Max Taylor-Davies

https://maxtaylordavi.es

Summary

I am a second-year PhD student in **robotics/ML** and **computational cognitive science**. My research is primarily concerned with **social cognition and learning** in humans and artificial agents. I am also interested in (multi-agent) reinforcement learning, moral cognition, and AI alignment (among other topics).

EDUCATION

• University of Edinburgh PhD, Centre for Doctoral Training in Robotics and Autonomous Systems Edinburgh, UK Sep 2022 – ongoing Imperial College London London, UK

MEng Molecular Bioengineering; 1st class honours

Oct 2017 – Jul 2021

Email: s2227283@ed.ac.uk

EMPLOYMENT

• UCL Visual Plasticity Lab Research Assistant London, UK Mar 2022 - Sep 2022

PolyAI
Software Engineer

London, UK
June 2020 – June 2021

MedEngineBerlin, GermanySoftware Engineer / Research EngineerJune 2018 - Oct 2019

Teaching

Seminar in Cognitive Modelling Teaching assistant + marker University of Edinburgh 2023-24

SELECTED PAPERS

- Taylor-Davies, M. & Lucas, C.G. Rational compression in choice prediction (International Conference on Cognitive Modeling 2024)
- Taylor-Davies, M. & Lucas, C.G. Balancing utility and cognitive cost in social representation (Information-Theoretic Principles in Cognitive Systems Workshop, NeurIPS 2023)
- McCallum, S., **Taylor-Davies**, **M.**, Albrecht, S. & Suglia, A. Is feedback all you need? leveraging natural language feedback in goal-conditioned reinforcement learning (Goal-Conditioned Reinforcement Learning Workshop, NeurIPS 2023, **selected for spotlight talk**)
- Taylor-Davies, M., Droop, S. & Lucas, C.G. Selective imitation on the basis of reward function similarity (CogSci 2023)

Invited talks

- Social learning and social groups @ Human Decision-making Lab, Edinburgh University (November 2023)
- Imitation, preferences and group identity @ Computational Cognitive Development Lab, Brown University (November 2023)

REVIEWING

- CogSci 2024
- CogSci 2023