

# Copilot for Xcode: Exploring AI-Assisted Programming by Prompting Cloud-based Large Language Models

Chee Wei Tan<sup>1</sup>, Shangxin Guo<sup>2</sup>, Man Fai Wong<sup>2</sup> and Ching Nam Hang<sup>2</sup>

<sup>1</sup>Nanyang Technological University

<sup>2</sup>City University of Hong Kong

cheewei.tan@ntu.edu.sg, {sxguo2-c,mfwong29-c,cnhang3-c}@my.cityu.edu.hk

## Abstract

This paper presents an AI-assisted programming tool called *Copilot for Xcode* for program composition and design to support human software developers. By seamlessly integrating cloud-based Large Language Models (LLM) with Apple’s local development environment, Xcode, this tool enhances productivity and unleashes creativity for software development in Apple software ecosystem (e.g., iOS apps, macOS). Leveraging advanced natural language processing (NLP) techniques, *Copilot for Xcode* effectively processes source code tokens and patterns within code repositories, enabling features such as code generation, autocompletion, documentation, and error detection. Software developers can also query and make “small” decisions for program composition, some of which can be made simultaneously, and this is facilitated through prompt engineering in a chat interface of *Copilot for Xcode*. Finally, we present simple case studies as evidence of the effectiveness of utilizing NLP in Xcode to prompt popular LLM services like OpenAI ChatGPT for program composition and design.

## 1 Introduction

The field of natural language processing (NLP) has witnessed remarkable achievements through the use of large language models (LLMs). These models exhibit remarkable skills in understanding and generating natural languages. Additionally, they employ feedback mechanisms, such as rewards or penalties, to improve their comprehension and fine-tune their future performance [Christiano *et al.*, 2017; Ouyang *et al.*, 2022]. The application of LLMs to AI-assisted programming has recently attracted considerable attention [Rajamani, 2022; Wong *et al.*, 2023], as it offers the possibility to embed advanced conversational agents in software development [Li *et al.*, 2022b; Chen *et al.*, 2021]. In fact, the emergence of LLM-driven tools like ChatGPT (Chat Generative Pre-Trained Transformer), Github Copilot, DeepMind’s AlphaCode resonates with the visionary ideas presented in Edsger W. Dijkstra’s seminal paper in [Dijkstra, transcribed 2007], illustrating the transformative potential of computers

in facilitating a seamless integration of code and creativity. By surpassing the boundaries of debugging, AI-assisted programming tools can embrace the harmonious combination of program composition and elegant design [Dijkstra, transcribed 2007].

Since the work in [Dijkstra, transcribed 2007], one of the earliest AI-assisted programming tool is the *MIT programmer’s apprentice*, which aimed to simulate a knowledgeable junior programmer and utilized natural language processing to acquire understanding of programming patterns, clichés, and interactions [Waters, 1982; Rich and Waters, 1988]. The “MIT programmer’s apprentice” played a pioneering role in introducing revolutionary concepts such as code generation (e.g., see [Handsaker, 1982]) and an early form of “prompt engineering” (e.g., see [Rich *et al.*, 1978]). These advancements were driven by the recognition of computer programming as a systematic process of abstraction and simplification [Dijkstra, transcribed 2007; Rich and Waters, 1982].

AI-assisted programming improves software productivity by automating tasks, detects errors, enhances code quality, promotes usability, improves reliability and accelerates the overall software development cycles [Wong *et al.*, 2023]. Rather than replacing human programmers, these tools empower them to unleash their creative potential. By automating repetitive and mundane tasks, AI-assisted programming frees up valuable time and mental energy for human programmers to focus on innovative problem-solving and designing elegant solutions with the help of predictive analysis. Furthermore, by incorporating natural language processing capabilities (i.e., via prompt engineering), these tools enable human programmers to interact with software systems in a more intuitive and human-like manner, thus streamlining the software development process [Dijkstra, 1972].

Cloud-based tools that leverage LLMs such as Codeium [Codeium, 2023], GitHub Copilot [Friedman, 2021], OpenAI ChatGPT [OpenAI, 2023a], and Amazon CodeWhisperer [Amazon, 2022], enable users to access their cloud-based LLM services and online resources through dedicated application programming interface (API) in an on-demand access. The pricing models for these tools vary depending on the complexity of the tool and the target audience. Some pricing models include enterprise pricing, subscription-based pricing, usage-based pricing, freemium (i.e., the tool is available for free, but additional premium

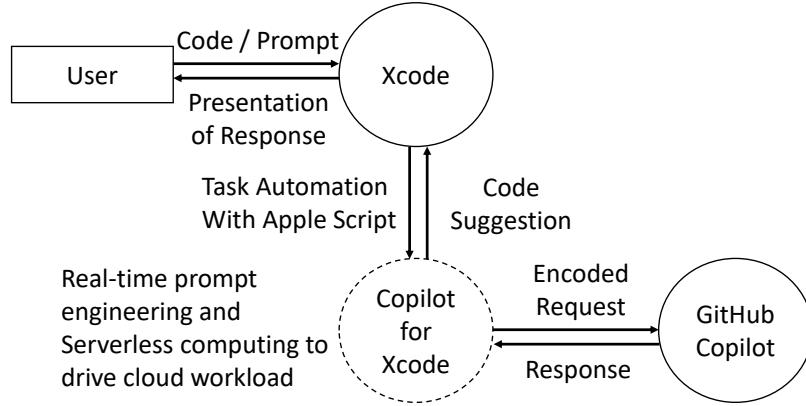


Figure 1: An overview of the AI-assisted programming application, *Copilot for Xcode* modeled as an intermediary software entity to connect user requests (e.g., prompt tokens) with cloud-based large language models.

features require payment), pay-per-use pricing or entirely free of charge. In fact, these pricing models and LLM-based services can be incorporated into existing systems like a local integrated development environment (IDE) or implemented via a Software-as-a-Service (SaaS) web interface, acting as a virtual service entity to meeting objectives and saving costs for the human programmer [Zheng *et al.*, 2015; Zheng *et al.*, 2016]. It is expected that the expanding reach and high demand usage of these LLM-based tools reflect the growing need for advanced NLP capabilities in software development. This trend aligns with Dijkstra’s visionary ideas as discussed in [Dijkstra, 1972; Dijkstra, transcribed 2007].

This paper presents *Copilot for Xcode*, an AI-assisted programming tool that was open-sourced on December 7, 2022, one week after OpenAI launched its ChatGPT on November 30, 2022.<sup>1</sup> Acting as an intermediary entity, as shown in Figure 1, it seamlessly integrates cloud-based large language model services with local IDEs like Xcode. This integration benefits software developers in the Apple ecosystem by streamlining AI-assisted programming service delivery and enhancing the accessibility of a myriad of cloud-based LLM applications. *Copilot for Xcode* enables real-time prompt engineering and efficient interaction between the human programmer and the large language models, offering the potential to integrate serverless computing capabilities with natural language processing in the cloud. The source code of *Copilot for Xcode* can be publicly accessed at <https://github.com/intitni/CopilotForXcode>.

## 2 Related Works

### 2.1 Language Models for Big Code Analysis

LLMs have surfaced as a promising approach to tackle challenges in computer programming, leveraging the software

<sup>1</sup>Apple’s recently-issued patent in [Siracusa *et al.*, 2023] dated June 27, 2023 suggests that they are actively exploring the integration of machine learning models into their software development system, specifically within Xcode, rather than relying solely on existing solution.

naturalness hypothesis [Hindle *et al.*, 2012]. This hypothesis posits that programming languages can be understood and manipulated in a similar fashion to how natural language processing techniques handle human languages. Since the introduction of the transformer architecture in 2017 [Vaswani *et al.*, 2017], LLMs trained on large-scale datasets of programs have shown significant benefits in code-related tasks by effectively learning programming language patterns and structures, which are collectively part of Big Code analysis [Vecchey *et al.*, 2016]. Recent LLMs such as T5 [Raffel *et al.*, 2020], BERT [Devlin *et al.*, 2018], GPT-4 [OpenAI, 2023b] and Palm 2 [Anil *et al.*, 2023] have demonstrated impressive capabilities in understanding and generating human-like text, opening up new possibilities for enhancing software engineers’ development experiences. These models undergo a two-step process involving pre-training and fine-tuning. Following these steps, prompt engineering can be applied to further optimize the model’s performance. As an integral part of AI-assisted programming, AI-based predictive analysis [Ji *et al.*, 2020] can anticipate potential issues in a software development life cycle. For example, it can proactively identify and flag critical incidents [Surameery and Shakor, 2023] before they manifest [Talamadupula, 2021].

### 2.2 AI-assisted Programming

AI-assisted programming is the incorporation of machine learning techniques and tools into the software development process [Mozannar *et al.*, 2022] to improve computer programming tasks. This concept shares similarities with pair programming [Bird *et al.*, 2022; Imai, 2022], whereby two human programmers collaborate to develop software by alternating between writing code (driver) and reviewing (observer) in a continuous switch. AI-assisted programming essentially replaces one of the two human programmers with an AI assistant, akin to the aforementioned “MIT programmer’s apprentice” [Waters, 1982; Rich and Waters, 1988]. The AI assistant automates tasks that can be broadly classified into two categories: generation and understanding. Generation tasks encompass activities such as code generation [Walderer and Lee, 1969; Manna and Walderer, 1971],

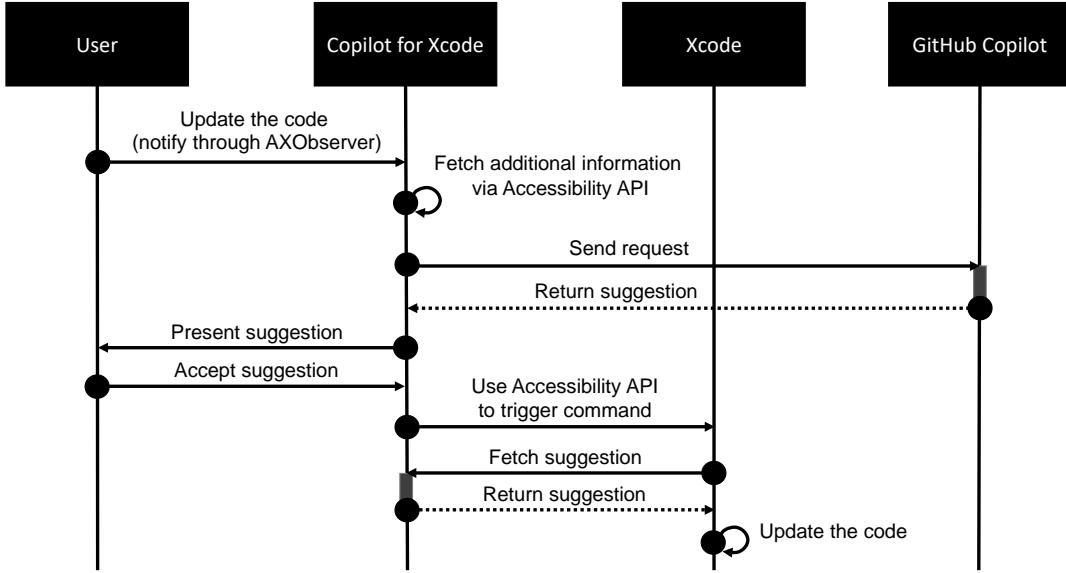


Figure 2: The sequence diagram illustrates the functionality of *Copilot for Xcode*, enabling real-time suggestions through integration with GitHub Copilot. When a user initiates a code update, *Copilot for Xcode* receives a notification and subsequently sends a request to the GitHub Copilot API. Upon receiving the suggestions from GitHub Copilot, the user has the option to adopt the recommendations and directly apply the changes within Xcode.

code completion [Robbes and Lanza, 2008; Bruch *et al.*, 2009], code translation [Acharya *et al.*, 2007; Allamanis *et al.*, 2014], code refinement [Saha *et al.*, 2017], and code summarization [Sridhara *et al.*, 2010; Sridhara *et al.*, 2011]. On the other hand, understanding tasks encompass activities like defect detection [Charniak, 1996] and clone detection [Kontogiannis *et al.*, 1996]. Improving the quality of large language models for these tasks focus on enhancing pre-training schemes [Li *et al.*, 2022b], expanding training corpora [Hussain *et al.*, 2019], and employing improved evaluation metrics [Chen *et al.*, 2021]. The GitHub Copilot [Friedman, 2021] is an example of an AI-powered programming tool that utilizes OpenAI Codex, which is based on GPT-3 LLM that has been trained on a vast amount of source code from the GitHub repository, totaling over 159GB [OpenAI, 2023a]. For further details on the latest advancements in AI-assisted programming, please see [Wong *et al.*, 2023].

### 3 Copilot for Xcode

#### 3.1 Xcode and its Offline Functionalities

Xcode [Apple, 2003] is an IDE created by Apple for developing software applications for the Apple ecosystem such as macOS and iOS. It provides a comprehensive set of tools, including editors, compilers, debugging tools, and interface builders, to help software developers create and maintain their applications. Xcode includes a source code editor with features like syntax highlighting, code completion, and refactoring capabilities. It supports multiple programming languages, including Swift, Objective-C, C, and C++, allowing developers to write code for a variety of Apple platforms. In addition to the code editor, Xcode offers a wide range of tools to assist in app development, such as an Interface Builder for

designing user interfaces visually, a graphical debugger for finding and fixing issues in code, and various performance analysis instruments. It also integrates with other developer tools, such as the iOS Simulator, which allows software developers to test their apps on virtual devices, and Instruments, a powerful profiling tool for measuring and optimizing app performance. Despite the extensive functionalities of Xcode, it has some limitations. For example, certain features depend on an offline rule-based system and may necessitate batch updates from Apple. Consequently, these services may not remain up-to-date consistently.

#### 3.2 The Copilot for Xcode Framework

The main limitation in Xcode is the sandboxing mechanism, which restricts plugin access to specific resources and prevents the launching of other programs. We address ways to overcome this limitation in order to enable the functionality of GitHub Copilot in Xcode. In particular, GitHub Copilot requires an additional program provided by GitHub to be executed alongside the plugin. Let us take the real-time suggestion feature as an example: the application first needs to bypass the sandbox in order to run the GitHub Copilot language server. This is accomplished by establishing communication between the Xcode source editor extension and a non-sandboxed XPC Service, which acts as a cross-process call service that facilitates the communication between the extension and the GitHub Copilot server. The server then presents suggestions in a user interface (UI) that is not managed by Xcode. To assemble a request for the language server, the application must gather sufficient information, but Xcode only provides the source code and file type. To obtain additional information without relying on Xcode, the application leverages the Accessibility API from the software devel-

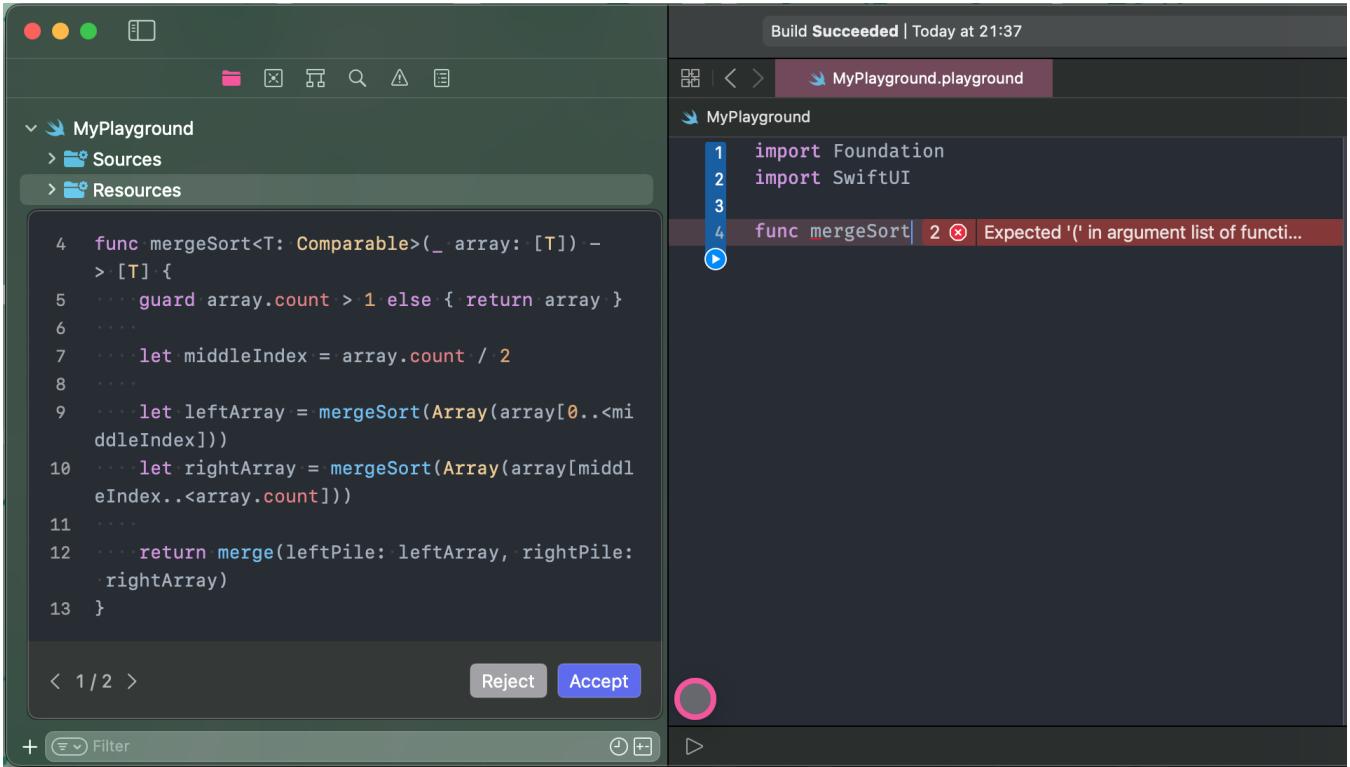


Figure 3: The user interface of *Copilot for Xcode* demonstrates the code suggestion capability, specifically showcasing the integration of GitHub Copilot for real-time suggestions related to the merge sort algorithm. In the accompanying Figure, the right-hand side displays the open source code editor within Xcode, focused on an interactive Swift playground. On the left-hand side, the cloud-based services deliver code suggestion responses.

opment kit. This particular API exposes information about each object within the application. Furthermore, to enable in-place code editing, the application executes extension commands programmatically. This is accomplished by utilizing the Accessibility API to interact with the menu bar items. These implementations thus allow Apple software developers to leverage GitHub Copilot [Friedman, 2021] and Codeium [Codeium, 2023] for code suggestions, while utilizing ChatGPT [OpenAI, 2023a] for code explanations, generation and natural language-based code modifications. The technical interaction of integrating Copilot with Xcode are depicted in Figure 2.

In addition, it facilitates the integration of an external chat panel that can access and read the user’s code. This chat panel serves as a connection point to leverage LLMs for functionalities such as code explanation and mutation using natural language. The chat panel can also be extended with plugins to offer additional features, including support for answering questions with real-time information from search engines. Some latest cloud-based LLM provide direct access through their official APIs for direct integration. In particular, *Copilot for Xcode* leverages the LangChain [Chase, 2022] framework, which facilitates the creation of customized LLMs tailored to specific use cases. This framework significantly enhances the prompt engineering process [Wu *et al.*, 2022; Poldrack *et al.*, 2023], allowing for the design of more effec-

tive prompts that can be utilized with the LLMs. This integration and framework combination optimize the functionality and usability of the LLMs, providing users with enhanced capabilities and improved prompt customization.

### 3.3 Code Suggestion

The code suggestion function offers a viable option for code completion and generation under diverse usage scenarios. Code completion, commonly referred to as auto-completion [Wong *et al.*, 2023], is an invaluable feature in software development that assists in completing unfinished code segments. On the other hand, code generation involves the automatic generation of source code from natural language input [Li *et al.*, 2022a], guided by user-defined constraints. This capability strengthens the efficiency on the development process by automating the creation of code based on linguistic specifications provided by the user.

In *Copilot for Xcode*, we offer real-time code suggestions that dynamically update whenever users modify their code. This capability, depicted in Figure 3, is powered by the integration of GitHub Copilot and Codeium, ensuring that the suggestions are specifically tailored to the files currently open in the workspace, thus enhancing productivity and accuracy while leveraging the capabilities of the code suggestion function. The feature offers two presentation modes for displaying suggestions, which includes two distinct modes. In the “Nearby Text Cursor” mode, suggestions are presented based

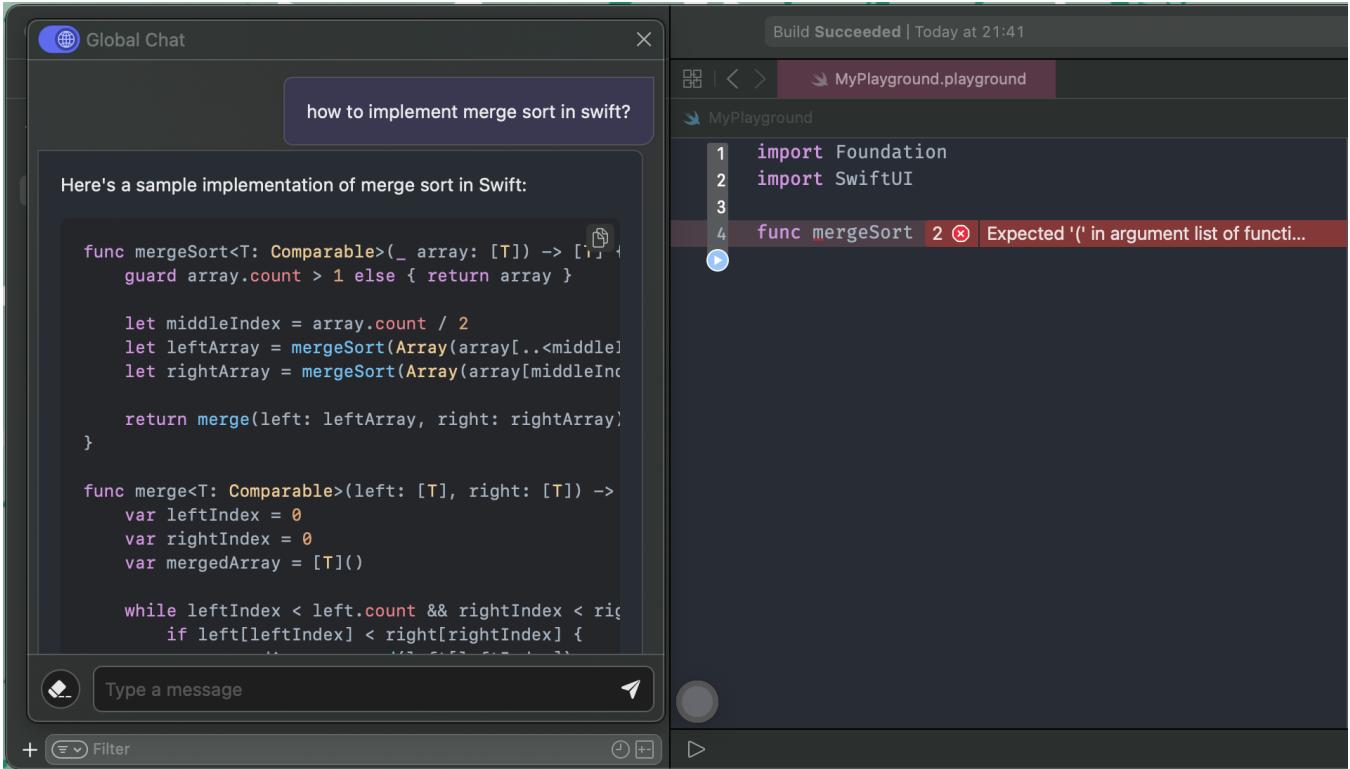


Figure 4: The user interface of *Copilot for Xcode* shows its Chat and prompt-to-code features, which enable code generation for the merge sort algorithm. These features are connected to ChatGPT, allowing for online prompt engineering and code generation within Xcode. The Figure illustrates the source code editor in Xcode on the right-hand side, while the chat conversational panel is displayed on the left-hand side.

on the current position of the text cursor. On the other hand, the “Floating Widget” mode displays suggestions in close proximity to the circular widget. When the user updates their code, the integrated application retrieves and integrates relevant suggestions for display within Xcode.

The software development experience is further enhanced by a range of pre-defined commands offered by *Copilot for Xcode*. The first useful command is Get Suggestions, which retrieves customized suggestions based on the current cursor position in the edited file on Xcode. In cases where multiple suggestions are available, users can conveniently navigate through them using the Next Suggestion and Previous Suggestion commands to choose the code suggestions based on their preferences. When incorporating suggested code, the Accept Suggestion command comes in handy to immediately select the code suggestion, while the Reject Suggestion command allows users to remove unnecessary suggestions along with their associated comments. Furthermore, there are two commands specifically designed for the usage of *Copilot for Xcode*. The Real-time Suggestions command, which can only be called by the *Copilot for Xcode* automatically, provides real-time suggestions after a successful retrieval so that the code suggestion can be presented in Xcode, while the Prefetch Suggestions, which can also be called by the *Copilot for Xcode*, command proactively fetches real-time suggestions in

the background, thus improving the overall responsiveness.

### 3.4 Chat and Prompt-to-Code for Code Generation

*Copilot for Xcode* provides its chat and prompt-to-Code features for code generation, as depicted in Figure 4. This functionality focuses on generating code from text inputs, enabling text-to-code generation within the IDE. By incorporating these advanced code generation capabilities, *Copilot for Xcode* enhances coding workflows, making them more efficient and intuitive. The chat function of our application, also powered by ChatGPT, complements these code generation features and offers additional enhancements for an interactive coding experience. Users can leverage specific features customized to their programming needs, such as extracting selected code in the active editor for reference and discussion of specific code snippets. Access to the relative path of the file being worked on facilitates easy navigation within the codebase. The chat or prompt-to-code functions also assist in capturing error and warning labels in the active editor, enabling swift issue resolution. Users can also obtain information about the text cursor location, facilitating precise discussions and context-aware conversations. These combined features empower users to engage in productive coding discussions and streamline their coding process, harnessing the capabilities of our AI-powered application.

The prompt-to-Code function offers a range of capabilities for code modification and creation using natural language. It is particularly beneficial when there is a need to update a specific section of code. This feature provides various use cases, such as enhancing code readability, rectifying code bugs, including documentation within the code, dividing extensive functions into smaller, manageable ones, generating code based on specific templates using custom commands, refining grammar and spelling errors in documentation, and facilitating the translation of localizable strings files. With "Prompt to Code," users can refactor existing code or write new code by harnessing the power of natural language.

## 4 Evaluation

We describe three case studies that illustrate the power of *Copilot for Xcode* in tackling real-world programming challenges through AI-assisted programming. The case studies presented here are based on real-world programming assignments given to undergraduate students. Furthermore, the case studies also highlight the significance of prompt engineering for code suggestion query and making “small” decisions for program composition and design in *Copilot for Xcode*.

### 4.1 Case Study: HCF of Two Numbers

The first case study considers computing the Highest Common Factor (HCF) or the Greatest Common Divisor (GCD) of two natural numbers [Dijkstra, 2007]. The HCF/GCD represents the largest positive integer that divides both numbers, leaving no remainder. Many existing approaches can be used to solve this problem, including subtraction, brute-force, and binary algorithm. One of the oldest algorithms to compute the HCF for two given natural numbers is the Euclidean algorithm. We tasked students to implement the Euclidean algorithm using the Swift programming language based on the observation that if  $r$  is the remainder when  $a$  is divided by  $b$ , then the HCF of  $a$  and  $b$  is equivalent to the HCF of  $b$  and  $r$ . Figure 5 depicts the brute-force approach for the HCF of two natural numbers, while Figure 6 provides a correct depiction of the HCF calculation using the Euclidean algorithm.

**Prompt:** HCF of Two Numbers

**Result:**

```
func findHCF(_ a: Int, _ b: Int) -> Int {
    let num1 = abs(a)
    let num2 = abs(b)

    var hcf = 1
    let smallerNum = min(num1, num2)

    for i in 1...smallerNum {
        if num1 % i == 0 && num2 % i == 0 {
            hcf = i
        }
    }

    return hcf
}
```

Figure 5: Code suggestion of HCF of two natural numbers using a brute-force algorithm without specific instructions in the prompt.

**Prompt:** HCF of Two Numbers by Euclidean Algorithm

**Result:**

```
func findHCF(_ a: Int, _ b: Int) -> Int {
    var num1 = abs(a)
    var num2 = abs(b)

    // Ensure num1 is greater than or equal to num2
    if num1 < num2 {
        let temp = num1
        num1 = num2
        num2 = temp
    }

    // Apply the Euclidean algorithm
    while num2 != 0 {
        let remainder = num1 % num2
        num1 = num2
        num2 = remainder
    }

    return num1
}
```

Figure 6: Code suggestion of HCF of two natural numbers with specific instructions on using the Euclidean algorithm.

### 4.2 Case Study: LCM of Two Numbers

The Least Common Multiple (LCM) of two natural numbers  $a, b$  refers to the smallest positive integer that is divisible by both numbers. Notably, this LCM is dual to the HCF/GCD [Dijkstra, 2007].<sup>2</sup> Typically, the LCM algorithm makes use of the HCF algorithm. However, in this assignment, a unique requirement is to develop an LCM algorithm that does not rely on the HCF algorithm. By default, the code suggestion in *Copilot for Xcode* assumes that developers have already implemented the HCF function without any additional instructions in the prompt. The tool thus generates a result that utilizes the HCF implementation as a helper function for the LCM as shown in Figure 7. However, to comply with the specific prompt of not relying on the HCF, Figure 8 presents a correct answer for the LCM calculation without requiring to compute the HCF.

**Prompt:** LCM of Two Numbers

**Result:**

```
func findLCM(_ a: Int, _ b: Int) -> Int {
    let hcf = findHCF(a, b)
    // LCM = (a * b) / HCF
    let lcm = (abs(a) * abs(b)) / hcf
    return lcm
}
```

Figure 7: Code suggestion of LCM of two natural numbers using the HCF without specific instructions.

<sup>2</sup>The product of two natural numbers is equal to the product of their respective least common multiple and greatest common denominator. This principle can be demonstrated using the Fundamental Theorem of Arithmetic in number theory or through an algorithmic method outlined in [Dijkstra, 2007]. From our observations, the ChatGPT Codex (GPT-3) was able to understand this concept (even able to demonstrate a plausible proof with prompt engineering), although it faced difficulties in extending the duality to encompass more than two natural numbers.

**Prompt:** LCM of Two Numbers without Using the HCF

**Result:**

```
func findLCM(_ a: Int, _ b: Int) -> Int {
    let maxNum = max(abs(a), abs(b))
    var lcm = maxNum

    while true {
        if lcm % abs(a) == 0 && lcm % abs(b) == 0 {
            return lcm
        }
        lcm += maxNum
    }
}
```

Figure 8: Code suggestion of LCM of two natural numbers without using HCF algorithm as a helper function.

### 4.3 Case Study: Navigating App on iOS

In the example below, we delve into a code generation scenario that effectively highlights the fundamental concepts of SwiftUI in iOS app development, specifically focusing on view navigation. This scenario is visually represented in Figure 9. The iOS app, built with SwiftUI, comprises two distinct views: HomeView and DetailView, as depicted in Figure 10. View refer to a crucial component that constructs the UI and plays a pivotal role in displaying and handling the visual elements that users observe and engage with on the screen. To effectively manage views, software developers are required to arrange them in a hierarchical structure and personalize each view by configuring different properties.

**Prompt:** Create a navigating views app with SwiftUI

**Result:**

```
import SwiftUI

struct ContentView: View {
    var body: some View {
        NavigationView {
            HomeView()
                .navigationBarTitle("Navigation App")
        }
    }
}

@main
struct NavigationApp: App {
    var body: some Scene {
        WindowGroup {
            ContentView()
        }
    }
}
```

Figure 9: This SwiftUI-based app consists of two screens: a home screen and a detail screen. The ContentView sets up a navigation with the entry of the app and provides a navigation bar title.

## 5 Conclusion

This paper introduced *Copilot for Xcode* that integrates cloud-based large language model services (Github Copilot and OpenAI’s GPT) with Apple’s integrated development environment, Xcode, for AI-assisted programming. We also discussed the efficacy of prompt engineering and possible strategies for AI-assisted programming using simple case studies

**Prompt:** Create the HomeView and DetailsView with SwiftUI

**Result:**

```
struct HomeView: View {
    @State private var showingDetail = false
    var body: some View {
        VStack {
            Text("Home Screen").font(.largeTitle).padding()
            Button(action: {
                showingDetail.toggle()
            }) {
                Text("Go to Detail").font(.title).padding()
            }.padding().sheet(isPresented: $showingDetail) {
                DetailView()
            }
        }
    }
}

struct DetailView: View {
    @Environment(\.presentationMode) var presentationMode
    var body: some View {
        VStack {
            Text("Detail Screen").font(.largeTitle).padding()
            Button(action: {
                presentationMode.wrappedValue.dismiss()
            }) {
                Text("Go Back").font(.title).padding()
            }.padding()
        }
    }
}
```

Figure 10: The home and detail screen for a navigating app. When a button on the home screen is tapped, the app navigates to the detail screen. Additionally, a back button on the detail screen allows the user to navigate back to the home screen.

to illustrate the practical application of this tool to program composition and design. When designing a program, making small decisions often involves breaking down complex tasks into smaller components manageable by the large language model. By carefully constructing prompts, programmers can influence the generation of code and steer the language model’s understanding towards the desired outcome.

As a software prototype, *Copilot for Xcode* has some limitations to consider during practical usage. For example, to bypass the sandboxing restrictions, it employs unconventional methods to retrieve information like file and project/workspace paths. As such, it is important to be aware that this might not always function seamlessly in future versions of Xcode. Also, the current code suggestions are presented as C-style comments in comment mode, which can inadvertently disrupt a user’s code if they are working on a format, e.g., JSON file, where such comments are not applicable.

By combining the capabilities of large language models and integrated tools for prompt engineering, *Copilot for Xcode* enhances and streamlines the software development process within Apple’s Xcode. The integration of *Copilot for Xcode* with other cloud-based services like Xcode Cloud can also improve the overall productivity and efficiency in software development, which is especially important to continuous integration (CI) and continuous delivery (CD) in the software development pipeline. As AI-assisted programming tools like Copilot get incorporated into more IDEs, it brings us closer to the realization of Dijkstra’s vision in [Dijkstra, transcribed 2007], fostering a symbiotic relationship between human programmers and AI-powered tools to achieve more efficient and reliable software development.

## References

- [Acharya *et al.*, 2007] Mithun Acharya, Tao Xie, Jian Pei, and Jun Xu. Mining api patterns as partial orders from source code: From usage scenarios to specifications. In *6th Joint Meeting of The European Software Engineering Conference and The ACM SIGSOFT Symposium on The Foundations of Software Engineering*, pages 25–34, 2007.
- [Allamanis *et al.*, 2014] Miltiadis Allamanis, Earl T. Barr, Christian Bird, and Charles Sutton. Learning natural coding conventions. In *22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering*, page 281–293. Association for Computing Machinery, 2014.
- [Amazon, 2022] CodeWhisperer Amazon. AI code generator - amazon codewhisperer. <https://aws.amazon.com/codewhisperer>, 2022. Accessed on June 1, 2023.
- [Anil *et al.*, 2023] Rohan Anil, Andrew M Dai, Orhan Firat, Melvin Johnson, Dmitry Lepikhin, Alexandre Passos, Siamak Shakeri, Emanuel Taropa, Paige Bailey, Zhifeng Chen, et al. Palm 2 technical report. *arXiv preprint arXiv:2305.10403*, 2023.
- [Apple, 2003] Xcode Apple. Xcode 15 - apple developer. <https://developer.apple.com/xcode/>, 2003. Accessed on June 1, 2023.
- [Bird *et al.*, 2022] Christian Bird, Denae Ford, Thomas Zimmermann, Nicole Forsgren, Eirini Kalliamvakou, Travis Lowdermilk, and Idan Gazit. Taking flight with copilot: Early insights and opportunities of AI-powered pair-programming tools. *Queue*, 20(6):35–57, 2022.
- [Bruch *et al.*, 2009] Marcel Bruch, Martin Monperrus, and Mira Mezini. Learning from examples to improve code completion systems. In *7th Joint Meeting of The European Software Engineering Conference and The ACM SIGSOFT Symposium on The Foundations of Software Engineering*, pages 213–222, 2009.
- [Charniak, 1996] Eugene Charniak. *Statistical Language Learning*. MIT press, 1996.
- [Chase, 2022] Harrison Chase. Langchain, 2022.
- [Chen *et al.*, 2021] Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, et al. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*, 2021.
- [Christiano *et al.*, 2017] Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 2017.
- [Codeium, 2023] Exafunction Codeium. Codeium - free AI code completion & chat. <https://codeium.com/>, 2023. Accessed on June 1, 2023.
- [Devlin *et al.*, 2018] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [Dijkstra, 1972] Edsger W Dijkstra. The humble programmer. *Communications of the ACM*, 15(10):859–866, 1972.
- [Dijkstra, 2007] Edsger Wybe Dijkstra. Defining the greatest common divisor. *E. W. Dijkstra Archive (EWD 1257)*, 2007.
- [Dijkstra, transcribed 2007] Edsger Wybe Dijkstra. A preliminary investigation into computer assisted programming. *E. W. Dijkstra Archive (EWD 237)*, (transcribed) 2007.
- [Friedman, 2021] Nat Friedman. Introducing github copilot: your AI pair programmer, 2021.
- [Handsaker, 1982] Robert E. Handsaker. Code generation in the programmer’s apprentice. Working Paper 233, MIT AI Lab, May 1982.
- [Hindle *et al.*, 2012] Abram Hindle, Earl T Barr, Zhendong Su, Mark Gabel, and Premkumar Devanbu. On the naturalness of software. In *2012 34th International Conference on Software Engineering*, pages 837–847. IEEE, 2012.
- [Husain *et al.*, 2019] Hamel Husain, Ho-Hsiang Wu, Tiferet Gazit, Miltiadis Allamanis, and Marc Brockschmidt. CodeSearchNet challenge: Evaluating the state of semantic code search. *arXiv preprint arXiv:1909.09436*, 2019.
- [Imai, 2022] Saki Imai. Is github copilot a substitute for human pair-programming? an empirical study. In *Proceedings of the ACM/IEEE 44th International Conference on Software Engineering: Companion Proceedings*, pages 319–321, 2022.
- [Ji *et al.*, 2020] Yangfeng Ji, Antoine Bosselut, Thomas Wolf, and Asli Celikyilmaz. The amazing world of neural language generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Tutorial Abstracts*, pages 37–42, 2020.
- [Kontogiannis *et al.*, 1996] Kostas A Kontogiannis, Renator DeMori, Ettore Merlo, Michael Galler, and Morris Bernstein. Pattern matching for clone and concept detection. *Automated Software Engineering*, 3(1-2):77–108, 1996.
- [Li *et al.*, 2022a] Yujia Li, David Choi, Junyoung Chung, Nate Kushman, Julian Schrittweis, Rémi Leblond, Tom Eccles, James Keeling, Felix Gimeno, Agustin Dal Lago, et al. Competition-level code generation with alphacode. *Science*, 378(6624):1092–1097, 2022.
- [Li *et al.*, 2022b] Zhiyu Li, Shuai Lu, Daya Guo, Nan Duan, Shailesh Jannu, Grant Jenks, Deep Majumder, Jared Green, Alexey Svyatkovskiy, Shengyu Fu, et al. Automating code review activities by large-scale pre-training. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2022.
- [Manna and Waldinger, 1971] Zohar Manna and Richard J Waldinger. Toward automatic program synthesis. *Communications of the ACM*, 14(3):151–165, 1971.
- [Mozannar *et al.*, 2022] Hussein Mozannar, Gagan Bansal, Adam Fourney, and Eric Horvitz. Reading between the lines: Modeling user behavior and costs in AI-assisted programming. *arXiv preprint arXiv:2210.14306*, 2022.

- [OpenAI, 2023a] OpenAI. Chatgpt: Optimizing language models for dialogue, Jan 2023.
- [OpenAI, 2023b] OpenAI. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.
- [Ouyang *et al.*, 2022] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744, 2022.
- [Poldrack *et al.*, 2023] Russell A Poldrack, Thomas Lu, and Gašper Beguš. AI-assisted coding: Experiments with gpt-4. *arXiv preprint arXiv:2304.13187*, 2023.
- [Raffel *et al.*, 2020] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 2020.
- [Rajamani, 2022] Sriram Rajamani. AI assisted programming. In *15th Annual ACM India Compute Conference, COMPUTE ’22*, page 5, New York, NY, USA, 2022. Association for Computing Machinery.
- [Rich and Waters, 1982] Charles Rich and Richard C. Waters. The disciplined use of simplifying assumptions. *ACM SIGSOFT Software Engineering Notes*, 7(5):150–154, December 1982.
- [Rich and Waters, 1988] Charles Rich and Richard C. Waters. The programmer’s apprentice: a research overview. *Computer*, 21(11):10–25, November 1988.
- [Rich *et al.*, 1978] Charles Rich, Howard E. Shrobe, Robert C. Waters, Gerald J. Sussman, and Carl E. Hewitt. Programming viewed as an engineering activity. AI Memo 459, Massachusetts Institute of Technology, January 1978.
- [Robbes and Lanza, 2008] Romain Robbes and Michele Lanza. How program history can improve code completion. In *23rd IEEE/ACM International Conference on Automated Software Engineering*, pages 317–326, 2008.
- [Saha *et al.*, 2017] Ripon K Saha, Yingjun Lyu, Hiroaki Yoshida, and Mukul R Prasad. Elixir: Effective object-oriented program repair. In *2017 32nd IEEE/ACM International Conference on Automated Software Engineering*, pages 648–659. IEEE, 2017.
- [Siracusa *et al.*, 2023] M. R. Siracusa, A. K. Katti, and et al. Integrating learning models into software development systems, June 27th 2023. US Patent and Trademark Office, US Patent 11,687,830.
- [Sridhara *et al.*, 2010] Giriprasad Sridhara, Emily Hill, Divya Muppaneni, Lori Pollock, and K Vijay-Shanker. Towards automatically generating summary comments for java methods. In *IEEE/ACM International Conference on Automated Software Engineering*, pages 43–52, 2010.
- [Sridhara *et al.*, 2011] Giriprasad Sridhara, Lori Pollock, and K Vijay-Shanker. Generating parameter comments and integrating with method summaries. In *2011 IEEE 19th International Conference on Program Comprehension*, pages 71–80. IEEE, 2011.
- [Surameery and Shakor, 2023] Nigar M Shafiq Surameery and Mohammed Y Shakor. Use chatgpt to solve programming bugs. *International Journal of Information Technology & Computer Engineering (IJITC) ISSN: 2455-5290*, 3(01):17–22, 2023.
- [Talamadupula, 2021] Kartik Talamadupula. Applied AI matters: Ai4code: Applying artificial intelligence to source code. *AI Matters*, 7(1):18–20, 2021.
- [Vaswani *et al.*, 2017] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017.
- [Vechev *et al.*, 2016] Martin Vechev, Eran Yahav, et al. Programming with “big code”. *Foundations and Trends® in Programming Languages*, 3(4):231–284, 2016.
- [Waldinger and Lee, 1969] Richard J Waldinger and Richard CT Lee. Prow: A step toward automatic program writing. In *1st International Joint Conference on Artificial Intelligence*, pages 241–252, 1969.
- [Waters, 1982] Richard C. Waters. The programmer’s apprentice: Knowledge based program editing. *IEEE Transactions on Software Engineering*, SE-8(1):1–12, January 1982.
- [Wong *et al.*, 2023] Man-Fai Wong, Shangxin Guo, Ching-Nam Hang, Siu-Wai Ho, and Chee-Wei Tan. Natural language generation and understanding of big code for AI-assisted programming: A review. *Entropy*, 25(6):888, 2023.
- [Wu *et al.*, 2022] Tongshuang Wu, Ellen Jiang, Aaron Donsbach, Jeff Gray, Alejandra Molina, Michael Terry, and Carrie J Cai. Promptchainer: Chaining large language model prompts through visual programming. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–10, 2022.
- [Zheng *et al.*, 2015] Liang Zheng, Carlee Joe-Wong, Chee Wei Tan, Mung Chiang, and Xinyu Wang. How to bid the cloud. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (SIGCOMM)*, pages 71–84, 2015.
- [Zheng *et al.*, 2016] Liang Zheng, Carlee Joe-Wong, Christopher G Brinton, Chee Wei Tan, Sangtae Ha, and Mung Chiang. On the viability of a cloud virtual service provider. In *Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science*, pages 235–248, 2016.