

# 6

## Discussion

This thesis describes a number of model-studies of evolutionary systems. All models are characterised by the concurrent evolution of a number of different types of replicators. Thus, in one form or another, the systems comprise coevolutionary processes. The genotypes of the replicators show different degrees of structure, ranging from simple but variably sized genotypes to genotypes composed of open-ended, functional structures with a non-linear genotype-phenotype mapping. Also, all model-studies are embedded in a spatial setting and incorporate different modes of local or global mixing.

During the studies described in the previous chapters we developed some concepts we considered to be important for understanding the evolutionary processes that occurred in the model-systems. In this final chapter of the thesis these concepts are discussed again and reviewed in the broader context of the thesis. First, however, we will give a short overview of the studies in which the different concepts were developed and thereafter we will discuss the concepts more thoroughly.

### 6.1 A short review

Chapter two is a model-study of coevolution in the context of evolutionary optimisation and introduces the concept of sparse fitness evaluation. In the model, the evolutionary process is driven by an externally defined, artificial fitness function. We compared the evolutionary dynamics resulting from a fitness evaluation procedure based on a full set of ‘fitness cases’ to fitness evaluation based on a small, coevolving set of ‘fitness cases’. The latter evaluation procedure is sparse in the sense that individual solutions “see” only a small proportion of all possible fitness cases during their lifetime. In the first evaluation procedure all individuals “see” all cases at every fitness evaluation; evaluation is said to be ‘full’. With finding complete solutions as ‘goal’ sparse fitness evaluation may not seem recommendable, this study, however, shows it to be effective in evolving complete solutions.

We use the term information integration to designate the accumulation of adaptations to the environment. This study shows that information integration can occur over many generations, and even lead to a complete solution, if the environment to which individuals adapt is presented only in a sparse manner. In fact, the evolved solutions comply better to the externally defined evolutionary goal if evolution occurs under sparse fitness evaluation than when it occurs under full fitness evaluation. In addition, a number of side-effects result from different forms of fitness evaluation, e.g. high phenotypic gen-

eralisability and low mutational stability in sparse evaluated individuals compare to full evaluated individuals. When the size, or complexity, of the evolutionary goal becomes larger sparse evaluation does not lead to a complete solution if evaluation is based on random samples of the ‘fitness cases’. Rather, a coevolutionary coupling between individuals and their environment, i.e. the solutions and the fitness cases, is necessary for successful information integration in these cases.

In chapter three we describe a model of the interaction between bacteria and colicins and study how the diversity of colicins evolves. Colicins are gene complexes that code for a toxin and a unique, corresponding antidote. Bacteria are killed by the toxins if they do not carry the proper antidotes. As a consequence, colicinogenic bacteria can invade sensitive bacterium populations. Here, it are the bacteria that ‘coevolve’ with the (ensembles of) colicins. In contrast to the previous model-system, here we employ a more realistic population dynamical view; population sizes are not fixed and thus populations can die out. On the other hand, we use a much simpler genetic structure to model the colicins; each colicin encodes two genes that are either active or inactive. On the level of the bacteria the colicins act as genes that can be acquired independently, i.e. a variable-sized bacterium genome.

We find that under all circumstances a high diversity of colicins evolves easily but this diversity can be expressed on one of two levels; the individual bacterium level or on the bacterium population level. In the first case individuals bacteria carry all antidote gene types that are present in the system but only a very limited number of active toxin genes. In the second case bacteria carry only complete colicin complexes, i.e. both toxin gene and corresponding antidote gene are active. Now, although the global diversity of colicin types is equal in both modes, the number of colicins per bacterium is much less than the total number present in the population. On which level the diversity of colicins is expressed depends on the growth-cost per antidote that is imposed on the bacteria, and the number of colicin types in the system. Here, we find that information integration can occur on different levels, the choice of which has a large impact on the population dynamics and further evolutionary dynamics.

In the fourth chapter a somewhat similar model is studied; the eco-evolutionary dynamics of restriction-modification (RM) systems in a bacterium-phage community. In this model we also use realistic population dynamics and simple genetics, as in chapter three. The bacteria and phages constitute a host-parasite system. The RM systems can serve as a defence mechanism for the bacteria that protects them to phage infections. Phage, however, can become insensitive to specific RM systems through acquisition of modification patterns. The system shows similarity to the colicin system; phage can be interpreted as ‘toxins’ and RM systems act as ‘antidote’, although the former can become ‘immune’ here as well. Also, contrary to the two genes of a colicin complex in the colicin system the phage and the bacteria act on independent levels in this system. Now, an environment on the bacterium level is defined in terms of the presence of phages, plus their modification patterns. But also an environment on the level of the phage is defined, based on the presence of bacteria and the RM systems that they carry. As a consequence, information integration now occurs on two levels rather than one as in the colicin-bacterium model. Again, the model is characterised by an individual-based mode and a population-based mode, but now as two attractors of a bi-stable system.

The relation between the individual-, and population-based modes and the sparse or

---

full evaluation of the environment is reversed in the two modes in this model compared to the model of the bacterium-colicin system. In the individual-based mode both bacterium and phage populations are homogeneous; both are fully equipped with RM systems, respectively modification patterns. As a result, bacteria experience full evaluation in terms of phage and their modifications and the phages experience full evaluation in terms of bacteria and the RM systems. In the population-based mode, on the other hand, the bacteria carry only a minimal number of RM systems and the phages are fully sensitive or even completely absent; in this mode evaluation is sparse on both levels.

In chapter five a coevolutionary process is studied between cellular automata and initial conditions. In this model we use again an artificial fitness function, i.e. the density classification task, to drive the evolutionary process, as in chapter two as well as structured genotypes and fixed population sizes. In a strict sense fitness evaluation in this model is necessarily sparse; the total number of initial conditions, i.e. the complete set of fitness cases, is  $2^{149}$ . In a local sense, however, evolutionary dynamics can enforce sparse fitness evaluation as well as full fitness evaluation on the individuals. We compare evolutionary dynamics that occurs in the context of spatial pattern formation to evolutionary dynamics occurring under global mixing of the populations. In this system we find optimisation, speciation, and red queen dynamics all in the same system for small changes of the parameters.

Although a complete solution is not possible in the context of this artificial fitness function, we find that under spatial pattern formation the cellular automata evolve toward generalised solutions, i.e. they optimise individual solutions. On the population level, the cellular automata and initial conditions show speciation, each to different extends. For instance, speciation of initial conditions occurs on at least two levels. Two main subpopulations of initial conditions evolve with densities below or above 0.5. Within these main ‘phylogonies’ many disparate taxons exist, with similar density values but far apart in genotype space. When we apply global mixing the system shows typical red queen dynamics; both populations show continued oscillations between two phenotypical states. As a result, the populations are mostly homogeneous and speciation, and also optimisation of the density classification task does not occur. Nevertheless, we find information integration to occur such that individuals can more easily chase the opponent population in the evolutionary red queen race.

## 6.2 Sparse fitness evaluation

Fitness evaluation in natural systems, i.e. selection events over a lifetime, often conveys only a very limited amount of information concerning all possible environmental contingencies. Thus, evolution seems to have little information available for devising competent responses. Despite the sparseness of fitness evaluation of the environment in evolution, we find beauty, complexity and diversity in nature, everywhere. Apparently, the natural evolutionary process is not hindered by the sparseness of the fitness evaluation. In fact, studies in this thesis suggest that it is *thanks to* sparse fitness evaluation that evolution can integrate information.

In chapter two, the coevolutionary optimisation model, the extent to which fitness evaluation is sparse or full is imposed on the individuals and the resulting evolutionary

dynamics of the two cases are compared. In the other models the extent to which evaluation is sparse or full is a result of the population and evolutionary dynamics. Both chapters two and five show that sparse evaluation, i.e. sparse presentation of the environment during fitness evaluation, does not stand in the way of the evolution of solutions that can cope with the *complete* environmental repertoire in a generalised manner. Also in chapter three we see that in the individual-based mode bacteria integrate and maintain all antidote genes although they hardly ever ‘see’ the various toxins. Below we will consider other possible outcomes of such an evolutionary ‘quest’, e.g. speciation or red queen dynamics, in more detail. Also, in chapter two, solutions that evolve under sparse fitness evaluation are shown to be more capable of generalising over novel environmental circumstances. Solutions that evolve under full fitness evaluation become specialised on the fitness cases presented, but show highly aberrant behaviour on novel cases. In chapters three, four, and five sparse fitness evaluation is found to occur both in a sparse manner as well as in a full manner, depending on the eco-evolutionary dynamics into which the system settles. Also, in the bacterium-colicin model sparse evaluation occurs in the individual-based mode, whereas in the RM model sparse evaluation occurs in the population-based mode; similar eco-evolutionary dynamics can result in different manners of evaluation. The density-classification model, chapter five, shows sparse and full evaluation at different levels simultaneously; in terms of density classes evaluation is fuller than in terms of evaluation of single ‘taxonomies’.

In the models that we studied in this thesis fitness evaluation can be sparse or full in a single system, depending on the parameter values or on behavioural regimes. In chapter five both modes of evaluation occur simultaneously. The extent to which fitness evaluation in natural evolutionary processes is sparse or full varies as well. The dichotomy between diffuse and pairwise coevolution shows the range of fitness evaluation in natural systems. Pairwise coevolution occurs for instance in some host-parasite interactions and some mutualisms. Such interactions can evolve such that participants always experience only one partner. In that case fitness evaluation may become relatively full. Diffuse coevolution on the other hand is exemplified by the co-occurrence of a variety of different interactions. If we compare diffuse coevolution to pairwise coevolution, we see the evolution of specialised organisms mainly in pairwise coevolution (e.g. (Rothstein, 1990)), whereas diffuse coevolution seems more often to lead to generalised organisms (e.g. (Boyes *et al.*, 1996)). In addition, sparse evaluation may trigger the evolution of complex systems. For instance, Huynen *et al.* (1999) have studied the tricarboxylic acid cycle (TCA) in organisms for which the complete genomes are published. The study shows that in different species the cycle is present in highly variable forms but is often incomplete. The observed variation shows that parts of the TCA have an independent functionality on their own, thereby suggesting how the TCA might have evolved. The complete metabolic pathway may be the result of the integration of different ecological and evolutionary circumstances rather than a solution to one, single ‘problem’.

---

### 6.3 Speciation, optimisation, and red queen dynamics

In chapter two we study evolution in the context of evolutionary optimisation. There, we consider speciation as the ‘mere’ occurrence of a variety of non-general, partial solutions and therefore we do not study speciation as such in chapter two. However, the results concerning the phenotypic generalisability suggest that speciation occurs more easily under full fitness evaluation than under sparse evaluation. If we were to make two sets of fitness cases that partially overlap and subsequently let two subpopulations evolve in the context of these two sets of cases they might speciate in terms of their respective behaviour in each others ‘habitat’. Individuals that undergo full evaluation evolve specialised behaviour in their own habitat and aberrant behaviour in unknown cases, i.e. the habitat of the second species. Sparsely evaluated individuals, on the other hand, may show phenotypically very similar behaviour because they generalise well over the cases they have not “seen” but that are part of the other’s set of fitness cases; speciation would not then occur.

Chapters three and four show how information integration can result in two dynamical modes: the individual-based mode and the population-based mode. The individual-based mode is characterised mainly by the homogeneity of the population; no speciation occurs in this case. The population-based mode, on the other hand, is characterised by the occurrence of many subpopulations, each with the same ‘genetic’ information content. The latter occurs easily, despite genetic information exchange between individuals through horizontal plasmid transfer, which plays a similar role to sex. The different species that occur in the system are, however, only a small subset of all possible species. Thus, speciation as such is not imperative, but population-based diversity is imperative. Speciation and the presence of a large number of species are ‘merely’ side-effects.

Chapter five, which describes the coevolution between cellular automata and initial conditions in the context of the density classification task, shows the full range of optimisation, speciation, and red queen behaviour. The latter only occurs under global mixing of the system. In that case, the populations are homogeneous and switch back and forth between two phenotypes. In all other cases the density classification task, as defined by the initial condition present at that moment, is solved either by the evolution of cellular automata that implement generalised solutions or by speciation (or niche differentiation) in the population of cellular automata such that different specialised ‘species’ *solve* different subsets of the initial conditions that are present in the population. In fact, the evolution of optimal solutions concurs with the presence of a large number of different species in the population of initial conditions and in the population of cellular automata. Speciation at this level seems to help the evolutionary quest for optimal solutions by giving rise to continued competition between alternative solutions.

### 6.4 Information integration

Chapter two concerns the process of information integration under sparse and full fitness evaluation. In the latter case, information integration can be interpreted as a type of tuning process; at every selection event individuals that conform most to the ‘evolutionary goal’ are selected. Any mutation that is detrimental with respect to any one of

the ‘fitness cases’ is eliminated. Under full evaluation, if we assume that the ‘complete’ environment does not change over time, evolution can be interpreted as occurring on a fitness landscape (Wright, 1967; Kauffman & Levin, 1987). A fitness landscape is a mapping from genotypes to fitness values and is generally depicted in a 2-dimensional graph, in which mutationally close genotypes lie close together on the genotype axes. In this landscape individuals evolve uphill, towards genotypes with higher fitness values (Calvin, 1986). When individuals “see” only a small number of all possible fitness cases the fitness landscape becomes dynamic. Genotypes that were previously located on hills or in valleys in the fitness landscape may now change their character and thereby change from being ‘attractive’ to being ‘repulsive’, and vice versa. As a result, the evolutionary dynamics may enter a regime of continuous change, i.e. red queen dynamics, in which previously acquired adaptations in individuals can get lost while these individuals adapt to new circumstances. Now, ‘information integration’ seems less likely to occur.

In chapter two, however, we show that sparse fitness evaluation can lead to complete information integration at the individual level and may even be beneficial rather than detrimental to the integration process. In addition, information integration has side-effects that depend on the particular form of the integration process. In chapter five, which concerns the coevolution between cellular automata and initial conditions, we actually show that complete information integration at the individual level can occur, i.e. optimal *individual* solutions are evolved in an evolutionary system in which other outcomes of a coevolutionary process occur as well, i.e. red queen dynamics and speciation. In fact, the system shows evaluation to be most sparse in the case where information integration leads to general solutions at the individual level.

Successful information integration, i.e. the accumulation of adaptations to the environment, is generally considered to occur at the individual level, as described in the previous paragraph. This thesis shows that information integration can occur at different levels in the same system, e.g. the individual level and the population level in chapters three and four. Chapter three, i.e. the bacterium-colicin system, shows that information integration can occur in an individual-based and a population-based mode. Although at the individual level bacteria have much less information in the population-based mode than in the individual-based mode, at the population level the number of colicin types in the two modes is the same. Note that whereas above the population-based mode was interpreted as speciation, here it is interpreted as information integration at the population level as an alternative to integration at the individual level.

Chapter four again shows the occurrence of an individual-based mode and a population-based mode, but now as a bi-stable attractor. Individual-based information integration occurs simultaneously in the bacterium and the phage population or in neither population, in which case both populations showed either population-based integration or the phages became extinct. Here, the systems can settle into the different modes of information integration not as a result of a change in parameter values but as a result of the population dynamics. Whereas ‘optimisation’ is often considered at the individual level only, chapters three and four show that information integration at the population level may actually be more beneficial for individuals. In chapter three bacteria in the population-based mode were less vulnerable to invasions by novel colicin types. In the RM model the phage population was much less viable when the bacteria were also in the population-based mode. Thus, individuals may do better by giving up individual ability.

---

However, information integration can occur at other levels as well. In chapter five we saw that cellular automata do not integrate a complete solution directly in the phenotype but they make it easily *accessible* at the level of the genotype via a very small number of mutations. In the same chapter initial conditions show an evolution towards ‘being difficult’, not only at the level of the density value but also at the level of gene-structures. Similarly, by evolving to different parts of a genotype landscape (Huynen & Hogeweg, 1994), or to different parts of an interaction graph (Hogeweg, 1994) individuals and populations or communities can change their local evolutionary surroundings so as to cope with environmental challenges in alternative ways.

## **6.5 Conclusion**

Information integration in evolutionary processes may at first seem to be centred around the question: “Did we end up with the optimal individual?”. This question has been posed before and has attracted much research. This thesis shows that the question posed as such is a meagre one, covering only a small part of the concept. We have shown that information integration does not occur only at the individual level, it can also occur at the level of the genotype or at the level of the population. We have also shown that information integration can occur under seemingly informationally poor conditions. In fact, the latter regime may favour generalised integration of information. Finally, we showed that different forms of information integration, resulting from sparse fitness evaluation or from full fitness evaluation, have different side-effects. Given that natural evolution tends to build and rebuild on the basis on what evolved before, different side-effects at this stage will influence the whereabouts of eco-evolutionary dynamics later.

---