ICTS Colloquium

Title : Sequence Alignment: Hidden Markov Models, Category Theory and all that jazz

Speaker : Soumyashant Nayak, University of Pennsylvania

Date : Tuesday, August 20, 2019

Time : 3:00 PM

Venue : Emmy Noether Seminar Room, ICTS Campus, Bangalore

Abstract : High-throughput deep-sequencing technologies such as RNA-Seq have become indispensable tools in the arsenal of biomedical scientists. RNA-Seq facilitates the investigation of alternatively spliced transcripts, post-transcriptional modifications, single nucleotide polymorphisms (SNPs), discovery of novel isoforms, etc. One of the first steps in the analysis of high-throughput RNA-Seq data is alignment to a reference genome or transcriptome. In this talk, we will trace the history of the sequence alignment problem. This problem not only shows up in the context of high-throughput sequencing data analysis but also in basic science problems in evolutionary genetics. We will discuss the famous BLAST algorithm and also some more recent methods for transcript-level quantification via `pseudo-alignment' which are much faster than the traditional methods and require fewer computing resources (without compromising too much on accuracy). Time permitting, we will touch upon the emerging field of applied category theory in the context of organizing ideas in this area.