An Extended Schelling Model of Segregation
T.A.N.S. - A.A. 2021/2022

Edoardo Grasso

Before we start

A brief description of the original Schelling model [1]

The system consists of a certain number of agents that live inside a neighborhood. The neighborhood is represented as a grid composed of $N * N$ tiles (houses) and $d * N * N$ agents, where $d < 1$ is the density. Each agent can be of one out of two possible ethnicities (white or black) and is characterized by a tolerance level (that is actually often set equal for all the agents). Agents can (and must) occupy one house at a time. The model is random initialized, placing an even proportion of agents of both ethnicities on the grid. The simulation evolves in discrete steps. At each step every agent activates its decisional process, that is: watch the neighbors and decide wether to move or stay. The neighbors it watches are the first neighbors, that are, in case of a central house, the ones above, below, on the left and on the right (there are just three neighbors in case of a side house and two in case of a corner house). Notice that some of the adjacent houses might be unoccupied, the agent always considers only the occupied ones as neighbors. The decision to move or stay is based on the ethnicities of the first neighbors and on the tolerance of the agent. In particular if it's true that:

$$\frac{Number\ of\ neighbors\ with\ different\ ethnicity}{Total\ number\ of\ neighbors} > tolerance$$

than the agent moves to a different (randomly picked) free house. The simulation runs for a predeterminated number of step or until the neighborhood becomes stationary, that happens if all the agents are happy or if the unhappy ones cannot find a good spot (it is not guaranteed that stationarity will always be reached).

Warning

The purpose of my project

The initial idea behind my project was to recreate the Schelling model in C++ adding two additional features:

- The number of possible ethnicities of the agents can be arbitrary large.
- The user can set a range of vision of the agents, that means setting the number of neighbors that every agent will consider during its decisional process.

When I finished coding my program I started to worry about ways to make benchmarks, so I started looking for articles that explored generalized Schelling models and I actually found one that analized indeed the effect of the range of vision. This is to say that the purpose of my work thereafter was to adapt my program to the one described in that article [2]. So I changed the geometry of my neighborhood to a thoroidal one and modified the core of the simulation (the iterate function). Also, more importantly, I decided that a good enough benchmark was to recreate the plot of the segregation coefficient (definition in the next section) vs the range of vision with two different values of tolerance, and since the only values for the range of vision that were used in that plot were the integers from 1 to 7, I constricted my model to take only one of these 7 values even though it could originally take any real value[1]. The model is still able to deal with more than two ethnicities even though there's no benchmark on the segregation results that it obtains in that case.

Segregation Coefficient

The Segregation Coefficient is defined as follows:

$$S = \frac{1}{dN^2}\left[\sum_{j,white}\frac{f_j - f_w}{1-f_w}\sum_{k,black}\frac{f_k - f_b}{1-f_b}\right]$$

$$0 \leq S \leq 1$$

Where d is the density, that is set equal to 0.9. N is the size of a side of the grid, which is set to 50 houses. f_w and f_b represent the expected fraction of white or black neighbors respectevely in a complete random initial society, both are equal to 0.5 if the concentration of each ethnicity is the same like in our case. The first summation is made over every white agent, f_j is the real fraction of white neighbors of agent j, same for the second summation but for black agents. For preciseness's sake I point out that the neighbors considered in these summations are the so called Von Neumann neighbors, this just means that they are the 4 adjacent houses (the first neighbors of the original Schelling model) indipendently on the value of the range of vision.
S, as stated in [2], is a good estimation of the ghettoization of our neighborhood.

[1]This has big performance advantages since I am able to store the addresses of all the 1 to 7 neighbors for each house in 7 different objects once and for all, this avoids the necessity of determining the proper neighbors for each house at every instantiation of a neighborhood (that now is much faster).

Brief description of the model

The code itself is pretty dense with comments that explain in appropriate detail every important function, it also contains a Description.txt file that provides detailed descriptions of each class that make up the program, as well as a README.txt file that is basically a user guide that illustrates all the functionalities aviable through the user interface. To keep this report as concise as possible, I'll limit myself here in making a very brief description of the launch_macro function that is used during validation, with the only purpose of providing some context before showing the actual validation results.
The cited function has the objective of launching 14 different simulations. Every simulation differs only from the range of vision and the tolerance of the agents. Each simulation consists of:

1. Randomly initialize a Neighborhood, that means:
   - create 2250 agents (that are $50 * 50 * 0.9$ where 50 is the size of a side of the grid and 0.9 is the density of the Neighborhood);
   - assign each of them to a random house;
   - calculate the initial state of happiness for every agent.
2. Call the iterate function, that makes the unhappy agents move to a better house, until the Neighborhood reaches stability or until a maximum number of iterations is reached.
3. Calculate the Segregation Coefficient of the Neighborhood and fill an histogram with its value.
4. Repeat the steps from 1 to 3 for 30 times.

Each simulation produces an histogram from which we obtain a mean value of the Segregation Coefficient $< S >= \frac{1}{30} \sum_{i=1}^{30} S_i$ and an estimate of the standard deviation $\sigma_S = \sqrt{\frac{1}{30} \sum_{i=1}^{30} (S_i - < S >)^2}$
We'll plot each of these 14 mean values with relative errors vs the respective values of range of vision and tolerance, and see how this plot rensembles the one obtained in [2].

The Results

A run of the launc_macro function provided the following results:

R	S	σ_S
1	0.63	0.01
2	0.85	0.01
3	0.92	0.01
4	0.95	0.01
5	0.96	0.01
6	0.96	0.01
7	0.97	0.01

Tolerance = 0.5

R	S	σ_S
1	0.52	0.02
2	0.36	0.05
3	0.16	0.10
4	0.01	0.02
5	0.00	0.02
6	0.00	0.01
7	0.00	0.01

Tolerance = 0.7

In these tables we clearly see the two different behaviours of the Segregation Coefficient, one of the main results of [2] was that, if we encourage the agents in having a larger range of vision and not consider only the very adjacent neighbors, even a slight increase in the tolerance of the agents can bring to a completely integrated neighborhood. Unfortunately the authors provide almost exclusively qualitative results, very few values of S for fixed values of range and tolerance are given, and some of them without error, anyway I'll list them in the table below so we can appreciate how all these values are compatible with the ones I found:

R	tol	S	σ_S
1	0.5	0.62	NaN
5	0.5	0.97	NaN
3	0.7	0.16	0.04
5	0.7	0.03	0.03

The rest of the validation will be more of an intuitive kind, since we'll have to limit ourselves in comparing by sight the two plots of S vs R:

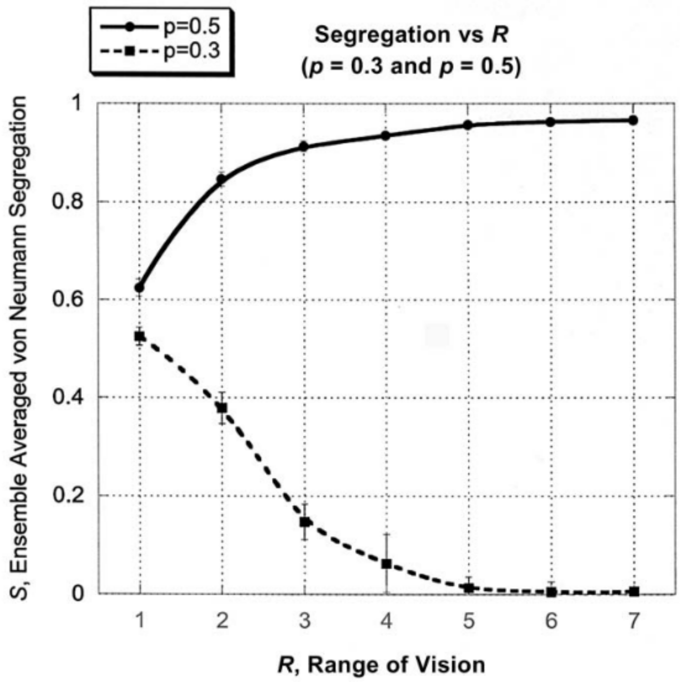

Figure 1: Results of [2]. Notice that with p, the authors mean the minimum percentage of neighbors of the same race for an agent to be happy, so their p=0.3 is equal to our tol=0.7

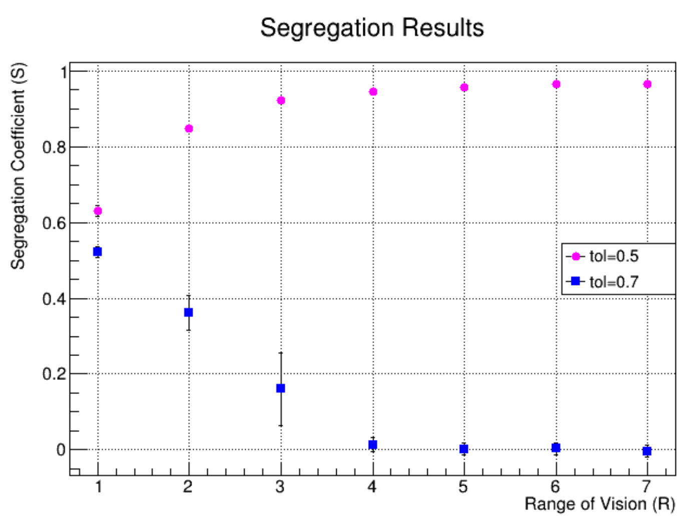

Figure 2: Results of our simulations

Conclusions

All the points seem to line up pretty well except for (R=4, tol=0.7) that, anyway, falls within the relative uncertainties. We can also observe other coomon factors, like the fact that the points for tol=0.5 have short error bars while for tol=0.7 the system seems to be less deterministic. I won't dig further into the implications of my model because that's not the purpose of this report.

References

[1] Schelling, Thomas C. "Models of Segregation." The American Economic Review, vol. 59, no. 2, 1969, pp. 488–93. JSTOR, http://www.jstor.org/stable/1823701. Accessed 4 Jun. 2022.

[2] 1. Laurie AJ, Jaggi NK. Role of "Vision" in Neighbourhood Racial Segregation: A Variant of the Schelling Segregation Model. Urban Studies. 2003;40(13):2687-2704. doi:10.1080/0042098032000146849