CAsimulations: Modeling of topological dynamics in a disease using cellular automata

Jorge Andrés Ibáñez Huertas¹, Carlos Isaac Zainea Maya²
¹Departamento de Matemáticas, Universidad Central, Bogotá, Colombia. jibanezh@ucentral.edu.co
²Departamento de Matemáticas, Universidad Externado de Colombia, Bogotá, Colombia. carlos.zainea@uexternado.edu.co

Abstract—The prediction of the behavior of the disease, the level of affection in a population and the ways to control it are the most important aspects studied by epidemiology using tools such as historical data and mathematical models. So, our objective is (1) to provide a tool capable of analyzing epidemiological phenomena starting from the most common social interactions within a group of individuals, (2) To provide a methodology to build epidemiological models from patterns and logical rules, (3) Determine the impact of social interactions on the spread of a disease. This paper describes the logical construction of two epidemiological models in cellular automata together with two of their variations based on topological and dynamical principles.

I. INTRODUCTION

Today, there are many mathematical models to project the progress of a disease. Contingencies such as the one experienced with COVID-19 promote a greater understanding of these models, and new theories are created around this problem from different perspectives. Among these models, the compartmental models allow studying the evolution of the different states considered. For example, with SIR the evolution of susceptible, infected and recovered individuals is studied separately; A peculiarity of this model is that it projects the growth of individuals in each of its compartments, and it is possible to establish at what point we are before the disease becomes endemic, however, it does not take into account the possible factors related to the care, risk or exposure of some individuals.

In this work we will show, based on simulation techniques, a possible approach to compartmental models that take into account the closeness with other individuals. To achieve this, we use cellular automata theory and an extension of it, using topological neighborhood concepts and fundamental neighborhood systems. In theorem [1] we show how the impact of one cell on the others constitutes a fundamental system of neighborhoods, and after we define some evolution rules for SI, SIS and SIR models for each cell. Finally, an example that allows us to study the models described in this paper in a village of 49 people.

In the development of this work, a Python package called CAsimulations was made, the documentation of this package can be consulted at https://grupo-de-simulacion-con-automatas.github.io/CAsimulations-Modelacion-de-dinamicas-topologicas-en-la-propagacion-de-una-enfermedad-usando-CA/

II. PRELIMINARIES

A. EPIDEMIOLOGICAL STUDY

One of the most important objects of study in epidemiology is to determine the endemic nature of the disease, that is, whether the disease will affect the population for a long time or whether it will gradually disappear. Usually, this is determined by the indicator \( R_0 \) known as the basic reproduction number and which represents the number of individuals that were infected by patient zero in a susceptible population. Normally, if \( R_0 < 1 \) the disease will slowly disappear and if \( R_0 > 1 \), we may have an endemic case.

Heesterbeek and Dietz defining the basic reproduction number as:

\[
R_0 = \int_0^\infty b(t)F(t)dt,
\]

where \( b(t) \) represents the average number of new infections that an infected individual will produce during time \( t \) and \( F(t) \), known as the survival function, represents the probability that a newly infected individual will remain in that state for at least time \( t \) [5], [6].

Generally, when we talk about epidemiological models, we consider three states or classes into which we can divide the population over time: those who can contract the infection, those who are infected and those who are recovered. If the recovered populations do not acquire permanent immunity, we have a SIS model, in which the susceptible population can contract the infection and once they recover they come back to susceptible state. On another way, if the recovered individuals generate disease immunity, we have an SIR model [3].

B. THE SIS MODEL

The SIS model considers two possible states, susceptible (S) and infected (I). Variations between states are given by new infections and individuals recovering from the disease. In addition, each state is affected by the mortality/birth rate and/or the parameter describing death caused by the disease.

Normally, when talking about epidemiological models with disease deaths, four parameters are considered:

- The infection rate \( \beta \), which represents the probability of a susceptible individual acquiring the disease after contact with an infected individual.
• The **recovery rate** \( \alpha \), which can be understood as the probability that an infected person will recover from the disease. It sometimes represents the average time it takes for an infected person to recover from the disease [4].

• The **birth/mortality rate** \( \mu \), which in the case of traditional models are assumed to be equal. The birth rate indicates the number of individuals entering the space, and the mortality rate represents the individuals who die from causes other than disease.

• The **death rate due to disease** \( \theta \), which indicates the probability that an infected person is likely to die from the disease.

In our work, we will use a constant population and thus \( S' + I' = 0 \). With this in mind, we can describe the model from a system of differential equations as follows:

\[
\begin{align*}
S' &= \mu(1 - S) + (1 - \theta)\alpha I - \beta SI, \\
I' &= \beta SI - (1 - \theta)\alpha I - \mu I.
\end{align*}
\]  

(2)

In [4] they use the equation 1 to determine in an analogous way that \( R_0 = \frac{\beta}{\alpha(1 - \theta) + \mu} \) so we will not go deeper into this process.

C. THE SIR MODEL

For this model, the state of immunity (R) to the disease is considered. Unlike the SIS model, in the SIR model there is no interaction from state I to state S, since it is assumed that individuals who recover from the disease will not be able to contract it again, so they will pass to state R. In this way,

\[
\begin{align*}
S' &= \mu(1 - S) + \alpha\theta I - \beta SI, \\
I' &= \beta SI - \alpha I - \mu I, \text{ and} \\
R' &= \alpha I - \alpha\theta I - \mu R.
\end{align*}
\]  

(3)

Using an argument similar to the one used to determine the value of \( R_0 \) in the SIS model, it is possible to determine that for the SIR model, \( R_0 = \frac{\beta}{\alpha + \mu} \).

D. CELLULAR AUTOMATAS

Cellular automata (or CA) were born with the work of Von Neumann in the late 1940s with his paper "The General and Logical Theory of Automata" in which he first put forward ideas for a machine capable of self-replication. He worked on a discrete two-dimensional system to develop fairly complex ideas for a machine capable of self-replication. We will make use of tools such as the fundamental systems of neighborhoods.

II. EPIDEMIOLOGICAL MODELS IN CA

Let us think for a moment that if an individual susceptible to a disease has contact with many infected individuals, it can get sick much more easily than an individual who has contact with few infected individuals. Now, how many interactions with infected individuals, out of all those that a cell can have, are enough to generate a considerable probability of contagion? To answer this we must first determine the nature behind the possible relationships between cells, and for this we will make use of the minimal neighborhoods and some of its properties, for this we will use concepts set forth in [10], [2], [9], [11].

Definition: A minimal neighborhood of \( x \) is the intersection of all the neighborhoods of \( x \). We denote it as \( U_x \).

A. SOCIAL INTERACTIONS AND IMPACTS

Unlike the work done in [1] in which each cell represented a region, we will consider each division as a single individual that will be provided with a set of qualities such as health status, age, neighbors, etc. These qualities will be given by the needs of the model we are developing, for example in the simplest SIR and SIS models it will not be necessary to endow the cells with ages, but we will have to take into account their neighborhoods and their health states.

Let us now think about the characteristics behind the close relationship between individuals or cells. We will denote the relationship between cells with the symbol, \( \sim \) and having said that we have that:

• All cells are in contact with themselves, so for each cell \( x \) it is satisfied \( x \sim x \).

• If a cell were in contact with some other cell, then that cell would be in contact with the first cell, i.e., \( x \sim y \) implies \( y \sim x \).

• If a cell interacts with two other cells, it does not necessarily imply that they interact with each other, so \( x \sim y \) and \( x \sim z \) do not imply that \( y \sim z \).
Let us consider for example the set \( X = \{a, b, c\} \) and the topology \( \mathcal{A} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{a, c\}, \{a, b, c\}\} \). The interaction relations present in \( \mathcal{A} \) are:

\[
\begin{align*}
a &\sim a, \quad b \sim b, \quad a \sim c, \quad \text{and} \quad b \sim c,
\end{align*}
\]

together with their equivalent symmetric relations.

The notion of relation in the previous example of cell interaction can be taken a bit further. Let us think for a moment about the impact that a behavior on cell \( b \) can have on cell \( a \), although these two cells do not interact with each other, the relationship that each of them has with cells \( c \) and \( d \) can have an impact on the other. Whereupon, we define the following interaction relationship:

**Definition:** We define the degree of impact between two points \( a \) and \( b \) as the least amount of interactions needed to get from \( a \) to \( b \). To recognize the degree of impact between two points, we will use the notation \( a \sim_n b \) where \( n \in \mathbb{N} \) denotes the least amount of interactions between \( a \) and \( b \).

If we return to the example, we can identify the following degrees of impact:

\[
\begin{align*}
a &\sim_0 a, \quad b \sim_0 b, \quad a \sim_1 c, \quad a \sim_1 b, \\
b &\sim_1 a, \quad b \sim_2 c, \quad c \sim_0 a
\end{align*}
\]

**Theorem:** The degrees of impact of an \( x \) cell define a fundamental system of neighborhoods.

**Proof.** Let \( x \in \mathcal{L} \) and \( \mathcal{V}(x) \) be a family of neighborhoods of \( x \). Define the set \( A_0 \) as the set of points with degree of impact with \( x \) is equal to zero and recursively to the sets \( A_k \), whose elements have degree of impact with \( x \) is equal to or less than \( k \). Clearly, \( A_i \subseteq A_j \) for \( 0 \leq i \leq j \) and thus \( A_i \in \mathcal{V}(x) \) for \( i = 0, 1, \ldots, n \).

Let us now see that the set \( A_0 \) is the minimal neighborhood of \( x \). Consider \( y \in U_x \), in particular, \( y \in \bigcap \mathcal{V}(x) \). If the degree of impact of \( y \) with \( x \) is greater than zero by definition of impact grade, we can assert that there exists \( z \in \mathcal{L} \) such that \( z \sim x, z \sim y \) and \( x \not\sim y \) and thus \( x \) and \( y \) are separable points.

Consider \( y \in A_0 \) and by the definition of zero impact degree, we claim that the points \( x \) and \( y \) are not separable, so \( y \in \mathcal{V}(x) \) and with this we conclude the proof. \( \square \)

One thing to keep in mind is that the degree of impact alone does not provide us with a measure of the impact of state changes of “distant” cells (or of the degree of impact greater than zero). We can define and understand these impact measures as the probability that a state change will affect the cell with which we are making the comparison. Thus, the impact indices will be values between 0 and 1 that can be given by any type of function that has as its domain the set of impact degrees.

**B. EVOLUTION RULES**

Before we start defining our evolution rules, we will define the following notations:

- The state of the cell \( x \) at a time \( t \) will be denoted as \( \pi^t(x) \).
- The number of individuals with impact degree \( g \) and state \( K \) of a cell \( x \) at a time \( t \), will be represented as \( \sigma^t_{g,K}(x) \).
- To represent the number of individuals with an impact degree \( g \) we will use the symbol \( \Delta_g \).
- We will use the symbols \( \mathcal{S}^t, \mathcal{I}^t, \mathcal{R}^t \) and \( \mathcal{D}^t \) to denote the sets of susceptible, infected, recovered and dead cells respectively in the space \( \mathcal{L} \) at time \( t \). Formally

\[
\mathcal{S}^t = \{x \in \mathcal{L} : \pi^t(x) = S\},
\]

and analogously, we define the sets \( \mathcal{I}^t, \mathcal{R}^t \) and \( \mathcal{D}^t \). Note that

\[
\mathcal{S}^t \cup \mathcal{I}^t \cup \mathcal{R}^t \cup \mathcal{D}^t = \mathcal{L} \quad \text{for all time} \quad t.
\]

In the following, we will define the rules for the simplest versions of the SIS and SIR models, for this we have to take into account that both models share the passage from state \( S \) to state \( I \) so we will first define the evolution rule for an SI model as follows:

**Definition:** For a cell \( x \) in a space \( \mathcal{L} \) we define the SI rule as:

\[
\phi^t_{SI}(x) = \begin{cases} 
S & \text{if } \pi^t(x) = S, \\
I & \text{if } \pi^t(x) = I \text{ and } \rho > i(t), \\
\pi^t(x) & \text{otherwise},
\end{cases}
\]

with \( \rho \in \mathcal{U}[0,1] \), \( P(g) \) the impact rate of degree \( g \) and

\[
i(t) = \frac{\beta}{\alpha} \sum_g \sigma^t_{g,S}(x) \cdot \frac{\sigma^t_{g,I}(x)}{\Delta_g} \cdot P(g),
\]

where the factor \( \frac{\beta}{\alpha} \) indicates the affectation caused by the disease and the remaining factor corresponds to the behavior of the cells and its impact on the cell \( x \).

The above definition allows us to naturally define the evolution rules for SIS and SIR models in cellular automata:

**Definition.** Given a cell \( x \) in a set \( \mathcal{L} \) we define respectively the evolution rules for the SIS and SIR models as follows:

\[
\phi^t_{SIS}(x) = \begin{cases} 
\phi^t_{SI}(x) & \text{if } \pi^t(x) = S, \\
I & \text{if } \pi^t(x) = I \text{ and } \rho > \alpha, \\
S & \text{if } \pi^t(x) = I \text{ and } \rho \leq \alpha.
\end{cases}
\]

\[
\phi^t_{SIR}(x) = \begin{cases} 
\phi^t_{SI}(x) & \text{if } \pi^t(x) = S, \\
I & \text{if } \pi^t(x) = I \text{ and } \rho > \alpha, \\
R & \text{if } \pi^t(x) = I \text{ and } \rho \leq \alpha \text{ and } \rho \leq \alpha, \\
R & \text{if } \pi^t(x) = R,
\end{cases}
\]

where \( \rho \in \mathcal{U}[0,1] \).

For the models with birth and mortality, we will start from the hypothesis that each individual in a system has a probability of dying from non-disease causes that depends on its own age.

To implement the notion of age of a cell, we will initially have to modify the domain of our evolution rule. Specifically,
we will have to take into account the state of the central cell together with its age and the state of its neighbors, so that

\[
\text{Dom}(\phi) = \Sigma_x \times K \times \Sigma \times \cdots \times \Sigma,
\]

with \( K = \{1, 2, \ldots, 100\} \),

if we assume that the age of the cell \( x \) can range from 1 to 100 time units (weeks, months, years, etc.).

Another feature that we will implement will be that of cell aging, this with the objective of analyzing the impact of a disease on the individuals of a system at different stages of their "life". The way we will approach this idea will be with the following adjustment to the range of our rule

\[
\text{Ran}(\phi) = \Sigma_x \times K.
\]

Since we are working on populations of constant size, the way we will interpret the birth of a cell will be with the occupation of the space left by one that dies. For this we will identify the spaces left by the cells that die with state \( M \) and deaths \( D \), will identify the spaces left by the cells that die with state \( M \) and deaths \( D \) will identify the spaces left by the cells that die with state \( M \). We define the evolution rule with births and deaths for \( M \) and deaths for months, years, etc.). We define the evolution rule with births and deaths for months, years, etc.). We define the evolution rule with births and deaths for months, years, etc.).

\[
\mu_{M,T}(x) = \begin{cases} 
D, 0 & \text{if } t \equiv 0 \pmod{T}, \\
\pi^t(x) & \text{if } t \equiv 0 \pmod{T}, \\
S, 1 & \text{if } t \equiv 0 \pmod{T}, \\
\phi^t_M(x), E^t(x) & \text{if } t \equiv 0 \pmod{T}, \\
\phi^t_M(x), E^t(x) + 1 & \text{if } t \equiv 0 \pmod{T},
\end{cases}
\]

where \( \omega_k \) is the probability of dying from non-disease causes for ages in the \( k \)-th partition of the interval \([0, 100]\), \( b \) is the birth rate, \( \phi^t_M \) is the evolution rule of the epidemiological model, \( E^t(x) \) denotes the age of the cell \( x \) at time \( t \) and \( \rho \in U_{[0,1]} \).

**Observation**: The probability distribution with which \( \rho \) is chosen may not be uniform and instead depend on the phenomenon being modeled. For the purposes of this project, we will focus only on the uniform distribution and hope to further analyze the level of impact that the choice of different probability distributions may have in future research.

Finally, for models with death by disease we will partition over the interval \([0, 100]\) and define a probability \( \theta_k \) that indicates the probability of dying from the disease in the \( k \)-th partition, similar to what was done for models with birth and mortality. Thus,

**Definition.** Let \( x \) be a cell in a set \( L \), \( M \) be an epidemiological model (\( SIS, SIR, \) etc.) and \( T \) be a temporal unit (days, months, years, etc.). We define the evolution rule with death by disease for \( M \) as:

\[
\theta^t_{M,T}(x) = \begin{cases} 
D, 0 & \text{if } \pi^t(x) = I \text{ and } \rho \leq \theta_k, \\
\mu^t_{M,T}(x) & \text{otherwise.}
\end{cases}
\]

Where \( \theta_k \) is the probability of dying from the disease for individuals with an age in the \( k \)-th interval of the interval partition \([0, 100]\), \( \mu^t_{M,T} \) is the evolution rule for models with births and deaths; and \( \rho \in U_{[0,1]} \).

**IV. APPLIED EXAMPLE**

Suppose we discover the beginning of a flu outbreak in the hospital of a village with 49 people, and we want to analyze the behavior of the disease in order to establish control measures for future variations of the same flu.

In the village we can identify five different places where the cells interact with each other: the school, the offices, the market, the hospital and the homes of each person. We will assume that a census was taken in the population, and we will make use of the ages of each individual, their place of work (or where they spend most of their time) and the people they live with. They are distributed as follows:

- **School (E)**: It is known that there are 9 children and 2 teachers in the village.
- **Offices (O)**: It has a staff of 16 individuals.
- **Market (M)**: 8 workers were identified.
- **Hospital (H)**: Between doctors, nurses and patients a number of 14 individuals are identified.

We will assume that all children in the village are only children and that households with three individuals have at least one child. In addition to this and according to the above information, we will assume that in the village the population lives in one of three types of housing: the C1 housing where only one adult lives; the C2 housing where two individuals live, either two adults or one adult and one child; and finally the C3 housing, where 3 people live: two adults and one child.

![Fig. 1: Ages and occupations in the village.](image)

Now we must establish the degrees of impact according to the way in which our population is distributed. We will assume that people who live together have an impact degree of zero among them, that with those who have contact in the place of occupation they have an impact degree equal to one and that with the other cells they have an impact degree of two. Figure 2 describes the different neighborhood systems along with the degrees of impact generated for each cell according to the type of house it lives in (C1, C2 or C3) and the occupancy of it and its closest neighbors.
In the case of the disease, we will assume that it lasts an average of 5 days, so we will take $\alpha = \frac{1}{5} = 0.2$ and, on the other hand, we will take an infection rate $\beta = 0.3$. We will assume birth and death rates equal to 0.02 and 0.005 respectively and finally, we will assume from previous influenza outbreaks that the case fatality rates of the disease are:

| Age range | Rates |
|-----------|-------|
| 1 - 15    | 0.005 |
| 16 - 48   | 0.01  |
| 49 - 55   | 0.1   |
| 56+       | 0.25  |

**TABLE I**: Estimated flu case fatality rates in village.

We will analyze disease behavior over a 30-day period, taking different impact rates. This will give us an insight into the implications of close relationships with individuals outside the minimal neighborhood in the event of disease spread.

In the exercise shown in figure 3 it can be observed that the disease ends up disappearing in a period of 16 days, it should be noted that at some point the entire population contracted the disease and that the number of cases began to decrease from day 3. Figure 4 shows the number of individuals per state for a period of 30 days, taking as impact rates for grades 0, 1 and 2 the values $P(0) = 1$, $P(1) = 0.5$ and $P(2) = 0.25$.

If we now instead assume that within the village the relationships are a bit more distant due to e.g. measures such as alternation periods, we can assume that the impact rates are lower than in the previous exercise. In this scenario we will take $P(0) = 1$, $P(1) = 0.005$ and $P(2) = 0.0025$ and thus,

It is worth noting that despite the fact that the population is more distant, the disease has a similar behavior to that of the first scenario. This is due to the nature described by the parameters of the disease and its basic reproduction number $R_0 \approx 1.46 > 1$ (1). In the figure 6 we can observe the average evolution over 100 simulations of the disease taking into account different impact rates:

Considering that the previous example shows that different impact rates affect the evolution curves of the model, we can ask ourselves if something similar happens with different initial conditions. Figure 7 shows the evolution of the infected population in the village for different initial conditions, each one located in one of the work sites mentioned above.

So far we have been working around a similar scenario with a value of $R_0 > 1$ and in all the exercises we have been dealing with, it is evident that the disease reaches a peak (or maximum point of number of active infections) at a similar time. This is because despite considering the same parameters of the disease, we are modifying the properties of
the cell space itself, and this in turn allows us to state that the initial condition will not depend only on the parameters \( \alpha, \beta \) and the rates \( \mu_k, \theta_k \) as in the classical model, but instead the ways in which the cells may interact will have to be taken into account.

V. CASEMULATIONS

For the development of this work, it was necessary to design and implement a Python library that would serve as a tool to develop the analyses presented throughout this document. As a tool for the reader, different examples were developed in which the concepts defined in this document are applied. Additionally, a detailed documentation on each of the functions that are part of the library is provided.

VI. CONCLUSIONS

In this work, the dynamic nature of the models based on compartments that describe the behavior of the states that characterize a propagative phenomenon in an established set was understood. Likewise, the properties of cellular automata that allow describing large-scale spatial behaviors, from basic concepts of topology such as the fundamental systems of neighborhoods, the partial order relations that define the sets \( V(x) \) given any topology, among others. The understanding of these concepts allowed us to establish a relationship between the interactions of the individuals of some finite system with the fundamental systems of neighborhoods, and this in turn gave us a clear vision of the incidence of the nature of the interactions and the frequencies with which they occur in the evolution of a disease.

The conclusions that can be stated from this work are:

- By making use of the properties of cellular automata to describe spatial behaviors and of fundamental neighborhood systems, it is possible to model the social relationships of a given set of individuals.
- The initial conditions on how cells interact do not affect the equilibrium points of the curves describing the behavior of the disease affect the equilibrium points of the curves describing the behavior of the disease. However, as observed in the examples performed, changes in the initial condition can affect the rate of propagation of the same disease.
- There is evidence that limiting and/or reducing the intensity of social interactions is an effective measure to reduce the number of infected individuals.
- Although the rules proposed allow the analysis of characteristics that were not possible with the classical models, a limitation is evident in that a maximum capacity of individuals in the system is assumed.
- The methodology used for the design and implementation of the evolution rules described in this work provides a clear path for the definition of rules that model the behavior of more general epidemiological models.

REFERENCES

[1] A. Holkoa, M. Medrek, Z. Pastuszak, and K. Phusavat. Epidemiological modeling with a population density map-based cellular automata simulation system. **ELSEVIER**, 2016.
[2] J. A. Barmak. *Algebraic Topology of Finite Topological Spaces and Applications*. Springer, 2011.
[3] F. Brauer and C. Castillo-Chavez. *Mathematical Models in Population Biology and Epidemiology*. Springer, second edition, 2012.
[4] D. de Pereda Sebastián. Modelización matemática de la difusión de una epidemia de peste porcina entre granjas. Master’s thesis, Universidad Complutense de Madrid, 2010.
[5] J. Heesterbeek and K. Dietz. The concept of \( R_0 \) in epidemic theory. **Statistica Neerlandica**, 1996.
[6] R. J. Heffernan, J. M. and Smith and L. M. Wahl. Perspectives on the basic reproductive ratio. **The Royal Society**, 2005.
[7] A. G. Hoekstra, J. Kroe, and P. M. Sloot. *Simulating Complex Systems by Cellular Automata*. Springer, 2010.
[8] F. F. S. Miguel and R. M. Jaime. *Autómatas celulares y su aplicación en computación*. VIII Congreso Internacional de Ingeniería Mecánica y Mecatrónica, 2014.
[9] J. R. Munkres. *Topología*. Prentice Hall, 2002.
[10] F. C. Segura and A. T. Mascarúa. Elementos de topología general. *Topology and its Applications*, 2015.

[11] C. M. N. U. *Topología General*. Springer, 2017.