The Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI-24)

Navigating Uncertainty in Epidemic Contexts with Reinforcement Learning

Elizabeth Akinyi Ondula

Thomas Lord Department of Computer Science
University of Southern California
Los Angeles, CA 90089 USA
ondula@usc.edu

Abstract

My research integrates stochastic epidemic models with reinforcement learning to develop effective strategies or policies to inform operational decisions. The objective is to refine policies that are attuned to diverse outbreak dynamics and to offer a tool for informed planning in real-world settings.

Introduction

Reinforcement learning (RL) has been widely applied to simulate and optimize policies for epidemic control across domains like healthcare, economics, and mobility. Its ability to dynamically adjust strategies in response to changing conditions makes it particularly suited to addressing challenges in an epidemic context.

Traditionally, response strategies have been static, while essential, often demonstrating a lack of flexibility in adapting to rapidly changing epidemic conditions. This has revealed a significant gap between reactive measures and the dynamic nature of epidemic spread. Static strategies, once effective, now face the necessity of reevaluation and adaptation in the face of these evolving challenges. For instance, in educational environments, where the balance between safety and uninterrupted learning is critical, the need for innovative and adaptable approaches becomes particularly evident. However, this requirement extends beyond the education domain to encompass other sectors, such as healthcare, where the management of patient flow and resource allocation during an epidemic is crucial, or in economics and mobility, where the impact of an epidemic can drastically alter market dynamics and transportation patterns. This approach embodies a dynamic recalibration of strategies, rooted in data-driven outcomes and real-time responsiveness that can be adapted in these sectors with a focus on safeguarding health and nurturing human-to-human interactions.

This research focuses on creating a simulation environment for understanding and studying localized policies generated by RL. The environment serves as a platform where RL algorithms and stochastic epidemic models interact offering a framework to capture disease-spread behaviors and policy evaluation as they occur in the real world. It's designed to be extensible, enabling the incorporation of a variety of models from the wider research community. The objective is to refine policies that are attuned to diverse outbreak dynamics, effectively managing the spread of diseases while preserving essential human-to-human interactions.

Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Related Work

Reinforcement learning has been widely applied across various domains like healthcare, economics, and mobility for epidemic control to simulate disease spread and optimize policies ((Arango and Pelov 2020), (Ohi et al. 2020), (Feng et al. 2022)). For instance, (Oikawa et al. 2022) provides valuable findings on the effectiveness of reducing class sizes to enhance social distancing and its consequential reduction in flu outbreaks. This aligns with our exploration of safety measures such as in schools. Similarly, the studies by (Fukumoto, McClean, and Nakagawa 2021) and (Wu et al. 2022) delve into the effects of school closures on COVID-19 spread and the implications of such measures on children and society. This resonates with our aim to identify strategies using reinforcement learning. Furthermore,(Kaiser, Kretschmer, and Leszczensky 2020) examines the efficacy of cohorting strategies within schools, emphasizing the need for well-thought-out interventions to minimize virus transmission, an aspect that our paper seeks to address through simulation-based policy evaluation. (Endo et al. 2022) also contribute to this discourse by evaluating the efficacy of various interventions, highlighting the potential limitations of certain approaches.

Modeling

We present a proof of concept for the education domain. Consider a classroom scenario with N students attending sessions over W weeks. During an ongoing epidemic, students face the risk of infection both off-campus and on-campus. Off-campus infections are considered to be an exogenous random process, where each student has an independent and identical probability c_w of being infected off-campus during week w, termed the community risk. On-campus infections result from infected students spreading a virus to other students.

State, Action, State Evolution The state is defined by a tuple representing community risk and the expected infected $(c_w, E[I_w])$ in a given week. The action, A_w denotes the number of students permitted in a class. The reinforcement learning problem thus is to formulate a policy to decide based on the state, the number of students to allow in a classroom each week w. We consider both off-campus and on-campus infections. If no students are allowed in the classroom, only off-campus infections occur. However, if students are allowed, both types of conditions are considered. The state evolution is governed by the dynamics of a stochastic epidemic model such as an approximate SIR model applied in (Ondula and Krishnamachari 2022) or indoor classroom model demonstrated by (Hekmati et al. 2022) that includes initial infection probability, weekly infection likelihood, and expected classroom infection. The model also quantifies the viral dose from an infected student, total dose, transmission probability, and overall post-interaction infection risk.

Reward The reward function is such that a higher reward is given to the RL agent when the expected number of infected students is low. The total reward is a weighted difference between the number of allowed students and the expected number of infected students

Experiments

We hypothesize that some variants of RL algorithms can effectively generate sensible policies that prescribe decisions such as classroom occupancy (in the education domain) based on infection counts and community risk patterns. We also posit that the reward design such as using a weight parameter (to balance the educational benefit of allowing more students and managing infection risk) will play a crucial role in shaping the RL agent outcomes and the precision of a policy. Our objective is to explore how the algorithm prioritizes various trade-offs like educational benefits by 'allowing more students' and infection risk minimization, thereby calibrating the matrix to align with varying epidemic scenarios. Preliminary results suggest that through iterative training, a Q-learning agent refines its policies. Furthermore, by adjusting a reward parameter, different Pareto-optimal trade-offs can be achieved between minimizing student infections and maximizing on-campus student presence. Future experiments will highlight the challenges and nuances of integrating stochastic epidemic modeling with reinforcement learning, emphasizing the importance of accurate parameter estimation and appropriate reward mechanism definition.

Anticipated Contribution and Timeline

Present - Dec 2023: We have a proof of concept software platform that incorporates a module for adding infection models as well as a blueprint for integrating a myriad of RL algorithms.
Jan 2023 - Feb 2024: By the consortium time we will have some preliminary evaluations based on the education scenario.
Mar 2024 - May 2024: Re-evaluation and incorporation of feedback from the consortium.
Jun 2024 - Dec 2024: Conduct Experiments, Evaluations, and Testing extensions to the environment.
Jan 2025 - May 2025: Work towards Thesis proposal.
Jun 2025 - May 2026: Work towards Thesis Defense. Plan to graduate by May 2026.

References

Arango, M.; and Pelov, L. 2020. Covid-19 pandemic cyclic lockdown optimization using reinforcement learning. *arXiv preprint arXiv:2009.04647*.

Endo, A.; Group, C. C.-. W.; Uchida, M.; Liu, Y.; Atkins, K. E.; Kucharski, A. J.; and Funk, S. 2022. Simulating respiratory disease transmission within and between classrooms to assess pandemic management strategies at schools. *Proceedings of the National Academy of Sciences*, 119(37): e2203019119.

Feng, T.; Xia, T.; Fan, X.; Wang, H.; Zong, Z.; and Li, Y. 2022. Precise mobility intervention for epidemic control using unobservable information via deep reinforcement learning. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2882–2892.

Fukumoto, K.; McClean, C. T.; and Nakagawa, K. 2021. Shut Down Schools, Knock Down the Virus? No Causal Effect of School Closures on the Spread of COVID-19. *medRxiv*, 2021–04.

Hekmati, A.; Luhar, M.; Krishnamachari, B.; and Matarić, M. 2022. Simulating COVID-19 classroom transmission on a university campus. *Proceedings of the National Academy of Sciences*, 119(22): e2116165119.

Kaiser, A.; Kretschmer, D.; and Leszczensky, L. 2020. Social network-based strategies for classroom size reduction can help limit outbreaks of SARS-CoV-2 in high schools. A simulation study in classrooms of four European countries. *medRxiv*, 2020–11.

Ohi, A. Q.; Mridha, M.; Monowar, M. M.; and Hamid, M. A. 2020. Exploring optimal control of epidemic spread using reinforcement learning. *Scientific reports*, 10(1): 22106.

Oikawa, M.; Tanaka, R.; Bessho, S.-i.; and Noguchi, H. 2022. Do Class Size Reductions Protect Students from Infectious Diseases? Lessons for COVID-19 Policy from a Flu Epidemic in the Tokyo Metropolitan Area. *American Journal of Health Economics*, 8(4): 449–476.

Ondula, E. A.; and Krishnamachari, B. 2022. Using Reinforcement Learning for Operating Educational Campuses Safely during a Pandemic (Student Abstract). In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, 13025–13026.

Wu, J. T.; Mei, S.; Luo, S.; Leung, K.; Liu, D.; Lv, Q.; Liu, J.; Li, Y.; Prem, K.; Jit, M.; et al. 2022. A global assessment of the impact of school closure in reducing COVID-19 spread. *Philosophical Transactions of the Royal Society A*, 380(2214): 20210124.