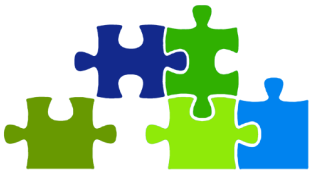
Science Atlantic
Computer Science
Mathematics & Statistics
Conference



Mathematics, Statistics, and Computer Science Conference

UNB Fredericton

October 13–15, 2017

UNB
EST. 1785
Fredericton

Atlantic Association for Research in the Mathematical Sciences

ACENET
accelerate discovery

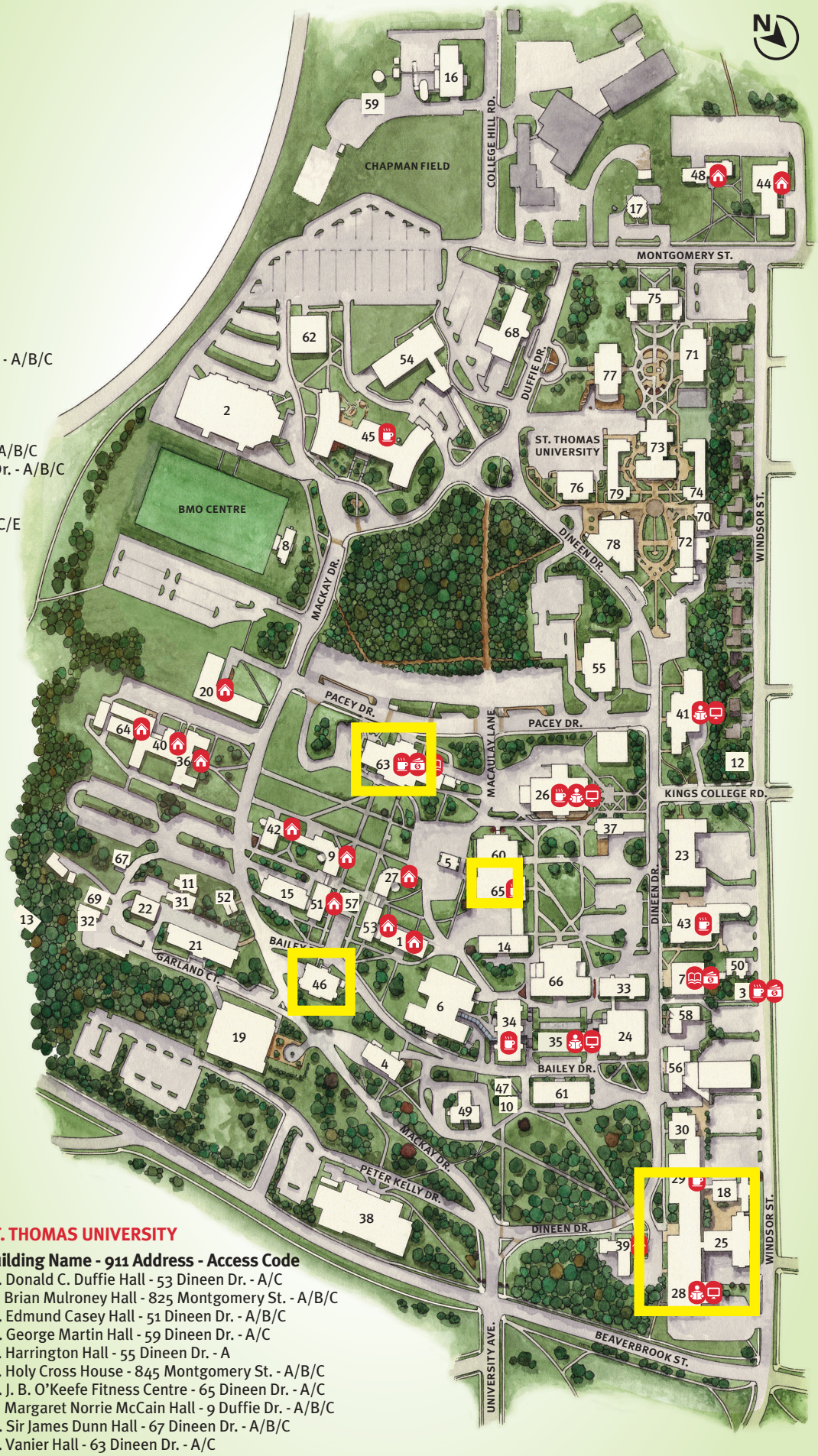
UNB
EST. 1785
UNIVERSITY OF NEW BRUNSWICK
Computer Science
Fredericton

UNB
EST. 1785
UNIVERSITY OF NEW BRUNSWICK
Science
Fredericton

UNIVERSITY OF NEW BRUNSWICK
FREDERICTON CAMPUS

Building Name - 911 Address - Access Code

1. Aitken House - 14 Bailey Dr.
2. Aitken University Centre - 20 Mackay Dr. - A/C
3. Alden Nowlan House - 676 Windsor St. - C
4. Alumni Memorial Building - 13 Bailey Dr. - A
5. Annex C - 13 Macaulay Lane - A/C
6. Bailey Hall - 10 Bailey Dr. - A/B/C
7. Bank/Bookstore Building - 29 Dineen Dr. - A
8. BMO Centre - 25 Mackay Dr. - A/C
9. Bridges House - 45 Mackay Dr.
10. Brydone Jack Observatory - 5 Bailey Dr.
11. Building #7 - 6 Garland Ct.
12. Burden Academy - Windsor St.
13. Campus House - 11 Garland Ct.
14. Carleton Hall - 19 Macaulay Lane - A/B/C
15. C.C. Jones Student Services Centre - 26 Bailey Dr. - A/B/C
16. Central Heating Plant - 950 College Hill Rd.
17. College Hill Daycare - 850 Montgomery St. - A/C
18. Computer Science Information Technology Centre - 550 Windsor St. - A/B/C
19. CURRIE CENTER - The Richard J. - 15 Peter Kelly Dr. - A/B/C
20. Elizabeth Parr-Johnston Residence - 34 Mackay Dr. - A/B/C
21. Enterprise UNB Building #1 - 2 Garland Ct.
22. Enterprise UNB Building #2 - 8 Garland Ct.
23. Facilities Management - 767 Kings College Rd. - C/E
24. Forestry & Geology Building - 2 Bailey Dr.
25. Gillin Hall - 540 Windsor St. - A/B/C
26. Harriet Irving Library - 5 Macaulay Lane - A/B/C
27. Harrison House - 12 Macaulay Lane
28. Head Hall - 15 Dineen Dr. - A/B/C
29. Head Hall/Old Civil Engineering - 17 Dineen Dr. - B
30. Head Hall/Electrical Engineering - 19 Dineen Dr.
31. Header House - 4 Garland Ct.
32. Hut #5 - 5 Garland Ct.
33. I.U.C. Forestry - 28 Dineen Dr. - B
34. I.U.C. Physics & Admin. - 8 Bailey Dr. - A/B/D
35. I.U.C. Science Library - 4 Bailey Dr. - A/C
36. Joy W. Kidd House - 42 Mackay Dr. - A/B/C
37. Keirstead Hall - 38 Dineen Dr. - A/B/C
38. Lady Beaverbrook Gym - 2 Peter Kelly Dr. - A
39. Lady Beaverbrook Residence - 9 Dineen Dr. - A
40. Lady Dunn Hall - 40 Mackay Dr. - A/B/C
41. Ludlow Hall - 41 Dineen Dr. - A/B/C
42. MacKenzie House - 43 Mackay Dr. - A/E
43. MacLaggan Hall - 33 Dineen Dr. - A/B/C
44. Magee House - 780 Montgomery St. - A/B/C
45. Marshall d'Avray Hall - 10 Mackay Dr. - A/B/C
46. McConnell Hall - 19 Bailey Dr. - A/B/C
47. McCord Hall - 7 Bailey Dr.
48. McLeod House - 810 Montgomery St. - A/B
49. Memorial Hall - 9 Bailey Dr. - A/C
50. Muriel McQueen Fergusson Centre - 678 Windsor St. - A/B/C
51. Neill House - 22 Bailey Dr.
52. Neville Homestead - 58 Mackay Dr.
53. Neville-Jones House - 16 Bailey Dr.
54. New Brunswick Community College Fredericton Campus - 26 Duffie Dr. - A/B/C
55. NRC Institute for Information Technology - e-Business - 46 Dineen Dr. - A/B/C
56. Provincial Archives - 23 Dineen Dr. - A/B/C
57. Residence Administration - 20 Bailey Dr. - E
58. R.N. Scott Hall - 25 Dineen Dr. - A/B/C
59. Salt Storage Building - 948 College Hill Rd.
60. Singer Hall - 7 Macaulay Lane - A/C
61. Sir Howard Douglas Hall - 3 Bailey Dr.
62. South Gym - 16 Mackay Dr. - A/B/C
63. Student Union Building - 21 Pacey Dr. - A/B/C
64. Tibbits Hall - 40 Mackay Dr. - A/B/C
65. Tilley Hall - 9 Macaulay Lane - A/B/C
66. Toole Hall - 30 Dineen Dr. - A/C
67. UNBEA Building 10 - 10 Garland Ct.
68. Wu Centre/College of Extended Learning - 6 Duffie Dr. - A/B/C
69. Yellow Building - 7 Garland Ct.



ST. THOMAS UNIVERSITY

Building Name - 911 Address - Access Code

70. Donald C. Duffie Hall - 53 Dineen Dr. - A/C
71. Brian Mulroney Hall - 825 Montgomery St. - A/B/C
72. Edmund Casey Hall - 51 Dineen Dr. - A/B/C
73. George Martin Hall - 59 Dineen Dr. - A/C
74. Harrington Hall - 55 Dineen Dr. - A
75. Holy Cross House - 845 Montgomery St. - A/B/C
76. J. B. O'Keefe Fitness Centre - 65 Dineen Dr. - A/C
77. Margaret Norrie McCain Hall - 9 Duffie Dr. - A/B/C
78. Sir James Dunn Hall - 67 Dineen Dr. - A/B/C
79. Vanier Hall - 63 Dineen Dr. - A/C

Access Key

A Level or ramp entrance
B Elevator
C Accessible restroom
D Assistance needed from parking lot *(no curb break)*
E One step up to entrance

For the most current information on accessibility call 506-453-4830

Bookstore | Library | Residence | Food services | Bank machine | Computer lab

Contents

Welcome and General Information

Welcome to the University of New Brunswick.

This time of the year is ideal for visiting Fredericton: a triumph of beautiful colours awaits you, lulled by the equally beautiful Wolastoq, the Saint John river. You have a very full program, between math and programming contests, talks, and the AARMS workshop; but if you can spare some time I encourage you to explore more of our campus and discover some of its wonderful green spaces.

I wish you a productive scientific exchange and an enjoyable stay in Fredericton and at the UNB.

Dr. Luigi Benedicenti, Dean of Computer Science

On behalf of our faculty and staff in Math and Stats Department, I'd like to welcome you to this unique undergraduate conference. Among all the events here, you will have the opportunity to listen to three plenary lectures by distinguished mathematical scientists. All three lectures have a bearing on the greatest scientific discovery of the century so far, the detection of gravitational waves. One of the lectures is on curvature. This is a particularly auspicious topic, for these waves are ripples in the curvature of spacetime. It is amazing that they were predicted by Einstein 100 years ago with only pencil and paper. And it is beyond stunning that the billion dollar equipment used to detect them has a sensitivity of 1 part in 10^{21} . To visualize what this means, imagine measuring the 4.25 light year distance to our nearest star, Proxima Centauri, to an accuracy of the width of a human hair! This grand achievement came about with mathematics, physics, massive computing power, statistical analysis, data visualization, and precision engineering. It leaves us with so much more to learn. I hope you enjoy the conference and leave more inspired.

Dr. Viqar Husain, Acting Chair, Dept. of Math & Stats

General Information

For **Friday's banquet**, please make your way to the Student Union Building for 5:45pm, to be seated by 6pm. A plenary talk will follow the meal. Afterward, attendees are welcome to head to the Grad House.

Saturday's lunch will be in the McConnell Hall dining hall (see campus map; volunteers will be available to direct you). Meal tickets for Saturday's lunch will be available from student volunteers as you enter McConnell Hall. **Please wear your name tags** so the volunteers will be able to identify that you're with the conference.

Prize winners will be announced following Saturday's final plenary talk, with a pizza party to follow.

Schedule At A Glance

Locations: With the exception of the Friday evening banquet in the Student Union Building, and Saturday's lunch in McConnell Hall, all events on Friday and Saturday will be in the Head Hall / Gillin Hall / CS ITC building complex. Sunday's AARMS Workshop will be in Tilley Hall 404.

Friday

11:00–12:00	Check-in – Lobby
11:50–12:30	Programming Competition Meeting – H-C9
12:30–5:30	Programming Competition
2:30–5:30	Mathematics Competition – G-C122
2:45–3:45	Joint meeting – H-C11
4:00–5:15	Math meeting – H-C11
4:00–5:15	CS meeting – H-C9
6:00–7:30	Banquet – SUB
7:30–8:30	**Sedgwick Lecture (CS)**– SUB: *Developing and Generalizing Augmented Visual Perception* – David Flatla

Saturday

8:30–10:00	Contributed talks – H-C11 (Math), H-C10 (CS), H-C9 (Stats)
10:00–10:15	Break – Lobby/G-C122
10:15–11:00	Contributed talks – H-C11 (Math), H-C10 (CS)
11:00–11:10	Welcome by Eddy Campbell – Auditorium
11:10–12:10	**Fields lecture (Stats)** – Auditorium: *Multiple change-point detection: Some recent developments and their applications.* – Amy Wu
12:10–1:40	Lunch – McConnell Dining Hall
1:40–3:15	Contributed talks – H-C11 (Math), H-C10 (CS), H-C9 (Math)
3:15–3:30	Break – Lobby/G-C122
3:30–4:00	NSERC info session – Auditorium
4:00–5:00	**Blundon lecture (Math)** – Auditorium: *How well do we understand curvature?* – Niky Kamran
5:00-5:20	Closing – Auditorium
5:30	Pizza party – Lobby

Sunday

AARMS Workshop on Statistical Learning and Health Data Analytics – Tilley 404

9:00–9:50	*Overview of Statistical Learning Methods* – Hugh Chipman
9:50–10:00	Coffee Break
10:00–11:30	*Health administrative data in New Brunswick* – Ted McDonald
11:30–12:00am	*Leveraging Medical Lexicons to Improve Health Data Analytics* – Samuel Stewart
12:00–1:00	Lunch – McConnell Dining Hall
1:00–1:30	*A graph-based change-point test with application in tracking cell division* – Amy Wu
1:30–2:00	*Statistical learning methods in emergency diagnoses and human microbiome analysis* – Hong Gu
2:00–2:30	*Analysis of correlated data* – Renjun Ma
2:30–3:00	AARMS CRG Discussion Continued

Keynote Speakers

David Flatla – *University of Dundee, Scotland*

Developing and Generalizing Augmented Visual Perception.

Abstract: Technology has a tendency to be used for purposes other than what it was originally designed for, and technology for improving accessibility is no exception. In this talk, I will describe my lab's work into developing accessible technology to improve visual perception, but with an eye on how this work can be generalized more broadly. In particular, I will focus on my own research areas of simulating and overcoming Impaired Colour Vision, as well as our recent work on overcoming situationally-induced visual impairment, as well as supporting speechreading (lipreading) acquisition.

Bio: Originally from Canada, Dr. David Flatla is a Lecturer and Dundee Fellow within Computing at the University of Dundee, Scotland. David leads the DAPRlab (Digitally-Augmented Perception Research lab), which seeks to enhance the day-to-day perceptual abilities of everyone - both people with impaired sensory capabilities as well as those with typical abilities. David's personal research focusses on helping people with impaired colour vision (ICV), both by developing new techniques for improving colour identification, as well as creating simulation techniques to help everyone else understand ICV. David's work has been awarded at both CHI and ASSETS.

Niky Kamran – *McGill University*

How well do we understand curvature?

Abstract: How does one describe the curvature of a geometric object, like a curve, a surface or a higher-dimensional continuum? What does the curvature of a geometric object tell us about its other properties, such as its degrees of symmetry and regularity? These questions have been studied by some of the great mathematicians of the past, including Newton, Gauss, Riemann and Cartan, and have led to the development of differential geometry, an important branch of mathematics in which the geometry of a space is studied using tools of mathematical analysis (differentiation, integration and comparison). Similar questions on the links between curvature, symmetry and regularity also come up in Physics, through Einstein's formulation of the relativistic field equations of gravitation in terms of the curvature of space-time. There have been some major recent advances in our understanding of curvature, although many questions still remain open. I will give a general description two such advances, namely the sphere theorem, proved by Brendle and Schoen in 2007, and the Willmore problem, settled in 2012 by Marques and Neves. I will also mention some open problems and work that is being currently carried out towards their solution. This talk is aimed at a general audience.

Bio: Originally born in Belgium, Dr. Niky Kamran is James McGill Professor of Mathematics at McGill University. His research concerns geometric analysis, differential geometry, and mathematical physics. He received a licentiate in mathematics from the Université Libre de Bruxelles and his Ph.D. in Mathematics from the University of Waterloo. In addition to McGill, Dr. Kamran has held positiions at the University of Waterloo, the CRM at l'Université de Montréal, and the Institute for Advanced Study, Princeton, School of Mathematics. He was the first Recipient of the CMS André Aisenstadt Prize in Mathematics in 1992, a Killam Fellow from 2006 to 2008, and winner of the CRM-Fields-PIMS Prize in 2014.

Amy Wu – *York University*

Multiple change-point detection: Some recent developments and their applications.

Abstract: Multiple change-point problems can be found in many areas of science and engineering. To detect all change-points in a data sequence is of great importance. A statistical analysis without considering their existence may lead to an incorrect or improper conclusion. We will present some numerical examples to illustrate a change-point problem, and show the importance to include a change-point in data modelling. Owing to the rapid development in model selection methods, a multiple change-point detection method can be built upon by converting a multiple change-point detection problem into a variable selection problem via proper segmentation of a data sequence. We will discuss recent developments in multiple change-point detection using this methodology and their applications in real problems.

Bio: Dr. Amy Wu is a professor in the Department of Mathematics and Statistics, York University. Dr. Wu is leading active and productive research programs in several areas including high-dimensional data analysis, model selection, clustering analysis, change-point analysis, spatio-temporal modeling, Bayesian statistics and statistical finance. These programs are supported by multiple external research grants. Since graduating with her PhD degree in statistics at the University of Pittsburgh in 1989, Dr. Wu has published more than 100 publications in both leading refereed statistical journals and book chapters and has supervised many graduate and undergraduate students. In addition, Dr. Wu has been actively involved in departmental and university committees and services. She had served as Associate Coordinator, Statistical Consulting Service, ISR, York University, Canada, 1995-1996; Director, Statistics Section, Department of Mathematics and Statistics, York University, Canada, 2001-2003; and Director, Graduate Program in Mathematics and Statistics, York University, Canada, 2003-2006.

Contributed Talks

Computer Science Session 1 – 8:30-10:00, Room H-C10

Determining if this word is used like that word: Predicting usage similarity with supervised and unsupervised approaches. – Milton King, Faculty of Computer Science, UNB (*Collaborators: Dr. Paul Cook*)

Abstract: Determining the meaning of a word in context is an important task for a variety of natural language processing applications such as translating between languages, summarizing paragraphs, and phrase completion. Usage similarity (USim) is an approach to describe the meaning of a word in context that does not rely on a sense inventory–a set of dictionary-like definitions. Instead, pairs of usages of a target word are rated in terms of their similarity on a scale. In this thesis, we evaluate unsupervised approaches to USim based on embeddings for words, contexts, and sentences, and achieve state-of-the-art results over two USim datasets. We further consider supervised approaches to USim, and find that they can increase the performance of our models. We look into a more detailed evaluation, observing the performance on different parts-of-speech as well as the change in performance when using different features. Our models also do competitively well in two word sense induction tasks, which involve clustering instances of a word based on the meaning of the word in context.

Undergraduate-Led Computer Science K-12 Outreach – Anna Luise Frankfurt, School of Mathematical and Computational Sciences, UPEI (*Collaborators: Patricia Kibenge, Dr. Andrew Godbout*)

Abstract: This report considers the first offering of an undergraduate-led computer science outreach program. This program focused on junior high students (grades 7-9), utilizing a week-long after-school boot camp. The undergraduate mentors organized, planned, and executed the entire program wherein the students had an introduction to computational thinking, computer programming, and a diverse group of undergraduate computer science students with a broad range of interests.

The program required minimal involvement from the university faculty and staff and was almost entirely run by the undergraduate computer science mentors recruited from the university student body.

The goals of the program were to give an opportunity for junior high students to learn about computer science and to share with them the image of a computer scientist, where otherwise these students may have never been exposed.

The program was beneficial to the undergraduate students by offering an opportunity to showcase skills such as but not limited to curriculum planning, teamwork, communication, and teaching.

This program was successful enough that it is now a part of ongoing university engagement planning. This report presents experience running the program as well as survey results from our participants.

SpaceLaunch: A faster app launcher that supports spatial memory – Manasi Shah, Faculty of Computer Science, UNB (*Collaborators: Nathaniel Brewer, Dr. Scott Bateman, Dr. Carl Gutwin*)

Abstract: In this talk, I will describe the importance of spatial memory and our new app launcher that exploits it: 'SpaceLaunch'. Mobile phones have become an essential part of our lives; from managing daily tasks to social connections, from tracking fitness to finding the nearest restaurant or ATM, we depend on mobile applications. With the increasing number of apps, finding our apps is increasingly becoming a problem. At least part of the problem is that app launchers (the interfaces used to find and launch apps) do not support our ability to remember app locations. Current app launchers provide page-based (i.e., apps placed on multiple pages) and folder-based (i.e., apps arranged within folders) mechanisms to access apps. However, these two common interfaces create obstacles to people building their spatial memory. Spatial memory is a form of memory that allows people to remember locations over time and retrieve them quickly. Current app launchers break people's spatial memory by hiding applications inside folders or to the pages which are not visible without slow interactions. To better support the development of spatial memory, we designed an app launcher - called SpaceLaunch - that makes all apps visible at once. Through our study, we found that SpaceLaunch better supports building spatial memory and is faster than other approaches for finding and launching apps.

Improving Clouds and Language Runtimes – Panagiotis Patros, Faculty of Computer Science, UNB (*Collaborators: Dr. Kenneth Kent, Michael Dawson*)

Abstract: The advent of the Internet of Things, Web Applications and Big Data Analytics has increased the pressure on Pay-As-You-Go Clouds, such as IBM Bluemix and Google App Engine. Cloud computing abstracts the underlying hardware and/or software resources that are made available by a provider to its clients. Also, to minimize costs and improve efficiency, they transparently share resources among their tenants, which is referred to as multitenancy. Clouds can also dynamically scale the provided resources such that performance is maintained in acceptable levels as the load fluctuates.

High-level languages, such as Java and Node.js, are commonly used; they provide a number of attractive features, such as Just-in-Time compilation and automatic memory deallocation. However, they do not run on the machine directly; instead, a Language Runtime executes their instructions.

Various research opportunities arise in the intersection of Clouds and Runtimes: First, can multitenancy be expanded to the level of the application server? Second, to what extent does multitenancy cause interference, is there a standardized way to measure this and how does it affect scaling? Third, Runtimes produce dynamic artifacts when they run: Could they be securely and efficiently shared across instances of the same

application? Fourth, how do behind-the-scene operations performed by Runtimes affect the tight on-time requirements, which are particularly important for high-velocity Big Data applications? Fifth, should all cloud-service requests be treated equally or could they be reordered under the right circumstances? Finally, how can a cloud application voluntarily restrict its resource consumption during non-critical times such that its interference is mitigated?

Who are you Wearing? – Hailey LeClair, School of Mathematics and Computational Sciences, UPEI (*Collaborators: Dr. Andrew Godbout*)

Abstract: If you've ever watched the red carpet interviews at an awards show then you have likely heard the phrase "Who are you wearing?". The implication being which designer created the dress or tuxedo or clothes that are being worn. While image classification and object recognition has advanced greatly over the past number of decades, algorithmically determining properties regarding clothing is still a difficult computational challenge. In this talk I will present work in progress towards exploring the idea of automatically classifying and recognizing clothing and fashion related items. I will explain how famous algorithms such as the Scale Invariant Feature Transform (SIFT) operate and why they struggle with the nuances of clothing and finally I will describe the direction of my research and how deep learning and convolutional networks may finally answer the eternal question of "who someone is wearing?".

Visualization of online advertisement URLs by deep learning – Xichen Zhang, Faculty of Computer Science, UNB (*Collaborators: Dr. Arash Habibi Lashkari, Dr. Ali A. Ghorbani*)

Abstract: The exploding development of World Wide Web has witnessed the growth of a new and profitable business model - online advertising. However, malicious advertisements have become one of the major issues to distribute scamming information, click fraud and malware. The existing advertising blocking systems are vulnerable to the evolution of new attacks and can cause time latency issues by querying remote servers. In this study, we propose a lightweight detection system depending only on lexical-based features. The statistical analysis of different URL sources shows that lexical-based features can provide sufficient information for online advertisement representation. Deep learning algorithms are used for online advertising classification and visualization. After optimizing the architecture of the deep neural network, our proposed approach can achieve satisfactory results with false negative rate as low as 1.31%. We also design a novel unsupervised method for data visualization. With the implementation of AutoEncoder for feature preprocessing and t-SNE for visualization, our model outperforms other dimensionality reduction algorithms and can generate clear clustering for different URL families.

Computer Science, Session 2 – 10:15-11:00, Room H-C10

Fluctuation of Wi-Fi Received Signal Strength Indicator Readings – Patricia Kibenge, School of Mathematical and Computational Sciences, UPEI (*Collaborators: Dr. Qiang Ye*)

Abstract: Current GPS-based technologies used in smartphones that allow for continuous or frequent localization quickly drain the battery of these devices. A convenient, energy-efficient alternative relies on Wi-Fi fingerprinting, where a training phase is used to populate a database of locations and corresponding Wi-Fi Received Signal Strength Indicators (RSSIs) before those values are compared to the target location's RSSI values in the locating phase. Crowdsourcing is an efficient way to complete the training phase if this paradigm is used in a large area. The consistency of the RSSIs collected by all devices used in crowdsourcing the training phase is integral to the accuracy of the Wi-Fi fingerprinting method. In addition to dependence on the strength of the signal created by the access point, RSSI levels are dependent on the design of the antenna, which may differ from one device to the next. This research attempted to determine whether a consistent mathematical relationship exists between RSSI readings collected by three different Android smartphones. No reproducible mathematical relationship was found within the results but all smartphones recorded RSSIs that correlated with their distance from the access points, which is essential for this localization method. Therefore, adjusting RSSI values with a mathematical formula is not an acceptable method of RSSI normalization between readings from different devices. However, using Wi-Fi RSSIs for mobile localization is still a viable alternative to GPS.

Understanding the Design of Non-Verbal Communication Tools through Studying Games – Jason Wuertz, Faculty of Computer Science, UNB (*Collaborators: Dr. Scott Bateman*)

Abstract: Communication and collaboration are central aspects of many everyday tasks, both physically and digitally. Multiplayer games are no exception, and with the stakes for professional gamers climbing higher and higher, insight into the design and use of collaboration and communication tools in distributed multiplayer games have become of increasing interest. Multiplayer games are among the most successful software applications for allowing people to communicate and coordinate and therefore studying these games can help us design better communication tools in both games and groupware applications (multiuser applications, like Google Docs) can provide valuable new insights. Our work focuses on non-verbal communication and our entry point into this area was finding what information the players of DotA 2 share using in-game gestures. This provided us with insight into the communication needs of teams in multiplayer games. To further this understanding we conducted a second study. Via a grounded theory analysis on distributed, multiplayer games we created a design framework for automatic, team awareness cues. This work focused on the information that games automatically make available to players in order to foster teamwork and collaboration. To validate and further our previous findings, we aim to build a generalizable, open-source MOBA (Multiplayer, Online Battle Arena) as an experimental environment. The aim of this work is to enable researchers to quickly create and test alternative communication tools in a multiplayer environment. Our research program has been designed to identify how fast and effortless communication can be designed into common, everyday digital collaborations.

Experimental Verilog Synthesis Features for Odin & Alternative Multiplier Hard-Block for FPGA – Jean-Philippe Legault, Faculty of Computer Science, UNB (*Collaborators: Dr. Kenneth Kent, Panos Patros, Aaron Graham*)

Abstract: The Odin project is part of a Verilog to routing workflow. The FPGA industry already possesses a plethora of proprietary software to synthesize a logic system onto an FPGA. The goal of the VTR project is to bring an open-source alternative to the table and permits the designer to also model the FPGA with an XML file. Odin is part of the tooling of VTR and is responsible for parsing Verilog files and an XML architecture description of an FPGA to output a netlist for further processing. Odin allows the user to simulate the execution of a circuit and verify correctness.

This thesis will explore the input driven shift operation in regard to FPGA placement performance, size, routing, and path delay. We propose a replacement for the current multiplier hard-block that enables both input driven shift operations and multiplication with better performance than current multiplier and soft logic input driven shifter. The proposed model is fracturable for both shifting and multiplier making it a direct replacement to the current multiplier.

Computer Science, Session 3 – 1:40-3:15, Room H-C10

The Development of SpaceLaunch: A Faster App Launcher Developed in Unity – Nathaniel Brewer, Faculty of Computer Science, UNB (*Collaborators: Dr. Scott Bateman, Manasi Shah, Dr. Carl Gutwin*)

Abstract: In this presentation, I will discuss my experience developing SpaceLaunch -a smart phone application launcher that allows people to learn app locations more easily than conventional app launchers. Previous work has demonstrated that retrieval of digital items is faster when people are provided with an overview of an entire information space. SpaceLaunch provides an overview of all apps, allowing users to utilize spatial memory when learning the locations digital items. Current methods of organizing smart phone applications (e.g., folders and pages) don't allow users to take full advantage of spatial memory. SpaceLaunch is designed to utilize this important aspect of memory to decrease application location learning and retrieval time. To empirically test SpaceLaunch, we developed two other application launchers; one mimicking the folders organizational strategy and the other mimicking the pages strategy.

In this talk I will focus on my experience using the Unity game engine to develop and test the SpaceLaunch prototype. I will outline our use of Unity in the context of Human-Computer Interaction (HCI) experimental prototype development, discussing issues that arose and how they were overcome. SpaceLaunch and its competitors were developed using the Unity game engine. Unity is a popular game engine with a large on-line community that supports 2D and 3D games. It also allows for deployment to over 25 platforms (e.g., Android, iOS, HoloLens, PS4, etc.). The comprehensive tools (including IDE, community, Room C# support, and build tools) associated with Unity provide opportunities for HCI researchers to rapidly develop, deploy and test experimental HCI systems.

Improving Estimage: Image to Crop Yield Estimation using Convolution Neural Networks – Tanya, Acadia University (*Collaborators: Dr. Daniel L. Silver*)

Abstract: Convolutional Neural Networks have advanced the state of the art in many machine learning applications such as computer vision, speech recognition and natural language processing. Object counting in natural images is a challenging problem that has recently received much attention in computer vision and machine learning. The proposed research will develop and empirically compare new methods for object counting using convolution neural networks. The talk will review our plans to improve an existing system that employs an Android application and associated cloud service to capture an image of a crop such as grapes on the vine and estimate its yield (by volume or weight) from the image.

Modeling Synthetic Perception for Virtual Agents – Suzannah Smith, Faculty of Computer Science, UNB (*Collaborators: Sven Seele, Dr. Rainer Herpers*)

Abstract: Representing humans in virtual environments, like games and simulations, is necessary in order to make the experience more immersive or even realistic to users. However, creating a virtual agent that exhibits seemingly human-like behavior is a very complex task. This presentation explores how to improve one of the aspects involved in achieving this task: getting an agent to perceive its synthetic environment similarly to how a real human would perceive the real world.

Previous work on the project included implementing a virtual agent and simulating its perception. The agent navigates through a scene while gathering information by attending to salient objects. Virtual visual and audio sensors gather information, which is transferred to the agent's memory. An attention process determines which stimuli to focus on, and gradually builds up a working memory of the scene. Attended stimuli are stored as percepts in short-term and possibly long-term memory.

We developed an evaluation scenario that puts human subjects into a virtual space, using an Oculus Rift with an integrated professional eye tracker by SMI (https://www.smivision.com/). Head and gaze direction data was recorded for each subject in a preliminary user study for two reasons: (1) The data can be analyzed to infer realistic parameters for the same situation that an agent is facing. These parameters include, e.g., dwell times, fixation points, and relative object saliencies. (2) In future work, the data can be recorded for both humans and agents to evaluate whether the simulated agent model realistically emulates human perception.

Evaluating Systems for Data Science – Alex Watson, Faculty of Computer Science, UNB (*Collaborators: Dr. Suprio Ray*)

Abstract: The volume of data that is generated each day is rising rapidly. There is a need to analyze this data and produce applicable results efficiently. Data science offers a formal methodology for processing and analyzing data. It involves a work-flow with multiple stages, such as, data collection, data wrangling, statistical analysis, machine learning analysis as well as modelling and communicating meaningful information

from the data. In this talk, I will introduce a data science benchmark, to evaluate systems with data processing and analytics tasks. I will describe and evaluate a variety of current commercial and open-source data analytics systems. These systems differ significantly in terms of available features, performance and scalability.

Assistance Techniques in Video Games – Jawad Jandali Refai, Faculty of Computer Science, UNB (*Collaborators: Dr. Scott Bateman, Dr. Michael Fleming*)

Abstract: Video games can be challenging. For some, that challenge is part of the fun of playing the game. For others, those challenges can make the game frustrating or even un-playable. To bridge the gap between skilled players and those who are less skilled, or to make a single player game more accessible to novices, developers often add options to make parts of the game easier for the player.

The problem is that games are diverse, with unique mechanics and using different controls. So, how can a game designer know exactly how to provide the right type of assistance to a player? Previous work has investigated assistance techniques for specific genres, such as racing and first-person shooters. Those papers provide effective ways of providing assistance, and the pros and cons of different assistance techniques. However, previous work has focused on only one aspect of games (e.g., controls and not mechanics) or on only one setting for gaming (e.g., multiplayer games and not single-player games).

In our research, we propose a compilation of all assistance techniques spanning 'core' tasks of gaming (the basic unit of interaction with a game). Thus far we have identified 42 assistance techniques across 11 core tasks, only 15 of which have been explored in previous research. We are planning to evaluate the newly discovered assistance techniques in mini-games that we build. The result of this research will be a design framework to guide game developers in incorporating assistance techniques in the games.

Improving gait through sensored cane feedback – Ian Smith, Faculty of Computer Science, UNB (*Collaborators: Dr. Scott Bateman, Dr. Erik Scheme, Dr. Satinder Gill*)

Abstract: Mobility impairment is a major concern for aging individuals, requiring many to use assistive devices such as canes to improve stability and mobility. However, there is a disconnect between how people use canes, and how their use has been prescribed by health care professionals. For example, patients are typically instructed to offload a target percentage of their body weight while an affected limb is in contact with the ground. However, it is difficult for a patient to consistently know how to offload specific a specific amount of weight to a cane. Training an individual to comply with these constraints is, understandably, an inaccurate process requiring the continued attention of a trained therapist. This means that canes are most often not used as prescribed. To address such problems, we have developed a multi-sensor smart-cane for monitoring usage and providing detailed feedback. The SmartCane system collects kinematic and loading information, providing a detailed description of how patients use their canes. Collected usage data can be analyzed and presented back to patients and health care professionals in two ways. First, historical metrics can be easily accessed by patients and health care professionals to track patient progress over time. Second, feedback can be provided directly to patients in real-time to correct gait and loading problems as they use their canes. In this presentation, we will provide an overview of the development and evaluation of the SmartCane and its potential to improve the use of this common assistive device.

Mathematics, Session 1 – 8:30-10:00, Room H-C11

Strongly Connected Node Reliability – Emily Wright, Dept. of Mathematics, MSVU (*Collaborators: Dr. Danielle Cox*)

Abstract: Let G be a digraph and assume that the vertices operate independently with probably $p \in [0, 1]$. The *strongly connected node reliability* of digraph G, scNRel(G, p) is the probability that all the operational vertices can communicate. In this talk we will discuss the relationship to the traditional node reliability model, compute reliability polynomials for some of families, present preliminary results on the shape of the strongly connected node reliability polynomial and plans for future work. This is joint work with Danielle Cox.

The Fire Break Problem on Graphs (Part 1) – HongChang Bao, Dept. of Computer Science, UNBSJ (*Collaborators: Jennifer McNichol, Dr. Tim Alderson, Dr. Andrea Burgess*)

Abstract: We let G be a connected graph and assume that a fire will break out at one vertex of G. So which vertices do we need to protect in order to save the maximum number of other vertices? In this talk I will discuss the case where we know the fire is going to break out. My talk will focus on Path, Room Cycle, Wheel, Complete bipartite graph and K-regular bipartite graph. I will also talk about a special K-regular bipartite graph with which we can save almost half of the vertices when we only protect two vertices. This is a joint work with Jennifer McNichol and is supervised by Dr. Tim Alderson and Dr. Andrea Burgess.

The Fire Break Problem on Graphs; Part 2 – Jennifer McNichol, Dept of Mathematics and Statistics, UNBSJ

Abstract: A fire will break out on a vertex of a connected graph and immediately spread to all adjacent vertices. We do not know which vertex the fire will originate at. Given that we are able to protect a fixed number of vertices, which vertices do we choose to protect in order to save as much of the graph as possible? In this talk we discuss solutions to this problem on a variety of graphs.

Applying Graph Theory to Network Defense Strategies – Andrew Moreira, Dept. of Mathematics and CS, Mt A (*Collaborators: Dr. Margaret-Ellen Messinger*)

Abstract: This research, inspired by the eternal vertex cover problem, examines the properties of a graph parameter $\varepsilon_m^\infty(G)$. We begin with the formulation of general bounds and the classification of $\varepsilon_m^\infty(G)$ as an N=2 m-eternal vertex cover. We then introduce the static condition, $\alpha_2(G)$,

and open the investigation between the static and dynamic condition. Additionally, we examine upper and lower bounds for paths P_n, cycles C_n, complete graphs K_n, complete bipartite graphs $K_{m,n}$, and trees T. Finally, we classify upper bounds for some Cartesian products G=($H \square L$), where *H,L* are graphs belonging to specific classes.

A Link Between Music Theory and Graph Theory – Jordan Barrett, Dal (*Collaborators: Dr. Jason Brown*)

Abstract: What makes a musical rhythm sound good? This fundamental question in music theory can be modeled by discrete methods in mathematics. Consider a hit song with an underlying beat. This beat normally cycles in groups of 16 short intervals, and so we could represent this beat as a directed cyclic graph C with 16 vertices, and with a set S in V(C) consisting of all the "hits" in the beat. In this talk we will explore 6 of the most popular rhythms in music history and determine which of these 6 dominates popular music. This will lead to a graph theoretic interpretation of the fundamental question: what makes a good rhythm good. We will see a connection between the "catchiness" of a rhythm and the "evenness" of points on the corresponding cycle. We will also ask the larger question of how to choose vertices on an arbitrary graph in the most "even" way possible, which will tie into the infamous facility location problem from computational geometry.

No knowledge of music theory is needed, and only a basic core understanding of graph theory will be assumed.

A Philanthropic Variant of Parallel Chip-Firing – Maggie MacPhee, Dept. of Mathematics and CS, Mt A (*Collaborators: Dr. M. Messinger, Dr. J. Howell, Rebecca Ryan*)

Abstract: Imagine you have a network representing people and their relationships with one another. Each person initially has some amount of money. Now the richest person, or people, give a dollar to each of their neighbours and then the process is repeated. No money is ever added, which means the system is closed, so it is natural to ask the question: "What is the long term behaviour of the system?" In order to answer this, we must look at the system more formally. If you let each person be represented by a vertex on the graph, you can assign each vertex a finite integral number of chips to represent money. Now, in each round, the vertices with the largest number of chips would "fire" to their neighbours. This is a variant of Parallel Chip-Firing we have called Maximal Chip-Firing. Our primarily goal is to look at the long-term behaviour of the Maximal Chip-Firing game on graphs.

Mathematics, Session 2 – 10:15-11:00, Room H-C11

The Damage Number of Circulant Graphs – Marina Muberwa, MSVU (*Collaborators: Dr. Danielle Cox*)

Abstract: Cops and Robber is a game that was first examined by Quilliot in his doctoral thesis and separately by Winkler & Nowakowski. It has two players: one controls the set of cops and the other controls the robber. In this presentation theorems regarding the damage number of circulant graphs will be discussed.

Eternal Domination of $6 \times n$ Grid Graphs – Dylan Coakley, Dept. of Math, Stat, Room CS, StFX (*Collaborators: Dr. Stephen Finbow, Dr. Martin van Bommel*)

Abstract: Consider the vertices of a set D of vertices of a graph as positions of guards. An eternal dominating set has the property that allows eternal protection of the graph in the following way: when a vertex v not in D is attacked, a guard at a vertex in D adjacent to v must move to v in order to defend it from the attack; the rest of the guards either remain stationary, or move to an adjacent vertex such that they produce a new set D' with the same property of eternal protection. Results for eternal domination have been established for the graph products $P_3 \times P_n$, $P_4 \times P_n$, and $P_5 \times P_n$. This talk will discuss findings for the graph $P_6 \times P_n$.

The Watchman's Walk Problem on Graphs – Brittany Pittman, Dept. of Math, MUN (Grenfell) (*Collaborators: Dr. Rebecca Milley*)

Abstract: The Watchman's Walk Problem attempts to find a minimum closed dominating walk (MCDW) in a connected graph. This walk can be thought of as an optimal route through which a single watchman can see every vertex in the graph. This problem was initially introduced by Hartnell, Rall, and Whitehead in 1998. Since then, variations of the problem have emerged for when multiple guards are available, and either the maximum unseen time of the vertices or the number of available guards are fixed. In this talk, we will consider the traditional Watchman's Walk problem, as well as its variations. First, we will consider a variation of the problem in which the number of available guards is fixed, and the objective is to find a route that minimizes the maximum unseen time of the graph's vertices. For this variation, we will also consider routes that minimize the average unseen time of a subset of vertices in an acyclic graph. Finally, we will consider the properties of graphs that remain as leave graphs formed by removing the edges of a minimum closed dominating walk in another graph, the multiplicity of MCDWs in outerplanar graphs, and graphs that can be decomposed into Watchman's Walks.

Mathematics, Session 3 – 1:40-3:15, Room H-C11

Peaceably Coexisting Armies of Queens Game, and a Variation – Katie MacEachern, Dept. of Math, Stats, Room CS, StFX (*Collaborators: Dr. Martin van Bommel, Dr. Stephen Finbow*)

Abstract: The *Peaceably Coexisting Armies of Queens Game* is a two player game in which each player takes turns placing queens on an $n \times n$ chessboard such that no two queens from opposing armies may attack each other. Play continues in this manner until one player is unable to

place a queen, making their opponent victorious. Analysis of this game, which focuses on determining who wins, and in how many moves, shows the first player can always be victorious, regardless of board size. As a game that should always end in a first player victory is uninteresting, a slightly different variation is proposed, that of *Peaceably Coexisting Armies of Queens Game with a Single King*. This game is played similarly to the original, but on their first move the first player must place a single king that is a member of their army. Analysis of this variation proves to be a more compelling exercise.

The Game of Sandwich Cops and Robber – MacKenzie Carr, Dept. of Mathand Stat, Acadia (*Collaborators: Dr. Nancy Clarke*)

Abstract: The game of Cops and Robber is a vertex-to-vertex pursuit game played on a simple, undirected graph. The game of Sandwich Cops and Robber is a variation of this game in which the robber is located on the edges of the graph. In order to capture the robber, the cops must move so that at least one cop occupies each endpoint of the edge on which the robber is located. We compare this new game with the original game of Cops and Robber, examining the relationship between the copnumber and the sandwich copnumber of particular classes of graphs, including graph products. We identify certain properties of graphs that cause the sandwich copnumber to differ from the copnumber.

Matroids: Generalizing Independence – Malhar Mukhopadhyay, MUN (Grenfell) (*Collaborators: Dr. Yevgeniy Vasilyev*)

Abstract: In this talk, I will introduce the notion of matroid, a structure that generalizes the concept of linear independence. We begin with a basic example of a linear matroid, one that can be represented by the set of columns in a real valued matrix or a matrix over a finite field. Next, we describe how matroids arise in graph theory. Moreover, we explain why a graph matroid is linear, using an example associating a matrix to a graph matroid, and introduce dual matroids (co-graphic matroids). We look at the linear matroid known as Fano plane, and show why it cannot be represented by a graph. We take a brief look at the Vamos matroid as well, and explain why it is not linear.

Non-linear Cartesian differential categories (part I) – Hayley Reid, MtA (*Collaborators: Jonathan Bradet-Legis, Dr. Geoff Cruttwell*)

Abstract: Cartesian differential categories are a recent idea that formalize mathematical settings which contain an operation resembling the derivative. The standard example is ordinary differentiation, but many other examples have also been identified, including the "differential lambda-calculus" and some aspects of Goodwillie's "Calculus of functors".

An almost-example in Abelian group theory has also been identified: it satisfies all of the Cartesian differential category axioms but one. In this talk, we look at what the almost-example is, and how it leads to the more general notion of a "non-linear Cartesian differential category".

Non-linear Cartesian differential categories (part II) – Jonathan Bradet-Legris, MtA (*Collaborators: Hayley Reid, Dr. Geoff Cruttwell*)

Abstract: It is possible to define differential forms and their exterior derivative in the general setting of a Cartesian differential category. However, by subtly modifying these definitions, we show it is also possible to define a similar structure in non-linear Cartesian differential categories. As with the ordinary exterior derivative, the resulting structure forms a cochain complex. This result may give an new invariant for Abelian groups.

Calderón reproducing formula in dimension three – Finlay Rankin, Dal (*Collaborators: Dr. Keith F. Taylor*)

Abstract: If H is a closed subgroup of $\mathrm{GL}_3(\mathbb{R})$, the group of 3×3 invertible real matrices, we can use the action of H on $\mathbb{R}^3$ to form a new locally compact group $\mathbb{R}^3 \rtimes H$. If there is an open free orbit when H acts on $\mathbb{R}^3$, then H is called strongly admissible and such groups give rise to generalizations of the continuous wavelet transform via the natural action of the semi-direct product group $\mathbb{R}^3 \rtimes H$ on $\mathbb{R}^3$. The strongly admissible subgroups of $\mathrm{GL}_3(\mathbb{R})$ have recently been classified. We will present the Calderón reproducing formula in 16 different cases, each giving a different generalization of the continuous wavelet transform to three dimensions.

Mathematics, Session 4 – 1:40-3:15, Room H-C9

Entropy and entanglement in quantum gravity: a delightful romp through mathematical physics – Courtney Allen, Dept. of Math & Stats, UNBF (*Collaborators: Dr. Viqar Husain*)

Abstract: Quantum gravity is a physical theory that endeavours to provide a quantum mechanical description of gravity. The purpose of this project was to determine the gravitational field produced by the quantum state of a system, and how gravity affects a property of these states known as entanglement. An entangled system is one whose particles are intertwined and cannot be thought of separately. The central question this project aimed to answer was: if a system begins in a product state–a system that we think of as non-entangled–does the system become entangled over time? The Newton-Schrödinger equation was used to model this situation. This equation couples the Schrödinger equation, which describes the time-evolution of a quantum system with an initial state ψ, with the Poisson equation, which determines the Newtonian potential φ from the quantum state. Both the quantum state of the system and the gravitational field evolve over time under the Newton-Schrödinger equation. The Crank-Nicholson scheme, a method used to approximate derivatives in the Schrödinger equation, was used to produce numerical solutions to the Newton-Schrödinger equation. After plotting how the system evolves over time, the entanglement of the system can be determined by calculating its entropy. A product state has zero entropy, while an entangled state has some non-zero entropy. Thus, by calculating the entropy of the system, the entanglement can be determined.

An Introduction to Quantum Computing – Marzieh Bayeh, Dal

Abstract: This is an introductory talk on quantum computing and the current research project related to this topic at Dalhousie University.

Modelling the Water Circulation in the Bras d'Or Lakes – Sarah Walsh, Dept of Math, Room CBU (*Collaborators: Dr. Bruce Hatcher*)

Abstract: The Bras d'Or Lakes is a complex estuarine system centered in Cape Breton Island. The Bras d'Or Lakes is home to a number of diverse species and has many unique physical properties. To better understand the water circulation, and to eventually aid in the prediction of species distribution and the spread of diseases, we turn to the coastal modelling process. To understand the modelling process, we must obtain a solid grasp of the underlying forces involved in moving the water. We also spend time describing the 3D primitive equations which govern the model. Other topics of interest include determining boundary and initial conditions, and which grid and discretization method to use. Finally, we briefly consider the appropriate coordinate system, stability, and the balance between resolution and computing resources. The current model we focus on is based off the CANDIE ocean model. This model was created in 2007 and uses the 3D primitive equations and a nested-grid. This model is a free surface, open top, and z-level model that can simulate passive particle tracking as well as reasonably reproduce salinity and temperature observations. Recently, there has been discussion in creating a new model based off the Regional Ocean Modelling System and to eventually run it on ACENET.

The Quantum Monodromy of the Kodaira-Thurston Manifold – Zoé McIntyre, Dept. of Math & CS, MtA (*Collaborators: Dr. Mark Hamilton*)

Abstract: The Kodaira-Thurston manifold M_Θ is an object arising in symplectic geometry that can be described as the quotient of $\mathbb{R}^4$ by the action of a group Γ of affine transformations, or equivalently, as a torus bundle over the 2-torus T^2 with projection $\pi : (x, y, z, t) \mapsto (x, t)$ and monodromy in the x direction given by $(y, z) \mapsto (y, z + y)$. Classical monodromy is a topological invariant that acts as an obstruction to the existence of globally defined action-angle coordinates, although such coordinates can always be defined locally. This 'barrier to globality' is reflected in the spectrum of quantum states obtained through quantization of a phase space, with *quantum* monodromy acting as an obstruction to the existence of a globally defined lattice of quantum states. In order to investigate the quantum monodromy of M_Θ, we use the recipe of geometric quantization to compute action coordinates a_1 and a_2 on the base space T^2 of π, then identify points on the torus for which $(a_1, a_2) \in \mathbb{Z}^2$. These points (the quantum states) form a local lattice for which we define a basis $\{u_1, u_2\}$, but traversing a loop in the x direction gives us a new, different basis $\{v_1, v_2\}$ for this same lattice. Defining quantum monodromy to be the linear transformation mapping u_j to v_j (for $j = 1, 2$), we find that both the classical and quantum monodromy of M_Θ can be associated with the matrix $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$.

What is a Fractafold and how can I make one? – James Eckst, Dept of Mathmatics, Dal (*Collaborators: Dr. Dorette Pronk*)

Abstract: Fractafolds can be defined as being to a fractal what a manifold is to a Euclidean half-space. We are looking at building closed fractafolds in subspaces of $\mathbb{R}^2$ glueing IFS fractals, in particular Sierpinski's Gasket. This talk will go over the rules to build these fractafolds, how we can represent them using graphs, and why Sierpinski's Gasket is ideal for closed fractafold.

Statistics, Session 1 – 8:30-10:00, Room H-C9

Multistate models for examining adherence to cervical cancer screening in Ontario – Faith Lee, Dept. of Statistics and Actuarial Science, Waterloo (*Collaborators: Dr. Rinku Sutradhar, Dr. Lawrence Paszat*)

Abstract: Cervical cancer screening has decreased cervical cancer incidence rates by allowing for the early detection of abnormalities. In July 2015, 36% of the eligible women in Ontario were overdue for screening. There is a need for the implementation of appropriate statistical methods to understand factors associated with screening adherence so that target populations requiring encouragement to participate in screening can be identified.

We studied individual- and physician-level characteristics associated with screening adherence through the implementation of a time non-homogeneous multi-state model on a province-wide longitudinal cohort (N = 1,156,720). The data was extracted from Ontario Health Insurance Plan Registry (OHIP) with index dates from 1991 to 2013. Characteristics were incorporated into the model as time-varying covariates. We were interested not only in how these covariates were associated with the rate of becoming adherent but also the disparity between immigrants and Canadian residents.

Individuals with fewer comorbidities, higher income, prior visits to a primary care physician, and residential stability had higher rates of becoming adherent. The rate of adherence was 6.9% lower among immigrants compared with non-immigrants. After including prior screening history into our regression model, we found that women who had smaller proportions of prior time spent non-adherent had higher rates of becoming adherent in the future.

The implementation of a multi-state model allows us to monitor trends longitudinally and to provide insights so we can mobilize healthcare resources efficiently to minimize the time spent non-adherent in the appropriate populations.

Testing for Genetic Association in the Presence of Competing Risks – Lucas Hynes, Dept. of Math and Stats, MUN (*Collaborators: Dr. Candemir Cigsar, Dr. Yildiz Yilmaz*)

Abstract: Competing risks arise in a study when subjects may experience different types of events and experiencing an event of one type precludes the observation of events of other types. The statistical analysis of time-to-event data in the presence of competing risks requires specific methods. Therefore, competing risks analysis is usually considered as an extension of the classical survival analysis. In this talk, I will first briefly discuss some concepts and methods in the survival analysis and some issues related to them in the presence of competing risks, and then review important methods developed for the analysis of competing risks data. I will consider competing risk analysis methods in genetic association studies and compare performance of the methods through a simulation study.

Finding Common Ground in the Human Gut: Batch Effects in Genomic Data – Molly Hayes, Dept. of Math and Stats, Dal (*Collaborators: Dr. Hong Gu, Dr. Morgan Langille*)

Abstract: Communities of microbes inhabit all areas of the environment, including the body cavities and exterior surfaces of larger organisms. These communities vary widely in composition; in the human gut microbiome, richness, diversity, and distribution of bacterial taxa are highly disparate among individuals. However, mounting evidence suggests that the gut microbiome interacts with host systems in myriad ways, and that understanding these associations could have profound implications for our ability to predict, diagnose, or treat certain pathologies. The most common way to characterize the bacterial microbiome is to determine the abundance of each taxon present in fecal DNA samples using 16S targeted-gene sequencing, which allows bioinformatic algorithms to use the ubiquitous 16S rRNA gene to identify taxa by comparing the observed sequences to those in a reference database. Unfortunately, these data are rife with obstacles to classical analyses: perhaps most critically, our conclusions are limited by the fact that data from similar studies-even those using the same samples-cannot be compared to one another due to dominating 'batch' effects that arise from differences in reagents, sequencing platforms, bioinformatic pipelines, and other unknown variables. In this presentation, I will 1) introduce the clinical importance of developing robust statistical analyses for 16S data from the human gut microbiome, 2) propose common-factor analysis as a framework to facilitate meaningful 16S data analysis in microbial ecology and translational microbiomics, 3) provide the results of our promising preliminary work, and 4) outline the next steps of our research.

Non-Taxonomically-Restricted Variable Selection for Prediction in Microbiome Data Analysis – Lihui Liu, Dept. of Math and Stats, Dal (*Collaborators: Dr. Hong Gu, Dr. Toby Kenney, Dr. Johan Van Limbergen*)

Abstract: Recent advances in sequencing technology have allowed the collection of extensive data sets describing the abundance of all microbes present in a sample. This leaves a serious challenge for the data analysis — sorting through these large data sets to determine the differences between samples from different groups (such as patients vs. controls). Microbial marker gene data usually contains information across multiple taxonomic levels with hundreds or thousands of different OTUs measured for each sample. Analysis is often performed at genus or phylum level. Here we present a method for selecting which of the microbes represent the key differences between the groups. Our method is able to select the significant microbes at varying levels of generality, depending what is implied by the data. For example, it is able to select an entire phylum of bacteria, or to identify a single species (or even a single strain if data are available at this level of detail). The method is called Subsampling Ranking Forward selection (SuRF). It is based on LASSO penalised logistic regression, and uses a subsampling approach to rank predictors, then a forward selection approach on the ranked list by a sequential permutation tests . We present simulations in multiple sparse settings to demonstrate our approach performs better than other existing approaches in recovering the true variables. We applied SuRF to pouchitis IBD data and the well-known moving picture of human microbiome data and we found SuRF can provide a better or comparative prediction while controlling the false positive rate.

Gene Clustering Based on Co-occurrence with Correction for Common Evolutionary History – Chaoyue Liu, Dept. of Math and Stats, Dal

Abstract: Homologous genes in prokaryotes can be described using phylogenetic profiles which summarize their patterns of presence or absence across a set of genomes. Phylogenetic profiles have been used for nearly twenty years to cluster genes based on measures such as the Euclidean distance between profile vectors. However, most approaches do not take into account the phylogenetic relationships amongst the profiled genomes, and overrepresentation of certain taxonomic groups (for example, pathogenic species with many sequenced representatives) can skew the interpretation of profiles. We propose a new approach that uses a coevolutionary method defined by Pagel to account for the phylogenetic relationships amongst target organisms, and a hierarchical-clustering approach to define sets of genes with common distributions across the organisms. The clusters we obtain using our method show greater evidence of phylogenetic and functional clustering than a recently published approach based on hidden Markov models. Our clustering method identifies sets of amino-acid biosynthesis genes that constitute cohesive pathways, and motility / chemotaxis genes with common histories of descent and lateral gene transfer.

AARMS Workshop

Sunday's workshop on Statistical Learning and Health Data Analytics is organized by an AARMS Collaborative Research Group (CRG), Statistical Learning for Dependent Data under the Administration of Ying Zhang. Contact Ying Zhang, Acadia University ying.zhang@acadiau.ca for any questions.

About our Collaborative Research Group (CRG)

Real-world problems are constantly challenging us to invent new statistical methods. Our group members are actively involved in methodological research and have numerous multidisciplinary collaborations tied to real-world problems. All the members have active interdisciplinary research programs funded externally by individual and/or partnership grants, are active in HQP training (undergraduate/graduate/PDF), and are publishing high quality papers. For the past decade, the group members have been solving problems involving complex dependent data, big data, and statistical learning, in important application areas such as environmental sciences, manufacturing quality control, public health, medical science research and bioinformatics. There have been active and successful collaborations among the members, modelling longitudinal, spatial, and clustered dependent data (Ma, Hasan, Yan, and Sneddon); testing trend in time series data (Cabilio and Zhang); data mining (Gu, Kenney, and Chipman); and bioinformatics/biostatistics (Gu, Kenney, Chipman, Zhang, and Peng).

The primary objective of the CRG is to develop a collaborative research program to share resources and coordinate activities in order to

(a) address emerging statistical learning methods and computing issues motivated by multidisciplinary collaborations related to big data with complex dependency structure.

(b) develop and distribute novel statistical software, popularizing our research tools in both statistics and application areas.

(c) encourage collaboration, the exchange of ideas and joint supervision of HQP at all levels (undergraduate, graduate, PDF) among universities and in partnership with external organizations.

List of group members (alphabetical order): Paul Cabilio (Acadia), Hugh Chipman (Acadia), Hong Gu (Dalhousie), Tariqul Hasan (UNB), Toby Kenney (Dalhousie), Renjun Ma (UNB), Jianan Peng (Acadia), Gary Sneddon (MSVU), Guohua Yan (UNB), Ying Zhang (Academic Administrator, Acadia)

Sponsors

We gratefully acknowledge the generous support of the following associations:

- **Science Atlantic**
- **The Atlantic Association for Research in the Mathematical Sciences**
- **ACEnet**

and the following offices of UNB Fredericton:

- **The Dean of Computer Science**
- **The Dean of Science**
- **The Department of Mathematics and Statistics**
- **The School of Graduate Studies**
- **The Vice President Academic**
- **The Marketing Office**