

Multilevel evolution: replicators and the evolution of diversity

P. Hogeweg

Bioinformatica, University of Utrecht, Padualaan 8, 3584 CH, Utrecht, The Netherlands

Abstract

In this paper we report on an initial attempt at studying evolution as a multilevel process, in which interactions of different space/time scales mutually determine the evolutionary potential. Side effects of processes on one scale may form a substrate for evolution at another scale.

We study the generation and maintenance of diversity in a multilevel replicator system and focus on the role of local information processing and pattern formation. We show how the conflicting requirements of the generation and maintenance of diversity are met in different ways in completely mixed (stirred) and poorly mixed (diffusive) systems. In the former diversity is maintained primarily by non-invadability, and operates at the lowest level; in the latter case diversity is maintained by locally contained invasions which are subsequently eliminated by higher level processes.

1. Introduction

Four lines of reasoning converged on the research reported in this paper.

(1) Evolution as a multilevel process.

Evolution is most often studied as a one level process involving informatic entities which can mutate in various ways and are selected by the environment through differential replication or decay (“survival of the fittest”). In this framework evolution is studied most often as an optimization process. Indeed, Genetic Algorithms exploit this paradigm successfully as a method for optimization. Realising that most of the relevant environment of biotic systems may consist of other biotic systems, i.e. systems which may change their characteristics at a similar time-scale, this paradigm is extended to include

coevolution. Coevolutionary processes are most often modelled as (one level) dynamical systems, rather than optimization systems, and the attractors are characterized.

Although the important insights in evolution so far have been obtained by virtue of squeezing the multiple levels into one level, we think the time has come to start exploring evolution as a multiple-level process in which patterns generated at one level may determine the evolutionary dynamics and the evolutionary potential at other levels. Indeed we think that such a multiple-level approach is a prerequisite for studying the crucial questions of evolution, i.e. how novel structures and functions can arise. In the above mentioned one level models either the coding structure and the target are fixed (optimization models), or the variables and interactions are fixed (dynamical systems). By modelling evolution as a multiple-level process the ‘side effects’

of one level can modify the targets, the coding structure and/or the interaction structure at other levels [10,11].

Extension of one-level models to multiple levels can be done starting from the optimization viewpoint, starting from the coevolutionary viewpoint, or from both.

1.1. Evolution of the coding structure in optimization models

By recognising that the ‘targets’ of evolution (i.e. continued existence) are such that they can be realised as effectively in many different ways, we can study which ‘solutions’ are preferred by the dynamics of the evolutionary process. In other words we can study how the choice between ‘neutral peaks’ in the fitness landscape is not a ‘neutral choice’, and how such non-neutral choices affect the coding structure and thence the future evolutionary dynamics. In this vein we have studied the generation of multiple or single coding under respectively high and low mutation rates [11] and generation of patterns in RNA sequences by evolutionary dynamics leading to relatively smooth or rugged parts of the RNA secondary structure landscape [12].

1.2. Evolution of new levels of selection in coevolutionary models

By embedding coevolutionary systems in space, the interaction between the selfreplicating units may form spatial structures. The dynamics of these spatial structures may determine the evolutionary fate of the selfreplicating units which generate them and therewith the low level definition of the (co)evolutionary system [1–4]. In other words the ‘emergent properties’ of a (one-level) coevolutionary system may generate a new level of selection.

1.3. Combining coevolution and coding

Interesting combinations of these two approaches include the Tierra system of Ray

[17,18] which studies the coevolution and coding structure of entities which evolve the coding of the replicative process itself, and the (spatial) game dynamics studied by Lindgren [14,25] which studies the coevolution of interaction ‘strategies’ in relation to the spatial structures generated by these strategies under different parameter regimes. In this paper we combine coevolution and coding the other way around, i.e. we study a system in which the coding structure of the higher level entities is modelled in terms of a coevolutionary process of (self replicating) molecules. The higher level entities optimize a very general optimization criterion, i.e. the diversity of the underlying coevolutionary system. We show that generation and maintenance of diversity in the system is achieved in very different ways dependent on pattern formation or the absence thereof in the lower level coevolutionary system.

(2) The cell as the environment for an evolving genome

Possibly even more important than interacting with other biotic systems, a biotic system has to deal with itself in its most direct environment. More specifically, the genome interacts with its own products (i.e. the cell), rather than with the environment, and the interaction of the cell with the environment just ensures continued existence of the system. Any new ingredient of the system in the first place should be sufficiently compatible with what is already there. Moreover genes are a passive form of information storage; only transcribed genes play any role in the system. Thus the cellular environment should not only be undisturbed by a new gene, it should also activate it. Gene regulation has been modelled with Boolean networks by Kauffman [13] and the amenability of such networks to optimization, and coevolution has been studied by him in models in which these processes are quenched into a powerful and extremely simple representation.

In this paper we intend to focus on the internal

aspects of evolution by including explicit intracellular dynamics in an evolutionary system which takes a ‘solipsistic’ view of a cell. Whatever a cell ‘is’ or ‘does’ is fine as long as it can continue doing it. However the external environment is represented in the form of an optimization criterion: the more ‘information’ the cell carries the more it might be able to “do”. Thus we maximize diversity of the system.

(3) Generation and maintenance of diversity in selfreplicating systems

Selfreplicating systems which interact through competition for a common resource tend to exclude each other. In an ecological context this phenomenon is called “competitive exclusion”, in evolutionary context “survival of the fittest”. Although the former term has definitely a more negative connotation than the latter, the maintenance of diversity is a crucial question in models of evolution as well. This is in particular the case because the amount of information which can be stored in one replicator is limited by the existence of a so called ‘error threshold’ (or ‘information threshold’): only by reducing the mutation rate can more information be included in one replicator. However, the reduction of the mutation rate is presumably only possible by more sophisticated replication mechanisms (e.g. proofreading), which require more information to be stored in the genome. Therefore an important open question in (prebiotic) evolution is how the information threshold can be crossed and how information accumulation can occur. Eigen and Schuster’s classical proposal to cross the information threshold by coevolution of entities catalysing each other’s selfreplication, is hampered by the difficulty of maintaining high diversity in selfreplicative systems. In well-mixed systems only relatively small cyclic interaction structures can persist, and these structures are vulnerable to ‘parasites’, i.e. molecules which receive but do not give catalysis to other molecules. Such parasites are likely to arise in mutating systems.

Embedding the coevolution in a spatial context does alleviate some of the difficulties occurring with the well-mixed systems: temporal fluctuations are stabilized, noncyclic interaction structures are stable, and the system can be resistant to parasites through the generation of higher levels of selection. Nevertheless information accumulation appears to be limited also in these systems, and in some cases we have shown that evolutionary processes tend to lead the system to an area of parameter space where, e.g., they are not resistant to parasites [1–5,8,16]. Maintenance of diversity in non-selfreplicating systems is less of an issue. Fontana [7] therefore initially constrains his system excluding selfreplication of his primary building blocks. In fact ecological systems as well as evolutionary systems and also other selfreplicating systems such as the immune system, are not purely selfreplicating systems: influx from the ‘external’ environment (other areas, seed banks, bone marrow/thymus), or by mutations occurs. Indeed such an influx tends to stabilize the system, as is shown by Stadler et al. [21,22] in replicator systems including mutations.

In this paper we study the generation and maintenance of diversity as mediated by higher level selection and influx. The higher level processes maximising diversity select interaction topologies and can control the influx of the ‘environment’ (in our system the ‘genome’).

(4) De-coupling of dynamics and information storage

Complex information processing can occur in dynamical systems when there are

- (1) stable configurations which provide for information storage;
- (2) moving configurations (for information transfer)
- (3) nonlinearity (decision making) (e.g. [24]).

Clearly these requirements are met in some dynamical systems, i.e. those systems which in CA are called type IV systems, and occur near the onset of chaos. Interesting as this notion is,

biological systems may store information in other ways, e.g. by ‘using’ long lived molecules. DNA can be considered such a long lived analogue of RNA. Likewise memory cells in the immune system may be used for immunity, rather than network configurations. Information storage in the form of long lived molecules decouples information storage to a certain extent from the parameters of the dynamical system.

In this paper we study the potential of this type of information storage, which takes the form of the control of information influx in a replicator system.

2. The model multilevel evolution system

We investigate the role of information storage, local information processing and pattern formation for the potential to evolve species-rich catalytic networks of selfreplicating units. For this purpose we embed random catalytic networks of selfreplicating units in a spatially extended system, i.e. in a Cellular Automaton (CA), add a second layer to the CA containing long lived counterparts of these replicators for information storage, and embed the CA in compartments and study a population of these under selection for ‘diversity’ (= number of species per ‘compartment’).

Thus we have a 4 level evolutionary system: (1) the individual low level replicators, which are represented as the local states of the cellular automata, (2) the meso-scale patterns generated in the CA by the dynamics of the catalytic network, which (co)determine the fate of the replicators as discussed above, and (3) the population level which causes only the most diverse CA’s to survive and replicate, and therefore codetermines the evolutionary fate of mesoscale patterns as well as the replicators and the information storage thereof. In addition there is a ‘genome’ in which information is stored. This four level evolutionary system uses both paradigms for studying biological evolution, i.e.

as a dynamical system of coevolving entities and as optimization process, using genetic algorithms. The optimization with genetic algorithms may be considered as a convenient and powerful way to investigate the properties of the low-level dynamical system [20], or as an extension of the biotic system under investigation, i.e. to a biotic systems consisting of ‘cells’.

The CAs harbour a collection of replicators which are semi-isolated from the environment. Each cell (automaton) of the CA can contain just one of these replicators, or is empty. In this baseline study the structure of the replicators is not considered and their interactions are assigned randomly. The building blocks of the replicators are not considered; the space limitation introduced automatically by the CA formalism is assumed to represent any limitation which might occur. There is a fixed set of possible replicators in the universe; in most simulations shown there are 160. This set is apriori fixed to facilitate the comparison of different selection regimes. The replicators can catalyse the replication of other replicators. These interactions are fixed at the beginning of the simulation. We study the effects of varying the interaction frequency.

The CA’s transitions are defined in 4 processes.

(1) *Catalysis*: the amount of catalysis each replicator (or its long lived counterpart in the genome layer) gets, is calculated as the sum of catalytic interactions it has with its 8 neighbours in the replicator layer.

(2) *Replication and Decay*: if a CA cell in the replicator layer is empty (i.e. is in state = 0) its next state is equal to the state of its neighbour which gets maximum catalysis; its neighbourhood consists of the 8 neighbouring cells in both CA layers. If none of the neighbours gets catalysis the cell remains empty. If the cell is in any other state (i.e. contains a replicator) it has a certain (species specific) probability to decay and become empty.

(3) *Mixing*: replicators move either by diffusion (i.e. they move in a random walk fashion to

neighbouring cells, using the diffusion algorithm of [23]) or they move by complete stirring: each replicator is moved to an arbitrary location in the CA, i.e. we model the mean field situation. Replicators can diffuse out of the CA (i.e. the CA has zero boundary conditions). We compare the evolutionary consequences of these two different mixing regimes.

(4) *Somatic mutations*: in addition ‘somatic mutations’ can be added to the CA’s transition rule by randomly changing replicator-types. We studied high and low mutation rates (50 or 1 mutations/CA-generation) as source of nonselective information and/or disruption of the system.

A second layer of the CA is passive and represents the ‘genome’, i.e. contains long-lived counterparts of the replicators. In analogy with the situation for DNA and RNA, the long lived elements of the genome layer can be transcribed into the upper layer of the CA forming the corresponding replicator. The reverse process is not taken into account in the experiments described here. Transcription takes place through catalysis by replicators, analogous to the catalysis of replicators among themselves, i.e. the non-zero elements in the genome layer get catalysis by the replicators in the neighbouring positions in the replicator layer, and if they get most catalysis of any of the neighbours of an empty position in the replicator layer, they are transcribed into this layer. Genome elements do not catalyse replication or transcription. They thus act as a passive information store, activated only by their collective transcription products. The same network of catalysis is used of replication and transcription.

At initialization the first layer of the CA contains a number of replicators and each replicator type has just one representative in the genome layer; they are randomly distributed over this layer, which is empty elsewhere. New genome elements are added during replication of the entire CA (see below), by a process analogous to gene duplication with mutation, or mutation alone. This process is implemented by

simply adding a randomly chosen replicator-analogue to a position in the genome layer. This position is chosen as follows: a random position (x, y) is selected; when no position in a (49-) neighbourhood of x, y contains a replicator-analogue, a new one is added at position x, y (i.e. gene duplication + mutation takes place); otherwise the new replicator-analogue replaces an existing one in this neighbourhood (i.e. mutation without duplication takes place); if moreover position x, y contains a replicator-analogue, it is deleted (genedeletion). This procedure ensures that genome size is limited: once the genome layer becomes densely populated, replacements and deletions predominate. However in the initial stages there is a strong bias to an increase of information storage. In this way we make the build-up of information relatively ‘easy’. The size of the neighbourhood (49) is an arbitrary parameter of the model, chosen such that the genome will build up to about 120 elements during our runs.

Replication of the entire CA occurs periodically. The CA’s with the highest number of replicator-types are preferentially selected for replication, those with the smallest number are replaced. During replication crossover can occur: in that case two cells fuse, by combining part of one CA with the complementary part of the other, thus combining both replicators and genomes.

3. Local vs. global competition

The dynamics of replicator systems are well studied in terms of differential equations [6,9,19,21,22]. We focus here on replicator systems with only positive interactions (sometimes called Hypercycles). In contrast to differential equation models we study replicator systems under discrete, finite conditions and with local rather than global competition. Local competition arises because only the replicators surrounding an empty position compete for its occupation

(and do so in a nonlinear way) whereas in the usual ODE models global resource competition is generally assumed (e.g. chemostat condition). Global resource competition ensures that competition is local in terms of interactions of replicators: only those being catalysed by the same other replicators compete directly, and those getting the largest (net) catalysis will ultimately survive. In contrast local resource competition makes competition in interaction space a global affair. For example giving catalysis gives now a greater disadvantage to the replicator doing so than to the others because it is in its own local surroundings that it generates a stronger competitor. One of the consequences is that in contrast to differential equation models, systems with parasites (i.e. replicators which receive but do not give catalysis) can be very stable ([1,2]; see results). Local interactions lead to spatial pattern formation. Thus pattern formation and local interactions are most often observed together. In this paper we differentiate the two by introducing in a local interaction system two different mixing regimes, i.e. local diffusive mixing (i.e. each replicator can move into neighbouring patches in between every interaction step) and complete mixing, i.e. in between interaction steps all replicators are moved to an arbitrary patch in the system. It turns out that both types of mixing markedly affect the strategies for information generation and conservation open to such systems and/or preferentially exploited by them.

4. Observables of the system

The primary observable of this system is of course the value of the fitness criterion attained under the various selection regimes. The fitness criterion, i.e. the number of replicators simultaneously present in the upper layer of the CA, is, however, but one of the ways in which the diversity of the system can be measured. The size of the genome is another interesting observ-

able, as is the number of replicators active over a period of time rather than simultaneously. The shape and stability of the attractor attained under selection at the population level was assessed by continuing its dynamics for 5000 steps in isolation, measuring the mean number of replicators present over time as well as the total number and by visual inspection of space-time plots (see plates). Another important observable is the topology of the attained networks (see figs. 1, 2). Indeed these other measures are more informative about the various evolution scenarios than is the fitness criterion itself.

5. Initial observations and default parameter settings

Initial observations led to the default settings of the parameters. In this section we discuss these shortly.

Initialization. We initiate the population in each CA with 30 replicators (and their long lived analogues) drawn randomly from the 40 lowest numbered replicators. Thus we ensure a moderate diversity of the initial population, as well as the necessity to incorporate information not available initially into the system. These randomly initiated CAs most often degenerate to very simple systems (2–3 replicators) or die out completely. This is in particular the case for lower connectivities and for the well-mixed systems: in the latter often only a few CAs ‘survive’ to the first replication. Nevertheless the population level optimization is able to generate diverse networks eventually.

Population size. Initially a population size of 36 was used, i.e. the maximum convenient population size. It appeared that halving this size did not affect the results provided the initial population contained some ‘surviving’ members. We therefore used a population size of 18 throughout the reported runs. Where the initial population did not survive until the first CA replication

a few other initiations were executed, before deciding that the system was not ‘viable’.

CA replication rate. Population level optimization is more successful when replication of the CAs happens frequently, i.e. before it has settled in its attractor. Fusion of CAs (i.e. the crossover operator) is more effective during transients than when stable attractors are reached. In the latter case most often the result is just that one attractor out-competes the other one. At present we have not analysed this interesting phenomenon further, but just used the observation to choose the default CA replication rate of once in 125 time steps; replication rate of once in 250 time steps appeared to be as efficient, but requires of course more simulation time.

Population level selection. The dynamics of the system introduces much temporal variation in the value of the optimization criterion, i.e. number of species ‘active’ in the system at the time of the CA replication. Therefore strong selection could be used without constraining the search space of the genetic algorithm too much. Therefore both “reproduction of the fittest” as well as “non-survival of the non-fittest” were used. Thus, CAs to be killed were selected uniformly on the basis of $(\text{Max } N - N)^3$; and CAs to reproduce were selected uniformly on the basis of N^3 (N = number of species > 50 × present in the CA; $\text{Max } N$ = maximum of N in the population).

Decay rate and strength of catalysis. In order to avoid ‘ties’ and very slow drift of competing replicators, variation in decay rate and catalysis strength were set to moderate levels, i.e. decay rate is drawn uniformly between .15 and .25, and catalysis between 80 and 120. It turns out that the average decay and catalysis in the final networks are close to the average value, showing that fast decay and low catalysis can also be ‘advantageous’ for replicators (compare [2]).

Replicator reproduction. Several rules for replicator reproduction were tested, including stochastic rules and rules in which non-catalysed replication was allowed with low probability. We

finally settled on the deterministic rule explained above in order to avoid unnecessary variability. The detailed properties of the system do depend on the particular choice of the replication rule and the size of the catalysis neighbourhood. Nevertheless, our observations gave us confidence that the major conclusions do not crucially depend on this choice.

CA size. The size of the CA codetermines the type of patterns which can persist. If the scale of a pattern is big it might not fit into a small space. This is the case, for example, for spiral patterns formed by cyclic catalytic interactions among a large number of species [1]. The limitation of the size of the system to 100×100 cells may indeed prevent the formation of interesting larger scale patterns, but this may also be the case in the spatial confines of cells. Nevertheless, visual observation in our systems did not reveal many large scale patterns if larger automata were used. We did our comparative runs on a size of 100×100 , and limited diffusion to one step per generation. For display purposes, Plates I and II show experiments on a CA of size 160×160 : the equivalent runs on 100×100 revealed similar patterns.

Duration of observations. Population size is only slowly increasing after $t = 25\,000$; stability still does increase in the spatial systems. Longer runs were sometimes used to check stability trends, but the main comparisons were done after 25 000 steps (= 200 CA generations).

6. Requirements and strategies for generating and maintaining diversity

In order to generate and maintain a large number of replicators in the network, partially contradictory requirements should be met and different strategies can be used to meet them. Under different circumstances (i.c. mutation rates, and connectivities) the difficulty of meeting these requirements varies. It turns out that the two mixing regimes differ significantly in

their ability to meet the requirements and therefore tend to apply different strategies under the same circumstances. The opposing requirements for generating and maintaining a high information content (high diversity) are

Incorporation of information: new entities should be able to ‘invade’ in order to be able to ‘test’ their suitability for inclusion in the repertoire.

Elimination of information: those entities which are ‘harmful’ for maintaining the diversity should be eliminated from the system.

Preservation of information: entities should be available whenever ‘needed’. Active replicators are not only needed for their direct contribution to diversity, but also for the preservation of other replicators, for the incorporation of new replicators and the elimination of harmful ones.

Under high somatic mutation rates preservation of information is automatically achieved: all information is present all the time in the form of mutants (because the average lifetime of replicators is 5 time steps, 50 mutations per time step maintains a store of 250 replicators independent of network interactions). The crucial issue under these circumstances is to be able to eliminate those replicators which can out-compete all others and therefore ‘kill’ the network. Contrariwise, under lower mutation rates preservation of information is the crucial issue, because competitive exclusion tends to limit the amount of information drastically. The ‘genome’ allows selective information conservation. However this information contributes to diversity only if it is activated and is able to invade. A high degree of connectivity increases the possibilities for information to be incorporated in the network: there are more ‘entry points’. If information can enter into the system easily, preservation as well as elimination of information becomes less of an issue: there are no really harmful replicators because inclusion of others is not preventable by anyone and preservation is not necessary because there are enough new recruits. Pattern formation also increases the incorporability of

information: the spatial heterogeneity generates a large number of different ‘opportunities’ for invasion. At the same time pattern formation can increase the possibilities for the elimination of replicators which did invade successfully: waves can ‘push’ information out of the system. However, unlike a high degree of connectedness, pattern formation does not prevent harmfulness itself.

7. Strategies selected

There are three types of attractors which maintain diversity far above the average diversity attained by an arbitrary network from arbitrary initial conditions. These are:

(1) ‘Invariant attractors’.

These attractors preserve (almost) all information in the form of replicators which are present continuously in the system in fairly large numbers. Thus the total number of replicators occurring in the system is equal to the number of replicators present at any time. In our experiments such networks consist of a maximum of about 10 replicators in a number of interlocked cycles of length 3–4, and usually have one or more ‘coupled parasites’ on this ‘core’.

An example. A typical example of such networks is shown in Fig. 1a; its dynamics are shown in Fig. 1b. The lower panel contains the core species which have relatively low densities and small fluctuations and the upper panel displays the parasites. The attractor shown remains invariant notwithstanding the presence of a genome which contains counterparts of other replicators: these replicators cannot invade; although they may be catalysed by the members of the network they are always out-competed by them. However, not for all species is the attractor uninvadable: ‘harmful’ species are excluded from the genome through population level selection; exposing the attractor to an influx of 50 random replicators per time step causes it to

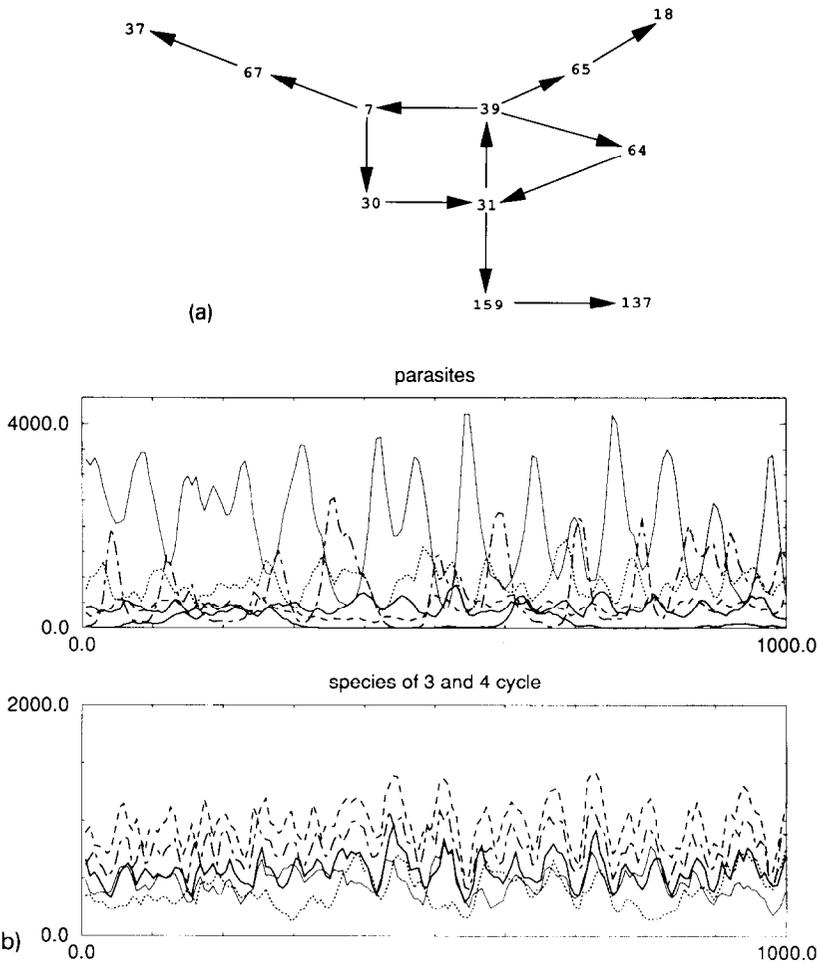


Fig. 1. Attractor of well-mixed system with connectivity = 4/160. (a) The selected subnet, consisting of a core of an interconnected 3/4 cycle and several parasitic chains. (b) Time plot of densities: Upper panel; parasites; Lower panel 3/4 cycle.

collapse rapidly to 2 species only. Indeed the genome contains only 46 different ‘genes’ after 25000 generations while the average influx of genes is over one hundred during this period. The attractor shown is also not self-sustaining: the coupled parasite number 37 (upper panel) is crucial for the maintenance of the attractor but is not preserved by the dynamics in this (finite) system. It is provided by the genome whenever needed. If the influx is blocked the system destabilizes and all species die out within 2000 time steps. Without no. 37, its ‘mother’ parasite eventually out-competes species 30 thus breaking the 4 cycle; the remaining 3 cycle is not resistant

to the remaining, now stronger, parasites and is killed, which in its turn kills the parasites as well.

General properties and circumstances in which they are formed. Other, similar attractors are globally uninvadable; they are strongly selected in systems with high rates of somatic mutations. Globally uninvadable attractors occur also in genomic systems without somatic mutations. The globally uninvadable attractors accumulate up to twice as many genes in the same period as the ones which have to battle harmful genes. These accumulated genes are, however, not expressed unless the attractor is destabilized. These attractors are very stable in isolation. Fusion (recombi-

nation) of two of these attractors which are out of phase may cause however destabilization. Such destabilization will cause the expression of hitherto unused genes which were accumulated only because they were ‘harmless’ under normal circumstances. Invasion of these gene products (i.e. new replicators) may lead the system back to the attractor, or may cause extinction after a long transient. Notwithstanding the long transient relative to the mean life span (i.e. a few thousand vs. ca. 500 time steps), population level

selection can lead the population back towards fixation of uninvadable attractors. These occasional, most often harmful ‘excursions’ lead to the build-up of the larger uninvadable attractors from the smaller ones (2–4) species which are initially selected.

Well mixed systems with a genome and with a connectivity of up to 7 catalytic links per node almost always evolve such an ‘invariant attractor’ of 7–10 species under selection for diversity (see Fig. 2). Likewise well-mixed systems with high

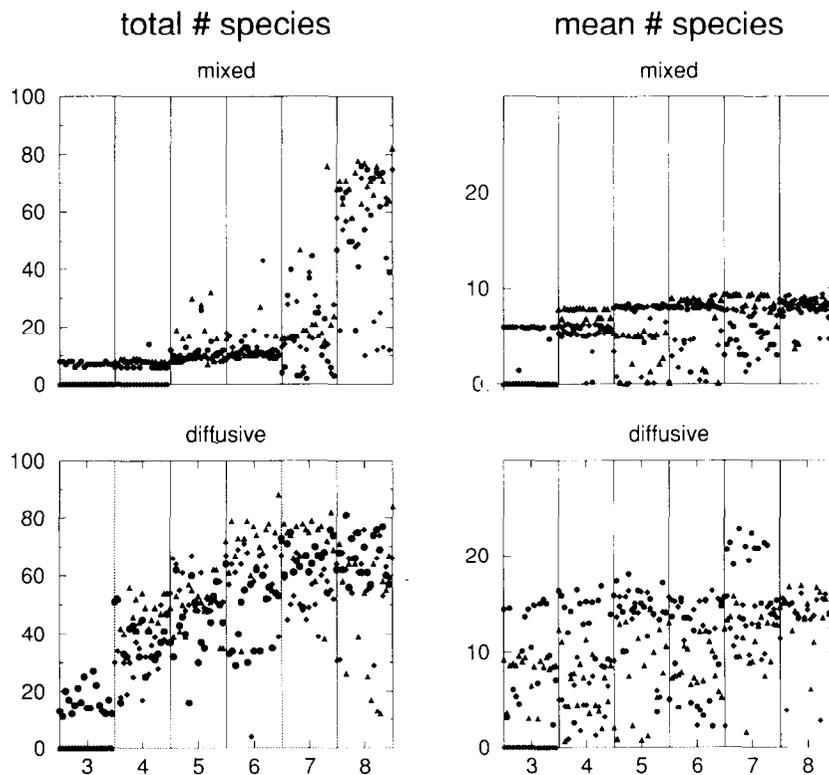


Fig. 2. Diversity of attractors for different mixing regimes and different connectivities. Diversity is measured in isolated CA's during a period of 5 000 timesteps, after a selection period of 25 000 timesteps in which the population of CA's reproduce and fuse (crossover). The lefthand panels show the total number of species present in the period of 5 000 timesteps, the righthand panels the average number of species present simultaneously (the number of species present simultaneously is the selection criterion). Connectivities range from (on average) 3 to 8 connections per species (x -axis). 3 populations (of 18 CA's) are included for each connectivity. The attractors of the well-mixed systems (upper panels) with connectivities upto 7 are of the ‘invariant’ type: the total number of species and the average number of species are typically similar, although some are unstable which results in a larger total number of species and a smaller average number of species (i.e. they die out after a transient). For connectivity 8 ‘percolation’ occurs with a large total number of species, but an average number of species equal to that of lower connectivities. The attractors of the diffusive systems (lower panels) have a permanent core which is successively invaded by different species: the total number of species is larger than the average number of species. The average number of species in the stable attractors are similar for all connectivities (but for higher connectivities a larger proportion of the population of CA's is stable), the total number of species increases with increasing connectivity. Percolation occurs at $c = 8$: the absence of a permanent core is not reflected in either the total number of species or the average number of species.

rates of somatic mutations can attain such invariant attractors which are globally uninvadable for all mutants, or nearly enough so for population level selection to eliminate harmful mutants from the population. Well-mixed systems with low mutation rates are generally not viable when initiated with a small number of species (as is our default): the incoming mutants kill them. Neither is a low mutation rate (one mutation per CA step) sufficient for preserving enough information for maintaining stabilizing second parasites in such attractors. In the pattern formation systems these invariant attractors are not found. Their higher invadability prevents such an 'isolistic' strategy and their potential for dynamically containing invaded species makes it unprofitable. Only for very low connectivities evolved attractors of diffusive systems resemble the attractors described in this section.

(2) Invasions and the preservation of an invariant 'core'.

In evolved diffusive systems only a minority of the species which do occur during a 5000 time step period are present continuously during that period (see Fig. 2). Most species just appear once in a while. There is however a core of species which dynamically maintains itself. This core of species allows the invasion of potentially very long chains (networks) of other species. This 'parasitic' network is not stable, and will eventually die out and free the space it occupied; subsequently the core species recolonise this area.

Examples. Color plate 1 shows such an attractor for a diffusive system without a genome but a high rate of somatic mutations (50 per time step) ensuring species preservation. Panel 1 shows a space-time plot, panel 2 a species profile, panel 3 the number of species present over time and the last panel, four snapshots of the CA. The catalytic network of the species present is shown in Fig. 3a. The attractor is dominated by the 2-cycle of species 12–18 (coloured grey). This 2 cycle is the 'sink' in which the stirred counterpart of this system falls. Due to pattern formation another

2-cycle, i.e. 45–122 (red and yellow) can invade (these species are catalysed by 12 and 18, respectively). This 2-cycle can however easily be invaded and extinguished by a variety of parasitic species, which eventually also come to a dead-end themselves. The dominant 2 cycle invades the open area again, etc.

Color plate 2 and fig. 3b show the network as evolved by the same universe, and the same initial conditions, but without somatic mutation and with information conservation in the genome. A quite different attractor is selected by the system, which is not resistant to a high influx of arbitrary species by somatic mutation, but which can maintain itself with a genome. Without genome the system persists in impoverished form: the 10 'core' species (i.e. those permanently present in the system) remain (colour plate 2b.) The core consists of 2 domains: one dominated by a 2-cycle; the other consists of a number of interlocked 3-cycles and parasites. The latter domain is 'fed' by the first domain: a parasite of the first domain is a member of a 3-cycle of the second domain. The second domain is eventually killed by waves of its parasitic species. The empty space is then reoccupied by the 2-cycle, and the process continues as before. Thus a fairly regularly large scale space time pattern is generated.

It is clear from the picture that the space-time behaviour of the full system is very much shaped by the permanent core species: the genome only provides some additional transient species which make detours or shortcuts between the phases, and thereby make a more irregular pattern. Color plate 2a is generated by initially mixing the evolved pattern. This disturbance killed species 16, i.e. the catalytic connection from domain 1 to domain 2 so that the domains remain connected only through a common parasite. The system then becomes dominated by domain 1 which is regularly locally extinguished by its parasite, but rapidly re-invades the barren area left when the parasite dies out. Domain 2 persists as well, and can occasionally out-compete domain 1 when the latter is 'suffering' from a parasite which does

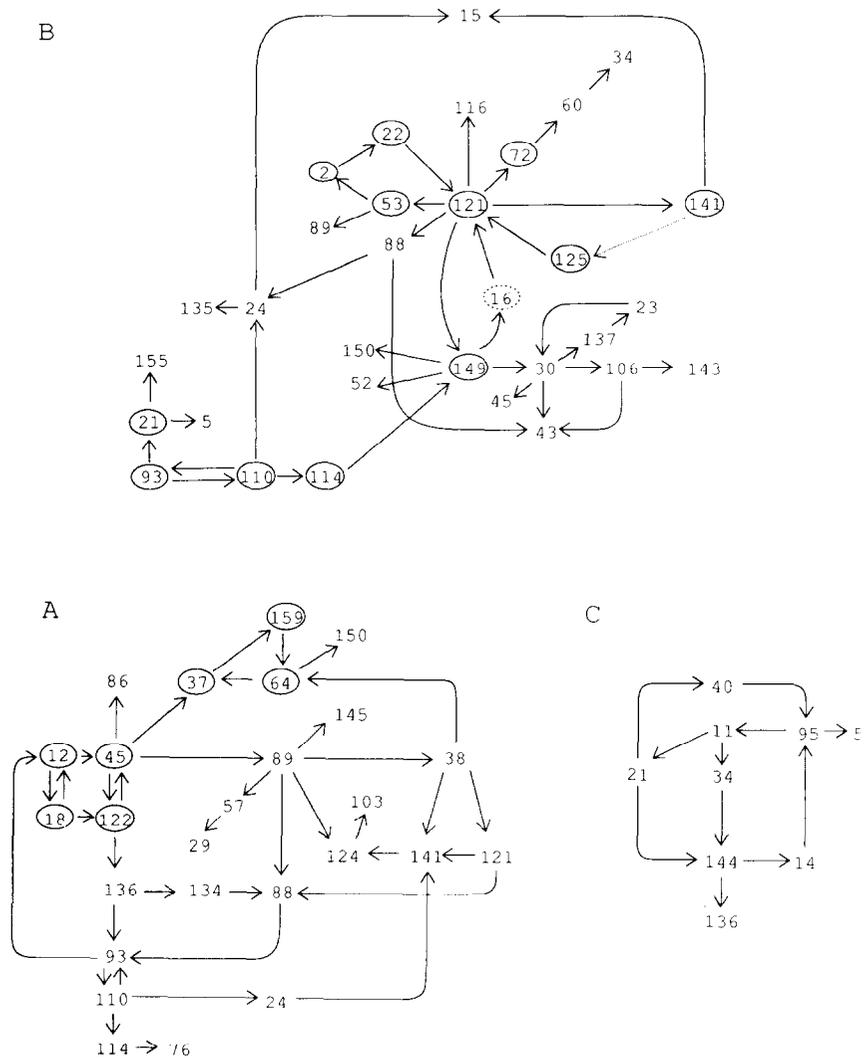
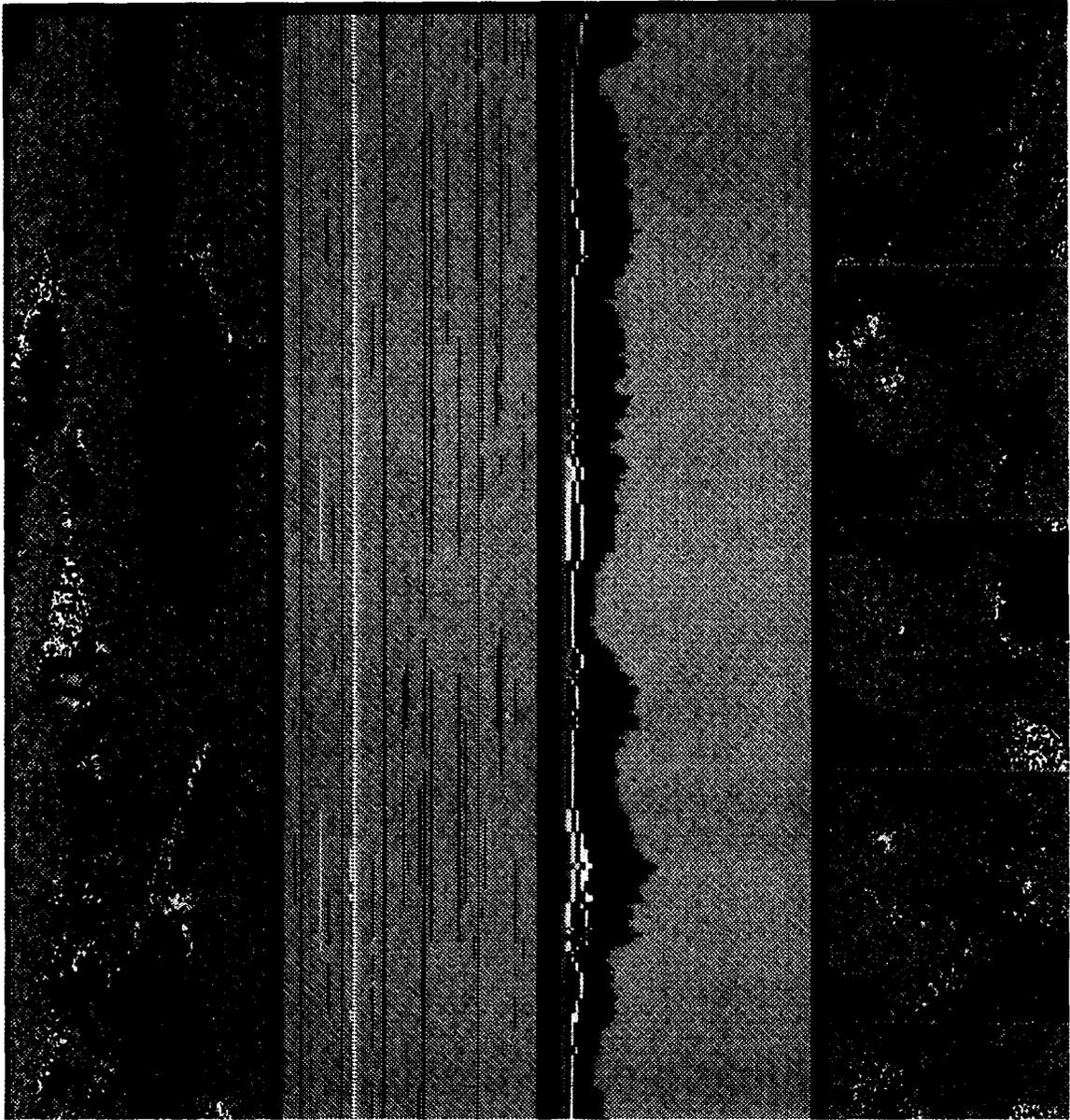


Fig. 3. Selected subnets in diffusive system with connectivity = 4/160. The encircled species are permanently present. All three are subnets of the same work. (a) Subnet selected in system with high somatic mutation rate and without a genome. (b) Subnet selected in system with genome without somatic mutations. (c) For comparison: subnet selected in the same system under global mixing.

not affect the former. The surprising stability of systems consisting of subnets interacting through a common parasite was also noted by [3].

General properties and circumstances in which they are formed. In these systems with a permanent core which tolerate and contain parasitic invasions, the problem of maintaining a large (instantaneous) diversity is ‘solved’ by selecting an attractor with a large domain of attraction,

but with slow return dynamics. This type of ‘solution’ is found mainly in diffusive systems with a genome, and occasionally in diffusive systems without a genome and with a high rate of somatic mutations. In order to be generated an influx of information and selection at the population level are necessary in the early stages. Once formed they are stable in absence of population level selection; the core is even



Color plate 1. Attractor of subnet selected in diffusive system with high rate of somatic mutations (Fig. 3a). (a) (lefthand panel): space-time plot of crosssection of CA; vertical axis: time (downwards); horizontal axis: space. (b) (2nd panel): species profile through time. (c) (3rd panel): number of species present through time. (d) (righthand panel): 4 snapshots of CA taken at time of their lowest edge.

maintained without information influx from the genome or through somatic mutations. Although the populations converge to such a self-maintaining attractor, they do contain a large number of variants (created by genomic mutation or by

recombination (fusion)) which are not long term stable and are eliminated by population level selection. These non-stable variants decrease the mean number of species per CA; therefore the mean diversity (whether measured at any one



Color plate 2. Attractor of subnet selected in diffusive system with genome (Fig. 3b.) The figure shows 3 space/timeplots starting in the same initial state (i.e. the state of one CA after 25 000 timesteps of selection). (c) (righthand panel) shows the space/time pattern in the full system, with genomic influx. (b) (middle panel) shows the space/time pattern in the system when the genomic influx is blocked (only permanent species survive). (a) (lefthand panel) shows the space/timeplot in the system when genomic influx is blocked and the initial spatial pattern is randomized. Only an improvised form survives in which the 2 subnets are connected via a common parasite as species 16 is not present (see Fig. 3b).

time, or over a period of 5 000 time steps) is not much larger in these attractors than in the case of the invariant attractors occurring in well-mixed systems although the repertoire used in both the stable and the unstable individuals far

exceeds the repertoire of the invariant systems (see fig. 2).

Diffusive systems without a genome and low rates of somatic mutations were not observed to form selfmaintaining systems of this type: they

still rely on population level selection for their maintenance after the observation period and the mean diversity in these populations does not exceed 10 species. Containment of invasions by spatial localization is, of course, not possible in well-mixed systems. Nevertheless a limited form of allowing and containing invasions was observed in well-mixed systems as discussed above. In both cases the containment of invasions depends on both the availability of enough information to do so (i.e. by secondary ‘harmless’ invasions) and on the avoidance of some types/magnitudes of invasions (by selection at the genomic level).

(3) Percolation and long transients.

High connectivities with and without pattern formation allow the system to wander all over the 160 dimensional state space if information preservation is ensured, i.e. under high mutation rates or when enough information has accumulated in the genome. Thus the total amount of information is accessed over time in both the stirred and in the diffusive system. In these systems we have crossed a ‘percolation threshold’. New mutants can always enter the system, and so diversity is maintained. In the diffusive system up to more than 30 types of molecules are present simultaneously in more than 50 copies in a 100×100 field, whereas in a completely mixed system not more than 12 molecules are present at any one time. Dynamically information is not preserved in either case: none of the species is present all of the time. At a high somatic mutation rate percolation-like behaviour also occurs at relatively low connectivities: in those cases however the trajectory is eventually trapped in a state which is not invadable anymore, e.g. dies out.

8. Cycle lengths in fixed random networks

The experiments reported in this paper are intended as a baseline for future research. The closed universe of random interactions limit its

potential for selecting specific network topologies: e.g. the distribution of minimal cycle lengths is determined by connectivity of the entire network of 160 species. By selecting subsets of species the profile of cycle lengths can be modified but only to a limited extent. The transitions between attractors with and without a permanent core occurs for connectivities between 6 and 9, i.e. where the majority of species attains a minimal cycle length of 2–3, instead of 3–4 or longer as is the case for lower connectivities. The interaction topologies of the invariant attractors of well mixed systems always consists of interconnected cycles of length 3–4, with some short parasitic chains attached. The permanent core of the invadable attractors of the diffusive systems are dominated by 2-cycles, notwithstanding the scarcity of 2-cycles in the overall network for the (relatively low) connectivities in which they are found. Notwithstanding these opposite ‘choices’ of the well-mixed and the diffusive system, stability of both appears to depend on the absence of a too strong (self reinforcing) parasitic ‘shadow’. In the well mixed percolating systems species occurring simultaneously form parasitic chains but most often do not form cycles: indeed cycles are not expected for subsets of ca. 10 species for the connectivities used (maximal 10/160). In the diffusive percolating systems diversity is large (up to 30) and cycles do occur, but none of them is resistant to invasions.

In future research we will use pattern matching between ‘genotype’ or ‘phenotype’ (e.g. secondary structure) as the basis of interaction between replicators. This enables us to extend the number of species, without necessarily changing the profile of cycle length drastically; e.g. a new species ‘similar’ to an existing one will have similar catalytic connections and thus add connectivity without altering cycle lengths.

9. Discussion and conclusions

One may object to our experiments that our

systems ‘do not do’ anything with the information which they maintain. This is indeed the case. However, I think that maintenance of information may be the limiting factor. Information which can be preserved will, because of its preservation, be ‘selected’ as substrate for the maintenance of other information, and thus obtain, secondarily, a ‘function’. Because biotic systems do not have any *a priori* goals they can simply do what they are able to do.

We have shown that maximalization of the number of species in replicator systems can be achieved in three different ways:

- (1) invariance and isolation;
- (2) permanence and contained invasions;
- (3) percolation and long transients.

Diversity by invariance and isolation is selected in well-mixed systems and, once formed, depends for its maintenance often only weakly on influx and higher (population) level processes; indeed its stability is mainly hampered by population level recombination. Only a small repertoire of species is expressed. The concept ‘evolutionary stable strategy (ESS)’ [15] refers to population level processes. The isolistic ‘evolutionary stable strategy’ operating through regulation and intra-compartment competition, found here, appears to be an alternative which deserves to be used as search image in the study of biotic evolution. This ‘strategy’ may be responsible for the occurrence of species which have maintained their characteristic over excessively long periods of time (‘living fossils’). Interestingly these often occur in circumstances where we would expect relatively high mutation rates. As we have shown high mutation rates favour this type of attractors in well mixed low connected systems.

Long transients and percolation occur in systems with high connectivities and/or under high nonselective influx. Even if percolation leads eventually into a sink, population level selection may maintain relatively high diversity in such systems due to the long transients. Although percolation indeed ‘solves’ the problem of main-

taining high diversity according to our fitness criterion as well as according the total number of species present, its inability to focus the system on some permanent information seems to render it a less attractive solution for evolutionary systems. Whether selection for permanent information can attain species rich attractors with a preserved ‘core’ in the circumstances which now selects percolation is an open question.

Intermediate between invariance and isolation on the one hand and percolation on the other hand, the ‘solution’ of maintaining diversity through contained invasions while preserving a permanent autocatalytic core seems the most powerful strategy found. It is selected always in diffusive systems with a genome and low to intermediate degree of connectivity and sometimes in diffusive systems without a genome but with relatively high mutation rates. In these systems the existence of spatial diversity permits and constrains temporal diversity and vice versa. For assessing the persistence of such systems invadability criteria, and related concepts, are useless (compare [4]). This ‘strategy’ employs all selection levels (when available): genomic, replicator network, meso-scale patterns and CA population. It generates as *side-effects* of maximising our fitness criterion (i.e. a large number of different replicators present at any one time), a large repertoire of species over longer periods, and processes reminiscent of adaptability, excitability and signal transduction: i.e. properties which are often taken as (co)-‘defining’ living systems.

Acknowledgements

I thank Maarten Boerlijst for his contributions to understanding evolutionary consequences of pattern generation in replicator systems. Mike Dale reviewed the manuscript and gave useful suggestions. The conceptual framework leading to this research is shaped by my long term collaboration with Ben Hesper.

References

- [1] M.C. Boerlijst and P. Hogeweg, Spiral wave structure in prebiotic evolution: hypercycles stable against parasites, *Physica D* 48 (1991) 17–28.
- [2] M.C. Boerlijst and P. Hogeweg, Selfstructuring and selection, in: *Artificial Life II*, C.G. Langton, C. Taylor, J.D. Farmer and S. Rasmussen, eds. (1992) pp. 255–276.
- [3] M.C. Boerlijst and P. Hogeweg, Evolutionary consequences of spatial pattern formation, in: *Spatio-temporal organization in nonequilibrium systems*, S.C. Muller and T. Plesser, eds. contributions to the Dortmund Dynamische Systemen Woche, June 1992 (Project Verlag, 1992) pp. 26–28.
- [4] M.C. Boerlijst, M.E. Lamers and P. Hogeweg, Evolutionary consequences of spiral waves in a host parasitoid system, *Proc. R. Soc London Ser. B* 253 (1993) 15–18.
- [5] H.N. Comins, M.P. Hassel and R.M. May, The spatial dynamics of host parasite systems, *J. Anim. Ecol.* 61 (1992) 735–748.
- [6] M. Eigen and P. Schuster, *The hypercycle: A principle of selforganization* (Springer, Berlin, 1979).
- [7] W. Fontana, Algorithmic chemistry, in: *Artificial Life II*, C.G. Langton, C. Taylor, J.D. Farmer and S. Rasmussen, eds. (Addison Wesley, Reading, MA, 1991) pp. 159–209.
- [8] M.P. Hassel, H.N. Comins and R.M. May, Spatial structure and chaos in insect population dynamics, *Nature* 353 (1991) 255–258.
- [9] J. Hofbauer and K. Sigmund, *The theory of evolution and dynamical systems* (Cambridge Univ. Press, Cambridge, 1988).
- [10] P. Hogeweg and B. Hesper, Evolution as pattern processing: TODO as a substrate for evolution, in: *From Animals to Animats*, J.A. Meyer and S.A. Wilson, eds. (MIT Press, Cambridge, MA, 1991).
- [11] P. Hogeweg and B. Hesper, Evolutionary dynamics and the coding structure of sequences: Multiple coding as a consequence of high mutation rates, *Comput. Chem* 16 (1992) 171–182.
- [12] M.A. Huynen and P. Hogeweg, Pattern generation in Molecular evolution; exploitation of the variation in RNA landscapes, *Mol. Ecol.* submitted; in: Thesis Utrecht University (1993).
- [13] S.A. Kauffman, *The Origins of Order; selforganization and Selection in Evolution* (Oxford Univ. Press, Oxford, 1993).
- [14] C. Lindgren, in: *Artificial Life II*, C.G. Langton, C. Taylor, J.D. Farmer and S. Rasmussen, eds. (Addison Wesley, Reading, MA, 1991) pp. 295–312.
- [15] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge Univ. Press, Cambridge, 1982), 224pp.
- [16] M. Nowak and R.M. May (1993) The spatial dilemmas of evolution, *Int J. of Bifurcation and Chaos* 3 (1993) 35–78.
- [17] T.S. Ray, An approach to the synthesis of life, in: *Artificial Life II*, C.G. Langton, C. Taylor, J.D. Farmer and S. Rasmussen, eds. (Addison Wesley, Reading, MA, 1991) pp. 371–408.
- [18] T.S. Ray, Evolution, complexity, entropy and artificial reality, *Physica D* 75 (1994) 239, these proceedings.
- [19] P. Schuster and K. Sigmund, *Replicator dynamics*, *J. Theor. Biol.* 100 (1983) (Reading, MA, 1991) 533–538.
- [20] K. Sims, Poster 1st European conference on Artificial Life (Ecal I), Paris (1991).
- [21] P.F. Stadler and P. Schuster, Mutation in autocatalytic reaction networks, an analysis based on perturbation theory, *J. Math. Biol.* (1992) pp. 579–631.
- [22] P.F. Stadler, W. Fontana and J.H. Miller, Random catalytic reaction networks, *Physica D* 63 (1993) 378–392.
- [23] T. Toffoli and N. Margolus, *Cellular Automata machines: a new environment for modelling* (MIT Press, Cambridge, MA, 1987).
- [24] S. Wolfram, Universality and complexity in cellular automata, *Physica D* 10 (1984) pp. 15–57.
- [25] C. Lindgren, Evolutionary dynamics of spatial games, *Physica D* 75 (1994) 292, these proceedings.