**词频统计需求分析实验报告**

**1.引言**

**1.1 编写目的**

①该文档是对于英语词频统计器的功能和性能的描述，重点描述了统计器的功能。

②该项目的预期读者包括：需求分析人员、设计人员、开发人员、项目管理人员、测试人员、用户、老师。

**1.2 项目背景**

项目名称：英语词频统计器

项目提出者：华中农业大学信息学院

项目开发者：华中农业大学咖喱gaygay团队

用户：学生，教师

项目实施单位：咖喱gaygay团队

与其他系统的关系：本系统独立运行

**1.3 参考资料**

《软件工程方法与实践 第三版》，机械工业出版社

《Eclipse从入门到精通》，清华大学出版社

《UML参考手册》，机械工业出版社

《J2EE设计模式》，中国电力出版社

《The J2EETM Turial 中文版》，中国铁道出版社

**1.4版本信息**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 修改编号 | 修改日期 | 修改后版本 | 修改位置 | 修改内容概述 |
| 1 | 2017.6.10 | 1.0 | 全部 | 完成第一次编写 |

**2、任务概述**

**2.1 系统定义**

2.1.1 项目来源及背景

该统计器是华中农业大学信息学院提出的对大二学生进行暑期综合实训的项目，而统计器可以供给老师、学生对英语文档进行各方面的统计。

2.1.2 项目要达到的目标

①可以对TXT、Doc格式的文档进行统计；

②统计文档中出现的所有单词及词频数，将单词和词频导出为Excel或文本文件；

③按词频统计结果对单词进行词频升降序、字母升降序排序，结果导出为Excel或文本文件；

④统计文档中的超纲词汇及词频；

⑤统计文档中单词的各种变化形式的数目，包括复数、不规则、过去分词、现在分词、比较级和最高级；

⑥提取某个单词长度、提取词头和词尾含有指定字符串的词，统计数目；

⑦不考虑单词变形统计两篇文章中总词汇数目及相同词汇数目，显示相同词。

2.1.3 系统各部分组成、与其他部分关系、各部分接口等

本系统是独立的系统，不需要与其他系统连接

**2.2 运行环境**

2.2.1 设备环境

普通PC即可

2.2.2 硬件环境

可以正常使用并运行的PC

2.2.3 软件环境

该软件适用于WindowsXP、Win7、Win10等操作系统，如有特殊需求可开发支持Linux或者Mac OS平台的版本

2.2.4 操作环境

计算机程序操作

**2.3 条件限制**

2.3.1 列出本软件开发工作的假定和约束，如经费限制、开发期限等

本项目要求在Eclipse上进行开发，需要JDK的支持。本项目由15级学生开发，经验不足，需要在教室指导下完成。开发经验较少。

2.3.2 列出本软件最终用户、用户的教育水平和技术专长

最终用户一般为学生与老师，没有特定的教育水平与技术专长的限制，有需求即可使用。

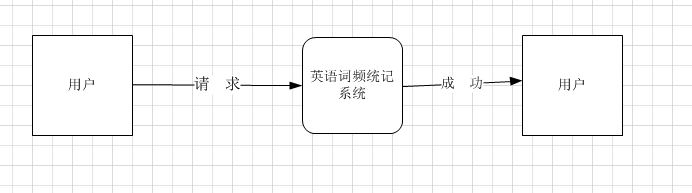
2.3.3 列出本软件的预期使用频度

本软件主要是统计单词功能，对于英语老师来说使用频度较高；而对于其他老师及学生而言频度相对较低

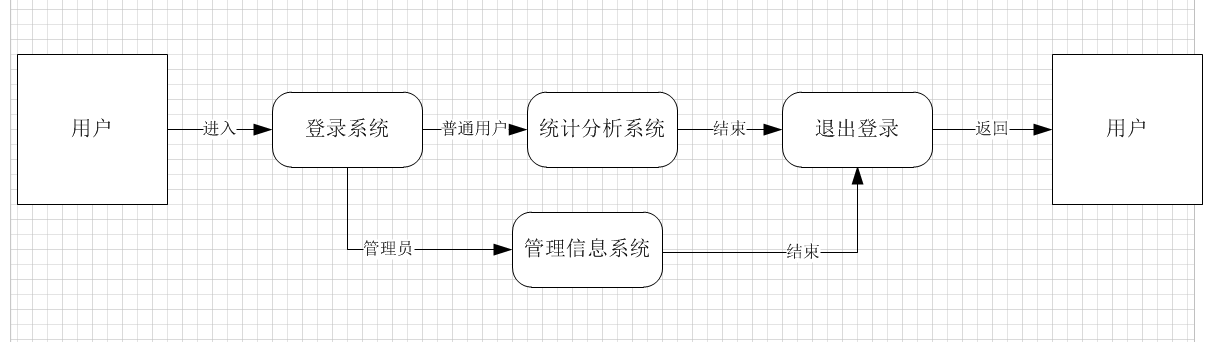
**3.数据描述**

**3.1数据流模型**

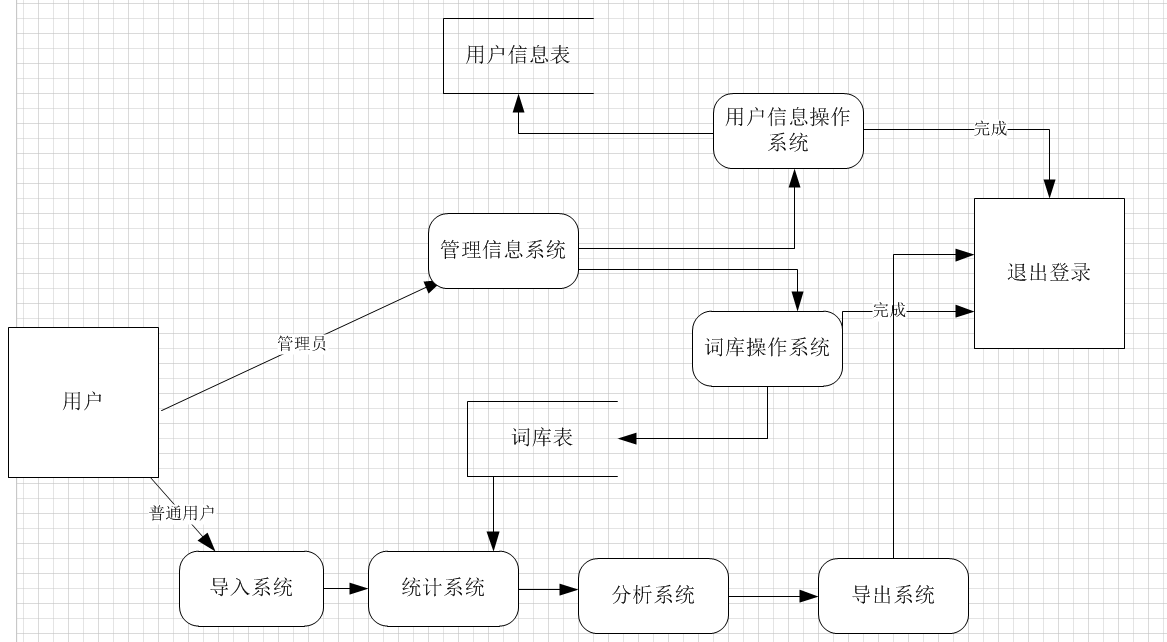
3.1.1顶层设计



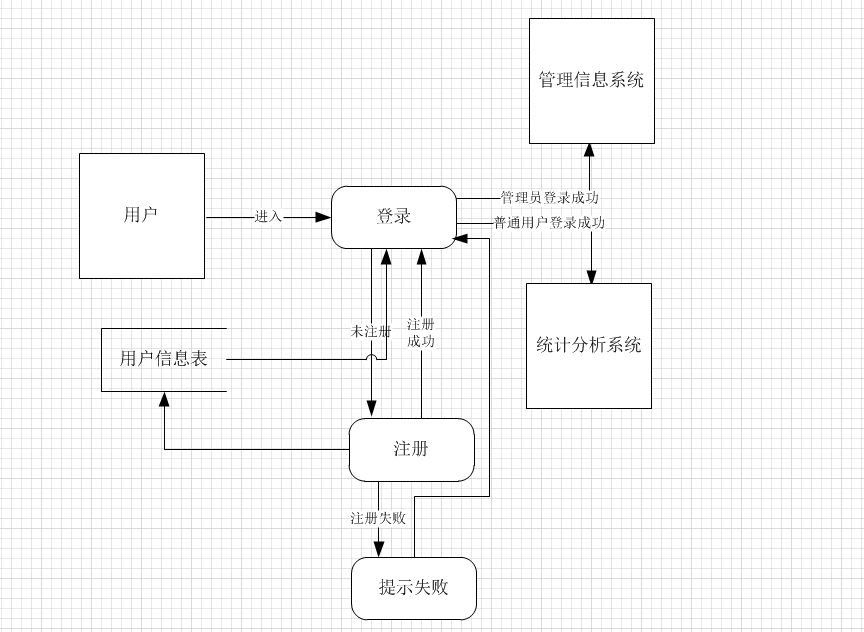
3.1.2第一层设计



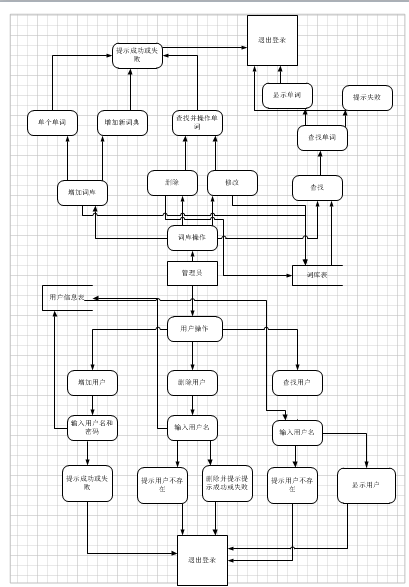
3.1.3第二层设计



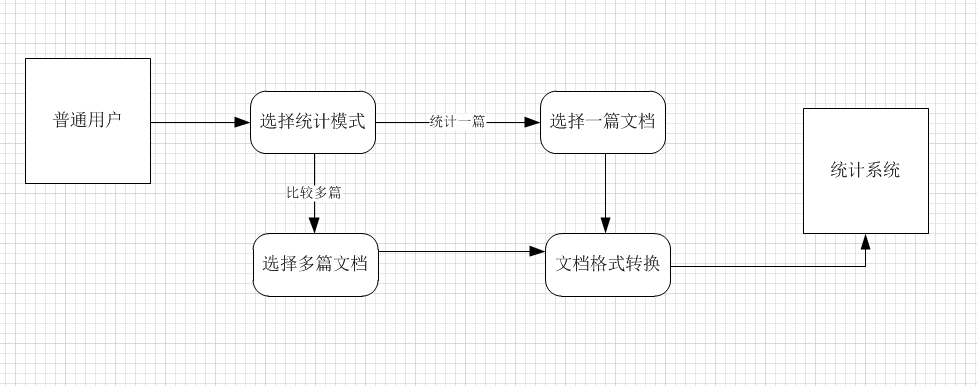
3.1.4登陆系统



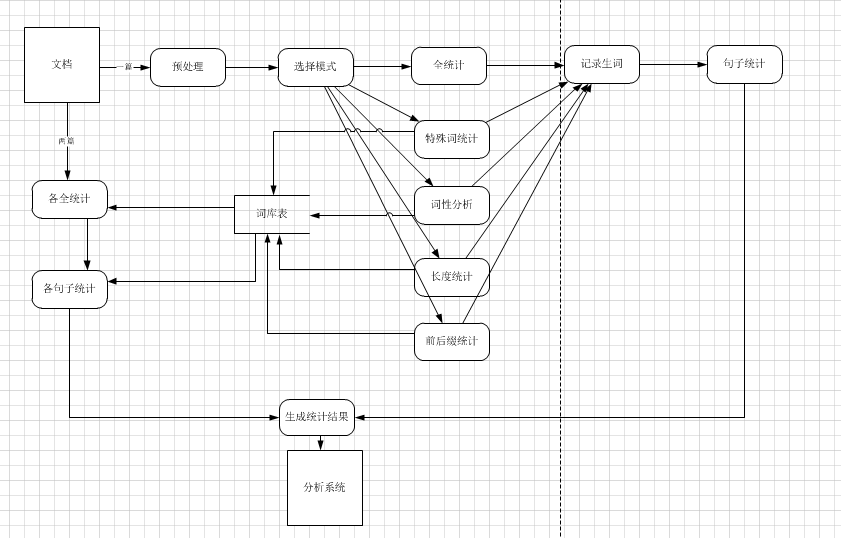
3.1.5管理信息系统



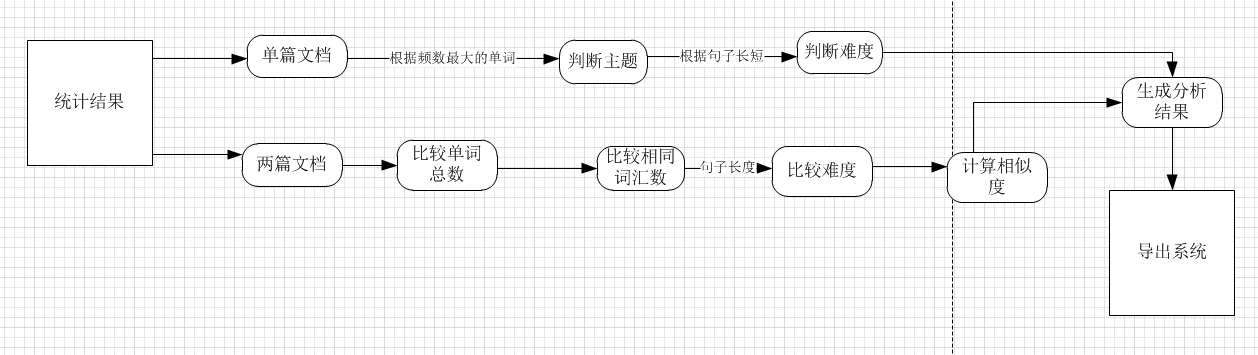
3.1.6导入系统



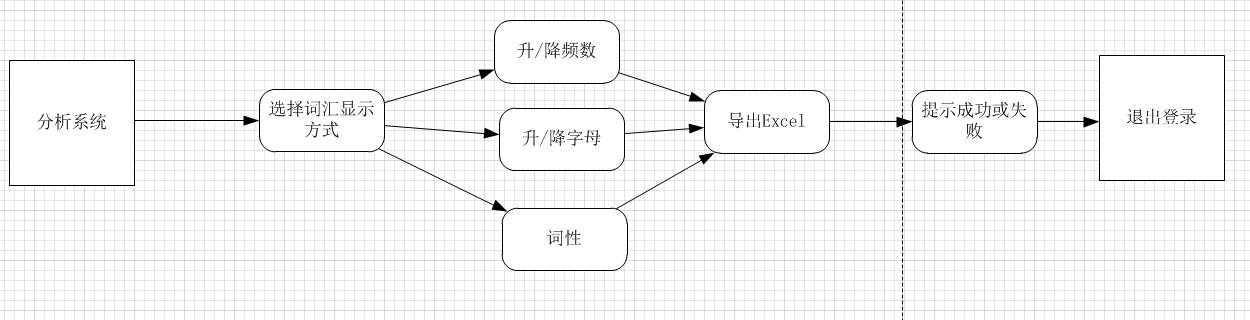
3.1.7统计系统



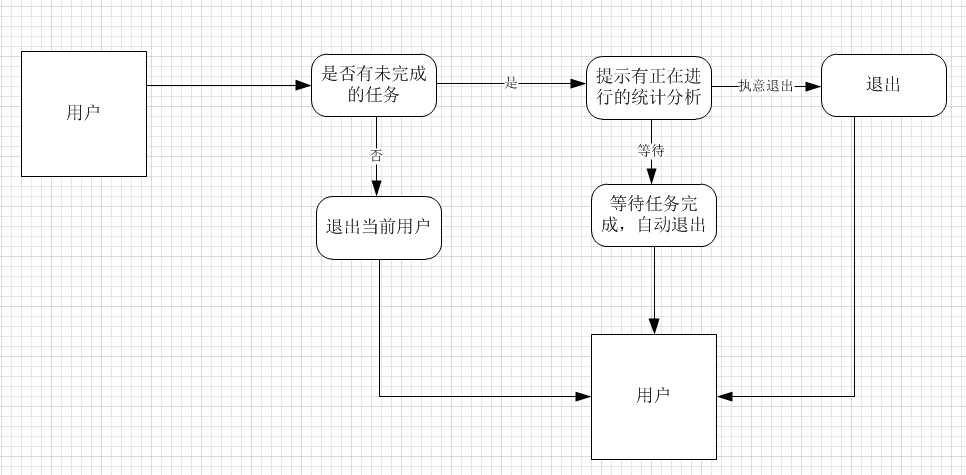
3.1.8分析系统



3.1.9导出系统

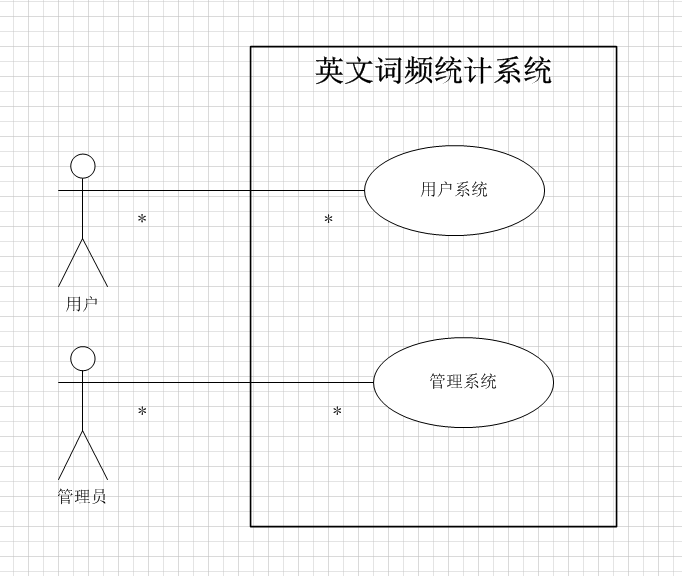


3.1.10退出登陆系统

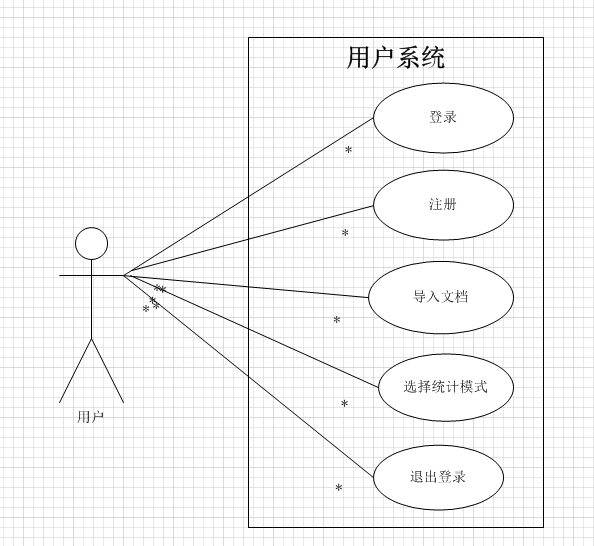


**3.2. 用例建模**

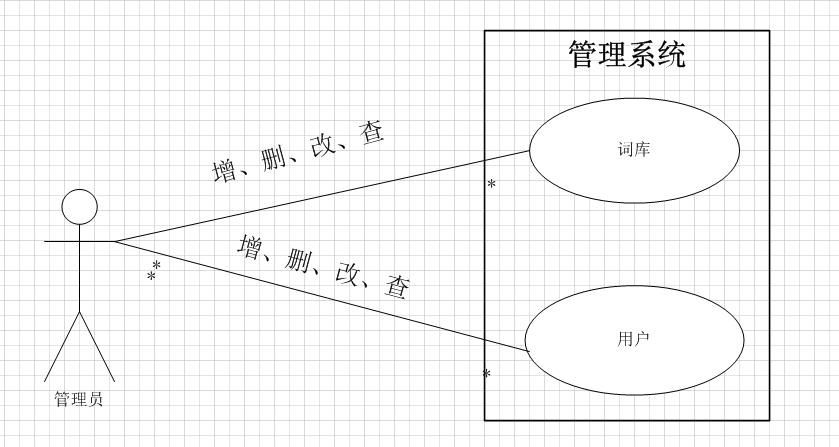
3.2.1英语词频统计总系统：管理员应用管理系统管理，用户应用用户系统。



3.2.2用户系统：用户在用户系统中可进行登录、注册、导入文档、选择统计模式、退出登录操作。

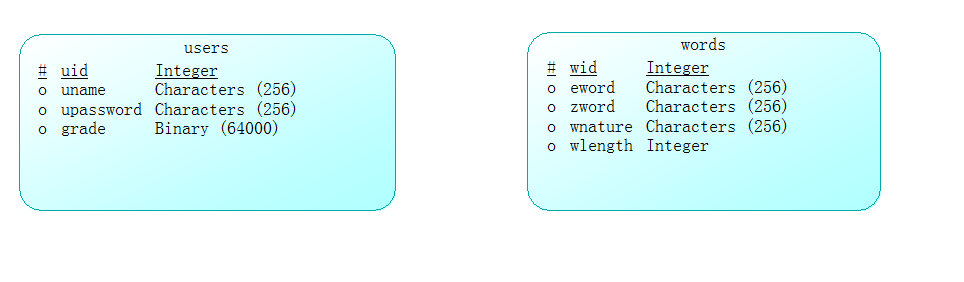


3.2.3管理系统：管理员在管理系统中可对用户信息、单词信息进行增加、删除、修改、查询

****

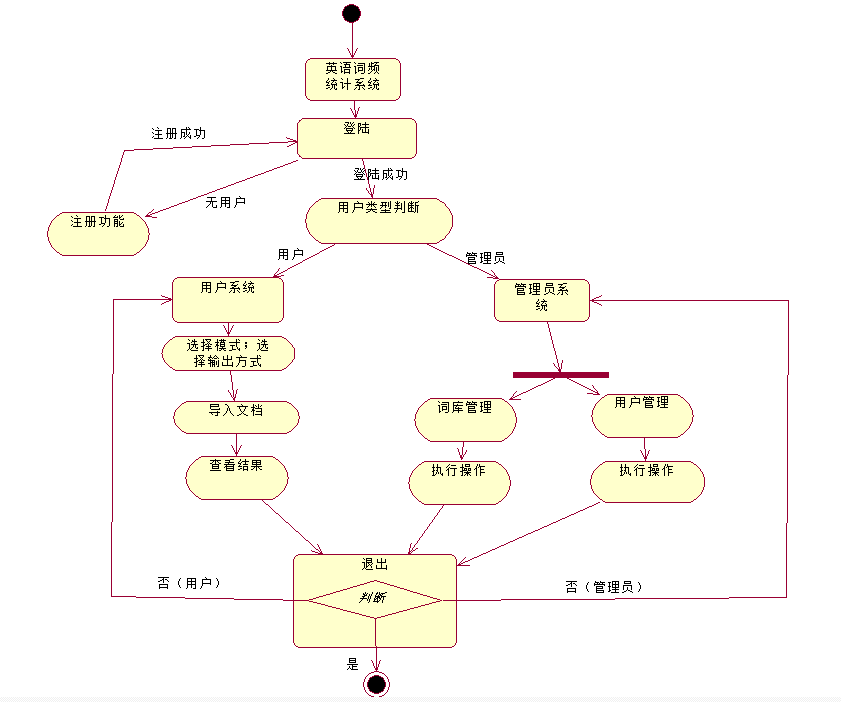
**3.3 系统类图**

本系统中共有两个类分别是用户类，和词库类



**3.4. 动态建模**

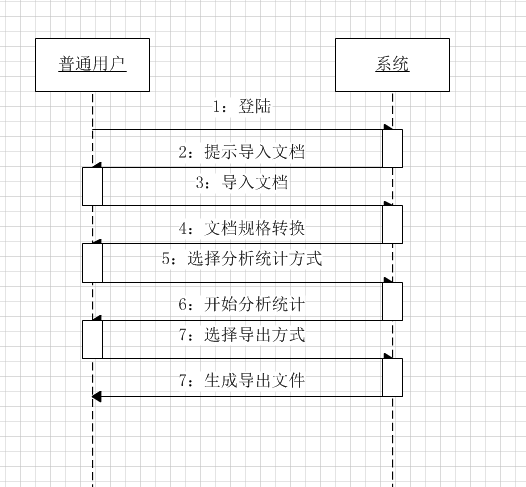
**3.4.1活动图**



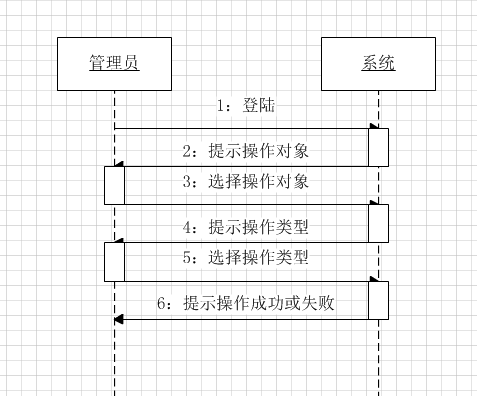
**3.4.2状态转移图**

**3.4.3顺序图模型**

a 普通用户与系统顺序图



b 管理员与系统顺序图



**3.5 数据库描述：**

**MySQL 是一个[关系型数据库](https://baike.so.com/doc/5294416.html" \t "_blank)，** MySQL 是一种关联数据库管理系统，关联数据库将数据保存在不同的表中，而不是将所有数据放在一个大仓库内，这样就增加了速度并提高了灵活性。

**3.6 数据字典：**

1. 用户信息表：users

|  |  |  |  |
| --- | --- | --- | --- |
| **字段名称** | **字段类型** | **说明** | **备注** |
| uid | int(50) | 主键 |  |
| uname | char(50) | 账号 |  |
| upassword | char(50) | 密码 |  |
| grade | int(10) | 级别 | 0代表普通用户  1代表管理员 |

1. 单词信息表：words

|  |  |  |  |
| --- | --- | --- | --- |
| **字段名称** | **字段类型** | **说明** | **备注** |
| wid | int(50) | 主键 |  |
| eword | char(100) | 英文单词 |  |
| zword | char(100) | 中文意思 |  |
| wnature | char(50) | 词性 |  |
| wlength | int(50) | 长度 |  |

**3.7. 数据采集：**

通过Internet从网站上收集一些词库的资料，四级、六级、专四、专八、托福、雅思、GRE等词库导入数据库词库表中。

**4.总结**

需求[分析](http://www.unjs.com/Special/dangxingfenxicailiao/)是[项目](http://www.unjs.com/Special/xiangmushishifangan/)开发的基础，基础打的牢不牢直接关系到后面所有的[工作](http://www.unjs.com/Special/gongzuosixianghuibao/)，是项目[实施](http://www.unjs.com/Special/xiangmushishifangan/)成败的关键，所以我们小组非常重视需求分析，我们以老师给的文档为依据，从用户的角度来分析，逐步展开了面向过程的需求分析，其中包括功能需求、性能需求、运行需求及其他需求。还有面向对象的需求分析，包括用例建模、系统类图以及动态建模。通过以上分析我们对英语词频统计系统有了更加全面具体的认识，从用户的角度出发在老师的功能要求基础上又增加了我们自己的想法，例如根据单词的数量、长度等统计出文章的难易程度。在这个项目的初步阶段我们四个人分工明确又互相讨论、帮助，最终出色的完成了这份需求报告，相信后续的项目进展会更加精彩。