School of Computer, Data and Mathematical Sciences

WESTERN SYDNEY UNIVERSITY

MATH3016 Mathematics Project

Potential Projects and Supervisors – Spring 2022

(as of July 21, 2022)

- Projects will normally be done in groups of two students.
- Project allocation is made by the subject coordinator in consultation with students and supervisors.
- Some projects may be possible to do jointly with a partner enrolled in Discovery Project, subject to agreement from the Supervisor and Subject Coordinators!

Algebra and Topology

1 What is Lattice theory?

Supervisor: Prof Roozbeh Hazrat (r.hazrat@westernsydney.edu.au) [PTA]

In many areas of mathematics, the relationship between two objects is of vital importance. This relation would also give a certain "order" to the subject in hand. One can then compare the objects via the order, partition and rearrange them. This is where the notion of lattice comes in. Lattice theory provides a unifying framework for unrelated developments in many mathematical disciplines, algebra, topology, geometry, etc.... It is a very abstract (and yet simple and beautiful) language to capture relations and orders. In this project we will try to understand the basics of lattice theory and apply it to areas we are familiar with.

Prerequisites:
- MATH1006 Discrete Mathematics
- MATH3001 Abstract Algebra
- Interest in abstraction and formal proofs

References:
- Priestley, H. A. *Introduction to Lattices and Order*, Cambridge University
- Nation, J. B., *Notes on Lattice Theory.* Available online.

2 Infinite symmetric groups

Supervisor: A/Prof James East (j.east@westernsydney.edu.au) [PTA]

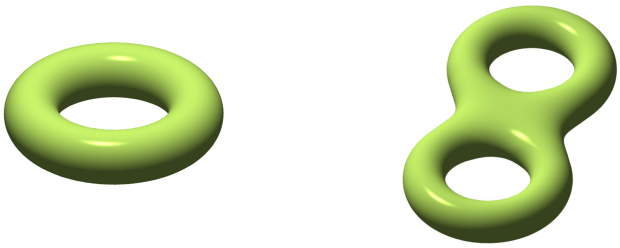
This project looks at symmetric groups over infinite base sets. The behaviour of these groups is very different from the finite case, and incorporates famous paradoxes from set theory. The project will focus in particular on properties of generating sets. For example, we'll prove famous results such as: (1) the length function with respect to any generating set is bounded (Bergman), and (2) any countable set of permutations is contained in a 2-generated subgroup (Galvin).

Prerequisites:
- MATH1006 Discrete Mathematics
- MATH3001 Abstract Algebra
- Interest in abstraction and formal proofs

3 What is Topology?

Supervisor: Prof Roozbeh Hazrat (r.hazrat@westernsydney.edu.au) [PTA]

Topology is about studying shapes and their deformations. For example an object with one hole is not similar to an object with two holes (see the pictures below).



In this project we will look at this branch of mathematics. We start by studying metric spaces which give us examples of topological spaces. We then study the abstract definition of a topological space and investigate the notions of "topological equivalence" of shapes (homeomorphism), compactness and connectedness of topological spaces, and so on.

Topology appears in many subjects outside mathematics such as physics, robotics and biology.

In this project we will look at this branch of mathematics. We study when two shapes are similar (homeomorphism), when they are compact, connected and so on.

Prerequisites:
- MATH1014 Mathematics 1A
- **Interest in abstraction and formal proofs**

4 From *Mad Vet Puzzles* to Algebra and Back

Supervisor: Prof Roozbeh Hazrat (r.hazrat@westernsydney.edu.au) [PTA]

This project demonstrates how mathematics can be used to explore a seemingly non-mathematical problem.

There is an interesting family of puzzles called *Mad Vet puzzles* which pose questions on how a collection of animals can be transformed by the Mad Vet machines into other collections.

For example, assume that there are three machines:

- Machine 1 converts a cat into two dogs and five mice (and vice versa).
- Machine 2 converts a dog into three cats and three mice (and vice versa).
- Machine 3 converts a mouse into a cat and a dog (and vice versa).

You can ask questions like the following:

1. Can you convert one cat into seven mice?
2. Can you convert one cat into a pack of dogs, with no mice or cats left over?

In this project, we explore puzzles of this kind, and along the way we will use (and learn) some algebra.

One of the goals is to illustrate a practice that is common in mathematics: namely, answering a question in one area by recasting it in another area, answering the recast question there, and then using that result to answer the original question.

Prerequisites:
- MATH1006 Discrete Mathematics
- MATH3001 Abstract Algebra
- **Interest in abstraction and formal proofs**

5 Semigroups in phylogenetics

Supervisor: Chad Clark (with Andrew Francis) (chad.clark@westernsydney.edu.au) [PTA]

This project will explore a new way to represent phylogenetic trees ("trees of life") as elements of a semigroup based on so-called Brauer diagrams. This makes it possible to do interesting things like multiply trees together and study whether semigroup results can have some meaning in trees. In this project you'll need to learn a little about phylogenetic trees, and more about semigroups. There may be opportunity to explore some un-answered questions, for instance, what trees are idempotents?

Reference:

- Andrew Francis and Peter D Jarvis, Brauer and partition diagram models for phylogenetic trees and forests, *Proceedings of the Royal Society A*, 2022, 478: 20220044

Prerequisites:
- MATH3001 Abstract Algebra
- **Interest in abstraction and formal proofs**

Applied Statistics

6 Impacts of the COVID pandemic on regional internal migration

Supervisor: Dr. Russell Thomson (russell.thomson@westernsydney.edu.au) [on-line]

The project will involve downloading regional internal migration estimates from the ABS in 2021 and 2019 in regions around Sydney. Regression Models will be fit to looking at associations between migration and distance from population centres, income, distance to beaches, etc. The project should incorporate some data visualisation.

Prerequisites:
- At least one statistical subject.
- Experience in programming and the use of R statistical software.

7 A meta-analysis of marine ecology data

Supervisor: Dr. Russell Thomson (russell.thomson@westernsydney.edu.au) [on-line]

The students will examine the meta-analysis technique and apply to actual (already collected) data. Study goals: (1) – what is the overall affect, across all the studies , (2) – are the study results heterogeneous and (3) – is there an aspect of each study that affects the results (meta-regression).

Prerequisites:
- At least one statistical subject.
- Experience in programming and the use of R statistical software.

Mathematics in Data Science

8 Splines and Deep Neural Networks

Supervisor: Prof Paul Hurley (p.hurley@westernsydney.edu.au) [PTA]

A relationship has been established between max-affine spline operators and deep neural networks:

- R. Balestriero and R. G. Baraniuk, "A Spline Theory of Deep Networks", https://proceedings.mlr.press/v80/balestriero18b/balestriero18b.pdf
- A talk on the topic: https://www.youtube.com/watch?v=5AMxhdj-96Q

The goal of the project is to understand and reinterpret this work, and ideally to demonstrate it through examples in a Jupyter notebook written in Julia or Python.

Prerequisites:
- MATH2010 Linear Algebra
- MATH3003 Analysis
- Programming skills in python will be helpful.

9 Testing and understanding jackknife approach to mutual information estimation

Supervisor: Prof Paul Hurley (p.hurley@westernsydney.edu.au) [PTA]

The paper Jackknife approach to the estimation of mutual information https://www.pnas.org/doi/10.1073/pnas.1715593115, describes a method to calculate the mutual information, a measure of how two random variables are dependent on each other.

The goal of this project is to understand this technique and implement it in Julia or python.

Prerequisites:
- Programming skills in python or Julia will be helpful.

10 The robustness of the BERT (transformer) model

Supervisor: A/Prof Yi Guo (y.guo@westernsydney.edu.au) [PTA]

BERT is the state-of-the-art language model at the moment widely used by tech giants such as Google, Facebook and so on. There are many automated document processing systems are built on BERT. However, when these systems are built on some specific purpose, say screening/selection, they can be attacked by varying a few words to gain advantages. This is called adversarial attack. A model robust to these attacks is desirable. The goal of the project is to understand the robustness of BERT in terms of classification, i.e. how stable it is if the input changes at a few places, say a couple of words are changed, and also visualise the robustness via gradient with respect to input.

Key references:

- Attention Is All You Need: https://arxiv.org/abs/1706.03762
- BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding: https://arxiv.org/pdf/1810.04805.pdf and
- code: https://github.com/google-research/bert
- Certified Robustness to Text Adversarial Attacks by Randomized [MASK]: https://arxiv.org/abs/2105.03743

Prerequisites:

- MATH2010 Linear Algebra
- Programming skills in python will be essential.

11 Universal approximation of the transformer models, full model and sparse model

Supervisor: A/Prof Yi Guo (y.guo@westernsydney.edu.au) [PTA]

Transformer is a new model proposed in 2017 for language modelling. In essence, it is a sequence to sequence (S2S) approximation, mapping a sequence of N vectors to another sequence of N vectors. So it is a functional, providing a family of such mappings. The question is how well it approximates a given S2S mapping in the form of data? There are several papers discussing this so-called universal approximation property of transformers. The goal of the project is to understand the theory behind the proofs of universal approximation property of transformer models and explore the possibilities of accommodating to extra long sequences, e..g a sequence of vectors more than 100k.

Key references:

- Attention Is All You Need: https://arxiv.org/abs/1706.03762
- Are Transformers universal approximators of sequence-to-sequence functions: https://chulheey.mit.edu/wp-content/uploads/sites/12/2019/12/yun2019transformers.pdf
- O(n) Connections are Expressive Enough: Universal Approximability of Sparse Transformers: https://proceedings.neurips.cc/paper/2020/file/9ed27554c893b5bad850a422c3538c15-Paper.pdf

Prerequisites:

- MATH2010 Linear Algebra
- Programming skills in python will be helpful.

12 Key words detection in audio signal

Supervisor: A/Prof Yi Guo (y.guo@westernsydney.edu.au) [PTA]

The goal of this project is to detect specified keywords, for example, car, bus, truck etc. in audio signal (waveform) ideally in real time, and count the occurrences of each word. We will use wav2vec2 for this project and start with the Facebook implementation provided below.

- Attention Is All You Need: https://arxiv.org/abs/1706.03762
- wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations: https://arxiv.org/abs/2006.11477
- Code: https://huggingface.co/anton-l/wav2vec2-base-keyword-spotting

Prerequisites:

- COMP2025 Introduction to Data Science
- MATH3007 Predictive Modelling
- Programming skills in python will be essential.