Alun Stokes

An M.Sc. student studying number theory at

McMaster University
1280 Main St. W.
Hamilton, ON, L8S 4L7
Canada

My research primarily concerns two topics: the 4-category equivalence of which most people know the dessin d'enfant, and the expressability of the double parameterisation of discrete metric spaces by their additive and (this is the novel bit) multiplicative hyperboolicity.

The first topic, supposing it has an answer to the question of the action of the absolute Galois group, would change the landscape of arithmetic as we understand it, full stop. My work is both computational and theoretical in this domain.

The second, as present, is just kind of interesting. However, how we tell whether we can or have found the idal embedding of a metric space so as to analytically say anything meaningful is a problem that then becomes **very** interesting, espcially in the contexct of our parameterisation.

My specialty is in addressing these traditionally analytic and algebraic problems not only as they come, but by using methods in numerical mathematics to solve otherwise intractible problems, and then regenerating exact solutions from approximations.

Personal Details

Highest Level of Education: Bachelor's Degree
Total years of experience: 5

Work Experience

Graduate Research and Teaching Assistant

McMaster University - Hamilton, ON
September 2021 to Present

Listed in order of recency, with double line breaks delineating semesters in which the appointments occurred. All positions were undertaken at McMaster University.

Graduate Topics in Risk Management (MFM 763)
Introductory Number Theory (MATH 3H03)

Numerical Linear Algebra (MATH 3NA3)
Introductory Linear Algebra (MATH 1B03)

My research focuses on both the computational and theoretical aspects of dessins d'enfants and their associated categorical equivalences, so that we may better understand the structure of the full absolute Galois group of the rational numbers. Current work is on investigating and characterizing parameter sets for which families of dessins are given rise to by group actions on particular projective spaces. Ultimately,

we intend to characterize the moduli spaces to which they correspond in a meaningful way. Supported by internal research grants, an OGS, and departmental funding. Supervised by Dr Cameron Franc.

Research Assistant

McMaster University - Hamilton, ON
May 2021 to August 2021

Investigated the use of standard machine learning techniques in determining the membership of matrices in non-congruence subgroups of the modular group, and identifying these groups in general. Morphed into a project focused on developing very fast homotopy continuation software in Julia in order to determine Belyi maps from dessins d'enfants with specific ramification data and normalizations determined. Compared to the established databases of degree 10 maps, we were able to compute maps of degree exceeding 20. Supported by an NSERC USRA Grant. Supervised by Dr Cameron Franc.

Undergraduate Teaching Assistant

McMaster University - Hamilton, ON
January 2021 to April 2021

Introduction to Discrete Mathematics (COMPSCI 1DM3)

Data Scientist

Statistics Canada Consumer Prices Division - Ottawa, ON
June 2020 to August 2020

Here I built-out various NLP methodologies for the purpose of mapping hierarchical item-classification data structures between one another, based on semantic descriptions of products in each category. The ultimate purpose of all this being to increase the degree of automation available for use in calculating the consumer-price-index (CPI), such that it can be faster and more accurate, also requiring many fewer human-hours spent cataloguing. I managed to report a greater accuracy than previously possible in the literature (which to be fair, was pretty minimal) using a hybrid strategy of skip-gram embeddings of a product's descriptional lineage followed by gradient-boosted decision trees for pairwise parental prediction to reconstruct a new hierarchy of product categories. This was supervised by Serge Goussev, MSc.

Research Assistant

McMaster University - Hamilton, ON
May 2020 to August 2020

Investigated the use of various strategies, including machine learning and more standard analytic and multi-dimensional scaling techniques to reduce embedding roughness and distortion of finite metric spaces (in the forms of finite graphs) into L^p spaces of various dimension. Surprisingly effective results were observed when using graph convolutional network with novel (but sparse) feature vectors to predict the Gromov hyperbolicities of both Erdos-Renyi and Barabasi-Albert graphs at a range of scales. The work was support by a Stewart Grant from McMaster University. This was supervised primarily by Dr George Dragomir (Columbia University), in addition to Dr Andy Nicas (McMaster University). Work continued with the collaboration with Dr David Rosenthal (St. John's University).

Research Assistant

McMaster University - Hamilton, ON
May 2019 to May 2020

Used both standard computer-vision segmentation techniques and then more recent neural-network-based strategies to segment and classify breast tumour (or non-tumour) photoacoustic tomographic

images, for the purposes of increasing diagnostic accuracy and speed-as the device was to be handheld and used intrasurgically to improve resurgence outcomes. In the course of this work, I developed a robust framework for batching off thousands of jobs to Compute Canada servers (allowing me far greater experimentation than the master's student I replaced) as well as novel network designs. Now sometimes used, my architecture made use of stacked U-net models with skip -connections between each pair of corresponding layers in each U. In conjunction with a special auto encoder loss, this model well outperformed standard segmentation models, even after significant transfer learning. This was performed in conjugation with a laboratory at Western university, and supervised by Dr Ned Nedialkov (McMaster University).

Education

Master's Degree in Pure Mathematics (Number Theory)

McMaster University - Hamilton, ON

September 2021 to April 2023

Honours Bachelor's Degree of Integrated Sciences in Mathematics & Statistics

McMaster University - Hamilton, ON

September 2017 to April 2021

International Baccalaureate/Secondary School in Higher Level: World Literature, Biology, Chemistry; Standard Level: French, Economics, Mathematics

Turner Fenton Secondardy School - Brampton, ON

September 2013 to June 2017

Skills

- Python (7 years)
- Julia (2 years)
- PyTorch (3 years)
- CUDA (2 years)
- GNU/Linux (6 years)
- JavaScript (4 years)
- LaTeX (4 years)
- SageMath (3 years)
- MATLAB (2 years)
- Macaulay2 (2 years)
- Java (5 years)
- Distributed systems (2 years)
- SQL (3 years)
- AWS (1 year)
- Git (5 years)
- Natural language processing (2 years)

the-art piece-wise matching exceeding 98% was achieved for as small as we used for reconstruction, making the work highly extensible as a drop-in for other piece-wise matching strategies.

Dean’s Honour List (4 years)
April 2021

Awarded all 4 years of undergraduate education for particularly meritorious GPA in each 3 consecutive trimesters per year.

1st Place - CANDEV Data Challenge
January 2020

Among 88 teams, and across two presentations on the topic, we won first place in a Government of Canada competition to solve legitimate problems faced by various departments. We used tuned a transformer-based model to cluster and identify redundancies in course offerings by the Canadian School of Public Service, using t-SNE for MDS, and then building a tool around the program to integrate its usage into suggestions for when new courses were being proposed.

McMaster Steward Award
May 2020

Awarded for a study into the use of quasi-hyperbolic embeddings into L^p spaces, as mentioned in the project above. Worth of $5,000.

McMaster President's Award
September 2017

One-time entrance scholarship of $2,500 for secondary school final GPA above 95%.

Groups

American Mathematical Society: Providence, Rhode Island, US
September 2021 to Present

GRID: grid.298859.7
American Mathematical Society: Providence, Rhode Island, US http://www.ams.org/

Society for Industrial and Applied Mathematics (SIAM)
July 2022 to Present

GRID grid.294116.9
Society for Industrial and Applied Mathematics, Philadelphia, US http://www.siam.org/

Publications

A Minimal-Input Multilayer Perceptron for Predicting Drug-Drug Interactions
https://journal.stemfellowship.org/doi/10.17975/sfj-2020-006
September 2020

Predictive models in the drug discovery industry have an essential role that cannot be understated. With the sheer volume of potentially useful compounds that are considered for use, it is becoming more difficult to investigate overlapping interactions between two drugs. Given that recreational drugs lack the rigorous warnings of prescription drugs, it is important for the layperson to know which drugs

can and cannot mix. Other methods are necessary to bridge this knowledge gap in the absence of deterministic experimental results for every drug combination. Ideally, such methods would require minimal inputs, have high accuracy, and be computationally feasible. We have not encountered a model that meets all the above criteria. In light of this, we propose a minimal-input multi-layer perceptron that predicts the interactions between two drugs. This model has a great advantage of requiring no structural knowledge of the molecules in question, and instead only uses experimentally accessible chemical and physical properties; in particular, 20 per compound. These 20 were selected as they seemed they may be pertinent to the authors, although given this was an exploratory model, there was not a rigorous method behind the choice. Using a set of known drug-drug interactions and associated properties of the drugs involved, we trained our model on a dataset of about 650,000 entries. We report an accuracy of 0.968 on unseen samples of interactions between drugs on which the model was trained and an accuracy of 0.942 on unseen samples of interactions between unseen drugs. We believe this to be a promising and highly extensible model that has the potential for high generalized predictive accuracy with further tuning.

Automatically Solving Square-Piece Jigsaw Puzzles using Convolutional Neural Networks with Gradient Boosted Decision Trees

https://gua.soutron.net/Portal/Default/en-GB/RecordView/Index/61

December 2020

In the square jigsaw puzzle problem, one must reconstruct a complete image from a set of unordered, square, non-overlapping puzzle pieces. When pieces are non-square, this problem can be solved with O(n^2) time complexity - but the specification of square pieces makes this task NP-hard. With no implicit matching criterion in the form of nodules and matching holes on the sides of pieces, this becomes a question of identifying continuity between parts of an image. Here, a fully automated solver is proposed, which makes no use of clues, anchors, or knowledge of the constructed image - which are each often required to adeptly solve puzzles computationally. To do this, a greedy solver is implemented to make placement decisions based on a compatibility criterion given by pair-wise determination by a convolutional neural network (CNN) with gradient boosted decision trees (GBDT). This compatibility criterion is calculated by using a CNN to perform feature extraction on a set of paired pieces, the output of which is then classified using a GBDT. A pairwise accuracy of 0.983 is achieved on 32x32 pixel pieces. Given the piece-wise solving method, this can be applied to arbitrarily large puzzles, both in number of pieces and pixel size, with the accuracy of the solver determined as a function of the number of pieces.

Additional Information

Note that, despite the slew of languages with which I have experience, my primary interests are in Python (and its libraries), and in particular, Julia. While I use GNU/Linux (primarily Ubuntu et. derivatives) most often, I am very comfortable with both MacOS and Windows (<11), in that order. All this said, I am rarely uncomfortable learning even unfamiliar technologies and softwares.