

A Parallel Proposal for SEIR model using Cellular Automata

Facundo Casares, Cristian Tissera, Fabiana Piccoli

LIDIC. Universidad Nacional de San Luis,
Ejército de los Andes 950 - 5700 - San Luis - Argentina
{ptissera, mpiccoli}@unsl.edu.ar

Abstract. Cellular Automata have been used with success in simulations of simple and complex systems belonging to different scientific areas, such as chemistry, biochemistry, economy, physics, etc.. In this work, we propose to use it in order to specify and implement a simulation model that allows to investigate behavioural dynamics for seasonal flu. This work presents a general solution where parallel programming techniques of shared memory are applied. Finally some experimental results about performance and flu behaviour are showed.

1 Introduction

Throughout times, the diffusion and spread of disease were a major concern of the human being. There were cases where a disease caused the disappearance of an entire population and important demographic changes, some of them were the plague (Europe, XIV century), yellow fever (Buenos Aires, XIX century) and cholera (Asia, XIX century). Today the situation continues, there are diseases monitored in certain regions such as malaria, and other are new as Influenza A or persistent as AIDS. For all this, it is important and priority the study and control of these diseases and their mode of transmission or contagion. One way to address the problem is analysing how the disease is distributed in a specific population. The study and analyses of complex real systems like this can be done through some simulation models.

When it comes to simulate discrete dynamical systems, cellular automata (CA) has been successfully used in simulation of diffusion process and, it is a valid alternative when we work with discrete dynamic systems which have complex behaviours from a simple set of rules. These rules allow to specify the new state of a component based on its state and its neighbourhood. In this way, it is possible to model complex dynamic systems from the specification of the local dynamics of each component. Besides, the state of each of them can be calculated simultaneously, i.e. in parallel. The behaviour of CA can display graphically the system evolution, allowing an easy comprehension of the studied dynamics. Besides, CA have demonstrated to be very useful simulation tools at the time of constructing artificial scenes, mainly in domains not suitable for other approaches. In other word, CA can be used to simplify complicated relationships

by means local interactions. For example, transmission of rumor, diseases or computer viruses can be reduced to local interactions among individuals/computers, as the case. In this work, we focus in spread of diseases, particularly the influenza or flu[1–3].

Different techniques have been developed to study the spread of diseases. This technique divide population into different types considering the characteristics of the disease: susceptible, exposed (with or without symptoms), infected, infectious, recovered, vaccinated, isolated, diagnosed, etc.. According to the disease and its infectious agent, an individual can be in some of above states. From these states and the dynamics governing their compartments, different mathematical models arise, some of them are: SI (Susceptible-Infected), SIS (Susceptible-Infected-Susceptible), SIR (Susceptible-Infected-Recovered), SEIR (Susceptible- Exposed-Infected-Recovered), and other variants (SIRS, SEIRS, SEIQR, etc.). Particularly in this work we focus on the SEIR model. It is the most suitable for epidemics whose infectious agent is a virus. Generally when infectious agents are viruses, the individuals recovered or cured can achieve a state of resistance for same virus [4, 5].

Taking into account the previous aspects and the parallel nature of CA, we present a parallel solution to simulate the flu propagation using CA and SIRS models. This solution will allow to study, in short time, the spread of influenza or any virus mutation in different environments, considering type of population, its distribution and other characteristics; and to take decisions in consequence, such as vaccination campaigns, isolation, quarantine, etc.

The paper is organized as follows: the next sections describe all the previous concepts. Sections 3, 4 and 5 sketch the characteristics of seasonal flu, our parallel proposal, and its empirical performance. Finally, the conclusions and future works are exposed.

2 Previous Concepts

In this section, we explain the main concepts to develop this work.

2.1 Cellular Automata

In their research about the machines with auto-replication capabilities, John Von Neumann and Stanislaw Ulam were the first in formulate the Cellular Automata(CA). But was in 1970 when these systems received special attention, Jhon H. Conway proposed the *game of life*, the most known CA. Since that date, the CAs have grown in popularity within of the scientific community and actually they are considered to solve problems with different nature[6].

A CA is a mathematical system with discrete values in space, time and state. It has different characteristics, some of them are *auto-replication*, universal computation capabilities and *auto-organisation* effects. This last property plays a very important role at the time of explaining certain kind of behaviours observed in physical and biological phenomena [7, 8], in consequence, the CA have

been used, for example, to simulate different phenomena as chemical reactions, diffusion processes, hydrodynamic, mechanic, filtration, chaos theory and others.

A CA is a discrete dynamic systems with capacity to develop complex behaviours from a simple set of rules. It represents a grid of locally connected finite automata, each of them produces an output from several inputs, next state is a result to apply a transition function. The upcoming state of a CA cell depends of its own current state and the states of its neighbouring cells [9].

Intuitively, we consider a CA as a system composed by a array of cells A . Each cell c_i in A represents a finite automaton with a set of states Q , an input alphabet σ and a transition function $\delta : Q \times \sigma \rightarrow Q$. The input alphabet σ is given by all the possible combination of the cell states of the adjacent (neighbouring) cells. If we denote N_{c_i} to the set of cells that we consider as neighbour of cell c_i , and $|N^{c_i}| = n$ is the number of adjacent cells, then the input alphabet is $\sigma \equiv Q_n$. Usually, a cell c_i and its adjacent are considered and represented as a unique set $N = \{c_i\} \cup N^{c_i}$. N is referenced as the *neighbourhood*.

By all above exposed, a CA is defined as a 4-tuple $M = \langle A, Q, \delta, N \rangle$ where:

- A is a D -dimensional array, and each component (cell of the array) has associated a *finite automaton*.
- Q is a finite set of states (of the automaton) of a cell.
- N is the specification of which cells are included in a neighbourhood, $N \equiv \{c_i\} \cup N^{c_i}$ such that N^{c_i} are the adjacent cells to c_i .
- Let $\sigma \equiv Q^n$ where $n = |N^{c_i}|$ is the number of adjacent cells to c_i . The *transition function* of states, $\delta : Q \times \sigma \rightarrow Q$, is a mapping such that if $q_i \in Q$ is the state of the cell c_i in the time t and $q_{i+1}, q_{i+2}, \dots, q_{i+n} \in \sigma$ are the states of the adjacent cells to c_i , and

$$\delta(q_i, q_{i+1}, q_{i+2}, \dots, q_{i+n}) = q'_i$$

where $q_i, q_{i+1}, q_{i+2}, \dots, q_{i+n}$ are the states of the central cell and its neighbours at the time t and q'_i is the state of the central cell at the time $t + 1$.

In some cases, it is possible to specify probabilistic transition rules, where an arbitrary probability p can be associated to a transition rule. The semantic of this kind of rules establishes that, always that a cell matches the configuration of the specified neighbourhood in a probabilistic rule, the cell will have at time $t + 1$ the new state specified in such rule with probability p .

2.2 SIR Model and Derivated

In the modelling of diseases, several considerations must be taken into account, among them are important to consider: the infectious agents (they are responsible for transmitting the disease and condition the states through which passes an individual affected) and transmission modes (they can be person-to-person, by the environment, by some vectors such as insects or agents, or among animals of the same or different species).

Because all the numerous factors involved in a disease, it is impossible to study them of same way. A start point is to classify states in that an diseased individual can be. A set of possible states is:

- *S*: Healthy individuals and *Susceptible* to be infected.
- *E*: *Exposed* individual to disease, infected but not infect others (i.e., the disease is latent).
- *I*: *Infected* individuals who infect others.
- *R*: *Resistant* individuals to diseases (normally, it happens after that a person recovers from illness or vaccinates).

In a same time, an individual can be in a single stage of the disease, therefore for a population of N persons, if we consider the above set of state, the following equation must be satisfied:

$$S + E + I + R = N$$

Kermack and McKendrick in 1927a formulated a simple model, SIR model, which consists of three stages: *Susceptible*, *Infected* and *Recovered*. The SIR model is easily written using ordinary differential equations (ODEs), this implies a deterministic continuous model. It assumes encounters among infected (I) and susceptible (S) individuals at a rate proportional to their respective numbers in the population.

The analytical techniques are good to address problems in a basic way. But, in the case of disease epidemic study, the system is complex, and, in consequence, more realistic solutions with high level of detail are necessary.

2.3 Multi-threaded Parallel Programming

Multi-threading and parallel programming are concept different, in this work we apply them together to obtain good performance in a CA solution.

The multi-thread programming imply a single process and this process generates many threads. All of them share the same space memory, and they can be able to execute independently and at the same time: in parallel. The traditional multi-threading was used to do time-slicing or take advantage of the CPU idle time, i.e. while one of the threads waits, another thread could execute.

By its side, parallel programming allows explicitly breaks the task down into smallest units, where each unit can be executed in parallel on a single CPU core. When it is possible divide the task and its sub-tasks share the same memory space and run in parallel, the problem can be solved applying Multi-threaded Parallel Programming.

To solve a problem by applying Multi-threaded Parallel Programming, we can use OpenMP(Open Multi-Processing)[10]. It is an API that supports multi-platform shared memory multiprocessing programming, it achieves parallelism via multi-threading and shared-memory.

In this work, we present a simulation system for flu transmission using CA and SEIR as models and multi-threaded parallel programming techniques.

3 Seasonal Flu

Influenza is a viral infectious disease that affects, primarily, the respiratory tract of humans. Usually it accompanied by other symptoms such as sore throat, weakness, dry cough, fever, and muscle aches, of stomach and head . In some cases, it may be complicated and derive in pneumonia becoming fatal. This can occur in certain age groups, such as young children and elderlies. There are three types of seasonal influenza: A, B and C. The influenza virus A and B are the most common, they are classified into subtypes according to the combination of two proteins in virus surface (H and N)[11, 12].

Virus transmission is done person-to-person, mainly through particles ejected when a sick person coughs, sneezes or talks. Also, it can be transmitted by means blood or contact with surfaces or objects contaminated. Besides, flu virus is resistant in a dry and cold environment, this property allows its rapid spread mainly in autumn and winter, seasons when it becomes seasonal epidemic. The virus can keep its infections level by about one week at body temperature, however, there are patients that require 15 days of recovery. Most people recover without medical treatment. Antibiotics are only useful if there is a bacterial infection.

A infected person with the flu virus goes through an incubation period (approximately from two to four days). The contagious period begins one day before that person has symptoms (this is a serious problem, a person could be spreading the influenza without knowing who is sick). After a week, the transmission power is reduced, even it disappears. The figure 1 summarizes how the disease evolves in a person, from he/she is susceptible until his/her recovering or, in the worst case, death.

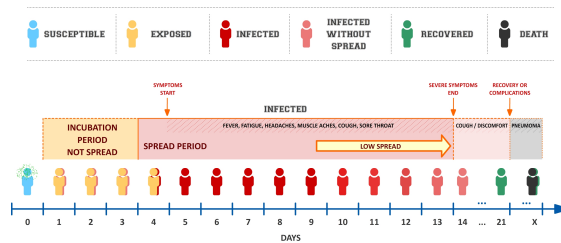


Fig. 1. Influenza progression in people

The most effective way to prevent the flu and its serious consequences is vaccination. In healthy adults, it can provide reasonable protection, while in elderlies can reduce its severity, the incidence of complications and deaths. As *C* influenza cases are much less frequent, generally the vaccines try influenza *A* and *B*. They are usually trivalent, they contain purified and inactivated proteins of the three strains most common in the following epidemic: two subtypes of influenza *A* and one *B*. The vaccination effectiveness depends on the match between the vaccine virus and surrounding virus. Moreover, a vaccine made one

year may not be effective to the next by two reasons: the virus change and mutate rapidly, and the strains have variable dominance[12].

4 Parallel Simulation of Seasonal Flu Epidemic

The epidemiology studies the factors of potentially harmful infectious agents that affect a particular population, and tries to explain and predict how the disease evolves in the time. As explained earlier, the simulation models that use CA concepts are ideal to represent this type of real systems. Through setting the main features of the problem, CA recreates a virtual world that comes alive and gives an approximation of what would happen in the real world.

In this section, we describe a model based on epidemiological model with SEIR approach and CA. By means this model, it is possible to analyse the effects of seasonal influenza in a population with a certain territorial distribution.

4.1 SEIR-CA Model

The *SEIR – CA* model has a cellular space defined by a finite two-dimensional lattice. Each cell of the automata is a place busy by only one person. In this work, the cellular space (2-D array) represents a social space in which the individuals can interact, this means, two adjacent cells occupied represent two people in touch. To perform the simulation, it is necessary to establish the following considerations:

- **Neighborhood:** For this systems, we consider the known Moore neighborhood: eight cells surrounding the central cell define the neighborhood . With this, every individual interacts with at most 8 people by once.
- **Cell State:** a cell is in one state of $Q = \{B, F, S, E, I, W, R, D\}$ where:
 - *B*: Automata limits (when borders are absorbent).
 - *F*: Free Cell (There is not any person and it can be selected to occupy).
 - *S*: The person is susceptible to contract flu.
 - *E*: When the person is in incubation period but not spread.
 - *I*: The person is infected and, he/she spreads the flu
 - *W*: The person is infected but does not spread (Infected without spread).
 - *R*: When the person is recovered.
 - *D*: Dead person. Generally, the deaths can be by any complication.

The transition through each of the states is shown in Figure 1.

- **Initial Configuration:** Before simulation begins, it is necessary to set relevant information such as size of population surface (CA Size), how the population is distributed in the surface, which is the infection percentage of population and their ages.
- **Virtual Clock:** Time is discrete, at the beginning of simulation, an interval of time is set. During this interval, a person can move to any of the neighboring cells and relate with other people. Generally, this movement follows some probabilistic pattern.

– **Model Evolution Rules:** There are three kinds of rules which are: the rules related to CA, with spreading of diseases, and those associated with persons movement inside CA. Each one of rules are:

1. **CA Rules:**

- If a cell is in state B (CA with absorbent borders), its state does not change at no time of the simulation.

2. **Diffusion and Spread Rules:** A cell occupied by a person in time t , also will be occupied in time $t+1$. It can change its state according to:

- If at time t , the central cell is susceptible (S) and, some of its adjacent cells are infected (I), the central cell will be incubating the influenza (E) at time $t+1$, but will not spread. The probability of state change is proportional to the number of adjacent cells I ($S \rightarrow E$).
- If the cell is incubating flu (E) and, the time t is the end of the asymptomatic period, at time $t+1$ the cell will be in infected state I ($E \rightarrow I$).
- After 8 days, a cell infected (I) in time t will pass to state infected without spread (W) at time $t+1$ ($I \rightarrow W$).
- In time t , a cell infected (I or W) could become recovered (R) or dead (D) state (time $t+1$). The selection between two stages depends of a probability function ($I \rightarrow \{R, D\} | W \rightarrow \{R, D\}$).

3. **Rules of Person Motion:** A person can move to a neighbour cell if it is free (state F). All free neighbour cells have the same probability to be occupied.

An important aspect to note is that the matrix does not necessarily represent a spatial universe. The model represents the people interrelation, a movement in the surface is an abstraction, it can mean that a person moves to speak with other (for example an officemate) or, he/she goes to a business and relates with a vendor. When a person comes in a neighbourhood of another, this means an interaction between two people. As we assume an uniform distribution of probabilities (a person can move equally to any free adjacent cells), this model simplifies the problem.

The movement is not physical, it models the interaction between a person with its neighbourhood.

Once defined all CA characteristics, in next section we explain the main issues of our proposal: a parallel SEIR-CA using shared memory.

4.2 Parallel Solution

To solve the *SEIR-Flu* in parallel, we consider the shared memory paradigm, in consequence it is necessary to consider synchronization mechanisms to access to memory or, other programming techniques. For the implementation, some characteristics are:

- OpenMP is selected as application programming interface (API).
- In each timestep, two CA are necessary, one represents the state in time t (input CA), and the other is the output (CA in time $t+1$).

- Each thread takes a particular cell and works over it. When it finishes, takes the next cell and starts again.

The solution is structured in three stages, each stage does:

- *First Stage*: In this stage, we calculate for each cell: the next state and a list of movement intentions. For that we apply the CA rules. When this stage finishes, the new CA for time $t + 1$ is obtained and each free cell has a intention list to be occupied by one of its neighbour.
- *Second Stage*: Taking into account the intentions list of each free cell, one of candidates is selected in random form. To improve the solution and save memory, we use a bitmap technique for intention lists. Each list is represented by 9-bits to the Moore neighbourhood plus the central cell (fifth bit). If some position is set to 1, this means that central cell wants to move there.
- *Third Stage*: After the second stage, the data structures of CA are inconsistent state. This stage carries to a safe state, all cells and their positions are adjusted. In time $t + 1$, the input and output CA exchange their roles.

Every stage are made in parallel, independent and sequential sections without using synchronization mechanisms.

Other characteristic is which is the breakpoint of simulation. It can end by two reasons, whichever first occurs, they are:

- Maximum simulation time: A time limit for simulation is established, if this is reached, the simulation ends.
- Some stable or starvation state: The simulation is based on spread from person to person, once the last sick person is cured or dies, none new flu infection in the current population can be developed. For example, when population is small, the rate of spread is not enough for the disease persists over 4 months, usually a stable state is reached and the simulation ends before 120 days.

The next section some experimental results of this parallel *SEIR-Flu* are displayed.

5 Experimental Results

In this section, we show and analyse the experimental results for the parallel solution of *SEIR-Flu*.

The environment of simulations, sequential and parallel, were in a multi-core computer whose characteristics are: 2 processors AMD Opteron 6272, 2.1GHz, 16 cores per processor, RAM memory of 64GB Memory (16x4GB), 1333MHz, OpenMP 4.0 and Debian 8.

In order to obtain the results, different scenarios of simulation are considered. Each of them is a combination of next parameters:

- The maximum time of simulation: 120 days. The stationary flu has more propagation in winter, in others seasons its infection decreases drastically.

- Virtual clock: A timestep is equivalent to 1 hour.
- Initial infection factor: 1% of population has flu.
- Population: 50% of cells of grid.
- Size of Square Grid: 50, 100, 200, 400, 800, 1600 and 3200 by side. For lack of space, we only report the more representative results .

We consider a stationary type of population pyramid [13], each individual has an age between 1 to 90 years. The degree susceptibility to infection and the mortality rate is determined according of individual age[14, 12]. In all populations, there are three groups of person: *Children* (Up to 6 years old), *Young-Adult* (7 to 60 years) and *Elderlies* (Greater than 61 years), each of them has a susceptibility equal to 35%, 20% and 35% respectively.

For mortality, we recognize the following groups and each mortality rate: *Under 3 years* (8%), *4 to 10 years* (5%), *11 to 18 years* (2%), *19 to 50 years* (0.5%), *51 to 60 years* (2%), and *more than 60 years* (8%).

Each reported value is the average of 10 executions of *Parallel SEIR-Flu*. For parallel solution and each scenarios, we use 2, 4, 8, 16, 32 and 64 threads. To sequential solution, we consider the same solution but for only one processor.

In first time, we evaluate if our parallel proposal works well and its performance is better than sequential solution. The figure 2(a) shows the speedup reached for every simulation scenario and different numbers of threads. In majority of scenarios, we achieve speedup. Although this is not close to optimal, we can reduce significantly the computational time of simulation. Similar behaviours and earnings are observed in Figure 2(b) for the same size of population (50%) but the 10% of them are infected initially.

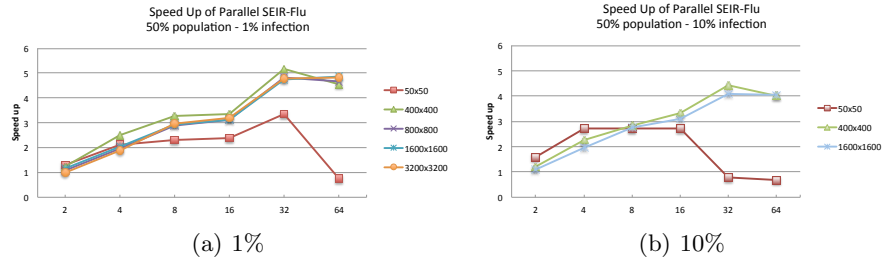


Fig. 2. Speedup of Parallel SEIR-Flu for 50% population with 1% and 10% initial infection

Besides of performance results, we can observe some behaviour of the flu infection process. For example, a higher level of initial infection implies less time to reach a state free of flu in the population of each grid, the Figure 3 shows this situation.

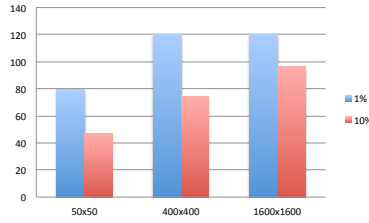
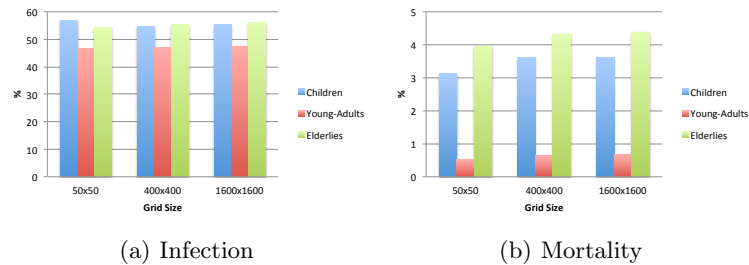


Fig. 3. Days of SEIR-Flu for 1% and 10% initial infection

In Figure 4, we sketch the percentage of infection(4(a)) and mortality (4(b)) for each populations group: Children, Young-Adults and Elderlies. The grid occupation is 50% and the 10% are infected.



(a) Infection

(b) Mortality

Fig. 4. Percentage of Mortality and Infected Individuals for 50% population and 10% initial infection

Figure 5 summarizes the speedups achieved for different levels of grid occupation when 1% of initial population is infected (Figures 5(a), 5(b) and 5(c)). Besides, Figure 5(d) displays how many days are necessary to reach a state free of flu.

6 Conclusions and Future Work

We presented a parallel model capable to simulate the propagation of seasonal flu. The proposed model uses the CA and SEIR concepts to analyse the effects of seasonal influenza in a population with a certain territorial distribution modelling the people interaction.

One of objectives of this work is to develop and select techniques of high performance computing in order to perform and execute large-scale complex simulations of diffusion processes. While our proposal is not yet complete, the results of our multi threading implementation are encouraging and give us an

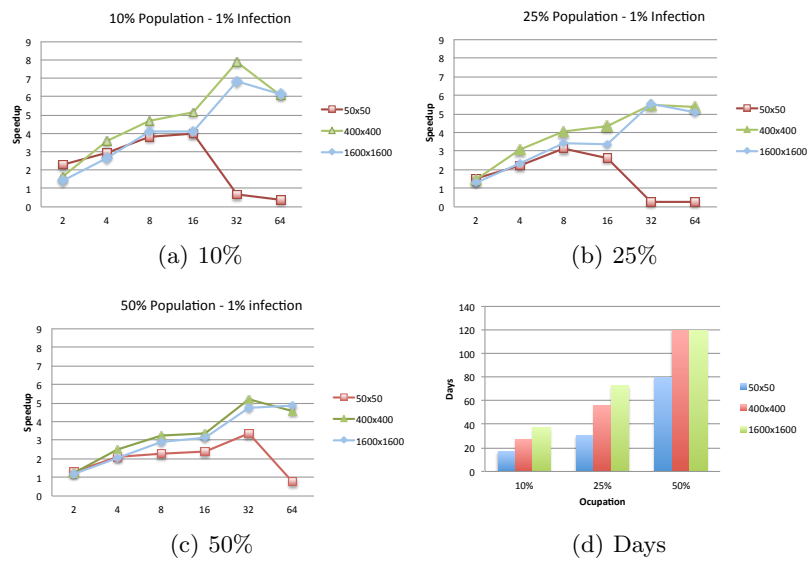


Fig. 5. Speedup and Days of SEIR-Flu for different levels of occupation with 1% initial infection for and

initial framework for future works in the characterization of diffusion processes like signal propagation, avalanches, rumours spread, etc.

As future works, it is important to decide the optimal size of set of threads needed to solve the model taking into account the performance parameter to try characterize this kind of applications and this experience will be transferred to others developments of more complex parallel applications.

References

1. F. Chierichetti and A. Panconesi. Rumor spreading in social networks. *Theoretical Computer Science*, 412(24):2602–2610, 2011.
2. M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili. Theory of rumour spreading in complex social networks. *Physica A: Statistical Mechanics and its Applications*, 374(1):457–470, 2007.
3. S. White, A. del Rey, and G. Sánchez. Modeling epidemics using cellular automata. *Applied Mathematics and Computation*, 186(1):193–202, 2007.
4. A. Huppert and G. Katriel. Mathematical modelling and prediction in infectious disease epidemiology. *Clinical Microbiology and Infection*, 19(11):999–1005, 2013.
5. T. Johnson and B. McQuarrie. Mathematical modeling of diseases: Susceptible-infected-recovered (sir) model. In *University of Minnesota, Morris, Math 4901 Senior Seminar*, 2009.
6. D. Reyes. Descripción y aplicaciones de los autómatas celulares, 2011. Technical Report, Autónoma de Puebla.

7. S. Kauffman. Emergent properties in random complex automata. *Physica D: Nonlinear Phenomena*, 10(1):145–156, 1984.
8. S. Wolfram. Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, 10(1):1–35, 1984.
9. J. Klüver and C. Klüver. *On communication. An interdisciplinary and mathematical approach*, volume 40.
10. B. Chapman, G. Jost, and R. van der Pas. *Using OpenMP: Portable Shared Memory Parallel Programming*. Number v. 10 in Scientific Computation Series. MIT Press, 2008.
11. C. Beauchemin, J. Samuel, and J. Tuszynski. A simple cellular automaton model for influenza a viral infections. *Journal of theoretical biology*, 232(2):223–234, 2005.
12. World Health Organization. Influenza (seasonal), 2014. Fact sheet N211.
13. P. Hernández. *Demografía y antropología demográfica*. CONACULTA-INAH, 2004.
14. Argentina Ministerio de Salud. *Manual para el fortalecimiento de la vigilancia de la enfermedad tipo Influenza utilizando la estrategia de Unidades Centinela*. 2011.