Technische Universiteit Delft

Modelleren 2A

Segregation

C. Barendrecht
G. Jin
S. Moerman
N. Snijder

Under supervision of:
Dr. P. Cirillo

This page is left intentionally blank.

Abstract

An extended version of the 'Shelling Tipping Model' is built...
This report is written by Casper Barendrecht, Guanyu Jin, Stijn Moerman and Nand Snijder. We are second year theoretical math students at Leiden University. We studied segregation of different groups of people in any discrete area, such as a city, corridor or even a flat. This study is strongly based on the 'Shelling Tipping Model' of 1978, which we later extended to include height. In this study, we are supervised by Dr P.Cirillo, and we'd like to thank him for his useful tips.
Casper Barendrecht, Guanyu Jin, Stijn Moerman en Nand Snijder. Delft, April 2017

Contents

1 Introduction

In 1978, Thomas C. Schelling developed his tipping model by placing pennies and dimes on a chess board and moved them according to various rules. By viewing the pennies and dimes as two types of people, the rule of moving as a preference for the individuals, and the chess board as a city, he soon discovered that segregation is formed on the board, even when the preference of the individuals is very subtle.

Based on this idea, we first built a basic model which consists of an 8×8 board with 40 individuals that are divided in two types. The individuals are moved according to their 'Happiness' in the current place. For the basic model, an individual is considered happy if $\frac{1}{3}$ of his/her second order neighbours (a person has at most 8 neighbours) is of the same type. Otherwise, an individual is considered unhappy and will be moved to the nearest place such that his/her happiness is strictly higher, which will be referred to as the 'Happiness Rule'. After that, we extended the basic model by changing the parameters such as the size of the population, board and number of types. We also included an option for random displacement: an individual which is not happy will be moved to a empty location randomly. (That is to say, that individual will be placed to any empty spot with equal probability.)

For both the basic and the extended model, we ran 500 simulations several times and investigated how different values of the parameters affected the segregation pattern. In order to formulate our research goals precisely, the following definitions are important:

1. **Generation**: A population is said to have entered a next generation if the happiness of every individual has been checked once.
2. **Equilibrium**: The population is said to have reached an equilibrium if there are no individuals that have moved in the past generation.
3. **Segregation time at n%**: The segregation time at n% is defined as the number of generations such that n% of the population has all his/her neighbours of the same type.

For this project, we focussed on the following main questions:
1. How do the parameters affect the equilibrium? Does the population always reach an equilibrium? How many generations on average does it take to reach an equilibrium? What's the probability distribution of the number of generations to reach an equilibrium?
2. What fraction of the individuals is happy after the equilibrium? Can we optimize that by changing the size of the bord?
3. What is the segregation time for 60% and how does the definition of the happiness affect it?

Short abstract about main findings...
how the report looks like further on...

2 Equilibrium

A board has reached quilibrium after g generations if, in the $g+1$-th generation, no one has moved.
One of the research questions was: will the board reach an equilibrium (that is, no more moves)? Based on intuition, you'd expect this to be true. An individual who moves, does this to a place where there are relatively more neighbours of her type, so most of the time there are more individuals that gain happiness than those who lose happiness. In the basic model however, this is almost always true, but sometimes we get a periodic solution.
Counterexample. Consider the following board. The numbers stand for the turn order (1 is selected first,

Figure 1: Counter example: 37 and 38 will move periodically.

then 2, etc.). Red and black stand for the 2 types. After some checkwork, we see that individuals 1 to 36 are all happy, but 37 is not. In 37's turn, we see that 37 has a happiness of 0, and the closest empty spot has happiness of $\frac{1}{7} > 0$, so 37 will move to this spot. New board:

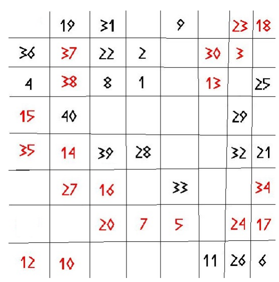


Figure 2: Counter example: 37 and 38 will move periodically.

Next, it's 38's turn. 38 has a happiness of $\frac{2}{7}$, which is less than the required $\frac{1}{3}$. The closest spot with greater happiness is the nearby corner spot with happiness $\frac{1}{3}$. Now the board looks like this:

Figure 3: Counter example: 37 and 38 will move periodically.

The others will remain happy and will not move. When it's 37's turn again, 37 has happiness $\frac{1}{7} < \frac{1}{3}$. The closest empty spot has happiness $\frac{1}{6} > \frac{1}{7}$, so 37 will move to that spot. Now 37 and 38 have swapped position:

Since 37 and 38 are of the same type, those 3 moves will repeat: we have a periodic solution.

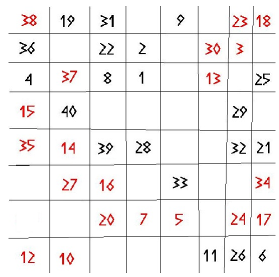


Figure 4: Counter example: 37 and 38 will move periodically.

Now what went wrong? After the first move, both 37 and 38 will gain happiness, but on the third move, 38 will lose all of her happiness. We get an endless loop.

Note that, for larger boards, this periodic solution can still appear, because we can keep the individuals near the periodic solution the same, or even put everything in the upper left corner.

3 Average number of generations until equilibrium

There are lots of parameters that may affect the average generations until equilibrium is reached. We chose to look at the influence of the Happiness Rule(HR) on the average generations. For each HR ranging from 0 to 1 with increment 0.01, we ran 500 simulations and calculated the average, maximum and minimum generations it took to reach equilibrium. The results are shown in figure 5.

In figure 5, we observe that the average number of generations (blue graph) is increasing with the Happiness Rule, which is expected since a higher HR requires more neighbours of the same type. We also note that at around a HR of 0.7, the average number of generations becomes more or less constant. That is also expected, because the requirement for the neighbours will be the same. For example, with a happiness of 0.8, a person with 3, 5 or 8 neighbours would require respectively 3, 5 or at least 7 of the same types. This requirement won't change much if a HR of 0.9 is applied. The same argument also explains why the average number of generations is constant at very low HR or for HR with little difference. So it would be sufficient if we ran the simulations with increment for example 0.1.

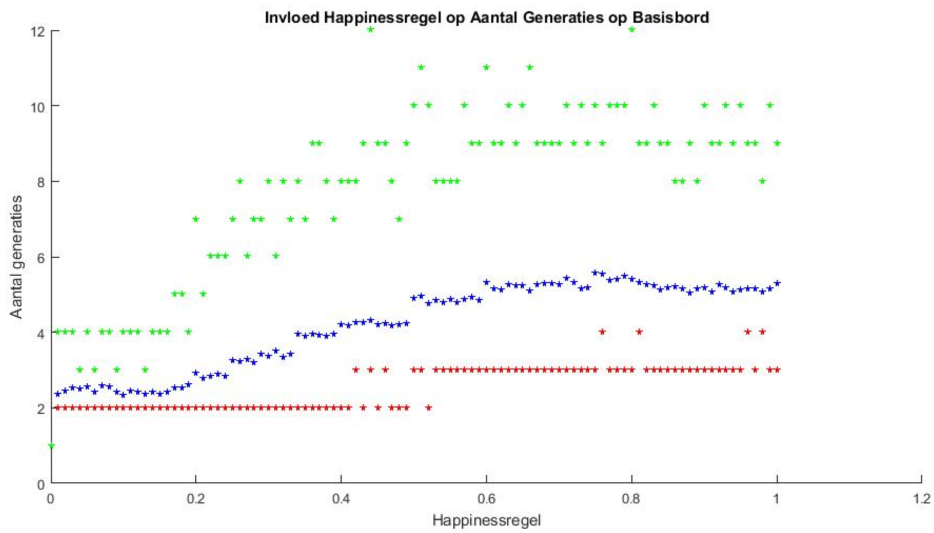
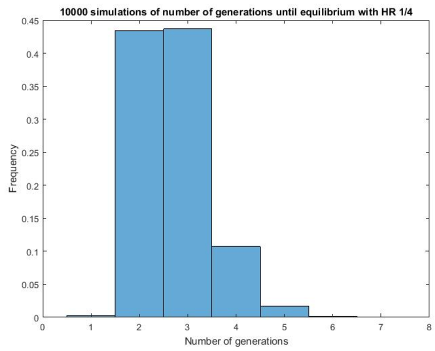
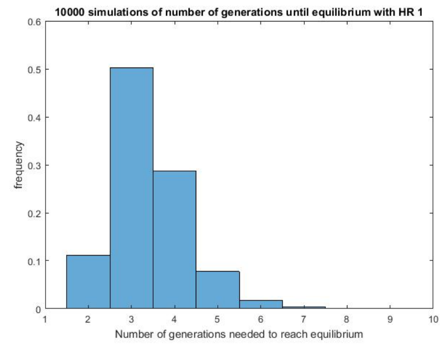


Figure 5: Influence of happiness rule on the number of generations until reaching the equilibrium. The green graph shows the maximal generations, the blue graph shows the average generations and the red graph shows the minimal generations

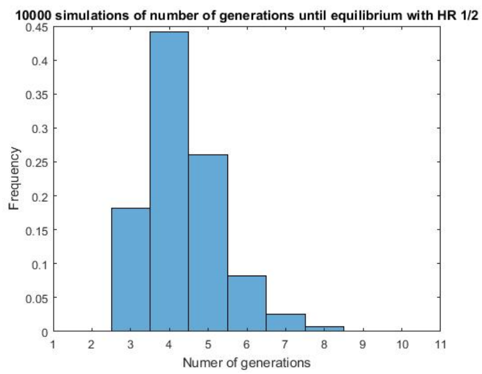
It is also interesting to know how the random variable Y, which we denote as the number of generations it takes to reach the equilibrium, is distributed. Since Y depends on the position of each individual, whose position also depends on other individuals, the probability distribution of Y might be complicated. Therefore, we approximated the distribution of Y with a histogram (figure 6) of bin size 1, which is reasonable because Y only takes integer values. Also, to look at the influence of the HR on the distribution of Y, we made histograms for HR of 1/4, 1/3, 1/2 and 1. We believe that the histograms are a good model for the distribution for Y, since accroding to the law of large numbers, the histogram will converge to the actual distribution since it is obatained after 10000 simulations.



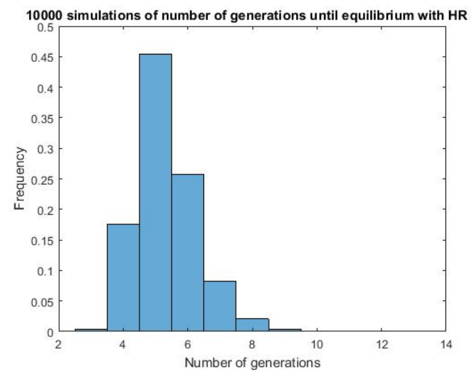
(a) 10000 simulations of number of generations until equilibrium with HR 1/4



(b) 10000 simulations of number of generations until equilibrium with HR 1/3



(c) 10000 simulations of number of generations until equilibrium with HR 1/2



(d) 10000 simulations of number of generations until equilibrium with HR 1

Figure 6: The probability distribution of Y (the number of generations until reaching the equilibrium) is approximated with a histogram of bin size 1. For each HR(1/4, 1/3, 1/2 and 1), 10000 simulations were ran.

From figure 6, we see the distribution of Y is quite similar to the binomial distribution. But we don't think that it is binomially distributed because more simulation does not give a more symmetrical distribution. It is not clear for us yet how it is distributed.

Interesting to note though, is that both the mean and the variance increase with the happiness rule. In line with the following section, we see a relatively small mean and relatively often large number of generations.

Partial explanation. While not immediately intuitive, the relatively small mean cannot be immediately explained. However we do have an explanation for the relatively high frequency for the large number of generations. If, after a few generations, we do not have equilibrium, moving to a better location gets harder (see next section for precise details). Thus, reaching equilibrium eats relatively much time.

4 Is every individual happy after equilibrium?

As it is shown in the previous section, equilibrium is not always reached within 10000 generations, or at all. But when it does, it is natural to question whether every individual is happy and how happy they are. To answer this question, we ran 500 simulations for each happiness rule ranging from 0 to 1 with increments of 0.01 and we calculated what the average, maximum and minimum happiness of the population is for each happiness rule. We ran the simulations for an 8×8 bord and 40 individuals and 20 per type. The result is shown in figure 7.

In figure 1, we observed that the minimum happiness is almost always equal to or greater than the happiness rule, which means that every individual is happy after the equilibrium. We also saw that maximum happiness is always 1, which means there is always one invididual who has all his/her neighbours of the same type. Furthermore, we see that the average happiness is always closer to the maximum happiness than to the minimum happiness and that at around a happiness rule of 0.8, we see that every individual has happiness 1. This implies that happiness 0.8 guarantees full segregation.

It is not very surprising that every individual is happy after the equilibrium, since it would otherwise mean that one individual is unable to find a place that better meets his/her desire. On a board with 24 empty spaces, this does not seem very likely. Another remarkable result is that figure 7 shows the minimum happiness is higher than the happiness rule and that the average happiness is closer to 1 if the happiness rule higher than around 0.8. At the first glance this might seems impossible because the heavy requirement of being happy, but on the other hand, it shows that the strong 'need' for segregation actually leads to higher happiness.

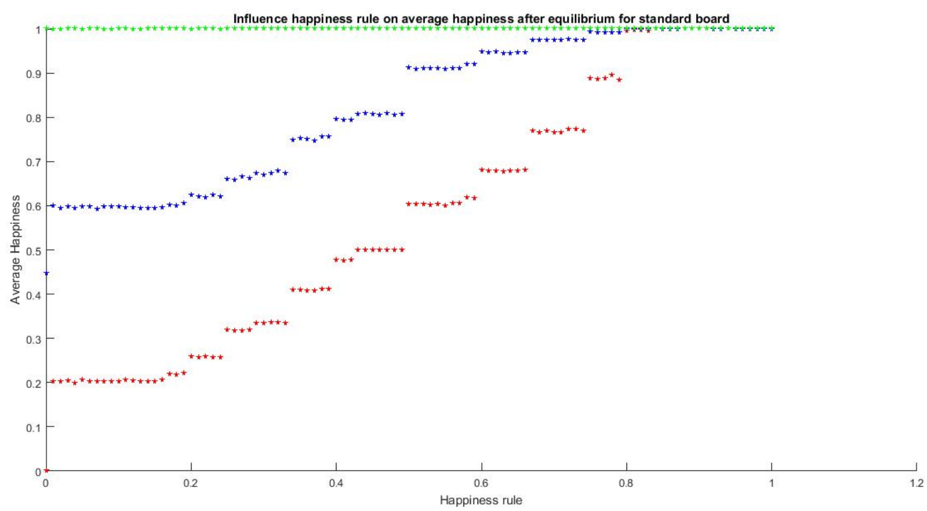
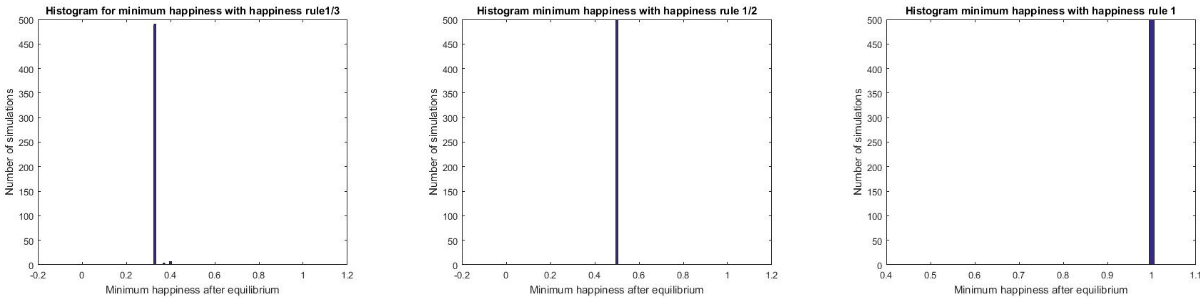


Figure 7: Influence of happiness rule on the happiness of the individuals. The green graph shows the maximal happiness, the blue graph shows the average happiness and the red graph shows the minimal happiness

Just to illustrate that the variance of every above calculated average happiness is practically zero, we only selected three happiness rules as an example and made a histogram in figure 8. In figure 8, we see that the probability of every individual is happy is practically 1.



(a) histogram minimum happiness with happiness rule 1/3

(b) histogram minimum happiness with happiness rule 1/2

(c) histogram minimum happiness with happiness rule 1

Figure 8: Histogram for minimal happiness after 500 simulations. The applied happiness rules are 1/3, 1/2, 1

5 Average segregation time

As mentioned earlier, the segregation time of $n\%$ is defined as the number of generations until at least $n\%$ of the population on a board lives in homogenous groups. Where person i is said to live homogenous if for any neighbour j of i, we have $\text{Type}(j) = \text{Type}(i)$.
This gives immediate rise to questions concerning the relation between the choice of n and the average segregation time at $n\%$. Furthermore, it is unclear if segregation at $n\%$ is guaranteed before a board reaches an equilibrium and what the effect is of the happiness boundary on the existence of a segregation time.

To research any of the given questions, we will first have to formalise our choices of board as well as the questions proposed.

5.1 Formalisations

Prior to starting any test or properly formalising our research questions however, we note that segratation at $n\%$ does not necessarily have to happen: If we consider $n = 100$ on the standard board with happiness 1/3. We will nearly never have total segregation before the board reaches an equilibrium. Therefore one might instead consider the average fraction of segregation at equilibrium,for any given happiness fraction.

Furthermore, the average segregation time as function of the segregation fraction should theoretically be a strictly increasing function since for any given board we have:

$$\begin{aligned}&n\% \text{ lives in homogenous groups after } k \text{ generations } \Rightarrow \\ &m\% \text{ lives in homogenous groups after } k \text{ generations, for any } 0 \leq m \leq n\end{aligned}$$

Having noted these facts, we can now properly formalise the research questions.
The following two questions are proposed:

1. ***What is the relation between the average segregation time and n.***

2. What is the average segregated fraction of the population after a board reaches equilibrium for given choices of happiness.

To establish results regarding these questions, we consider different setups in testings. We will be testing two different boards. The first board to be analysed is the standard board. The second board is a larger "4-Type" board. The details are specified below:

Table 1: My caption

	Standard Board	4-Type Board
Number of types:	2	4
Length:	8	10
Width:	8	10
Happiness:	1	1
Population per type:	20	16

The 4-Type board is constructed to maintain the same ratio of inhabited and uninhabited spots as the standard board. The choice of happiness on these boards is 1 unlike the usual $\frac{1}{3}$. This guarantees that for any $n \leq 100$, segregation at $n\%$ takes place prior to the board reaching an equilibrium. To observe the average segregation time $n\%$ for any n, 500 simulations will be ran per board and averaged out in order to give an approximation for the average segregation time at $n\%$. Likewise the average segregated fraction will be estimate by the average of the segregated fraction of an equilibrium from 500 simulations with given happiness q.

5.2 Results

The results regarding the first question are shown below:

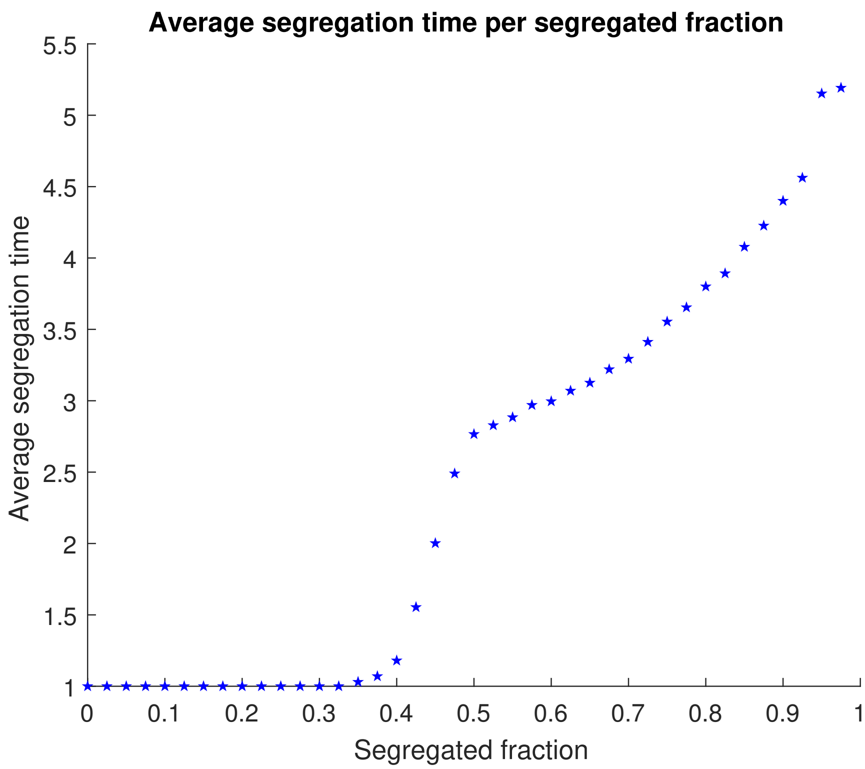


Figure 9: Average segregation time on the standard board

Most notable of figure 9 is that it is neither linear nor exponential, which is what one might initially expect. Instead, it appears to be partially exponential and partially lineair. From the figure, we note that the average segregation time increases fastest between 0.4 and 0.5. Which is to say that it takes three times longer for 50% of the population to live in homogenous groups than for 40%.
Also note that the 'lift off' is approximately at $\frac{1}{3}$, but it is unlikely that this has anything to do with the standard happiness rule of $\frac{1}{3}$.
Partial explanation. The first part of the graph is easy to understand. Because of the initial random placement of the individuals, the odds are quite good that few individuals are already homogenous.
For the second part between 0.4 and 0.5, it appears to take a relatively long time to transform from the initial chaos to 50% segregation. A 'boundary' has to be taken, comparable with the activation energy for chemical reactions.
If 50% segregation has been reached, it becomes more easy to segregate even further, because individuals can easily move from one homogenous group (of the other type) to another (of their own type), which explains the part between 50% and 70%. One'd expect this to continue and it'd be done in another one or two generations, but with 70% segregation, it gets harder for individuals to move to a location that is homogenous.

Next, we consider the same question but this time for the 4-Type board, and we get the following result.

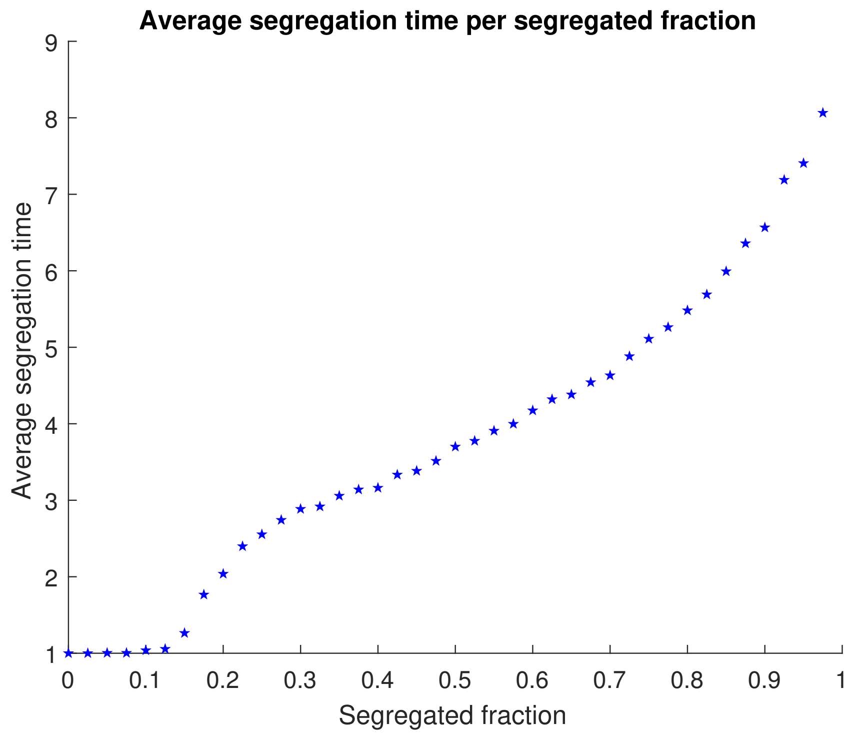


Figure 10: Average segregation time on the 4-Type board

The most notable difference between this figure and the figure 9 is that the 'lift off' lies much earlier than in figure 9. Now based on the number of types, you'd expect this to be around half as soon.

Now why is the fraction at which figure 10 lifts off not exactly half, but sooner than the lift off of figure 9? Although this difference feels intuitive, we will give a formal explanation. Consider the probability that any individual i has a neighbours of the same type, and no neighbours of another type, given $k - a$ empty spots, where k is the maximum number of possible neighbours for the location of i. In the basic model, this probability equals

$$p_b = \frac{19}{39}\frac{18}{38}\cdots\frac{19-a}{39-a} = \frac{\frac{19!}{(19-a)!}}{\frac{39!}{(39-a)!}} = \frac{19!(39-a)!}{39!(19-a)!}$$

In the 4-Type model, this probability is

$$p_4 = \frac{15!(63-a)!}{63!(15-a)!}$$

which is less than half of p_b.
Also note that an individual is more likely to be homogenous if it is placed at a corner or edge. This is because a corner or edge spot has less possible neighbours than an interior spot, thus increasing the probability of only having neighbours of your own type. The 4-Type board has relatively fewer edge positions in comparison to the standard board. This effect also decreases the fraction of segregation prior to displacement of individuals. However the people/boardspace ratio is almost the same in both models and thus should not have too much of an impact on the results.

Another effect is the rise of the number of generations, that is, the time it takes before the board reaches full segregation. However, this effect is less strong than one might initially expect. On the standard board, it takes on average 5.5 generations to reach full segregation. On the 4-Type board however, it takes 8.5.

The remaining part of figure 10 is quite similar to figure 9: quick rise, followed by a slightly decreasing ascending relation. The explanation is similar to that of the basic model.

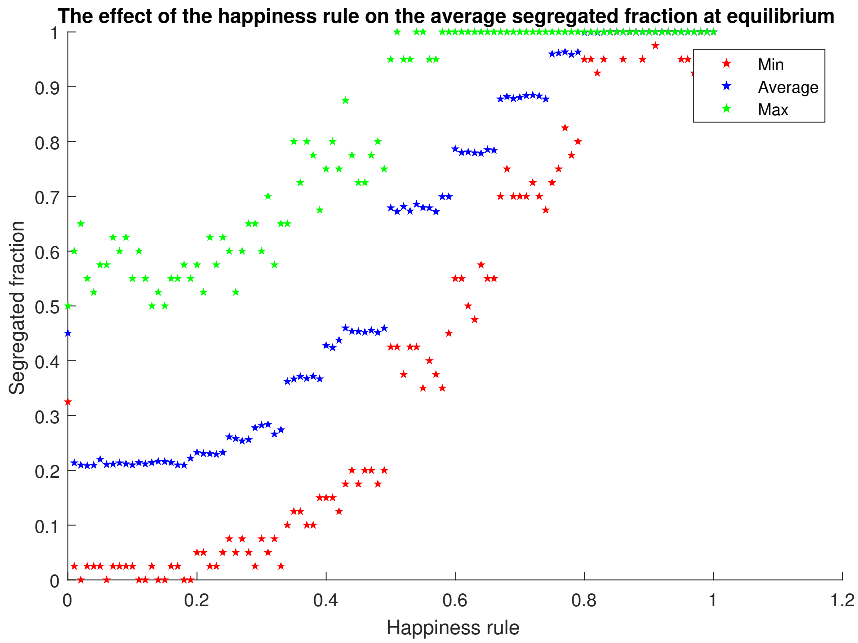


Figure 11: Average segregated fraction as a result of happiness on the standard board

Figure 11 displays three scattered values. The red and green dots represent the minimum and maximum segregrated fraction of 500 boards for the given happiness. Notable in this picture is that the average segregated value is greater than the happiness. Another thing worth noting is that the segregated fractions tend to appear in different groups seperated by relatively large percentages.

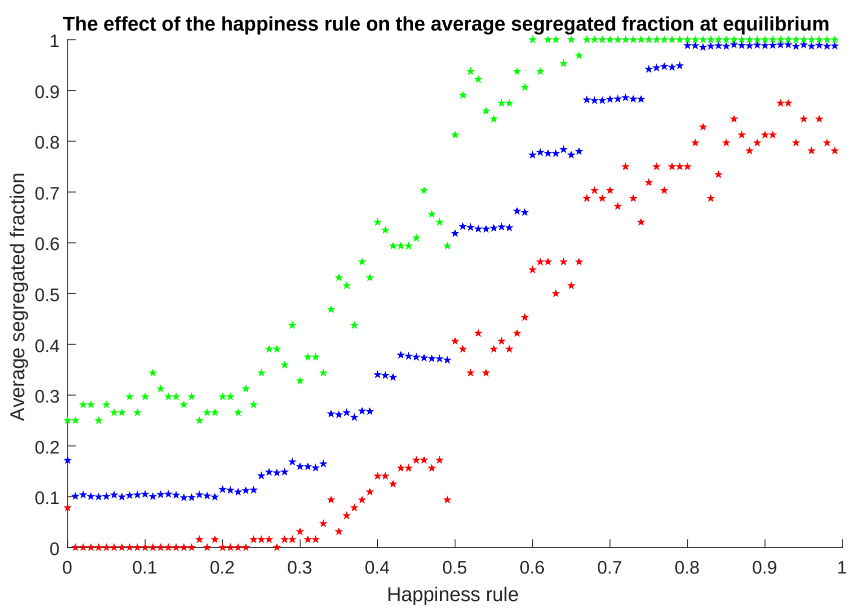


Figure 12: Average segregated fraction as a result of happiness on the 4-Type board

Figure 12 shows that contrast to the standard board, this board does not respect the property that the average segregated fraction is always greater than or equal to the happiness. A thing of interest however, is that the averages of both figure 11 and 12 seem to cluster.