SIR model
Analysis and extensions

Hristo Vrigazov
Iliya Zhechev
Viktor Bozhilov

June 2016

Contents

Chapter 1

Introduction

1.1 First section

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Citation of Einstein paper [**?**].

1.2 Second section

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.2.1 First subsection

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.2.2 Second subsection

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

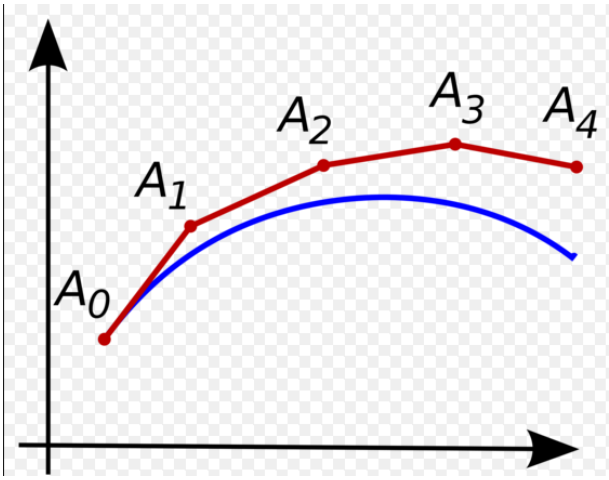
Chapter 2

Euler Method

2.1 Introduction

In this chapter we will introduce the method used for solving the differential equations used in the SIR model. We will start by familiarizing you with the Euler's method and then explaining in detail how we used in to compute the epidemic spread. Euler's method is one of the simplest numerical procedures used for solving ordinary differential equations given an initial value. It is also first order method, which means that the global error is proportional to the step size. The method is named after Leonhard Euler and is often used in the construction of more complex and accurate procedures.

Figure 2.1: Example of the approximation by using Euler's method



2.2 About the method

The basic idea is that knowing an initial value you can approximate the curve of the function satisfying the equation. You start by calculating the gradient with the initial value and making a step dt using the slope. This should approximate the function, since the gradient is the direction of change of the function. This gives you a second, approximated point that is on the curve of the solution. You can now using this new point calculate the next. Doing this iteratively you accumulate set of points approximating the curve you are looking for. The main problem of the method is the level of approximation. Figure 2.1 shows how the

error is accumulated as the step size and the number of iterations are increasing. The smaller the step you are using the greater the accuracy but as the steps are getting smaller the computing power you need to invest to calculate the same interval of the curve becomes greater. There is an inherit tradeoff between the level of approximation and the number of iterations you need to run, so you need to find a balance there.

2.3 Illustrating the Euler method with an example

Let's say we have a differential equation

$$dF = 2dFdt$$

and we have an initial point

$$F(0) = A_0$$

We can now find an approximate second point by calculating

$$A_1 = A_0 + 2A_0\epsilon$$

, where ϵ is the level of approximation. After computing A_1 we can calculate $A_2 = A_1 + 2A_1\epsilon$. Now we see the pattern and generalize the idea with the following recursive relationship which can then be plotted (see Figure 2).

$$A_n = A_{n-1} + 2A_{n-1}\epsilon$$

$$A_0 = 1$$

Chapter 3

The modeling software

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Chapter 4

Cellular automaton

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

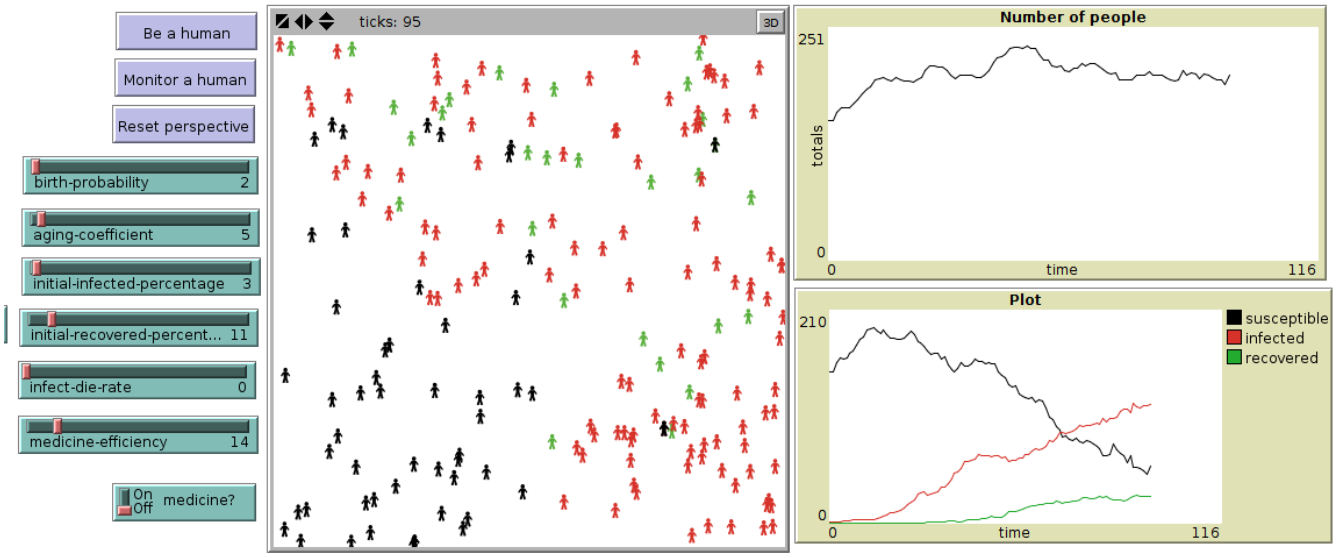
Chapter 5

Agent-based SIR model

5.1 Introduction

Computational models are becoming popular for modeling real-world processes and events. They are often defined by simple rules, but that does not mean that their behavior is simple [1]. Their advantage over mathematical equations is that they are more flexible, and are easier to use. In this chapter, we will explore an extension of the SIR model, and we will represent the three compartments and the transitions between them without mathematical equations, but with a computational, agent-based model. The model uses a number of autonomous people on a limited territory. A person can infect other person, an infected person can recover or die. We will observe how the different parameters of the model impact the number of people that are alive.

Figure 5.1: Parameters - on the left, simulation in the middle and on the right - the total number of people and the number of people in each compartment



5.2 The model

5.2.1 Description

We are observing three compartments as in classical SIR model: susceptible, infected, removed. The territory of interest in represented by a grid, which consists of squares. Each square can be:

1. Empty
2. With a susceptible person in it
3. With an infected person in it
4. With a recovered person in it

Every person has energy as property, and when the energy of a person is 0, he dies and is removed from the simulation. The time is represented a discrete list of moments, each moment is called a "tick". On each tick, a person loses some amount of energy and moves in a random direction. During the first tick, a number of people are initialized randomly on the grid. When a person is created, he can be either susceptible, infected or recovered, with a probability for each category. The disease is characterized by die-rate, which is the probability that an infected person will die. On each tick, a new set of people are born (one can control the birth rate). When infected, a person can heal himself with probability 0.01, otherwise to heal himself, a medicine should be applied. The medicine has efficiency. The medicine is applied globally - for all people. When medicine is applied, on each tick a person can go from infected to recovered. The probability for that event happening can be controlled using the medicine efficiency.

5.2.2 Variables

The variables of the model could be controlled from the user interface. All of them are a number between 0 and 100, (except for the medicine? which is a binary variable), since they all represent a certain probability. Here is the full list of variables of the model, with description on what each one does:

1. birth-probability - this is the variable that controls the probability that new people would be born on each tick. Setting all other parameters to zero and the birth-probability to a positive number makes the population grow almost exponentially, so there is a bound for too many people, after which.
2. aging-coefficient - this variable controls how likely is it for a person to lose energy. On each tick, every person loses energy amount between 0 and the aging-coefficient.
3. initial-infected-percentage - this variable controls (indirectly) the percentage of the initial number of infected people. It is used only on the first tick. When a person is created during the first tick, he can either remain susceptible or become infected with a probability equal to initial-infected-percentage.
4. initial-recovered-percentage - this variable controls (indirectly) the percentage of the initial number of recovered people. Those are the people who have immunity to the disease. It is used only on the first tick. When a person is created during the first tick, he can either remain susceptible or become susceptible with a probability equal to initial-recovered-percentage.

5. infect-die-rate - this variable controls how likely is a person to die (go from infected state to die). On each tick, every person who is infected, can die with a probability equal to the infect-die-rate.

6. medicine-efficiency - this variable controls how likely is a person to be healed when a medicine is applied. On each tick, if a person is infected and medicine is applied, he can heal (from infected state to recovered) with a probability equal to medicine-efficiency

7. medicine? - this variable is boolean and indicates whether a medicine is applied (a medicine is either applied for everyone, or for no one).

These are the variables that control the rules which the people on the grid follow. There are also other user interface elements, but they do not impact the model and are briefly mentioned in the conclusions section.

5.2.3 Goals

We want to observe how the number of people in the population changes in time, using different model parameters. The other variables we want to observe mainly by comparing their values in time are the number of people in every category. We will try to answe the following questions:

1. When we exclude the medicine, the birth and the death, what results does our model produce compared to the classical SIR model. What are the similarities and the differences?

2. How does our model perform with data from famous pandemics, like the Spanish flu from 1918?

3. What does the model predict for the spreading of Ebola in Liberia?

5.3 Observations

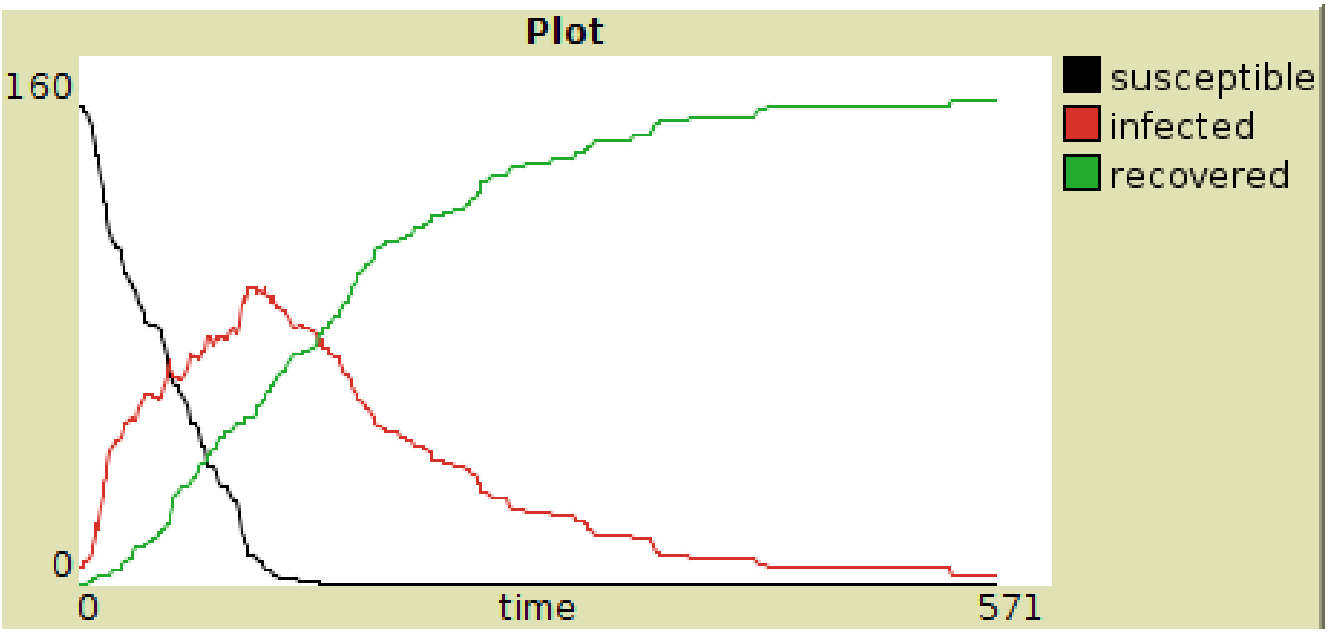
5.3.1 SIR model

Let us set the parameters birth-probability, aging-coefficient, initial-recovered-percentage and medicine-efficiency to 0. They are not taken into account in the classical SIR model. Let us set the initial-infected-percentage to a random number different from 0, say 3. We will also play around 20 simulations in each experiment, to see whether what we see in an instance of a simulation is a pattern. In all simulations, we notice the same pattern (see figure 5.2): in the beginning, almost all people are susceptible, but with time the number infected rises until the number of infected hits its peak. This is when we have pandemics. Then the number of infected starts decreasing, and the number of recovered start increasing, until almost all are recovered. In all simulations, we notice that same pattern, which is similar to the SIR model with differential equations, described earlier, which is quite interesting. Of course, they are not exactly the same, the agent-based model is more oscillating, but they follow the same pattern. This is a sanity check for our model, which he passes.

5.3.2 Spanish flu

In 1918, the humanity experienced one of the deadliest diseases in human history, the spanish flue [5]. Using historical data, we will compute the parameters for our agent-based model, run the simulation and

Figure 5.2: Classical SIR model using agent-based simulation



then compare the people that actually died to these that died in the simulation. We will also compare the number of recovered people.

We will ignore the birth rate and the aging coefficient, since we were not able to find information in Wolfram Alpha or any other internet source about the birth rate and the mortality rate. Here is how we computed the other parameters of the model:

1. initial-infected-range - there were 1.84 billion people on Earth in 1918 [7], and the flue infected 500 million people [5]. So the percentage of infected people was:

$$\frac{5.10^8}{1,84 * 10^9} \approx 0,27$$

   Since the number of infected people is 27%, we will set initial-infected-range to 27.

2. infect-die-rate - the die rate of the Spanish flu was estimated to be between 10% and 20% [5]. We will do 5 experiments for each value between 10 and 20 of the infect-die-rate variable.

3. medicine-efficiency - the people tried different vaccines and ways to heal, and between 400 million and 450 million survived [5]. For the purposes of our model, we will think of all ways to heal the disease (medicines, vaccines, natural healing) as one global medicine. Since there were 500 million infected, and between 400 million and 450 million survived, the efficiency of the medicine is between 80% and 90%, so we will also have 5 experiments for each different value in this range.

The results we want to observe is what is the number of people, who were infected and died, and then the number of people who were infected, but then recovered. Let's see what do we expect for each of those:

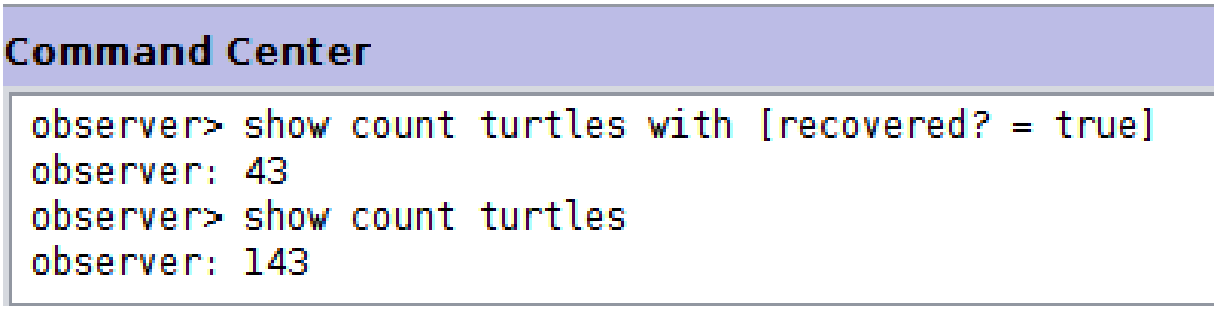
1. Percentage of died people - the number of people died during the 1918 flu pandemics were between 3% and 5% of the whole population [5].

2. Percentage of recovered people - the number of people who got infected and then healed is between 21% and 24% of the whole population, calculated based on the data from Wikipedia [5].

Those are the result we expect. Let us now perform a number of experiments. The total number of people on the grid is 150. The first results of the experiments are given in the following table:

Initial-infected	Infect-die	Medicine-efficiency	Died	Recovered
27	10	80	5	43
27	10	80	1	65
27	10	80	6	56
27	11	80	5	45
27	11	80	6	50
27	11	80	6	56
27	12	80	10	48
27	12	80	8	44
27	12	80	7	44
27	13	80	11	55
27	13	80	6	39
27	13	80	7	37

Figure 5.3: A sample result, obtained from the command center. For this experiment, the number of recovered is 43, and the number of people left is 143, which means that 150-143=7 people died



The table has 1200 records and was created using automation, so we will not show the whole table here. The average value for "Died" column is 8.6, and for the "Recovered" column is 44.7. Let us compare the results from the model with the expected results:

1. Died - in reality, 3-5% of the population died. In our simulations, on average the percentage of people died is 5.7%, which is close.

2. Recovered - in reality, 21-24% of the population recovered. In our simulations on average the percentage of people who were infected and then recovered is 29.8%, which is reasonably close, since we made a lot of roundings and strong assumptions.

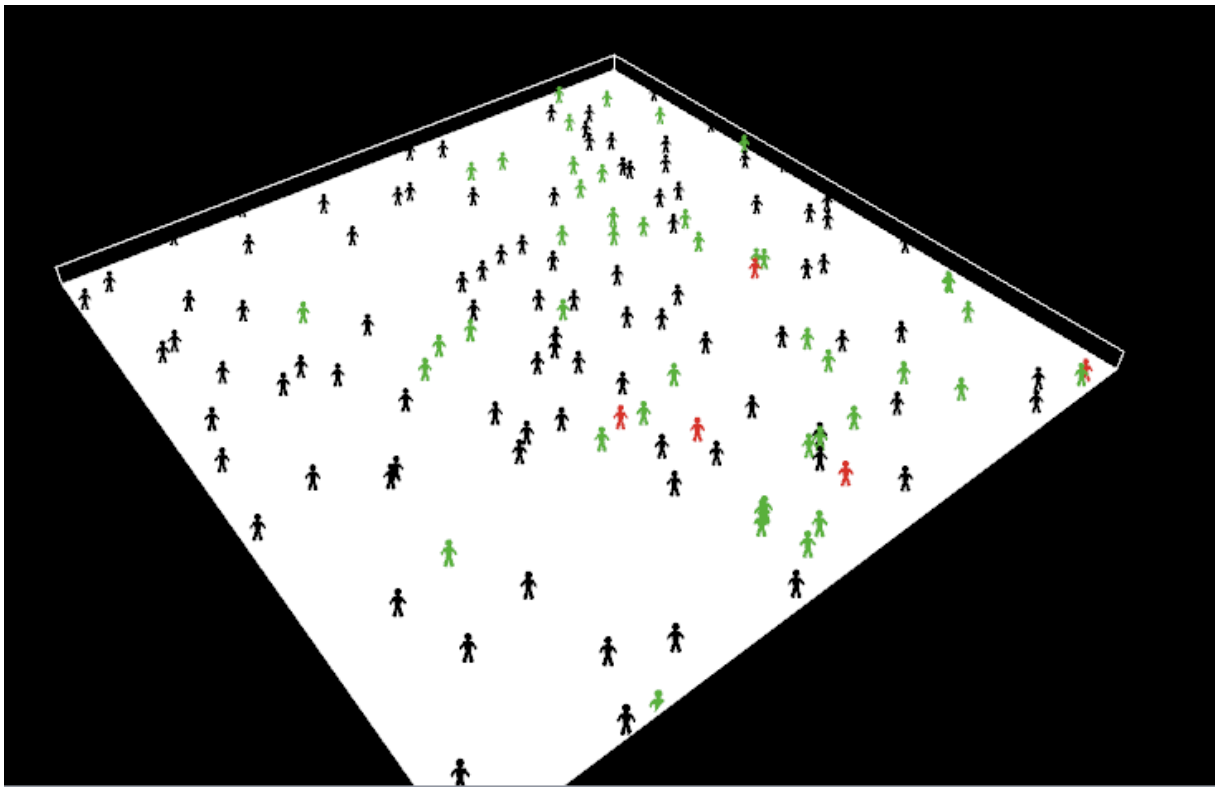
5.3.3 Ebola in Liberia

to be done

5.4 Conclusions

Based on what we observed in the experiments, the agent-based model created here is fully capable of modeling real world spreading of diseases. We showed that it behaves is a similar way to a classical SIR

Figure 5.4: 3D view of a simulation



model when we did not use its extensions, and we observed how he can predict reasonably well how many people would die and how many would recover in pandemics like the famous Spanish flu.

5.5 Implementation

The model was implemented in Netlogo [2], IDE and domain specific language for agent-based modeling. The source code is available online [3]. The model itself is also available online for experimentation [4].

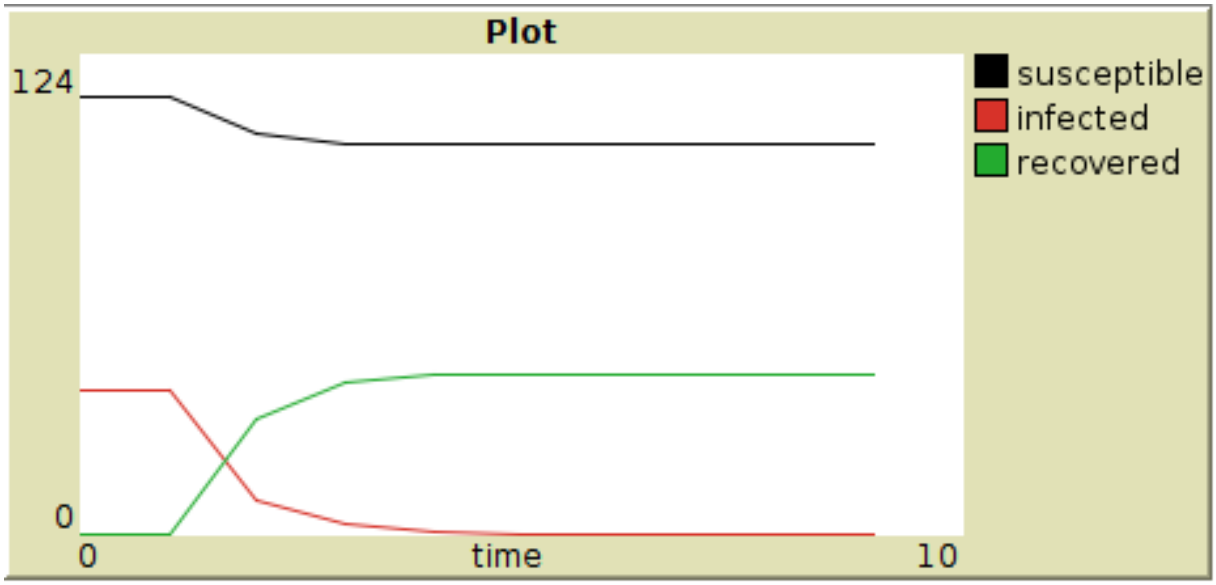
5.6 User Interface notes

TO DO: write cool stuff about the UI

5.7 Possible extensions

TO DO: let your imagination go wild

Figure 5.5: Susceptible, infected, recovered during one experiment. Note that the moment in the spreading of the disease we chose for initial is after the peak, so the graphics show similar graphs to the end of the classical SIR model



Bibliography

[1] Stephen Wolfram *A New Kind of Science*. Wolfram Media, 2002.

[2] Netlogo
`https://ccl.northwestern.edu/netlogo`

[3] SIR model code repository
`https://github.com/ichko/SirModel`

[4] Netlogo model code repository
`https://github.com/ichko/SirModel`

[5] 1918 flu pandemic - Wikipedia
`https://en.wikipedia.org/wiki/1918_flu_pandemic`

[6] Ebola virus in Liberia - Wikipedia
`https://en.wikipedia.org/wiki/Ebola_virus_epidemic_in_Liberia`

[7] Wolfram Alpha
`https://www.wolframalpha.com/`