



MATH-IMS Joint Applied Mathematics Colloquium Series
The Chinese University of Hong Kong

This MATH-IMS Joint Colloquium Series is organized by Center for Mathematical Artificial Intelligence (CMAI), under Department of Mathematics and Institute of Mathematical Sciences (IMS) at The Chinese University of Hong Kong. The colloquium series focuses on mathematics and applications of artificial intelligence, big data and related topics.

Date: Oct 28, 2022 (Friday)

Time: 14:00-15:00 (Hong Kong Time)

Zoom Link: <https://cuhk.zoom.us/j/92775210812>

Learning to Untangle Genome Assembly with Graph Convolutional Networks

Speaker: Prof. Xavier Bresson
Head of Graph Machine Learning at Sea Group

Abstract: A quest to determine the complete sequence of a human DNA from telomere to telomere started three decades ago and was finally completed in 2021. This accomplishment was a result of a tremendous effort of numerous experts who engineered various tools and performed laborious manual inspection to achieve the first gapless genome sequence. However, such method can hardly be used as a general approach to assemble different genomes, especially when the assembly speed is critical given the large amount of data. In this work, we explore a different approach to the central part of the genome assembly task that consists of untangling a large assembly graph from which a genomic sequence needs to be reconstructed. Our main motivation is to reduce human-engineered heuristics and use deep learning to develop more generalizable reconstruction techniques. Precisely, we introduce a new learning framework to train a graph convolutional network to resolve assembly graphs by finding a correct path through them. The training is supervised with a dataset generated from the resolved CHM13 human sequence and tested on assembly graphs built using real human PacBio HiFi reads. Experimental results show that a model, trained on simulated graphs generated solely from a single chromosome, is able to remarkably resolve all other chromosomes. Moreover, the model outperforms hand-crafted heuristics from a state-of-the-art de novo assembler on the same graphs.

Bio: Xavier Bresson is an Associate Professor in the Department of Computer Science at the National University of Singapore (NUS). His research focuses on Graph Deep Learning, a new framework that combines graph theory and neural networks to tackle complex data domains. In 2016, he received the USD 2.5M NRF Fellowship, the largest individual grant in Singapore, to develop this new framework. He was also awarded several research grants in the U.S. and Hong Kong. He co-authored one of the most cited works in this field (10th most cited paper at NeurIPS) and has significantly contributed to mature these emerging techniques. He has organized several workshops and tutorials on graph deep learning such as the IPAM'23'21 workshops on "Deep Learning and Combinatorial Optimization", the MLSys'21 workshop on "Graph Neural Networks and Systems", the IPAM'19 and IPAM'18 workshops on "New Deep Learning Techniques", and the NeurIPS'17, CVPR'17 and SIAM'18 tutorials on "Geometric Deep Learning on Graphs and Manifolds". He has been a regular invited speaker at universities and companies to share his work. He has also been a speaker at the NeurIPS'22, KDD'21, AAAI'21 and ICML'20 workshops on "Graph Representation Learning", and the ICLR'20 workshop on "Deep Neural Models and Differential Equations".