

# 北京理工大学研究生课程教学日历

课程名称 语言信息处理

课程代码 0700002 课程性质 考查

主讲教师 黄河燕 史树敏 2017-2018 学年第 1 学期

辅导教师 \_\_\_\_\_ 计算机 学院

授课对象 \_\_\_\_\_

时数 教学/实际	全总 学时 期数	学时分配				每周 学时 数
		讲 授	实 验	习 题	考 核	
教学计划	48	40	3	2	3	3
实际上课						

周次	上课方式	时数	授 课 内 容	课外阅读和书面的作业		学习检查		参考书名和章节
				时数	内 容	检查方式	所需时间	
1	课堂讲授	3	课程介绍及语言信息处理引论	2	课外阅读：《数学之美》等入门读物			数学之美
2	课堂讲授	3	语言学基础知识及知识表示					语言学教程
3	课堂讲授	1	语言学基础知识及知识表示	1	课外阅读：语言学相关文献，特别是词、句、篇和语言分析相关内容			语言学教程
3	课堂讲授，课堂练习及讨论	2	文法基本理论	0.5	回顾形式语法相关内容	课堂练习及讨论：上下文无关文法推导	0.5 课时	自然语言理解，自然语言处理的形式模型
4	课堂讲授，课堂练习及讨论	3	文法基本理论 专题：谓词逻辑与基于逻辑的 NLP	0.5	回顾谓词逻辑相关内容	课堂练习及讨论：一阶谓词逻辑表达语义	0.5 课时	自然语言理解，自然语言处理的形式模型
				1	书面作业：哪些形式可以用来表示语义？解释其如何表达语义？	作业打分		

5	课堂讲授, 课堂练习及讨论	3	语言分析技术: 预备知识	1	课外阅读: 统计学习模型相关内容	课堂练习及讨论: 贝叶斯理论及隐马模型在NLP中的应用	1 课时	统计学习方法
6	课题讲授	3	语言分析技术: 词法分析技术	1.5	书面作业: 怎样将汉语分词任务转化为序列标注问题?	作业打分		统计学习方法, 统计自然语言处理基础, 统计自然语言处理
7	课堂讲授, 课堂练习及讨论	3	语言分析技术: 句法分析技术			课题练习及讨论: 怎样为CFG推导引入统计信息	0.5 课时	统计自然语言处理, 自然语言处理综论
8	课堂讲授	3	语言分析技术: 语义分析技术	1	书面作业: 给定原始文本语料, 如何做无监督的汉语分词?	作业打分		统计学习方法, 统计自然语言处理基础, 统计自然语言处理
9	课堂讲授, 课堂讨论	3	专题: 图模型在NLP中的应用 专题: NLP中的机器学习	1	课外阅读: 图模型相关知识	课堂讨论	1 课时	Pattern Recognition and Machine Learning

10	课堂讲授	3	篇章分析与会话处理技术					语言学教程，统计自然语言处理
11	课堂讲授	2	语言生成技术	1	课外阅读：语言生成技术相关文献			自然语言处理综论
11	课堂讲授	1	机器翻译	1	课外阅读：机器翻译技术相关文献			统计机器翻译
12	课堂讲授及课堂讨论	3	机器翻译专题：NLP和机器翻译中的深度学习	2	课外阅读：深度学习相关文献	课堂讨论：深度学习的优劣势及迁移方法	1课时	统计机器翻译，深度学习
13	课题讲授	3	信息检索与问答系统	1	课外阅读：信息检索与问答系统相关文献			信息检索
14	课堂讲授	3	信息抽取与自动摘要	1	课外阅读：信息抽取相关文献	课堂讨论：多文档摘要的难点和解决方法	0.5课时	统计自然语言处理
15	专家讲座	3	知识图谱及其应用技术					
16	主题讨论	3	运用自然语言处理技术进行类人答题	16	期末项目及报告	课堂演示及书面报告	3	

## 一、 教学目的

通过本课程的学习，使研究生：

- 1、了解语言信息处理的主要研究内容与发展历程；
- 2、了解文法基础理论，掌握典型的自然语言形式模型；
- 3、掌握语言信息处理，特别是词法、句法、语义、语用等多层次语言分析的基本原理与主要技术方法；
- 4、了解主要的自然语言生成技术；
- 5、掌握语言信息处理，特别是机器翻译、信息检索、信息抽取等工程应用的核心问题与主要技术方法；
- 6、了解语言信息处理技术的前沿及发展趋势。

## 二、 授课方法和方式

课堂讲授，材料自学与课堂讨论，专题讲座相结合。

## 三、 成绩评定方式

成绩以百分制衡量。

成绩评定依据：开放式问题讨论成绩占 40%，平时作业成绩占 30%，期末项目及报告成绩占 30%。

## 四、 教材和必读参考资料

- 1、 Allen, J.F. Natural Language Understanding. The Benjamin/Cummings Publishing Company, Inc., Redwood City, CA. 1995. (自然语言理解 (第二版), James Allen 著, 刘群等译, 电子工业出版社. 2005)
- 2、 Jurafsky, D. and J. H. Martin. Speech and Language Processing-An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall. 2000. (自然语言处理综论 (第二版), Daniel Jurafsky, James H. Martin 著, 冯志伟、孙乐译, 电子工业出版社. 2005)
- 3、 Christopher D. Manning, Hinrich Schütze. Foundations of Statistical Natural Language Processing. Cambridge, MA: MIT Press. 1999. (苑春法等译. 统计自然语言处理基础. 北京: 电子工业出版社. 2005.)
- 4、 Philipp Koehn. Statistical Machine Translation. New York, NY: Cambridge University Press. 2009. 宗成庆, 张霄军译. 统计机器翻译. 北京: 电子工业出版社. 2012.
- 5、 冯志伟. 自然语言处理的形式模型. 安徽: 中国科学技术大学出版社. 2010.
- 6、 宗成庆. 统计自然语言处理(第二版). 北京: 清华大学出版社. 2013.
- 7、 李航. 统计学习方法. 北京: 清华大学出版社. 2012.
- 8、 胡壮麟. 语言学教程(第四版). 北京: 北京大学出版社. 2013.
- 9、 吴军. 数学之美. 北京: 人民邮电出版社. 2012.

任课教师: 黄河燕、史树敏 \_\_\_\_\_年\_\_月\_\_日

教学院长 \_\_\_\_\_年\_\_月\_\_日

注:

1. 此教学日历由授课教师填写, 教学院长签字后执行, 学院留存一份。
2. 任课教师应将教学日历提供给上课的研究生, 课程完成后填写实际上课的学时数。