CONVERGENCE ANALYSIS OF A MULTIAGENT COOPERATION MODEL

Markus Eberling and Hans Kleine Büning
Department of Computer Science, University of Paderborn, 33095 Paderborn, Germany

Keywords: Cooperation, Multiagent systems, Value propagation, Imitation-based learning.

Abstract: Cooperation between autonomous and rational agents is still a challenge. The problem even gets harder if the agents follow different policies or if they are designed by different companies that have contradicting goals. In such systems agents cannot rely on the cooperation willingness of the other agents. Mostly, the reason for receiving cooperation is not observable as it is a result of the private decision process of the other agent. We deal with a multiagent system where the agents decide with whom to cooperate on the basis of multiple criteria. The system models these criteria with the help of rated propositions. Interaction in our system can only occur between agents that are linked together in a network structure. The agents adapt their values to the best performing neighbor and rewire their connections if they have uncooperative neighbors. We will present an imitation-based learning mechanism and we will theoretically analyze the mechanism. This paper also presents a worst case scenario in which the mechanism will fail.

1 INTRODUCTION

Agents in multiagent systems are designed to behave rational and autonomous (Ferber, 1999; Wooldridge, 2009). Therefore, they have to decide with whom to cooperate on their own. This process can be influenced by different factors which may not be observable for other agents. The problem is that the individual rational choice may be different from the social rational choice. In this context an agent cannot rely on receiving cooperation whenever it is needed. However, cooperation is essential in many multiagent systems if the agents should achieve a global goal.

Cooperation in everyday life can be found in groups of humans or in companies that are organized in a network structure besides other examples. In most scenarios cooperation leads to higher benefit for the whole group and to higher benefit for the individuals. Mostly, the group members have a common goal but different motivations to join the group (Pennington, 2002) or to stay in it (Buchanan and Huczynski, 1997). Companies build networks to achieve their goals (Peitz, 2002) and moreover good supply chains are helpful to produce qualitative products. Reciprocal behavior is one of the characteristics of such networks (Sydow, 1992). Another aspect is altruism which is on the one hand helping others without being payed for (Berkowitz and Macaulay, 1970) and which can produce costs on the other hand (Krebs, 1982; Wispé, 1978). The decision to cooperate is often based on different criteria like kin selection or social cues.

We model the process of determining cooperation partners with the help of propositions which are rated by the agents. Based on the distances of these ratings they determine the agents they are willing to cooperate with. Each proposition leads to a criterium that has to be fulfilled. If all criteria are fulfilled, the agent will cooperate with an agent asking for help.

In (Eberling, 2009) and (Eberling and Kleine Büning, 2010b) a local learning algorithm was proposed that favors the determination of cooperation partners. This paper will give a convergence analysis of this approach. The agents adapt to the best neighbor by imitating its proposition ratings and reach high levels of cooperation. The agents in the system only have local knowledge as they are only aware of those agents they are linked to. There exist similar models but most of them lack theoretical analysis under which preconditions convergence to cooperative behavior will emerge. This paper will give a theoretical analysis of the adaptation mechanism and will show that there are examples where the system does not converge to cooperation. However, it is claim that these examples are very rare and that the assumptions that have to be

Eberling M. and Kleine Büning H..
CONVERGENCE ANALYSIS OF A MULTIAGENT COOPERATION MODEL.
DOI: 10.5220/0003139901670172
In *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence* (ICAART-2011), pages 167-172
ISBN: 978-989-8425-41-6

made for them are very specific as good results have been observed in previous work.

In literature one can find similar models based on observable markers such as tags which are evolved over time. In the work of Riolo et al. cooperation can only occur between two agents a and b if $|\tau_a - \tau_b| \leq T_a$ holds, where τ_a is the tag value and T_a is a similarity threshold (Riolo et al., 2001). Hales also made experiments based on this mechanism to determine cooperation (Hales, 2002; Hales, 2004). The difference to our work is that on the one hand adaptation means copying the value and threshold as well as the strategy of another agent. In our scenario the agents are only allowed to imitate the values without purely copying. Another aspect is that we will deal with a set of such inequations that all have to be fulfilled. Hales and Riolo et al. lack formal analysis why the cooperation emerges but only give experimental results. We will formally show different cases where cooperation may and may not emerge.

De Weerdt et al. (de Weerdt et al., 2007) calculate task allocations using a distributed algorithm in a social network. A social network is a graph where the nodes represent the agents and the edges model possible interaction links. The tasks are assigned to agents which have limited resources. They show that the problem of finding an optimal task allocation, which maximizes the social welfare, is NP-hard. In contrast to the work presented here, their model does not consider cooperation costs and the agents also know about all tasks before the decision process is started which is also different to the work presented here. Another difference is the static social network structure. In contrast, we analyze dynamic networks and show that the challenges of those networks favors the cooperation between the agents.

2 SCENARIO DESCRIPTION

In this section we describe the formal model used in this paper. Due to page limitations we will only describe the features of the model and omit the formal definitions. They can be found in (Eberling and Kleine Büning, 2010a). We will first define the basic model and then describe the considered scenario.

The agents in our model are linked together and form a so called interaction network IN. Basically, the interaction network $\mathrm{IN} = (\mathcal{A}, \mathcal{N})$ is an undirected graph with a finite set of agents $\mathcal{A}$ as the nodes and a set of links $\mathcal{N}$. The links between the agents represents the neighborhood relationship. Therefore, agents a and b are able to interact iff there exists an edge between them in the interaction network, i.e. $\{a, b\} \in \mathcal{N}$. An interaction network is called *dynamic* if the graph can change between successive simulation steps. Note that due to the interaction network the agents' view of the system is local only.

In our system the agents have to fulfill different jobs consisting of smaller tasks. Each task requires a specific skill out of a skill set $s_t \in \mathcal{S}$ and leads to a non-negative payoff $q_t \in \mathbb{R}_0^+$ if the task is fulfilled. Therefore, a task t can be modeled as a pair $t = (s_t, q_t)$. Let $\mathcal{T}$ be the finite set of all possible tasks. Then $\mathcal{J} \subseteq Pow(\mathcal{T})$ is the set of all jobs. Hence, a job $j \in \mathcal{J}$ is a set of tasks $j = \{t_1, \ldots, t_n\}$ where $t_{\min} \leq n \leq t_{\max}$ with $t_{\min}, t_{\max} \in \mathbb{N}$ denote the minimum and maximum number of tasks a job consists of and n the number of tasks. The payoff for a job is the sum of the tasks' payoffs if it is fulfilled, i.e. if all tasks are fulfilled, and zero otherwise.

The environment env the agents are situated in is a tuple $\mathrm{env} = (\mathcal{S}, \mathcal{P}, \mathrm{IN}, \mathcal{J})$ where $\mathcal{S}$ is a finite, non-empty set of skills, $\mathcal{P} = \{p_1, \ldots, p_m\}$ is a set of propositions, $\mathrm{IN} = (\mathcal{A}, \mathcal{N})$ is an interaction network and $\mathcal{J}$ is a finite set of jobs. The set of propositions are a mean to model the decision process to determine cooperation partners based on many criteria. The agents share the set of propositions that are part of the environment. These propositions can be opinions about the overall world state or the evolution of the environment. As we do not concentrate on the modeling of such propositions we do not provide a formal definition. A proposition p can represent anything like "The road is clear" in the context of a taxi-driving agent or "The color blue is prettier than black". For our purposes it is enough to know that there are propositions that may influence the behavior of the agents. More details can be found in (Eberling and Kleine Büning, 2010a).

An agent $a \in \mathcal{A}$ is a tuple $a = (\mathcal{S}_a, \mathcal{N}_a, \mathcal{C}_a, \mathcal{V}_a, \Theta_a)$ where $\mathcal{S}_a \subseteq \mathcal{S}$ is the set of skills agent a is equipped with, $\mathcal{N}_a \subseteq \mathcal{A}$ is the agent's neighborhood defined by the interaction network, $\mathcal{C}_a \subseteq \mathcal{N}_a$ is the set of neighbors, agent a is willing to cooperate with, $\mathcal{V}_a \in [0, v_{\max}]^m \subset \mathbb{Q}^m$ is a vector giving values to the propositions and finally $\Theta_a \in (0, \Theta_{\max}]^m \subset \mathbb{Q}^m$ is a threshold vector. To keep the agents as simple as possible, only the proposition-values are modeled as observable properties. All other parts of the agents (i.e. skills, thresholds and neighbors) are not visible to other agents and constitute private knowledge. Based on the values the agents give to the propositions their cooperation partners are determined. The set of cooperation partners $\mathcal{C}_a$ of agent a are all neighbors $b \in \mathcal{N}_a$ for which the following holds:

$$\forall p \in \mathcal{P} : |\mathcal{V}_a(p) - \mathcal{V}_b(p)| \leq \Theta_a(p) \tag{1}$$

This means that for the cooperation partners the distances between the ratings for the propositions have to be less or equal to the thresholds. We also define a cooperation relation $\mathcal{C} \subseteq \mathcal{A} \times \mathcal{A}$ based on the sets of cooperation partners:

$$b \in \mathcal{C}_a \Leftrightarrow (a,b) \in \mathcal{C} \tag{2}$$

According to this definition, it is easy to see that the relation $\mathcal{C}$ is not symmetric in general.

As the agents should learn to select their cooperation partners we endowed them with the possibility of adaptation to other agents. This adaptation affects the proposition values of the agents. They change their values to imitate a better performing neighbor. The intention for this is, that better performing agents are believed to gain high performance based on better values for the propositions. The adaptation step is defined through Equation 3:

$$\mathcal{V}_a^{t+1} = \mathcal{V}_a^t + \eta \cdot (\mathcal{V}_{a^*}^t - \mathcal{V}_a^t) \tag{3}$$

where a^* is the best performing agent out of agent a's neighborhood and $\eta \in [0,1] \subset \mathbb{Q}$ is an exogenous adaptation strength.

Algorithm 1: Simulation.

1: Initialize $|\mathcal{A}|$ agents and their neighborhoods randomly
2: **loop**
3: Generate $10 \cdot |\mathcal{A}|$ jobs and allocate them to randomly chosen agents
4: **for all** agents $a \in \mathcal{A}$ **do**
5: $\mathcal{E}_a \leftarrow$ set of best agents of $\mathcal{N}_a \cup \{a\}$
6: **if** $a \notin \mathcal{E}_a$ **then**
7: $a^* \leftarrow$ best agent of $\mathcal{N}_a$
8: $\mathcal{V}_a \leftarrow \mathcal{V}_a + \eta \cdot (\mathcal{V}_{a^*} - \mathcal{V}_a)$
9: with probability $P_{\mathcal{N}}$: replace r uncooperative neighbors by r randomly chosen agents
10: **end if**
11: **end for**
12: **end loop**

Algorithm 1 describes our simulation which incorporates the adaptation mechanism. In each step $10 \cdot |\mathcal{A}|$ jobs are generated and assigned to randomly chosen agents with uniform distribution (line 3). This leads to an assignment of 10 jobs on average per agent. The jobs are dynamically generated and separately assigned to the agents and processed sequentially by the agents. This leads to a fundamental property of our system: the agents are not able to reason about the whole set of jobs and to select the most beneficial ones. We decided to do this because we concentrate on the cooperation aspect and not on the aspect of most efficient task allocations as it is done in

Figure 1: Simple MAS composed of two agents.

similar models (de Weerdt et al., 2007). Another reason for this is that we want to concentrate on agents that are as simple as possible.

The agents work on the jobs and every fulfilled job is rewarded with the job's payoff for the allocated agent, only. Cooperative agents that helped others to fulfill their jobs are punished with a payoff of $-0.25 \cdot q_t$ for every task they processed. Both, the uncertainty about the next jobs as well as the cooperation costs make it impossible to apply common coalition formation techniques (Branzei et al., 2008) to our considered system.

The second phase of the algorithm is the considered approach for imitation-based learning (lines 4-11) and consists of two sub-phases. The first subphase is the *adaptation part*. First, the best performing agents are determined locally (line 5) and these agents build the elite set $\mathcal{E}_a$. If the agent is not in this set, then it is said to be unsatisfied and adapts its value-vector to the vector of the locally best performing agent (lines 7-8). The second sub-phase is called *social networking* and effects the interaction network. With some probability $P_{\mathcal{N}}$ the agent replaces r uncooperative agents with randomly chosen agents out of $\mathcal{A}$ (line 9). For all settings with $P_{\mathcal{N}} > 0$ we have a dynamic interaction network.

Note, that the agents are not able to sense the threshold values of their neighbors. Therefore, they are not able to compute which neighbor is not willing to cooperate with them. But it is possible for the agents to keep a history of previous behavior of their neighbors. This can be used as an approximation of the set of uncooperative neighbors which is not considered here.

3 CONVERGENCE ANALYSIS

In this section we analyze the convergence behavior of our adaptation mechanism. To ease the analysis we will only concentrate on static interaction networks with very small agent sets.

3.1 The Simplest Scenario

Let us consider a very simple system composed of two agents and a single proposition. For better readability $\mathcal{V}$ denotes the rating of this single proposition as a rational number instead of an one-dimensional vector. The system is illustrated in Figure 1. We denote with $\text{profit}(a)$ the profit that an agent a earned in one simulation step. In the scenario with two agents the job phase can produce the following three different profit distributions:

1. $\text{profit}(a) = \text{profit}(b)$
2. $\text{profit}(a) > \text{profit}(b)$
3. $\text{profit}(a) < \text{profit}(b)$

Case 1 is very simple since no adaptation takes place. As case 2 and 3 are symmetric we will concentrate on case 2 in the remainder of this section. Then agent b will always adapt to a by the adaptation rule provided in Equation 3 which can be transformed to:

$$\mathcal{V}_b^{t+1} = \mathcal{V}_b^0 \cdot (1-\eta)^{t+1} + \eta \cdot \mathcal{V}_a^0 \cdot \sum_{i=0}^{t} (1-\eta)^i \quad (4)$$

Lemma 1. *Let* $\text{dist}(a,b,t)$ *be the distance of the proposition ratings of two agents a and b in step t with* $\text{dist}(a,b,t) = \left|\mathcal{V}_a^t - \mathcal{V}_b^t\right|$. *In a scenario with just two agents, the distance never increases, i.e.* $\forall t$: $\text{dist}(a,b,t+1) \leq \text{dist}(a,b,t)$.

The proof can be found in (Eberling and Kleine Büning, 2010a).

We now want to know how many steps are needed until both agents are willing to cooperate with each other, i.e. after how many steps $a \in \mathcal{C}_b$ and $b \in \mathcal{C}_a$ holds. As agent b adapts to agent a, only, we suppose that agent a is less tolerant, i.e. $\Theta_a < \Theta_b$. This means that agent a determines the number of steps needed, since $a \in \mathcal{C}_b$ will follow first. From the proof of Lemma 1 (see (Eberling and Kleine Büning, 2010a)) we have:

$$\text{dist}(a,b,t) = (1-\eta) \cdot \text{dist}(a,b,t-1) \quad (5)$$

Through simple transformations we get:

$$\text{dist}(a,b,t) = (1-\eta)^t \cdot \text{dist}(a,b,0) \quad (6)$$

Thus, we are searching for the smallest t that satisfies $(1-\eta)^t \cdot \text{dist}(a,b,0) \leq \Theta_a$. This is true for

$$t' = \left\lceil \frac{\ln(\Theta_a) - \ln(\text{dist}(a,b,0))}{\ln(1-\eta)} \right\rceil . \quad (7)$$

Therefore after step t' $a \in \mathcal{C}_b$ and $b \in \mathcal{C}_a$ will hold.

We only considered case 2, where agent b adapts to agent a in every step. If case 2 does not hold in every simulation step, we will have to deal with interleaved cases. If case 3 holds we have the symmetric situation that will lead to the same result in the end. However, case 1 can slow down the development as no adaptation takes place, if both agents reach the same profit. But eventually it will hold that case 2 or 3 will again occur and the process is continued. Although we have seen good results in previous work, we cannot ensure convergence in every setting.

Figure 2: Simple MAS composed of three agents.

Lemma 2. *The adaptation cannot ensure convergence. There are settings in which the system will fail.*

Proof. Subsection 3.2 gives an example where the adaptation does not converge. ■

3.2 A Simple Scenario without Convergence

Consider the following very simple multiagent system in Figure 2. We have three agents. For the agents and the interaction network we consider the following formal specification:

- $\text{IN} = (\{a,b,c\}, \{\{a,b\},\{b,c\}\})$
- $\mathcal{S} = \{1,2,3,4,5\}$
- $t_{\min} = t_{\max} = 3$ and $q_t = 1$ for all tasks t
- $\mathcal{S}_a = \{1\}$, $\mathcal{S}_b = \{3\}$, $\mathcal{S}_c = \{5\}$
- $\mathcal{P} = \{p_1\}$
- $\mathcal{V}_a = 0$, $\mathcal{V}_b = 50$, $\mathcal{V}_c = 100$
- $\Theta_a = 2$, $\Theta_b = 100$, $\Theta_c = 2$

As we have only one proposition we will use the simplified notation from the previous subsection. Now, consider the following profit distribution:

$$\text{profit}(a) > \text{profit}(c) > \text{profit}(b), \text{ for odd } t \quad (8)$$

$$\text{profit}(c) > \text{profit}(a) > \text{profit}(b), \text{ for even } t \quad (9)$$

This profit distribution can be the result of the relative intolerant agents a and c and the very tolerant agent b. This can lead to an alternating adaptation of agent b to agent a in odd simulation steps and to agent c in even simulation steps. As the agents a and c only have a single neighbor, agent b, and this agent is always the worst performing one, they never adapt. Therefore, the length of the value-interval remains constant.

If we set the adaptation strength $\eta = 0.5$ and let agent b adapt in the alternating way as described above, we can calculate the proposition value $\mathcal{V}_b^t$ for every time step t. The value of agent b changes in every step and we observed in previous work that it does not converge to a single value but it oscillates between $33\frac{1}{3}$ and $66\frac{2}{3}$ (Eberling and Kleine Büning, 2010a). For both directions it holds that in every simulation step the minimal distance is $33\frac{1}{3}$. Therefore, agent b never receives help from the other two agents. The only possibility for agent b to gain profit is the fulfillment of a job containing three times skill 3. But this situation is very rare because the probability of getting such a job is 0.8% with the given parameters. However, as agent b is very tolerant it always helps the other two agents if they ask for help. That's why agent b is punished very often in contrast to the other two agents.

However, this construction is very artificial. In scenarios that have been considered in previous work (Eberling, 2009; Eberling and Kleine Büning, 2010b) this problem does not occur or at least it does not lead to significant performance losses. We dealt with 1000 agents and neighborhood sizes of 15 to 20 agents in a random network. Because of the results in (Eberling, 2009; Eberling and Kleine Büning, 2010b), we assume that in random networks the probability of having situations without convergence is very low and might be close to zero. One very strong assumption we made in this subsection is, that agent b adapts to its neighbors in an alternating way. If the agent adapts to one neighbor only, we will get a similar convergence behavior as in the scenario considered in subsection 3.1. Assume, that only the case occurs in which agent a is the overall best agent. Then we can apply Equation 7 to calculate the time steps t' needed until agent a and b will mutually cooperate, if the adaptation strength is set to $\eta = 0.5$:

$$t' = \left\lceil \frac{\ln(2) - \ln(50)}{\ln(0.5)} \right\rceil = 5 \tag{10}$$

This means that after five adaptation steps it holds that $(a,b) \in \mathcal{C}$ and $(b,a) \in \mathcal{C}$. Especially it holds that $\text{dist}(a,b,t) \leq \Theta_a$ for all $t \geq 5$ which means that b will receive cooperation from agent a and this lets agent b perform better than agent c, after time step 5. As a consequence c will adapt to b and the interval between the values of agent a and agent c will diminish.

3.3 Empirical Analysis

Figure 3 shows some simulation results with 1000 agents over 200 simulation steps. The results are means over 30 independent simulations. The jobs are randomly generated and allocated with a uniform distribution. As can be seen the simulations reach high cooperation rates for static scenarios ($P_{\mathcal{N}} = 0.0$) and even better results for dynamic scenarios ($P_{\mathcal{N}} > 0.0$). This shows that the non-converging behavior does not take place in these scenarios or at least has no great influence on the performance. For detailed experimental analysis see (Eberling and Kleine Büning, 2010a).

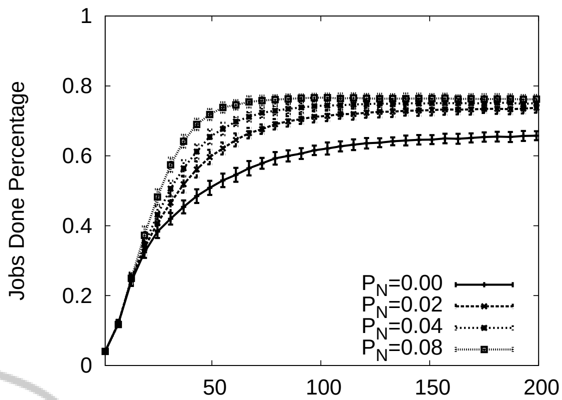

Figure 3: Percentage of completed jobs for 200 simulation steps.

4 CONCLUSIONS AND FUTURE WORK

Cooperation decisions in everyday life between humans are based on many criteria which may not be observable by others. With the help of ratings for propositions this process has been modeled in a multiagent system as a multidimensional decision process. This paper analyzed the decision process and its ability to converge to cooperative behavior. The model fits well to systems where the cooperation willingness is not necessarily part of the designed agents. Compared to other approaches towards self-organization one can see that the agents in the presented model do not need to be very complex and they do not require much knowledge as only the values for the propositions have to be observable for the agents.

The local learning algorithm is able to produce high rates of cooperation in the considered multiagent systems. This paper has provided a theoretical analysis and has shown that the considered approach does not always lead to convergence to the intended behavior. However, using randomly generated systems with uniform distribution have not shown non-convergence in previous experiments. The assumptions that had to be made in the presented analysis are very specific and do not seem to hold in uniformly at random generated scenarios. The question arises how this behavior can be detected and avoided in a local way without too much computational effort. This is left for future work.

Additionally, the influence of other network structures and the convergence behavior in such systems should be examined. Also the influence of trust mechanisms should be analyzed. In such settings different trust mechanisms that favor the process of creating cooperative structures should be analyzed.

REFERENCES

Berkowitz, L. and Macaulay, J. (1970). *Altruism and Helping Behavior: Social Psychological Studies of Some Antecedents and Consequences*. Academic Press, New York.

Branzei, R., Dimitrov, D., and Tijs, S. (2008). *Models in cooperative game theory*. Springer Verlag, Heidelberg.

Buchanan, D. and Huczynski, A. (1997). *Organisational Behaviour: An Introductory Text*. Prentice Hall, London.

de Weerdt, M. M., Zhang, Y., and Klos, T. B. (2007). Distributed task allocation in social networks. In Huhns, M. and Shehory, O., editors, *Proceedings of the 6th International Conference on Autonomous Agents and Multiagent Systems*, pages 488–495, Bradford, UK. IFAAMAS, Research Publishing Services.

Eberling, M. (2009). Towards determining cooperation based on multiple criteria. In Mertsching, B., Hund, M., and Aziz, M. Z., editors, *KI 2009*, volume 5803 of *LNCS*, pages 548–555, Heidelberg. Springer.

Eberling, M. and Kleine Büning, H. (2010a). Convergence analysis of a multiagent cooperation model (extended version). Technical Report TR–RI–10–321, University of Paderborn.

Eberling, M. and Kleine Büning, H. (2010b). Self-adaptation strategies to favor cooperation. In Jedrzejowicz, P., Nguyen, N. T., Howlett, R. J., and Jain, L. C., editors, *KES-AMSTA (1)*, volume 6070 of *LNCS*, pages 223–232, Heidelberg. Springer.

Ferber, J. (1999). *Multi-Agent Systems: An Introduction to distributed artificial intelligence*. Addison Wesley Longman Inc., New York, USA.

Hales, D. (2002). Evolving specialisation, altruism, and group-level optimisation using tags. In Sichman, J., Bousquet, F., and Davidsson, O., editors, *MABS 2002*, volume 2581 of *LNCS (LNAI)*, pages 26–35, Heidelberg. Springer.

Hales, D. (2004). The evolution of specialization in groups. In Lindemann, G., Moldt, D., and Paolucci, M., editors, *RASTA 2002*, volume 2934 of *LNCS (LNAI)*, pages 228–239, Heidelberg. Springer.

Krebs, D. (1982). Psychological approaches to altruism: An evaluation. *Ethics*, 92(3):447–458.

Peitz, U. (2002). *Struktur und Entwicklung von Beziehungen in Unternehmensnetzwerken*. Deutscher Universitätsverlag, Wiesbaden.

Pennington, D. C. (2002). *The Social Psychology of Behavior in Small Groups*. Psychology Press, New York.

Riolo, R. L., Cohan, M. D., and Axelrod, R. (2001). Evolution of cooperation without reciprocity. *Nature*, 414:441–443.

Sydow, J. (1992). *Strategische Netzwerke: Evolution und Organisation*. Gabler, Wiesbaden.

Wispé, L. (1978). *Altruism, sympathy, and helping: Psychological and sociological principles*. Academic Press, New York, New York.

Wooldridge, M. (2009). *An Introduction to MultiAgent Systems - Second Edition*. John Wiley & Sons, West Sussex, UK.