On the Nature of Culture and Communication. A Complex Systems Perspective

Jürgen Klüver and Christina Stoica

Abstract

The article describes the foundations of a new theory of communication as the generative core of cultural systems. We define the concepts of meaning and information in terms of complex systems theory; communication is considered as the interplay of social and cognitive dynamics, i.e., interactions between speakers and receivers, which are determined by social and cognitive rules. The theory is validated by several computational models developed to analyze some important aspects of the theoretical foundations. The computational models themselves are validated by some empirical social experiments with student groups.

1 Introduction

Neither in sociology nor in anthropology the term of "culture" is defined in a way that most scholars would agree upon. We do not wish to enumerate the many theoretical approaches to give a general definition of this concept. To make matters worse, "culture" is often identified with "social" and accordingly research on society deals with the different concepts of "culture" and "social structure" in the same way. Yet if one wishes to model certain features of social processes, i.e., social dynamics, in a precise way, it is necessary to distinguish between "cultural" and "social" and to define both concepts as different, although interdependent, aspects of societies.

Habermas [1981, II] has given such a definition that is suitable for modeling purposes. He defines social structure as "the set of all rules that are valid in a particular society" [cf. also Giddens 1984] and culture as "the set of all knowledge that is held true in a particular society" [cf. also Geertz 1973]. To be sure, these definitions are very general and must be explained. For example, "knowledge" must contain concepts such as worldviews, belief systems, values, and the like. It is also necessary to distinguish between different kinds of social rules [cf. Klüver 2000]: There are on the one hand rules of interaction that determine particular forms of behavior if the respective situation demands it. On the other hand there are "topological rules", i.e. rules that define "who can interact with whom". This second type of rules defines a social topology. Yet as a basis for this article, these general definitions

1

2 We are quite aware that we do not quote Habermas in the exact way he stated his definitions. But for the purpose of this article this variation of Habermas is best suited.
are sufficient. Therefore, cultural systems will be characterized in the manner of Habermas, i.e., as sets of true knowledge – "true" in the eyes of the members of a society. But how are such systems generated as a dimension of a society? In other words, by quoting the classic study of Berger and Luckmann [1966], how can "the social construction of reality", which is done in the framework of a particular culture, be understood?

At last since Mead – although he does not systematically use the term of communication – it is a truism that all social processes are at their core communicative ones. As the social construction of reality, i.e., the construction of the knowledge by which we define and interpret reality, is a process of social interaction, the first answer to our question is: Culture is generated by the socially determined processes of communication between the members of a particular society. Therefore, the next question is: What is communication?

Having come so far, every social scientist is sadly aware of the fact that there is no general theory of communication that has been formulated in a precise, i.e. mathematical, way. There are doubtless numerous theoretical attempts to define communication; the theories, e.g., of Habermas and Luhmann [1984] are just two examples of the different possibilities for handling this problem. But no general definition of communication exists and certainly, since the famous theoretical approach by Shannon and Weaver [1969], no serious attempt has been made to develop a general mathematical theory of communication that can be of use to the social sciences [cf. the overview in Favre-Bulle 2001]. The "mathematical theory of communication" of Shannon and Weaver cannot fulfill this task, as has been frequently noted. Their "thermo-dynamical" definition of "information" and their deliberate neglect of "meaning" are only two reasons for this deficit (see below).

Because of this general deficit and because of the fundamental importance of the concept of "communication" for all of the social sciences, we try to sketch in this article the foundations of a mathematical theory of communication. This theory, in contrast to the Shannon-Weaver approach, is suited for the social sciences and in this sense makes a contribution to the foundations of social science in general. In this paper, we give some general definitions to illustrate these purposes and in addition, we describe several computational models based on these theoretical ideas. The models already contain some of the theoretical considerations discussed in the following sections; in particular, they can and must be understood as the computational operationalization of the theoretical foundations.
2 Communication as a Dynamical Process: Some Theoretical Definitions

Communication is a (dynamical) process that consists of at least two communicators, A and B, who perform communicative acts. The well-known speech acts described in the theories of Austin and Searle may be taken as examples although communication in the sense defined here is not restricted to verbal communication. Each communicator must be considered as a complex cognitive system; the communicative acts generate a cognitive dynamic in each of the respective communicators. The communicative acts are to be understood as the transfer and exchange of meaning and information; the acts are regulated by social rules and rules for the production of signs that are often coded as symbols. Therefore, a communicative situation - or communicative system respectively - consists of different communicators, certain social rules that regulate the interactional dynamics between the communicators, semiotic production rules that regulate the combination of the signs or symbols respectively that are used in the communicative situation and cognitive rules that regulate the cognitive dynamics of the communicators. The communicative process begins by defining a communicative theme that generates certain initial cognitive states of the communicators.

Consider, for example, the communicative situation of an examination at a university. The social rules are characterized by the different roles of the communicators, e.g., a professor and a student. Socially, the situation is defined as an asymmetrical one; i.e., the status of the professor is higher than that of the student and accordingly the student has to answer the questions of the examiner and is usually not allowed to ask questions himself. The semiotic production rules are defined by the themes of the examination. Neither the examiner nor the student is allowed to generate symbols, i.e., concepts and sub-themes that do not belong to the theme of the examination. During an examination on, e.g., evolutionary biology or computational sociology discussions about the works of Shakespeare are not admissible, even if both communicators wish so. The cognitive dynamics that are generated within each of the two communicators can be understood as the generation of certain semantic fields or networks. The questions of the examiner cause the student to generate certain associations of concepts that belong, in the opinion of the student, to a correct answer; accordingly, the answers of the student cause certain associations within the examiner that give rise to new questions and so forth.

The communicative process, then, must be considered as a two-dimensional dynamical process. On the one hand, the social rules of interaction, including the social topology, generate a particular interactional dynamics between the communicators; i.e., questions, answers, evaluation of the answers, new questions in the case of the
examination. On the other hand, cognitive rules and cognitive topologies generate certain cognitive dynamics; i.e.,
the cognitive processing of the messages. In our simple example this consists of the generation of certain concepts
suitable as answers or of new questions. Finally, the whole communicative process and its particular dynamics are
regulated by a mutual interdependency between the interactional and the cognitive dynamics. We may therefore
characterize the dynamics of communication as the result of two kinds of interpolating dynamics.

Readers who are acquainted with the general semiotic theory of signs will immediately perceive the
relationship of this general definition to the three-dimensional model of signs by Morris [1970]. Social rules and
topologies refer to the pragmatic dimension and the concept of semiotic production rules refers to the syntactical
dimension of signs as well as to the semantic dimension. In a rather abstract sense our general definition may be
understood as the transformation of the classical model of Morris into the framework of complex dynamical
systems.

Some additional remarks regarding these very general and abstract definitions are necessary. We use the
framework of complex systems theory to define communicative processes. There are several reasons for this
theoretical approach [cf. Klüver 2000; Klüver et al. 2003]. One of its main advantages is that it allows us to model
communication, cognition and social interaction quite naturally within one and the same theoretical framework: On
the one hand the individual communicators are modeled as complex cognitive systems; on the other hand the
exchange of communicative acts is modeled as a complex social interactional system and last but not least the whole
communicative process is understood as a complex communicative system, whose communicative dynamics is
generated by the interplay of social interactional dynamics and cognitive dynamics.

An important consequence of this theoretical approach is that one can use certain kinds of formal models to
capture the important characteristics of communication for all dimensions of the communicative process; i.e.,
models such as cellular automata, Boolean nets and neural networks [Klüver et al. 2003]. Although these formal
models greatly differ on a first sight it can be shown that they all offer the important possibility to capture complex
processes in a similar manner: One can translate the respective empirically valid rules of social or cognitive
interactions in a direct manner into formal rules of interaction of the formal models. Therefore, this theoretical and
methodological approach is a very general one. No other assumptions about specific characteristics of actors or
society are needed than this framework, i.e., that the social and/or cognitive systems can and must be understood as complex adaptive systems.

A model of the examination process described above, therefore, would define the social situation of the examination as a social or communicative system, respectively, in which the process of formulating questions, answers, evaluations and new answers is considered to be part of the dynamics of the situation; i.e., changes in the system's states. Accordingly, the communicators are modeled as dynamical cognitive systems whose dynamics, i.e., the generation of concepts, is regulated by the semiotic rules and the social rules of interaction on the one hand and certain cognitive rules on the other hand.

"Topology" and "rules of interaction" refer to the distinction we made above [see also Klüver 2000]. The dynamics of complex systems depends not only on (local) rules of interaction that are applied in particular situations, that is if the conditions of application are fulfilled, but also on "topological rules"; that is, rules that regulate "who interacts with whom" [cf. Cohen et al. 2000]. Therefore, both kinds of rules have to be taken into account. The same distinction is important for the modeling of cognitive systems.

The central concepts for each definition of communication are without doubt "meaning" and "information". It is not possible for us to discuss here the countless attempts to define both concepts. In terms of complex systems theory we define the "meaning of a signal" as an (point) attractor in the cognitive system of the receiver. In other words, a signal causes a certain cognitive dynamic to be activated in the receiving cognitive system. If the dynamics generate a simple point attractor, this state or sub-state is literally the meaning of the signal. For example, consider the famous dog of Pavlov. The conditioned learning of the dog – producing saliva after the hearing of a bell – can be interpreted as a newly learned meaning of the bell; i.e., an attractor state in the brain that in turn generates the saliva production. It can easily be demonstrated that much more complex forms of meaning can be defined this way. In this manner the meaning of a signal also defines the memory of the signal. When the dog receives the signal of the bell and generates the newly learned attractor state, then the dog remembers; i.e., it combines with the reception of the signal the meaning the dog has attached to it during the process of conditioning.
By the way, it is important that the cognitive system generates a point attractor when receiving a signal. If the system is only able to generate an attractor of period \( k > 1 \), then the system would literally oscillate between different states and would be unable to act, for example producing saliva. As fatal result the system would not survive in the long run. Therefore, the ability to generate meaning as point attractors and to remember them, i.e. to generate the attractor states again is essential for surviving.

For another example we may imagine encountering an aunt that we have not seen for some time. In the meantime, the aunt has a new hairstyle. We generated in the past an attractor, "aunt with the old hair-style". Despite the visual difference between the aunt with the new hairstyle and the memory of the old one, we immediately recognize her again; i.e., we actualize the memory of her. The explanation for this ability, in terms of complex systems theory, is quite simple. Our cognitive network that generates the attractor "aunt" when seeing or hearing about her must be considered to be systems of Wolfram class 2 [Wolfram 2002].

Wolfram classes are a general taxonomy for complex systems: Class 1 is the simplest case, containing only systems with point attractors, i.e., attractors consisting of one point in the state space. Class 2 consists of systems with point attractors and "simple" attractors with a period \( k > 1 \), but not very large periods. Class 3 is the class of "chaotic systems that is systems with no attractors or only so called strange attractors (consisting of a certain region in the state space). Class 4 contains the "real" complex systems, i.e., systems whose trajectories contain all kind of attractors but are not chaotic in the sense of class 3.

A Wolfram class 2 system now means that the system will generate the same attractor when receiving different inputs that generate different initial states; the set of these different initial states or different inputs respectively is the so called basin of attraction for that attractor. (That also is the case with systems of class 1, but for specific reasons these systems would be too simple for the cognitive tasks described in this section.) We can explain our ability to recognize the aunt despite the change in her hair-style by assuming that the different initial states generated by the perceptions, "aunt with the old hair-style" and "aunt with the new hair-style", belong to the same basin of attraction of the "aunt" attractor (for that particular aunt, of course). A simple picture may illustrate this:
By the way, it is often mentioned, although never explained, that one of the great advantages of artificial neural nets is their tolerance with respect to faults; i.e., they recognize a signal they have learned even when the signal is distorted in some ways. This capability can be explained, of course, in the same manner. Artificial neural nets generate, as a result of their training processes, a topology that causes the behavior of the nets to act like a system of Wolfram class 2. There are reasons to assume that many learning processes in human and artificial systems result in the generation of Wolfram class 2 systems.

In a parallel way it is possible to explain our ability to recognize the new hairstyle as a new hairstyle. To explain this one has to assume that the perception "aunt with the new hair-style" generates not only one attractor, the same as for the aunt with the old hair-style, but also two attractors at the same time. Such phenomena often happen in sufficient large systems. The cognitive system then compares the two attractors and the difference is understood as the changing. That is the new hair-style. Apparently, it is possible to give the concept of "meaning" an exact definition that allows the modeling of communication, insofar as it is also understood as an exchange of meaning, in a precise and formal manner.

This definition of "meaning", by the way, has the advantage that it takes into regard the fact that the same perceptions, experiences and the like do not necessarily have the same meaning for different persons. The exact meaning of the perception of the mentioned aunt has of course different meanings for her nephew and for a stranger. The subjective aspect of meaning is obviously captured with this definition. The semantical aspect of meaning, on the other hand, is contained too because the meaning as some point attractor refers to an object outside of the cognitive system. Therefore, the meaning of "aunt" is not arbitrary but linked to the object of our perception. The
subjective aspect of this concept of meaning lies in the fact that perceptive processes are always dependent on the biography of the perceiver.3

For the definition of ”information” we refer to the famous definition by Shannon and Weaver. Informally, according to Shannon and Weaver, the degree of information of a message or signal is greater the more improbable the message is and vice versa. Their definition, however, is not literally applicable to social communicative processes because it presupposes an ”objective” concept of probability: In the definition of Shannon and Weaver the informational content is always the same for senders and receivers. This assumption is not valid in usual communicative processes. But it is possible to take over the basic idea of this classical definition and define the amount of information in a similar manner.

Imagine two communicators A and B. When they speak about international politics and A tells B that country X has declared war on country Y, then the degree of probability of this message for B depends on his/her knowledge about this political situation. If B is the chief of an Intelligence Service then B may have expected the message. It is very probable to him and therefore the degree of information is low. If, on the other hand, B is a political layman and he believed the talks of peace from the leaders of country X, then the degree of information is high because B did not expect it.

Accordingly, we define the amount of information of a message as a measure of the difference between the message given by the speaker and the message expected by the receiver. It is important to note that this difference relies on the knowledge the receiver had about the respective subject before hearing – or reading – the message. This expectation could be obtained, for example, via an interview with the receiver before the communication. When both messages are coded as vectors and their components coded as real numbers between 1 and 0 then the amount of information can be measured as a vector distance. The factual message then may be expressed as

\[ X = (x_1, x_2, ..., x_n), \]

where as the expected message is given by

\[ Y = (y_1, y_2, ..., y_n). \]

3 The same definition is given for artificial neural nets by McLeod et al. 1998.
To be sure, the more a message is unexpected, i.e., improbable for the receiving system, the greater the difference between the x- and y-components and the greater the amount of information of the message. Therefore we may define

$$pi = 1 - \left| (xi - yi) \right|$$

to be the expectation probability of the component $xi$; i.e., the deviation from the expected component $yi$. Note that this is a "subjective" probability, i.e., it need not be the same for both communicators, in contrast to the Shannon-Weaver probability.

By summing up over all components of the respective vectors and by using the Shannon-Weavers assumption that the degree of information of a message is the negative probability of the message we obtain

$$I = -\sum_i pi$$

and by including the logarithm (dualis) for normalization purposes we obtain precisely the Shannon-Weaver definition

$$I = -\sum_i pi \ast \log_{pi}.$$  

The difference between this definition and the Shannon-Weaver formula is, of course, the different concept of probability that is being used. But this is necessary because factual communicative processes also depend on the factual knowledge of the communicators. Therefore, we had to introduce a different probability concept.

The main reason for this definition of information with its orientation to the classical Shannon-Weaver formula is that we believe that Shannon and Weaver introduced an important insight with their definition. The necessary step beyond their definition, so as to make it suitable for application to real communicative processes between human communicators, is the replacement of their physical definition with a subjective cognitive definition; i.e., a definition that is orientated to knowledge and therefore to the cognitive biography of the communicators. It is not a trivial result that the resulting mathematical definition has literally the same form as obtained with Shannon-Weaver’s objective approach.

It is a truism that each theory of communication needs – minimally – precise definitions for "meaning" and information". Our demonstration of the possibility of making such definitions shows that a new mathematical theory of communication suitable for the social sciences is a real possibility.
3 A parameter for the classifying of communications

We defined communication as a two-fold dynamics, namely a social-interactional one and a cognitive one. Although the cognitive processes of the different communicators play an important part in the resulting communicative dynamics it seems undeniable that factual communicative processes are mainly dependent on the social situation they occur in. "Social situation" means, of course, the social rules of interaction that characterize a communicative situation as a particular social situation.

Imagine a corporal at an army and a recruit who receives different orders by the corporal, e.g., throwing himself into the mud. The cognitive processes of the recruit are irrelevant for the communicative process; in particular if he likes the orders or if he silently insults the corporal. The social structure of the situation is such that the recruit can communicate only with "yes, Sir" and act according to the orders. The same dependency on the social rules can be observed in the example of the examination, where the factual communication is predetermined by the rules of such social situations. Yet not only the factual communicative process is determined by the social rules but the cognitive processes of the communicators too: the recruit fares best if he only thinks about the carrying out of the orders; the student in the examination has to try to think only about the subjects the examiner asks him. Such hierarchical rules of interaction force the cognitive processes in strictly determined directions.

Social situations that are more symmetric than these two examples, on the other hand, allow more freedom for the respective cognitive processes and accordingly more freedom for the resulting communicative ones. Two friends who are communicating as equals are obviously at liberty to chose different themes of communication, to answer or not, as they like, and so on. To be sure, communicative interactions between friends have also to obey certain rules such as, e.g., not to insult the other. Yet these rules do not determine the factual communication to the degree as the rules of the other two examples. We may therefore conclude that the social rules of interaction, insofar as they express more or less hierarchical relations, determine both communicative and cognitive processes in a less or more restricted manner.

The social rules in these examples express social relations in the dimension "higher – lower". In former publications we have developed the concept of the dimensionality of social systems (cf. Klüver 2000; Klüver et al. 2003) and have shown that social systems can be characterized as three-dimensional social spaces. One dimension is the "higher – lower" one that can be visualized as a vertical dimension. The poles of "belonging to – strange" that may be visualized as a horizontal dimension give another dimension. The pair "active – passive" that completes our
space of three dimensions gives the third dimension. It is important to note that these dimensions are basically nothing else than a mathematical translation of the well known three types of social differentiation [cf. Habermas loc. cit.; Luhmann loc. cit.]: "Strange versus belonging to" corresponds to the segmentary differentiation, "higher versus lower" to the stratificatory differentiation and "active versus passive" to the functional differentiation. Insofar as these three types of differentiation are sufficient to completely characterize the basic structure of a society it is possible to describe each social situation by the measure the according social rules exhibit in each of the different dimensions.

If we characterize a certain set of social rules by a numerical value in each of the dimensions then the simplest way is to define one pole as 1 and the other as 0. As the social difference in the vertical dimension between the corporal and the recruit is very large the according measure for this dimension would be approximately 1. Because they belong to the same military unit their difference in the horizontal dimension is very small and therefore the measure is near 0. Their difference in the functional dimension is again very large – only the corporal is active – and accordingly the measure in this dimension is approximately 1. The two friends on the other hand obtain for all dimensions very low values because their social difference is very small, at least with respect to their friendly communication, they are belonging to the same group of friends and they have the same active parts in the communication. These considerations allow us to define the communicative degree, \( cd \), of a communicative situation, dependent on the respective social rules:

\[
\begin{align*}
\text{cd} &= \sum_{i=1}^{3} \text{dim val}_i \\
\text{where dimval}_i \text{ is the numerical value of the social rules for the dimension i. A particular value of cd can in this sense be understood as a linear combination of the respective unit vectors of each dimension; yet the unit vectors are something like the "ideal types" of Max Weber because they seldom characterize real communicative situations.}
\end{align*}
\]

The exact measurement of the cd-values for empirical situations is, of course, not always easy and sometimes they must be the result of estimating processes. Yet as we shall demonstrate in the next section it is often possible to get satisfying results with such estimations.

By considering the different examples given above the following general communicative hypothesis seems rather evident:
The lower the cd-value of a certain situation is the greater is the degree of freedom for the communicative and cognitive processes, i.e. the more complex are the resulting communicative and cognitive dynamics respectively.

The converse holds too.4

Another factor must be taken into account that does not strictly belong to the social characteristic of a situation but is also determining the factual cognitive and communicative processes. This factor is simply the time the communicators have for communication before some non-communicative action has to occur. In the case of the corporal and the recruit there is no time at all, because the corporal expects immediate action and the recruit is well aware of that. In the case of friends or lovers who have literally "all the time of the world" the value of the time factor is on a much larger scale. Because both cognitive and interactive processes have much time at their disposal these processes can be much more complex performed. If we again measure this time factor on a scale from 0 to 1, 0 meaning "no time at all" and 1 meaning "as much time as desired" then the definition of the cd-parameter must be enlarged to

\[
cd = \sum_i \text{dim val}_i \cdot \frac{1}{ta}
\]

with \(ta\) – "time to act" – as the time factor. The general hypothesis, of course, remains the same.

4 A Basic Model

According to these general theoretical considerations, we constructed one first basic model of social communication. Although it is at present not possible for us to capture all theoretical aspects mentioned above in one model, even this first still rather simple model yields interesting results. The model was implemented into a computer program by two of our students, Simon Cohnitz and Christian Dinnus, who also proposed valuable improvements.

The computer system, "COMMUNICATOR", consists of about ten communicating actors. Each actor is represented by his semantic network; i.e., a graph consisting of concepts (knots) and edges; i.e., the connections between the concepts. The connections are weighted. That is, the strength of the logical or semantic relations between the concepts is represented by a real number between 0 and 1. The weights represent the associative strength between the concepts. Usually these weights are different for different actors. "Associative strength" is

\footnote{The measure of complexity of a complex dynamical system is obtained by the belonging of the rules of the systems to a certain Wolfram class (see above).}
defined as the frequency or probability by which the cognitive system of a communicator combines different concepts when "speaking" to another communicator. For example, when a communicator wants to speak about concept A and this concept has, e.g., connection weights \( cw(A,B) = 0.5, \ cw(A,C) = 0.9 \) and \( cw(A,D) = 0.1 \) then he will utter the sentence \((A,B,C)\) more frequently than the sentence \((A,D)\). Technically speaking, in each communicative act the probability for choosing \((A,B,C)\) is higher than the probability of \((A,D)\).

The social structure for this communication group is also represented as a graph with weighted connections. For two communicators, A and B, \( w(A,B) > w(B,A) \) means that A is higher in the social hierarchy than B. For normalization purposes we define \( w(A,B) = 1 - w(B,A) \). Then it is possible to define a deviance parameter, dev, as a structural characteristic of the group:

\[
\text{dev} = \frac{\sum_{i} \sum_{j} (w(i,j) - w(j,i))/n}{n}
\]

for all group members \( i \) and \( j (i \neq j) \) and \( n \) the number of group members. "Deviance" means the measurement of the difference between a group with total equality – \( w(i,j) = w(j,i) \) for all pairs \( (i,j) \) – and the factual group.

It is possible with a simple algorithm to generate a social structure according to a particular dev-value. Note that the mapping between a group structure and its dev-value is not bijective because different group structures may have the same dev-value.

The basic purpose of this model is to analyze the influence of particular social structures on the understanding processes of the communicators; i.e., on the degree of understanding they obtain from the communication. The "degree of understanding" is defined in the model as the degree of correspondence among the concepts of the different communicators. The concepts are distributed at random at the beginning of a run; each communicator gets a set of thirty concepts out of a total set of a hundred concepts. Each concept is connected with the other concepts according to a particular strength of connection, \( cw \). In the case of \( cw(A,B) = 0 \) there is no connection between the concepts A and B. The \( cw \)-values are also distributed at random for each communicator; i.e., the \( cw \)-values are mostly different. "Communication" means that one communicator A utters one "sentence", i.e., a sequence of three concepts. A chooses at random one concept and generates the "sentence" by adding the two concepts that are the most strongly associated with the first one (see above). The receiving communicator B then "answers"; that is, he also generates a sentence. If his semantic network contains all the concepts A has produced, then B will repeat these.
If B's network contains only one or two of the concepts A has produced, then B will add those concepts that are most strongly associated with the concept(s) A has used. If B has none of A's concepts, B selects an answer at random.

"Learning" in this model consists of taking over concepts from a speaker and vice versa. The new concepts are added into the respective networks with connection strengths according to the social relations between the communicators. In other words, if $w(A,B) = k > w(B,A)$, if A utters a sentence $(X,Y,Z)$, and if Z is unknown to B, then B enlarges its semantic network by adding Z with $cw(Y,Z) = k$. The reason for this rule is the well known fact that most people are more willing to learn from a person with higher authority and will generate new cognitive connections according to the social authority of the other.

Learning depends mainly on the social positions of the communicators. We assume that a person A, who is socially higher than B, will not immediately learn from someone significantly lower. Only if a highly ranking person A obtains a new concept several times from persons who are ranking more lowly in the social hierarchy then A will take over the new concept. If the more highly ranking person, A, has learned from a more lowly ranking person, B, then the w-value between A and B will be lessened in favor of B; i.e., B will obtain a position more nearly equal to that of A. In other words, if learning takes place then the social differences between two persons will change to greater equality. The reason for this is the often observed fact that people who are regarded as experts in some fields will more likely be treated as equals in factual communication even though their official position may be low.

Lower placed persons always learn from the higher placed communicators. In this case the social difference between these persons will be increased; i.e., the higher placed person gets even more authority.

Because the main question being addressed with this model is the dependency of the degree of understanding on the social structure, we show the result of three runs with different social structures – different, that is, with respect to the dev-parameter.

---

5 One may argue about the empirical validity of this rule. Do more lowly ranked persons always learn from their social superiors? In any case, a lowly ranking person must be interested to listen to a higher authority and to implement the information obtained into his memory, i.e. his semantic network, because his superior will expect it. The converse is not the case because the more highly placed person has not to fear negative consequences if he does not remember the information obtained by the more lowly ranking person.
Fig. 3: Initial values semantic correspondence $C = 0.19$, $\text{dev.} = 0.0$ (a radical democratic structure)

One sees that the effects of the learning processes change the democratic structure. After 250 runs $C = 0.41$, $\text{dev} = 0.17$; after 500 runs $C = 0.51$, $\text{dev} = 0.32$; after 750 runs $C = 0.53$, $\text{dev} = 0.44$ and after 1000 runs $C = 0.56$, $\text{dev} = 0.51$. Additional runs showed that these values do not change further – the whole system has reached an attractor state.

Fig. 4: Initial values $C = 0.19$ (the same group as in picture 1), $\text{dev} = 1.0$ (a strictly hierarchical structure). This is again an attractor state (in both dimensions).

Fig. 4: Initial values $C = 0.19$ (the same group as in picture 1), $\text{dev} = 1.0$ (a strictly hierarchical structure). This is again an attractor state (in both dimensions).
Fig. 5: Initial values: \( C = 0.2, \text{dev} = 051 \) (a mixture between an egalitarian structure and hierarchical components). Further runs showed that \( C = 0.55 \) and \( \text{dev} = 0.85 \) is an attractor state.

The figures show the changes of the initial values of deviance and correspondence dependent on the specific initial values. One sees that not only the degree of correspondence is changing, that is increasing, but also the social structure. Even the implementation of a radical democratic structure in fig. 3 generates social hierarchies.

As a general result, we found that the rules of COMMUNICATOR apparently favor hierarchical group structures, which one may see as another version of the famous principle of Matthew. In other words, the learning processes lead to social differentiation along the dimension, high-low. Even in the case of a totally hierarchical initial structure (picture 3), the dev-values are only decreased by a small amount. In most cases the C-values never transcended the threshold of 0.6 (except, of course, where we deliberately generated initial values of \( C > 0.6 \)). Total correspondence or understanding, respectively, is apparently very difficult to obtain in groups where the social structure has effects on the mutual learning process activated by communication and where the initial understanding is not high. It is up to empirical investigations of communicative group processes to see if these results are valid for factual groups (see below). Our everyday experience confirms most of these results.

The crucial factor in this model is apparently the asymmetric character of learning. While more lowly ranking persons always learn, i.e. take over new concepts, more highly ranking persons learn only when the new concepts are several times repeated. More symmetric forms of learning may obtain better results with respect to understanding. Yet in our experiments with groups of totally democratic structure the C-values also only seldom transcended the thresholds shown in figure 3. "Democratic" structure mean in our model that the learning processes are symmetric, i.e. A learns from B and vice versa.

5 Expanded Models

In several respects this basic COMMUNICATOR is not sufficient with regard to our theoretical basis and to everyday knowledge about communicative processes. On the one hand, the semantic connections between the concepts are initially set at random and are different for the communicators – each communicator has a unique semantic network that can be totally different from those of the other communicators. On the other hand, each communicator may take up new concepts during his turn in the communication if he is not able to construct sentences with the concepts or related concepts of his predecessor in the communicative process. Neither rule is realistic because a "culture", defined as the knowledge of a group [cf. Habermas 1981], also regulates the
connections between different concepts. Accordingly, semantic networks are not totally different for two members of a social group with a common culture. In addition, a communicator is not simply allowed to talk about any subject he wishes to discuss. On the contrary, if a group has chosen a theme, then the communicators have to deal with it. Changing a theme is only allowed in certain cases; in particular, if the communicator has a rather high social position or if the group as a whole has agreed to change the theme. Other deficits of the COMMUNICATOR-model will be discussed below with respect to more advanced models.

According to these two deficits we enlarged the rules of COMMUNICATOR by setting the connections between the concepts that constitute the group culture to be the same for each communicator. The communicators still differ with respect to their initial concepts. Several concepts are "clustered"; i.e., they are connected by connections with cw-values \( \neq 0 \). Such clusters are defined as a "theme"; i.e., as the set of all concepts that are directly or indirectly connected with cw-values \( \neq 0 \). In other words, a theme consists of concepts that can be combined in a "sentence" with a probability \( p > 0 \). Note that the connections of the concepts of a theme can be very different. Two simple examples may illustrate this:

![J_Kluver_Figure6]

Fig. 6: Two different structured themes, consisting of three concepts

In the first case, A is connected with C only via B. In the second case, all concepts are directly connected. Accordingly, we defined the semantic compactness \( sc \) of a theme by the ratio of the factual connections and the possible connections (in the case of \( n \) concepts the number of possible connections is, of course, \( (n^2 - n)/2 \) if one counts only a connection from A to B and not the connection from B to A). This gives us a semantic parameter, \( sem \), by defining the compactness of the whole culture as the (arithmetical) mean value of all themes of this culture:

\[
sem = \frac{\sum_{i} sc_i}{k}
\]

for \( k \) themes \( i \).
In addition, we defined for each theme a degree of rigidity, which is simply the probability by which a communicator is allowed to change a theme. The reason for this definition is the observation that different themes allow, in differing degrees, their enlargement or change. Talking about politics may lead to themes of history, prejudices, gender, sex, and so on. Talking about mathematics usually forces the communicators to stay with that theme or to persuade the others if, and how, to change it.

In the expanded model the communicators interact in the following way. The person of highest rank starts the communication by choosing a theme and generating a sentence, i.e., a triple of two to three concepts from this theme. The next communicator has to "answer"; i.e., he repeats the first concept (or another concept from the sentence if he does not have the first concept) and adds one or two new concepts from the same theme. If he is not able to answer with such a sentence, he either may choose a new theme, although only under rather restricted conditions, or remains silent and the communication passes to a third communicator.

In other words, we expanded the first model by introducing certain semiotic production rules (see above) and looked for the effects on the degree of understanding of the whole group. As a general trend we obtained that the semantic compactness of the group culture plays a decisive role insofar as high values of sem generate higher values of C. Yet because the first results obtained with the expanded model are still preliminary, and because we are already working on more sophisticated models, we give here no detailed results.

Even this expanded model is still not sufficient in several additional respects. On the one hand are the communicators just modeled by directed graphs that can be enlarged and changed during the communicative process. Yet these graphs are not such dynamical systems as "real" i.e. human communicators. To overcome this deficit another expansion of the COMMUNICATOR-model substitutes the simple directed graphs by a certain combination of different neural nets, namely hetero-associative feed-forward networks (HN) and so called self-organized Kohonen maps (KM). In other words, each artificial communicator consists of one network of both types. The HN has the task to associate the characteristics of a certain concept with the concept itself. The reason for this sub-model is that communication often consists in the naming of characteristics and the adding of the respective concept. If a child, e.g., says, "look, something small with four legs and it makes noises" and the answer is "this is a dog" then the child will associate these characteristics with the concept "dog". Conversely, if on hears the concept "dog" one associates characteristics like those the child has told. The HN in our model performs just these tasks.
The KM is able to generate a semantic net from the concepts and their respective characteristics in a self-organized manner. Therefore, the artificial communicators in our expanded models do not need to get their semantic networks by the user of the program but they are able to construct their networks themselves. It is obvious that this model allows not only the study of the emergence of group understanding and in this way the emergence of a certain culture but also the study of individual learning processes in dependency of the communicative processes in the according group. The model, which is implemented by Gregor Kaczor as part of his MA-thesis, is not yet finished. Results of experiments with this model will be given in later publications.

Another shortcoming of the basic COMUNICATOR-models is the assumption that each communicator is able and willing to communicate with every other communicator. Real groups do not behave this way (see below) because, e.g., if communicator A does not like communicator B then A will try to avoid communications with B. Other factors of "who will communicate with whom" play according roles. Another expanded version of the COMMUNICATOR, implemented by Thorsten Toellner, takes this into account by substituting the simple rules of interaction of the basic model with rules bases on a socio-matrix. This matrix contains numerical values that represent the feelings of the different communicators to each other. A little example may illustrate this:

\[
\begin{array}{ccc}
A & B & C \\
A & 0 & 1 & -1 \\
B & 0 & 0 & 1 \\
C & 1 & -1 & 0 \\
\end{array}
\]

**Fig. 7: A socio-matrix of a group of three**
A likes B (1) and dislikes C (-1), B is neutral with respect to A and likes C, and C likes A but not B. The main diagonal of such a matrix always contains only zeros because the feelings of a group member towards himself do not matter.

The changed rules of this model are now that the communicators try to communicate only with persons they like (feeling value = 1) or at least do not dislike (0). The interesting question with respect to this model is the generation of subcultures based on mutual feelings of like and dislike. The respective results will also be reported after the finishing of this model.

---

6 A Kohonen map is able to learn in an "unsupervised" manner, that means, generally speaking, that it is able to order the concepts it obtains by their own logic, i.e. by their relations. Because a "dog" is more related to "cat" than to "fish", a KM will place "dog" nearer to "cat" than to "fish".
6 Empirical Validations

Formal models of complex processes are by themselves quite interesting if they demonstrate how precise, i.e. mathematically formulated theories can be constructed. Yet the decisive proofs of the validity of such models are of course empirical tests. We undertook such tests by experimenting with communicative groups of students at the University of Duisburg-Essen and compared the empirically observed communicative behavior of the students with the results of the basic COMMUNICATOR-program that is the model with the addition of the semiotic production rules. This comparative study, which is described in more detail in the MA-thesis of one of our students, Jochen Burkart, was quite satisfactory.

The experiments were carried through with three different groups of students from different curricula – computer science, communications science and education science. Each group had the task to discuss three different themes for approximately twenty minutes. The whole communication, therefore, lasted an hour. Before the start of the discussions each student had to construct a semantic net for each of the themes in about half an hour. The concept of the theme should be placed in the center of a page of paper and then the other associated concepts should be placed in concentric circles. Quick associations should be marked by placing the concept into the circle nearest to the theme concept, later associations in the other circle. Consider the example of the theme "Paris":

![J_Kluver_Figure8]

Fig. 8: A semantic network ordered in concentric circles
At the end of the whole discussion the students had to construct new semantic networks of the discussed themes in the same way.

This ordering of the concepts was done in order to give the COMMUNICATOR model values of the semantic connections between the concepts. For example the concept "Paris" has a connection value to "Eiffel-Tower" of 0.8, and to "Chirac" of 0.5. Because we could not ask the students to define the values of the connections themselves in a short time, we choose this way of constructing semantic networks.
The initial semantic networks of the students were then inserted into the COMMUNICATOR-model and it performed approximately as many interactions of the artificial communicators as the students did. Because the communicative behaviors of the three groups were rather different the total number of interactions was counted and the program got this number. Each discussion group, by the way, was recorded by video. The question was if the groups obtained a similar degree of understanding that is correspondence, as the COMMUNICATOR-program.

The comparison of the program with the first group showed a high degree of similarity. The C-value of the program was \( C = 0.6 \), the student group obtained \( C = 0.55 \). One can say that the program prognosticated the group behavior in a quite satisfactory manner. The simulation result of the second group was far worse: The program prognosticated again \( C = 0.6 \), but the group just obtained \( C = 0.35 \). The program was obviously not able to predict the factual group behavior. The simulation of the third group, in contrast to the second, obtained again satisfactory results: \( C = 0.63 \) was the result of the program, \( C = 0.57 \) the result of the group.

A study of the videotapes of the three groups quickly explained these different effects. The communicative behavior of the first group was characterized by a nearly equal distribution of communicative participation. All students had approximately the same number of "speech acts. This corresponds to the rules of the COMMUNICATOR-program that give each communicator the same chances to participate. Therefore, we could conclude, the COMMUNICATOR is able to predict the communicative behavior of real groups if the participation of the members is approximately equally distributed. Because this was the case with the third group too, although not as equally distributed as the first group, the success of COMMUNICATOR with respect to the third group finds the same explanation.

The second group strongly differed from the other groups: Of nine participants one student did not say anything at all during the whole discussion and four other students spoke only seldom. Only four students continuously participated during the whole discussion and even in this subgroup two speakers were dominant. To be sure, we did not expect such a behavior and the failure of COMMUNICATOR with respect to this group is certainly due to this unusual behavior of the group members. We assume, by the way, that COMMUNICATOR and we were the victims of a classical "Hawthorne Effect": All students of the second group are in the phase of their final examinations with us as examinators. Probably several of the students were anxious not to say anything we could dislike. Therefore they decided to stay on the safe side and preferred to say nothing. The students of the other group
were not in that situation. By "Hawthorne Effect" we mean a well-known problem of social experiments: Sometimes the conditions of the experiment will generate the observed effects that would not have occurred under other conditions. In other words, the experiment operates as a kind of self-fulfilling prophecy.

The failure with respect to the second group is cause to think again about the empirical validity of some rules of COMMUNICATOR, in particular the assumption of equal distribution of the participation. Perhaps in analogy to the socio-matrix group members should attach themselves before the discussion numerical values that represent their willingness to participate at this discussion. According to these values then the rules of COMMUNICATORS could take into account an unequal distribution.

Yet in those cases where the assumption of equal distribution is valid, COMMUNICATOR is able to make rather sound predictions. These results show that even such a simple model is empirically valid in the sense that it captures basic features of social reality.

5 Final Remarks

It is obviously possible to implement the general theoretical considerations of section 2 into a computer model, although the model that was empirically tested still lacks important characteristics of real social processes. Therefore, besides systematic experiments with this model, one of the main tasks will be to expand the model in other ways. Some of them have been described in the previous section, although it is still too early to give systematic results of the expanded models. But despite its initial simplicity, the basic model already demonstrates some important aspects of our general theoretical definitions and we expect that first fundamental results about the "logic of communication" will emerge in the next time.

The more advanced models described in the last section obviously are able to combine the different types of dynamics whose interplay characterizes communicative processes. That is in particular the case with the model that consists of different neural nets that interact. Here learning takes place in a rather elaborate manner and communication is much more than the simple exchange of concepts. But even such sophisticated models are, of course, not enough to capture all aspects of communicative processes. Therefore, these models are a beginning: they are a demonstration that mathematical and computational theories of communication are a real possibility. The fact that communication must always be understood and modeled as a two-dimensional process, i.e., as the interplay of
two types of dynamics, just explains why even our most elaborated models, that take this into account, are still not elaborate enough.

We explained in the introduction to this article why a theory of communication is necessary as one of the general foundations of the social sciences. In particular, only with a theory of information will it be possible to integrate social and cognitive processes as the generating forces of culture. The general theoretical framework presented in this article, the results of experiments with more sophisticated models, and other empirical validation experiments will be published in a larger study on communication.

References


