苏州大学实验报告

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 院系 | 计算机学院 | | 年级专业 | | 21计科 | | 姓名 | 方浩楠 | 学号 | 2127405048 |
| 课程名称 | | 编译原理课程实践 | | | | | | | 成绩 |  |
| 指导教师 | | 王中卿 | | 同组实验者 | | 无 | | 实验日期 | 2023.11.7 | |

|  |  |
| --- | --- |
| 实 验 名 称 | 基于PLY的LaTex解析 |

1. 实验目的

理解并实践将LaTeX文档转换为PDF的过程。

掌握解析LaTeX文件并生成抽象语法树（AST）的技术。

学习如何将AST进一步转换为HTML和PDF格式。

提升Python编程和处理文本文件的能力。

1. 实验内容

使用ply库来解析LaTeX文件并创建AST。

通过自定义的Node.py模块建立和操作AST。

使用AST2HTML.py将AST转换为HTML格式。

使用HTML2PDF.py将HTML格式转换为PDF文件。

分析并调试转换过程中的潜在问题。

1. 实验步骤和结果

项目结构如下:

表格

描述已自动生成

项目运行方式:

要运行此项目，需要安装以下依赖项：

**ply~=3.11**

**pytest~=7.4.3**

**fpdf2~=2.7.6**

您可以通过运行以下命令来安装这些依赖项：

**pip install -r requirements.txt**

使用方法

方式1

使用命令行参数指定要转换的文件路径，例如：

**python run.py data/example1.tex output/example1.pdf output/example1.html**

这条指令会读取`data/example1.tex`文件，并将其转换为`output/example1.pdf`和`output/example1.html`文件。

只要是形如

**python run.py {读取的latex文件位置} {输出的pdf文件位置} {输出的html文件位置}**

的指令均可以运行

其中html文件作为中间文件，可以不指定路径.这样就不会生成html文件,而是直接将latex文件转换为pdf文件

方式2

也可以直接打开run.py,然后修改run.py的这一部分:

**if \_\_name\_\_ == "\_\_main\_\_":**

**tex\_filename = ""**

**pdf\_output\_filename = ""**

**html\_output\_filename = ""**

修改这一部分也可以实现转换功能

标签实现情况如下:

|  |  |
| --- | --- |
| \begin{document}...\end{document} | 实现 |
| \title | 实现 |
| \author | 实现 |
| abstract | 实现 |
| \section | 实现 |
| \subsection | 实现 |
| itemize | 实现 |
| item | 实现 |

1. 实验总结

通过本次实验，我加深了对LaTeX文档结构的理解，学会了如何使用Python来解析和转换文件格式。遇到的挑战包括解决依赖项冲突、调试AST转换中的错误，以及优化HTML到PDF的布局转换。实验过程中，我学会了更高效地调试代码，同时也认识到了代码的模块化和文档编写的重要性。对于未来的工作，我计划改进错误处理机制，并提高转换工具的用户友好性和健壮性。