

季紫薇

政治面貌：中共党员 ◇ 邮箱：zjiad@connect.ust.hk ◇ 个人网站：ziweiji.github.io ◇ 电话：15927277932

教育经历

香港科技大学	中国香港
博士候选人于电子与计算机工程系 (自然语言处理, 导师: Pascale Fung)	2019.9 - 至今
华中科技大学	中国武汉
学士于电子科学与技术专业 (排名: 1/156, GPA: 3.97/4.0)	2015.9 - 2019.6

部分奖项

国家奖学金 (3 次, 0.2%)	2016, 2017, 2018
全国大学生数学竞赛湖北省二等奖	2017
华中科技大学校三好 (3 次, 3%)	2016, 2017, 2018
海外交流先进个人 (UC Berkeley)	2017
本科特优生 (1%)	2016

项目经历

减少开放领域对话中的幻觉	2022.8 - 2023.1
--------------	-----------------

- 对话系统可以利用经过预训练语言模型和知识来产生流畅和信息丰富的回答。然而, 这些模型仍然容易产生输入源不支持的幻觉反应。我们选取 OpenDialKG 数据集研究, 其包括 15k 个基于知识图谱的开放式对话。
- 为了处理外部知识和对话之间的异质性, 并减少幻觉, 我们提出了 RHO: 1) 结合文本嵌入和相应的知识图谱嵌入的本地知识注入。2) 通过多跳推理能力的注意机制实现全局性知识注入。3) 基于在知识图谱上路径的应答重新排序技术, 以实现更好的对话推理。
- 实验结果表明, 我们的方法在自动和人工评估上都大大超过了 SOTA (FeQA 提升 17.54%)。

人工智能电影	2021.2 - 2022.2
--------	-----------------

- 为了提供一个定制的电影制作工具并为专业的电影人提供灵感, 我们与中央美术学院合作提出了一个自动、实时、可控的电影生成系统。
- 我们采用分层结构, 首先生成情节梗概, 然后生成剧本及其视觉呈现: 1) 设计了一个片型可控、情节引导的电影剧本生成系统。2) 从社交媒体收集视频数据库, 并根据剧本检索视频剪辑。3) 开发用户交互界面, 我们的系统允许用户选择片型并输入起始词, 以控制生成电影的主题和发展。
- 实验结果表明, 我们的框架在自动和人工评估方面都优于基线, 特别是在片型控制方面。
- 在平遥国际电影节和“徐冰的语言”艺术展上展出, 并发表于 AACL Demo。

探究视觉-语言预训练中的物体幻觉	2022.9 - 2023.1
------------------	-----------------

- 大规模的视觉-语言预训练 (VLP) 模型在基于视觉信息生成文本时, 容易产生物体幻觉, 即输出不存在的视觉对象。
- 我们系统地研究了物体幻觉的问题: 1) 研究了最近最先进的 VLP 模型, 表明它们仍然经常产生幻觉, 而在标准指标 (如 CIDEr) 上取得较好分数的模型可能更频繁地出现物体幻觉。2) 研究了 VLP 中不同类型的图像编码是如何影响幻觉的, 包括基于区域、基于网格和基于补丁的。我们发现基于补丁的特征表现最好, 较小的补丁分辨率可以明显减少物体幻觉。3) 将各种 VLP 目标函数解耦, 并证明 token-level 的图像-文本对齐和可控生成对减少幻觉至关重要。在此基础上, 我们提出了一个简单有效的预训练损失 ObjMLM, 以进一步减轻物体幻觉。
- 结果表明, 在两个基准测试中 (COCO Caption 用于域内评估, NoCaps 用于域外评估), 提出的损失函数可以减少高达 17.4% 的物体幻觉。
- 发表于 EACL。

- 建立了具有新的物理机制的模拟模型，与实际的实验结果吻合并给出了理论解释：1) 研究电离平衡对弱电解质在纳米通道中的电动力学流动的影响。2) 研究膜表面电荷对纳米孔发电装置中离子传输的影响。
- 本科期间在 Nanotechnology 和 Applied Physics Letters 上发表两篇论文。

代表论文

RHO(ρ): Reducing Hallucination in Open-domain Dialogues with Knowledge Grounding	under review
Ziwei Ji , Zihan Liu, Nayeon Lee, Tiezheng Yu, Pascale Fung, et al.	
Survey of Hallucination in Natural Language Generation	ACM Computing Surveys 2022
Ziwei Ji , Nayeon Lee, Rita Frieske, Pascale Fung, et al. Get 105 citations	
VScript: Controllable Script Generation with Visual Presentation	AAACL Demo 2022
Ziwei Ji , Yan Xu, I-Tsun Cheng, Pascale Fung, et al.	
Plausible May Not Be Faithful: Probing Object Hallucination in Vision-Language Pre-training	EACL 2023
Wenliang Dai, Zihan Liu, Ziwei Ji , Dan Su, Pascale Fung	
Model Generalization on COVID-19 Fake News Detection	AAAI CONSTRAINT 2021
Yejin Bang ¹ , Etsuko Ishii ¹ , Samuel Cahyawijaya ¹ , Ziwei Ji ¹ , Pascale Fung	
Multi-hop Question Generation with Graph Convolutional Network	EMNLP 2020 Findings
Dan Su, Yan Xu, Wenliang Dai, Ziwei Ji , Tiezheng Yu, Pascale Fung	

技能及其他

子任务	对话生成、故事生成、图像描述、问题生成、命名实体识别、假新闻检测等文本与图像领域
学术服务	ACL、EMNLP 审稿
编程语言	Python, C, Java, JavaScript, MATLAB
技能	Pytorch, TensorFlow, Linux, Git, SVN, DeepSpeed
语言	中文（母语）、英语（雅思 7 分, CET-6）