南京邮电大学毕业设计(论文)开题报告

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 题　　目 | | Commit信息自动生成工具的研究与实现 | | | | |
| 学生姓名 | | 王达 | 班级学号 | B15041718 | 专业 | 软件工程（分段培养） |
| 提纲(开题报告2000字以上)：   1. 对指导教师下达的课题任务的学习与理解；   由于在日常的软件开发当中，经常会出现代码的更改，但是想要理解这些编译的基本原理，就需要对两次代码进行细致的比较。在软件开发过程中，软件工件的更改托管在控制版本系统（CVS）中，这些更改可以通过使用提交消息（例如，commit notes或commit comments）进行部分记录。 提交消息背后的预期目的是描述更改并帮助这些更改背后的编码基本原理。 开发人员稍后可以使用这些描述来理解和验证更改，定位和重新（分配）错误报告，以及跟踪对其他工件的更改。 通常，commit消息是开发人员在处理软件维护任务时所依赖的信息，知识和文档的重要来源。但是，主要是因为软件开发人员日常活动的数量和性质，提交消息可以是非信息性的（例如，“initial commit”，“last commit before lunch”）或几乎是空的。缺乏描述性/有用的提交消息的另一个可能的解释是考虑到使用基于行的差异工具生成的更改和更改的代码单元的详细信息足以理解更改。根据Buse和Weimer的观点，原始差异并不总是足以作为关于变化的一些问题的总结，因为原始差异仅是报告文件的两个版本之间的文本差异，这通常很长且令人困惑，并且 没有为开发人员提供许多高级问题的答案。 因此，基于行的差异不能提供足够的上下文来理解更改背后的原因。  因此，一个可以通过自动生成提交消息来协助开发人员提交更改的工具显得尤为重要。   1. 阅读文献资料进行调研的综述（10篇左右）； 2. 自动生成提交消息的工具   在软件维护任务期间，提交消息是开发人员依赖的重要信息，知识和文档来源。 但是，日常活动和中断的数量和性质可能会影响生成的提交消息的质量。 这份正式的演示文稿介绍了ChangeScribe，它是一种自动生成提交消息的工具。 ChangeScribe可从http://www.cs.wm.edu/semeru/changescribe（Eclipse插件，说明，演示和源代码）获得。   1. 通过对源代码更改的总结自动生成提交消息   尽管版本控制系统允许开发人员描述和解释提交消息中代码更改背后的基本原理，但实践状态表明大多数情况下此类提交消息非常短或甚至是空的。实际上，在最近对23K + Java项目的研究中，已经发现只有10％的消息是描述性的，并且与典型的英语句子（即15-20个单词）相比，超过66％的消息包含更少的单词。但是，总结软件更改的准确完整的提交消息对于支持许多开发和维护任务非常重要。提出了一种方法，创建为Change Scribe，旨在从变更集中自动生成提交消息。更改Scribe通过考虑提交构造型，更改类型（例如，文件重命名，仅对属性文件进行的更改）以及基础更改的影响集来生成自然语言提交消息。   1. 基于交互的程序差异工具   差分工具与一系列软件工程任务高度相关，包括分析开发人员   1. 发行说明的自动生成   本文介绍了ARENA（Automatic RElease Notes generAtor），一种自动生成发行说明的方法。 ARENA从源代码中提取更改，对其进行汇总，并将它们与版本控制系统和问题跟踪器中的信息集成在一起。 它是根据1,000份现有发行说明的手动分析而设计的。 为了评估ARENA发行说明的质量，我们进行了三项实证研究，共有53名参与者（45名专业开发人员和8名学生）。 结果表明，ARENA发行说明是开发人员生成的非常好的近似值，通常包括手动生成的发行说明中缺少的重要信息。   1. 自动捕获nl查询的源代码上下文，用于软件维护和重用   随着软件系统的不断发展和发展，维护和重用任务的代码变得越来越困难。 使用自然语言查询的现有静态代码搜索技术几乎不提供帮助，以帮助开发人员确定搜索结果是否相关，并且很少有人建议使用替代词来帮助开发人员重新构建不良查询。 在本文中，我们提出了一种新方法，可以自动从源代码标识符中提取自然语言短语，并将短语和搜索结果按层次结构进行分类。 我们的上下文搜索方法允许开发人员在一个软件中探索单词的使用，帮助他们快速识别相关的程序元素以便进行调查，或者快速识别用于查询重构的替代单词。 对22位开发人员进行的实证评估表明，我们的语境搜索方法在努力和效果方面明显优于最密切相关的技术。   1. java类的自然语言概要的自动生成器   JSummarizer是一个Eclipse插件，用于自动生成Java类的自然语言摘要。 摘要基于类的构造型，它隐式编码类的设计意图，并由JSummarizer自动推断。 该工具使用一组预定义的启发式方法来确定将在摘要中反映哪些信息，并使用自然语言处理和生成技术来形成摘要。 生成的摘要可用于重新编写代码，并帮助开发人员更轻松地理解大型和复杂的类。   1. 总结修改效果的工具   描述一个工具，它采用两个版本的过程并生成一个报告，总结它们之间的语义差异。 与基于程序依赖图的比较的现有工具不同，我们的工具根据过程的可观察输入 - 输出行为而不是其句法结构来表达其结果。 并且因为分析是真正的语义，它不需要语法组件的先前匹配，并且产生更少的虚假差异，因此正确地确定保留意义的变换（例如重命名局部变量）以使其没有可见效果。 应用于大型实时系统代码的修改的初步实验表明该方法是实用的  3. 根据任务书的任务及文献调研结果，初步拟定的执行（实施）方案（含具体进度计划）。  1.学习并了解控制版本系统（CVS）的基本使用方法。 2 周  2.学习并了解Git以及基于Git的代码托管平台，重点了解github的使用：  https://github.com/ 2 周  3.学习并了解Commit信息自动生成相关工具（参考资料）： 3 周  4.复现已有的Commit信息自动生成工具，对其生成的commit信息做出评估。 2 周  5.对已复现的Commit信息自动生成工具做出评估，分析其各自利弊。 2 周  6．整理资料，论文写作，准备答辩 2 周 | | | | | | |
| 指导教师批阅意见 | 指导教师(签名)：  年  月  日 | | | | | |