高思林　Silin Gao

个人主页 (Homepage) | Google Scholar | GitHub | 领英 (LinkedIn) | 电子邮箱 (Email)

教育经历　Education

洛桑联邦理工, 计算机与通信科学学院 瑞士, 洛桑
计算机与通信科学博士 2021 年 9 月至今
École Polytechnique Fédérale de Lausanne (EPFL), School of Computer & Communication Sciences Lausanne, Switzerland
Doctor of Philosophy (PhD) in Computer and Communication Sciences Sep 2021 - Present

清华大学, 计算机科学与技术系 中国, 北京
人工智能研究助理 2020 年 9 月至 2021 年 6 月
Tsinghua University, Department of Computer Science and Technology Beijing, China
Research Assistant (RA) in Artificial Intelligence Sep 2020 - Jun 2021

清华大学, 电子工程系 中国, 北京
电子信息科学与技术工学学士 2016 年 8 月至 2020 年 6 月
Tsinghua University, Department of Electronic Engineering Beijing, China
Bachelors of Engineering (BE) in Electronic Information Science and Technology Aug 2016 - Jun 2020

- **绩点 (GPA): 3.88/4.0　　排名 (Rankings): 8/262**

部分发表或在投论文　Selected Publications or Under-Review Papers

Augmenting LLMs' Reasoning by Reinforcing Abstract Thinking
Silin Gao, Antoine Bosselut, Samy Bengio, Emmanuel Abbe
Preprint. Under review.

Robust Reasoning with Contextualized Visual Representation Learning
Wenkai Chen*, **Silin Gao***, Albert Gatt, Antoine Bosselut
Preprint. Under review.

VinaBench: Benchmark for Faithful and Consistent Visual Narratives [Paper]
Silin Gao, Sheryl Mathew, Li Mi, Sepideh Mamooler, Mengjie Zhao, Hiromi Wakaki, Yuki Mitsufuji, [Website]
Syrielle Montariol, Antoine Bosselut [Code], [Data]
Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2025

Efficient Tool Use with Chain-of-Abstraction Reasoning [Paper]
Silin Gao, Jane Dwivedi-Yu, Ping Yu, Xiaoqing Ellen Tan, Ramakanth Pasunuru, Olga Golovneva, [Poster]
Koustuv Sinha, Asli Celikyilmaz, Antoine Bosselut, Tianlu Wang
Proceedings of the 31st International Conference on Computational Linguistics (COLING 2025)

DiffuCOMET: Contextual Commonsense Knowledge Diffusion [Paper]
Silin Gao, Mete Ismayilzada, Mengjie Zhao, Hiromi Wakaki, Yuki Mitsufuji, Antoine Bosselut [Code]
Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024)

PeaCoK: Persona Commonsense Knowledge for Consistent and Engaging Narratives [Paper]
Silin Gao, Beatriz Borges, Soyoung Oh, Deniz Bayazit, Saya Kanno, Hiromi Wakaki, Yuki Mitsufuji, Antoine Bosselut [Code], [Data]
Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL 2023)
Outstanding Paper Award

ComFact: A Benchmark for Linking Contextual Commonsense Knowledge [Paper]
Silin Gao, Jena D. Hwang, Saya Kanno, Hiromi Wakaki, Yuki Mitsufuji, Antoine Bosselut [Code], [Data]
Findings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)

End-to-End Task-Oriented Dialog Modeling with Semi-Structured Knowledge Management [Paper]
Silin Gao, Ryuichi Takanobu, Antoine Bosselut, Minlie Huang [Code]
IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)

HyKnow: End-to-End Task-Oriented Dialog Modeling with Hybrid Knowledge Management [Paper]
Silin Gao, Ryuichi Takanobu, Wei Peng, Qun Liu, Minlie Huang [Code]
Findings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL-IJCNLP 2021)

Paraphrase Augmented Task-Oriented Dialog Generation [Paper]
Silin Gao, Yichi Zhang, Zhijian Ou, Zhou Yu [Code]
Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)

荣誉和获奖情况 Honors and Awards

ACL 2023 杰出论文奖 (Outstanding Paper Award) ACL 2023 最佳论文委员会 (Best Paper Committee)
授予 ACL 2023 会议提交论文总数的 1% 2023 年 7 月
Granted to 1% of the total submissions in ACL 2023 Jul 2023

2020 年度毕业设计杰出论文 (Outstanding Thesis in Diploma Projects) 清华大学 (Tsinghua University)
电子工程系图像与信号研究所前 2 名毕业论文 2020 年 7 月
Top 2 thesis in Image and Signal Research Institution at Electronic Engineering Department Jul 2020

2017, 2018 和 2019 年度学术优秀奖 (Academic Excellence Awards) 清华大学 (Tsinghua University)
2016 至 2017, 2017 至 2018 和 2018 至 2019 学年绩点前 5% 2017 年 11 月，2018 年 11 月和 2019 年 11 月
Top 5% on GPA in the 2016-2017, 2017-2018 and 2018-2019 academic year Nov 2017, Nov 2018 and Nov 2019

2016 年度新生奖学金 (Freshmen Scholarship) 清华大学 (Tsinghua University)
2016 年天津市高考第 3 名（总人数 35169） 2016 年 8 月
Ranked 3rd among 35,169 students in 2016 Tianjin College Entrance Exam Aug 2016

工作实习经历 Industry Experience

通过强化抽象思维提升大语言模型推理鲁棒性的方法 (苹果) 2025 年 2 月至今
Improving Reasoning Robustness of LLMs by Reinforcing Abstract Thinking (Apple) Feb 2025 - Present
公司指导人 (Supervisor): Emmanuel Abbe, AIML-MLR, Apple 瑞士, 苏黎世 (Zurich, Switzerland)

- 提出了一个抽象化推理的框架（**AbstraL**），该框架利用强化学习来训练大语言模型生成推理问题的抽象化表述（abstractions）。这些抽象化表述可以经由符号化工具（例如数学方程求解器）稳定地推导出问题的答案。
  Proposed an abstract reasoning framework, **AbstraL**, which trains LLMs with reinforcement learning to generate abstractions of reasoning problems. The abstractions can be connected to symbolic tools to stably derive the solutions.
- 所提出的抽象化推理方法 AbstraL 有效地提高了大语言模型在数学推理上的鲁棒性，使得大语言模型能够更好地泛化于自然性（例如改变问题的输入条件）和对抗性（例如加入无关的干扰条件）的分布迁移。
  AbstraL effectively improves LLMs' reasoning robustness on mathematics, when encountering both natural (e.g., altering input conditions) and adversarial (e.g., adding distracting conditions) distribution shifts.
- 此研究项目的论文成果已投稿。
  An academic paper is under review.

面向鲁棒多模态推理的情境化视觉表征学习 (Meta) 2024 年 9 月至今
Contextualized Visual Representation Learning for Robust Multimodal Reasoning (Meta) Sep 2024 - Present
公司指导人 (Supervisor): Tianlu Wang, FAIR, Meta 瑞士, 洛桑 (Lausanne, Switzerland)

- 使用对比学习的方法训练一个新的视觉编码器 (**C-ViT**)。该视觉编码器能够将视觉输入（即图像）和情境上下文（例如用户的请求或问题）进行早期融合，从而学习到能够灵活适应不同情境的动态视觉特征表示。
  Used contrastive learning to train a new vision encoder with early fusion of visual inputs (i.e., images) and contexts (e.g., user requests or queries), **C-ViT**, which can learn dynamic visual representations flexibly adapted to different contexts.
- 将训练得到的情境化视觉编码器接入到下游的多模态推理大模型，动态的视觉编码提升了多模态模型在测试时对于训练分布外（OOD）样本的泛化性能。
  Plugged our contextualized vision encoder C-ViT into multimodal LLM reasoners, where the dynamic vision encoding enables better test-time generalization to out-of-distribution (OOD) queries.
- 此研究项目的论文成果已投稿。
  An academic paper is under review.

面向大语言模型智能体高效工具使用的抽象推理方法 (Meta) 2023 年 6 月至 2024 年 3 月
Efficient Tool Use with Abstract Reasoning for LLM Agents (Meta) June 2023 - March 2024
公司指导人 (Supervisors): Tianlu Wang, Jane Yu and Asli Celikyilmaz, FAIR, Meta 美国，西雅图 (Seattle, USA)

- 提出了一种链式抽象 (**CoA**) 推理的方法，该方法训练大语言模型智能体用抽象的占位符来进行推理链的规划。这种抽象推理规划在迁移至训练分布外的知识时具有良好的鲁棒性，并使得大语言模型智能体能够更容易地进行并行化的文本解码和工具调用。
  Developed chain-of-abstraction (**CoA**) reasoning, where LLM agents learn to plan reasoning chains with abstract placeholders, which are robust to out-of-distribution knowledge shifts, and enable decoding and tool calling in parallel.
- 在数学推理和基于维基百科的问答任务中，使用 CoA 推理方法的大语言模型智能体同时实现了更精确的推理和更高效的工具使用，并且展示出了更好的零样本泛化性能。
  LLM agents with CoA achieve more accurate and efficient reasoning with tools in both mathematical reasoning and Wikipedia-based QA, which also present better zero-shot generalization performance.
- 相关研究论文（第一作者）发表于 COLING 2025 国际会议。
  Published a 1st-author paper to the 31st International Conference on Computational Linguistics (COLING 2025).

学术经历 Academia Experience

开放式生成任务中的推理及其与世界知识的对齐 (博士项目) 2021 年 9 月至今
Reasoning with World Knowledge Alignment in Open-Ended Generative Tasks (PhD) Sep 2021 - Present
导师 (Advisor): Antoine Bosselut, Natural Language Processing (NLP) Lab, EPFL 瑞士, 洛桑 (Lausanne, Switzerland)

- 提出了一个新的基准，**ComFact**, 它面向考虑上下文语境的常识知识检索任务。相比于传统的启发式方法，用 ComFact 数据训练得到的常识知识检索器在 F1 评测指标上取得了约 34 个百分点的绝对提升。同时，更好的常识知识检索也为下游的开放域对话生成任务带来了约 9.8% 的相对提升。研究论文（第一作者）被 EMNLP 2022 国际会议接收入 Findings。
  Developed a new benchmark, **ComFact**, for contextual commonsense knowledge linking. Knowledge retrievers trained on ComFact achieve ∼34% absolute F1 boost over heuristics, which also yield ∼9.8% relative improvement on a downstream dialogue generation task. Published an academic paper as 1st author, which is accepted to EMNLP 2022 findings.
- 提出了一个新的常识知识图谱，**PeaCoK**，它大规模地表征了以社会角色为中心的世界知识。PeaCoK 使得常识知识生成器能够可靠地学习有关社会角色属性的推理，并用推理得到的知识有效地增强了下游的叙事（例如对话）生成，提升了叙事的一致性与参与度。研究论文（第一作者）发表于 ACL 2023 国际会议。
  Proposed a new commonsense knowledge graph, **PeaCoK**, that represents world-level persona knowledge at scale. PeaCoK promotes the learning of reliable persona inference generators, which also enables more consistent and engaging downstream narrative modeling. Published an academic paper as 1st author, which is accepted to ACL 2023.
- 提出了一系列基于扩散模型结构的，考虑上下文语境的常识知识生成模型，**DiffuCOMET**。相比于传统的基于自回归模型结构的常识知识生成器，DiffuCOMET 在训练时考虑了上下文情境，并利用了扩散模型的优势，因而能更好地平衡生成知识的多样性和质量。研究论文（第一作者）发表于 ACL 2024 国际会议。
  Proposed a series of knowledge models, **DiffuCOMET**, which uses the diffusion method to generate contextual commonsense knowledge. DiffuCOMET outperforms auto-regressive knowledge generators, with better diversity and quality trade-off. Published an academic paper as 1st author, which is accepted to ACL 2024.
- 提出了一个新的基准，**VinaBench**，以应对视觉叙事生成任务的挑战，即视觉叙事需自洽且忠实于输入的文本叙事。VinaBench 使视觉生成模型能够学习叙事中潜在的常识和论述的约束，从而有效地提高视觉叙事生成的自洽性及其与输入文本叙事的一致性。研究论文（第一作者）发表于 CVPR 2025 国际会议。
  Developed a new benchmark, **VinaBench**, to address the challenge of faithful and self-consistent visual narrative generation. VinaBench enables learning the underlying commonsense and discourse constraints in visual narratives, which effectively improves the consistency of visual narrative generations and their alignment to the input textual narrative. Published an academic paper as 1st author, which is accepted to CVPR 2025.
- **下一步工作：** 面向上一个工作（VinaBench）所提出的带有约束的视觉叙事生成任务，开发在测试时可（根据用户需求）动态泛化至新约束的视觉叙事模型。计划将此研究项目的论文成果投稿至 ICLR 2026 或 CVPR 2026。
  Next Step: Test-time generalizable modeling for visual narrative generation with constraints (based on VinaBench). Planning to submit an academic paper to ICLR 2026 or CVPR 2026.

面向端到端任务导向型对话的半结构化知识管理（研究助理项目） 2020 年 9 月至 2022 年 2 月
Semi-Structured Knowledge Management for End-to-End Task-Oriented Dialog Modeling (RA) Sep 2020 - Feb 2022
导师 (Advisor): Minlie Huang, Conversational Artificial Intelligence (CoAI) Lab, Tsinghua University 中国，北京 (Beijing, China)

- 定义了一个带有半结构化知识管理的任务导向型对话任务，并开发了面向所定义任务的数据集 Mod-MultiWOZ。
  Defined a task of modeling task-oriented dialog with management of semi-structured knowledge, and developed a modified version of MultiWOZ dataset (Mod-MultiWOZ) facing the defined task.
- 提出了一系列端到端的任务导向型对话系统来解决所定义的任务，该系统使用一种扩展的信念跟踪方法来管理半结构化的任务知识库，并在 Mod-MultiWOZ 数据集上展现了优异的任务型对话性能。
  Proposed several end-to-end task-oriented dialog systems to address the defined task, which use an extended belief tracking to manage semi-structured knowledge and achieve strong performances on the Mod-MultiWOZ dataset.
- 两篇相关研究论文（第一作者）被分别收录于 ACL 2021 国际会议 Findings 和 TASLP 国际期刊。
  Published two academic papers as 1st author, which are accepted to ACL 2021 findings and TASLP, respectively.

基于转述的面向任务导向型对话生成的增强方法（暑期研学项目） 2019 年 6 月至 2020 年 7 月
Paraphrase Augmented Task-Oriented Dialog Generation (Summer Internship) Jun 2019 - Jul 2020
导师 (Advisor): Zhou Yu, Department of Computer Science, University of California, Davis 美国，戴维斯 (Davis, USA)

- 提出了一种在任务导向型对话中挖掘用户话语潜在的对话功能的方法，并基于挖掘的对话功能构建了对话样本的转述，用以数据增强。
  Proposed a method of excavating potential dialog function of user utterances in task-oriented dialog modeling, which is used to construct dialog paraphrases for data augmentation.
- 提出了一种使用转述数据来增强任务导向型对话生成的方法，并以此方法提高了对话模型 TSCP 和 DAMD 在 CamRest676 和 MultiWOZ 数据集上的任务完成能力。
  Proposed a new framework of paraphrase augmented task-oriented dialog generation, which improves the task completion ability of state-of-the-art dialog models TSCP and DAMD on CamRest676 and MultiWOZ datasets, respectively.
- 相关研究论文（第一作者）发表于 ACL 2020 国际会议。
  Published an academic paper as 1st author, which is accepted to ACL 2020.

面向自动语音识别的混合特征跨维随机场语言模型 (本科毕业设计项目) 2019 年 3 月至 2020 年 5 月

Mixed-Feature Trans-dimensional Random Field Language Models for Speech Recognition Mar 2019 - May 2020

导师 (Advisor): Zhijian Ou, Speech Processing & Machine Intelligence (SPMI) Lab, Tsinghua University 中国，北京 (Beijing, China)

- 提出了混合特征跨维随机场语言模型（Mixed TRF LM），它将离散 n 元语法特征和基于神经网络的特征集成在单步的语言模型的构建中。
  Proposed mixed-feature trans-dimensional random field language models (Mixed TRF LMs), which integrate both discrete n-gram features and neural-network-based features in a single-step model construction.
- 应用动态噪声对比估计（动态 NCE）训练了所提出的混合特征跨维随机场语言模型，在 n-最佳列表重新评分任务中，所提出的模型在 PTB 和 Google 十亿字数据集上的表现超越了其它语言模型。
  Applied dynamic noise-contrastive estimation (dynamic NCE) to train Mixed TRF LMs and proved its outperformance in n-best list rescoring tasks on both PTB and Google one-billion-word datasets.
- 相关研究论文（第一作者）发表于 ICASSP 2020 国际会议。
  Published an academic paper as 1st author, which is accepted for oral presentation at ICASSP 2020.

掌握技能 Skills

编程语言 (Programming): Python, C/C++, HTML, Shell, MATLAB, SQL, Verilog

机器学习和人工智能工具 (ML/AI): PyTorch, Transformers, NLTK, DeepSpeed, TensorFlow, Scikit-Learn

语言 (Languages): 汉语（母语）Chinese (Native), 英语（流利）English (Fluent)

已修课程 Coursework

计算机科学相关: 机器学习, 强化学习, 分布式信息系统 (包含信息检索和自然语言处理), 数据与算法 (包含数据结构和算法), 数据库原理, 计算机程序设计 (C/C++), 媒体与认知 (包含人工智能), 高级 Matlab 编程及其应用, 操作系统, 数字逻辑和处理器基础 (包含计算机体系架构和汇编语言)

Computer Science: Machine Learning, Reinforcement Learning, Distributed Information Systems (information retrieval, natural language processing), Data and Algorithm (data structures and algorithms), Database Concepts, Computer Program Design (C/C++), Media and Cognition (artificial intelligence), Advanced Matlab Programming and Its Application, Operating System, Fundamental of Digital Logic and Processor (computer architecture and assembly language)

其它: 概率论与随机过程, 线性代数, 离散数学, 微积分, 认知心理学
Others: Probability and Stochastic Processes, Linear Algebra, Discrete Mathematics, Calculus, Cognitive Psychology