Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Lecture with Computer Exercises:
Modelling and Simulating Social Systems with MATLAB

Project Report

Social aggregation in a multicultural environment

Camillo Balerna
Nicolas Lanzetti
Giovanni Moscato
Marc Neumann

Zurich
December 2015

Agreement for free-download

We hereby agree to make our source code for this project freely available for download from the web pages of the SOMS chair. Furthermore, we assure that all source code is written by ourselves and is not violating any copyright restrictions.

Camillo Balerna   Nicolas Lanzetti   Giovanni Moscato   Marc Neumann

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

SOCIAL AGGREGATION IN A MULTICULTURAL ENVIRONMENT

Authored by (in block letters):

For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
BALERNA	CAMILLO
LANZETTI	NICOLAS
MOSCATO	GIOVANNI
NEUMANN	MARC-PHILIPPE

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date

ZURICH, 9.12.2015

Signature(s)

Moscato

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.

Contents

1 Abstract

In the presented project a combination of the Schelling and Axelrod models was implemented. The agent can decide whether to move to another place or stay, depending on their satisfaction level with the neighbours, computed with Axelrod. Assuming that an agent decides to stay, some features can be shared, meaning that an interaction between the agents takes place. This algorithm was implemented in order to investigate the effect of a possible movement in a pure social interaction scenario. Results showed that the agent's happiness increases if they are allowed to move.

2 Individual contributions

The entire project has been carried out as a team, consisting of four mechanical engineering students. The group was further subdivided in small teams: one was more responsible for the mere Matlab implementation of the code, while another one wrote the report and organized the campaign of simulations. But it must be underlined that the discussion and interpretation of the results was always done in weekly meetings, where the knowledge and ideas of the single individual could have inspired the other team members for new scenarios to be investigated.

3 Introduction and Motivations

3.1 The Schelling Model

"People get separated along different lines and in different ways. There is segregation by sex, age, income, language, colour, taste, comparative advantage and the accidents of historical location. Some segregation is organized; some is economically determined; some results from specialized communication systems; and some results from the interplay of individual choices that discriminate. [...] My ultimate concern of course is segregation by colour in the United States; but at the level of abstraction of this paper, any twofold distinction could constitute an interpretation." [2]

An important economic and sociological phenomenon where individuals cluster together according to some shared similarities such as race, nationality, income or education level, is segregation. In the end of the 60s, Thomas Schelling (winner of a Nobel Prize in economics), suggested a simple model to explain this phenomenon: people, living on a lattice, chose where to live on the basis of the colour of their neighbours. It was shown that even if people only have a very mild preference for living with neighbours of their own colour, since they decide to move in order to satisfy these preferences, complete segregation might occur. Such a result is considered surprising and has generated a large literature. It can be applied to many fields: in fact the Schelling result is of interest to economists because it illustrates the emergence of an aggregate phenomenon, that is not directly foreseen from the single person behaviour and because it concerns an important economic issue, i.e. that of segregation.

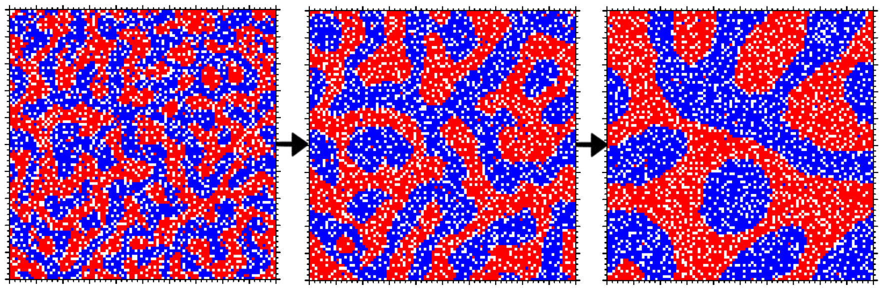

Figure 1: An example of the Schelling model with increasing tolerance threshold.

3.2 The Axelrod Model

"Despite tendencies toward convergence, differences between individuals and groups continue to exist in beliefs, attitudes, and behaviour. An agent-based adaptive model reveals the effects of a mechanism of convergent social influence. The actors are placed at fixed sites. The basic premise is that the more similar an actor is to a neighbour, the more likely that that actor will adopt one of the neighbours traits."[1]

Robert Axelrod introduced the word 'culture' to denote "*the set of individual attributes that are subject to social influence*". He supposed that culture "*is something people learn from each other*", and hence evolves through social influence. In order to study the process of cultural diffusion, Axelrod built a model based on two simple assumptions: first, people sharing many of their cultural attributes are more likely to interact with others who share the same attributes, and second, these interactions tend to increase the number of cultural attributes they share, making them more likely to interact again. Another example said by Axelrod is language: "*a person is more likely to talk to someone who speaks a similar language than one who speaks a dissimilar language, and the very act of communication tends to make their future patterns of speech even more similar*". The objective of Axelrod model is to investigate why cultural diversity persists even though people tend to approach and interact with each other culturally through a selfreinforcing mechanism, which can be summarized as following: the more people interact, the more they will be similar. Furthermore Axelrod's model also had a significant impact in the scientific world and research, since several authors have analysed the model and some of its structural assumptions in depth.

3.3 Motivation

After having introduced how the Schelling and Axelrod model work, the motivation of our project is now introduced.

What really seems interesting to be analysed and understood are the cultural and social motivations behind the aggregation and segregation phenomenon of the individuals. In any heterogeneous social environment (such as big cities, like metropolis) we can observe both isolation of groups of people (e.g. "China Town") and cultural integration. It must be underlined that social separation has always existed, but it has gained particular importance in the last decades thanks to everincreasing people's propensity to migrate in another region.

The aim of our investigation is the one of combining the features of Axelrod and Schelling models together: in fact Schelling only takes into account whether someone is going to move or stay depending on its satisfaction, while in Axelrod model the agents are fixed, they cannot move, and they decide whether to interact and share an attribute or not. So what would happen if an individual could be able to decide to move depending on its satisfaction level and interact with his neighbourhood at the same time?

Here are some open questions that we will try to answer, implementing a code that both takes into account Schelling and Axelrod at the same time:

- How easy is it for a group of heterogeneous cultures to successfully interact and coexist with each other?
- In what circumstances do people prefer to relocate rather than adapt to other's culture? How does the exigency of the single agents influence the output of our simulation?
- To what extent do we observe ghettos?
- How does population density play a factor in these processes?

4 Description of the Model

The behaviour of the agents is ruled by a satisfaction index p, which is computed according to Axelrod's probability computation:

$$p_i = \frac{1}{F \cdot n} \cdot \sum_N \sum_F \delta_{f,j} \qquad \text{with } \delta_{f,j} = \begin{cases} 0 & \text{if the feature is equal,} \\ 1 & \text{if the feature is different,} \end{cases} \tag{1}$$

where F is the number of features, N are the neighbours[1] and n the number of neighbours. The agent acts according to the following inequalities:

- $p < t_{\text{low}}$: the agent moves;
- $t_{\text{low}} \leq p < t_{\text{up}}$: the agent interacts;
- $p \geq t_{\text{up}}$: the agent is happy.

From this description one may notice that the Schelling and the Axelrod models are nothing but special cases of our model. For $t_{\text{low}} = 0$ one obtains an Axelrod model with tolerance parameter t_{up}, while for $t_{\text{up}} = 0$ one gets a Schelling model with tolerance parameter t_{low}. To validate our model both cases were tested.

5 Implementation

5.1 Initialization

The grid of size $N \times N$ is stored as a multi-dimensional array $N \times N \times F$. For every cell G_{ij} the culture vector $\sigma_{i,j}$ of size $F \times 1$ initialized with random values. The culture vector contains the traits of the agent at the (i, j)-th position, therefore

$$0 \leq \sigma_{i,j,1\ldots F} \leq Q.$$

If a cell (i', j') is empty then

$$\sigma_{i',j',1\ldots F} = 0.$$

The probability for a cell to be empty ruled by a model parameter called `vacancies`. Once the grid is initialized we may add some additional initial conditions such as (see figure 2):

- spots of a single culture ("`leopard`").
- ring of a single culture ("`circle`").

[1]As neighbours N of an agent at position (i_0, j_0) we consider the Moore Neighbourhood of range 1:

$$N_{(i_0,j_0)} = \{(i,j) : |i - i_0| \leq 1, |j - j_0| \leq 1, \{i \neq i_0, j \neq j_0\}\}.$$

.

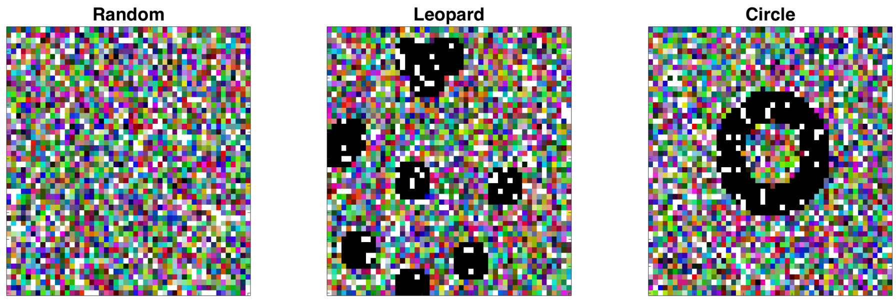

Figure 2: Initial conditions for the simulation.

5.2 Simulation

At each elementary timestep a random agent (i, j) is selected. Its satisfaction index $p_{i,j}$ is calculated according to equation 1. The following choice is made:

- if $p_{i,j} < t_{\text{low}}$ the agent is considered very unsatisfied, hence it moves in a free cell within the grid, according to the function `moveCheck` (Section 5.2.1);
- else if $t_{\text{low}} \leq p_{i,j} < t_{\text{up}}$ the agent is willing to adapt, hence it interacts with a randomly selected neighbour, according to the function `interact` (Section 5.2.2);
- else the agent is totally satisfied, therefore nothing happens.

This is repeated at every timestep for all agents. At the end of each timestep t the ratios of satisfied, unsatisfied and adapting agents, as well as the count of unique culture vectors τ_t is stored for plotting and result analysis (Section 5.3).

5.2.1 Movement algorithm

This algorithm specifies how agents move within the grid. The moving agent (i, j) picks a random empty position (i', j'):

- if $p_{i',j'} \geq t_{\text{low}}$ the agents moves to position (i', j');
- otherwise it picks another free position.

This is reapeted until a suitable position (i', j') is found. However, after 100 attempts, the loop interrupts and the agent moves to (i', j'), even though its satisfaction may not have improved. This means that an agent will always try to improve its satisfaction so that it will not need to move again.

5.2.2 Interaction algorithm

This algorithm describes how agents exchange cultural traits with each other. The selected agent picks a random neighbour[2] (i', j') and a random feature f and substitutes its fth feature to match its neighbour's, such that $\sigma_{i,j,f} = \sigma_{i',j',f}$.

5.3 Data visualization

If the data visualization of Schelling's model is straightforward, the same does not hold true for multi-feature agents. To achieve a clear visual representation of the results three kinds of plots are used: the agent grid plot, the satisfaction grid plot and the numerical data plot. All the plots are generated by the function `plotGridDirect`.

5.3.1 Agent grid plot

In order to unambiguously represent culture similarity colours were used. For this purpose F was set[3] equal to 3 and a primary colour `(R,G,B)` was assigned to each single feature. The final colour is built as a combination of (ξ_1, ξ_2, ξ_3) as follows:

$$\xi_k = 1 - \frac{\sigma_{i,j,k}}{Q}, \tag{2}$$

where $(\xi_1, \xi_2, \xi_3) =$ `(R,G,B)`; by construction empty cells are white.
Once the `RGB` triplet of every agent is determined, it is stored in a $N \times N \times 3$ matrix which is plotted using the built-in MATLAB function `image`.

5.3.2 Satisfaction plot

This plot shows a grid with four colours:

- empty cells are white;
- cells containing unsatisfied agents are red;
- cells containing interacting agents are blue;
- cells containing satisfied agents are green.

Notice that by construction this plot is suitable to represent models for any given F.

[2]Again, as neighbours N of an agent (i_0, j_0) we consider the Moore Neighbourhood of range 1:

$$N_{(i_0,j_0)} = \{(i,j) : |i - i_0| \leq 1, |j - j_0| \leq 1, \{i \neq i_0, j \neq j_0\}\}.$$

.

[3]Notice that this data representation is not suitable for $F > 3$.

5.3.3 Numerical data plot

Goal of this plot is to depict quantitative evolution over time of most relevant state variables of the system:

- percentage of satisfied agents α_t (green);
- percentage of interacting agents β_t (blue);
- percentage of unsatisfied agents γ_t (red);
- ratio of unique cultural vectors to total possible cultural states: $\frac{\tau_t}{Q^F}$ (black).

As for the satisfaction plot, the plot is suitable for any given value of F.

6 Simulation Results and Discussion

In this section, interesting simulation results are presented and discussed. It must be underlined, that in the following, convergence designates the scenario where the most of the agents are satisfied and the grid is mostly stable.

6.1 Moving agents and not moving agents

Firstly, it was investigated how the movement of the agents influences the simulation. Therefore two models were analysed: one with moving agents (Figure 3) and one with non moving agents[4] (Figure 4). Three situations may occur:

- both models converge;
- only the model with movement converges (as illustrated in the figures 3 and 4);
- both models do not converge, which means that the tolerance thresholds are too strict.

In general, the possibility to move within the grid increases the agents' happiness. On the other hand it is clear that the agents tend to be more alike and to create ghettos. Another conclusion of this simulation is that the interaction alone does not always lead to convergence and therefore it is not always easy for a group of heterogeneous cultures to coexist in case their movement is constrained.

[4]Note that the initial condition is equal for both simulations. In this case we chose a random grid with isolated spots of single culture

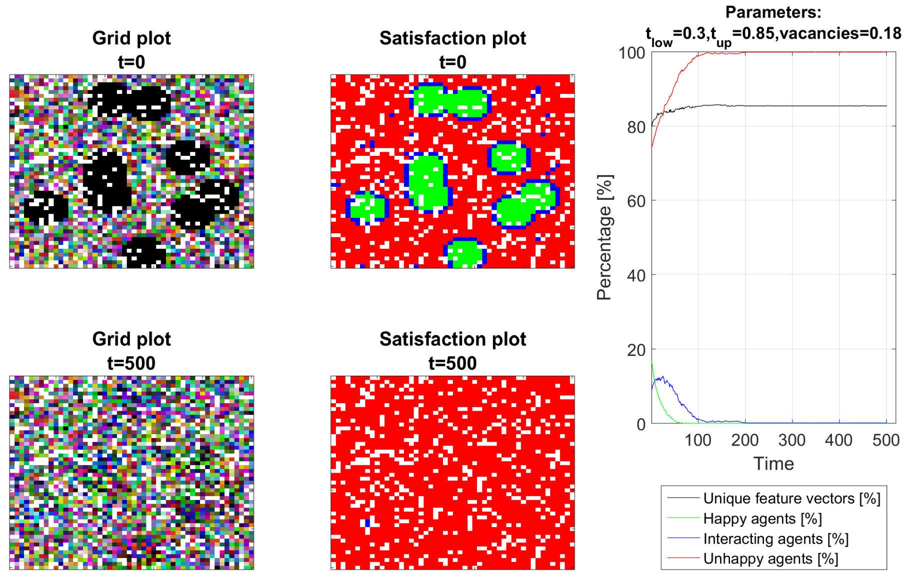

Figure 3: Simulation with only interacting agents.

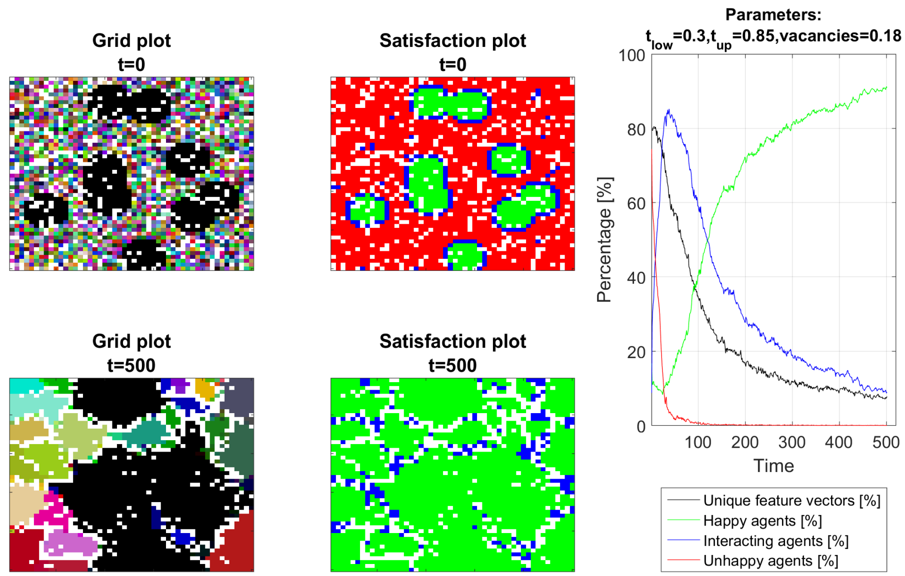

Figure 4: Simultion with moving (and interacting) agents.

6.2 Parameter correlation

A correlation between the lower tolerance value (t_{low}) and the final number of happy agents was discovered. In the conducted simulations, this made the difference between converging and diverging.
This phenomenon can be described in a social environment very clearly: if the agents are too exigent it gets more difficult for them to become happy.
As one can see from the figure 5, if t_{low} is too small, the number of happy agents after a certain amount of timesteps is limited. A reasonable explanation might be the following: if the agents start interacting too early, it takes more time for them to become fully satisfied, since only one trait per timestep changes, as for moving agents the change is drastic. Starting from a certain t_{low} the number of happy agents gets closer to a saturation, and after that it falls to zero, which shows the point of too exigent agents. At this point the model diverges.

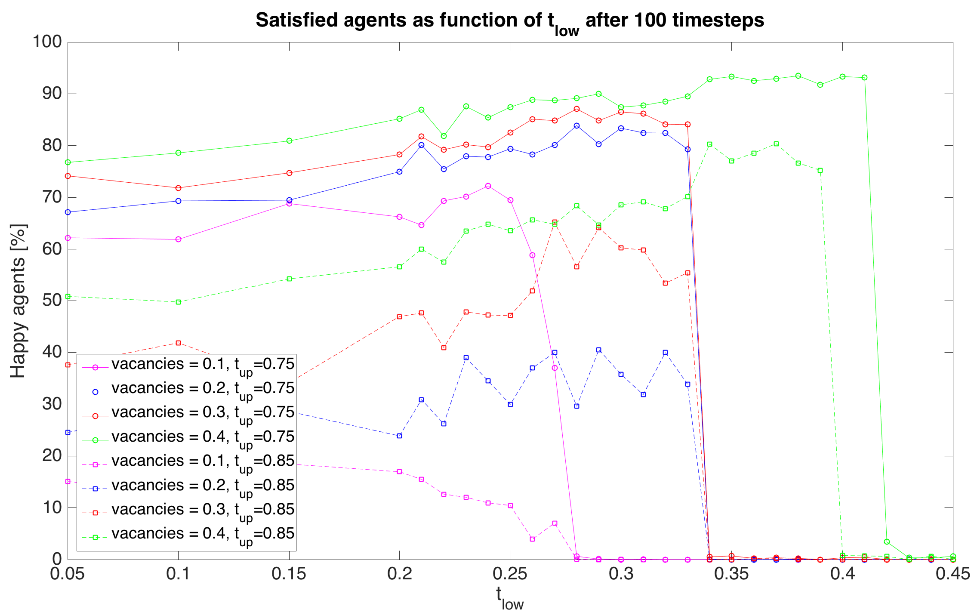

Figure 5: Correlation of t_{low} and the percentage of happy agents.

6.3 Ghetto formation

As long as the model converges ghettos can be observed. In general there are three types of delimitations between them (see figure 6):

- if two completely different colours are close-by an empty film forms in between;

- if two quite similar colours are adjacent a layer of interacting agents that keep switching their culture appears;
- if two very similar colours are adjacent there is no border in the satisfaction plot.

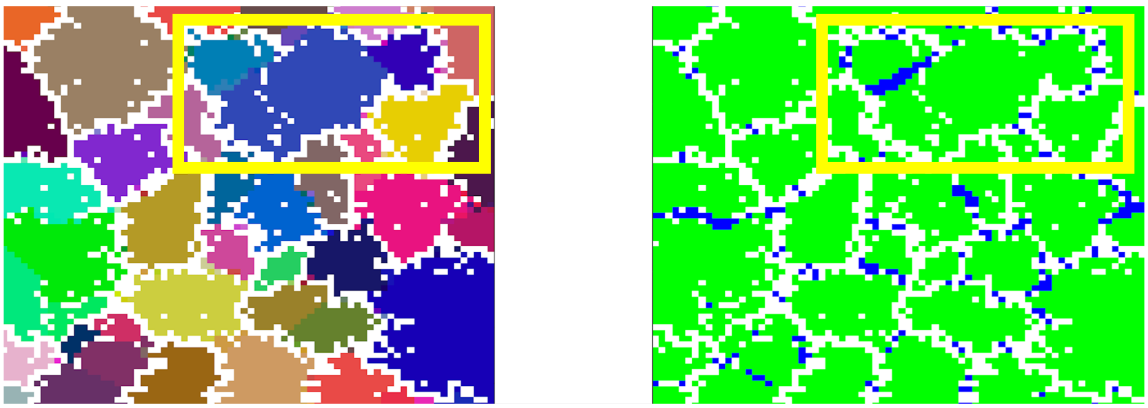

Figure 6: Types of borders.

To further analyse the ghetto formation the three initial conditions have to be considered separately, i.e. totally random grid, `circle` and `leopard`.
If a totally random grid is set, the formation of small clusters of agents sharing the same culture is observed. Obviuosly, if the tolerance parameters are too strict, the simulation will produce with a non converging result.
More interesting are the final patterns when the grid is initialized with one of the above described initial conditions. Setting the `leopard` or the `circle` grid as initial condition the simulation starts with a group of happy agents close to each other. Under this set up the formation of bigger ghettos is observed, especially around pre-existing close spots (see figure 4).

6.4 Vacancies

A correlation between the percentage of vacancies and the final result of the simulation can be noticed. In grids presenting a high percentage of empty cells, which can be interpretated as a city with a low density population, it is easier to reach a stable solution and therefore a general happiness of the agents. On the other hand high density population (low vacancies percentage) makes the convergence more difficult to be reached.

7 Summary and Outlook

7.1 Summary

At the beginning of the project some realistic and plausible assumptions about the results of our model were made. To recall these expectations: it was supposed that the strength in spatial separation would increase as the cultural differences enlarge, and that the integration would get easier in a more homogeneous cultural environment.
Looking at the results obtained from the simulations, it can be observed that the final pattern of the grid was always quite similar, unless a certain perturbation is inserted, i.e. an initial condition for the distribution of the agents' features.

It is also clear that the possibility to move, as a general trend, increases the outcoming happiness of the agents.

As already mentioned, the goal of the investigation was to get answers to the questions presented in section 3.3: in general, the expectations have been met.

7.2 Outlook

Future developments of the code might include weighted features and additional initial conditions. For a further analysis it may be of interest to investigate the influence of the parameters F and Q on the system.

A List of relevant variables

A.1 Fixed model parameters

These are the parameters fixed for the whole simulation:

`N`	Grid size
`time`	Simulation time
`F`	Number of features
`Q`	Traits per feature
`vacancies`	Ratio of empty cells
`tolInterval(1)`	Lower tolerance value (t_{low})
`tolInterval(2)`	Upper tolerance value (t_{up})
`radius`	Radius of spots or of the ring of single culture

A.2 Timestep state variables

These are the variables updated at every step:

`satisfactionIndex(i,j)`	Matrix containing whether the agent (i, j) is satisfied
`featureCountHistory(t)`	Count of unique cultural vector at time t (τ_t)
`graphVisualization(1, t)`	Percentage of satisfied agents at time t (α_t)
`graphVisualization(2, t)`	Percentage of interacting agents at time t (β_t)
`graphVisualization(3, t)`	Percentage of unsatisfied agents at time t (γ_t)

B References

[1] R. Axelrod, *The dissemination of culture - A model with local convergence and global polarization*, Journal of Conflict Resolution, Vol. 41 No. 2, 1997, pp. 203-226, Available at .

[2] T. Schelling, *Dynamic models of segregation*, Journal of Mathematical Sociology, 1971, Vol. 1, pp. 143-186.

[3] S. Balietti, 2011, *Schelling's model,* [MATLAB function], Available at `https://github.com/msssm/lecture_files/tree/master/evolutionary` (Accessed December 2015).

[4] *Axelrod's model of dissemination of culture*, Available at `http://jasss.soc.surrey.ac.uk/12/1/6/appendixB/Axelrod1997.html` (Accessed December 2015).

[5] D. Vinkovic, A. Kirman, 2006, *A physical analogue of the Schelling model*, Proceedings of the National Academy of Science, Available at `http://vinkovic.org/Projects/Schelling/Vinkovic_Kirman_PNAS_final.pdf` (Accessed December 2015).