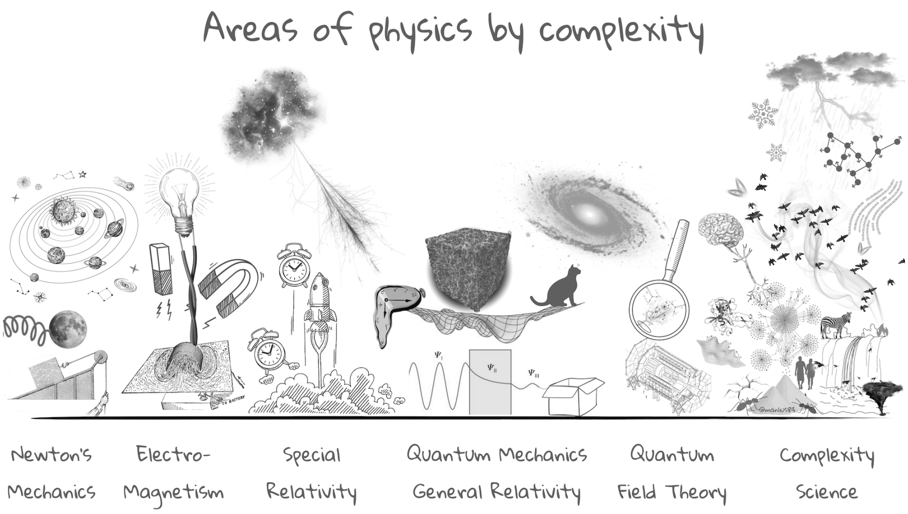
Final Report

Physics of Complex Networks: Structure and Dynamics

Università degli Studi di Padova



Project # 31,34,45

Jules, Vandendriessche

Last update: June 23, 2024

Contents

1 | 31. Schelling segregation model

Task leader(s): *Jules Vandendriessche*

1.1 | Introduction to segregation

This task focuses on an idea first introduced in 1973 by Schelling [4]. In this model, a slight aversion to dissimilar agents may lead to drastic segregation. This model will be implemented 3 times in this task, each time adapted according to the proposed paper.

The original model proposed by shelling can be viewed as N people sitting on a 1-D line next to each other (this analogy because agents are for now free to place themselves anywhere freely). The agents have a property, that distinguishes the population in 2 groups. The neighbourhood exists out of the agents that affect an agent in question. In this analogy, it would be the agent left and right of it, however the neighbourhood can be extended. The agents are happy when the fraction of similar neighbours in their neighbourhood is above the threshold (f), which is usually taken as 1/2. However, for now the threshold will be defined as the number of dissimilar agents in the neighbourhood that are tolerated. The dynamics of this first case goes as follows: An unhappy agent is selected at random, and it looks for the closest positions in between other agents where it will be happy. If there is no position where his wishes are satisfied, he will lower his threshold by one and try again. If it is still not possible, he prefers to not move and will remain unhappy. This goes on until an equilibrium stage has been reached. For this part, null boundaries are considered, meaning that agents on the ends only have 1 neighbor.

In the paper written by D'all'Asta et al [2], some changes to the dynamics are proposed. First of all, empty sites are introduced, hence the analogy to a 1-D street with houses and agents that move in between them is of order. Agents are free to move to any house (empty site) that fulfils their needs. Furthermore we will take a look at an unconstrained model, where agents are allowed to move as long as their utility does not decrease. Lastly, a comparison is made between non binary and binary dynamics. Meaning that agents who have a neighbourhood filled with agents that are alike, and agents that have a neighbourhood just above the threshold will be distinguished in the non binary case, and viewed as the same in the binary case. The threshold (f) is now defined as the magnetisation of the agents neighbourhood multiplied by the spinstate of the agent. Periodic boundaries are considered in this case.

The third paper by Henry et.al [3] introduces agents on a network. Agents have a certain continuous property and they are connected through links, with a length defined according to the metric that is defined as:

$$d_K(x, y, K) = \begin{cases} \frac{1}{K} & \text{if } x = y, \\ \frac{\lceil K \cdot \frac{|x-y|}{2} \rceil}{K} & \text{if } x \neq y. \end{cases}$$

with x and y the property of the agent connected through the edge, and K a parameter such that only a certain amounts of lengths occur. K will be taken to be 4. These links will be 'short' if it connects similar agents, and will be 'long' if it connects dissimilar agents. At each turn an edge is selected and will be cut with a probability related to the length of that edge. If cut, the agent will form a link with a random agent. It can be proven, and we will show numerically, that the resulting length distribution of the links is independent on the initial network.

1.2 Results and discussion

Regarding the simple model of Schelling's original model, one expects the same dynamics when the fraction $\frac{\text{neighbourhood}}{\text{threshold}}$ is the same. This is generally seen in Figure 1.1. However, defining the magnetisation $\frac{N_+ - N_-}{N_+ + N_-}$, with N_+ and N_- indicating the number of agents in each group and denoting the tuple (threshold, neighbourhood), one sees a dissimilarity between (1,2) and (2,4); the first being not segregated, the latter segregated. This relates to the phase space of happy configurations: -++ and ++- for (1,2) and −+++, -+++-, +++−, +-+-+, +-++-, and -++-+ for (2,4). Even though the agents are less stringent about the state of their neighbourhood, the dynamics produce a more segregated configuration. Due to the additional rule of lowering the threshold, the result is an almost bi-diagonal matrix. Two exceptions are noted. When the magnetisation is too small with respect to the threshold, the minority is too sparsely distributed, and a configuration where an agent would be happy does not exist. Note that if cooperation is allowed, they could group together and be happy. Secondly, when the magnetisation is zero, the data matrix starts becoming tri-diagonal, indicating that segregation occurs even when agents are happy in a neighbourhood with more dissimilar agents.

Regarding the second paper, we'll first focus on the constrained model, meaning that the nodes only move when they are unhappy. The neighbourhood is defined as the two sites next to the agent. Spin states are used, with -1 and +1 indicating agents belonging to each group, and an empty site denoted by 0. Utility is defined as the sum of the states of the agent's neighbourhood. The empty site density, p, is defined as $\frac{N_0}{N_+ + N_- + N_0}$. Denote by u the density of unhappy agents. In Figure 1.2, u is plotted against m and p. Visualizing the figure, several points can be noted. First, u_∞ is supposed to vanish at $m = 0$ and $m = 1$ and for large enough p. This occurs because, as long as there is one empty site, nodes can gradually improve their utility, and a configuration where everybody is happy is always obtainable (assuming $m = 0$). For $m = 1$, by definition, there are no unhappy nodes. For large enough p, it vanishes because nodes, in the worst-case scenario, live on their own, which is a happy configuration when $f = 0$. The reason for the peak in unhappiness is due to the decrease of the minority. The minority will be, on average, less satisfied the smaller it becomes, but this effect is negated by the increase in the happy majority. Hence, on average, more agents are happy.

In the analysis of unconstrained dynamics, a comparison is made between neighborhood sizes of 2 and 4, with the threshold set at zero and slightly above. This means that agents are happy if their neighborhood is equally split in the first case, and require a slight majority in the latter. This comparison is done for both binary (Figure 1.3) and non-binary (Figure 1.4) utility functions. The plots show how the dynamics of a random initial configuration evolves depending on the neighborhood and the threshold, in function of p. The density of unhappy agents and the density of interfaces (density of opposite spin-states neighboring each other) are also shown. For the non-binary utility function, focusing on the highest values of p, it is seen that a larger neighborhood induces more segregation than a smaller neighborhood, similar to what was seen in Schelling's original model. A transition seems to happen in the case of the smaller neighborhood when pp increases. Local structures form but do not persist. Note also that segregation happens equally for all spin-states, in contrast to the binary utility function where segregation to opposite spin-states is seen, along with random patches of empty sites. Taking our focus to the binary utility function and still looking at the high density of empty sites, segregation remains when the threshold is raised above 0. This is expected since the only happy state for a neighborhood of 2 would be +++, and for 4 it would be +++++, +-+++, and their permutations. In this latter case, the opposite spin would not be happy and would move. Furthermore, this raise in threshold can constrain the unconstrained model when the density of empty sites is too low. Regarding the densities of unhappy agents and of interfaces, these converge rather fast, at about 100 iterations, after which they fluctuate, with 2 exceptions (Figure 1.4, plot 2, p=[0.5,0.3]). In general, the higher the p, the faster the convergence. The oscillations are caused by agents moving to empty sites, lowering the magnetization of the neighborhood of their old neighbors. In the case of smaller p, and as long as the dynamics are still unconstrained, complete segregation happens with a buffer of empty sites.

Regarding the third paper, a simulation is ran for 100 nodes, who have an attribute randomly sampled from U(-1,1). Four graph are used, of which Erdos-Renyi and Barabasi-Albert. The two other graphs are denoted as 'long links' (LL) and 'short links' (SL). The LL graphs connects all nodes if the corresponding edge would have maximal length, SL if minimal. At each time-step, a random edge is selected, and will be cut with probability $p \times length$, with $p \in (0, 1]$. It is seen that convergence is faster, the higher p, but the end distribution does not depend on p.

The general trend is found back, all four graphs their edge-length distributions converge to similar values. However, finite size effects are present, and the distributions are averaged over multiple realizations.

1.3 | Appendix

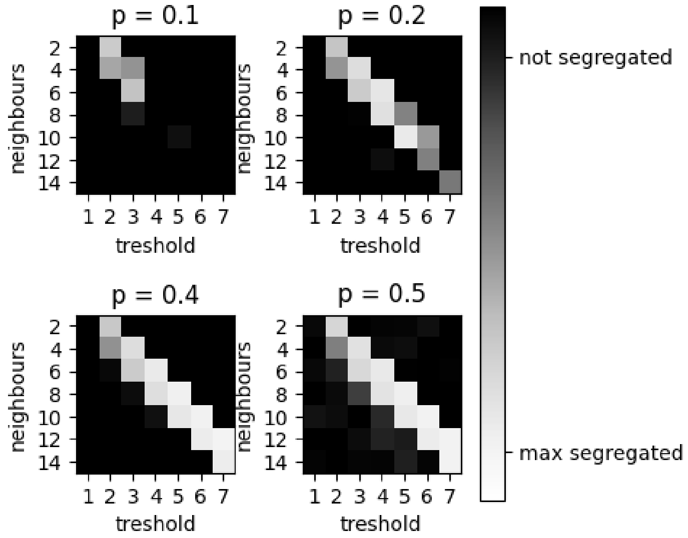


Figure 1.1: *Segregation in function of threshold. Simulation done with 100 agents. Segregation scale is linear and normalised to minimal possible segregation (maximal possible groups) in system with given magnetisation.*

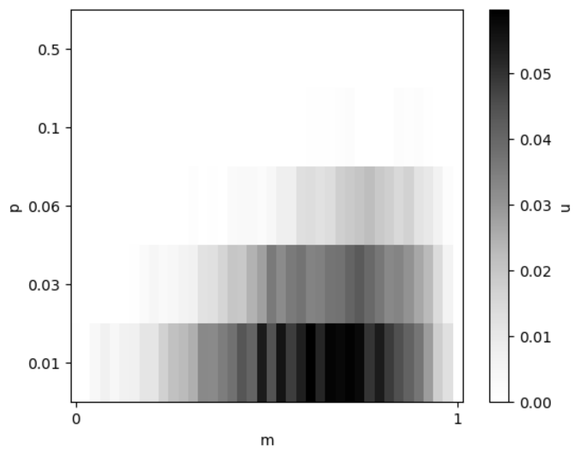


Figure 1.2: *Density unhappy agents in function of m and p. 1000 agents are considered and averaged over 30 realisations.*

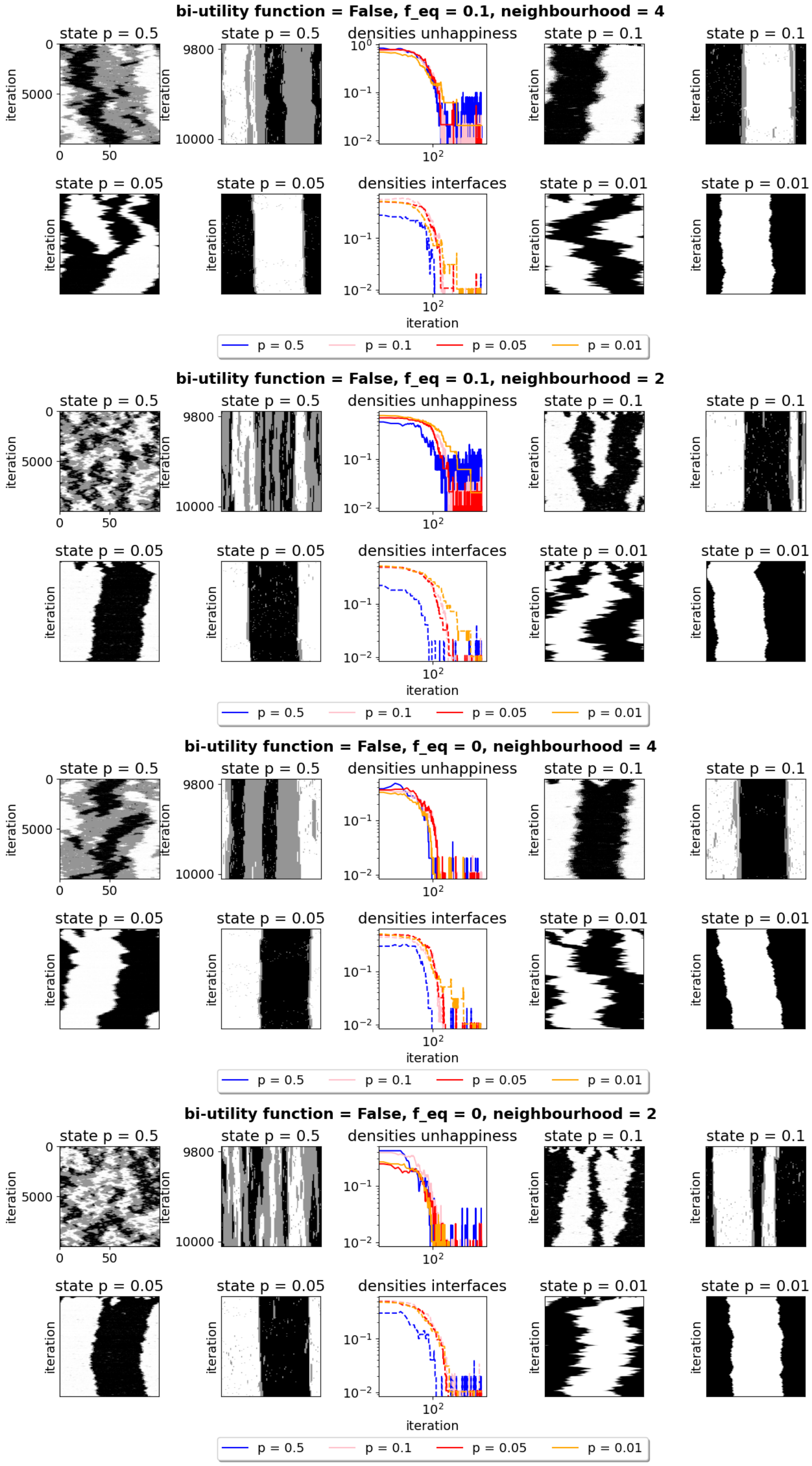


Figure 1.3: *Dynamics according to unconstrained model in function of threshold (f_{eq}) and size neighbourhood for non binary utility function.*

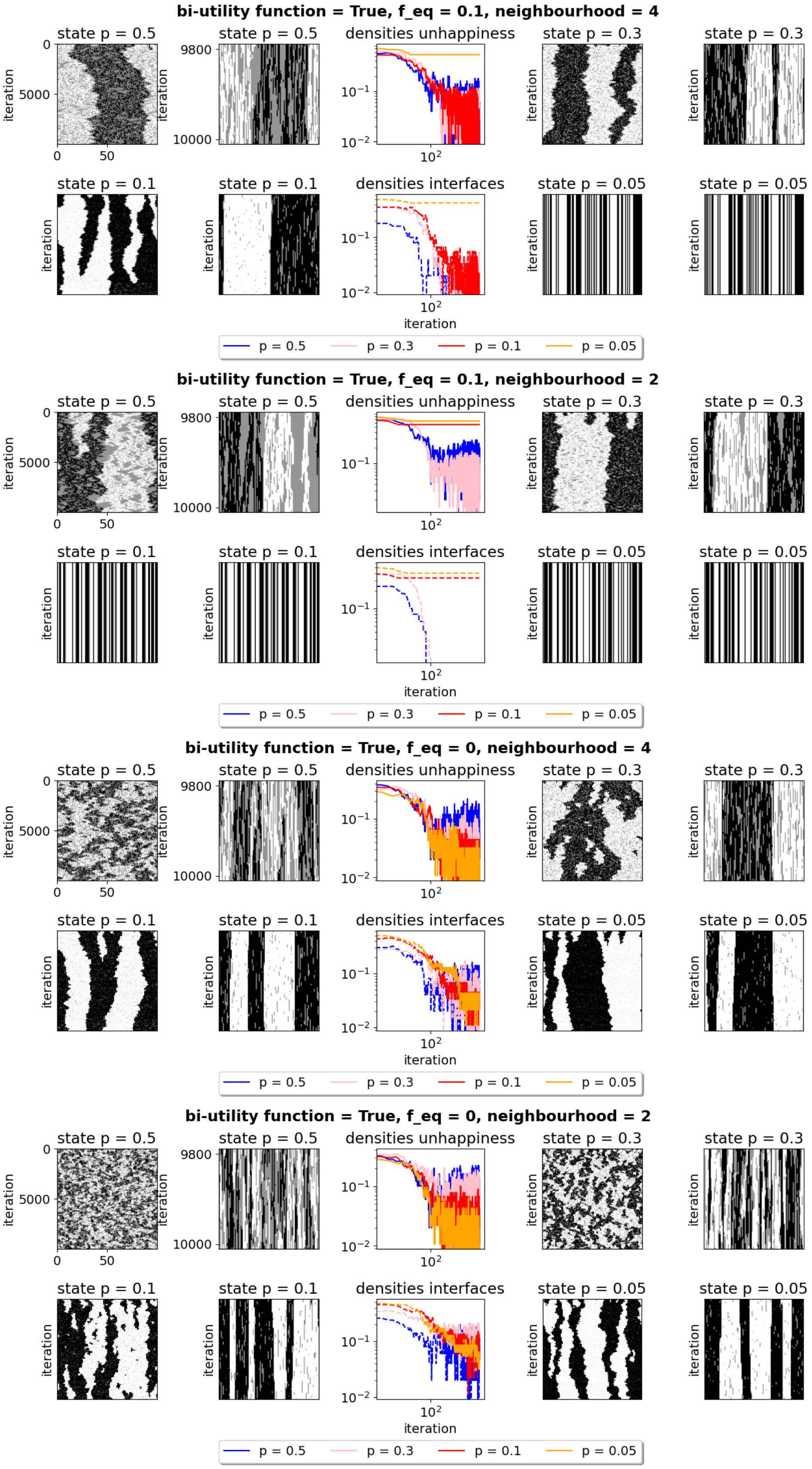


Figure 1.4: *Dynamics according to unconstrained model in function of threshold (f_{eq}) and size neighbourhood for binary utility function.*

2 | 34. Game theory on networks

Task leader(s): *Jules Vandendriessche*

2.1 | Introduction to social game theory

In this task, game theory on graphs will be examined. Agents are connected through a network and play games with their neighbours, according to a certain strategy. During the games they acquire utility. After a round of games is over, they will update their strategy. Hence, agents have tree properties: strategy (SRAT), update rule (UR) and payoffs. Numerical experiments will be discussed for 2 different games and 3 update rules.

One of the two games, used by Cardillo et. al. [1], is the weak prisoners dilemma, The second, used by Sinatra et. al [5], is the weak ultimatum game.

The weak prisoners-dilemma (PD) is a game defined according to the following payoff matrix, where the strategies are cooperate (C) and deviate (D) and b>1, such that it induces agents to deviate.

	C	**D**
C	1	0
D	b	0

Table 2.1: The weak prisoner's dilemma payoff matrix. Payoffs are symmetric with the exception of (D,C), where the player playing D obtains payoff b.

This problem can be translated into a spin-state problem. Given the adjacency matrix $\mathbf{A}$ and the strategy vector of the $\mathbf{s}$. The sum of the strategies of an agents neighbours can be calculated according to the vector equation:

$$\mathbf{n} = \mathbf{A}\mathbf{s}$$

Then, the payoff vector is given by:

$$\mathbf{\Pi_i} = (1 - \mathbf{s_i})\mathbf{n_i} b + \mathbf{s_i}\mathbf{n_i}$$

The ultimatum game (UG) is a game where each player chooses a strategy tuple $\mathbf{strat} = (p, q) \in [0, 1] \times [0, 1]$, where p is the offer value and q is the acceptance value. Each player plays a game with all their neighbors. For each game between player i and player j, the payoffs are determined as follows: If $p_i > q_j$, then player i receives $1 - p_i$ and player j receives p_i. Otherwise, players receive nothing. If $p_j > q_i$, then player j receives $1 - p_j$ and player i receives p_j. Otherwise, players receive nothing.

REP is an update rule. The agent (i) picks a random agent (j) to whom he is connected and it compares its payoff (f). If $f_j > f_i$ then agent i will adopt his strategy with probability $\frac{f_j - f_i}{Nmax(k_i, k_j)}$ with k denoting the degree of the agents and N = 2 for UG and N = b for PD.

UI is an update rule. The agent compares its payoff with its neighbour with the largest payoffs. If its payoff is bigger, it will copy its strategy.

MOR is an update rule. The agent (i) copies the strategy of their neighbour (j) with probability $\frac{f_j}{\sum_l (A_{il} f_l)}$.

2.2 Results and discussion

The PD is played by agents with 2 properties, a strategy and an update rule. These will co-evolve together, meaning that when an agent copies a strategy, it also copies its update rule. A comparison is made between 2 UR at a time: REP vs MOR, REP vs UI and UI vs MOR, in function of the initial distribution of the UR and the value b. This is done for both for ER and SF networks, both with <k> = 6. The initial distribution of the strategy is distributed according U(0,1), cooperation indicating 1, deviation 0. The analysis is done for a limited number of nodes (100) and averaged over 10 realizations. The game is played for 1000 iterations or until the 30 last values of both UR and STRAT have a standard deviation of less then 0.0005. Payoffs are not accumulated.

The results are shown in figure 2.1. Discussing the strategies first, one sees that REP dominates MOR in both SF and ER, and gets dominated by UI, except in the case of high initial condition. In all four cases, very little dependence of b is is seen. In contrast with UI vs MOR for the UR, where non trivial behaviour is seen for ER and in lesser degree for SF. In case of the strategies, ER shows clear behaviour in funtion if the inital UR distribution and b, while SF is more "over the place" indicating that the game might had to be played on for longer. Behaviour is shown to be rather similar in all 3 cases. For ER, Agents cooperate until a certain threshold has been reached. Only for UI vs MOR is a small dependence seen on the initial distribution. As said earlier, SF networks show a more independent behaviour of the parameters, only playing D more often in the case of UIvsMOR, indicating that agents are more likely to cooperate when PER is available.

The UG is played by agents who have 2 properties, p and q as defined earlier. For sake of simplicity, q = 1-p. Note that this means that payoff will only be obtained if pi + pj > 1, for any 2 connected agents. This problem has 2 important Nash equilibria. The first one where all agents play 0.5. The second one, less trivial, where half plays p = 1 and the other half plays p = 0, because neither have reasons to deviate. The last one is sub optimal. Note that being completely altruistic (playing p = 1) does not increase the payoff.

Finally, a social rule will be introduced, where the poorest agent and its neighbours are replaced by new agents without payoff and new strategies. This has as consequence that agents have to care also about the well-being of their neighbours.

In absence of the social rule, one sees from the figure 2.2 that the distribution converges quickly , after 25 % of the iterations the distribution remains nearly constant. A large hub is clearly beneficial for reaching the optimal Nash Equilibrium and it happens faster for SF than for ER. Of the 3 update strategies, UI works the fastest (as can be seen by the lack of blue in the plots) but it does have the drawback that all agents might switch to a value different from 0.5 on the same turn, and the ideal strategy dies out.This could be solved by introducing the probability that agents change their strategy at random, such that all strategies are kept alive. This is however not done. Unsurprisingly, REP is the slowest of the 3, as an agents is more prone to remain unchanged.Regarding ER, a slight dependence on the mean degree is seen for the convergence . The same trends are seen for the other two UR, UI is the fastest while REP is the slowest.

When implementing the social rule, the role of the central hub is dependent on the UR. During the initial stages of the game, opportunity takers have an advantage. Because of the rule UI, once a large hub has become an opportunity taker, the other agents will be troubled to show altruistic behaviour. The reverse is seen for MOR. Altruism is seen, as the dominant strategy is playing p =1. REP is again the slowest of the 3, with the strategies more widespread but showing clear altruism.

The distributions of ER show that convergence has yet to be reached. Note that altruism almost always goes together with opportunity takers.

2.3 | Apendix

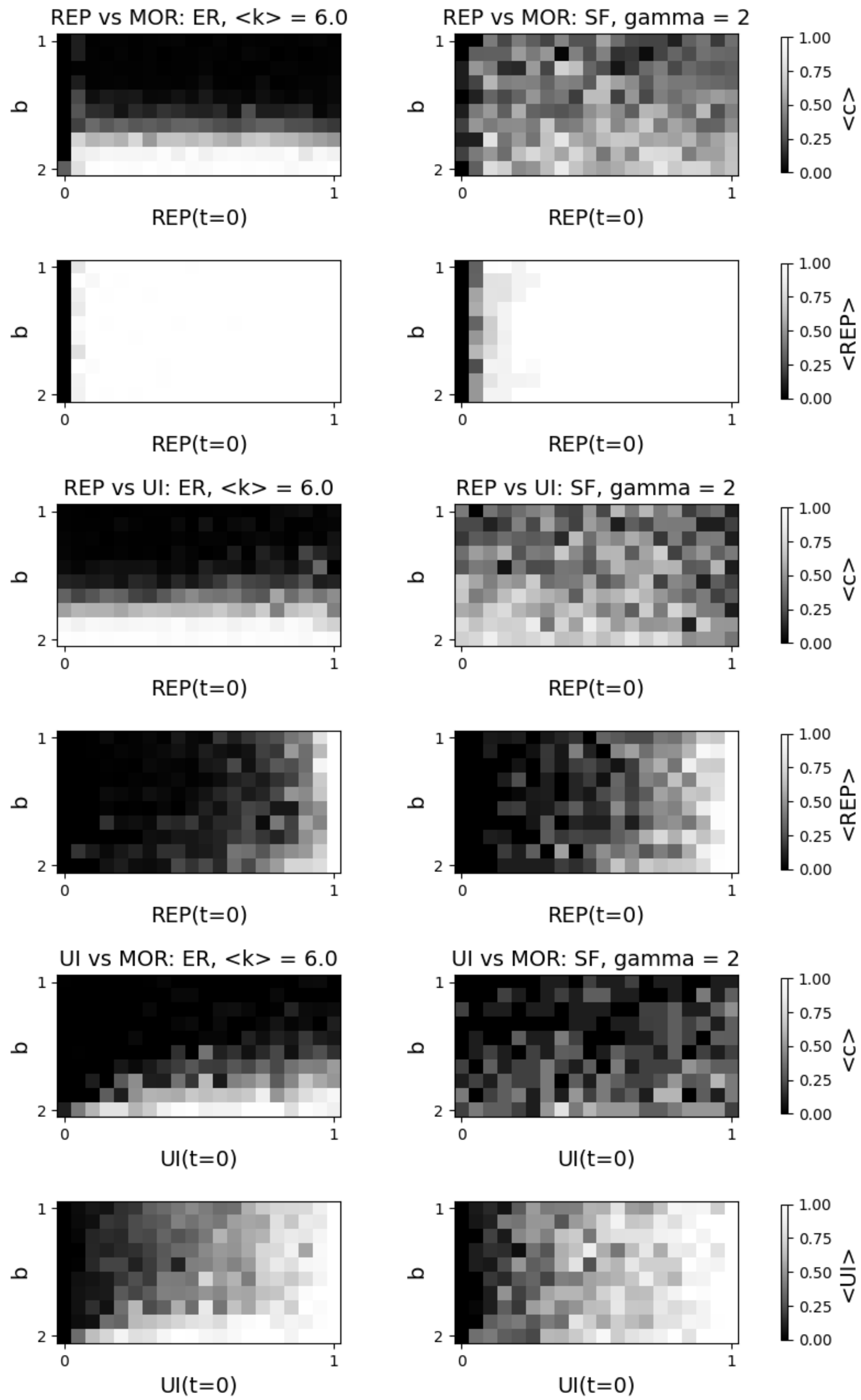


Figure 2.1: *Weak prisoners dilema played by 100 nodes. Played for 1000 iterations or until convergence of UR and STRAT. Averaged over 10 realizations.*

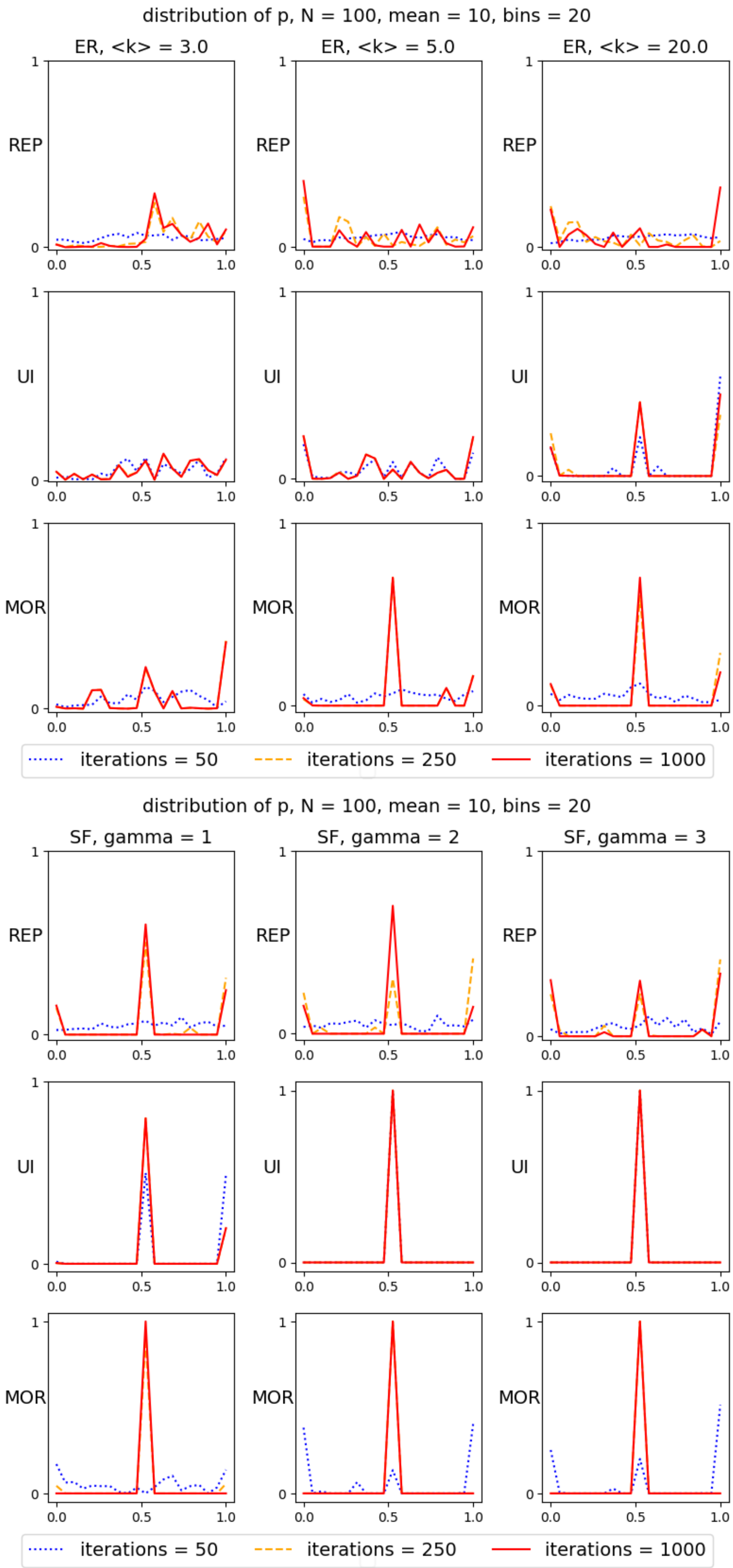


Figure 2.2: *Ultimatum game without social penalty. Game played with 100 agents and averaged over 10 realizations.*

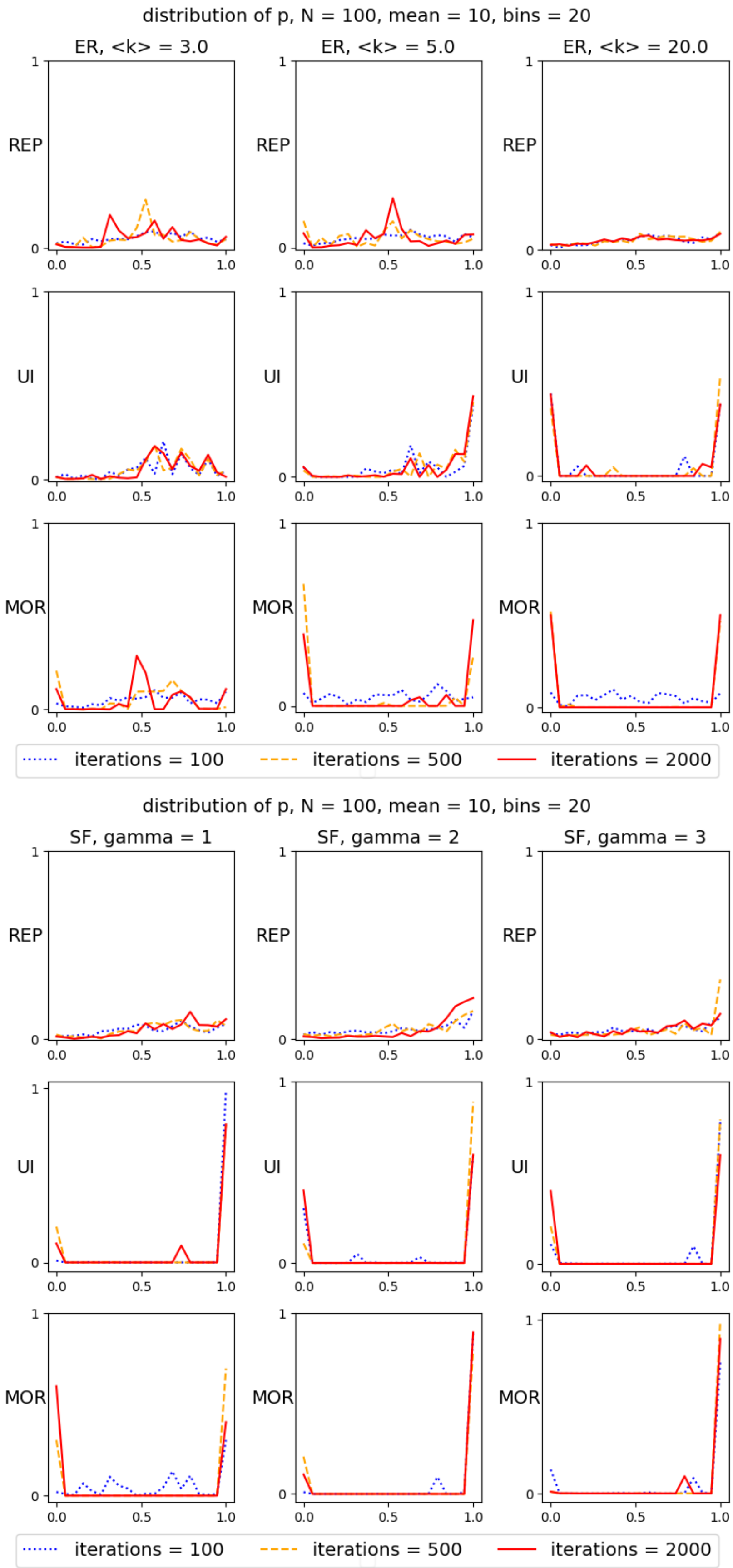


Figure 2.3: *Ultimatum game with social penalty. Game played with 100 agents and averaged over 10 realizations.*

3 | 45. European transportation networks I

Task leader(s): *Jules Vandendriessche*

3.1 | Introduction

Since this is a data project, the focus lays more on how the output is achieved than on its meaning. The data is from EuroGlobalMap and comes with a data catalogue.

In this dataset, one finds a selection of European countries and Europe as a whole. For each of this countries files are provided for different topological attributes, of which 2 are of our main interest, node-like and edge-like. Examples of node-like would be stations and airports while examples of edge-like are country borders and roads. Node-like data sets exists out of a geological point with attributes, while edge-like has 2 points and attributes. According to the catalogue, every edge has to start and finish at a gives node in one of the node files.

The goal of this project is to reconstruct the rail network of the given countries and of Europe.

3.2 | Results and discussion

Before constructing the network, one has to ask itself what kind of network it is interested in. In the case of a train network, one can be interested in the network decoupled from the physical topology. Meaning that all nodes are stations and an edge between nodes exists if the stations are directly connected without going through another station and disregards everything else. However, a train network has more than only stations, level crossings, merges and ports do play a role in the network. In this task, I opted to go for the second one.

Practically this mean a reversal of the problem, instead of connecting the stations through the edges, the edges start and finish points are taken as nodes and stations are connected to these nodes. These nodes can be a number of other things. From the catalogue one finds that these are: Exits, ferry stations, level crossings and airports.

Furthermore, in the data one finds that "Railroad lines can only touch at their ends and must not overlap each other." No file is found to pinpoint these points. One could make use of the exclusion principle and name the unknown points 'rail merges', but I opted to call them unknown. Note that rails are given 3D information by their property 'LLE', it indicates on which level the rail can be found. When a rail is raised, the LLE goes up by one number. This is done by cutting the edge and starting a new one at the appropriate level.

Figure 3.1: *Satellite picture of 50°53'28.0"N 4°25'19.7"E. Example of a crossing which might give trouble when reconstructing the "flow" of the network.*

To reconstruct the network, the following algorithm has been applied:

1. Use the locations and the LLE of the begin and startpoints of the provided edges to construct all possible nodes. Give these a unique ID.
2. Check if the node locations appear in the other provided files for the nodes. If so, label the node accordingly.
3. Edges that are raised are cut in the data. These have to be linked back together. This is done by selecting all the nodes whose location appear twice but with different LLE. Then map one ID on the other, both in the nodes file and in the edge file and delete the nodes that are mapped away.
4. After that the data can be cleaned up. Nodes labeled "unknown" that only connect 2 edges are deleted as are double edges and self loops.

The algorithm can be adapted such one obtains a network with only stations. It has been found that the dataset is not complete, as can be seen in picture 3.2.

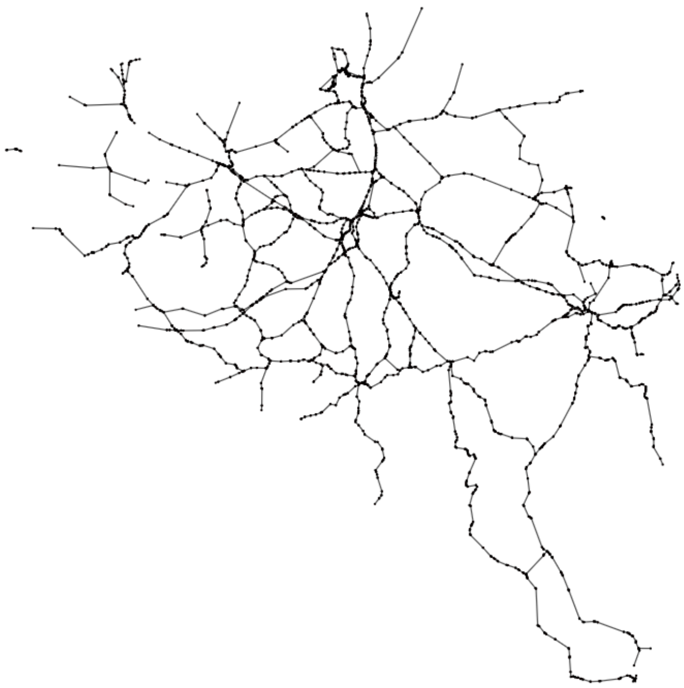


Figure 3.2: *Train network of Belgium. Note the 3 disconnected components in the north-east and the lone nodes in the west.*

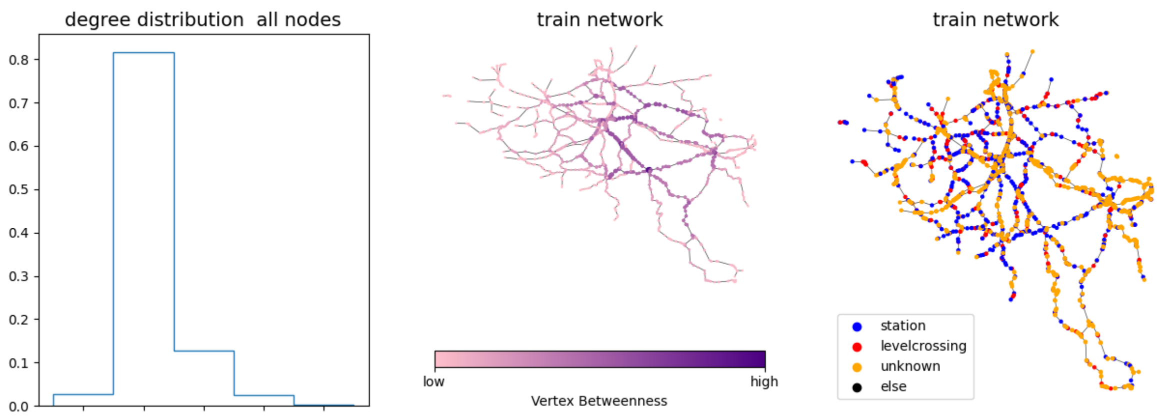


Figure 3.3: *Train network of Belgium after applying before mentioned algorithm with node degree and betweenness.*

4 | Bibliography

[1] Alessio Cardillo, Jesús Gómez-Gardeñes, Daniele Vilone, and Angel Sánchez. Co-evolution of strategies and update rules in the prisoner's dilemma game on complex networks. *New Journal of Physics*, 2010.

[2] L Dall'Asta, C Castellano, and M Marsili. Statistical physics of the schelling model of segregation. *Journal of Statistical Mechanics: Theory and Experiment*, 2008.

[3] Zhanh Henry, Pralat. Emergence of segregation in evolving social networks. *PNAS*, 2010.

[4] T. C. Schelling. Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1971.

[5] R Sinatra, J Iranzo, J Gómez-Gardeñes, L M Floría, V Latora, and Y Moreno. The ultimatum game in complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2009.