Robust and Scalable Inference of Cancer Progression Pathways using Conjunctive Bayesian Networks

Sayed-Rzgar Hosseini

Department of Mathematical Sciences, the University of Texas at El Paso (UTEP)

Various machine learning methods have been widely used to analyze the increasingly available cancer multi-omics data. However, the clinical impact of purely data-driven approaches is hindered mainly by their limited generalizability and lack of mechanistic interpretability. To address these challenges, hybrid mechanistic/data-driven models are critically needed in the field of computational oncology. Furthermore, due to the extensive tumor heterogeneity, disease mechanisms are highly context-dependent, which necessitates development of context-aware computational approaches for disease modeling and prediction of therapeutic outcomes in precision cancer medicine. To meet these critical needs, we will develop novel strategies by leveraging the unique ability of a probabilistic-graphical model called Conjunctive Bayesian Networks (CBNs) for quantifying cancer evolution directly from cross-sectional genomic data. Previously, we showed that the CBN model can be used to estimate the predictability of cancer evolution as it is able to reflect the underlying cancer fitness landscapes directly from genotypic data. However, the reliability of the inferred pathway probability distributions has not yet been ascertained, which motivates the need for a robust inferential framework. I have recently extended this CBN-based framework to enable robust and scalable inference of cancer evolutionary pathways, which will serve as a basis for uncovering cancer progression mechanisms, and for interpretable representation of disease context and cellular states. In this work, by analyzing synthetic, simulated and real data, I have rigorously compared R-CBN with previous CBN models including CT-CBN, H-CBN, and B-CBN, and the results indicate a superior robustness of the R-CBN model in various settings. Furthermore, I have devised a dynamic programming approximation algorithm, which renders the model amenable to scalability. Thus, R-CBN has the potential to be broadly utilized as a reliable framework to infer cancer-driving evolutionary trajectories, and to distill mechanistic insights from cross-sectional cancer genomic data.