Table of contents

The CODATA-RDA Research Data Science Applied workshops on Extreme sources of data, Bioinformatics and IoT/Big-Data Analytics | (smr 3178)

Tuesday 25 July 2017

Bioinformatics - Adriatico Guest House - Informatics Laboratory (09:00-18:00)

Abstract: Next Generation Sequencing (NGS) technologies have led to discoveries of new diagnostic, prognostic and therapeutic targets. Despite these discoveries, treatment of cancer patients, detection of cancer biomarkers and prediction of therapy response remain largely unsolved problems. These difficulties are hindering the realisation of effective approaches to personalized medicine; and data needs to be better exploited to systematically elucidate the mechanisms and causes underlying cancer origination and development.

Cancers accumulate genetic mutations that allow their cells to proliferate out of control. Mutations occur randomly, are inherited through cell divisions, and orchestrate cancer initiation and development with accumulation patterns differing between individuals. NGS technologies are routinely used to detect mutations in tumoral biopsies, and free-access large collections of cancer datasets are now available. Cancer mutation profiles are incredibly heterogenous, and we observe few common mutations across patients even if their cancers have similar histological classification. Tumor Heterogeneity (TH) is intimately related to Cancer Evolution, and is considered to lead to the emergence of drug-resistance mechanisms, relapse and failure of treatments. Quantification of TH across cancer types and patients is of the utmost importance in modern cancer research.

I will present a causal framework to infer, from DNA sequencing data, Graphical Models that recapitulates the progression of the tumors (i.e., evolutionary models). This inference problem has several formulations, according to the type of NGS data that we have access to. I will discuss an approach that combines Statistics, Machine Learning and Formal Methods to infer models from single-sample data; and then I will move on to the problem of studying Cancer Evolution from multi-samples of the same individual. These two problems are orthogonal, and I will discuss attempts at defining a unique framework to study Cancer Evolution. Example applications with real data will be presented and discussed.

time	title	presenter
09:00	Data Science approaches to infer Cancer Progression Models	CARAVAGNA, Giulio
10:30	Coffee break	
11:00	Data Science approaches to infer Cancer Progression Models	CARAVAGNA, Giulio
12:30	Lunch break	
14:00	Lab Session: Data Science approaches to infer Cancer Progression Models	CARAVAGNA, Giulio
15:30	Coffee break	

16:00	Lab Session: Data Science approaches to infer Cancer Progression Models	CARAVAGNA, Giulio