The Dynamics of Trust – Emergence and Destruction

Dominik Klein
TiLPS
Tilburg University
d.klein@uvt.nl

Johannes Marx
Political Science Departement
University of Bamberg
johannes.marx@uni-bamberg.de

Abstract

We study the emergence and evolution of trust in larger societies. We focus on the thin notion of trust, that is the trust needed for interacting with hitherto unknown individuals encountered for just a single interaction. Our model builds upon well established theoretical knowledge of the determinants of trust. These works identify parameters such as the existence of networks, the level of mobility or the percentage of trust-abusing agents in a society. While the influence of each of these factors individually is well-established by empirical work, a precise account of the interplay of these factors is lacking. To bridge this gap, we devise a multi agent computer simulation that allows a fine grained analysis of the dynamic processes governing the emergence of trust and its dependencies upon these parameters. We model agents using a bayesian learning framework about the value of trust, taking both individual and social information into account.

1 Introduction

Trust is a crucial ingredient for the well functioning of human interactions. It appears in such diverse situations as trust in the functioning of institutions, trust in the content and sender of some message or trust in the future behavior of some business partner.

Trust describes a belief in (future) actions of others in the absence of compelling external reasons to do so. Recent conceptual work [Castelfranchi and Falcone 2001] gives a more fine grained analysis of the beliefs relevant for placing trust in others and subsequently delegating a certain action to them. In the analysis of [Castelfranchi and Falcone 2000], trust refers to an entire package of beliefs such as the belief in the other person's intention to execute the task delegated to her as well as her capability of doing so. Further analysis [Castelfranchi and Falcone 2001] adds additional parameters, such as external obstacles to the successful execution of the task. Equally, also the trustee's willingness to execute the task delegated to him can be influenced by several factors, such as his relationship to the trustee, but also the fear of legal prosecution or therelike.

The stepwise process in placing trust in others, together with the relevant vulnerabilities are modeled in game theoretic trust games. The first player, the trustor has to make a decision whether to trust the trustee or not. By doing so, he becomes vulnerable towards the trustee returning rather than exploiting the trust placed in him. In this paper, we depict trusting interactions as the play of a trust game between the two agents. Of course, the actual interaction the agents are involved in might be much more complex. The trust game merely mirrors the trust-related component of such an interactive situation.

The literature in the social sciences distinguish between a thick and a thin notion of trust. *Thick* trust is a general attitude towards individuals who we have regular interactions with and with whom we build up a general

level of dependency. On the other hand *thin* trust refers to the attitude we need for singular interactions with other members of the society that we haven't encountered before nor expect to see again. It is this attitude of trust that features in the highly circulated current PEW-report[1], showing a generational decline in trust by eliciting people's attitudes towards the claim "'Most people can be trusted" .

In this paper we are interested in the emergence and dynamics of the thin notion of trust. Thin trust by its very definition, precludes any type of personalized learning or reputation functions about individuals. It is the notion of trust relevant for buying something from an unknown vendor on craiglist, not the trust slowly build up towards your bank advisor through iterated interaction. Consequently, we do not model any type of reputation function - neither for individuals nor for groups. In this approach, we furthermore restrict the ways in which agents can learn about the expected value of trust to their own experiences as trustors or trustees. In our approach we follow [Falcone and Castelfranchi 2004], giving a detailed discussion on how a trustee's experience might influence his trust level, i.e. his potential behavior as a trustor.

Notably, we are not primarily interested in a faithful representation of the dynamics of thin trust in actual societies. We rather aim at a conceptual understanding of the interplay between various factors present and relevant for the emergence of trust. We hold that our model is informative about many processes in the actual and digital world while at the same time being rich enough for assessing the functioning of various trust-related explanations to be found in the social science literature. However we will say little about the latter aspect in this paper.

To the best of our knowledge, the notion of thin trust is the subject of several informal explanations in the social sciences, but hasn't been attacked with formal or computational models yet. Somehow orthogonally, many common approaches study the impact and design of reputation functions and personalized learning, see for instance [Birk 2001] or [Nooteboom, Klos and Jorna 2001] for simulation approaches. On a more conceptual level, [Jonker and Treur 1999] offer a conceptual framework for discussing trust updating based on experience, anticipating the bayesian updating process used for our framework (p.10).

Empirical research identifies various factors relevant for the emergence of trust, such as mobility (social and geographical [Putnam 95]), cultural background [Fukuyama 2006, Guiso, Sapienza and Zingales 2008] and network structures [Portes 1998, Burt 2000]. Based on these parameters, we set up an agent based computer simulation to gain a fine grained understanding of the dynamic processes governing the emergence of trust and the interplay of the mechanisms involved. At the same time this simulation constitutes a validity check for certain theoretical models in the literature. The simulation is based on the above mentioned theoretical findings combined with some recent findings from bayesian learning theory.

2 The Setup

Following the analysis of [Castelfranchi and Falcone 2000], the decision to trust somebody is a multidimensional decision problem, incorporating assessments of parameters such as the beliefs that the potential trustee is committed and that he is capable, but also upon risk dependent decision thresholds. These beliefs are based upon prior encounters with the person in question or others that are judged similar, but also upon indirect testimony collected in reputation functions or general knowledge of the societal embedding, such as the risk of legal prosecution for trust abusers.

On the other hand, the trustee's behavior is guided by similar factors. It is driven by his general mental makeup, but also by the fear of legal prosecution, considerations for his reputation, normative influences and many others. For this study we make two idealizing assumptions. First, we assume that by the nature of the interaction modeled, the above cited multidimensional conceptualization of trust can be collapsed to a one-dimensional bayesian decision problem, that is agents combine their assessments in a singe-dimensional measure on a [0-10]-scale and then act upon a decision threshold. In a somehow sloppy way, we refer to this single-dimensional measure as *the expected value of trust*. Second, we assume that the factors driving the trustee's behavior, be it normative reasoning, fear for prosecution or others, do not change throughout the simulation. Consequently, every individual trustee's behavior remains the same throughout the simulation.

Furthermore, we equip our agents with a very limited access to new information about relevant parameters for trust. There is no communication passing between the agents, the only information agents can access is their prior experience as both, trustor and trustees. This approach renders trust an even more fragile concept than it already is. Scarcity of information combined with the lack of social embedding increases vulnerability towards frustration that can arise through a short series of unsuccessfull interactions. A moderate cascade of consecutive

---

[1]retrieved on 10-03-2014 from http://www.pewsocialtrends.org/files/2014/03/2014-03-07_generations-report-version-for-web.pdf

trust abusive encounters can transform even the most positive agent into a skeptic, unwilling to engage in any more trust games. On the other hand, agents with a negative trust expectation refuse to assume the role of a trustor in future interactions, thereby depriving themselves of the main source for reassessing the expected value of trust. The only informational source these agents have to overcome their skepticism is indirect learning. Even if unwilling to act as a trustor, these agents might still be chosen as trustee. If an agent experiences trust placed in him by another person he learns about the trust level of his partner, thereby gaining indirect information about the trustworthiness in society.

3 The Model

We set up an agent based Netlogo simulation, accomodating a set of 1500 agents each individually setting out to learn the expected value of trust. These agents roam around a two-dimensional grid of size 51×51. To homogenize the model, the grid is wrapped to a torus, i.e. agents moving across the top edge reaper in the bottom part and vice versa - the same holds for left and right. Our simulation is round based, with each round consisting of 3 different phases. In the first moving phase all agents relocate to a new spot, followed by a partnering phase in which agents team up with neighboring agents[2] to trustor-trustee pairs. Each field can only occupied by one agent at a time and each agent can only be member of one pair per round, so some agents end up without a partner. Finally, in the playing round trust games between trustor and trustee are performed and the appropriate updates of the agent's trust memories are calculated. The parameters chosen for this simulation are informed by empirical research on the determinants of trust as well as by bayesian learning theory. In the following we describe the input parameters of the program

Mobility denotes the distance each agent covers in the moving phase. The higher this parameter the less likely a society is to develop local inhomogeneities. There is a crucial difference between a mobility of 0, describing static societies with constant neighborhood relations and a positive mobility, describing a dynamic society. In this paper we exclusively focus on positive levels of mobility.

The second parameter is **%-defectors**, the actual share of trust abusing agents within a society. As defended above, we assume for this simulation that the individual agent's behavior as trustee remains invariant throughout the simulation, thus %-defectors is constant in time.

Starting-trust-memory is the mean value of the agents' initial assessment on the value of trust. The individual agents starting-trust is drawn from a random distribution around this mean value.

The fourth parameter, **learning** denotes the relative epistemic weight an agent attributes to newly incoming information relative to his previous estimates. The higher this parameter, the stronger an agent's reactions to newly incoming information accelerating his learning process, but also making him more susceptible to lose trust when presented with a short stream of consistently negative experiences. This parameter is the least manipulable by a market designer, as it is hard wired in the agents' updating process. Empirical studies determine this parameter to lie anywhere in between 3% and 10%, see [BME 2011]. All subsequent simulations are a straight average for *learning* values between 3 and 10.

Memory is a binary parameter, denoting whether agents have a limited capacity to learn about the behavior of individual others. Empirical results identify the existence of memory or networks as a relevant parameter for the emergence of trust. If the parameter memory is set true, agents do have a limited memory allowing them to remember trust returning interactions. If agents are asked to pick a partner for a trust game, they preferably pick partners they remember as trustworthy, if such partners are available. In our basic simulation memory are disabled. We use this parameter only for an extension of the model, inquiring whether memory have a positive effect.

Social-Factor. Agents in our simulations draw their information about the expected value of trust from two different sources. The first, direct source is the experience gained while acting as a trustor. The second, indirect source, described in more detail below, is the behavior of other trustor's approaching this agent as a trustee. Besides epistemic facts, the agents can also have other considerations towards information of the second type: A taste for social uniformity can put additional weight on the social information, while agents might be cautious

---

[2] We work with von-Neumann neighborhoods, i.e. each field has 4 neighbors

towards second order information if they are doubtful towards the sincerity or learning competence of their peers. The social factor describes the relative weight an agent attributes to his indirect learning experiences relative to direct information. *Social Factor* assumes values between 0, indicating that no social learning takes place and 2, putting a strong emphasis on the value of social information. Unless noted otherwise, the social factor is set to 1, thus treating direct and social learning on par.

3.1 Agents and their learning process

The primary value we are interested in is the agent's trust, that is their expectation that a randomly chosen agent from society returns trust rather than abusing it. We model this expectation by a variable **trust-memory** in the interval $[0; 10]$, where 0 denotes the expectation that trust will certainly be abused while 10 stands for the belief that placing trust is beneficial, no matter what. A rational agent will engage in trust games if his expectation to meet a trusting agent is high enough, that is if *trust-memory* is above a certain threshold. In general, this threshold is the product of a complicated process, taking into account the exact payoff structure, but also the agents risk aversion, his absolute stakes and many other parameters. In this model we abstract away from the individual background parameters, combining them to a general decision rule whether the agent is willing to engage in trust games or not. We set the threshold variable to 5, thus we apply the following learning rule

$$\text{Play trust if } trust - memory \geq 5, \text{ else do not play.}$$

At the beginning of the simulation, agents are initiated with a trust-memory drawn from a random distribution around the global input parameter *starting-trust-memory*. The agents' trust memory then gets updated with every new piece of experience they gain through trust games. For each interaction, agents are paired up in trustor-trustee pairs. Depending upon his trust memory, the trustor decides whether to place trust or back out. In the former case he learns about the trustee's trustworthiness. The trustee, in turn, receives information on whether the trustor was prepared to place trust or not. These information give rise to the following updates on *trust-memory*.

In case the trustor decides not to engage in a trust game he does not receive any new evidence and his trust-memory remains unchanged. If the trustor engages in a trust game he learns whether the trustee cooperated ($E = 10$) or defected ($E = 0$). The trust memory is then updated by the weighted average

$$\text{trust-memory}_{new} = (1 - \beta) \cdot \text{trust-memory}_{old} + \beta \cdot E, \tag{$*$}$$

where β is the input parameter *learning*, the weight the agent is willing to attribute to the newly incoming information.

On the other hand, the trustee receives the information whether the trustor was willing to place trust in him ($E = 10$) or not ($E = 0$). Assuming that this willingness reflects the trustor's informational state, the trustee updates his beliefs upon this information. We refer to the trustee's learning experience as *indirect* or *social* learning. Again, the updating is done through a formula similar to $(*)$.

$$\text{trust-memory}_{new} = (1 - \delta \cdot \beta) \cdot \text{trust-memory}_{old} + \delta \cdot \beta \cdot E \tag{$**$}$$

where β again is the *learning* and δ is the parameter *social factor*, that is the relative weight the agent is willing to attribute to his second order information relative to his direct learning experience. For a distrusting agent, social learning is the only means of acquiring new evidence about the system.

4 Robustness and Pretests

The robustness of the input parameters was checked in several pretests. For the geometry of the state space we cross checked on a larger board size of 201×201 – as well as for a board size of 51×51 with added geographical inhomogeneities. Both did not change the results significantly, with one exception about clusterings mentioned below. In a further experiment we also checked against different sizes of the population, running from 900 to 2100 agents. Also here, no significant impact of the number of agents could be found, hence the restriction to a medium population of 1500 agents. Furthermore, the states of universal trust and universal distrust are global

Table 1: Impact of *%-defectors*

percent-defectors	all	mobil. = 1	= 2	= 5	= 10
		Share of *trusting* sim.			
30	0.584	0.583	0.589	0.578	0.589
32	0.545	0.536	0.563	0.542	0.547
34	0.521	0.510	0.521	0.521	0.516
36	0.504	0.5	0.5	0.5	0.505
38	0.5	0.5	0.5	0.5	0.5
40	0.5	0.5	0.5	0.5	0.5
42	0.5	0.5	0.5	0.5	0.5
44	0.5	0.5	0.5	0.5	0.5
46	0.5	0.5	0.5	0.5	0.5
48	0.499	0.495	0.5	0.5	0.5
50	0.499	0.484	0.5	0.5	0.5
52	0.498	0.453	0.5	0.5	0.5
54	0.487	0.438	0.474	0.49	0.49
56	0.465	0.411	0.469	0.474	0.474
58	0.436	0.365	0.432	0.443	0.433
60	0.415	0.323	0.417	0.427	0.417
average	0.497	0.475	0.498	0.498	0.499
total number	15272	1459	1529	1531	1532

attractors: All simulations converge towards one of these states. Furthermore, both states are stable, i.e. once a simulation converged towards one of these states it will not leave them again. As in turns out, the entire convergence process has been completed within 1000 rounds. Hence, taking final measures after 1000 rounds of interaction measures the final emergence of trust or distrust in the system.

5 Results

Unless noted otherwise, all final measures have been taken after 1000 simulation rounds of moving and potential interactions. The results presented here are based on 30720 simulation runs, 10 for each combination of parameters. Using the full product space of parameters gives rise to certain improbable starting conditions such as agents that are almost certain about the trustworthiness of others, while at the same time the number of trust abusing agents exceeds the trustworthy ones by far. Such improbable data points blur the quantitative outcome, but we hold that these artifacts do not affect the qualitative picture obtained. Our primary target of interest is the functionality of the underlying social group, that is the share of agents that are willing to engage in social interaction by assuming the role of a trustee. By our decision rule, these are exactly the agents with a trust-memory larger than 5, thus we choose as our output measure:

$$trust\ level = \frac{\text{number of agents with trust-memory} \geq 5}{\text{number of agents}}$$

As noted above, almost all our simulations converge towards the extreme values of *trust level*, that is either a state of universal trust or universal distrust. To simplify notation, we call a simulation *trusting* if *trust level* ≥ 0.8. As experience shows, simulations with a *trust level* above this threshold run inevitably towards a stable state of universal trust.

Our first results evaluate the validity of our model. As could be expected, both the average percentage of trust returning agents (table 1) as well as the initial amount of trust at the beginning of the simulation (table 2) correlate positively with *trusting*.

After validating the initial model, the next influence to be checked is the relationship between mobility and trust. *Mobility* determines how many steps the agents move per round. It thus measures the velocity in which the social context of the agents changes. Mobility in our model can thus represent geographical as well as social mobility or other related changes in the agents' environment. Current literature predicts that a high mobility

Table 2: Impact of *starting trust degree*

starting-trust	all	mobil. = 1	= 2	= 5	= 10
		Share of *trusting* sim.			
3	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0
5	0.58	0.49	0.64	0.53	0.59
6	0.927	0.838	0.922	0.938	0.934
7	0.999	0.98	1	1	1
8	0.999	0.982	1	1	1
average	0.497	0.475	0.498	0.498	0.499
total number	15272	1459	1529	1531	1532

Table 3: Effects of mobility

mobility	Share of *trusting* sim.	mobility	Share of *trusting* sim.
1	0.475	11	0.498
2	0.499	12	0.497
3	0.497	13	0.498
4	0.499	14	0.496
5	0.499	15	0.495
6	0.499	16	0.503
7	0.500	17	0.498
8	0.495	18	0.499
9	0.498	19	0.499
10	0.499	20	0.499

has a negative impact on the emergence of trust, [Putnam 95]. Our simulation does not support this claim. In fact, a mobility of 1 proves detrimental to the emergence of trust, while higher values of mobility do not have a significant influence on the emergence of trust, see table 3. We conjecture that this effect is due to local clustering. A mobility of 1 is low enough to allow for the development of local clusters of trust or distrust, while a higher mobility impedes such local effects. This claim is supported by an analysis of the local inhomogeneity (modified Bray-Curtis index of similarity), measured for a division of the arena into 3×3 square districts of equal size. The measures of dissimilarity are displayed in table 4. In particular, our simulation predicts that local inhomogeneities are detrimental to the emergence or maintenance of trust. Since distrust is more stable than trust, local clusters of distrust are contagious to neighboring fields, gradually spreading out to the entire society.

We conjecture that this clustering is not exclusively caused by the mobility value of 1, but can explained by the interplay between mobility and the relatively moderate field size of 51×51. For larger field sizes of 201×201, we could replicate similar effects with a mobility of 2.

The main effects of mobility appear at moderate starting trust degrees of 5 or 6. For other values, as well as for extremal values of the number of defecting agents the dice are too heavily loaded towards one of the sides to allow for a significant effect of mobility. That is, the influence of mobility is significantly weaker than the other parameters tested.

We conclude from our simulation that ceteris paribus, a comparatively lower level of trust in highly mobile societies cannot be explained through a direct effect of mobility. However, an indirect influence of mobility is still conceivable. Different levels of mobility can impact *other* relevant input parameters of our simulation. In particular the trustworthiness of agents might vary with their mobility due to, for instance, different probabilities of being held accountable for transgressions. In our basic simulation agents do not learn anything about the underlying society other than the expected value of trust. Arguably, this assumption is very unrealistic. Agents might learn about the behavior of individuals as well as about group features related to trustworthiness (see [Falcone and Miceli 2013]). The knowledge gained can be either explicit or merely implicit, only showing in the agents' behavior. In an extension of our framework we equip our agents with a limited learning algorithm. If,

Table 4: Effects of Mobility on Dissimilarity

mobility	average dissimilarity after 50 rounds	after 100 rounds
1	0.307	0.447
2	0.199	0.224
3	0.119	0.284
4	0.119	0.247
5	0.128	0.207
6	0.088	0.183
7	0.107	0.208
8	0.093	0.232
9	0.091	0.229
10	0.099	0.184

Table 5: Impact of memory

number of simulations	nr. of *distrusting* sim[3]	nr. of *trusting* sim.
simulations without memory	15272	15446
simulations with memory	15276	15442

to an agent, some partner proved trustworthy twice within a short time span (at most nine other encounters in between), this partner is added to his memory. Whenever agents are asked to pick a trustee, they preferably pick somebody from within their memory list, if available. Thus, memory can also be interpreted as the creation of a limited trust network. Since only trustworthy can enter the memory lists, we expected memory to have a positive impact on the general trust level. Surprisingly, this does not hold true. As can be seen from table 5, memory have no significant effect on the emergence of trust. Unless noted otherwise, all other results presented in this paper are based on simulations without memory.

We model trust as a purely subjective expectation about behavior in trust games. Such expectations can be affected by actual learning experiences, but also by shocks, public external signals produced for instance through rumors, news coverage, advertisement campaigns or therelike. While the effects of such shocks can be enhancing or diminishing trust, we are interested in negative shock only. Under which condition can a single strong public signal influence the long term behavior of the system? We are interested in whether the existence of memory reduces the vulnerability towards external shocks. After 200 rounds of simulation we introduce a shock that reduces each agents trust-memory by a random number between 0 and 5. While we saw in the last paragraph that memory alone does not have any significant influence on the limit behavior of the system, they impact the vulnerability to shocks drastically. The strongest impact occurs for a low mobility of 1. There, the introduction of memory increases the share of simulations converging towards a universal state of trust from .305 to .345. Surprisingly, for higher values of mobility the effect reverses. As can be seen from table 6, the introduction of memory increase the vulnerability from shocks from a certain level of mobility onwards. We leave it to future analysis to gain a more fine grained understanding of this phenomenon.

Our model incorporates two types of learning, direct learning through experiences as a trustor as well as social information gained as a trustee, both having an individual impact on the emergence of trust. The relative impact of these two process is guided by the parameter *social factor*. Indirectly, this parameter mirrors the interplay between various individual factors such as the agents taste for social uniformity or their trust that other agents are competent and sincere in their learning endeavors. In our initial simulation first and second order learning were treated on par, thus we worked with a second order factor of 1. In the extreme case of a social factor of 0, the agents do not take into account second order information at all. Thus agents can never regain trust after having lost it once. Most of our simulations start with an average trust-memory larger than 5, i.e. with more trusting than distrusting agents. In this situation, second order learning has a positive average effect on the trust memory, thus we expect a high second order factor to foster the creation of trust. This

---

[3] *Distrusting* is defined analogously to *trusting* as denoting simulations with *trust level* $\leq$ 0.2

Table 6: Memory and Shock Vulnerability

| | Share of *trusting* sim | |
mobility	without netw.	with memory
1	0.305	0.345
2	0.348	0.374
3	0.362	0.365
4	0.365	0.371
5	0.364	0.366
6	0.364	0.365
7	0.374	0.362
8	0.382	0.368
9	0.376	0.361
10	0.376	0.365

Table 7: Impact of second order learning

social factor	Share of *trusting* sim
0.0	0.0
0.5	0.286
1.0	0.696
1.5	0.726
2.0	0.762

expectation is satisfied, as can be seen from table 7.

Current conceptual work examines the relationship between trustworthiness and trust, see [Falcone and Miceli 2013, Castelfranchi and Falcone 2010]. In the current simulation we do not implement any relationship of this kind. We assume the agents' behavior as trustees to be constant throughout time. Furthermore, the processes determining the agents trust-memory, both the initial distribution of trust memory as well as the learning algorithm, are independent of the player's trustee type. Nevertheless, our simulation does display a correlation between trustee type and average trust-memory at the final stage of the simulation, see table 8 for details. Especially for a mobility of 1 up to 75% of the simulations display a higher average trust-memory of trustworthy agents relative to their untrustworthy peers. We do not have a convincing explanation for this phenomenon yet. One possible explanation is that the trust type feeds back in the trust expectation of the surroundings. By rewarding trust, trustworthy agents increase the trust expectation of agents around them and thereby increase their own chance of having a positive experience as trustees. However further simulations are necessary to thoroughly evaluate this explanation.

Table 8: Correlation between trust-memory and trust type

mobility	Share high	mobility	Share high
1	0.743	6	0.569
2	0.643	7	0.578
3	0.628	8	0.596
4	0.592	9	0.562
5	0.599	10	0.62

Share high = Share of simulations with avg trust of *trustee = 1* > avg trust of *trustee = 0*

6 Conclusion and Future Work

We were interested in understanding the dynamics and evolution of trust of human agents in situations characterized by high mobility and informational scarcity. We opted for an informational minimally setting with agents only learning through their own experience. To this end we set up an agent based NetLogo simulation based upon a bayesian learning paradigm. Our simulation incorporates various parameters identified in the current social capital literature. We hold that the factors used for simulation apply equally to real societies as to digital societies as created by for instance the user base of a digital market place. Our base line model behaves as predicted, displaying a positive correlation between both, initial trust of the agents and average trustworthiness on one hand and the creation of trust on the other hand. We take this as confirmation for our model. Our extended model, incorporating memory, mobility and external shocks displays some surprising effects showing that our theoretical knowledge on the determinants of trust in a dynamic setting is incomplete and misleading. Corresponding empirical results might hinge on some yet undiscovered variables or relationships that remain to be revealed. Lastly, we hold that the fine grained understanding of trust gained through our simulation is easily integrable with other theoretical models, for instance on changes in trustworthiness. We hope that this paper helps bridging a gap between a long tradition of theoretical and empirical research in the social sciences and formal models of trust in logic and computer science. We also take it to be a good showcase example for the use of computational models in the social sciences.

In future work we plan to expand this model in various directions. For once we aim at incorporating a dynamic model of trustworthiness, allowing the agents to alter their behavior as trustees in line with as they learn about the behavior of others.

A second factor we aim to understand better is the impact of the geometry of the underlying space. Mobility in our setting is construed as both, geographical and social mobility in the real world, but also the velocity in which the interactive surroundings change in digital settings. We are both interested in the influence of geographical inhomogeneities within a two-dimensional state space as well as influences of the size and dimensionality of the space itself.

Furthermore, we are interested in a more fine grained understanding of the impact of parameter combinations rather than studying individual parameters in isolation. In a first study, we have shown that the the effects of the basic parameters %-*defectors* and *starting trust memory* are not independent, but their interplay for the creation of trust is more complex.

References

[Bereby-Meyer, Erev 1998] Y. Bereby-Meyer, I. Erev. On Learning to Become a successful Loser: A Comparison of Alternative Abstractions in the Loss Domain. *Journal of Mathematical Psychology*, 42(2-3):266–286, 1998.

[Bicchieri 2006] C. Bicchieri. *The Grammar of Society*. CUP 2006.

[BME 2011] C. Bicchieri, X. Erte and R. Muldoon. Trustworthiness is a Social Norm, but Trusting is Not. *Politics, Philosophy and Economics*, 10(2):170-187, 2011.

[Birk 2001] A. Birk Learning to trust *Lecture Notes in AI: Trust in Cyber-societies*, 2246:133–144, 2001.

[Burt 2000] R. Burt. The Network Structure of Social Capital. *Research in Organisational behavior*, 22:345–423, 2000.

[Castelfranchi and Falcone 2000] C. Castelfranchi, R. Falcone Trust is much more than subjective probability: Mental components and sources of trust 2nd Hawaii International Conference on System Sciences - Mini-Track on Software Agents, Maui, 2000

[Castelfranchi and Falcone 2001] C. Castelfranchi, R. Falcone Social Trust: A Cognitive Approach *Journal of Mathematical Psychology*, 42(2-3):266–286, 2001.

[Castelfranchi and Falcone 2010] C. Castelfranchi and R. Falcone *Trust theory: A socio-cognitive and computational model* Chichester, UK: Wiley

[Falcone and Castelfranchi 2001] R. Falcone, C. Castelfranchi The socio-cognitive dynamics of trust: Does trust create trust? *Lecture Notes in AI: Trust in Cyber-societies*, 2246:55-72, 2001.

[Falcone and Castelfranchi 2004] R. Falcone, C. Castelfranchi Trust dynamics: How trust is influenced by direct experiences and by trust itself *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent*(2), 740-747, 2004.

[Falcone and Miceli 2013] R. Falcone and M. Miceli Relationships between Trusting and Being Trustworthy Trust 2013

[Fukuyama 2006] F. Fukuyama. *The End of HIstory and the Last Man.* Free Press 2006.

[Guiso, Sapienza and Zingales 2008] L. Comer, P. Sapienza and L. Zingales. Alfred Marshall Lecture Social Capital as Good Culture. *Journal of the European Economic Association*, 6(2,3):295–320, 2008.

[Jonker and Treur 1999] C. Jonker, J. Treur. Formal analysis of models for the dynamics of trust based on experiences *Autonomous Agents '99 Workshop on "Deception, Fraud and Trust in Agent Societies"', Seattle, USA, May 1*, pp. 81-94.

[McKenzie 2007] A. McKenzie. *The Structural Evolution of Morality.* CUP 2007

[Nooteboom, Klos and Jorna 2001] B. Nooteboom, T. Klos, R. Jorna Adaptive Trust and Co-operation: An Agent-Based Simulation Approach *Lecture Notes in AI: Trust in Cyber-societies*, 2246:83-110, 2001.

[Portes 1998] A. Portes. Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology*, 24(1):1–24, 1998.

[Putnam 95] R. Putnam. Tuning in, Tuning out - the Strange Disappearance of Social Capital in America. *Political Science and Politics*, 28(4):664-683, 1995.