

# Research and Application of Automatic Programming Technology Based on GPT

Wenjiang Chen Yang Pan

Shenzhen Bytecode Information Technology Co., Ltd., Shenzhen, Guangdong, 518000, China

## Abstract

This paper introduces the research and application of automated programming technology based on GPT. Firstly, it introduces the concept and characteristics of automatic programming, and then expounds the principle and implementation method of automatic programming technology in detail. On this basis, this paper discusses the application of automatic programming technology in software development, including document generation, code generation, automated testing and other aspects, and combined with actual cases. The results show that the automated programming technology based on GPT can significantly improve the efficiency and quality of software development and has wide application prospects.

## Keywords

GPT; automated programming; document generation; test automation generation

## 基于 GPT 的自动化编程技术研究及应用

陈文奖 潘杨

深圳字节码信息科技有限公司, 中国·广东·深圳 518000

## 摘要

论文介绍了基于GPT的自动化编程技术的研究及应用。首先介绍了自动化编程的概念和特点,接着详细阐述了自动化编程技术的原理和实现方法。在此基础上,论文探讨了自动化编程技术在软件开发中的应用,包括文档生成、代码生成、自动化测试等方面,并结合实际案例进行了分析和验证。研究结果表明,基于GPT的自动化编程技术能够显著提高软件开发效率和质量,具有广泛的应用前景。

## 关键词

GPT; 自动化编程; 文档生成; 自动化测试生成

## 1 引言

随着信息技术的快速发展,软件在各个领域的应用越来越广泛,而软件开发的效率和质量是影响软件应用效果的重要因素之一。传统的软件开发过程中,编写需求文档、设计文档及编写代码环节是耗时且容易出错的环节之一。为了提高软件开发效率和质量,自动化编程技术应运而生,尤其是基于 GPT 的自动化编程技术,能够显著提高软件开发效率和质量,具有重要的应用价值和研究意义。

目前,国内外学者已经开展了大量关于自动化编程技术的研究工作,包括代码生成、模板方法、模型驱动、元模型等方面<sup>[1]</sup>。

## 2 自动化生成代码的介绍与应用

随着各行各业对软件的需求不断增长,软件开发的效

率及质量已成为人们关注的焦点。为解决这些问题,自动化代码生成已成为软件工程中的一种重要技术。其中,GPT 和 DSL 以及领域驱动设计作为核心方法,在自动化代码生成中有初步的应用。

GPT 是一种基于预训练模型的自然语言处理技术,其主要目的是使用语言模型来生成逻辑语句或代码;能够辅助需求分析人员自动编写需求文档及设计文档。DSL 是一种领域特定语言,它提供了一种用于表示某个领域特定问题的语言结构;能够简化复杂的业务的一种表达方式。领域驱动设计则是一种应对复杂软件开发的方法,它将软件开发过程与业务领域的结构和需求紧密结合起来。

### 2.1 GPT 在代码生成中的应用

GPT 作为一种自然语言处理技术,能够将根据人类给出的提示词返回需求的内容。通过 GPT,我们可以在不具备编程技能的情况下编写出复杂的文档及代码。例如,在用户界面上,我们可以设计类似于聊天室的界面,用户通过与系统进行对话,系统将会自动生成相应的文档及代码以满足

【作者简介】陈文奖(1979-),男,中国广东梅州人,本科,高级工程师,从事自动化编程、原生低代码等研究。

需求。

GPT 生成的代码质量有赖于语言模型的质量。因此，语言模型的训练是一个重要的研究领域。例如，我们可以使用深度学习技术来训练语言模型，以提高代码的质量和可读性。

### 2.2 DSL 在代码生成中的应用

DSL 是一种面向特定领域的语言，它以该领域特定的概念为基础。DSL 可以大大提高代码的可读性和可维护性。在特定领域中，我们可以使用 DSL 来表达视图模型、实体模型、业务逻辑等，并通过编译解释及利用微模板技术来生成高质量的文档及代码。

### 2.3 领域驱动设计在代码生成中的应用

领域驱动设计是软件开发中的一种方法，它强调将软件系统开发的过程与业务领域的结构和需求紧密结合起来。在代码生成中，尤其是 DSL 的开发中，领域驱动设计可以提供在设计领域特定语言时所需的高级抽象。它可以确保设计的领域模型与业务领域紧密相连，同时也可以帮助团队更好地了解业务模型并避免设计不必要的模式和实现。

在领域驱动设计中，我们应该注意到领域建模的重要性。设计 DSL 之前，我们需要深入了解特定领域，确定领域中的核心概念和相关特性，并将其抽象为领域模型。此外，领域驱动设计的另一个重要方面是系统的架构和实现。我们应该选择最适合该领域的架构，以及最佳的实现方法，以确保生成的代码的质量和可读性。

总之，前面介绍了 GPT、DSL 以及领域驱动设计在自动化代码生成中的应用。它们也有一个共同的目标：提高软件工程质量 and 效率。因此，在实际应用中，我们可以将它们的优点结合起来，利用自动化代码生成来缩短整个软件开发周期，降低开发成本，提高软件质量和可维护性。

## 3 自动化编程技术的原理和实现方法

### 3.1 自动化编程技术的基本原理

自动化编程技术的基本原理是利用机器学习、自然语言处理、人工智能等技术来自动生成和修改文档及代码。具

体而言，自动化编程技术通过分析 DSL 中的结构和语法规则，学习代码的模式和模板，并利用这些模式和模板生成新的代码。同时，自动化编程技术也可以通过对代码的学习和推理来自动完成代码的修复、重构和优化。

字节码软件工厂是一种自动化编程实现；该软件工厂由 VsCode 事件风暴插件，Hi 语言转换服务、中间语言转换服务等组成。字节码软件工厂架构如图 1 所示。

### 3.2 自动化编程技术的实现方法

自动化编程技术的实现方法包括以下几种：

①代码生成器：代码生成器是一种基于模板和规则的代码自动生成工具，它可以根据预设的模板和规则生成代码<sup>[2]</sup>。代码生成器的优点是速度快、易于使用，但是它的生成能力受限于预设的模板和规则。

②基于机器学习的代码自动生成：利用机器学习技术来分析源代码的模式和模板，学习代码的结构和语法规则，并利用这些模式和模板自动生成新的代码。基于机器学习的代码自动生成可以适应不同的代码场景，但是需要大量的训练数据。

③基于 GPT 的代码自动生成：利用 GPT 大模型技术来生成代码，可以通过调用现有 GPT 大模型来生成代码，并可以根据输入的代码生成新的代码。基于 GPT 的代码自动生成技术可以生成复杂的代码，且仅需要少量的训练数据及提示语。

### 3.3 自动化编程技术的应用范围和限制

自动化编程技术的应用范围非常广泛，包括但不限于以下领域：①文档编写：自动化编程技术可以实现需求文档、概要设计文档、详细设计文档的自动编写。②软件开发：自动化编程技术可以大大提高软件开发的效率和质量，减少代码错误和瑕疵。③自动化测试开发：自动化编程技术可以用于测试工具的开发，生成自动化测试用例和测试脚本，提高测试效率和准确性。

然而，自动化编程技术也存在一些限制：①精度限制：自动生成的代码可能存在一些语法错误和逻辑错误，需要经过人工检查和修正。②适用场景限制：自动化编程技术适用

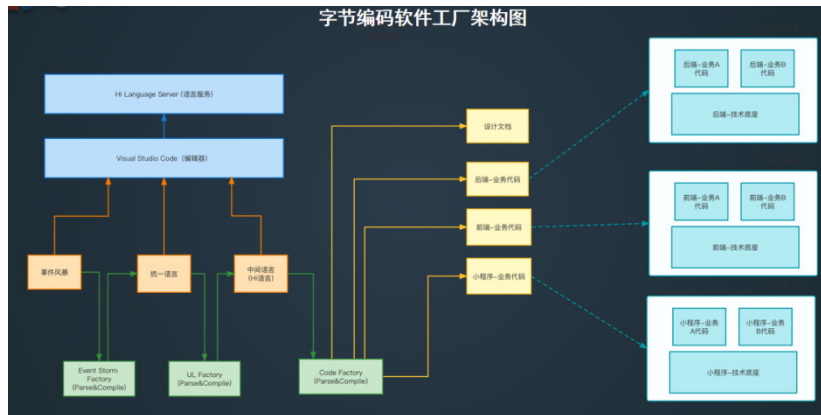


图 1 字节码软件工厂架构

于一些简单和重复性较高的编码任务，但对于一些复杂和创新性较强的编码任务，仍需要人工编写代码。③难以应对变化：由于软件需求和技术不断变化，自动化编程技术可能无法适应这些变化，需要不断更新和调整模型。

### 4 基于 GPT 的自动化编程技术的应用

#### 4.1 代码生成

##### 4.1.1 代码生成技术的概述

代码生成技术是指利用计算机程序自动化生成符合特定要求的文档及代码的技术。常见的代码生成技术包括基于模板的代码生成、基于规则的代码生成和基于 GPT 的代码生成等。其中，基于 GPT 的代码生成技术是近年开始的研究热点，利用 GPT 等技术可以生成高质量的文档及代码，大大提高了编写文档及代码的开发效率<sup>[3]</sup>。

##### 4.1.2 基于字节码软件工厂的代码生成技术的实现方法

基于字节码软件工厂的代码生成技术结合了 GPT、DSL 以及领域驱动设计；并可以通过以下步骤实现：①给出软件项目的简单描述，通过调用 GPT 及 L1 微模板库生成 L1 模型语言（即 DSL）及业务规则。②根据 L1 模型语言及业务规则生成项目的概要设计文档。③根据 L1 模型语言及业务规则调用 GPT 及 L2 微模板库并经过解析及编译生成 L2 模型语言及 L2 业务代码。④根据 L2 模型语言及 L2 业

务代码生成项目的详细设计文档。⑤根据 L2 模型语言及 L2 业务代码调用 L3 微模板库，解析及编译生成 L3 目标语言。

⑥通过编译 L3 目标语言，生成可运行应用程序。⑦其中 L2 模型语言及 L3 目标语言需求实现模型等价，行为等价及图灵完备；这就可以实现最大化自动代码生成；既可以完全不需要手动编写目标代码，又可以实现所有定制化开发。

字节码软件工厂实现原理如图 2 所示。

##### 4.1.3 实际案例分析和评价

以代码生成为例，基于 GPT 的自动化编程技术已经得到了一些成功的应用；目前在智能家居、电子商务、商超批发、大宗商品交易、供应链金融、避难监控等领域有初步应用。

通过实际案例的分析和评价可以看出，基于字节码软件工厂的代码生成技术在提高软件开发效率和质量方面具有显著的优势。

#### 4.2 自动化测试

自动化测试技术是指利用计算机程序自动执行测试用例来验证软件功能和性能的技术。自动化测试可以大大提高测试效率和准确性，节省人工测试的时间和成本。常见的自动化测试技术包括单元测试、集成测试、端到端测试和性能测试等。

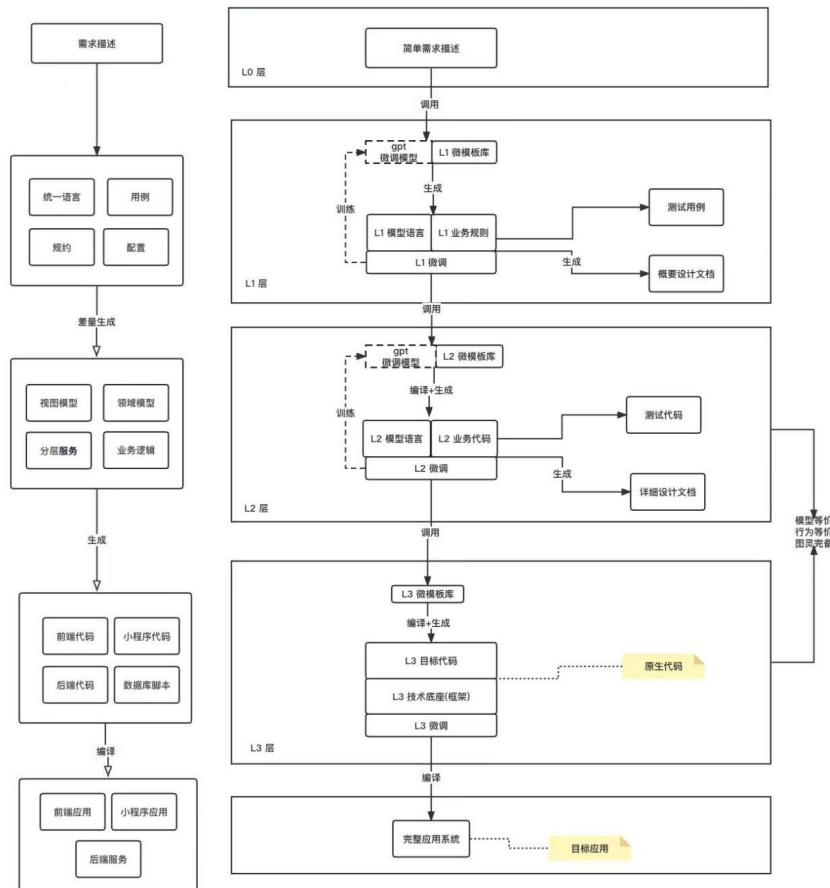


图 2 字节码软件工厂实现原理

#### 4.2.1 基于字节码软件工厂的自动化测试技术的实现方法

基于字节码软件工厂的自动化测试技术的实现方法，主要是通过 AIGC 的生成测试用例描述，再根据测试用例描述实现自动生成和执行测试用例代码的功能。具体实现步骤如下：①根据需求，使用 GPT 技术生成测试用例描述，再根据测试用例描述生成被测软件的测试代码。②对生成的测试代码进行修改和优化，以提高测试执行的效率和准确性。③使用集成工具，自动执行测试用例并生成测试报告。

#### 4.2.2 实际案例分析和评价

基于字节码软件工厂的自动化测试技术已经在实际的软件开发项目中得到初步应用。例如，一家互联网公司在开发新产品时，采用了基于字节码软件工厂的自动化测试代码生成技术实现可以达到 10 以上原生代码开发效率，自动生成的测试文档质量高于传统手写的测试文档。

与传统的手工测试相比，基于字节码软件工厂的自动化测试技术也具有以下优势：①可以自动生成和执行测试用

例，大大提高了测试效率和准确性。②可以通过自研工具的优化和修改，提高测试的覆盖率和执行效率。

## 5 结语

总的来说，基于字节码软件工厂为代表的 GPT 自动化编程技术和自动化测试技术是当前软件开发和测试领域的重要发展趋势。这些技术可以帮助分析设计人员、开发者和测试人员更加高效地完成各自的工作，提高软件的质量和可靠性。尤其是在大型软件系统的开发和测试中，这些技术的应用能够显著减少工作量，提高效率，并且避免人为的错误。

## 参考文献

- [1] 王艳平.LDPC码及其在中继通信环境下协作编码技术中的应用研究[D].成都:西南交通大学,2009.
- [2] 谭明新.Turbo码在第三代移动通信中的应用[J].哈尔滨工程大学学报,2003,2(18):14-15.
- [3] 柴先明,黄知涛.信道编码盲识别问题研究[J].通信对抗,2008(2):4.