

# 如何撰写毕业论文

宗成庆

模式识别国家重点实验室

---

**E-mail:** cqzong@nlpr.ia.ac.cn      **Home Page:** <http://www.nlpr.ia.ac.cn/cip/english/zong.htm>

**Add.:** No.95, Zhong Guan Cun Dong Lu, Beijing 100190, China

# 有言在先

---

**本报告中所列举的反面例子是我归纳编造的，并无具体所指，切勿对号入座，若有雷同，纯属偶然。**

# 提 纲

---



1. 关于论文题目
2. 论文的整体布局
3. 常见的主要问题
4. 认真对待每一个细节
5. 结束语

# 1. 关于论文题目

## ◆ 基本要领

- ✧ 措辞简单明了，通俗易懂；
- ✧ 含意明确，不易太大，但能概括全文；
- ✧ 尽量有特色，不易太长。

## 举例：

统计机器翻译新方法研究

 太大了！

面向××××的若干问题研究

 太泛了！

融合源语言句法知识的统计翻译模型研究

 棒！

# 提 纲

---

1. 关于论文题目



2. 论文的整体布局

3. 常见的主要问题

4. 认真对待每一个细节

5. 结束语

## 2. 论文的整体布局

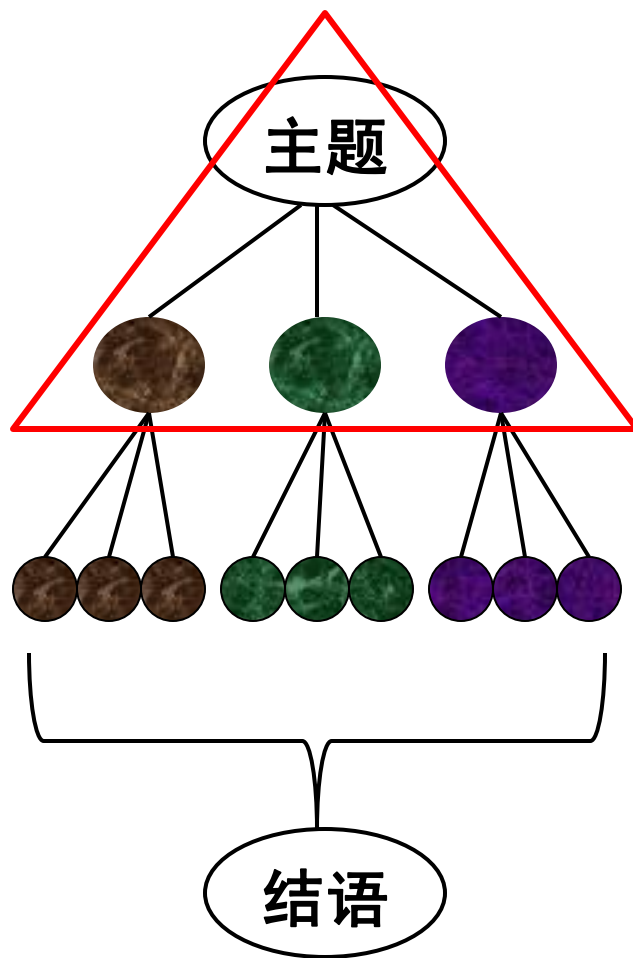
### ◆ 基本要领

- ✧ 开门见山，尽早破题；
- ✧ 全文各部分是有机联系的整体，构成一个完整的“故事(story)”，切忌将已发表的几篇论文简单地拼凑、堆砌在一起，各章内容互不相干；
- ✧ 主题鲜明，重点突出，所有章节的内容均围绕主题展开。

# 2. 论文的整体布局

## ◆ 阐述方式

问题提出  
↓  
问题分解  
↓  
技术创新  
(理论方法、  
实现技术、实  
验对比、分析)  
↓  
归纳总结



起(起因)

承(过程)

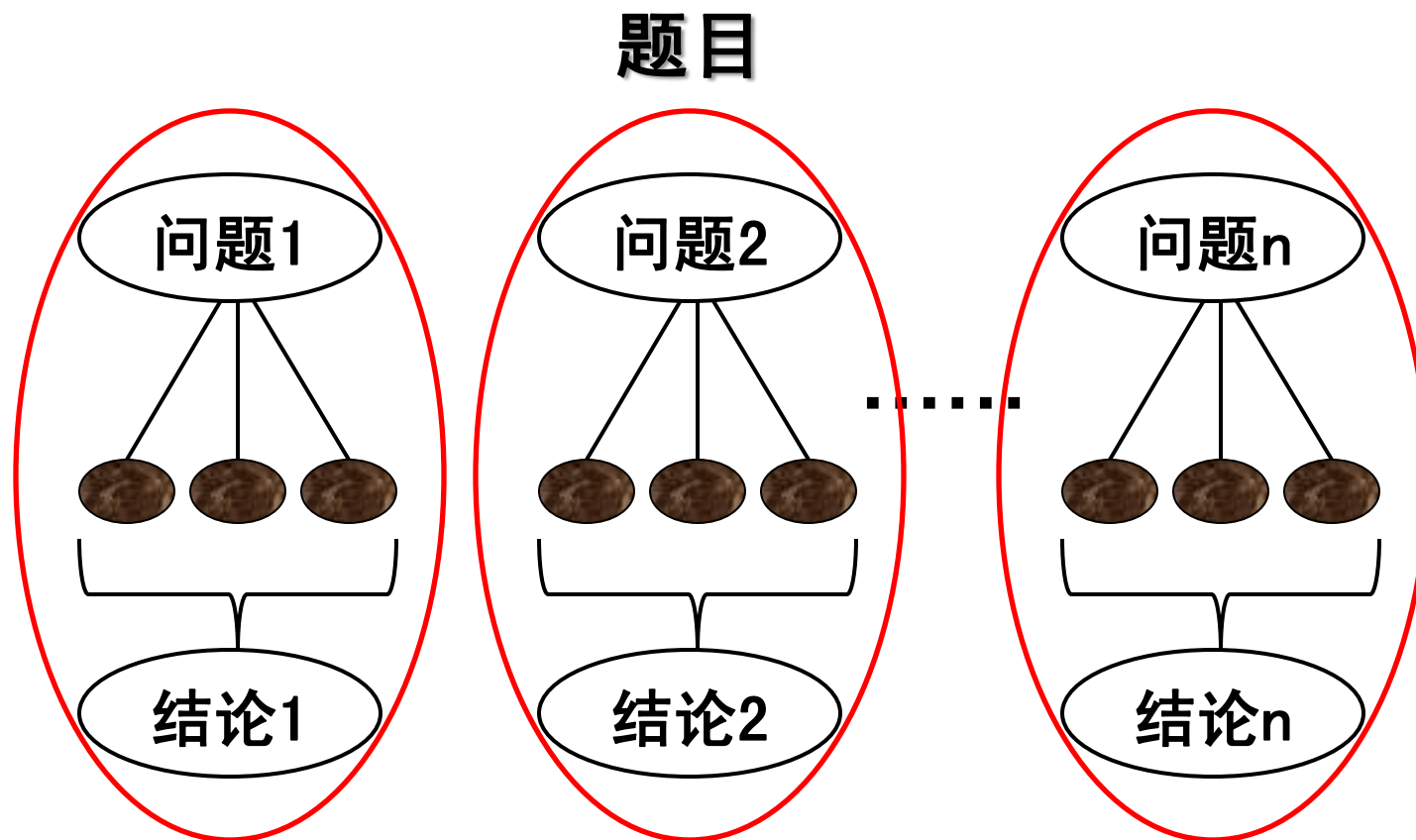
转(转折)

合(结论)



自顶向下(Top-Down)瀑布式的展开方式

## 2. 论文的整体布局



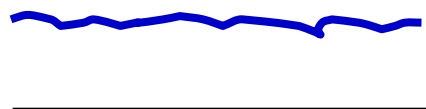
**不要出现千岛湖式的布局**



# 2. 论文的整体布局

## ◆ 内容分配

平淡无味型

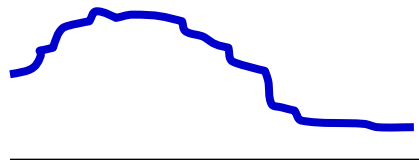


各章内容平均用墨，重点不突出，索然乏味。



(1)

虚胖臃肿型

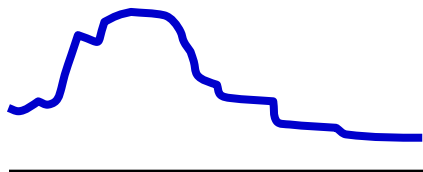


介绍前人的工作太多，阐述自己工作的部分过于简略。



(2)

龙头蛇尾型

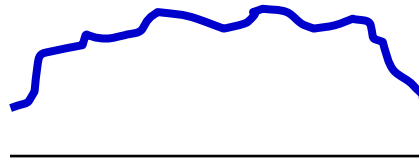


问题提得很大、很悬，自己的工作却没有什  
么实质性内容。



(3)

虎头豹尾型



对前人的工作分析深入，问题犀利，观点(方法)  
新颖，结语画龙点睛。



(4)

## 2. 论文的整体布局

### ◆ 常见的提纲

#### 第1章 绪论：简明扼要

破题：基本概念和主要研究内容 (做什么?)

意义：理论意义和应用价值 (为什么?)

结构：各章节的安排

以博士学位论文为例，一般应在 3~4页(A4 纸，正文用小4号字，宋体，1.25倍行距，下同)，大约 3000 ~ 5000字。

## 2. 论文的整体布局

### 第2章 国内外相关研究综述：深入浅出

综述：针对拟研究的问题已有哪些工作和方法？

分析：已有方法的优缺点对比，存在哪些缺陷？

归纳：方法分类，问题归纳（预示本文的工作）。

动因：阐述自己的思路，进行问题分解。

**切忌：** 1) 罗列已有的工作，面面俱到，没有自己的观点；  
2) 追溯历史太久远，涵盖范围太宽泛；  
3) 照抄别人的文献。

一般地，应在 15~20页，约 1.8万 ~ 2.4 万字。

**注：** 也可将本章与第一章合并。

## 2. 论文的整体布局

第3~5章 阐述自己的创新工作：理论与实验并重

- 阐述自己最重要的创新；
- 注意几个创新点之间的关联；
- 有充分的实验对比和结果分析；
- 建议给自己的方法起个响亮的名字。



重心

**注意：**1) 博士和学术型硕士学位论文是对科学问题研究的创新成果阐述，不是实验报告或项目报告；  
2) 创新点不易太多，一般为3个。

一般地，应在 50~60页，约 6 ~ 7 万字。

## 2. 论文的整体布局

---

### 第6章 结束语：画龙点睛

- 归纳全文的创新点；
- 指出下一步需要进一步开展的工作。

**切忌：1) 重复前面的摘要；  
2) 对下一步工作的展望太虚、太大。**

一般地，应在 3页左右，约3000字。

# 提 纲

---

1. 关于论文题目

2. 论文的整体布局

 3. 常见的主要问题

4. 认真对待每一个细节

5. 结束语

# 3. 常见的主要问题

- ① 题目太大、太泛，没有特色；
- ② 主题不聚焦，研究内容缺乏整体性，各章内容相互独立；
- ③ 综述前人的工作时只是简单地罗列，缺乏深入的分析 and 归纳总结；
- ④ 大篇幅地引用他人的工作，引用的图、表、数据等不标明来源出处；
- ⑤ 贡献不明确，创新点不突出；
- ⑥ 实验对比的方法或系统不是本领域公认的，或者实验对比不充分，缺乏深入分析。



**危险!**

# 提 纲

---

1. 关于论文题目
2. 论文的整体布局
3. 常见的主要问题

 4. 认真对待每一个细节

5. 结束语



## 4. 认真对待每一个细节

① 文字表达应通顺，含意明确，避免使用华丽的辞藻。

✧ 句子不顺畅，含意模糊不清，例如：

这些比较的都是应用在检索方面，是否在分类任务中，其他的模型是否能发挥相应的优势还没...

✧ 整段所表达的意思不明确。

✧ 出现半截句，如：

“本论文提出如何利用条件随机场解决……问题”

→ 本文提出了如何利用条件随机场解决……问题的方法。

✧ 口语化严重，过多地使用“我们”。

✧ 用词过分华丽，尤其在致谢中。

## 4. 认真对待每一个细节

### ② 避免出现错别字，差错率应不超过1/10000。

- ✧ 常被忽视的错误：“的”与“地”不分
- ✧ 不可容忍的错误：人名错误，split → spit, from → form, window → widow
- ✧ 不可忽视英文摘要。

### ③ 尽量避免使用直接截图、粘贴的编辑方式，不要汉英混用。

- ✧ 直接从已发表的论文中截取图表，粘贴到自己的论文中
- ✧ 图表中使用英文，且无任何汉语注释。

## 4. 认真对待每一个细节

### ④ 字体、字号应统一，格式应规范。

- ✧ 建议每章的标题用小三号黑体，标题下空一行；每节的标题用四号黑体，标题下与正文之间空0.5行；正文汉字用小4号宋体，英文用Times New Roman体，1.25倍行距；算法描述用“楷体”等。
- ✧ 段首空两个汉字的位置。
- ✧ 公式中的变量字母大/小写、正/斜体要一致(公式内一致和与正文内一致)。

### ⑤ 图、表、公式的编号应统一，与正文中的引用一致。

- ✧ 建议采用每章独立计数方式，如“图2-1”、“表3-4”，公式:(3-1)等。

## 4. 认真对待每一个细节

### ⑥ 术语使用要规范、统一。

✧ 使用正确的术语，英文缩写首次出现时应给出全拼，除了个别用语，英文词首字母应小写：如：

...条件随机场(conditional random field, CRF)...

...隐马尔可夫模型(hidden Markov model, HMM)...

✧ 切勿杜撰、硬造，而无严谨、明确的定义。

### ⑦ 可以给自己的方法起个响亮、好记的名字，适当地拔高，但切勿口出狂言。

✧ 名字的英文缩写不宜太长，一般不超过5个字母。

✧ 尽量不要使用“提出了...理论体系”、“解决了...问题”，“首次提出了...”等类似的表述。

## 4. 认真对待每一个细节

⑧ 参考文献信息应详细，格式统一，正文中的引用应在参考文献中找到唯一的对应，切勿遗漏。尽量引用本领域内国际上公认的有代表性的重要之作，尤其是近几年的。一般而言，博士论文的参考文献不应少于100篇。

◇ 参考文献列表建议使用如下格式：

- 英文的文献在前面，中文的文献在后面，按作者的发音顺序排列；
- 作者相同时，按文献发表的时间先后顺序排列；对于英文文献，唯一作者的文献排在前面，两位作者的文献排在其次，三位或三位以上作者的文献在后；对于中文文献，唯一作者的文献在前，两位和两位以上作者的文献随后；

## 4. 认真对待每一个细节

- 对于英文文献，可用如下格式：

[作者姓，年份]作者姓□作者名. 年份. 文献题目. 详细出处(会议名称、地点等或期刊卷、期号；页码)。

年份相同时，在年份后面加a、b、c等英文字母区分。

对于两位作者的文献，“作者姓”采用如下形式：

“第1作者的姓 and 第2作者的姓”；对于三位或三位以上作者的文献，采用如下形式：“第1作者的姓 et al.”。

请见下面的例子：

## 4. 认真对待每一个细节

[Zhang and Zong, 2013]Zhang, Jiajun and Chengqing Zong. 2013. Learning a Phrase-based Translation Model from Monolingual Data with Application to Domain Adaptation. *Proc. ACL*, Sofia, Bulgaria, August 4-9, 2013. Pages 1425-1436

[Zhang et al., 2013a]Zhang, Jiajun, Feifei Zhai and Chengqing Zong. 2013a. Syntax-Based Translation with Bilingually Lexicalized Synchronous Tree Substitution Grammars. *IEEE TASLP*, 21(8): 1586-1597

[Zhang et al., 2013b]Zhang, Jiajun, Feifei Zhai and Chengqing Zong. 2013b. A Substitution-Translation-Restoration Framework for Handling Unknown Words in Statistical Machine Translation. *Journal of Computer Science and Technology*, 28(5):907-918

可以写成: **Zhang, J., F. Zhai and C. Zong**

# 4. 认真对待每一个细节

说明：按照国家新闻出版署的规定，英文人名应写成如下形式：

[Brown et al., 1993] Brown P F, Della Pietra S A, Della Pietra V J, and Mercer R L. 1993. The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computational Linguistics*. 19(2):263-309.

**原名：Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, and Robert L. Mercer.**

[宗成庆，2013]宗成庆. 2013. 统计机器翻译. 北京：清华大学出版社

[宗成庆等，2008]宗成庆，高庆狮. 2008. 中国语言技术进展. 中国计算机学会通讯，4(8): 39-48

[宗成庆等，2009]宗成庆，曹右琦，俞士汶. 2009. 中文信息处理60年. 语言文字应用，(4): 53-61



# 提 纲

---

1. 关于论文题目
2. 论文的整体布局
3. 常见的主要问题
4. 认真对待每一个细节



**5. 结束语**

# 5. 结束语

- ◆ 评价一篇论文优劣最重要的标准是看其解决的问题是否重要、创新点是否突出、实验是否充分可靠
- ◆ 论文的结构是否清晰、逻辑是否清楚、表述是否流畅、准确，具有非常重要的意义
- ◆ 我们提倡规范、严谨、朴实的写作风格，“标新立异”应体现在创新思想上，而不是在写作形式上
- ◆ 应遵循不同学科领域的论著写作规范
- ◆ 上述观点和建议只代表我本人的体会，仅供参考，不强求千篇一律。任何学生都应首先尊重自己导师的意见。

---

*Thanks!*

谢谢!