Segregation

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

This page is left intentionally blank.

Abstract

An extended version of the 'Shelling Tipping Model' is built...

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 takes place on the board, even when the preference of an individual is very subtle.

Based on this idea, we created a basic model which consists of an 8×8 board with 40 individuals that are divided evenly into 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) shares his type. Otherwise, an individual is considered unhappy and will move to the nearest place where his/her happiness is strictly higher. This will be referred to as the 'Happiness Rule'. An example of how this model goes in effect is given below. It is easily observed that not every individual in figure 1a is happy. After segregation the individuals have relocated themselves as in figure 1b, there every individual is happy.

(a) Situation before segragation
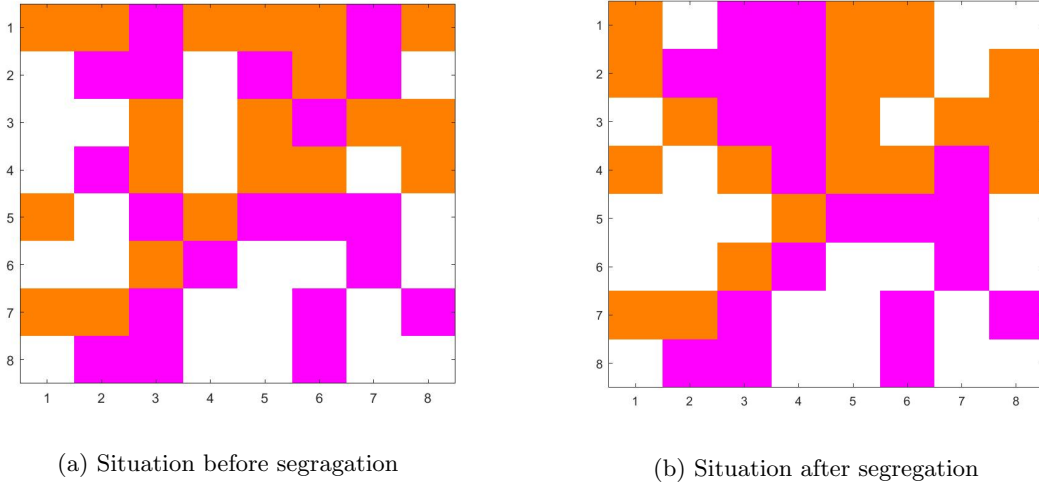
(b) Situation after segregation

Figure 1: An example of segregation in our basic model: type 1 is orange and type 2 is pink

After that, we extended the basic model by changing the parameters such as the size of the population, the boardsize and the number of types. We also included an option for random displacement: an individual which is not happy will be moved to a empty location chosen at random. (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 multiple times and investigated to what extent certain parameters affected the segregation pattern. In order to formulate our research goals precisely, the following definitions are introduced:

1. **Generation**: A population is said to have entered the next generation if the happiness of every individual has been checked once.

2. **Equilibrium after g generations**: The population is said to have reached an equilibrium after g generations if no individual has moved during the $g + 1$-th generation.

3. **Homogeneity**: A person is said to live homogenous if all of his/her neighbours share his type.

4. **Segregation time at $n\%$**: The segregation time at $n\%$ is defined as the number of generations after which $n\%$ of the population is homogenous.

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 this be optimised by changing the boardsize?

3. What is the average segregation time at 60%, 80% and 100%. How is this segreggation time distributed with Happiness Rule 1 and how does the Happiness Rule affect the average segregation time?

2 Our Model

As mentioned in the introduction, we will mostly be investigating the basic model: segregation of 20 individuals of type 1 and 20 individuals of type 2 on an 8×8 board. In this model an individual is considered unhappy if less than $\frac{1}{3}$ of his/her neighbours shares his/her type. If an individual is unhappy, this individual will move to the nearest place where he/she would be happier. Any person without neighbours is by definition considered unhappy (so his/her happiness is 0), since people prefer living with others over living alone. Furthermore, the basic model uses the second order neighbourhood, The maximum number of neighbours a person can have is 8. The model however can be changed by preference in several ways.

First of all, the boardsize can be varied. If no other parameters are altered, this will only imply that the board becomes fuller or emptier. Secondly, we can change the number of types on the board. For instance, one can consider a 12×12 board with 10 types of individuals and 10 individuals per type: see figure 2. As one can see, far more individuals have been moved to another place on the board and segregation takes place.
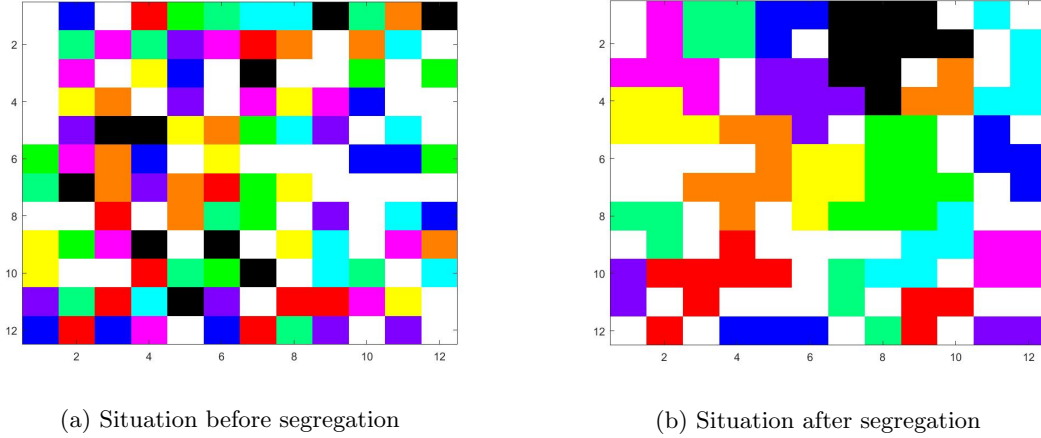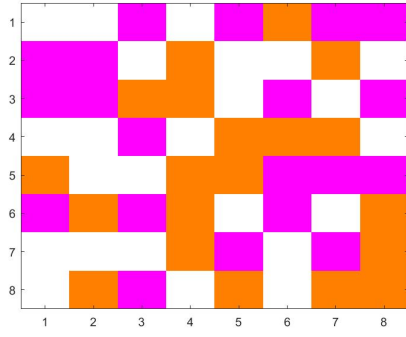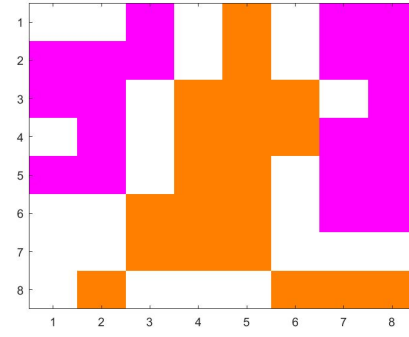

(a) Situation before segregation

(b) Situation after segregation

Figure 2: An example of segregation in a model on a 12×12 board with 100 individual of 10 different types

Subsesquently, we can vary our 'Happiness Rule': the minimal fraction of neighbours that has to share an individuals type, in order to be considered happy. The expectation is that the time to reach equilibrium increases as the Happiness Rule is strengthened. This makes sense as every individual needs to be surrounded by relatively more neighbours of his/her own type. In consequence, larger homogenous groups will be formed after segregation with a higher Happiness Rule. In figure 3 and 4 the effect of a Happiness Rule of respectively $\frac{2}{3}$ and 1 on the board after segregation is illustrated.
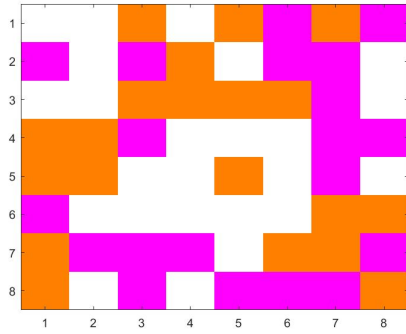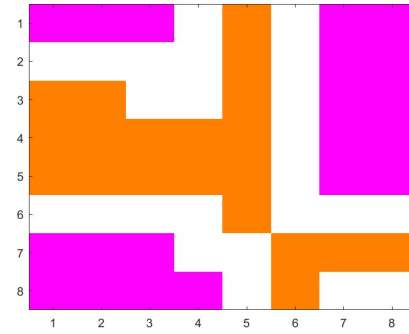
(a) Situation before segregation

(b) Situation after segregation

Figure 3: An example of segregation in a model with happinessrule $\frac{2}{3}$

(a) Situation before segregation

(b) Situation after segregation

Figure 4: An example of segregation in a model with happinessrule 1

In addition, the type of neighbourhood can be changed. The neighbourhood is by default a second order neighbourhood, in which both the adjacent and diagonal people are considered neighbours. There exists only one smaller type of neighbourhood: the first order neighbourhood, in which only adjacent neighbours are accounted for; A person has up to four neighbours. The way the different types of neighbourhood are defined is illustrated in figure 5. Here 0 represents the individual himself and every other place displays which order of of neighbourhood it belongs to. We will not research the effect of the type of neighbourhood on the segregation in our model in this report. However, some interesting graphs are included in the appendix.

Figure 5: Neighbourhoodorder

Moreover, we included the possibility for individuals to move to a random place on the board if he/she is not happy, rather than moving to the closest. In the basic model he/she will move to the nearest place where he/she will become happier, With random displacement, we expect that it takes longer to reach an equilibrium. Again, we will not discuss the impact of the random displacement in our report, but we added some nice results in our appendix.

Furthermore, we expanded the model with the so-called 'criminals'. One can choose to have c out of the n individuals behave as if they were a criminal. Any person who is not a criminal himself, does not like criminals in his neighbourhood, and thus his/her happiness will be 0, overruling any positive effect of friendly neighbours, if a criminal lives in his/her neighbourhood. As a result, all criminals should end up living together after segregation. In figure 6, one can see an example of the basic model with 10 criminals among

7

the 40 individuals. It is visible that our expectation is not always the outcome. In this case, equilibrium is reached, because every unhappy individual cannot move to another placen, since every empty location is in the neighbourhood of a criminal. So the unhappy individuals will not become happier on these places.

(a) Situation before segregation
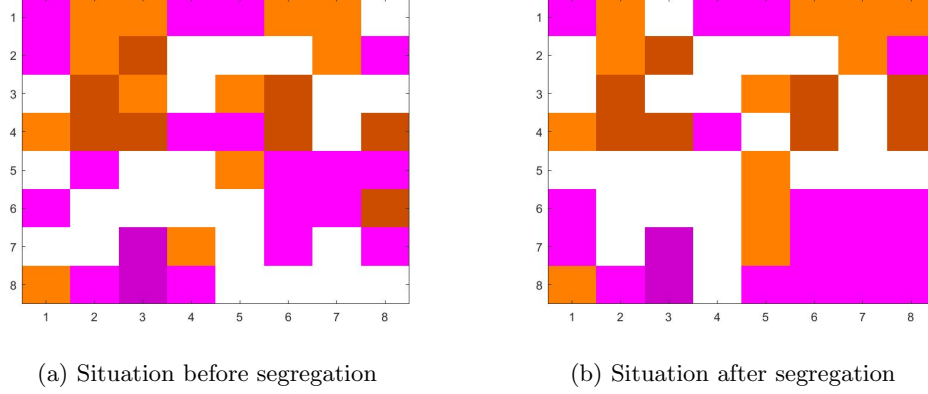
(b) Situation after segregation

Figure 6: An example of segregation in the basic model with 10 criminals (the slightly darker places)

Last but not least, we made it possible for individuals to switch to another type. In our model there are two types of switching. The first is only possible if the only individuals are of type 1 and type 2. In this version of switching, every person has in every generation a chance p of switching the other type. In the other version, every unhappy individual will change type under the influence of his/her neighbours. This means that the chance to become of type t is the fraction of neighbours of type t. This is a correct probability distribution, because this sums up to 1. Individuals without neighbours will not switch type, because they have no neighbours who influence them. We will investigate the influence of the second type of switching further in section 7.

3 Equilibrium

As mentioned in the introduction, a board has reached quilibrium after g generations if no person has moved in the $g + 1$-th generation.

One of the research questions was: will the board reach an equilibrium (that is, no more moves)? Based on intuition, this is expected 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:

(a) The start situation. Individual 37 is on her turn.

(b) Individual 37 is moved to the nearest place that better meets her desire.

(c) Individual 38 is on her turn and is moved to the nearest place that better meets her desire.

(d) A generation has passed, 37 is on her turn and is moved to the nearest place for higher happiness. Because 37 and 38 are of the same type, this process repeats.

Figure 7: A possible configuration of the standard board in which the equilibrium will not be reached. The numbers indicate the selecting order of an individual. The two different colour indicates the two types. The equilibrium is not reached due to the periodic movements of individuals 37 and 38. The movements of invduals 37 and 38 are tracked and shown in subfigure a to d in a chronological order.

In figure 7, a possible configuration in which the equilibrium will not be reached is shown. The numbers stand for the turn order (1 is selected first, 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 that spot, which is indicated in figure 7b.

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 $\frac{1}{3}$. So 38 will now move to the place which is shown in figure 7c.

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, as it is shown in figure 7d.

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

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.

4   Average number of generations until equilibrium

There are lots of parameters that may affect the average number of 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 8.

In figure 8, 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, and thus less probability that the selected individual is happy, making it more likely that she will move. We also note that at around a HR of 0.7, the average number of generations becomes more or less constant. This can be explained noting that the requirement for the neighbours will practically 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 type. This requirement won't change much if a HR of 0.9 is applied. The same argument explains why the average number of generations is constant at very low HR or if we raise the HR with a sufficient small amount like 0.01.
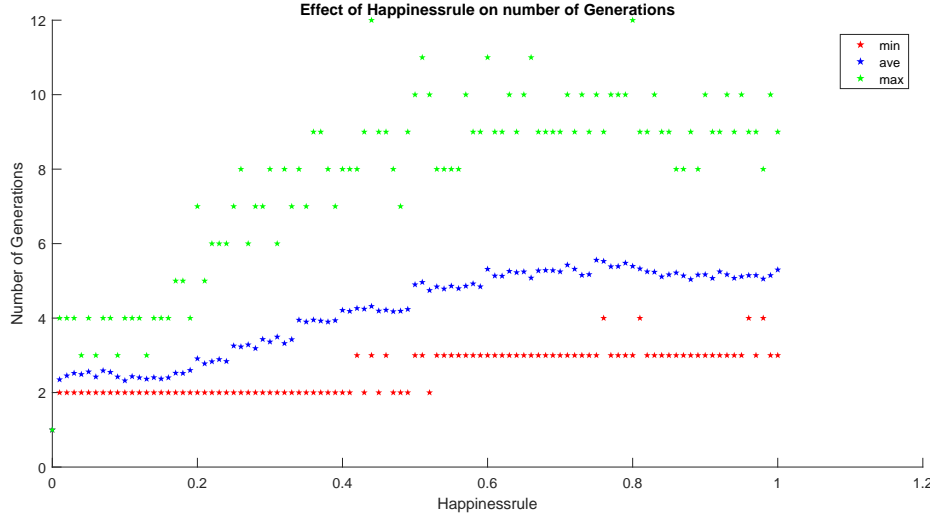

Figure 8: Influence of happiness rule on the number of generations until reaching the equilibrium. The green graph shows the maximum number of generations, the blue graph shows the average number of generations and the red graph shows the minimum number of generations

It is also interesting to investigate how the random variable Y_j, which we denote as the number of generations it takes to reach the equilibrium under HR j, is distributed, and how the distribution is influenced by the HR. For that reason, we ran 10000 simulations for each HR(1/4,1/3,1/2 and 1) and plotted some histograms(figure9). We chose bin size 1 for each histogram because Y_j is a random variable that only takes integer values. So a histogram of bin size like 0.5 will not give any extra information and a loss of lots of information will occur if bin size 2 is chosen.
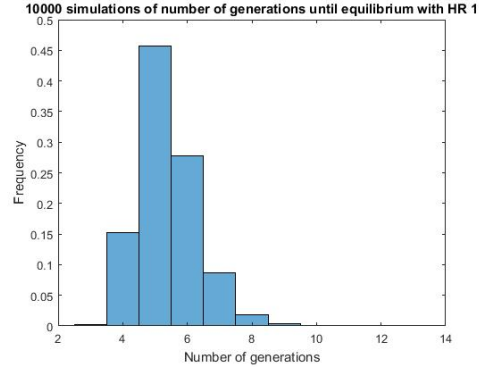
11
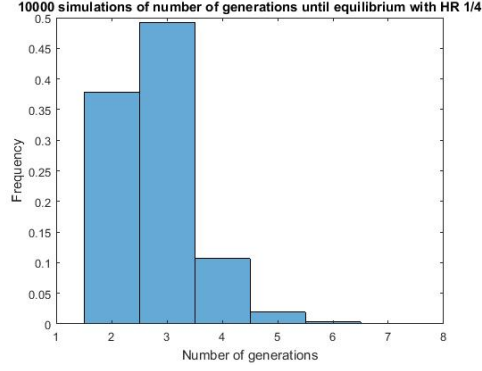
(a) 10000 simulations of number of generations until equilibrium with HR 1/4
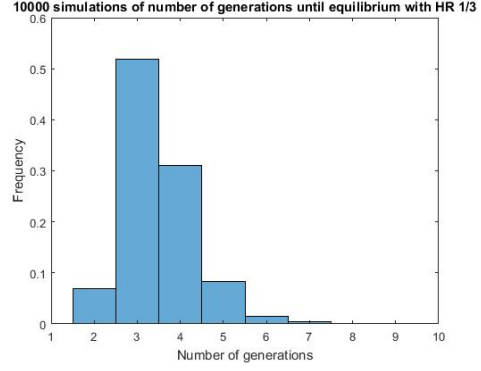
(b) 10000 simulations of number of generations until equilibrium with HR 1/3
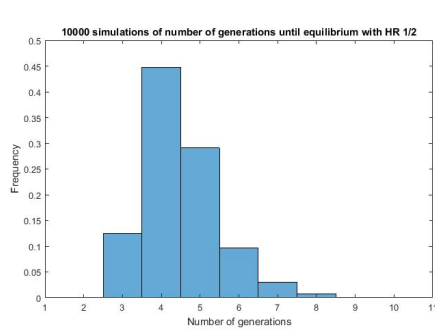
(c) 10000 simulations of number of generations until equilibrium with HR 1/2

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

Figure 9: 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.

Looking at the histograms, all Y_j's ($j \in \{1/2, 1/3, 1/4, 1\}$) might come from the same distribution, but with different parameters, since their sample means are obviously different.

In an attempt to find a good(confidence level 95%) fit for the distribution of $Y_{1/3}$, we've used the chi-squared goodness of fit test(chi2gof) to see whether $Y_{1/3}$ is Poisson distributed. We first used the 'fitdist'-function of Matlab to calculate the maximal likelihood estimater for $\hat{\lambda}$ and then performed the chi-sqaured test. The test has rejected the null-hypothesis $H_0 : Y_{1/3} \sim \text{Poisson}(\hat{\lambda})$ with significance level $\alpha = 0.05$.

With the same method, we've also tested whether $Y_{1/3}$ is binomial or negative binomial distributed. But again, they have been rejected by the chi-sqaured test. So we can only conclude that $Y_{1/3}$ is not Poisson, binomial and negative binomial distributed.
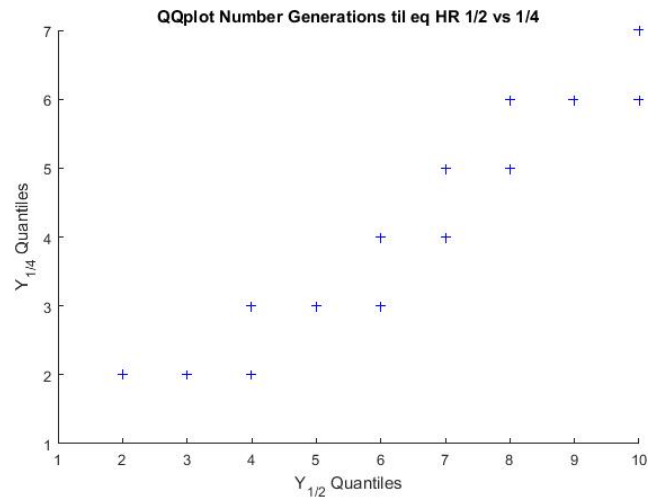
In order to make a better comparison of the distribution between $Y_{1/3}, Y_{1/4}, Y_{1/2}$ and Y_1, we made some QQ-plots, which is shown in figure 10.
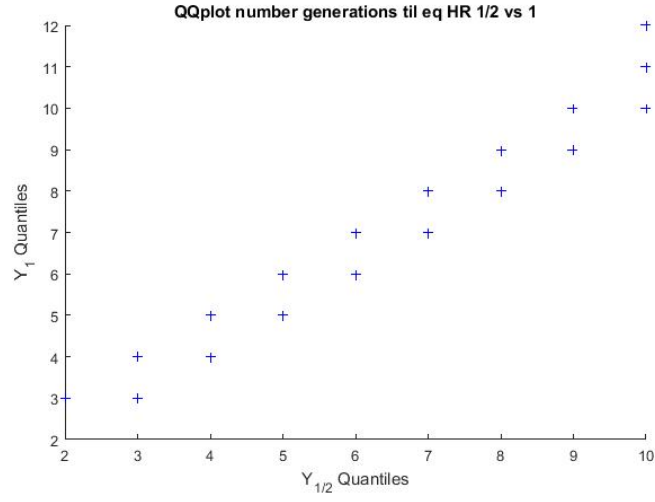
(a) QQ-plot of Y_1 vs $Y_{1/2}$
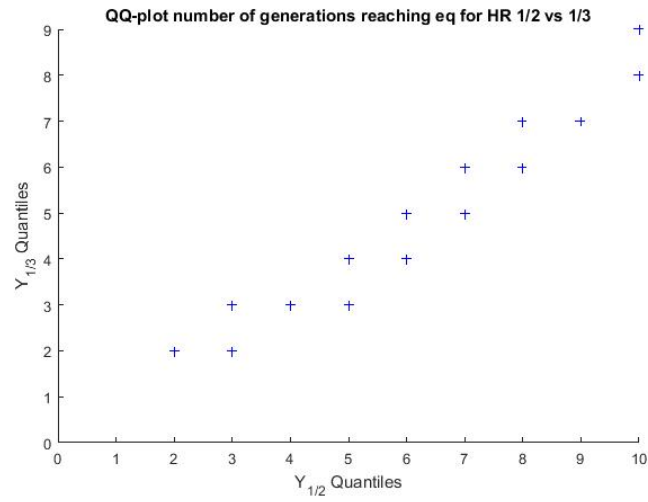
(b) QQ-plot of $Y_{1/3}$ vs $Y_{1/2}$

(c) QQ-plot of $Y_{1/4}$ vs $Y_{1/2}$

Figure 10: QQ-plot of Y_1 vs $Y_{1/2}$, $Y_{1/3}$ vs $Y_{1/2}$ and $Y_{1/4}$ vs $Y_{1/2}$

As expected, the QQ-plots have a stairwise pattern. This is because our data came from a discrete random variable. We see that the plots appears linear, especially the plot of Y_1 vs $Y_{1/2}$. The plot indicates that these 4 Y_j's might belong to the location-scale family of some distribution.

With the two-sample kolomogorov-smirnov test(kstest2) of significance level $\alpha = 0.05$ in Matlab, which has null-hypothesis $H_0 : F(X) = G(Y)$ for F, G the CDF of X and Y, it can be concluded that these Y_j's don't come from the same distribution with the same parameters. Unfortunately, we weren't able to find a test that tests whether these Y_j's comes from the same family distribution(i.e. same distribution but different parameter values).

5   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 11.

In figure 11, 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 11 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.
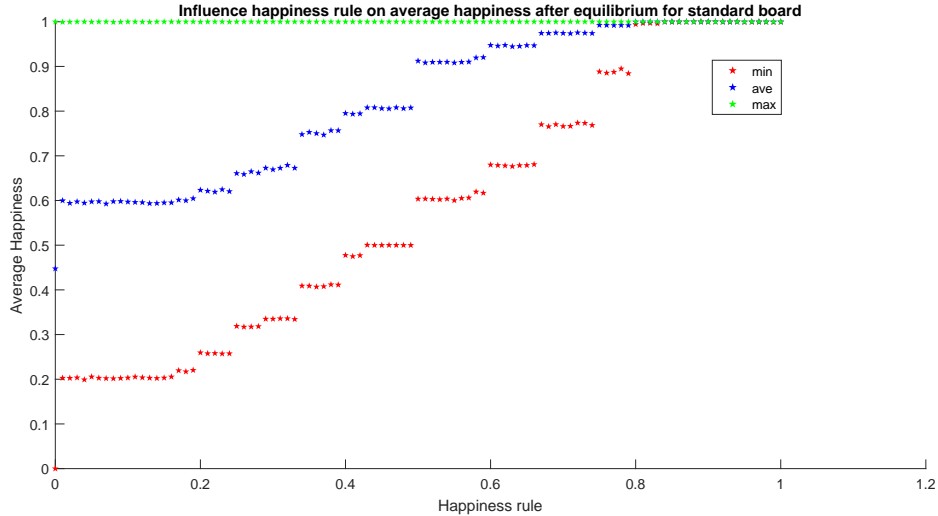
Figure 11: Influence of happiness rule on the happiness of the individuals. The green graph shows the maximal happiness, the blue graph shows the average happiness and the red graph shows the minimal happiness

Just to illustrate that the variance of every above calculated average happiness is practically zero, we only selected Happiness Rule 1/3 as an example and made a histogram in figure 12a. In figure 12a, we see that the probability of every individual is happy is practically 1.
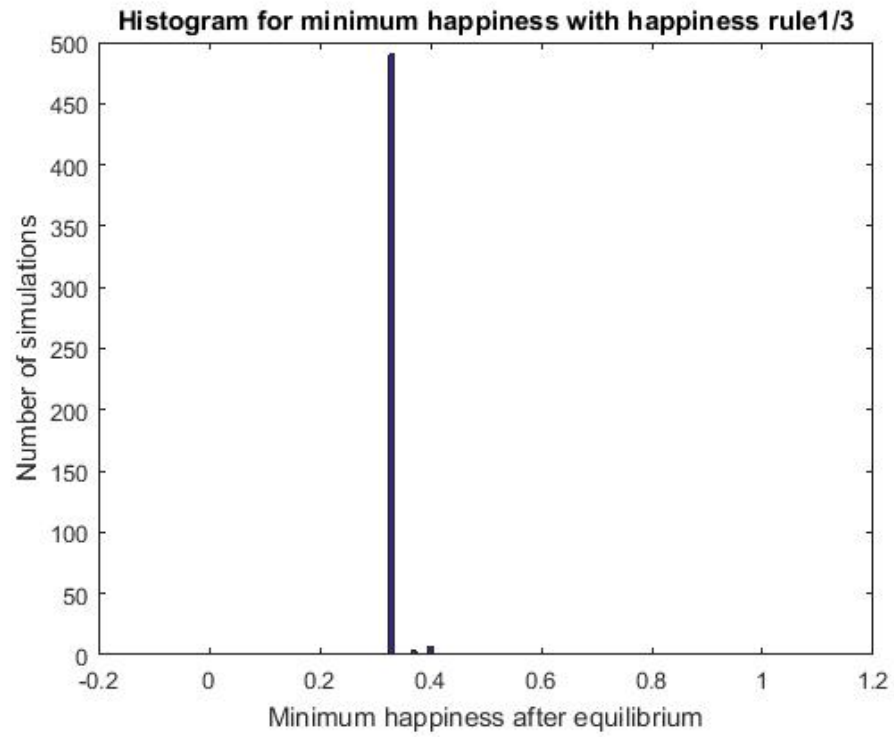
(a) Histogram for minimum happiness after 500 simulations with HR=1/3

6 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.

6.1 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:

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

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

3. For any Happiness Rule ranging from 0 to 1, how often do we reach at least 60%, at least 80% or even full (100%) segregation?

4. For 60% and 80% segregation, what distribution do we get for the segregation time, if $HR = 1$? Are these comparable?

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: Specs of the two considered boards

	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.

6.2 Results

6.2.1 Question 1

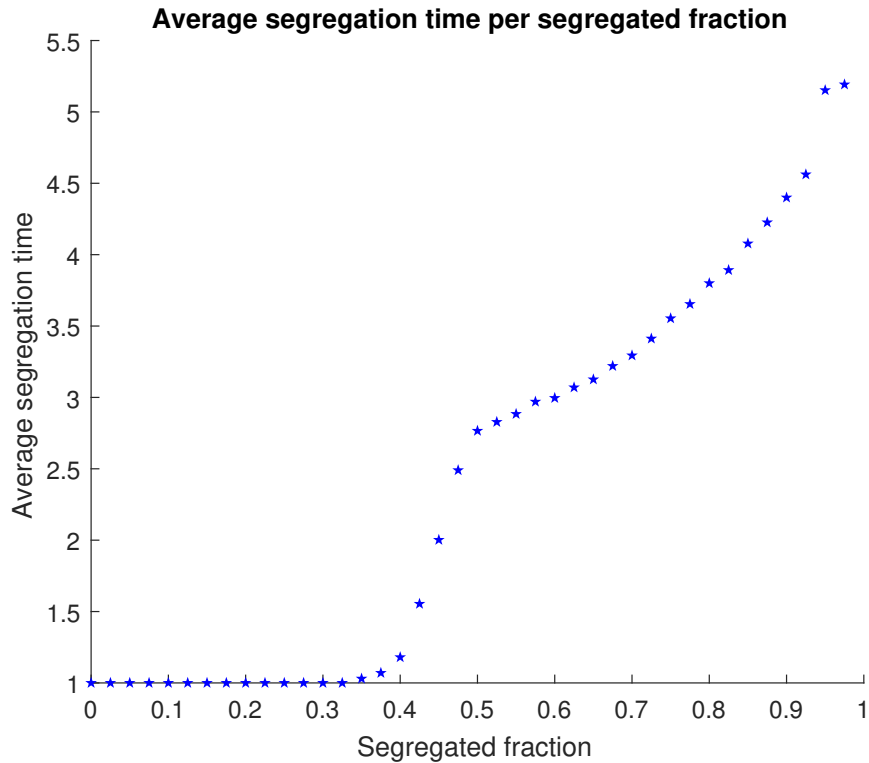The results regarding the first question are shown below:

Figure 13: Average segregation time on the standard board

Most notable of figure 13 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.

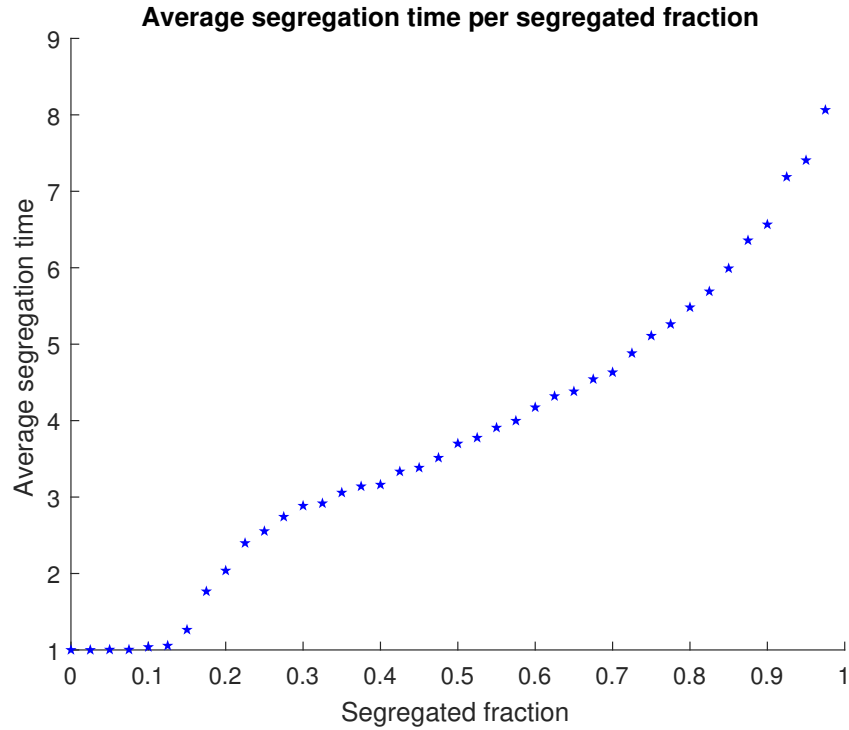Next, we consider the same question but this time for the 4-Type board, and we get the following result.

Figure 14: Average segregation time on the 4-Type board

The most notable difference between this figure and the figure 13 is that the 'lift off' lies much earlier than in figure 13. 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 14 lifts off not exactly half, but sooner than the lift off of figure 13? 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

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

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

19

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 14 is quite similar to figure 13: quick rise, followed by a slightly decreasing ascending relation. The explanation is similar to that of the basic model.
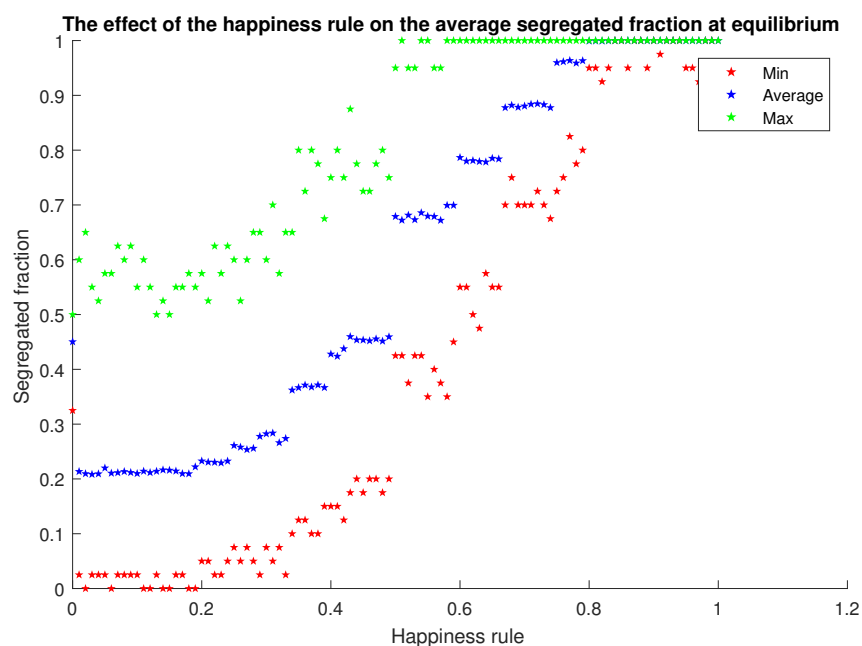
6.2.2 Question 2

Figure 15: Average segregated fraction as a result of happiness on the standard board

Figure 15 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.
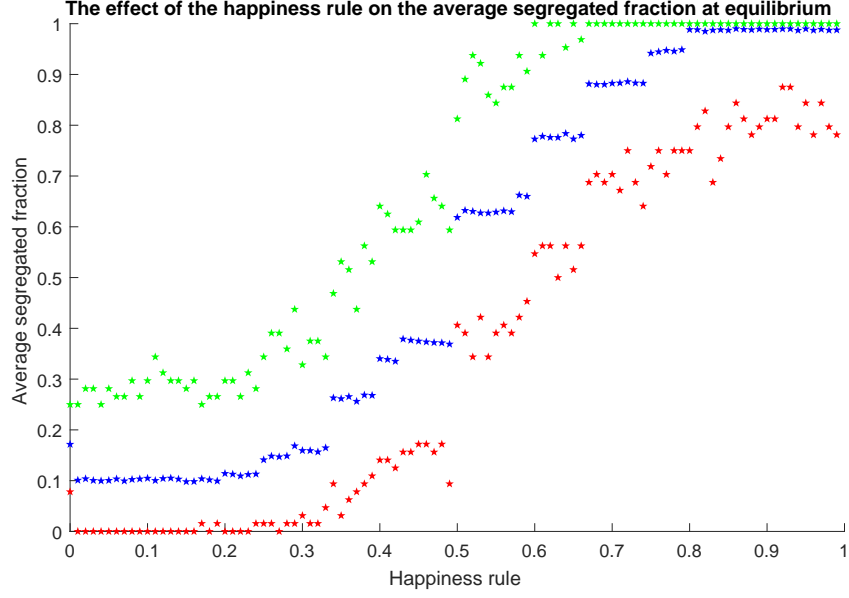
Figure 16: Average (blue) segregated fraction as a result of happiness on the 4-Type board. Also included max (green) and min (red).

Figure 17 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 15 and 17 seem to cluster. This is because, as mentioned earlier, the effect of the Happiness Rule is quite discrete when using second order neighbourhood.

6.2.3   Question 3

In order to answer the third question, we ran 500 simulations for each HR ranging from 0 to 1 with increment 0.025, and for each segregation %, 60, 80 and 100. The results are shown below.
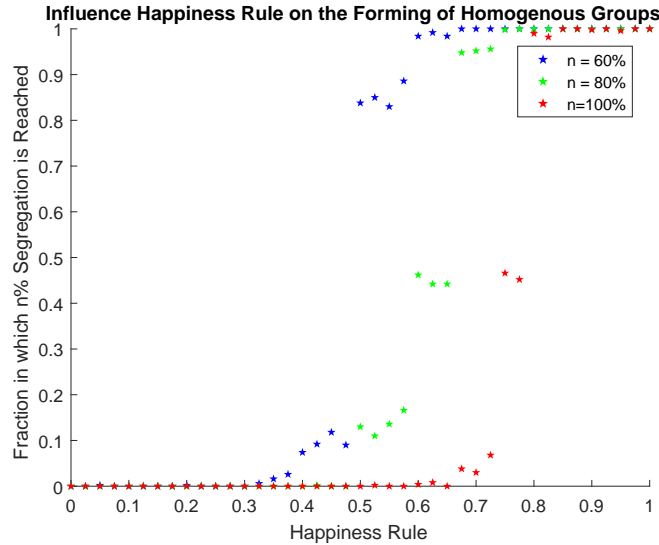

Figure 17: For each $HR = 0, 0.025, ..., 1$, the fraction of the simulations in which 60% (blue),80% (green) and 100% (red) segregation is reached.

6.2.4 Question 4

We will now focus on a more relevant question, namely the distribution of the 60% and 80% segregation time. The plots are shown below.

7 The effect of switching types

Segregation is a phenomenon taking many different forms. Rather than limiting the scope to racial segregation, we decided to broaden our view. In example, one might delve into the segregation in studies with respect to friendship. In this case, for any person k. The neighbours of k represent the k's friends or the persons with whom k spends most of his time. In this case, k's type, can represent either the classes he is currently taking, or the kind of sports person k is doing, or even what kind of cellphone k has. In any of these cases, the type of k is not necessarily set indefinitely. In these cases, the type of k might switch, depending on the types of his friends.

This gives rise to a new modification of the model:
For any person k, prior to moving to a new location k has a multinomial chance to switch to a different type, where

$$\mathbb{P}(\text{NewType}(k) = t) = \begin{cases} \frac{\#\text{Neighbours of type } t}{\text{Total number of neighbours}} & \text{if total number of neighbours} > 0 \\ \mathbb{1}_{\{t = \text{Type}(k)\}} & \text{if total number of neighbours} = 0 \end{cases}$$

Note that this is well defined, since for total number of neighbours > 0, we have $\mathbb{P}(\text{NewType}(k) = t) \geq 0$ and $\sum_{i=0}^{\text{types}} \mathbb{P}(\text{NewType}(k) = t) = 1$. An example of how this may effect the equilibrium, is illustrated below:
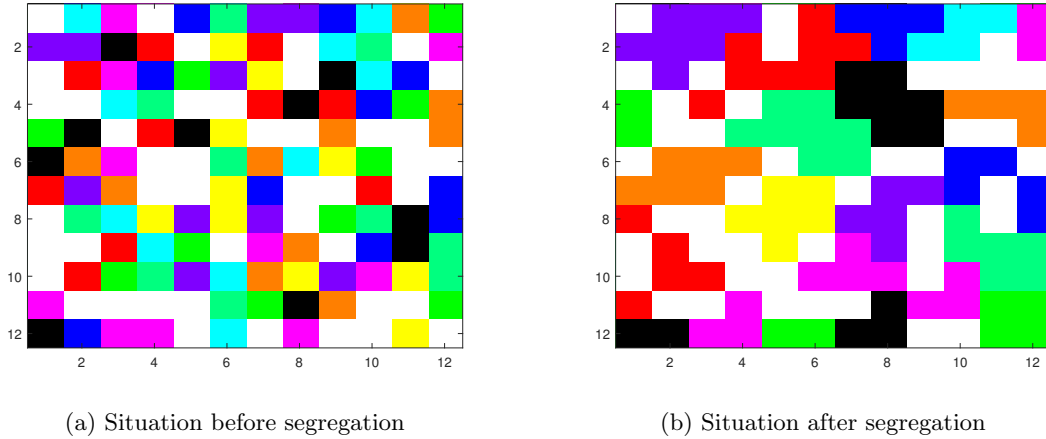

(a) Situation before segregation

(b) Situation after segregation

Figure 18: An illustration of the effect of switching types on the 12×12 Board

A crucial difference between equilibrium with and without switching, is that if switching is allowed, the concentration of the types may differ between the starting board and the equilibrium status. This can be seen in figure 18 as the number of people of the green type decreased from 10 to 8, the yellow type even decreased from 10 to 6. As a matter of fact, allowing switching might lead to the extinction of several types. This is not generally the case for lower bounds on the happiness rule, but if the happiness rule is increased, several types might cease to exist. Despite this being extraodinarily interesting, the extinction of types is not included in this report.

Appendix

Here are some of our results we did not use in our report, because we did not investigate them. However, these graphs may be really interesting to investigate further.