Using Probabilistic Planning to Model the Spread of COVID-19 in Kingston, Ontario

Bruce Chidley¹, Christian Muise¹

¹Queen's University 17bjc4@queensu.ca, christian.muise@queensu.ca

Abstract

Modeling and understanding the spread of disease has been a topic of much focus for epidemiological researchers in recent years due to the effects of the COVID-19 pandemic. High levels of global attention and an abundance of recently col-

- 5 lected data have created an environment for epidemiological models to be highly detailed and impactful. The best of these models incorporate ideas and research from a range of domains, with clarity and ease of consumption being key focuses so that they have the highest chance of impacting pub-
- lic policy. Here, probabilistic planning is leveraged to understand the spread of COVID-19 at a regional level in the city of Kingston, Ontario, through an agent-based model implemented in RDDL. This data-driven model operates with the functionality of introducing mask and vaccine mandates as sparingly as possible, as per policies created by JaxPlanner, while attempting to ensure that hospitals are operating below capacity. Mimicking the variability of real data, a variety of model configurations are experimented with, and the resulting simulation differences are noted. The observed dynamics
- are well in line with the theoretical models, and the intervention methods utilized proved to be successful at the task of mitigating hospital burden, though the policies decided upon were inconsistent at times.

Introduction

- 25 Understanding how a disease spreads is necessary when attempting to minimize the damage done to a population. The intricacies of exactly how a disease spreads are often quite difficult to identify, and it requires significant effort to truly capture minute details. However, this type of work is essen-
- tial for knowing when and how certain intervention methods should be deployed in a population. The basis for these decisions lies in model forecasts made by researchers; forecasts that are founded in whatever data is available. Oftentimes, the parameters comprising a model have their values based
- on insufficient or inferred data. These vague parameter estimates can result in stochastic model simulations having large variability, causing optimal intervention policies to be inconsistent. An important result of this is related to the economic burden that can arise from intervention methods. Em-
- 40 ploying intervention methods sub-optimally can cause the population to incur not only a significant economic loss but can also lead to a reduced quality of mental health for affected citizens (Liu et al. 2020). So, it is crucial for policy-

makers to intervene only when necessary - a task that is only possible through reliable data-driven models. It is also very important that the capacity of hospitals is not exceeded to ensure that new patients are able to receive the treatment they require. Thus, optimal intervention implementation would be to minimize intervention while continually keeping hospitals operating below capacity.

Due to the complex nature of epidemics, the models that describe them are oftentimes rather complicated - decreasing generalizability and risking poor adaptation to new discoveries or data. These disease models are typically constructed with the functionality of intervention techniques 55 in order to see how spread can slowed or stopped (Adiga et al. 2020). From this, policymakers have been able to incorporate strategies based on many such models, proving their viability in handling these kinds of highly complex scenarios (Kerr et al. 2021). However, even among the 60 most successful and detailed of models, there are still limitations that can cause significant differences in behaviour to emerge between the model's predictions and real-world outcomes. While general data availability and quality issues are common throughout these models, many of these limita-65 tions have to do with location and location-based data such as regional population, organization, and infrastructure. This work presents a model using RDDL and JaxPlanner, which aims to minimize intervention uptime while keeping hospitals operating below capacity through the use of real geo-70 graphical and regional data from the city of Kingston, Ontario, as an exploration of the power of RDDL in modeling disease dynamics and intervention, and as an avenue to demonstrate the necessity for accurate data in producing reliable models of this kind. This not only serves as a baseline 75 for future work in modeling disease spread in RDDL, but it also examines JaxPlanner's effectiveness at solving problems of this nature.

Background

Understanding SIR modeling is fundamental in the field of epidemiology. The basic SIR model has many variations, and this agent-based model is built on the SEIRS model. The SEIRS model splits a population into four classes, which are used to track and simulate the spread of disease. All individuals in a population are in exactly *one* of these classes at a time:

50

- 1. Susceptible: Individuals who may contract the disease at any given point
- 2. Exposed: Individuals who have contracted the disease but
- are not symptomatic and are not able to spread the disease

90

- 3. Infectious: The individuals in this class may or may not be showing symptoms but are now able to spread the disease
- 4. Recovered: Here, individuals are fully recovered from 95 the disease, meaning that they cannot spread it and they cannot contract it



Figure 1: The classes that an individual moves through in an SEIRS model

As shown in Figure 1, an individual begins in the susceptible class and may move through the classes linearly until reaching the recovered class, where they become suscepti-100 ble after a given period of time. These periods of time are not constant from person to person and can vary due to an individual's genetics and the severity of infection, amongst other factors.

We must also define the R_0 value here, which represents 105 the average number of susceptible individuals that an infectious individual will spread disease to. The R_0 value is calculated as $R_0 = \beta / \gamma$, where β represents the number of individuals the infection is spread to per time step, and γ represents the rate of recovery. 110

Related Work

Mathematical Modeling Often serving as the basis for more complex agent-based models, mathematical models seek to explain the spread of disease from a compartmental standpoint. The states a person can be in (SEIR) relative to 115 the disease and the rates of change of each of these states in a given population over time are modeled and can be readily tweaked to account for a variety of situations. Individuals themselves are not tracked in these models, but rather the population as a whole is estimated on each time step. Sig-120 nificant work was done early into the pandemic in Sweden, furthering our understanding of the disease behaviour in a population with real, up-to-date data (Britton 2020; Sjödin et al. 2020; Kamerlin and Kasson 2020). While these mod-

els performed well for certain tasks, the simplicity and gen-125 eralizations made by the models produced results that were often not in line with observed data.

Agent-Based Modeling Built from strong foundations in mathematical modeling, agent-based modeling has seen increased usage in recent years; specifically with respect 130 to disease modeling. Early into the COVID-19 pandemic, many older agent-based models were adapted to model

COVID-19's spread and resistance to intervention methods (Ferguson et al. 2020). While these sorts of models were arguably the most advanced at the time, they often had lim-135 itations with regard to population behaviour and the implementation of intervention methods (Adiga et al. 2020). Building on these ideas, popular models such as OpenABM-Covid19 and Covasim emerged, which have been used by public health officials worldwide to aid in decision support, 140 taking detailed population behaviour into account and tracking how a variety of intervention methods affect the spread of COVID-19 (Kerr et al. 2021; Hinch et al. 2021). These are seen as state-of-the-art COVID-19 models and are being consistently improved upon. GSAM is another powerful 145 model with very high scalability, allowing billions of agents to act distinctly at the cost of specificity with regard to their environments. This type of model is good for understanding broad disease dynamics but does not perform well when trying to understand fine details such as the effects of re-150 gional population distributions (Parker and Epstein 2011). There are also models like BioWar, where nefarious agents are introduced who act to intentionally spread disease, giving way to some interesting dynamics (Carley et al. 2006). In addition to these, there are a plethora of other models that 155 exist, which specialize in different tasks (Balcan et al. 2009; Venkatramanan et al. 2019; Silva et al. 2020; Mahmood et al. 2022) - many of them being open source and regularly updated.

Automated Planning Modeling the spread of disease is a 160 long-standing exercise and problem in the field of Planning. New strategies for handling probabilistic settings are regularly emerging, opening the door for more complex and realistic models. Specifically, RDDL is beginning to see use as a tool for developing intervention plans to hinder the spread 165 of disease, with Harmanani modeling COVID-19 spread in enclosed spaces, grounded in real data (Harmanani 2023). By allowing the planner to intervene by way of masking and vaccinating, the results produced by the model were found to quite accurately reflect real-world data. While this 170 attempt was successful at describing behaviour in enclosed spaces, there has been little attempt to model this kind of behaviour on a large scale such that the impact of disease on regional population dynamics could be analyzed in a meaningful way. 175

Methods

Modeling the environment

Space Initialization It was important that the space that the agents were to move around in represented Kingston, Ontario as closely as possible. By utilizing the OSMnx 180 Python package (Boeing 2017), accurate geographical data was ingested via a simple script. This package grants access to the coordinates of all registered buildings and the types of buildings that they are, breaking them down into the following categories:

• Residential: "House", "Apartment", "Dormitory", or "Residential", encapsulating the leftover residential buildings

190

195

210

220

230

- Work: "Office", "Commercial", "Industrial", "Retail", or "Warehouse'
- Commercial"Commercial" or "Retail"
- Education: "School", "College", or "University"

It is important to note that no building can belong to more than one building type. By running a simple Python script to retrieve this building data, the information was able to

be stored and ingested into the problem files that fueled the disease simulations.

Agent Initialization Agents are created with a set of attributes, and they have the ability to move between the defined locations. Each agent is assigned to an age bracket 200 based on the Kingston age distribution (Government of Canada 2021). The age brackets are defined as 0, representing ages 0-9, 1, representing ages 10-19, and so on until 8, which represents the 80+ age range. These age brackets determine how agents interact with COVID-19, how they move 205

around the space on each time step, and how they react to interventions (planner actions).

Every step in the simulation represents half a day. On the first step of each day, agents move about the problem space, interacting with other agents and potentially spreading disease. On a day's second step, they return to their homes and interact with the agents they live with. This notion of time is tracked via a simple counter that flips between 0 and 1 on each step. We also incorporate days of the week, with agent

behaviour varying depending on what day of the week it is. 215 This is tracked by a similar counter, ranging from 1 to 14 (as each day consists of two time-steps).

As hinted at, agents are all assigned homes, picked from the extracted list of residential buildings at random. Each home contains 1-5 agents, filled per the probabilities derived from Kingston data (Government of Canada 2021). This will be where agents go on the second time step of every day. With respect to disease interaction, if an agent enters the infectious class, they have a chance to self-isolate, and they will stay at home for their entire infectious period. If they

225 are hospitalized, they will not go home or to work/school until they move into the recovered state.

Agents are also assigned a place they go during the weekdays. If an agent is aged 0-19, they attend school. The school they attend is chosen such that the Euclidian distance between their home and the school is as small as possible (geographically, the closest school to their home). Some agents in the 20-29 age range will attend post-secondary school,

whereas other agents in the 20-29 range and all agents 30-69 will attend work during the day (proximity not being a factor 235 here). Agents aged 70 and above are assigned a commercial building that they visit during the day, which is chosen to be closest geographically in the same way schools are selected.

The agents in the 20-29 range who attend post-secondary institutions in Kingston are placed according to each institu-240 tion's user-defined populations. Agents are sequentially assigned to each institution until their prescribed population numbers are reached. The homes of these agents are selected based on their proximity to their given school, placing them in the closest residential buildings possible such that they

are not above capacity. By this algorithm, dormitories will

Parameter	Value
R_0 - Basic Reproduc- tive Ratio	3.32 ¹
β - Average number of agents infected on each time step	$R_0\gamma$
$1/\sigma$ - Time spent in exposed class	Normal(9, 2) ²
$1/\gamma$ - Time spent in infectious class	Mild: Normal(16, 4) ³ Severe: Normal(36.2, 8) ⁴ Critical: Normal(36.2, 8) ⁵
$1/\omega$ - Time spent in recovered class	Varied
Probability of mild, severe, or critical case	Age Dependent ⁶
Probability of self- isolation	Varied
Probability of mask- ing	0.7 7
Probability of vacci- nating	Age Dependent ⁸

Table 1: Parameters comprising the main disease dynamics.

be filled up first, followed by other accommodations. The location these agents visit during the day is set to be a random building belonging to the institution they are attending.

Finally, each agent is given a place they go during the weekends. If the agent is aged 0-9, they will remain at home. Otherwise, the agent will go to two of the geographically closest commercial buildings throughout the weekend.

Transit between locations will not be accounted for in this model, with agents instantly moving from location to loca-255 tion.

Disease Mechanics Agents in this model flow through the SEIRS chain whenever they are exposed to COVID-19. The most involved dynamics here surround the movement of agents from the susceptible class to the exposed class. At 260 each location, agents have a probabilistic chance of contracting disease from any given agent who is at the same location and in the infectious state. We define the chance of the chance of one infectious agent passing COVID-19 to one susceptible agent to be equal to

> $(R_0 / (\text{time spent in infectious class}))$ (number of susceptible agents at the location)

¹(Ontario 2020)

³(Wölfel et al. 2020)

⁴(Verity et al. 2020)

⁵(Verity et al. 2020)

⁶(Verity et al. 2020; Ferguson et al. 2020)

⁷(Ontario 2022)

⁸(of Canada 2023)

250

265

²(Pung et al. 2020; Nishiura, Linton, and Akhmetzhanov 2020)

where R_0 is defined per the value in Table 1. This will result in each infectious agent spreading COVID-19 to a total number of agents equal to R_0 on average. So, on each time step, a susceptible agent at a location contracts the disease equal to the above probability, calculated and checked for

all infectious agents at their location. Once an agent is out of the susceptible class, they move through the other classes linearly, remaining in each class for the lengths of time outlined in Table 1.

270

- As previously mentioned, agents have a chance of self-275 isolating, but they are also hospitalized depending on the severity of their infection. If the case is mild, they are not hospitalized and either self-isolate or continue to go about their routine. If the case is severe or critical, then the agent
- is hospitalized in a regular or ICU hospital bed respectively. 280 In both hospitalization cases, they are effectively removed from the simulation and cannot infect any other agents.

Agents can also be masked or vaccinated based on actions taken by the planner. The effectiveness of masks and

- vaccines against COVID-19 is something that is constantly 285 being studied and debated, and as such, there is no single correct number reflecting this. So, we decided to approximate these values based on data describing the generally observed phenomena. We let an agent being masked multi-
- ply the chance of transmission by 0.8 (Gibson et al. 2023) in 290 both directions (multiplying by 0.8 if the susceptible agent is wearing a mask and also multiplying by 0.8 if the infectious agent is wearing a mask), and we let an agent being vaccinated multiply the chance of transmission by 0.4 (CDC
- 2023) in the receiving direction (multiplies by 0.4 if the sus-295 ceptible agent is vaccinated, but does not affect the rate at which the infectious spreads disease).

Putting this all together, the following snippet from the RDDL domain, which is a portion of the calculation deter-

mining whether or not an agent (denoted ?a here) is in the 300 exposed class, shows exactly how one agent is put into contact with all other contacts at its current location, and how the agent may contract COVID-19:

```
305
    sum_{?a2 : agent} [
     if ((AGENT_JOB(?a) == AGENT_JOB(?a2)) ^
    infectious(?a2) ^ is_working(?a2))
        then
           Bernoulli(
310
           ((if (masked(?a)) then MASK_FACT else 1) *
           (if (masked(?a2)) then MASK_FACT else 1) *
           (if (vaccinated(?a)) then VACCINE_FACT else 1)) *
              R_NAUGHT / infectious_total_time(?a2) /
315
                 (
                    1 + (sum_{?a3 : agent} [
                       (AGENT_JOB(?a) == AGENT_JOB(?a3)) ^
                       susceptible(?a3) ])
                 )
320
           )
           )
        else
           false ]
```

This snippet looks at the case when the agent, ?a, is at 325

their job, hence the AGENT_JOB(agent) fluent. The code for when an agent is at home or a store is the exact same as the above, but AGENT_JOB(agent) is changed to AGENT_HOME(agent) or AGENT_STORE(agent) respectively. First, all other agents in the simulation are iterated 330 over, using summation to aggregate them. For each agent that is at the same location as our target agent, if the agent is infectious, then the Bernoulli function is used to calculate the probability of the target agent contracting COVID-19, per the previously outlined infection contraction equation. 335 An agent being masked or vaccinated simply multiplies this probability by some number between 0 and 1 (0.8 and 0.4, respectively, for our setup). If this summation is greater than or equal to 1, then the initial agent is said to have contracted the infection. 340

Planner Actions and Goal

Mask Intervention One set of actions the planner can take has to do with mask implementation. The two actions in this category are to enforce a mask mandate among all agents, or to only enforce it among post-secondary students. When 345 either of these actions is taken, the affected agents apply a mask at a rate of 0.7 (Ontario 2022). When a mask action is taken, it must be active for a minimum of 14 time steps, at which point the planner can re-apply the mandate or remove it. This forces the mask intervention methods to be applied 350 on a "per week" basis, more closely mimicking real intervention when compared to the possibility of switching it on and off on any time step.

The following snippet shows exactly how this masking is tracked for an agent, ?a:

```
if ((all_mask ^ (~mask-implemented)) |
(student_mask ^ AGENT_STUDENT(?a)
(~mask-implemented-students)))
    then Bernoulli (MASK CHANCE)
                                                              360
else if (mask-implemented | mask-implemented-students)
    then masked(?a)
else
                                                              365
    false
```

Here, the boolean fluents all_mask and student_mask denote the planner actions to mask all agents or only students, and the boolean fluents mask-implemented and mask-370 implemented-students denote whether or not a mask mandate has been applied in the previous 14 time-steps for both groups. So, agents apply masks at some probability (0.7 here) when a mask action is first introduced, and that choice is held constant throughout the 14 time steps.

Vaccine Intervention The other set of actions the planner can take revolves around vaccine mandates. In a very similar fashion to the mask intervention implementation, a vaccine mandate can be implemented for all agents or only for post-secondary students. When the planner takes one of these actions, agents become vaccinated at a given rate based on their age bracket (of Canada 2023). Once an agent is vaccinated, they remain so for the entire duration of the simu-

355

375



Figure 2: Each graph shows agent numbers observed in a specific class at each time step over 1000 trials - each trial initialized with 93 susceptible and 12 infectious agents. The maximum and minimum values observed at each time step are shown in a dark colour along with a random sample of 100 trials in a lighter colour. Each graph has a black line corresponding to the agent numbers at each time step predicted by solving the system of differential equations sharing the same initial configuration. The initial jump in agents on the bottom two graphs has to do with the 12 infectious agents having the same infectious period, causing them to move out of the class at the same time.

lation. In reality, vaccination effectiveness wanes over time, but since our experiments examine a relatively short period 385 of time, this is not something that is considered.

Planner Goal The goal of the planner is to maximize a reward function, defined by the following equation:

```
Reward = (Students Masked * Student Mask Penalty)
390
              (Others Masked * General Mask Penalty) +
              (Students Vacc * Student Vacc Penalty) +
              (Others Vacc * General Vacc Penalty) +
              (Hospital Beds Exceeded * Regular Bed Penalty) +
              (ICU Beds Exceeded * ICU Bed Penalty)
```

Each penalty is a negative value, meaning that this equa-395 tion has a maximum 0, which the planner will try to be as close as possible to. On each time step, the number of agents who are masked or vaccinated is multiplied by the respective penalty value, with post-secondary student penalties being lower than general population penalties. In the problem file, 400 a number of regular and ICU hospital beds that mark the "hospital capacity" are defined. If either of these capacities is exceeded, then a penalty is incurred per bed over the limit on every time step this holds. The penalty values were defined in a way to elicit realistic behaviour from the planner. 405 They can be modified based on what is deemed to be important.

Evaluation

Base Model A base model was first established, where planner actions have no effect, and agents do not self-isolate 410 or become hospitalized. Stochasticity was removed where reasonable, having time spent in the exposed class equal to 9, time in the infectious class equal to 16, and time in the recovered class equal to 14. One problem file was created, which was used to run 1000 randomly seeded trials, and ranges 415 for the number of agents in each SEIR class at a given time were identified. Figure 2 shows the dynamics of each class of this model with 105 (initially with 93 susceptible and 12 infectious) agents over 300 time-steps with a black line superimposed representing theoretical agent numbers created 420 through solving a system of ordinary differential equations (ODEs) over time. This system of ODEs was initialized with

the same parameters as with the base model in RDDL. It is important to note that this ODE model assumes that the pop-

- ulation in question is evenly distributed in the space, and it 425 does not operate on the agent level. Therefore, it is a rather crude estimate but can be used as a starting point to compare the quality of an agent-based model. While the theoretical values are within the maximum and minimum observed
- values from the agent-based model at each time step, they 430 seem to reach equilibrium values that do not exactly line up with the output of the agent-based simulations - this is likely due to the agent-based model being built on the specific geography of Kingston, leading to unique simulation
- behaviours. For example, higher concentration areas (such 435 as post-secondary institutions) will cause disease to spread faster and more effectively, leading to a larger population becoming infected compared to a population that is evenly distributed about some space. These two models being gen-
- erally similar in behaviour indicates that the RDDL model 440 accurately simulates disease dynamics, with the four SEIR classes generally oscillating towards a reasonable equilibrium. It is also important to note that even with the same parameter configuration, the ranges observed for each SEIR class are quite large. This is likely due to the parameter val-445
- ues driving these simulations being imprecise, having been configured based directly on the available data, which is typically incomplete or lacking in quality.

Generating Results A series of simulations were run, and data was extracted from the outcomes. Each simulation op-450 erated in a world with around 100 agents (varies slightly depending on random initialization of agents in each problem file) and was run for 100 time-steps. Due to this scale, the regular hospital bed capacity was set to be 2, and the

- ICU bed capacity was set to be 1. With this configuration, 455 dynamics were observed that are consistent with the theory backing up the creation of the model, and we were able to draw meaningful conclusions from the actions taken by Jax-Planner.
- To understand the model's performance in depth, a num-460 ber of configurations were compared with each other. With actual COVID-19 dynamics, time spent in the recovered class can be several months (Diani et al. 2022). However, seeing as how our model only runs for 100 time steps, this
- value was varied between 2, 14, and 28 time steps. Another 465 parameter that was varied was the rate of self-isolation. The rate of self-isolation is a parameter that is particularly difficult to estimate due to a lack of reliable data, and it is often the case that individuals may continue their day-to-day ac-
- tivities unaware that they are infected at all. For this reason, 470 we vary the values of this parameter between 0, 0.3, 0.5, and 0.7. With respect to the actions that JaxPlanner can take, we looked at cases where the planner did not have masking/vaccinating capabilities in addition to the cases where it did to see how much of an impact it actually had on the reward cal-475
- culation. Simulations were run for every possible parameter configuration from the options listed above (i.e. the crossproduct of all parameter combinations).

For every configuration, simulations were run several times over, and their results were aggregated. First, only 480

the time spent in the recovered class was varied, with the model having no self-isolation or intervention. For each of these three configurations, 20 problem files were generated that the simulations drew from. 20 simulations were run per problem file, with the simulation seeds being tracked along with the relevant data produced. The other configurations were then incorporated, where self-isolation rates and planner actions vary, with simulations drawing upon the previously generated problem files with respect to time spent in the recovered class. Using the saved seeds, simulations for 490 these configurations were run 20 times on each of their respective 20 problem files.

Figure 3 and figure 4 are examples of how typical simulations played out, and they both have the exact same parameter, agent, and geography configuration. The only dif-495 ference is the seed used for each trial. In figure 3, we see that the planner elected to vaccinate students around time step 30, coinciding with increasing cases and agents becoming hospitalized. As soon as this action was taken, the active cases began to diminish, showcasing the effectiveness 500 of agent vaccination in this model. Around time step 40, the planner masked students, lining up with a spike in hospitalizations. Note that, like most behaviours in this model, hospitalization is stochastic. So, even though the planner may have performed the optimal actions at a given time, the num-505 ber of regular hospitalizations reached 5, which is 250% of the specified capacity of 2. On a larger scale, this stochasticity would not have such a pronounced effect. Now, in figure 4, we do not see the same actions taking place, and in fact, no actions are taken at all over the course of the simulation. This 510 is quite significant because it implies that even under the exact same conditions, the planner policy (determined at the outset of the simulation) can vary quite substantially. In the case of disease models, this means the difference between masking and/or vaccinating a population, and not doing so. 515 This variability can be attributed to a number of reasons such as the simulation length, and agent numbers and distribution. However, it is likely that the primary reason for this variability is the high level of uncertainty associated with parameter values, which can greatly affect simulation outcomes. 520 Though, in order to say this concretely, extensive sensitivity analysis would need to be performed on this model.

Evaluating JaxPlanner Performance Once all simulations were run, the resulting data was compared for each configuration. Focusing on the total reward gives an idea 525 of how well each configuration performed relative to each other. To visualize the rewards for all configurations, figure 5 shows a series of box plot graphs - one for each recovery rate that was chosen. Each configuration's box plot is built using the average total reward for each of the 20 problem files, 530 for which 20 simulations were run. We can visually deduce from this that self-isolation and intervention improved the model performance, with a combination of both generally leading to the best total reward. Table 2 summarizes this in a different way, showing the average percent improvement of different configurations when compared to the base model with no intervention or self-isolation. This implies that the goal of the model was reached. By implementing interven-

485



Figure 3: Outcome for a typical simulation, with planner action on the left, an SEIR graph in the middle, and hospitalizations on the right.



Figure 4: Another outcome for a simulation with the exact same configuration that produced the output seen in figure 3

tion methods, the planner was able to reduce the severity at which hospitals operated above capacity, all while activating 540 these interventions as little as possible. The improvement of the reward function for all configurations when compared to the base model is evidence of this.

Configuration		% Improvement
Rate of Self-Isolation: 0	No intervention	Base
	Intervention	29.4
Rate of Self-Isolation: 0.3	No intervention	21.7
	Intervention	41.6
Rate of Self-Isolation: 0.5	No intervention	49.2
	Intervention	57.1
Rate of Self-Isolation: 0.7	No intervention	66.8
	Intervention	68.3

Table 2: The percent improvement for each configuration over the base model.

Discussion

- Limitations The most stand-out limitation of this research 545 was the operating capacity of JaxPlanner. When JaxPlanner is searching for a plan, there is a maximum memory allotment that cannot be exceeded. This is acceptable for most small-scale models, but with a model as complex as this one,
- the memory limit was quickly reached. It is for this reason 550 that the simulations were run with around 100 agents - significantly lower than the population of Kingston, which is around 100,000. This number was the most that JaxPlanner would allow for, given 100 time-steps (50 days). Despite

this, meaningful results were obtained, with dynamics mim-555 icking that which is found in the real world.

Another limitation was the lack of high quality data that was available. Much of the data with respect to disease dynamics is being actively researched, with results varying quite significantly between publications. Rough estimates 560 were obtained, but these are often based on studies that operate all around the world - far from Kingston, Ontario. It is believed that this lack of specificity resulted in the majority of the model and policy variation. With nearly all parameters having quite a wide value of ranges it may take on, it is 565 no surprise that simulations varied quite dramatically from each other, even with identical configurations.

There were also assumptions made with regard to the disease dynamics themselves. In reality, people have a chance of dying due to COVID-19 complications. A more realis-570 tic model would incorporate births and deaths and may even see individuals in the population moving through age brackets. The rate of transmission also acts differently in the real world, being affected by age and the type of infection (symptomatic or asymptomatic), amongst many other factors. 575

Future Work Aside from collecting and refining the data driving model simulations, future work may add more complex disease mechanics, noting the subtle differences that would emerge. There are near endless possibilities for modifications in disease models, and the model described in this paper was kept relatively simple in order to clearly demonstrate the capabilities of RDDL in solving this type of problem. Fine-tuning on this existing model could also be attempted, utilizing optimization techniques to improve the



Figure 5: Each graph shows box plots for each configuration based on the total reward for a whole simulation (a number smaller in magnitude is better). Each configuration was run on 20 times on 20 problem files (identical problem files and simulation seeds per graph), and the average total rewards for each problem file are the values that comprise each box plot. The following configurations are in each graph: no intervention and a self-isolation rate of 0 (Base), no intervention and a self-isolation rate of 0.3 (OSI 0.3), no intervention and a self-isolation rate of 0.5 (OSI 0.5), no intervention and a self-isolation rate of 0.7 (OSI 0.7), intervention and a self-isolation rate of 0 (OIn), intervention and a self-isolation rate of 0.3 (InSI 0.3), intervention and a self-isolation rate of 0.7 (InSI 0.7).

- performance and scale allowable via JaxPlanner. As an alternative, other planners may be tested to see how performance and results may differ. It may be the case that some planners perform better than others, and that should be explored to get a better understanding of the problem space. Finally,
- given that data is almost never perfect, one may attempt to verify the quality of existing data and identify which parameters affect the outcome of simulations the most via modern sensitivity analysis techniques in conjunction with different planners. This has the possibility to assist data collectors in terms of what to focus on in order to improve the quality of disease models.

Conclusion

This paper describes an agent-based COVID-19 model implemented in RDDL based on the geography of Kingston, Ontario, incorporating complex disease mechanics. Through the use of JaxPlanner, intervention methods were added in the form of masking and vaccination actions, which were successfully executed to ensure that hospitals were operating at or below capacity as often as possible while mini-

- mizing the time that these intervention methods were active.
 Primarily, this is both an exploration into the power of Jax-Planner in complex multi-agent domains, and of the versa-tility of RDDL in large-scale simulations. This model also underscores the greater issue of data quality, highlighting an
 important domain that is strongly affected.

615

References

Adiga, A.; Dubhashi, D.; Lewis, B.; Marathe, M.; Venkatramanan, S.; and Vullikanti, A. 2020. Mathematical models for covid-19 pandemic: a comparative analysis. *Journal of the Indian Institute of Science*, 100(4): 793–807.

Balcan, D.; Colizza, V.; Gonçalves, B.; Hu, H.; Ramasco, J. J.; and Vespignani, A. 2009. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the national academy of sciences*, 106(51): 21484–21489.

620 Boeing, G. 2017. OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks. *Computers, Environment and Urban Systems*, 65: 126–139.

Britton, T. 2020. Basic prediction methodology for covid-19: estimation and sensitivity considerations. *MedRxiv*, 625 2020–03.

Carley, K. M.; Fridsma, D. B.; Casman, E.; Yahja, A.; Altman, N.; Chen, L.-C.; Kaminsky, B.; and Nave, D. 2006. BioWar: scalable agent-based model of bioattacks. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Sys-* 630 *tems and Humans*, 36(2): 252–265.

CDC. 2023. CDC Covid Data tracker.

Diani, S.; Leonardi, E.; Cavezzi, A.; Ferrari, S.; Iacono, O.; Limoli, A.; Bouslenko, Z.; Natalini, D.; Conti, S.; Mantovani, M.; et al. 2022. SARS-CoV-2—the role of natural immunity: a narrative review. *Journal of Clinical Medicine*, 11(21): 6272.

Ferguson, N. M.; Laydon, D.; Nedjati-Gilani, G.; Imai, N.; Ainslie, K.; Baguelin, M.; Bhatia, S.; Boonyasiri, A.; Cucunubá, Z.; Cuomo-Dannenburg, G.; et al. 2020. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team. *Imperial College COVID-19 Response Team*, 20(10.25561): 77482.

Gibson, G. C.; Woody, S.; James, E.; Weldon, M.; Fox, S. J.; 645 Meyers, L. A.; and Bhavnani, D. 2023. Real time monitoring of COVID-19 intervention effectiveness through contact tracing data. *Scientific Reports*, 13(1): 9371.

650

Government of Canada, S. C. 2021. Census profile, 2021 census of Population - Profile Table.

Harmanani, M. 2023. Modelling the Spread of COVID-19 in Indoor Spaces using Automated Probabilistic Planning. *arXiv preprint arXiv:2308.08190*.

Hinch, R.; Probert, W. J.; Nurtay, A.; Kendall, M.; Wymant, C.; Hall, M.; Lythgoe, K.; Bulas Cruz, A.; Zhao, L.; 655 Stewart, A.; et al. 2021. OpenABM-Covid19—An agentbased model for non-pharmaceutical interventions against COVID-19 including contact tracing. *PLoS computational biology*, 17(7): e1009146. Kamerlin, S. C.; and Kasson, P. M. 2020. Managing coronavirus disease 2019 spread with voluntary public health measures: Sweden as a case study for pandemic control. *Clinical Infectious Diseases*, 71(12): 3174–3181.

Kerr, C. C.; Stuart, R. M.; Mistry, D.; Abeysuriya, R. G.;

- Rosenfeld, K.; Hart, G. R.; Núñez, R. C.; Cohen, J. A.; Selvaraj, P.; Hagedorn, B.; et al. 2021. Covasim: an agent-based model of COVID-19 dynamics and interventions. *PLOS Computational Biology*, 17(7): e1009149.
- Liu, J. J.; Bao, Y.; Huang, X.; Shi, J.; and Lu, L. 2020. Mental health considerations for children quarantined because of
 COVID-19. *The Lancet Child & Adolescent Health*, 4(5):
 347–349.
 - Mahmood, I.; Arabnejad, H.; Suleimenova, D.; Sassoon, I.; Marshan, A.; Serrano-Rico, A.; Louvieris, P.; Anagnostou,
- A.; JE Taylor, S.; Bell, D.; et al. 2022. FACS: a geospatial agent-based simulator for analysing COVID-19 spread and public health measures on local regions. *Journal of Simulation*, 16(4): 355–373.
- Nishiura, H.; Linton, N. M.; and Akhmetzhanov, A. R. 2020. Serial interval of novel coronavirus (COVID-19) infections.
- International journal of infectious diseases, 93: 284–286.

of Canada, P. H. A. 2023. Covid-19 vaccination: Vaccination coverage.

Ontario, P. H. 2020. Respiratory Virus Outbreaks during the

- 2020-21 Season: Considerations for Public Health Planning. Ontario, P. H. 2022. Association between Mask Mandates and Population-level COVID-19 Outcomes – What We Know So Far.
- Parker, J.; and Epstein, J. M. 2011. A distributed platform
 for global-scale agent-based models of disease transmission.
 ACM Transactions on Modeling and Computer Simulation (*TOMACS*), 22(1): 1–25.

Pung, R.; Chiew, C. J.; Young, B. E.; Chin, S.; Chen, M. I.; Clapham, H. E.; Cook, A. R.; Maurer-Stroh, S.; Toh, M. P.;

695 Poh, C.; et al. 2020. Investigation of three clusters of COVID-19 in Singapore: implications for surveillance and response measures. *The Lancet*, 395(10229): 1039–1046. Silva, P. C.; Batista, P. V.; Lima, H. S.; Alves, M. A.;

Guimarães, F. G.; and Silva, R. C. 2020. COVID-ABS: An agent-based model of COVID-19 epidemic to simulate

health and economic effects of social distancing interventions. *Chaos, Solitons & Fractals*, 139: 110088.

700

715

Sjödin, H.; Johansson, A. F.; Brännström, Å.; Farooq, Z.; Kriit, H. K.; Wilder-Smith, A.; Åström, C.; Thunberg, J.;

Not Söderquist, M.; and Rocklöv, J. 2020. COVID-19 healthcare demand and mortality in Sweden in response to nonpharmaceutical mitigation and suppression scenarios. *International journal of epidemiology*, 49(5): 1443–1453.

Venkatramanan, S.; Chen, J.; Fadikar, A.; Gupta, S.; Higdon,

710 D.; Lewis, B.; Marathe, M.; Mortveit, H.; and Vullikanti, A. 2019. Optimizing spatial allocation of seasonal influenza vaccine under temporal constraints. *PLoS computational biology*, 15(9): e1007111.

Verity, R.; Okell, L. C.; Dorigatti, I.; Winskill, P.; Whittaker, C.; Imai, N.; Cuomo-Dannenburg, G.; Thompson, H.; Walker, P. G.; Fu, H.; et al. 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet infectious diseases*, 20(6): 669–677.

Wölfel, R.; Corman, V. M.; Guggemos, W.; Seilmaier, M.; Zange, S.; Müller, M. A.; Niemeyer, D.; Jones, T. C.; Vollmar, P.; Rothe, C.; et al. 2020. Virological assessment of hospitalized patients with COVID-2019. *Nature*, 581(7809): 465–469.