

2024学年开学组会

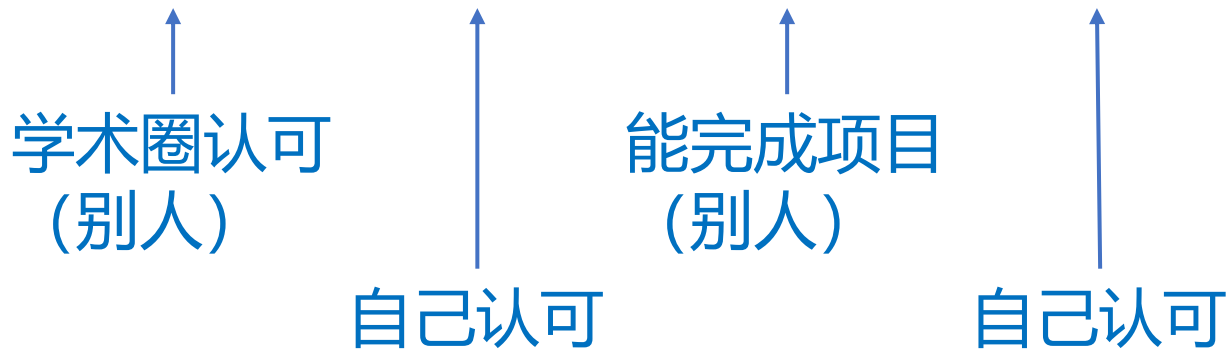
南方科技大学，计算机科学与工程系
计算语言学与意识科学实验室

Computational Linguistics and Consciousness Science Lab
2024年9月6日

我们是谁？

- 计算语言学与意识科学实验室
- Computational Linguistics and Consciousness Science Lab

- 我们要做什么？
- 做**有意义、有意思、有价值、有品质**的研究



为什么进组？

- 研究生，博士生，本科生
- 工作，毕业，论文，升学，思考人生

每天干什么？（研究生、博士生）

- 第一年：
 - 上课、做助教、读论文、做实验，思考人生.....
- 第二年：
 - 做实验、开题、写论文，思考人生.....
- 第三年：
 - 做实验、写论文、答辩，思考人生.....

如何过好研究生的这几年？

- 多阅读
- 多思考
- 多提问（有效提问）

- 谦虚学习
- 相信自己

- 从小事做起（比如说第一个project、第一篇paper）
- 不一定做好，但是要做完

读论文 (paper)

- 开题答辩需达到的阅读量，支撑开题报告的核心文献数量
- 本科生：8
- 硕士生：16
- 博士生：32

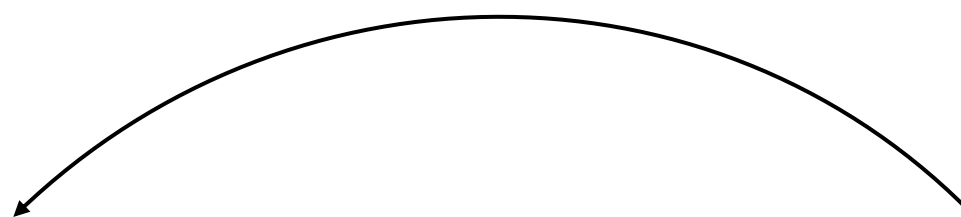
写论文 (paper)

- 易：新建文件夹很容易
- 难：指数增长，细节爆炸

- 初期：问题+结果 驱动
- 后期：DDL 驱动

原则上：
硕士答辩前需开启第一轮
博士答辩前需至少完成一轮

- 第一轮完整闭环很重要：初稿 -> 终稿 -> 审稿人意见 -> 再稿



发论文

- 已ACL系列会议为主：ACL、EMNLP、NAACL、COLING
- 亦可尝试其它AI相关会议：ICLR、NeurIPS、ICML、AAAI
- 期刊：可作为会议论文的增长版本
- Transactions of ACL、Computational Linguistics、Cognition

计算资源管理

- 仔细阅读微信群里发布的守则！

服务器 Huggingface 使用手册

如何设置共享 huggingface 缓存

原贴：<https://benjijang.com/posts/2024/01/shared-hf-cache/>

原贴代码：<https://github.com/bstee615/shared-hf-cache>（上文作者提供）

RTX A6000 IP地址：10.16.21.70

L40 IP地址：10.16.117.170

链接A6000 服务器上的HF共享缓存路径为 /data1/huggingface

∴ L40 服务器上的HF共享缓存路径为 /data1/huggingface

(以上缓存路径已经设置完成，不需要再重复设置！)

组会要求

- 研究生要求：不迟到，不早退，非特殊情况不要请假
- 每人每学期至少作一次正式报告，准备PPT，和听众讨论
- 本科生要求：尽量参加组会

学长扶帮带

- 硕士生：
 - 第二学年，每个学期带一组创新实践同学，带一名本科毕设同学
- 博士生：
 - 第二、三学年，每个学期带一组创新实践同学，带一名本科毕设同学