

陈宏申

✉ chenhongshen@ict.ac.cn · ☎ (+86) 188-0115-5908

📁 研究方向

深度学习, 自然语言处理, 机器翻译

🎓 教育背景

中国科学院计算技术研究所, 北京

2012 -- 至今

在读博士研究生 计算机科学与技术, 预计 2017 年 6 月毕业

导师: 刘群

四川大学, 成都

2008 -- 2012

学士 计算机科学与技术

👨‍💻 专业经历

新加坡科技设计大学

2015 年 12 月 -- 2016 年 6 月

研究助理 (全职) 合作导师: Yue Zhang

工作: 用于多标注数据的神经网络模型 (发表在 NLP 顶会之一 EMNLP 上)

- 使用 soft-attention 机制, 从字向量中学习词向量
- 使用 Bi-LSTM CRF 基线模型
- 提出了基于神经网络的 stacking 模型
- 提出了基于神经网络的多视角联合模型
- 联合神经网络 stacking 与多视角模型, 将 CTB5 的词性标注准确率从 94.10 提升到 95.53
- 开源地址: <https://github.com/chenhongshen/NNHetSeq>

中国科学院计算技术研究所

2012 年 9 月 -- 至今

研究助理

工作: 基于依存边转换的统计机器翻译模型 (发表在 NLP 顶会之一国际计算语言学会议 Coling 上)

- 提出了一种基于分析 - 转换 - 生成的句法统计机器翻译模型
- 提出了依存边转换翻译规则抽取方法
- 提出了基于图分析 - 柱搜索的句子生成翻译解码算法

工作: 依存边转换翻译规则生成器 (2016 年 10 月获全国计算语言学会议最佳论文奖)

- 使用神经网络压缩器和编码器结构生成基于句法的翻译规则
- 前馈神经网络将翻译规则的源端压缩成源端信息综合向量
- 两个 RNN 与 MLP 分别生成翻译规则目标端的词并判别位置关系

💻 IT 技能

- Python, C/C++, Theano, Linux

📄 发表论文

- Hongshen Chen, Yue Zhang, Qun Liu, Neural Network for Heterogeneous Annotations, To Appear In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP) 2016
- Hongshen Chen, Qun Liu, A Dependency Edge Transfer Translation Rule Generator, In Proceedings of The Fifteenth China National Conference on Computational Linguistics (CCL), 优选并推荐至《中国科学 F 辑: 信息科学》发表 (300 余篇中仅 3 篇) 2016

- Hongshen Chen, Jun Xie, Fandong Meng, Wenbin Jiang, and Qun Liu, A Dependency Edge-based Transfer Model for Statistical Machine Translation, In Proceedings of *Conference on Computational Linguistics (COLING)*, Dublin, Ireland 2014

💡 专利

- 基于依存树的统计机器翻译方法及系统 (No.: 201410389422.2)

📊 评测经历

China Workshop on Machine Translation Evaluation (CWMT) 2013 年

- 核心成员，英汉新闻系列排名第一，英汉科技系列排名第二
- 负责多系统翻译融合

NIST Open Machine Translation Evaluation (NIST) 2012 年

- 负责层次短语模型，阿拉伯语到英语的翻译

🏆 获奖情况

- 2016，第十五届全国计算语言学会议最佳论文奖
- 2016，中国科学院大学，三好学生
- 2014，自然语言处理组 2013 年度勇猛精进奖
- 2013，自然语言处理组 2012 年度最佳新人奖
- 2012，四川大学，本科优秀毕业设计三等奖
- 2010，国家励志奖学金
- 2010，四川大学与美国劳伦斯技术大学，视觉机器人比赛 “Robofest-2010” 一等奖
- 2010，四川大学奖学金，综合二等奖
- 2009，四川大学奖学金，综合三等奖
- 2009，四川大学，计算机学院，学习进步奖

📌 其他工程项目

- 横向项目 -统计机器翻译系统（核心开发人员）
- 横向项目 -命名实体识别系统（核心开发人员）