付振新 (Zhenxin Fu)

Peking University

- School of Electronics Engineering and Computer Science (EECS)
- Institute of Computer Science and Technology (ICST)

Phone: +86 15650708568
Email: fuzhenxin95@gmail.com
Github: https://www.github.com/fuzhenxin
Homepage: https://zhenxinfu.com

Personal

I am a second year graduate student at EECS and ICST, majoring in Computer Science and Natural Language Processing. I develop research interests in style transfer, dialog system and multi-task learning. I am also a member of the student cluster competition team of Peking University, knowing commonsense on popular developing tools.

Education

- **B.S. Computer Science, EECS, Peking University, 2014-2018** : Major in computer science, developing interest in network, database and parallel computing.

- **M.S. Candidate. Natural Language Processing, ICST, Peking University, 2018-Present** : Major in Natural Language Processing, developing interest in style transfer, open-domain conversation.

Competition & Internship

- **Student Cluster Competition at SC16, SC17, ASC19, & SC19** : I am in charge of reproducing results of a given paper on our own cluster. We are selected to contribute to a special issue of the journal, Parallel Computing, as one of the top teams from reproducibility challenge. In ASC19 and SC19, I worked as a coach.

- **Tencent Beijing 2017.4–2017.8** : Test Engineer (Intern) for virtual assistants & name entity recognition task using sequence labeling.

- **Alibaba Group 2018.8–Present** : Research Intern for Natural Languge Processing in AliMe Team & DAMO Academy.

Publications

- **Joint Learning of Word Representation and Sense Representation for Unsupervised Word Sense Disambiguation**, AAAI Student Abstract and Poster Program, 2020.
  *Jie Wang, **Zhenxin Fu**, Moxin Li, Haisong Zhang, Dongyan Zhao, Rui Yan.*
  Unsupervised WSD by learning the sense representation through the supervision from word embedding.

- **Query-bag Matching with Mutual Coverage for Information-seeking Conversations in E-commerce**, CIKM (Short), 2019.
  ***Zhenxin Fu**, Feng Ji, Wenpeng Hu, Wei Zhou, Dongyan Zhao, Haiqing Chen, Rui Yan.*

A bag containing lots of questions corresponds to a same answer. Ranking the bag directly helps the information-seeking conversation.

- **Semi-supervised Text Style Transfer: Cross Projection in Latent Space**, EMNLP-IJCNLP, 2019.
  *Mingyue Shang, Piji Li, **Zhenxin Fu**, Lidong Bing, Dongyan Zhao, Shuming Shi, Rui Yan.*
  Semi-supervsed text style transfer with cross projection in latent space: cross training autoencoder, style transfer and latent space projection.

- **Multilingual Dialogue Generation with Shared-Private Memory**, NLPCC (English), 2019. **Outstanding Paper Award**.
  *Chen Chen, Lisong Qiu, **Zhenxin Fu**, Junfei Li, Rui Yan.*
  Multilingual dailog generation with shared-private memory inspired from multi-task learning.

- **Automated ICD Coding Based on Word Embedding with Entry Embedding and Attention Mechanism (**基于融合条目词嵌入和注意力机制的自动 **ICD** 编码**)**, NLPCC (Chinese), 2019.
  *Hongke Zhang, **Zhenxin Fu**, Qianping Ren, Hui Xu, Dongyan Zhao, Rui Yan.*
  A classification model for International Classification of Disease (ICD) with item embedding and label attention.

- **Find a Reasonable Ending for Stories: Does Logic Relation Help the Story Cloze Test?** AAAI Student Abstract and Poster Program, 2019.
  *Mingyue Shang, **Zhenxin Fu**, Hongzhi Yin, Bo Tang, Dongyan Zhao, Rui Yan.*
  Natural language inference is used to promote story cloze test.

- **Student Cluster Competition 2017, Team Peking University: Reproducing Vectorization of the Tersoff Multi-Body Potential on the Intel Broadwell Architecture**, Parallel Computing (Journel), 2018.
  ***Zhenxin Fu**, Lei Yang, Wenbin Hou, Zhuohan Li, Yifan Wu, Yihua Cheng, Xiaolin Wang, Yun Liang.*
  The report of reproducibility challenge for the student cluster competition at SC17.

- **Learning to Converse with Noisy Data: Generation with Calibration**, IJCAI-ECAI, 2018.
  *Mingyue Shang, **Zhenxin Fu**, Nanyun Peng, Yansong Feng, Dongyan Zhao, Rui Yan.*
  Noise training cases harm trainnig process of generation-based dialog system. So we integrate instance weight into dialog system to overcome the problem.

- **One "Ruler" for All Languages: Multi-Lingual Dialogue Evaluation with Adversarial Multi-Task Learning**, IJCAI-ECAI, 2018.
  *Xiaowei Tong, **Zhenxin Fu**, Mingyue Shang, Dongyan Zhao, Rui Yan.*
  We apply multi-task learning in open-domain dialog evaluation.

- **Style Transfer in Text: Exploration and Evaluation**, AAAI, 2018.
  ***Zhenxin Fu**, Xiaoye Tan, Nanyun Peng, Dongyan Zhao, Rui Yan.*
  We propose unsupervised sequence to sequence model to transfer text from style A to style B. For example, converting paper title to news title.

- **ParConnect Reproducibility Report**, Parallel Computing (Journal), 2017.
  *Lei Yang, Yilong Li, **Zhenxin Fu**, Zhuohan Li, Wenbin Hou, Haoze Wu, Xiaolin Wang, Yun Liang.*
  The report of reproducibility challenge for the student cluster competition at SC16.

Others

- **Reviewer**: AAAI-2019; ACL-2019; EMNLP-2019; INLG-2019; AAAI-2020.

- **Teaching assistant**: Introduction to Natural Language Processing: Foundation, Theory and Application, Spring 2019.

- **Scholarship**: Wang Xuan Scholarship, first prize for the undergraduate internship at ICST.

<div align="center">

Last updated: December 3, 2019

`https://zhenxinfu.com/zhenxinfu.pdf`

</div>

付振新

北京大学

- 信息科学技术学院
- 计算机科学与技术研究所

手机: +86 15650708568
电子邮件: fuzhenxin95@gmail.com
Github: `http://www.github.com/fuzhenxin`
Homepage: `https://zhenxinfu.com`

简介

北京大学计算机科学技术研究所硕士二年级在读，研究方向为自然语言处理。感兴趣的研究领域包括文本风格迁移，开放域对话系统和多任务学习。同时，作为北大超算队队员，参加过 SC16、SC17 和 ASC19 大学生超算比赛。

教育经历

- 学士学位 北京大学 信息科学技术学院 2014-2018：专业为计算机科学与技术，对网络，数据库，并行计算等方向比较有兴趣.

- 硕士在读 北京大学 计算机科学技术研究所 2018-今：专业为自然语言处理，对文本风格迁移，开放域对话系统比较有兴趣.

比赛和实习

- 大学生超算比赛 **(SC16, SC17, ASC19, SC19)**：在前两次比赛中负责论文复现部分（在自有集群上复现一篇 SC 会议论文）。由于该项目上较高的成绩，我们受邀在期刊 Parallel Computing 上展示我们的复现报告。以教练的身份参加 ASC19 和 SC19。

- 腾讯北京 **2017.4–2017.8**：智能助手测试，命名实体识别.

- 阿里巴巴 **2018.8--**今: 自然语言处理方向研究型实习生，阿里小蜜 & 达摩院团队.

论文

- **Joint Learning of Word Representation and Sense Representation for Unsupervised Word Sense Disambiguation**, AAAI Student Abstract and Poster Program, 2020.
  *Jie Wang, **Zhenxin Fu**, Moxin Li, Haisong Zhang, Dongyan Zhao, Rui Yan.*
  通过词表示监督中间词义表示来进行无监督的词义消歧：词表示由词义表示组成。

- **Query-bag Matching with Mutual Coverage for Information-seeking Conversations in E-commerce**, CIKM (Short), 2019.
  ***Zhenxin Fu**, Feng Ji, Wenpeng Hu, Wei Zhou, Dongyan Zhao, Haiqing Chen, Rui Yan.*
  Bag 有一些列问题组成，这些问题对应相同的回复。通过用户问题和 Bag 的匹配增强对话系统。

- **Semi-supervised Text Style Transfer: Cross Projection in Latent Space**, EMNLP-IJCNLP, 2019.
  *Mingyue Shang, Piji Li, **Zhenxin Fu**, Lidong Bing, Dongyan Zhao, Shuming Shi, Rui Yan.*
  半监督的文本风格迁移，交叉训练风格迁移，自编码器，隐变量映射。

- **Multilingual Dialogue Generation with Shared-Private Memory**, NLPCC (English), 2019. **Outstanding Paper Award**.
  *Chen Chen, Lisong Qiu, **Zhenxin Fu**, Junfei Li, Rui Yan.*
  基于共享-私有 Memory 的多语言对话模型，受多任务学习启发。

- **Automated ICD Coding Based on Word Embedding with Entry Embedding and Attention Mechanism (**基于融合条目词嵌入和注意力机制的自动 **ICD** 编码**)**, NLPCC (Chinese), 2019.
  *Hongke Zhang, **Zhenxin Fu**, Qianping Ren, Hui Xu, Dongyan Zhao, Rui Yan.*
  病例自动化编码，利用条目表示和类别注意力的分类模型。

- **Find a Reasonable Ending for Stories: Does Logic Relation Help the Story Cloze Test?** AAAI Student Abstract and Poster Program, 2019.
  *Mingyue Shang, **Zhenxin Fu**, Hongzhi Yin, Bo Tang, Dongyan Zhao, Rui Yan.*
  利用自然语言推理辅助故事结尾预测任务。

- **Student Cluster Competition 2017, Team Peking University: Reproducing Vectorization of the Tersoff Multi-Body Potential on the Intel Broadwell Architecture**, Parallel Computing (Journel), 2018.
  ***Zhenxin Fu**, Lei Yang, Wenbin Hou, Zhuohan Li, Yifan Wu, Yihua Cheng, Xiaolin Wang, Yun Liang.*
  为 SC17 大学生超算比赛的论文复现报告。

- **Learning to Converse with Noisy Data: Generation with Calibration**, IJCAI-ECAI, 2018.
  *Mingyue Shang, **Zhenxin Fu**, Nanyun Peng, Yansong Feng, Dongyan Zhao, Rui Yan.*
  对话数据含有大量的噪声，数据质量对模型效果有较大影响。本文提出将 Instance Weight 应用到对话生成，使不同质量的数据样本对模型更新有不同影响。

- **One "Ruler" for All Languages: Multi-Lingual Dialogue Evaluation with Adversarial Multi-Task Learning**, IJCAI-ECAI, 2018.
  *Xiaowei Tong, **Zhenxin Fu**, Mingyue Shang, Dongyan Zhao, Rui Yan.*
  将 Multi-task 思想引入到对话评测当中，不同语言之间互相促进。

- **Style Transfer in Text: Exploration and Evaluation**, AAAI, 2018.
  ***Zhenxin Fu**, Xiaoye Tan, Nanyun Peng, Dongyan Zhao, Rui Yan.*
  本文提出一种无监督的序列到序列模型，将一句话从一种风格转换为另一种风格。并且提出了两个评价指标作为对该研究方向的促进。

- **ParConnect Reproducibility Report**, Parallel Computing (Journal), 2017.
  *Lei Yang, Yilong Li, **Zhenxin Fu**, Zhuohan Li, Wenbin Hou, Haoze Wu, Xiaolin Wang, Yun Liang.*
  为 SC16 大学生超算比赛的论文复现报告。

其他

- 审稿人: AAAI-2019; ACL-2019; EMNLP-2019; INLG-2019; AAAI-2020.

- 助教: 自然语言处理导论: 基础，理论与应用，2019 年春.

- 奖学金: 北大计算所王选奖学金（一等）; 北京大学专项奖学金.