Alan Turing Institute Scoping Workshop

Topological Data Analysis

Oxford, 10-11 September 2015

	Thursday		**Friday**
		09:00-09:55	Guest Lecture: **Herbert Edelsbrunner**
		10:00-10:55	**Ezra Miller**
10:30-11:25	*Registration*		*Coffee*
11:30-12:25	Exec Intro: **Robert Ghrist**	11:30-12:25	Software: **Michael Kerber**
	Lunch		*Lunch*
13:30-13:55	**Vitaliy Kurlin**		
14:00-14:25	**Jacek Brodzki**	13:45-14:25	Software Demo: **Nina Otter**
14:30-14:55	**Mason Porter**	14:30:14:55	**Subramanian Ramamoorthy**
15:00-15:25	How we need your help... **Chas Bountra, Bill Oxbury**	15:00-15:25	**Wojtek Chacholski**
	Tea		*Tea*
16:00-16:25	**Michael Faber**	16:00-16:25	**Vidit Nanda**
16:30-17:25	**Konstantin Mischaikow**	16:30-17:25	Panel discussion
18:30-21:00	Dinner		

The workshop dinner on Thursday evening will be held at the Cherwell Boathouse, 50 Bardwell Rd, Oxford OX2 6ST.

A light sandwich lunch will be provided on Thursday. For lunch on Friday participants may want to make use of the cafeteria in the AWB or one of the sandwich shops and cafes close by.

Lectures will be held in L4 of the Andrew Wiles Building, Oxford OX2 6GG, UK.

Abstracts and titles

Chas BountraTDA_A
Challenges in my area. How we need your help....

Jacek Brodzki
Title: Synchronisation problems on graphs and holonomy
Abstract: A synchronisation problem on a graph can be stated as follows. Given a function from the edge set of the graph to a chosen group, can one find a function from the vertex set of the graph with values in the same group such that, for every edge, the "difference" of the group elements associated to the ends of the edge equals the element associated with the edge? Starting with some very interesting ideas and results of Amit Singer and coworkers, I will introduce a very natural notion of holonomy that is helpful for deciding when such a problem has a solution. I will describe some natural generalisations of this problem based on work in progress jointly with Mukherjee and Gao.

Wojtek Chacholski
Title: Noise and multi persistence
Abstract: In this talk I will present a new method for extracting persistence information out of multi graded modules. This method builds on understanding mathematical properties of what one might regard as noise. Denoising can be then thought as the localisation away from the noise. In the talk I will focus mainly on stability aspects of these new invariants and give several illustrating examples. This is a report on joint work of the topological data analysis group at KTH in Stockholm that consists of Anders Lundman, Ryan Ramanujam, Martina Scolamiero, Sebastian Öberg and myself.

Robert Ghrist
Title: Local-to-Global Data: Applied Algebraic Topology
Abstract: Many contemporary challenges in the engineering sciences concern the inference of global features from local data. This passage from local- to global- data is as subtle as it is fundamental; however, it is not unprecedented. In the mathematical sciences, several types of local-to-global challenges were overcome with new techniques – from topology, homological algebra, and sheaves. This talk will outline both the vision and the first steps of exporting homological and topological tools to the data sciences, with an abundance of examples.

Herbert Edelsbrunner
Title: Metrics, Homology, and Persistence.
*Abstract:*The general pipeline for analyzing point data with persistent homology uses a notion of distance to convert the points into a filtration of complexes, it uses the homology functor to get a tower of vector spaces, and it finally converts the tower into bars or, equivalently, points in the birth-death plane. In this talk, we will focus on the first step, arguing that the pipeline also works for Bregman divergences, which are generally not symmetric and do not satisfy the triangle inequality. Examples are the Kullback-Leibler divergence commonly used for text and images, and the Itakura-Saito divergence preferred for sound data. Along the way, we will touch upon popular concepts in topological data analysis, including Delaunay complexes and discrete Morse theory.

Michael Faber
Title: Topology of Large Random Spaces
Abstract: Large random simplicial complexes are high-dimensional generalisations of random graphs, they can be used for mathematical modelling of large databases and complex networks in applications. In my talk I will describe topological properties of large random simplicial complexes focusing on their Betti numbers and fundamental groups. The key role plays the uniform hyperbolicity property (in the sense of M. Gromov) which allows verifications of global topological features by performing some local tests.

Michael Kerber
Title: Algorithms for topological data analysis

Abstract: This talk gives an overview of the algorithmic steps of the task of applying persistent homology to real data sets. It discusses the current state of the art as reflected by current software libraries - a practical demonstration of these libraries will follow the talk within the workshop. The talk also surveys some ongoing work which leads to future speed-ups and extensions of the computational pipeline.

Vitaliy Kurlin
Title: Applications of Topological Data Analysis to Computer Vision
Abstract: We will review for a broad audience several hard and important problems of Computer Vision that can be solved using rigorous methods of Topological Data Analysis. The talk is based on author's past work about data skeletonisation and on the joint work with Andrew Fitzgibbon during the EPSRC-funded secondment at Microsoft Research Cambridge.

Ezra Miller
Title: Biological applications of persistent homology: interactions with statistics, geometry, and algebra
*Abstract:*Applying persistent homology to biological problems can lead to fresh perspectives on the relevant topology and geometry (and on the underlying biology, too). The examples here come from two datasets: magnetic resonance images of cerebral arteries and photographic images of fruit fly wings. The first part of the talk explains what we have learned about the geometry of blood vessels in aging human brains, along with lessons this exploration has taught us about applications of persistent homology in general. The second part concerns current investigations in evolutionary biology, especially geometric questions that arise from statistical analysis in the context of multiparameter persistence.

Konstantin Mischaikow
Title: Building a Database of Dynamics for Regulatory Networks
Abstract: Consider a regulatory network presented as a directed graph with annotated edges that indicate if the first node is up-regulating or down regulating the second node. What kind of dynamics can this network generate? While this may seem to be an inadequately posed question it arises fairly often in biological contexts. Our motivation for addressing it arises from gene regulatory networks where we assume that the nodes represent genes and act as switches. However, we do not assume that we know the appropriate parameter values let alone the nonlinear reactions that govern the switches. Nevertheless, as I will describe in this talk, we can give a mathematically justifiable computationally tractable description of the global dynamics at all parameter values for moderate sized networks.
We call the output of our approach a Database of Dynamic Signatures in that we produce a potentially large, but queriable combinatorial descriptions of the global dynamics that are valid over explicitly defined regions of parameter space.
I will focus on the mathematics behind our approach and discuss a few idealized examples.

Vidit Nanda
Title: Reconstructing manifolds and functions from random samples
Abstract: We survey the work of Niyogi, Smale and Weinberger which provides explicit bounds on the number of uniform random samples required to reconstruct the homotopy type of an underlying compact Riemannian manifold with high confidence. We also describe an alaogous result for Lipschitz functions between such manifolds: one can recover the action on homotopy of such a function using sampled points on the domain, the codomain, and values of the function restricted to the sampled domain points.

Nina Otter
Title: Software for persistent homology: overview and demonstration
Abstract: Persistent homology is a method from algebraic topology used in topological data analysis to study qualitative features of data. It is robust to small perturbations in measurement, independent of dimension and scales and provides a compact representation of the outputs. Despite recent progress, the computation of persistent homology remains an open area of research with numerous important and fascinating challenges. It is a rapidly evolving field, with new algorithms and implementations being developed and released at a fast pace.
In this talk I will give an overview of the state-of-the-art algorithms and software for the computation of persistent homology, outline some of the challenges related to the computation of persistent homology, and give a demonstration of the software capabilities. All scripts used

for the demonstration will be made available online. This talk is based on joint work with M. Porter, U. Tillmann, P. Grindrod and H. Harrington.

Bill Oxbury
Challenges in my area. How we need your help....

Mason Porter
Title: Topological Data Analysis of Contagion Maps for Examining Spreading Processes on Networks
Abstract: Social and biological contagions are influenced by the spatial embeddedness of networks. Historically, many epidemics spread as a wave across part of the Earth's surface. However, in modern contagions long-range edges --- for example, due to airline transportation or communication media --- allow clusters of a contagion to appear in distant locations. We study the spread of contagions on networks through a methodology grounded in topological data analysis and nonlinear dimension reduction. We construct 'contagion maps' that use multiple contagions on a network to map the nodes as a point cloud. By analyzing the topology, geometry, and dimensionality of manifold structure in such point clouds, we reveal insights to aid in the modelling, forecast, and control of spreading processes.

Subramanian Ramamoorthy
Title: Topological Trajectory Classification for Dynamic State Estimation
Abstract: A common feature of models of dynamics that are used in application domains such as computer vision and robotics is hierarchy. With the emerging interest in obtaining such models of dynamics from data, there is also the need to learn hierarchical representations and utilise them in the construction of predictive models. While there has been significant recent activity in this area, such as deep neural networks for automatic feature learning in computer vision, the question of how best to learn action-relevant symbols and hierarchical representations is relatively less well studied.

In this talk, I will describe the use of tools from computational topology, e.g., persistent homology, within this problem setting. I will present a sampling-based approach to trajectory classification which enables automated high-level reasoning about topological classes of trajectories, along with empirical examples from humanoid robot data. I will then show how this multi-scale representation can be utilised within nonparametric state estimation algorithms, to simultaneously compute the estimate over progressively more coarse trajectory classes and the detailed state. This enables novel ways to fuse information arising from heterogeneous sensory modalities, such as position measurements and linguistic instructions about those positions.

I will conclude with some suggestions of other domains with more challenging forms of dynamics and sources of uncertainty where a similar approach may be fruitful.