HTTP://WWW.AGRIC-ECON.UNI-KIEL.DE/ABTEILUNGEN/II/VEROEFFENTLICHUNGEN.HTM

ARTIFICIAL EMERGENCE OF SPONTANEOUS INSTITUTIONS: A SURVEY OF RECENT RESEARCH[1]

Contributed paper: VIII EAAE Congress, Edinburgh, 3.-7. September 1996

Rolf A.E. Mueller (Department of Agricultural Economics, University at Kiel)[2]

ABSTRACT

To understand the development and to contribute to the design of institutions conducive for sustainable agricultural resource use are important tasks for agricultural economists. In this paper an innovative approach to the study of spontaneous institutions is reviewed and its use for generating invisible-hand explanations of the emergence of spontaneous institutions is discussed. The approach consists of modelling the interactions of economic agents as Prisoner's Dilemma games and of tracing in computer simulations the behaviour patterns of populations of economic agents over many generations in complex and often co-evolutionary environments. For many environments or scenarios such simulati[3]ons trace processes that lead to robust cooperative behaviour constituting a social institution.

RÉSUMÉ

Comprendre l'évolution des institutions promouvant un usage à long terme des ressources agricoles et contribuer à leur mise en place sont devenues des tâches importantes pour les économistes agraires. Dans cet article, nous passons en revue une approche innovante des institutions spontanées et argumentons de sa capacité à fournir des explications implicites à l'émergence d'institutions spontanées. Cette approche consiste à représenter les interactions des agents économiques par le dilemne du prisonnier et à simuler par ordinateur les modes de comportement de populations d'agents économiques sur plusieurs générations au sein d'environnements complexes et souvent co-évolutifs. Pour de nombreux environnements ou scenari, ces simulations décrivent des évolutions menant á des modes distincts de comportement social.

1. Introduction

The capacity of mankind to propagate is developed to an extent where it might exceed man's capacity to adapt to an impoverished natural environment (Arrow et al. 1995). If nature was the only environment in which farmers make their food production decisions, then, as Malthus had feared, population growth would indeed outpace growth of food production. Also part of the food production environment are, however,

---

[1] Paper presented at the VIII Congress of the European Association of Agricultural Economists, 3-7 September 1996, Edinburgh, Scotland.

[2] This paper has benefitted considerably from comments by Wilhelm Brandes and Brian Hardaker.

complex artificial systems that have been created by man or which are the consequences of human action. Two important subsystems are the technology system that provides new, more productive agricultural technologies, and the social system of rules governing the interactions among resource users, such as tenure laws or customary grazing rights, that facilitate cooperation among resource users.

Population growth and some new agricultural technologies have contributed to an increased demand from agriculture for scarce natural resources and signs of serious resource degradation abound in rural areas the world over. In agricultural economics research these changes have led to an increased interest in the interactions among farmers who compete for scarce natural resources and the impact on natural resources of these interactions. In all societies are the interactions among people governed by systems of rules, of which coordination through the authority of government or market competition have received the most research attention. This self-imposed limitation of the scope of research interest left many alternative systems of rules that govern the use of natural resources unexplored and only recently have economists shown a revived interest in other institutions than government or the market. Furthermore, the preoccupation with government interventions and markets as coordination mechanisms has stifled the contribution economists might otherwise have made to institutional innovation and to default on their duty "... to model structures of interaction 'that might be' " (Buchanan 1992, p. 150).

If an increasing number of agricultural economists accept as their object of study the rules governing the interactions among the users of agricultural resources, the demand for innovative research approaches and methods will increase. This paper is an attempt to cater for such prospective demand. In particular, suitable research methods from Artificial Life will be presented together with some relevant applications to the study of conditions conducive for cooperation. Artificial Life is a new interdisciplinary research area which accommodates researchers from such diverse fields as biology, computer science, physics, economics, mathematics and political science. In agricultural research circles, however, there are few signs of artificial life and one purpose of this paper is to draw the attention of agricultural resource economists to this new research area. The other purpose of the paper is to remind ourselves of an important tradition in institutional economics, a tradition which considered institutions neither as artefacts that have been created for a particular purpose nor as natural things that can have an existence without the actions of people. Rather, this tradition or paradigm regards institutions as the results of human action but not of human design. According to Hayek (1967), the intellectual roots of this tradition can be traced back to Adam Ferguson in the 18th century. The paradigm is particularly suited to provide the backdrop for the methods from Artificial Life that offer themselves for the study of institutions.

2. Institutions

2.1. What are institutions?

Like many terms that are used in everyday life as well as in the sciences, the use of the term "institutions" is not uniform and some explication of how the term is to be used in the context of this paper is necessary. Most students of economic institutions can probably agree with Bardhan (1989b, p. 3) who regards 'institutions' as a collective term for "... the social rules, conventions, and other elements of the structural framework of social interaction." or with North's (1990, p. 4) characterisation of institutions as being "... perfectly analogous to the rules of the game in competitive team sport." Further, few are likely to object against the characterisation of the functions of institutions, which, according to North (1981, p. 201) "... provide the framework within which human beings interact. They establish the cooperative and competitive relationships which constitute a society and more specifically an economic order." A more precise definition of institutions has been given by Schotter (1981, p.11) who defined a social institution as "... a regularity in social behaviour that is agreed to by all members of society, specifies behaviour in specific recurrent situations, and is either self-policed or policed by some external authority."

Institutions may be categorised along several dimensions. For one, because institutions are "... humanly devised constraints that shape human interaction" (North 1990, p. 3) people may not voluntarily adhere to them and some institutions have to be enforced through an external authority, typically the state. An example for institutions in need of external enforcement are rules of property which prevent thieves and the government from realising their ideas of a just distribution of wealth. Other institutions are self-enforcing, as is the case with the convention to stay on the right when driving on the continent.

Further, a distinction is frequently made between formal institutions, such as positive law and codified commandments, and informal ones, such as conventions or codes of conduct. Akin to informal institutions are Arrow's (1974) "invisible institutions" such as the principles of morality and ethics. Lack of formality is, however, unrelated to the importance of informal institutions and North (1990, p. 138) concedes that "... they appear to have a pervasive influence on the institutional structure."

Without question the most important dimension for categorising institutions is the way in which institutions have come into existence. This dimension has been introduced by Menger more than one hundred years ago. Menger (1985, p. 133) distinguishes between "... some social phenomena [that] are the results of a *common will* directed toward their establishment ... while others are the unintended result of human efforts aimed at attaining essentially *individual goals* (the unintended results of these)." The distinction has also been adopted by North (1990, p.4), who realises that institutions "...may be created, as was the United States Constitution, or they may simply evolve over time, as does the common law." The most elaborate and detailed discussion of the distinction has, however, been provided by Hayek (1967) who distinguished three classes of phenomena: (i) natural phenomena that are wholly independent of human action; (ii) artificial things, that are the product of human design and (iii) a middle category of unintended patterns and regularities that are the results of human action but not of human design. It is difficult to speak about something which does not have a proper name. For Menger (1985), there was no difference between "organic" and "unintended" social phenomena. Because "organic" has become a frequently used attribute of breakfast cereals and refuse, it is unlikely to gain much currency in the context of things social. The attribute "unintended" has the disadvantage that it can only be inferred but not observed. Calling something a "middle category" is uninformative unless it is made clear what the extreme categories are. Ullmann-Margalit (1978) refers to phenomena from the "middle category" as "social" phenomena. However, the term "social" is already loaded with so many connotations that it conveys little information. For lack of a better alternative I shall use the term "spontaneous" for those institutions that are the results of human action but not of human design, and I distinguish them from designed institutions that are the results of action designed to create the institution in question.

2.2. Approaches to the study of institutions

The approaches to the study of institutions, including the institutions governing the use of agricultural resources, are many. Only a small subset will be sketched here to provide a perspective for the characterisation of the computer methods that have been employed to simulate the emergence of spontaneous institutions.

2.2.1. Synthesis of design heuristics

In the past, when government and the market were the coordination mechanisms that attracted the most research attention, and when institutions governing agricultural resource use were considered as being exogenous to the economy, progress in the knowledge about such institutions was slow. As a consequence, there is still much to be learned from empirical experiences with particular institutional arrangements. The synthesis of an inventory of empirical accounts of institutional arrangement may then yield valuable insights and heuristics useful for guiding resource policies. A recent example for this approach is Ostrom et al. (1994) who distilled several design heuristics for common property institutions.

2.2.2. Theories of designed institutions

Bardhan (1989a) identified two approaches to the explanation of designed institutions: the imperfect information theory of institutions, that emphasises the information problems in contracting, and the transaction cost school of economic institutions, that is represented by Coase, North, Williamson, and others. Typically, explanation of designed institutions are based on the identification of whatever interests of the designer of the institution are served by the institution. These interests then explain why the institution was implemented. For example, the drive to institute farmers' rights to crop genetic resources can be interpreted as an attempt by the promoters of such rights to secure for themselves or their clients rents that are at present left in the public domain. Similarly, rules regulating the spacing of wells or the capacity of irrigation pumps may be explained in terms of the social losses that are avoided from reducing the overexploitation of fugitive groundwater resources.

2.2.3. Explanation of spontaneous institutions

The question that Menger (1985, p. 146) considered as "...a noteworthy, perhaps the most noteworthy, problem of the social sciences [is]: "*How can it be that institutions which serve the common welfare and are extremely significant for its development come into being without a* ***common will*** *directed toward establishing them?"* The question may be answered by means of invisible-hand explanations. Ullmann-Margalit (1978) has identified five key characteristics of invisible-hand explanations:

- Invisible-hand explanations typically displace initially plausible intentional design explanations by providing an account of the spontaneous institutions in question as the outcome of "... a process involving the separate actions of many individuals who are supposed to be minding their own business unaware of and a fortiori not intending to produce the ultimate overall outcome" (Ullmann-Margalit 1978, p. 267).
- The explanation invokes an invisible-hand process that has the independent and disperse actions of the individuals as its input; the output of this process then is the regularity in aggregate behaviour that is to be explained.
- The initial stage of the invisible-hand process can be completely described in terms of the characteristics and circumstances of the individuals participating in the process.
- The individuals whose actions fuel the invisible-hand process are assumed neither to foresee nor to intend to bring about the aggregate results of their actions.
- Invisible-hand processes are typically represented as stories that provide a rational reconstruction of how the spontaneous institution could have emerged; they do not have to be factually accurate accounts of how the institution in question actually has emerged.

Invisible-hand stories have been told using different rhetorical styles. Some, such as Adam Smith (1937) or Hayek (1967) use plain language. Recently, however, invisible-hand accounts are often presented in the form of evolutionary games and told in a language enriched by mathematical notation and phase diagrams (e.g. Sugden 1986). The computer simulations for the study of the emergence of spontaneous institutions, with which we will be concerned in the remainder of this paper, can be regarded as a variant of evolutionary game theory where numerical solutions generated with computers replace the analytical procedures of evolutionary game theory. This characterisation misses, however, an important point. The new evolutionary methods are more than a new way of telling in great numerical detail old invisible-hand stories. As we shall see, in addition to tracing invisible-hand processes, evolutionary methods can also be used to invent new invisible-hand stories and lend themselves as a means for inventing institutions 'that might be'.

3. Artificial Life and its methods

3.1. What is "Artificial Life"?

The term Artificial Life (AL or ALife) was coined by C.G. Langton of the Santa Fe Institute for an area of research concerned with questions and phenomena that tend to belong to the union of biology, computer science, and complex adaptive systems. Although the term eludes definition, a popular account describes this new area of research as being "devoted to the creation and study of lifelike organisms and systems built by humans" and which hopes to create life *in silicio* (Levy 1993, p. 5). ALife has spilled over into the social sciences and researchers have begun to attempt the simulate the behaviour of societies on their computers (Flam 1994).

ALife may be characterised in terms of the questions addressed, its communication infrastructure, and its research methods. The questions addressed by AL range from questions that have their origin in the biological sciences (e.g. questions about the nature of life), to questions typical for computer sciences, such as questions about machine learning and the adaptability of computer programs. As increasing numbers of researchers from conventional disciplines migrate into the new research territory, the set of questions addressed in AL is, however, still in a state of flux. The communication infrastructure of the field comprises a considerable number of ALife-conferences that have been held, several books and specialised scientific journals, and, which is perhaps the most telling indicator of a viable communication infrastructure, ALife sites have sprung up in the bitsphere of the World Wide Web.

3.2. Artificial life methods

As indicated by the term 'artificial', research in the ALife area is synthetic and usually relies on computer simulations for the study of life and life-like phenomena, in particular the study of evolutionary processes that may involve self-replication, mutation, competition, and selection, as well as the study of the emergence of spontaneous order in decentralised systems through self-organisation. Typically, ALife-simulations employ dynamic and adaptive models of local interactions amongst large numbers of artificial agents of low complexity to investigate and explore the aggregate behaviour of complex systems.

Research methods employed in AL-simulations are (i) cellular automata, (ii) evolutionary computation methods, such as evolutionary strategies, evolutionary programming, genetic algorithms, and genetic programming, (iii) classifier systems, and (iv) autonomous adaptive agents. Some authors (e.g. Gutowitz 1995) also regard artificial neural nets as an AL-method. Until now, only cellular automata, evolutionary programming and genetic algorithms have been employed to provide invisible-hand accounts of the spontaneous emergence of institutions and need to be described here.

3.2.1. Cellular automata

Cellular automata, which have been invented by von Neumann (1963), are a class of discrete and dynamic mathematical systems that are composed of simple elements but which are able to exhibit complicated aggregate behaviour (Wolfram 1984). They can be regarded as being 'ALife' because complex behaviour emerges from the local interactions of identical structures, and because a cellular automaton simulates reproduction if some of its cells (the "living" cells) can spawn new cells. A cellular automaton consists of a grid or lattice with identical cells at the interstices. The cells can assume a finite number of states. In the simplest case a cell may be in one of two states, indicating whether a cell is dead or alive. The cells may change their states from one period or generation to the next according to transition rules that determine the state of a cell in the next generation on the basis of its own present state and the present states of some or all of its immediate neighbours in the present generation. All cells are governed by the same transition rules and all cells update their states in synchrony so that the entire cellular automaton changes its appearance in discrete steps. In more complex cellular automata the cells in the grid may also be updated asynchronously (Lindgren and Nordahl 1994) or transition rules may evolve rather than remain fixed (Sipper 1994). Gerhardt and Schuster (1995) report several applications of cellular automata in various sciences, such as the study of the behaviour of ideal gases in physics, the explanation of colour patterns of animals in biology, or simulating the diffusion of a pest in a corral reef in ecology. In agricultural economics a cellular automata approach has been used by Balmann (1994) to simulate land allocation and change in regional farm size distribution.

3.2.2. Evolutionary programming

Evolutionary programming involves the creation of a random population of finite state machines that transform a sequence of input symbols into sequence of output symbols. Such machines are then exposed to an environment consisting of a sequence of symbols and their performance in predicting as yet unobserved symbols is assessed. A payoff or fitness is determined on the basis of prediction performance. Offspring machines are created by randomly mutating parent machines, which are the machines that scored the highest fitness. Mutation involves the change of one of the properties defining a finite state machine, i.e. output symbols, states, and state transition (Fogel 1995). Examples for applications of evolutionary programming are solving travelling salesmen problems or designing neural networks for pattern recognition and control systems (Fogel and Fogel 1996).

3.2.3. Genetic algorithms

Genetic algorithms (GA) are probably the most mature AL-method and they have become part and parcel of the optimisation tool box used at graduate schools of engineering (Lloyd 1996). Their applications defy enumeration and involve, besides solving hard optimisation problems, a variety of things such as engineering design, pattern recognition, financial forecasting, or ecological modelling (Goldberg 1994; Forrest 1993). The basic components of a genetic algorithms are a population of "chromosomes", i.e. bit strings representing problem solutions, and the operators of selection according to fitness, reproduction with crossover, and random mutation of new offspring (Forrest 1993; Mitchell 1996). The initial population of a simple genetic algorithm consists of a population of random bit strings of fixed length created by the computer. Each bit string represents a solution to the problem that is to be solved with the genetic algorithm. The GA then searches for good solutions in this population and uses the better ones to breed a population of improved solutions. Search of the GA is guided by the fitness of each bit string in the initial population, where fitness is a measure of the problem solving performance of a bit string. The selection operator of the GA then biases reproduction towards bit strings of high fitness and away from bit strings with low fitness. Before the new crop of bit strings enters the next generation cycle, bit strings mate randomly and swap sections (or "genes") of their bit strings. Furthermore, in some bit strings may mutate (i.e. change individual bit values). Finally, the fitness of the bit strings in this new generation, which consists of crosses and mutants of the fittest bit strings of the previous generation, is assessed. This cycle is repeated until some exogenously specified number of generations have been simulated.

4. Simulating the emergence of spontaneous institutions with Artificial Life methods

Artificial life stories of the emergence of spontaneous institutions have their origin in Axelrod's (1980a) computer tournament of iterated Prisoner's Dilemma (IPD) strategies. All other simulations reported here are variations of this theme. Before the variations, their results and implications are reviewed, the prisoner's dilemma game and Axelrod's initial tournaments are briefly recapitulated.

4.1. Genesis: Axelrod's ecological simulation of iterated Prisoner's Dilemma games

In the standard prisoner's dilemma (PD) game there are two players. The players have to choose one of two actions: to cooperate or not to cooperate with the other player. The consequences of choice are payoffs measured in some unit of utility for the players. The players may not communicate with each other (i.e. the game is a non-cooperative one) and they have to make a choice (i.e. no player has the option not to make a choice). Furthermore, there is no indication that the game will be repeated in the future.

The game can be represented in its standard form:

		Player B	
		Strategy 1 **C**ooperate	*Strategy 2* **D**efect
Player A	*Strategy 1*: **C**ooperate	R; R	S; T
	Strategy 2: **D**efect	T; S	P; P

Such a game is a PD-game if the structure of the payoffs satisfies the conditions:

$T > R > P > S$ and

$2R > T+S$

with:

- T: the temptation payoff (e.g. 5 utils);
- R: reward from mutual cooperation (e.g. 3 utils);
- S: the sucker's payoff (e.g. 0 utils) and
- P: the punishment for mutual defection (e.g. 1 util).

The first relation assures that defection is the dominant strategy for both players. The second relation indicates that there are social gains from cooperation. If the game is repeated, this relation also has the interpretation that the players will not do better by playing a non-mutually cooperative strategy, i.e. toggle out of synchrony between cooperation and defection, rather than by mutually cooperating in each round.

If played only once, the outcome of the PD-game played by rational players should be mutual defection and the gains from cooperation, which amount to 2R-T-S, are not realised. Mutual defection is also the best strategy if the players repeat the game for a known number of times. Rational players of the repeatedly played iterated Prisoner's Dilemma (IPD) game should, however, find continuous cooperation in their best interest if the probability p that the game is played on round i (i.e. the probability that the game stops on trial i is 1-p) is sufficiently high (Telser 1987, p. 217).

The PD-game can be used to model conflicts among resource users, such as the conflicts arising among herders or irrigators in the "tragedy of the commons" (Ostrom et al 1994). Similarly, the decision not to contribute to the maintenance of some resource conservation measure, such as a check dam in a ravine or contour bunds along an eroding slope, may be cast in a PD-game mould. Obviously, many real-world situations contain more detail than the PD-game, if only that more people or players are involved who meet repeatedly in similar situations. The PD-game and its extensions - the multiplayer isolation paradox and the multiple-round IPD-game -, nevertheless capture much of the drama of conflicts among agricultural resource claimants.

Tragedies happen but they need not be the rule and farmers dissolve regularly conflicts about agricultural resources. The unpalatable outcome of the PD-game played by self-interested, rational players may be avoided if additional rules are imposed on the players by an outside authority or if the game is changed in other ways such that the players have more than two strategies available. Explaining the absence of widespread tragedy and conflict as the result of rules imposed on the players by an outside authority could yield an account of the creation of a designed institution but would not yield a invisible-hand explanation of such institutions. A invisible-hand story of a spontaneous institution arising from a PD-situation proceeds in three steps: First, the repeated interaction among resource claimants is modelled as an IPD-game with an unknown but not necessarily infinite number of rounds. Second, in a game with n-iterations each player has available 2^n strategies. If players have some memory of the moves in previous rounds of the game, their choice of strategies may depend on their partner's behaviour in previous rounds. Third, if cooperative behaviour is chosen by one and reciprocated by the other players, and if mutual cooperation is robust against invasion by uncooperative players, cooperative behaviour may become the rule and can be considered an institution as defined by Schotter (1981).

Axelrod (1980a) has shown how such institutions may arise in iterated PD-games. He evaluated in a round-robin tournament a number of strategies that had been suggested by experienced game theorists. The outcome of the tournament is legend: the nice, provocable, and forgiving Tit-For-Tat-strategy won the tournament without being the winner in anyone single contest. The round-robin tournament lacks, however, some of the ingredients necessary for an evolutionary account of the emergence of an institution because no new strategies were generated during the tournament.

Taking his inspiration from biologists, Axelrod (1980b) extended his tournament approach into an ecological tournament. In this simulation, the strategies represented a population of homogenous individuals, a round of plays constituted a generation, and the payoffs from the rounds of PD-games were interpreted as the number of offspring. The number of individuals using a particular strategy was therefore determined by the weighted average payoff of a strategy, where the weights depend on the reproductive success in previous generations of the other strategies or players. After one thousand generations, when the proportion of strategies had stabilised, Tit-For-Tat was the dominant strategy in the population. The lessons that Axelrod (1980b) drew from this ecological tournament and which are important in the context of the emergence of institutions are two: (i) there is no best strategy independent of the environment but some strategies are more robust than others and do well in various environments; (ii) nice, provable, and forgiving strategies tend to be robust. Taken together the lessons imply that institutions fostering cooperation tend to emerge in situations characterised by potential conflict among self-interested and rational economic agents because often some robust cooperative strategies are regularly chosen by many players.

4.2. Variation of strategies

Because the performance of a strategy in an ecological tournament depends on its environment, which comprises the population of strategies that were sufficiently successful to avoid elimination, the spontaneously emerging pattern of strategies that avoid elimination may depend on the strategies in the initial population. To overcome the selection bias that may be present in a non-random sample of PD-strategies, Nowak and Sigmund (1992) carried out an ecological tournament with reactive, stochastic PD-strategies. Such strategies were defined as "...given by a triple (y, p, q), where y is the probability to cooperate in the first round and p and q are the conditional probabilities to cooperate after a C (respectively D) of the other player" (Nowak and Sigmund 1992, p. 252). For the tournament populations were formed that consisted of 99 randomly sampled strategies and a strategy that closely resemble Tit-For-Tat was seeded into the population. Reproduction according to fitness was introduced by adjusting the frequencies of the strategies in the population according to the ratio of the average payoff of a given strategy to the average payoff in the population. In the tournaments, uncooperative strategies tended to increase in frequency initially and nice, cooperative strategies became rare. However, once the suckers had been eliminated reciprocating Tit-For-Tat-like strategies staged a comeback only to be replaced by strategies close to Generous-Tit-For-Tat, i.e. strategies that always respond with cooperation to a cooperative move but that do not always retaliate an uncooperative move with defection. Once Generous-Tit-For-Tat had taken over, however, the structure of the population stabilised. Generous-Tit-For-Tat can therefore be considered a spontaneous institution that emerges in an IPD-environment of strategies with single-step memories that is seeded with a small fraction of brave, nice, and reciprocating strategies, such as Tit-For-Tat.

In another experiment, Nowak and Sigmund (1993) enlarged the information set of the reactive strategies to also include the strategies' own moves in the previous round and strategies were defined by conditional probabilities that depended on the outcome of the previous move. Furthermore, the initial population consisted only of random strategies and randomly chosen mutants were introduced into the population at regular intervals. Again, the frequencies of strategies were adjusted according to their relative performance and strategies whose performance fell below a certain threshold were discarded. Hence, in contrast to the earlier experiments where reproduction according to fitness drove the evolution of strategies but where no new strategies could be created during the experiment, mutation also allows new strategies to enter the population. The winning strategy in this experiment was cooperative but at the same time able to exploit unconditional cooperators. Furthermore, the average payoffs in the simulations were either close to the value for mutual cooperation or to the value for mutual defection, with extremely long runs at either value and very short transitory periods. The implication of this experiment is that mutual cooperation is not necessarily the dominant institution and that mutual cooperation may easily and spontaneously turn into mutual defection for no obvious reasons.

A more sophisticated mechanism than the one employed by Nowak and Sigmund (1993) had already been used by Axelrod (1987) who used a genetic algorithm to generate new, unforeseen IPD-strategies. The strategies in this Artificial Life experiment were deterministic and the choice of a move in any round of the game depended on the outcomes of the last three encounters with a given other strategy (i.e. players had a memory of three rounds). Two different environments were specified. One environment consisted of the eight strategies that had performed best in his earlier ecological tournament among prespecified strategies (Axelrod 1980b). The other environment consisted of the evolving population of strategies itself and therefore simulated coevolution among strategies. Strategies reproduced according to their performance, and crossover and mutation operators were applied to form a the next generation of strategies. Similar to Nowak and Sigmund (1993), Axelrod (1987) also found that the population of strategies initially evolved away from cooperation. After some ten to twenty generations, however, cooperation emerged again when strategies had evolved that were able to discriminate between partner strategies that reciprocate cooperation from those that do not. As in Axelrod's (1980b) earlier experiments, the emerging institution again was blissful mutual cooperation.

In contrast to the players in a standard PD-game, free people may have the choice not to play the game at all. For example, given a choice, people may withdraw from unpleasant communities of players who are either uncooperative or even prone to exploit cooperators. When the option to opt-out is included in evolutionary IPD-games defection on cooperators becomes less attractive because it may drive other players into isolation. Batali and Kitcher (1994) investigated the opt-out option in an evolutionary experiment with the extended PD-payoff structure $T>R>W>P>S$, where W is the payoff when opt-out is chosen. They found that the option not to play the PD-game increases the cooperation in the population but that it also leads to less stable group behaviour. In particular, the results of the experiments were characterised by pronounced cycles reminiscent of chaos where periods with high proportions of defectors in the population were followed by periods with high proportions of cooperators. Common property situations have been modelled as PD-games with an opt-out payoff of $W > 0$ (Ostrom et al. 1994, p. 57f). The results of Batali and Kitcher (1994) therefore suggest that the tragedy of the commons is not an inescapable fate and that blissful mutual cooperation may be a rather fickle evolved institution for the governance of the commons.

In order "...to learn what conditions favour the development of norms so that cooperation can be promoted where it might not otherwise exist or be secure" Axelrod (1986, p. 1097) studied coevolutionary games similar to n-person PD-games but in which players had the option to hurt defectors who may be detected with a certain probability. Players who hurt defectors, must, however, bear some enforcement cost.

Given these options, players were characterised by (i) their boldness B to defect even though defection may be observed with a given probability S, and (ii) their vengefulness V or probability that a player will punish a defector. In evolutionary games, in which strategies were allowed to mutate and to procreate in proportion to fitness, average boldness in the population initially declined rapidly. Once boldness had fallen to low levels, vengefulness also fell to levels close to zero and boldness could increase with impunity and became the stable state. That "sad" outcome could be prevented when metanorms where introduced into the game which implied that players ought to punish also those players who did not punish defectors. When this metanorm was introduced and when average vengefulness was not too low in the initial population, the metanorms could prevent the breakdown of the norm to be vengeful.

4.3. Variation of the payoff-structure

Fogel (1993) and Angeline (1994) modified Axelrod's (1987) approach and employed evolutionary programming techniques to generate and select strategies for IPD-games. Furthermore, they also investigated the impact of variations in the value of the temptation payoff T on the emergence of stable cooperation. For values of T that satisfied the condition $2R>T+S$ the results of the evolutionary programming experiments were similar to the results obtained in the genetic programming experiments. Fogel (1993) varied T in steps of 0.25 within the range $5<T\leq 6$ with the other payoffs held constant at $R=3$, $S=0$, and $P=1$. He observed in all trials initial defection and falling total payoffs, followed by increasing payoffs. In experiments with values of $T=6$, i.e. when joint maximum payoff could be obtained by behaviour other than mutual cooperation, the increase in total payoff was not due to spreading mutual cooperation as many of the best strategies in the evolved population were prone to defect on cooperators. Fogel (1993) therefore believed that cooperation can only be trusted to evolve if the difference between the temptation payoff T and the reward for mutual cooperation R is small. Angeline (1994), however, observed the emergence of wide-spread non-mutual cooperation in experiments where the temptation payoff was set at $T=6$ but a lower tendency toward non-mutual cooperation when $T>6$. The most common form of non-mutual cooperation in Angeline's (1994) experiments was of the type where one player played the sequence {CDCD} whilst the other responded with {DCDC}. Hence mutual cooperation may not be the only institution evolving from the repeated interaction of self-interested adversaries but the rule to take turns in mutual exploitation may also become a viable institution.

4.4. Emergence of cooperation in different environments

Several authors have reported the results of Artificial Life simulations of IPD-games in which the environment of the game was modified in one of two ways: (i) varying the size of the population of strategies, and (ii) introducing space, i.e. allowing games to be played only among the strategies that are part of some neighbourhood.

Fogel (1993) and Bankes (1994) considered the impact of the number of players on the emergence of cooperation. Fogel (1993) ran his evolutionary programming experiments with population size ranging from 50 to 1000 individuals but found little impact of population size on the emergence of cooperation. Bankes (1994) used a genetic algorithm framework and the N-player generalisation of the PD-game suggested by Schelling (1978) to explore the impact of the number of players on the emergence of spontaneous institutions. The main qualitative results of Bankes' (1994) ALife-experiments were two: Although the average level of cooperation was unrelated to specific properties of the payoff functions for cooperation and defection, the ease with which cooperation emerged was quite sensitive to the minimum size of a coalition of players that can benefit from cooperation and significant cooperative behaviour was never achieved when this number was greater than three. In addition, similar to the results of Nowak and Sigmund (1993), high levels of cooperation and defection with intermittent short transition periods occurred in all runs.

Beginning with Axelrod (1984), several simulation studies of the evolution of cooperation among players whose location in space is taken into account have been conducted. These studies, which have recently been reviewed by Lindgren and Nordahl (1994) often employ cellular automata techniques. When space is accounted for in the simulations two broad qualitative results obtain: (i) In spatial models a larger number of strategies can persist through the formation of regions with stable strategies; and (ii) varying the payoffs affects the dynamic properties of the emergence of stable structure and some payoffs structures result in spatio-temporal chaos of strategies. These observations were corroborated by Oliphant (1994) who studied the impact of spatial arrangements in simulations of non-iterated PD-games where pairs of players were chosen with and without taken their position in space into account. When position in space was not taken into account, the population of players quickly evolved towards defection. When the opponents for a given individual were selected from a Gaussian distribution around the individual, then cooperation quickly emerged as the dominant strategy in the population. Cooperation was, however, not always stable. In particular, in runs with high differences between the temptation payoff T and the reward for cooperation R, cooperation remained stable but when the standard payoff values of the PD-game were used cooperation was interrupted by sharp transitions into periods of defection.

Huberman and Glance (1993) cast some doubt on the relevance of the spatial PD-games. Their criticism is based on the observation that in social systems the elements of the system rarely update their states in synchrony whereas the artificial systems assumed synchronous updating. For the same simulation scenario that Nowak and May (1992) had used earlier to demonstrate the impact of space on the emergence of cooperation and spatiotemporal chaos, Huberman and Glance (1993) demonstrated that spatial PD-games quickly evolve into a state of global defection when updating of the interacting players is continuous and asynchronous.

5. Implications, potential and limitations

Given the considerable number of published ecological and evolutionary simulations that trace the emergence of spontaneous institutions, a tentative assessment of the use of this approach to the study of institutions governing agricultural resource use may be attempted. This assessment, which is presented in form of short theses, also summarises the paper.

- Artificial life methods are suitable for generating invisible-hand processes that can explain the emergence of spontaneous institutions.
- Artificial emergence studies of spontaneous institutions provide an important complement to conventional explanations of institutions that are biased toward the explanation of designed institutions.
- The main insight from the artificial emergence studies is that institutions conducive for cooperative resource use can emerge among self-interested economic agents in a wide range of circumstances.
- The emergence of cooperative institutions often appears to depend on sometimes rather subtle conditions, which are not yet fully explored and which probably never will be explored in full detail.
- Conflict of interest among economic agents was always modelled as a Prisoner's Dilemma game. Conflicts among resources claimants may, at times, be better modelled using other games than the PD-game, e.g. the "Battle-of-the-Sexes" game, the "Hawk-and-Dove" game, or other games that have been suggested in the context of institutions for the governance of natural resources (Runge 1981; Sugden 1986; Ostrom et. al. 1994). Exploring the dynamic evolutionary properties of these games would be an obvious next step in research.
- The possibility of institutions being the result of chaotic processes with long stable phases should encourage analysts of institutions to take a sufficiently long view when studying resource institutions and it should remind institutional reformers of the futility and dangers of premature intervention.
- The artificial emergence literature is well integrated with the evolutionary games literature but it is as yet uninformed by empirical studies of institutions governing natural resources. Artificial emergence research could benefit from a closer integration with empirical research of institutions if the results and insights from empirical research were used to constrain the parameters of the scenarios or environments of evolutionary simulation experiments.

References

Angeline, P.J. 1994. An alternative interpretation of the iterated prisoner's dilemma and the evolution of non-mutual cooperation. Pages 353-358 in: Brooks, R.A. and Maes, P. (eds). *Artificial Life IV. Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems.* Cambridge, Mass.: MIT Press.

Arrow, K., Bolin, B., Constanza, R., Dasgupta, P., Folke, C., Holling, C.S., Jansson, B.-O., Levin, S., Mäler, K.G., Perrings, C. and Pimentel, D. 1995. Economic growth, carrying capacity, and the environment. *Science* 268:520-521.

Arrow, K.J. 1974. *The limits of organization.* New York, NY: W.W. Norton & Company.

Axelrod, R. 1980a. Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution* 24(1):3-25.

Axelrod, R. 1980b. More effective choice in the prisoner's dilemma. *Journal of Conflict Resolution* 24(3):379-403.

Axelrod, R. 1984. *The evolution of cooperation.* New York, NY: Basic Books.

Axelrod, R. 1986. An evolutionary approach to norms. *American Political Science Review* 80(4):1095-1111.

Axelrod, R. 1987. The evolution of strategies in the iterated prisoner's dilemma. Pages 32-41 in: Davis, L. (ed.). *Genetic algorithms and simulated annealing.* London: Pitman.

Balmann, A. 1994. *Pfadabhängigkeiten in Agrarstrukturentwicklungen.* Dissertation, Fachbereich Agrarwissenschaften, Georg-August-Universität Göttingen.

Bankes, S. 1994. Exploring the foundations of artificial societies: Experiments in evolving solutions to iterated n-player prisoner's dilemma. Pages 337-342 in: Brooks, R.A. and Maes, P. (eds). *Artificial Life IV. Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems.* Cambridge, Mass.: MIT Press.

Bardhan, P. 1989a. The new institutional economics and development theory: A brief critical assessment. *World Development* 17(9):1389-1395.

Bardhan, P. (ed.) 1989b. *The economic theory of agrarian institutions.* Oxford: Clarendon Press.

Batali, J. and Kitcher, P. 1994. Evolutionary dynamics of altruistic behavior in optional and compulsory versions of the iterated prisoner's dilemma. Pages 343-348 in: Brooks, R.A. and Maes, P. (eds). *Artificial Life IV. Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems.* Cambridge, Mass.: MIT Press.

Buchanan, J.M. 1992. *Better than plowing.* Chicago, IL: Chicago University Press.

Flam, F. 1994. Artificial-life researchers try to create social reality. *Science* 265:868-869.

Fogel, D.B. 1993. Evolving behaviors in the iterated prisoner's dilemma. *Evolutionary Computation* 1(1):77-97.

Fogel, D.B. 1995. *Evolutionary computation.* Piscataway, NJ: IEEE Press.

Fogel, D.B. and Fogel, L.J. 1996. An introduction to evolutionary programming. Pages 21-33 in: Alliot, J.-M., Lutton, E., Ronald, E., Schoenauer, M. and Snyers, S. (eds). *Artificial Evolution. Selected Papers, European Conference, AE 95 Brest, September 4-6, 1995.* Berlin: Springer Verlag.

Forrest, S. 1993. Genetic algorithms: Principles of natural selection applied to computation. *Science* 261:872-878.

Gerhardt, M. und Schuster, H. 1995. Das digitale Universum. Braunschweig: Vieweg & Sohn Verlagsgesellschaft.

Goldberg, D.E. 1994 Genetic and evolutionary algorithms come of age. *Communication of the ACM* 37(3):113-119.

Gutowitz, H 1995. Artificial life simulators and their applications. http://alife.santafe.edu/topics/simulators/dret/dret.html

Hayek, F.A. 1967. *Studies in philosophy, politics, and economics.* Midway reprint 1980. Chicago, IL: Chicago University Press

Huberman, B.A. and Glance, N.S. 1993. Evolutionary games and computer simulations. *Proceedings of the National Academy of Science USA* 90:7716-7718.

Levy, S. 1993. *Artificial life.* London: Penguin Books

Lindgren, K. and Nordahl, M.G. 1994. Cooperation and community structure in artificial ecosystems. *Artificial Life* 1(1/2):15-37.

Lloyd, S. 1996. Complexity simplified. *Scientific American* 274(5):88-92.

Menger, C. 1985. *Investigation into the method of the social sciences with special reference to economics. (Untersuchungen über die Methode der*

Sozialwissenschaften und der Politischen Ökonomie insbesondere). New York, NY: New York University Press.(Originally published in 1883).

Mitchell, M. 1996. *An introduction to genetic algorithms.* Cambridge, Mass.: MIT Press.

Neumann, J. von 1963. The general and logical theory of automata. Pages 288-328 in: *Collected works, Vol. V. Design of computers, theory of automata and numerical analysis.* Oxford: Pergamon Press.

North, D.C. 1981. *Structure and change in economic history*. New York, NY: W.W. Norton & Company.

North, D.C. 1990. *Institutions, institutional change and economic performance.* Cambridge: Cambridge University Press.

Nowak, M. A. and May, R.M. 1992. Evolutionary games and spatial chaos. *Nature* 359:826-829.

Nowak, M.A. and Sigmund, K. 1992. Tit for tat in heterogeneous populations. *Nature* 355:250-253.

Nowak, M. and Sigmund, K. 1993. A strategy for win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 364:56-58.

Oliphant, M. 1994. Evolving cooperation in the non-iterated prisoner's dilemma: The importance of spatial organization. Pages 349-352 in: Brooks, R.A. and Maes, P. (eds). *Artificial Life IV. Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems.* Cambridge, Mass.: MIT Press.

Ostrom, E., Gardner, R., and Walker, J. 1994. *Rules, games, and common-pool resources.* Ann Arbor: University of Michigan Press.

Runge, C.F. 1981. Common property externalities: Isolation, assurance, and resource depletion in a traditional grazing context. *American Journal of Agricultural Economics* 63(4):595-606.

Sipper, M. 1994. Studying artificial life using a simple, general cellular model. *Artificial Life* 2:1-35.

Schelling, T.C. 1978. *Micromotives and macrobehavior.* New York, NY: W.W. Norton & Company.

Schotter, A. 1981. *The economic theory of social institutions.* Cambridge: Cambridge University Press.

Smith, A. 1937. *An inquiry into the nature and causes of the wealth of nations.* Edwin Cannan (ed.). New York, NY: Modern Library.

Sugden, R. 1986. *The economics of rights, cooperation and welfare.* Oxford: Basil Blackwell.

Telser, L.G. 1987. *A theory of efficient cooperation and competition.* Cambridge: Cambridge University Press.

Ullmann-Margalit, E. 1978. Invisible hand explanations. *Synthese* 39(2):263-291.

Wolfram, S. 1984. Computer software in science and mathematics. *Scientific American* 251:140-151.