Agent-based dynamic modeling of residential segregation

Beta Version

Group C30
Avinash Pathapati (s3754715)
Dirk Jelle Schaap (s2745135)
Joppe Boekestijn (s2754215)
Lex Koelewijn (s2757036)

October 13, 2019

Abstract

Collective behavior of a group of agents with individual residential preferences was modeled in this research. Small deviations in preferences among individual agents results in complex group behaviour and possibly an equilibrium solution. Three groups of agents are modeled each with their own preferred neighborhood and on a per agent basis their individual tolerance with respect to the other groups. Group composition is dynamic as agents will progress between the three groups, new agents will be created and old agents will die. These temporal group composition changes approximate reality more closely than previous research and introduce new effects and collective behaviors. [*to be expanded in the Final version*]

1 Introduction

1.1 Problem

The problem focused on in this article is that of residential preferences in which a simplified model of the real world will be used in order to simulate agent behavior. Agents will adhere to their personal preferences in their behavior to optimize their own satisfaction. Behavior of individual agents is traceable in small numbers however when the population grows, complex and dynamic patterns of behavior emerge where these patterns are the result of the preferences of each individual agent. Thus individual preferences, which may seem insignificant, have a strong influence on group behavior and in this research some determining factors will be investigated, especially the degree of segregation that occurs over time between the agents.

1.2 State of the art

Schelling [Sch71] created a model in which there were two groups of agents whom all had individual preferences with respect to in which neighborhood they want to live as well as the ratio of the other group that they will tolerate in their neighborhood. In the paper the range of tolerance of individual agents with regards to the other group as well as the total amount of both groups are varied and it is shown how this can potentially result in a equilibrium state. Depending on all these variables complete segregation tends to emerge even when the majority of agents have at least some tolerance with regards to the other group. Hence the goals of the individual agents generally do not coincide with the collective end result, so fairly radical solutions tend to be found which are hard to explain based on agents' individual preferences.

1.3 New idea

Schelling created a model using two groups of agents in which there was one preferred neighborhood where both types of agents would prefer to live. The approach used in this paper is that there will be three groups of agents: young, adults and elderly where the assignment of agents to each group is dynamic in the sense that agents age and thus progress from one group to the next. Elderly agents will die after some time and adult agents can create new young agents, so the composition of the agent population is more dynamic. Furthermore whereas there was only one preferred neighborhood for all agents in [Sch71], there will be buildings in neighborhoods which dictate the preference of where agents want to reside. There will be three types of buildings, where only one can be in a neighborhood, and each building will be attractive to one age group of the agents. For example, elderly agents might want to live near a hospital. Each group of agents will have its own preferred neighborhood, however as in reality real estate is limited so agents can only move to their preferred location if there is space available and their personal preferences and tolerances allow for it. Lastly since the paper by Schelling was published in the 1960s models were solved mathematically and thus only a limited amount of models were tested. Seeing as the modeling can now be done digitally a much larger number of models with the slightest of modification can be tested. Moreover a visualization of the transgression of agents over time rather than only showing the resulting equilibrium will be possible using this technique.

2 Method

2.1 Simulation model

The model created is based on the state-of-the-art reference article [Sch71]. The basic idea is very similar, but several modifications were implemented in order to expand on the original research. The basic version consists of a grid where each coordinate can house one agent where each agent demands to live near similar agents. If an agent is unhappy, it will move to a randomly selected empty tile. This will cause a migration of agents and eventually, some final state which can be seen as the "solution". In this research a more complex approach is taken with multiple new elements introduced that will influence the migrations behavior of agents. A major side effect is that as opposed to the model created by [Sch71] this newly introduced model will not have a final state and can keep going indefinitely in theory.

A brief description of the model created in this paper will be given: At the start of the simulation a percentage of the tiles are kept empty and the remaining tiles are equally divided among the three groups of agents: young, adult and elderly agents. Seeing as agents can also age within a group each agent is given a random age as well within the category which thus results in an even distribution for every age in the simulation among agents. The initialization step is (pseudo-)random and is designed to generate the even distribution among all ages.

Analogous to the model created by Schelling, the agents want to live near agents of the same type. Two factors are introduced in the model which influence this preference: buildings and aging. Buildings will be placed in the simulation and each age group will be attracted to the tiles surrounding a building, so the adults might want to live near a office building for example. This means that an agent might be happy when it lives near its building of effect even though its neighbors are all different. The second aspect is that of aging: the agents age and thus progress from one age group to the next over time. This means that agents their preferences with regards to their neighbors as well as buildings will change over time and will likely cause them to move again once they reach a new age group.

Another feature introduced, which creates the possibility of an indefinite simulation, is the addition of death and birth in the simulation. Elderly agents will die after a certain number of (pre-specified) steps of the simulation and agents of the adults group will at random given the reproduction rate cause the birth of new agents belonging to the young age group. In this way there will be a continuous stream of agents entering and exiting the simulation and this is why there will (usually) not be a final state. These modification allow for the investigation of certain effects over a longer period of time, for example what will happen with the population when a baby boom is simulated.

2.2 Implementation details

The simulation has been implemented from scratch seeing as that provides the greatest level of freedom which is important because of the extra functionalities that are implemented on top of earlier research. Existing projects implementing a Schelling simulation are available online but these would have to be adapted considerably to fit the aim of this research so hence it was built from the ground up. The programming language used in the implementation is `Python3` and because the simulation is approached as a manipulation of a matrix no complicated libraries were needed. The two packages that are being used are a basic graphical user interface module called `Tkinter` and the extremely common plotting module `Matplotlib`. `GitHub` is used for version control.

2.3 Experiment design

The modified simulation allows for many different types of experiments because of the extra functionalities added, so in actuality there are a lot of potential problems that can be investigated using the model introduced. Executing all possibilities is infeasible so a selecting of problems that will be investigated was made:

- *Is it possible to simulate a constant migration of agents?*
  As explained, the modifications made should cause the simulation to loop endlessly and give insight into how agents migrate over the map. An interesting example is when, for example, a young agent is happy, it lives near three other young agents. However, when it turns into an adult its demands change. This does not necessarily mean it will move, since there is a chance three out of the five non-young neighbours are adults as well.

- *Do points-of-interest (buildings) change the migration patterns?*
  Buildings have been introduced into the simulation. One type of building, with areas of effect, for each age group. Agents will want to live as close as possible to the buildings for their respective age groups. Will this cause more segregation? Can buildings be placed tactically in order to combat segregation?

- *Is it possible to simulate something similar to a baby-boom?*
  The chance of reproduction for adults could permanently or temporarily be increased or

reduced. A possible application of this is to simulate a baby boom and observe the effects of it. This may cause exaggerated segregation when combined with the points-of-interest and make it harder for all agents to become happy.

3 Results

3.1 Experiment findings

Due to an issue with the visualization in terms of performance, results of the underlying data are available but not the actual migration patterns just yet. These should be available in the Final version.

Buildings

In order to test the effect of buildings the simulation was ran for 500 epoch with a grid with the size of 20 by 20, a density of 0.66, a homophily rate of 2, an ageing rate of 3 and a reproduction rate of 0.33333.

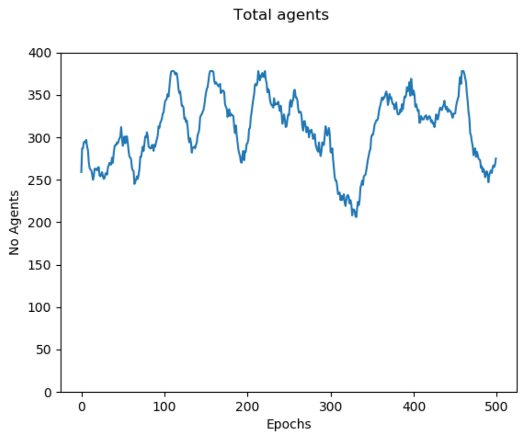

Figure 1: Total agents with buildings

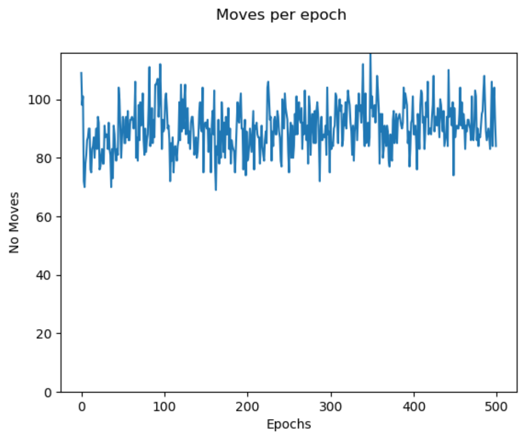

Figure 2: Moves per epoch with buildings

Babyboom

3.2 Interpretation of findings

Is it possible to simulate a constant migration of agents?

This first experiment turned out to be an essential problem to solve before other queries could be tested using the system. Due to the nature of the simulation it was important to find an expression for the required parameters of the model that would result in an even amount of deaths as well as births. Because the simulation goes on indefinitely the parameters of the model need to be picked such that an equilibrium between births and deaths is reached. If these values are out of balance then the number of agents in the simulation will either reach zero when agents die faster than are born, or it will completely fill up the grid when births occur at a higher rate than deaths. Either way, the simulation will stop seeing as no more moves are possible so these are the only two final states this model can reach and both are

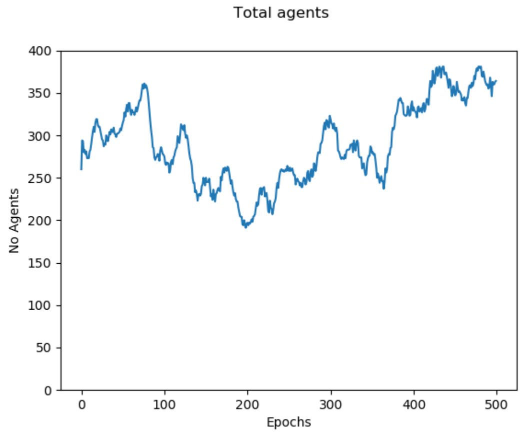

Figure 3: Total agents without buildings

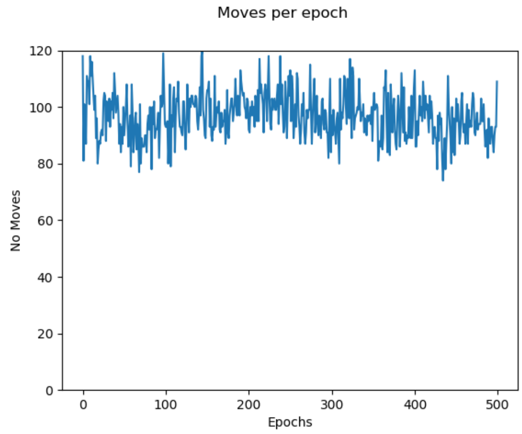

Figure 4: Moves per epoch without buildings

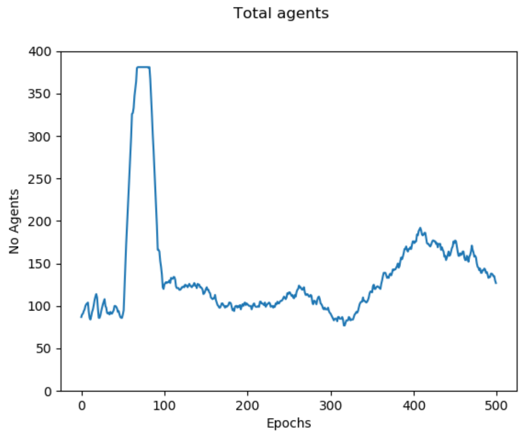

Figure 5: Babyboom Scenario 1

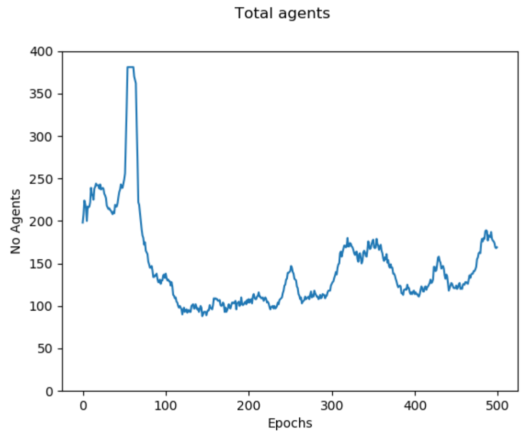

Figure 6: Babyboom Scenario 2

undesirable, so the parameters of the model have to be set up in a specific way adhering to a number of rules. It seemed important to find the formal relation between the parameters of the model as that would allow easy testing rather than relying on trial and error. As a starting point a formula for the theoretical amount of free spaces of the grid at each iteration was created which also yielded formulas for the rate of death and birth of the simulation.

$$Freespaces = D^2 * (1-d) + D^2 * d * \frac{1}{3} * \frac{1}{a} - D^2 * d * \frac{1}{3} * r$$

$$Death = D^2 * d * \frac{1}{3} * \frac{1}{a}$$

$$Birth = D^2 * d * \frac{1}{3} * r$$

Where D is the dimension of the grid, d is the agent density, a is the ageing rate of the agents and r is the reproduction rate.

Please note that this is the theoretical formula describing the behavior of the agents, meaning that over long runs this proves to be true. However due to the random nature of births and deaths and the fact that these effects can compound quite quickly over iterations the simulation might still show unbalanced behavior in the total amount of agents. This is inevitable and an essential part of the simulation but it is important to keep in mind seeing as an average of runs might be required. The equations above imply that in order to keep a constant amount of agents in the simulation in the long run the following must hold:

$$Death = Birth$$

$$D^2 * d * \frac{1}{3} * \frac{1}{a} = D^2 * d * \frac{1}{3} * r$$

$$\frac{1}{a} = r$$

Now this last relation is an easy formula relating the ageing rate to the reproduction rate while keeping the dimensions and the agent density constant. The optimal density in order to maintain a constant rate of death and birth seems to be twice the reproduction rate however it has yet to be uncovered why this holds seeing as theoretically the agent density should not be of influence. In actuality it is of course seeing as the density determines the amount of free spaces available in the simulation and currently births will not happen if there are not enough empty tiles left meaning that this will influence the theoretical equilibrium as well. So intuitively the importance of agent density in keeping a constant amount of agent in the simulation makes sense, but a formal description has yet to be found.

So, to answer the question: yes. It is possible to simulate a constant migration of agents. An important prerequisite for that is that the amount of agents should stay roughly similar or at least not approach zero or the limit of the simulation. The required relation in order for births and deaths to be in balance was formulated and provides a key building block to test the constant migration pattern. Due to the fact that agents their preference change over time seeing as they get older and new agents are constantly born, agents will move constantly over time. So it is possible to simulate the constant migration of agents however it is much harder to reach segregation or another form of equilibrium as was the case in Schelling's paper. [Sch71]

Do points-of-interest (buildings) change the migration patterns?

I am unable to interpret the migration pattern seeing as I have not been able to observe that at this point due to the work in progress on the GUI. I can say something about the underlying data.

Buildings do seem to have an effect on the behavior of the agents at this stage. When looking at figure 2 you see the amount of moves all agents make in 1 epoch when buildings are present and in figure 4 you see the same except that now there are no buildings in the simulation. It is interesting to see that the amount of moves per epoch, thus indicating that the agent is unhappy and not living near a building of interest, is lower in the simulation with buildings as opposed to the run without buildings. In the simulation with buildings the average amount of moves per epoch seems to lie around 90 whereas in the simulation without buildings this seems to be around 100 if not higher. In the simulation whiteout buildings you can see that the pattern is less predictable, there seems to be more variance in terms of the amount of moves per epoch.

Is it possible to simulate something similar to a baby-boom?

Two scenarios have been introduced and tested, the resulting effects on the total population in the simulation can be seen in figure 5 and 6. For the first scenario 500 epochs, a grid of 20 by 20, agent density of 0.2, a homophily rate of 2, an ageing rate of 10 and a reproduction rate of 0.1 was used. From epoch 50 to epoch 60 the reproduction rate is temporarily increased to 0.85. Figure 5 shows that the population grows rapidly with the increased reproduction rate and the artificially introduced cap on the maximum amount of agents in the simulation is quickly reached. Thus the actual amount of births due to the babyboom is limited once the grid fills up which happens rather quickly seeing as the death and reproduction rate is severely out of balance. The most significant lasting effect of the babyboom is that the overall level of the population seems to be higher on average than before the babyboom when situation is back in equilibrium. Intuitively this makes sense because the babyboom will have caused a general and relatively long term increase in the population. For the second scenario 500 epochs, a grid of 20 by 20, agent density of 0.5, a homophily rate of 2, an ageing rate of 4 and a reproduction rate of 0.25 was used. From epoch 50 to epoch 60 the reproduction rate is temporarily increased to 0.85, from epoch 60 to 80 the reproduction rate was lowered to 0.20 and after 80 epochs the reproduction rate settles at its initial rate of 0.25. The results can be found in figure 6. The period of a reduced reproduction rate was introduced to exaggerate the effect as well as the real world idea that reproduction might slow temporarily after a spike. This scenario shows some interesting results: the population growths during the baby boom but this effect is negated in the period of the relatively minor decrease in the reproduction rate. Seemingly the effect of a 0.5 drop in the reproduction rate for 20 epoch is more significant than an increase of 0.6 for 10 epochs, which might make sense given the short lifespan of agents in this setup. The overall effect however is that the population reaches an equilibrium that is lower than the initial situation after the babyboom and reduced reproduction.

4 Conclusion

4.1 Discussion

[to be added in the Final version]

4.2 Relevance

The article [Sch71] this research took inspiration from, showed that the inaccessible problem of segregation of US neighborhoods could be abstracted successfully by using dynamic models. This new simulation expanded on it by introducing agents with dynamic demands and by introducing buildings affecting the agents. Instead of agents having certain demands and sticking with them, the requirements of the agents change as they age. Moreover, there are now two ways for an agent to become happy, either by its neighbours or a building. The provisional results of this paper have shown that it is possible to simulate a constant migration of agents. The relevance of this is that it might be possible to apply this technique to real-world problems. A possible application of this can be to model a city district and try to recreate the migration patterns. This can give an insight into the origin of unwanted segregation or could help to determine the optimal placement of new public service buildings. *[to be expanded in the Final version]*

4.3 Team Work

The project proposal and paper analysis were done in equal parts. This report is mainly written by Lex & Dirk Jelle, while the bulk of the coding is done by Avinash & Joppe. Choices with regard to implementation and experimental setup were decided collectively, either over WhatsApp or in the weekly meetings. Ideally, everyone contributes equal parts to the report and to the code, but we found that goal difficult to achieve. This could be a point of improvement for us.
For the Beta version all four of use worked on the code. Avinash focusing on the GUI, Dirk Jelle on the internal structure of the code, Joppe on adding buildings and Lex on the plotting of experiments as well as debugging. The report was written/improved by Lex (and some by Dirk Jelle). Lex also did the preliminary experimenting.

49

References

[Sch71] Thomas C Schelling. "Dynamic models of segregation". In: *Journal of mathematical sociology* 1.2 (1971), pages 143–186. DOI: `https://doi.org/10.1080/0022250X.1971.9989794`.

Comment Summary

Page 1

1. Avoid passive and drop the superfluous "in this research". Write: "We model collective ..."
2. result (singular!)
3. we model ...
4. superfluous, change is always temporal ;-)
5. Your abstract should at least mention your main result.
6. Replace this long phrase with as simple "We consider residential ..."
7. who (no m)
8. this is slightly misleading: consider "i want to live in Reitdiep" vs "I want to live where rich people live".

Page 2

9. the main point of Schelling is: even when most agents are _very_ tolerant, segregation happens.
10. really? everyone wants to move to the same place? don't you mean that each type of agent prefers to live next to agents of the same kind?
11. ah! this is an important new idea! mention it in the abstract!
12. start a new paragraph here
13. remove this.
14. not always have
    i. or never?
15. be explicit here: what numbers do you use? what intervals are young, etc.?
16. type, i.e. age category.

Page 3

17. agents change their
18. remove this.
19. how are they placed in the grid?
20. -s
21. to provide
22. 3.6 or 3.7 or ... ?
23. Git and GitHub are different things...
24. it
25. good questions :-)

Page 4

26. okay, thank you for clarifying this.
27. what is an epoch? define it or just say "steps"
28. explain what these numbers mean, or use the same terms in your explanation above.
29. Is this different than without buildings? Why did you make these two graphs in particular? Which research questions do they answer?
30. you mean: find good values for the parameters
31. only two kinds of final states

Page 5

32. ah, better put all figures on one page?
33. use the same range in the second axis as in the figure with buildings to make them comparable!
34. maybe consider making a stacked graph showing the three different age groups above each other?
35. use \text{Frespaces} or \mathsf{Freespaces}

Page 6

36. note
37. you mean: the empirical results should approximate this formula
38. This is a confusing sentence, please rephrase it.
39. Remove "So,". Please avoid informal phrases.
40. wrong word order
41. Please provide results for this claim! Did you keep track of the amount of clusters over time? Can you should the grid at say 10, 50, 100 steps etc.?
42. Are you working alone? Ask your team members! Work together!
43. to really compare figures 2 and 4 please ensure they have the same y-axis range and please put them on the same page.
44. This sounds vague. Please calculate it and do not make wild guesses just by looking at the graphs.

Page 7

45. Avoid passive. Say "We introduced and tested ..."
46. Figures
47. start a new paragraph here

48. were used

    or, even better: We used ...

Page 8

49. Only one reference is a bit disappointing. Please check whether similar studies have been done! Segregation with age groups seems too obvious to not exist already. Also, make sure to mention also any other references you have read or used as inspiration during the project!