POSSIBILITIES IN THE SCHELLING MODEL *

M. VAN ZEE

Department of Information and Computing Sciences
e-mail: m.vanzee@students.uu.nl
Utrecht University

Contents

* This mini-thesis is written as the final deliverable for the master course Adaptive Agents given by Gerard Vreeswijk. Several revisions have been made after receiving feedback from the concept version halfway the course. Therefore, the last page of this document contains a list of revisions, while this is not intended to be part of the article.

Abstract

This minithesis presents Schelling's [13] model of segregation and a variation on it. By moving dimes and pennies on a checkerboard, Schelling demonstrated that segregation emerges and persists even if every person in the society possesses equal preferences for both racial groups in its neighborhood. After analyzing this informal approach, we will proceed by discussing a formal extension to the model. The dynamics of segregation will be explored using the model by Zhang [18] that makes use of techniques recently developed in evolutionary game theory to determine stochastical stable states within the model. We will discuss several aspects of this technique, originally proposed by Young, in more detail. Finally the conclusion proposes several points of further research.

Keywords

Residential segregation; Local interactions; Global behavior; Markov chains; Random pertubations; Stochastic stability; Agent-based simulation.

1. Introduction

During the 1960s the economist Thomas Schelling researched segregation and racial preferences. While the major metropolitan areas showed very strong segregation between white and black people, the individual attitudes did not seem to match with it. There seemed to be more segregation than was actually preferred. Schelling's intuition is clearly reflected in a survey that has been done on this subject in 1982 [4]. The General Social Survey asked both black and white people about their preference for the racial coloring of their neighborhoods. It turned out that 55.3% preferred to live in a mixed neighborhood, so half-black, half-white. In 1990 and 1996 a similar question was asked, and 60% did not oppose living in a fully integrated neighborhood meaning they either responded "do not care"; "favor" or "strongly favor".

Table 1: *Percentage of blacks who live in a half-half neighborhood*

Metro area	Neighborhoods in sample	Half-half neighborhoods	Percentage (%)
Baltimore	581	25	3.80
Buffalo	289	12	2.49
Chicago	1885	71	2.91
Cleveland	829	34	2.97
Detroit	1273	40	2.56
Milwaukee	428	16	2.74
St. Louis	456	20	4.24
Washington	911	63	5.96

These results contrast sharply with the real world society. As shown in Table 1, the percentage of blacks living in a half-half neighborhood is far below 50%, actually even below 10%.

It is clear that there is a strong difference between the motives of an individual and the behavior of the system. It seems as if the distribution of the races within the population is very segregated. This is what the model of segregation of Schelling is about, why this *"distribution is so U-shaped that it is virtually a choice of two extremes"*. While Schelling his ultimate concern is segregation by color in the United States, he admits that because of the level of abstraction of his model any twofold distinction could constitute an interpretation. Schelling has discussed his model in three different publications. In the first article a one-dimensional model of segregation is discussed [12]. Later, this is extended to a two-dimensional model including a more in-depth analysis [13]. Finally, this model is presented in chapter 4 of the book *"Micromotives and macrobehavior"*, where Schelling elaborates on the relation between the motives of an individual and the global behavior of the system [11]. I will only consider the model of segregation on a two-dimensional plane, relying mostly on the second publication [13]. In this version the neighborhood preferences are more worked out[1], and it is the best-known version.

2. Defining segregation

The model of Schelling is about the kind of segregation that can result from differences in individual behavior. This can be in religion, age, color, etc. It examines some of the individual incentives that can lead collectively to segregation. Furthermore it examines to what extend this individual behavior can differ for segregation to occur. By taking this approach, Schelling admits that not all processes of segregation are covered:

- Organized (*e.g.* governmental) action can also lead to segregation, or prevent it. A recent variation on the model by O'Sullivan [1] extends the Schelling model using public policy, but we will not discuss it in this article;

[1] In the model of 1971 an individual always wants a percentage of its neighbors to be of its own color, while in the model of 1978 this depends on the number of neighbors.

- The socioeconomic differentials between whites and nonwhites can lead to segregation. Evidently color is correlated with income, and income with residence; so even if residential choices were color-blind, whites and blacks would not be randomly distributed among residences. This is also referred to as *economically induced segregation*. A publication by Sethi and Somanthan [15] explores the effect of income on segregation.

While these two points might be of greater relevance than the individual choice, Schelling argues that in a matter as important as racial segregation, even a third place deserves attention.

3. Separating mechanisms

From now on I will use the terms *whites* and *blacks* when I am referring to the two different races within a certain population. Before actually going to the model, it is worth seeing how much we can find out by simple reasoning. Whites and blacks might not mind each other's presence, in fact they might even prefer integration, but they nevertheless might want to avoid being a minority. If white and blacks neither want to be a minority, then except of a mixture of 50%-50% one group will always be dissatisfied. We can imagine that as soon as an individual of this minority group leaves the neighborhood, the remaining ones of the minority group will even be more dissatisfied, which will in the long run cause all individuals of this color to leave the neighborhood. Hereafter, it is unlikely that an individual from one color will settle in a neighborhood that is completely populated by the other color. Moreover, when an individual relocates it is likely that his color will be a majority in the new neighborhood. The relocation will only increase this majority, which will make the minority group more dissatisfied with the new situation. Putting these facts together, we can assume that in the long run complete segregation occurs. This simple observation allows us to reintepret the results of the enquiry as described in the introduction, and make us realize that these results are not that striking, since it is logical that segregation occurs when two groups want to avoid being a minority.

If we define the neighborhood preferences as a range instead of a fixed number, the limits of this range will define the mixtures that will survive. For example, if both blacks and whites can stand a 25% minority, the initial mixtures ranging from 25% to 75% will survive, but mixtures more extreme will all become of one color. Also, those who leave this extreme mixture might arrive in a neighborhood that increases its majority there and cause the other color to evacuate.

Finally it is to be noted that complete segregation is a stable equilibrium; nobody has an incentive to move, and it is persistent against small perturbations. Once the segregation has completed, it seems like the process is irreversible.

4. Schelling's spatial model

Schelling presents his results in an informal way. This is no doubt the strength of the model, or as Schelling states it: *"they (*the results*) are crude and abstract but have the advantage that anyone can reproduce them using materials that are readily available"*. I will therefore not do this any harm by making it more complicated; I will leave this for following sections. If this section might appear to be a bit too trivial from time to time; try to imagine that this article was written in 1971 and was one of the first attempts into this way of thinking.

Assume a population more of less evenly divided into two groups, represented as stars and zeros, randomly spread out over a 8×8 grid (Figure 1, left image) (these are the settings in the model of 1978 and exactly the dimensions of a chessboard, in the model of 1971 Schelling used a 13x18 grid. It is clear that he wanted to persuade the reader to start playing with it, for which he deserves credits). A person can only move into a vacant space and when he moves he leaves a space vacant. In order that people be able to move there must be some vacant spaces; it turns out that 25%-30% vacancies is a good compromise.

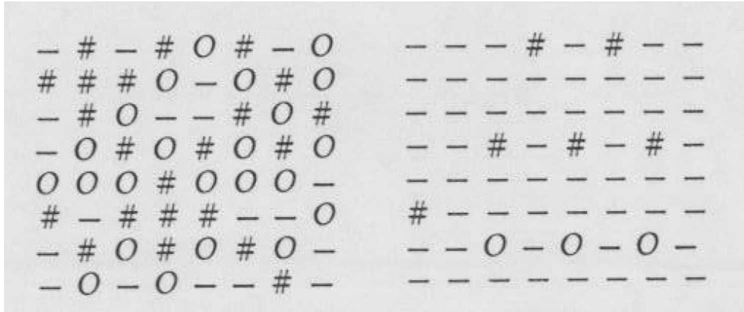

Figure 1: *Schelling model, initial state*

Everybody only cares about his neighborhood, which is defined as the eight surrounding squares (the so-called *Moore neighborhood*, see Figure 4). The rule of movement is that an individual discontent with his own neighborhood moves to the nearest vacant spot that surrounds him with a neighborhood that meets his demands. Furthermore it is worth to mention that because all experiments have been done by hand and eye, no strict rule for the order of moves has been used. According to Schelling: *"the particular outcome will depend very much on the order in which discontented stars and zero are moved, the character of the outcome not very much"*.

Table 2: *Schelling model, preferences of an individual*

Total amount of neighbors	Number of neighbors of the same color
1	1
2	1
3,4,5	2
6,7,8	3

Suppose an individual is satisfied when the preferences in Table 2 are met, meaning that he is happy when the amount of neighbors of the same color is at least the given number, for a certain amount of total neighbors. The pattern in the left image of Figure 1 does not look strongly segregated. The right image shows the individuals who are dissatisfied with their current position, which is only a fraction (9 out of 45). The problem is where to move them. Anybody who moves leaves a blank cell that somebody can move into. Also, moving will most likely leave behind several neighbors of the color of the moving agent, who may become discontent after the move. In the new neighborhood the amount of agents of the color of the moving agent will increase, possibly dissatisfying someone of the other color.

Schelling calls this the "unraveling" process. Everybody who selects a new environment affects the environments of those he leaves and those he moves amoung. There is a chain reaction.

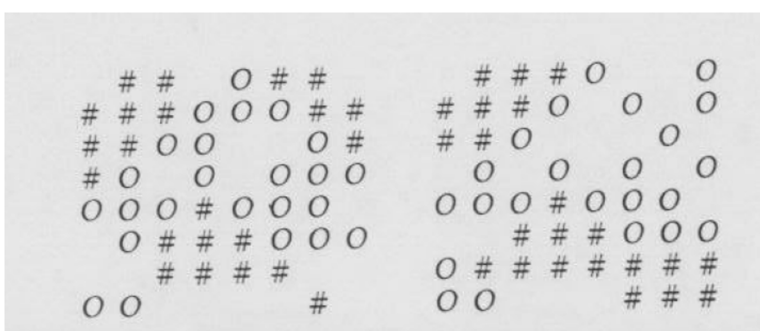

Figure 2: *Schelling model, two outcomes with different ordering*

Two possible outcomes (with different orderings) are depicted in Figure 2. They look more segregated and this can be made clear by making a comparison as well. Before the segregation, only 9 individuals where not satisfied, and the rest was satisfied in accordance to their demand in Table 2. After the segregation in Figure 2, on the left side image the average neighbors that every individual has of its own color is about triple as much as they demand according to Table 2, and on the right image it is even four times are much!

What can we conclude from this? Schelling is quite modest and states: *"(...) that observable aggregate phenomena could be compatible with types of 'molecular movement' that do not closely resemble the aggregate outcomes that they determine"*. I would say: that small demands by individuals might lead to extrapolated outcomes for a group.

Now that we have a good idea of what the model of Schelling entails, we will look at an extension. The goal is to, rather than the inductive character of Schelling his model, find a mathematical model that will rigorously prove that segregation will occur.

5. Towards a mathematical model

Recently the mathematical model is being increasingly used as a basis for simulations on a computer. It allows for models to be explored and played with in a way that previously would have only been possible for skilled mathematicians [5].

Edmonds and Hales argue that a simulation should be seen as a formal model of intermediate generality, one which needs to be treated not as an analytic model but more like a partly understood phenomena. The simulation does encode a theory which, along with an interpretation can be used to represent some aspects of social phenomena. However, the nature of this theory is one that is only accessible via the experimentation of running the simulation. To use a simulation as a model of observed phenomena one also needs a theory of how the simulation works.

In the case of the Schelling model, a mathematical model might provide some insights that can help us better understand the real world phenomenon of segregation. Before we move on to this it is worth looking a at the article by Edmond & Hales in more detail.

In their paper, Edmond and Hales state that simulations need to be replicated before they can start to be trusted, and that one should take a critical standpoint towards the specifications of a simulation. A model cannot be considered on its own but needs the context of the intended interpretation to make sense.

I agree with this point of view, were it not that the example (the Schelling model) that Edmond and Hales provide to prove their points is incorrect. First, they refer to the first article (Schelling, 1969) and reproduce a two-dimensional space, while the article that they refer to only considers a one-dimensional game. Next, a variable c is used to define the percentage of the neighborhood that an individual prefers to be of the same color, following Schelling (1971) his approach. This variable is then plotted for values between 0 and 1 (Figure 3).

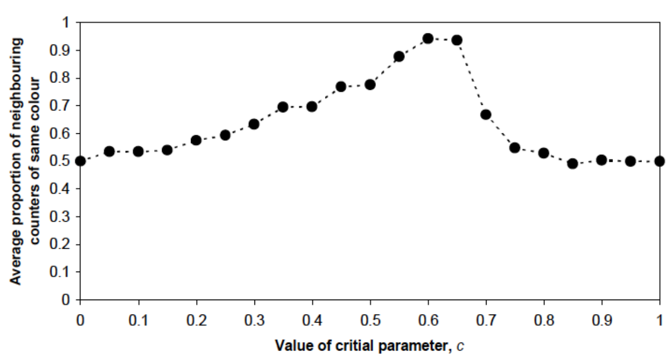

Figure 3: *Segregation resulting from different levels of intolerance*

From this it is concluded that the Schelling model demonstrates an initially counter intuitive result because segregation actually decreases for $c > 0.65$, while Schelling did not intend this to happen (it should continue to increase). According to the authors, the reason why the simulation shows this behavior is because movement of individuals is defined as moving to a randomly chosen free square within the neighbourhood. This is incorrect. In the article that the authors refer to, Schelling literally writes: *"If fewer than half are his color, he moves in either direction to the nearest point (...) at which half his eight nearest neighbors are the same color as he"*. He does not use the concept of random movement at any point. All the other articles of Schelling state similar rules. The authors, while stressing to be critical, made errors while reproducing the model that should prove their point.

6. Extending the model

Even though the finding of Schelling were presented informally, they have been influential in different disciplines of social sciences. His work inspired many "tipping point" models and agent-based simulations (e.g. [6]). It is even said to be the basis of a famous book called "The tipping point, how little things can make a big difference" [7], although this is refuted by the author[2].

Given the model of Schelling, the first question that arises is how individual decisions of agents may affect the macro behavior of the system; and the second, how these decisions may organize themselves in space, in other words, what kind of spatial patterns may be observed on the global level [2]. However, because proper analytical tools were unavailable, there was no important development following Schelling's original model.

Although it was Young [16] to first mention that techniques developed in stochastic dynamical systems can be used to analyze the Schelling model, it was one of his students called Zhang who actually attempted to do this [18]. I will now describe this attempt, which uses a variant of Schelling his model.

Consider an $N \times N$ lattice graph embedded on a torus. V is the set of vertices, and every $v \in V$ is occupied by either a black or a white agent. Notice that this implies there are no vacant spaces in the model, which makes patterning (*i.e.* defining the borders of a seggregated area) more easy because there can be no ambiguity in to what color a vacant space belongs.

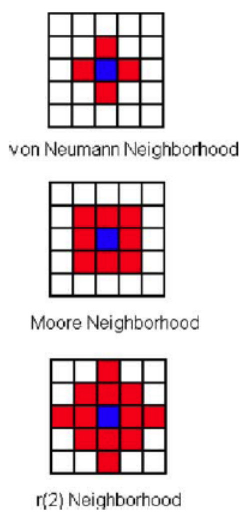

Figure 4: *Neighborhood definition*

[2]For an interview on this subject by the author, see: http://www.nypress.com/print-article-18144-print.html

We are using a Moore Neighborhood (Figure 4), so any agent considers 8 agents around his as his neighborhood. Let E be the collection of all pairs (i, j), where agent i and j are neighboring agents:

$$E = \{(i, j) | i \text{ and } j \text{ are neighboring agents}\} \tag{1}$$

An agent receives a payoff in every round that can be interpreted as how much he or she likes to pay for a certain location. In other words: the value that an agent ascribes to a certain location. The payoff consists of two parts: a deterministic term u that depends on the amount of like neighbors in the neighborhood, and a random term ϵ.

The deterministic term is assumed to be a kinked curve (Figure 5). As we can see, every agent prefers to have half of his neighborhood of his own color. This implies that if the neighborhood size is 1, there are 4 like neighbors and 4 not-like neighbors. If we also count the agent himself to the neighborhood then it consists of 5 agents of his own color and 4 not. In conclusion: although the agent prefers integration, this optimal state will result in majority in the neighborhood of its color. Notice that the line on the left side of n is relatively steeper than the one on the right side, which implies that an agent rather lives with a majority of his own color than a majority of the other color. This doesn't sound like an unrealistic claim. Note that Z and M are constants, and x stands for the number of like-color neighbors an agent has.

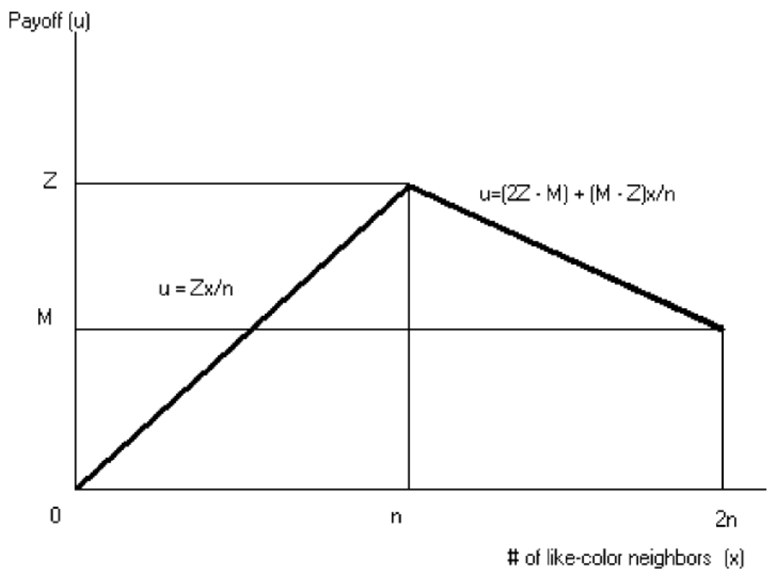

Figure 5: *Payoff profile*

We can write the deterministic part u of an agent's payoff function as:

$$u = \begin{cases} \dfrac{Zx}{n} & \text{if } x \leq n, \\ (2Z - M) + \dfrac{(M - Z)x}{n} & \text{otherwise,} \end{cases} \qquad Z > M > 0. \tag{2}$$

And an agent's total payoff is written as

$$\beta u(\cdot) + \epsilon,$$

where β is a positive constant determining the relative importance of the random term. Every agent has the same β. Because there are no vacant spaces, agents can move by switching location. In every round, two random agents are chosen. If the mutual expected payoff after the switch exeeds the sum of their current payoff, a switch will occur. We denote the payoff for switching locations for an agent i as $u_i(\cdot|\text{switch})$ and for not switching as $u_i(\cdot|\text{not switch})$. The sum of the mutual payoff of two picked agents that do not switch is

$$\begin{aligned}&(\beta u_1(\cdot|\text{not switch}) + \epsilon_1) + (\beta u_2(\cdot|\text{not switch}) + \epsilon_2)\\ &= \beta[u_1(\cdot|\text{not switch}) + u_2(\cdot|\text{not switch})] + (\epsilon_1 + \epsilon_2) = \beta U + \eta \end{aligned} \tag{3}$$

The sum of the payoffs when two picked agents do switch locations is

$$\begin{aligned}&(\beta u_1(\cdot|\text{switch}) + \epsilon_1') + (\beta u_2(\cdot|\text{switch}) + \epsilon_2')\\ &= \beta[u_1(\cdot|\text{switch}) + u_2(\cdot|\text{switch})] + (\epsilon_1' + \epsilon_2'\} = \beta V + \xi \end{aligned} \tag{4}$$

Thus a switch will happen only if $\beta U + \eta < \beta V + \xi$. We are able to derive a rule that can determine the probability a switch will occur. For this we use the theorem of Holman and Marley to derive the choice probabilities using the multinominal logit [10][3]. The theorem of Holman and Marley is as follows:

Theorem 1. (Holman and Marley) *Assume that the ϵ_i are i.i.d. according to the double exponential distribution*

$$F(x) = \Pr(\epsilon_i \le x) = \exp - \left[\exp -(\frac{x}{\mu} + \gamma)\right], \tag{5}$$

where γ is Euler's constant ($\gamma \approx 0.5772$) and μ is a positive constant. Then the resulting choice probabilities are given by

$$P_A(i) = \frac{\exp(u_i/\mu)}{\sum_{j=1}^{n} \exp(u_j/\mu)}, i = 1 \cdots n. \tag{6}$$

If we want to apply this theorem to our model, we first need to be sure that the assumptions are consistent. Therefore we assume that our random variables η and ξ are i.i.d. distributed according to the double exponential function as described in the theorem. To simplify things, Zhang assumes that the constant γ is set to 0 and μ to 1. Thus, the resulting cumulative distribution function is

$$F(x) = exp(-e^{-x}). \tag{7}$$

To derive the switch probability, we need to know the density function that corresponds to the distribution function we have chosen. The density function (or as Zhang calls it: the probability distribution function) is a function that describes the relative likelihood for a random variable to occur at a given point. Let X be the domain of all possible choices for x, and let x_i be an arbitrary element in this domain. The density function $f(x_i)$ will give the probability that x_i will occur. If we sum up all the probabilities for all values of x up till x_i, we have the cumulative distribution $F(x_i)$. The relation between the density function and the cumulative distribution function is thus that the first is the derivative of the latter. The density function corresponding to (5) is given by [10]

$$f(x) = \frac{1}{\mu}\left[\exp -(\frac{x}{\mu} + \gamma)\right]\left\{\exp - \left[exp - \left(\frac{x}{\mu} + \gamma\right)\right]\right\}. \tag{8}$$

Applying (8) to our model results in

$$\begin{aligned} f(x) &= \frac{1}{1}\left[\exp -(\frac{x}{1} + 0)\right]\left\{\exp - \left[exp - \left(\frac{x}{1} + 0\right)\right]\right\}\\ &= (\exp -x) \cdot (\exp -[\exp -(x)]\\ &= (\exp(-x - e^{-x}) \end{aligned} \tag{9}$$

which corresponds with the probability distribution function as given by Zhang.

[3]This is a different approach that the one taken in the paper, where the method developed by McFadden [9] is used. The method of McFadden uses slightly more restrictive hypotheses, but gives similar results.

The probability that a switch will occur can be rewritten to

$$\mathrm{Pr}(\text{switch}) = \mathrm{Pr}(\beta U + \eta < \beta V + \xi) = \mathrm{Pr}(\eta < \beta V - \beta U + \xi). \tag{10}$$

Equation (2.25) of [10] states: Assume that the ϵ_i are i.i.d. Let F be the common cumulative distribution function of the ϵ_i, and let f be the corresponding density. For any given realization x of ϵ_i, alternative i will be chosen with probability density $\prod_{j\neq i} F(u_i - u_j + x)$. Accounting for all possible realizations, we have

$$P_A(i) = \int_{-\infty}^{\infty} f(x) \prod_{j\neq i} F(u_i - u_j + x)\mathrm{dx}. \tag{11}$$

We simplify Equation (11) to account for only two choices (namely, *switch* and *not switch*), this results in

$$P_A(i) = \int_{-\infty}^{\infty} f(x)F(u_i - u_j + x)dx, \tag{12}$$

Where u_i and u_j represent the measured utility. Filling (7), (9) and (10) into (12) gives

$$\begin{aligned}
P_A(switch) &= \int_{-\infty}^{\infty} f(\xi) \cdot F(\beta V - \beta U + \xi)\mathrm{d}\xi \\
&= \int_{-\infty}^{\infty} \exp(-\xi - e^{-\xi}) \cdot \exp(-e^{-\beta V + \beta U - \xi}) \cdot \mathrm{d}\xi \\
&= \int_{-\infty}^{\infty} \left[-\xi - e^{-\xi}\left(\frac{e^{\beta U} + e^{\beta V}}{e^{\beta V}}\right)\right] \mathrm{d}\xi \\
&= \int_{-\infty}^{\infty} [-\xi - e^{-\xi}e^{\phi}]\mathrm{d}\xi \\
&= \exp(-\phi) \cdot \int_{-\infty}^{\infty} \exp[-(\xi - \phi) - e^{-(\xi-\phi)}]\mathrm{d}(\xi - \phi) \\
&= \exp(-\phi) \cdot \int_{-\infty}^{\infty} f(\xi - \phi)\mathrm{d}(\xi - \phi) \\
&= \exp(-\phi) \cdot 1 = \frac{\exp(\beta V)}{\exp(\beta U) + \exp(\beta V)}, \qquad \text{where } \phi \equiv \ln(\frac{\exp(\beta U) + \exp(\beta V)}{\exp(\beta V)}
\end{aligned}$$

which is equivalent to Equation (6), since it was assumed that $\mu = 1$. Filling the values of U and V from (3) and (4) into this equation gives

$$= \frac{e^{\beta[u_1(\cdot|\text{switch}) + u_2(\cdot|\text{switch})]}}{e^{\beta[u_1(\cdot|\text{not switch}) + u_2(\cdot|\text{not switch})} + e^{\beta[u_1(\cdot|\text{switch}) + u_2(\cdot|\text{switch})]}}, \beta > 0. \tag{13}$$

We call this rule the log-linear switch rule. This kind of behavioral rule is commonly used in the literature (*e.g.* [3], [16]). Although the derivation is quite complex, what it simply says is that if a switch increases the expected payoff of an agent, it is more likely to occur. Note that it is still possible that a switch will not occur, or that a switch will occur even though it decreases the payoff of both agents.

Now that we have acquired a switch rule, we are able to turn this model into a Markov process. The next section will explain this process in more detail.

7. Interpretation as a Markov process

There are several advantages of adding a random element to a model [17]:

- It plays a role similar to mutations in biology by injecting variability into agents' behavior. This variability can also cause populations to "tip" over;
- The persence of perturbations implies that the evolutionary dynamic never settles down completely. This proves to be a useful analytical tool for analyzing the long-run behavior.

The reason why the random element was added to the model of Zhang is because of the second point. It is possible to see the model with the switch rule as a Markov process of very large dimensionality. A Markov process consists of states and the transistion probabilities between these states. A state of the system is specified by the location of the agents on the lattice. The transition probabilities are specified by the probability that a switch will occur between two agents, which will change the state of two cells on the lattice.

More formally: let A_N denote the $N \times N$ lattice of the model. A state $x \in X$ is then defined as a function $x : A_N \rightarrow \{black, white\}$, which allocates a color to each location, and x^t represents the state at time t. The Markov process is denoted as a transition probability matrix P^{β}, where β represents the importance of the perturbation. P_{xy} is the probability of moving from state x to y in one period. This Markov process holds several characteristics; we will go through them using the article of Young [17]:

- *Perturbed process:* agents do not always make "correct" decisions, depending on the value of β. Small values of β imply big perturbations but when β approaches infinity the process becomes almost completely deterministic;
- *Irreducible:* there is a positive possibility to move from any state to another in a finite period. This is because the process is perturbated; it ensures that it is theoretically possible (although this change may be very small) to reach every state. An example of an irreducible Markov process is given in Figure 6. The dashed red transitions represent perturbed transitions which ensure that every state be reached.

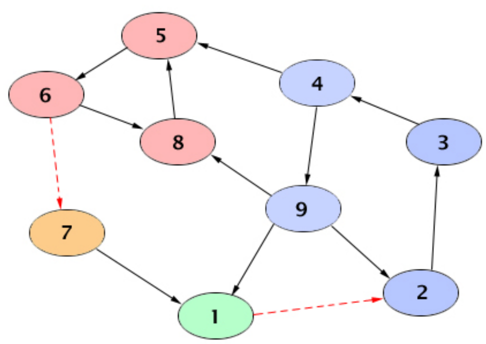

Figure 6: *Irreducible Markov process by perturbation.*

- *Ergodic:* this characteristic is not mentioned in the article of Zhang, but it is easy to derive. Ergodicity holds if, and only if, the process has a unique recurrence class. A recurrence class implies that every state in the class is recurrent. A state is recurrent if and only if, once the process has entered it, the probability of returning to it is one. Otherwise, the process is transient. Figure 7 gives an overview of both types of classes. Once the process enters the recurrence class, it is impossible to leave it. Equivalently, a process with one unique recurrence class is ergodic if and only if the states can be divided into two disjoint classes A and B such that 1) there is a positive probability of moving from any state in A to some state in B; 2) there is a positive probability of moving from any state B to another state in B; 3) there is zero probability of moving from any state in B to any state in A. An irreducible process is a particular instance of this class where A is empty and B constitutes the entire state space (Figure 6).

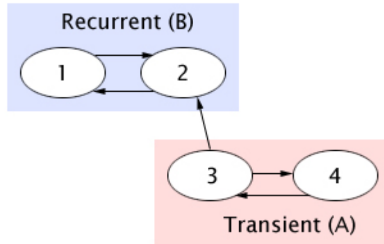

Figure 7: *Recurrence and transient classes.*

- *Aperiodic:* besides irreducibility we need a second property of the transition probabilities, namely the so-called aperiodicity, in order to characterize the ergodicity of a Markov chain in a simple way. Zhang and Young both do not elaborate on this notion, but I believe it is insightful to explain this in more detail to gain more understanding in the process. What follows is a prove that the period of every state in an irreducible Markov process is equal.[4].

The period d_x of the state $x \in X$ in the Markov chain is given by $d_x = \gcd(\geq 1 : P^n_{xy} > 0)$ where *gcd* denotes the greatest common divisor. In other words: the period of a state is the largest number that divides all the possible ways in which a state can return to itself. A state $x \in X$ is said to be *aperiodic* if $d_x = 1$. The Markov process is called aperiodic if all states are aperiodic. I will now show that the periods d_x and d_y coincide if the states x and y are both in the same recurrence class. For this purpose we introduce the notation $x \to y[n]$ if $P^n_{xy} > 0$, which means that the process can travel from state x to state y in n steps. Before we turn to the prove, we need two inequalities which are derived from the *Chapman-Kolmogorov equation* [8]. For arbitrary $n, m, r = 0, 1, ...$ and $x, y, z \in X$ it holds that:

$$P^{n+m}_{xx} \geq P^n_{xy} P^m_{yx} \tag{14}$$

$$P^{r+n+m}_{xy} \geq P^r_{xz} P^n_{zz} P^m_{zy}. \tag{15}$$

Theorem 2. *If the states $x, y \in X$ both belong to the same recurrence class, then $d_x = d_y$.*

Proof. If $x \to x[n], x \to y[k]$ and $y \to x[m]$ for certain $k, m, n \geq 1$, then the inequalities of (14) and (15) imply that $x \to x[k+m]$ and $x \to x[k+m+n]$. Thus, $k+m$ and $k+m+n$ are dividable by d_x (because d_x is the greatest common divisor of all possible paths from x to x). As a consequence the difference $n = (k+m+n) - (k+m)$ is also divisible by d_x. This shows that d_x is a common divisor for all natural numbers n having the property that $P^n_{xy} > 0$, i.e. $d_x \leq d_y$. For reasons of symmetry the same argument also proves that $d_y \leq d_x$ and thus we can conclude that $d_x = d_y$. □

From this we can conclude that all states of an irreducible Markov process have the same period. This is consistent with our findings, since in every state there is a positive probability that we will remain in the same state (no switch will occur). This means that it is possible to return to the current state in one step, which implies that the period of every state is 1.

The standard approach to analyse the behavior of a Markov process (to find out what the long-term behavior will result in) is to solve the stationary distribution algebraically. Let μ be a probability distribution on X and consider the system of linear equations:

$$\mu X = \mu, \text{where } \mu \geq 0 \text{ and } \sum_{x \in X} \mu(x) = 1. \tag{16}$$

For the process we just created (an ergodic Markov process), this solution always has exactly one solution. Because this process is both ergodic and aperiodic, not only does its *average* behavior converge to the unique stationary distribution μ^β, so does its probabilistic behavior *at each point*

[4]I have used the lecture notes of Schmidt [14] as a reference.

in time t when t is sufficiently large [17]. In other words: it does not matter at what point in time we are looking; if t is sufficiently large, $\mu^{\beta}(x)$ is the probability that state x will be observed most of the time. Now we are ready for the definition of a *stochastically stable state*:

Definition 1. (Foster and Young, 1990): A state $x \in X$ is *stochastically stable* relative to the perturbed process if $\lim_{\beta\to\infty} \mu^{\beta}(x) > 0$. The *stochastically stable set* is the smallest set that contains all the stochastically stable states.

What this actually means is that a stochastically stable set will have a positive probability to be visited in the long run. Therefore it will be observed much more frequently than a state that is not stochastically stable. As $\beta \to \infty$ and $t \to \infty$, it is very likely that the system is to be found in the stochastically stable set.

Looking back, we have turned our model into a Markov process with the characteristics to predict the long-run behavior of the system, but we have no idea yet what this behavior will be. We have no way of measuring the behavior of the system, for we have no indicator that defines the state of the system as a whole.

8. The long-run behavior

Zhang introduces a function ρ which is a measure for the segregation of the population. The idea is to count the total number of neighbors that are of different color. Recall from (1) that we used the set E to store the neighboring agents in. We define a set ED (Edges Different) as

$$ED = \{(i, j) \in E | i \text{ and } j \text{ have different colors}\}. \tag{17}$$

The function ρ then contains the number of neighbors of different color:

$$\rho = |ED|. \tag{18}$$

This function measures the degree of exposure (degree of potential contact) between the members of the two groups. It has another interesting and useful property, namely that it is a *potential function* ρ for this spatial game[5]. A game is a potential game if the changes in every player's payoff can be characterized by the first difference of a function. On other words: every time the payoffs of the players change, the potential function should also change.

Actually, $-\rho$ is a potential function of this game, because utility-improving switches for agents will always reduce the value of ρ (and so increase the value of $-\rho$). To state it differently: if a black and white agent gain from trading, ρ will decrease because this will always result in more like-color neighbors. This fact is also intuitively clear: a white agent will be happier moving from an 80% white neighborhood to a 60% white neighborhood. However, this agent will never be able to find a black trading partner because the black agent has to move from a 40% black neighborhood to a 20% black neighborhood. The white agent cannot compensate this loss of satisfaction because of the asymmetry in Figure 6. Function ρ decreases as the switch leaves fewer black-white neighboring pairs in the society, reflecting the loss of utilities in the whole population as the society moves a step further toward segregation.

Now that we have a way to describe the dynamics of the system, we can combine this with the previously described Markov process to do a prediction on the long-run behavior.

[5] I will not go into the proof of the validity of the potential function, but the main idea is that utility-improving switches will *always* reduce the total number of black-white neighboring pairs. The proof can be found in the article of Zhang.

Recall that we defined X as the set of all states of the Markov process. Define S as a set of states that minimize the value of the potential function ρ:

$$S = \{x | \forall_{x^t \in S} \rho\{x\} \leq \rho(y)\}. \tag{19}$$

Zhang now uses a special case of theorem 6.1 of Young (1998) to derive the following proposition:

Proposition 1. *In the artificial world we just described, S is stochastically stable, i.e.,* $lim_{t\to\infty} lim_{\beta\to\infty} Pr\{x^t \in S\} = 1$.

In general, for any potential game with the log-linear revision rule, the set of all the states that maximize the potential function is stochastically stable.
In our model, the states that maximize $-\rho$ (minimize ρ, remember that $-\rho$ was the potential function) are stochastically stable. We know that ρ measures the contact between different colors. A state that minimizes ρ thus minimizes inter-racial contact. Hence it represents a segregated configuration. It then follows:

Proposition 2. *In the long run, if β is large, residential segregation is observed almost all the time.*

9. Results of the simulation

I will use this section to briefly summarize the results of the simulations performed by Zhang and his conclusions.

We run a simulation on a 100×100 landscape with a Moore neighborhood. We use the utility function from Equation (1) with (arbitrary) values $\beta = 10$, $Z = 1$ and $M = 0.6$. This results in the following payoff function:

$$u = \begin{cases} \dfrac{x}{n} & \text{if } x \leq n, \\ 1,4 - \dfrac{0.4x}{n} & \text{otherwise.} \end{cases} \tag{20}$$

The simulation starts with a checker board configuration, where every agent is living in a 50-50 neighborhood (Figure 8). In this "Pareto" state, everyone is getting the highest payoff and the configuration is socially optimal. However, as we know, this state is not stable, because the cost of deviation is very low. Once some agents accidentally change location, their neighbors become less satisfied than before, so they will move too, as soon as they have a change.

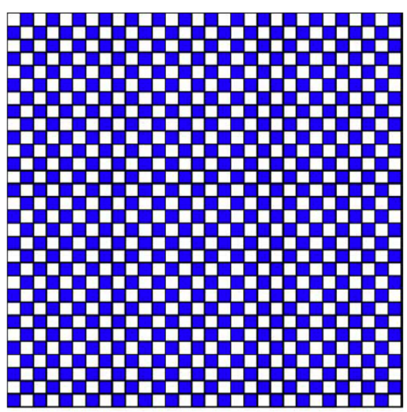

Figure 8: *The pareto optimal equilibrium*

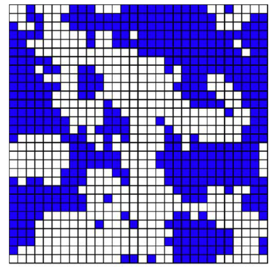

Figure 9: *A snapshot in the short run.*

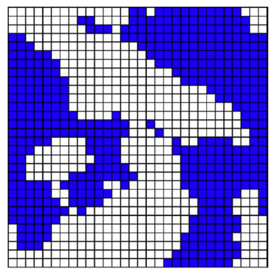

Figure 10: *A snapshot in the long run*

In the simulation of Zhang, we see that segregation starts to emerge as time goes on (Figure 9). If we wait long enough, we see that in Figure 10, blacks and whites are almost completely separated.

Proposition 2 states that segregation emerges in the long run, thus it is important to know how long this actually takes. Besides, it makes sense to ask what affects the speed of convergence. In the simulations performed by Zhang, segregation never occurres for $\beta < 2$. The waiting time decreases quickly for values close to 2, and at $\beta = 4$, simulations end fairly quickly.

Agent-based models almost always consider small neighborhoods, much smaller than the real world notion of a neighborhood [6]. Logically, if every agent views the whole lattice as a single neighborhood, then moving does not alter the composition of the neighborhood and it will become a random decision, This implies that the time before segregation occurs is infinitely long. From this we could reason that a smaller neighborhood implies a shorter waiting time.

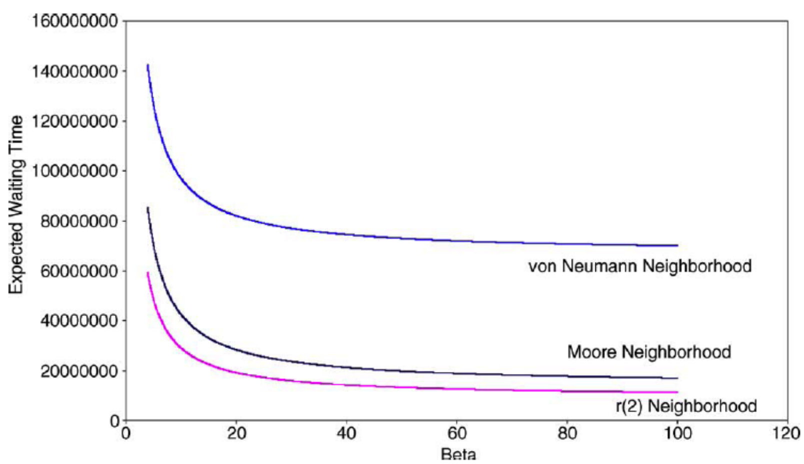

Figure 11: *Waiting time*

Interestingly, when the simulation is ran under different neighborhood definitions, we find that a bigger neighborhood actually decreases the waiting time for segregation (Figure 11). This can be explained quite easily: when agents have fewer neighbors, they are less differentiated by neighborhood characteristics. This means that it will take longer for an advantage trading pair to be formed, because an agent finds many other agents that live in the same kind of neighborhood. On the other hand, once a pair is matched, the potential gain is higher and the switch is more likely to happen, which will increase the potential function by a larger amount. It seems that the impact of the first effect is bigger than that of the second.

The conclusion of Zhang is that the simulation allows us to try more variations of the model. According to Zhang, the mathematical approach has improved upon Schelling in at least three different ways:

- The model proposes an analytical measure of segregation. Normally segregation is always presented as a visual effect, while in this model we have a clear measure: the potential function ρ;
- This model defines agents' utilities, which allows for a social welfare function. As a consequence, we saw that individually optimal actions may lead to a socially sub-optimal outcome. This was recognized by Schelling as well, without needing a social welfare function. Still, one could argue that it might be useful to have this measure in further research;
- This model changes the notion of segregation from an *emergent and persistent phenomenon* to a measure that is embedded in a game-theoretic framework in which a wide range of existing results and analytical tools are readily applicable.

10. Conclusion

Following Schelling [13], we have constructed a dynamic model to understand the evolution of residential segregation using the model of Zhang [18]. While Schelling his approach is inspiring, his inductive approach was merely able to produce simple results. By mathematically analyzing the model, segregation is characterized as a stochastically stable state that tends to emerge and persist in the long run regardless of the initial state. Several aspects in the articles we have discussed have been unattended, and provide interesting possibilities for further research.

Firstly, the payoff profile as given in Figure 5 assumes that individuals prefer a perfectly integrated neighborhood, and experience anything else as less desirable. I would propose a payoff profile similar to the one in Figure 12. As discussed in the observations by Schelling, if agents have a range in which they are indifferent about the amount of like-color neighbors, this might lead to a mild version of segregation. Moreover, this scenario does no seem very unrealistic to me; I believe that in real-world situations individuals are often unaware of the exact mixture of their neighborhood.

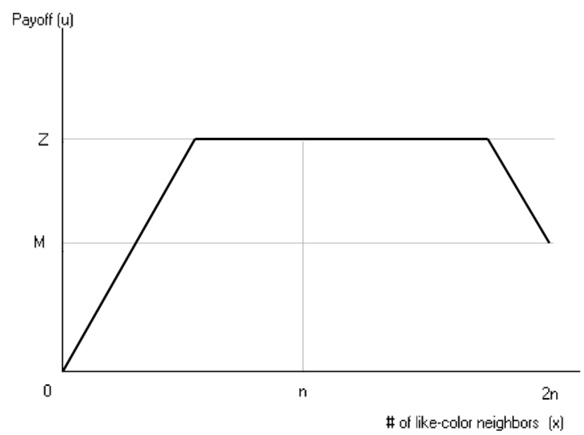

Figure 12: *Proposed payoff profile*

Secondly, the bid rent curve used by Zhang is assumed to be identical for every individual. Not only do both whites and blacks possess identical preferences, even the persons within a racial group are assumed to be identical. The idiosyncratic term does cause small perturbations in the behavior, but the deterministic term is assumed to be identical. The potential function depends on this notion of similarity, and thus the overall behavior of the system as well. If we would drop this assumption and introduce different payoff profiles for the racial groups, the resulting scenario could be quite different. To my knowledge this has not yet been done in a rigorous manner, and I believe it would be a suitable way to extend the model of Zhang.

Thirdly, it has to be said that plenty other research has been done on the Schelling model, which I have left out of consideration. I have deliberately chosen to focus on one extension of the model 1) because this is the first rigorous analysis of the Schelling model which I believe to be convincing, 2) I believe an in-depth analysis of one extension is more insightful than globally scanning several extension and 3) space is limited.

References

[1] O'Sullivan Arthur. Schelling's model revisited: Residential sorting with competitive bidding for land. *Regional Science and Urban Economics*, 39(4):397–408, 2009.

[2] Juan Miguel Benito and Penelope Hernandez. Modelling segregation through cellular automata: A theoretical answer. Working Papers. Serie AD 2007-16, Instituto Valenciano de Investigaciones Economicas, S.A. (Ivie), July 2007.

[3] Lawrence Blume. Population games. Game Theory and Information 9607001, EconWPA, July 1996.

[4] T.W Davis J.A, Smith. *General Social Surveys 1972-1998: Cummulative Codebook*. 1999.

[5] Bruce Edmonds and David Hales. Computational simulation as theoretical experiment. *Journal of Mathematical Sociology*, 29:83–86, 2005.

[6] Joshua M. Epstein and Robert Axtell. *Growing artificial societies: social science from the bottom up*. The Brookings Institution, Washington, DC, USA, 1996.

[7] Malcolm Gladwell. *The Tipping Point: How Little Things Can Make a Big Difference*. Back Bay Books, January 2002.

[8] Boris Vladimirovich Gnedenko. *The theory of probability, by B.V. Gnedenko. Translated from the Russian by George Yankovsky*. MIR, Moscow,, 1969.

[9] D McFadden. *Conditional logit analysis of qualitative choice behavior*, volume 1, pages 105–142. Academic Press, 1974.

[10] Gareth Myles. Discrete choice theory of product differentiation: Simon p. anderson, andre de palma and jacques-francois thisse (the mit press, cambridge, 1992) pp. 423. *International Journal of Industrial Organization*, 12(2):291–292, 1994.

[11] T C Schelling. *Micromotives and Macrobehavior*. Norton, 1978.

[12] Thomas C Schelling. Models of segregation. *American Economic Review*, 59(2):488–93, May 1969.

[13] Thomas C Schelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(May):143–186, 1971.

[14] Volker Schmidt. Markov chains and monte-carlo simulation, 2006.

[15] Rajiv Sethi and Rohini Somanathan. Racial inequality and segregation measures; some evidence from the 2000 census. Working Papers id:2104, esocialsciences.com, 2009.

[16] H. Young. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, 1998.

[17] H. Peyton Young. Social dynamics: Theory and applications. In Leigh Tesfatsion and Kenneth L. Judd, editors, *Handbook of Computational Economics*, volume 2 of *Handbook of Computational Economics*, chapter 22, pages 1081–1108. Elsevier, 00 2006.

[18] Junfu Zhang. Residential segregation in an all-integrationist world. *Journal of Economic Behavior & Organization*, 54(4):533–550, August 2004.

APPENDIX - Reconsiderations after concept version

- In the first version of this mini-thesis, I intended to discuss both the implementation of Zhang and the one of O'Sullivan. Because the concepts behind the version of Zhang provided my with enough material which I believed to be relevant, I have chosen to let the version of O'Sullivan out of consideration. Moreover, O'Sullivan does not really continue where Zhang has stopped, he uses quite a different model. I believed this would do the overall coherency of my story no good;
- In the peer review, it was suggested that I should make a clear distinction between opinion and facts. I have not chosen to do this, because 1) I do not express my opinion very often and 2) whenever I express my opinion I clearly mention that it is my opinion.