编号：



毕业设计(论文)任务书

课 题： 英语文本语篇关系提取算法设计

|  |  |  |
| --- | --- | --- |
| 学 院： | 计算机与信息安全学院 | |
| 专 业： | 软件工程 | |
| 学生姓名： | 薛天杭 | |
| 学 号： | 1400330129 | |
| 指导教师单位： | 计算机与信息安全学院 | |
| 姓 名： | | 黄桂敏 |
| 职 称： | | 教授 |

**题目类型： 软件开发**

2019年11月25日

注：1、本任务书一式两份，一份院或系留存，一份发给学生，任务完成后附在

说明书内。

2、任务书均要求打印，打印字体和字号按照《本科生毕业设计（论文）统

一格式的规定》执行。

**一、毕业设计（论文）的内容**

1、研究英语文本中的语篇关系原理；

2、研究英语文本语篇关系的提取方法；

3、选取最优英语文本语篇关系的提取方法，设计英语文本语篇关系的提取算法；

4、完成算法软件编程任务；

5、实验结果满足预期要求，实现英语文本的语篇关系提取。

**二、毕业设计（论文）的要求与数据**

1、设计英语文本语篇关系的提取算法；

2、以英语文本数据集为实验数据；

3、实验结果以满足要求，实现英语文本的语篇关系提取；

4、完成相关测试，完成毕业设计的总体要求。

**三、毕业设计（论文）应完成的工作**

1、外文翻译材料一份（包括不少于4万字符的英文原文和译文）

2、毕业设计开题报告一份。

3、毕业论文一份（2万汉字以上，附中英文摘要，其中英文摘要300-500单词）。

4、本系统软件及源程序清单一套。

**四、应收集的资料及主要参考文献**

收集资料：英语文本数据。

主要参考文献：

[1]练睿婷. 语篇标注语料库的建设研究[C]. 中国中文信息学会.第四届全国学生计算语言学研讨会会议论文集.中国中文信息学会:中国中文信息学会,2018:284-290.

[2]孙爱珍.语境共现词汇链的自动提取及与语篇衔接之关系(英文)[J].Chinese Journal of Applied Linguistics,2016,34(04):3-14+127.

[3]Li S, Wang L, Cao Z, et al. Text-level Discourse Dependency Parsing[C]. meeting of the association for computational linguistics, 2017: 25-35.

[4]Tu M, Zhou Y, Zong C, et al. A Novel Translation Framework Based on Rhetorical Structure Theory[C]. meeting of the association for computational linguistics, 2017: 370-374.

[5] Zhang L, Jia Y, Li A, et al. An interface research on rhetorical structure and prosody features in Chinese reading texts[C]. international symposium on chinese spoken language processing, 2016: 1-5.

[6] Zhang J J, Fung P. Automatic Parliamentary Meeting Minute Generation Using Rhetorical Structure Modeling[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2017, 20(9): 2492-2504.

[7] Li K, Zhang J, Yao C, et al. Automatic Relation Extraction from Text: A Survey[C]. the internet of things, 2016: 83-86.

[8] Chengcheng L. Automatic Text Summarization based on Rhetorical Structure Theory[C]. international conference on computer application and system modeling, 2018.

[9] Zhang S. Entity relation extraction to free text[C]. international conference natural language processing, 2017: 1-5.

[10] De Uzeda V R, Pardo T A, Nunes M D, et al. Evaluation of Automatic Text Summarization Methods Based on Rhetorical Structure Theory[C]. intelligent systems design and applications, 2018: 389-394.

[11] Marir F, Haouam K. Rhetorical structure theory for content-based indexing and retrieval of Web documents[C]. international conference on information technology research and education, 2016: 160-164.

[12] Jose M. Chenlo, Alexander Hogenboom, David E. Losada. Rhetorical Structure Theory for polarity estimation: An experimental study[J]. Data & Knowledge Engineering, 2015, 94(PB):135-147.

**五、试验、测试、试制加工所需主要仪器设备**

计算机1台

**任务下达时间：**

2019年11月25日

**毕业设计开始与完成时间：**

2019年11月28日至2020年4月26日

**组织实施单位：** 计算机与信息安全学院软件工程教研室

**教研室主任意见：**

无

签字  2019年11月27日

**学院领导小组意见：**

同意

签字  2019年12月28日