An Agent Based Model on Schelling's Segregation by Languages

Azadeh Ahkamiraad, Dieudonne N Ouedraogo
Hayford K.D. Adjavor, Mohammed Alhendi
Rajat Dhiman & Saraswathi Hathikal

***Abstract*—The purpose of this study is to show how people are distributed and settled down in cities based on their location in neighborhoods of their preferred language. Our Project could be defined based on the conventional Schelling's Segregation (1969). Briefly, an agent in a small region may change its location when his/her majority neighbors are not speaking the same language as the agent's language. We have developed a new dynamic model for our study which bear semblance to Schelling's model. Our model uses dynamic classes containing iterators and constructors. Also, we modeled and analyzed the interaction between agents and the macroscopic effect of the interaction; which leads to segregation based on agent's attribute (language). We make various assumptions and simulate scenarios to analyze the dynamic behavior and the evolution of emergence. To depict the level of clustering, we are using the mean distance of every node in the cluster**

I. Introduction

The Schelling's Segregation model was a breakthrough and a significant milestone in expressing social dynamics and how it is possible for the society to emerge (or show emergent behavior). More work on the Schelling's Segregation may lead us into potential inclusion of all forms of visible segregation (Caste, Gender, Color)[1]. However, we must be aware of the fact that Visible Segregation is not the only factor interpreting human behavior. The choice of company is determined by multiple factors, not necessarily discriminative, but emotional and cognitive as well. This may be regraded as the vision for an all inclusive model of human behavior and ABM is a perfect tool for the classical model.

II. Explanation of System Modeled

Assumptions for Classical Schelling's Model[2][3], are as follows:

1) The agents in Schelling's problem are based two attributes(race) are spread over 2D space.
2) The randomly chosen agent looks around in his/her neighborhood
   (a) If the agent finds other agents of the same type, he/she stays
   (b) Else, the agent moves to a random location in 2-D space

If we keep the threshold of neighbors speaking the same Language low enough, we would find localized segregation taking place. But a High Threshold (Indicating a low level of tolerance) would cause the agents to not settle down in forms of colonies and keep moving around.

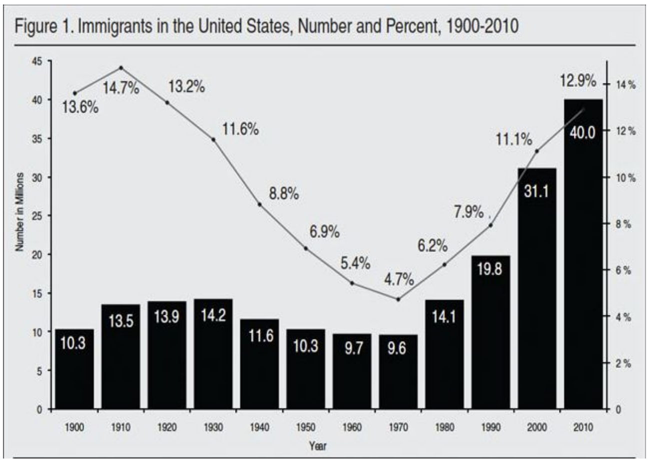


Fig. 1: Source: Decenial Censuses, 1990-2000, American Community Survery, 2010

III. Explanation of experiments and Mathematical Analysis

Thomas Schelling's model, does provide a apt explanation about how society decides to segregate and organize themselves. It also explains how people are inherently logical (but not necessarily rational at every instant), about the decision to aggregate with the same kind, or a community consisting of similar agents. However, as we see from social dynamics evolving over time, we find that time has caused societies to integrate much more than they used to. For instance, looking at the population distribution of the united states and finding the proportion of every ethnicity in the states using the histogram based on immigrant population alone, over the past century, (Figure 1) , we may conclude that the population is growing. The rise in population is 40 million in 2010 as opposed to 9.6 million in 1970 would indicate that individuals don't just segregate, but selectively integrate as well. This behavior may be termed as **localized segregation (or selective aggregation)**.

Based on this observation, we have introduced two key behaviors in our model. These assumptions are listed as follows:

1) **Adaptation**: People speaking the same language mostly come from the same culture. They tend to segregate, but **similarity** is not the sole metric of segregation. We

have also included the 'majority in numbers' as a factor influencing segregation & adaptation, in our model. Considering the population distribution of the United States as our modeling environment, we have the following four group that co-exist. We have simultaneously tried to observe their behavior:

(a) The Majority, i.e., English speaking population has the highest adaptation (or influence rate and population). Thus, over the years, the agents for other languages will try to adapt to the behavior exhibited by the English speaking population and hence convert to the same. They are represented by the color 'Blue'

(b) The second most common language observed is 'Spanish', represented by the color 'Red' and has the second highest adaptation rate and population

(c) We then define one of the minorities as 'Chinese', which is represented by the color 'Green' and have the third highest Adaptation/influence

(d) Lastly, we consider the Arabic population having the least adaptation/influence/population and are represented by the color 'Yellow'. All adaptation rates are kept constant.

2) **Reproduction based on Adaptation**: We also assume that the majority in numbers is a direct contributor to the overall increase in the number of agents, over the years. The number of agents belonging to a particular language community increase based on their adaptation rate and current population.

The behavior of these two factors may be expressed with the following equation:

$$T = a \cdot english + b \cdot spanish + c \cdot chinese + d \cdot arabic$$

Where, $a = 1.7$, $b = 1.1$, $c = 1.1$, $d = 1.05$ are the **fixed adaptation rates** for selected languages, and $english, spanish, chinese, arabic$ are the populations of the languages over the years. The spawn of agents and their locations over the map is implemented using **organized randomness**, so as to obtain different results while maintaining the order of majority. The addition of new agents each year is the expressed by the following equation:

$$New\_Agents = Adaptation \cdot Current\_Agents$$

This would reflect the exponential increase of the majority language over time.

Mean Distance: In order to analyze the level of segregation and sparsity, we have aggregated the positions and calculated the mean distance between individuals of one language. Individuals that are evenly spread out throughout the environment and have a centroid coinciding the approximate center of the environment, would not produce sufficient variation in the quantity. Whereas, multiple asymmetric clusters skewed across directions would cause the quantity to variate drastically.

3) **Mathematical Analysis**: We have simulated using two and four agents.

(a) Position of each agent is defined by x and y all in range of [0,1].

(b) Neighborhood is defined as the nearest element to the agent in term of Euclidean distance.

(c) We have also increased the neighborhood size to 40,60

(d) At the beginning of the process all agents are uniformly distributed in the square region [0,1]*[0,1]

(e) Agents move according to the following:

i) Go through all agents possible

ii) Pick a random position in in the square

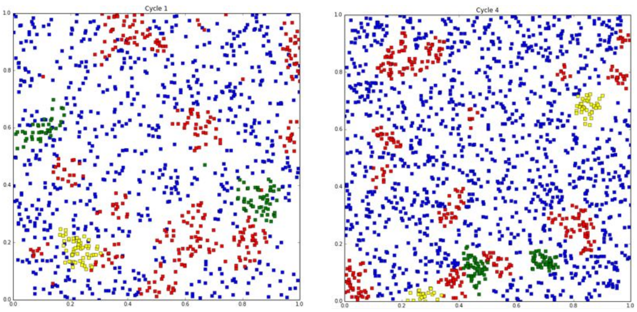
iii) If the agent is pleased with the position (tolerance level is met), agent relocates. Else, goto (i)

IV. Results

The results shown below emulate our projected outcomes appropriately. Change in Threshold and total number of agents, organized randomness imply that the mean position metric may be used to quantify the level of clustering. We observe the following:

A. Moderate Threshold

- The first case included taking agents in a similar manner to the population distribution of United States. The population is as orthodox as it is accepting. The details for first two years is shown in figures 2a 2b



(a) Year 1, Cycle 2 (b) Year 2, Cycle 4

Fig. 2: Trend of localized segregation/selective aggregation of the population for a moderate threshold

We further observed, the minorities are forced to aggregate. This maybe due to the inherent and moderately repulsive behavior. It shows induced rationality and process of intermingling. Figure 3

B. Low Threshold

We adjusted a threshold of 0.2 and obtained the following results for the first two years ref. Figures 4a and 4b below.

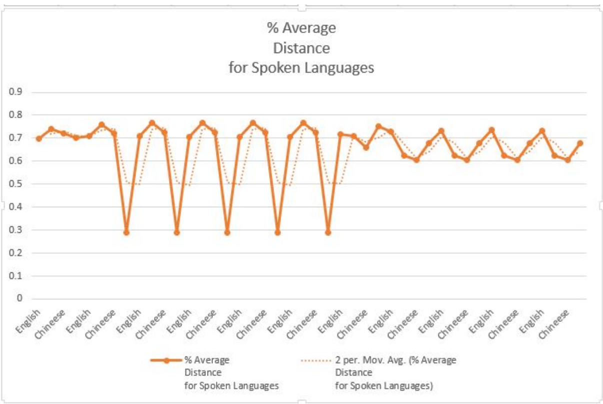
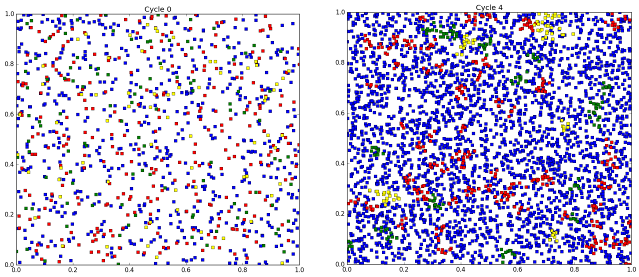


Fig. 3: It can be observed from multiple simulations that the mean distance for each language is affected by, the population size, population increase and the tolerance level

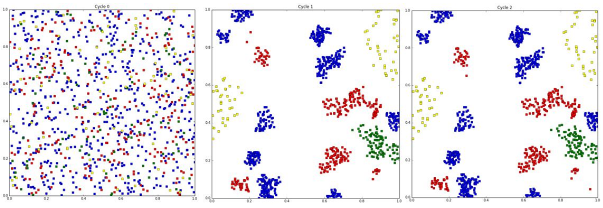


(a) Year 1, Cycle 0 (b) Year 4, last cycle

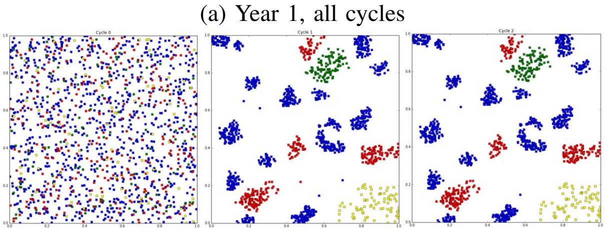
Fig. 4: Mild Segregation is observed - coincides with the classical model [3]

C. High Threshold

A threshold of 0.9 would produce high segregation, it also represents the case of a very orthodox population of individuals. The first two years have been shown as follows



(a) Year 1, all cycles



(b) Year 2, all cycles

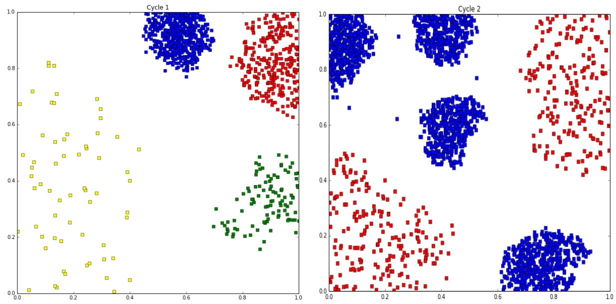
Fig. 5: Representation of a very orthodox population distribution - Ideal Case

It is visible from Figures 5a and 5b, that minorities are more spread out and are majorities are very closely aggregated. Visible change is only seen in the first two cycles of each year given by Mean Distance for all cycles. We also find that mean distance does not vary significantly after first two cycles due to tolerance criteria selected.

D. Variable Neighborhood

For every movement in the Moore neighborhood, we performed prior computations with 8 agents, just as Schelling had taken for the matrix based simulation, where every agent has eight neighbors. However, in a real society, this is not the case. Individuals look beyond just their own neighborhood, i.e. a broader community in order to make a choice to stay or leave, hence we assumed a larger neighborhood size (40 neighbors) and tried simulating the classic Schelling's model with this revised assumption. ref. Figure 6a, 6b, 7

The results are shown as follows:



(a) Cycle1 (b) Cycle 2

Fig. 6: Increasing the segregation neighborhood also leads to implementation of organized randomness. By implementing 20 neighbors and tolerance level of 0.5

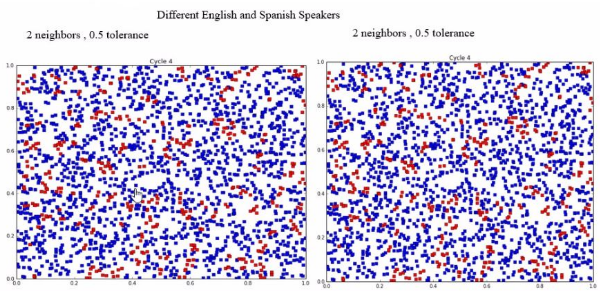


Fig. 7: Simulation for tolerance = 0.5, neighborhood = 2

We can clearly see from the cases above that even at a low tolerance level, if the neighborhood is expanded, high segregation is produced. As we raise the size of the neighborhood, the system partitions clearly into clusters of same language, segregation is very obvious. As the size of the neighborhood diminishes, segregation falls apart, the system tends to integrate, all other values remain constant.

If one holds values of the neighborhood constant or in small variation in certain interval, the perception is that even small tolerance level (and or population increase) leads to amplified segregation, but our set of simulations shows that a main contributor to segregation at the macroscopic level is the choice of the size of the neighbors, Schelling's assumption was based on values of neighborhoods confined in a region where segregation moves alongside tolerance. Therefore we have proved that Schelling's model was wrong.

There is a relation between size of neighborhood tolerance and segregation and our future goal is to quantify it.

V. Discussion of Results & Scope of future Study

Our results show the different case scenarios. We can also implement language learning process by integrating a list of languages that gets appended every time an agent moves to a neighborhood where there is a majority of people speaking a different language.

We would also like to express the mathematical relationship between neighborhood size and the tolerance level. We would like to use our model to explain migration between regions by including other variables such as employment, crime rate, climate etc

References

[1] C. T. Schelling, "Dynamic models of segregation," *Journal of Mathematical Sociology*, vol. 1, pp. 143–186, 1971.

[2] H. Sayama, *Modeling Complex systems*, 9. 2015, vol. 53, pp. 1689–1699, ISBN: 9788578110796.

[3] C. T. Schelling, "Models of segregation," *The American Economic Review*, vol. 59, no. 2, pp. 488–493, 1969.