



(12)发明专利申请

(10)申请公布号 CN 111190643 A
(43)申请公布日 2020.05.22

(21)申请号 201911398174.7

(22)申请日 2019.12.30

(71)申请人 苏州浪潮智能科技有限公司
地址 215100 江苏省苏州市吴中区吴中经济开发区郭巷街道官浦路1号9幢

(72)发明人 边智慧

(74)专利代理机构 北京集佳知识产权代理有限公司 11227
代理人 郗晨芳

(51)Int.Cl.
G06F 8/73(2018.01)

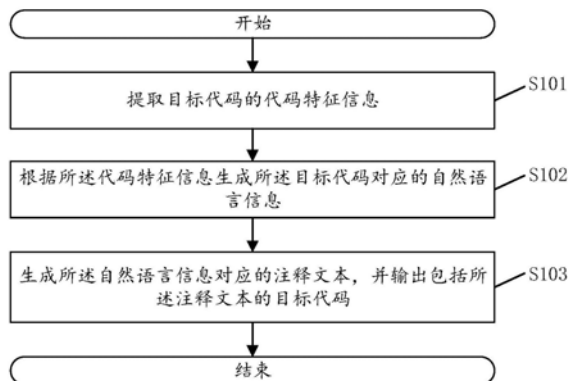
权利要求书2页 说明书7页 附图2页

(54)发明名称

程序代码注释的生成方法、系统、电子设备及存储介质

(57)摘要

本申请公开了一种程序代码注释的生成方法,所述生成方法包括提取目标代码的代码特征信息;其中,所述代码特征信息包括代码关键词、代码方法、代码动作和动作参数;根据所述代码特征信息生成所述目标代码对应的自然语言信息;生成所述自然语言信息对应的注释文本,并输出包括所述注释文本的目标代码。本申请能够自动生成程序代码的注释,提高代码注释效率。本申请还公开了一种程序代码注释的生成系统、一种电子设备及一种存储介质,具有以上有益效果。



1. 一种程序代码注释的生成方法,其特征在于,包括:
提取目标代码的代码特征信息;其中,所述代码特征信息包括代码关键词、代码方法、代码动作和动作参数;
根据所述代码特征信息生成所述目标代码对应的自然语言信息;
生成所述自然语言信息对应的注释文本,并输出包括所述注释文本的目标代码。
2. 根据权利要求1所述生成方法,其特征在于,在生成所述自然语言信息对应的注释文本之后,还包括:
记录所述注释文本中每一子文本与所述目标代码行数的对应关系;
相应的,输出包括所述注释文本的目标代码;
根据所述对应关系确定每一所述子文本对应的行数,根据所述子文本对应的行数输出包括所述注释文本的目标代码。
3. 根据权利要求1所述生成方法,其特征在于,所述提取目标代码的代码特征信息包括:
通过对所述目标代码递归遍历提取所述目标代码的代码特征信息。
4. 根据权利要求1所述生成方法,其特征在于,在根据所述代码特征信息生成所述目标代码对应的自然语言信息的过程中,还包括:
当检测到无法生成自然语言信息的异常代码特征信息时,利用代码解析池中的备选代码方法执行方法重构操作以便生成所述自然语言信息。
5. 根据权利要求4所述生成方法,其特征在于,所述利用代码解析池中的备选代码方法执行方法重构操作以便生成所述自然语言信息,包括:
确定所述异常代码特征信息中的方法的程序代码目的,将代码解析池中与所述程序代码目的对应的备选代码方法设置为目标备选代码方法;
将所述异常代码特征信息中的方法替换为所述目标备选代码方法得到备选代码特征信息;
利用所述备选代码特征信息执行解析操作,以便生成所述自然语言信息。
6. 根据权利要求1至5任一项所述生成方法,其特征在于,还包括:
当利用所述备选代码特征信息执行解析操作时,若检测到解析失败信息,则将所述异常代码特征信息中的方法替换为新的目标备选代码方法以便执行解析操作;
记录方法替换次数;
当所述方法替换次数大于预设值时,输出注释失败的提示信息。
7. 根据权利要求6所述生成方法,其特征在于,在输出包括所述注释文本的目标代码之后,还包括:
对通过方法替换得到的注释文本进行标记。
8. 一种程序代码注释的生成系统,其特征在于,包括:
特征提取模块,用于提取目标代码的代码特征信息;其中,所述代码特征信息包括代码关键词、代码方法、代码动作和动作参数;
自然语言生成模块,用于根据所述代码特征信息生成所述目标代码对应的自然语言信息;
注释模块,用于生成所述自然语言信息对应的注释文本,并输出包括所述注释文本的

目标代码。

9. 一种电子设备,其特征在于,包括存储器和处理器,所述存储器中存储有计算机程序,所述处理器调用所述存储器中的计算机程序时实现如权利要求1至7任一项所述程序代码注释的生成方法的步骤。

10. 一种存储介质,其特征在于,所述存储介质中存储有计算机可执行指令,所述计算机可执行指令被处理器加载并执行时,实现如上权利要求1至7任一项所述程序代码注释的生成方法的步骤。

程序代码注释的生成方法、系统、电子设备及存储介质

技术领域

[0001] 本申请涉及计算机技术领域,特别涉及一种程序代码注释的生成方法、系统、一种电子设备及一种存储介质。

背景技术

[0002] 随着IT行业的快速发展,开发者在开发程序的过程中的代码量也在逐步上升,很多程序员在开发的时候不愿意去添加注释,导致后来的开发或者维护人员在理解代码上耗费很大的精力。

[0003] 目前尚不存在具体的技术方案来完成实现程序代码注释的生成,相关技术中通过method refactoring(方法重构)来快速理解代码,这种方法是通过方法的调用来换一种方式表达,但是该方法也无法生成整个代码的注释。

[0004] 因此,如何自动生成程序代码的注释,提高代码注释效率是本领域技术人员目前需要解决的技术问题。

发明内容

[0005] 本申请的目的是提供一种程序代码注释的生成方法、系统、一种电子设备及一种存储介质,能够自动生成程序代码的注释,提高代码注释效率。

[0006] 为解决上述技术问题,本申请提供一种程序代码注释的生成方法,该程序代码注释的生成方法包括:

[0007] 提取目标代码的代码特征信息;其中,所述代码特征信息包括代码关键词、代码方法、代码动作和动作参数;

[0008] 根据所述代码特征信息生成所述目标代码对应的自然语言信息;

[0009] 生成所述自然语言信息对应的注释文本,并输出包括所述注释文本的目标代码。

[0010] 可选的,在生成所述自然语言信息对应的注释文本之后,还包括:

[0011] 记录所述注释文本中每一子文本与所述目标代码行数的对应关系;

[0012] 相应的,输出包括所述注释文本的目标代码;

[0013] 根据所述对应关系确定每一所述子文本对应的行数,根据所述子文本对应的行数输出包括所述注释文本的目标代码。

[0014] 可选的,所述提取目标代码的代码特征信息包括:

[0015] 通过对所述目标代码递归遍历提取所述目标代码的代码特征信息。

[0016] 可选的,在根据所述代码特征信息生成所述目标代码对应的自然语言信息的过程中,还包括:

[0017] 当检测到无法生成自然语言信息的异常代码特征信息时,利用代码解析池中的备选代码方法执行方法重构操作以便生成所述自然语言信息。

[0018] 可选的,所述利用代码解析池中的备选代码方法执行方法重构操作以便生成所述自然语言信息,包括:

[0019] 确定所述异常代码特征信息中的方法的程序代码目的,将代码解析池中与所述程序代码目的对应的备选代码方法设置为目标备选代码方法;

[0020] 将所述异常代码特征信息中的方法替换为所述目标备选代码方法得到备选代码特征信息;

[0021] 利用所述备选代码特征信息执行解析操作,以便生成所述自然语言信息。

[0022] 可选的,还包括:

[0023] 当利用所述备选代码特征信息执行解析操作时,若检测到解析失败信息,则将所述异常代码特征信息中的方法替换为新的目标备选代码方法以便执行解析操作;

[0024] 记录方法替换次数;

[0025] 当所述方法替换次数大于预设值时,输出注释失败的提示信息。

[0026] 可选的,在输出包括所述注释文本的目标代码之后,还包括:

[0027] 对通过方法替换得到的注释文本进行标记。

[0028] 本申请还提供了一种程序代码注释的生成系统,该程序代码注释的生成系统包括:

[0029] 特征提取模块,用于提取目标代码的代码特征信息;其中,所述代码特征信息包括代码关键词、代码方法、代码动作和动作参数;

[0030] 自然语言生成模块,用于根据所述代码特征信息生成所述目标代码对应的自然语言信息;

[0031] 注释模块,用于生成所述自然语言信息对应的注释文本,并输出包括所述注释文本的目标代码。

[0032] 本申请还提供了一种存储介质,其上存储有计算机程序,所述计算机程序执行时实现上述程序代码注释的生成方法执行的步骤。

[0033] 本申请还提供了一种电子设备,包括存储器和处理器,所述存储器中存储有计算机程序,所述处理器调用所述存储器中的计算机程序时实现上述程序代码注释的生成方法执行的步骤。

[0034] 本申请提供了一种程序代码注释的生成方法,包括提取目标代码的代码特征信息;其中,所述代码特征信息包括代码关键词、代码方法、代码动作和动作参数;根据所述代码特征信息生成所述目标代码对应的自然语言信息;生成所述自然语言信息对应的注释文本,并输出包括所述注释文本的目标代码。

[0035] 本申请向提取目标代码的代码关键词、代码方法、代码动作和动作参数等代码特征信息,由于代码特征信息描述了目标代码运行过程中的操作流程因此根据代码特征新生成对应的自然语言信息。本申请进一步生成资源语言信息对应的注释文本,进而得到了包括注释文本的目标代码。本申请能够自动生成程序代码的注释,提高代码注释效率。本申请同时还提供了一种程序代码注释的生成系统、一种电子设备和一种存储介质,具有上述有益效果,在此不再赘述。

附图说明

[0036] 为了更清楚地说明本申请实施例,下面将对实施例中所需要使用的附图做简单的介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域普通技术人

员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0037] 图1为本申请实施例所提供的一种程序代码注释的生成方法的流程图;

[0038] 图2为本申请实施例所提供的另一种程序代码注释的生成方法的流程图;

[0039] 图3为本申请实施例所提供的一种代码解析器的解析代码原理示意图;

[0040] 图4为本申请实施例所提供的一种程序代码注释的生成系统的结构示意图。

具体实施方式

[0041] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0042] 下面请参见图1,图1为本申请实施例所提供的一种程序代码注释的生成方法的流程图。

[0043] 具体步骤可以包括:

[0044] S101:提取目标代码的代码特征信息;

[0045] 其中,本实施例中所提到的目标代码可以为某个程序或多个程序的程序代码,该目标代码可以为Java语言编写的代码、可以为C语言编写的代码,本实施例不对目标代码对应的编程语言进行限定。在本步骤之前可以存在获取目标代码的操作,在得到目标代码后可以对目标代码执行代码特征提取操作得到代码特征信息。

[0046] 本实施例所提取的代码特征信息可以包括代码关键词、代码方法、代码动作和动作参数。代码关键词为编译器有特殊含义的字符串,是编译器和程序员的一个约定,可以利用代码关键词来告诉编译器其声明的变量类型、类、方法特性等信息。本实施例可以预先设置每一种编程语言对应的代码关键词表,在提取目标代码之前先确定目标代码的编程语言,进而在提取目标代码的代码关键词时可以将目标代码与代码关键词表执行匹配操作确定目标代码中的代码关键词。代码方法指程序代码中的函数,本实施例也可以预先设置函数表进而识别目标代码中的代码方法。代码动作指程序代码中的动作,如打印动作、输出动作等,动作参数为代码动作对应的参数如打印的地址、输出的端口等。作为一种可行的实施方式,本实施例可以通过对所述目标代码递归遍历提取所述目标代码的代码特征信息。

[0047] S102:根据所述代码特征信息生成所述目标代码对应的自然语言信息;

[0048] 其中,在得到代码特征信息之后可以通过解析得到对应的自然语言,具体的本实施例可以根据目标代码中代码特征信息的具体内容和对应的执行逻辑得到每一行目标代码对应的操作内容,即将编程语言转换为自然语言。通常的,自然语言信息可以为目标代码对应的英文信息。

[0049] S103:生成所述自然语言信息对应的注释文本,并输出包括所述注释文本的目标代码。

[0050] 其中,在得到了自然语言信息的基础上,本实施例可以将自然语言信息转化为特定语种的注释文本,例如转换为中文注释,并将注释文本与目标代码一同输出。本实施例还可以根据注释文本在开发过程中排查代码上的Bug,并生成对应的解决方法。

[0051] 本实施例向提取目标代码的代码关键词、代码方法、代码动作和动作参数等代码

特征信息,由于代码特征信息描述了目标代码运行过程中的操作流程因此根据代码特征新生成对应的自然语言信息。本实施例进一步生成资源语言信息对应的注释文本,进而得到了包括注释文本的目标代码。本实施例能够自动生成程序代码的注释,提高代码注释效率。

[0052] 作为对于图1对应实施例的进一步介绍,在图1对应实施例S103生成所述自然语言信息对应的注释文本之后,还可以记录所述注释文本中每一子文本与所述目标代码行数的对应关系,进而根据所述对应关系确定每一所述子文本对应的行数,根据所述子文本对应的行数输出包括所述注释文本的目标代码。

[0053] 进一步的,在根据所述代码特征信息生成所述目标代码对应的自然语言信息的过程中,当检测到无法生成自然语言信息的异常代码特征信息时,本实施例可以利用代码解析池中的备选代码方法执行方法重构操作以便生成所述自然语言信息。方法重构指通过方法的调用来换一种方式表达,通过分析试行各个方法来找到一个合适的易于理解的方法。

[0054] 作为对于图1对应实施例的进一步介绍,利用代码解析池中的备选代码方法执行方法重构操作以便生成所述自然语言信息可以包括以下步骤:

[0055] 步骤1:确定所述异常代码特征信息中的方法的程序代码目的,将代码解析池中与所述程序代码目的对应的备选代码方法设置为目标备选代码方法;

[0056] 步骤2:将所述异常代码特征信息中的方法替换为所述目标备选代码方法得到备选代码特征信息;

[0057] 步骤3:利用所述备选代码特征信息执行解析操作,以便生成所述自然语言信息。

[0058] 其中,代码解析池中可以包括多个备选代码方法,每一种代码方法都有其对应的程序代码目的,即对应的代码程序实现效果,本实施例可以预先设置代码解析池中每一备选代码方法对应的程序代码目的,在需要执行方法重构操作时可以从代码解析池中选择与异常代码特征信息中的方法的程序代码目的相同的备选代码方法设置为目标备选代码方法,执行代码方法替换操作得到备选代码特征信息,进而利用所述备选代码特征信息执行解析操作生成所述自然语言信息。

[0059] 进一步的,当利用所述备选代码特征信息执行解析操作时,当利用所述备选代码特征信息执行解析操作时,若检测到解析失败信息,则将所述异常代码特征信息中的方法替换为新的目标备选代码方法以便执行解析操作;记录方法替换次数;当所述方法替换次数大于预设值时,输出注释失败的提示信息。

[0060] 作为一种可行的实施方式,在输出包括所述注释文本的目标代码之后,还可以对通过方法替换得到的注释文本进行标记。由于通过方法替换得到的注释文本可能存在一定的偏差,因此可以在输出包括注释文本的目标代码时,对对通过方法替换得到的注释文本进行标记,以便进行人工核对,提高注释正确率。

[0061] 下面通过在实际应用中的实施例说明上述实施例描述的流程。请参见图2,图2为本申请实施例所提供的另一种程序代码注释的生成方法的流程图。本实施例面向Java语言来完成对代码的解析,进而自动生成注解。在识别代码块后,将其放入代码解析器中,然后提取其中的关键字段、动作以及参数,通过对它们的解析,用一些简短的语句将其描述,紧接着将描述语言输入文本编译器,最后生成注释。其中,解析器中包含各类识别代码的具体方法,最明显且有效的是调用递归遍历的方法来解析代码。

[0062] 首先,通过识别完代码块后将目标代码输入代码解析器解析,在此过程中提取目

标代码中的关键代码、方法、动作及其参数,然后通过具体解析生成简单的自然语言。然后,将代码解析器解析出来后生成的简单的自然语言全部传入到文本生成器中,将其进行汇总整合生成注释。在提取代码块关键字段过程中若存在未能解析的代码,则将未能解析的代码输入到解析池中进行方法重构最终得到注释文本。在遇到代码解析器无法解析的代码时,可以将无法解析的代码中的方法重构替换成相似的方法,若连续替换10词以上依然无法解析则可以输出解析失败的提示信息。

[0063] 请参见图3,图3为本申请实施例所提供的一种代码解析器的解析代码原理示意图,当向代码解析器输入目标代码后,代码解析器会选择关键词、方法、动作及参数。例如,解析器会提取第3行的interface Iinterface,通过解析可以判断出这是在定义一个接口;第6行代码中void method();声明接口的方法;第8行代码中的implements和Iinterface可以判断出这段代码是在实现接口;第10行的public void method(){}是在重写接口的方法。如上所述,此过程会提取并解析代码中的方法、动作等来完成对代码的解析,会结合被识别的整段代码进行解析。

[0064] 本实施例通过代码解析器提取关键代码的过程,其中主要是运用递归遍历来完成。在解析代码的过程中遇到无法解析的代码,应该运用方法重构,找到一个合适的方法来替换,循环解析,直至解析出最合适的注释。本实施例通过设计面向编程语言向自然语言转换的方法,来提高开发人员后续的开发进程,以及给后来的维护人员也带来了一定的方便,可以高效率的完成开发以及维护。本实施例通过代码解析器来做解析添加注释,相较于人工后续添加注释效率较高,节省了人力以及时间。

[0065] 请参见图4,图4为本申请实施例所提供的一种程序代码注释的生成系统的结构示意图;

[0066] 该系统可以包括:

[0067] 特征提取模块100,用于提取目标代码的代码特征信息;其中,所述代码特征信息包括代码关键词、代码方法、代码动作和动作参数;

[0068] 自然语言生成模块200,用于根据所述代码特征信息生成所述目标代码对应的自然语言信息;

[0069] 注释模块300,用于生成所述自然语言信息对应的注释文本,并输出包括所述注释文本的目标代码。

[0070] 本实施例向提取目标代码的代码关键词、代码方法、代码动作和动作参数等代码特征信息,由于代码特征信息描述了目标代码运行过程中的操作流程因此根据代码特征信息生成对应的自然语言信息。本实施例进一步生成资源语言信息对应的注释文本,进而得到了包括注释文本的目标代码。本实施例能够自动生成程序代码的注释,提高代码注释效率。

[0071] 进一步的,还包括:

[0072] 对应关系记录模块,用于在生成所述自然语言信息对应的注释文本之后,记录所述注释文本中每一子文本与所述目标代码行数的对应关系;

[0073] 相应的,注释模块300用于根据所述对应关系确定每一所述子文本对应的行数,根据所述子文本对应的行数输出包括所述注释文本的目标代码。

[0074] 进一步的,特征提取模块100用于通过对所述目标代码递归遍历提取所述目标代码的代码特征信息。

[0075] 进一步的,还包括:

[0076] 方法重构模块,用于当检测到无法生成自然语言信息的异常代码特征信息时,利用代码解析池中的备选代码方法执行方法重构操作以便生成所述自然语言信息。

[0077] 进一步的,方法重构模块包括:

[0078] 备选方法确定单元,用于确定所述异常代码特征信息中的方法的程序代码目的,将代码解析池中与所述程序代码目的对应的备选代码方法设置为目标备选代码方法;

[0079] 备选代码特征信息生成单元,用于将所述异常代码特征信息中的方法替换为所述目标备选代码方法得到备选代码特征信息;

[0080] 第一解析单元,用于利用所述备选代码特征信息执行解析操作,以便生成所述自然语言信息。

[0081] 进一步的,还包括:

[0082] 方法替换单元,用于当利用所述备选代码特征信息执行解析操作时,若检测到解析失败信息时,将所述异常代码特征信息中的方法替换为新的目标备选代码方法以便执行解析操作;

[0083] 计数单元,用于记录方法替换次数;

[0084] 失败提示单元,用于当所述方法替换次数大于预设值时,输出注释失败的提示信息。

[0085] 标记单元,用于在输出包括所述注释文本的目标代码之后,对通过方法替换得到的注释文本进行标记。

[0086] 由于系统部分的实施例与方法部分的实施例相互对应,因此系统部分的实施例请参见方法部分的实施例的描述,这里暂不赘述。

[0087] 本申请还提供了一种存储介质,其上存有计算机程序,该计算机程序被执行时可以实现上述实施例所提供的步骤。该存储介质可以包括:U盘、移动硬盘、只读存储器(Read-Only Memory,ROM)、随机存取存储器(Random Access Memory,RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0088] 本申请还提供了一种电子设备,可以包括存储器和处理器,所述存储器中存有计算机程序,所述处理器调用所述存储器中的计算机程序时,可以实现上述实施例所提供的步骤。当然所述电子设备还可以包括各种网络接口,电源等组件。

[0089] 说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。对于实施例公开的系统而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。应当指出,对于本技术领域的普通技术人员来说,在不脱离本申请原理的前提下,还可以对本申请进行若干改进和修饰,这些改进和修饰也落入本申请权利要求的保护范围内。

[0090] 还需要说明的是,在本说明书中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者

设备所固有的要素。在没有更多限制的状况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

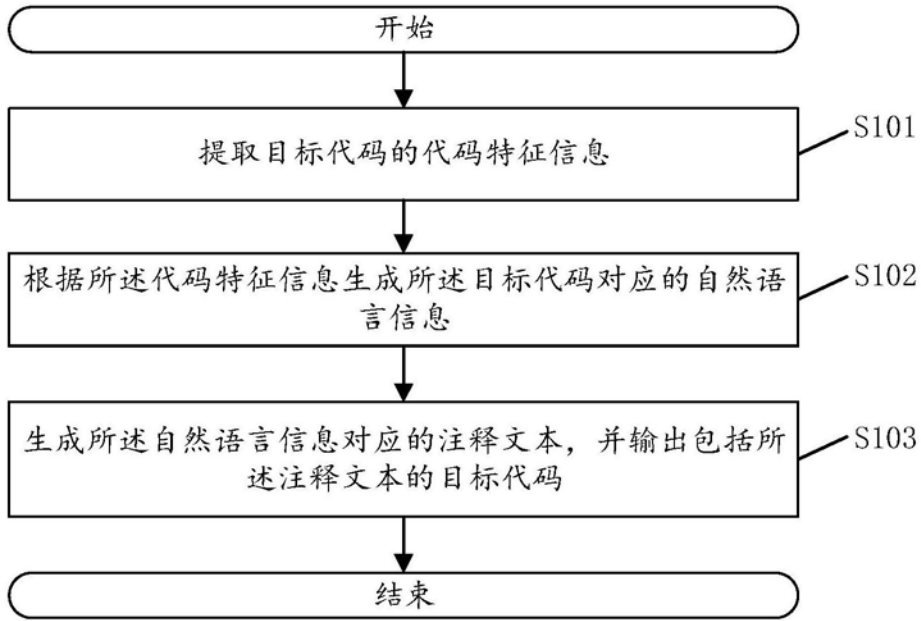


图1

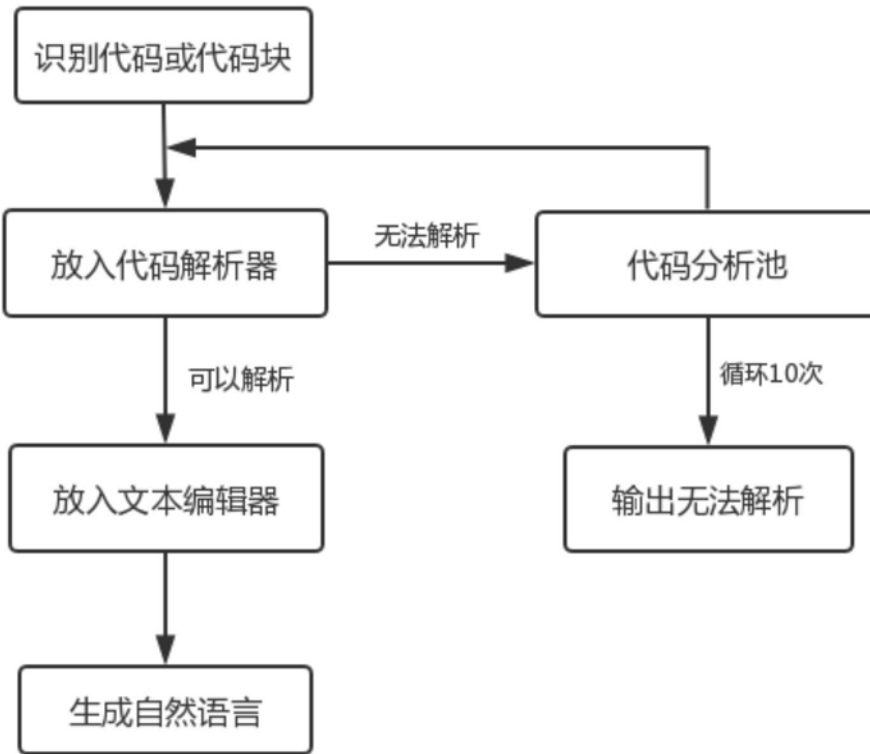


图2

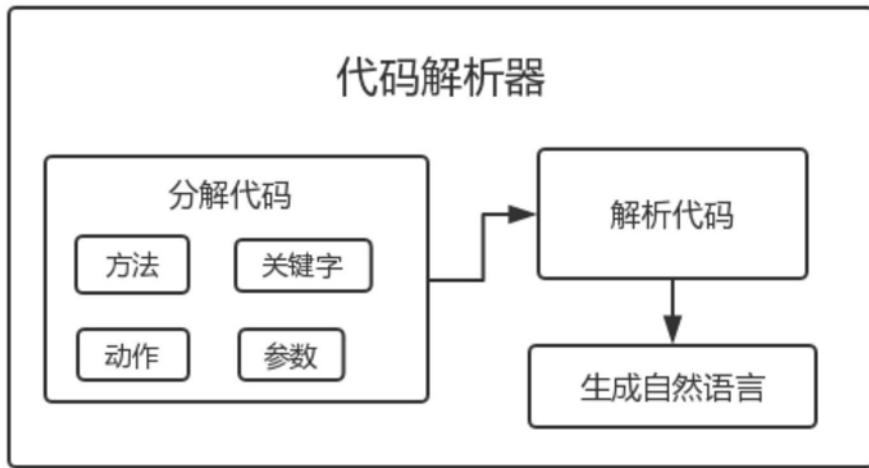


图3

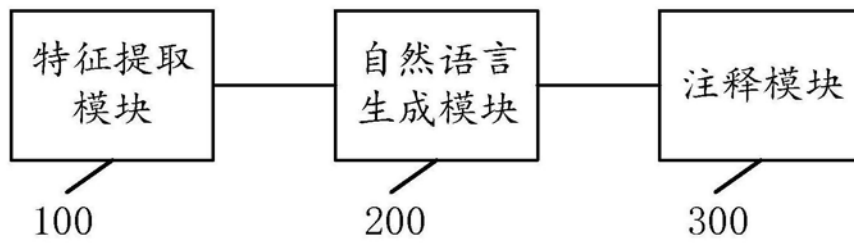


图4