Neuro-Symbolic Natural Language Processing

André Freitas, Marco Valentino, Danilo S. Carvalho

Website: `https://sites.google.com/view/nesynlp2025`

Despite the performance leaps delivered by Large Language Models (LLMs), NLP systems based only on deep learning architectures still have limiting capabilities in terms of delivering safe and controlled reasoning, interpretability, and adaptability within complex and specialised domains, restricting their use in areas where reliability and trustworthiness are crucial. Neuro-symbolic NLP methods seek to overcome these limitations by integrating the flexibility of contemporary language models with the control/interpretability of symbolic methods. This hybrid approach brings the promise to both enhance inference capabilities and to deepen the theoretical understanding of LLMs. This tutorial aims to bridge the gap between the practical performance of LLMs and the principled modelling of language and inference of formal methods. We provide an overview of formal foundations in linguistics and reasoning, followed by contemporary architectural mechanisms to interpret, control, and extend NLP models. Balancing theoretical and practical activities, the tutorial is suitable for PhD students, experienced researchers, and industry practitioners.

André Freitas, Research Group Leader at the Idiap Research Institute (Switzerland) and Associate Professor (Senior Lecturer) at the University of Manchester (UK).
email: `andre.freitas@idiap.ch`
website: `https://www.andrefreitas.net`
André Freitas is a and a Research Group Leader at the Idiap Research Institute (Switzerland) an Associate Professor (Senior Lecturer) at the Department of Computer Science at the University of Manchester and an AI Group leader at the National Biomarker Centre (CRUK Manchester Institute). He leads the Neuro-symbolic AI Lab. His main research interests are on enabling the development of AI methods to support abstract, flexible and controlled reasoning. In particular, he investigates how the combination of neural and symbolic paradigms can deliver better models of reasoning. He is an active contributor to the main conferences and journals in the AI/Natural Language Processing (NLP) interface (ACL, EMNLP, EACL, COLING, TACL, Computational Linguistics, AAAI, NeurIPs), with over

1

100 peer-reviewed publications.

Marco Valentino, Assitant Professor (Lecturer) at the University of Sheffield (UK).
email: `m.valentino@sheffield.ac.uk`
website: `https://www.marcovalentino.net/`
Marco is a lecturer in the School of Computer Science at the University of Sheffield. His research focuses on developing the next generation of AI systems that can use explanatory inference as a core mechanism for learning and reasoning in natural language, particularly in complex domains such as science, mathematics, and healthcare. To this end, he investigates the integration of neural and symbolic AI methods to enhance the robustness and faithfulness of AI-generated explanations and, ultimately, to uncover the principles governing the explanatory inference process in humans. He regularly contributes to major AI and NLP conferences, including AAAI, ACL, EMNLP, NAACL and EACL. Marco was involved in the organisation of workshops, including MathNLP (EMNLP 2022 and 2025, LREC-COLING 2024), and TextGraphs (COLING 2022 and ACL 2024), and tutorials, including *"Reasoning with Natural Language Explanations"* at EMNLP 2024.

Danilo S. Carvalho, Principal Clinical Informatician, National Biomarker Centre, University of Manchester (UK).
email: `danilo.carvalho@manchester.ac.uk`
website: `https://danilosc.com`
Danilo is a Principal Clinical Informatician (Researcher) at the National Biomarker Centre, Cancer Research UK - Manchester Institute, at the University of Manchester, working on *Safe and Explainable Artificial Intelligence (AI) architectures*, centred on generative AI for biomarker discovery and analysis. His main research subject is *Representation Learning*, from the meaning of words to gene interactions, for supporting controlled inference and discovery on large scale data, with emphasis on conceptual interpretation with controlled inference over complex concept spaces. He has experience in both industry and academia, having presented works at multiple AI/NLP international conferences over the past 10 years, such as EACL, COLING and ESANN.

2