On searching generic properties of non generic phenomena:
an approach to bioinformatic theory formation

Paulien Hogeweg
Bioinformatics Grp, Utrecht University
Padualaan 8, 3584CH Utrecht, the Netherlands.
Email ph@binf.biol.ruu.nl

Abstract

In this paper we first shortly review the current view of the evolution of complexity and novelty in biotic evolution. Next we show that the basic processes thereof do happen automatically and are generic properties of systems including the basic mechanisms of Darwinian evolution *plus* **local** as opposed to global interactions. Thus we show that the so generated multilevel evolution can be studied within the paradigm 'simple rules lead to complex phenomena'. We derive some results demonstrating the power of such multilevel evolutionary processes to integrate information at multiple space and time scales.

Nevertheless we also point out shortcomings of such an approach which necessarily uses a priori chosen and preferentially relatively simple interaction schemes. However, straightforward extensions towards more complex interaction schemes generally leads to ad hocness and over-determinedness, rather than fundamentally new behavior of the system, and often to less understanding of that behavior. Nevertheless biological theory formation needs a method to go beyond the generic behavior of simple interaction schemes.

We propose to use evolutionary optimization of very trivial fitness functions which are obtainable in many different ways to push back the necessary a priori choices and to zoom in to interesting non generic phenomena and their general properties. . We thus derive insights in relationships between sets of derived properties at several scales. We discuss how this approach can be used in biological theory formation. focusing on information accumulation and utilization in replicator systems and immune systems.

Introduction

Reasoning from a chemical point of view, de Duve (1995) portrays *'life as a cosmic necessity'*. Maynard Smith and Szathmáry (1995b; 1995a) , reconstructing the course of evolution, conclude that a limited number of major transitions shaped living systems as we know them today, and that these major transitions involved the processes of symbiogenesis, conflicts among levels of selection, division of labor and the transition from limited inheritance to universal inheritance. Studying evolution from a bioinformatic point of view, we have shown that the first three of these major transition defining processes are generic consequences of extending basic mutation and selection with local interactions. Thus we can also portray *'life as a local necessity'*.

Nevertheless, due to inheritance based information accumulation, we can hardly study e.g. an elephant as a generic property of matter or information: many of its properties appear to be arbitrary accidents. Although indeed chance is an inalienable part of life, there may be stronger constraints than now appears. Biological modeling usually either focuses on those phenomena which are 'generic', or simply aims at mimicking properties observed in a particular system. For better understanding biotic systems we have to face the difficult question how we can obtain generic theories of non generic phenomena.

In other words we usually study either how complex behavior is generated from simple rules, or how simple (in the sense of a priori definable) behavior is generated by complex rules. Understanding biological systems requires that we also face the difficult question of studying complex behavior generated by complex rules, without getting lost in arbitrary over-determinedness.

In this paper we present one approach for doing this. It involves focusing on 'side effects' of evolutionary optimization where the optimization criterion is extremely 'uninteresting', and can better be seen as a minimal condition than as 'goal'. We present two examples in which we employ our approach. Using diversity of entities as optimization criterion, we derive relationships between the topology of catalytic networks, self-structuring and information storage and utilization: self-structuring is a prerequisite for information stor-

age and utilization. Using recognition of pathogens as optimization criterion, we derive a relationship between genetic operators and immune system diversity, and thus obtain a hypothesis to explain difference between vertebrate and invertebrate immune systems. In all cases the observed patterns can only be observed in the evolved systems because the 'random' initial condition of the evolutionary optimization displays none of the features which we like to study.

Evolution of complexity in biotic systems

Biotic systems are multilevel systems. Indeed the interplay between partial independent processes at many space, time and organizational scales appears to be the preeminent hallmark of biotic complexity. Classical population genetic and evolutionary theory does not address the generation of complexity, and indeed its occurrence does not seem to follow automatically from a 'survival of the fittest' point of view (and indeed biotic systems can also become simpler in evolutionary time). Nevertheless complex multilevel systems did arise.

Reconstructing biological evolution Maynard Smith and Szathmáry (1995b; 1995a) conclude that such complexity arises by 'major transitions in evolution' of which the basic ingredients are:

1. Symbiogenesis, i.e. the process in which independent replicators give up their self-sufficiency and become 'parts of a whole'. Examples include eukaryotic organelles (mitochondria,chloroplasts) which evolved from prokaryote precursors, worker castes in social insects etc.

2. Conflicts among levels of selection. In the wake of symbiogenesis conflicts of levels of selection can arise, where the parts of a whole re-evolve partial independence, and deteriorate the 'whole'. Evolution of uniparental inheritance of organelles appears as one of many 'countermeasures' to such a process,

3. Division of labor, by which 'tasks 'or 'functions' initially performed by one type of entity, are later subdivided among a number of 'specialized' entities. This process occurs again at many levels: differentiation in 'germ-line and 'soma', and again in social insects in the evolution of specialized worker casts.

4. Transition from limited inheritance to universal inheritance, and therewith the generation of 'universal' coding schemes. As examples Maynard Smith Szathmáry (1995b) list the evolution from autocatalytic sets to template based replication. in early stages of evolution, the evolution of a dual inheritance system in metazoans and transition from signal based communication to universal grammar based languages in the later stages of evolution.

In their treatment they recognize the universality of these processes from a chemical and natural history perspective, and discuss some of the evolutionary consequences and constraints GIVEN the occurrence of these processes. As one of the important premises they take that any hypothesized intermediate structure should be selectionally advantageous *in the short run*.

Similar conclusions about major transitions were also independently (and earlier) derived by Fontana and Buss (1994b; 1994a) both on theoretical and natural history grounds. In their treatment they stress that it are especially these processes which 'would be repeated in alternate 'worlds". They studied random *non self-replicating* metabolic nets and show that the novel interactions are automatically created when independently evolved catalytic networks are brought into contact. Thus they derive 'organization for free'. However, because they utilize global interactions, only one 'organization' exists and the mechanisms of Darwinian evolution are absent.

In the next section we review results of replicator networks subject to simple Darwinian mutation selection processes in space and show that in space the processes associated with the major transitions are an automatic consequence of mutation and selection, due to the generation of higher levels of selection due to spatial self- organization. Moreover we show that in multilevel systems the assumption of necessary short term fitness benefits is relaxed.

Local interactions and the emergence of multiple levels of selection.

Self-structuring is an ubiquitous property of locally interacting systems. The relationship of self-structuring and Darwinian evolutionary processes can be seen in terms of

(a) self-structuring as constraint on achievable structures

(b) self-structuring as alternative to Darwinian selection

(c) self-structuring as a substrate for Darwinian selection.

The latter point of view is in our view the most fruitful. We will show that through self-structuring multiple levels of selection arise with novel interactions and novel fitness dimensions. In this sense self-structuring enhances rather than constrains the power of Darwinian evolution. (Whether or not it does so for engineering purposes is an open question, although some examples suggest that it does (Hillis 1992; Pagie & Hogeweg 1998b)).

The simplest way of defining an evolutionary process is to define some set of predefined interactions between replicators and subject one (or a few) of the parameters of the system to mutations (selection automatically ensues from the dynamics of the system). Evolutionary systems so defined can neither redefine their interactions nor redefine their genetic representations, both of which are important in open ended evolution. The dynamics of the system can however redefine the fitness of the replicators, which is also crucial for open ended evolution. Indeed we have shown that because of the latter feature, even such simple evolutionary systems can give rise to processes akin to those recognized by Maynard Smith and Szathmáry (1995b) as associated with major transitions in evolution, provided that the interactions between the replicators is defined locally. This is because local interactions will lead to the the formation of higher level structures (e.g. spiral waves, turbulence, patch like structures of different sizes) which constitute new levels of selection. We have shown this in a variety of coevolutionary systems, evolving e.g. strength of 'help' (e.g. catalysis) in cooperative systems (Boerlijst & Hogeweg 1991a; 1991b; Couwenberg & Hogeweg 1998), predation efficiency (Boerlijst, Lamers, & Hogeweg 1993; Savill & Hogeweg 1997) dispersal rate (Savill & Hogeweg 1998) or parasitoid aggregation strength (Savill, Rohani, & Hogeweg 1997). Here we summarize the major conclusions of this work.

- **Feedback of mesoscale entities on microscale entities through multilevel selection**
  The micro-scale replicators generate mesoscale entities, but, via mutation and selection, the reverse is also true: the mesoscale entities generate the microscale entitles by which they are made. This may lead to and maintain microscale entities which are non viable or less viable without the mesoscale entities. For example the microscale entities may evolve to a shorter lifespan, or minimize the catalysis it gets for self-replication, because this enhances the competitive strength of the mesoscale entity which they generate. This was shown in cases where oscillatory dynamics between the microscale entities leads to spiral waves in spatial systems. These spiral waves compete for space, and the fastest rotating spirals win. Shorter lifespan and less catalysis leads to faster rotating spirals and hence to microscale entities with these (seemingly) unfavorable properties. (Boerlijst & Hogeweg 1991a; 1991b).

- **Self-enhancement of mesoscale entities**
  The direction of selection enforced on the microscale

entities appears to be such that it tends to favor the competitive strengths of the mesoscale entities they generate. For example in a host parasitoid system (Savill, Rohani, & Hogeweg 1997), regions of spiral waves and regions of turbulence occur. In the spiral wave area the aggregation parameter evolves to lower values which favors the formation of spiral waves over turbulence and the reverse is true in the turbulence areas: there aggregation evolves to higher values for which spiral waves can not be formed or even maintained.

- **Symbiogenesis**
  The properties of local interacting, evolutionary systems mentioned in the items above embody a process reminiscent of 'Symbiogenesis' in that self-sufficiency is (partly) given up in favor of the larger scale entities.

- **Conflicts between levels of selection**
  Conflicts between levels of selection are inherent in this process; in fact it is the reversal of the direction of selection which within the system, defines the emergence of a new level of selection.

- **Division of labor**
  Division of labour is also inherent in the formation of mesoscale entities. All entities 'do only what there is to do' and for the microscale entities this depends on the position they occupy in the mesoscale entities. For example in spiral waves only the entities in the core of the spiral in the long run produce offspring, those in the spiral arms become extinct. Thus a kind of 'germ-line' and 'soma' differentiates.

- **Direct vs Indirect interactions**
  In contrast to globally defined ecoevolutionary system in which all replicators are interacting and competing with equal probability with all other replicators , in spatial systems with only locally interactions 'who out-competes whom' can not simply be assessed on the basis of a few of the direct links (those in which the competing entities are directly involved) of the interaction network of which they are part: all links may play a role as they help to define the mesoscale patterns and their properties. Thus local interactions in space generate new indirect interdependencies of the system.

- **Short term observations are insufficient to asses long term fitness** Short term fitness and long term fitness may be quite different. This is shown in fig 1. for the host parasitoid system studied by (Savill, Rohani, & Hogeweg 1997). Host parasitoid interactions are modeled by a spatial extension

of the classical Nicholson Bailey equations. Parasitoid aggregation strength is the evolving parameter. Over a time-span of up-to 50 generations (left panel) parasitoids with a larger aggregation tendency produce more offspring than those with less aggregation tendency, wherever they are located in space. Nevertheless in the long run (right panel) those with the weakest aggregation tendency 'inherit the world'. It is caused by the above mentioned property of spiral waves that only the entities in the core of spirals will give rise to offspring in the long run. Note that this is a property of the 'attractor' of the system and true for any short term vs longterm time-slice.

Long term information integration in evolutionary processes also is strikingly apparent in experiments which use coevolving populations of 'problems' and 'solutions' for function optimization (Hillis 1992; Pagie & Hogeweg 1998b) : the availability of only very sparse information per generation of the function to be optimized, even improves the chance of obtaining the globally correct solutions.

Experimental biologists surely can not be envied: assessing inclusive fitness seems a time consuming and in fact impossible task.

We conclude that even in these systems with inflexible information storage and transmission, major transitions involving the generation of new levels of selection is a basic property of mutation/selection processes *iff* interactions are local (as they indeed necessarily are!) Thus, as an addition to de Duve's assessment on chemical grounds of 'life as a global imperative', with respect to its multilevel properties we can regard 'life is a local imperative'.

We like to stress that the above listed properties all (a) were quite unexpected from a simple minded 'survival of the fittest' point of view, and appear to enhance the 'versatility' of Darwinian selection processes, in the sense that they they go beyond optimization of the behavior of predefined entities, by generating novel ones, and complex interactions between different levels of selection.

(b) can provide explanations for as yet not understood observations on many specific biological systems. For example, recently we have studied spatial eco-evolutionary systems which provide explanations for the shape of influenza phylogenies, which show stagnating evolution (flat phylogeny) for virus strains taken from birds, and rapid, progressive evolution (steep phylogeny) for virus strains taken from pigs or humans. Assuming a shorter lifespan, or shorter immunological memory for birds than mammals (or a smaller universe) the striking difference in phyloginies arises due to self-structuring into two-armed spiral waves in birds
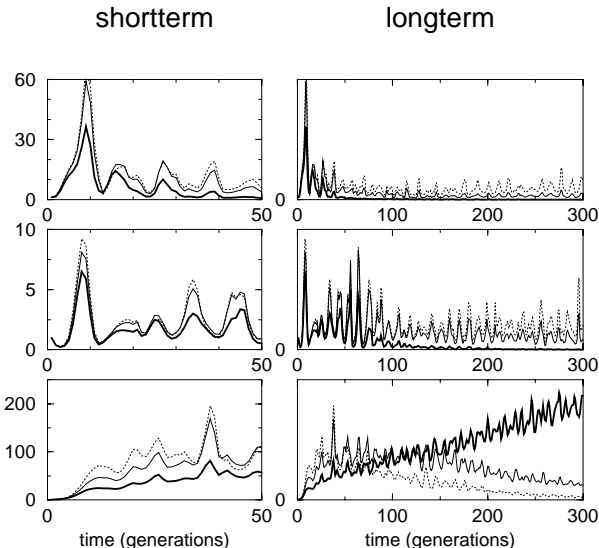

Figure 1: **Short term and Longterm fitness**
The left panels show short term fitness (50 generations) and the right panels longterm fitness (300 generations) at 3 different locations in space (top in spiral arm, middle in chaotic region, bottom in spiral core.) The vertical axis gives average number of offspring per individual parasite with different aggregation tendency towards their host: thick line low-, thin line intermediate - and dotted line high aggregation tendency.

At all locations high aggregation tendency gives the highest number of offspring over a time-slice of 50 generations. Nevertheless in the long run the fitness of parasites with the lowest aggregation tendency is highest. Note that this is the case in the eco-evolutionary attractor of the system, i.e. for any time-slice of 50 vs 300 generations.

vs chaotic waves in mammal. (Hogeweg 1998). An other example is the occurrence of a large variety of toxic plasmids in bacteria and fungi which has puzzled micro-biologists and has led to hypotheses on additional (as yet unidentified) functions of those plasmids which might benefit their hosts. We have shown that such diversity of plasmid is an automatic consequence of the information integration capabilities of local interactions in spatial systems (Pagie & Hogeweg 1998a).

(c) are a direct consequence of the formation of 'generic' patterns, e.g. spiral waves, turbulence, patches of different sizes etc.

Thus notwithstanding the novel insights in the dynamic potential Darwinian selection processes in space, they do not give us 'novel' entities, as biotic evolution undoubtedly has (e.g. elephants). Thus, we have not yet surpassed the stage for which Maynard Smith urged all evolutionary biologists to go once a year to the zoo, stand in front of the elephant and proclaim: 'elephant I believe you came about by random mutation and selection', even if now we can add 'plus local interactions creating new levels of selection'.

Beyond generic patterns: complex to complex mappings

Simple rules may give rise to complex behavior. This was an interesting issue 10 years ago (at the first AL-IFE conference) and has now become common place. In the time since then we have seen that evolutionary optimization most often leads to complex implementations when it is free to choose its implementation as in Genetic programming, even in the case when fairly simple coding would be possible.

Observation of biotic systems suggests a complex to complex mapping. Studying such complex to complex mappings without getting lost in over-determinedness and ad hocness therefore seems an important challenge for biological theory formation.

Obviously in the above observations simple and complex are not well defined, but can be operationalized with respect to our own (in)ability and/or willingness to a-priori conceive, make, understand or define it, although we may be able to observe it once it is there.

A complex to complex mapping thus means that neither the micro rules nor the macro behavior is a priori definable, whereas in the paradigms above at least one of them necessarily is a priori 'in hand'. and therefore restrict us to studying complex behavior which is generic for *some (a priori chosen)* set of simple rules, or studying implementations for *some (a priori chosen)* simple functions.

The most important contribution of the simple to complex paradigm for bioinformatic modeling of specific systems is the possibility (and in fact necessity) of 'non goal oriented modeling', i.e. a modeling approach in which we do not specify a priori the phenomena to be model-led. This is a necessary consequence of simple to complex paradigm, as explicit modeling efforts of the complex behavior would seldom lead us to very simple rules. In a non goal oriented modeling paradigm we therefore formulate a set of simple rules which aim to implement only some of the context in which the complex behavior in which we are interested occurs. Observing the so obtained systems we search for side-effects of the rules which among quite expected and unexpected phenomena may also represent some (traces of) of the phenomena we were interested in to begin with, and point at connections (via the simple rules) between the observed a priori apparently independent phenomena (Hogeweg & (1989) 1989; Hogeweg 1988). Examples are given in (Hogeweg & Hesper 1985; te Boekhorst & Hogeweg 1994a; 1994b) .

The approach to complex to complex mapping we propose is an extension of this modeling methodology which additionally uses evolutionary optimization to zoom in to non generic 'initial' conditions or 'not-so-simple' rules. The optimization criterion to be used in the evolutionary optimization should: (1) not represent directly the phenomena in which we are interested but only represent some kind of boundary condition for them and (2) should be realizable in (many) different ways. Moreover the coding used in the evolutionary optimization should be such that it is to some extent 'free' to choose a realization. We study the side effects of the so obtained 'not so simple' rules and 'not-so-general' initial conditions by observing the so obtained systems in a similar way as above.

We have applied this idea to investigate issues related to the potential for and the role of (more or less (un)limited) inheritable information accumulation and utilization, This relates to the fourth ingredient of the major transitions in evolution listed by Maynard Smith and Szathmáry (1995b), which was not addressed in our 'simple to complex' experiments discussed above (which did nevertheless display the three other ingredients). We discuss these experiments and the results obtained in the next section.

Bioinformatic theory formation of non-generic phenomena

We discuss two examples of using evolutionary optimization towards a rather 'uninteresting' target to zoom in on 'interesting' systems for bioinformatic theory formation. The background of the first example is basic bioinformatic theory related to complex repli-

cator networks (Eigen & Schuster 1979; May 1972) and the above discussion on the role of selfstructuring in evolution, whereas the second example relates to a more specific question and apart from addressing our questions about (un)limited heredity, confronts our approach to the more usual approach in theoretical biology.

The potential role of DNA in an RNA world:
unlimited inheritance needs self-organization and multiple levels of selection

In contrast to metabolic networks of non-selfreplicators (compare (Fontana & Buss 1994b; 1994a) networks of replicators generically cannot maintain high diversity. Exceptions are specific interaction topologies and systems with high mutation (or influx) rates (e.g. (Kaneko & Ikegami 1992; Forst 1997)). What are the properties of those replicator networks which can maintain high diversity, and what is the role of information storage and the occurrence of multiple levels of selection in this context? Moreover what is the relation between information storage and multiple levels of selection? We studied these questions by evolving a population of CA's in each of which locally interacting networks of catalytic replicators compete, with species diversity in the CA as fitness criterion. We performed these experiments with and without allowing pattern formation, and with and without allowing information storage (in the form of long lived non catalytic counterparts of the catalytic selfreplicators, i.e. in the form of DNA in an RNA world; DNA is 'transcribed' due to catalysis by the same RNA's which catalyze its RNA transcript). For more details see (Hogeweg 1994a; 1994b) The results show:

- After long evolutionary time, the temporal persistence of the evolved species diversity is indefinite, although the lifetime over which species diversity is contributing to individual fitness is relatively short. This is due to evolution toward a relative smooth part of the landscape (Huynen & Hogeweg 1994). In other words, this mechanism results in the resolution of conflicts between levels of selection: in contrast to arbitrary sets of replicators, none of the final individual replicator species takes over in spite of the fact that the higher level selection pressure is removed.
  This indefinite persistence of the evolved networks allows us to study the evolved networks independent of the top-level evolutionary dynamics, and thus use the CA evolution as tool to obtain non-generic, but also non-ad-hoc, and interesting initial conditions for studying the diverse replicator networks and the role
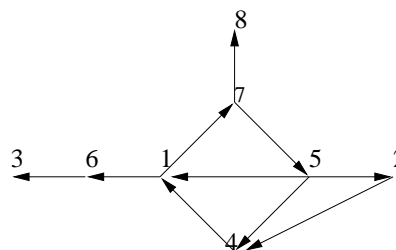

Figure 2: evolved network; node numbers correspond to to curves in next figure (bottom to top).
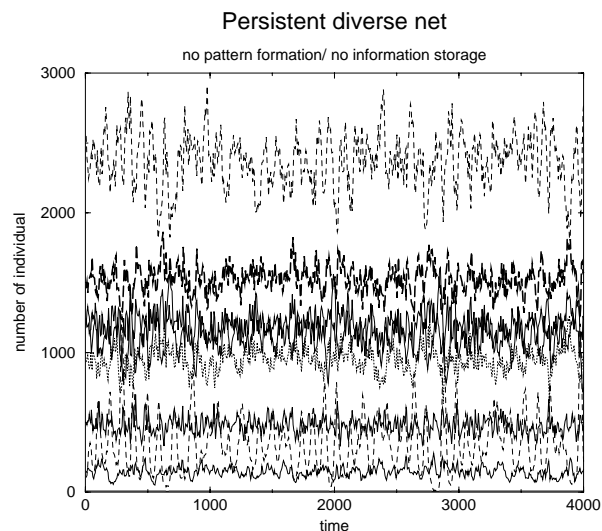

Figure 3: attractor of the evolved network; diversity is conserved also without information usage; parasitic chains do not destroy the network because of strong nonlinear catalysis, local interactions and interlocked cycles (see previous figure)

of information therein.

- The well mixed systems, i.e. those without pattern formation, evolved replicator networks with the following properties:
  (1)the networks contain a few short interlocked cycles with a few 'parasitic' chains with not more than 8-10 replicators in total. (see fig 2);
  (2) Information storage is necessary for evolution of species diversity clearly above that occurring in random networks (because it enables the simultaneous incorporation of more then 1 new species in the RNA network), but
  (3) The amount of information stored is minimized. Moreover

(4) The stored information in almost never used and indeed not needed for the maintenance of diversity. Thus all the information available is dynamically contained as the (chaotic) attractor of the RNA network (see fig 3)..

(5) Such an attractor with relatively many species is rare and is not found by random initialization even of replicators interacting with this particular network topology. Evolutionary optimization is needed as tool to find both the network and the initial conditions which lead to the attractor.

- The spatial systems with local diffusion, i.e. the systems with pattern formation evolved replicator networks with the following properties.

(1) The networks contain many more species and consist of (several) 2-cycles with long parasitic chains.

(2) Information storage is not necessary to evolve diverse networks, nevertheless when information storage is allowed,

(3) Information storage is maximized. Moreover

(4) The stored information is regularly transcribed and thus a greater variety of RNA species is maintained then contained in an attractor of the RNA network alone.

Fig 5 compares the dynamics of a network which has evolved with information storage, when this information is available or not: the 'transcription' of the stored information not only increases diversity due to the temporary presence of 'transient' RNA species, but also stabilizes the competition between mesoscale patterns, and thus allows more 'permanent' RNA species in the system.

Fig 4 contrasts the use of stored information in systems with and without pattern formation. As stated above, in the latter diversity is maintained in the attractor of the catalytic replicators, and the stored information is not (or barely) expressed, whereas the former crucially relies for its (larger) diversity on the stored information (iff available).

The major conclusion is:

In replicator systems (i.e. in the RNA world) transition from limited to unlimited inheritance is only evolutionary favored when multiple levels of selection develop due to spatial self-organization.

Genetic operators and immune repertoire diversity:
somatic recombination biases evolution towards complexity 'beyond need'

How large should the immune repertoire be? And why is it much larger in vertebrates then in invertebrates:

Diversity and Information storage
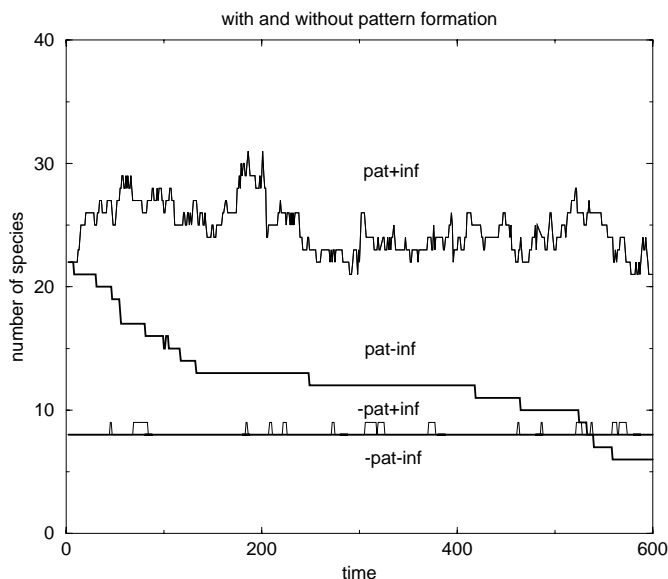with and without pattern formation

Figure 4: Dynamics of evolved networks. Information is stored during evolution, in all cases; information availability and pattern formation as shown in figure: only in the case of pattern formation is the stored information exploited

both seem to cope. Assuming random repertoires, the issue of repertoire size is previously studied in terms of allowing optimal distinction between self and non-self antigens, focusing on evolutionary time or physiological time and including a more or less explicit structure of the (vertebrate) immune system (respectively: (Percus & Perelson 1993; de Boer & Perelson 1993; Borghans & de Boer 1998). Within a partial implementation of the above sketched modeling methodology Takumi and Hogeweg (1998) studied this question by evolving immune systems using either a genetic coding as in vertebrates (i.e. including somatic recombination) or as in invertebrates (direct encoding). Using a fixed set of pathogens to be recognized by the immune system and ignoring the issue of self non-self discrimination we have shown:

- Vertebrates and invertebrates cope equally well with the same set of pathogens, but do so in a different way.

- The vertebrate immune system evolves large repertoires, and the invertebrate immune system evolves small repertoires.

- Moreover the evolved repertoires cope much better with evolving or random sets of pathogens than a

random repertoire does. This is true for both the vertebrate and the invertebrate case although the vertebrate case does somewhat better.

We conclude with respect to the issues of complex to complex mappings and limited vs unlimited inheritance:

- Answering biological questions by models which use 'random' initial conditions may lead to misleading results, because evolved systems may behave very non-random. In the case studied here:

- Questions of repertoire size (as well as immune system behavior in general) may not be answerable in terms of random repertoires: evolved repertoires behave very differently. Moreover

- Similar problems can be solved in entirely different ways. Genetic coding can determine which 'solution' is chosen, and therewith what behavior is obtained 'for free'.

- Differences in repertoire size (or other properties of biotic systems) do not necessarily reflect differences in the environment toward which they adapted, but rather are also a consequence of the way genetic information is stored and retrieved.

- Somatic recombination adds coding potential to a genome of limited length. In the above example this added potential is employed without rendering direct fitness benefit, i.e. it leads to autonomic increase in complexity which is not 'needed' in the system as defined. (but may pave the way for future evolution, in this case e.g. slef-nonself discrimination).

Discussion and conclusions

Nowadays we are not protected by technical limitations against formulating and to a certain extent studying systems including much specialized knowledge, and/or many special assumptions. The behavior observed/studied in many systems formulated accordingly, does not (or barely), go beyond that observed in more simple systems where we can study/understand the behavior much better. Moreover by studying simpler systems it is easier to focus on 'generic' properties occurring in many contexts (including more complicated systems).

We have argued that in addition to studying generic properties of minimally defined systems with random initial conditions, it is possible and seems profitable to try to derive general properties of relatively 'rare' and complicated systems, while minimizing a priori assumptions. One way of doing this is by studying the side-effects of evolutionary optimization towards a target which describes the boundary conditions of the systems in which we are interested, rather then those aspects in which we are interested. It is the ubiquity of side effects and the existence of alternative solutions which make this method fruitful. So derived 'special' but not (entirely) 'arbitrary' systems reveal general tendencies/relations absent in most arbitrary (be it knowledge poor or rich) systems, but important for the class of systems exhibiting the boundary conditions used in the optimization process. At the very least, as shown, such systems will provide counterexamples for unquestioned assumptions, which are often implicitly or explicitly derived from random (rather than evolved) systems.

In our examples we have used this method in addition to minimally defined locally interacting evolutionary systems, to study evolution itself, in particular questions on long term information storage, transmission, and utilization, i.e. inheritability. In this respect our results showed that self-structuring may be a prerequisite to exploit stored information (i.e. for the transition from attractor based ('limited') inheritance to 'storage based' ('less limited') inheritance. Moreover, when the actual use of stored information is assumed (as it usually is in evolutionary models) self-structuring allows long term information integration. Thus we conclude that self-structuring should neither be seen as an alternative for - nor as a constraint on - 'evolution' but as the substrate on which it operates. It is the interplay between self-structuring and mutation/selection processes which create the 'major transitions' (and novelty) in evolution.

Acknowledgements

References

Boerlijst, M. C., and Hogeweg, P. 1991a. Self-structuring and selection: spiral waves as a substrate for evolution. In Langton, C. G., ed., *Artificial Life II*, 255–276. Redwood City, CA: Addison-Wesley.

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

Boerlijst, M.; Lamers, M.; and Hogeweg, P. 1993. Evolutionary consequences of spiral waves in host parasitoid systems. *Proc. Royal Soc. London B* 253:15–18.

Borghans, J., and de Boer, R. 1998. How specific should memory be? (in prep).

Couwenberg, J., and Hogeweg, P. 1998. On the fate of cheaters in ecological and evolutionary timescales (in prep).

de Boer, R., and Perelson, A. 1993. How diverse should the immune system be? *Proc. R. Soc London B* 252:171–175.

de Duve, C. 1995. *Vital Dust: life as a cosmic imperative.* New York: Basic Books.

Eigen, M., and Schuster, P. 1979. *The hypercycle: a principle of natural self-organization.* Berlin, Heidelberg, New York: Springer.

Fontana, W., and Buss, L. 1994a. The arrival of the fittest: towards a theory of biological organization. *Bull. Math. Biol* 56:1–64.

Fontana, W., and Buss, L. 1994b. What would be conserved 'if the tape were played twice'. *PNAS* 91:757–761.

Forst, C. 1997. Molecular evoluton of catalysis. In Husbands, P., and Harvey., eds., *Ecal IV Proceedings,* 83–91. Boston, USA: MIT press.

Hillis, D. 1992. Coevolving parasites improve simulated evolution as an optimization process. In Langton, C. G. e. a., ed., *Artificial Life,* 313–324. Redwood City, CA: Addison-Wesley.

Hogeweg, P., and (1989), B. H. 1989. An adaptive, selfmodifying, non goal directed modelling approach. In Elzas, M.; Oren, T.; and Zeigler, B., eds., *Modelling and simulation methodology: Knowledge systems paradigms,* 77–92.

Hogeweg, P., and Hesper, B. 1985. Socioinformatic processes: Mirror modelling methodology. *J. Theor. Biol.* 113:311–330.

Hogeweg, P. 1988. MIRROR beyond MIRROR, puddles of LIFE. In Langton, C. G., ed., *Artificial Life,* 297–316. Redwood City, CA: Addison-Wesley.

Hogeweg, P. 1994a. Multilevel evolution: replicators and the evolution of diversity. *Physica D* 75:275–291.

Hogeweg, P. 1994b. On the potential role of dna in an rna world: Pattern generation and information accumulation in replicator systems. *Ber. Bunsengesel Phys. Chem* 98:1135–1139.

Hogeweg, P. 1998. Spatial selforganization and the shape of phylogenetic trees (in prep).

Huynen, M., and Hogeweg, P. 1994. Pattern generation in molecular evolution: exploitation of the variation in rna landscapes. *J. Mol.Evol* 39::71–79.

Kaneko, K., and Ikegami, T. 1992. Homeochaos:dynamic stability of a symbiotic network with populaton dynamics and evolving mutation rates. *Physica D* 56:406–429.

May, R. 1972. *Stability and complexity in model ecosystems.* Princeton: Princeton Univ. Press.

Maynard Smith, J., and Szathmáry, E. 1995a. The major transitions in evolution. *Nature* 374:227–232.

Maynard Smith, J., and Szathmáry, E. . 1995b. *The Major Transitions in Evolution.* Oxford: Freeman.

Pagie, L., and Hogeweg, P. 1998a. Coexistence of colicines and cost of immunity (in prep) .

Pagie, L., and Hogeweg, P. 1998b. Evolving adaptability due to coevolving targets. *Evolutionary computation (in press).*

Percus, J.K, O. P., and Perelson, A. 1993. Predicting the size of the t-cell repertoire and antibody combining region from consideration of efficient self-nonself discrimination. *PNAS USA* 90:1691–1695.

Savill, N. J., and Hogeweg, P. 1997. Evolutionary stagnation due to pattern-pattern interactions in a co-evolutionary predator-prey model. *Artificial Life* 3:81–100.

Savill, N. J., and Hogeweg, P. 1998. The evolution of dispersal in predator-prey waves: Speciation prevents extinction. *Proc. Royal Soc. B* 265:25–32.

Savill, N. J.; Rohani, P.; and Hogeweg, P. 1997. Self-reinforcing spatial patterns enslave evolution in a host-parasitoid system. *J. theor. Biol.* 188:11–20.

Takumi, K., and (1998), P. H. 1998. Evolution of the immune repertoire with and witout somatic dna recombination. *J. Theor. Biol (in press).*

te Boekhorst, I. J. A., and Hogeweg, P. 1994a. Effects of tree size on travelband formation in orang utans: Data analysis suggested by a model study. In Press, M., ed., *Artificial Life IV,* pp 119–129.

te Boekhorst, I. J. A., and Hogeweg, P. 1994b. Self-structuring in artificial 'chimps' offers new hypotheses for male grouping in chimpanzees. *Behaviour* 130:229–252.
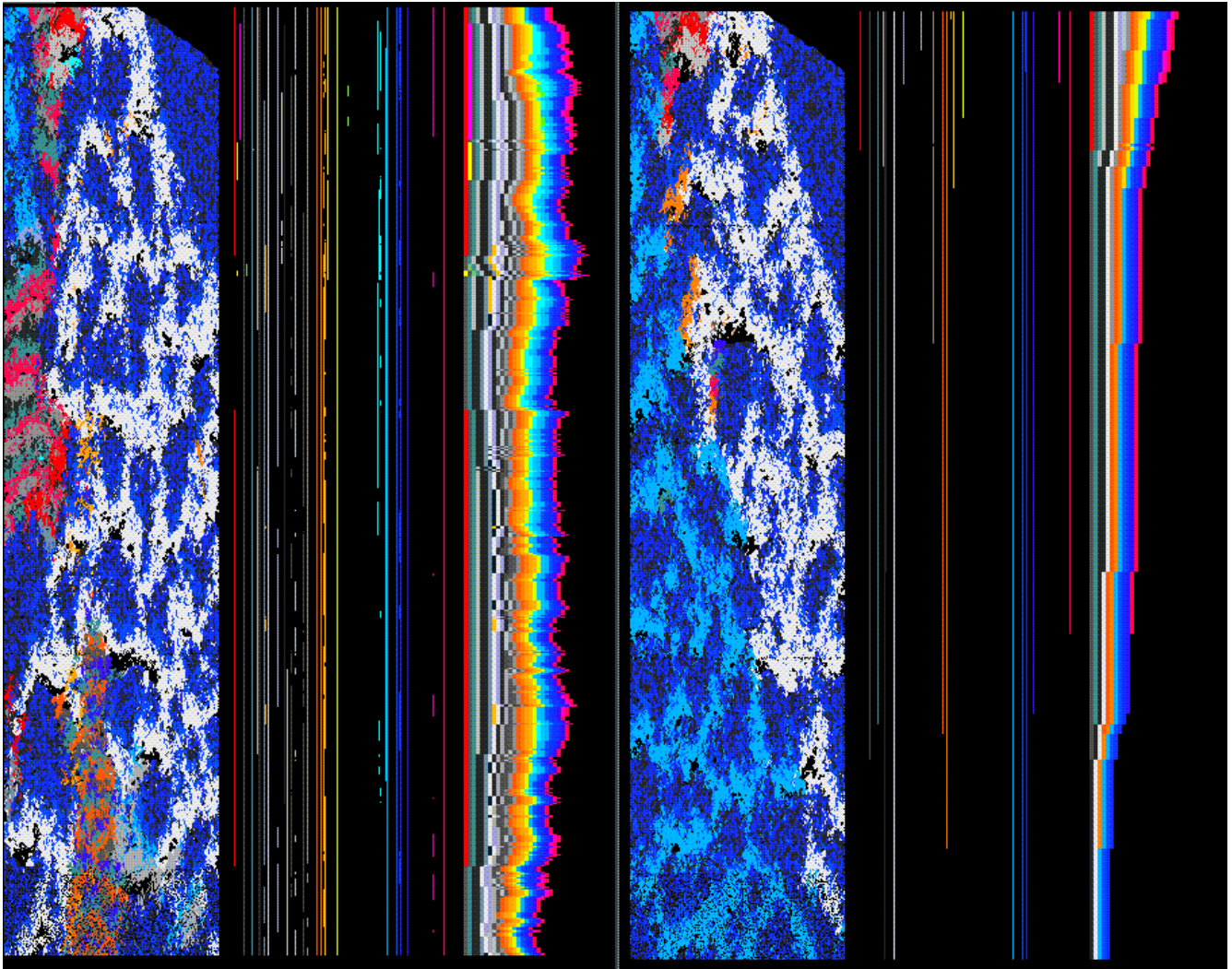
Figure 5: Spatial selforganization and the use of stored
information;
interaction network evolved with pattern formation
and with information storage.
three left hand panels: the stored information is avail-
able; diversity is high.
three righthand panels: the stored information is NOT
available; diversity is lost.
(left: space-timeplot; middle: profile of RNA species
present; right: number of species)