2019 Conference for Mathematics, Statistics and Computer Science

Dalhousie University, Halifax, NS – October 25th-27th, 2019

Platinum Sponsors:

CGI https://www.cgi.com

AARMS https://aarms.math.ca/

Faculty of Computer Science,
Dalhousie University
https://www.dal.ca/

Faculty of Science,
Dalhousie University
https://www.dal.ca/

Gold Sponsor:

ShiftKey Labs,
Dalhousie University,
https://shiftkeylabs.ca/

Silver Sponsor:

Morneau Shepell Ltd.
https://morneaushepell.com

Talks Schedule

	SESSION 1	SESSION 2	SESSION 3	SESSION 4
9:05-9:20am	Frédéric LeBlanc	Colin Vibert	Jennifer McNichol	Sarah Li
9:25-9:40am	William Taylor-Melanson	Joy Liu	Abby Anderson	Everett Patterson
9:45-10:00am	Nicholas Barreyre	Ben Wang	Samantha Bardwell	Adam Lucas
10:05-10:20am	Donné D'Arnall	John Marcoux	Lucas MacQuarrie	Dylan Ruth
1:35-1:50pm	Yong Yu	Brady Ryan	Ellen McCole	Charles Gerrie
1:55-2:10pm	Sarah Park	Kieran Bhaskara	Lawrence Daniel Doucett	Samuel Bauer
2:15-2:30pm	Kethireni Ajith Kumar Veera Raghavan	Jesse Preston	Graeme Zinck	Andrew Fraser
2:35-2:50pm	Maryam Tajjeddin	Emily Wright	Syed Zeeshan Ahmed	Uyen Dao
4:45-5:00pm	Samia Sifat	Caleb Jones	Mathieu Briedeau	Adam Smith
5:05-5:20pm	Deepkumar Shah	Justin Hughes	Simon Gauvin	Youssef Zaazou
5:25-5:40pm	Hardik Manek	Aaron Dwyer		Jeremy Peters

Table of Contents

Session 1

9:05- 9:20am	Fredéric LeBlanc **Bounding the Complexity of Classes of Decision Trees**
9:25- 9:40am	William Taylor-Melanson **A Wearable System for Gait Phase Prediction and Interactive Movement Feedback**
9:45- 10:00am	Nicholas Barreyre **Bitwise Conditional Controls Over Language Models for Music and Text**
10:05- 10:20am	Donné D'Arnall **Crop and Weed Stem Classification Using Recurrent Neural Networks**
1:35- 1:50pm	Yong Yu **Alignment Historical Aerial Images using Machine Learning**
1:55- 2:10pm	Sarah Park **Classification of Engraved Illustrations using a Statistical Machine Learning Approach**
2:15- 2:30pm	Kethireni Ajith Kumar Veera Raghavan **Predict the Donor Journey Using Deep Learning Models**
2:35- 2:50pm	Maryam Tajjeddin **Predicting Risk of Aggressive Responsive Behaviours among People Suffering from Dementia using Natural Language Processing (NLP) and Machine Learning (ML)**
4:45- 5:00pm	Samia Sifat **Facial Expression Recognition and Morphing with Machine Learning**
5:05- 5:20pm	Deepkumar Shah **Using Machine Learning in Cannabis Industry**
5:25- 5:40pm	Hardik Manek **Credit Card Fraud Detection**

Bounding the Complexity of Classes of Decision Trees

Presenter: LeBlanc, Fréderic
Université de Moncton
Supervisor: Marchand, Mario
Other authors: Leboeuf, Jean-Samuel
Department: Math

Supervised machine learning, in the context of data classification, is concerned with learning classifiers that predict labels well from a set of labeled data. The performance of classifiers is defined by their risk, which is the probability that they make a prediction error on data sampled from a fixed but unknown distribution. This risk can be bounded for classifiers taken from hypothesis classes with low complexity, as measured by the growth function or VC/Natarajan dimension. Decision trees are a kind of classifier that has the advantage of being easy to interpret by human experts; however, they tend to overfit the data. To address this problem, it is customary to prune decision trees, and we would like to prune using risk bounds giving a tradeoff between empirical risk on the training set, and complexity of the decision tree before and after pruning. In order to achieve this, we obtained an upper bound on the growth function of binary decision trees with continuous features which is tight for decision stumps, and defined recursively for more complex trees. Furthermore, we give an asymptotic characterization of the Natarajan dimension of classes of decision trees.

A Wearable System for Gait Phase Prediction and Interactive Movement Feedback

Presenter: Taylor-Melanson, William
University of Prince Edward Island
Supervisor: Godbout, Andrew
Other authors:
Department: Math

We present an interactive wearable system which operates as a function of the wearer. This system tracks continuous phases of a movement in real-time using a recurrent neural network. As part of this work, we present a wearable prototype that implements our design. The prototype is capable of tracking one's gait and producing interactive vibrational haptic feedback synchronized with one's walking strides. We deploy the system in a pilot study of seven walking subjects who interacted with the system. As the subjects walked on a treadmill, they received haptic vibrational stimulation synchronized with their walking strides. We outline a number of additional feedback modalities that are possible with our architecture. Further, we outline our data labeling system to produce data sets for supervised learning and present the recurrent neural network architecture that drives our system. Continuous phase tracking contrasts with other comparable systems that identify discretized phases or temporal landmarks within a movement. When compared to other recent attempts to continuously track the phase of walking movements, our system produces superior performance results with smaller training data sets. For our pilot study, our system achieves real-time operation without having seen data from any of our seven subjects.

Bitwise Conditional Controls Over Language Models for Music and Text

Presenter: Barreyre, Nicholas
Dalhouie Unviversity
Supervisor: Oore Sageev
Authors:
Barreyre, Nicholas *, **Meade, Nicholas ***, **Oore, Sageev**, **Keselj, Vlado**, **Lowe, Scott C.**
*** Equal contribution**
Department: Math

Neural networks are effective language models. In particular, Long Short-Term Memory Recurrent Neural Networks can learn to generate sequences of words. When doing so, however, it can be hard to control the model after the training procedure has finished. To address this, we explore how conditioning a neural network can influence the generation process. Using data from two domains: English text and classical piano performance (an event-based representation), we demonstrate the effects of control signals. A control signal is an interpretable vector that is concatenated to each example. It provides information about the target during training. For example, knowing the composer of a piano performance gives information about the probability distribution over the next possible musical events. This vector allows a user to control one or more conditions of the probability distribution at test time. The user might indicate that they are sampling from the neural network based on the condition that the composer is Bach in one circumstance and that the composer is Rachmaninoff in another. We demonstrate several examples where a small (i.e. few-bit) control signal can be used to effectively apply simultaneously large and intricate effects on the generated output.

Crop and Weed Stem Classification Using Recurrent Neural Networks

Presenter: D'Arnall, Donné
Acadia University
Supervisor: Mcintyre, Andy and Chipman, Hugh
Other authors:
Department: Math

Weeds reduce crop productivity and thus the profitability of farms. Currently, weeding is done through costly human labor or environmentally harmful herbicides. An alternative is to have a robot autonomously drive along crop rows identifying weeds and removing them. The goal of this project is to use image segmentation to identify stem emergence points from images captured by two high-resolution cameras mounted on the robot. When the emergence point of a weed is successfully identified, the robot can deploy the picking arm and pull the weed, then move on to find the next candidate point for removal. Accurate discrimination between weeds and crops, in addition to the identification of the precise locations where weed stems emerge from the soil, is critical to the weeding task. Of particular interest to the present study is the investigation into the utility of image sequences over time as opposed to individual snapshots when attempting to identify weeds and their emergence points. Convolutional and recurrent neural network architectures will be trained to detect weed stem origins using a large database of hand-labeled sugar beet crop imagery.

Alignment Historical Aerial Images using Machine Learning

Presenter: YU, Yong
University of Prince Edward Island
Supervisor: Godbout, Andrew
Other authors: Godbout, Andrew
Department: Math

This work proposes a solution to challenges with aligning historical aerial photographs, either with other historical aerial photographs or with modern satellite images. For example, we take an aerial photograph of Charlottetown, Prince Edward Island from 1935, and want to align the photo with a modern satellite photo of Charlottetown from 2019. There are many uses for aligning such images, which include monitoring shoreline erosion, comparing landscapes and even assessing light pollution. Typically, the solution for image alignment involves finding key feature points using SIFT and then aligning feature points across multiple images using RANSAC. We propose implementing a machine learning approach. The solution is inspired by recent developments in convolutional neural network architectures for geometric matching. We create a labeled dataset generated by selecting google map satellite images and transforming them slightly. For example, rotating, translating or warping the image. The labels for our supervised machine learning task become the transformation (homography matrix) we apply to one image to get the other. Given the two images (original and transformed) use a convolutional neural network architecture to learn the homography between them. We will train the system with a large algorithmically built dataset of satellite images. We will test the system on unseen satellite images and also historical aerial photographs. The solution should provide a computational speed up on traditional approaches and allow us use the knowledge learned from analyzing modern satellite images applied to historic aerial photographs.

Classification of Engraved Illustrations using a Statistical Machine Learning Approach

Presenter: Park, Sarah
Mount Allison University
Supervisor: Cormier, Michael
Other authors:
Department: Math

Many previous classification systems have aimed to classify images based on the contents of a scene. Classification based on image formation processes requires finer distinctions between image classes. Our current approach aimed to classify engraved illustrations based on the method in which they were created; as either woodcut or copperplate engravings. Woodcut engravings are created by cutting away the portions of an image that is unwanted, and the ink is carried by and transferred by the raised edges. Copperplate engravings are created by incising fine lines into a copper plate and these lines transfer the ink. Histograms of the magnitudes of the image gradients were used as features for classification; the gradient properties were expected to distinguish the underlying differences between two classes based on their formation processes. A Bayesian approach was chosen as it allows for degrees of uncertainty to be established, has an intuitive interpretation, and can be thresholded to produce a hard classification. Currently, our results have been very positive, and our system has revealed higher accuracy for classifying the illustrations as

copperplate or woodcuts than for distinguishing other content-based classes. This suggests that the use of gradient magnitudes intrinsically distinguishes and classifies engraved illustrations better than the contents of a scene.

Predict The Donor Journey Using Deep Learning Models

Presenter: Veera Raghavan, Kethineni Ajith Kumar
Acadia University
Supervisor: Lee, Greg
Other authors:
Department: Math

Fundmetric is Halifax-based company that uses machine learning models to predict the behavior of donors which helps the charities to identify, connect with and communicate with their donor. Fundmetric clients send email to donors through Fundmetric regarding the client,Äôs current issues or appeals (constituting an ‚Äòaction‚Äô) and the donor takes a sequence of actions which hopefully leads to a donation. But what if we could suggest the next action a client should take by consulting a deep learning model? Can we suggest actions that lead the donor to make larger donations than they previously have?
The data that we are dealing with is time-series, making RNN (Recurrent Neural Networks) / LSTM (Long Short-Term Memory) models the appropriate choice of machine learning algorithm to train for this problem. Our goals are to predict if a) a constituent will make donation or not and b) how much a constituent is likely to give. In order to suggest actions for the charity to take, we look at the previous five actions of a non-donor, and suggest the action most likely to lead to a donation. In other words, the model predicts the best sequence of actions taken by the donor which leads to higher donations.

Predicting Risk of Aggressive Responsive Behaviours among People Suffering from Dementia using Natural Language Processing (NLP) and Machine Learning (ML).

Presenter: Tajeddin, Maryam
Acadia University
Supervisor: Silver, Daniel
Other authors:
Department: Math

Patients with dementia will eventually experience significant loss of cognitive function. Many will have difficulty properly communicating life,Äôs challenges and instead become agitated, resulting in verbal or physical aggression. Monitoring the risk of a resident harming themselves or others due to aggressive behavior is a priority within a long-term care facility where dementia is present. Caregivers at Shannex regularly record resident health and behaviour using computing systems. Each of these systems digitally record information as either structured data or unstructured text, providing an on-going log of each resident,Äôs patient history. The objective of this project is to use natural language processing (NLP) and machine learning (ML) techniques to develop models that can predict the probability of resident exhibiting aggressive behaviours that may harm themselves or others within the next week. NLP techniques have been used to extract key features of a person,Äôs behaviour or social interactions from text such as emails, tweets, and Facebook entries. These features have then been used to build ML models that predict the

person,Äôs sentiment, propensity to purchase, and likely next action. We propose that it is possible to develop similar ML models to predict the probability of a resident,Äôs aggressive behaviour based on structured and unstructured data that Shannex has in their computing systems.

Facial Expression Recognition and Morphing with Machine Learning

Presenter: Sifat, Samia
Acadia University
Supervisor: Silver, Danny L
Other authors:
Department: Math

Automatic facial expression recognition and facial morphing with different expressions has been an area of research in the field of machine learning and computer vision for decades. This research area has proven to be a powerful tool for visual effects in social media, image transformation software. Many models have been proposed to simulate the transformation of facial expressions. We propose to design and implement an automatic web-based system for facial expression recognition and facial morphing with machine learning with possible transfer learning. The website is able to capture photo in a desktop, laptop or phone and automatically detect a face, capture and preprocess that portion of the image for predictive analysis, then send it to machine learning model. The neural network models classify the facial image expression into happiness, sadness, surprise, anger, fear, disgust and neutral. The deep learning network models also transform the given facial image into one of up to six other expressive states, one of: happiness, sadness, surprise, anger, disgust, fear, and neutral as requested by the user and send back the modified face to the website. An independent set of test image accuracy, precision, recall and F-measure is used as a measure of effectiveness for image classification. The image morphing performance is measured by the mean squared error (MSE) between pixels of an actual image of a person in the desired emotional state versus the predicted image.

Using Machine Learning in Cannabis Industry

Presenter: Shah, Deepkumar
Acadia University
Supervisor: Silver, Daniel
Other authors: Dr. Andrew McIntyre
Department: Math

Terpenes are an important class of organic compounds produced from fruits and vegetables that we consume every day[1]. For instance, Myrcene is found from mango, Humulene is found from coriander. Terpenes mainly contribute to taste and smell factors that we perceive when consuming everything from apples and grapes to cannabis, which has recently been decriminalized and has important medical and business (e.g., recreational) use cases. A desirable property of these products is brand-wise consistency; i.e., an apple of a particular brand ought to invoke a similar experience between batches or trees, across various seasonal growing conditions, annual harvests, etc. The current study investigates the utility of Machine Learning (ML) algorithms as analysis tools for classifying mass spectra (MS)[2] of cannabis samples over a variety of brands / products with the end goal of predicting the presence (or lack thereof) of specific terpene compounds and terpene compositions to pre-determine schedules for plant

development, treatment routines, and market potential. Where humans have conventionally been used in quality assurance tests (potato chip brand tests, milk spoilage, etc.)[3], ML algorithms combined with rapid, easy-to-use and portable MS technology and devices offer the potential for more consistent, bias-free experience analysis and a multitude of improved efficiencies for producers.

Credit Card Fraud Detection

Presenter: Manek, Hardik
Acadia University
Supervisor: Silver, Daniel
Other authors:
Department: Math

Credit Card Fraud Detection

With the advent of new technologies and Internet, cashless transactions have become effortless. But, for online transactions, one does not need to be personally present in a certain place where the transaction occurs hence making it vulnerable to fraudulent attacks. A cyber-attacker can pretend to be the owner of a credit card and make a fraudulent transaction. There are several techniques to determine the nature of the transaction, for instance, by comparing the current transaction with previous transactions. If the monetary difference between current transaction and previous transaction is too large, then there is a greater probability of current transaction being a fraudulent transaction. This method is not reliable for anomaly detection. In some countries like India and China, banks deploy a two-step verification process which, strengthens the security of the transaction. While in other countries, employees in the bank manually segregate the transactions to be fraud or not. These methods are highly dependent on human intervention. Machine Learning can be utilized to automate the process of anomaly detection. Supervised algorithms such as Logistic Regression can be used to build a model that will predict the output in the form of binary classes i.e. 0 for a valid transaction and 1 for a fraudulent transaction. Moreover, better testing accuracy will be obtained by using Autoencoder Neural Network which is one of the unsupervised algorithms. Two main features considered in the dataset are Time and Amount, other features are compressed using Principal Component Analysis.

Session 2

9:05- 9:20am	Colin Vibert **One-sided simultaneous confidence intervals for one-way layouts with unequal variances**
9:25- 9:40am	Joy Liu **Animal Risk Assessment and Disease Control**
9:45- 10:00am	Ben Wang **Financial Data Exploration and Analysis for Stock Screening**
10:05- 10:20am	John Marcoux **Total Variation Denoising of Diffusion MRI Images Using a Modified Monge-Kantorovich Norm**
1:35- 1:50pm	Brady Ryan **Evaluation of Designs and Statistical Methods Under Response-Dependent Two-Phase Sampling for Genetic Association Studies**
1:55- 2:10pm	Kieran Bhaskara **Fractions with a twist: Continued fractions and polynomial Pell equations**
2:15- 2:30pm	Jesse Preston **Source-Sink Diffusion**
2:35- 2:50pm	Emily Wright **Diffusion with Multiple Sources and Sinks**
4:45- 5:00pm	Caleb Jones **An Algorithm Approach to the Game of Surrounding Cops and Robbers on Graphs**
5:05- 5:20pm	Justin Hughes **Delayed Cops and Robbers on Graphs**
5:25- 5:40pm	Aaron Dwyer **Domineering in Misère Play**

One-sided simultaneous confidence intervals for one-way layouts with unequal variances

Presenter: Vibert, Colin
Acadia University
Supervisor: Jianan Peng
Other authors:
Department: Math

This presentation will present one-sided simultaneous confidence intervals for means in a one-way layout under the assumption of unequal variances. Four types of intervals based on fiducial generalized pivotal quantities, MaxT and MinP method will be introduced. A real data example will be used to illustrate these methods.

Animal Risk Assessment and Disease Control

Presenter: Liu, Joy
University of Prince Edward Island
Supervisor: Liu Kai
Other authors:
Department: Math

Agriculture, food supply and animal in general are vulnerable to bacterial, fungal and viral threats .With different kinds of animals existing under human impact, such as companion animal, zoo animal, laboratory animal, livestock animal and wildlife animal, we focus on companion and livestock animal because they are both much more closely related to human life. Livestock is exceptionally vulnerable to pathogenic threats due to its high susceptibility nature. The objective of the proposed research is to control the Animal disease under the insurance and finance perspective.

Financial Data Exploration and Analysis for Stock Screening

Presenter: Wang, Ben
University of Prince Edward Island
Supervisor: Alvarez, Alexander; Liu, Kai; Islam, Shafiqul
Other authors:
Department: Math

In this report, we perform data exploration and analysis with the objective of proposing suitable portfolio selection strategies that take into account quarterly financial statements of publicly traded companies (fundamental analysis). After downloading the data from the WRDS website, we continue to analyze the data using R software. Finally, we make conclusions on strategies for selecting stocks and investment methods.

Total Variation Denoising of Diffusion MRI Images Using a Modified Monge-Kantorovich Norm

Presenter: Marcoux, John
Acadia University
Supervisor: Mendivil, Franklin
Other authors:
Department: Math

Efficient methods for denoising various forms of data are contantly evolving and are crucial in fields like medical imaging where even small amounts of noise can affect a doctor's diagnosis. It has been proposed in a 2016 paper(D. La Torre, F. Mendivil, O. Michailovich, E. Vrscay) that using the Monge-Kantorovich metric, along with total variation and regularization, one could create an effective and hopefully efficient method for denoising diffusion MRI images. In this presentation this aforementioned method for denoising will be discussed as well as the advantages of a Monge-Kantorovich style norm versus a Euclidean or other type of norm. Finally the effectiveness of a python based implementation that was created summer 2019 will be discussed as a measure of the effectiveness of the algorithm.

Evaluation of Designs and Statistical Methods Under Response-Dependent Two-Phase Sampling for Genetic Association Studies

Presenter: Ryan, Brady
Memorial University
Supervisor: Yilmaz, Yildiz
Other authors:
Department: Math

In many genetic association studies, while the aim is to identify genetic variants associated with a trait (such as a disease trait), budgetary constraints prevent genotyping all individuals in a cohort. Selection of individuals for genotyping according to their quantitative trait value can improve cost-efficiency. We consider two-phase response-selective sampling designs. In the first phase, trait and inexpensive covariate values for all individuals in a cohort are obtained; in the second phase, genetic sequence data for a subset of individuals are obtained according to their trait values and possibly their inexpensive covariates. In this talk I will present our proposed likelihood and pseudo-likelihood methods, along with their performance under common, low-frequency, and rare variant analyses. I will also present a comparison of these methods and the most efficient response-dependent sampling design under each method. Finally, I discuss the robustness of the estimation methods and sampling designs under misspecified models. This is a joint work with Anathika Nirmalkanna (MSc), Dr. Yildiz Yilmaz and Dr. Candemir Cigsar.

Fractions with a twist: Continued fractions and polynomial Pell equations

Presenter: Kieran Bhaskara
Dalhousie University

Continued fractions are an important tool in solving many nonlinear Diophantine equations. In particular, solutions of Pell's equation x^2 – ny^2 = 1 can be determined by examining the continued fraction expansion of sqrt(n). In this talk, we give a general introduction to continued fractions and their use in solving a polynomial analogue of Pell's equation. We also present general solutions for a few specific polynomial Pell equations.

Source-Sink Diffusion

Presenter: Preston, Jesse
Mount Saint Vincent University
Supervisor: Cox, Danielle; Mullin, Todd.
Other authors: Wright, Emily
Department: Math

In this talk we will introduce Source-Sink Diffusion. Preliminary results will be presented, and the periodicity of the process will be discussed. It is known that Diffusion of all graphs have period 1 or 2. We will show that with the addition of source and sink that this is not always the case.

Diffusion with Multiple Sources and Sinks

Presenter: Wright, Emily
Mount Saint Vincent University
Supervisor: Cox, Danielle; Mullen, Todd
Other authors: Preston, Jesse
Department: Math

Consider the game of Diffusion, where at each time step, all vertices simultaneously send one chip to each neighbour with fewer chips. We introduce a variation of Diffusion, Source Sink Diffusion, that plays out as normal, but during each time step, a chip is given to the vertex adjacent to the source, and a chip is taken from the vertex adjacent to the sink. General characteristics of graphs with one source and sink in period 1 are discussed, and results on periodicity of graphs with multiple sources and sinks will be presented and compared to the results of Diffusion. Joint work with D. Cox, J. Preston, and T. Mullen.

An Algorithmic Approach to the Game of Surrounding Cops and Robbers on Graphs

Presenter: Jones, Caleb
Memorial University
Supervisor: Pike, David
Other authors:
Department: Math

In the standard game of cops and robbers on graphs, the least number of cops that have a winning strategy on a graph G is called the ‚Äúcop number‚Äù of G, and is denoted c(G). The game of surrounding cops and robbers is a relatively new variant of the original game in which the cops win by surrounding the robber, not by catching them. In this version of the game, the least number of cops that have a winning strategy on a graph G is called the ‚Äúsurrounding cop number‚Äù of G, and is denoted s(G). There exists a theorem and a corresponding algorithm that can find c(G) for any graph G (Bonato and Chiniforooshan, 2009). We alter the theorem and algorithm to instead find s(G) for any graph G.

Delayed Cops and Robbers on Graphs

Presenter: Hughes, Justin
Mount Allison University
Supervisor: Messinger, Margaret-Ellen
Other authors:
Department: Math

In the world of graph theory, one commonly discussed problem is the cops and robber problem, where several agents (the cops) take turns visiting neighbouring vertices to find another moving agent (the robber). The game ends when one cop occupies the same vertex as the robber (ie: when the robber is captured). In this game, the cops and robber play optimally (meaning the cops try to capture the robber, and the robber tries to evade capture). The cops and the robber can always see each other's current vertices. Any graph G has a minimum number of cops needed to capture the robber, which is referred to as the cop number of G.

In this talk, we consider two variations of this original problem:
Variant A: the robber can always see the cops' positions, but the cops can only see the position the robber was occupying in the previous round.
Variant B: same as Variant A, except if a cop occupied a vertex adjacent to the robber at the end of any round, then the cops can see the robber's current location.

We explore these variations of the problem, compute the cop number for different classes of graphs, prove that Variant B's cop number is equal to the normal cop number, and attempt to find a subgraph result for Variant A (a result such that for a graph G, any subgraph of G will not have a higher cop number than G itself).

Domineering in Mis√®re Play

Presenter: Dwyer, Aaron
Memorial University-Grenfell
Supervisor: Milley, Rebecca
Other authors: Willette, Michael
Department: Math

In combinatorial game theory, mis√®re play is a set of rules in which the first player who cannot make a move is the winner. Games under this rule set are complicated to analyse, as many of the algebraic properties of normal play, in which the first player who cannot move loses, are not present. This talk will take a look at the combinatorial game Domineering under mis√®re play, and explore both the difficulties encountered and methods used in finding results under certain restrictions.

Session 3

9:05- 9:20am	Jennifer McNichol **The Estimation of Calibration Coefficients in Quantitative Fatty Acid Signature Analysis**
9:25- 9:40am	Abby Anderson **Modeling of spruce budworm population on a windy island**
9:45- 10:00am	Samantha Bardwell **A Dynamic Individual-Based Model of a Population of People Who Inject Drugs**
10:05- 10:20am	Lucas MacQuarrie **An Introduction to Mathematical Biology and Modelling of Tumour Development**
1:35- 1:50pm	Ellen McCole **Benchmarking MDLR for OGI in the field in a way that doesn't directly involve CH4 release**
1:55- 2:10pm	Lawrence Daniel Doucett **SphereSkeltons: Sphere-mesh Fitting for Extracting Topologically Accurate Medial Skeletons from Point Clouds**
2:15- 2:30pm	Graeme Zinck **Opacity in Modular Systems**
2:35- 2:50pm	Syed Zeeshan Ahmed **Preserving consumer DNA privacy for finding relatives in a malicious two-party computation**
4:45- 5:00pm	Mathieu Briedeau **Improving heterogeneous distributed databases resiliency using fuzzy logic**
5:05- 5:20pm	Simon Gauvin **Vizwik: Lessons Learned from the Design and Realization of a Visual Dataflow Language for the Web**

The Estimation of Calibration Coefficients in Quantitative Fatty Acid Signature Analysis

Presenter: McNichol, Jennifer
University of New Brunswick - Saint John
Supervisor: Stewart, Connie
Other authors:
Department: Math

Quantitative Fatty Acid Signature Analysis (QFASA) has become a popular method of diet composition estimation, especially for marine predators. Along with fatty acid signatures for a particular predator and their respective prey, QFASA hinges on known values termed calibration coefficients, which account for the differential metabolism of individual fatty acids. In practice, calibration coefficients are not known and therefore must be estimated via feeding trials with captive animals. The main criticism of QFASA is that verifying the accuracy of calibration coefficients is nearly impossible and may introduce bias into the diet estimates. To resolve this issue, a new model was proposed which allows for estimation of calibration coefficients simultaneously alongside the diets. However, the proposed model has only a limited range of supporting evidence and has not yet been directly compared to QFASA. In this talk, results from a simulation study comparing the new model with QFASA will be discussed.

Modeling of spruce budworm population on a windy island

Presenter: Anderson, Abby
Memorial University-Grenfell
Supervisor: Vasilyeva, Olga
Other authors:
Department: Math

The spruce budworm is a destructive insect that feeds on balsam fir trees in coniferous forests. Most of the dozen budworm species are capable of destroying entire forests. Outbreaks of the insect population have been modeled by many mathematicians. We consider a modification of the reaction-diffusion model for spatial dynamics of spruce budworm population introduced by D. Ludwig, D. G. Aronson, and H.F. Weinberger, taking into account the movement bias caused by prevailing winds.

Mathematically, this modification takes the form of a reaction-diffusion-advection equation. We are interested in applying our model to the case of an island (e.g. Newfoundland), and thus we impose hostile (Dirichlet) boundary conditions.This presentation will focus on the determination of the critical values of advection (wind) speed, which, in turn, determine conditions under which the population survives or explodes (i.e. the population experiences an outbreak). The determination of the critical advection speeds is based on phase-plane analysis, integral calculus, and geometric techniques.

A Dynamic Individual-Based Model of a Population of People Who Inject Drugs

Presenter: Bardwell, Samantha
St. Francis Xavier University
Supervisor: Lukeman, Ryan
Other authors:
Department: Math

A recent health crisis affecting vulnerable populations across Canada, including Nova Scotia, involves the presence of fentanyl and carfentanil in heroin and other drugs, greatly increasing the risk of fatal overdose. In fact, overdose risk depends on a host of other factors, drawn from demographics, use history, and available interventions. In this work, a dynamic, stochastic mathematical model is used to study the population dynamics of people who inject drugs (PWID). The mathematical model combines individual-based and population-level compartmental analysis approaches. By specifying the numerous characteristic-dependent relationships that form a user‚Äôs risk, the model aims to quantify the leading effects on injection drug risk and allow for accurate predictions for future trends in overdose fatality, recruitment into the population, and intervention efficacy. In this talk, I will provide an overview of the model, its implementation in Matlab, and its specific calibration to the PWID population of urban Toronto. I will then present results that show predicted impacts of varying model parameters and intervention strategies, and what current assumptions about mortality and recruitment imply for PWID population dynamics across decades.

An Introduction to Mathematical Biology and Modelling of Tumour Development

Presenter: MacQuarrie, Lucas
University of Prince Edward Island
Supervisor: Islam, Shafiqul; Saad, Nasser
Other authors:
Department: Math

Biology has historically been considered a field too complicated to be modeled by mathematics, however this is far from the truth. Various mathematical tools have been applied to biological phenomenon to better understand and predict their process as well as provide insight as to how we can manipulate these processes for the benefit of mankind. Temporal networks, for example, have been used to model epidemics on small and large scales while various kinds of differential equations have been used to model tumour growth. This presentation will be an introduction to various models in mathematical biology with a focus on mathematical oncology and current research by JFC Alfonso et al.

Benchmarking MDLR for OGI in the field in a way that doesn't directly involve CH4 release

Presenter: McCole, Ellen
St. Francis Xavier University
Supervisor: Risk, Dave
Other authors:
Department: Math

Methane is an unwanted greenhouse gas whose global warming potential is 28-34 times more powerful than carbon dioxide on a 100-year horizon and whose emissions are higher than government inventories suggest. In the last year, Canada committed to cut their methane emissions in the oil and gas sector by 40%-45% below 2012 levels by 2025. To achieve this goal, regulators, researchers, and industry are interested in improving methane leak detection practices. Optical Gas Imaging (OGI) has now taken over as an efficient and safe close-range leak detection method. The performance of OGI varies significantly with weather conditions, and there is no empirical approach to define minimum detectable leak rate (MDLR) under field conditions. The goal of this project is to develop a field-deployable benchmarking system for OGI MDLR, that doesn‚Äôt directly involve dangerous or environmentally controlled releases of methane. We conducted a feasibility experiment to find a safe and readily available proxy gas that could replace methane for field benchmarking. Once this proxy gas was established to be visible under similar background temperature contrast conditions to methane by the FLIR GF320 OGI camera, further experiments could be carried out to benchmark MDLR using the proxy. Since OGI cameras define methane leaks on a pixel-by-pixel basis, where leak rate drives both the pixel intensity and number of pixels, we were able to develop a quantitative approach to image analysis that relies on pixel output.

SphereSkeletons: Sphere-mesh Fitting for Extracting Topologically Accurate Medial Skeletons from Point Clouds

Presenter: Doucett, Lawrence Daniel
St. Mary’s University
Supervisor: Jiju Poovvancheri
Other authors: Jiju Peethambaran
Department: Math

We investigate into the problem of 1D skeleton extraction from 3D scans of objects acquired via volumetric capture or LiDAR sensors. Skeletons are one of the most compressed shape representations which aid in simplifying and optimizing 2D/3D graphical objects in animations and modeling. We use sphere mesh representations to approximate the entire shape from a 3D scan. A sphere mesh is a hybrid geometric representation that exhibits the properties of an implicit surface as well as simplicial complex representations. The algorithm constructs the skeleton of a shape from edges of the sphere mesh that approximate the scan, where the sphere mesh is constructed through a set of maximally inscribed empty balls (MEB) fitted onto the point cloud and a nearest neighbor search. A minimal spanning tree is used to prune off extraneous edges from the skeleton. Edges of over-lapping spheres are given much lower weight in the graph as a stability metric. Skeletons, thus created consist of a set of spatial line segments that do not intersect except at the endpoints and capture the shape topology. This paper presents the 2D proof of

concept with a few preliminary results that will be extended to 3D. The reconstructed skeletons will further facilitate the recovery of the original shape from the point cloud by connecting the spheres via tangential planes, or tangent lines in 2D. This research will ideally allow for more efficient real time rendering and tracking of 3D objects, which is an integral part of modern VR/AR applications.

Opacity in Modular Systems

Presenter: Zinck, Graeme
Mount Allison University
Supervisor: Ricker, Laurie; Marchand, Hervé; Hélouët, Loïc
Other authors: Ricker, Laurie; Marchand, Hervé; Hélouët, Loïc
Department: Math

We use online services every day, and a growing number of internet of things devices are popping up in our homes. While these are useful, they can also expose our data: attackers can identify secret information, even with only a partial observation of a system. We use formal methods to model such systems and determine if they are secure. Specifically, we use labelled transition systems to verify a property called opacity. A system is opaque if an attacker cannot distinguish areas of confidentiality from areas of non-confidentiality. The verification of opacity becomes more challenging when considering a system composed of many interacting components. There are cases where individually opaque components reveal their areas of confidentiality when interacting with other modules. We are interested in a situation where opacity holds over the composition of system components: when an attacker observes the interface between components, individually opaque modules remain opaque in the combined system. We present algorithms for verifying and enforcing opacity for such system architectures.

Preserving consumer DNA privacy for finding relatives in a malicious two-party computation

Presenter: Ahmed, Syed Zeeshan
University of New Brunswick - Saint John
Supervisor: Kaser, Owen ; Lemire, Daniel
Other authors:
Department: Math

DNA analysis of two individuals can reveal the degree of relationship between them if they share a common ancestor. Consumers face data privacy risks when they use direct-to-consumer DNA testing. Sensitive data are often uploaded into the cloud and made publicly available. Customers interested to find individuals with common ancestry or build up a family-tree may be willing to share their genetic data and compromise on their data privacy. However, some consumers only want to verify if they are related to the other individual and preserve their data privacy. DNA data possess sensitive information regarding some known medical conditions as well. Therefore, certain individuals may be unwilling to engage in actions that may publish their DNA data. So, in this research, we are seeking a software implementation for establishing individuals' relatedness while preserving their data privacy. We are implementing and comparing two approaches. Both approaches use the established ‚Äúopposite homozygotes‚Äù technique to implement the genealogical DNA matching algorithm on genotyped data. The first one is a hashing-based approach and the other one involves secure two-party computation from cryptography. Both these

approaches involve the usage of an efficient communication protocol for the concerned parties. The objective of our work is to define, design and implement our intended approaches and then compare them in terms of security, execution times and correctness of the results (false negatives and positives). The overall performance of our conceived approaches will be extensively evaluated through simulations of individual and family DNA datasets gathered through publicly available portals.

Improving heterogeneous distributed databases resiliency using fuzzy logic

Presenter: Brideau, Mathieu
Université de Moncton
Supervisor: Hervet, Eric
Other authors: Hervet, Eric
Department: Math

The world of Big Data is becoming more complex, prompting businesses to seek effcient and relatively less costly solutions to integrate their multiple, distributed heterogeneous sources of data. Integrating such complex elements increases exponentially the potential for failure. In this presentation, we will explore a novel approach to improve the resiliency of multidatabases with techniques based on fuzzy logic. We also detail some implementation decisions and compromises that were made.

Vizwik: Lessons Learned from the Design and Realization of a Visual Dataflow Language for the Web

Presenter: Gauvin, Simon
University of New Brunswick - Saint John
Supervisor: Boley, Harold; Baker, Chris
Other authors:
Department: CS

Visual programming languages have been in existence for 40 years and have contributed to a wide range of innovations such as graphical user interfaces and their editors. Dataflow-based visual languages have been the most successful to date with commercial products such as LabView, and more recently, Google Tensor Flow. These languages have been notoriously complex to build due to the a wide range of technologies that need to be integrated. Technologies include assemblers, compilers, parsers, debuggers, virtual machines, renderers, and visual layout algorithms. In this talk we shall present results of a multi-year project that produced a cloud-based visual dataflow language and the practical design patterns that made it possible to deliver to the web.

Session 4

9:05- 9:20am	Sarah Li **Towards a Finite Presentation of Unitary Dyadic Operators**
9:25- 9:40am	Everett Patterson **Using Linear Algebra in Quantum Enganglement Theory**
9:45- 10:00am	Adam Lucas **A Look at Distributed Computing and Modal Logic**
10:05- 10:20am	Dylan Ruth **Quantifier Elimination in Divisible Abelian Groups and Their Expansions**
1:35- 1:50pm	Charles Gerrie **Multi-block analysis of Unixcrypt**
1:55- 2:10pm	Samuel Bauer **Recovering Matrix Algebras**
2:15- 2:30pm	Andrew Fraser **How Good is Your Approximation? – The Essential Role of Error Estimation and Control in Numerical Computation**
2:35- 2:50pm	Uyen Dao **On Scheffe's criteria**
4:45- 5:00pm	Adam Smith **Applications and Numerical Solutions of Integral Equations**
5:05- 5:20pm	Youssef Zaazou **Horizons as Boundary Conditions in Spherical Symmetry**
5:25- 5:40pm	Jeremy Peters **Geometric Algebra for Relativity**

Towards a Finite Presentation of Unitary Dyadic Operators" instead of "Towards a Finite Representation of Unitary Dyadic Operators

Presenter: Li, Sarah (Meng)
Dalhousie University
Supervisor: Ross, Neil Julien
Other authors: Ross, Neil Julien
Department: Math

An important problem in quantum information theory is the decomposition of unitary operators into a product of basic gates. This decomposition is typically not unique and it is desirable to find short representations because they result in cheaper physical implementations. In this presentation, we give an introduction and progress report on our study of dyadic unitary operators. These operators are of interest in quantum computation, because they correspond to well-studied quantum circuits. We will first introduce dyadic unitary operators along with the required algebraic prerequisites. Then, we will sketch a deterministic algorithm which performs operator decompositions. Finally, we will discuss our ongoing work in which we use this algorithm and the Cayley graph associated to dyadic unitary operators to find relations between the generators.

Using Linear Algebra in Quantum Entanglement Theory

Presenter: Patterson, Everett
Mount Allison University
Supervisor: Johnston, Nathaniel
Other authors:
Department: Math

As classical computing becomes limited by the laws of quantum mechanics, Quantum Information Theory has arisen as away to utilize these laws to embolden computational abilities. Much of this theory can be expressed using linear algebra, with emphasis on Hermitian matrices and tensor products. Within this Ô¨Åeld, entanglement plays a prominent role with Entanglement Witnesses (EW) acting as an operator on Quantum States. We examine what properties Entanglement Witnesses must satisfy,and further examine what these properties might tell us about certain families of matrices.

A Look at Distributed Computing and Modal Logic

Presenter: Lucas, Adam
Dalhousie University
Supervisor: Kishida, Kohei
Other authors:
Department: Math

Distributed computing studies questions surrounding when and how a collection of machines (the model of computation usually does not matter) might go about collectively deciding on a configuration of outputs given each machine his own input given a method of communication with one another.
Dynamic epistemic logic is a section of modal logic which attempts to model how rational beings reason

with limited or imperfect knowledge of a situation.
Last year, a paper was written which explained an approach to distributed computing using DEL. I will briefly explain the main ideas of these areas and talk about my research relating to this new approach.

Quantifier Elimination in Divisible Abelian Groups and Their Expansions

Presenter: Ruth, Dylan
Memorial University-Grenfell
Supervisor: Vasilyev, Yevgeniy
Other authors:
Department: Math

In model theory (a branch of mathematical logic studying mathematical structures from the point of view of first order language), a major role is played by the notion of elementary equivalence that is observed when two structures satisfy the same first order sentences. While this is a weaker property than being isomorphic, it can still provide some useful insight into mathematical structure. A common way of establishing elementary equivalence is by showing that there exists a family of partial isomorphisms between finitely generated substructures exhibiting the back-and-forth property. If this property holds for any two sufficiently large structures satisfying a certain list of axioms, then modulo those axioms, any first order property involving quantifiers is equivalent to one without quantifiers. Establishing quantifier elimination enables one to more easily analyze the complexity of a structure. In my talk , I will discuss the question of quantifier elimination in the cases of torsion-free divisible abelian groups (such as (R,+,-,0)) and their expansions (results of adding new relations to the language of groups), such as divisible ordered abelian groups (e.g. (R,+,-,0,<)) and some intermediate structures, where we omit the binary relation of order, but keep a predicate for a convex subset.

Multi-block analysis of Unixcrypt

Presenter: Gerrie, Charles
Dalhousie University
Supervisor: Selinger, Peter
Other authors:
Department: Math

This research introduces a method for multi-block analysis of the unixcrypt cipher. We extend a single-block algorithm previously developed for this cipher. Taking advantage of how block keys are related to each other, we collect information about the key used to generate them. We then use statistical methods to pick the most likely key. Thus with sufficiently many blocks we can recover the original key and decrypt the ciphertext completely.

Recovering Matrix Algebras

Presenter: Bauer, Samuel
Memorial University
Supervisor: Kochetov, Mikhail
Other authors:
Department: Math

Given a set of generating matrices along with degrees in an abelian group G, we can retrieve the grading in terms of the elementary part of the grading and the division part of the grading. This leverages a constructive proof of the Noether-Skolem theorem. This is used classify gradings up to triality.

How Good is Your Approximation? - The Essential Role of Error Estimation and Control in Numerical Computation

Presenter: Fraser, Andrew
Saint Mary's University
Supervisor: Muir, Paul
Other authors:
Department: Math

Computational Science is a central component, along with the traditional modes of experimental and theoretical investigation. Computational Science involves the development and solution of mathematical models, i.e., systems of equations, that represent approximations to real world phenomenon in a wide variety of areas of scientific investigation. These mathematical models typically do not have closed form solutions and thus the models are solved using computational software to obtain approximate solutions. Since these solutions are approximate, the question that must be addressed is "How Good is the Computed Solution?". This question is answered for a given numerical solution through the computation of a good quality error estimate. This talk will discuss current work on answering this question in the area of computational methods for differential equations that depend on time and/or one or more spatial dimensions. We will describe the use of collocation, a general numerical method that can be used for a wide range of problem classes, as well as our recent work in the development of efficiently computable error estimates for collocation solutions based on special types of interpolants. We provide results from numerical experiments to demonstrate the effectiveness of our approach.

On Scheffe's criteria

Presenter: Dao, Uyen
University of Prince Edward Island
Supervisor:
Other authors:
Department: Math

In solving differential equations, we might sometimes encounter solutions that involve the use of multiple-term recurrence formulas, which is generally viewed as undesirable for its complexity. Thus, there arises a natural need to establish a condition, or criteria, under which a¬† solution of a linear ordinary n-th order differential equation possesses a two-term recurrence relation.

In this talk, we are going to focus on the 2nd-order differential equations only, for it is most widely applicable to many natural sciences, but the result of the criteria can well be applied to the nth cases. The general idea is to utilize the method of Frobenius to transform the coefficients into power series, then put on suitable restrictions such that only two recurrence terms survive in the final formula. This would only require some basic first-year understanding of differential equations.

Applications and Numerical Solutions of Integral equations

Presenter: Smith, Adam
University of Prince Edwards Island
Supervisor: Shafiqul Islam
Other authors:
Department: Math

Closely related to differential equations, integral equations are commonly used in various areas of physics and mathematics. In fact, if you‚Äôve taken a course in differential equations you‚Äôve likely had interaction with integral equations with or without your knowledge. These equations are often applied to model concepts in science such as population growth and biological species living together. Although simple in appearance, these equations can prove to be difficult to find exact solutions of and often efforts are put towards finding approximations to illustrate the structure of the solution. A majority of the complexity behind finding solutions to integral equations falls behind the form the kernel function takes and for this reason a large focus of the integral equation is categorizing it by its kernel. With a look at some examples we will use the theory behind the concepts of maximum entropy and B-splines to approximate solutions to a category of integral equations with an emphasis on the flexibility on the presented method.

Horizons as Boundary Conditions in Spherical Symmetry

Presenter: Zaazou, Youssef
Memorial University
Supervisor: Booth, Ivan; Kunduri, Hari
Other authors:
Department: Math

The study of black hole dynamics is fundamental to understanding Einstein‚Äôs theory of general relativity in the strong field regime. Of interest are binary black hole mergers where several solar masses worth of energy is expelled in fractions of a second. Simulations of these mergers are very computationally costly, and it may take supercomputers several months to run seconds worth of simulation. Therefore, we propose a numerical scheme for simulating spherically symmetric spacetimes made dynamic by matter fields using a horizon-based initial value formalism. These simulations have the advantage of not requiring nearly as much time and computational resources as the fully dynamic system while nonetheless revealing interesting properties about black hole dynamics.

Geometric Algebra for Relativity

Presenter: Peters, Jeremy
Dalhousie University
Supervisor: Schnetter, Erik; Coley, Alan.
Other authors:
Department: Math

Geometric algebra is an elegant extension of vector calculus, and provides a unified description in many areas of physics. Geometric algebra is a potentially useful tool to simplify computations in, e.g., numerical relativity. In this talk, I will start by defining the geometric product of two vectors, and use this geometric product to build a geometric algebra on R^n. Preliminary applications of geometric algebra to rigid transformations, the quaternions, and the Pauli algebra in quantum mechanics will be briefly described. Then geometric algebra will be extended to geometric calculus, which is then used to rewrite Maxwell's equations in a single equation. Within this framework, I will show numerical solutions to Maxwell's equations to simulate a 1+1-dimensional oscillating dipole. Finally, I will describe a gauge-theory formulation of GR, which will use objects from geometric algebra. This will be the starting point for future applications in linear and non-linear gravity.