Society first and then minds: Self organisation of a social symbol system by learning agents

Sigmar Papendick[1], Jörg Wellner[2] , and Werner Dilger[2]

[1] University of Konstanz, Department of Sociology
D-78464 Konstanz, Germany
[2] Chemnitz University of Technology, Computer Science
D-09107 Chemnitz, Germany

We show results of testing a behaviouristic approach of symbol meaning emergence in multi-agent populations. In contrast to dominating approaches in cognitive and linguistic science which focus on advanced social cognition as a precondition of the emerged language, we present simulation results that strengthen the theoretical position of G. H. Mead, a sociologist who claimed that the emergence of symbols has preceded the development of individual minds and social intelligence. In our simulations, positive reinforcement signals due to successful cooperation control the individual learning of symbolic behaviour and allow the emergence of robust, shared symbol systems in scalable multi-agent populations.

Introduction

The usage of individually acquired symbols is considered as a central criterion in most definitions of communication-based human social systems. Accordingly, individualistic approaches in cognitive and social sciences often stress the importance of specialised mental preconditions in order to explain symbolically coordinated social phenomena. In this paper, we follow a quite opposite approach: We want to present results of multi-agent simulations, showing that the evolutionary emergence of a shared meaning of symbol-sets does not require any initial cognitive capabilities at all, but a proper environment.

The paper shows that agents may learn to play symbolically coordinated cooperation-games of at least six different necessary actions, which is more than most approaches of the evolution of communication documented (see e. g. [Di Paolo, 1997], [Noble, 1999], or [Wellner et al., 2001]). Further, we show that a symbol system may be learned by agents through a reinforcement signal. Although reinforcement learning may be seen as not suitable in similar simulations (e.g. in [Oliphant, 1999] agents learn by observing each other), the setting of our simulation generates a reliable reinforcement signal, what agents can use to adapt their actions and communications. Agents, generating a shared symbolic meaning of visual objects do not have such a signal, because there is no environment actions can related to (e.g. [Smith, 2001]).

The paper is structured as follows: In the next section we give a short overview of the main arguments of our sociologically motivated approach. Then we present technical details of the cooperation task as well as of agent interaction and learning. Two different variations of the cooperation game are considered. In the first version, exactly two agents play the game, in the second one, an agent is automatically replaced randomly by another agent after performing an action. The last section gives a discussion of the limits of our approach and concludes the paper.

Sociological Background

The usage of elaborated languages in order to coordinate social behaviour seems to be an exclusive characteristic of human social phenomena. Although we find hereditary usage of symbols in most animal species (beginning with social insects up to mammals), it is striking that individually learned usage of symbols combined with the emergence of manifold symbol-based cultures within one species seem to occur only in human societies. Accordingly, the striking complexity of human cognitive and linguistic performance have led to reversal conclusions that these individual cognitive capabilities must be an inevitable precondition of symbolically mediated social phenomena. This position is supported by the fact that the usage of any "simple languages" cannot be observed in other species ([Deacon, 1997]: p. 39). It seems to be obvious that he presumably exclusive human efficiency in processing symbolically encoded meanings is not only a precondition of elaborated languages, but also a precondition for the emergence of ontogenetically acquired usage of symbols at all. Put to an extreme position, this argumentation explains language as an epiphenomenon of the evolution of large brains or as a tool to explain information transfer (for an overview and critique of these positions see [Dunbar, 1998]).

In this paper, we submit an approach that finds its origins in an – often considered as outdated – flavour of early social science: Social behaviourism. We address a specific argumentation of the social-psychologist and sociologist George H. Mead, a well known representative of this position. In the first two chapters in his book "Mind, Self and Society. From the standpoint of a social behaviourist" ([Mead, 1967], first published 1934) he argues that individual minds and concepts of self and identity can be considered as a *consequence* of the need to coordinate cooperation with gestures: "Mind arises through communication by conversation of gestures in a social process or context of experience – not communication through mind." ([Mead, 1967]: p. 50)

Unlike theories that explain language as a phenomenon following the development of social intelligence ([Worden, 1998]), Mead's approach claims an exactly opposite position. Additionally, his approach does not conflict with the theoretical baseline of "social intelligence"-approaches when the development of human brain size has to be explained ([Dunbar, 1998]: p. 93).

The most interesting aspect in Mead's argumentation from a sociological point of view is, that it may demand us to change the design of modelling completely: If we want to enjoy the benefits of individual artificial "intelligence", we might first have to set up an environment that requires cooperation and communication in order to coordinate social behaviour, before we can expect "intelligent" behaviour arising. In this perspective, "intelligence" can of course be regarded as a means in order to fit environmental constraints, but it has to be interpreted as well as an result of individual adaptation, being initiated by structured social behaviour.

Mead regards social behaviour as an early evolutionary adaptation to an environment in which individuals gain benefits if they manage cooperation. In order to coordinate their cooperation, they may learn – only by reinforcement, without any advanced cognitive capabilities – to observe others and to discover coincidences between details of the observed behaviour preceding distinct actions (these details are further identified as "gestures" that might include vocal signals as well) and these actions themselves. As a result, effective cooperation occurs if the observation of gestures (as a stimulus) allow individuals reciprocal predictions of the behaviour of others and selections of own functional behaviour as a learned response. The linkage between the stimulus (observed gesture) and the response (own behaviour, including aspects that can be identified as gestures by others) is controlled by the success of their cooperation. In this functionalistic perspective, Mead claims that gestures have a super-individual meaning, as long as they are used by individuals that have learned to react to them in an appropriate way in order to survive in a given physical and social environment. The definite transition from "gestures" to "significant symbols" occurs if an individual – showing a gesture and evoking a response in another individual – would show this response by itself, if its gesture would have been presented to it by the other ([Mead, 1967]: p. 46).

This definition of symbols seems to be very suitable for simulation purposes, because it allows a clear distinction of functional aspects of symbol-based communication on the one hand and subjective, representatory aspects of symbolically shared meaning on the other: From a subjective point of view, Mead introduces "mind" as a set of dispositions to behave in the same way, as an individual expects others to behave after sending them "significant symbols". In the presentation of our simulation design below, we will use the term "symbol" to refer to behavioural characteristics of simulated individuals (agents), that can gain – as we will show – the same *social function* as gestures and significant symbols in Mead's theoretical framework. By using this term, we will of course miss all subjective – representational – implications of significant symbols that are introduced by Mead in order to explain the symbol-processing characteristics of the human mind.

The Game and Learning Task

Our paper is sociologically motivated, therefore, we focus on individual learning of agents. Communication can be used by agents in order to succeed in a simple cooperation task. The constraints of the cooperation task are: Agents are not per se accomplished with a shared symbol system, because, at the beginning of the simulation, all agents (and all further added agents) are randomly initialised and react on exchanged symbols differently. These "symbols" are results of behavioural selections as any other behaviour selected by the agents, with the exception that this "symbolic behaviour" does not effect the agents environment or other agents directly, but can be observed by other agents[1]. Individual reinforcement learning is used to acquire some confidence about how to interpret a received symbol and what symbol is to be sent with an accompanying environment-related action. The sensory capabilities are restricted in the way, that agents do neither sense the state of the environment nor the action of another agent. They only receive a symbol sent by another agent.

We used an instance of a game, that is quite common in similar experiments. Depending on a received symbol, an agent chooses an appropriate action. A successful action benefits both agents, whereas an inappropriate action is of disadvantage to both agents as well. Thus, agents have to cooperate in order to gain a benefit. Most studies of the evolution of cooperation focus on only one pair of interactions. However, human interaction often consists of several pairs. In our simulation experiments, a game is a sequence of action-communication pairs. Each agent has to perform an action and a communication at the same time. The communicated symbol is received by the other agent. That symbol is used by the receiving agent to choose itself an action and symbol.

The state e of environment ($e \in \mathcal{E}$, $\mathcal{E} = \{0, 1, 2, \ldots, e_{max}\}$) is defined by a natural number. At the beginning of a game, the environment is in state $e = 0$. Actions $a \in \mathcal{A}$ ($\mathcal{A} = \{0, 1, 2, \ldots, a_{max}\}$, $a_{max} = e_{max}$ by definition) agents can perform are labelled by natural numbers too. In general, action $a = i$ transforms the environment from state $e = i - 1$ into state $e = i$. Action $a \neq i$ has no effect on an environmental state $e = i - 1$. Therefore, action $a = 0$ never has any effect to the environment. The effect of an action is further restricted by the following constraint: no agent is allowed to change the state of environment twice in a row. That is, if agent A_1 performed successfully an action $a = i$, cooperation comes into play. Only agent A_2 may perform action $a = i + 1$ effectively, changing the state from $e = i$ to $e = i+1$. As long as the state is $e = 0$ it does not matter, whether agent A_1 or A_2 performs action $a = 1$ first, both acting agents would change the state of environment. A game is successful if the state of environment becomes $e = e_{max}$.

Every action a_i is accompanied by a communication σ_i ($\sigma_i \in \mathcal{S} = \{0, 1, \ldots, s_{max}\}$). Therefore, a game of two agents A_1 and A_2 can be specified by a sequence of action-communication pairs $\{(a_i, \sigma_i)\}$ for $i = 0, 1, \ldots$ with a_i and σ_i performed by agent A_1 if ($i \bmod 2 = 1$) or performed by agent A_2 otherwise. A game is finished in either case: a predefined threshold of a maximum number of action-communication pairs R is exceeded or the environment is transformed into the (final) state e_{max}.

Agents are endowed with a simple energy mechanism E, which has an impact on two different levels of the simulation. Firstly, the change of the individual energy value produces an reinforcement signal for the adaptation of selection probabilities of actions and messages. Secondly, the amount of new agents, entering the simulation, is controlled by the energy value of the existing agents.

[1]In Mead's theoretical framework, the randomly selected symbolic behaviour at the beginning of a simulation corresponds to gestures.

Performing an action reduces the energy level of an agent by E_{act}. An energy payoff E_{payoff} will increase the energy level in the case the game is successfully played. At the end of a game, an agent calculates the difference of its energy level before and after the game $E_{game} = E_{end} - E_{start}$, generating the reinforcement signal.

The simulation starts with a predefined number of agents. Each introduced agent has a birth energy E_{birth} which allows it to do several games with negative E_{game}. The number of agents will be increased by 1 if there are two agents A_1 and A_2 with $C_{1,2} > C_{sex}$ ($C_{sex} > 0$ is the necessary number of games an agents must have played before being considered as a parent agent) and $E_{1,2} > E_{birth}$. These two agents are regarded as parents for the new one. Each agent, selected to be a parent will loose some energy E_{breed}. Selection of parents via a roulette-wheel mechanism takes place on the energy level of the agents. There is no inheritance of any parameter or learnt behaviour. Therefore, the appearance of a new agent is controlled by the fitness of the population. The better agents might play the game, the more agents are able to enter the population.

Learning takes place only after a game is finished. Furthermore, agents learn – by definition – only if E_{game} is positive. An agent maintains two separate tables, one for learning (and therefore selecting) an action, another one for message symbols. The selection of a new message and a new action is done by a roulette-wheel mechanism, based on the (real) values, for all messages and actions in the appropriate row. The adaptation of the values (initialised with positive random values less than 0.1) is generally done by

$$ACT(\sigma_{k-1}, a_{k-2})^{a_k} := ACT(\sigma_{k-1}, a_{k-2})^{a_k} + \alpha * f(\overline{E^m}, E_{game}) \tag{1}$$

and

$$MESS(\sigma_{k-1}, a_{k-2})^{\sigma_k} := MESS(\sigma_{k-1}, a_{k-2})^{\sigma_k} + \alpha * f(\overline{E^m}, E_{game}) \ . \tag{2}$$

a_k is the action, taken at time k, and σ_k is the message sent at time k.

To force agents to improve learning results and to overcome local optima we define f as a function $f = f(\overline{E^m}, E_{game}$, with $\overline{E^m}$) being the medium of the last m positive payoffs:

$$f = \begin{cases} E_{game}, & \text{if } E_{game} > \overline{E^m} \\ 0, & \text{otherwise .} \end{cases} \tag{3}$$

Simulation Results

By trial and error agents will solve the cooperation problem if R (the maximum number of communication-action pairs) is large enough and if E_{payoff} is suitable to keep agents alive. For given E_{act} and R one can calculate an appropriate E^*_{payoff} for any $\mathcal{S}_{max}$ in order to enable randomly acting agents to survive. Agents, that exploit the possibility of associating a meaning to symbols, will survive with less than E^*_{payoff}, and produce an average length of a game close to e_{max}, i. e., doing the right action exactly once in a game at the right time.

Two-Actions Game

The most simple but still interesting case is an environment with a final state of $e_{max} = 2$. A similar version is used in most approaches to the evolution of communication or cooperation. Figure 1 shows the percentage of successful games and the average energy of agents for different but fixed number of agents[2].

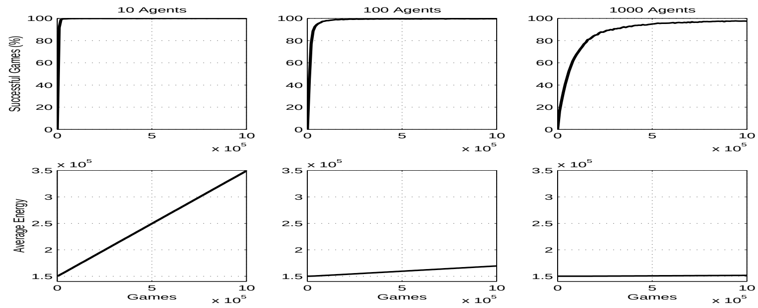

Figure 1: Three different simulation with fixed number of agents: Simulation parameters: $R = 2$, $e_{max} = 2$, $S_{max} = 9$, $E_{birth} = 150000.0$, $E_{act} = 1.0$, $E_{payoff} = 2.0$, $\alpha = 0.1$, $m = 10$.

The results indicate, that the learning algorithm is robust with respect to the number of agents. The increase of energy as an indicator of successful cooperation shows that agents are able to construct a shared symbol system and to exploit it (Note: random behaviour would result in 1/9 of successful games only and average $E_{game} = -0.78$).

All games finished after exactly two rounds ($R = 2$). A higher limit for the threshold could be favourable in order to give agents the chance to succeed in the third or forth round or so. But we have chosen $R = 2$ for the sake of comparability to other approaches. Furthermore, $R = 2$ means, that no agent selects a message or action based on a previously performed action (see Equations 4 and 5) , because the two agents playing a game each act only once in a game. Thus, a successful simulation indicates an established mapping $\emptyset \to \mathcal{S}$ and $\emptyset \to \mathcal{A}$ for the agent acting first and $\mathcal{S} \to \mathcal{A}$ for the second agent (the message of the second agent is of no interest because it is not received by another agent).

Six-Actions Game

If more than two actions are necessary to transform the environment from the initial to the final state ($e_{max} > 2$), the probability of succeeding by random behaviour decreases. An efficient and successful mapping of that kind is learnt by the population only if the average duration of a game is close to the optimal solution. For example, if the final state of the environment is e_{max} then the average duration of games should approach $r = e_{max}$ during a simulation in order to indicate the emergence of a shared symbol system and hence an establishment of a population wide meaning $\mathcal{M}$ of situations. However, every agent may use a different subset of $\mathcal{M}$ and it is not necessary that the tables for selecting a

[2] All results are average for 10000 consecutive games. All simulation results shown are typical for the given parameters.

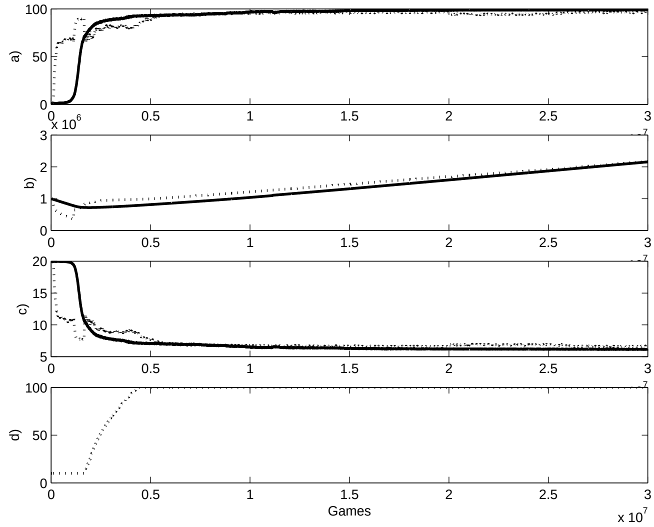

Figure 2: Simulations of a six action game. Solid lines is a simulation with fixed number of 100 agents, dashed lines is a simulation with a growing number of agents from 10 to 100. a) - successful games in %, b) - average energy, c) - average length of a game, d) - number of agents. Simulation parameters: $R = 20$, $e_{max} = 6$, $\mathcal{S}_{max} = 9$, $E_{act} = 1.0$, $E_{payoff} = 6.0$, $E_{birth} = 1000000.0$, $E_{breed} = 1000.0$, $C_{sex} = 20000$, $\alpha = 0.2$, $m = 20$.

message or action (in agent's head) are concurrent, favouring the same message or action in the same situation.

Figure 2 (solid lines) shows a simulation with a fixed number of 100 agents. The greatest energy gain in a game an agent can achieve is $E_{game} = 3.0$ (random behaviour would result in approximately 0.79% successfully played games of average length 19.98). A game of 12 rounds would keep the energy level of the agents constant. Because some games are longer than the average length which results in an energy loss, the average energy of agents is still decreasing for round length of about 12. After reducing the average length of a round to about 8 the average energy starts growing. At the end of the simulation the average length of a game approaches 7.0. That is not optimal but nearly optimal and is only possible when the agents have established a mapping for $\mathcal{M}$.

Dashed lines of Figure 2 are the result of a similar simulation, but with an increasing number of agents from 10 to 100. During the first part of the simulation (up to 200.000 games) the 10 starting agents build a shared symbol system, as indicated by increasing average energy and achieved small length of rounds. As long as new agents enter the population values for the percentage of successful games and length of a game oscillate, but never run out of an unfortunate range.

Four-Actions Game with permanently exchanged agents

The six-action simulations indicate that even difficult problems are learnable by the proposed agent architecture, even without sensing the actual environment (apart from messages). However, there is one argument, which says that a correct interpretation of messages is necessary only in the first interaction of a game in which agents signal whether they start first. Once, this is known by the other agent, the selection of actions might to be dependent only on the previously taken action. Hence, the particular symbol received would be of less importance for selecting the right action.

To investigate this argument, we changed the definition of a game in the way that every agent will be replaced randomly by another agent after performing a message-action pair. Thus, no agent is able to base any selection of the next message or action on its previously taken action. Now, every action and message selection is based solely on the received message. The learning equations, corresponding to 1 and 2 are:

$$ACT(\sigma_{k-1})^{a_k} := ACT(\sigma_{k-1})^{a_k} + \alpha * f(\overline{E^m}, E_{game}) \tag{4}$$

and

$$MESS(\sigma_{k-1})^{\sigma_k} := MESS(\sigma_{k-1})^{\sigma_k} + \alpha * f(\overline{E^m}, E_{game}) \ . \tag{5}$$

The payoff energy E_{payoff} for a successful played game is now divided by the length of a round and multiplied by the number of actions an agent did take part at that particular round. For example, if an agent did take part two times at a successful round of length 10, then it receives a payoff of $2 * E_{payoff}/10$.

Table 1 shows the result at the beginning and at the end of a simulation, starting with 5 agents for $e_{max} = 4$ (random behaviour would result in approximately 12,09% successful games of average length 9.8). The occurring of action-message combinations during the games is counted. For example, in the first games action 1 is accompanied by all possible messages to a certain degree. At the end of the simulation, only message 3 and 4 are used by agents to indicate performing action 1. Moreover, message 3 and 4 are not longer user with any other action, thus, signalling the receiving agent indeed that action 1 was performed. The last row of the table still does not show any structure, because the average length of games at the end of the simulation approached 4, therefore, the message, sent by an agent performing action 4 was almost never received by another agent.

The emerged mapping of communication-input to selected behaviour can be regarded as the emergence of significant symbols, if coordinated cooperation still works efficiently even if individuals can be randomly replaced by others already living in the population. Under these conditions, it does not make any difference on the level of performed cooperation if an individual plays all the cooperation-game on its own or has to interact with others. The fact, that there might exist a few "synonymous" symbols for the same action is of no relevance in terms of the definition of meaning, as long as there are no multiple action-dispositions for the same symbol.

Conclusion

The simulation results presented in this paper have shown, that cooperation, coordinated by individually learned symbols, is possible even without specialised symbol processing ca-

Table 1: Simulations of a four action game with permanently replaced agents. Table entries (in percent) indicate the usage of a message depending on the chosen action. The top part of the table covers games 10000 to 20000 (with 5 agents), the bottom part covers games 29990000 to 30000000 (now with 100 agents). $R = 10$, $e_{max} = 4$, $\mathcal{S}_{max} = 9$, $E_{act} = 1.0$, $E_{payoff} = 4.75$, $E_{birth} = 500000.0$, $E_{breed} = 100000.0$, $C_{sex} = 100000$, $\alpha = 0.5$, $m = 10$.

	message symbol									
action	0	1	2	3	4	5	6	7	8	9
games 10000 - 20000										
1	2.37	6.05	3.14	42.32	30.34	2.52	3.76	3.65	2.68	3.18
2	2.48	14.11	3.85	3.12	5.27	8.62	5.14	36.85	5.09	15.46
3	2.81	23.39	20.39	3.41	7.26	4.19	17.63	5.94	8.53	6.45
4	9.79	6.59	10.94	12.08	12.15	7.62	14.20	5.67	11.44	9.53
games 29990000 - 30000000										
1	0.01	0.02	0.00	48.97	50.98	0.00	0.00	0.01	0.00	0.01
2	0.01	0.01	0.00	0.00	0.00	0.01	0.00	47.89	0.01	52.07
3	0.01	16.50	22.23	0.00	0.00	17.11	25.93	0.03	18.19	0.00
4	10.32	9.99	9.23	10.78	9.99	10.31	10.97	9.75	9.73	8.91

pabilities. An evolutionary setting that requires cooperation, simple reinforcement driven learning mechanisms and the necessity to observe and perform a behaviour that does not directly affect cooperation seems to be sufficient to allow the emergence of shared symbol systems. An additional constraint, important in our approach, is an environmental setting that demands for cooperation due to an information gap: Because agents cannot sense the environment, they are forced to observe behavioural characteristics, developed later to symbols, in order to compensate the information gap. The learning task was structured in such a way that successful coordination of agent's action presupposes a shared symbol system of the population and vice versa. Once a shared symbol system was established, the usage of symbols was robust with respect to increasing population size, if the population's rate of growth is controlled by individual fitness as an indicator of individual adaptation to an emerged symbol system.

Mead's basic theoretical approach that assumes the individual acquisition of emerged symbols occurs prior to the evolution of mind, is supported by our simulation results as far as functional aspects of communication are concerned: individual representations especially designed to control symbolic or cooperational behaviour are not necessary at all on the chosen level of cooperational complexity tested here.

The restrictiveness of our approach does not seem to be a deficiency, as long as we do not claim to explain phenomena in human social reality – this is clearly not our intention. Following Mead's theoretical approach, we want to test if his functionalistic view on meaning as an emergent phenomena in cooperational contexts can be generalised and beneficially used to improve modelling of systems of distributed artificial intelligence.

Being successfully with this modelling approach does not mean having identified the causes of human language evolution. At its best, it might allow the conclusion that the principles identified in the modelling process might be necessary, but clearly not sufficient preconditions for explaining the evolution of human language.

Acknowledgements

This work is supported by the *Deutsche Forschungsgemeinschaft* under grant number DI 452/10-1. The paper benefited from the comments of three anonymous reviewers.

References

[Deacon, 1997] Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and brain.* W.W.Norton & Co., New York.

[Di Paolo, 1997] Di Paolo, E. A. (1997). An Investigation into the Evolution of Communication. *Adaptive Behavior*, 6(2):285–324.

[Dunbar, 1998] Dunbar, R. (1998). Theory of mind and the evolution of language. In Hurford, J., Studdert-Kennedy, M., and Knight, C., editors, *Approaches to the evolution of language.* Cambridge University Press, Cambridge.

[Mead, 1967] Mead, G. H. (1967). *Mind, self and society. From the standpoint of a social behaviorist.* University of Chicago Press, Chicago and London.

[Noble, 1999] Noble, J. (1999). Cooperation, Conflict and the Evolution of Communication. *Adaptive Behavior*, 7(3/4):349–369.

[Oliphant, 1999] Oliphant, M. (1999). The Learning Barrier: Moving from Innate to Learned Systems of Communication. *Adaptive Behavior*, 7(3/4):371–383.

[Smith, 2001] Smith, A. D. M. (2001). Esablishing Communication Systems without Explicit Meaning Transformation. In Kelemen, J. and Sosik, P., editors, *Advances in Artificial Life*, Berlin. Springer-Verlag.

[Wellner et al., 2001] Wellner, J., Papendick, S., and Dilger, W. (2001). Scalability and the Evolution of Normative Behavior. In Zhong, N., Liu, J., Ohsuga, S., and Bradshaw, J., editors, *Intelligent Agent Technology. Research and Development. Proceedings of the 2nd Asia-Pacific Conference on IAT*, New Jersey. World Scientific.

[Worden, 1998] Worden, R. (1998). The evolution of language from social intelligence. In Hurford, J., Studdert-Kennedy, M., and Knight, C., editors, *Approaches to the evolution of language.* Cambridge University Press, Cambridge.