

CSC 455+555 SoCDAI

P2 Instructions

This assignment will introduce you to agent-based simulation modeling to work on one of the most critical health crises of our times. You will be creating a simulation of a society that resembles the real world to study the mobility patterns of community and the spread of COVID-19. The central question we are trying to find an answer to through this simulation is, what are the impacts of following social distancing norms in a society and comparing it with a society that doesn't follow these norms.

One of the most prominent impacts of COVID-19 is reduced social gatherings in public places such as grocery stores, parks, and workplaces. This is necessary to control the spread of the virus but it also adversely affects the economy and especially worsens the situation for industries that rely heavily on social gatherings. Needless to point out, the boredom and psychological impacts of staying at home for prolonged periods make it desirable for most people to be able to get back to life as it was.

Simulation (Agent-Based Simulation Modeling)

Create a simulation environment to model a community showing the mobility patterns of people in a society. Use [MESA](#) (An Agent-Based Modeling framework in Python), which is easy to use and is well documented. Try to achieve a basic version of the simulation before trying to incorporate anything complicated.

The simulation environment should contain:

- People, modeled as agents, of two types
 - Initial population 1,000
- Homes
 - Each person is a resident of one home
 - Each home is the residence of four people
 - 250 homes in total to accommodate 1,000 agents
 - If an agent is infected and stays at home, it infects all other agents living in that home as well
- Parks
 - Two parks
- Grocery stores
 - Five grocery stores

Your simulation should support mobility of these agents across the places specified above. For instance, people may go visit grocery stores or parks. An agent gets infected (with a given probability) if it is in the same location as some other infected agent at the same time step.

Configuration for the simulation

1. Modeling Mobility Patterns

Use one day as one time step for your simulation. To simplify the problem you can assume each agent is only at one place throughout a given time step, so if someone visits a grocery store we can assume that the agent was in the grocery store for that entire time step (i.e. for the entire day).

Refer to the following table for probabilities of mobility across various infrastructure in the society at the start of each day (the probabilities remain constant on and across all days)

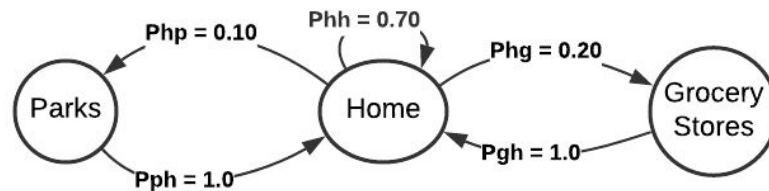


Figure 1. Transition probabilities for agents to visit different places in the society. After each visit to either grocery store or park the agent returns home (in the next time step)

2. Modeling the spread of the virus

Each person is either infected or non-infected. A non-infected agent transitions to an infected state when it comes in contact with an infected agent. For simplicity, assume an agent gets infected (i.e. it transitions to the infected-asymptomatic state) on the next time step. Infection has following substates: asymptomatic, symptomatic, critical, cured, and deceased.

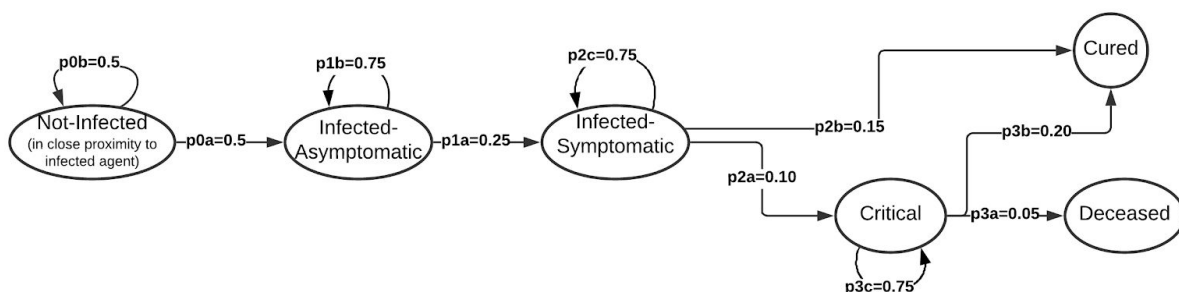


Figure 2. Transition probabilities across various disease states

Seeding the simulation

Assume three starting states with the following population as infected:

1. 10% as infected
2. 25% as infected
3. 50% as infected

Basic implementation [50 points]

Your first task is to create a simple simulation environment where agents (both infected and non-infected) move between places. With the simple environment, you should analyze the spread of COVID-19 in an artificial society. You should seed your simulation with the three starting states of 10%, 25%, or 50% (one at a time) of the population infected. For each simulation, answer the following questions based on results from your simulation.

Assuming everyone who is cured develops antibodies and can never get infected again (hence we have a termination criteria for our simulation, when no one has the infection, i.e., when all infected are either cured or deceased).

For each starting state, using your simulation, report the following:

- When does the simulation stabilize (reaches a termination criteria or experiences no further change)? How long will it take to reach that point (report your results in number of days)?
- When was the infection at its peak?
- How many people die by the time the simulation stabilizes?
- Plot graphs for the spread of the disease based on results from your simulation (with the number of active COVID-19 cases on the y-axis and time steps (days) on the x-axis)

Note that you may need to execute each simulation several times to mitigate the effect of randomness. The above answers could be based on an average of multiple runs.

Adding Quarantine Centers [25 points]

As a measure to contain the virus the authorities introduce '*quarantine centers*' where infected patients can stay until cured (or deceased) and avoid spreading the virus.

- Quarantine center
 - Create one quarantine center for the entire society
 - Max capacity is 100 (10% of the population)
 - If the quarantine center is completely occupied, infected agent continues to move to other places
 - Cured agents continue to move as before
 - Deceased agents are removed from the center
 - Assume no agent in the simulation gets infected in the quarantine center

Disease State	Probability of going to quarantine center	Explanation
---------------	-------------------------------------------	-------------

Infected Asymptomatic	$p_{\text{asympt-quarantine}} = 0$	Asymptomatic people are unaware of their illness and hence don't go to the quarantine center.
Infected Symptomatic	$p_{\text{symp-quarantine}} = 0.20$	20% of the symptomatic patients choose to go to the quarantine center.
Critical	$p_{\text{critical-quarantine}} = 1.0$	All critical patients are sent to quarantine centers.

1. Report how your answers will change for all parts of question 1 with the quarantine center in place?
2. Did the quarantine center reach its limit of 100 at any time step? If yes, at what time step does it happen? Show your results using a plot.

Enforcing Social Distancing Norms [15 points]

To further contain the spread of the virus the authorities have further introduced social distancing norms to be followed (like wearing mask and practicing social distancing in public places)

To incorporate the impact of following social distancing norms we reduce the transition probability from the state not-infected to infected ($p_{0a}=0.5$ in Figure 2). We assume it's 5 times less likely to be infected if one follows these social distancing norms, hence we use $p_{0a} = 0.10$ and $p_{0b}=0.90$ with social distancing norms in place.

Answer all parts of question 1 with this setting.

Does following social distancing norms improve the situation? Does it reduce the number of deaths? Use plots to highlight the difference compared to before.

Conclusion [10 points]

- Share other interesting findings you uncover using your simulation.
- How much of an impact did adding quarantine centers (question 1) add and what was the impact of following social distancing norms (question 3) in controlling the spread?
- Based on your understanding of the simulation, identify 2 influential factors in spreading the virus? Suggest changes to improve the situation based on your analysis.

Bonus Question [10 points]

Using real world data to seed your simulation

Now that you have a basic simulation model you may want to try this on real world data. Here is a link to a [Kaggle dataset](#) that you can use. It is an aggregated COVID-19 dataset that has the stats for the *spread* (i.e., confirmed cases) of COVID-19 and *mobility patterns* of people (in public places) from 133 countries. You can pick a country and run your simulation with data for that country in isolation (i.e. we assume no one moves in or out of a country). You may want to scale down the population and all public infrastructure to reduce run time and to simplify the problem.

For instance, following are a few things that you should consider experimenting with:

- Replace mobility numbers with numbers from the dataset for a specific country under study.
- You may want to include more public infrastructures like, visits to retail stores and workplaces (and other such places for which we have the data in the dataset)
- Replace the default number of four members in each house with a number proportional to the population density of that country. (Assume two members for each household for the US and use this as reference to calibrate household size for each country based on the given population density for that country)
- Decide on the capacity (per population) of hospitals based on the Health system index. (for reference assume US has the max capacity to treat 10% of its population at a time and compute numbers for other countries based on this)
- You may have to modify some of the probabilities in Figure 1 based on the data (as the data has the number of active cases and testing for each day). Specify clearly if you are changing any of the transition probabilities across diseases states (by real counts or percentage of populations).

The above list is not comprehensive and you can think of other parameters that you can either replace with some statistical estimate based on the data or learn it using your simulation. Specify clearly what have you changed to incorporate the real data in your simulation.

Report results for three countries for each of the categories below:

- Countries with high population (or you may choose a country with high population density)
- Countries with high GDP per capita
- Countries with most COVID-19 Testing per capita

Deliverables:

Due on: 10:00 pm ET, October 4, 2020

Please submit all your files in one zip folder which should include the following items:

1) **Report:** Your report will be the document that includes all the analysis, results and observations you make for this assignment, your report (PDF) should include:

- Results/Observations from your analysis.
 - Report results from all your analysis and answer each question.
- Back your arguments using the empirical evidence and include citations wherever relevant.
- Include Graphs, tables, Figure, or any additional material you think is relevant for your work.

2) **Source Code:** Your source code, along with a Readme with instructions and requirements to run your code. Your source code should include:

- Your *code* files: Make sure to fix all file paths, if your code breaks you lose points. (You may include Github repository link as submission but if you do please make sure to provide appropriate access to the TA)
- *requirements.txt*: Include a requirements.txt that includes all the dependencies/libraries with the correct versions. If possible include a script to setup the virtual environment to execute your code.
- *readme* that has all the steps to run your code, mention any dependencies not included in requirements, or if anything needs to be downloaded, provide a link here. If for any reason we are unable to run your code we are going to lose significant points, so make sure you have every detail included.

Software artifacts are only as good as their documentation, if we can't figure out how to run it, it doesn't matter how good it is.