Debjit Paul

github.com/debjitpaul debjitpaul.github.io linkedin.com/in/debjit-paul debjitpaulms@gmail.com

Education

Heidelberg University 2018 - 2022
Ph.D. in Computational Linguistics *Grade: Summa cum laude*
Advisor: Prof. Anette Frank

Saarland University 2014 - 2017
Msc. in Computer Science
Advisor: Prof. Ditrech Klakow

GuruNanak Institute of Technology 2010 - 2014
B.Tech in Computer Science and Engineering

Work Experience

Research Scientist at Huawei Noah's Ark Lab R&D, London March 2025- Present
- Working on designing agentic frameworks
- Creating benchmarks to evaluate agentic frameworks

Postdoctoral Researcher at EPFL April 2022- November 2024
Host: Prof. Boi Faltings, Prof. Antoine Bosselut, Prof. Robert West
- Developed methods to enhance the reasoning capabilities of LLMs
- Designed new benchmark datasets to evaluate the reasoning capabilities of LLMs
- Developed reinforcement learning methods to enhance text generation capabilities of current NLP systems

Applied Scientist Internship at Amazon Winter 2021
- Designed new real-world continual learning experiment setup for production
- Developed methods for class incremental learning, aiming to expand the features for spoken language understanding tasks within a voice assistant framework like Alexa.

Research Summary

My overarching research objective is to enhance the **reasoning capabilities of AI systems** by advancing their ability to understand the complex interplay between language, knowledge representation, and cognitive reasoning processes. To accomplish this goal, I have focused my investigations on three areas:

- **Aligning Models with Human or AI Feedback** - Text generation models are observed to display undesired and inconsistent behaviours, such as hallucination and unfaithful reasoning. My research has focused on designing methods to rectify the undesired behaviours of text generation models through interaction.
- **Neuro-symbolic Reasoner** - NLP systems must integrate structured, symbolic knowledge with neural representations to capture knowledge dynamics and generalise robustly to unseen situations. My research has focused on developing computational models that perform reasoning by combining the expressiveness of symbolic structures with the adaptability of neural architectures.
- **Learning to generate explanations for reasoning** - When a model explains its decision-making steps or logical steps it took to reach that conclusion, it enhances transparency and helps build trust in the system. My research has focused on designing computational models that generate explanations before making the final decisions.

Selected Recent Publications

Please see Google Scholar for the complete list of publications; * denotes equal contributions

- *A Logical Fallacy-Informed Framework for Argument Generation*
  *Association for Computational Linguistics: NAACL 2025 (**Outstanding Paper Award**)*
  *Luca Mouchel, **Debjit Paul**, Robert West, Antoine Bosselut, Boi Faltings*
- *Parity-Aware Byte-Pair Encoding: Improving Cross-lingual Fairness in Tokenization*
  arXiv 2025 (Under Review)
  *Negar Foroutan, Clara Meister, **Debjit Paul**, Joel Niklaus, Sina Ahmadi, Antoine Bosselut, Rico Sennrich*
- *Making Reasoning Matter: Measuring and Improving Faithfulness of Chain-of-Thought Reasoning*
  Findings of the Association for Computational Linguistics: EMNLP 2024
  Debjit Paul, Robert West, Antoine Bosselut, Boi Faltings
- *Creativity in AI: Progresses and Challenges*
  arXiv 2024 (Under Review)
  Mete Ismayilzada, **Debjit Paul**, Antoine Bosselut, Lonneke van der Plas
- *Entity Insertion in Multilingual Linked Corpora: The Case of Wikipedia*
  Association for Computational Linguistics: EMNLP 2024
  Tomás Feith, Akhil Arora, Martin Gerlach, **Debjit Paul**, Robert West
- *Exploring Defeasibility in Causal Reasoning*
  Findings of the Association for Computational Linguistics: ACL 2024
  Shaobo Cui, Lazar Milikic, Yiyang Feng, Mete Ismayilzada, **Debjit Paul**, Antoine Bosselut, Boi Faltings
- *REFINER: Reasoning Feedback on Intermediate Representations*
  Association for Computational Linguistics: EACL 2024
  ***Debjit Paul**, Mete Ismayilzada, Maxime Peyrard, Beatriz Borges, Antoine Bosselut, Robert West, Boi Faltings*
- *CRoW: Benchmarking Commonsense Reasoning in Real-World Tasks*
  Association for Computational Linguistics: EMNLP 2023
  *Mete Ismayilzada, **Debjit Paul****, *Syrielle Montariol**, *Mor Geva, Antoine Bosselut*
- *CRAB: Assessing the Strength of Causal Relationships Between Real-world Events*
  Association for Computational Linguistics: EMNLP 2023
  *Angelika Romanou, **Debjit Paul****, *Syrielle Montariol**, *Leo Laugier, Karl Aberer, Antoine Bosselut*
- *Language Model Decoding as Likelihood–Utility Alignment*
  Findings of the Association for Computational Linguistics: EACL 2023
  *Martin Josifoski, Maxime Peyrard, Frano Rajič, Jiheng Wei, **Debjit Paul**, Valentin Hartmann, Barun Patra, Vishrav Chaudhary, Emre Kiciman, Boi Faltings and Robert West*
- *COINS: Dynamically Generating <u>CO</u>ntextualized <u>I</u>nference Rules for <u>N</u>arrative <u>S</u>tory Completion*
  Association for Computational Linguistics: ACL 2021
  ***Debjit Paul**, Anette Frank*
- *Generating Hypothetical Events for Abductive Inference*
  Proceedings of the Tenth Joint Conference on Lexical and Computational Semantics (*Sem 2021)
  Debjit Paul, Anette Frank
- *Class Incremental Learning for Intent Classification with Limited or No Old Data*
  Proceedings of the First Workshop on Ever Evolving NLP (EvoNLP)
  ***Debjit Paul**, Daniil Sorokin, Judith Gaspers*

- *CO-NNECT: A Framework for Revealing Commonsense Knowledge Paths as Explicitations of Implicit Knowledge in Texts*
  Proceedings of the 14th International Conference on Computational Semantics
  Maria Becker, Katharina Korfhage, ***Debjit Paul****, Anette Frank*
- *Social Commonsense Reasoning with Multi-Head Knowledge Attention*
  Findings of the Association for Computational Linguistics: EMNLP 2020
  Debjit Paul*, Anette Frank*
- *Argumentative Relation Classification with Background Knowledge*
  Proceedings of the 8^{th} International Conference on Computational Models of Argument (COMMA 2020)
  Debjit Paul, Maria Becker, Juri Opitz, Graeme Hrist and Anette Frank
- *Ranking and Selecting Multi-Hop Knowledge Paths to Better Predict Human Needs*
  Association for Computational Linguistics: NAACL 2019
  Debjit Paul, Anette Frank

Skills

Languages: **Python**, Java, LaTeX
Tools: Git/GitHub, Unix Shell, VS Code, PyCharm, Atom
Libraries: pandas, NumPy, Matplotlib, Tensorflow, pyTorch, spaCy, Transformers library

Mentoring

PhD Researchers
Angelika Romanou co-advised with Antoine Bosselut — Sept 2022 - June 2023
MS Thesis
Mete Ismayilzada, Topic: Commonsense Reasoning — Sept 2022 - June 2023
Farouk Boukil, Topic: Faithful Reasoning — Sept 2024 - Present
MS Researchers
Rochat Mathieu Louis, Topic: Graph Continual Learning — Sept 2022 - April 2023
Gabriele D'Angeli, Topic: Reinforcement Learning for NLP — August 2023 - Present
Briki Farah, Topic: Reinforcement Learning for NLP — August 2023 - Present
Colin Baptiste Hofmann, Topic: Reasoning for NLP — August 2023 - Present
EPFL Summer Intern
Omar El Malki, Topic: Reinforcement Learning for NLP — June 2022 - Feb 2023
Luca Mouchel, Topic: Logical Fallacy — August 2023 - Present

Talks and Lectures

Faithful Reasoning with Language Models

* Guest Lecture at Topics in Natural Language Processing Course, EPFL, 2024
* Microsoft Research, 2024

Modern NLP with Large Language Models

* Invited Lecture at IFI Summer School PhD Student, University of Zurich, 2024

Textual Feedback to Improve Natural Language Reasoning

* Conference Presentation at EACL 2024
* 4th TAILOR Conference – Trustworthy AI from lab to market, 2024

Neuro-Symbolic Commonsense Reasoning in NLP

* Invited Talk at KU Leuven, 2022

Generating Contextualized Inference Rules for Narrative Story Completion

* Conference Presentation at ACL 2021

Multi-Head Knowledge Attention for Social Commonsense Reasoning

* Conference Presentation at EMNLP 2020
* Workshop Presentation at CODI 2020
* Invited Talk at Edinburgh Napier University 2020

Extracting Multi-Hop Knowledge Paths for Human Needs Classification

* Conference Presentation at NAACL 2019
* Invited Talk at AIPHES 2019

Honors and Awards

Outstanding Paper Award at NAACL 2025
Nominated as Best Student paper at COMMA 2020
Facebook Travel Award for EurNLP 2019, London,UK
Winner of HQ Hackathon 2017, at Trivago, Dusseldorf, Germany

Services

Area Chair: ACL 2025 & 2024, EMNLP 2023 & 2024, NAACL 2025, *SEM 2024
Program Committee: TACL 2023-2025, ACL 2023 & 2022, COLM 2024 & 2025, NAACL 2021, EMNLP 2022, 2021 & 2020, EACL 2023 & 2021, ARR, *SEM 2020 & 2021, CoNLL 2021, KI 2019, COIN 2019
Session Chair: EMNLP 2023

References

Prof. Anette Frank Email: frank@cl.uni-heidelberg.de
Prof. Antoine Bosselut Email: antoine.bosselut@epfl.ch
Prof. Boi Faltings Email: boi.faltings@epfl.ch
Prof. Robert West Email: robert.west@epfl.ch