

NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

C S S E M I N A R

Title: Superbubbles and Clumps

Speaker: Speaker 1: Professor Costas Iliopoulos
Department of Computer Science
King's College London
University of London

Speaker 2: Miss Ritu Kundu
PhD student
Department of Informatics (Faculty of Natural & Mathematical Sciences)
King's College London
University of London

Date/Time: 25 April 2016, Monday, 02:00 PM to 03:00 PM

Venue: Executive Classroom, COM2-04-02

Chaired by: Dr Wong Lim Soon, KITHCT Chair Professor, School of Computing
(wongls@comp.nus.edu.sg)

Abstract:

The information that can be inferred or predicted from knowing the genomic sequence of an organism is astonishing. String algorithms are critical to this process. This talk provides an overview of two particular problems - superbubbles and clumps - that arise during computational molecular biology research, and recent algorithmic developments in solving them.

Biodata - Speaker 1:

Costas Iliopoulos received his PhD in 1983 from Warwick University and he established the Algorithm Design Group at King's College London in 1991. His field of expertise is the theory and practice of algorithmics, in areas including to, pattern matching, repetition and regularity finding, text compression, and automata theory and applications. The research problems he works on have many practical applications in fields varying from molecular sequence analysis, computer vision, symbolic computation & music.

Costas and his colleagues developed the most efficient (time/storage) software for mapping "Next generation Sequences", handling the massive DNA data produced by the new

sequencing hardware (Illumina) and the fastest methods/software for pairwise sequence alignment. Also, they developed the best method for hairpin identification in sequences (tropical diseases), methods for predicting the functional consequences of non-synonymous DNA sequence variants. Furthermore, they developed the Transcriptome map of mouse isochores in embryonic and neonatal cortex as well as the mouse liver, muscles & brain. In Combinatorics on words, they developed new data structures for order-preserving matching, new methods for finding Abelian regularities, subtree repeats, quasiperiods, cubic runs, tandem repeats, maximum number of squares, covers and seeds in strings and sequences.

In handling Big Data, they developed efficient parallel algorithms for mapping degenerate and weighted sequences to a reference genome, large scale DNA sequence analysis, parallel approximate string matching for massive data, storing and querying a massive number of highly similar sequences and mapping short reads to a massive number of dynamically changing genomic sequences.

Costas is the editor-in-Chief of the Journal of Discrete Algorithms (Elsevier) and he has served as chair and member of numerous programme committees of international conferences and workshops. He has held visiting positions in a number Universities: Paris Est, Pretoria, McMaster, Rouen, Patras, Warsaw, Curtin, Stellenbosch etc. He has been supported by European Union, Royal Society, Institute of Maths, London Mathematical Society, Wellcome foundation, grants etc

Biodata - Speaker 2:

Ritu Kundu is a PhD student under the supervision of Prof. Costas S. Iliopoulos, in the Department of Informatics (Faculty of Natural & Mathematical Sciences), King's College London, where she previously completed her M.Sc. in Advanced Computing. She has also worked as software developer for a couple of years.

Ritu's research focusses on the design and analysis of efficient algorithms and data structures for sequential data (Stringology), which have many practical applications in fields varying from molecular sequence analysis, computer vision, symbolic computation and computational music. Her recently published works include several novel linear-time algorithms - for conservative degenerate pattern matching, for identifying superbubbles, and for discovering clumps. Some of her recently developed software-tools can be found at her GitHub profile.

She currently is a student member at Centre for Combinatorics on Words & Applications (CCWA <<https://ccwaperth.wordpress.com/> Murdoch University, Perth, Australia) and also is acting as the web-master for International Federation for Information Processing (IFIP) Working Group 10 <<http://www.dcs.kcl.ac.uk/staff/csi/IFIP1-10/index.html>. She is also a member of research teams of King's College London that collaborate with CSUCI and University of Helsinki, supported by the Royal Society. She has also reviewed for LATA 2016 and MACIS 2015.

She is being funded by EPSRC DTP. In addition, she has been awarded various grants - ERASMUS+, British Council's INSPIRE, and KCL's Global Research Grant etc.