and is awaited in the caller, it will change the semantics. Therefore, the developer’s understanding of code is required to fix these cases.

2. **Unnecessary async/await methods**: `Corrector` checks whether `async` method body has only one `await` keyword and this `await` is used for a `TAP` method call that is the last statement of the method. `Corrector` does not do this for `async void` (`fire & forget`) methods; because if it removes `await` from the last statement in `async void` methods, it will silence the exception that can occur in that statement.

   To fix these cases, `Corrector` removes the `async` from the method identifiers and the `await` keyword from the `TAP` method call. The method will return the `Task` that is the result of `TAP` method call as shown in the examples of RQ4.

3. **Long-running operations under async methods**: To detect these operations, `Corrector` looks up symbols of each method invocation in the bodies of `async` methods. After getting symbol information, `Corrector` looks at the other members of the containing class of that symbol to check whether there is an asynchronous version. For instance, if there is an `x.Read()` method invocation and `x` is an instance of the `Stream` class, `Corrector` looks at the members of the `Stream` class to see whether there is a `ReadAsync` method that gets the same parameters and returns `Task`. By dynamically checking the members, `Corrector` can also find asynchronous versions not only in the .NET framework but also in third-party libraries.

   `Corrector` also maps corresponding blocking and non-blocking methods which do not follow the `async` suffix convention (e.g., `Thread.Sleep` -> `Task.Delay`).

   `Corrector` avoids introducing asynchronous operations of
file IO operations in loops, as this could result in slower performance than the synchronous version.

After finding the corresponding non-blocking operation, ASYNCifier simply replaces the invocation with the new operation and makes it await-ed.

(4) Unnecessarily capturing context: CORRECTOR checks whether there is a statement that access a GUI element (read or write) in the call graph of async method. It inspects every object’s symbol if the symbol is from System.Windows or Microsoft.Phone namespaces. All GUI elements are in these namespaces; but all constructs in these namespaces are not GUI elements. It makes our analysis conservative.

If CORRECTOR does not find any GUI element access after await points in async methods, it simply puts ConfigureAwait(false) as following TAP calls. Even though it is enough to put ConfigureAwait for one TAP call in the async method, it is good practice to put it for every TAP call in the async methods.

5. EVALUATION

5.1 Quantitative

To evaluate the usefulness of ASYNCifier and CORRECTOR we answer the following questions by using our code corpus:

EQ1: Are they applicable?

We executed ASYNCifier over our code corpus. After each transformation, ASYNCifier compiled the app in-memory and checked whether compilation errors were introduced. 54% of the 1245 APM instances adhere to the preconditions set in section 4.1.2, which were all successfully refactored. By manually checking 10% of all transformed instances, randomly sampled, we verified that ASYNCifier refactors APM instances correctly. In the 46% of unsupported APM instances, ASYNCifier does not touch the original program.

The two main causes for unsuccessful refactorings are (1) instances that do not adhere to preconditions, and (2) tool limitations. The former consist mostly of instances that can not be refactored because of fundamental limitations of the algorithm. Examples are callback expressions that can not be refactored because of fundamental limitations. The latter consist of the examples given in section 4.1.5.

We also applied CORRECTOR to the full corpus. All instances of type 2, 3, and 4 misuses were corrected automatically.

EQ2: What is the impact of refactoring on code?

ASYNCifier touches 28.9 lines on average per refactoring. It shows that these refactorings need automation support because they touch many lines of code.

CORRECTOR touches one line per each misuse of type (3) and (4) in Section 4.2. It touches 2 or 3 lines per each misuse of type (2); 2.1 lines on average.

EQ3: Is the tool efficient?

For ASYNCifier, the average time needed to refactor one instance is 508ms rendering ASYNCifier suitable for an interactive refactoring mode in an IDE.

Because the detection and fixing of type (2) and (3) misuses is straightforward, we did not measure the execution time. However, detecting type (4) misuse is expensive, as it requires inspection of the call graph of the async method. We found that analyzing one async method for this misuse takes on average 47ms. This shows that CORRECTOR can be used interactively in an IDE, even for type (4) misuse.

5.2 Qualitative evaluation

To further evaluate the usefulness in practice, we identified the 10 most recently updated apps that have APM instances. We applied ASYNCifier ourselves, and offered the modifications to the original developers as a patch via a pull request. 9 out of 10 developers responded, and accepted each one of our 28 refactorings.

We received very positive feedback on these pull requests. One developer would like to have the tool available right now: “I’ll look forward to the release of that refactoring tool, it seems to be really useful.” [33] The developer of phoneti-tartab [4] said that he had “been thinking about replacing all asynchronous calls [with] new async/await style code”. This illustrates the demand for tool support for the refactoring from APM to async/await.

For CORRECTOR, we selected the 10 most recently updated apps for all type (2) and (3) misuses. We did not especially select 10 apps for the type (4) misuse; but CORRECTOR did fix this misuse in the selected apps. In total, we selected 19 apps because one app had both type (2) and (3). Developers of 18 apps replied and accepted our all patches, corresponding to 149 instances of type (2), 38 instances of type (3), and 98 instances of type (4) misuses. In total 18 apps accepted 285 instances of CORRECTOR transformation.

Response to the fixes that removed unnecessary async/await keywords was similarly positive. One developer pointed out that he missed several unnecessary async/await instances that CORRECTOR detected: “[…] I normally try to take the same minimizing approach, though it seems I missed these.” [32] The developer of SoftbuildData [6] experienced performance improvements after removing unnecessary async/await: “[…] performance has been improved to 28 milliseconds from 49 milliseconds.” Again, these illustrate the need for tools that support the developer in finding problems in the use of async/await.

Furthermore, the developer of the playerframework [5] said that they missed the misuses because the particular code was ported from old asynchronous idioms. It demonstrates the need for ASYNCifier as it can help a developer to upgrade his or her code, without introducing incorrect usage of async/await.

6. DISCUSSION

6.1 Implications

Our study has practical implications for developers, researchers, and language and library designers.

Developers learn a new programming construct through both positive and negative examples. Robillard and DeLline [35] study what makes large APIs hard to learn and conclude that one of the important factors is the lack of usage examples. We provide hundreds of real-world examples of all asynchronous idioms on http://LearnAsync.net/. Because developers might need to inspect the whole source file or project to understand the example, we also link to highlighted source files on GitHub [39]. We also provide negative examples anonymously, without giving app names. Language and library designers can learn which constructs and idioms are embraced by developers, and which ones are
tedious to use or error-prone. Because some other major languages have plans to introduce similar constructs for asynchronous programming, this first study can guide them to an improved design of similar language constructs for their languages. For instance, capturing the context might not be the default: developers are very likely to forget to use ConfigureAwait(false).

Tool vendors can take advantage of our findings on async/await misuse. IDEs such as Visual Studio should have built-in quick fixes (similar to ours) to prevent users from introducing misuse. For instance, if developers introduce a fire & forget method, the IDE should give a warning unless the method is the top level event handler.

Researchers in the refactoring community can use our findings to target future research. For example, as we see from Table 1, the usage of Task jumped to 8% from 1% in WP8. This calls for work on a tool that converts old asynchronous idioms of CPU-bound computations (e.g. Thread) to new idioms (e.g. Task).

6.2 Threats to Validity

Internal: Is there something inherent to how we collect and analyze the usage that could skew the accuracy of our results? First, the study is only focusing on static usage of asynchronous constructs, but one use of a construct (i.e., a call site) could correspond to a large percentage of execution time, making it a very asynchronous program. Likewise, the opposite could be true. However, we are interested in the developer’s view of writing, understanding, maintaining, evolving the code, not on the performance tools’ view of the code (i.e., how much of the total running time is spent in asynchronous code). For our purposes, static analysis is much more appropriate.

External: Are the results representative? First, despite the fact that our corpus contains only open source apps, the 1378 apps span a wide domain, from games, social networking, and office productivity to image processing and third party libraries. They are developed by different teams with 3376 contributors from a large and varied community. Our code corpus contains all windows phone apps from GitHub and Codeplex without doing any random sampling or selection. While we answer our research questions for the Windows Phone ecosystem, we expect they can cross the boundary from mobile to any platform written in C# (e.g. desktop, web). Asynchronous programming is similar on those platforms: developers have access to the same async/await language constructs, and similar APIs.

Reliability: Are our empirical study and evaluation reliable? A detailed description of our results with fine-grained reports are available online.

Because we used an internal version of Microsoft’s Roslyn, we had to sign an NDA, which prohibits us from releasing the binaries of any tool using it (AsyncAnalyzer, Asyncifier, and Corrector). We will be able to publish the tools based on a public release that we expect by late Fall ’13.

6.3 Future Work

Our study was limited to apps targeting the Windows Phone platform. However, we believe that the tools can also be used for apps targeting other C# platforms, such as desktop, web (ASP.NET) and console apps. Future work would entail a study of asynchronous programming on those platforms similar to the one presented in this paper.

The refactoring tool that replaces APM instances with async/await-based code has several limitations, as mentioned in section 4.1.5. We plan to remove those limitations, and we expect to be able to show that the success rate of the refactoring tool will increase to 65%.

As soon as there is a publicly available version of Roslyn, we plan to update and release all the now-unreleased tools.

7. RELATED WORK

Empirical Studies: There are several empirical studies [9, 20, 26, 29] on the usage of libraries or programming language constructs. To the best of our knowledge, there is no empirical study on asynchronous constructs and language constructs for asynchronous programming.

We have previously conducted an empirical study [28] on how developers from thousands of open source projects use Microsoft’s Parallel Libraries. There is a small intersection between asynchronous and parallel libraries: only Thread, Task, and ThreadPool constructs. In this paper, we studied these three constructs as 3 of 5 different approaches for asynchronous CPU-bound computations.

Refactoring Tools: Traditionally, refactoring tools have been used to improve the design of sequential programs. There are a few refactoring tools that specifically target concurrency. We have used refactoring [13, 14] to retrofit parallelism into sequential applications via concurrent libraries. In the same spirit, Wloka et al. present a refactoring for replacing global state with thread local state [40]. Schafer et al. present Relocker [36], a refactoring tool that lets programmers replace usages of Java built-in locks with more flexible locks. Gyori et al. present Lambdafacator [21], that refactors existing Java code to use lambda expressions to enable parallelism.

To the best of our knowledge, there is no refactoring tool that specifically targets asynchronous programming. In industry, ReSharper is a well-known refactoring tool, but it does not support async/await-specific refactorings [34]. Our refactoring helps developer design responsive apps, which is the area never explored so far [12].

8. CONCLUSION

Because responsiveness is very important on mobile devices, asynchronous programming is already a first-class citizen in modern programming environments. However, the empirical research community and tool vendors have not yet similarly embraced it.

Our large-scale empirical study of Windows Phone apps provides insight into how developers use asynchronous programming. We have discovered that developers make many mistakes when manually introducing asynchronous programming based on the modern C# language features async/await. We provide a toolkit to support developers in preventing and curing these mistakes. Our toolkit (1) safely refactors legacy callback-based asynchronous code to async/await, (2) detects and fixes existing errors, and (3) prevents introduction of new errors. Evaluation of the toolkit shows that it is highly applicable, and developers already find the transformations very useful and are looking forward to using our toolkit.

We hope that our study motivates other follow-up studies to fully understand the state of the art in asynchronous programming.
9. REFERENCES

[38] Don Syme, Tomas Petricek, and Dmitry Lomov. The F# asynchronous programming model. In *Proceedings of the 13th international conference on Practical

