

James Matthew Uygongco Young

matsagad[at]gmail[dot]com ◉  matsagad.com ◉  matsagad ◉  james-matthew-young

EDUCATION

Imperial College London October 2023 - September 2024
MRes Bioinformatics and Theoretical Systems Biology London, United Kingdom

- Distinction Honours

Imperial College London October 2020 - July 2023
BEng Mathematics and Computer Science London, United Kingdom

- Overall: 80% — First Class and Final Year Dean's List Honours

Relevant Modules: Applied Probability, Math Biology, Statistical Modelling, Multivariable Calculus & Differential Equations, Intro to Machine Learning, Logic-Based Learning, Computational Neuroscience, Symbolic Reasoning.

United World College of South East Asia August 2018 - May 2020
International Baccalaureate Diploma Tampines, Singapore

- Total: 44/45 points — 7s in HL Math, Further Maths, and Physics.

WORK EXPERIENCE

Illumina July - October 2023
Bioinformatics Scientist Intern - Genome Informatics Team (R&D) Cambridge, United Kingdom

- Developed a de-novo pipeline for the automatic correction of haplotype switch errors in phased diploid assemblies by leveraging trio data. Found a high precision on switched genotype variants against benchmarks.



Imperial College London October 2022 - March 2023
Personal Math Tutor and Undergraduate Teaching Assistant London, United Kingdom

- Led weekly tutorial sessions covering discrete maths, logic, and algorithms for first-year computing students.

Reckon Digital March - September 2022
Junior Software Engineer London, United Kingdom

- Contributed to new and existing client applications centred around the React and Django frameworks.



RESEARCH PROJECTS

Diffusion Posterior Sampling via Sequential Monte Carlo for Zero-Shot Scaffolding of Protein Motifs
Report  ◉ Code  ◉ Supervised by Dr O. Deniz Akyildiz. April - August 2024


- Formalised single-motif, multi-motif, and symmetric motif scaffolding as inverse problems and implemented diffusion posterior samplers to solve them with an unconditional protein diffusion model acting as a prior.
- Engineered a unified interface for scaffolding experiments with customisable configs.

Tackling the Inverse Problem of Turing Patterns with Modern Hopfield Networks
Supervised by Prof Robert Endres. January - March 2024

- Created joint embeddings for Turing patterns and their parameters through a mixture-of-experts multi-modal variational autoencoder. Inferred parameters of unseen patterns by coupling the pattern encoder and parameter decoder with a modern Hopfield network.

Evolving Promoter Architectures for Encoding Extracellular Information within Budding Yeast
Report  ◉ Code  ◉ Supervised by Dr Vahid Shahrezaei, Dr Peter Swain. December 2022 - June 2023

- Optimised variable-sized time-inhomogeneous continuous-time Markov chains via a Quality-Diversity approach on an unstructured archive defined by distance metrics on graph topology and simulated trajectories.
- Proposed a stochastic simulation using matrix exponentials paired with an SVM decoder to form an objective function estimating the mutual information between a model's gene activity and environmental stress.
- Inferred some structural properties of the combinatoric transcription factor signalling process within yeast.

Paired Comparison Models on Tennis Matches
Report  ◉ Supervised by Prof Axel Gandy. April - May 2022

- Reviewed the Bradley-Terry (BT) model, Elo, Glicko, Glicko2, and a proposed Markov chain model and evaluated them on men's ATP tennis data through ranking, likelihood, and over-fitting measures.
- Extended the BT model to gauge the effects of momentum between sets and experimented with Bayesian priors to address disconnected comparison graphs. Produced simulations and graphical aids in R.

ENGINEERING PROJECTS

XFold: Building The AlphaFold2 Inference Pipeline from Scratch

Code 

November 2024

- Wrote the AlphaFold2 inference pipeline in PyTorch. Components include various steps for embedding evolutionary or recycled information, the evoformer stack, and the structure module.

Histree: A Complex Genealogy Visualisation Tool

Code  ◦ *Supervised by Dr Anthony Field.*

October - November 2022

- Developed the Flask backend responsible for constructing and caching family tree networks and content-rich nodes. Efficiently fetched entities from WikiData and MediaWiki APIs through the SPARQL query language.

WACC Language Compiler


February - March 2022

- Implemented semantic analysis and ARM11 assembly code generation for a while-like language.
- Built with Kotlin on top of ANTLR under a continuous integration and development workflow.

Pintos Operating System

October - November 2021

- Set up priority scheduling, multilevel feedback queue scheduling, and an interface for running user programs.
- Built with C from an incomplete code base riddled with concurrency edge cases.

A few more of my projects can be found at github.com/matsagad  and matsagad.com/logs .

TECHNICAL SKILLS

Languages Experienced with Python. Comfortable with R, TS/JS, Kotlin, and Java.
Worked with C, Haskell, Prolog, SMT-LIBv2, and Z3.

Technologies Comfortable with Git, shell scripting, and interfacing with HPC/GPU clusters.
Often use PyTorch and accompanying libraries for ML engineering.

MISCELLANEOUS

Philippine National Records for Solving a Rubik's Cube in Fewest Moves (24 single, 28.67 average) 

DoC Haskell L-Systems Graphics Contest 2020 (First Place) 

Bloomberg BPuzzled: Global Finals 2023 (First Place in Europe); at Imperial College 2022, 2023 (First Places)

INTERESTS

De Novo Protein Design Generative Models Deep Learning Game Theory Puzzles