

电子科技大学

UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

博士学位论文

DOCTORAL DISSERTATION



论文题目 基于东拼西凑的电子科技大学毕业论文 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 模板
生成样例

学科专业 软件工程

学 号 2923000000

作者姓名 时富军

指导教师 某老师A

分类号 K825.1-64

密级 公开

UDC^{注1} 676.874

学 位 论 文

基于东拼西凑的电子科技大学毕业论

文L^AT_EX模板生成样例

(题名和副题名)

时富军

(作者姓名)

指导教师 某老师A、教授A、电子科技大学A

某老师B、教授B、电子科技大学B

某老师C、教授C、电子科技大学C

某老师D、教授D、电子科技大学D

(姓名、职称、单位名称)

申请学位级别 博士 学科专业 软件工程

提交论文日期 2013年1月13日 论文答辩日期 2013年1月13日

学位授予单位和日期 电子科技大学 2013年1月13日

答辩委员会主席 某老师

评阅人 某老师、某老师、某老师

注1：注明《国际十进分类法UDC》的类号。

ENGLISH TITLE OF MASTER THESIS

A Doctor Dissertation Submitted to
University of Electronic Science and Technology of China

Major: Software Engineering

Author: Zhang Xiaojun

Advisor: Mou Laoshi

School: School of Communication and Information

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

作者签名：_____

日期： 年 月 日

论文使用授权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后应遵守此规定）

作者签名：_____

导师签名：_____

日期： 年 月 日

摘 要

J. H. Wilkinson^[1]建立了非奇异矩阵的逆是矩阵元素的连续函数的理论。G. W. Stewart^[2]推出了矩阵的广义逆的连续性。为了得到Drazin逆的连续性,本文先给出了 M -矩阵、 H -矩阵类的逆的连续性。Campbell和Meyer^[3]也给出了Drazin逆的连续性性质,但没有给出明显的边界。

Drazin逆对扰动是很不稳定的。然而,在某种特定的扰动条件下,矩阵 $(A + E)^D$ 与 A^D 的接近程度能够得到量化且也能得到明显的相对误差边界。基于Drazin逆的不同形式,很多科学家和学者从事这一方面的研究。U. G. Rothblum给出的Drazin逆的以下的表达式:

$$A^D = (A - H)^{-1}(I - H) = (I - H)(A - H)^{-1}$$

其中 $H = I - AA^D = I - A^DA$ 。基于这个表达式,我们在本文中也给出了 $\|(A + E)^D - A^D\|_2 / \|A^D\|$ 和 $\|(A + E)^\sharp - A^D\|_2 / \|A^D\|_2$ 的范数估计,并与前人的成果进行了比较。

关键词: M -矩阵, H -矩阵, Drazin逆, Pseudo-Drazin逆, 条件数

ABSTRACT

The theory that the inverse of a nonsingular matrix is continuous function of the elements of the matrix was established by J. H. Wilkinson^[1]. The continuity of the generalized inverse A^+ of a matrix A was introduced by G. W. Stewart^[2]. In this paper, at first, the continuity of the special matrices inverse, such that M –matrices and H –matrices, respectively, are provided. Campbell and Meyer^[3] also established the continuity properties of Drazin inverse, but the explicit bound was not given.

The Drazin inverse is unstable with respect to perturbation. However, under some specific perturbation, the closeness of the matrices $(A + E)^D$ and A^D can be proved and the explicit bound the relation error can also be obtained. Based on the different representations of Drazin inverse, many scientists and scholars have worked it research. U. G. Rothblum gave the following representation of Drazin inverse:

$$A^D = (A - H)^{-1}(I - H) = (I - H)(A - H)^{-1}$$

where $H = I - AA^D = I - A^DA$. Based on the representation, we also obtain the norm estimate of $\|(A + E)^D - A^D\|_2 / \|A^D\|$ and $\|(A + E)^\# - A^D\|_2 / \|A^D\|_2$ and compare with the precedent results.

Keywords: M –matrices, H –matrices, Drazin inverse, Pseudo-Drazin inverse, Condition number

目 录

第一章 更新日志	1
第二章 模板使用说明	6
2.1 简介	6
2.1.1 面向没用用过 \LaTeX 的用户介绍	6
2.1.2 面向 \LaTeX 使用者的介绍	7
2.1.3 面向 \LaTeX 程序员的介绍	8
2.2 安装 \CTeX 套装	8
2.2.1 安装模板	8
2.2.1.1 从GitHub上下载模板	8
2.2.1.2 从 \CTeX (MiKTeX) 中安装模板 (不推荐)	9
2.2.2 更新模板	9
2.2.3 测试编译	9
2.3 模板结构	10
2.4 开始编写论文	10
2.4.1 确定学位信息	10
2.4.2 模板的其他选项	11
2.4.3 编写正文	12
2.4.3.1 新建一章	12
2.4.3.2 章节命令	12
2.4.3.3 修改论文	13
2.4.3.4 插入图片	13
2.4.3.5 插入表格	14
2.4.3.6 图表设计技巧	19
2.4.3.7 插入脚注	19
2.4.4 插入数学公式	20
2.4.5 定理环境等	21
2.4.6 交叉引用	21
2.4.7 枚举环境	21
2.5 其他内容	22
2.5.1 封面	22

2.5.2 摘要	23
2.5.3 参考文献	23
2.5.3.1 测试文献	24
2.5.4 攻读期间的研究成果	25
2.5.5 致谢	25
2.5.6 附录	25
2.5.7 外文资料	25
2.5.8 主要符号表和缩略词表	26
2.6 排版源代码	26
2.7 Linux中编译	32
2.8 使用 \LaTeX 时容易犯的错误	32
2.9 进一步了解 \LaTeX	33
2.10 修改本模板	33
第三章 示例：数学论文	34
3.1 Drazin逆计算方法的简介	34
第四章 示例：实现方案	36
4.1 设计目标与原则	36
4.1.1 设计目标	36
4.1.2 设计原则	36
4.2 系统方案	37
4.3 系统模块的实现	38
4.3.1 程序分解模块	38
4.3.2 构建特征向量模块	39
4.3.3 特征库构建模块	41
4.4 本章小结	45
致 谢	46
参考文献	47
附录 A 附录章	50
A.1 附录节	50
附录 B 附录另一章	51
B.1 附录另一章的一节	51
攻博期间取得的科研成果	52

图目录

图 2-1 WinEdt界面	10
图 2-2 菊花	13
图 2-3 子图效果	14
图 2-4 通用选项卡设置	27
图 2-5 格式化选项卡设置	27
图 4-1 系统结构	37
图 4-2 helloworld.apk的结构	38
图 4-3 程序分解流程	39
图 4-4 构建特征向量流程	40
图 4-5 特征库构建流程	41
图 4-6 Fisher 线性判别基本原理	43

表目录

表 2-1 模板结构	11
表 2-2 目前学术界广泛推荐使用的三线表	15
表 2-3 带辅助线的三线表	15
表 2-4 可自动换页的超长三线表	16
表 2-5 用Excel制作的表格	19
表 2-6 contents目录中文件用途	22
表 2-7 源代码中等宽符号替换	28
表 4-1 Ω 数据表定义	40
表 4-2 LDA算法变量定义	43

缩略词表

缩略词	英文全称	中文全称
SVM	Support Vector Machine	支持向量机

主要符号表

符号	说明	页码
Π	数学及物理学领域普遍存在的数学常数。	26

第一章 更新日志

2014年3月6日：v0.6.8

1. 在v0.6.1中将所有正文的英文字体改为Times字体，却忽略了公式中的字体也被修改为Times字体。现改回v0.6.1之前版本中公式的字体（CM）。

2014年3月4日：v0.6.7

1. 修复BUG：缩略词表和主要符号表各自的第二页页眉不正确的问题。
2. 添加bm宏包，以支持希腊字母在数学模式中的粗体。例如：

\mathbf{X} $\mathbf{\Lambda}$

2014年3月3日：v0.6.6

1. 修复BUG：按照规范，脚注应该每页重新编号。
2. 修复BUG：参照研究生院范例修正注脚线和内容的距离。
3. 针对脚注跨页的情况，在说明文档中加入第 2.4.3.7 节，说明解决方案。
4. 针对源代码排版段落在换段后会多出一个空行的问题，在说明文档的第 28 页加上一段话。提示大家在用Highlight转换出的源代码排版文件中，需要手动删除一个多余的换行命令。

2014年3月1日：v0.6.5

1. 修复BUG：子图有不居中的情况。
2. 修复BUG：当子图环境整体标题和子图标题都超过一行时，子图标题没有悬挂缩进的问题。

2014年2月25日：v0.6.4

1. 添加了研究生院规范中附件1,2,3的所有封面、中文扉页、英文扉页。具体使用方式见第 2.4.1 节。
2. 添加了两个选项，分别用于使模板不生成摘要之前内容或正文之外的所有内容。详见第 2.4.2 节。

2014年2月24日：v0.6.3

1. 修订硕博封面第2页，使之和学校范例尽量一样。

2014年1月20日：v0.6.2

1. 新增一个用于检查模板所依赖的全部宏包版本的packagecheck.tex。编译它即可输出当前版本的模板所需的全部宏包的所需版本和当前已安装的版本，或者找不到需要的宏包。

2014年1月20日：v0.6.1

1. 论文中所有正文中的英文修正为Nimbus Roman No9 L字体(wiki链接), 这个字体和Times New Roman字体高度相似。
2. 中英文摘要的关键词内容改为悬挂缩进。
3. 将中英文摘要和目录加入到pdf书签中。
4. 修正附录中的图表和公式的标号形式。

2014年1月16日: v0.6

1. 添加选项listoffigures, 用以打印图目录。
2. 添加选项listoftables, 用以打印表目录。
3. 添加命令\englishshcoolextraline, 用以在titlepage.tex文件中。当用\englishshcool命令填写学院英文名时, 英文名过长一行显示不下的情况下(例如通信学院), 用\englishshcoolextraline 命令将写不下的词写在下一行。
4. 根据反馈, 去掉所有超级链接的有色外框。
5. 将枚举环境的序号缩进到和正文一致。
6. 添加选项cv。开启后会用contents目录下的cv.tex替换publications.bib生成的攻读硕博期间发表的论文。cv.tex是按照一位通信学院的同学提供的一份往年的硕士论文中的格式做的模板。这个“个人简历及攻读硕士学位期间取得的研究成果”格式, 以及“个人简历”都不是学校规范中要求的。所以如果不是自己学院要求、自己导师要求, 建议不要用这个选项。
7. 将原有的主要符号表的实现机制改为更新的宏包, 同时支持了缩略词表。原有的主要符号表使用方式略有更改, 详见第 26 页。
8. 修正可能在\verb命令中或排版源代码中可能使用仿宋的问题。已全改为宋体。
9. 本说明文档在第 2.4.3.6 页补充了一小节关于图表制作技巧的内容。

2013年12月25日: v0.5.6

1. 本次更新不影响排版效果。仅仅修改了封面上校徽的颜色, 使橙色和蓝色与学校识别视觉系统一致。

2013年11月26日: v0.5.5

1. 在v0.5.3版本修改中, 子图的标题过长, 导致引用总标签失败。现修复了这个问题。并在第 14 页进行了一小段测试。

2013年10月29日: v0.5.4

1. 修改封面到摘要之前的部分实际的页码, 避免pdf阅读器中页码冲突。
2. 修改中英文摘要标题和关键词的字体, 改为加粗形式。并按照范例, 在“摘要”二字中间加空格。

2013年8月10日：v0.5.3

1. 根据2013年7月18日研究生院网站发布的新的《学位论文撰写格式规范》
<http://gr.uestc.edu.cn/article.shtml?id=6643>，修改了模板。但由于本科教务处学位科尚未修改规范，所以本科选项的排版规范暂未做大变化。

2013年6月7日：v0.4.8

1. 修正目录的标题字体。之前用了2008年本科规范的宋体小二，是错误的。

2013年6月3日：v0.4.7

1. 根据一位同学的需求，将源代码的排版改为等宽字体。为了实现这个功能，将highlight.sty的内容更改，并内置到模板中了。所以之前使用了排版源代码并用input命令加载了highlight.sty的同学需要删除那个input语句，否则会报错说一些命令已经定义过了。
不过我并没有做到完全自动化的等宽效果。原因和解决办法见更新后的2.6节。

2013年6月2日：v0.4.6

1. 修正BUG：页脚字号改为小五。

2013年6月1日：v0.4.5

1. 根据一位同学的需求，将本科论文中需要的外文资料翻译部分的章节标题做了调整，英文部分题注改为英文。详见2.5.7节。
2. 上一版示例中将获奖的条目类型拼错了，所以显示结果不对。
3. 我自己临时打算继续用LaTeX的Beamer幻灯片文类做答辩幻灯片，于是新建了一个幻灯片主题。但是因为没设置完，也没有说明文档，所以还没法交给新手直接用。所以不建议动手能力弱的同学现在使用。

2013年5月25日：v0.4.4

1. 修正摘要关键词的逗号之前会有一个多余的空格，而英文摘要关键词的逗号后没有空格的问题。
2. 修正可能会出现超大段距的问题。并且重新设置了一个默认为0的弹性段距，最大为10磅，最小为-1磅。有弹性段距会让排版得到更美观的结果。而且-1到10磅相对于行距来说并不明显。
3. 增加了一个参考文献类型ONLYNOTE，可以将参考文献的note字段直接作为参考文献条目内容。用于录入不规范的参考文献条目，例如攻读期间取得的成果中可能会要录入获奖[7]。而获奖并不是标准的参考文献类型。
4. 本科论文目录中加入外文资料原文和译文。

5. 本科论文的章标题序号改为阿拉伯数字序号。
6. 修正表格字号，按照规范应为五号字（英文10.5磅）。
7. 发现问题：如果`\citeup`命令之前是标点符号，请暂时手动将前面的字符同`\citeup`命令都放入`\mbox{}`命令的括号中，例如：《新华字典`\mbox{\citeup{xxx}}`》。这样可以避免在这个标点符号之后直接换行，而将上标的`[]`置于下一行行首。

2013年5月20日：v0.4.3

1. 修复BUG：枚举环境之前与前后文有额外的间距，现已去掉。
2. 修复BUG：之前模板没有按照规范要求将摘要等前置内容章起右页，现已修正。至于中文摘要在第I页，而英文摘要在第III页的问题，我认为是正常的，页码不应该因为左页没内容就跳过去。奇偶页码与左右页的关系不应该改变。
3. 修复BUG：由于我的疏忽，本科论文规范修订时将目录中的缩进打错成半角，这和硕博规范不一致。现已将本科论文的目录缩进修正和今年的规范一致，硕博不变。
4. 修复BUG：之前模板的参考文献和致谢的顺序是按照硕博规范设置的，刚刚发现本科规范和硕博规范在这个顺序上是反的。现已将本科论文的致谢置于参考文献之后。
5. 发现可能编译出错的问题，新安装的CTeX套装在安装目录下`\CTeX\tex\latex\ctex`中有一个旧版本的`ctex`宏包，应将这个目录删除，`ctex`才会下载最新的`ctex`宏包。我已在模板中添加了需要的`ctexbook`版本。

2013年5月14日：v0.4.2

1. 修复BUG：此版本之前，当目录只有一页时目录的页眉显示成主要符号表。
2. 上一版本在使用WinEdt等自带的编译功能时不显示LOGO，但是会引起封面多出一页。此版本装饰了这个问题，在不使用正确编译的编译脚本编译时在LOGO处显示一段说明文字。

2013年5月4日：v0.4.1

1. 将封面所需的两个png文件去掉。采用`pstricks`宏包生成的矢量图，并嵌入到cls文类中。感谢周光耀同学帮忙！
生成所需的特殊编译过程已写入到批处理脚本中。这个过程是：
用`filecontents`环境将`pstricks`图片嵌入到cls文件中，令其生成`logo.tex`文件。
然后在文类中利用`write18`命令正确编译`logo.tex`，生成`logo.pdf`。这个`write18`

命令默认是不允许使用的，所以在WinEdt中执行编译是不会生成logo.pdf的，也就不会显示Logo。

2013年4月4日：v0.4

1. 添加english选项，使用后将所有预定名变为英文。供我校留学生使用。这个需求是研究生院老师提出来的，同时也有研究生同学向我询问可否支持的。使用方法是在master或doctor 选项后加上,english。

例如：`\documentclass[master,english]{uestcthesis}`

2. 添加定理环境的结束符语句。但是目前所有结束符还设置为空。也就是没有任何变化。因为学校的规范中没有定义各种环境的结束符应该是什么。这个需求是一位数学学院的同学提出来的。需要用的同学可以暂时自己改，例如，将uestcthesis.cls文件的661行改为`\theoremsymbol{【证毕】}`，即可在证明环境后自动加上这个标志。注意修改cls 文件请使用notepad++等支持无BOM 头的UTF-8编辑器。

2013年2月25日：v0.33

1. 设置定理环境的缩进。

2013年2月21日：v0.32

1. 将参考文献中英文条目的作者名改为：名在前缩写，姓在后全拼。

2013年2月21日：v0.31

1. 去除枚举环境的所有垂直附加距离。感谢不中同学反馈BUG。
2. 已确认还未解决的问题：参考文献中英文作者名的录入规则和学校要求不符。缺少获奖类型的参考文献录入。

2013年2月13日：v0.3

1. 在模板内部修复了参考文献的遗留问题，不再需要sed工具。
2. 模板已经提交到CTAN，并已被MiKTeX收录。所以本模板已经可以通过MiKTeX的宏包管理器安装，不再需要typesetting目录。更新说明中安装一节。
3. 在pdf元数据中填写了论文的信息，并加入了XMP高级元数据。设置了用Adobe Reader 打开时的默认阅读方式为书籍方式。
4. 加入了排版源代码的方案说明，见2.6节。

2013年1月14日：v0.2

1. 完成本硕博三种格式论文所需的全部结构。

第二章 模板使用说明

2.1 简介

2013年1月14日更新后的模板已经支持本硕博三种标准。并且本人已经和教务处实践科交流过，实践科对本模板是支持的。同时实践科也表示可以接收pdf格式的电子档。并且已经同意将本科毕业论文的排版规范改为和研究生标准统一化，所以也简化了本模板的开发工作。本次更新后已经针对本科和硕博的不同要求，可以分别编译出不同结构的论文。

本模板以《研究生学位论文（含研究报告）撰写规范（2012年修订版）》（以下称学校标准）为标准。对于其中未提及的规范，参照中国高等学校自然科学学报编排规范(修订版)》。参考文献格式符合学校标准的基础上，采用国标GB/T7714-2005。编译结果已经可以符合学校标准，如有不完全符合的地方，请大家分享反馈回来，我尽量修改。改不了的地方，请大家看看历年的硕博论文，本模板的排版质量已经非常非常高了。

2.1.1 面向没用用过 \LaTeX 的用户介绍

Q: 什么是 \LaTeX ?

A: \LaTeX 是 \TeX 的一种扩展。 \TeX 是科技论文排版的首选软件， \TeX 程序从1989年至今没有过任何修改，可见这个程序的可靠性之高。在用 \TeX 写文章时，需要边写内容边定义格式。而 \LaTeX 在 \TeX 基础上，将排版和写作分开，使得写作的人不用管格式的处理，从而专注于文章的内容!

Q: 我怎么没听说过这东西?

A: 确实 \LaTeX 在国内普及度不高，学校对本科生的论文很少要求到 \TeX 可以达到的高度。但是研究生和博士生论文一般都需要用 \LaTeX 排版。全世界著名大学都要求用 \LaTeX 排版毕业论文。国内大部分的重点大学也都有 \LaTeX 学位论文模板(点击查看列表)。比如清华大学的校内论坛还有专门的 \TeX 讨论分区。

Q: 使用本模板比使用Word好在哪?

A:

1. 不用设置任何字号、字体、行距、页眉、页脚、页边距、目录，使用者根本就不需要知道这些命令。

2. \TeX 诞生的最初目的之一就是排版数学公式， \TeX 排版出的数学公式比Word在细节上要好太多了。全世界所有的数学期刊包括中国的数学期刊，都（只）接受 \LaTeX 源文件。
3. 文章中的所有序号，第X章、第X节、图X-X、表X-X都是自动生成的，绝不会出错！如果要引用某个图，只需要给图起一个别名，引用这个别名即可， \TeX 会自动将它变成正确的序号。
4. 全自动的行距、字距微调，确保不该换行的地方不换行（比如参考文献的引用角标不会到下一行开头去），不该换页的地方不换页（比如图和图题之间），能放下的图片不会因为一点点距离不够就放到下一页。
5. 参考文献的格式全自动处理成模板设置好的格式，绝不会出错！参考文献的输入内容大多学术网站都查询后可导出，比如Google学术搜索。
6. 生成的PDF文件可全文查找，所有的引用都自动生成超链接。

Q: \LaTeX 有什么缺点？

A: 编写过程中不能像Word一样所见即所得， \LaTeX 是一种所思即所得的排版系统。 \LaTeX 需要在编译时才能确定文章中的引用序号和各种图片表格的位置，所以虽然不能边写边看结果，但是得到的结果质量比Word高。

Q: 多久能学会使用 \LaTeX ？

A: 使用本模板写作毕业论文，使用者是作者的角色，并不需要学习如何编写 \LaTeX 模板。作者只需要写文章内容，用几个简单的命令指明哪段文字是章标题，哪段文字是节标题。或者用参考文献引用的命令引用参考文献。以及插图命令、表格命令。文科论文大概只需要对照说明文档即可立刻投入写作。理工科论文由于要输入数学式等特殊内容，需要多学一些专门输入数学式的方法。各种复杂的数学符号有专门的网站可以手写识别出LaTeX中需要输入的命令。

2.1.2 面向 \LaTeX 使用者的介绍

本模板提供了插入单张图片、子图、三线表、自动换页三线表、截图表格，以满足完成论文的大部分的图片和表格插入需求。详见后文介绍。

编译时使用`pdflatex`编译，暂时没有设置`hyperref`对各种编译器自动兼容。`hyperref`的兼容性本来就非常糟糕，暂时不动它了。

正确编译的方式见模板根目录下的“`compile.bat`”文件内容。其中编译过程的原因是：由于硕博论文使用了`multibib`宏包以支持显示攻读期间的研究成果，所以 \CTEX 套装自带的WinEdt默认编译方式会少编译这一部分。

在撰写论文的过程中可以不用这个“`compile.bat`”，按照常规的编译方式只会缺少“攻博/硕期间发表的论文”。

其余内容见后文介绍。

2.1.3 面向 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 程序员的介绍

许多 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 用户并没有区分出 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 和 $\text{T}_{\text{E}}\text{X}$ 的重要区别： $\text{T}_{\text{E}}\text{X}$ 是内容和排版混在一起的，而 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 是将内容和排版分开的。如果没有注意到这个区别，使用 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 只是比 $\text{T}_{\text{E}}\text{X}$ 多了些高级命令而已。

本模板直接将所有排版内容都写在了文类中，并包装了一些作者命令。希望今后的维护者能够注意到这个特点，不要破坏它。国内许多其他高校的 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 模板大多没有采用文类，许多设置的方式也不是最好的。所以借鉴时需要谨慎。

目前本模板试图和标准的 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 命令完全兼容，插图和表格的包装命令中并没有前后间距的设置。所以如果用标准命令插入图片和表格，也不影响符合学校标准。但是注意，非常不推荐模板使用者这样做。

2.2 安装 $\text{C}^{\text{T}}\text{E}^{\text{X}}$ 套装

$\text{C}^{\text{T}}\text{E}^{\text{X}}$ 是Windows系统下面向中文优化的一套 $\text{T}_{\text{E}}\text{X}$ 工具的集合，包括编译器、编辑器、阅读器等一系列工具。安装后即可立刻使用。

从网页<http://www.ctex.org/CTeXDownload>下载 $\text{C}^{\text{T}}\text{E}^{\text{X}}$ (当前版本2.9.2.164，大小203M)。可以在网页上看到这个203M的版本不包含完整的宏包（插件），所以安装这个版本在第一次使用时需要联网才能正常使用。也可以直接下载1.31G的完整版（不推荐）。安装完成后，前者比后者会节约大概2GB的磁盘空间。

安装时建议不要改变除安装路径以外的其他选项，即完整安装。由于安装包中包含非常多文件，所以某些笔记本硬盘可能会安装的比较慢。这是正常现象。安装完成后扩展名为`tex`的文件应该会被自动关联。双击`thesis.tex`应该会自动用WinEdt打开。

2.2.1 安装模板

2.2.1.1 从GitHub上下载模板

推荐从GitHub上的本项目主页（<https://github.com/shifujun/UESTCthesis>）中下载本模板。页面上有一个zip按钮，可以无需安装git程序下载最新的模板。解压开这个zip，可以得到本说明文档和一个example目录。example目录既是本说明文档的源文件也是模板。example相当于一个人的论文撰写工作目录。关于它的结构，见下文的2.3节。

2.2.1.2 从CTeX (MiKTeX) 中安装模板 (不推荐)

此安装模板的方法可以将模板安装在CTeX宏包库中, 这样在任何位置都可以使用模板。不过我们撰写论文一般只撰写一份, 所以这样做好处不是很大。而且这样的安装方式并不总能获得最新的模板。

运行开始菜单→ CTeX → MiKTeX → Maintenance (Admin) → Package Manager (Admin)。这个程序是管理CTeX 系统中的宏包用的。点击菜单栏中的Repository → Change Package Repository。选择第一项 “Package shall be installed from the Internet”, 点击下一步, 选择服务器。在学校使用可以选择China的mirrors.ustc.edu.cn。这是中科大的源, 同在教育网中, 速度会比较快。然后点完成。之后程序应该会立刻开始同步数据。

同步完成后, 在工具栏右侧的Name:处填写uestcthesis, 然后回车。搜索出来的就是本模板。右键点击本模板的条目, 选择Install。完成后点Close 关闭。

安装完成后运行开始菜单→ CTeX → MiKTeX → Maintenance (Admin) → Settings (Admin)。依次点击按钮 “Refresh FNDB” 和 “Update Formats”, 然后确定。

2.2.2 更新模板

在正式打印论文之前一定要访问项目主页检查模板是否有更新。一般来说每次更新都会修正bug, 所以推荐尽早更新模板。更新模板时无需改动已撰写的论文内容。只需替换模板的两个排版设置文件uestcthesis.cls和uestcthesis.bst。这两个文件一般在工作目录中, 也可能由CTeX安装在系统中。如果从GitHub上下载模板, 指的就是example目录中的这两个文件。从GitHub上下载新的模板, 用其中的cls和bst文件替换现有的文件即可。

如果是用CTeX或MiKTeX安装的模板, 请用Update程序更新模板。但不一定能获得最新的模板, 因为模板首先在GitHub上更新。

2.2.3 测试编译

在v0.6.2版本中新增加了一个packagecheck.tex文件, 编译它即可输出当前版本的模板所需的全部宏包的所需版本和当前已安装的版本, 或者找不到需要的宏包。缺失的宏包可以通过用管理员权限编译该文件, 使MiKTeX自动安装缺失的宏包。版本过旧的宏包需要通过MiKTeX的更新程序更新宏包。

用WinEdt打开thesis.tex如图 2-1 (双击即可)。点击图中红色箭头指向的黑色实心小三角, 在下拉菜单中选择PDFLaTeX。然后点击绿色箭头指向的按钮将此

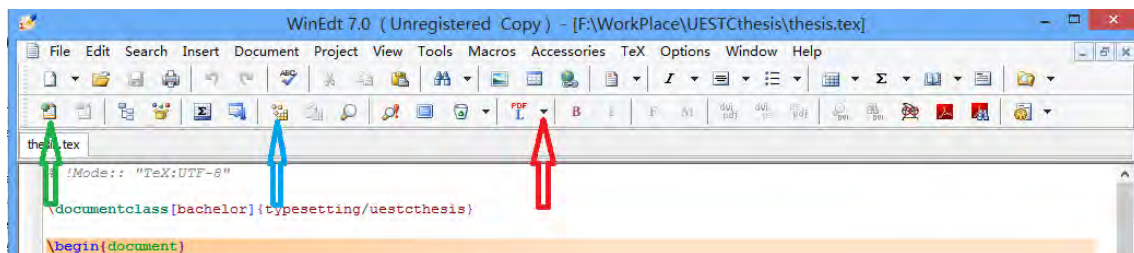


图 2-1 WinEdt界面

文件设置为主文件。最后点击蓝色箭头指向的按钮，或按键盘上的F9按键，编译论文。用这种方式编译论文比较方便，但会缺少研究生论文所需的“攻读期间发表的成果”一章。所以正式编译时必须用compile.bat 编译论文才能得到正确的结果。

通过将此文件设置为主文件之后，以后只要thesis.tex在WinEdt中打开了，即使当前编辑的文件不是thesis.tex，编译时也会编译thesis.tex文件。这样可以避免每次都切换回thesis.tex 编辑窗口。

如果安装的是Basic版的CT_EX套装，则第一次编译时会自动联网下载缺少的宏包，所以会比较慢。编译完之后会自动弹出编译好的pdf 论文。

如果安装的是Full版的CT_EX套装，当前版本应该没有问题。但也应该联网更新宏包。选择菜单TeX, MikTeX, MikTeX Update Wizard。然后选择“I want to get updated packages from a remote package repository”, “Let me choose a remote package repository”。同样可以选择中科大的源。

2.3 模板结构

从GitHub上下载的zip解压后，只有example目录是一份模板。source目录是模板的源代码。example就相当于一份论文，也叫工作目录。你可以重命名这个目录。

工作目录中应包含3个目录和3个文件，见表 2-1。表中说明为Windows环境下的情况。Linux和Mac 请将bat 文件换为sh文件即可。example 目录中还有两个zip文件：windows下编译脚本.zip和linux,mac下编译脚本.zip。选择其中一个解压到工作目录中。

2.4 开始编写论文

2.4.1 确定学位信息

打开thesis.tex文件，第三行如下：

```
\documentclass[doctor]{typesetting/uestctthesis}
```


表 2-1 模板结构

名称	目录/文件	说明
chapters	正文目录	使用者可复制其中的DuplicateMe.tex，创建论文的各章子文件。
contents	其他内容目录	使用者可以通过编辑其中的文件填写论文结构的非正文部分。
pics	插图目录	使用者在编写论文时要插入的图片放在此目录中。
thesis.tex	主文件	最终得到论文需要编译这个文件，这个文件的内容应该只包含章的载入顺序。使用者可以重命名这个文件，但须将compile.bat中第一行的thesis改为相应名字。
compile.bat	编译批处理	双击编译论文，得到正确的thesis.pdf文件。
clean.bat	清理批处理	双击可删除目录中全部编译过程生成的中间文件。

修改第三行中间的中括号中的doctor为其他的选项。bachelor代表学士学位；master代表学术型硕士学位；doctor代表博士学位；engineeringmaster代表工程硕士学位；professionalmaster代表专业硕士学位；researchreport代表留学生研究报告；engineeringdoctor代表工程博士学位；

这些选项选择好之后，模板会自动添加相应的封面、扉页等前置必备页面。

然后打开contents目录中的titlepage.tex。填写其中的信息。本科生注意要删除其中后一部分的信息，否则编译会出错。

2.4.2 模板的其他选项

本模板还有如下可选选项：

选项listoffigures，用以生成图目录。

选项listoftables，用以生成表目录。

选项cv。开启后会用contents目录下的cv.tex替换publications.bib生成的攻读硕博期间发表的论文。cv.tex是按照一位通信学院的同学提供的一份往年的硕士论文中的格式做的模板。这个“个人简历及攻读硕士学位期间取得的研究成果”格式，以及“个人简历”都不是学校规范中要求的。所以如果不是自己学院要求、自己导师要求，建议不要用这个选项。

选项withoutforepages。启用后不编译摘要之前的部分（封面、中文扉页、英文扉页、独创性声明和论文使用授权）。

选项onlychapters。启用后只编译正文部分。

最后两个选项都是方便撰写论文时加快编译用的。为了这个目的，还可以使用draft选项。这个选项会不真正插入图片。如果在撰写论文正文中经常需要

编译查看改动结果，在不涉及图片、参考文献等附加内容的情况下，建议开启`onlychapters`和`draft`选项。同时只需要运行一次`pdflatex`命令编译即可。编译脚本执行的正确编译会连续编译3次，所以会比较慢。

以上这些选项的启动方式形如：

```
\documentclass[doctor,onlychapters,draft,listoftables]{typesetting/uestcthesis}
```

2.4.3 编写正文

2.4.3.1 新建一章

论文正文的每一章都是一个单独的`tex`文件，存放于`chapters`目录中。由于本模板采用UTF-8 编码，所以不能用Windows 记事本新建的UTF-8文本文件创建`tex`文件。因为Windows的UTF-8 格式文件头和我们用的不一样。所以为了方便起见，`chapters` 目录下有一个`DuplicateMe.tex`文件，它是一个写好文件头的空文件。新建一章时只需要将它复制一份，改一个新的文件名即可。文件名最好可以代表这一章的内容。

`chapters`目录中的文件并不会都出现在最终的论文中，比如`DuplicateMe.tex`放在`chapters` 中就不会有影响。将一章加入到论文中，需要在`thesis.tex`中加入`\include{chapters/DuplicateMe}`这样的语句。`\include{}`的顺序决定章的顺序。在写作中，可以将前面已经写好并检查排版没问题的章暂时从`thesis.tex`去掉，比如在`\include{chapters/DuplicateMe}` 前面加一个`%`，把这行注释掉。

为了跨平台着想，本模板直接采用了UTF-8编码。每个文件的第一句话都是`% !Mode:: "TeX:UTF-8"`，`%`后面都是注释，不会出现在文章内容中。这句话是为了让一些文本编辑器自动选择UTF-8 编码打开文件。在Windows 中，文本文件的默认编码是ANSI，虽然可以另存为UTF-8，但是依然有区别。Windows记事本的UTF-8 格式多了一个BOM文件头。

2.4.3.2 章节命令

本模板使用的全是标准的章节命令。有：

```
\chapter{}\section{}\subsection{}\subsubsection{}
```

这四个级别的章节命令。如果不懂怎么用，可以打开`tuition.tex`看看。

2.4.3.3 修改论文

CT_EX套装自带的PDF阅读器SumatraPDF，也就是按F9编译后自动弹出的PDF阅读器，它可以自动在tex源文件中找到对应的段落。在SumatraPDF中双击指定段落，即可自动转到tex源文件中对应的位置。另外，可以用这个方法查看范例论文中的图片、表格、公式等源文件编写方法。

2.4.3.4 插入图片

为了控制全文的图片插入形式一致，包括居中、前后间距、题注字体等，本模板定义了几个命令来插入图片：

`\pic[htbp]{图标题}{width=0.6\textwidth}{文件名}`

这条命令用于插入单个图片，第一个中括号是可选参数，用于向系统建议图片

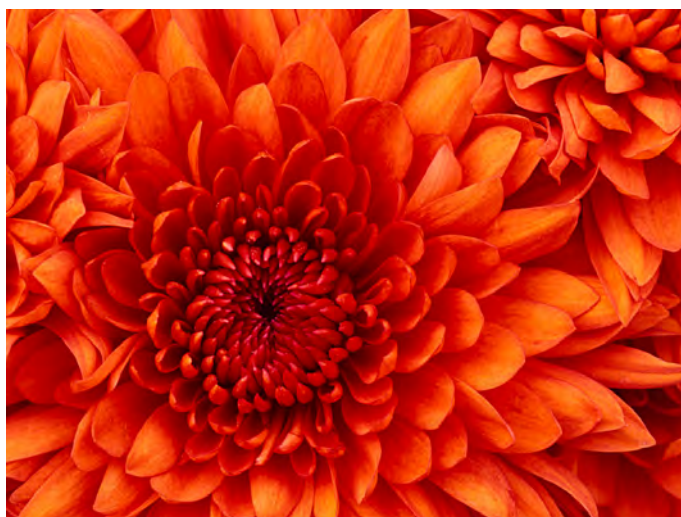


图 2-2 菊花

的位置。默认参数htbp就是让系统考虑放在当前位置（h）、本页底部（b）、次页顶部（t）、单独一页（p）。图片不会浮动到引用位置之前，只会向后浮动。另外可以设置[H]，命令系统不考虑版面美观，就放在当前位置。也可以向[]中加入“!”，告诉系统不用考虑版面限制，但此时还只是一种建议。

第三个参数中可以设置图片的大小，最大不能超过width=\textwidth。因为这是设置图片按宽高比不变放大或缩小到宽度等于一行文字的宽度。如果超过这个大小，图片将超出版心。例子中的0.6是乘以0.6倍的意思。也可以设置height=5cm等参数。最后一个参数既是文件名也是标签名（标签名用于引用图片，如图X-X这种），这个文件名不需要扩展名，有也没关系。图片必须放在pics目录下。同时建议不同扩展名的图片，也不要起一样的名字。图片格式支持png,pdf,jpg,jpeg。

还可以插入多幅图片组成的子图组，这些图共用一个图编号，各自又有自己的子编号。

```
\begin{pics}[htbp]{整体标题}{整体标签（英文数字组成）}  
\addsubpic{子图标题}{width=0.4\textwidth}{文件名}  
\addsubpic{企鹅}{width=0.4\textwidth}{Penguins}  
\end{pics}
```

这四行命令中，第一行和最后一行定义了一个子图环境，在其中可以插入若干

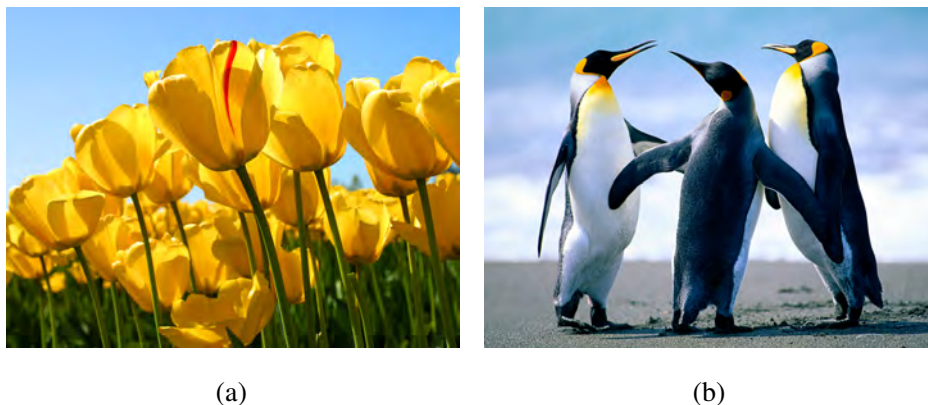


图 2-3 子图效果

(a)郁金香；(b)企鹅

条`addsubpic`命令。当一行图片排满后可以换行到下一行。注意图片间有间距，所以两张`0.5\textwidth`宽的图片不能排在一行。

测试引用图片标签。单张图片图2-2是一副菊花。子图效果如图2-3所示，其中有两幅子图。子图2-3(a)是郁金香，子图2-3(b)是企鹅。

2.4.3.5 插入表格

\TeX 系统对表格的支持可能不够出色。这种说法主要是因为微软的Excel实在是太出色了，以至于人们经常用Excel制作结构非常复杂看起来很好看的表格。其实大多数表格都可以简化。《中国高等学校自然科学学报编排规范(修订版)》中讲到：表应精心设计，为使表的结构简洁，建议采用三线表，必要时可加辅助线，如表 2-3。

插入普通三线表可以用我定义的命令：

```
\threelinetable[htbp]{标签}{0.4\textwidth}{lcr}{表标题}  
{项目&频率&技术\\  
}{
```

```

802.11b\tnote{a}&2.4GHz&DSSS\\
蓝牙&2.4GHz&FHSS\\
HomeRF&2.4GHz&FHSS\\
}{
\item[a] 这里还可以添加脚注!
}

```

表 2-2 目前学术界广泛推荐使用的三线表

项目	频率	技术
802.11b ^a	2.4GHz	DSSS
蓝牙	2.4GHz	FHSS
HomeRF	2.4GHz	FHSS

^a 这里还可以添加脚注!

插入带辅助线的三线表可以将头部改为:

```

{&\multicolumn{总列数减1}{c}{内容}\\ \cmidrule[0.05em]{辅助线开始列号-辅助线终止列号}
项目&频率&技术\\
}

```

具体请查看tuition.tex中的这部分源代码。

表 2-3 带辅助线的三线表

项目	参数	
	频率	技术
802.11b ^a	2.4GHz	DSSS
蓝牙	2.4GHz	FHSS
HomeRF	2.4GHz	FHSS

^a 这里还可以添加脚注!

仔细观察可以发现，threelinetable是一个有8个参数的命令。第一个参数和插入图片的一样，是一个可选参数，功能也一样。第一行的参数只有那个{lcr} 要说一下，大括号里的字母可以是lcrX四个字母。每个字母代表表中一列的对齐方式。lcr分别是左对齐、居中、右对齐。X是将剩余空间都分配给这列，并且这列内容可以自动换行。X可以有多个，那样多列宽度平均分配。

第二行和第三行的一对大括号中是表头内容。第三行和第七行的一对大括号中是表格的内容。这两部分的内容使用的都是 TeX 的标准表格格式：每个单元格间用 $\&$ 符号分开，每行结束要用 $\backslash\backslash$ 表示换行。

在802.11b旁边有一个 $\backslash\text{tnote}\{a\}$ 命令，它是用来添加表格脚注的命令。 a 是手工输入的编号，因为你可以表格中加许多个 a 。在最后三行的大括号中，要解释这个脚注 a 的内容，就像倒数第二行的内容那样。如果表格中没有脚注，将最后一个大括号内容空置即可。

这个普通的三线表也是一个浮动体，和图片一样，它的位置也是系统决定的。并且表格和表格的标题（题注）不会分在两页，这一点Word就差很多了。

我们有时还需要打印比较长的表格，我们希望它可以换页。这时可以用这个命令：

```
\lengthreelinetable{标签（英文数字组成）}{表标题}{列数}{lcr}
{
项目&频率&技术\backslash
}{
802.11b&2.4GHz&DSSS\backslash
}
```

表 2-4 可自动换页的超长三线表

项目	频率	技术
802.11b	2.4GHz	DSSS
蓝牙	2.4GHz	FHSS
HomeRF	2.4GHz	FHSS
802.11b	2.4GHz	DSSS
蓝牙	2.4GHz	FHSS
HomeRF	2.4GHz	FHSS
802.11b	2.4GHz	DSSS
蓝牙	2.4GHz	FHSS
HomeRF	2.4GHz	FHSS
802.11b	2.4GHz	DSSS
蓝牙	2.4GHz	FHSS
HomeRF	2.4GHz	FHSS
802.11b	2.4GHz	DSSS
蓝牙	2.4GHz	FHSS
HomeRF	2.4GHz	FHSS

接下页

接上页

表 2-5 用Excel制作的表格

总利润		¥ 3,590	¥ 3,800	¥ 5,050	¥ 3,700	¥ 4,130
费用	趋势					
薪金费用		¥ 100	¥ 180	¥ 130	¥ 80	¥ 220
工资费用		¥ 230	¥ 110	¥ 70	¥ 140	¥ 120
外部服务		¥ 230	¥ 200	¥ 30	¥ 160	¥ 100
供应品（办公和运营）		¥ 190	¥ 40	¥ 70	¥ 140	¥ 220
维修和维护		¥ 110	¥ 110	¥ 170	¥ 120	¥ 20
广告		¥ 20	¥ 160	¥ 60	¥ 130	¥ 110
汽车、交付和旅行		¥ 80	¥ 170	¥ 110	¥ 110	¥ 210
会计和法律		¥ 50	¥ 130	¥ 60	¥ 150	¥ 190
租金		¥ 80	¥ 40	¥ 230	¥ 250	¥ 100
电话		¥ 250	¥ 20	¥ 120	¥ 250	¥ 100
水电费		¥ 160	¥ 190	¥ 90	¥ 160	¥ 130
保险		¥ 120	¥ 90	¥ 160	¥ 190	¥ 250
税款（房地产等）		¥ 160	¥ 130	¥ 100	¥ 70	¥ 130
利息		¥ 30	¥ 20	¥ 190	¥ 210	¥ 130
折旧		¥ 80	¥ 70	¥ 60	¥ 70	¥ 70

2.4.3.6 图表设计技巧

学校规范中规定了图表中的文字应用宋体五号字。而图片中的字号本模板没有办法代为解决。但是我要介绍如何做到这一点。无论用户使用什么制图工具，一般都可以先设置画布的大小设置为426pt*672pt或者149mm*235mm。这样大小的画布和论文页面版心的大小基本相同，所以在这张画布上绘图可以以100%大小插入到论文中。也就是可以将\pic命令中指定图片大小的的大括号空置。这样在绘图的时候插入文字如果注意设置好字体为宋体五号字，在插入到论文后也能够保持字号了。

关于图片的格式转换。推荐大家下载开源的工具Inkscape: <http://www.inkscape.org/zh/>。在模板中请尽量插入矢量图，大部分理工科所用的制图工具都支持矢量图的导出。虽然矢量图格式会有所不同，但是Inkscape都能处理，并且都可以转换为pdf格式。在本模板中插入矢量图都需要先转换为pdf格式。

如何去白边是个常见的问题。去白边是指去掉图片四周多余的空白。用Inkscape打开图片，然后打开文件菜单，选择文档属性（快捷键Shift+Ctrl+D）。在页面选项卡中有自定义尺寸方格，点击“缩放页面到内容”的加号。最后点击“缩放页面到绘图或选区”按钮。最后保存文档即可。如果图片四周依然有很大的空白，则需要仔细检查图片的背景是否本身就是一块白色的图案。在矢量图中一般可以选择到，并删除它。

在注意了以上问题之后，其实用矢量图制作出来的表格效果完全没有任何问题，因此推荐复杂的表格用这样的方法制作。

2.4.3.7 插入脚注

本模板已经设置了脚注的格式，使之符合研究生院撰写范例中要求的格式。每页注脚不能超过10个。

通常注脚不会跨页。但是如果 \LaTeX 没有更好的办法了，它会把注脚分成两部分，后一部分排版到下一页的注脚区内。虽然这样的情况发生的概率非常小，但是如果出现之后，又不想注脚跨页，可以选择两种办法：

- 1.最好的办法是重新组织文字，使插入注脚的位置向前或者向后移动至少1行。
- 2.在主文件的`\documentclass`行之后加入一行：

```
\interfootnotelinepenalty=10000
```

这样 \LaTeX 会强制换页将注脚所在行移到下一页去，在前一页留下空白。

2.4.4 插入数学公式

本模板对数学式没有做过多修改，仅添加了`amsmath`和`kmath`宏包，以提供更美观的数学字体。关于数学式的编写规则请阅读“`TeXbook`中文版.pdf”中的16至19章。这本书是 \TeX 的作者高德纳写的 \TeX 权威著作，网友xianxian翻译。

行内公式：单个`$`包围的行内公式。例如：

若 $A \in F^{n \times n}$ ，且对任意非负整数 p ，有
编译结果是：

若 $A \in F^{n \times n}$ ，且对任意非负整数 p ，有
这就是行内公式。

行间公式：

```
\begin{equation}
```

```
A^D = \lim_{\varepsilon \rightarrow 0} (A^{p+1} + \varepsilon I)^{-1} C_A^{(p)}
```

```
\end{equation}
```

其中第一行和最后一行是公式环境的标志，中间依然按照数学式的输入规则输入。

编译结果为：

$$A^D = \lim_{\varepsilon \rightarrow 0} (A^{p+1} + \varepsilon I)^{-1} C_A^{(p)} \quad (2-1)$$

若无需显示标号，则将第一行和最后一行中的`equation`都替换为`equation*`。

\LaTeX 是数学学界广泛使用的工具，大部分数学工具例如`matlab`都支持导出 \LaTeX 公式。详情请自行在互联网上搜索。各种数学符号的输入方法可以在一些在线网站上通过手写查询出来：

<http://detexify.kirelabs.org/classify.html>

http://webdemo.visionobjects.com/equation.html?locale=zh_CN

互联网上任何有关 \LaTeX 数学式的技巧应该都适用于本模板。

2.4.5 定理环境等

本模板提供了四种分别编号的环境：定义、定理、公理、引理。和一种不带编号的环境：证明。这五种环境的环境名都是它们的拼音全拼。例如：

```
\begin{dingli}
```

通过任意闭合曲面的电通量等于该闭合曲面所包围的所有电荷量的代数和与电常数之比。

```
\end{dingli}
```

定理 2.4.1 通过任意闭合曲面的电通量等于该闭合曲面所包围的所有电荷量的代数和与电常数之比。

2.4.6 交叉引用

在正文的任意位置可以用命令`\label{标签名}`在此设置一个标签，然后在论文的其他位置可以使用命令`\ref{标签名}`和`\pageref{标签名}`分别显示该标签位于的章节号和页码。例如：此例见第 21 页的第 2.4.6 节。在输入`\ref{标签名}`命令时，如果其前后有是中文，可以在其与中文之间加入一个~符号，使其空一格更美观。

2.4.7 枚举环境

枚举环境就是如下的东西，使用方法见此处的源代码。可以引用其中编号，例如 6(e)vD。

1. 内容。
可以另起一段。
2. 内容。
3. 内容。
4. 内容。
5. 内容。
6. (a) 内容。
(b) 内容。
(c) 内容。
(d) 内容。
(e) i. 内容。

- ii. 内容。
- iii. 内容。
- iv. 内容。
- v. A. 最多到此4层嵌套。
 - B. 内容。
 - C. 内容。
 - D. 内容。

2.5 其他内容

其他内容指的是：中文摘要、英文摘要、目录、参考文献、鸣谢、外国原文、中文译文，还有可以省略的附录等，见表 2-6。这些内容的格式和顺序都是固定的，所以将它们单独放在一个目录中。使用者只需要编辑它们的内容，并不用选择按什么顺序载入它们。

表 2-6 contents 目录中文件用途

文件名	说明
titlepage.tex	封面信息
Cabstract.tex	中文摘要
Eabstract.tex	英文摘要
acknowledgements.tex	致谢
appendix.tex	附录
original.tex	外文资料原文，研究生可忽略此文件，不用编辑
translation.tex	外文资料译文，研究生可忽略此文件，不用编辑
reference.bib	参考文献
publications.bib	攻读期间的研究成果，本科生可忽略此文件，不用编辑

2.5.1 封面

我们校毕业论文的封面是由文印中心统一提供的，模板里的封面只是为了论文完整一些而设计的。

编辑contents/titlepage.tex即可输入封面的信息。

2.5.2 摘要

分别编辑两个摘要文件。其中：`\begin{Cabstrack}\{\}\{\}\{\}`是Cabstrack摘要环境的起始语句，从它到它对应的`\end{Cabstrack}`之间的部分就是这个环境的内容，也就是中文摘要的内容。`\begin{Cabstrack}`旁边有五个大括号，其中填写关键词。按照要求关键词应选择3–5个，所以至少填前三个。括号可以空着（其中不能有空格），但不能删除，必须有5对大括号。英文摘要也是一样的，只是环境名变成Eabstract了。

2.5.3 参考文献

用文本编辑器打开contents目录中的reference.bib文件（可用鼠标拖动文件到WinEdt窗口上打开）。模板带的reference.bib文件是一份示例文件，其中包含所有参考文献样式的示例各一条。可以按照其中的示例编写。其中多余3个作者的条目无需按照示例中的写法在最后手工写“等”或“et al”，参考文献模板的新版本已经解决了这一问题。

还可以通过Google学术搜索<http://scholar.google.com/>直接导出bibtex格式。打开Google学术搜索，点击页面上的“设置”，将其中的参考书目管理软件设置为显示导入BIBTEX的链接。之后搜索条目下方就会出现“导入BIBTEX”。将内容复制到reference.bib中即可。但是需要注意的是，Google等网站提供的内容有一些字段不完整的情况。使用时应对照原reference.bib中相同的参考文献类型，核对是否缺少某一字段。中文的参考文献要令`language = {Chinese}`（非空即可）。

注意bib文件中每一条目的第一行例如：

@ARTICLE{turongting2007,

注意后面的“,”不要删除。其中turongting2007是这条参考文献的标签，用于在文章中引用这个参考文献，以显示它的标号。标签必须是英文和数字组成的。Google学术搜索导出的条目标签是自动用作者名+年份+标题第一个词生成的，英文的当然没问题，但是中文文献的标签需要手工修改成一个英文和数字组成的标签。而ARTICLE是条目的类型。

在正文中引用参考文献时有两个命令可以使用。`\cite{标签名1, 标签名2……}`和`\citeup{标签名1, 标签名2……}`分别是正常字号的中括号引用和上角标形式的中括号引用。两个命令的大括号中可以写多个标签名，用英文逗号分隔开来。

本模板生成的参考文献格式符合GB/T7714-2005，但有一处错误需要手动修正。生成的英文析出文献的文献类型标识码之后会有一个句号（GB/T7714-2005

规定没有)。这个错误比较不明显,出现的频率也不多。可以只在最后提交正式论文时再行修复。

关闭打开的thesis.pdf文件。双击执行“compile.bat”即可自动修复这个错误并重新编译出thesis.pdf。

2.5.3.1 测试文献

A.1 普通图书

[4], [5], [6], [7], [8], [3], [2], [1], [9], [10], [11]。

A.2 论文集、会议录

[12], [13], [14]。

A.3 科技报告

[15], [16]。

A.4 学位论文

[17], [18]。

A.5 专利文献

[19], [20], [21]。

A.6 专著中析出的文献

[22], [23], [24], [25], [26], [27], [28]。

A.7 期刊中析出的文献

[29], [30], [31], [32], [33]。

A.8 报纸中析出的文献

[34], [35]。

A.9 电子文献

[36], [37], [38], [39], [40], [41]。

附加测试

B1.连续出版物4.3

[42], [43], [44]。

顺序编码制数字的压缩性测试

[3-8, 11, 22, 34], [2], [1], [9], [10], [11]。

B2. CAJ-CD B/T 1-2005

中国学术期刊(光盘版)检索与评价数据规范

Data Norm for Retrieval and Evaluation of Chinese Academic Journal(CD) (修订版试行稿)

CAJ-CD

中国学术期刊(光盘版)技术规范CAJ-CD B/T 1-2005

代替CAJ-CD B/T 1-1998

[45, 12], [46], [47],[48],[49], [50],[51],[52]

2.5.4 攻读期间的研究成果

攻读期间的研究成果文件是`publications.bib`，其格式和编辑方法同参考文献完全一致。在WinEdt中使用编译按钮编译的结果不会显示这一章是正常现象。需要使用`compile.bat`编译才能得到包含这一章的正确结果。

如果需要录入非参考文献标准格式的内容，例如获奖，可以使用ONLYNOTE类型的参考文献。这个类型是直接原文输出note条目的类型。可见示例[7]。

2.5.5 致谢

致谢就是简单的编辑`contents/acknowledgements.tex`就可以了。

2.5.6 附录

附录中如果分章，需要全部写在`contents/appendix.tex`文件中，即其中写多个`\chapter{}`。如果将`appendix.tex`中所有内容删除，最后的论文将不会出现附录。

2.5.7 外文资料

外文资料原文和译文分别是`contents/original.tex`和`contents/translation.tex`。它们都和普通的正文一样。

在v0.4.5版本更新时，我将外文资料的两章，章号重置为1，但是章标题不显示章号。而节标题从1.1开始编排。同时，外文资料原文中，图表题注都改为了英文。从示例文件中可以看到，这两章的章命令改为使用标准的`\chapter{}`命令，不再使用带星号的命令。

2.5.8 主要符号表和缩略词表

主要符号表和缩略词表的开启都不需要特别设置。只要使用了声明符号和缩略词的命令，并且使用了它们，就会自动出现在主要符号表和缩略词表中。如果只声明，而不使用则不会出现在两个表中。如果某个表没有内容，则不会出现这个表。

在v0.6版本更新后，新加入了缩略词表。并且原先实现主要符号表的机制也有所改变。简单说，之前只使用`\fuhao`命令设置符号就可以了。而现在依然保持了这个命令的定义不变，但是还需要在使用这个符号的位置调用引用符号命令，才能在主要符号表中显示这个符号。

定义一个符号使用命令：

`\fuhao{符号}{说明}{标签}`。这里标签还被用于在主要符号表中按字母顺序排序。

定义一个缩略词使用命令：

`\newacronym[description=中文全称]{标签}{缩写形式}{英文全称}`

然后在需要使用符号或缩略词的地方用命令：

`\gls{标签}`

使用举例：我先在这里定义一个符号`\fuhao{Π}{数学及物理学领域普遍存在的数学常数。}{pi}`，然后在需要用到的地方用命令`\gls{pi}`调用它。例如，圆周率的符号是 Π 。然后再定义一个缩略词

`\newacronym[description=支持向量机]{svm}{SVM}{Support Vector Machine}`

在需要使用这个缩略词的地方用命令`\gls{svm}`引用它。需要注意这个效果，当我第一次引用它时，它是先全称加括号缩略词的方式，而之后再引用它则只有缩略词的形式了。例如：

有一类算法叫做`\gls{svm}`，`\gls{svm}`的效果非常显著。

有一类算法叫做Support Vector Machine (SVM)，SVM的效果非常显著。

有关这两个表的更复杂的用法，请查阅glossaries宏包的相关手册。主要是如何定义缩略词和符号的特殊复数形式，以及如何灵活的引用出符号或缩略词的复数形式，首字母大写形式等。

2.6 排版源代码

许多学院的论文都需要插入若干程序源代码。经过多方比较，我推荐一个使用外部工具的方案。这个工具比现有的 \LaTeX 宏包解决方案要更美观，对中文的支持更好，几乎支持所有的语言。

从<http://www.andre-simon.de/zip/download.html>下载这个工具。这个工具就叫highlight，是开源软件，有各个平台的版本，完全不需要自行编译。Windows用户就直接下载Windows (W32) installer的连接就行。下载后进行完全安装。

这个工具界面是中文的，很容易使用。打开后，设置通用选项卡如图 2-4。

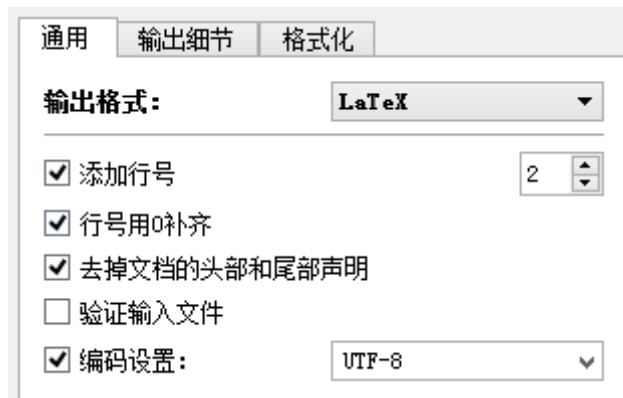


图 2-4 通用选项卡设置

输出细节选项卡不用动。设置格式化选项卡如图 2-5。其中高亮主题在模板的v0.4.7以后已经不用选择了。模板中内置了plain黑白主题。



图 2-5 格式化选项卡设置

然后设置输出路径，可以设置到chapters目录中。然后选择输入文件（文件路径中不能包含中文），可以批量转换很多源文件。最后点转换文件即可。转换好的文件是以源文件名后再加.tex 扩展名命名的。在论文中需要插入的位置写入：

```
\input{chapters/Place_has_TrainDA0.java.tex}
```

在2013/06/03更新的v0.4.7版本后，不再需要使用input命令导入highlight.sty。模板

中已经内置了这个文件的内容，采用黑白等宽字体。使用之前版本input了highlight.sty文件的同学需要删除这个导入命令，否则会报错。

在生成的tex文件中倒数第4行应该是排版的源代码的最后一行。这一行最后的换行命令\\通常是多余的、不该有的，应该删去。否则换段时会多出一个空行。这和我们不应该在一个段落的最后加上\\的道理是一样的。如果有大量的源代码需要处理，建议使用本文替换工具将多余的\\连同最后三行一起匹配，然后替换为仅留最后三行的内容。

如果需要完全的等宽效果，还需要手工修正一些符号。因为L^AT_EX对符号有特殊的处理，并不是普通的字体。需要按照表2-7在Highlight生成的tex文件中全文替换。

表 2-7 源代码中等宽符号替换

源代码中存在符号	搜索	替换为
ˆ	\textasciicircum _注	\^{} _注
	\textbar _注	\symbol{124}
\	\backslash \$	\symbol{92}
<	\$<\$	\symbol{60}
>	\$>\$	\symbol{62}
~	\$\sim \$	\symbol{126}
_	\textunderscore _注	\symbol{95}
{	\{	\symbol{123}
}	\}	\symbol{125}

_注 注意命令后有一个空格，连这个空格一起替换掉。下同。

示例结果如下：

```
01 !@#$%^&*()_+ = - ' ~ { } [ ] \ | : " ; ' , . < > / ?
02 abcdefghijklmnopqrstuvwxyzabcdef
03 QWERTYUIOPASDFGHJKLZXCVQWERTYUIO
04 12345678909876543211234567890987
```

```
01 package dbcontrol.dao;
02
03 import dbcontrol.vo.Place_has_TrainVO;
04 import dbcontrol.vo.TrainVO;
05 import dbcontrol.vo.VO;
```

```
06 import java.sql.ResultSet;
07 import java.sql.SQLException;
08 import java.util.LinkedList;
09 import java.util.logging.Level;
10 import java.util.logging.Logger;
11
12 /**
13  * Place_has_Train表DAO类
14  *
15  * @author Shifujun
16  */
17 public class Place_has_TrainDAO extends AbstractDAO {
18
19     Place_has_TrainVO vo;
20
21     @Override
22     public void createPreparedStatements() {
23         try {
24             insertStmtStr = "insert into Place_has_Train (
25                 Place_ID,Train_ID) values (?,?)";
26             insertStmt = dbutil.createPreparedStatement(
27                 insertStmtStr);
28
29             updateStmtStr = "";
30             updateStmt = dbutil.createPreparedStatement(
31                 updateStmtStr);
32
33             deleteStmtStr = "delete from Place_has_Train
34                 where Place_ID = ? and Train_ID = ?";
35             deleteStmt = dbutil.createPreparedStatement(
36                 deleteStmtStr);
37
38             selectStmtStr = "select * from Place_has_Train "
```

```
39         where Place_ID = ? and Train_ID = ?";
40         selectStmt = dbutil.createPreparedStatement(2
41         selectStmtStr);
42
43         selectListStmtStr = "select * from 2
44         Place_has_Train where Place_ID = ?";
45         selectListStmt = dbutil.createPreparedStatement(2
46         selectListStmtStr);
47     } catch (Exception ex) {
48         Logger.getLogger(Place_has_TrainDAO.class.2
49         getName()).log(Level.SEVERE, null, ex);
50     }
51 }
52
53 @Override
54 public void insertDAO() throws SQLException {
55     insertStmt.setInt(1, vo.getPlace_ID());
56     insertStmt.setString(2, vo.getTrain_ID());
57     insertStmt.executeUpdate();
58 }
59
60 @Override
61 public void deleteDAO() throws SQLException {
62     deleteStmt.setInt(1, vo.getPlace_ID());
63     deleteStmt.setString(2, vo.getTrain_ID());
64     deleteStmt.executeUpdate();
65 }
66
67 @Override
68 public void loadDAO() throws SQLException {
69     selectStmt.setInt(1, vo.getPlace_ID());
70     selectStmt.setString(2, vo.getTrain_ID());
71     ResultSet rs = selectStmt.executeQuery();
```

```
72         if (rs.next()) {
73             this.vo = new Place_has_TrainVO();
74             vo.setPlace_ID(rs.getInt("Place_ID"));
75             vo.setTrain_ID(rs.getString("Train_ID"));
76         } else {
77             vo = null;
78         }
79     }
80
81     @Override
82     public void setVO(VO vo) {
83         this.vo = (Place_has_TrainVO) vo;
84     }
85
86     @Override
87     public void updateDAO() throws SQLException {
88         //这个表不需要更新。
89     }
90
91     /**
92      * 查找和this.vo的地点相同的所有VO。
93      *
94      * @return
95      * @throws SQLException
96      */
97     public LinkedList<Place_has_TrainVO> getTrainListByPlace(
98     ) throws SQLException {
99         selectListStmt.setInt(1, vo.getPlace_ID());
100         ResultSet rs = selectListStmt.executeQuery();
101         LinkedList<Place_has_TrainVO> list = new LinkedList<
102         Place_has_TrainVO>();
103         while (rs.next()) {
104             Place_has_TrainVO vo = new Place_has_TrainVO();
```

```

105         vo.setPlace_ID(rs.getInt("Place_ID"));
106         vo.setTrain_ID(rs.getString("Train_ID"));
107         list.add(vo);
108     }
109     return list;
110 }
111
112 @Override
113 public VO getVO() {
114     return this.vo;
115 }
116 }

```

2.7 Linux中编译

本模板已经在Linux系统中测试过，可以正常编译。在TeXLive完全安装的基础上，所有宏包包括CJK都安装了，还需要另外安装宋体和黑体。本模板使用的宋体和黑体是Windows自带的，也是Word使用的。需要在Windows中C:

Windows

Fonts 目录中复制simhei.ttf和simsun.ttc到Linux 系统中，如果依然提示缺少某个字体就一样复制过去。最简单的方式是将两个字体文件放到UESTCthesis工作目录中。

编译脚本已经写好在zip包中，解压到工作目录即可使用。

2.8 使用L^AT_EX时容易犯的错误

1. 以为回车就是换行。L^AT_EX中换行的命令是“\\”。
2. 以为换行就是另起一段。L^AT_EX中另起一段的命令是“\par”，也可以连续输入两个以上的回车。
3. 以为空格就是空格。L^AT_EX中中文与中文之间的空格或回车都会被忽略掉，英文与英文之间的空格和回车有多少个都只算作1个。这是为了避免输入时录入不正确数量的空格。如果一定要输入连续多个空格，可以使用符号~，或者命令\hspace{~}。后者是不可换行的空格。比如英文人名Donald Ervin Knuth，如果将其中的空格改为\hspace{~}，则系统会优先保证这三个

第三章 示例：数学论文

M. P. Drazin^[2]在1958给出Drazin逆的定义, 其定义如下:

若 A 是复 n 阶方阵, 满足下列等式的 X 称为Drazin逆:

$$XAX = X, AX = XA, A^{k+1}X = A^k \quad (3-1)$$

其中 k 是 A 的指标, 即使得 $\text{rank}(A^{k+1}) = \text{rank}(A^k)$ 的最小整数。记为 A^D 。当 $k = 1$ 时, 称为 A 的群逆。记为 A^\sharp 。

此后, Drazin逆被广泛运用于很多方面。如: 奇异微分差分方程; Markov链; 迭代法; 数值分析等等。特别地, 这种广义矩阵在奇异线性常微分方程组以及奇异差分方程组求解问题中有重要的应用^[3-8, 11, 22, 34]。所以有必要对其进一步研究。

3.1 Drazin逆计算方法的简介

很明显, 一般通过(3-1)求 A 的Drazin逆比较困难。于是, 在此之后有很多人执立于这方面的研究。比如: C. D. Meyer^[34], U. G. Rothblum^[3], Y. Wei^[22]等等。其中C. D. Meyer早在1974年就给出一个用极限的形式求, 其表达如下:

若 $A \in F^{n \times n}$, 且对任意非负整数 p , 有

$$A^D = \lim_{\varepsilon \rightarrow 0} (A^{p+1} + \varepsilon I)^{-1} C_A^{(p)} \quad (3-2)$$

其中 $F^{n \times n}$ 为豪斯多夫拓扑空间上的 n 阶方阵, $C_A^{(p)} = A^{p+1}A^D$ 。

特别地, 若 $p \geq \text{ind}(A)$, 则

$$A^D = \lim_{\varepsilon \rightarrow 0} (A^{p+1} + \varepsilon I)^{-1} A^p$$

当 $p = \text{ind}(A) = k$ 时, 则

$$A^D = \lim_{\varepsilon \rightarrow 0} (A^{p+1} + \varepsilon I)^{-1} A^k$$

U. G. Rothblum在1976年也给出了Drazin逆的另外一种表达形式:

若 A 是一任意方阵, H 是 A 的本征投影, 则 $A - H$ 为非奇异阵且

$$A^D = (A - H)^{-1}(I - H) = (I - H)(A - H)^{-1} \quad (3-3)$$

其中 $H = I - AA^D = I - A^DA$

目前，比较常用的是建立在将矩阵 A 的约当标准型的基础之上的表达形式，该形式也是诸形式当中比较简单的一种。该形式是Campbell和Meyer^[2]在1979年给出的，其形式如下：

若 A 的约当标准型为

$$A = P \begin{pmatrix} D & 0 \\ 0 & N \end{pmatrix} P^{-1}$$

其中 D 是秩为 r 的非奇异阵， N 为指标为 k 的幂零矩阵，则

$$A^D = P \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \quad (3-4)$$

此外，Campbell^[2]也给出了另外一种表达形式：

$$A^D = A^k (A^{2k+1})^+ A^k \quad (3-5)$$

其中 A 是Moore-Penrose逆。

1996年Y. Wei在前人的基础上，给出了以下Drazin逆的表达形式：

若 $A \in C^{n \times n}$ 且 $\text{ind}(A) = k$ ，则

$$A^D = (\tilde{A})^{-1} A^k \quad (3-6)$$

其中 $\tilde{A} = A^{k+1}|_{R(A^k)}$ 是 A^{k+1} 在 $R(A^k)$ 上的限制。

基于第一类广义逆的基础上，程云鹏^[2]给出了以下的表达形式：

$$A^D = A^k (A^{2k+1})^{(1)} A^k \quad (3-7)$$

其中 $(A^{2k+1})^{(1)} \in A^{2k+1}\{1\}$ 。

该方法的优点是，仅需要一个 $\{1\}$ ，即可得到 A^D 。当 A 的指标容易求得并且较小时，可采用此法。

当矩阵 A 的阶数较高时，求 A 的指标是不容易的。另一方面，当 A 的病态严重时，求 A 的较高幂次又会使病态更严重。这时，采用式（3-7）进行计算是不适合的，最好采用Cline给出的逐次满秩分解的方法，该方法每一步都作较小阶矩阵的满秩分解，有限步后可以确定出矩阵的指标和Drazin逆。

设 $A \in C^{n \times n}$ 。令 A 的满秩分解为

$$A = B_1 G_1$$

而 $G_i B_i$ 的满秩分解为

$$G_i B_i = B_{i+1} G_{i+1} \quad (i = 1, 2, \dots)$$

第四章 示例：实施方案

本章首先介绍了本系统实现的设计原则与目标，然后描述了系统的整体架构与总体设计方案，接着阐述了系统的各个模块实现时的解决方案，最后给出了系统在 Android 具体实现时的实现方法和使用说明。

4.1 设计目标与原则

4.1.1 设计目标

本系统能够检测已知恶意软件及其变种，并能通过模糊检测发现具有相似恶意行为的未知恶意软件，为 Android 平台这样的开放式移动平台提供安全保障，可广泛用于各种型号的 Android 设备。系统具体设计目标如下：

1. 适用于目前主流的 Android 平板及手机，至少可运行于3.0版本系统。
2. 能够检测用户指定的程序是否为恶意程序。
3. 能够自动检测设备上的所有程序，并可定时检测。
4. 能够监控设备的程序安装行为，自动检测安装的程序是否为恶意程序。
5. 能够保证本程序自身的特征库不被破坏，并能及时修复和更新特征库。
6. 能够保证用户使用方便。

4.1.2 设计原则

从安全产品的特点出发，本系统设计与实现将遵循下列一些设计原则：

1. 高效性

系统运行效率高，可快速实现对目标程序的特征提取与检测，并将结果用最清晰简洁的方式告知用户。

2. 灵活性

为用户提供的各种功能具有可选性，用户可根据自己的需要选择其中的功能，而且用户可自行决定如何处理检测结果为恶意的程序。

3. 实用性

本系统应按照实用性原则进行设计，在保证对程序检测的同时，力求用户界面简洁友好。

4. 可扩展性

目前本系统只检测程序的 Java 实现部分，但是还有极少数程序代码是用 C 语言编写的，在后续的开发过程中，可以在不改变程序结构的前提下，实现这一部分的检测功能。

5. 健壮性

系统应具有应对非法操作的能力，并且当针对于本系统的恶意攻击到来时，可以及时防御，防止自身特征库遭到损坏。

4.2 系统方案

本系统由两部分构成，第一部分是产品部分，即 Android 应用程序，采用 Java 作为编程语言。第二部分是检测算法模型构建部分，采用 Matlab 实现，其输出的模型数据供产品部分作为特征库使用。故本系统的特征库构建于 PC 端，而对目标程序的检测运行于 Android 端。从而将构建过程中包含的巨大计算量留在 PC 端。其具体结构如图 4-1 所示。

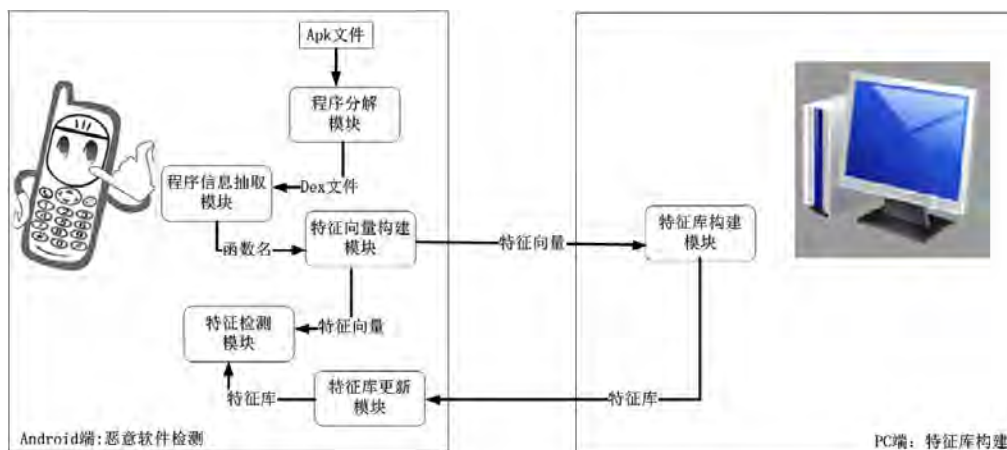


图 4-1 系统结构

其中程序信息抽取模块是本系统检测恶意软件的基础，特征检测模块是系统的核心与实现难点。特征检测模块中又分为变种检测模块和模糊检测模块，变种检测模块检测待检软件是否是已知恶意软件的变种，而模糊检测模块实现的是本系统从人脸匹配中引入的新型检测手段，可对特征库中不存在的未知恶意软件进行检测。

4.3 系统模块的实现

4.3.1 程序分解模块

Android 的程序文件为 APK 格式，APK 文件是 Android 最终的运行程序，是 Android Package 的全称。类似于 Symbian 操作系统中 sis 文件，APK 文件其实是 Zip 文件格式，但后缀名被修改为 APK。通过解压，可以看到 Dex 文件。Dex 是 Dalvik VM executes 的全称，即 Dalvik 虚拟机可执行文件，并非 Java ME 的字节码而是 Dalvik 字节码。

一个 APK 文件结构为：

1. META-INF\———签名信息，用来保证 apk 包的完整性和系统的安全，jar 文件经常可以看到；
2. res\———资源文件夹，包括程序中使用的图片，布局文件等；
3. AndroidManifest.xml\———项目配置清单，但不是明文的 XML 格式，无法直接打开阅读；
4. classes.dex\———Dalvik 可执行二进制文件，在运行时被动态优化为 dex 文件并由 Dalvik 虚拟机解释执行；
5. resources.arsc\———编译后的二进制资源文件，资源文件打包而成，字符串值（源码中的 /value/Strings.xml）就在其中；
6. lib\———动态链接库文件；
7. assets\———原始文件文件夹，其中的文件不会被压缩，也不能像 res 目录下的资源文件一样通过资源类引用。

图 4-2 是我们解压缩 helloworld.apk 文件后看到的内容，可以看到其结构跟工程结构有些类似。

```
01. | -- AndroidManifest.xml
02. | -- META-INF
03. |   | -- CERT.RSA
04. |   | -- CERT.SF
05. |   | -- MANIFEST.MF
06. | -- classes.dex
07. | -- res
08. |   | -- drawable
09. |   |   | -- icon.png
10. |   |   | -- layout
11. |   |     | -- main.xml
12. | -- resources.arsc
```

图 4-2 helloworld.apk 的结构

classes.dex 文件是 Java 源码编译后生成的 Java 字节码文件。但由于 Android 使用的 Dalvik 虚拟机与标准的 Java 虚拟机是不兼容的，Dex 文件与 Class 文件相

比，不论是文件结构还是 opcode 都不一样。目前常见的 Java 反编译工具都不能处理 Dex 文件。Android SDK 中提供了一个 Dex 文件的反编译工具 dexdump。用法为，dexdump -d -f -h xxx.dex。

指令参数解释：

- d：反编译程序段
- f：从文件头显示摘要信息
- h：显示文件头详情
- C：反编译低级符号名
- S：只计算大小

在知道了程序安装包是 Zip 编码之后，我们就可以通过遍历 Zip 包中包含的项目名，找到程序的二进制文件，即 Dex 文件。并在内存中建立一段缓冲区，可将 Dex 文件读入内存，再写到指定的临时文件中。分解流程如图 4-3 所示。

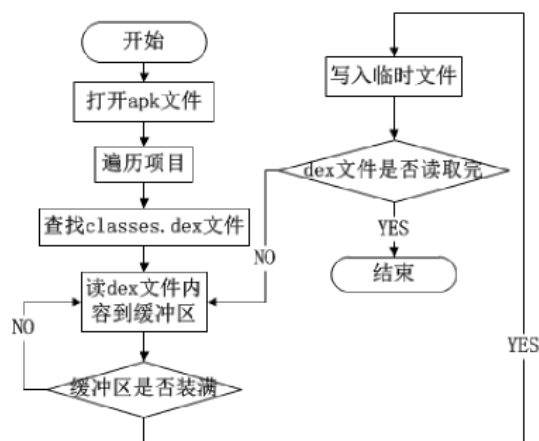


图 4-3 程序分解流程

4.3.2 构建特征向量模块

特征向量是数据挖掘中的一个概念，一个数据集中的每个数据实例都可以用一组属性值来描述，每一个数据实例都具有一个特殊的目标属性，称为类属性，它表征每个数据实例归属的类。这一组属性值即是代表这个数据实例的特征向量。在我们的检测问题中，Android 系统 API 和 Java 标准函数就是我们定义的属性，而恶意和非恶意就是类属性。这一过程如图 4-4 所示。

我们将特征向量空间 (Ω) 存储在数据库中，数据表的定义见表 4-1 所示。

构造特征向量时，先初始化一个全为 0 的特征向量，然后利用 SQL 查询语句确定代表某一函数名的维度序号：

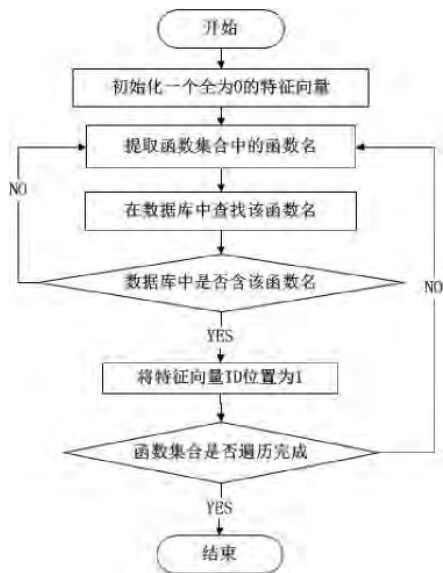


图 4-4 构建特征向量流程

表 4-1 Ω 数据表定义

字段	主键	类型	是否为空	备注
ID	是	Int	NOT NULL	特征向量维度序号
MethodName	否	Text	NOT NULL	函数名

select ID from Omega where MethodName = “待查函数名”;

并将特征向量 (ω) 的第 ID 位设置为1。见式(4-1)。

$$\omega_i = \begin{cases} 1 & i = ID \\ 0 & else \end{cases} \quad (4-1)$$

4.3.3 特征库构建模块

特征库构建模块实现于 PC 端，恶意样本来自各大权威机构公布的数据，详见第41页4-1小节，其流程如图 4-5 所示，特征库数据模型构建方法如下：

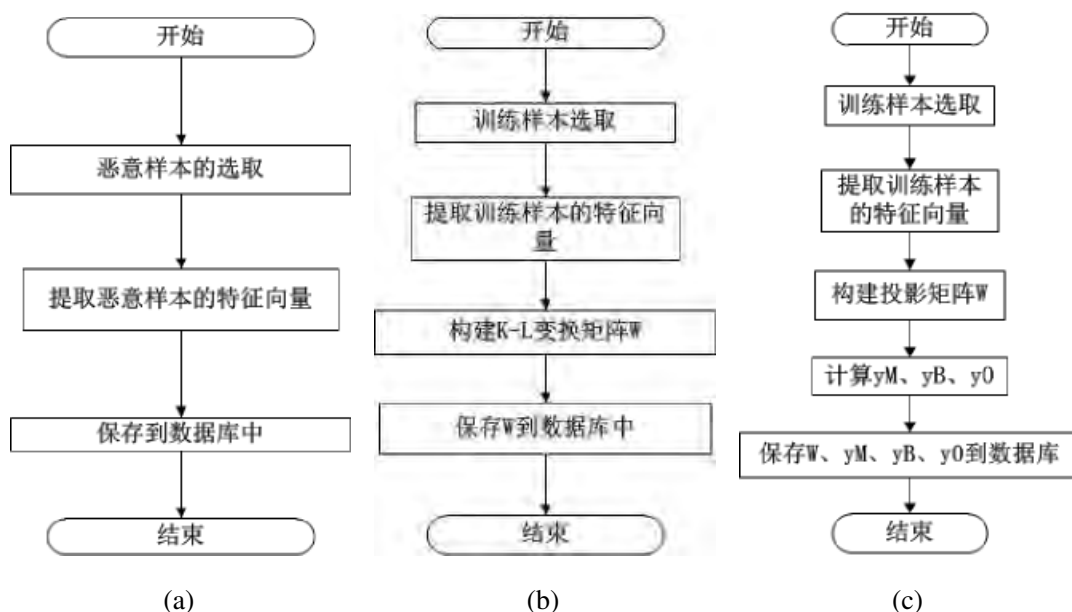


图 4-5 特征库构建流程

(a)KNN特征库构建流程；(b)K-L 变换矩阵构建流程；(c)LDA 投影矩阵构建流程

1. 最近邻居(KNN)算法

KNN 算法的特征库就是恶意程序样本的特征向量集合，我们将这些特征向量存储到数据库中。

2. 主成分分析（PCA）算法

考虑到文献^{[8][7]}中的适用条件，由于我们的特征向量维度远大于样本数量，所以需要去掉冗余数据，使训练数据矩阵为可逆矩阵才能使用线性判别分析（LDA）算法。

PCA 方法主要是通过对协方差矩阵进行本征分解，以得出数据的主成分（即本征矢量）与它们的权值（即本征值）。PCA 提供了一种降低数据维度

的有效办法；如果分析者在原数据中除掉最小的本征值所对应的成分，那么所得的低维度数据必定是最优化的（也即这样降低维度必定是失去信息最少的方法）。

我们的目标是把高维的数据集 Ω_B 和 Ω_M 变换成具有较小维度的数据集 Y_B 和 Y_M 。 Y_B 和 Y_M 是矩阵 Ω_B 和 Ω_M 的 Karhunen–Loève 变换（K-L 变换）。即 $\mathbf{Y} = \text{KLT}\{\mathbf{X}\}$ 。

计算特征向量平均值见式(4-2)。

$$u = \frac{1}{N} \sum_{\omega \in \Omega_B \cup \Omega_M} \omega \quad (4-2)$$

从 Ω_B 和 Ω_M 中减去平均值 u 见式(4-3)。

$$B = \begin{bmatrix} \Omega_B \\ \Omega_M \end{bmatrix} - hu \quad (4-3)$$

其中 h 是全为1的列向量。

求协方差矩阵 C 见式(4-4)。

$$C = B \cdot B^T \quad (4-4)$$

计算 C 的特征值和特征向量，提取不为0的特征值所对应的特征向量，构成 K-L 变换矩阵 W 。

所以， Y_B 和 Y_M 可由式(4-5)计算。

$$\begin{aligned} Y_B &= \Omega_B \cdot W \\ Y_M &= \Omega_M \cdot W \end{aligned} \quad (4-5)$$

最后我们将 K-L 变换矩阵 W 存储在数据库中。

3. Fisher 线性判别分析（LDA）

应用统计方法解决模式识别问题时，一再碰到的问题之一是维数问题。在低维空间里解析上或计算上行得通的方法，在高维空间里往往行不通。因此，降低维数有时就成为处理实际问题的关键。在数学上总是可以把高维空间样本投影到一条直线上，形成一维空间，即把维数压缩到一维。但是投影方向有无数种，若把样本投影到一条任意的直线上，可能使几类样本混在一起无法区分，如图 4-6(a) 所示。。但在一般情况下，总可以找到某个方向，使在这个方向的直线上，样本的投影能分开得最好，如图 4-6(b) 所示。。问题是如何根据实际情况找到这条最好的、最易于区分的投影线。这就是 Fisher 法所要解决的基本问题。

表 4-2 LDA算法变量定义

变量	定义	变量	定义
S_b	样本类间离散度矩阵	x	一个程序
S_i	样本类内离散度矩阵	X_M	恶意程序集合
S_w	总类内离散度矩阵	X_B	非恶意程序集合
W	投影方向向量	y_M	恶意样本的投影值
$J_F(W)$	Fisher 准则函数	y_B	非恶意样本的投影值
M	恶意 (Malice) 的缩写	y_0	识别阈值点
B	非恶意 (Benign) 的缩写		

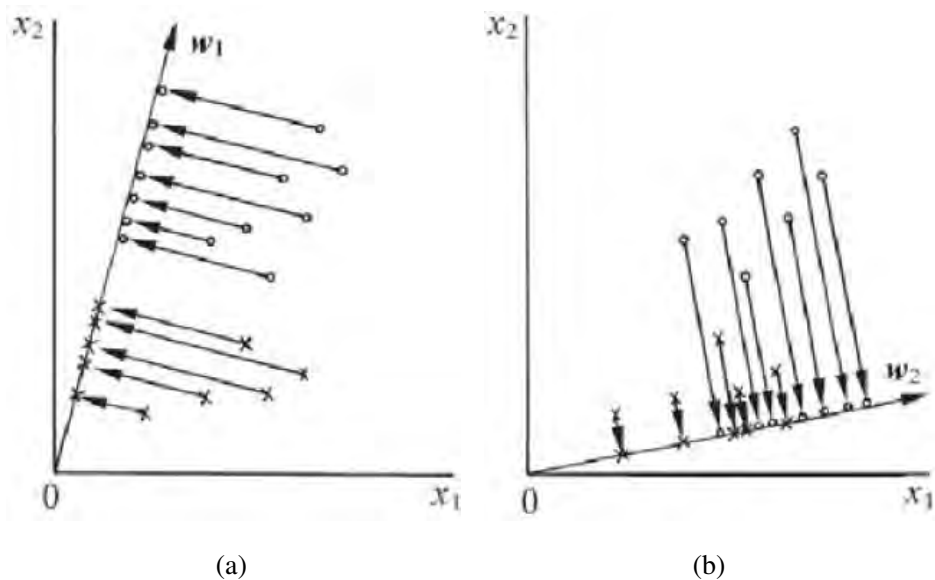


图 4-6 Fisher 线性判别基本原理

(a)最优方向投影; (b)K-L 任意方向投影

描述 LDA 算法前，首先定义几个基本变量，变量定义见表 4-2。LDA 算法步骤如下：

(a) 计算样本均值向量 m_i ：

$$m_i = \frac{1}{N_i} \sum_{y \in Y_i} y, i = B, M$$

(b) 计算样本类内离散度矩阵 S_i 和总类内离散度矩阵 S_w ：

$$S_i = \sum_{y \in Y_i} (y - m_i)(y - m_i)^T, i = B, M \quad (4-6)$$

$$S_w = P(x|x \in X_B)S_B + P(x|x \in X_M)S_M \quad (4-7)$$

(c) 计算样本类间离散度矩阵 S_b ：

$$S_b = P(x|x \in X_B)P(x|x \in X_M)(m_B - m_M)(m_B - m_M)^T$$

$P(x|x \in X_B)$ 和 $P(x|x \in X_M)$ 是恶意程序和非恶意程序的先验概率，根据目前 Android 市场的情况，我们取 $P(x|x \in X_M) = 0.001$ 。

(d) Fisher 准则函数为：

$$J_F(W) = \frac{W^T S_b W}{W^T S_w W} \quad (4-8)$$

为求函数取极大值时的 W^* 。可用拉格朗日乘数法，定义拉格朗日函数为：

$$\frac{L(W, \lambda)}{W} = S_b W - \lambda S_w W$$

另偏导数为零，得

$$S_b W^* = \lambda S_w W^*$$

其中 W^* 就是 $J_F(W)$ 的极值解。因为 S_w 可逆，等式两边左乘 S_w^{-1} ，可得

$$S_w^{-1} S_b W^* = \lambda W^*$$

所以求 W^* 即求矩阵 $S_w^{-1} S_b$ 的特征值问题。在我们这个特殊情况下，只有两种类别，故

$$S_b W^* = (m_B - m_M)(m_B - m_M)^T W^*$$

其值为一标量，所以对 W 投影方向无影响。忽略这个标量的比例因子可得，

$$W^* = S_w^{-1} (m_B - m_M)$$

(e) 求出 W^* 后即可计算：

$$y_M = \text{mean}(W^*T \cdot Y_M) \quad (4-9)$$

$$y_B = \text{mean}(W^*T \cdot Y_B) \quad (4-10)$$

$$y_0 = \frac{m_B + m_M}{2} + \frac{\ln(P(x|x \in X_B)/P(x|x \in X_M))}{N_B + N_M - 2} \quad (4-11)$$

最后将LDA变换矩阵 W 、 y_M 、 y_B 、 y_0 保存在数据库中。

4.4 本章小结

本章介绍了 AFace 系统的实现细节，它是一个由 Android 端和 PC 端两部分组成的系统。我们首先介绍了系统的设计原则。然后介绍了我们在此原则下设计的检测流程及其背后的检测原理。最后，作为一款完整的产品，还介绍了系统的产品软件架构以及产品的界面设计。

致 谢

历时将近两个月的时间终于将这篇论文写完，在论文的写作过程中遇到了无数的困难和障碍，都在同学和老师的帮助下度过了。尤其要强烈感谢我的论文指导老师—XX老师，她对我进行了无私的指导和帮助，不厌其烦的帮助进行论文的修改和改进。另外，在校图书馆查找资料的时候，图书馆的老师也给我提供了很多方面的支持与帮助。在此向帮助和指导过我的各位老师表示最中心的感谢！

感谢这篇论文所涉及到的各位学者。本文引用了数位学者的研究文献，如果没有各位学者的研究成果的帮助和启发，我将很难完成本篇论文的写作。

感谢我的同学和朋友，在我写论文的过程中给予我了很多你问素材，还在论文的撰写和排版灯过程中提供热情的帮助。由于我的学术水平有限，所写论文难免有不足之处，恳请各位老师和学友批评和指正！

参考文献

- [1] International Federation of Library Association and Institutions. Names of persons: national usages for entry in catalogues[M]. 3rd ed. London: IFLA International Office for UBC, 1977.
- [2] W. Crawford, M. Gorman. Future libraries: dreams, madness, & reality[M]. Chicago: American Library Association, 1995.
- [3] 赵耀东. 新时代的工业工程师[M/OL]. 台北: 天下文化出版社, 1998[1998-09-26]. [http://www.ie.nthu.edu.tw/info/ie.newie.htm\(Big5\)](http://www.ie.nthu.edu.tw/info/ie.newie.htm(Big5)).
- [4] 广西壮族自治区林业厅. 广西自然保护区[M]. 北京: 中国林业出版社, 1993.
- [5] 蒋有绪, 郭泉水, 马娟, 等. 中国森林群落分类及其群落学特征[M]. 北京: 科学出版社, 1998.
- [6] 唐绪军. 报业经济与报业经营[M]. 北京: 新华出版社, 1999:117-121.
- [7] 赵凯华, 罗蔚茵. 新概念物理教程:力学[M]. 北京: 高等教育出版社, 1995.
- [8] 汪昂. (增补)本草备要[M]. 石印本. 上海: 同文书局, 1912.
- [9] J. A. O'Brien. Introduction to information systems[M]. 7th ed. Burr Ridge, Ill.: Irwin, 1994.
- [10] H. J. Rood. Logic and structured design for computer programmers[M]. 3rd ed.[S.l.]: Brooks/Cole-Thomson Learning, 2001.
- [11] 昂温 G, 昂温 P S. 外国出版史[M]. 陈生铮, 译. 北京: 中国书籍出版社, 1988.
- [12] 中国力学学会. 第3届全国实验流体力学学术会议论文集[C]. 天津:[出版者不详], 1990.
- [13] E. M. Rosenthal. Proceedings of the Fifth Canadian Mathematical Congress, University of Montreal, 1961[C]. Toronto: University of Toronto Press, 1963.
- [14] V. G. Ganzha, E. W. Mayr, E. V. Vorozhtsov. Computer algebra in scientific computing: CASC 2000: proceedings of the Third Workshop On Computer Algebra in Scientific Computing, Samarkand, October 5-9.2000[C]. Berlin: Springer, 2000.
- [15] U. S. Department of Transportation Federal Highway Administration. Guidelines for handling excavated acid-producing materials, PB 91-194001[R]. Springfield: U.S. Department of Commerce National Information Service, 1990.
- [16] World Health Organization. Factors regulating the immune response: report of WHO Scientific Group[R]. Geneva: WHO, 1970.
- [17] 张志祥. 间断动力系统的随机扰动及其在守恒律方程中的应用[D]. 北京: 北京大学数学学院, 1998.
- [18] R. B. Calms. Infrared spectroscopic studies on solid oxygen[D]. Berkeley: Univ. of California, 1965.

- [19] 刘加林. 多功能一次性压舌板: 中国, 92214985.2[P]. 1993-04-14.
- [20] 河北绿洲生态环境科技有限公司. 一种荒漠化地区生态植被综合培育种植方法: 中国, 01129210.5[P/OL]. 2001-10-24[2002-05-28]. <http://211.152.9.47/sipoasp/zlijs/hyjs-yx-new.asp?Recid-01129210.5&leixin>.
- [21] A. Koseki, H. Momose, M. Kawahito, et al. Compiler: US, 828402[P/OL]. 2002-05-25[2002-05-28]. <http://FF&p=1&u=netahtml/PT0/search-bool.html&r=5&f=G&l=50&col=AND&d=PG01&sl=IBM.AS.&.OS=AN/IBM&.RS=AN/IBM>.
- [22] 白书农. 植物开花研究[M]//李承森. 植物科学进展. 北京: 高等教育出版社, 1998:146–163.
- [23] 国家标准局信息分类编码研究所. GB/T2659-1986世界各国和地区名称代码[S]//全国文献工作标准化技术委员会. 文献工作国家标准汇编:3. 北京: 中国标准出版社, 1988:59–92.
- [24] 韩吉人. 论职工教育的特点[G]//中国职工教育研究会. 职工教育研究论文集. 北京: 人民教育出版社, 1985:90–99.
- [25] BUSECK P R, NORD G L, Jr., VEBLEN D R. Subsolidus phenomena in pyroxenes[M]//C. T. Prewitt. Pyroxense.[S.l.]: Mineralogical Society of America, c1980:117–211.
- [26] M. E. Fourny. Advances in holographic photoelasticity[C]//American Society of Mechanical Engineers. Applied Mechanics Division. Symposium on Applications of Holography in Mechanics, August 23-25, 1971. University of Southern California, Los Angeles, California. New York: ASME, c1971:17–38.
- [27] 裴丽生. 在中国科协学术期刊编辑工作经验交流会上的讲话[C]//中国科协学术期刊编辑工作经验交流会资料选. 北京: 中国科学技术协会学会工作部, 1981:2–10.
- [28] G. Martin. Control of electronic resources in Australia[M]//L. W. PATTLE, B. J. COX. Electronic resources: selection and bibliographic control. New York: The Haworth Press, 1996:85–96.
- [29] 李炳穆. 理想的图书馆员和信息专家的素质与形象[J]. 图书情报工作, 2000(2):58.
- [30] 陶仁骥. 密码学与数学[J]. 自然杂志, 1984, 7(7):527.
- [31] 亚洲地质图编目组. 亚洲地层与地质历史概述[J]. 地质学报, 1978, 3:194–208.
- [32] D. J. Des Marais, H. Strauss, R. E. Summons, et al. Carbon isotope evidence for the stepwise oxidation of the Proterozoic environment[J]. Nature, 1992, 359:605–609.
- [33] J. A. Hewitt, J. A. Hewitt. Technical services in 1983[J]. Library Resource Services, 1984, 28(3):205–218.
- [34] 丁文祥. 数字革命与竞争国际化[N]. 中国青年报, 2000-11-20(15).
- [35] 张田勤. 罪犯DNA库与生命伦理学计划[N]. 大众科技报, 2000-11-12(7).
- [36] 江向东. 互联网环境下的信息处理与图书管理系统解决方案[J/OL]. 情报学报, 1999, 18(2):4[2000-01-18]. <http://www.chinainfo.gov.cn/periodical/abxb/qbxb99/qbxb990203>.

- [37] 萧钮. 出版业信息化迈人快车道[EB/OL].(2001-12-19)[2002-04-15]. <http://www.creader.com/news/20011219/200112190019.html>.
- [38] M. Christine. Plant physiology: plant biology in the Genome Era[J/OL]. Science, 1998, 281:331-332[1998-09-23]. <http://www.sciencemag.org/cgi/collection/anatmorp>.
- [39] S. W. Metcalf. The Tort Hall air emission study[C/OL]//The International Congress on Hazardous Waste, Atlanta Marriott Marquis Hotel, Atlanta, Georgia, June 5-8,1995. impact on human and ecological health. 1995[1998-09-22], <http://atsdrl.atsdr.cdc.gov:8080/cong95.html>.
- [40] D. L. Turcotte. Fractals and chaos in geology and geophysics[M/OL]. New York: Cambridge University Press, 1992[1998-09-23]. <http://www.seg.org/reviews/mccorm30.html>.
- [41] Scitor Corporation. Project scheduler[CP/DK]. Sunnyvale, Calif.: Scitor Corporation.
- [42] 中国地质学会. 地质论评[J].1936,1(1)-. 北京: 地质出版社, 1936-.
- [43] 中国图书馆学会. 图书馆学通讯[J].1957(1)-1990(4). 北京: 北京图书馆, 1957-1990.
- [44] American Association for the Advancement of Science. Science[J].1883,1(1)-. Washington, D. C.: American Association for the Advancement of Science, 1883-.
- [45] 朱一玄. 聊斋志异资料汇编[G]. 郑州: 中州古籍出版社, 1985:177-178.
- [46] 叶剑英. 叶委员剑英关于安平事件调查结果的声明:1946-09-09[B]. [出版地不详]: 中央档案馆, 1946.
- [47] . 国务院外国专家局的报告:1958-12-11[B]. 呼和浩特: 内蒙古自治区档案馆(全宗252,目录1,卷宗57), 1958:65-67.
- [48] 沈括. 梦溪笔谈[O]. 大德九年茶陵刊本.[出版地不详]: 北京图书馆珍藏, 1070.
- [49] 纪昀. 纪文达公遗集:卷十六[O]. [出版地不详]: [出版者不详], 1800.
- [50] 刘江. 假如陈景润被量化考核[N/OL]. 新华每日电讯, 2004-03-12(7)[2004-04-04]. <http://search.cnki.net/ccnd/main-frame.asp?encode=gb&display=chinese>.
- [51] 万锦坤. 中国大学学报论文文摘(1983-1993)[DB/CD]. 英文版. 北京: 中国大百科全书出版社, 1996.
- [52] DAI S, FEI H, WANG L, et al. Reliability Test and Statistical Analysis[M]. Beijing: National Defense Industrial Press, 1983. (in Chinese).

附录 A 附录章

如果将appendix.tex中所有内容删除，最后的论文将不会出现附录。

A.1 附录节

附录 B 附录另一章

B.1 附录另一章的一节

攻博期间取得的科研成果

- [1] C. Cong, H. Wang, Y. Hu, et al. Synthesis and antibacterial activity of novel 4''-O-benzimidazolyl clarithromycin derivatives[J]. European Journal of Medicinal Chemistry, 2011, 46(7):3105 – 3111. <http://www.sciencedirect.com/science/article/pii/S0223523411002893>.
- [2] X. Ma, L. Zhang, R. Wang, et al. Novel C-4'' modified azithromycin analogs with remarkably enhanced activity against erythromycin-resistant Streptococcus pneumoniae: The synthesis and antimicrobial evaluation[J]. European Journal of Medicinal Chemistry, 2011, 46(10):5196 – 5205. <http://www.sciencedirect.com/science/article/pii/S0223523411005721>.
- [3] C. Liu, C. Hao. Morinda Officinalis Polysaccharides' Impact on the Immunologic Balance of Obstructive Jaundice Rats' Lymphocyte T[J]. Chinese Journal of Cellular and Molecular Immunology, 2011, 27(6):678 – 679.
- [4] C. Liu, C. Hao. Morinda Officinalis Polysaccharides' Impact on the Expression of Proteins MCP-1 and IL-8 in Hepatic Stellate Cells[J]. Chinese Journal of Gerontology, 2011, 31(14):2713 – 2714.
- [5] J. Liu, C. Liu, W. He. Fluorophores and their applications as molecular probes in living cells[J]. Current Organic Chemistry Accepted.
- [6] H. Li, C. Liu, J. Li, et al. Bifunctional Silicon for Cyclic Ether Opening Catalyzed by Nanoporous Gold[J]. Manuscript in preparation.
- [7] XXX, XXX, XXX, XXX, 王升. XXX的陶瓷研究. 四川省科技进步三等奖, 2003年12月.