COVID-19 Trajectory Reinforcement Learning Project Proposal

Andrew Massoud, Devon Williams, Shengwei Zhang, Siyue Huang, Matthew Taj, Dr. Tom Woolf

Introduction and Background:

The COVID-19 pandemic has impacted millions of people globally, from patients and their families to hospital workers and clinicians. The rise in severe cases has resulted in a strain on hospital capacities as many patients require extensive treatment. To combat this, the trajectory of the COVID-19 disease caused by the SARS-CoV-2 virus can be explored through patient data in multiple states (*Figure 1*) [1].

The National COVID Cohort Collaborative (N3C) is a secure data enclave through which researchers can access billions of rows of patient data [2]. Within this data are features that can be extracted to determine which decisions impact the states of COVID patients to the greatest extent. Patient data recorded in clinical settings can serve as a data source for machine learning models to optimize clinical decision-making policies. Reinforcement learning (RL) algorithms are relatively new to the medical field, but due to the availability of data and ability to map into state-action pairs, we aim to apply RL to the problem of understanding the COVID-19 disease trajectory to learn what decisions are most impactful and how to craft an optimal policy.

Study Design:

Section 1: Data Source

The main data source we will use for this project is Electronic Health Records (EHR) data. It is a systematic aggregation of all the patients' health records in digital form. EHR contains a range of fields for each patient, including demographics, medical history, medication and allergies, observations and procedures, health care plan and billing information. The data is organized in chronological order so that we can use EHR to track the disease trajectory for each patient. Thus, we see the possibility to apply reinforcement learning in this process to help clinicians design better policies (i.e., treatment plans) for patients at different stages of COVID progression. Considering the difficulties in accessing and manipulating real patient data, we will firstly work on synthetic COVID patients' data as the training set for our RL algorithms. We will use Synthea to generate this data set.

Section 2: Generating Synthetic COVID Patients Data

Section 2.1: Why Synthea: Synthea is an open-source synthetic patient generator that can simulate the medical history of a patient population. This software will output realistic but not real patient data and associated health records covering every aspect of health. We would like to use Synthea to generate COVID patients specifically because we have seen some success in deploying the Synthea model for COVID patients last year in academia [3]. Using synthetic patient data in the early stage, we are free from cost, privacy, and security concerns, and after we develop a mature RL algorithm, we will transform to N3C platform. The data format will be exactly the same as what Synthea produces, and we will be able to easily validate our algorithms on the real patient data.

Section 2.2: Synthea Framework: The synthetic data generated by Synthea depends on the utilization of the Publicly Available Data Approach to the Realistic Synthetic EHR, PADARSER. The assumption for PADARSER is that access to the real EHR is impossible or undesirable. In order to achieve privacy preservation, PADARSER used publicly available data gathered from aggregate health incident statistics, clinical practice guidelines (CPGs), and medical coding dictionaries in the generation process and applied them to a time-based model for each synthetic patient [4]. The output EHR has sufficient properties of practicality to substitute the demand for the real EHR in many secondary uses[4].

Section 2.3: Data Processing - Random Forest on Feature Selection: Random forest is an ensemble learning method, consisting of a multitude of decision trees, each of them randomly extracts observations from the dataset and randomly extracts features. Every tree makes a class prediction, and the class with majority votes becomes selected output. Random forests generally outperform decision trees, however data characteristics can affect their performance[5]. Due to the potential high dimensionality of our data, we tend to use random forest to preemptively select for features that are most relevant to our problem.

Section 3: Reinforcement Learning

Section 3.1: Conceptual Introduction: Reinforcement learning revolves around an interaction between two entities, an agent and its environment, and this interaction can be described in 3 variables: actions, rewards and states (*Figure 2)*. *Actions* are decisions that the agent makes that influence the state and result in some reward. *Rewards* are scalar values that the agent hopes to maximize over some number of trials. *States* can be thought of as a summary of our agent at some particular time (or any function of the history of our agent given past actions, rewards, or states). The types of problems that reinforcement learning can handle is seen in the Reward Hypothesis, which states that all goals can be described by the maximization of expected cumulative reward.

Section 3.2: Categories of Reinforcement Learning: There are 3 properties that reinforcement learning algorithms can mix and match suiting the problem: models, policies, and value functions. *Models* are intuitively some implementation of the information you have about the agent's environment. For example, if we were trying to solve a specific maze a map of the maze would be our model. *Policies* are rules or regulations that your agent follows at some specific state(s). Finally, we have *value functions* which are rewards given at a particular state.

Section 3.3: Codebases and Implementation: We plan on relying more heavily on policy based algorithms as we hope to find clinical decisions that can be optimized to minimize the number of infections/fatalities. Our ideal model would be a perfect description of human physiology; because that is quite hard to produce (or update) we will tend to more policy based outcomes. These will be optimized with offline reinforcement learning methods instead of relying on a model. In offline reinforcement learning, instead of changing your policies and re-interacting with the environment you are only able to learn on data collected from one particular instance. We are currently working with two different codebases, Deepmind and OpenAI Gym. Both have methods and implementations of reinforcement learning techniques which we hope to leverage and combine with more recent developments in offline reinforcement learning (NeurIPS 2020).

Timeline:

While all roles overlap to a certain degree, Devon and Matthew have the initial leads on working through RL codebases, Nicky and Andrew are working on extracting patient data and defining state-action pairs, and Shirley will explore applying random forest algorithms to preprocess data to optimize feature selection. By early March we anticipate that half the work will be spent on the backend of developing RL algorithms, and the other half will be on processing and mapping the data to them. We hope to develop a working model with Synthea's data before moving to real patient data from N3C to develop optimal policies regarding decisions that affect patient states during the course of COVID-19's disease trajectory.

Appendix

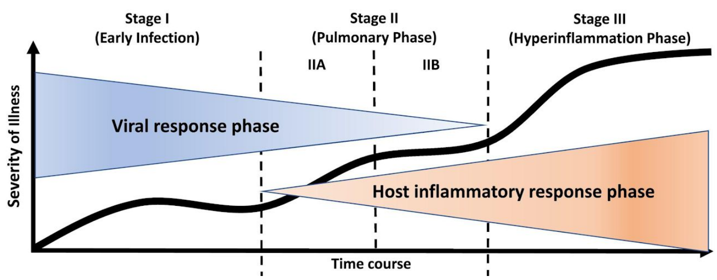

Figure 1: COVID-19 Disease Trajectory Classification by State [1]

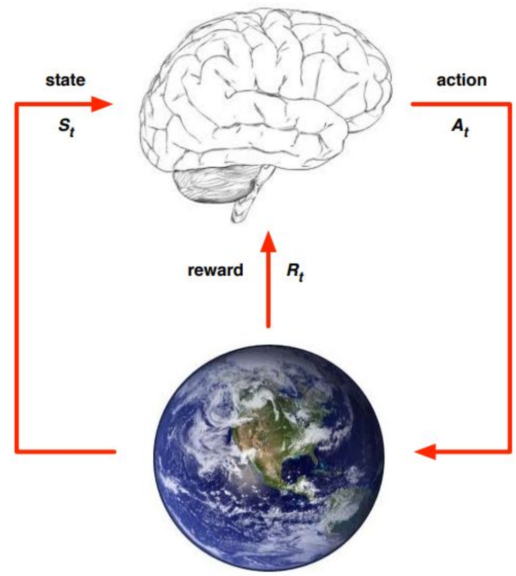

Figure 2: Highlighting the interaction between the agent (the brain) and the environment (Earth) being dependent on the three variables in RL. [6]

References

[1] Siddiqu HK, Mehra MR. “COVID-19 Illness in Native and Immunosuppressed States: A Clinical-Therapeutic Staging Proposal” Journal of Heart and Lung Transplantation, doi: 10.1016/j.healun.2020.03.012.

[2] “Home,” N3C. [Online]. Available: https://covid.cd2h.org/. [Accessed: 23-Feb-2021].

[3] Walonoski, Jason et al. “Synthea™ Novel coronavirus (COVID-19) model and synthetic data set.” Intelligence-based medicine vol. 1 (2020): 100007. doi:10.1016/j.ibmed.2020.100007

[4] Walonoski, Jason. “Synthea: An Approach, Method, and Software Mechanism for Generating Synthetic Patients and the Synthetic Electronic Health Care Record.” Journal of the American Medical Informatics Association, 30 Aug. 2017, doi:10.1093/jamia/ocx079.

[5] Piryonesi S. Madeh; El-Diraby Tamer E. (2020-06-01). "Role of Data Analytics in Infrastructure Asset Management: Overcoming Data Size and Quality Problems". Journal of Transportation Engineering, Part B: Pavements. 146 (2): 04020022. doi:10.1061/JPEODX.0000175.

[6] Silver, David, “Introduction to Reinforcement Learning with David Silver” 13 May 2015, https://deepmind.com/learning-resources/-introduction-reinforcement-learning-david-silver

Sutton, Richard S. *Introduction to Reinforcement Learning*. MIT Press.