Discrete Mathematics and Data Science Research Team Seminars

**2018 – 2021**

|  |  |  |
| --- | --- | --- |
| Date and Time | Speaker | Seminar |
| 20 April 2021  14:00-15:00 | Dr Vadim Zverovich (University of the West of England, Bristol) | **Prevalence of Braess' Paradox?**  The well-known Braess’ paradox illustrates situations when adding a new link to a transport network might not reduce congestion in the network but instead increase it. This is due to individual entities acting selfishly/separately when making their travel plan choices and hence forcing the system as a whole not to operate optimally. Deeper insight into this paradox from the viewpoint of the structure and characteristics of networks may help transport planners to avoid the occurrence of Braess-like situations in real-life networks.  A generally accepted belief is that Braess’ paradox is widespread. This was confirmed by some researchers who claimed that the likelihood of the paradox is 50%, or even higher under some assumptions. In this talk, we will discuss our recent results devoted to the probability of Braess' paradox to occur in the classical network configuration introduced by Braess. |
| 27 November 2020  14:00-15:00 | Dr Angela Mihai (School of Mathematics) | **Likely instabilities in liquid crystal elastomers**  In this talk, I will present stochastic material models described by strain-energy densities where the parameters are characterised by probability distributions at a continuum level. To answer important questions, such as “what is the influence of probabilistic parameters on predicted mechanical responses?” and “what are the possible equilibrium states and how does their stability depend on the material constitutive law?”,  I will focus on likely instabilities in nematic liquid crystal elastomers. I will discuss the soft elasticity phenomenon where, upon stretching at constant temperature, the homogeneous state becomes unstable and alternating shear stripes develop at very low stress, and some classical effects inherited from the underlying polymeric network, such as necking, cavitation, and shell inflation instabilities. These fundamental problems are important in their own right and may stimulate related mechanical testing of nematic materials. |
| 24 November 2020  14:00-15:00 | Dr Angela Mihai (School of Mathematics) | **Likely instabilities in liquid crystal elastomers**  In this talk, I will present stochastic material models described by strain-energy densities where the parameters are characterised by probability distributions at a continuum level. To answer important questions, such as “what is the influence of probabilistic parameters on predicted mechanical responses?” and “what are the possible equilibrium states and how does their stability depend on the material constitutive law?”,  I will focus on likely instabilities in nematic liquid crystal elastomers. I will discuss the soft elasticity phenomenon where, upon stretching at constant temperature, the homogeneous state becomes unstable and alternating shear stripes develop at very low stress, and some classical effects inherited from the underlying polymeric network, such as necking, cavitation, and shell inflation instabilities. These fundamental problems are important in their own right and may stimulate related mechanical testing of nematic materials. |
| 28 April 2020  15:10 - 16:00 | Dr Pavel Skums (Georgia State University) | **Methods of discrete mathematics in molecular epidemiology and outbreak analysis**  The COVID-19 pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is continuing its global spread and straining or overwhelming health care systems around the world. At the same time, long-standing epidemics caused by HIV, Hepatitis C (HCV) and other pathogens continue to be major causes of morbidity and mortality in the world. In the quest for adequate answers to those public health challenges, data science is becoming indispensable. Recent advances in biotechnologies brought to life the discipline of computational molecular epidemiology that uses the analysis of viral genomes to investigate outbreaks and understand epidemiological and evolutionary dynamics of pathogens. In my talk, I will discuss the problems and algorithmic results that arise from applications of graph theory and discrete optimization methods in molecular epidemiology studies of HIV, HCV and SARS-CoV-2. |
| 17 March 2020  14:00 to 15:00 | Dr Angela Mihai (School of Mathematics) | **Likely instabilities in stochastic elasticity**  The study of material elastic properties has traditionally used deterministic approaches, based on ensemble averages, to quantify constitutive parameters. In practice, these parameters can meaningfully take on different values corresponding to possible outcomes of the experiments. From the modelling point of view, stochastic representations accounting for data dispersion are needed to improve assessment and predictions. In this talk, I will discuss stochastic-elastic material models described by a strain-energy density where the parameters are characterised by probability distributions at a continuum level. To answer important questions, such as what is the influence of the probabilistic parameters on the predicted elastic responses? and what are the possible equilibrium states and how does their stability depend on the material constitutive law?, I will consider the cavitation and finite amplitude oscillations of hyperelastic spheres, and the soft elasticity of liquid crystal elastomers. These problems, for which the solution is tractable analytically, can offer significant insight into how stochastic-elastic models can be integrated into the nonlinear field theory. Similar approaches can be developed for other mechanical systems. |
| 12 November 2019  14:10 to 15:00 | Dr Penny Holborn (University of South Wales) | **Use of NLP techniques and machine learning algorithms for binary classification**  Natural Language Programming techniques are widely used in many applications from speech recognition to modern chatbots integrated into many of the devices and online services we use today. The recent success of NLP in solving complex real-world problems is mainly due to the availability of huge amounts of data and advances in machine learning algorithms. This talk will introduce fundamental knowledge and popular techniques utilised for gaining insight from textual data.  Two real world case studies will be presented, the first will look to introduce binary classification of real-world customer reviews. The aim being, to build a robust NLP text classification model that can classify customer reviews in real-time. The second, utilising word embedding for understanding talent data. Here, inferences are made on publicly available data to help understand current talent demographics and skills to help drive business strategy. |
| 31 October 2019 | Dr Pawel Dlotko (Swansea University) | **Topological data analysis: a tiny bit of theory, lot of intuition and a few applications**  Shape is one of the most fundamental concepts known to the humanity. We can recognize it in a number of non-related instances. But the language to quantify shape is largely unknown in the scientific community. In this talk I will try to fill in this gap by proposing a number of shape descriptors provided by Topological Data Analysis. While keeping the theory to absolute minimum, we will see a lot of intuition, implementation as well as real examples from physics, material science, neuroscience all the way to economy and political sciences. |
| 28 May 2019  14:10 to 15:00 | Dr George Theodorakopoulos  School of Computer Science & Informatics (Cardiff University) | **Privacy for location histograms: How to look like a tourist in your hometown**  A location histogram comprises the number of visits by a user to each location in a region of interest (restaurants, hospitals, cinemas, etc.). Such histograms are useful in location analytics for product recommendation and advertising, and also more generally for clustering and classification. However, disclosing a histogram may lead to inference of sensitive information about, e.g., the user's wealth level.  This talk will present joint work on protection algorithms for location histograms. We introduce two new privacy notions for individuals: sensitive location hiding and target avoidance/resemblance. The former aims to conceal all visits to a certain subset of locations that are deemed sensitive, whereas the latter aims to modify the histogram to make it look like any desirable histogram (e.g. a tourist's typical histogram) or to make it look as dissimilar as possible to a given histogram. For each privacy notion, we formulate an optimization problem that aims to maximize the corresponding notion, appropriately quantified, subject to a constraint on the acceptable quality deterioration of the histogram. We solve these problems optimally using a constrained shortest path algorithm, and we present heuristics that speed up the computation by at least two orders of magnitude while still being almost as effective as the optimal solution. |
| 18 March 2019 | Joe Paat | **The proximity function for IPs**  Proximity between an integer program (IP) and a linear program (LP) measures the distance between an optimal IP solution and the closest optimal LP solution. In this talk, we consider proximity as a function that depends on the right hand side vector of the IP and LP. We analyse how this proximity function is distributed and create a spectrum of probabilistic-like results regarding its value. This work uses ideas from group theory and Ehrhart theory, and it improves upon a recent result of Eisenbrand and Weismantel in the average case. This is joint work with Timm Oertel and Robert Weismantel. The proximity functions for IPs. |
| 25 February 2019 | Oded Lachish (Birkbeck, University of London) | **Smart queries versus property independent queries**  In the area of property testing, a central goal is to design algorithms, called tests, that decide, with high probability, whether a word over a finite alphabet is in a given property or far from the property.  A property is a subset of all the possible words over the alphabet. For instance, the word can be a book, and the property can be the set of all the books that are written in English; a book is 0.1 far from being written in English if at least 0.1 of its words are not in English. The 0.1 is called the distance parameter and it can be any value in [0,1]. The input of a test is the distance parameter, the length of the input word and access to an oracle that answers queries of the sort: please give me the i'th letter in the word.  The quality of a test is measured by it query complexity, which is the maximum number of queries it uses as a function of the input word length and the distance parameter, ideally this number does not depend on the input length. Tests that achieve this ideal for specific properties have been discovered for numerous properties. In general, tests that achieve the ideal for different properties differ in the manner in which they select their queries. That is, the choice of queries depends on the property.  In this talk, we will see that for the price of a significant increase in the number of queries it is possible to get rid of this dependency. We will also give scenarios in which this trade-off is beneficial. |
| 11 December 2018 | Dr Angelika Kimmig | **Probabilistic Logic Programming**  Reasoning with relational data, learning, and dealing with uncertainty are central to many aspects of Artificial Intelligence. Their combination is studied under a variety of names, and a broad range of languages and tools have been developed. Probabilistic logic programming achieves this combination by extending the representation and reasoning capabilities of logic programming to settings with uncertain data. This talk provides a gentle introduction to the field, and also touches upon applications and challenges. |
| 20 November 2018 | [Dr Andrei Gagarin](https://www.cardiff.ac.uk/people/view/279669-gagarin-andrei) | **Embedding graphs containing K5-subdivisions on the torus and proving non-toroidality**  Given a graph G, a classic problem is how to determine whether it is possible to draw G in the plane (on the sphere) with no edge crossings.  Such drawing of G in the plane would be a planar embedding. The torus is the sphere with a handle, i.e. an orientable topological surface of genus 1, which is closest to the sphere.  A similar problem is how to determine whether it is possible to draw G on the torus with no edge crossings, i.e. to obtain a toroidal embedding of G.  A toroidality testing algorithm usually starts with a (non-planar) sub-graph of G isomorphic to a subdivision of K5 or K3,3 and tries to extend one of its embeddings on the torus to an embedding of the whole graph G in all possible ways. We have shown a modification of this approach - non-planar graphs which don't contain certain types of K3,3-subdivisions are much easier to decide on their toroidality by decomposing them in accordance with the 'edges' of K5 in the K5-subdivision, testing planarity of resulting components, and eventually considering rearrangements of planar embeddings.  This provides an efficient method to handle the case of an initial K5-subdivision sub-graph in the graph. For general non-planar graphs containing K3,3-subdivisions, we show some particular examples to decide on their toroidality in an ad-hoc way. |
| 16 October 2018 | [Dr Martin Caminada](https://www.cardiff.ac.uk/people/view/473184-caminada-martin) | **Some Open Issues in Formal Argumentation Theory**  Formal argumentation theory has been one of the main topics in the area of non-monotonic reasoning for the last two decades. The idea is, roughly, construct arguments (which are defeasible inferences) from an underlying knowledge base (which consists of inference rules). Some of these arguments will then attack other arguments (for instance by having an opposite conclusion). The resulting directed graph (in which the vertices represent the arguments, and the edges represent the attack relation) is called an argumentation framework. Given such an argumentation framework, one then needs a graph-theoretical principle to determine which set (or sets) of arguments is justified. This principle (and there are many of them) is called an argumentation semantics. Once the set (or sets) of justified arguments has been determined, the justified conclusions (that is, the resulting logical inference) will be the conclusions of the justified arguments.  In the current talk, I will discuss some open issues when it comes to applying formal argumentation theory for the purpose of meaningful logical entailment. We will especially look at the consistency of the overall outcome, as well as what is called crash resistance and non-interference. Much of the presentation is based on one of the author's book chapters in the Handbook of Formal Argumentation. |
| 9 October 2018 | [Prof. Coralia Cartis (University of Oxford)](https://www.maths.ox.ac.uk/people/coralia.cartis) | **Dimensionality reduction techniques for global optimisation**  We show that the scalability challenges of Global Optimisation (GO) algorithms can be overcome for functions with low effective dimensionality, which are constant along certain linear subspaces. Such functions can often be found in applications, for example, in hyper-parameter optimisation for neural networks, heuristic algorithms for combinatorial optimisation problems and complex engineering simulations. We propose the use of random subspace embeddings within a(ny) global minimisation algorithm, extending the approach in Wang et al (2013). Using tools from random matrix theory and conic integral geometry, we investigate the success rates of our low-dimensional embeddings of the original problem, in both a static and adaptive formulation, and show their independence on the (large) ambient dimension of the problem. We illustrate our algorithmic proposals and theoretical findings numerically, using state of the art global solvers.  This work is joint with Adilet Otemissov (Turing Institute, London and Oxford University). |