Abhimanyu Pallavi Sudhir

AI researcher working on program markets in the context of AI and bounded rationality.

Formal education

- University of Warwick · PhD Computer Science · 2022-26 – supervisor: Long-Tran-Thanh
- Imperial College London · Undergraduate Masters (MSci) Math · 2018-22 – 1st class honors

Internships

- Goldman Sachs · AI Research Intern · Jan-Aug 2021, London – Developed and implemented novel methods in NLP and recurrent neural networks for financial forecasting

Research

Markets and AI (PhD work)

My primary work focuses on topics at the intersection of markets and AI, especially AI alignment.

- Abhimanyu Pallavi Sudhir and Long-Tran Thanh (2024), "Betting on what is neither verifiable nor falsifiable", arxiv.org/abs/2402.14021
- Abhimanyu Pallavi Sudhir (2021), "A mathematical definition of property rights in a Debreu economy", arxiv.org/abs/2107.09651

> *Related write-ups and talks.*
>
> — Lesswrong post (2024): "Reinforcement learning from market feedback, and other uses of information markets"
> — LessWrong post (2023): "Betting on what is un-falsifiable and un-verifiable"
> — Poster at the Co-operative AI Foundation (CAIF) summer workshop, 2023: abhimanyu.io/legacy_writing/PhD_presentations/caif.pdf
> — LessWrong post (2022): "Meaningful things are those the universe possesses a semantics for"

Consistency checks and forecasting (2024)

Developed a Consistency Benchmark for LLM forecasters, along with a principled arbitrage-based metric for inconsistency and a novel consistency calibration method similar to Platt scaling.

- Abhimanyu Pallavi Sudhir*, Alejandro Alvarez, Adam Shen, and Daniel Paleka* (2024), "Consistency Checks for Language Model Forecasters" *Workshop paper, accepted to:* Agentic Markets Workshop at ICML 2024; NextGenAISafety Workshop at ICML 2024; Oxford ELLIS Robust LLMs Workshop 2024

Scalable Oversight Benchmark (2024)

Ongoing collaboration with a team supervised by Arjun Panickssery and Nina Rimsky, to develop a comprehensive benchmark for Scalable Oversight protocols.

General mathematics (Undergraduate work and prior)

- Abhimanyu Pallavi Sudhir (2019), "Infinitesimal translations and a multivariate Grünwald-Letnikov calculus", arxiv.org/abs/1904.02710
- Abhimanyu Pallavi Sudhir (2019), "Generalisations of the determinant to interdimensional transformations: a review," arxiv.org/abs/1904.08097
- Abhimanyu Pallavi Sudhir (2018), "The generalized Cauchy derivative as a principal value of the Grünwald-Letnikov fractional derivative for divergent expansions," arxiv.org/abs/1809.08051
- Abhimanyu Pallavi Sudhir (2014), "On the Determinant-like function and the Vector Determinant," *Advances in Applied Clifford Algebras* (24-3: 805-807), doi:10.1007/s00006-014-0455-3

- Abhimanyu Pallavi Sudhir (2013), "Defining the Determinant-like function for m by n matrices using the exterior algebra," *Advances in Applied Clifford Algebras* (23-4: 787-792), doi:10.1007/s00006-013-0416-2

Academic service

- *Teaching Assistant for CS255: Artificial Intelligence (Warwick)* · 2024
- *Reviewer for NextGenAISafety Workshop at ICML 2024* · 2024
- *Teaching Assistant for CS141: Functional Programming (Warwick)* · 2023
- *Reviewer for Advances in Applied Clifford Algebras (Springer)* · 2020

Courses and workshops attended

- *Co-operative AI Foundation* · Jul 2023 · workshop on AI and cooperative game theory

Other projects

Costly (2024)

Wrote the Python package `costly` for estimating costs and running times of complex LLM workflows/experiments/pipelines in advance before spending money, via simulations.

Project page: github.com/abhimanyupallavisudhir/costly
Install: `pip install costly`

Equivariant learning (2021-22)

Final-year MSci project with Professor Jeroen Lamb at Imperial College London exploring equivariant learning and causal DAGs.

Report: abhimanyu.io/legacy_writing/Imperial_reports/m4r.pdf

Lie theory (2019)

Undergraduate research project with Professor Richard Thomas at Imperial College London on Lie groups and algebras.

Report: abhimanyu.io/legacy_writing/Imperial_reports/urop.pdf
Presentation: abhimanyu.io/legacy_writing/Imperial_presentations/lie_theory.pdf

Lean (2018-19)

Computerized formal proving in Lean with Professor Kevin Buzzard at Imperial College London.

- Wrote the FilterProduct.lean and Hyperreal.lean modules for the Lean math library
- Formalized the first-year "Foundations of Analysis" module exam Blog post: xenaproject.wordpress.com/2019/05/06/m1f-imperial-undergraduates-and-lean/

PhysicsOverflow (2014-15)

Co-founded PhysicsOverflow, a postgraduate-level physics Q&A site and open peer review system. See en.wikipedia.org/wiki/PhysicsOverflow for more details.

- Abhimanyu Pallavi Sudhir and Rahel Knoepfel (2015), "PhysicsOverflow: A postgraduate-level physics Q&A site and open peer review system," *Asia-Pacific Physics Newsletter* (4-1: 53-55), doi:10.1142/S2251158X15000193

Links

- Email: abhimanyu.pallavi-sudhir@warwick.ac.uk

- Contact: +44-7771824896
- Website: abhimanyu.io
- Blog: TheWindingNumber.blogspot.com
- Google Scholar: scholar.google.com/citations?user=lb38BjYAAAAJ
- Github: github.com/abhimanyupallavisudhir
- LessWrong: lesswrong.com/users/abhimanyu-pallavi-sudhir