

# 清华大学信息技术研究院语音和语言技术中心 (CSLT) 研究生管理条例\*

——2009~2010 学年度第 17 次中心主任会通过——

## 1、前言

为了增强研究生(含联合培养研究生和博士后)研究工作的针对性、系统性和深入程度,按照学校、院(系)对研究生培养管理工作的最新要求,经过 CSLT 主任会研究讨论确定,制定本条例。

理论研究、技术研发和工程开发都是科研活动的重要组成部分,本中心的所有研究工作都围绕上述活动展开。

## 2、管理模式

根据研究工作的特点,本中心采取目标管理与过程管理相结合的方式。管理遵循下面的原则:

- **任务驱动:** 以任务的完成情况作为研究生工作的主要衡量标准。要求在本中心(或联合机构)工作的时间,每周不少于 40 小时(导师确认的学习、调研、开发、活动等计入工作时间)。在工作时,不得利用网络以任何形式进行与分配的任务无关的事情。本中心实行作息打卡制度,以每周一到周五上午 7 点到晚上 10 点间为有效工作时间,周六、周日计算的有效工作时间不超过 8 小时。

- **纪律严明:** (1) 研究生需要请假时, Email 请假无效。1 天以内的请假,分别向所属实验室的负责老师请假。2~5 天的请假,先由所属实验室的负责老师准假,再由实验室主管副主任批准。5 天以上的请假,由校、院(系)批准。所有请假必须事先以书面形式进行,表格见附件 1。急病假可后补,但须出具急诊证明及病假条;紧急事假须先电话或短信告知,主管老师认可批假后再休假,然后后补假条。(2) 研究生拟申请参加到企业的实习时,必须遵循的原则:所在实习单位应与清华大学签订相关协议;实习活动应经中心主管副主任批准。如果研究生未经批准擅自参加实习活动,中心将视情节严重程度给予扣发补助、通报批评、上报学校开除学籍等不同形式处分。

- **团队合作:** 研究生既要有合作的意识,又要实实在在地互相合作,互相帮助。

- **资源共享:** 在不违反清华大学与本中心(或联合机构)关于科研成果的保密规定的前提下,本中心实行论文和研究成果等资源的共享,以避免不必要的重复劳动。

- **严格保密:** 属于本中心(或联合机构)的任何成果(包括其修改版和增强版),任何人不得以任何手段泄露、分发、销售、转让或从事其他任何可能给本中心带来损失的行为。

- **奖优罚劣:** 中心对表现优秀的研究生给予奖励,对表现差的研究生给予批评。每连续三次受到批评的研究生将被延长毕业时间半年一次。

## 3、指导方式

本中心实行导师直接指导和导师指定副导师(或辅导教师)指导的方式来指导所有研究生的研究工作。由导师(指定的副导师或辅导教师)根据本中心(或联合机构)需要,安排和指导研究生的相关科研工作或科研活动。

## 4、例会制度

根据中心《关于建立和完善 CLST 学术交流制度学术的决议》(以下简称《学术交流制度》)的

\*2010 年 6 月 8 日第十三次修订。2010 年 3 月 16 日第十二次修订。2008 年 9 月 16 日第十一次修订。2007 年 6 月 4 日第十次修订。2006 年 7 月 10 日第九次修订。2006 年 5 月 24 日第八次修订。2006 年 3 月 12 日第七次修订。2005 年 6 月 27 日第六次修订。2005 年 5 月 12 日第五次修订。2002 年 9 月 1 日第四次修订。2001 年 5 月 24 日第三次修订。2000 年 3 月 20 日第二次修订。1998 年 10 月 18 日第一次修订。1998 年 1 月 8 日生效。

要求，研究生按学术方向归属到相应的研讨小组（分别是语音识别 ASR，声纹识别 VPR，语言理解 NLU）。研究生应按《学术交流制度》规定，提前做好例会和研讨会的准备工作。

- (1) 例会报告内容应以学术报告、综述报告为主，编程技巧和基础知识介绍为辅。主要目的是通过做报告达到锻炼能力和加强交流的目的；
- (2) 对学术报告，时间严格限定在 15 分钟以内，以锻炼在有限时间内讲清报告内容的的能力；其他类型报告时间可以相对宽松，以讲清楚为目的；
- (3) 加强中心不同方向的学术交流，建议在不同方向的学术报告中给出必要的背景内容介绍，便于不同研究方向的同学都能从中受益；
- (4) 报告的组织方式宜采用比较灵活的轮排制，由学生班长征求导师意见后进行协调，每次例会报告数不少于 2 次。
- (5) 对每个例会报告，建议由导师进行质量把关，尽量安排有实质内容的报告。

为锻炼研究生的组织能力和服务意识，每次中心例会安排一名学生协助中心例会的主持进行会议准备，主要包括：提前借会议室并发通知；提前半小时准备好例会要用的投影仪和/或笔记本电脑；提前做好报告电子文档的保存工作；等等。

## 5、月度小结

研究生应根据中心《学术交流制度》的规定，每 4 周在中心例会上汇报一次月度科研工作。

此外，研究生需要在每月的最后一天 24:00 之前通过电子邮件交一份小结（详细要求与文档模板见附 2）给中心所有老师，篇幅不超过一页 A4 纸。

## 6、学术报告

研究生应在中心例会上进行一定次数的报告：第一学期至少一次中文报告，最后一学期至少一次英文报告，其余每学期至少一次中文报告和一次英文报告。研究生的报告数将在每学期末进行检查。

如果在相应学期中报告数目没有满足上面要求，则自动认为其整个研究计划后延一个学期。

关于组内报告的具体要求如下：

- 文献阅读报告只能用中文，科研开发成果报告原则上要求用英文。
- 英文报告中的有关术语要求注明中文含义，属于特别重要的语句可以全句注上中文。
- 报告要求重点突出，讲清思路和关键点；要使用投影仪。
- 根据需要，必要的、重要参考文献必须附在报告后面。
- 必须提前三天（不计周末和报告当日）将报告（要附相关的参考文献）发给指导小组。如有反馈意见，则应根据情况调整报告的重点。
- 原则上，中英文报告的时间为 15 分钟。（不计提问时间）

如果此次报告不符合要求（不计入完成报告总数），需要重做，但不算批评。

研究生在学术会议上发表论文，不管是否参加会议，要提前一个月在组内讲一次，计入统计。

## 7、论文发表

研究生的毕业年限与论文发表情况直接相关。

(1) 硕士研究生：至少发表（或录用）一篇 EI 检索论文的，可以申请两年毕业；至少发表（或录用）一篇中心核心期刊的，可以申请两年半毕业；至少要有一篇在本中心指定的期刊或会议上录用的文章，方能申请三年完成硕士学位论文毕业。

(2) 博士研究生：参照学校和系里的论文发表要求。

研究生应重点在以下本领域有影响的会议和期刊上发表学术论文。

会议：

- (1) ICASSP: EI 收录
- (2) Interspeech (ICSLP, EuroSpeech)

- (3) ISCSLP
- (4) IEEE workshop on Automatic Speech Recognition and Understanding (ASRU)
- (5) NCMMS
- (6) ACL
- (7) COLING
- (8) ACM SIGIR
- (9) ACM SIGKDD

其他会议，一会一议，需要导师（或指定的副导师或辅导教师）决定，否则不承担会议注册费

#### 刊物：

- (1) IEEE Transaction on Audio, Speech and Language Processing: SCI 收录
- (2) Speech Communication: SCI 收录
- (3) Computer, Speech and Language: SCI 收录
- (4) Computational Linguistics: SCI-E、SSCI 收录
- (5) ACM Transaction of Asian Language and Information Processing
- (6) 声学学报: EI 收录
- (7) 中文信息处理学报
- (8) 清华大学学报: EI 收录

其他刊物，一刊一议，需要导师（或副导师或指定的辅导教师）决定，否则不承担论文版面费。

为今后联系方便，所有发表论文的作者联系方式，一律要求有教师的联系方式。

## 8、生活补贴

本中心研究生在读期间，学习态度端正、科研工作认真的，可以每月领取相应的生活补贴。

## 9、设备管理

中心分配给各研究生的用于科研工作的相关设备，研究生应按中心关于设备管理的有关规定妥善使用和保管，不得随意拆解变更。因老化而不可用的设备必须按学校规定执行“报废”流程。研究生毕业前清点设备，如有损遗，按学校规定进行赔偿。

## 10、毕业工作交接

每个毕业研究生应在答辩前 2 个月，按附件 3 要求，完成相关源码、数据和文档等的整理，提交中心核查。经核查符合要求，才准予进入毕业答辩流程。

## 11、附则

对于不能按本条例要求进行工作学习的研究生，中心将按学校有关规定，劝其改正、延期毕业，直至上报学校予以劝退处理。

本管理条例若有与学校、院（系）等冲突的条款，以学校、院（系）的规定为准，其他条款仍然有效；但对于本中心提高要求的部分，以本条例要求为准。

本管理条例的解释权归中心主任会。

2010 年 3 月 16 日  
语音和语言技术中心 CSLT

附件 1：请假申请表

CSLT 学生请假申请表

申请日期： 年 月 日

请假人	组别	姓名
请假类别	<input type="checkbox"/> 事假 <input type="checkbox"/> 病假 <input type="checkbox"/> 其他（请说明）	
请假日期	从 月 日至 月 日共 天	
负责老师意见		签字：
主任意见		签字：
备注		

---

附件 2: 月度小结要求

### 月度小结要求

1. 按下页 WORD 模板填写(注意页眉), 每月最后一天 24:00 为最后期限, 不交或过期递交将按空白递交处理。
2. 小结不超过一页。对没有内容的项, 可以空白处理。
3. 小结仅涉及学术、科研、开发等方面的问题, 其他方面的问题由其他 E-Mail 单独反映。
4. 对于需要深入描述或作报告说明的部分, 无需展开, 只需要嵌入一个引用作简单说明。
5. 各项的详细说明:
  - 5.1 分配的任务和完成情况: 指导教师或研究小组组长制定并经过确认的研究或开发计划以及对计划的实施情况。自选任务原则上不在此条; 但可以由其他 e-Mail 汇报而作特殊处理, 或经导师确认后追认为分配的任务。如果你是在读文章, 或在思考问题, 那么必须在作完报告的那个月才能认为是完成任务。
  - 5.2 遇到的问题和解决方案: 指在完成分配的任务中遇到的问题以及提供解决方案。
  - 5.3 下月或下一步的工作计划: 如题。
  - 5.4 为技术中心其他同学所作的工作: 指“团队协作”的执行情况, 本项限于在分配的任务完成过程中的互相帮助。
  - 5.5 非学位论文的发表以及程序库的规范化和上交情况:
    - (1) 非学位论文指刊物、杂志、学术会议等方面的学术论文, 而不是毕业设计、硕士论文和博士论文。

考虑到文章有书写、递交、看清样、定稿等诸多过程, 本项仅涉及文章递交之前的书写工作(这是工作量最大的部分), 而且每篇文章只报告一次, 即在文章书写完成的时候在本项中报告一次。

由于写文章是对研究生的总体要求, 文章的递交部分不放在“分配的任务”处报告。
    - (2) 程序库的规范化和上交鼓励研究生及时把研究开发过程中可以抽象、提取、总结、归档的部分按本中心的要求作规范化处理, 然后上交, 由中心负责安排供大家共享。
  - 5.6 对本月个人表现的自我评价: 如题。
  - 5.7 对现有工作和任务的改进建议; 其它待说明情况: 本项不包括导师和组长制定的工作和任务改进计划, 而仅限于在本月的研究和开发中受启发而得到的额外改进意见和建议。详细的说明另以 E-Mail 或其它方式与导师或组长汇报沟通。
6. 文件名命名规则: yymm-MR-name.doc, 其中 yy 表示年, mm 表示月, name 表示姓名首字母缩写, 如熊振宇递交的 2001 年 6 月份的小结名称是 0106-MR-XZY.doc, 该文档作为邮件的附件, 邮件的标题为“【月度小结 熊振宇 0106】”; 如有姓名缩写相同者, 则姓要写全, 如 0506-MR-LiuJ、0506-MR-LiJ。

---

## 月度工作总结 (标准模板)

张三 2007年6月

### 1. 分配的任务和完成情况

- (1) 分配的任务一, 完成情况
- (2) 分配的任务二, 完成情况

### 2. 遇到的问题 and 解决方案;

- (1) 遇到的问题一, 解决方案
- (2) 遇到的问题二, 解决方案

### 3. 下月或下一步的工作计划

- (1) 下月或下一步的工作计划

### 4. 为其他相关人员所做的工作

- (1) 帮别人测试调试、录音标注、其它

### 5. 非学位论文的发表以及程序库的规范化和上交情况

- (1) 非学位论文的发表

### 6. 对本月个人表现的自我评价

### 7. 对现有工作和任务的改进建议; 其它待说明的情况

## 附件 3: 文档与源码递交要求

### 一、科研文档

文档递交, 要求至少含有以下部分(分目录存放, 下文中加粗带括号的文字为目录名称):

1. **[组内报告]** 组内报告 PPT (及相关的附件, 如 PPT 所附参考文献的电子版)。PPT 命名符合组里规定。
2. **[学术论文]** 参加学术会议相关材料, 如论文原文、Poster 原件、组内演练时所用的 PPT, 以及发表在学术刊物上的论文。在“学术论文”目录下, 不同学术会议和学术刊物建不同目录, 以名称惯用缩写命名, 如 ICASSP2007、SpeCom 等。
3. **[开题报告]** 开题报告相关文档, 如各种表格的电子版、WORD 格式的报告、PPT 格式的报告等。
4. **[学位论文]** 学位论文相关文档, 如论文最终上交学校的电子版、答辩决议和答辩过程记录(在答辩秘书处)等。
5. **[科研实验]** 科研实验(工程项目)相关文档。每个(种)实验一个子目录。在各目录中, 至少应有:
  - (1) **实验报告**(实验目的、实验数据、实验流程与操作步骤、实验结论), 报告是自我完备的, 不能用“结果参见 XX 论文”来进行引用; 报告也应正确记录实验过程, 即实验结果应该是可以重现的。
  - (2) **[实验工具]** 如 PYTHON 源码、EXE 程序等等, 其中 EXE 程序的源码, 按本规范下文中关于源码递交的要求, 另行保存, 但此处需要说明 EXE 程序是由哪个工程编译链接得到的(在 readme.txt 中说明此信息)。
  - (3) **[实验数据]** 这部分通常会很大, 要注意完整性。重要的可复用的中间结果数据也应包含, 例如从语音原始文件经过计算后得到的特征数据文件, 从特征文件训练得到的模型文件等。
6. **[系统工具]** 演示系统和实用工具的相关文档。每个系统/工具一个子目录。在各个子目录中, 至少应包含以下内容:
  - (1) **[技术文档]** 包括系统/工具的完整软件设计文档(概要设计和详细设计)、实用手册和测试报告, 具体格式可参阅所附的文档内容说明。
  - (2) **[程序源码]** 按照不同功能分类, 分别组织成不同的工程子目录。相关内容应与上面的设计文档和使用手册保持一致。
  - (3) **[测试数据]** 给出使各个系统/工具能准确运行的相关实验数据和标准结果, 便于对相关系统/工具的处理能力进行评测。

以上各级目录均需含一个 readme.txt 文件, 说明目录内容。

### 二、程序源码

递交程序源码时, 应将源码中的各个工程存放在相互独立的目录中。所有内容则存放在名为 Projects.NAME 的目录中。以下内容, 要求与源码一起, 分层次存放。递交时要逐款进行检查。

#### 1、工程(任务)说明---**第一级注释**

一个单独的文本文件, 名字为 ProjectsList.txt。其内容是对研究期间所有的涉及编程的任务, 分别简要地说明相应工程的以下信息:

- <A> 背景与意义
- <B> 起止的时间
- <C> 主要的结论

<D> 应用的环境，得出结论的前提，其它注意事项

<E> 工程对应的目录

## 2、 每个工程的具体说明---[第二级注释](#)

一个单独的文本文件，名字为 `Readme.txt`。实际上 VC 会生成此文件，只要修改其内容即可，它必须包含如下内容：

<A> 工程的用途（或背景和意义）

<B> 工程的文件清单

<C> 工程的配置要求：如要包含什么库，要在工程中定义什么常量，特殊的包含路径，特殊的条件编译选项

## 3、 涉及算法的重要文件内的具体说明---[第三级注释](#)

这些说明要求写在 `CPP` 文件的头部，内容包括：

- 文件名
- 作者（原始作者，修改者）
- 时间（各次稳定版本的形成时间）
- 版本、版权信息
- 关于本文件功能或作用的总体说明
- 涉及主要算法的重要函数的清单，及相应简要描述
- 其它注意事项说明（如条件编译设置等）
- 对包含的重要头文件的作用的说明

## 4、 涉及算法的重要函数的功能说明---[第四级注释](#)

这些说明要求写在函数的头部，要求在“函数申明”（`H` 文件）和“函数定义”（`CPP` 文件）中都要进行一致的说明，内容包括：

- 函数的接口说明（指参数含义，输入要求，应用条件）
- 完成的功能（指计算什么，处理什么，输入什么等等）
- 函数应具有防止崩溃情况出现的能力，否则应改写。

## 5、 涉及算法的重要函数内部流程的说明---[第五级注释](#)

建议在写程序时写。写完代码后，应进行全面的深入思考，对关键环节作一个清晰的说明。说明分布在函数内部的一些流程关键点处。主要包括（并不限于）：

- 算法出处，参考文献
- 条件编译选项的意义与使用说明
- 内部流程的阶段说明
- 重要变量的意义说明

## 6、 其它注意事项

- 为调试或单元功能测试而加入的代码，要进行明确说明，以方便维护者理解这些代码，并且明白这些代码在必要时是可以删掉的。作者应保证这些代码的删除不会影响程序功能的正确性。
- 程序功能的正确实现不应与调试代码或测试代码相关（即删除这些代码后程序功能仍然正确，运行仍然正常）。
- 源文件中不能有被注释掉的代码，这些代码应删除。
- 如果需要表明某段有相当长度的代码的整体性，建议在代码块前用一种符号，在代码段后用另一种与之相对应的符号。如：

(I)

```
// xxxx -->  
代码  
// xxxx <==
```

(II)

```
// -----VVVVV-----  
代码  
// =====AAAAA=====
```