Bálint Gyevnár

balint.gyevnar@ed.ac.uk | gbalint.me | github.com/gyevnarb

Education

University of Edinburgh — Edinburgh, UK
Centre for Doctoral Training in Natural Language Processing — *Sep. 2021 – May 2025 (est.)*
Supervisors: Stefano V. Albrecht, Shay B. Cohen, and Christopher G. Lucas

University of Edinburgh — Edinburgh, UK
Integrated Master of Informatics — *Sep. 2016 – May 2021*
Supervisor: Maria Wolters

Nanyang Technological University — Singapore
Undergraduate Exchange in Computer Science — *Aug. 2018 – May 2019*

Experience

Teaching Assistant — Sep. 2020 – present
University of Edinburgh — *Edinburgh, UK*

- Delivering and moderating online tutorial sessions of ∼12 students for introductory level Machine Learning course.
- Coursework and exam marker for courses in the School of Informatics, including Reinforcement Learning, Computer Systems, and Machine Learning.

Research Intern — May 2020 – Oct. 2020
Five AI Ltd. — *Edinburgh, UK*

- Developed IGP2, a goal-based interpretable prediction and planning system for autonomous vehicles with intuitive explanations.
- Scenario-based and open-world testing and evaluation of IGP2.
- Publication at International Conference on Robotics and Automation (ICRA), 2021.

Current Projects

Explainable Autonomous Vehicle Intelligence — Sep. 2021 – Present

- Generating causally justified explanations of actions of AV motion planning and prediction.
- Integrating generated explanations with dialogue systems to deliver relevant natural language explanations.
- Evaluating of methods with human participants to measure the effect of explanations on trust and understanding.
- Leading and managing a team of 5 people with the Autonomous Agents group.

Acquisition & Communication of Colour Naming Systems — Sep. 2021 – Present

- Measuring the communicative efficiency and acquisition patterns of human colour naming systems using information-theoretic measures.
- Simulating acquisition patterns of colour term learning using self-organising maps and the World Colour Survey.
- Lead author on team of 5 with journal submission to *Entropy* currently under revision.

Developing and maintaining open-source repository (with 48 stars) for interpretable prediction and planning system for autonomous vehicles. — May 2021 – Present

Other Skills

Research Interests: explainable AI, autonomous vehicles, natural language processing, causal reasoning.
Programming Languages: Proficient in **Python**. Experienced with **C#**, **C++**. Some Java, Bash, and R.
Natural Languages: English (fluent), German (advanced), Japanese (intermediate), Hungarian (fluent).
Libraries: PyTorch, TensorFlow, Pandas, Matplotlib, Django, ggplot2, etc.

Publications

Conference

- S.V. Albrecht, C. Brewitt, J. Wilhelm, **B. Gyevnar**, F. Eiras, M. Dobre, S. Ramamoorthy. Interpretable Goal-based Prediction and Planning for Autonomous Driving, *International Conference on Robotics and Automation (ICRA)*, 2021

- C. Brewitt, **B. Gyevnar**, S. Garcin., S.V. Albrecht. GRIT: Fast, Interpretable, and Verifiable Goal Recognition with Learned Decision Trees for Autonomous Driving, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021

Journal

- **B. Gyevna**r, G. Dagan, C. Haley, S. Guo, F. Mollica. Communicative Efficiency or Iconic Learning: Do communicative and acquisition pressures interact to shape colour-naming systems?, *PsyArXiv preprint*, 2022

Workshop

- **B. Gyevnar**, M. Tamborski, C. Wang, C.G. Lucas, S.B. Cohen, S.V. Albrecht. A Human-Centric Method for Generating Causal Explanations in Natural Language for Autonomous Vehicle Motion Planning, *IJCAI Workshop on Artificial Intelligence for Autonomous Driving*, 2022

- C. Brewitt, S.V. Albrecht, J. Wilhelm, **B. Gyevnar**, F. Eiras, M. Dobre, S. Ramamoorthy. Autonomous Driving with Interpretable Goal Recognition and Monte Carlo Tree Search, *Workshop at RSS on Interaction and Decision-Making in Autonomous-Driving*, 2020