Detecting Repeated Evolutionary Trajectories with Inverse Reinforcement Learning

Description

Tumor progression is an evolutionary process of accumulating selectively advantageous mutations in cells. By sampling the tumor cells and analyzing their genomes, one can reconstruct the mutational history of a tumor and characterize it using a mutation tree structure, where the nodes represent the mutations and are connected according to their order of occurrence. However, mutation trees can vary considerably between any two patients, making it challenging to detect repeated trajectories, let alone reliable predictions of future evolutionary steps. Assuming that the complex behaviors of cancer cells follow a Markov Decision Process, the trajectories/paths in the mutation trees can be considered as realizations of an optimal policy given an unknown reward function. With inverse reinforcement learning, it is possible to estimate the reward function given a set of mutation trees. The estimated reward function can then be used to infer optimal/repeated evolutionary trajectories for a given cancer type.

Goal

A recent method combines inverse reinforcement learning with a pop-up restaurant process, where the number of reward functions is not specified a priori [1]. The proposed tasks in this project are:

- Understand the method in [1] and implement it using Python (and PyTorch or Jax);
- Apply the method to several other tree data sets (e.g. [2, 3]);
- Compare the inferred trajectories with the results obtained with alternative methods (e.g. [4]);
- If time allows, investigate how cancer treatments change the reward functions.

Educational experience

- Improve your understanding of cancer evolution and reinforcement learning.
- Learn the best coding practices (code reviews, static code analysis) and modern ML frameworks.
- Gain hands-on experience in analyzing real cancer data sets.

References

1. Kalantari, J. et al. (2020). *The unreasonable effectiveness of inverse reinforcement learning in advancing cancer research*. In Proceedings of the AAAI Conference on Artificial Intelligence, 34(1), 437-445.
2. Morita, K. et al. (2020). *Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics*. Nature communications, 11(1), 1-17.
3. Jamal-Hanjani, M. et al. (2017). *Tracking the evolution of non–small-cell lung cancer*. New England Journal of Medicine, 376(22), 2109-2121.
4. Caravagna, G. et al. (2018). *Detecting repeated cancer evolution from multi-region tumor sequencing data*. Nature methods, 15(9), 707-714.

Time effort

Lab rotation or semester project.

Contact details

Luo, Xiang Ge <xiangge.luo@bsse.ethz.ch>
Czyż, Paweł Piotr <pawelpiotr.czyz@ai.ethz.ch>
Beerenwinkel, Niko <niko.beerenwinkel@bsse.ethz.ch>