An investigation into the evolution of communicative behaviors.

Ezequiel A. Di Paolo
School of Cognitive and Computing Sciences
University of Sussex, BRIGHTON BN1 9QH, U.K.
Email: `ezequiel@cogs.susx.ac.uk`

Contents

Abstract

Communication is a phenomenon with many different aspects, and it has attracted the attention of a variety of scientific disciplines, biology being perhaps the one with the best chance of providing a good theoretical backbone by addressing the unifying theme that underlies the different views on the subject. However, recent work in the evolution of communication have tended to evade rather than embrace this task. I provide a critical analysis of the reasons for this situation, which are, for the most part, methodological and conceptual, and manifest themselves in the way biologists characterize the phenomenon, as well as in the tools they use to research it. I present an alternative characterization in terms of autopoietic theory, and show that not only is it possible to work with it, but that it also addresses issues of interest to other disciplines. By choosing as my object of study a game of interactions, I intend to provide some continuity with traditional approaches and the view of communication presented here. Traditional tools, such as game theory, are not blindly discarded, but are extended in order to go beyond equilibrium studies into the nature of the evolutionary dynamics. Further extension involves the use of a computational model, so some of the methodological issues that arise by its use are discussed. Within this model, communication evolves in a society of artificial agents even in the presence of costs *against* it, and this is explained in terms of selective mechanisms acting within the constraints provided by other factors such as spatial organization. A complex network of mechanisms is explored by studying the phenomenology of emergent self-regulating unities in the spatial distribution of agents. Dialogic communication also evolves non-trivially in a similar game in which agents share all the relevant environmental information and, by coordinating their actions, they are able to perform tasks beyond their individual cognitive capabilities, showing that the concept of information has to be used with care, and providing a metaphor for the evolution of cognition as rooted in social activity. Conclusions are drawn both on the general subject of explaining complex processes with many interacting causal factors, and on the relation of these results to the evolution of natural communication.

1 Introduction.

The variety of behaviors that we tend to group under the label “communication” can range from the simplest iconic display to the most complex forms of human interaction. Not surprisingly, this subject has attracted the interest of a quite heterogenous group of disciplines, ranging from evolutionary biology and ethology, to psychology, psychotherapy, sociology, philosophy of language, epistemology and media studies, among others. This diversity suggests an immediate question about how related the corresponding subjects of these disciplines falling under the same label really are. It is apparent that the phenomenon of communication can be focused from many angles, so that the very existence of a constant underlying theme comes into question.

Perhaps the people who have done the greatest effort in the formalization of their own focus on the subject have been the theoretical biologists. In their case, the objects of study have been the inter- and intra-specific signalling systems used by animals, their functions and their evolution. However, they have had a difficult time in defining exactly what constitutes an act of communication, without borrowing terms and concepts (such as “signal”, “information”, etc.) from other disciplines and contexts, sometimes ill-defined themselves. The resulting characterization of the phenomenon was still so poor, that biologists went one step further by incorporating the notion of adaptation into their definitions, creating an interesting problem from a methodological point of view, on which I will elaborate in Section 2 of this paper.

My opinion is that, if there is an unifying theme at all relating the different focuses on the subject, biologists have shown a tendency to move away from it and, by so doing, they have diminished the chances of building a good theoretical backbone on which other disciplines could also rely, a goal that biology is in a privileged position to achieve. In this paper I propose to defend the thesis that such a central common concept can indeed be stated clearly and, in fact, *has* been stated clearly, and that it is possible to work with it.

To support such a thesis, I will try to show that the problem of the origin and evolution of communication can be tackled using a different and clearer characterization of the phenomenon, while maintaining some terminological continuity with previous approaches; also that the evolution of communication can be so explained, and that the explanation involves natural selection as well as other mechanisms that provide a context in which natural selection can work; and, finally, that such an investigation provides results and guidelines for future research that are relevant to other disciplines such as Developmental Psychology and Cognitive Science.

Part of this work will deal with the scientific tools for addressing this problem. Traditional approaches in the form of mathematical models will be extended in order to account for the dynamics of the process, and a computational model will be used in order to further the understanding of those dynamics. This latter line has been increasingly used in recent years, also on the subject of the evolution of communicative behaviors (MacLennan & Burghardt, 1994; Steels, 1997), and on the related subject of social intelligence (Dautenhahn, 1995). In addition, I will discuss some methodological issues about the construction of this sort of models and what kind of knowledge one can expect to gain from their use.

In Section 2 of this paper I present some critical comments on the way the subject has been approached in recent studies from theoretical biology. The uses of the notions of *selection* and *information*, as part of the characterization of the phenomenon of communication, are criticized, the first from a methodological and the second from a conceptual point of view and some consequences of this situation are analysed. In Section 3 an alternative characterization of *communication as coordinated activity* is presented, within which the above criticisms are avoided. I have found that the most concise way of presenting these concepts is by using the language provided by the theory of autopoiesis (Maturana & Varela, 1980, 1988), so I have included some introductory comments and pointers to the relevant literature. In spite of this, communication characterized as coordination of actions is not the sole property of autopoietic theory; many other converging views from different disciplines are also mentioned. These views suggest that the adoption of such a standpoint could provide a natural framework inside which problems like the evolution of communication could be extended into problems like the evolution of sociality, the evolution of language and the evolution of cognition, thus providing the unifying theme for the many existing focuses on the subject.

In Section 4 I discuss the "methodological map" of this work. I choose as the object of study a simple game of communication similar to those used in previous approaches, in order to provide, when possible, some continuity in terms of language and concepts. However, the same game, and the corresponding modelling that follows can be also be accommodated within the adopted conceptual framework originating in autopoietic theory. For the same reasons of continuity, I start with a classical game-theoretic approach to the problem, which I extend in order to account for the evolutionary dynamics, because these constitute the real object of investigation. This extension shows that the traditional approach is limited, because it is based on finding what conditions lead to an equilibrium situation (an Evolutionarily Stable Strategy), and, in this particular problem, such an equilibrium may be unreachable or many viable equilibria may co-exist. The possibility of a further study into the dynamics of the problem suggests that the most practical way is the computational approach, so general issues about using this technique as a proper scientific tool are also discussed.

The mathematical model that results from extending the game-theoretic approach is presented in Section 5. Here I try to address the question about how important the dynamics of an evolutionary process are and when and how they can complement a stationary study of the problem. The results provide an answer to this question, but unfortunately not much insight into the original problem of the origins and evolution of communication. I claim that this is because many interesting mechanisms are still left out or simplified in this model, so that we are left with two choices, either to extend the model in order to account for such mechanisms, incrementing its complexity and also extending the number of assumptions, or to build a bottom-up computational model of the same problem in which all the features of interest are included from the start[1]. A simple potentiality vs. cost analysis suggests the second alternative as the best one.

The computational model itself is discussed in Section 6. The obtained results are very interesting. They show that communication evolves even in cases that go against our immediate intuitions, and, most importantly, they provide us with a way of explaining why this evolution occurs. Coordinated activity is explained by the interplay of the spatial organization of the population and selection mechanisms. I present a first order analysis of the complex network of mutual influence between these and other mechanisms to explain some features of self-regulation and change in spatial structures and their influence on the degree of communicative behavior.

Also in this section I present a modified version of the communication game, in which the subject of the misuse of the concept of information is addressed again. Contrary to traditional views on the subject, we see in this game communication evolving in a non-trivial manner even when all agents have the same access to environmental information. While performing their coordinated activity, agents are able to achieve tasks that are impossible for them at the individual level, suggesting an interesting metaphor with some affinity to theories that view the evolution of cognition as rooted in the evolution of social activity.

Finally, in Section 7 I discuss the relevance of these results both in terms of the contribution to the way that scientific research can face the study of complex processes, and in terms of their relevance to the problem itself. Some parallels between emerging features of our model and some existing hypotheses about communication in primates are also mentioned along with some limitations of the model and some interesting avenues for further exploration.

---

[1] Note that this is different from still another possibility: to build a computational simulation that *extends* the mathematical model.

2 Preconceptions about communication.

The fact that in almost every new approach to the problem of the origins and evolution of communication the authors have believed it necessary to provide a new definition of the phenomenon, or at least to refine existing definitions, is sufficient proof of how slippery this concept can be. The strategy usually adopted in such a situation is to try to view the subject from the perspective of an already existing body of research. Consequently the subject suffers transformations and/or reductions in order to ensure the applicability of the research tools and methods currently available. Evidence of these transformations can be exposed in the fact that different studies on the same subject reveal the origins of the research tradition in which they are embedded by the presence of different preconceptions about that subject. By preconception I mean the unquestioned application of concepts and theoretical constructs that have proved successful in other domains to the new subject. It is important to notice that this transference of concepts is not done as in the case of a working hypothesis, where the possibility of revision exists, but rather as *a priori*s that need no further questioning and are "accepted as such" (Stolzenberg, 1984). On the other hand, it would be a mistake to think that any preconception is necessarily a *mis*conception. The latter has to do with how well our ideas fit[2] into the world, while the former refers to a problem that arises in the methodology of our scientific activity.

These preconceptions have obvious consequences for research. Not only do they shape the way in which specific questions are addressed but they also constrain the range of questions that make sense to ask. It is my purpose in this section to briefly analyse two of the most important preconceptions about the problem of the evolution of communication found in the theoretical biology literature that we have learned to accept and how they have affected our views on the subject.

The best place to start looking for these preconceptions is right at the beginning, where the subject is delineated, i.e. in definitions.

2.1 The role of selection.

There is little doubt among researchers that the origins of communication must be sought in the domain of biological phenomena. Many biologists have addressed the problem from the point of view of behavioral ecology and evolution. With very few exceptions definitions of the phenomenon coming from these studies have included preconceived ideas about the causes and effects of communicative behavior. The main preconception found in most definitions is that communication is necessarily beneficial for some of the participants (at least probabilistically), otherwise selection would not have favoured it.

Wilson defines communication as the altering by one organism of the probability pattern of behavior in another organism in a manner adaptive to either one of them or to both (Wilson, 1975). Lewis and Gower define communication as "the transmission of signals between two or more organisms where selection has favoured both the production and reception of the signal(s)" (Lewis & Gower, 1980). Krebs and Davies define it as the "process in which actors use specially designed signals or displays to modify the behaviour or a reactor". They later make it clear that they understand "specially designed signals" as those that have been favoured by natural selection (Krebs & Davies,

---

[2] For a good discussion of the meaning of *fit* in this context see (von Glasersfeld, 1984).

1993). Burghardt defines it as a behavior that is "likely to influence the receiver in a way that benefits, in a probabilistic[3] manner, the signaller or some group of which it is a member" (Burghardt, 1970; MacLennan & Burghardt, 1994). Maynard-Smith and Harper define a signal "as an action or structure that increases the fitness of an individual by altering the behaviours of other organisms detecting it, and that has characteristics that have evolved because they have that effect" (Maynard-Smith & Harper, 1995).

All these definitions (and many more coming from the biological literature) are questionable on simple methodological grounds: *they mix a characterization of the phenomenon with a possible (and, admittedly, plausible) explanation of it.* In other words, these definitions resolve *a priori* the problem of why communication has been established in animal societies, leaving no room for alternative/complementary explanations or their rebuttal, while at the same time providing a poor characterization of the phenomenon. We may think of a similar situation in which a physicist defines a "phase transition" as a "change in structure observable in certain systems or aggregations of matter that occurs when energy is continuously added to or taken away from that system". There are two things that are patently wrong with this definition. First, if we have never seen a phase transition before this definition doesn't help us much. What kind of change in structure? How is energy supplied? Secondly, it is known that phase transitions can occur in energetically isolated (adiabatic) systems when pressure changes. So, the explanation that has been built into the definition is not unique. Should we refine our definition? Should we define a different phenomenon corresponding to the alternative explanation?

Analogous questions are applicable to the definitions of communication mentioned earlier. Some biologists might respond to the first kind of consideration by saying that everyone intuitively knows what a signal is and that we all agree on that. But this is certainly not true and evidence of this can be found in the fact that almost every single approach to the problem of the evolution of communication have presented a new refinement of modification to previous definitions. Is a predator hiding in order to catch a prey emitting a signal, as would be implied by some of the above definitions (for instance, Bughardt's)? And a prey developing a mimetic characters in order to confuse a predator (as defended in Maynard-Smith & Harper, 1995)? This is clearly a problem that arises from the attachment of the idea of benefit in a description of the phenomenon that should (and can) be characterized independently of it. If we saw a group of animals committing collective suicide after a call given by one of them, we still would like to describe this as a case of communication. The fact that we rarely see such behaviors should perhaps be used to support selective *explanations*, but not to redescribe the phenomenon.

The second kind of question might be answered by many biologists by showing that there is no possible explanation for the existence of communication other than its being selected because of its benefits. So, it does not matter if the working definitions leave no room for alternative or complementary explanations. Unfortunately this is not relevant. Even if I were a die-hard adaptationist I would be forced to recognize that a description of a phenomenon and an (or *the*) explanation of it belong to different domains and, therefore, should not be mixed. We define wings as appendages that are used by some organisms for aerial locomotion. We do not define them in terms of their selective

---

[3] The words "likely" and "probabilistic", which abound in these definitions, are confusing, unless a context is specified and random processes can be found to dominate in that context.

advantages, even if we explain their presence in those terms. Moreover, if there is no other possible explanation apart from selective advantages what is the point of stressing that fact at the very definition of the phenomenon?

Biologists have appealed to this kind of definition in order to rule out obvious cases of behaviors that are not communicative and which fall within the characterization given by broader definitions such as "an interchange of signals". However, as we have seen, the problem of poor characterization remains and the proposed solution is methodologically questionable.

What are the consequences of this mixing of explanation and characterization of communicative behavior? The first obvious consequence is that alternative or complementary mechanisms are out of the question under this functionalist focus. Every single aspect of communication should be explainable in terms of their selective advantages otherwise we would not be talking about *communication*. Given its self-explaining character the evolution of communicative behavior should pose not problem at all. But, there are problems and, not surprisingly, these problems arise from the contradiction of observed phenomena and the assumption of self-benefit prevalent in explanations invoking natural selection, for instance honest vs. dishonest signalling (Zahavi, 1975, 1977; Grafen, 1990; Hurd, 1995), manipulation vs. cooperation (Krebs & Dawkins, 1984), etc. I am not saying that problems in the understanding of communication will suddenly be eliminated if we propose some mechanism other than selection, what I am saying is that the *possibility* that other mechanisms, acting perhaps in conjunction with natural selection, might transform our view of these problems is simply ignored.

In the Section 3 a definition of communication will be given that does not presuppose the presence of selective advantages and later in this paper it will be shown that some aspects of communication, within the proposed model, cannot be explained in terms of selective advantages *alone*.

2.2 Signals and information.

Another preconception about communication that is often found in the biology literature and, to a lesser extent, in psychology studies, is that it involves the transmission of information from a sender to a receiver. This is an understandably popular view in our age of mass-media, fax machines and computer networks. Information is a concept that has penetrated our views of developmental processes (Oyama, 1985) and, in general, of the mechanisms underlying the dynamics of complex systems and, for some people, information should even have the same ontological status as matter, or space. This view of communication as informational exchange implies that there is "something" that is being transmitted through some channel, although few researchers specify what it is. For instance, in (Maynard-Smith & Harper, 1995) the term "information" is used to help define a taxonomy of signals while right from the beginning they make it clear that they "mean information in the colloquial, rather than technical, sense".

An immediate criticism to this use of the term is simply that there is not such a *thing* as information. Information is not an objective quantity independent of an observer. Theories such as Shannon's (Shannon & Weaver, 1949) are not really about our everyday idea of information but about the "reliable transmission of signals over unreliable channels" (von Foerster, 1980). These theories can only account for content-free, uncertainty reducing mechanisms as the authors themselves make clear. Information, as they present it, is a probabilistic concept that has little to do with the kind of information that is often invoked when explaining communicative behavior[4] (Oyama, 1985, pages 64 - 71).

All an observer can say when witnessing an act of communication depends only on the activity of the participants and on the observer herself (particularly, on her ability to draw inferences from her observations). A given behavior identified as a signal can trigger different responses in the participant identified as the receiver and can be interpreted differently by the observer in different circumstances. There is no actual, observer-independent, ground for identifying any informational content in a signal. Information does not exists before the activity and it cannot be separated from it. It cannot be searched for before the signal is emitted, so it cannot be an objective, independent quantity that exists "out there":

> "... We believe that information can even be stored and then, later on, retrieved: witness the library, which is commonly regarded as an information storage and retrieval system. In this, however, we are mistaken. A library may store books, microfiches, documents, films, slides and catalogues, but it cannot store information. One can turn the library upside down: no information will come out. The only way to obtain information from a library is to *look* at those books, microfiches, documents, slides, etc. One might as well speak of a garage as a storage of and retrieval system for transportation. In both instances a potential vehicle (for transportation of for information) is confused with the thing it does only when someone makes it do it. *Someone* has to do it. *It* does not do anything." (von Foerster, 1980)

The idea of something being transmitted in communicative interactions is probably an extension of our idealizations of human language, in which we think that every utterance carries some informational content from emitter to receiver about a given state of affairs. And this is possible because we suppose that both participants share the same cognitive properties and therefore can interpret the message in the same way as it happens with man-made communication devices. However it is known that this is a questionable idea when applied to the way language is *actually* performed in everyday activity (see Section 3.3). So, without any objective way to grasp *the* interpretation of the message, on what grounds can we speak of the information contained in it? If anything at all, what we call information "in the colloquial sense" is not only dependent on, but is in fact a *posterior* interpretation of, the activity that generates it and, therefore, it cannot be used to explain how that activity originated.

Another criticism to the idea of communication as an exchange of information is that it implies the view signals must denote something. They must refer to some feature of the world, or stand for some state of affairs. Otherwise they are meaningless and out of the scope of the explanatory mechanism of natural selection[5]. But this is certainly not true. Reference and denotation are not necessary for communication to make a difference for those organisms that use it. Again, wings are not "about" anything, yet they are fundamental for animals that depend on them, "meaningful" if we wish. Even if we can describe a given communicative behavior as being about something, this doesn't mean that this "aboutness" is a necessary feature of communication. In fact, this "aboutness"

---

[4] A definition of *information transfer* that comes close to the intuitive sense of the word is presented in Section 3.3.

[5] As the reader may have guessed the preconceptions criticized here are not really independent.

is more a feature of the description rather than a feature of the behavior itself and, therefore, it "belongs" to the observer.

Let us examine some of the consequences of the metaphor of the exchange of information in the way that the problem of the origin of communication has been addressed. This view has led many researchers to assume that a necessary condition for communication to arise is that not all relevant aspects of the environment are equally known to the different organisms that constitute a group, (it is interesting to see how this idea has influenced the computational approach to the problem, MacLennan & Burghardt, 1994; Werner & Dyer, 1990). If all the relevant "information" is readily available to everyone, why "should" communication arise? Given that something is being transmitted, it must be worth the effort of transmitting it. Thus, communication can only be understood if there is some relevant feature of the environment whose conspicuousness has to be enhanced by a signal (such as a predator, or food), or if some internal state needs to be publicized. Unfortunately, this assumption leaves out as non-communicative many interesting phenomena involving the coordination of behaviors between organisms even when they share all there is to be aware of. Examples of these include many behaviors in which a group of interacting organisms achieve results impossible to obtain at the individual level, such as the formation of hunting patterns in wolf packs and other predators in order to kill large preys and other examples of teamwork [6]. In these cases, all the relevant "information" is readily available to all the participants. Other examples can be found in most species of social mammals in which hierarchical structures exist. In these societies communication is used to help in the generation and maintenance of the social structure. A rebel member who receives a punishment by the leader of the group, is not just being given a piece of information about "who's the boss here", it is being forced to reintegrate into the social structure or else. Signals are used as actions rather than packages of information. Do such behavioral coordinations necessarily entail the presence of communication? I will claim that they do, although this may not fit into our everyday notion of communication in which there is a definite "signal" that is "intended" to be "transmitted". But it is precisely the use of this everyday notion within a scientific approach that I am criticizing here.

A different view of communication will be presented in the following section together with some benefits for research that come from adopting it, not the least important of which is the extendibility of these ideas into other problems such as the evolution of sociality, cultural phenomena and the evolution of human language.

3 Communication as social coordinated activity

Is it possible to define communication without appealing to concepts such as selective advantages or information? More importantly, can we work with such a definition? Here I will try to answer these questions. The view of communication as a biological phenomenon that will be presented in this section is not new and similar aspects of it have been developed by different researchers in different fields. I will use here the language of autopoietic theory (Maturana & Varela, 1980), although similar ideas have been expressed, sometimes in different terms, in the fields of cybernetics (von Foerster, 1980; Pask, 1980) and certain branches of psychology and family therapy (Bateson,

---

[6] Another concrete example is given by the experiments described Section 6.4

1972; Watzlawick, Beavin, & Jackson, 1968; Watzlawick & Beavin, 1977). From a philosophical perspective many traditions converge to similar views (Austin, 1962, 1971; Dewey, 1958; Habermas, 1979; Wittgenstein, 1963; Heidegger, 1962).

3.1 Ontogeny and phylogeny in autopoietic systems

It is far beyond the scope of this section to give an introduction to autopoietic theory and the reader is referred to (Maturana & Varela, 1980, 1988) for a complete account of this important field[7]. Certain concepts will be needed for the definition of communication that I wish to present and they will be briefly explained. This is not an easy task because concepts in this theory are linked by a very compelling logic and it is difficult to present isolated ideas without doing some injustice to the theory as a whole. I hope that readers knowledgeable in this area will understand this point, while others will recur to the mentioned sources if they are looking for a deeper account.

Autopoiesis is a theory of the organization of living organisms as composite, autonomous unities. An autopoietic system is a system whose organization is maintained as a consequence of its own operation. Autopoietic systems in a given space produce their own components and boundaries and, as a result of the network of processes (of production, transformation and destruction) realized by the interactions between these components, the organization that maintains them (both the components and their relations) is specified. All living organisms are autopoietic systems that inhabit physical space. Autopoiesis is a property of the organization of the system; a given autopoietic organization is embodied in a particular *structure* or physical realization, and each state of such a system is determined *only* by that structure and a previous state. This seems almost trivial, but it is a fundamentally important point. It implies that any behavior of an organism that we can witness as observers is a direct result of the organism's own structure and of its history. Thus, autopoietic systems are a subset of the larger set of operationally closed systems[8].

Any autopoietic system exists in a medium with which it interacts and, as a result of that interaction, its trajectory in state-space (its history) changes, although its operation as a dynamic system remains closed. Put crudely, we can think of this as a system whose states are determined by a set of differential equations and as a result of interactions (coupling) with another system (the medium) some parameters in this set of equations are perturbed. The state of the system will still be determined by the equations but the trajectories will generally differ if the perturbations are different. The structure of the system, then, determines its *domain of perturbations*, that is, what are the possible trajectories that can be triggered by interactions with the medium given a certain initial state without destroying the system. If the system undergoes changes of state that result in plastic changes of structure, and therefore changes in its domain of future perturbations, and all this happens without disintegration or loss of its autopoiesis, then the system is said to undergo a process of *structural coupling* with the medium.

---

[7] A good introduction can be found in chapter 4 of (Winograd & Flores, 1986)

[8] "Closed" is used here in the mathematical sense. Organizational closure can be defined when the organization is characterized by processes such that: "the processes are related as a network, so that they recursively depend on each other in their generation and realization of the processes themselves, and 2. they constitute the system as a unity recognizable in the space (domain) in which the processes exist" (Varela, 1979, page 55). It is important not to confuse this notion with that of a system being "closed" to interactions with its medium.

In the analogy with the set of differential equations not only would some parameters be perturbed but the form of the equations themselves would suffer a change.

If the medium is also a structurally plastic system then both systems may become structurally interlocked, mutually selecting their plastic changes, and thus defining a history of plastic interactions that for the organism is its *ontogeny.* As long as autopoiesis is maintained during this history, the organism is said to be adapted to the medium:

> " Adaptation, then, is always the trivial expression of the structural coupling of a structurally plastic system to a medium. ... It follows that, in the operation of living systems as autopoietic unities in a medium, the coincidence between a given structure of the medium (place in the medium) and a given structure in the living system is always the result of the history of their mutual interactions, while both operate as independent, structurally determined systems. Furthermore, as a result of the structural coupling that takes place during such a history, history becomes embodied both in the structure of the living system and in the structure of the medium, even though both systems necessarily, as structure-determined systems, always operate in the present through locally determined processes." (Maturana, 1978, page 39)

In slightly different terms Ashby arrives at the same definition of adaptation in terms of stability and homeostasis: "... a form of behaviour is adaptive if it maintains the essential variables ... within physiological limits" (Ashby, 1960, page 58).

This view of structural coupling can not only account for changes in the individual during its lifetime, but also for phylogenic changes during evolution. *Phylogeny* is the result of the history of structural coupling of a series of autopoietic unities connected sequentially by reproduction during which adaptation is conserved. Selection acts negatively when, as a result of interactions with the medium, autopoiesis is lost, but it also acts through the process of structural coupling between medium and the organisms. It is important to remark that this concept of adaptation implies that all living systems are and have been equally adapted to their environments. They have succeeded in conserving their adaptation without disintegrating or going extinct, therefore, there are no grounds to compare the adaptation of two different organisms either living at the same time or at different points in evolutionary history. This marks a difference with the traditional neo-Darwinian view of evolution in which the existence of multiple mechanisms is not denied but their relevance is downplayed against the role of natural selection and, therefore, the structure of organisms is accounted for as a result of a process that optimizes fitness. Researchers within this tradition do not hold the naïve view that organisms optimize as the "rational man" of classic economic models does, but they use optimization as an explanatory tool (see Maynard-Smith, 1978). However, "no global optimal fitness scheme apparently suffices to explain evolutionary processes. There are, to be sure, local genetic agents for, say, oxygen consumption or feather growth, which can be measured on some comparative scale where optimality may be sought, but no single scale will do the job for all processes", (Varela, Thompson, & Rosch, 1991, page 194). In the words of the same authors, the autopoietic view of adaptation requires a switch from a "prescriptive logic to a proscriptive one, that is, from the idea that what is not allowed is forbidden to the idea that what is not forbidden is allowed" (Varela et al., 1991, page 195), and, furthermore, it requires the departure from the idea that adaptation can be measured by observer-dependent scales and that evolution proceeds in accordance with those measures.

Another important issue that is explicitly accounted for in this view of evolution is the mutual specification of changes of structure both in the organism and in its environment, particularly when this environment is partly constituted by other organisms, as in the case of communication that concerns us here. This has been recognised, at least implicitly, by researches from the neo-Darwinian tradition in their use of game-theoretic approaches to account for frequency-dependent effects on fitness. However the previous criticism still applies[9]. These issues has also been addressed independently (see Lewontin, 1982, 1983, 1984).

3.2 Communication in structure-determined, closed systems.

An organism undergoing a process of structural coupling with the medium may act recursively over its own states if the plastic deformations of the medium have been triggered by the organism's previous actions and at the same time this deformations will provoke future perturbations in the organism. In the particular case in which the medium includes another autopoietic system their individual ontogenies become coupled. Maturana and Varela speak of a *network of co-ontogenies* (Maturana & Varela, 1988). A domain of interlocked triggering of states between the organisms participating in the network of co-ontogenies is established as long as the coupling subsists. This is called a *consensual domain.*

Behaviors in a consensual domain are mutually orienting behaviors. They are both arbitrary and contextual:

> "The behaviors are arbitrary because they can have any form as long as they operate as triggering perturbations in the interactions; they are contextual because their participation in the interlocked interactions of the domain is defined only with respect to the interactions that constitute the domain." (Maturana, 1978, page 47)

An observer can describe behaviors in a consensual domain as a case of coordinated activity. *Communication* is then defined as *the behavioral coordination that we can observe as a result of the interactions that occur in a consensual domain* (Maturana & Varela, 1988). Behaviors within a consensual domain have also been referred to as *linguistic behaviors*[10] (Maturana, 1978; Maturana & Varela, 1980).

It is important to notice that all behaviors that arise from coordination of actions in a consensual domain are included in this definition of communication and not only those that can be described in semantic terms by an observer. Therefore, behaviors such as grooming, playing, the building of social relationships and the formation of hunting patterns *are* communicative behaviors.

Through the history of structural coupling with the medium, a correspondence can be identified by an observer between features, or situations in the medium and the behaviors which are coordinated and oriented in their presence as a result of communication. In these cases the observer may speak of certain actions as being signals that denote or describe a certain state of affairs. However it is a mistake to consider communication to be just a set of descriptive interactions used by the organisms to pass on information

---

[9] As a symptom of this, we may consider the complications with these models that arise if the effects of phenotypic fitness, besides those of frequency-dependent fitness are included, (see Repka & Gross, 1995)

[10] Later on, the "linguistic" adjective has been reserved for those higher order recursive communicative behaviors in which the actions being coordinated are in themselves communicative actions.

in order to handle a particular situation. This is to ignore the fact that all interactions arise as a consequence of structural coupling between unities and that the relevant behaviors arise in each unity as a result of its own structure and are not prescribed by the perturbations that the unity suffers. Besides, denotations and descriptions cannot be primitive operations in the consensual domain. They require the existence of previous agreement. Therefore they cannot, by themselves, give rise to what we have called communication:

> "Language must arise as a result of something else that does not require denotation for its establishment, but that gives rise to language with all its implications as a trivial necessary result. This fundamental process is ontogenic structural coupling, which results in the establishment of a consensual domain.
>
> Within a consensual domain the various components of a consensual interaction do not operate as denotants; at most, an observer could say that they connote the states of the participants as they trigger each other in interlocked sequences of changes of state. Denotation arises only in a metadomain as an a posteriori commentary made by the observer about the consequences of operation of the interacting systems" (Maturana, 1978, page 50)

From this point on this view can be extended in order to address many features of natural language including the behaviors of humans as observers, which, naturally, occur *in* language, but I will stop here, as the concepts of main relevance to the present work have already been presented.

3.3 Consequences and similar views

The most obvious consequence of adopting this point of view about communication is a broadening of the scope of phenomena that can be considered communicative. As the idea of a composite autopoietic unity that interacts with its environment through a process of structural coupling does not allow an account of its behavior in terms of inputs and outputs, at first sight, this account of communication will tend to be more complex than accounts that reduce the phenomenon to an informational exchange. What is gained, then, by adopting the former instead of the latter?

Independently of what one may think about theories of biological organization (of which the number is very small, partly because of the difficulty of the subject and not because of its importance), there are some characteristics of the autopoietic view of communication that, in my view, can only improve the current state of research. First of all, the scheme is unifying both "horizontally" and "vertically". I have mentioned that behaviors that cannot be accounted for in terms of information, such as group-behavior, playing, and many other social behaviors are included within this view along with more "traditional" communicative behaviors. Also, the same view can characterize communication at different levels of complexity. I will argue in the rest of this section how similar views have been used for studying the psychology of human interactions as well as to support certain philosophical views about human language.

Secondly, the focus on behavior, history and structure (both of the organism and the medium) places communicative behaviors in its adequate context, permitting the formulation of questions about the relationships that these entities hold and how they affect the phenomenon itself. More traditional approaches tended to consider only simpler mechanisms ("what is the cost of this behavior? How much is gained by it?"), but the reality is that, while these mechanisms still need to be considered, their functioning is embedded in a set of constraints that most of the time is ignored. In the autopoietic view, ignoring these relationships is much more difficult, as they play a fundamental role in the definition of the phenomenon, so that one is compelled to address them.

The other advantage of the autopoietic point of view is that it explicitly addresses the relevance of the observer, by pointing at things that arise only in the domain of descriptions in contrast to features that are inherent to the process. This introduces the need to be careful and self-aware about the use of certain terms we take for granted.

Although the language of autopoietic theory was needed to make a concise presentation of this idea of communication as a process of interactions and recursive orientations of behaviors, this same view has also been expressed in a variety of forms in different fields. For instance, viewing the issue from a cybernetic perspective, it is possible to identify the most relevant exponents of this position: Pask's Conversation Theory (Pask, 1980) and von Foerster's eigen-behaviors (von Foerster, 1977, 1980) applied to communication.

Von Foerster builds his concept of eigen-behaviors based on the simple idea that perceptions and actions in an organism in possession of a nervous systems are circularly defined: actions determine future perceptions and vice versa[11]. This fact was already mentioned as one possible consequence of structural coupling in which an organism becomes recursively coupled with its own states. Put very simply in von Foerster's notation, if s_i stands for the activity in the sensorium and m_i for the motor activity at time i, this idea can be illustrated in the following way:

$$s_0 = M(m_0)$$

$$m_1 = S(s_0)$$

$$m_1 = S(M(m_0)) \doteq Op(m_0)$$

where S and M can be thought of as mappings from sensory to motor activity and vice versa. This implies the recursive application of the operator Op indefinitely until we arrive at:

$$m_\infty = Op(m_\infty)$$

where m_∞ is called an eigen-behavior of the organism which is the stable result of the infinite mapping of Op onto itself.

In communication we have two organisms mutually triggering each other's behaviors so that:

$$m_{\infty 1} = Op_2(m_{\infty 2})$$

$$m_{\infty 2} = Op_1(m_{\infty 1})$$

---

[11] This is also an extension of the Piagetian concept of equilibration (see also von Foerster, 1977; Varela, 1979).

and then;

$$m_{\infty 1} = Op_2(Op_1(m_{\infty 1})) = Op_{21}(m_{\infty 1})$$

$$m_{\infty 2} = Op_1(Op_2(m_{\infty 2})) = Op_{12}(m_{\infty 2})$$

which is the bi-stable result of a communicative interaction (von Foerster, 1980)[12]. Note that the activity of one organism depends both on the combined operators of each individual. This process is similar to the given description of structural coupling, the difference is that the operators remain constant in this illustration, while in structural coupling the operators would be modified as a result of perturbations from the medium and from the interaction (giving room for a more dynamic interpretation of the same phenomenon, in which behaviors do not necessarily have to converge and history becomes involved and embodied in the operators, i.e. in the structures). Von Foerster also sees communication as a process of mutual triggering of behaviors in which stable patterns emerge that appear to us "as communicabilia (signs, symbols, words, etc.)".

Conversation theory (see for example Pask, 1980) is also concerned with describing the communicative interactions between systems whose structures necessarily differ. Its scope of application is very wide and has been used to study the interactions both between humans and between humans and machines, and also in problems of learning complex subjects (Pask, 1976) and machine assisted instruction. One of the main focuses of this theory is the idea that agreement can be reached over descriptions which do not have to have an objective grounding in external reality, but rather, they emerge from the process of conversational interactions. According to this, Pask uses the Petri-Holt concept of information which differs from probabilistic approaches (Holt, 1972), and, as far as I know, is the best attempt to define our intuitive notion of it. One of the definitions of *information transfer* that he gives is the following: the "*appearance* of dependence between several, otherwise independent systems" within a certain given temporality or time frame. Note that the role of the observer is implied in the word "appearance" (emphasis in the original). This concept differs from the way the term is generally used, mainly in that one refers to the activity of interacting systems and the other to a thing of feature of the world With this definition information transfer and the notion of communication as presented in this section can be reconciled. However I will proceed with care in the use of this term.

When addressing the issue of human communication and human language, many different philosophers coincide in an instrumental and pragmatic view similar to the one that can be derived from the extension of the notion of communication as coordinated activity into the domain of human interaction. These views differ from the predominant position that considers language as an object, or system of representations. For example, Dewey, expresses very strongly that language is a form of concerted action:

> The heart of language is not "expression" of something antecedent, much less expression of antecedent thought. It is communication; the establishment of cooperation in

---

[12] In a more concise and slightly stricter but less illuminating notation: $(m_{\infty 1}, m_{\infty 2}) = Op(m_{\infty 1}, m_{\infty 2})$, where Op now stand for the combined interplay of the structures of the participants

> an activity in which there are partners, and in which the activity of each is modified and regulated by partnership. To fail to understand is to fail to come into agreement in action; to misunderstand is to set up action at cross purposes." (Dewey, 1958, page 179)

Later on, Dewey compares language with tools ("the tools of tools"), and makes the point that language is not only a means but an end in itself ("it is consummatory as well as instrumental"). This reminds us of Vygotsky's "semiotic tools of mediation" (Vygotsky, 1988) and Wittgenstein's "tool box" (Wittgenstein, 1963, p. 6e) which has also been compared to the concept of Bahktin's "social language" (see Wertsch, 1991). It also reminds us of Austin's performative utterances (Austin, 1962), those sentences whose truth value we cannot really decide and, furthermore, it is irrelevant, but whose effect is to perform an act such as declaring, promising, insulting, etc. It is important to notice that Austin first drew this distinction to separate these sentences from those that he called constative, or "statements", but later on he realized that this separation was not possible neither by grammatical, logical nor semantical means, after all to "state" a fact is still an act. Actual language, as opposed to idealizations of it, must be understood mainly as a set of illocutionary forces.

Also, along these lines, Rorty compares the pragmatic view of language of both the later Wittgenstein and the early Heidegger (the Heidegger of *Being and Time*, first Division):

> "...The early Heidegger saw as clearly as the later Wittgenstein that the present-at-hand was only available in the context of pre-existing relations with the ready-to-hand, that social practice was the presupposition of the demand for exactness and for answers that could be given once and for all. Both saw that the only way in which the present-at-hand could explain the ready-to-hand was in the familiar unphilosophical way in which evolutionary biology, sociology, and history combine to give a causal explanation of the actuality of one particular social practice rather than another. Early Heidegger and late Wittgenstein set aside the assumption (common to their respective predecessors, Husserl and Frege) that social practice - and in particular the use of language - can receive a noncausal, specifically philosophical explanation in terms of conditions of possibility." (Rorty, 1991)

All these viewpoints coincide in interpreting language as a case of coordinated activity generated in everyday use, of mutually tuned interactions, or "dance" (von Glasersfeld, 1988) prior to anything else we may consider it to be "in terms of conditions of possibility". This is good news for our approach to communication, for it is the continuity that communication as coordinated activity has into the realm of human language that makes it such an attractive position from which to study the evolution of communicative behavior. Traditional views place non-human communication and human language into two very different, almost non-intersecting, domains, creating with this action an unsurpassable divide that has made the evolution of language a taboo subject until very recently[13] and seemingly immune to the naturalizing process. By contrast, both human and non-human communication can be understood as coordinated activity. The differences between the two remain evident and this is not ignored by proponents of this view (see Maturana, 1978; Maturana & Varela, 1980); one could interpret these

---

[13] This question and its relation to the evolution of human cognition has been avoided by scientific research and even "officially" banned since 1866 by the Linguistic Society of Paris (Bickerton, 1995).

differences to be, in fact, their main concern. Therefore, it is possible to claim that the position presented in this section sets a good basic background for the understanding of communication in its widest scope. Unfortunately, as it has already been said, to pursue this subject further is to step beyond the limits of the present work.

4 Methodology of this work

In this work a computational model will be used to study some questions about the evolution of communication. This model will be aimed at reproducing some idealization of how this evolutionary process might have occurred. At this point it is important to consider what kind of knowledge can be gained for this computational approach. Even though problems in physics, biology and cognitive Science have been approached in this way for some time now, their status as pieces of "genuine" scientific research is still not as highly regarded by the scientific community as other computational approaches such as Monte Carlo modelling.

The root of this problem lies in the fact that it is not quite clear what sort of methodology is being used in this kind of work. If some time is dedicated to ask this question repeatedly for each individual presentation in any Artificial Life conference in the last few years, we will discover that in a significant number of cases the answer will not be defined, and in many of those cases in which there is an answer, the variation will be considerable.

This problem has already been identified and some guidelines to correct it offered (Miller, 1995), and some of those guidelines criticized (Di Paolo, 1996b). However, the problem remains. It is not my intention to address this issue for the general case here, but I will try to get at a clearer definition of the problem in order to try to solve it, at least, for the present piece of work. Let me start by dividing the problem into three parts.

1. The first consideration is one of usefulness. How necessary is to build a computational simulation of a particular process in order to understand it? Why can't we just use our traditional tools? If we can't, have we tried to modify them? Can we (should we) redefine our problem? These are fair questions and ideally we should try to answer them (in each individual case) before embarking on any research.

2. Once we are convinced that a computational approach is necessary, the question of praxis should be addressed. Will this approach complement more traditional ones and build upon their results? Or will it be used radically in order to escape from problems implicitly created by those traditional approaches? How will answers be sought? How should simulations be built?

3. The third consideration, the one I consider that poses the hardest questions, is the problem of explanation. Once we have our simulation results, how are we to make any clear conclusions out of them? How can we identify the answers to our original problem from the implementation dependent effects of our simulation? And most importantly, how can our results be communicated convincingly to the scientific community in order to be reused by it in further research?

It is possible to claim an analogy between these questions and the most typical and simplified cycle of scientific research: identifying and delimiting a research problem, working out answers to many aspects of that problem, and producing a coherent account of those answers and their implications to be shared with the rest of the scientific community only to start the cycle again. As is the case with this particular cycle, the three groups of questions above are highly interdependent and it is possible to imagine many different causal arrows between them. The decomposition, then, serves the objective of clarifying the problem as a first step towards a solution, and does not pretend to be dogmatic in any way. Now it is possible to apply all these considerations to the present work.

The first step is to try to define as clearly as possible the problem that will be studied. In the case of the evolution of communication there are many interesting issues that could be addressed, but it is my opinion (which this work will try to justify) that the main problem is not so well understood and that it is hard to distinguish valid issues from problems arising as a result of partial explanations, problems that would disappear if these explanations were modified or replaced.

Therefore this work will be aimed at (some aspects of) the following question: *How does communication evolve?* Of particular concern will be the identification of mechanisms that allow/constraint the evolution of communication and how do they work. In Section 2 some prior questions, arising from methodological concerns, were raised about the traditional view of communication, one of them specifically addressed the question of the role of selection as the only explanatory mechanism in this process. Is it possible for structuring mechanisms such as spatial organization, to constrain or complement, through an interplay with the effects of selection, the role of the latter? This will be one of the main questions in this work.

This implies that, in the present approach, dynamical considerations will be particularly important because the focus will be on a process that is dynamical by definition. As was mentioned before, many approaches to the problem of how communicative systems evolved have focused on explaining global conditions as a consequence of which certain communicative behaviors represent stable solutions in the equilibrium state. A typical exponent of this kind is the application of game theory in order to formalize some verbal arguments about a particular feature of communication (Grafen, 1990; Hurd, 1995). This tool gives us a good idea of how animals would tend to behave if they could optimize their behaviors taking into account frequency-dependent effects. A key concept in these studies is that of an Evolutionarily Stable Strategy or ESS (Maynard-Smith, 1982). This is defined as the strategy that, if adopted by the majority of the population, will not be invaded by any other strategy, and the typical way to find it is by functional maximization at least within a certain region of the strategy space. By definition, this is an equilibrium concept, so it has to extended in order to apply it to the dynamical questions of this work. Dynamical considerations in game-theory are not new and some examples can be found in (Maynard-Smith, 1982). However, as seen in Appendix D of that book, these models explicitly include the (local) optimization assumption. In accordance to what has been discussed in Sections 2 and 3, I will present a dynamical system without this explicit assumption and in which the environment is also included. This will constitute a "satisfying" framework[14], necessarily modelled at a lower level. If

---

[14] This means that selection will only act negatively, eliminating those agents unable to adapt to their environments. It will be easy to see, that under other simplifying assumptions the dynamics will lead the system to perform local optimization, basically as a result of unequal differential rates of reproduction combined with global accessibility to resources. However, these assumptions will be removed in the

the findings of this modification in the mathematical tools of game theory do not differ much from the original approach, then the latter should be used because the dynamics do not add much to the understanding of the problem. If, however, there are significant differences, then dynamical approach will be justified.

The computational approach, however, is still unjustified so far. This will be done at the end of Section 5 where the results of the mathematical modelling will be available. For now it will be broadly justified by stating that the mathematical modelling is itself very limited particularly in those aspects that may prove more relevant.

As a consequence of the definition of communication presented in Section 3, the computational model will be used in order to look at the mechanisms that affect the emergence of coordinated activity in a society of agents that evolves from an initial state of uncorrelated interactions. This will be done by simulating evolution under a scheme very similar to the one used in the mathematical model minus some simplifying assumptions. Computational simulations will be used to study spatio-temporal patterns and structures emerging in the geographical distribution of individuals, as well as correlations between signalling behavior and other activity.

Explanations provided by this approach will not have the character of general laws. They will rather address the importance of effects usually unaccounted for, such as structure, contingency and stabilizing mechanisms, all within the constraints provided by the satisfying scheme. Also, correlated activity will be shown to emerge even between agents sharing all the relevant information, proving that the metaphor of information exchange is inaccurate, and suggesting an interesting avenue of further research.

In other aspects the model will remain as simple as possible. A feature that will not be considered in the present work is the adequate modelling of the agent's structure. The focus will be primarily in global structuring patterns that emerge from, and constrain, the evolutionary process. No ontogenic model will be included, so no claim will be made about the very important effects of ontogenic structural change during the coordination of actions with the medium. Agents will be seen as simple unities for most of the time. This is a strong simplification for a model which is partly based on concepts derived from autopoietic theory, however, I maintain that the framework provided by this theory is still very useful for the design and understanding of the present study.

Explanations will be built upon emerging properties of the model and, hopefully, we will be able to point to analogous properties in real ecological systems, so enabling us to formulate clear hypothesis about the role of those properties in such systems.

Other methodological considerations will be presented when the models (mathematical and computational) are described in detail.

5 A game of action coordination: mathematical model

In order to study the origins of communication as orienting activity I propose the following "game" to be played by pairs of agents living in a shared environment[15]. We will consider an *agent* as an unity that is able to act in the environment and, as a consequence of its actions, it can receive certain payoff in a given currency that we may call *energy* and also spends a certain amount of its own accumulated currency. For

---

computational model with important differences in the results.

[15] This game is similar to the Basic Action-Response Game discussed in (Hurd, 1995).

most parts of this work agents will be seen as simple rather than composite unities, so that the focus will be more on global patterns of behavior rather than on the structural features of individual unities, for reasons that were explained in the previous section. When a certain level of energy is reached the agent is able to reproduce (offsprings may differ from parents as a result of the introduction of random mutations), and likewise, when this level falls below a certain minimum the agent dies. Energy can be accessed by the agents if they perform a correct action (or a series of correct actions) on an energy container or *food source*, of which there can be various types, each one of them requiring different actions in order to extract part or all of the energy. The total environmental energy contained in these sources is constantly renewed at a fixed rate.

There are two "components" to each agent's actions: the *effective component*, upon which the allocation of payoff is decided, and the *external manifestation* of the action, which is not directly relevant to the allocation of payoffs. This means that for an agent to get a certain payoff the effective component of its action must match the actions required by the particular food source it is dealing with, while the external manifestation is irrelevant. Behaviors that are required to get a certain amount of food in natural organisms, such as reaching for some fruit, shaking the branch of a tree or digging the ground, can be thought of as the effective component, and the appearance of the movements implied in that behavior to another organism as an example of one possible external manifestation of that behavior. Others may be sounds, gestures, etc. While in real cases it may be hard to decouple these two components in a single action, for simplicity's sake, we will suppose that, in principle, any effective component can be found with any external manifestation.

If they have a piece of food in front of them agents may be able to perceive what type of food it is. They are also able to perceive the external manifestation of the actions of other agents. At each time step agents are selected to play the following game, which is the simplest version of the games that will be used in the present work.

1. The selected agent, who will play the *first role* (A_1), selects at random another different agent in its vicinity, who will play the *second role* (A_2).
2. A food source is randomly selected from the first agent's vicinity.
3. A_1 perceives the type of the food source.
4. A_1 acts.
5. A_2 perceives the external manifestation of A_1's action, but not the type of food.
6. A_2 acts.
7. The payoff is distributed. If both agents performed the correct action the total amount of energy is equally distributed in halves. If only one of them performed the right action, that agent receives a proportion c of the total energy ($0.5 \leq c \leq 1$), the other receives no payoff and the rest of the energy remains in the food source.

The game is played indefinitely or until the population becomes extinct. All agents have the same chance of being picked as A_1. The possibility exists that effective components and external manifestations of actions (at first arbitrarily coupled) may become correlated in such a way that agents playing the second role may use them as a prompt to act correctly over the piece of food that they are not able to see. However this may be against the immediate interests of the first agent who may receive a lesser payoff. The possibility of signals evolving out of non-signalling behaviors has been postulated by approaches from theoretical biology (see for instance Krebs & Dawkins, 1984) and can also be accommodated as an observer's description of a historical process of structural coupling. So, for convenience, I will speak of "signals" and "signalling" whenever I refer to the external manifestation of actions in the following paragraphs without attempting to make this a strict definition.

An interesting feature of this game is the temporality that is introduced in the assignment of roles. This marks a difference between this game and others, such as the Prisoner's Dilemma, in which agents perform their actions simultaneously. Also, as we can see, this game includes the feature of hidden information, as the agent playing the second role is not able to see the food type that it is dealing with. I have criticized the assumption that this feature is necessary for communication to evolve and I will present in Section 6.4 another version of this game in which this is actually proved to be unnecessary.

Some features of the evolving signalling behavior in this game can be analysed with a simple mathematical model provided that all interactions are assumed to occur globally (vicinity = whole environment). This means that geographical effects are not relevant. Also, it will be assumed that the effective component of the agent's action is always correct if it is able to perceive the food type. In order to study the dynamics of this system I will present a dynamical model built upon purely game-theoretic considerations with some modifications and simple rules of time evolution. It has been proved (Zeeman, 1980) that if there is an ESS this will be manifested as a fixed-point stable attractor in the system's state-space and global convergence is assured if there are no other attractors. However, there may be attractors that are not ESS's.

An important difference in this model, as already mentioned in Section 4, is that optimization will not be explicitly assumed, as in other cases of dynamical considerations in game-theory (Maynard-Smith, 1982, page 183, eq. D.1). As long as individuals have enough energy they will survive regardless of the condition of other individuals. Only as an after-effect of the dynamics of the system, through a coupling with environmental variables, will different "pressures" be made tangible for the different strategies that the agents adopt. It is easy to see that a satisfying scheme, with global interactions and global accessibility to environmental resources is, in fact, equivalent to an locally optimizing scheme. So, all this may seem as an unnecessary complication, but the fact is that it will make a difference in the computational model in which some assumptions will be relaxed.

Let P_i be the size of a population of identical players of the game sharing an environment with other such populations ($i = 1, \ldots, n$). For simplicity we may call P_i a population of "communicators" when any two agents of this population succeed in obtaining their share of energy from a piece of food regardless of their roles in the game; in short, if they use signals reliably (again this is not intended to replace the definition of communication given before). Let $N = \sum_i P_i$ be the total number of agents.

Let F be the instantaneous amount of energy present in the environment which is being renewed with a constant rate R and it is distributed randomly over F_Q food sources of different types, and $\bar{F} = F/F_Q$ is the average energy contained in one piece of food, (F_Q is not relevant to the model, but it is used because it facilitates comparisons with the computational implementation).

The amount of energy that will be consumed by the individuals in populations i will depend on the frequencies of encounter with agents of the same and other populations. If we suppose that any agent of any population can play the first or second role in the game with equal probability (and this occurs because interactions are global), then these frequencies p_{ij} can be calculated:

$$p_{ij} = \begin{cases} \frac{P_i P_j}{N(N-1)} & \text{if } i \neq j \\ \\ \frac{P_i(P_i-1)}{N(N-1)} & \text{otherwise} \end{cases}$$

Even when trivially $p_{ij} = p_{ji}$ it is important to remember that an interaction $i \rightarrow j$ is not the same as an interaction $j \rightarrow i$.

Let $L_i = L_i(p_{ij}, F_Q)$ be the proportion of the available environmental energy that is gained by individuals of the population i per unit of time. The general case is:

$$L_i = \frac{N_c}{F_Q}(\sum_j c^i_{ij} p_{ij} + \sum_j c^i_{ji} p_{ji})$$

where c^k_{ij} is the payoff obtained by an agent of type k in a game $i \rightarrow j$. N_c is the number of contests per unit of time, in our model this number is instantaneously equal to N so that there are as many contests as the total number of agents present at any given time.

If D is the energy consumption per unit of time for an individual agent, and B is the number of individuals that are born per unit of net energy gained by the population, the quantity $B(FL_i - DP_i)$, if positive, expresses how many individuals are born into the population i per unit of time. If we suppose that individuals do not accumulate energy, and that all the energy gained is used to give birth to new individuals, then the same quantity, when negative, will express the number of individuals that will die due to loss of energy at the population level. The following distinction is made:

$$Q_i = \begin{cases} B(FL_i - DP_i), & \text{if } B(FL_i - DP_i), \geq 0 \\ \\ 0 & \text{otherwise} \end{cases}$$

$$M_i = \begin{cases} B(FL_i - DP_i), & \text{if } B(FL_i - DP_i), < 0 \\ \\ 0 & \text{otherwise} \end{cases}$$

The following equations apply:

$$\frac{dF}{dt} = R - F \sum_i L_i$$

$$\frac{dP_i}{dt} = Q_i(1 - \mu_i) + M_i + \sum_{j \neq i} \mu_{ji} Q_j$$

where μ_{ij} is the rate of mutation of newborn individuals of kind i to a different kind j and $\mu_i = \sum_j \mu_{ij}$. B, D and μ_{ij} are assumed to be constant for all populations and over time.

I will consider an environment in which there are only two relevant actions ("A" and "B") that the agents can perform in order to extract energy from a piece of food. Agents playing the first role in the game will emit one of two signals given the action required for a particular piece of food. Accordingly, they may emit α or β to "signal" "A" or "B", or they may emit δ or γ for the same purpose. For reasons of simplicity we will not consider mixed strategies. The agents playing the second role will act in accordance with the signal received and their own structure, some of them performing actions "A" or "B" when detecting α or β and some of them acting correspondingly when detecting δ or γ. Tables 1 and 2 show the behavior of the four possible types of agents in this scenario.

Action	Type 1	Type 2	Type 3	Type 4
A	α	δ	δ	α
B	β	γ	γ	β

Table 1: External manifestations associated with effective components A and B for the actions performed by each type of agent when playing the first role.

Sign	Type 1	Type 2	Type 3	Type 4
α	A	A	-	-
β	B	B	-	-
δ	-	-	A	A
γ	-	-	B	B

Table 2: Effective components (A, B) of the actions performed by each type of agent when playing the second role, depending on the perceived external manifestation (Sign) from the first player; "-" means a non-relevant action.

Agents of types 1 and 3 may be called "communicators". When confronted with individuals of the same type these agents will emit signals that will cause the second player to act correctly, consequently sharing the energy contained in the food source. Agents of types 2 and 4 behave deceptively towards agents of types 1 and 3 respectively. When they play the second role with agents of these types they take advantage of their signals and act correctly, but when they play the first role they emit different signals and agents of types 1 and 3 will therefore act incorrectly, so agents of types 2 and 4 will get a share of the food energy that will be equal to or greater than 50%. We may call these types "anti-communicators" because not only do they not communicate with the same type, but they also deceive a certain communicator type.

But anti-communicators are deceived too. When confronted with the communicator type that they do not specifically deceive (type 3 for type 2 and type 1 for type 4) their signals will be interpreted correctly and they will not get the extra payoff when they play the first role. This creates an interesting circle of influences in the interplay of the four species, as each one of them deceives another one, and is deceived by a different one. However, this feature is not intentional, but it is a consequence of using the simplest unbiased choice of possible behaviors. With this information we may calculate the c_{ij}^k for this game, (Table 3):

$i \rightarrow j$	c^1	c^2	c^3	c^4
$1 \rightarrow 1$	1	-	-	-
$1 \rightarrow 2$	1/2	1/2	-	-
$1 \rightarrow 3$	c	-	0	-
$1 \rightarrow 4$	c	-	-	0
$2 \rightarrow 1$	0	c	-	-
$2 \rightarrow 2$	-	c	-	-
$2 \rightarrow 3$	-	1/2	1/2	-
$2 \rightarrow 4$	-	1/2	-	1/2
$3 \rightarrow 1$	0	-	c	-
$3 \rightarrow 2$	-	0	c	-
$3 \rightarrow 3$	-	-	1	-
$3 \rightarrow 4$	-	-	1/2	1/2
$4 \rightarrow 1$	1/2	-	-	1/2
$4 \rightarrow 2$	-	1/2	-	1/2
$4 \rightarrow 3$	-	-	0	c
$4 \rightarrow 4$	-	-	-	c

Table 3: Payoff matrix (c_{ij}^k) for the agent types 1, 2, 3 and 4. All numbers indicate the proportion of the average energy per food piece, c is the payoff for not communicating ($0.5 \leq c \leq 1$).

We must include one further constraint in this model: in order to model the μ_{ij}'s correctly it is necessary to characterize the species with a simple genetic model. The simplest is a 2-gene haploid model in which the first position indicates the signals that the agent emits and the second how signals are interpreted, with only two possible alleles. It is easy to see that, under these circumstances, μ_{ij} will be significantly greater between "neighboring" types than between "non-neighboring" types, (type 1 and type 3 are non-neighboring types, both alleles would have to be changed to go from one to the other). Otherwise, the μ_{ij}'s will be the same for each gene and for analogous mutations.

It is possible to prove that the system has a stable fixed point attractor in:

$$P_{1a} = P_{3a} = c,$$

$$P_{2a} = P_{4a} = N_a/2 - c,$$

$$F_a = \frac{2RF_Q(N_a - 1)}{N_a(N_a(1-c) + 2c)},$$

where $N_a = R/D$. In this scenario, the whole population is divided between agents of type 2 and 4 ($c \cong 0$). This is equivalent to say that a mixed strategy in which agents behave half the time as anti-communicators of type 2 and half the time as anti-communicators of type 4 is a stable strategy.

The following figures present some results from the numerical resolution of this model in order to see if anything can be learnt from its dynamics. Unless otherwise stated, the values of the parameters are: $R = 500, D = 0.25$ and μ_{ij} is 0 for non-neighboring types and 0.001 for neighboring types, (these figures were chosen to resemble analogous parameters in the computational implementation).

The evolution of types 1 and 2 is shown in Figure 1 for $c = 0.6$, $B = 0.01$ and for the initial condition $F = 2500, P_1 = P_3 = 1000, P_2 = P_4 = 0$. Here we see how $P_1 \rightarrow c \cong 0$ and $P_2 \rightarrow N_a/2 - c \cong 1000$. The behaviors of P_3 and P_4 is analogous. A population consisting uniquely of communicators is invaded by anti-communicators. Changes in parameter B only affect the time constant which is linear in $1/B$[16]. Changes in c affect both the asymptotic value and the time constant which varies with c non-linearly.

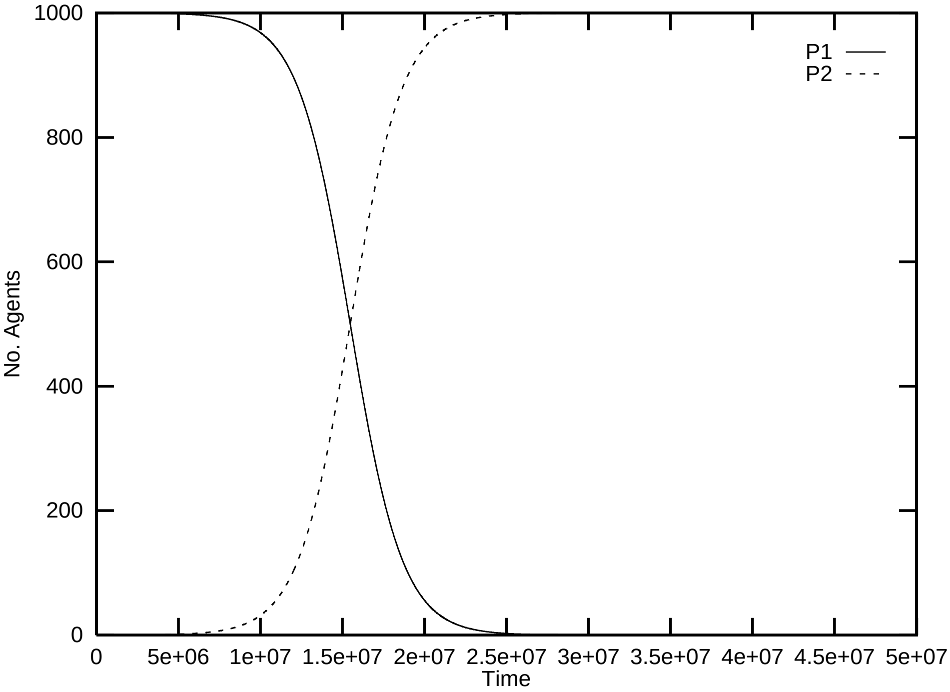


Figure 1: Time evolution towards fixed attractor. $c = 0.6$, $B = 0.01$ and initial conditions: $F = 2500, P_1 = P_3 = 1000, P_2 = P_4 = 0$.

Does the same situation hold for other points in the initial conditions space? Explorations show that for many similar conditions, especially symmetric ones, the behavior remains basically the same: an evolution towards a fixed point attractor.

What happens with asymmetrical initial conditions? Figures 2 and 3 show the behavior for the same parameters and $F = 2500, P_1 = 1999, P_2 = P_3 = P_4 = 0$. The behavior is periodic, oscillating between periods of communicator and anti-communicator prevalence. Note that P_3 and P_4 show the same behavior as P_1 and P_2 respectively, only with a half-period shift. The time scale is also significantly different for the same set of parameters.

This periodic behavior can be seen with many other asymmetrical initial conditions. In Figure 4 we can see the shape of this periodic attractor[17] in the space $P_1 - P_3 - P_4$.

---

[16] This can be easily shown by dimensional analysis.

[17] Strictly speaking the attractor is toroidal because F varies with a different frequency, which happens

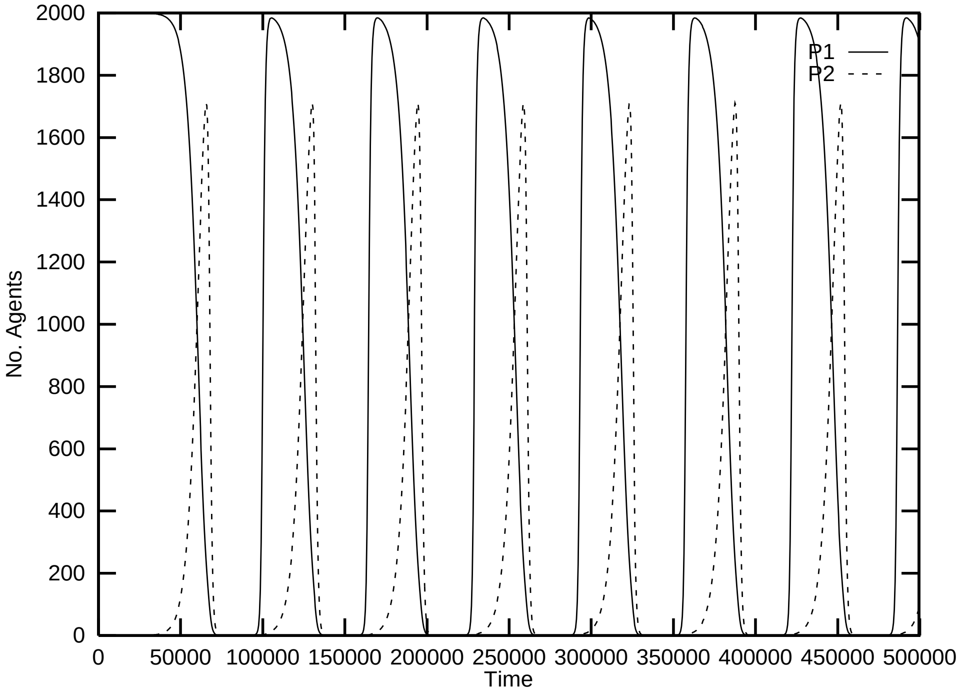


Figure 2: Evolution of P_1 and P_2 for initial conditions $F = 2500, P_1 = 1999, P_2 = P_3 = P_4 = 0$.

This shape remains the same for variations in c and in an important range of variations in B, though the frequencies of oscillations vary in both cases. For $B \cong 1$ and greater, the attractor suffers interesting changes of shape, but remains periodic.

These results lead to some important conclusions. As it has already been said, most (if not all) of the occasions in which the game-theoretic machinery is applied this is done to obtain a first-order, static result such as a fixed-point solution. For many problems this is almost enough. But for many others it is not. Evolution, by definition, is a dynamical process. If sufficient "ingredients" in this process can be considered to be static, then many non-linearities and dynamical effects are gracefully reduced and the game-theoretic approach is the best way to understand the problem.

The question implicitly asked in this section was: *is this the case for the evolution of communicative behaviors?* In other words, is the game-theoretic approach enough to understand our problem? And the answer is simply *no*. It has been found that whether the system reaches an equilibrium depends on initial conditions, i.e. on historical factors. Depending on such factors the system may evolve into a periodic regime of oscillations of very significant amplitude (almost the whole range of population size is covered in turn by each species, as seen in the figures). Once within the regime, the system will remain in it permanently; therefore, the ESS state will never be reached. This has been recognised as "an obvious weakness of the game-theoretic approach to evolution", (Maynard-Smith, 1982, page 8).

More rigorous proofs of these conclusions for a general case can be found in (Zeeman, 1980), who showed that global convergence to an ESS is assured only in the absence of other attractors, which may exist and not be ESS's themselves.

---

to be the double of the frequency of variation of the P_i's, so the attractor remains one-dimensional.

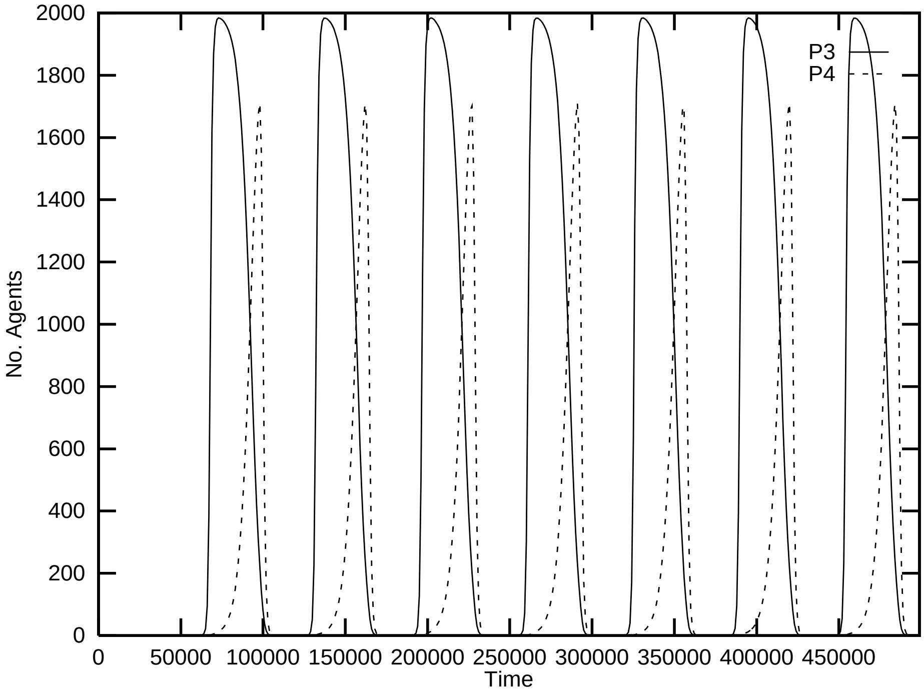


Figure 3: Evolution of P_3 and P_4 for initial conditions $F = 2500, P_1 = 1999, P_2 = P_3 = P_4 = 0$.

To sum up: a dynamical approach to the problem is justified in general because:

1. Whether an evolutionary system reaches an equilibrium state may be contingent on the history of the system.
2. An ESS may be unreachable.

In the particular model presented here, both 1. and 2. hold.

Important as they are, these conclusions do not seem to tell us much about our main problem. Is it to be expected that a society of interacting agents will fluctuate between periods of communicative and non-communicative behaviors? At least with the game-theoretic approach it was possible to go beyond (either to confirm or to discard) simple verbal arguments (the "zero-order" approach). But, so far, the dynamical approach doesn't further our understanding of the problem. It has only told us that game-theory has its limitations, something most researchers are aware of. There are many simplifying assumptions in this model that could be relaxed, and it is easy to see that the effect of such relaxations will most probably have important consequences for the evolution of communication. I will mention some of the most evident:

1. *Spatiality.* Global accessibility does not represent interactions among real organisms fairly. In real life, neighbours tend to remain neighbours, at least for some period of time. The habitat of the offsprings tends to be the same habitat of the parents. All this drastically changes the encounter frequencies p_{ij} adding an important inertia to their time evolution. In fact, selecting a partner in the game ceases to be a purely random process at all.

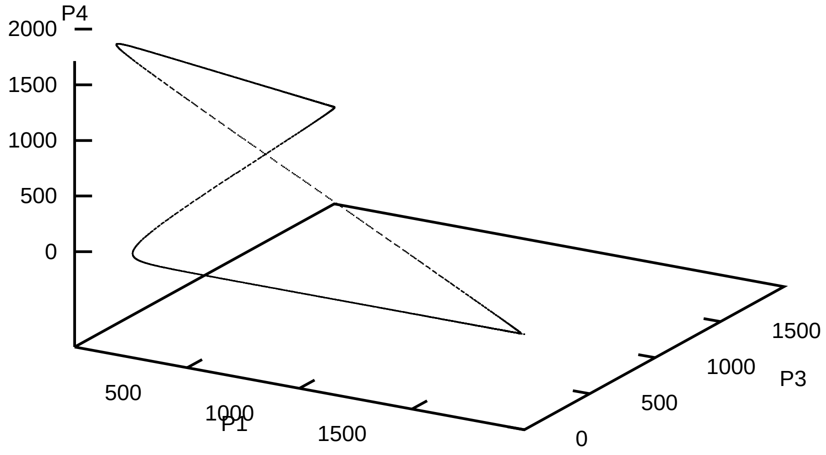


Figure 4: Periodic attractor in the space $P_1 - P_3 - P_4$.

2. *Symmetry* This model is too symmetrical, partly as a consequence of panmictic interactions (which has an uniforming effect on the diagonal terms of the p_{ij}), and partly because the nature of the different species, which, in turn, is a consequence of adopting the simplest unbiased scenario for this game. Any ordered, even permutation of indexes leaves the resulting dynamics unchanged for any set of initial conditions.

3. *Synchronicity* All interactions occur concurrently. Even in games without memory or voluntary choice of partners such as this one, this is a strong simplification, especially if spatial effects were to be taken into account.

4. *Sexual reproduction* Reproduction is asexual. It is to be expected that, with local interactions, the removal of this assumption will add a (local) uniforming factor with direct effect on the p_{ij} matrix.

The model could be extended to include these factors. Each one of them should be modelled in turn, extending perhaps the number of assumptions. For example, a clustering model could be use to consider effects of spatiality, a stochastic update model could address the problem of synchronicity, sexual reproduction could be included and symmetry could be broken by the addition of agents that sometimes behave as communicators but not the rest of the time. Each one of these factors will surely be affected by the way the others are modelled but, with good fortune, after many steps of iterative modelling, it may be possible to arrive at a final model.

This is an interesting enterprise, however, after some pragmatic considerations I decided to turn to a computational approach. Not only is the modelling much easier (it is not necessary to make the above mentioned assumptions in the first place) but the possibilities for studying the problem are greater due to the inherent flexibility of this approach.

6 A game of action coordination: computational model.

In this section I present some computational investigations into the problem of the evolution of communication. The game described in the previous section together with some variations of it, will be the focus of the computational modelling.

6.1 The model

The dynamics of this model are very similar to those described in Section 5. Agents live, interact and die in a shared environment. The physical environment is represented as a toroidal grid of size X_{max} by Y_{max} with an Euclidean metric as is common in work in simulation of adaptive behavior. Agents are situated in this environment in a position which, in principle, does not change with time; agents are not mobile. A number, F_Q, of energy repositories, or food sources, are also distributed within the environment and remain fixed, ($F_Q = X_{max}Y_{max}$, so that there is, on average, one food source per cell in the grid). Food sources distribution can be uniform or random, this has proven not to be relevant to the results. These sources are created at the beginning of the simulation run and their position and associated food type remain unchanged. A food source cannot be created nor destroyed during the rest of the simulation, but its energy content may run out, as a tree runs out of fruit. Food types are identified with numbers "0", "1", etc.

Neighborhoods are simply defined as squares centered around an agent. Agents can interact with other agents and access food sources only within their neighborhoods. Initially agents are randomly distributed in the environment. Each agent possess an internal energy counter that specifies the energy available for self-maintenance. Every time an agent is born its energy level is set at a value chosen following a Gaussian distribution $\sigma(E_i, \Delta_e)$ centered on E_i. (Typical values are $E_i = 150$, $\Delta_e = 10$.) Energy flow is strictly accounted for. The environment is subject to a constant flux R (500 per time step), which is manifested in a stochastic increase of the energy stored in the food sources. Agents derive their initial energy from their parents, and during their lifetime, from the food sources. This ensures energy conservation. Besides, agents spent a fixed amount of energy D (0.25 per time step) every time they act (successfully or not), which does not return to the environment, so as to make sure that the second law of thermodynamics is observed.

At each time step a number of agents equal to the size of the current population is randomly selected to perform the first role in the interactional game. The updating is performed asynchronously although, on average, each agent will always be selected for this role once per time step. A second agent and a food source within the neighborhood of the first agents are randomly selected (Figure 5). If no agent is found after a finite number of trials, the first agent looses its chance to play the game, and the energy cost is discounted anyway. Each time an agent is selected its age is incremented by one, regardless of the role it plays.

The payoffs of the players will be constituted by shares of the energy contained in the food source that has been selected. They will be assigned in the same way as

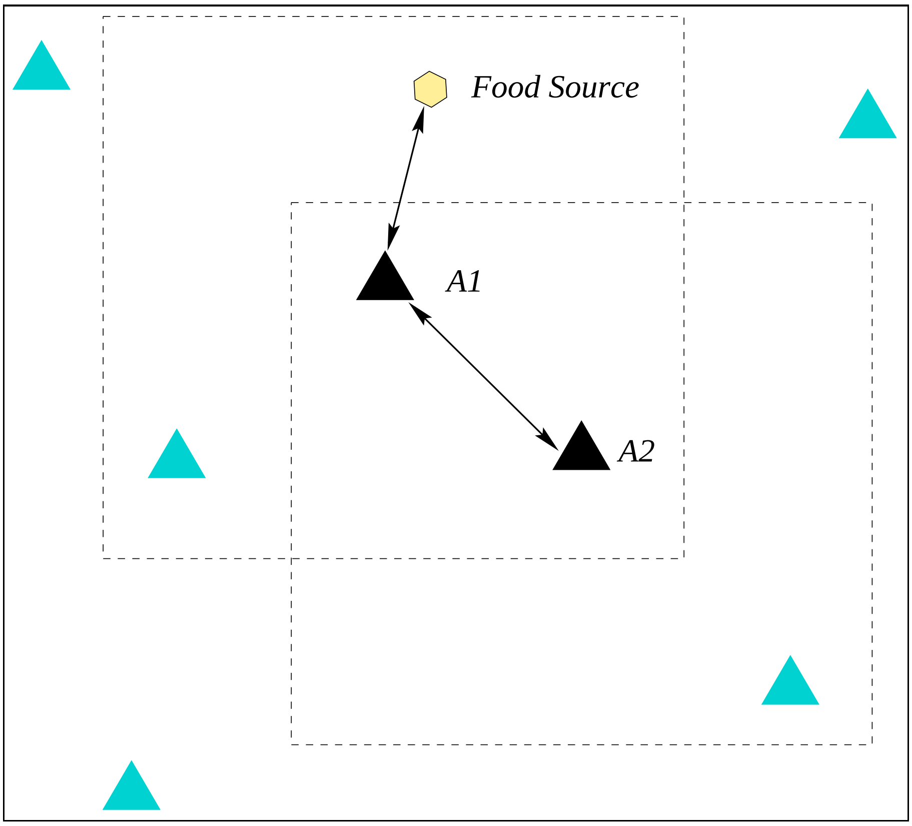


Figure 5: Selection of a second player (*A2*) and a food source within the neighborhood of the first player *A1*.

described in Section 5: in equal parts when both of them act correctly and, when they do not, a proportion $c \geq 0.5$ of the energy for the agent acting correctly and nothing for the other. When an agent has accumulated enough net energy it is able to reproduce. The corresponding level of energy is selected according to the distribution $\sigma(2E_i, \Delta_e)$. Reproduction is sexual, and selection of partner is based only on the locality condition. The initial energy of the offspring is supplied by the parent whose high energy level triggered the reproduction event in the first place.

Those agents unable to sustain their costs will eventually run out of energy and die. Costs are the same for all agents, with the exception of very "old" ones. After a certain age has been reached, costs are mildly incremented linearly with age. This is the only developmental feature included in this model, and it is done in order to avoid the presence of "immortals". This mechanism replaces the lack of a good model of developmental rules, with which it could be expected that typical lifespan would result from an interplay between population dynamics, availability of resources and evolution of those rules.

As a consequence of energy based reproduction and death, the size of the population is variable in time.

The structure of agents is that of a state-less machine. As I said, the focus of this work is on the global mechanisms that allow or constraint the evolution of communication as a first step towards an understanding of the problem. Undoubtedly, the area of agent's structure and development will prove to be of much relevance for future research, but in the present work these features have been reduced to a minimum.

Actions depend on the perceived food type and on the perceived external manifestation of the partner in the game, when these are visible (an example of a behavioral matrix is shown in Table 4). The results presented later, with the exception of those in Section 6.4, correspond to simulations in which the game is played as described in Section 5, that is, with non-sharing of information about the environment. Agents playing the second role will only act according to the perceived external manifestation of the actions of the first agent, and on the base hypothesis that the food is type "A". Correspondingly, the actions of the first player will depend on the perceived food type and on the base hypothesis that they perceive an external manifestation α. This means that, for this particular game in which there is only one interactive step, only one column and one row of the behavioral matrix will be used. Nevertheless, the whole matrix representation is kept as it will be used completely in the experiments described in Section 6.4.

	Perceived Signal			
Food Type	α	β	$\cdots$	ν
"0"	B α	A α	$\cdots$	C δ
"1"	C γ	D β	$\cdots$	A β
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
"n"	B α	A α	$\cdots$	C δ

Table 4: Example of look-up table that encodes the behavior of an agent. Capital letters represent effective components and Greek letters external manifestations of actions. Numbers represent food types. In this example if the agent perceives food type "1" and an external manifestation from the other agent is α, it will perform an action in which the effective component is C and the external manifestation is γ. Each position in the table is encoded by two genes.

The behavioral matrix of the machine is encoded in a haploid genome, represented by a bit-string of length G, of which, depending on the game, only G_a places will code for relevant behavior. Offsprings receive their genome from the result of a uniform crossover operation on their parents genotypes, plus certain probability of mutation μ per place.

Once started, the simulation proceeds until the population is extinct or for a predefined number of time steps. Parameters were chosen in order for the environment to have an effective carrying capacity of 500 - 1000 agents. This means that the environment is sufficiently benign so that agents do not suffer strong pressures to resolve their communicative strategies in either way. This is a satisfying scheme in which selection acts negatively. There is no fitness function to optimize, neither are there any special rewards nor punishments for behaving in an specific way apart from the rules of the game.

A problem derived from the use of this scheme is the lack of obvious measures of evolution, which is hardly surprising, given that all living agents are, and have been, equally adapted (see Section 3.1). Many variables were monitored, the size of the population, the amount of instantaneous environmental energy, the average number of offsprings, etc. But, for our purposes, the simplest way to monitor the evolution of communication, is to look at changes in the average activity success of the first and second players, and the average success in coordinated activity, and correlations between all these.

6.2 Success in coordination and correlations.

Results discussed in this section were obtained from two sets of simulations in which only four different types of food were included in the environment, in one of them each food type had a different associated action, and in the other a given action was correct for two food types and, therefore, only two actions were relevant (even when agents were able to perform more). Simulations with more food types and actions were also carried out, and the results were similar, though much more expensive to obtain and analyse.

In contrast with the mathematical model of Section 5, agents are initialized totally at random. This means that they will also have to evolve an association between food types and the effective components of actions when they play the first role, (in Section 5 we supposed that the first agent always acted correctly). This is a task that is consistently and very rapidly achieved in all the simulations, and its explanation in selective terms does not present any problem, given that each new correct association that is incorporated into the agents provides a net benefit for them and does not interfere with previous associations. In other words, the specific genes that code for these associations have no epistatic links with other genes in the genotype.

Both in simulations with four food types and four actions, and simulations with four food types and two actions (henceforth: "4-4" and "4-2" simulations), a baseline case was run in order to understand what kind of activity emerged even when agents weren't able to perceived each other reliably. In these runs, the external manifestation of the first agent's action was replaced with a random number when perceived by the second agent. We may call the proportion of cases in which both agents acted correctly simultaneously, the *coordination ratio*. In 4-4 baseline simulations, this ratio stabilized at 25 % and in 4-2 simulations at 50 %. Since no food type is predominant, the best "guess" a second agent can make given that there is no correlation between the "signal" it perceives and the particular type of food it is dealing with, is to perform any of the four actions in the 4-4 simulations, and any of the two relevant out of the four possible actions in the 4-2 simulations. The achievement of these levels of coordination is independent of the parameter c

This means that any success in coordination of behaviors will be manifested with a greater coordination ratio than those observed in the baseline cases. Only then will it be possible to say that second players are taking advantage of any correlation between effective component and external manifestation that may exist in the actions of the first players.

Figure 6 shows the evolution of the success ratios for the whole population in a typical run, (for a same set of parameters variations between different runs were small). This quantities should not be confused with any measure of "fitness" of the population. *R1* indicates the average proportion of successful action for first agents, which, as predicted, climbs steadily up to a level near 100 %, (perfect success is not necessarily achieved due to the effects of random mutations). *R2* shows the same quantity for second players, and *R12* for both roles simultaneously (the coordination ratio). We see that both *R1* and *R2* start at a level of 25 % which corresponds to a random guessing behavior. Note that the correlation between *R2* and *R12* increases with time and eventually they show the same behavior. This can be explained if we consider that as *R1* approaches 100 % the number of times that the second agent acts correctly and the number of times that both of them do, tend to coincide. This correlation alone (Figure 7) does not tell us

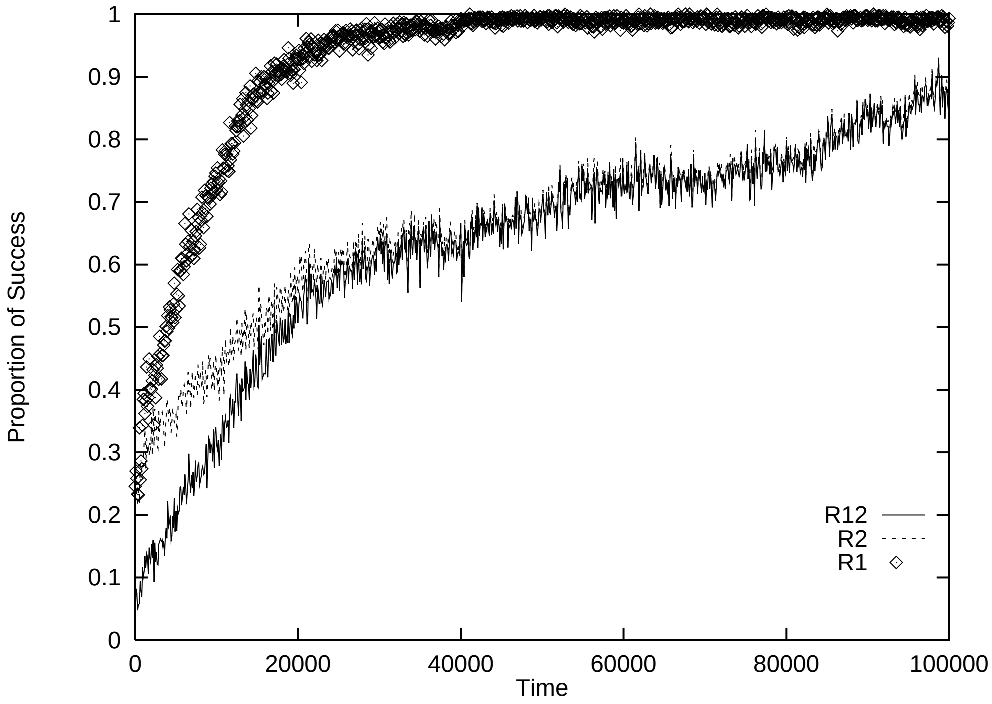


Figure 6: Proportional success in action for the first player (top line, *R1*), second agent (*R2*), and both simultaneously (*R12*).

that coordinated activity is evolving. However if both this correlation and a greater than baseline level of coordination are observed, then the sufficient and necessary conditions are met to claim that coordinated activity, or communication, has evolved. It can be said that this evolution achieves a relatively steady state and this is a very different behavior to the two possibilities observed in the mathematical model (either oscillations or no communication at all).

By observing the resulting behaviors once a relatively stable state has been achieved it is possible to determine that signalling behaviors presents a strong correlation with action and not with food types. Therefore agents in 4-2 simulations agents evolve a "two-signal" code for the two relevant actions that can be performed.

A study of the effect of parameter c on *R12* was performed to see how the extra payoff against coordination affected the level of simultaneous success. The intuitive expectation is that this level will decrease as c increases from 0.5 to 1.0. Figure 8 shows the resulting behaviors of *R12* for different values of c for 4-2 simulations; analogous results were observed in 4-4 simulations.

We see that the proportion of coordinated activity for the whole population decreases as c is incremented from 0.5 to 0.7, and for greater values the baseline case is reproduced. This means that the level of communication goes from almost perfect for $c = 0.5$ (neutral cost) to coordination by guessing when $c \cong 0.7$. The fact that for a certain range of costs *against* communicating, coordinated activity evolves anyway is in contradiction with the intuitions that can be gained from the mathematical model or from simplified selective arguments. Explanations for this phenomenon should be sought first in those assumptions that were made in the mathematical model but were unnecessary in the

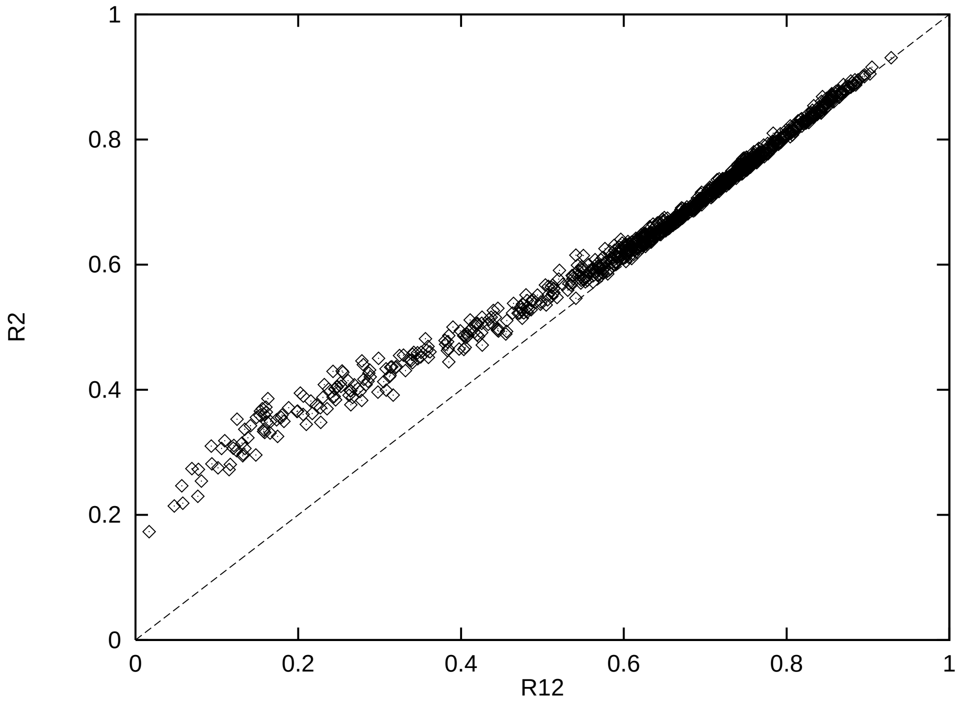


Figure 7: Correlation between second player success and coordinated activity for the same run shown in Figure 6.

computational model.

The first question that makes sense to ask is how different are individual communication levels from *R12*, which is just an average over the whole population. If there is much variation, can we attribute it to any structure in the population? Answers to these questions can be found in a study of the spatial "behavior" of the evolving system.

6.3 Spatial structures.

A simple inspection of the resulting data in all simulations shows that the individual history of coordination success can differ significantly from agent to agent and from the value of *R12* at that time. For instance, groups of agents achieving 90 % of coordination success can coexist with other groups that achieve 60 % both in a stable state during the same simulation run. Even though an individual historical average is qualitatively different to an instantaneous population average, one would expect the resulting numbers not to differ much, especially if *R12* has been stable for some time. In fact, if all the individual levels of historical coordination success are averaged over the whole population, the result in very similar to *R12* even during transients. What makes, then, some individual so different from others?

The answer lies in the spatial distribution of agents in the environment. In order to understand why, first we must understand why communication evolves at all even for values of $c > 0.5$, and what the influence of the spatial organization on this evolution is.

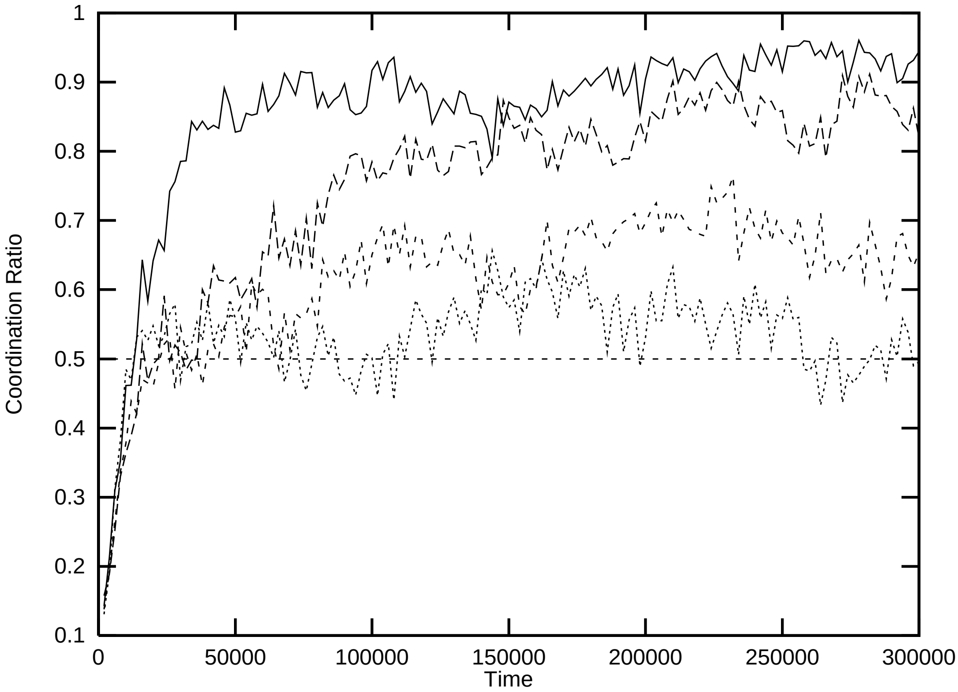
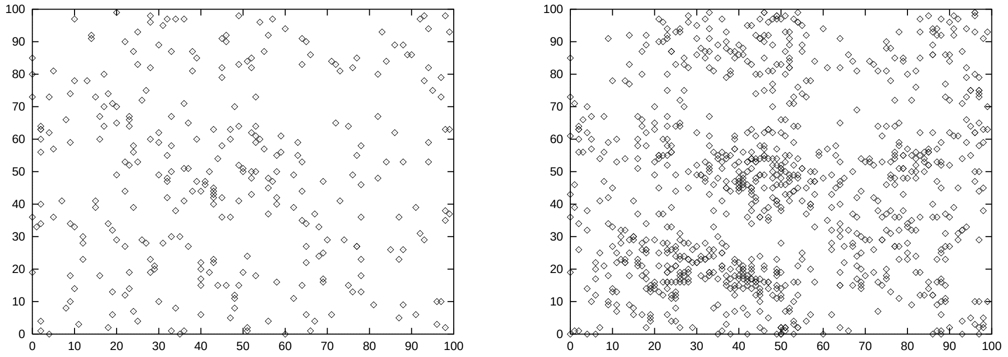


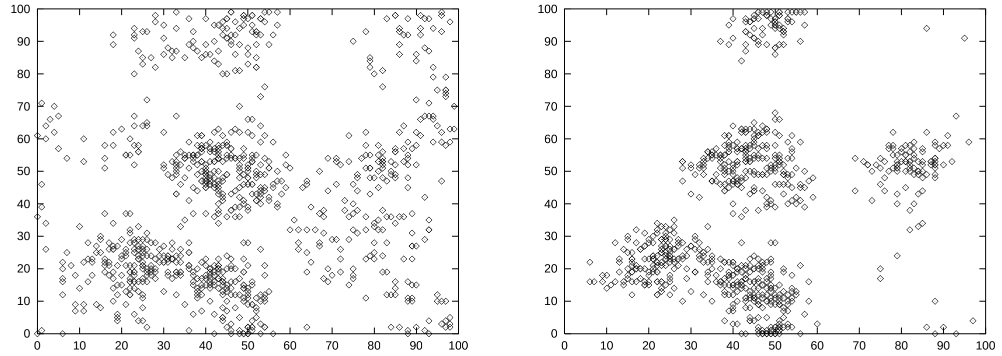
Figure 8: Whole population coordination ratio (*R12*) for different values of c in a 4-2 simulation. The top (full) line between 90% and 100% corresponds to $c = 0.5$, that is, no cost against communication, the line between 80% and 90% corresponds to $c = 0.55$, the one between 60% and 70% to $c = 0.6$ and the one between 50% and 60% to $c = 0.65$. Values of c beyond 0.7 result in a level of 50% which is the baseline case for this game.

6.3.1 Cluster formation and stability.

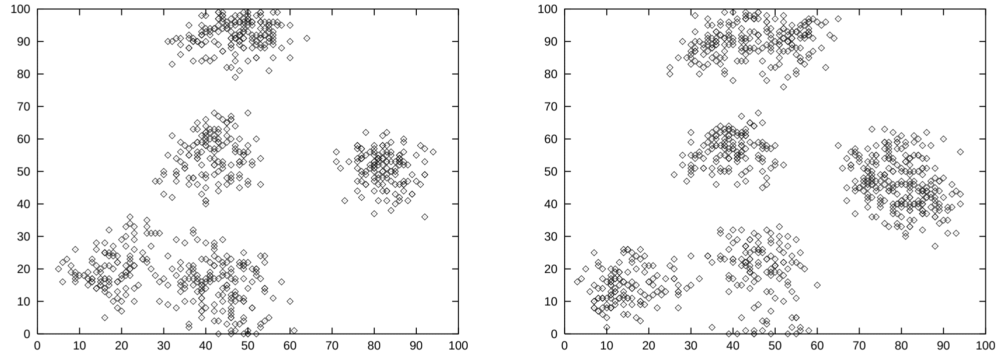
Figure 9 shows the first stages in the evolution of agent spatial distribution. Initially agents are distributed randomly across the whole of the environment, in which boundary effects have been avoided with the use of periodic boundary conditions. We can observe how this initial symmetry is broken and how agents show a tendency to aggregate into clusters. Symmetry breaking is caused by minor differences in the initial distribution in positions and also due to asynchronous updating; so that some agents will be more successful than others just because they have a few more agents to interact with or were called to act a few more times and, therefore, they have a slightly greater chance of accumulating enough energy for reproduction. As reproduction is also a local process and the position of the offspring does not differ much from the position of the the parents, there is a positive contribution to the same effect, namely the accumulation of agents in these particular areas. On the other hand, agents who are relatively isolated will have less chance of interaction and will tend to die sooner, also contributing positively to increase the isolation of agents in similar positions. Evidence supporting this explanation can be seen in Figure 10, in which the number of agents that have not been able to interact because of isolation is plotted against time. This number decreases sharply as the clustering process advances, and then increases as clusters grow and, finally, it stabilizes at a level above zero because, even when clusters are already formed, some agents, especially those in the cluster periphery, will not find a game partner within their neighborhood in the alloted time. As long as the effect is not cumulative, these



(a) t = 0 (b) t = 2000



(c) t = 4000 (d) t = 8000



(e) t = 16000 (f) t = 28000

Figure 9: Formation of clusters. Each figure shows the distribution of agents in a 100x100 toroidal environment at different times during the simulation. Neighborhood size: 10x10.

agents will find a partner in another time step, and therefore they will still have a chance to survive. As far as the previous explanation goes this process ends once clusters are formed, and there are no agents that do not "belong" to a cluster.

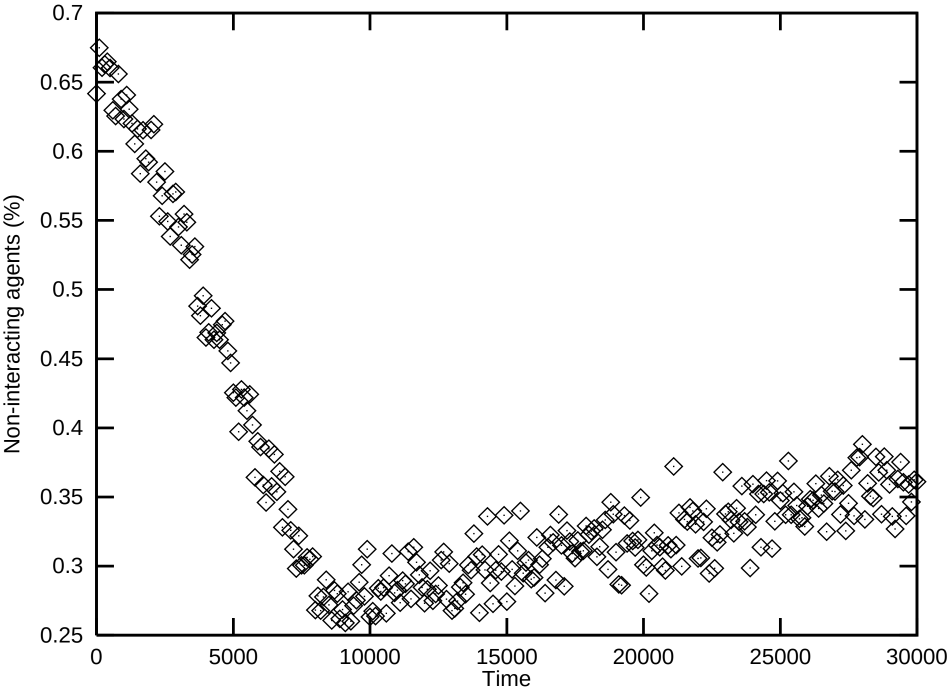


Figure 10: Proportion of agents that are unable to find a partner in less that 10 trials, and therefore do not participate in an interaction at that time. This figure correspond to the same run as in Figure 9.

It is important to note that the rules of the game play an important part in the formation of clusters. It is possible to conceive an evolutionary process in which all features are local as in the present case, but in which there is no specific interaction between agents, but the criterion for payoff allocation would depend exclusively on the individual behavior of the agents. Such a case could also shown interesting spatial patterns but, in general it will not be possible to distinguish spatial unities, as variations will tend to be continuous (see, for instance, the distribution of species in Di Paolo, 1996a).

Unfortunately, cluster formation mechanisms do not provide a satisfactory explanation of why clusters remain more or less stable structures as it is generally observed. Why don't they expand or shrink or break or fuse? The fact is that these behaviors *have* been observed, but they are the exception rather than the rule, and they can be explained once an understanding of the general case of cluster quasi-stability is achieved.

Clusters possess their own phenomenology as composite spatial structures, which is very interesting in itself, though not all of it is directly related to the problem of communication[18]. Many of their behaviors can be understood in terms of the global effects of agent-agent interactions, which, in a sense, define them as composite unities. In order to understand why clusters are stable it is necessary to know what the "forces"

---

[18] For instance, clusters can follow food gradients, define and maintain their own frontiers and self-regulate by changing their size.

involved are and how they reach an equilibrium and what happens when they don't. This task implies looking at these structures in more detail.

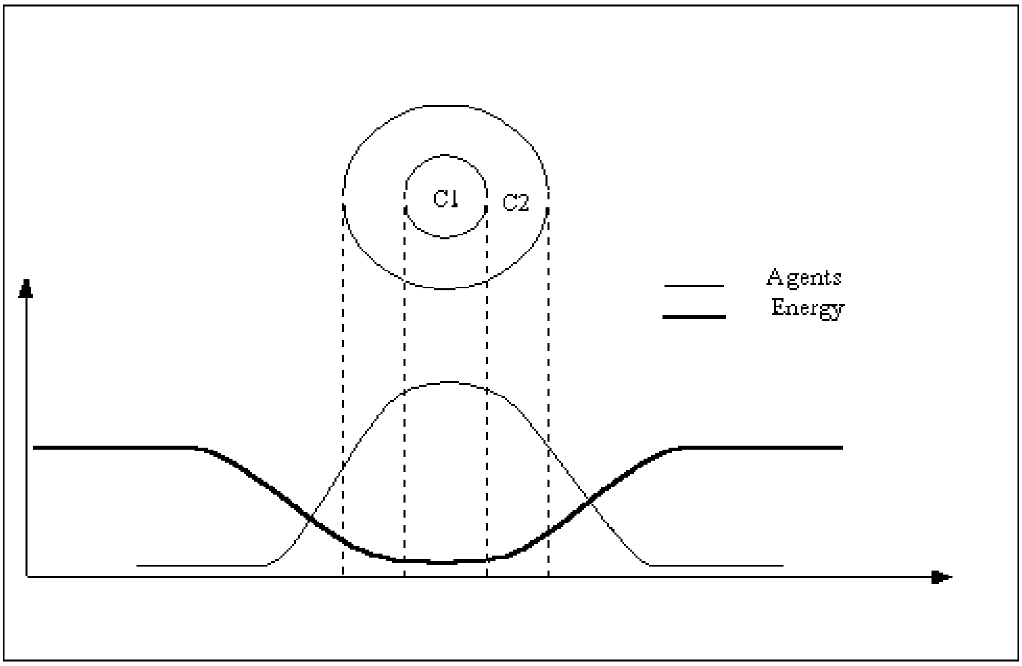


Figure 11: Energy and population distribution inside and around a cluster.

Tendency to expand. Figure 11 shows qualitatively a typical observed distribution of environmental energy and density of agents inside and in the vicinity of a cluster. It is easy to see that resources will be more frequently used in more populated areas toward the center of the cluster than on peripheral areas in which the population is more sparse, so that the amount of available energy will, on average, decrease towards the center as shown in the figure. Agents living in the periphery will have access to resources of greater quality, and the average energy gained per game played will be greater than that of agents in the center region. So, at first glance, peripheral agents would seem to be better off and have a better chance of having more offsprings, therefore the cluster would seem to experience a tendency towards *expansion*.

Tendency to collapse. On the other hand, while all agents have the same probability of being selected to play the first role in the game (once per time step on average), the frequency with which an agent *actually* plays the game depends on the surrounding population density. The more densely populated the area, the higher the chances of finding a partner to play with. For this reason, agents living near the center of a cluster will play the game more frequently. Moreover,the probability of an agent playing the second role also depends on its position within the cluster. Given that the density distribution is not uniform, certain regions of an agent's neighborhood will be more densely populated than others, which means that second agents will be chosen more frequently from those regions. As a result of the circular geometry of the cluster and of the increasing density of agents towards its center, agents in this particular region will benefit from this effect in a cumulative way, and will, on average, play the second role more times than they play the first role. Conversely, agents living near the periphery will be chosen less frequently for playing the second role because their distribution is more sparse, and therefore, they will, on average, play the first role more times than the second role. The ratio between the frequency of being the second player with respect to the frequency in which agents act as first players have been observed to range from a minimum of 0.9 at the periphery to a maximum of 1.25 at the center of the cluster. In short, agents living in the populated areas near the center will play the game more frequently and, besides, they will perform the second role more frequently than agents living in sparse areas so they will stand a better chance of receiving (by communication, or just by guessing) more energy per unit of time. This provides the cluster with a tendency to *collapse*.

The equilibrium of both these tendencies determines, in principle, the size of the quasi-stable cluster. At a certain size the extra energy gained by the peripheral subpopulation will match the energy lost by their being more sparsely distributed than the center subpopulation (regions *C2* and *C1* in Figure 11). Resulting clusters have been observed to have a typical radius of 1 or 2 neighborhood sizes.

Energy transport. There is also a dynamical effect that reinforces the achievement of stability when the cluster is far from equilibrium and explains the shape of the population density within the cluster (Figure 11). Any agent playing the first role can only have access to food sources in its neighborhood, but agents playing the second role can receive energy from outside its neighborhood. This is a consequence of the fact that the neighborhoods of the first and second players must intersect but not necessarily coincide. Assuming that the geometry of a cluster is circular (or at least convex), we see that a uniform density within the cluster cannot be stable. This is because, agents in the center will play the second role more frequently than their peripheral neighbors (initially only because of geometrical reasons) and, consequently, they will be accessing energy contained in peripheral, as well as center, food sources more frequently than agents in the periphery, translating in a higher rate of reproduction and a local increment of density. This inequality will cease to be, when the peripheral population becomes so sparse that the energy transport is neutral: an equilibrium situation. This transport, provides an extra negative feedback if the force towards expansion should overcome the force to collapse. If for any reason the periphery gets more populated, i.e. the cluster is expanding, the energy flux will increase and as a consequence the distribution on the periphery will become more sparse because more energy will be diverted towards the center.

Genetic homogeneity. Due to the nature of the cluster formation process and the homogenizing effect of sexual reproduction, clusters tend be inhabited by agents who are very similar genetically. This will still be the case even when new mutations appear and become fixed. After a sufficiently long time (in practice not very long) there will be agents bearing this new mutation distributed across the whole cluster. The reason for this is that it is impossible, because of the typical size of the cluster, to keep two regions of it isolated so that certain genotypes will be found in one region and not in the other. Even if two different regions within a cluster were found to have radically different environmental features, the quasi-homogeneous allocation of offsprings added to the effect of sexual reproduction will represent an uniforming factor over the whole cluster population. This has consistently proven to be the case in all simulations. Agents within a cluster are genetically very homogeneous, except for genes encoding neutral behaviors[19], while agents from different clusters may, in general, differ. This is an

---

[19] Neutral behaviors are found in 4-2 simulations when different agents use different signals for the

important point to bear in mind in the argument that will follow.

6.3.2 Why does communication evolve?

I said that clusters as quasi-stable structures result from an interplay of two opposing tendencies in the spatial organization that arise from the fact that conditions differ at the center and at the periphery of the cluster. The next obvious question that must be examined is if this difference of conditions has any effect on the evolution of coordinated activity. The answer is yes.

Environment at the periphery of a cluster. Agents are subject to two qualitatively different micro-environments which emerge as a result of the spatial organization and, ultimately, as a result of their own activity[20]. We saw how agents living in the periphery have access to resources of better quality (Figure 11) and how these agents tend to play the second role less than they play the first one. Besides, they also tend to interact fewer times in absolute terms because of their sparse distribution (i.e. they have a greater probability of "missing" a turn because finding a partner in the alloted time is harder for them). This means that, for these agents, to communicate is particularly costly because every time they communicate they lose one of their fewer opportunities to obtain a greater absolute payoff compared to agents living in the center of the cluster[21]. If we were asked what would be an ESS in such an environment, the answer clearly would be: "not to communicate".

Environment at the center of a cluster. Agents inhabiting the center of a cluster have access to poorer, much more frequently used, resources, therefore the individual gain for not communicating is not too high in absolute terms, though still positive. On the other hand, these agents tend to play the second role more times than they play the first one, which means that there may be cases in which agents will have a positive individual gain if they *do* communicate, simply because they will be acting as second players more frequently (Darwin's principle of "use and disuse" (Darwin, 1902, Chapter V)). Provided that the cost against communication c is not too high, agents living in such an environment will tend to experiment a stronger selective "pressure" to be good interpreters, than the "pressure" to be bad signallers. While it may pay a little extra to be a "deceiver", the situation may be that once a deceiver lineage starts to grow, it will pay more to "break the deceiver's own code", rapidly re-establishing the level of coordinated activity.

Selection. Ideally this situation would be resolved by having two distinct subpopulations, one of (almost) total communicators in the center of the cluster and one of total non-communicators in the periphery, but this is not possible as selection must act within the constraints provided by the spatial organization. Given that the sizes of each subpopulation are comparable, and a newborn agent has comparable probabilities of being placed in any of the "two" regions (almost independently of the position of its parents), then there is no ground for selection to be very specific about which of the extreme behaviors to choose. Therefore, *surviving agents will tend to be able to partially*

---

same food type, but agents playing the second role respond in the same way in both cases, and the used signals do not conflict with signals corresponding to other food types.

[20] Or a continuum of micro-environments to be more precise.

[21] It must be remembered that the extra payoff for an agent that communicates is given by $(c - 0.5)\bar{F}$ where $\bar{F}$ expresses the *local* average value of environmental energy.

satisfy the conditions of both micro-environments and, consequently, they will necessarily possess the ability to communicate up to a certain level.

This is consistent with the observed variation of the coordination ratio for the whole population with c (Figure 8). The level of successful coordination is above the baseline case for $c > 0.5$ and decreases until it reaches this level for $c \cong 0.7$. According to the above description at this value of c the advantages of being good interpreters for the center population are overcome by the advantages of not communicating even within their poorer-resource micro-environment.

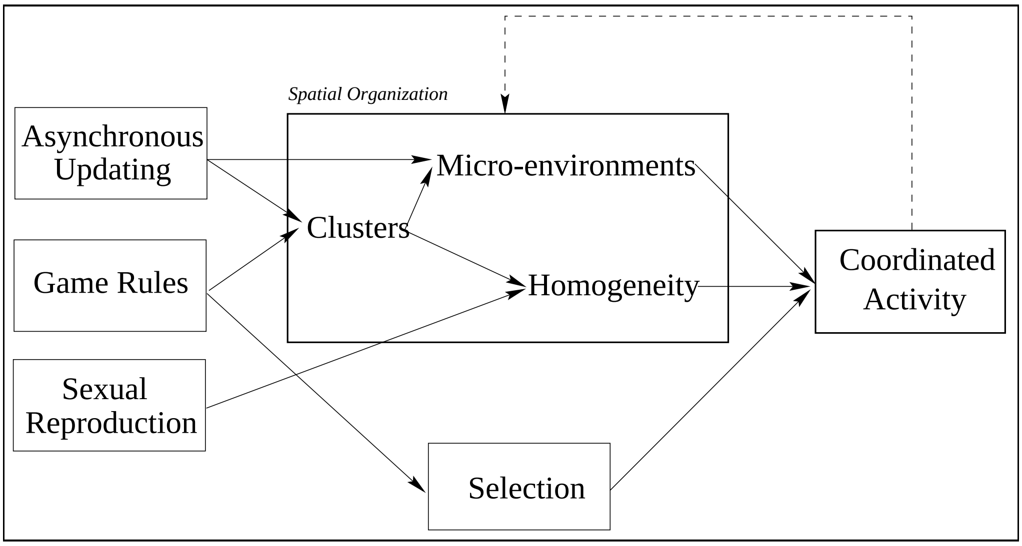


Figure 12: Explanatory mechanisms involved in the evolution of coordinated activity in the computational model.

Thus, the evolution of communication in this model can be explained by the interplay of self-organizing and selective mechanisms. A diagram showing the relation between the components of this explanation can be seen in Figure 12, (the dashed line will be explained below). An alternative explanation could be attempted in terms of the mechanism of *kin selection*. Given the relatedness of agents inhabiting the same cluster, and given that coordinated activity results in greater payoffs at the population (if not at the immediate individual) level, then it would make sense for an individual to communicate with related individuals thus spreading the benefits over a number of other agents carrying the same genes. This is one of the possible mechanisms in which cooperative action can evolve within the neo-Darwinian framework. While kin selection may be playing an important factor in this case, this explanation cannot account for the stability of communicative behaviors in our model just by itself. If we ignore spatial factors, kin selection can explain stability only when the population gene pool has a very narrow distribution near an adaptive peak surrounded by a valley of low fitness. In such a case, deceivers cannot become fixed in the population, because this implies the survival of intermediate mutations lying in the valley of low fitness. This is not the case in our model. In fact deceivers (agents that take advantage of the signals of others, while emitting wrong signals themselves) are just one mutation away. The only requirement for a communicating agent's offspring to become a deceiver is to flip one bit of any gene coding for an already correlated external manifestation, thus becoming a good interpreter and a bad signaller, and therefore collecting the benefits of the existing communicating behaviors of others without contributing to those behaviors. Kin selection would work as an explanation only in the case of deceivers being sufficiently far in genotype space, so that the only way they could be reached is by the Dobzhansky-Muller (see Orr, 1995) mechanism of speciation, that is, by reproductive isolation.

In contrast, the explanation in terms of spatial organization can account for the stability of partial coordinated behaviors because it cannot be subject to the same criticism that the kin selection explanation. A lineage of deceivers *can* be born into a cluster with good communication level, and indeed these deceivers may do better than their communicating neighbors. But eventually, if not right at the beginning of the lineage, deceivers will have to inhabit a micro-environment in which communicators do better. As a consequence, the deceiver lineage cannot grow beyond a certain limit.

However, this explanation is itself, a first order approach. It has been found that, in general, coordination levels differ importantly among clusters. We still haven't explained the first question we set about to answer: why are there differences in the communicative behaviors of agents.

6.3.3 Many satisfying possibilities.

I have called the state achieved by the clusters quasi-stable. The fact is that clusters do change over time, but not very rapidly; the velocity of change is related to the speed of reproduction and the longevity of the agents. The more obvious kind of change is change in position. Clusters can move diffusively or following slight spatial gradients in energy distribution. It is not usual, though not impossible, for clusters to collide and fuse. This can be explained by the fact that energy depletion in the area between two nearby clusters will, in general, cause these clusters to be repelled from each other, moving away towards greater energy resources.

Clusters can also grow and, more rarely, shrink; these are interesting phenomena because they are related to the level of coordinated activity within them. The following correlation has been consistently observed: *clusters with high level of success in coordination are larger in size and more populated than clusters with a lower level.* In general, the former can have a radius of up to 3 neighborhood sizes, while the latter have a radius of 1 neighborhood size or even less. An example of this distribution can be seen in Figure 13 which corresponds to a 4-2 simulation after 80000 time-steps with $c = 0.52$. At this time the average instantaneous success in coordinated activity for the whole population is stable about 75 %.

The very co-existence of qualitatively different clusters in the same simulation run is evidence that more that one "solution" is allowed within the satisfying scheme for evolution implemented in our model. Spatiality allows for relative isolation which combined with the lack of prescriptive rules, such as explicit fitness functions, allows for variety.

Why a particular cluster is small with a low level of coordinated activity while another is big with a good level of coordination must mainly be attributed to contingencies in their respective histories. We cannot look for general reasons because all agents evolve under the same general laws. All we can do is to describe, in terms of feedback mechanisms, certain tendencies that appear once a cluster is already embarked in a particular historical path. It comes as no surprise that this kind of cybernetic explanation will focus mainly on the "loop" defined when the results are seen as modifying the causes

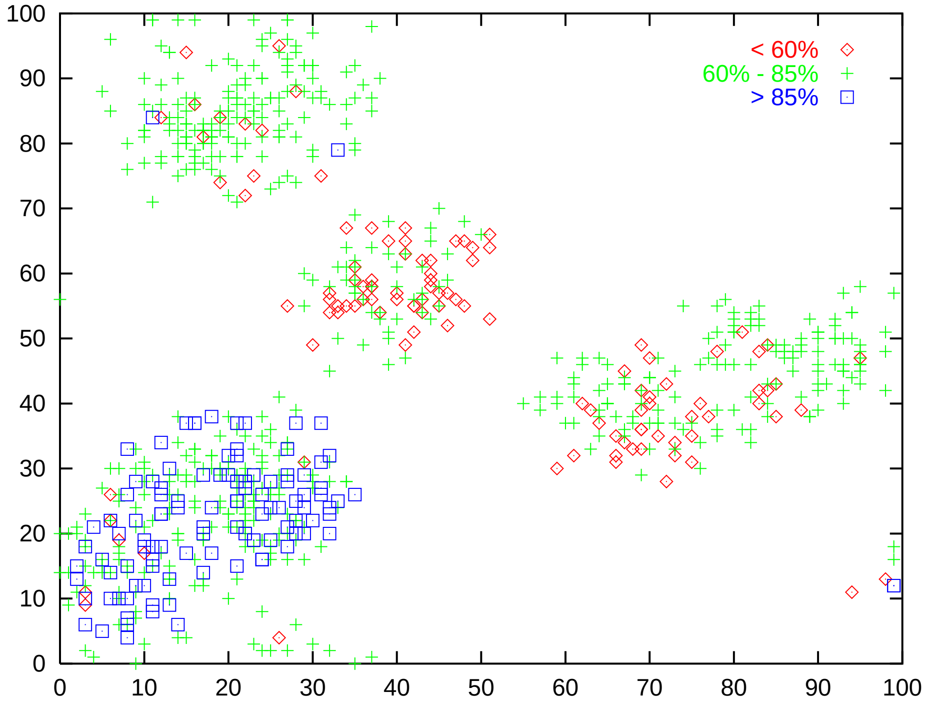
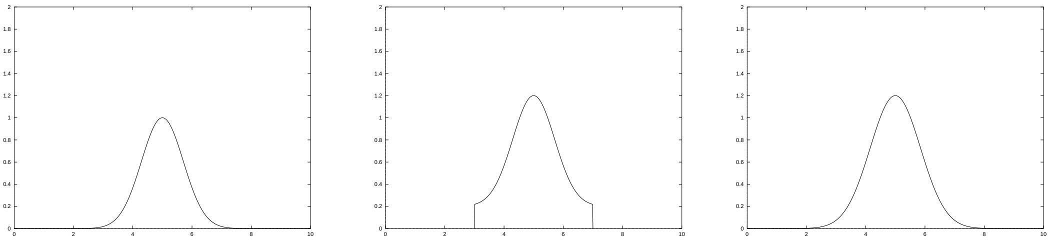


Figure 13: Qualitatively different clusters in a same simulation run. Note how most of the agents in the cluster near the center of the figure have a low index of coordinated activity (below 60%), while in the neighbor cluster in the lower left corner of the figure most agents show a much higher level (around 85%).

(the dashed line in Figure 12).

In general, it is observed that big clusters either grow from small ones, or are relatively big at the formation stage, so that there are two phenomena in need of explanation: 1) why do initially big clusters favour coordinated activity, and 2) why an increase in coordinated activity can provide a tendency towards cluster growth. In explaining these two phenomena I will make explicit an assumption that I have been using so far: interaction among different clusters is negligible. In some cases this assumption cannot be maintained.

Understanding how a big cluster can favour coordinated activity is straightforward if we accept the explanation of why this activity evolves at all. If we look again at Figure 11, we will see that the levels of food and agent densities outside the cluster reach a stable value (F_{ext} and 0, respectively) which, since there is no interaction with other clusters, are independent of the cluster's size. In fact these levels are the boundary conditions that must be satisfied in all, sufficiently separated, clusters. This means that very near the periphery there is not much difference if the cluster has found its equilibrium at a certain size or other. However, if we assume a constant shape the distribution of agents (this assumption has been numerically supported by the results), then in the center there will be an important difference: the number of neighbors that may call a given agent will increase as the cluster size increases, thus incrementing the frequency with which this agent is confronted with the task of interpreting a signal relative to the frequency with which it has to emit a signal. Therefore, the resulting level of success in communication will evolve towards a greater value than in the case in which this relative frequency is



(a) (b) (c)

Figure 14: Changes in cluster density as the level of communicated activity increases by a small amount: (a) indicates the initial state, (b) shows the increase in density as agents consume more energy, and (c) shows the expansion stage (see text).

lower, that is when the cluster is smaller.

This explains why initially bigger clusters have more chance of evolving a higher level of coordinated activity. It remains to be explained what happens if this level suddenly increases in a small cluster. First of all, it is easy to see that this is an unstable situation simply because more coordinated activity implies a bigger drainage of energy from the space occupied by the cluster, so that either the cluster changes, or the level of communication is restored to a previous value, or the cluster disappears. The first possibility has been observed rather frequently, the second one is very hard to observe and not much can be said about it in the present implementation of the model, and the third one has been observed once in a case in which a cluster was constrained in its growth due to the presence of neighbor clusters, and so it broke into two smaller pieces with a lower level of communication. In this particular case the assumption of cluster isolation did not hold.

In the first alternative we have a small cluster in which the level of coordinated activity increases as the population is (geographically) invaded by a "more communicative" mutant. We suppose that for invasion to occur the net increase in communicative behavior has to be rather small, otherwise the lineage would die out very quickly. The first consequence of this geographical invasion, will be a reduction in the level of resources in the area within the cluster, followed by a slow replacement in the population constituted by bad communicators. As the cost of producing an offspring is constant for all agents, the increase in the rate of energy intake will translate in an increase in agent density because the new population will tend to reproduce faster. Consequently the level of resources will be further decremented. At this point it is both possible and necessary for the cluster to grow. Density increase and energy shortage will also be manifested near the periphery, meaning that the "force" that impeded cluster growth will decrease (this tendency to collapse was a consequence or the need of agents to have a good number of neighbors to play with), so momentarily the tendency towards expansion, itself incremented by the shortage of food, will prevail, and the cluster will grow in size until a new equilibrium is reached (see Figure 14).

These explanations are in fact simplifications of very complex dynamical processes in which more ingredients than those mentioned may play an important part. For instance, all the cases discussed above follow a quasi-static approach, in which inertia has been unaccounted for. The complex effects arising from cluster interaction have also been ignored. The reason for this is that, interesting as these phenomena may be, they do not much further our understanding of how communication evolves due to spatial organization and selection in this model. The relations between the evolution of communication in this model and the evolution of communication in the real world will be discussed in Section 7.

6.4 Coordinated activity with "information" sharing.

Using the same mechanics, the following variation of the original game was implemented. Instead of requiring an unique action, access to energy in the food sources requires the performance of a specific sequence of actions by both agents. Energy is released partially depending on an action being correct at the required step of the sequence.

In this case the actions of *both* players depend on the perceived food type and the perceived external manifestation of the other agent in the previous step. This means that the second agent also has access to information about the food type. If, for instance, food type '2' requires the sequence of actions "A, B, C, D" the following would be an example of two agents successfully coordinating their behaviors in this game:

$$\begin{array}{llcl}
\text{Time} & & \text{Actions(``2'')} & \\
 & A_1(\text{``2''},\emptyset) \rightarrow (A,\alpha) & \text{A} & \\
\downarrow & & \text{B} & (B,\beta) \leftarrow A_2(\text{``2''},\alpha) \\
 & A_1(\text{``2''},\beta) \rightarrow (C,\delta) & \text{C} & \\
 & & \text{D} & (D,\gamma) \leftarrow A_2(\text{``2''},\delta)
\end{array}$$

where A_i is the behavioral operator of each agent, the result of which is a pair (effective component, external manifestation); and $\emptyset$ is the null manifestation.

This game has a more dialogic structure and it resembles the kind of interaction described by von Foerster's eigen-behaviors (Section 3.3). The problem is somehow more complicated than the simple action-response game, because not only must the external manifestations of one agent and effective components of the other become correlated, but also the external manifestations of both of them through the whole sequence. Not only must the second player rely on the signal of the first player to act correctly, but this signal must also select in the second player another signal that in turn will select the correct behavior in the first player for the next step in the game. We see that this is a recursive task much more in accordance with the nature of communication that I have defended in this work. We also see that a description of this task in terms of traditional notions of information is useless. If by information we mean information about features of the environment, we find that these are equally accessible to both participants, if we mean information about the state/intention of the agents, they haven't got any[22].

A simulation was run with four different types of food, two of them requiring a sequence of actions such as "A, B, C, D" where the first player must perform "A, C" alternating with the second player who must perform "B, D" and the two others requiring the sequence "C, D, A, B" which means that each agent must revert the order of its own actions. Payoffs are allocated after the first two actions, and then again after the last two actions, in the same manner as described in the previous game.

---

[22] By contrast, the Petri-Holt concept of information transfer proposed by Pask fits into the nature of this game.

Figure 15 shows the evolution of the average instantaneous success in dialog over time for two typical runs with different values of c (variations between runs with the same set of parameters were not qualitatively significant). This quantity is similar to *R12* with the difference that it is defined as the proportion of cases in which the *whole* sequence of actions is performed correctly in a given interaction.

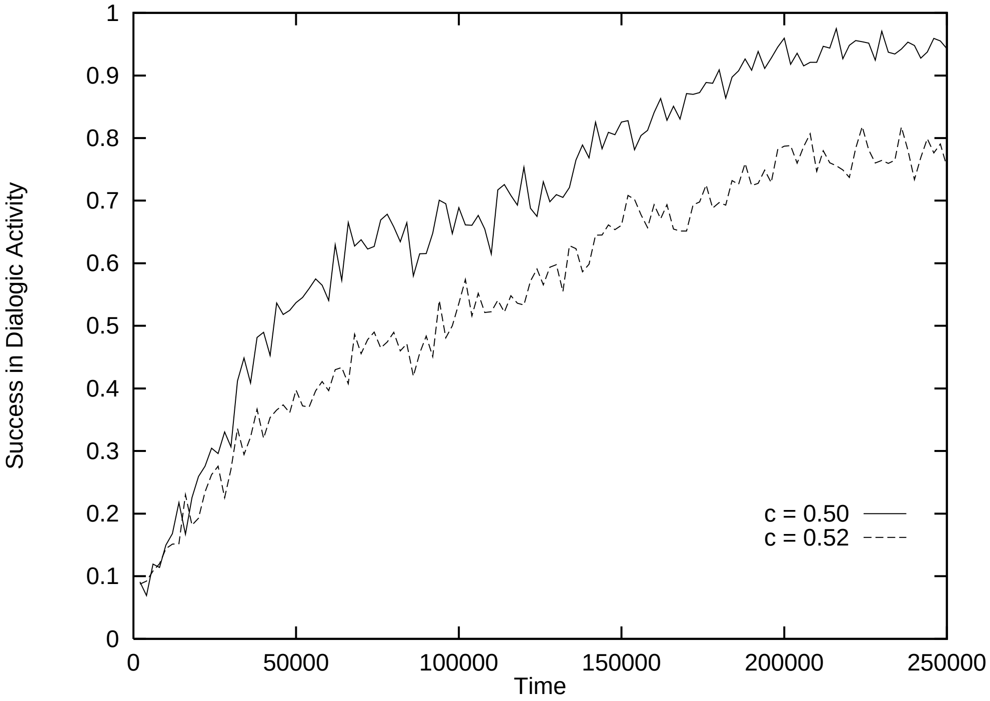


Figure 15: Evolution of dialogic coordinated activity for two values of c in two typical runs.

Similar considerations mentioned for the previous game apply here as well. The level of dialogic success depends on the parameter c in a similar fashion as the level of coordinated activity did for the previous game; agents also form clusters and these clusters can differ in their levels of dialogic success.

We may observe the resulting behaviors by analysing the evolved behavioral structures. For example, an evolved behavioral matrix within a highly successful cluster (above 90% of dialogic activity) looks likes this:

	Perceived Signal			
Food Type	α	β	δ	γ
"0"	A γ	C α	B γ	C δ
"1"	A δ	D γ	C γ	B β
"2"	B β	D δ	A γ	C γ
"3"	B δ	A γ	D α	C α

Here, agents must produce the sequence A, B, C, D in the presence of food types "0" and "1" and the sequence B, A, D, C in the presence of food types "2" and "3". The column corresponding to the signal α encodes the first action of the first agent ($\alpha = \emptyset$ in this case). As a contrast with the previous game, in this case the whole matrix is used and not just one column and one row. Assuming that this particular agent belongs to a rather homogeneous cluster, as it does, it is easy to see that, when interacting with a behaviorally identical agent, it will achieve a level of dialogic coordination of 93.75% (just by counting for each food type the success levels that this agent would achieve in a dialog with itself). The actual level achieved by this agent was in fact 96.2%, probably because other agents in its neighborhood are slightly different, or because there is a relatively smaller number of food sources of type "0" surrounding it (success when the food is of type "0" is only 75%, note that the effective component D is missing in the corresponding row). If the dialog is continued for further steps than those required by the game, we see how very soon agents arrive at their eigen-behaviors, which, in fact, are not very meaningful, because all the payoffs were assigned during the "transient" stage. In fact, it could be said that, as dynamical systems, agents avoid falling into simple eigen-behaviors too quickly, in order to be able to produce a four-step sequence during the transient period.

These results show how agents coordinating their activity are able to perform tasks beyond their individual abilities. This particular game requires that both agents perform a sequence of behaviors in the presence of an external environmental feature that remains unchanged with time. Given that agents are state-less machines, this is something impossible for them to do individually. However, pairs of interacting agent can achieve an important level of success in this task, each one taking advantage of the presence of the other "using" them as the internal states they lack. An analysis of the evolution of these behaviors is more complicated than in the previous game, because the notions of cost and benefit of a specific action become even more diluted and context-dependent. It is my opinion that in these circumstances, the best framework from which to view this game, and more complex versions of it, will be the extension of the communication concept into that of "languaging" as presented in autopoietic theory (Maturana, 1978). This is intended to be done in a future approach to this problem. However, as I will discuss below, the results of this game still provide an interesting metaphor for hypotheses that relate the evolution of cognition with the evolution of communicative behaviors and sociality.

7 Conclusions.

The objective of this work has been to support the thesis presented in the Introduction, which stated that the many focuses on the subject of communication do have a basic common denominator, and that biology was in a privileged position to provide an understanding of such a unifying theme by studying the evolution of communicative behaviors in animal societies but, that in order to do so, biologists had to transform the way they characterize the phenomenon itself. Such re-characterization was performed and justified by the adoption of a conceptual framework underlying many disciplinary views on the subject, and it was shown that it was possible to work with it.

Based upon this framework a computational approach was methodologically justified and implemented in order to study the evolution of communication as coordinated activity. Concerning the explanations provided in this work, it has been my implicit assumption throughout this paper that the origins and evolution of communication belong to that realm of phenomena in which no single explanatory mechanism will be sufficient if we want to obtain an understanding of their nature. In order to grasp some aspects of that nature it is necessary to accept the existence of a network of many interacting, sometimes conflicting, sometimes cooperating, processes underlying the phenomenon and, after that, proceed with the addition of hypotheses and simplifying assumptions.

Traditional approaches to the evolution of communication have shown a tendency to apply the simplifying step right at the beginning, proposing a single explanatory principle, natural selection, even a the definition level, and consequently, trying to understand the actions of this principle almost in a decontextualized manner.

This has been the reason for criticizing the "beneficial-information-exchange" approach to communication; because it is necessary, in order to view the phenomenon in its whole complexity, to deconstruct the simplifying steps of traditional approaches and rebuild our understanding of what constitutes communication and what does not on a more solid basis. Fortunately, autopoietic theory provided both a good candidate and a concise language for this task. Thus, communication was defined as the coordination of the actions of more than one organisms that arise during their structural coupling in a consensual domain, a description that parallels many other approaches to communication and language in cybernetics, psychology and philosophy. Also, a definition one can work with.

However, my critical stance does not compel me to blindly ignore the technical and methodological aspects of previous work in the area. This is why my first step was the choice of a game similar to those proposed in the theoretical biology literature on the evolution of communication, in which a common language in terms of costs, payoffs, etc., provided a continuity with these and the present works. Also, I did not ignore the tools, such as game-theory, used in these traditional approaches to understand conditions that allow equilibrium situations, but rather showed that the benefit of their use was limited even when modifications were introduced in order to account for dynamical effects as well. In this way the need for a computational approach was justified.

The computational model has demonstrated that coordinated activity can evolve even in those cases in which the static and dynamic mathematical models showed it would not. And the reason for this difference has been mainly the possibility that the computational model can provide to study the actions of natural selection in the context of other concurrent processes such as spatial organization. Of course, nothing can stop us from changing the roles as to which mechanism should be viewed as the explanatory principle and which others as the context, except for questions of clarity. All we can really say is that evolution of communication in this model, can only be accounted for as the result of more than one interacting factor, and that such interaction can be explored and grasped, at least partially, thanks to the flexibility of the computational approach. The fact that these many concurrent processes influence each other is not enough to say that one is the cause of the other, thus it can be claimed that spatial organization is influenced by selection, and vice versa, but these are, in fact, different processes and not necessarily different aspects of a more fundamental principle. Sometimes the mesh of influences can be so intricate, that very complex self-regulating structures emerge. Clusters of interacting agents have shown to be examples of such structures.

It is not a bold claim to say that the existence and importance of many causal factors underlying the evolution of communication in the computational model, which is itself a simplification of a much more complex real process, implies that the evolution of real communication in animal societies has also been the result of the interdependent activity of many mechanisms. It could be argued that this is all that this model has to say about the evolution of real communication. However, this model also suggest specifically which factors could have had an important influence in the real process. One such factor is, undoubtedly, *spatiality*.

An interesting case, in which the researchers arrive at very similar conclusions in another context, can be found in (Boerlijst & Hogeweg, 1991). This work looks at the importance of spatial self-structuring occurring in the pre-biotic evolution of catalytic cyclical ensembles ("hypercycles"). Mathematical models of these hypercycles are subject to the criticism of being unstable against the introduction of parasites which take advantage of cooperative interactions without being cooperative themselves. Vulnerability to such parasites is in fact a consequence of studying the problem in an totally mixed medium (no spatiality). Once a spatial model is introduced, dynamic structures emerge which prove to be very robust against parasitic invasion. It seems that spatial organization can provide a very strong influence on the outcome of an evolutionary process, resulting sometimes in far from intuitive characteristics.

There is not much novelty in stating that finite, bounded populations of interacting organisms are subject to a set of environmental features which is a direct consequence of their spatial and social organization, and which can be very different from the kind of environments found in first-order approaches to the problem. To this day, however, research along these lines, has been rather scarce.

A broad analogy can be drawn between the position of an agent in a cluster in our model, and the degree of participation in a social organization for a natural organism. For instance, animals spending more time in social activities as a consequence of their developmental stage, or their social status will be analogous to agents living in the center of a cluster. Such an analogy would suggest that many answers to the question of natural communication could be sought in the nature of group structure, or social hierarchies.

Interestingly, one of the correlations found in this study, the correspondence of cluster size and degree of coordinated activity, has a very suggestive parallel in primate societies: that of typical group size and relative neocortical size (Sawaguchi & Kudo, 1990; Dunbar, 1992). It has been suggested (Dunbar, 1993) that language evolution has been deeply influenced by the growth in group size in humans as compared with other primates, and a functionalist explanation was advanced in terms of the role of language as a bonding mechanism in the maintenance of stability in large groups. The way that the analogous correlation was explained in the computational study has been rather different, more in terms of structural constraints rather than functional adaptations, but the parallel is worth noting.

I have also mentioned as one of the key features of the autopoietic view on communication its natural extendibility into the domain of human language. In that area we find that dialogic activity such as the game of Section 6.4 brings into question a very important aspect of communication as coordination of actions that has not been explored in detail in this paper: that of the relation of the evolution of communication and the evolution of cognitive abilities. Much can be said about this, especially in the case of human communication, but without a good model of ontogeny, the present work cannot penetrate very far into this area. I consider these experiments as a metaphor that suggests a possible way of understanding the evolution of cognition as rooted in social interactions (see for example Humphrey, 1976). If social coordinated activity leads to more complex cognitive achievements, as we saw they can, this opens two possible paths for cognitive evolution. One of them leads roughly to highly structured societies of organisms in which the complexity of the social action involves the whole in such a way that individual behaviors have become more and more specialized, rigid and subordinated to it. And the other leads to a process of plastic and/or phylogenic assimilation of certain social modes of behavior in individual organisms possibly leading in turn to higher complexity in individual and social action. Examples of the first case are insect societies, and of the second case, some mammal societies. This later path was very much discussed by Baldwin as a case of social heredity in gregarious animals (Baldwin, 1896). Interestingly enough, in the case of humans, this is how Vygotsky accounts for experimental evidence in the development of higher cognitive functions in the child (Vygotsky, 1988)[23].

Evidently, any attempt at understanding the nature of such a process will have to take into account many of the features that were left out of our model, such as agent's autonomy, development and plasticity. It is to be expected, that the interactions between these and other factors will render the dynamics of the process almost unintelligible in their totality, even if they could be reproduced in a simulated computer environment. This doesn't mean that we will be unable to answer *some* questions about this process by performing such simulations, and hope to find the guidelines for answering still others. As I said, the methodological issue in the performance of such a task is, given our current state of affairs, almost as important as the task itself, so we cannot afford to ignore it if we want this kind of computational approach to be as respected as other, already established, scientific practices.

8 Acknowledgements

Thanks to Phil Husbands and Inman Harvey for their comments on this work. The author is grateful to the *Consejo de Investigaciones Científicas y Técnicas de la República Argentina* and the Argentine Ministry of Education for their support.

Reference

Ashby, W. R. (1960). *Design for a Brain: The origin of adaptive bheaviour* (2nd edition). Chapman and Hall, London.

Austin, J. L. (1962). *How to Do Things With Words. The William James Lectures delivered at Harvard University in 1955.* Oxford at the Claredon Press.

Austin, J. L. (1971). Performative - constative. In Searle, J. (Ed.), *The Philosophy of Language*, pp. 13–22. Oxford University Press.

Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, *30*, 441 – 451.

Bateson, G. (1972). *Steps towards an Ecology of Mind: collected essays in anthropology, psychiatry, evolution and epistemology.* Intertext Aylesbure.

---

[23] Baldwin speaks a bit loosely of imitation while Vygotsky speaks of internalization of social mediated activity as the mechanisms underlying this process, but the similarities are stronger than the differences.

Bickerton, D. (1995). *Language and Human Behaviour.* University College London Press.

Boerlijst, M. C., & Hogeweg, P. (1991). Spiral wave structure in pre-biotic evolution: hypercycles stable against parasites. *Physica D*, *48*, 17 – 28.

Burghardt, G. M. (1970). Defining 'communication'. In Johnson, Jr, J. W., Moulton, D. G., & Turk, A. (Eds.), *Communication and Chemical Signals.* Appleton-Century-Crofts, New York.

Darwin, C. (1902). *The origins of species by means of natural selection.* Murray.

Dautenhahn, K. (1995). Getting to know each other - artificial social intellingence for autonomous robots. *Robots and Autonomous Systems*, *16*, 333 – 356.

Dewey, J. (1958). *Experience and Nature.* Dover.

Di Paolo, E. A. (1996a). A computational model of speciation in non-uniform environments without physical barriers. Cognitive science research paper 412, School of Cognitive and Computing Sciences, University of Sussex.

Di Paolo, E. A. (1996b). Some false starts in the construction of a research methodology for Artificial Life. In Noble, J., & Parsowith, S. (Eds.), *The 9th White House Papers: Graduate Research in Cognitive Science and Computing Sciences at Sussex.*

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *20*, 469 – 493.

Dunbar, R. I. M. (1993). Co-evolution of neocortex size, group size and language in humans. *Brain and Behavioral Sciences*, *16*, 681 – 735.

Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology*, *144*, 517 – 546.

Habermas, J. (1979). *Communication and the Evolution of Society*, chap. "What is Universal Pragmatics?". Heinemann, London.

Heidegger, M. (1962). *Being and Time.* Harper and Row, New York. Trans. J. Macquarrie and E. Robinson.

Holt, A. W. (1972). Comments. In Bateson, M. C. (Ed.), *Our own metaphor.* A. Knopf.

Humphrey, N. K. (1976). The social function of intellect. In Bateson, P. P. G. (Ed.), *Growing points in Ethology*, pp. 303–317. Cambridge University Press.

Hurd, P. L. (1995). Communication in discrete action-response games. *Journal of Theoretical Biology*, *174*, 217 – 222.

Krebs, J. R., & Davies, N. (1993). *An Introduction to Behavioural Ecology* (3rd edition). Blackwell, Oxford.

Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind reading and manipulation. In Krebs, J. R., & Davies, N. B. (Eds.), *Behavioural Ecology: An Evolutionary Approach* (2nd edition)., chap. 15, pp. 380–402. Blcakwell, Oxford.

Lewis, D. B., & Gower, M. D. (1980). *Biology of Communication*. Blackie, Glasgow.

Lewontin, R. (1982). Organism and environment. In Plotkin, H. C. (Ed.), *Learning, development and culture: essays in evolutionary epistemology*. Wiley Chichester.

Lewontin, R. C. (1983). The organism as the subject and object of evolution. *Scientia*, *118*, 63 – 82.

Lewontin, R. C. (1984). Adaptation. In Sober, E. (Ed.), *Conceptual issues in evolutionary biology: an anthology*. The MIT Press, Cambridge Mass.

MacLennan, B. J., & Burghardt, G. M. (1994). Synthetic ecology and the evolution of cooperative communication. *Adaptive behabior*, *2*(2), 151–188.

Maturana, H. (1978). Biology of language: The epistemology of reality. In Miller, G. A., & Lennenberg, E. (Eds.), *Psychology and Biology of Language and Thought. Essays in Honor of Eric Lennenberg*, pp. 27–63. Academic Press Inc. New York.

Maturana, H., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel Publishing, Dordrecht, Holland.

Maturana, H., & Varela, F. J. (1988). *The tree of knowledge: the biological roots of human understanding*. Shambhala, Boston, Mass.

Maynard-Smith, J. (1978). Optimization theory in evolution. *Annual Review of Ecology and Systematics*, *9*, 31 – 56.

Maynard-Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.

Maynard-Smith, J., & Harper, D. G. C. (1995). Animal signals: models and terminology. *Journal of Theoretical Biology*, *177*, 305–311.

Miller, G. E. (1995). Alife as theoretical biology: how to do science with computer simulation. Cognitive science research paper 378, School of Cognitive and Computing Sciences, University of Sussex.

Orr, H. A. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, *139*, 1805 – 1813.

Oyama, S. (1985). *The Ontogeny of Information*. Combridge University Press.

Pask, G. (1976). Conversation techniques in the study and practice of education. *British Journal of Educational Psychology*, *46*, 12 – 25.

Pask, G. (1980). Developments in conversation theory - part 1. *International Journal of Man-Machine Studies*, *13*, 357 – 411.

Repka, J., & Gross, M. R. (1995). The evolutionarily stable strategy under individual condition and tactic frequency. *Journal of Theoretical Biology*, *176*, 27 – 31.

Rorty, R. (1991). *Essays on Heidegger and others: Philosophical Papers Vol. 2*, chap. "Wittgenstein, Heidegger and the Reification of Language". Cambridge University Press.

Sawaguchi, T., & Kudo, H. (1990). Neocortical development and social structure in primates. *Primates*, *31*, 283 – 290.

Shannon, C., & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.

Steels, L. (1997). Synthesising the origins of language and meaning using co-evolution, self-organization, and level formation.. In Hurford, J. (Ed.), *Evolution of Human Language*. Edinburgh University Press.

Stolzenberg, G. (1984). Can an inquiry into the foundations of mathematics tell us anything about mind?. In Watzlawick, P. (Ed.), *The Invented Reality*, pp. 257–308. W.W. Norton.

Varela, F. J. (1979). *Principles of Biological Autonomy*. Elsevier, North Holland, NY.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. The MIT Press.

von Foerster, H. (1977). Objects: tokens for eigenbehavior. In Inhelder, B. (Ed.), *Hommage à Jean Piaget*. Delachaux et Niestel, Neuchatel.

von Foerster, H. (1980). The epistemology of communication. In Woodward, K. (Ed.), *The Myths of Information: Technology and Postindustrial Culture*. Routledge.

von Glasersfeld, E. (1984). An introduction to radical constructivism. In Watzlawick, P. (Ed.), *The Invented Reality*. W. W. Norton, New York.

von Glasersfeld, E. (1988). The reluctance to change a way of thinking. *The Irish Journal of Psychology*, *9*, 83 –90.

Vygotsky, L. S. (1988). *Thought and Language*. The MIT Press, Cambridge, Mass.

Watzlawick, P., & Beavin, J. (1977). Some formal aspects of communication. In Watzlawick, P., & Weakland, J. H. (Eds.), *The Interactional View: Studies at the Mental Research Institute Palo Alto, 1965-1974*. W. W. Norton.

Watzlawick, P., Beavin, J. H., & Jackson, D. D. (1968). *Pragmatics of Human Communication: a Study of Interactional Patterns, Pathologies and Paradoxes*. Faber.

Werner, G. M., & Dyer, M. G. (1990). Evolution of communication in artificial organisms. Technical report UCLA-AI-90-06, UCLA.

Wertsch, J. V. (1991). *Voices of the Mind: a Sociocultural Approach to Mediated Action*. Harvester Wheatsheaf.

Wilson, E. O. (1975). *Sociobiology: The new synthesis*. Belknap Press of Harvard University Press, Cambridge, Mass.

Winograd, T., & Flores, F. (1986). *Understanding Computers and Cognition.* Norwood NJ, Ablex.

Wittgenstein, L. (1963). *Philosophical Investigations* (2nd edition). Oxford: Basil Blackwell and Mott.

Zahavi, A. (1975). Mate selection - a selection for handicap. *Journal of Theoretical Biology*, *53*, 205–214.

Zahavi, A. (1977). The cost of honesty (further remarks on the handicap principle). *Journal of Theoretical Biology*, *67*, 603–605.

Zeeman, E. C. (1980). Population dynamics for game theory. In Nitecki, Z., & Robinson, E. (Eds.), *Global Theory of Dynamical Systems.* Springer Verlag, Berlin.