Schelling Model with Central Tolerance

Thomas Gaehtgens (ist186809), Pedro Duarte (ist187347), Ana Filipa Valente (ist190376)
IST (*Network Science*) — 2021/22

Abstract

This project builds upon Schelling's model of segregation by considering that each individual agent has a preference to inhabit a position as close to the center of the grid as possible. This is achieved by adding a tolerance with dependence on the distance to the center of the grid for each agent. We show that for this setup a specific structure emerges about the center of the grid where the majority group occupies the center in a circular shape and the minority group gets pushed away to the the outer regions.

I. Introduction

Early cities arose in several regions, motivated by agricultural productivity and economic scale. They provided a number of benefits for its people, bringing them all together in that one spot. As these grew, so did the competition for the available, limited space and resources.

The branch of sociology that studies these effects is human ecology - the study of man's collective interaction with his environment, analogous to an ecosystem. Robert Ezra Park and Ernest W. Burgess, the first to develop the theory of human ecology, postulated that cities were much alike the environments found in nature, governed by the same forces of Darwinian evolution. According to them, society arranges itself spontaneously because of competition processes comparable to those found in the struggle for survival in nature. Park and Burgess suggested that "scarce urban resources, especially land, led to competition between groups and ultimately to the division of the urban space into distinctive ecological niches or "natural areas" in which people shared similar social characteristics because they were subject to the same ecological pressures" [1]. Hence, they were fundamentally interested in the effect of spatial and temporal position upon human behaviour and institutions.

As stated by Encyclopedia Britannica, "ecology in the social sciences is the study of the ways in which the social structure adapts to the quality of natural resources and to the existence of other human groups", taking into account the biological, environmental, demographic, and technical conditions of the life of the citizens [2]. All these factors represent ultimately value. In our analysis, we will study these dynamics considering the value attributed to each zone within the bounds of a city.

As a motto for our work, a quote from the Burgess and Park's seminal work "The City": "Society is made up of individuals spatially separated, territorially distributed, and capable of independent locomotion" [3].

The model studied is as simple as can be: we define a radial city with geographical value and desirability decaying linearly with the distance from the centre, implying a "city centre + suburbs" structure, similar to Burgess' concentric zone model.

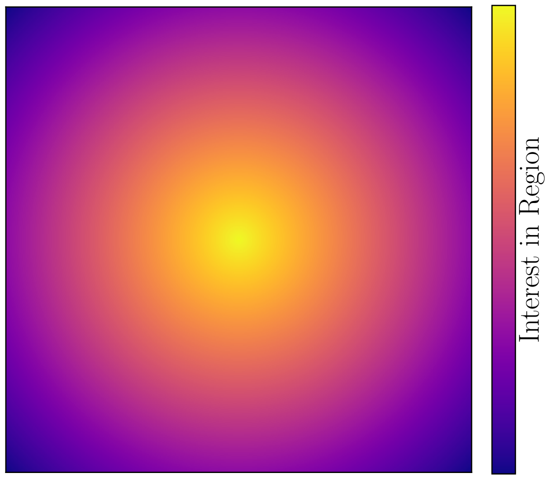


Figure 1: Graphical representation of the model used where the yellow center represents peak interest and blue corners represent minimum interest.

The creation of the aforementioned "ecological niches" is precisely the driving mechanism of segregation, quite literally, the separation of groups of people with different characteristics.

Human ecology doesn't view political context and racial inequality as the driving force behind metropolitan spatial patterns. Rather, racial and class segregation are seen as a natural result of the competition for the more desirable locations within the city.

It will be tested if a simple geographical value attribution is enough for these social structures to arise.

For this purpose, this study will work within the Schelling Model framework, to simulate the natural evolution of the city and the motion of its inhabitants.

I.1. Schelling's Model of Segregation

Proposed by Schelling in 1971, the Schelling model [4] [5] is an agent based simulation that tries to describe the kinds of segregation that can result from discriminatory individual behaviors. A discriminatory behaviour stems from conscious or unconscious actions based on the perception of difference among a population and ultimately leads to the emergence of collective segregation, based on individual preferences, the strengths of those preferences, and the means for exercising them. In order to recreate this effect in a generalized way, the model is defined by the following rules.

It starts with an $N \times N$ grid and a population comprised of two distinct agent groups: group **R** and group **B**. The grid positions can either be populated or have empty space. The fraction of occupied grid positions is given by ρ and is made up of part p_R agents of group **R** and part p_B agents of group **B**.

Each agent starts on a random position in the lattice and faces two possible outcomes based on the neighbouring population. It can either be satisfied and maintain it's position, if a fraction of it's neighbours belonging to the same group is larger than a certain threshold T, or dissatisfied, and change to a random location on the lattice. As such, an agent is more tolerant the lower the value of the threshold T is.

$$\begin{cases} F \geq T \implies \text{Satisfied.} \\ F < T \implies \text{Dissatisfied.} \end{cases}$$

This process is repeated until every agent is satisfied in it's position, if the model converges.

II. Methods

The work developed for this project builds on the basic Schelling model by considering that the agents prefer to be as close to the center of the grid as possible. As such, their tolerance increases (and the value of the threshold T decreases) the closer in euclidean distance an agent is to the center of the grid.

The modeled society is composed of a fixed set of agents $A = \{a_1, a_2, ..., a_n\}$ that can belong to one of two distinct groups, group **R** and group **B**, labeled by the colors red and blue respectively. To each of these groups is assigned a numerical value, 1 for B and -1 for R.

The size of the set A and the group to which a specific agent belongs are invariant over time.

Each agent occupies a unique position defined by (x, y) on an $N \times N$ toroidal lattice (a lattice with periodic boundary conditions) i.e. two agents cannot occupy the same position.

An agent a_i has access to it's Moore Neighbourhood (that is, the eight closest points to its position) at any given time t, defined by $N_t(a_i)$.

The variables used to specify the model are summed on table (I).

Description	Variable
Grid size	N
Percentage of the population that belongs to group R	p_R
Percentage of the population that belongs to group B	p_B
Fraction of the grid occupied by agents	ρ
Minimum threshold value on the grid	T_{min}
Maximum threshold value on the grid	T_{max}

Table I: Description of the variables being used to describe the Schelling model.

II.1. Time Evolution

The toroidal lattice is populated by placing agents belonging to a group chosen at random following a Bernoulli distribution with p = *0.5* (while there are agents from both groups available to be placed) on a random position following a uniform distribution.

Once the grid is in its starting position, it is ready to be iterated. Each iteration begins by determining the satisfaction of each agent and relocating those who are dissatisfied to a random empty position.

An agent a_i of group g_i on iteration t is satisfied if a fraction of it's neighbours given by

$$F_{i,t} = \frac{|\{a_j \in N_t(a_i) : g_j = g_i\}|}{|N_t(a_i)|}$$

is greater than its threshold value.

II.2. Threshold Value

The model used for this project strays from the original Schelling model when it comes to the threshold values used for each agent. While the original model has the same fixed tolerance for every agent, we argue that when considering cities, preference for the center should be contemplated by reducing the threshold value the closer an agent is to the city center.

A given agent a_i located at (x, y) in a $N \times N$ grid has a distance to the center given by the euclidean distance:

$$r = \sqrt{\left(x - \frac{N}{2}\right)^2 + \left(y - \frac{N}{2}\right)^2} \tag{1}$$

Assuming a **linear** dependence to the distance to center, the threshold for an agent with a distance r is given by:

$$T(r) = \frac{T_{max} - T_{min}}{R_{max}} r + T_{min} \tag{2}$$

where R_{max} is half the square grid's diagonal length, the maximum distance to the center.

As such, $T(r = 0) = T_{min}$ and $T(r = R_{max}) = T_{max}$.

II.3. Metrics

The study of the system at hand requires establishing a set of metrics that will help answering our questions.

II.3.I. Average r by Group

The first metric defined to study our system is the average r (that is, the average euclidean distance of an agent to the grids center) of all agents belonging to a given group. It is defined by

$$\bar{r}_k = \sum_{i=1}^{N'_k} r_{ki}/N'_k \tag{3}$$

where k represents the group (R or B), N'_k acts as the number of agents belonging to k and r_{ki} stands for the r value of agent a_i.

This will allow us to perceive whether or not members of a certain group are preferably populating the center or the edges of the grid. This is an especially important metric as our tolerance depends on the distance of each agent to the center of the grid.

II.3.II. Moran's Index

As proposed by [6], another metric we will use to study our system is Moran's index I, a measure of spatial autocorrelation used to tell us if our system is clustered, dispersed (uniformly distributed) or randomly distributed. This will be the metric used to study how segregated our system is.

It is defined as:

$$I = \frac{N'}{\sum_{i=1}^{N'} \sum_{j=1}^{N'} w_{ij}} \frac{\sum_{i=1}^{N'} \sum_{j=1}^{N'} w_{ij}(g_i - \bar{g})(g_j - \bar{g})}{\sum_{i=1}^{N'} (g_i - \bar{g})^2} \quad (4)$$

with N' being the number of agents in our system (of both group R and B), $\bar{g}$ being the average group of those agents ($\bar{g} = \sum_{i=1}^{N'} g_i/N'$) and w_{ij} being 1 if agent j is in the Moore neighbourhood of agent i or 0 otherwise.

Its value varies on a scale from -1 to 1, with -1 symbolizing a uniformly distributed (dispersed) pattern, 0 symbolizing a random pattern and 1 symbolizing a clustered (segregated) pattern.

We will compute I for the entire system as it converges to an equilibrium configuration and will compute I for several different sections of our system at equilibrium.

II.4. Model Parameters

The study of an agent based simulation requires testing how changes to the models parameters influence its behaviour. We chose a fixed value for the grid size $N = 50$ as it is big enough to show the models dynamics but small enough to have a reasonable computation time[1]. The remaining parameters were kept free with constraints given by table(II).

Parameter	Range
p_R	$]0, 1[$
ρ	$]0, 1[$
T_{min}	$]0, 1[$
T_{max}	$]T_{min}, 1[$

Table II: Description of the variables being used to describe the Schelling model.

III. Results

We now present the results obtained in this study. We start by showing three final equilibrium states of our system for different values of p_B in figure 7.

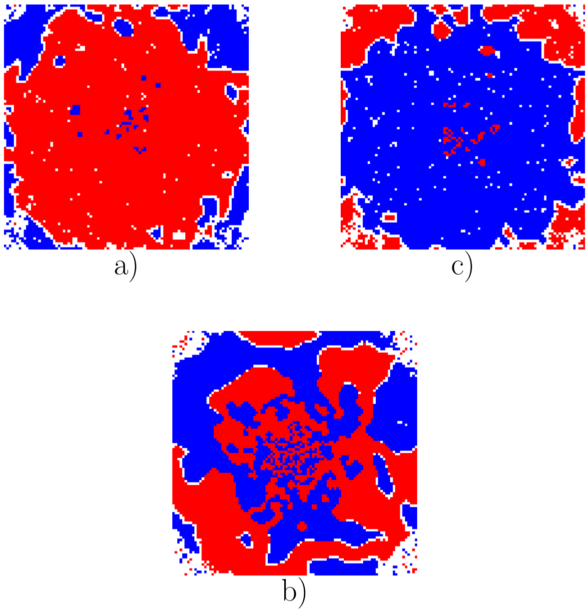


Figure 2: Equilibrium state of simulation runs with: **a)** $p_B = 0.2$; **b)** $p_B = 0.5$; **c)** $p_B = 0.8$. Parameters: $N = 100$; $\rho = 0.9$; $T_{max} = 0.9$, $T_{min} = 0.1$.

It is visible that in a) and c), figure 7, for $p_B \neq 50\%$, what tends to happen is the majority stays concentrated about the center of the grid and the minority gets "pushed" towards the edges. However, since the threshold value in the center is minute, small communities of the minority group emerge.

For $p_B = p_R = 50\%$, the pattern observed is heterogeneous. If one takes an r slice of the grid system and look at the groups occupying it, one would see the same amount of blue and red groups, each with similar sizes. However, since the tolerance of the agents increases towards the center of the grid, the groups formed in the center of the system are much smaller than those formed at the edges of the system.

III.1. Effect of Population Density

We begin by assessing whether population density ρ is a relevant parameter in changing the behaviour of the systems equilibrium configurations. It is found that the higher the density values the more iterations it takes to reach the equilibrium state - see figure 3.

[1]For certain plots, we used $N = 100$ to accentuate certain characteristics of the equilibrium distributions.

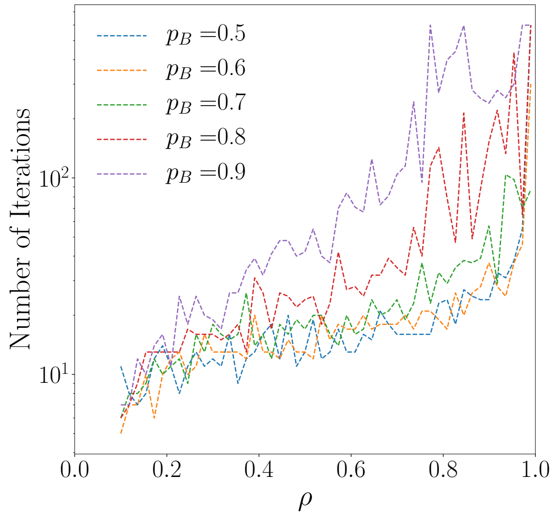


Figure 3: Influence of density ρ on the number of iterations required for the system to converge for different values of p_B. For each value of p_B and each value of ρ a simulation is ran and the number of iterations is plotted. Parameters: $N = 50$; $T_{max} = 0.9$; $T_{min} = 0.1$; Nb. runs per $p_B = 200$.

If the density is too low, agents of the same group will have more space available to maneuver into satisfactory configurations, leading to a quicker convergence. This maneuvering space decreases as the grid becomes more densely populated which increases the number of iterations required for the system to converge. Therefore a less densely populated lattice is favoured as it saves computation time.

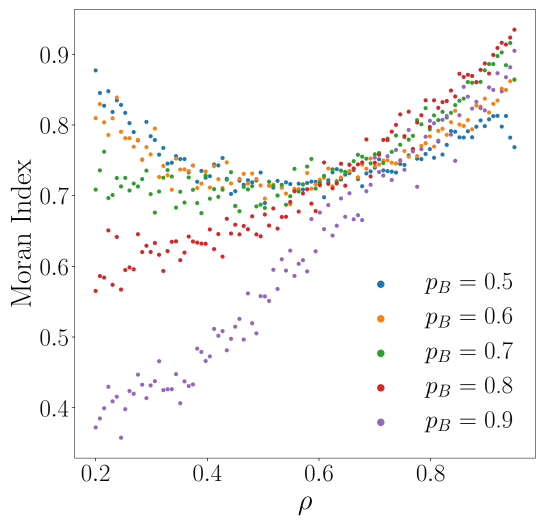


Figure 4: Influence of density ρ on the Moran index of the equilibrium system for different values of p_B. For each value of p_B and each value of ρ a simulation is ran and Moran index of the final matrix is computed and plotted. Parameters: $N = 100$; $T_{max} = 0.9$; $T_{min} = 0.1$; Nb. runs per $p_B = 100$.

We are interested in finding the minimal value of parameters that give rise to segregation and by making use of the Moran Index (figure 4) we find that for $\rho \gtrsim 0.7$, values obtained (for different p_B) are in the same range, growing linearly with ρ. Because of this, we chose to work with values of density superior to 0.7.

III.2. Size of the Majority

We will now study how the size of the majority affects the equilibrium state of our system. We'll start by calculating the average r of all agents of a certain group for different p_B.

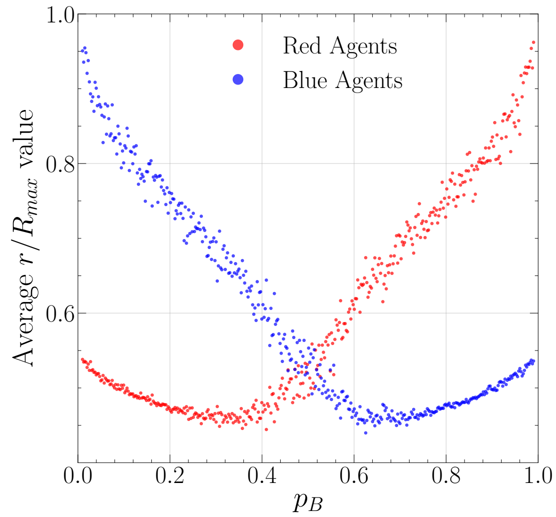


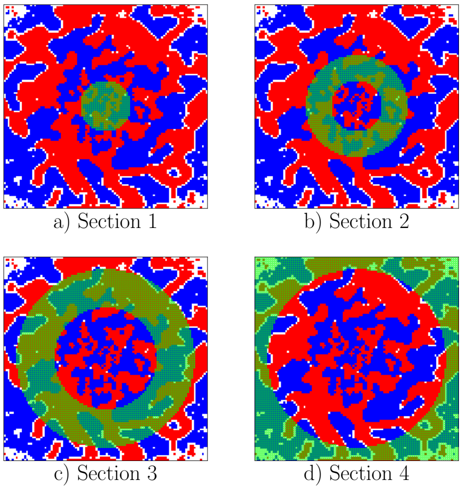
Figure 5: Influence of p_B on the average value of r. For each value of p_B a simulation is ran and the average values of groups **R** and **B** are computed and plotted for the equilibrium configuration. Parameters: $N = 50$; $\rho = 0.9$; $T_{max} = 0.9$, $T_{min} = 0.1$; Nb. of runs = 400.

Looking at the blue curve of figure 5, we see that, when the blue group is the minority ($p_B < 50\%$), $\bar{r}$ decreases approximately linearly with p_B, that is, that the minority tends to move towards the edges of the system and the smaller the minority population is, the more they move away from the center.

When the blue group is the majority ($p_B > 50\%$), for $p_B < 60\%$, we see once again that the average r decreases approximately linearly with p_B, which shows a tendency of the majority to populate the center of the grid. That would explain why the minority seems to populate the edges of the system: the majority prefers to populate the center and the minority is "pushed away".

However, for $p_B > 60\%$, we see that $\bar{r}$ increases (much more slowly and clearly not linearly) with p_B. This indicates another phenomenon is occurring and will be discussed ahead, in section III.3..

To study different locations of the grid, the grid is divided into four sections, consisting of the green sections highlighted in figure 6). As the tolerance depends only on r the sections are concentric rings with their center at the center of the grid, the first one being a circle and the last one extending itself to the edges of the grid.



a) Section 1 b) Section 2 c) Section 3 d) Section 4

Figure 6: Plots representing the circular ring-like sections highlighted in green used to study different locations of the system. Their radius are $r_1 = N/8$, $r_2 = N/4$ and $r_3 = 3.5N/8$.

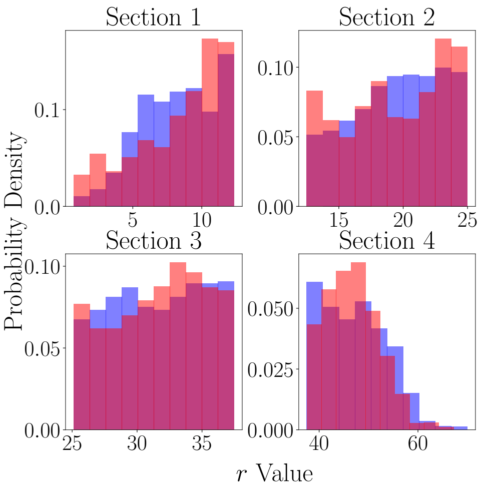


Figure 7: Histograms of red/blue agents' frequency by r value calculated for the sections of figure 6, using $p_B = 0.5$. Parameters: $N = 100$; $\rho = 0.9$; $T_{max} = 0.9$, $T_{min} = 0.1$.

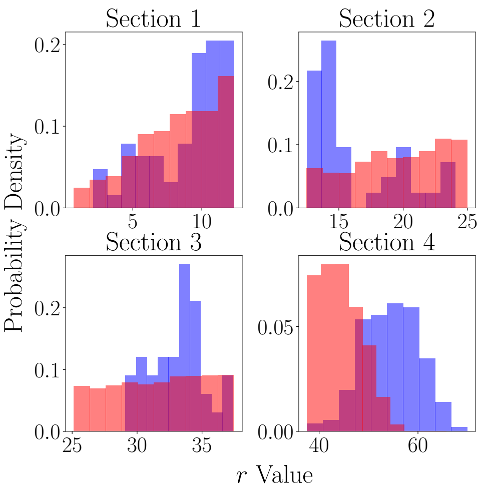


Figure 8: Histograms of red/blue agents' frequency by r value calculated for the sections of figure 6, using $p_B = 0.2$. Parameters: $N = 100$; $\rho = 0.9$; $T_{max} = 0.9$, $T_{min} = 0.1$.

Figure 7 displays the r value distribution for the heterogeneous system with $p_B = 0.5$ and figure 8 displays the r value distribution for a segregated system with $p_B = 0.2$. For the heterogeneous case it is noticeable that both groups follow the same distribution on every section as is expected, however this is not the case when $p_B = 0.2$. Here, the most internal region displays the same distribution but as one gets farther away from the center, the r distributions become differentiated culminating on two bell shaped curves with different $\bar{r}$ values.

III.3. Population Geometry

As stated in the previous section regarding figure 5, for $p_B > 60\%$ we can see that $\bar{r}$ increases slowly and not linearly with p_B. To understand what was occurring, we decided to examine the equilibrium configurations of the data.

Looking at some equilibrium configurations for $p_B \neq 50\%$, a pattern tends to arise about the center of the grid. The picture being painted by the data is that of a circle: the majority gathers around the center of the grid and shapes up approximately like a circle - see plots a) and c) of figure 7.

As such, it is of interest to study whether or not the increase on $\bar{r}$ with p_B is due to the increase in size of the quasi-circle created about the center of the grid.

For this, an empty grid is created and filled it with $N'_{\text{majority}} = N^2 \rho p_{\text{majority}}$ agents (that is, the same number of agents belonging to the majority in the standard grid). The grid is filled placing one agent at a time, making sure its r value was the smallest possible. Afterwards, the $\bar{r}$ for the system is calculated - see figure 9.

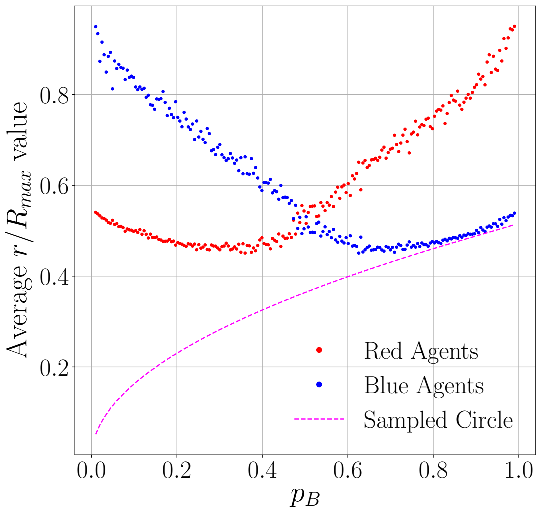


Figure 9: Influence of p_B on the average value of r VS. the curve obtained for a sampled circle about the center. For each value of p_B a simulation is ran and the average values of groups **R** and **B** are computed and plotted for the equilibrium configuration; additionally, a circle about the center of the grid is sampled with the appropriate amount of agents (based on p_B) and the value of $\bar{r}$ is computed and plotted. Parameters: $N = 50$; $\rho = 0.9$; $T_{max} = 0.9$, $T_{min} = 0.1$; Nb. of runs $= 400$.

As figure 9 shows, for $p_B > 80\%$, the majority (blue) curve seems to follow the sampled circle's curve. Additionally, the ratio between the sampled circle and the blue agents' curve was calculated for $p_B > 80\%$ and it was found to be approximately constant. This indicates that the increase in $\bar{r}$ with p_B for large p_B is due to the growth of the circular shape formed by the majority about the center of the grid.

III.4. The Impact of Distance

It is now possible to study the impact of our r-dependant tolerance the system's equilibrium configurations. For this, we'll use both metrics discussed previously: the average r value for each group and the Moran index computed at different locations in the grid.

To study the degree of segregation of our system in different locations, we calculated the Moran index for different sections of our grid, the sections described earlier and depicted in figure 6. In order to do this, we fixed all our parameters, ran the simulation many times, computed the Moran index for the different sections for each run and plotted the resulting data in an histogram, figure 10. Note that we chose quite different extreme threshold values ($T_{max} = 0.9$; $T_{min} = 0.1$) so as to to exacerbate the influence of the r-dependant tolerance on our system.

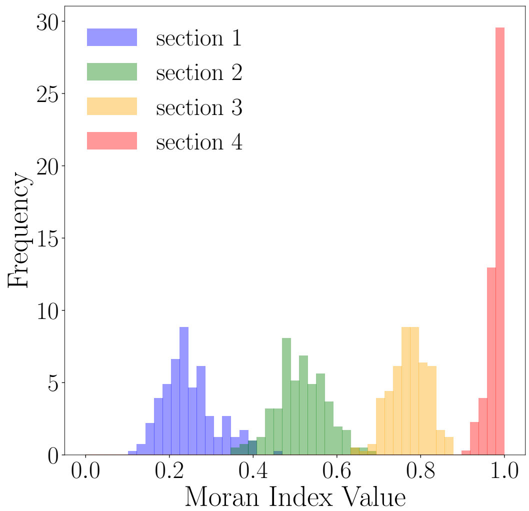


Figure 10: Influence of r-dependant tolerance on the Moran index value for different sections of the matrix (section 1 being the closest to the center and section 4 being the furthest from the center). The simulation was run 200 times and the Moran index was evaluated at different sections of the equilibrium matrix. Parameters: $N = 50$; $\rho = 0.8$; $P_B = 0.7$; $T_{max} = 0.9$; $T_{min} = 0.1$; Nb. of runs = 200.

The histogram shows that, the further a section is from the center, the higher their average Moran index is ($I_1 \approx 0.2$, $I_2 \approx 0.5$, $I_3 \approx 0.8$ and $I_4 \approx 1$). As described earlier, a Moran index of zero indicates a random distribution and an index of one indicates a clustered (segregated) distribution. As such, the further we are from the center, the more segregated our system becomes. This is expected, as the tolerance of our agents decreases with an increase in distance to the center.

We will now study the impact of the strength of our r-dependant tolerance (that is, the impact of the parameters T_{min} and T_{max}) on the equilibrium configuration of our system. For this, we'll assess how the average r of the majority and minority groups is affected by a change in the T_{max} parameter while keeping the T_{min} parameter constant. This is shown in figure 11.

Figure 11 presents a majority and a minority curve. It is expected that these curves complement each other: as we have a finite grid with high density, if the majority occupies a certain space, the minority will surely occupy the remaining space.

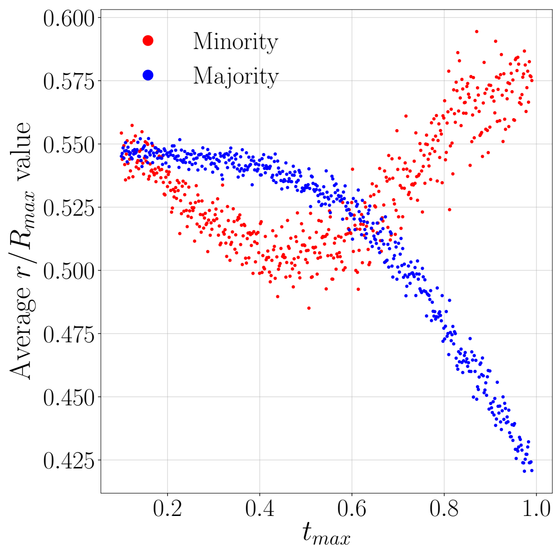


Figure 11: Influence of T_{max} (influence of r-dependant tolerance) on the average value of r. For each value of T_{max} a simulation is ran and the average values of groups **R** and **B** are computed and plotted for the equilibrium configuration. Parameters: $N = 100$; $\rho = 0.9$; $p_B = 0.7$; $T_{min} = 0.1$; Nb. of runs: 500.

We start by analyzing figure 11 through the majority curve. For $T_{max} < 40\%$, we can see that the average r for the majority remains approximately constant and $\approx R_{max}/2$. This is because, for such overall high tolerance (low threshold values), the system converges quickly to an approximately random configuration - see figure 12.

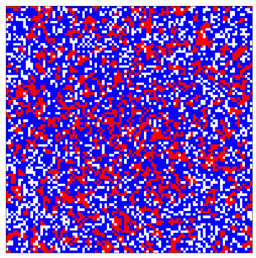


Figure 12: Equilibrium configuration for the parameters: $N = 100$; $\rho = 0.8$; $p_B = 0.7$; $T_{min} = 0.1$; $T_{max} = 0.3$.

For $T_{max} > 40\%$, $\bar{r}$ decreases with T_{max}, slowly at first, but rapidly for $T_{max} > 70\%$. This could be due to the fact that, since the agents near the edges get less tolerant, the majority tends to cluster together in a large circle about the center of the grid, reducing their average r.

Now looking at the minority curve, for $T_{max} > 40\%$ we see that $\bar{r}$ increases with T_{max} - as discussed before, as the majority makes a circle about the center of the grid, the minority is "pushed" away towards the edges. For $T_{max} < 40\%$, the average r decreases approximately linearly with T_{max}. This might be because, since T is overall small, it becomes quite easy for the system to converge, especially for the positions of the majority - as discussed before, the majority will form a random pattern (figure 12). However, when a tolerance gradient is present, no matter how small, there will still be a tendency for dissatisfied agents to move towards it. Since the majority will be satisfied with a random pattern, the minority will be able to move towards the center more easily, reducing its $\bar{r}$.

IV. Conclusion

The original Schelling model showed that a spatially-segregated society appeared naturally from individual biases and preferences of its constituents. Similarly, the altered Schelling model suggested shows the emergence of a specific structure in the society's spatial distribution, stemming from a city center preference. For a strong enough preference to inhabit the center, the majority will naturally occupy a circle about the center of the system and the minority will naturally be pushed away the edges of the city. This is shown by the polarization of the average distance to the center of each one of the groups.

Today's cities are vast and filled with opportunities, but most of and the best opportunities are found in its center so the emergence of this structured is troublesome for the minority, who won't have access to the same amount of opportunities.

An effort has to be put in place to contradict this structure in order to achieve a more heterogeneously distributed and fairer society.

As further research we suggest studying the effects of implementing quotas, where there has to be a certain percentage of the minority group inhabiting the more economically advantageous regions and the implementation of income distribution on both groups.

References

[1] Brown, Nina. *Robert Park and Ernest Burgess: Urban Ecology Studies, 1925.* Center for Spatially Integrated Social Sciences.

[2] Encyclopaedia Britannica, *Human Ecology*, `https://www.britannica.com/topic/human-ecology`, [Accessed at 2021-11-05].

[3] Park, Robert, Ernest W. Burgess and Roderick D. McKenzie. The City. Chicago: University of Chicago Press, 1925

[4] McCown, F., *Schelling's Model of Segregation* [Online]. Available from: `http://nifty.stanford.edu/2014/mccown-schelling-model-segregation/` [Accessed at 08/11/2021].

[5] Schelling, T.C., *Dynamic models of segregation.* Journal of Mathematical Sociology, 1971. **1**(2): p. 143-186. Available from: `http://norsemathology.org/longa/classes/stuff/DynamicModelsOfSegregation.pdf` [Accessed at 08/11/2021].

[6] Urselmans, L. and S. Phelps, *A Schelling model with adaptive tolerance.* PloS One, 2018. **13**(3): p. e0193950. Available from: `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0193950` [Accessed at 08/11/2021].