Scaling up by symbolically generalized media*

Jörg Wellner
Chemnitz University of Technology, Computer Science
D-09107 Chemnitz, Germany
jwe@informatik.tu-chemnitz.de

Sigmar Papendick
University of Konstanz, Department of Sociology
Universitätsstr. 10, Postfach 5560
D-78464 Konstanz, Germany
Sigmar.Papendick@uni-konstanz.de

Werner Dilger
Chemnitz University of Technology, Computer Science
D-09107 Chemnitz, Germany
wdi@informatik.tu-chemnitz.de

Abstract

In this paper we present an evolutionary approach for developing an agent system, consisting of a large and varying number of agents. We start off by describing in short a sociological approach to problems of coordination in a society of a huge number of members. Elaborated cognitive explanations of handling individual information are rejected and the concept of symbolically generalized communication media is suggested instead. In a first attempt we have modeled an agent system based on this concept. Simulation results show that agents may coordinate their actions even though they have no individual representations of each other. Simulation starts with a small group of agents and evolves a system of several hundred agents which base their actions mainly on exchanged messages.

1 Introduction

This paper addresses an important aspect of scaling up multi-agent systems having a huge and varying number of agents. Many coordination approaches for agent systems rely on mechanisms which include detailed knowledge of an agent architecture. It is essentially knowledge of another agent in order to cooperate with it. It is personalized in the way that an agent knows goals, skills, or beliefs ([10]) for different opponents. To simplify the usual situation one can state that the more potential partners for an interaction an agent has the more agent specific knowledge it has to cope with. This is one reason why current logic based agent approaches scale so badly, because keeping track of information about other agents is an expensive matter. In the next section we consider in more detail a concept, developed by sociologists, to answer questions concerning the coordination of individuals in a society of a huge number of members, namely symbolically generalized communication media. In Section 3, we discuss the relevance of this concept for multi-agent systems. A first approach of

*This work is supported by the *Deutsche Forschungsgemeinschaft* under grant number DI 452/10-1.

modeling the proposed mechanisms is presented in Section 4, together with a detailed analysis. Section 5 concludes the paper, indicating that one can reasonably base multi-agent systems on the proposed sociological concepts in order to achieve a good scaling.

2 The Concept of Symbolically Generalized Media

Humans faced a problem similar to the one above during the development from small groups to modern societies. In small groups it is possible for each individual to keep in mind relevant facts about other members of the group. Different strategies were developed to keep one's knowledge about each other up-to-date, e. g. gossip [3]. As groups became larger, personalized coordination mechanisms became less efficient, due to the necessary increase of cognitive capabilities which are - however - constrained, and due to the fact that the number of possibilities for an interaction grew exponentially, and likewise the chance to deny an interaction. In different stages of societal development different coordination mechanisms predominated and ruled the society. Different authors emphasized different mechanisms. Among them is Talcott Parsons ([9]) who claimed generalized media of interchange, namely money, political power, influence, and value commitments. Another author (influenced by Parsons), Niklas Luhmann ([6], [5]), claimed symbolically generalized media of communication. Among them are power, money, love, art, and law. These mechanisms are characterized first of all by generalization and symbolization. Various social subsystems emerged based on different media, like the economic system or the political system. Because the proposed media are incorporated in a different manner for different reasons, in the subsequent we concentrate only on some media proposed by Luhmann.

A medium like money or power is coded in a binary yes-or-no manner in the way that one either has money or power or not. The generalization is realized in many different respects. Money can be used more than once, it is reproducible. In the process of accepting money, for example for a product or for performed labor, it does not matter what it is later used for. Furthermore, it does not matter from whom someone gets money or who offers something to buy. The symbolic character of these media is also prevalent, especially for money. It counts for a value but money itself has no value. Only because everybody believes, that he will get what he wants that it can be bought, everybody accepts money as a transmitter medium. But apart from these functions on the level of the social system, what is the function of such a medium on the interactional level? As Luhmann states, symbolically generalized media of communication reduce complexity, especially complexity outside an individual. A very complex situation may collapse into a simple yes-or-no question. By means of these media, a lot of information and therefore complexity is excluded from interest, and an interaction just focuses on a yes or on a no, or in other words, accepting something or accepting something not. For example, if one individual offers something to buy, another individual just thinks about buying it or not, but it does not think about the reason why the other individual may sell it, or what product it wants to get for it later. A buyer has no reason to care about the money when it is once spent, and a seller does not care about where the money comes from. Luhmann points out that symbolic generalized media are based on their embodyness. They are related to physical entities, e. g. money is related to needs, or power is related to physical force and pain

Why do media reduce complexity during an interaction? They serve as structures of expectation, they order and structure with respect to what should be expected in which situation. On the one hand, media reduce possibilities, and, on the other hand, they open possibilities for further communication or the prospect of further communication. Thus, media reduce external complexity to expectable internal complexity.

Some concepts related to symbolic media are already dealt with on a large scale in multi-agent research, especially norms (see e. g. [7] or [2]) and market based coordination mechanisms (see e. g. [8] or [1]). All the mechanisms based on these concepts, first of all norms, have the goal of enabling interactions between agents which do not know much about each other, but do know something in general. Norms are condensed expectation structures. A population wide norm makes actions of agents expectable.

3 Benefits of Media in Agent Systems

It is obvious that mechanisms like symbolically generalized media of communication may play an important role in scaling huge agent systems. The largest utility of media is concerned with their knowledge reducing effect for agent interactions. Interactions, controlled by a medium, are structured in a straightforward way. They do not ensure that an interaction always succeeds, but they ensure that

- agents know in advance on what aspect(s) negotiation should be limited,
- agents need not to know each other,
- agents can be black boxes to each other (they can't look inside their head), but coordination may still succeed, and
- agents know in what stage an interaction currently is, and when it should be stopped.

Every agent is only concerned with its own beliefs or goals, there is no need to take into account elaborated reasoning mechanisms about beliefs or goals of other agents, since they become immediately apparent to each other during an interaction to some extent. Whatever an agent wants or believes it will be canalized by a medium to another agent. A medium does not reveal an agent's goal or its beliefs, but it offers a way to achieve a goal or to verify or to strengthen its own beliefs.

Different media are based on different techniques to support inter-agent coordination. Power may an agent force to do something which is not in its own best interest, but may serve someone else. Especially the symbolized power by many agents may another agent force to confirm and to comply population wide norms. Symbolizing power does not necessarily mean to apply power, it just symbolizes the potential application of power. However, in the case power is applied the receiver suffers in some way.

The medium money is related to resources, especially to the shortage of one or more certain resource. Because money is the second coding of assets (of resources) (see [6]) it serves as a exchange medium for all kinds of resources. Therefore, an agent without a direct access to a certain resource but with a need for it may get it by spending money. That is, if two agents negotiate about the access to a resource they do it by negotiating the value for it. As a result they do not negotiate on the basis of who has a more urgent need for the resource, or which agent has a greater impact on the hole population.

4 A first Approach to the Evolution of Symbolized Media

4.1 The Simulation Scenario

In this paper we will not answer questions about how to incorporate all aspects of different media inside an agent architecture. As Luhmann points out, media are evolved, during societal development, together with a functional differentiation of the society. In the rest of this paper we present results of an approach to evolve a large population of agents which coordinate their actions in a media-like way. The medium we have in mind is power.

A simulation consists of a large number of trials of a cooperation game which we called the "Planter-and-Harvester-Game" for simplicity. We introduce two different types of agents, with respect to their ability to change the environment. There are also two types of actions that change the state of the environment in an effective way, namely "planting" and "harvesting" complementing each other. Plant agents, called *Planter* (or agent type 0) can perform only plant actions effectively, harvest agents, called *Harvester* (or agent type 1) can perform only harvest actions effectively. At the beginning of a game the environment U is always in state $U_s = 0$. A plant action $Plant_I$ – performed by a *Planter* – transforms the environment into state $U_t = 1$, a harvest action $Harvest_I$ – performed by a *Harvester* – transforms it into the final state $U_e = 2$. In more complicated games the final state may be $U_e > 2$ assuming action sequences $Plant_I$, $Harvest_I$, $Plant_{II}$, and so on. Action $Plant_I$ in state $U = 1$ has no effect with regard to the state of the environment, similarly action $Harvest_I$ in state $U = 0$. In general, the transformation of the environment state U_t at time t that a plant agent may perform by action a_t is defined by

$$[(U_t = i) \wedge ((i \bmod 2 = 0) \wedge (type = 0)) \wedge (a_t = i)] \Longrightarrow U_{t+1} := i + 1$$

and for harvest agents by

$$[(U_t = i) \wedge ((i \bmod 2 = 1) \wedge (type = 1)) \wedge (a_t = i)] \Longrightarrow U_{t+1} := i + 1\ .$$

In any other situation the environment remains in state i ($U_{t+1} = i$).

At the beginning of a game two agents are randomly selected from the population, one of them is the start agent. This agent begins by sending a message M_0. The other agent receives this message and does both, it – the first time – performs an action a_1 and sends another message M_1 to the start agent. Then, the first agent performs an action a_2 and sends a message M_2 to the second agent, and so on. A round is defined by a successive sequence of performing one action and generating a message for each of the two agents. Both types of agents have the same repertoire of actions regardless of the efficiency: apart from plant and harvest actions they have a *Null*-action without any effect, a sanctioning action *Bite*, an action *Exit*, and an action *Replace*. The later action affects the opponent agent in the way, that it gets replaced by a other agent, randomly selected from the population. This may increase the general possibility for a successful coordination. A game may end by three different outcomes: an agent performed the *Exit* action, the environment reached the final state U_e, or the number of rounds in the games exceeded the defined threshold *rounds*.

There is a predefined set of symbols $\mathcal{S} = \{0, 1, 2, ..., S_{max}\}$. A message consists exactly of one of these symbols. A symbol itself has no meaning to an agent, there is no predefined semantics at all.

A game ends successfully if the environment was transformed into the final state. In this case, the last two agents, participating in the game, get a certain amount $E^\star$ of "energy". In other cases there is no energy payoff. Every action that an agent performs consumes a specified amount of energy of the agent. There are low cost actions (*Null*, *Exit*, and *Replace*) and high cost actions ($Plant_x$, $Harvest_x$). For a low cost action the agent consumes energy $E_l > 0$, for a high cost action $E_l + E_h$, $E_h > 0$. The cost of the action *Bite* is $E_l + E_b$, $E_b > 0$. This action affects the other agent in the way that the "bitten" agent looses pain energy $E_p > 0$. At the beginning of an agent's life time its energy is set to $E = E_s > 0$, its start energy. If E ever falls below 0, the agent dies, i. e., the agent is removed from the population.

An agent does not know its own type nor perceives the type of another agent. They are black-boxes to each other. An agent perceives the message of another agent and – perhaps – some sensory input like the state of the environment or the fact of being bitten. In any case not all relevant aspects of the environment are known in the same way to all the participants. Agents must test different actions at different times and the only hint to whether an action or message was appropriate or not is given by a reward signal. This signal is always generated by the agent itself, based on the energy difference between two consecutive actions. A sigmoid function 1 generates the reward signal r based on the energy difference e_d; a positive energy difference results in a positive reward, a negative difference results in a negative reward. The reward generating function f_r is parametrized by two values, a and b:

$$f_r = 2 * a * \frac{1}{1 + exp(-b * e_d)} - a\ . \tag{1}$$

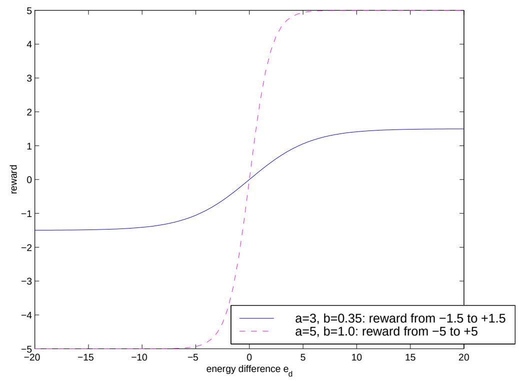

Figure 1: Two examples of the reward generating function.

Figure 1 shows two examples of the function. Thus, the individual learning of an agent is of reinforcement-learning type. This definition of a reward signal is a weak one, since it does not assume any intelligent observer (outside the agent) who generates a reward signal based on its knowledge about correct actions.

Beside an energy value agents have an age A, which at the beginning of an agents life time is set to 0. Any time an agent gets selected to play the game, its age will be incremented by 1. If the age reaches an individual maximum, A_{max}, the agent will be removed immediately from the population. At the start of the simulation, the population P consists of a certain number of agents P_s. The number of agents during the simulation may shrink or grow, depending on the fitness of the agents. An agent may enter the population if there are at least two agents, whose age is above value A_{sex} and whose energy value is above a value E_{sex}. The two "parents" are selected by a "Roulette wheel" [4] from all possible parent agents based on their energy value. Once a successful breeding occurred, the two parent agents are prevented from reproduction for a certain period of time t_{pause}. The amount of gene material parents pass on to their child is described in the next section. The general schedule of the simulation is:

1. initialize start population P_s
2. do forever
   - select randomly two agents, $Agent_1$ and $Agent_2$, and $t := 0$, $U_0 := 0$
   - $Agent_1$ generates and sends a start message M_0
   - do
     - $t := t + 1$
     - $Agent_2$ receives previous message M_{t-1} and generates action a_t and message M_t
     - $t := t + 1$
     - $Agent_1$ receives previous message M_{t-1} and generates action a_t and message M_t

     until $U_t = U_s$ or $a_t = Exit$ or $t = 2 + rounds$
   - remove from or add agents to the population

Whenever the number of agents in the population P_t falls below P_s, agents are randomly added to the population until $P_t = P_s$. If this situation happens it can be interpreted as restarting a simulation.

4.2 The Agent Architecture

We focused explicitly on one certain aspect of media, namely the relevance of expectations in choosing an appropriate answer to a received message. Thus, we combine an internal state with the expectation of a received message. This results in a frame-like structure which will be executed on two levels. In a first step a set F_i of frame structures is chosen based on the state of the environment. This step will be performed without any learning by the agent and is totally determined by the environment. In a second step the agent chooses one frame structure from the previously chosen set F_i. The selected frame will be executed resulting in an action a_{t+1} and a new message M_{t+1}. A frame F is defined with respect to a received message $M_r = M_t$ in the following way:

$$\begin{array}{llllll} \text{if} & M_r = M_{e1} & \text{then} & a := act_1 & \text{and} & M := mes_1 \\ \text{elseif} & M_r = M_{e2} & \text{then} & a := act_2 & \text{and} & M := mes_2 \\ \text{else} & & \multicolumn{4}{l}{\text{execute trouble frame} \quad F^T\ ,} \end{array}$$

where $a_{t+1} = a$ and $M_{t+1} = M$. The "trouble frame" F^T will be executed in the case that the received message was neither M_{e1} nor M_{e2}. This frame has a special structure, because it does not check the occurrence of a certain message, rather it checks whether the agent was bitten or not in order to determine the new action and message:

$$\begin{array}{llllll} \text{if} & bitten = true & \text{then} & a := act_{T1} & \text{and} & M := mes_{T1} \\ \text{else} & & \text{then} & a := act_{T2} & \text{and} & M := mes_{T2}\ . \end{array}$$

For every state of the environment the agent has $n_f \geq 1$ frames. The selection of a frame at time t will be guided by a Q-value Q_F, i. e., reinforcement learning ([11]) takes place in order to choose an appropriate frame in a given (environmental) situation. The entire collection of frames for an agent by a given final state U_e of the environment is:

state (depends on U)	frame set $F_i = \{F_{(i,0)}, \ldots, F_{i,n_f-1}\}$
0: $U^\star$	$F_{(0,0)}, \ldots, F_{(0,n_f-1)}$
1: $U = 0$	$F_{(1,0)}, \ldots, F_{(1,n_f-1)}$
2: $U = 1$	$F_{(2,0)}, \ldots, F_{(2,n_f-1)}$
$\vdots$	$\vdots$
U_e: $U = U_e - 1$	$F_{(U_e,0)}, \ldots, F_{(U_e,n_f-1)}$
U^T	$F_0^T, \ldots, F_{n_f-1}$

$U^\star$ is the frame set for an agent when it starts the communication by generating just a message M_0, and U^T is the frame set for the trouble state.

Evolution is based on frames, agents do not change frames during their life time, they are just able to change the Q-value of a frame with respect to other frames inside the same frame set. At the beginning of the simulation, all frames of all agents are initialized randomly. In particular, variables M_{e1}, M_{e2}, mes_1, mes_2, mes_{T1}, and mes_{T-2} get randomly chosen values from $\mathcal{S} = \{0, 1, 2, \ldots, S_{max}\}$, and variables act_1, act_2, act_{T1}, and act_{T2} get randomly chosen values from $\mathcal{A} = \{Null, Bite, Exit, Replace, Plant_I, Harvest_I, Plant_{II}, \ldots\}$. Inheritance happens on the frame level, i. e., cross-over takes place *between* frames, not inside a frame (but inside a frame set). Individual parts of a frame are subjected to mutation. Therefore, e. g. part M_{e1} or act_2 may get a new random value during mutation process. Q-values are not passed on to offspring, and are set to a small random value at the beginning of an agents life time.

4.3 Simulation Results

Figure 2 shows the general outcome of a simulation based on frame type agents. The maximum number of agents was set to 1024. The simulation started with 3 agents and as long as the number of agents was below 15 a higher energy pay off $E^\star$ was given for success than indicated in the figure (to support an onset of evolution). The number of agents grows rapidly until the limit is reached. Later, evolution still takes place optimizing the frame structures. This may result for example in changing cooperation sequences, or in a "competition" of different sequences (cf. Figure 3). A sequence is defined by M_0, M_1 a_1 M_2 a_2 ..., that is, M_0 is the start message of the first agent, M_1 the answer message and a_1 the action of the other agent and so on. The coding

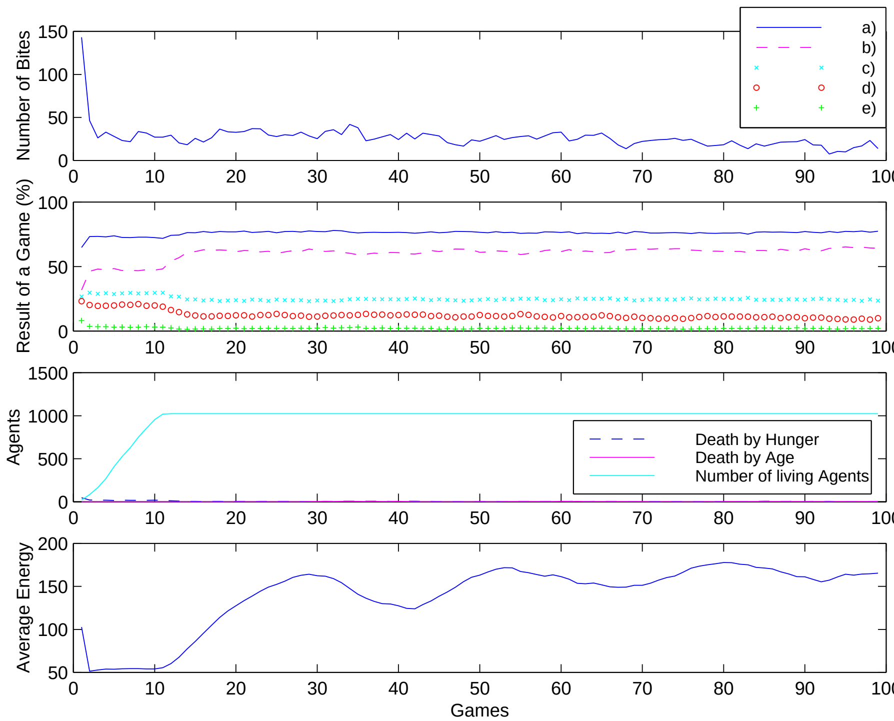

Figure 2: Simulation of 1000000 games of agents of the frame type. Results averages 1000 games (apart from the *Bites*-graph, which shows the total number of bites in 1000 games). Result of the simulation (legend on top for second graph): a : maximum possible success (counting the occurrence of a "correct" pairing of the agents); b : the actually achieved success; c : correctly performed *Exit*; d : *Exit* in wrong situation; e : stopped, because maximum rounds exceeded. $U_e = 4$, $S_{max} = 3$, $rounds = 10$, $E^\star = 10.0$, $E_l = 0.5$, $E_h = 2.5$, $E_b = E_p = 0.1$, $E_s = 50.0$, $A_{max} \in \{550, \ldots, 800\}$, $A_{sex} = 20$, $t_{pause} = 20$, $a = 5.0$, $b = 1.0$, $n_f = 2$.

of actions is: 0 - *Null*, 1 - *Bite*, 2 - *Exit*, 3 - *Replace*, 4 - *Plant*$_1$, 5 - *Harvest*$_I$,.... Because we analyzed only sequences which did not contain a *Replace*-action, and which were successful, all these sequences end with action 7. Figure 3 shows the eight most frequent sequences of the entire simulation. In detail, they are:

number (cf. Fig. 3)	number of occurrence	sequence M_0 M_1 a_1 M_2 a_2 ...
1	160877	1 0 4 0 5 1 6 2 7
2	66551	2 0 4 0 4 0 5 1 6 2 7
3	37402	0 0 4 0 5 1 6 2 7
4	26721	0 1 5 0 4 0 5 1 6 2 7
5	19039	2 1 5 0 4 0 5 1 6 2 7
6	7118	0 0 4 0 4 0 5 1 6 2 7
7	6453	2 0 5 0 4 0 5 1 6 2 7
8	5734	2 1 7 0 4 0 5 1 6 2 7

The first sequence occurs 160877 times, out of 346727 successful sequences, without a *Replace*-action.

The communicative behavior of agents becomes more and more regular. Because we have chosen $n_f = 2$ it is obvious that a frame set is assumed to contain exactly one appropriate frame for *Planters* and one for *Harvester*. An individual only has to explore which one is better suited. A detailed analysis of the communicative behavior reveals indeed that communication controls the behavior of agents. There are many

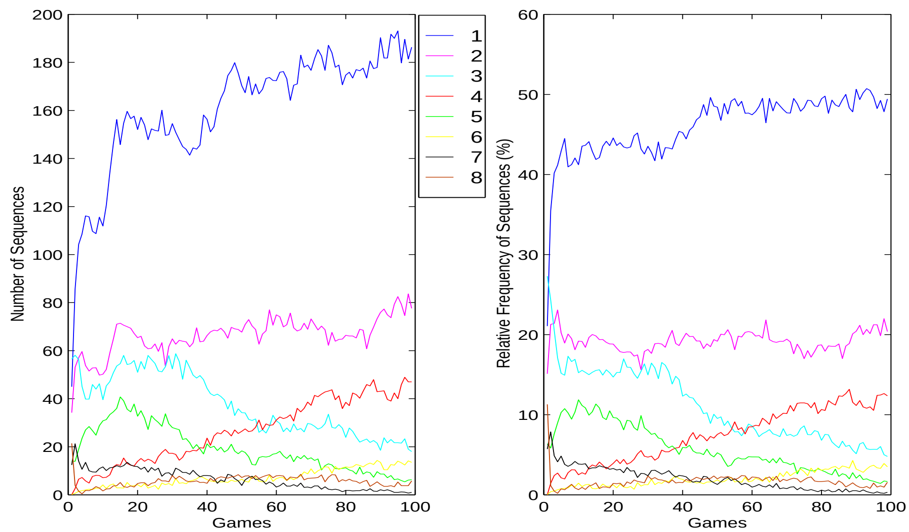

Figure 3: The eight main sequences of the frame based evolution. Left: Absolute occurrence of sequences, right: relative occurrence of the sequences (in relation to 346727 successful sequences). The eight sequences occur 329895 times. See text for further explanations.

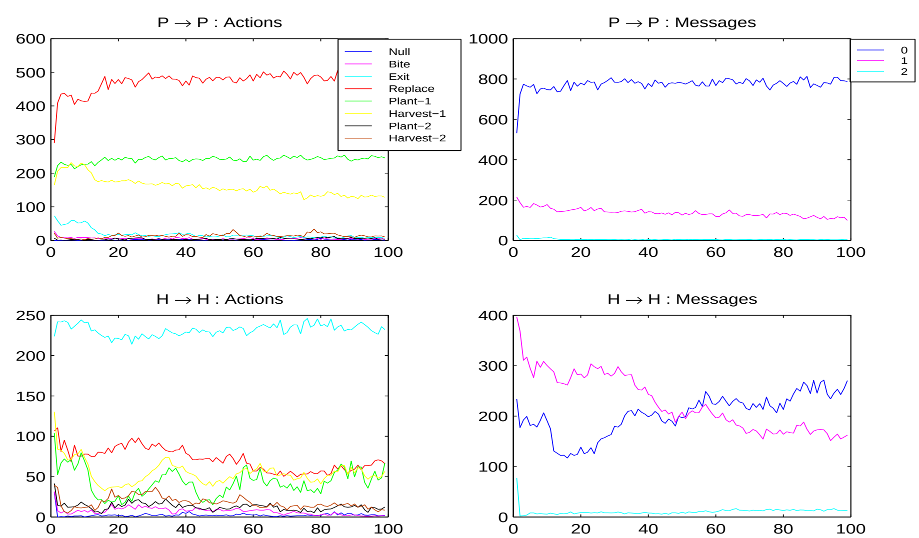

Figure 4: Performed actions and sent messages of *Plant* agents and *Harvester* agents with respect to unfavorable agent grouping. *Planters* (top row) prefer the *Replaces* action, *Harvester* (bottom row) prefer the *Exit* action. Both types of agents realized that the other agent was not a "correct" partner and that without having any direct access to the type of an agent (neither to the own nor to the type of another agent).

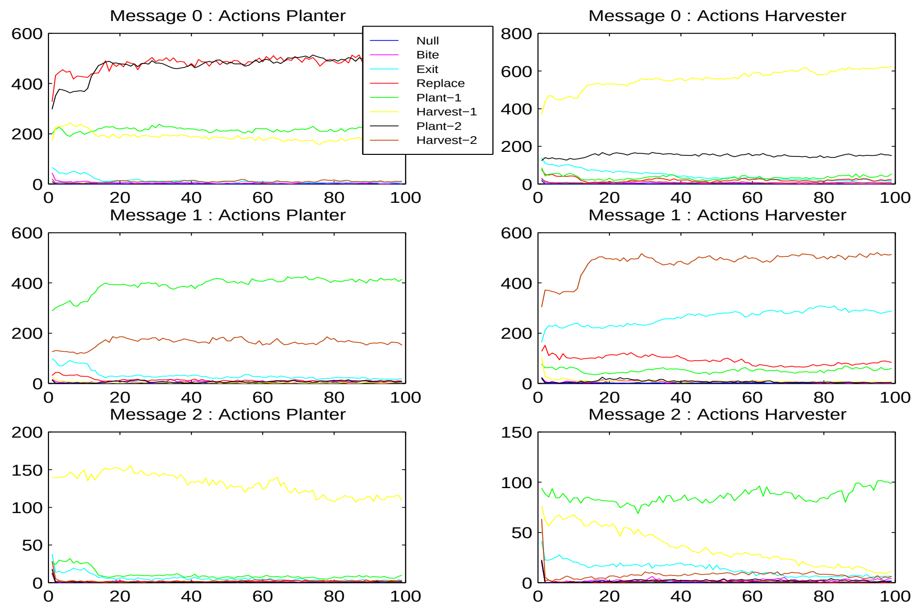

Figure 5: The actions of *Plant* agents and *Harvest* agents which follow a certain message.

different aspects which may be chosen (apart from the previous sequence analysis) to investigate the communicative and environment related behavior. Thus, we may ask for type specific actions, i. e., in the case when two *Planters* or two *Harvesters* are grouped together, what action is chosen (see Figure 4)? Another question may concern the question of what action is performed on the basis of a received message (see the analysis in Figure 5). Further investigations are certainly possible, especially a more detailed analysis with respect to a combination of two or more dimensions of interest. For example, in which situation (state of the environment) a *Planter* reacts with which action after receiving message 0 (cf. left graph on top of Figure 5)?.

As the results indicate, the agents were able to set up a population wide semantics for the exchanged symbols. The meaning of a symbol depends – of course – on the environmental state, however symbols became functional crucial for the agent's choice of the next message or action. Although not shown here, simulations are easily adapted to cases where several thousand agents may evolve, still acting in a coordinated manner.

5 Conclusion

In this paper we have shown, that a growing population of agents may act in a coordinated manner even in the case when the cognitive capabilities of the agents are limited and, moreover, when agents do not know anything about each other (apart from received messages). We started by questioning what kind of mechanisms human society evolved in order to cope with a growing number of individuals. We found an interesting answer in the work of sociologists like Parsons and Luhmann. A key answer of them are the proposed symbolically generalized media. We have modeled one of them (*p*ower) in a first approach. However, our simulation is still too simple to establish all aspects of a symbolic medium.

Nevertheless, Luhmanns suggestions regarding symbolic media, especially the aspect of structuring a situation by expectations, turned out to be useful. Luhmann, as well as Parsons, focused on the question of how the *p*roblem of doubled contingency may be solved on the social dimension. This problem, which can be described in a simplified way by the black box metaphor (see Section 2), taking seriously also in multi-agent research, may lead to new approaches regarding problems related to scaling up agent systems. The work described in this paper, may be seen as one such new approach. In subsequent work, we will deal with a more elaborated model of a symbolic medium. Further, the impact of more than one medium has to be analyzed, especially their potential for a more heterogeneous agent society and more complex problems to be solved by the agents.

References

[1] G. Ballot and E. Taymaz. Technological change, learning and macro-economic coordination: An evolutionary model. *Journal of Artificial Societies and Social Simulation <http://www.soc.surrey.ac.uk/JASSS/2/2/3.html>*, 2(2), 1999.

[2] K. Binmore. *Game Theory and the social contract*, volume 1: Playing fair. Cambridge, Mass.: MIT Press, 1994.

[3] R. Dunbar. *Grooming, Gossip, and the Evolution of Language.* Harvard University Press, Cambridge, Mass., 1996.

[4] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley Publishing Company, Inc., Reading, Mass., 1989.

[5] N. Luhmann. The Autopoiesis of Social Systems. In F. Geyer and J. van der Zouwen, editors, *Sociocybernetic Paradoxes*, page 172ff., London, 1986. Sage.

[6] N. Luhmann. *Social Systems.* Stanford University Press, Stanford, Ca., 1995.

[7] M. Paolucci and R. Conte. Reproduction of Normative Agents: A Simulation Study. *Adaptive Behavior*, 7(3/4):307–322, 1999.

[8] S. Park, E. H. Durfee, and W. P. Birmingham. Emergent Properties of a Market-based Digital Library with Strategic Agents. In Y. Demazeau, editor, *Third International Conference on Multi-Agent Systems (ICMAS98)*, pages 230 – 237, Los Alamitos, Cal., 1998. IEEE Computer Society.

[9] T. Parsons. *The Structure of Social Action.* Free Press, New York, 1968.

[10] A. S. Rao and M. P. Georgeff. Modeling Rational Agents within a BDI-Archtiecture. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484, Cambridge, Mass., 1991.

[11] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction.* The MIT Press, Cambridge, Mass., 1998.