Euan Ong

✉ elyro2@cam.ac.uk | ⌂ ong.ac | ◉ euanong | in euanong | 🐦 @euan_ong | 🎓 Euan Ong

move slow and fix things

Education

Trinity College, University of Cambridge   *Cambridge, UK*

B.A. (HONS) IN COMPUTER SCIENCE (FIRST CLASS HONOURS WITH DISTINCTION)   *2020 - 2023*

- Ranked **1st / ~120 students in all three years**. Awarded the **CST Department Prize for Best Student** every year, and the university-wide **Winifred Georgina Holgate Pollard Memorial Prize** in 2023.
- Awarded the college-based Percy and Jeremy Pemberton Prizes (for the two Trinity undergraduates "who have most distinguished themselves in their studies"), Examination Prize, and Top of Tripos Prize. Elected to a Senior Scholarship at Trinity College after first year.

Publications

Learnable Commutative Monoids for Graph Neural Networks
   Euan Ong, Petar Veličković
   The First Learning on Graphs Conference, 2022

Probing the foundations of neural algorithmic reasoning
   Euan Ong
   Tech. rep. UCAM-CL-TR-990, 2023 (won the CS department's **Best Dissertation Award** for 2023)

Image Hijacks: Adversarial Images can Control Generative Models at Runtime
   Luke Bailey*, **Euan Ong***, Stuart Russell, Scott Emmons *(* indicates equal contribution)*
   The Forty-first International Conference on Machine Learning, 2024

Successor Heads: Recurring, Interpretable Attention Heads In The Wild
   Rhys Gould, **Euan Ong**, George Ogden, Arthur Conmy
   NeurIPS ATTRIB (Oral), 2023*; The Twelfth International Conference on Learning Representations*, 2024

Research Projects

Probing the Foundations of Neural Algorithmic Reasoning

BACHELOR'S THESIS (ADVISED BY DR. FERENC HUSZÁR AND DR. PETAR VELIČKOVIĆ)   *Oct 2022 - ongoing*

- As part of an **original, independent project** (which I submitted as my Bachelor's thesis), I explored how NNs learn to imitate algorithms by probing a fundamental claim of neural algorithmic reasoning: that neural networks trained to imitate algorithms in high-dimensional space perform better as 'algorithmic modules' than differentiable versions of the algorithms themselves.
- I tested this claim through statistically robust ablations, and found evidence to refute it. Indeed, based on my observations, I discovered a way to construct differentiable algorithms that preserves their efficiency and correctness guarantees while avoiding performance bottlenecks, thereby achieving a long-standing goal of neural algorithmics.
- This thesis earned the **highest-score in the year**, and won the CS department's **Best Dissertation Award** for 2023.

Image Hijacks: Adversarial Images can Control Generative Models at Runtime

INTERNSHIP, CENTER FOR HUMAN-COMPATIBLE AI (SUPERVISED BY SCOTT EMMONS)   *Jul 2023 - Sep 2023*

- During an internship at CHAI, we developed a way to craft *image hijacks* – adversarial images that can hijack the behaviour of vision-language models (VLMs) at runtime – and trained image hijacks that force VLMs to e.g. leak their context window or comply with harmful instructions.
- I derived a novel algorithm for training images that force the VLM to match the behaviour of a specific prompt under all possible user inputs (thought to be impossible by prior work), which I used to make a VLM believe the Eiffel Tower is in Rome.
- This work was **presented to the UK's AI Foundation Model Taskforce**, and was **accepted to ICML 2024**.

Successor Heads: Recurring, Interpretable Attention Heads In The Wild

INDEPENDENT (ADVISED BY ARTHUR CONMY)   *Jul 2023 - Sep 2023*

- As part of an **independent project with friends**, we discovered *successor heads*: attention heads present in a range of LLMs that increment tokens from ordinal sequences, such as numbers, months and days.
- By learning a commutative diagram, I isolated a common numeric subspace within embedding space, that for any given token (e.g. 'February') encodes the index of that token within its ordinal sequence (e.g. months). We also found that numeric token representations can be decomposed into interpretable features representing the value of the token mod 10, which are amenable to vector arithmetic.
- This work was **accepted as an oral presentation at NeurIPS ATTRIB 2023**, and was **accepted to ICLR 2024**.

Learnable Commutative Monoids for Graph Neural Networks

INDEPENDENT (ADVISED BY DR. PETAR VELIČKOVIĆ)   *Jun 2022 - Sep 2022*

- As part of an **original, independent project**, I explored the expressivity bottlenecks of graph neural network (GNN) aggregators through the lens of *commutative monoid homomorphisms*.
- Using ideas from abstract algebra and functional programming, I built a new *learnable commutative monoid* GNN aggregator, which beats the state of the art on complex aggregation problems (especially out-of-distribution), while remaining efficient and parallelisable on large graphs.
- This result was **published at the Learning on Graphs conference 2022, the flagship conference in graph machine learning**.

Personality Machine

CS GROUP PROJECT, CENTRE FOR GENDER STUDIES (SUPERVISED BY DR. ELEANOR DRAGE) *Jan 2022 - Apr 2022*

- In collaboration with the Centre for Gender Studies at Cambridge, we built a tool highlighting the questionable logic behind the use of AI-driven personality assessments often used in hiring.
- Our tool demonstrates how arbitrary changes in facial expression, clothing, lighting and background can give radically different personality readings, and was **featured in the BBC and the Telegraph**.

Dissecting Transformers for Systematic Generalisation

HACKBRIDGE STUDENT SOCIETY (SUPERVISED BY ALBERT Q. JIANG) *Oct 2020 - Sep 2021*

- Through a student research society, we investigated whether transformers could *systematically generalise* when learning to recognise formal languages such as PARITY and 2-DYCK, empirically corroborating various theoretical claims about transformer generalisation. Inspired by our observations, I derived a parallel, stackless algorithm for recognising 2-DYCK, which is (in principle) implementable by a transformer with a constant number of attention layers.
- As part of this project, **I derived a positional encoding which was later independently invented by DeepMind**, who found that it improves length-generalisation on many abstract tasks.

Object Detection in Thermal Imagery via Convolutional Neural Networks

MAGDALEN COLLEGE SCHOOL (SUPERVISED BY PROF. NIKI TRIGONI AND DR. PEDRO PORTO BUARQUE DE GUSMÃO) *Jan 2019 - May 2019*

- For a high-school research project, I trained a Faster R-CNN object detection network to identify landmarks (e.g. doors and windows) in thermal images of indoor environments, with applications in the development of navigational aids for search and rescue operations.
- The project was submitted to the **TeenTech Awards 2020**, where it **won the Data Science project award in the senior category (age 17-19)**.

Industry Experience

Jane Street Capital (Software Engineering Intern) *Jul 2022 - Sep 2022*

- I developed in-house systems in OCaml, gaining an appreciation for the use of functional paradigms such as monads, GADTs, existential types and type witnesses in large-scale software engineering: I prototyped an email processing pipeline with a rich filtering DSL, and designed a system for autogenerating first-class field accessors for a complex message datatype. **I received an offer to join as a full-time software engineer.**

Focal Point Positioning (Research & Development Intern) *Jul 2021 - Aug 2021*

- I analysed and evaluated the performance of machine learning models in Python when used in a pipeline for pedestrian dead reckoning.

Cisco Ensoft (Software Development Micro-Intern) *Jul 2021*

- I implemented features and visualisations in Python for an internal Cisco dashboard.

Awards & Achievements (Computing)

International Olympiad in Informatics (IOI) *Singapore / Azerbaijan*

BRONZE MEDALLIST, TEAM UK *Sept 2020 / Aug 2019*

- Selected as one of the **top 4 computer scientists in the UK** to compete in the 32nd International Olympiad in Informatics, held online due to COVID-19. **Won a bronze medal, placing 91st out of 343 contestants.**
- Also represented the UK as part of a team of 4 in the 31st International Olympiad in Informatics, held in Baku, Azerbaijan in June 2019.
- Deployed an online judge and training portal for finalists to train on British Informatics Olympiad problems.

Y Combinator

FORMER CO-FOUNDER (YONEDA LABS) *2023*

- **Accepted to Y Combinator's W24 batch** as part of *Reagent AI* (now *Yoneda Labs*), a startup using Bayesian methods for optimising chemical reactions. **I have since left this startup in order to pursue academic research.**

Cybersecurity Contests

VARIOUS AWARDS *2017 - ongoing*

- In 2023, **won first place as part of team cheriPI** in the finals of **pwnEd** (run by University of Edinburgh) and **LakeCTF** (run by EPFL).
- In 2022, **won first place as a team** in the **international Country-to-Country CTF**.
- In 2022, **invited to Meta's BountyCon EDU conference** for university students skilled in application security engineering
- In 2021, **placed in the top 15 teams in NYU's CSAW CTF**, qualifying for the finals.
- In 2021, **won the Global CyberPeace Challenge CTF IT track**, and placed fourth in the OT track (a SANS Grid NetWars tournament)
- In 2020, **placed second on the course CTF** for the **SANS SEC504: Hacker Tools, Techniques, Exploits, and Incident Handling** course.
- In 2018, **won a ticket to the Black Hat conference** at the **National Finals of the UK Cyber Security Challenge**.
- In 2018, **won the Best Individual Performer award** for the final stage of **HM Government's Cyber Discovery programme**, by placing first in a real-world Internet of Things penetration testing scenario and earning the highest individual score in the CTF.

Hack Cambridge Atlas

OVERALL WINNER (HEARSIGHT) *Jan 2022*

- Competed in a team of 4 at the University of Cambridge's annual hackathon to develop *HearSight*, a wearable headset for the visually impaired rendering a 3D binaural 'sound map' of the world around them. The project **won the Main Challenge of the hackathon**.

Google Code-In

Dec 2017

- **Selected as one of 50 grand prize winners from over 3,500 contestants** for my contributions to the open-source organisation Sugar Labs: porting programs from desktop to mobile/web-based clients, and building a server backend, moderation portal and project sharing site for an online block-based programming language (Music Blocks).
- Acted as a mentor for Google Code In 2018 and 2019, helping students new to open source development learn how to contribute and get to grips with the frameworks used by Sugar Labs. Assessed student work and created tasks for students to complete.

Awards & Achievements (Other)

UKMT IMO Winter School
Hungary

INVITED STUDENT *Dec 2018*

- Selected as one of the **top 20 UK mathematicians** based on BMO1 performance for an intensive, 8 day long olympiad training camp.

UKMT Initial Olympiad Training Camp
Oxford

INVITED STUDENT *Aug 2018*

- Selected as one of the **top 24 UK mathematicians** in or under year 11 based on BMO/IMOK performance for an intensive, week-long olympiad training camp. Ranked in the top third (**top 8**) in the Oxford Olympiad. Selected for the Advanced Mentoring Scheme, "for those students working closest to IMO level".

2020	**Team UK**, International Olympiad on Astronomy and Astrophysics (IOAA)
2020	**Winner**, Magdalen College School Coronation Cup ("most outstanding achievement … during the year")
2020	**Distinction**, UKMT British Mathematical Olympiad Round 2
2019	**Finalist (top 17)**, UK Linguistics Olympiad
2019	**Roentgenium Award (top 56)**, Cambridge Chemistry Challenge (C3L6)
2019	**Winner (GCSE category)**, British Education Awards
2019	**Distinction**, ABRSM Grade 8 Piano

Community & Outreach

Cambridge ML Collective (CMC), Cambridge PL Collective (pl@cam)

FOUNDER *2021 - present*

- Founded machine learning (ML) and programming languages (PL) collectives at Cambridge to grow the undergraduate research community.
- Facilitated reading groups and project collaborations, several of which resulted in workshop and conference acceptances.
- Organised, planned and executed Fully Connected, Cambridge's first machine learning conference for undergraduates, in conjunction with Cambridge University AI Society.
- Co-organised a Cambridge hub for the Learning on Graphs conference 2022.

[scale down]

CO-FOUNDER *2023 - present*

- Co-founded `[scale down]`, a community / co-working space for students of both the sciences and humanities to work on passion projects which they otherwise wouldn't have time to do during their degree.
- Student projects include understanding structure in neural networks, modelling the economic effects of a delay in World War 1, and writing concept albums.

Cambridge University Computing and Technology Society (CUCaTS)

PRESIDENT *2020 - 2023*

- Worked to restart CUCaTS, Cambridge's computer science and technology society, after a hiatus in 2019-20.
- Ran events and puzzlehunts for Cambridge computer scientists and the wider STEM community.
- Used a CP-SAT solver to organise study groups for undergraduates, taking into account factors such as availability, college and study preferences.

Cambridge University Cyber Security Society (CyberSoc)

PRESIDENT *2021 - 2022*

- Spearheaded the revival of CyberSoc, Cambridge's cybersecurity society, after a 2 year hiatus.
- Helped to organise events and workshops for beginners to CTFs.
- Revived and grew cheriPI (Cambridge's CTF team), which now regularly qualifies for final rounds of national and international CTFs.

University of Cambridge Competitive Programming Society (UCCPS)

JUNIOR TREASURER, CO-FOUNDER *2020 - 2021*

- Managed society finances and administrative tools, and helped to develop the society website.
- Helped to organise an introductory competitive programming workshop and contest series (with over 100 participants), and a lecture series with well-known speakers from competitive programming.