MATHEMATICAL MODELS OF EVOLUTION

A THESIS
SUBMITTED TO THE DEPARTMENT OF INDUSTRIAL
ENGINEERING
AND THE INSTITUTE OF ENGINEERING AND SCIENCES
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Hakan Özaktaş
November, 1992

MATHEMATICAL MODELS OF EVOLUTION

A THESIS
SUBMITTED TO THE DEPARTMENT OF INDUSTRIAL
ENGINEERING
AND THE INSTITUTE OF ENGINEERING AND SCIENCES
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Hakan Özaktaş
November, 1992

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Béla Vizvári

Assoc. Prof. Béla Vizvári(Principal Advisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Halim Doğrusöz

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Cemal Dinçer

[illegible]

Approved for the Institute of Engineering and Sciences:

M. Baray

Prof. Mehmet Baray
Director of Institute of Engineering and Sciences

ABSTRACT

MATHEMATICAL MODELS OF EVOLUTION

Hakan Özaktaş
M.S. in Industrial Engineering
Supervisor: Assoc. Prof. Béla Vizvári
November, 1992

Two categories of evolutionary models are analyzed. The first category is the so-called *autogenesis* phenomenon. The emergence of self-organization, which has been discussed previously by Csányi and Kampis is verified. The model is extended to an interrelated multi-level autogenesis system. Similarly, self-organization is observed in a hierarchical order for each level. The second category is the optimization model of evolution. An ongoing process of consecutive LP runs associated with random perturbation of the parameters at each step, is designed to simulate the evolutionary mechanisms (mutations, variations and selection) and the population dynamics of a hypothetical ecological system. Two different LP approaches for Lotka-Volterra systems are compared and contrasted. A brief history of evolution and some mathematical models that have been constructed up to date are also described in the beginning chapter.

Key words: Evolution theory, evolution models, neo-Darwinism, neutralism, mutation, selection, autogenesis, population dynamics, modelling theory, linear programming, duality.

ÖZET

EVRİMİN MATEMATİKSEL MODELLERİ

Hakan Özaktaş
Endüstri Mühendisliği Bölümü Yüksek Lisans
Tez Yöneticisi: Doç. Béla Vizvári
Kasım, 1992

Evrim modelleri iki kategoride incelenmektedir. Birincisi *kendindenyaratılış* adı verilen sürecin modelidir. Bu tür modellerde var olan ve daha önce Csányi ve Kampis tarafından basit bir modellemesi de yapılmış bulunan, zaman içinde 'kendi kendini düzenleme' doğrulanmaktadır. Model çok düzeyli kendinden yaratılış sistemi biçiminde daha kapsamlı bir hale getirilmektedir. Benzer bir kendi kendini düzenleme süreci, her düzeyde hiyerarşik bir biçimde meydana gelmektedir. İkinci kategoride, evrimin optimizasyon modeli incelenmiştir. Burada, parametrelerin her defasında rassal bir biçimde biraz değiştiği; peşi sıra tekrar eden doğrusal programlama algoritmalarıyla, bir ekolojik sistemin dinamik yapısı ve evrimsel mekanizmaların benzetimi yapılmaktadır. Lotka-Volterra sistemleri için doğrusal programlama yaklaşımları incelemesi ve karşılaştırılması yapılmaktadır. Evrim teorisinin tarihi ve bugüne kadar yapılmış olan bazı matematiksel modeller raporun başında kısaca özetlenmektedir.

Anahtar sözcükler. Evrim teorisi, evrim modelleri, yeni-Darvinizm, nötralizm, mutasyon, doğal seçme, kendindenyaratılış, nüfus dinamiği, modelleme teorisi, doğrusal programlama, ikillik.

"...it at once struck me that under these circumstances favorable variations would tend to be preserved, and unfavorable ones to be destroyed. The result of this would be the formation of a new species."

Charles Darwin[1]

---

[1] "The Autobiography of Darwin and Selected Letters.", Dover Publications, Inc.

ACKNOWLEDGEMENT

I am mostly grateful to Béla Vizvári for suggesting an unimaginable research topic full of enthusiasm, and who has been supervising me with patience and everlasting interest and being helpful in any way during my graduate studies.

I am also indebted to Cemal Dinçer and Halim Doğrusöz for showing keen interest to the subject matter and accepting to read and review this thesis. Their remarks and recommendations have been invaluable.

I have to express my gratitude to the technical and academical staff of Bilkent University. I am especially thankful to my classmates of my first graduate year, Hakan Demirel, İhsan Durusoy, Aslı Sencer, Burçkaan Gürgün and to my former and present officemates, Ceyda Oğuz, Cemal Akyel, Esra Doğan, Mohamud Mohamed, Vedat Verter, Muhittin Demir for helping me in any way during the entire period of my M.S. studies.

It is almost unlikely that I could have got out of the computer mess-work by myself. I would like to thank to Erkan Uçar who has aided me with any of the difficulties I have had in general —particularly my UNIX and SUN troubles. I would also like to thank to Akif Eyler for helping me in implementation of the *Sun Pascal*. I am also thankful to Oktay Yırtıcı for showing interest to my M.S. work and lending me the necessary style and assembly files for LaTeX, and to Haldun Özaktaş for aiding me with the LaTeX figures and compilation of the whole text. Finally I have to express my sincere gratitude to anyone who have been of help, which I have forgotten to mention here.

Contents

Chapter 1

Introduction

Evolution has been a field of bitter controversy since the time of the invention of the concept. The discussions on the issue have previously been based on the contradictions with the existing beliefs related to creation. Later on, criticisms on the plausibility of the theory have also been made and alternative opinions to Darwinism have been placed.

Evolution has been a field of interest to many scientific disciplines, The primary concern of evolution is related to biology, but the field has been popular for physicists, applied mathematicians and statisticians as well. The contributions of these disciplines, even to the purely biological aspects of evolution have been mostly indispensable.

The popularity of evolution is even though surprising. Probably not any other field of science has taken so much recognition by the entire world of scientists. It is even a fact that this much attention to this specific concept may become harmful to the original idea itself. In general evolution is an overused term, mostly it is used indicating some progress or change—as it does also for Darwinian evolution—so one has to be careful when interpreting the written material upon the issue.

The approaches to the problems related to evolution have been diverse.

There is not a generally accepted mathematical model for evolution, the models which have been constructed so far are able to describe simple population dynamics and some biological phenomena such as mutation.

Despite the considerable developments of the mathematical models of evolutionary problems the field is still in its infancy [Feistel and Ebeling, 1989]. The reason is probably the just stated fact; there hasn't been a unified approach, an outlined subdiscipline for evolution in general, instead there have been many different attempts to solve some subproblems of the field. So despite the immense literature contained, the point that has been reached is not too far, though significant results have been obtained. Of course modeling of this very complicated process is not trivial and immediate developments are not to be expected.

A brief history and underlying concepts of the theory of evolution and natural selection is given in the next chapter. Also some mathematical models which have been dealt so far are discussed shortly. However the ones that are described are those which are most popular and it has to be reminded that a comprehensive summary of the evolution models is almost impossible. This is both due to the widespreadness of the topic and the fact that most models in literature are small, very specific cases of interest. Evolution is a topic of concern to physicists, biochemists, mathematicians, system theorists, statistical physicists and operations researchers as well as biologists. Some related topics, self-organization, chaos, thermodynamics are briefly outlined. It has to be reminded that these issues have extreme importance for the further developments in evolutionary theory. It wouldn't be surprising if a significant progress in evolution is achieved within the next decade.

The models that have been constructed can be divided into two classes. The first is the models of molecular (primitive) biological structures. All the living structures from viruses to elephants carry their genetical identities in their chromosomes, in the form of DNA and RNA chains. For an individual this genetic code is unique. It is evident that life has begun as an accidental gathering of certain molecules, which have formed the simple DNA and RNA molecules and thus triggered the process of formation of the living structures. The evolution of such molecules have been studied by Eigen mainly, in [Eigen, 1971] where the biochemical foundations of the process is given and in [Eigen and Schuster, 1977], [Eigen and Schuster 1978a], [Eigen and Schuster, 1978b] where the kinetic model is constructed.

A certain autogenesis model has been defined by Csányi and Kampis in [Csányi and Kampis, 1985] where the model foundations are described and in [Kampis and Csányi, 1987] where the implementation of the model is made. The essence of the model is the emergence of a living structure, throughout the process of forming of an informational organization of the molecules, a functional structure by means of itself, i.e. self-organization. The genetic code stored as a form of DNA or RNA molecules is a sort of functional organization. So the idea descends from the appearance of the RNA and DNA molecules by self-organization. Similar models have existed in literature and are discussed in [Csányi and Kampis, 1985] and [Kampis and Csányi, 1987]. Though each of them are different models the idea in all of them is similar. Formation of a primitive organization of molecules enables the emergence of life.

A similar program to the simulation model described in [Kampis and Csányi, 1987] has been written and implemented. The results obtained have been quite similar. After that the program was extended to the multi-level structure of autogenesis process, which was described in the formal models in [Kampis and Csányi, 1987]. The model universe has been slightly altered for this model. The discrete grid space was replaced by a set of randomly distributed components, with an initial random formation of a functional structure.

As a result of this new model; a similar development of an organized structure was viewed. Additionally the formation of more complicated structures were also analyzed, by the use of this hierarchical system. The results are demonstrated in the following chapters.

The second model of interest has been a simple model of an ecological environment, the descendant of a Lotka-Volterra model of population dynamics.

The core part of this model is the linear balance inequalities of limited resources; the feasible set of the population sizes for all of the species that are of concern. Within this system the cumulative value of Darwinian fitness (the survivor of the fittest principle) for different populations, is tried to be maximized, as a measure of the total fitness of the entire ecosystem. Thus an LP model is adopted, because of its simplicity and its reasonable representative characteristic of the dynamics of the problem. The LP optimization is considered at a fixed point in time where the parameters of the problem are assumed to be stable. Thus the idea at such a time is to determine the objective targeting position of population sizes, that would be tried to attained, in the next equilibrium point.

The next task that has been done was led from the different approaches of the dynamics modeling. Apart from the consumption balance equations, one might consider the minimal benefit inequalities for the whole ecosystem to survive. Here of course, those species are seeking the ways of exploiting natural resources to survive, and the objective should be to minimize the total loss that is given to Nature in this way, so Nature is trying to minimize the increase of entropy throughout the exploitation of resources.

So the model remains the same except the instant of optimization which finds out the best value of the ecosystem that would be a future target state. The two approaches are compared and contrasted. It is left as an open question whether two types of optimization have something common in reality and this sort of similarity can be generalized to some economical models.

Some of the terminology used may be unknown to the persons who are not involved with biology. A short glossary of the fundamental terms and several concepts (such as the Hardy-Weinberg equilibrium) that have been used mostly, has been put at the end of the report, to facilitate the understanding of the text.

Chapter 2

History

In this chapter a brief description of the developments achieved in the subject of evolution models will be presented. Also some underlying concepts of evolution will be introduced to serve as a background for the discussions that will be made in later chapters.

As an initial discussion, a short history of the emergence of evolutionary thought will be given, starting from Darwin's theory of natural selection. Following this, some mathematical models of the evolution problem will be briefly discussed.

2.1 Darwin's Principle of Selection

In this section a summary of Darwin's field research and his inference of the theory of evolution is given. It is not intended to cover all details of his work, but only to remind (or make acquaintance if the topic is not familiar at all) the fundamental facts that are important for understanding the present work. The details can be found in [Starr and Taggart, 1989].

Darwin was led to his theory of evolution after decades of studies, traveling around the world. He had noticed two things which were striking, during his trip to South America. The first was the glyptodont fossils which he had encountered in Argentina. It was an animal that had been extinct, but it was queer that the fossils indicated a surprising similarity between the glyptodonts and the armadillo, an animal existing in South America. It was interesting that two such similar animals had been living in the same part of the world, though in quite different time intervals. Another fact was the variation of traits of local populations of the same species; the famous example he had observed was the finch species of the Galápagos Islands, where the local populations were confined to several islands, separated with some distance. Each local population had differentiation in characteristics such as the beak shape, coloration, etc. He had seen nothing like this in England, which was a small island where environmental conditions did not vary much.

After years of study, Darwin published his theory in 1858, after having received a paper from Alfred Wallace, who also concluded in a similar theoretical framework. Wallace had also carried out field investigations in South America and Malay Archipelago and obtained similar results with those of Darwin.

While formulating his theory of natural selection, he based his observations on artificial selection; which was used for selective breeding of several domestic animals such as dogs, cows, or pigeons. To illustrate an example of artificial selection , if a pigeon type with black tail is desired, only those with black tails are permitted to mate. Continuing on in this manner after 3 or 4 generations, the whole population becomes black tailed.

The principle of natural selection is similar. A certain species population exhibits a diverse variety of characteristics. The competition for survival takes place both within and between the populations of species. Because of scarce resources the struggle for life enables only those species with favorable characteristics to survive and reproduce properly—and thus inherit those favorable characteristics. The remaining ones die out, in time. Thus the frequency of those favorable characteristics and traits within a population (or throughout the entire ecosystem) increases gradually. So evolution occurs. Here it can be seen that Nature is mainly the determining factor.

The source of variation that arouses new characteristics is genetic mutation. There are other sources of variation as well, such as crossover effect or chromosomal aberration, but the determinant factor is mutation. Though mutations are rare events, they can cause an incredible amount of variation, by shuffling existing alleles by other effects [Starr and Taggart, 1989].

Now it is possible to understand why so much different species have ever existed in the world. Furthermore, the theory of evolution states that life has originated as a simple form of organizational structure, and then it has evolved to more complicated organisms, and thus different forms of living things have spread out in this way.

To illustrate the formation of differentiation within a certain population we may consider the example of the finch species of Galápagos Islands. Most probably the species had originated on the mainland of South America, and were spread to the islands from there. Now each of these islands are places where environmental conditions vary considerably and even more important is that the islands are separated by a significant distance so that the local finch populations are not in interaction with each other. The reasoning that leads to the conclusion that all of these different finch species have descended from a single species which had been originated on the mainland, millions of years ago. At the time of their formation the finches were a single population. Throughout time some groups of those finches have migrated to those islands where each group became *isolated*. Evolution has occurred independently in each of the isolated groups according to the environmental conditions, thus after some considerable time has passed (up to now) different formations of characteristics and traits have developed. So today we have several different finch species in Galápagos Islands [Welch and Arnon, 1976].

The process can be generalized like this. At the beginning there is a certain population in a certain geographic region. Either a portion of this population migrates to other regions which are geographically apart or some groupings occur, so that interaction between these groups become impossible. Due to this isolation adaptations of different groups to different environmental conditions through mutations and natural selection will result with the development of thoroughly different phenotypic characteristics. At a certain time the local populations will become so apart that successful mating between them will become impossible. By this time those populations will be considered as different species [Welch and Arnon, 1976].

Darwin's theory of evolution has aroused everlasting discussions and always been challenged. Fossil evidence that have been obtained supports the theory, though the subject matter is still open to discussion, and objections to the weaknesses of the theory are still made. It is another fact that at Darwin's time the state of biology was quite subdeveloped when compared to that of now. Mendel's genetic inheritance principle, the structure of the chromosomes, the DNA and RNA chains were yet unknown. So after these developments the theory had to be revised under the light of these new achievements, but his fundamental principles had remained the same. The revised evolutionary thought is also named as neo-Darwinism.

The principles of selection have been applied to some other biological systems, which have similarities to living organisms with such attributes as metabolism, self-replication, mutation etc. The prerequisites for the so called Darwinian systems have been listed as [Eigen and Schuster, 1977];

- The system has to be self-selective. It has to stabilize certain structures at the expense of others.
- The nature of this stability is dynamic.
- To evaluate this stability, there has to be some feedback mechanism which ensures that advantage is reflected to those dynamic properties which are responsible for amplification (self reproduction).

2.2 Criticisms of Natural Selection and Additional Discussions

As stated the theory of natural selection, has always been criticized and challenged with alternative views of evolution. The controversies are not several but quite numerous and it is beyond the scope of this research to give an extensive study of these. Some of these criticisms and opposing ideas will be discussed in this section.

The main controversy with the evolutionary thought arises from the importance of chance effects in nature; is this a major factor determining the evolution of species or is it only an insignificant factor unaffecting the principles of natural selection? The conflict is also known as neutralism versus selectionism [Karlin, 1984]. To make the controversy clearer the phenomenon of random genetic drift will be shortly explained.

Random genetic drift is the change of allele frequency in a population, due to chance effects. In a population the frequency of an allele is constant but random fluctuations occur, due to the dynamic structure of the population. Sometimes, it is possible that these fluctuations occur in one direction only, so that the frequency of an allele changes after some considerable amount of time. It is clear that the effect of genetic drift is more significant when the population size is small, and it is mostly effective for isolated small populations. Sometimes the effect of chance leads to so much differentiation that new species populations may arise. This is called the Founder effect [Starr and Taggart, 1989]. It is claimed by some neutralists that the major part of evolution is occurring through random fixation, not selection [Kimura, 1990].

Some evolutionists claim that evolution takes place mostly due to chance effects than the survival of the fittest principle of natural selection. According to them much of the variations occur during the bottleneck periods, where the population sizes are reduced due to catastrophic events; and evolution is led by genetic drift. Thus there arises another controversy; whether the evolutionary changes are gradual (as stated by the Darwinists), or sudden (as stated by the neutralists, at times of bottlenecking periods).

Even in the case that the gradualist (Darwinist) view is adopted; the effect of sudden changes during the bottlenecks of time can not be overlooked; since in the history of the world there have been some important bottleneck periods following some catastrophic events, such as the formation of oxygen in the atmosphere leading to the destruction of the unaerobic living structures, or the presumed meteoric clash which led to the extinction of the dinosaurs and many other species.

It has been of great interest for many scientists to apply the findings of biological evolution to other areas as well, such as sociology and economics. However it is clear that the analogies between the principles of evolution related to these areas are very limited. Some system theorists have claimed that biological game theory is a sort of economics [Rapoport, 1985]. But it is quite doubtful that economical principles are applicable to biology and evolution.

More important than these controversies of evolutionary theory is the fact that the topic has been so widespread and diversified—mostly independent between these disciplines—that there aren't much accepted standards, approaches to the theory, and the huge amount of related publications with so much distinctions have already become intractable. When a physicist writes a paper on evolution, this may be quite unrelated to the author of another applied mathematics paper in a biology journal. So under the field of evolution, there are so many researchers studying, most being unaware of the things that are done by their contemporary scientists.

Darwin's theory of biological evolution is based on Newtonian mechanics; a population at Hardy-Weinberg equilibrium is unchanged, unless it is outdated by selection—just as an object stays at rest until a force is exerted. However this view has been criticized by biologists like von Bertalanffy, in that the living systems have thermodynamic character and the classical views of physics is not quite applicable. Atomistic view is only suitable for Newtonian physics and obviously not for biological or social systems [Kampis, 1989]. The discussion is made in some detail in the next section. Still today the relation between biological evolution and physical or thermodynamical evolution is not clear, there is no interdisciplinary consensus in the field of evolution. The issue is a topic of debate [Brooks et al. 1984].

Another counter-argument that has been put forward related to selection is the following. The neo-Darwinian view regards natural selection as an external force to the biological system, leading to genetic variations. The *coevolutionary view* that has been placed in the seventies, claims that evolutionary forces (selection) can not be defined above the living populations, but rather within close relation with them [Kampis, 1991].

2.3 Chaos, Self-Organization and Thermodynamics

The theory of chaos is dealing with the question; is disorder in nature due to chance effects only, or is there a certain order behind this 'disorder'? The origin of the theory goes back to the analysis of meteorological events, and Edward Lorenz was the first to formulate this theory.

There is a certain disorder in Nature, the shapes and positions of clouds etc., a certain snapshot in time is never repeated. The theory tries to reveal the uncertainties, insufficiencies of theoretical newtonian physics, such as entropy and turbulence. Several related matters concerning the chaotic theory are meteorological events, the shape of snow crystals and the fluctuation of the population sizes of several wild animals.

The behavior of some simple linear models of motion are nonproblematic. However with nonlinear systems (such as the motions of fluids), the dynamics are totally unknown in advance. The principle feature of chaos is that simple deterministic systems can generate random, unpredictable behavior [Crilly, 1991]. Such systems are very sensitive to initial conditions, if these initial conditions are slightly modified, then the behavior of the system is drastically affected.

The Lorenz model of a dynamic environment is a system of differential equations and the characteristic of such a model is the existence of a certain chaotic attractor[1], and an orbit forming around this attractor. An initially perturbed point, follows quite a different orbit around this attractor. This type of unpredictable behavior is called chaotic.

This dynamic model (such as the atmosphere) is very sensitive to small changes in the environment (weather). This fact is known as the 'Butterfly Effect'; i.e. the flap of a butterfly has drastic effects on the meteorological events, a single flap of a butterfly here, may cause a cyclone in the Pacific Islands after several weeks.

Until the 19th century physics was dealing with a mechanistic world, in which natural phenomena were directed according to strict, previously defined rules. This determinism led eventually to Laplace's famous claim that given enough facts, one could not merely predict the future but retrodict the past as well.

However, the rise of thermodynamics in the 19th century—the only branch of physics realizing the existence of an 'arrow of time', as stated in the preface of [Feistel and Ebeling, 1989]—challenged the timelessness implied by the mechanistic picture of the universe. So if the world was a big machine, this new field stated, that it was running down, using its useful energy and increasing its entropy [Prigogine and Stengers, 1984].

The theory of evolution presents an immediate difficulty, by contradicting the Second Law of Thermodynamics. This problem has for long, been puzzling and even still it is a topic for debate. The difficulty is the following. The Second Law states that entropy is a nondecreasing quantity through *time*, thus disorder should be increasing, whereas the theory of evolution states that

---

[1] Attractor is defined for fictitious point(s), center(s) for the chaotic system where the orbiting trajectory is located around, for a set of initial starting points.

information and complexity is increasing and thus better organized species are forming and evolving.

To make the contradiction clearer, if the imagery of the world as a running machine is reconsidered, the Darwinian paradigm introduces a contradictory thought. Despite the fact that the world machine is running down, losing its energy and organization, some subsystems of it—here biological subsystems—are running up, increasing their organization [Prigogine and Stengers, 1984].

The difficulty with the Second Law has been counteracted by the idea that, the evolving systems are not systems in thermodynamic equilibrium. This is an open system which takes in a flux of energy and may use it for its self-organization, so that by this assumption the law does no more conflict with the structure of these systems [Feistel and Ebeling, 1989], [Haken, 1988].

It is clear that the closed system assumption is not adequate to explain evolutionary phenomena. The concept of an open system is necessary to facilitate the comprehension of this behavior. In their celebrated bestseller book 'Order Out of Chaos' Prigogine and Stengers argue that emergence of order and increasing complexity is possible where the Second Law is interpreted in a different way than that of the classical view, and thus increase of organization and decrease of entropy can occur, under certain assumptions.

To restate briefly, in thermodynamics a system is approaching thermal equilibrium, with increase of disorder and elimination of complex structures. In biological evolution on the other hand, the system structure is increasing its complexity and organization, therefore entropy is decreasing. So the open system assumption is put forward, i.e. these systems have a huge amount of energy input, at the expense of their surroundings. However, some physicists claim that this type of reasoning is not plausible and thermodynamics is an unrelated issue to biological evolution [Brooks et. al. 1984].

The final word in this section will be on self-organization. Apart from man-made systems, some systems in Nature exist [Haken, 1988], which have been produced by theirselves. Or as stated in [Jumarie, 1987]:

> A self-organizing system is a system which can spontaneously modify its internal structure, that is to say the hierarchical internal relation between its elements in order to achieve its own objective which is defined by itself.

The significance of the Darwinian thought is evident for this type of systems, since they have been evolving with increasing organization and complexity.

Quite clearly, the biological systems might be seen as self-organizing. Indeed, there are examples of self-organization in nature. To give an example, the autopoiesis model of the Chilean scientists Varela and Maturana is a self-organizing system. The autogenesis model, which is described in the next chapter is also a system of self-organization.

The discussion on the related issues of evolution can be extended indefinitely, but it is not our intention in this research work to deal with these physical concepts in detail. Only an overview of the principles relevant to evolution have been of interest; and a slight introduction of the views of some physicists and philosophers has been sufficient.

2.4 Mathematical Models of Evolution

The appearance of mathematical models of evolution in literature dates back to the 20's and 30's, more than 60 years after the appearance of Darwin's book 'On the Origin of Species by Means of Natural Selection'. The pioneers of this development are, Volterra (1931), Fisher (1930) and Wright (1932). However it was not before the 50's that these works have been recognized. The next attempts are clustered around in the 70's, namely the famous works of Dobzhansky (1970), Eigen (1971), Eigen and Schuster (1977), Maynard-Smith and Price (1974), etc. [Feistel and Ebeling, 1989].

Here some of these models will be introduced briefly, to serve as a background to the discussions that will be made in the following chapters.

Before starting, we will classify the types of approach to evolution modeling in three groups. We may state these as:

- Equilibrium dynamics approach: These are models that constitute of a set of differential equations, representing the rate of increase (or decrease) of the populations concerned. Here the equilibrium points and stability are of interest. These models represent a certain correspondence with real life evolution problems; since natural selection drives the population quantities of the universe from one equilibrium to another, and the behavior of the evolutionary dynamics depends on the stability of the current solution. In dynamical systems three major types of solution (large scale behavior) is possible: fixed points, limit cycles and chaotic behavior. For nonlinear systems the behavior of the systems is unstable and mostly chaotic [Kampis, 1991]. The dynamics of the population, thus either converges to a stable or an asymptotically stable equilibrium or remains at an oscillating limit cycle, unless the system is chaotic, depending on the structure of the underlying differential equations. Most of the models that have appeared in the literature are of this type.

- Game theoretic approach: It was Maynard-Smith and Price (1974) who had first introduced a formal model of the game theoretical approach to evolution modeling. Here animal conflicts are treated using game theory. Instead of representing the population dynamics, strategy frequency equations are considered, and the concept of evolutionary stable strategies—those strategies where even small deviations are penalized—are of main concern.

- Optimization approach: This is an approach that seems interesting, but up to now some criticisms have been made. Here the question of what is to be optimized is a bit vague. Also it might be demanded that whether Nature is concerned with optimization or even if that is so, what are the relevant objectives. Regarding the problem as a hill climbing procedure for each species population, does each population have the means of determining somehow, the optimal search direction. It is claimed that

> Darwinian evolution in this sense, is myopic [Simon, 1981]. The tendency is towards a local optimum or a satisficing result. More will be said about this approach in the following chapters. The reason of these criticisms are probably due to the fact that no standard models have been accepted for optimization models of evolution. In the following chapters an LP-type optimization model of evolution will be constructed.

It has to be remarked that these approaches are neither exclusive nor exhaustive. Actually Maynard Smith's game theoretic approach is also of the equilibrium dynamics type, though this doesn't mean that a game theoretic approach should be also of this type.

Now we will describe briefly some mathematical models of the existing literature. Only very famous examples are given in this section, further propositions will also be discussed later. As it was just said the models are not confined to a single type that have been described above, so the descriptions are not presented in a classified order.

2.4.1 Fisher's Selection Equations

Fisher was one of the pioneering mathematicians who made a breakthrough by applying mathematics to the evolutionary theory. He was the first to relate genetical behavior to the theory of natural selection [Bürger, 1983]. He formulated the genetic allele frequencies, as a set of differential equations. His contemporary Wright was the first mathematician to have produced an alternate theory to natural selection, the effect of random drift [Karlin, 1984]. The Wright-Fisher stochastic model is of great interest and will be briefly discussed at the end of this section.

Fisher's differential equations represent the dynamics of allele frequencies (a genotype configuration of a chromosome) in a certain population of species, under the assumption that the state of the ecosystem is at Hardy-Weinberg equilibrium.

So the classical selection equation of Fisher is the following:

$$\dot{x}_i = x_i\left(\sum_j w_{ij}x_j - \sum_{rs} w_{rs}x_r x_s\right)$$

Here $i = 1 \ldots n$ is the index of an allele in the population, w_{ij} is the *fitness* of a certain genotype A_iA_j, and $\sum w_{rs}x_r x_s$ is the mean fitness of the population which is monotonically increasing.

A related model of interest is the Wright-Fisher model of genetic frequency. Assume that two alleles at a gene locus, A and a, are of concern with initial respective frequencies i and $N - i$ in a constant haploid population of size N. We have discrete periods $n = 0, 1, 2, \ldots$ each representing a certain generation in the population, and for each generation the frequency of alleles are determined by binomial sampling from the previous generation, with parameters N and p_i. Let X_n be the number of A type alleles (the state space is finite between 0 and N). The stochastic process is a finite state Markov chain. Another useful process is $Y_n = X_n/N$; and it represents the fraction of A type alleles in the population at nth generation [Klebaner, 1988].

The expected fraction of A type alleles is [Karlin and Taylor, 1981]

$$p_i = \frac{(1+s)[i(1-\alpha)+(N-i)\beta]}{(i+s)[i(1-\alpha)+(N-i)\beta]+[i\alpha+(N-i)(1-\beta)]}$$

where initial frequencies are i and $N - i$ for A and a types respectively; α, β are the mutation parameters, and s is the selection parameter.

In this typical model, a mutation means, an allele of type A being converted into an allele of type a or vice versa. α is the rate of transitions from A to a, and β from a to A. By selection, it is meant that allele A is selectively superior to a; and the survival abilities of A and a are in proportion $(1 + s)/1$, where s is a small positive value.

The neutral model ($s = 0$) is shown [Karlin and Taylor, 1981] to converge to a diffusion process as $N \to \infty$ and is relatively easy. For s different from 0 however, this is not the case. Nevertheless for large populations, where s is small the genetic dynamics still can be approximated by a diffusion model [Schuster and Sigmund, 1989].

The single population two allele is relatively an easy model. Extended models have also been studied; such as multi-locus models. A two population two allele haploid model example (with migration) is analyzed in [Asmussen, 1983].

2.4.2 Volterra's Ecological Equations

The general Lotka-Volterra equations, model the population dynamics of an ecosystem which is characterized by the mutual relationships between the species. In an ecosystem these interspecies, relations can be defined by a mutual consumption matrix, and such a matrix is capable of representing predation, competition, symbiosis and possibly more complicated types of relationships.

The LP model which is defined and discussed in the next chapter is also based on this type of ecological dynamics.

The differential equations represent the dynamics of a group of several species populations interacting. We have;

$$\dot{y}_i = y_i \left(b_{i0} + \sum_i b_{ij} y_j \right)$$

where y_i represents the population size of species i and b_{i0} is for the external effects (nature) and b_{ij} represents the benefit obtained by the species i from the presence of one unit of species j. These equations describe mutual relations between the different populations; prey-predator, symbiotical, parasitical or more complicated food chain structures.

Simple Volterra models including 2 populations are easy to be dealt with, all possible cases have been analyzed, for predation competition, combat etc. (see for example, [Beltrami, 1987]) and the solutions are asymptotically stable, except for the prey-predator model where unstable oscillations can occur. To give an example, for the linear competitive model, there is either stable coexistence, a single equilibrium where both populations exist with certain amounts or mutual exclusion where there are two equilibria—which the problem can converge—so that in each, only one of the populations survives [Munoz and Selgrade, 1989].

For higher dimensions, only special cases are analyzed; a 4 population model with 2 prey 2 predator system of Lotka Volterra dynamics is given in [Kirlinger,1989]. It has been shown that for $n \geq 3$ limit cycles occur and the behavior of models of higher dimensions is chaotic [Schuster and Sigmund 1983].

2.4.3 Eigen's Molecular Hypercycle Equation

A hypercycle is defined to be a catalytic (sometimes autocatalytic) reaction process, in which at some step one or more of the products is identical with some of the reactants and are thus continuously entering the reaction cycle. Some of the famous hypercycles are the carbon cycle (a reaction process which turns hydrogen into helium) and the citric acid cycle (a typical process of oxidation of fuels). Hypercycle is the essence of vital functions. The self-catalytic cycle is crucial for self-reproduction of biological structures. Without hypercycles, the replication of DNA or RNA molecules and other life processes, would have been impossible [Eigen and Schuster, 1977].

The concept of hypercycle was first introduced in [Eigen, 1971], where the biochemical analysis is made. The kinetic model of hypercycle is given in three consecutive papers [Eigen and Schuster, 1977], [Eigen and Schuster, 1978a], [Eigen and Schuster, 1978b]. In its simplest form the hypercycle equation is:

$$\dot{x}_i = x_i \left(k_i x_{i-1} - \sum_j k_j x_j x_{j-1} \right)$$
$$x_0 = x_n$$

In these equations x_i represents the concentration of a certain macromolecule and the cumulative concentration of the system is fixed to be one. Specifically for $n \leq 4$ there is an asymptotically stable optimum so that all trajectories starting from the feasible set converge to that point. If $n \geq 5$, then the solutions are limit cycles. The assumptions outlined in this model is of extreme importance for models of self-organization. This will be apparent, when the autogenesis models in the next chapter are introduced.

2.4.4 Maynard-Smith's Strategy Dynamics

The application of game theory to evolution has been criticized by some authors and has to be regarded with some carefulness. The game between nature and the living beings is a conflict but unlike game theory in the evolution game, the players are not free to choose their strategies, (but they inherit them) and the outcome of the game is determined only through an endless series of conflicts—if a stable solution exists [Vincent, 1985].

In this game theoretical approach, the differential equations represent the dynamics of a certain strategy, with frequency $(x_1, \ldots, x_n)$. A strategy called evolutionary stable (ESS) is sought, which is a stable solution of the system. Here;

$$\dot{x}_i = x_i \left(\sum a_{ij} x_j - \sum_{rs} a_{rs} x_r x_s \right)$$

and it is quite similar to Fisher's equations.

The ESS concept introduced by Maynard-Smith and Price (1974) characterize strategies that when adopted by a large majority of the population, they become the optimal strategy for each individual of the society—by maximizing the expected fitness of the strategy maker. However, it has to be stated that natural selection does not always favor the formation of ESS strategies [Liberman, 1988]. An 'ESS population strategy' is the uninvadable strategy, if it is the strategy that is applied by the population in general, except some rare deviant strategies which are disfavored by nature and selected in time [Liberman, 1988]. Therefore the equilibrium is stable; and in fact there is some resemblance between an ESS point and the Nash equilibrium [Crawford, 1990].

An evolutionary stable point (behavior, strategy) is optimal, and that's why it is of interest because that means that no alternative mutant strategy (phenotype) can invade. But it is dubious that whether ESS really exists, even in small ecological systems [Vincent and Brown, 1984].

A formal definition of an ESS can be given as [Lessard, 1990]

$$\begin{aligned} \text{either} \quad & A(s^*, s^*) \geq A(s, s^*) \\ \text{or} \quad & A(s^*, s^*) = A(s, s^*) \text{ and } A(s^*, s) \geq A(s, s) \end{aligned}$$

where

$$A(s, s^*) = \sum_{i,j-1}^{N} s_i a_{ij} s_j^* = s^T A s^*$$

and $a_{ij} = A(i,j)$ is the expected payoff to an individual adopting a strategy i, to an opponent adopting strategy j. $A = \|a_{ij}\|_{i,j=1}^{N}$ is the payoff matrix.

So ESS guarantees protection of population against invasion by mutants (deviated phenotypes) and this can be easily shown, for example in [Rapoport, 1985]. When $n = 2$ every nontrivial matrix A has at least one ESS, but for $n \geq 3$ this need not be so [Haigh, 1988].

As can be seen the set of equations, though for quite distinct problems, are somewhat similar. For example Fisher's equations given above is a special case of the Maynard-Smith contest equations. It has been stated that Eigen's model of evolution of macromolecules, is formally equivalent to the haploid mutation-selection model [Bürger, 1989]. The so called replicator dynamics equations have been defined to cover all these [Schuster and Sigmund 1983]:

$$\dot{x}_i = x_i \left(F_i(x_1 \ldots x_n) - \phi/c \right)$$

The above equations is a generalization of the stated for models of differential equations.

2.5 Evolution and Optimization

Before closing the chapter on the background issues and some of the models of evolution, a brief discussion about the relation of evolution and optimization will be given. The question whether nature is concerned with optimization when evolving itself is debatable and is put forward in the next chapter when an LP type of modeling is made. Though discussed later, the maximization of a certain potential function (which can be regarded as the overall Darwinian fitness) has been of concern to the physicists dealing with evolution [Eigen, 1971], [Schuster, 1989]. Some of the researchers devoted to the field, have defined evolution as a hill climbing procedure in the phenotypic landscape of living species [Feistel and Ebeling, 1989] and some have remarked that this hill climbing process converges to a local optimum [Simon, 1981].

It is true that evolution is understood as a sort of progress but is it really true that biological species evolve from primitive to complicated structures? Whether evolution has a direction is a question of debate as well. Some scientists claim that progress in this direction is not necessarily true. An algorithmic example of a simulation program (which is designed to play checkers) is given in [Kúrka, 1987] in which the intelligence level of the program decreases within a period of subsequent playings of the game, throughout the 'bad learning process' which is attained of course unwillingly.

Several topics of evolution have been of interest to OR scientists. Topics like evolution of body size, where the optimal energy allocation between physical growth and reproduction is sought [Ziólko and Kozlowski, 1983] have been dealt.

A graph theoretical model has been constructed for certain protein sequences of animal species, to show that these are tied by phylogenetic trees, and thus the theory of evolution holds true [Foulds, 1986]. The topic has been quite popular and similar analyses have been carried out by other researchers as well.

The objective criterion of these models is the maximum parsimony principle. The assumption of this principle is the phylogeny trees of the related species (nodes of the graph) are built so that the overall length is minimized as possible [Erdös and Székely 1991].

An interesting paper on natural selection, which is tightly related to mathematical programming is [Galar, 1989]. Most commonly used algorithms are myopic descent/ascent procedures; which search improvement at each step, aiming to a local optimal point so that gap crossings are not possible. In the related paper a search algorithm where gap crossings are possible, is analyzed.

Chapter 3

Models of Molecular Evolution

In this chapter some models related to evolution will be discussed. Since the whole picture of even a relatively simple ecosystem is quite complex, our concern will be with those simple models representing elementary organisms, at molecular levels. However similarities with broader ecological structures are evident and extensions of the results to larger problems are almost straightforward (with several simplifying assumptions, of course).

The models described differ in their mathematical sense than those summarized previously. Still, we are concerned with equilibrium points and how to reach them, but the methodology is different. It was seen in the former discussions that, the set of differential equations were not easily solvable at all.

In the autogenesis model the tool is computer simulation. This model represents a simple organization consisting of a set of components; which undergoes a process of imperfect replication. The system reaches an equilibrium of organization throughout time.

For the LP model the approach consists of consecutive runs of linear optimization, each time with slightly modified data. The model represents a small ecosystem subject to formation of new species through mutation.

3.1 Autogenesis

Genetic information is stored in DNA, RNA chains written in a certain genetic code, which is realized by the biochemical process of replication, transcription and translation; and this code determines the phenotypic properties of the organism. Mutation can be regarded as a kind of an imperfect copying procedure of the genetic information The essence of life is replication, a copying procedure of genetic information [Kůrka, 1987].

The hypercycle model constructed by Eigen (1971), represents the dynamics of molecular evolution, the autocatalytic cycles which govern the life processes of the self-organizing DNA and RNA molecules. The model that is described in this section also is an attempt to model the life processes of such molecules. The model is mainly concentrated on the replication phenomena (i.e. self-reproduction) of these.

One of the main concerns of this thesis study has been the modeling of the so-called autogenesis phenomenon. This phenomenon is discussed in detail by the inventors of the term, in the papers 'Autogenesis: The Evolution of Replicative Systems', [Csányi and Kampis, 1985] and 'A Computer Model of Autogenesis', [Kampis and Csányi, 1987]. Here we present a brief description of the model structure.

Autogenesis, in the light of above papers can be described in the following manner. The system is assumed to be consisting of different levels of subsystems, each somewhat similar in structure. The process taking place in one of these subsystems is considered as *Autogenesis*.

Autogenesis is the evolution of a replicative system where the replication is an imperfect copying process of its functional information. The system is actually a component space and this information describes the organization of this space.

The main idea of such a modeling approach, is to represent the development of an organization, a working structure by time; through the so-called

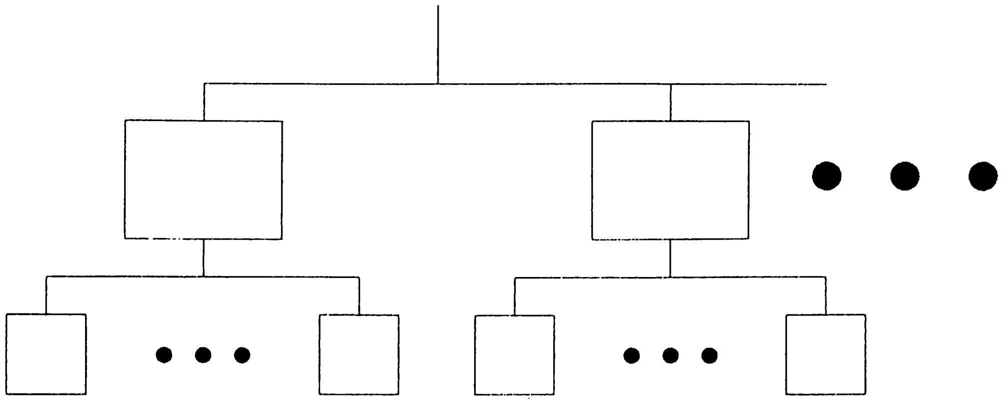


Figure 3.1: System: Autogenesis takes place in each of the subsystems.

process of imperfect replication. The elements of the model are; components: which are areas of local heapings or certain populations, functions: which represent the relations (of the transportable quantities) between the components. The current organization of the system is represented by this model universe consisting of a set of components and the interrelating functions (two dimensional in the examples discussed here). The notion of functional information can be visualized as the total figure outlining these components and functions; which are assumed to be forming this organization. Throughout the process of imperfect replication as time goes on, the whole organization (the functional structure) undergoes some changes, and at the end reachs some sort of a steady state; thus representing the development of a self-organizing system.

The analyses have been pursued in two steps. The first one was confined to a single level autogenesis problem (representing the autogenesis phenomenon in one subsystem only). The next step was to consider the system as a whole, where each subsystem is involved with distinct but interrelated autogenesis.

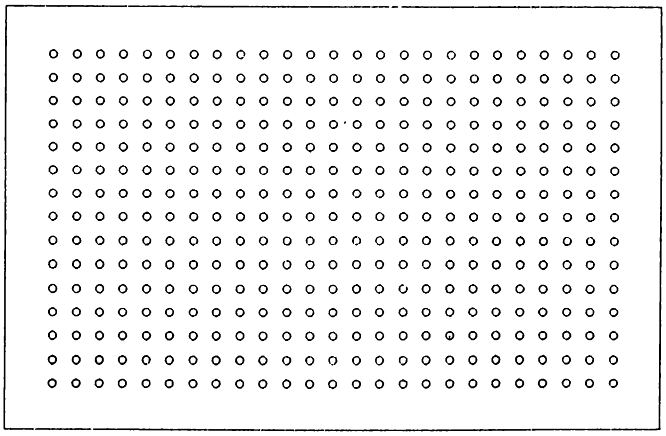


Figure 3.2: Autogenesis: Grid space

3.1.1 Single Level Autogenesis

The fundamentals of this model have been taken from the exemplary model described in [Kampis and Csányi, 1987]. The model is not unrelated to a former autopoiesis model of Varela, Maturana and Uribe (1974), which is a finite two dimensional grid space, where each grid point is reserved for a distinct component of the system. The dynamics of the system is governed by transformation of a certain amount of material between these components; so that the total mass of the system is conserved. Now it is assumed that functional relations exist between adjacent pairs (not necessarily symmetric) of components. If a function from a certain component to another exists, this means that material flow is possible. The essence of this *material flow* is that, it represents the informational flow (organization) within the molecule (for example the RNA molecule) that is of concern. At each iteration the functional picture is changed slightly, this being achieved by a probabilistic addition of new functions (arrows) to the grid space and removal of the existing functions of the components whose quantities have just become zero. This is the reason that this evolutionary process is declared as imperfect replication of the system at each iteration. After this revision a certain amount of material displacement is made. For that, the required quantity is removed randomly from the system, and then it is replaced according to the current functional picture of the universe. A portion of the removed quantity is replaced nonsystematically to

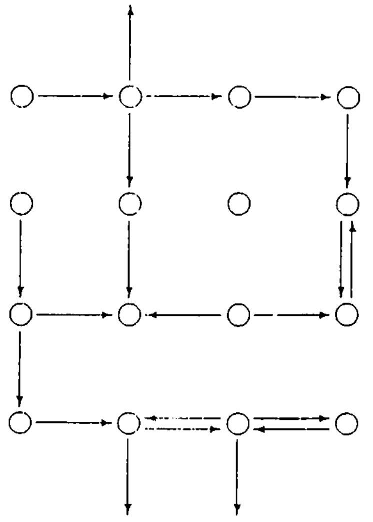


Figure 3.3: Autogenesis: Functions

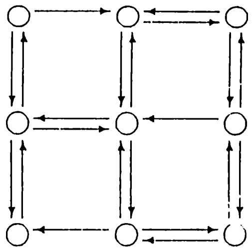


Figure 3.4: AGSP

be considered as random effects. Evolution comes to an equilibrium with error free replication, that is without mutations [Schuster, 1989].

The beginning quantities of the components are either distributed randomly or the initial state of the system is donated by an inhomogeneous structure representing a closed organization which is called the auto-genetic system precursor (AGSP). Simulation runs have shown that the system converges to a state where the material is deposited in high amounts to a certain region of the grid space, thus having almost closed loops (organizations) of arrows and in this way forming *sinks*. The formation of these regions (compartmentalization) is facilitated by the existence of an AGSP in the initial state, however even for a completely random initial state, compartmentalization takes place, maybe in more time. If for an unoccasional case two compartments occur, one of them is contracted after a sufficient amount of time elapses, resulting with a single compartment. Our simulation program written in *Turbo Pascal* has also produced similar results.

3.1.2 Description of the Process and Illustration of Results

To illustrate the process of autogenesis, one needs to consider the probabilistic formation of functions. Initially, a uniform space of components (which is yet unrelated, i.e. without functions) is donated with a beginning set of functions; so that each component can have communication only with its neighbouring components (in north, south, east and west directions) and following this, at each iteration, some set of material flow is made with respect to this functional figure, but formation of new functions is also allowed at each iteration, though with a small probability throughout the process. Thus the functional figure is perturbed at each time unit, and in addition, the amount of masses of each components change. The total amount of mass is conserved in the grid space.

The redistribution process functions at each iteration in the following way. A certain amount of material is removed from the entire space randomly, and redistributed regarding the functional picture. To represent the random noise effect, some amount of random displacement is made as well.

The formation and deletion of functions (arrows) are carried out in the following manner. At each iteration with a certain amount of probability (for example $p = 0.1$) for each component an arrow is placed to its neighbouring 4 components; if it has some certain amount of material. If the component has zero mass, then all the arcs emanating from it (in the previous figure) are deleted.

Thus the process continues with the formation of a self organized figure. The functional figure (formation of arcs) determines the new distribution of the total mass of the space, in which a concentrated, organized region occurs. As the process continues on, the formation of these arcs, lead to (as expected) some formation of denser regions (by the material flow through these arcs) and as the process outlined is a *greedy* one, the most favorable region (with more arcs leading flows inside) tends to collect the majority of the mass. Since the exemplary models used in this research, are relatively small, it is not surprising that only a single region of concentration is formed, at the end. The formation of the figure resembles a somewhat closed loop area, suspended from the outside by flowing arcs, until the steady state is reached; and this functional region becomes a sort of closed (though not fully, due to random effects) organization. The process can be visualized by a typical simulation run which is illustrated in Figures 5. As stated, the existence of an AGSP may speed up the formation of a concentrated (organized) region; the existence of 2 or more AGSP's may lead to tentative formation of several concentrated regions, but eventually the process favors, the survival of one of these regions. Detailed discussions and examples are illustrated in [Kampis and Csányi, 1987].

3.2 Multilevel Autogenesis Model

After having implemented a model for a single level autogenesis problem, the next attempt was to extend the model, so that it would cover several levels of subsystem organizations (the entire system). However before doing that the autogenesis model of the subsystems is slightly modified.

Here we are concerned with still a 2-dimensional, but continuous universe instead of the discrete grid space. Here the components are spread out in a more sparse way. They are distributed at random initially, and are assumed to be fixed after then. Now for a subsystem, there corresponds a single component in an upper level of a larger system. Thus the model universe is something like that shown in Figure 3.6. and it is symmetric in the sense that, if the upper level has 20 components, then we have 20 subsystems beneath that level.

the total mass of the space, in which a concentrated, organized region occurs. As the process continues on, the formation of these arcs, lead to (as expected) some formation of denser regions (by the material flow through these arcs) and as the process outlined is a *greedy* one, the most favorable region (with more arcs leading flows inside) tends to collect the majority of the mass. Since the exemplary models used in this research, are relatively small, it is not surprising that only a single region of concentration is formed, at the end. The formation of the figure resembles a somewhat closed loop area, suspended from the outside by flowing arcs, until the steady state is reached; and this functional region becomes a sort of closed (though not fully, due to random effects) organization. The process can be visualized by a typical simulation run which is illustrated in Figures 5. As stated, the existence of an AGSP may speed up the formation of a concentrated (organized) region; the existence of 2 or more AGSP's may lead to tentative formation of several concentrated regions, but eventually the process favors, the survival of one of these regions. Detailed discussions and examples are illustrated in [Kampis and Csányi, 1987].

3.2 Multilevel Autogenesis Model

After having implemented a model for a single level autogenesis problem, the next attempt was to extend the model, so that it would cover several levels of subsystem organizations (the entire system). However before doing that the autogenesis model of the subsystems is slightly modified.

Here we are concerned with still a 2-dimensional, but continuous universe instead of the discrete grid space. Here the components are spread out in a more sparse way. They are distributed at random initially, and are assumed to be fixed after then. Now for a subsystem, there corresponds a single component in an upper level of a larger system. Thus the model universe is something like that shown in Figure 3.6. and it is symmetric in the sense that, if the upper level has 20 components, then we have 20 subsystems beneath that level.

```
00000.00..0....0...00....       ...00.00..0.....0.0....00
..000000..00.00..0..0....       .0000000+000000...00...0.
00.0000.....0.00...0.000.       00000.00000000000.00..0.0
0000..000000..000...00.0.       .00..0..000...00...000000
.000.0...0.0..00.0...0.00       .00.00.00000.000.000.0.00
..0.00.0.0...00.000..0000       0000.000000..00.0000.0000
00.......0....00..0000.00       .000.000.00.00..00.000.0.
0000.00.00....0...0..000.       0....0..00..00.0..0.0.0..
...00.000..0..0000..00000       .0.........0.........0...
..00....00.0.0000.00..0..       .0.........0.............
......0.000.....0..0.....       ......0.0................
......000.......0..00..0.       ..0...................0.
.....0...0....00.0..0....       ...0.............0.......
.......0..............0..       .........0......0........
...0......0..0........0..  2    ......0......0...........  7

...00.0.00000....00..0...
..00..0000XH000..00.0000.
...0..000+XH00.0..000.0..
....0..000+000.00000....0
....0...0.0.0...0........
.....................0....        (.)    deserted components
.....0.................0.         (0)    between 1 and 5
.....................0....        (+)    between 6 and 10
...0...................0..        (H)    between 11 and 15
....0...0.......0........         (X)    more than 15
0.0......................
0...0....0.......0....00.
.0......0..0.............
.0.......................
0........................  24
```

Figure 3.5: Autogenesis: The states of the grid space after 2nd, 7th and 24th iterations. A closed loop organization occurs in the upper middle part of the grid space. The model had started to iterate in a uniform grid space. The redistribution parameters were set to $n = 30.0$ (for random exchange) and $m = 150.0$ (for functional redistribution). The initial probability of placing a functional flow between two components was set to $p1 = 0.5$ for the first iteration and the initial values of each point had been set to 1.0. After the first iteration, new functional flows were placed between nonzero components, with a probability of $p2 = 0.1$. The amount of functional redistribution is equivalent to the fourth of the amount of material in the component from where the functional flow emanates. The functions emanating from the newly deserted points are removed at the end of the iteration. The detailed description of the algorithm is given in [Kampis and Csányi, 1987].

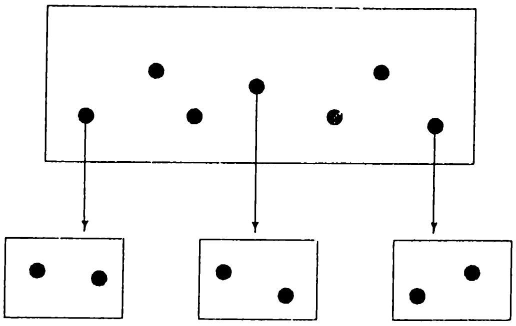


Figure 3.6: Multilevel Autogenesis: System Structure

The principles of an autogenesis evolution are same with the previous model. The imperfect replication of the functional structure, removal and redistribution of the components are similar.

Here the procedure is as follows: The process starts at the lower levels, at first. The convergence criterion is the formation of some peaks on the lower level. When a certain percentage of the lower level subsystems have enough peak points, then a similar autogenesis process is started in the upper level system. It is assumed that, only those components in the upper level, whose corresponding fields satisfy the criterion stated above are active.

As expected, the results indicate (for 2 levels) that certain peaks occur, the landscape in the lower fields present steep and high peaks whereas moderate and smooth peaks develop in the upper level which is not independent of the lower fields.

3.2.1 Description of the Process of the Multi-Level Model

Now the idea of a single level operation, in this new model is exactly the same; the formation of functions between the components and material flow, with respect to this formation. However instead of a grid space, there are randomly distributed components in a 2-D continuous space, and functional relationship between any two components is possible.

The effect of running the procedure for a single level model is the formation of several peak points, depending on the overall field mass size and the values of the redistribution parameters. To state in other words the observed phenomenon is the concentration of the majority of the mass in several components, if there is significant amount of redistribution and if the field mass is not comparatively, very small. Before studying the reasons of this behavior, one should look at the effects of the functional figure.

The model space, in this case, is less denser when compared to the grid space environment and the functional relationship is somewhat inversely proportional to the physical distance between the two components (in our model). The effect of this is the possible formation of dense arcs (in which flow intensity is high) between nearby components.

The possible local events that may be encountered are:

- Sink formation: Accumulation occurs at a certain point, a single component. Heaping of a huge amount of mass occurs at the point as steady state is reached.
- Formation of transient points: The components where the entering flow to a certain set of components is balanced by the emanating arcs, in the long run. The mass of the components remains roughly the same, it is neither too large nor too small.

  There is also the possibility of a set of components (being in a transient state) for which the total inputing flow is not balanced by the outgoing flow. For this case, either desertation or sink formation occurs, when steady state is reached.
- Loop formation (Closed organization): A closed loop is formed, also supplied by ingoing flows for a sufficient amount of time, similar to the case in 2-D grid model.

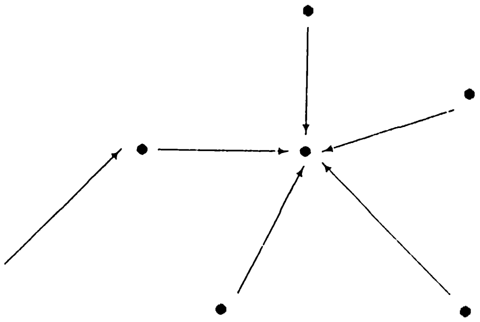


Figure 3.7: Sink formation

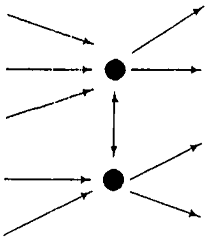


Figure 3.8: Transient points

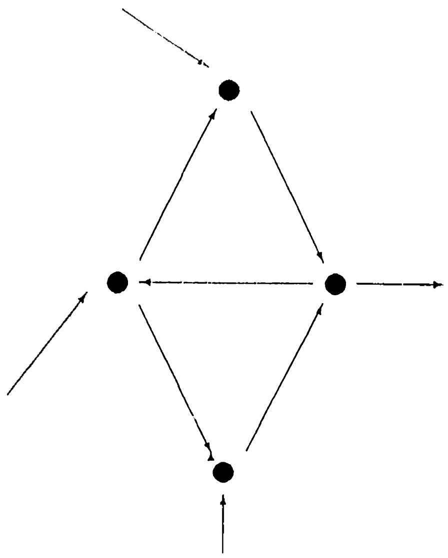


Figure 3.9: Loop formation

For the sake of formation of self organized structure, interest is focused on the formation of closed loops, in this model. It will be discussed in the following section and the formation of loops, hence self-organized structures will be illustrated with examples.

Now in the multi-level model, we have considered a 2-level problem, with one main level, which is tied to some lower level systems. The description of the process can be stated in the following way. Independent single level autogenesis runs are started in the lower levels. At each iteration, one iteration of the single level model is implemented, in each of the lower levels. Additionally, a slight amount of material flow is allowed between these independent spaces. The net effect of this process is a rescaling process of the entire space. (If the space gains or loses some amount of material, each component increases or decreases in the same proportion.)

To enable a certain level space of component space to start it's self-organization the sublevels of the space should have some certain amount of organization. As will be stated later, the development of an organization is enabled by some amount of peak formation at several points, in a certain level. After some sufficient amount of organization occurs in those lower level structures (at least half of the sublevels); a similar run is started in the main level. It has to be recalled that each component in this level represents a certain predefined lower level, at any time the sublevels that does not satisfy the above condition indicate the corresponding component in the upper level is *inactive*. If a certain percentage of the sublevels are inactive, then the iteration of the main level is not performed at that time. Independently of the upper level operations, some amount of mass exchange between the sublevels are also possible (so that the total mass of a certain level is conserved), and the mass changes are reflected to the corresponding upper field components as well.

Procedure parameters are as defined in the Pascal program are (the letter symbols represent the predefined program variables):

- n: the problem size, the number of subfields, the number of components in each subfield and the main field as well.
- *cntup*: the total number of iterations.
- *ken*: the amount of functional redistribution.
- *kem*: the amount of random redistribution.
- *mass*: mean value of randomly assigned initial mass, for each component in a subfield.
- *p1*: the probability of a flow occurring between any pair of components in subfields or in the main field.
- $c \times sc$: the amount of possible flow between two pairs of components within one unit of distance (or less).
- t: condition parameter for the mass of a certain component.
- *con*: condition parameter for the sum of certain components.
- b: the least fraction of active subfields necessary, to operate the main field.

The exact algorithm of the program may be given in the following way.

Step 0. INITIALIZE; (the program parameters have to be set in advance) each subfield is assigned to n components and each of these are randomly distributed to a 2-D field space.[1] Initial mass is uniformly distributed with a mean value parameter *mass*. Main field components are also placed randomly; their mass is set equal to their corresponding subfield mass. The functional structure is initialized for all fields independently. The component pairs are picked up randomly and a flow is placed with a probability of pf, with a value equal to $e \times sc/dist^2$ where *dist* represents the distance between the components.

After initialization the following steps are iterated for *cntup* iterations;

Step 1. AUTOGENESIS; for each subfield the following single iteration of autogenesis takes place. A certain amount of material (ken+kem) is removed randomly from the field. *ken* is redistributed functionally by selecting pairs of components ($[i, j]$) randomly and if there is a flow between them (from i to j), then a flow value of the product of the mass of the first component and the flow intensity ($f[i, j] * x[i].qty$) is added to the mass of the second component. *kem* is redistributed randomly, disregarding the functional structure.

Step 2. FIELDEXCHANGE; an amount of *ken* is collected and randomly redistributed, within the subfields.

Step 3. CHECKCONDITION; for each subfield the following condition is checked. For all components which have the sum of their present and previous mass greater than $2 \times t$, if the present sum of such fields is greater than *con*; then the field is denoted as active.

Step 4. MAINFIELD; If $100b$ percent of the fields are active, then the autogenesis iteration is carried out for the main field. The redistribution parameters are adjusted as, *nken* and *nkem*.

---

[1] It has to be recalled that the distribution is assumed to be uniform in the continuous X-Y space. Due to roundoffs, two or more components may be assigned the same coordinates. This means that these components are very near, but they have to be distinguished.

The process is similar to the single level operation, however the exchanged masses in this upper level are reflected to the lower level immediately, (where the independent runs are still continued), and vice versa, but the effect of the previous operation is more dominant; since the mass exchanges in the upper level have drastic effect on the related subfields. The process in the upper field stops when the stated condition is violated.

3.2.2 The Interpretation of the Results

As stated our multi-level structure carries some sort of symmetry; so that in each field (no matter it is a subfield or a higher level field) a certain predefined number of components exist. We denote the number of components in a field as the size of the problem. Each component in a subfield is characterized by a subfield which is of secondary level to the related field (and of course the mass of the component is identical with the total mass of its corresponding subfield); so the changes in the values of either of them is directly reflected to the other; throughout the process.

Our analysis results have been confined to two level structured model examples, with sizes of 20 and 100; mainly. The program code was written in *Turbo Pascal*, implemented on PC (for size 20) and also simulation runs were made with *Sun Pascal* (for sizes up to 150) as well.

There is one fact common to all implemented models, with varying sizes and levels; and that is the formation of *peak components*, in which the total mass increases abnormally. The formation of these peak points is caused by the initial configuration of the functions and random effects.

Although peaks are visualized depending relatively on their surrounding components, some generalizations are possible. Upon the first observation of the component quantity data, the peak points are apparent. However an operational definition of a *peak* may also be given, before tabulating the results. Regardless of the fact that it is a subfield or a main field, after a thorough analysis of the field landscapes, the following criteria can be given for whether a component can be considered as a peak or not. It has to be stated that this sort of classification is valid only for fields which have an overall mass above a critical value (for $n = 20$ above 40 or 50, and for $n = 100$ above 100), and the criteria should not be interpreted in a strict sense. For field masses below 1000 units, the criterion should be above 2.5 or 3 times the *field mass/size* ratio. For fields around a mass of 2000-3000 the components above the stated ratio are sufficient. For fields 10,000 or larger, again components with masses greater than 2 times of this ratio can be depicted as peaks.

The stated degrees in the above paragraph is a theoretical way of generalizing the peak definition. For the practical purpose of identifying the peaks, the following critical values were used. For subfields of size 20 the value was set to 10, for size 100 this was about 30. For main fields of size 20, components with values above 100 were considered, and of size 100 components above 1000 were spotted as peaks.

The significance of the peaks is related to the fact that, peak formation leads the development of an organization, if the functional picture is favorable and if the field size is not too low (10 or less). The accumulation of mass at a certain region enables loop formation and thus a closed organization, and this process may be triggered by the initial appearance of relatively higher components, namely the peaks. As the process continues the peaks which have a large influx of material can grow up and redistribute some portion of the mass they own to their surrounding components which they supply by intense or semi-intense flows.

When the problem size is relatively small (that means less crowded fields of components) the formation of sharper peaks is favored in the upper field and also in the subfields where a significant amount of mass is present. In larger problems (such as size 100); there are still formations of peaks, yet not as sharp as those observed in problem size 20; and there are more peaks (many instead of one very high peak). The structure is similar in the subfields (many peaks) where significant amount of mass exist; however the peaks are sharper than

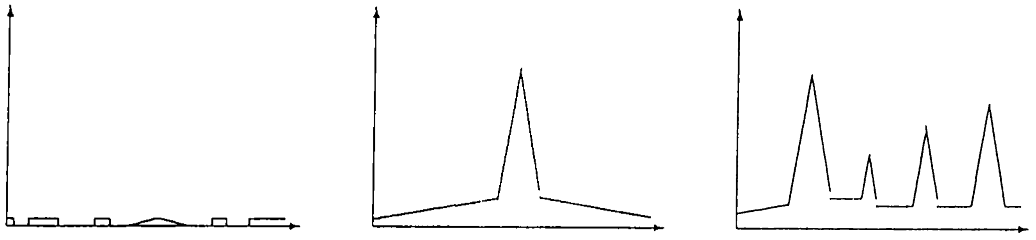
Prob. size	Total mass	Stab. iter.	Number of peaks	Peak masses (individual)	Peak percentage
20	1992	5	1	1496	75%
20	1998	5	3	1039, 322, 130	75%
20	1999	6	6	498, 381, 258, 175, 159, 103	78%
100	50067	33	10	10367, 8106, 3892, 3297, 2867, 2768, 2613, 1900 1462, 1052	77%
100	49964	34	10	17390, 4538, 4088, 3062, 3055, 3031, 2452, 1187, 1082, 1077	76%
150	112412	48	11	15702, 11285, 13418, 4348, 4975, 11290, 2911, 8815, 7136, 5800, 2701	78%

Table 3.1: Main field analysis. For problem size of 20, the components above 100 are considered to be peak points, and for size 100 the components above 1000 are selected. These points amount to about 75 % of the overall mass of the subfield.

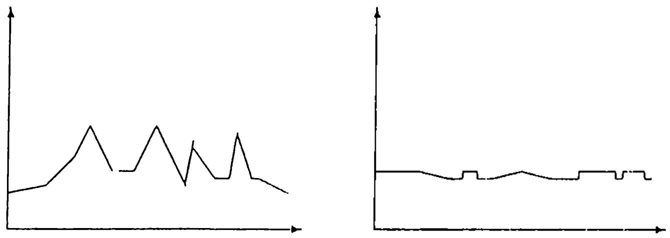
those of the main level in the subfields. Typical simulation results, and the formation of the peak points for the main field and the subfields are summarized in Tables 3.1 and 3.2 respectively. Independently, for 12 runs of size 20 the frequency of the number of peaks that are encountered in the subfield analysis is listed in Table 3.3. Typical patterns of subfield landscapes are illustrated in figure 3.10. So there are two important factors in the formation of a given *field landscape*: the first is the functional picture which underlies the entire flow process, and the second is the total mass located in that field, if the field doesn't contain sufficient amount of material, then random effects dominate the process since the random redistribution has an absolute effect regardless to the subfield mass. (Actually, even for fields with a small mass, despite the effect of random distribution the functional picture also favors some of the components to be relatively higher than the others and which are more or less stable, but these are not interpreted as peaks since their sizes are not very high. In general for such fields the presence of these *relatively higher* components do not enable the emergence of self-organization.) If the amount contained in the field is very high then formation of sharp peaks and deserted points requires

Prob. size	Total mass	Subfield mass	Subfield peaks	Type of distribution
20	1986	1459	168, 145, 177, 125 etc.	rough
20	1998	70	60	one peak
		71	29, 20	several peaks
		318	128, 24, 19, 47, 13, 14	rough peaked
		1083	442, 139	rough peaked
		145	116, 11	several peaks
100	50067	807	24, 133, 29, 187, 90, 62 34, 87	rough peaked
		190	150	one peak
		14	—	complete desertation
		10367	524, 350, 318, 291 etc.	rough
		277	128, 91	several peaks
		865	23, 106, 31, 38, 111, 86 247, 69	rough peaked
		141	39, 16, 35	several peaks
		3297	146, 320, 110 etc.	rough
		354	51, 97, 61, 53, 32	several peaks
150	112412	4975	—	rough
		425	68, 87, 77, 42, 25, 45	several peaks
		15702	—	rough
		111	52	one peak
		68	31	one peak
		2911	227, 196, 612, 142, 128	rough peaked
		1007	68, 132, 95, 92, 98, 80	several peaks
		32	—	complete desertation

Table 3.2: Subfield analysis. For size 20, components above 10, and for size 100, components above 100 are taken as peaks. The table provides sampled information of several subfields, not the entire system. See figure 10 to get an idea about the possible subfield landscapes.



A: Total desertation B: One peak C: Several peaks



D: Rough peaked E: Rough

Figure 3.10: Subfield landscapes. These are the typical landscapes encountered in the subfields. In general, as the total mass of the subfield increases, the pattern of the landscape changes from total desertation to one peak, then to several peaks, then to rough peaked and finally to the rough landscape.

# of peaks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
obser. quant.	69	12	3	5	1	1	5	1	1	-	1	2	-	1	-	2

Table 3.3: Subfield peak behavior. For 12 runs of problem size 20, the frequency of number of peaks observed in the subfields are tabulated. It is seen that in more than half of the subfields single peaks are observed. The observation of several peaks is more rare (in general 2 to 9 peaks). The rough peaked structures (about 10 to 15) and the rough landscapes (a dense landscape with no apparent peaky structure at all) are even much more rare. The rest of the subfields (not listed) which are nearly the one-fourth of the examined ones, experienced an almost desertation. See also the typical subfield landscapes, in figure 10.

much more amount of time than that of the simulation time, so the landscape remains more uniform. Of course the problem size has direct effect on both the functional picture and the sparsity of the field; and the total amount of mass per field. To illustrate a typical functional structure an example model space for the main field after stabilization has occurred is given in figure 3.11 and table 3.4. Thus the highest level of the exemplary models are somewhat independent from it's subfields; because the total mass does not change and the simulation time is shorter, since stabilization occurs after some time. As a matter of fact, as one gets into higher levels the continuity of the process is dependent on the subfields (for the required condition to be fulfilled) so in general, fewer number of iterations occur in the higher levels. So in a sense the higher levels are more conservative and slowly progressing, than their subfields due to the stated fact that the process is stopped in those fields when the percentage of active subfields is below the critical value.

The above fact is observed in the simulation results. The subfields are subject to drastic changes of landscape, by the imposed mass changes occurring on the higher level fields. The effect of these changes is a rescaling effect on the subfield; however if the field mass is reduced or increased considerably, then as discussed above, the field structure changes.

The development of the organizational progress can be summarized in the

Indice	Col	Row	Iter	Qnew	Qold
1	32	10	5	30.53	30.53
2	78	19	5	16.83	16.83
3	46	16	5	29.56	27.56
4	41	6	5	35.48	35.48
5	62	6	5	15.03	14.03
6	51	5	5	188.49	187.49
7	25	19	5	17.16	17.16
8	35	16	5	57.16	57.16
9	36	17	5	588.77	589.77
10	53	9	5	217.72	218.72
11	46	17	5	291.56	293.56
12	45	14	5	240.77	240.77
13	47	18	5	29.87	29.87
14	7	8	5	34.19	35.19
15	42	9	5	40.94	40.94
16	52	7	5	66.65	64.65
17	14	20	5	27.68	27.68
18	3	1	5	19.64	20.64
19	5	5	5	27.43	27.43
20	15	1	5	16.43	16.43

Table 3.4: The state of the main field, after stabilization has occurred. The values of the material that is existing at each component is tabulated versus its X-Y coordinates. The parameter values for this typical run were set as the following (also see the algorithm of the program): $p1 = 0.2$, $e = 0.25$, $sc= 40.0$, $n = 20$, $ken= 16.0$, $kem= 8.0$, $b = 0.5$, $mas= 4.0$, $vari= 2.0$, $con= 24.0$, $cntup= 50$. The functional picture can be seen in Figure 3.11.

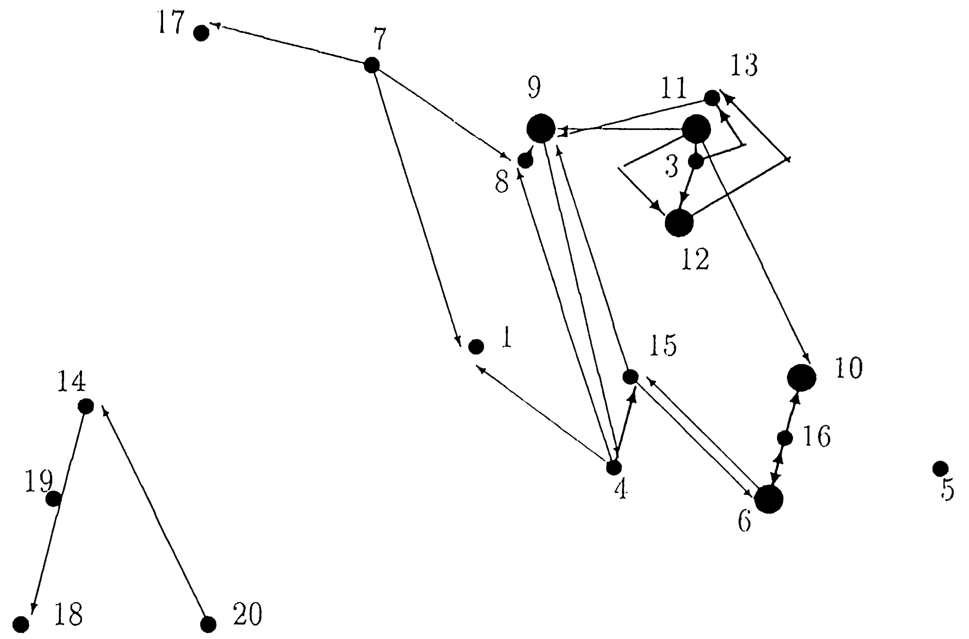


Figure 3.11: Self-organization of the main field. The peaks components (large circles) are placed around the functions with intense flows (dark arrows). A major part of the material of the system is settled around these self-organized regions. The isolated points which are connected by weak flows have quite low quantities of material.(Component *number 2* is not illustrated since it does not fit into the page.)

following way for the two-level system. At the initial iterations some formation of slight peaks become apparent which satisfies the condition stated, and as enough fields have become *active*, the iterations for the upper field is started. As discussed there are also field exchanges in small quantities between the subfields, however they have only negligible effect on the upper field operations. As the autogenesis process takes place in the upper field, peak formation and desertation occur at some components, which immediately effect the subfields. The desertation of an upper field component destructs the functional organization of the corresponding subfield, and the peak formation in a component results with a new structure of the concerned subfield. As a result the appropriate condition defined for the subfield level is violated after some time, thus the process for the upper level is stopped, and there are only several components which are active. Thus the implementation of the upper field operations, after *enough organization* has taken place in the sublevel fields have emerged, destructs this organization through the slowly progressing organization of the upper level, which stops completely after the original condition is violated. This stabilization occurs around 7 iterations for problem size of 20, and around 34 for size 100. After then, in the upper level slight changes are occurring, those resulting from the material exchanges between the lower level fields. The generalization of this process is discussed in the following section.

3.2.3 Formation of Higher Level Structures

Though the analysis was confined to two level problems by using the several results of the process; the possibility of emergence of higher level structures were also investigated.

The simulation data indicates that the formation of new structures is quite possible, as sufficient number of peaks are occurring at the main level almost unexceptionally. Therefore, the condition would be satisfied by the most, at least half (which is the critical value) of the tested models even at the initial iterations of the process.

The discussion made at the end of the previous section can be extended to the multi-level autogenesis models (3 or more levels). It was told that the structure formation at the second level destructed the structure already formed at the first level, and the new structure at the second level was sufficient enough to satisfy the similar condition (with adjusted parameters) of starting the iterations of a further higher level. The implementation of this fact would, similarly, result with a structural formation of a higher level organization but on the other hand the structure of the second level fields would be altered and soon the process in this third level would be stopped, with no further evolution. Thus the whole autogenesis process continues in this manner for the formation of higher level fields, until a point is reached so that no further descent to a higher field organization is possible.

3.3 Possible Characterizations of Autogenesis Models

The importance of an autogenesis model is that, it represents the development of an organizational structure; beginning from simple primitive structures, and the emergence of more complicated structures progressively. The functional picture (in a single field) itself is a set of stored information representing a self-organizing (living) structure. This *functional information* is slightly altered through an imperfect copying process. The similarities with biological development are evident. The genetic information in the chromosomes of a living cell is stored in double helix chains (1 dimensional set of information) subject to changes through mutations, variations, genetic drift, etc. The gradual growth of complexity can be interpreted as the building up of any biological species; the smallest unit being the cell, then a group of cells forming up the tissues, and the tissues forming up the organs, finally reaching as a chain formation to the entire living body. Since the genetic structure is unique; at each level the structural information is somewhat similar. Of course the autogenesis model discussed in this research are quite simple.

Though the primary model of concern is as described above, the autogenesis models can be used to simulate some other phenomena as well. There are alternative ways of interpreting these. Two alternative approachs of interpreting the model will be stated here. The first approach is to view a demographic map, where the landscape represents the distribution of a population (a certain species or human beings) in a geographical environment and the functional picture indicates the underlying flow (migration for example) relationships between them. The second is to characterize a phenotype space; in which each point represents a slightly modified entity, a biological phenotype characteristic visualized on a 2-D space (continuous or discrete). It has to be stated that the 2-D assumption in this case is only arbitrarily made; there is no necessity of the space being two dimensional. This type of illustrations (again in 2-D) are also used in two of the previously discussed references, [Feistel and Ebeling, 1989] and [Allen and Mc Glade, 1988]. In the first approach however, the 2-D space makes a better and understandable ecosystem, where living structures have been somehow distributed.

3.4 Foundations of the LP Model

An LP model for the problem of evolution has been developed, and in this section it will be described briefly. Also some aspects of the model will be discussed.

The dynamic structure of a certain ecosystem is governed by the Lotka-Volterra differential equations. As may be recalled, these equations represent the mutual interactions of all populations in which mutations and appearances of new species are ignored.

The ecological system is similar to that of the Volterra's equations, the coefficient matrix representing the inter-species relations is similar. However, instead of locating the equilibrium points, we are concerned with solving consecutive LP problems and the equilibrium being reached in this way. Thus the process continues by jumping from one equilibrium to another with respect to changes in the environmental conditions or internal structure of the ecosystem.

One of the basic assumptions of an LP model is that, it assumes static environment, unchanging conditions. However the evolution problem is dynamic in nature, and this is the main difficulty when considering LP modeling approach. Even if we can formulate certain objectives and constraints , the parameters are continuously changing in real life. So one of our assumptions is to cut certain slices of time, where optimization takes place. To prolong the model in time, we introduce consecutive runs of LP, so that it becomes an ongoing process.

As it was noted previously, the logic under the optimization type approach to evolution is debatable. It is questionable; whether nature is concerned with optimization or not. Even if it is so, still it is quite dubious that the whole ecosystem is running towards a global (or local) optimal value, gradually.

The idea of optimization however has its justifiable grounds when the minimum energy principle of thermodynamics is considered. A further question can be put on the meaning of optimization in a certain ecological system, whether optimization in a system of mutual conflicts makes sense or not. The adopted view is the survivor of the fittest or luckiest principle, so the optimization of the Darwinian system of the overall system will be of concern.

Obviously each species tries to survive and reproduce, thus it is natural to think that they are all competing, trying to follow the optimal paths, for their populations. Though the individual behavior of a species is quite different than that of the whole population. The tendency of the populations whether towards egoism or altruism is a debatable issue.

The game theoretic approach of Maynard-Smith and Price is based on the idea of finding local optimal (stable) values of the available strategies. We will not be considering individual objectives, but an objective covering all the species populations existing, to formulate the problem in the LP fashion.

Let us consider a fixed point in time, and assume that all parameters are available. Now all species are interacting, in the Lotka-Volterra sense, having prey-predator, parasitical, symbiotical or some other form of behavior.

Consider two species and let us index them as i and j. These two may have prey-predator type of relationship or symbiosis. Or they may not be affecting each other, at all. Let us assign a numerical value of the effect of the existence of one unit of j in the ecosystem to species i. If the existence of j is favorable to i, then we have a positive impact, else a negative one.

This idea leads to two ways of viewing the mutual interactions between the species. The first is to consider the consumption rate of a certain species by another species, and vice versa. The second way is to consider the benefit gained by a species population, through the existence of another species population. The model that will be considered in this section will be the first approach, but modeling in the other sense will also be discussed later. Though they are two opposing ways of regarding the problem, an exact opposite correspondence between the related parameters is not necessarily the case however a sort of negative correlation exists between two different models of the same problem. This discussion will be made clearer in the next chapter.

To provide a limited structure of resources, we choose an index as that of the average consumption rate of species i by species j. (However it can be seen from the above discussions that, this is not the only way to formalize the relationship between the species.) Let us call it a_{ij} thus forming a mutual consumption matrix A. As seen the parameters are nearly the same with those of the Volterra equations indicating the growth and decay parameters with respect to mutual relations. Now let b_i $i = 1, \ldots, n$ denote the limitations on the resources by that time. So;

$$Ax \leq b \qquad (1)$$

will give the set of feasible points, within the ecosystem.

Up to now we have discussed the mutual relations of populations and it is our general assumption that only these are of interest when the LP model is concerned. It is true however that there are mutual relations inside the populations which can affect the overall population dynamics, so it's worth mentioning about, at this instance.

Cannibalism for example, is a feeding pattern in which the predating species may also consume individuals of its own. This phenomenon is possible especially under tragical situations such as food scarcity or extreme crowding within certain living groups and has been a topic of interest [Stenseth, 1985].

The evolution of reciprocity is possible in small groups where mutual cooperations are frequent, however it is very difficult for larger populations. Some simulation studies have been made in this issue, repeated plays of prisoner's dilemma game (which is a noncooperative game) indicate that mutual cooperation evolves through time, however for n-person version of that game (with n much larger) this is not the case [Boyd and Richerson, 1988].

The model assumes that the species act also as resources to the ecosystem, and matrix A defines the mutual resource relationship between the species. However there are also other resources, either inorganic (such as water, oxygen) or some organisms which are beyond of our interest in the ecosystem (maybe some bacteria) and they may also have some usage limitations. Thus in the model an extended matrix will be adopted to account for these constraints as well, an $m \times n$ matrix where n represents the total number of species and m the total number of resources (of course the species being contained as a subset in the resource set,). The size of the RHS vector is also adjusted appropriately. So we redefine the constraints on resources as;

$$A'x \leq b' \qquad (2)$$

where, A' is $m \times n$ and $m \geq n$. Here, of course (1) is a subset of (2). Now to generalize an objective, we consider that of maximizing overall fitness of the universe. Let c_j represent the average fitness contained by one unit of species j. So with respect to the above constraints the objective may be stated as;

$$\max cx.$$

It has to be noted that in biological systems where evolution is taking place; the species which will originate in the system can not be known in advance. Only the principle of evolution by means of mutation and selection is known to us. Most of the mathematical models in literature assume a fixed number of species (known to the modeler a priori) possible such as the Volterra models or other molecular evolution models.

This is an aspect that can be criticized, especially for ecological models including evolution, because it doesn't provide sufficient examples of evolution of unpredictable living structures.

The basic aspects of our LP model, have been summarized in a research report [Vizvári, 1991]; stressing the fact that unpredictable, unwanted types of species can occur by means of selection.

3.5 Generalizing a Simulation Procedure for the LP Model

As stated before, the problem of evolution is quite dynamic, whereas the LP model assumes a static problem and an unchanging environment. Therefore we consider time instants and construct a particular LP model (either the one described in the previous section or another suitable model) assuming that the existing species and environmental conditions at that time are known in advance.

The LP optimization is assumed to determine the state of the universe at the next equilibrium position. Thus optimal x_j values which become zero by the optimization process are expected to be extinct within a certain period of time.

To account for the variations and appearances of new species (mutations) the model will also allow for slight modifications of the existing parameters. Thus throughout the process, favorable mutations will have a chance to survive, others will die out in a short period.

Before describing the process, some assumptions of the ecosystem which is modelized will be stated. The system is, in general assumed to be in equilibrium, but there are certain unusual, external events (catastrophes) which alter this equilibrium. At these moments struggle for life becomes important. The Lotka-Volterra interrelationships are significant in determining which of them survive the catastrophe and which of them become extinct. This is determined by the LP-optimization at a time interval, which selects the species with sufficient amount of Darwinian fitness. This time interval is very small, much smaller than the unit time taken as a basis of mutual consumption relationships, in the LP-problem.

The model adopts the 'neutralist' opinion of evolution in this sense. As may be recalled the neutralists claim that the significant amount of variations occur during very short intervals of time, where the population sizes have been reduced and altered significantly (due to some catastrophic event, in our example model), so that the effect of random drift becomes important. Otherwise the ecological equilibrium is mostly stable and evolution does not take place.

After the catastrophic event (LP-optimization) some of the populations become extinct. Here we define the concept of a 'subspecies'. When a population becomes extinct, the column of that population is not deleted from the LP problem. Instead, a subspecies is formed in its place (which might be visualized as a very crude, primitive form of a mutant descending from the extinct species) and its parameters are randomly generated with a mean value of 0. So they have minimal effect on the mutual Lotka-Volterra relationships; but they are not considered as resources of a similar type to that of their ancestors. If one of these subspecies, after a future catastrophe are viable and thus selected to be fit enough, then this newly formed species replaces that resource.

The populations which survive experience some variations in this short, bottlenecking period. The new equilibrium is restored at the end of this period and remains fairly constant until another catastrophic event is observed.

So a possible way of iterating the evolving system is given as follows:

- Step 0. Initial configuration, and a feasible basis. Let $x = (x_1, \dots x_n)$ represent the quantity vector of species at a certain time instant. Let the matrix A' and vectors b' and c are as those stated previously, and assume that they are initialized appropriately.
- Step 1. Catastrophic event; solve the optimization problem;

$$\max c'x$$

$$\text{s.t. } A'x \leq b' \ , \ x \geq 0$$

- Step 2. x_j values → 0 are expected to be extinct within a certain time. Allow for certain variations of the original populations and thus the appearance of new species. Subspecies are formed in place of the extinct populations, but they are not recognized as resources yet. Some amount of variation on the original parameters are made randomly. For subspecies the parameters are regenerated with mean zero.
- Step 3. Equilibrium is restored. Adjust the A', b', c' matrices, regarding the newly formed species. Reconstruct the LP model. The equilibrium remains unchanged until the next catastrophe.
- Step 4. Solve the LP (just as in Step 1) and continue the procedure similarly.

The process can be illustrated as that of in figure 12. The foundation of the model is based on a process; where at a certain instant with fixed parameters (Lotka-Volterra relations) an optimization is taking place to project the solution to the next future equilibrium position. This represents a certain catastrophic event, in which the population sizes are changed drastically, some species becoming extinct. During the following bottlenecking period a significant amount of variations will occur and new species which may have chances to survive will be possibly produced.

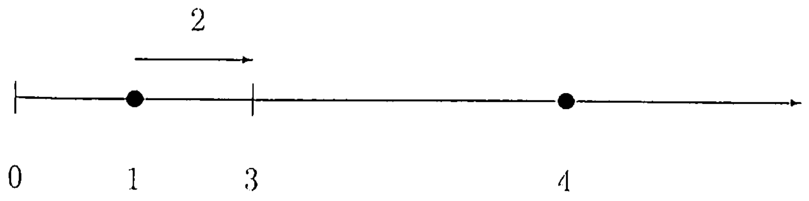


Figure 3.12: The dynamic structure of the ecology. 0 is the initial state of the populations. At point 1 a catastrophic event occurs (LP optimization takes place) and within the bottlenecking period 2 an important amount of variations occur. The new equilibrium is reached at point 3 with newly formed species as well. After this, at point 4 another catastrophic event happens (second optimization), where most of the species become extinct, and the process continues on, in a similar fashion.

The model is advantageous in some respects, when compared to the previous models discussed. The most important is that, it represents many phenomena that are related to the evolution of the universe; genesis (step 0), mutations and variations of species (step 2), catastrophic events, where a certain amount of existing species die out (step 1), followed by bottlenecking periods (steps 2 and 3), in which the effect of mutations become most important—in a smaller population variations of inheritable traits are more likely to survive—and then the immediate appearances of new species.

The final word will be about setting the b vector. It is debatable whether this should be set exactly to the previous optimal x vector (noting that b is a subvector of b' corresponding to resources which are also species in the system) or something completely unrelated to this result. Our belief is that, even the exact setting might not be appropriate, b should be somehow related to the previous x^*, therefore either we have $b = x^*$ or at least $b = f(x^*)$. In concordance with the model assumptions, such an $f(x^*)$ may be defined as $x^* + X'$ for $x^* > 0$ and 0, otherwise. This means that within the period following the restored equilibrium, the population sizes may grow or deteriorate somehow, independent of the catastrophic events.

We have already assumed that, after the equilibrium is restored (the optimal solution is determined); the population sizes remain constant, but it is actually debatable to have done so. Nevertheless for a realistic model, the catastrophic incidents may be represented by the LP optimization and the population dynamics of the periods in-between (where new variations do not occur) can be characterized by the Lotka-Volterra differential equations.

3.6 An Example Model

A typical LP model has been designed and implemented for small ecosystems of molecular structures, possibly some artificially constructed systems for certain bacteria types for laboratory purposes. The linear optimization, as stated, is confined to certain points in time. The motivation of this type of modeling comes from the Lotka-Volterra models (some set of differential equations), and the foundations are summarized in [Vizvári, 1991].

The simulation routine starts with an initial ecosystem of n species, which are also the resources to this system; therefore at a certain time there is a limited supply of them. In this model, these n species are still the fundamental living entities, but variations can occur through mutations; these variations occur in a random (previously unknown) fashion. There are also some natural resources which are not considered as species, possibly some inputing material fed from outside to the system.

The initial model is optimized and the solution is stored. For each fundamental being (the n species), a new column is generated (by perturbation) with regard to the original parameters of it. New species are also formed, one from each of the n population. The model is reoptimized again; and the new solution indicates, whether new variations of the original species have chances to survive. After this step the original parameters are readjusted by taking the weighted average of the like species (the original one and its descendants), and new variations are made by perturbing these new 'average' columns.

The process continues in this way. At each iteration n new species are generated from the original families of each population of the initial configuration. The x values which are set to 0; are considered to be extinct but they are not deleted from the problem; since they are not accounted for when the weighted averages are taken. At their places the so-called subspecies originate which have slight effect on the ecosystem, but they are not depicted as resources, until they obtain nonzero values in the following optimization processes. A sample run of the program is given in the appendix.

3.7 The Significance of an Unbounded Solution

An interesting fact after obtaining several simulation results, is that some runs have unbounded solutions. It is important to understand the meaning of an unbounded solution, at a certain iteration of these types of models.

It has to be recalled that, the ecosystem which is of concern doesn't have to satisfy the mass conservation laws; it is seldomly a closed self-sufficient system; furthermore, mostly it represents only a partial system of a larger ecosystem, which is of real interest to the modeler. (A person may be interested only in a certain number of species populations.) Thus the subsystem is most probably a system which is fed from outside.

In light of these explanations, an unbounded solution can be interpreted as follows. At that instant of time, there is not a certain fixed equilibrium point, which nature favors; but at least some of the species of concern can survive with respect to certain feasibility requirements, but in indefinitely large amounts.

In the above simulated model, we considered n species (which serve also as resources), in which the mutual relationship matrix was generated randomly by the simulation program itself. In a small ecosystem, the food chains can be represented in a simpler way, (in a sparser matrix); this is the fact in nature as well, the food chains are relatively short. On the other hand, there is an immense number of resources and most of them can not be considered as a living thing (or may not be of interest as a sort of species to the biological modeler). In the next chapter, examples of this type will be given; where a detailed analysis of the ecological relationships is made.

Chapter 4

More On LP Models of Evolution

In this chapter different approaches to the LP optimization problem faced previously will be of main concern. A preliminary discussion on this topic has already been made in the previous chapter. It has to be reminded that linear optimization is confined to a very small interval of time, in the model constructed previously and alternative approaches to only this portion of the model are discussed. Otherwise, the procedure described in the previous chapter remains the same.

Further interest has been put on the issue for the following two reasons. The first is the simple fact that different approaches represent the different aspects of the original problem. The second reason is that the similarities of the different ways of constructing the LP might lead to some results that may be quite interesting in modeling theory, if generalized to a broader class of problems.

The notation used in the previous chapter will also be adopted here. The mutual interaction matrix of the species is $A_{n\times n}$ but the whole matrix of consumption may be recalled to be as $A'_{m\times n}$. The remaining submatrix other than A will be denoted as B so that $A' = \left[\frac{A}{B}\right]$. Similarly b' is the extended RHS vector b. In the transposed problem the former decision vector x will be replaced by x', which represents not only the quantity of the species but of the resources as well. The x' values which represent only natural resources will be denoted as x_0, so that $x' = \left[\frac{x}{x_0}\right]$. The dual variable will be represented as y'.

The first approach was discussed in the previous chapter. Here it will be restated briefly. Then the dual of this model and its significance will be illustrated. After these, a different approach and a new model will be presented. Following this section, the dual of this problem will be given as well. Later the models will be compared and contrasted, and some possible improvement ideas will be discussed.

4.1 The Primal Problem of Evolution

In this section the model discussed in the latest chapter will be revisited. The idea can be recalled as maximizing the total fitness of the ecosystem that is modeled subject to the constraints of limited resources in nature.

As stated previously food chains are shorter in reality (thus the A matrix is mostly sparse) and there is a huge amount of resources when compared to the number of species being of concern to the ecological modeler. Some resources are found in abundant quantities and need not be accounted as limited, though. For example there is an abundant supply of sunshine in most places (which is essential for producing vitamin D in the body), however in some northern countries there is not enough sunshine for this purpose and living beings should purchase vitamin D from other resources as well.

Recalling the definition of the problem given in the previous chapter, the ecological optimization problem is stated as follows:

$$\begin{aligned} \max \quad & cx \\ \text{s.t.} \quad & A'x \le b' \\ \text{and} \quad & x \ge 0 \end{aligned}$$

where;

a'_{ij} : average consumption of resource i by one unit of species j.
b'_i : available quantity of resource i.
c_j : fitness of one unit of species j.
x_j : the amount of species j.

The stated problem is the species problem of maximizing the overall Darwinian fitness values of its constituting species. The constraints are the limited supplies of resources—including the species themselves. Here the optimal decision is made according to the survival abilities of the species, and the population sizes of the species (that will be calculated for a future equilibrium point in time) are the decision variables. Here the evolutionary mechanisms, mutations and selection are assumed to be external forces to the populations living in the ecosystem. In the following subsection, the dual of the problem and its significance of the dual variables will be given.

4.1.1 The Dual Problem

The main importance of the dual models will be clear in the following sections where a different LP formulation of the same problem is to be discussed. The initial idea originated, is to view some similarities with these models; where some theoretical results can be inferred and these may be generalized to serve as new concepts in modeling theory.

As was stated, with the dual model we mean the dual of the LP problem, which is considered for an instant of time, not the entire process discussed in the previous chapter. So the only difference is with the instantaneous linear optimization. Instead of the primal problem, its dual can be considered;

$$\begin{array}{rl} \min & b'^T y \\ \text{s.t.} & A'^T y \geq c^T \\ \text{and} & y \geq 0 \end{array}$$

where;

a'_{ij} : average consumption of resource i by one unit of species j.

b'_i : available quantity of resource i.

c_j : fitness of one unit of species j.

y_i : fitness contribution, per unit of resource i.

y_i may be considered as the marginal contribution to the cumulative fitness of the system by increasing resource i by one unit, provided that the optimal basis doesn't change. A possible interpretation of the dual problem can also be given in the following manner. The primal problem was concerned with maximizing, the cumulative fitness of the entire populations with respect to certain resource constraints. That is the problem represents the view of the species which constitute the ecosystem. The dual problem on the other hand, is the problem faced by Nature—thus the whole process can be regarded as a game between nature and the entire species populations—, and the objective is to determine appropriate (minimal) fitness per resource values in such a way that each species should be kept content with their donated fitness values.

4.2 The Transposed Problem

The dual problem in the previous section represented a different view of the problem but since its variables have the shadow price meaning, quantifying the values of the original variables can be made in an indirect way. In this section, another LP problem which has a similar objective and constraints to the dual of the former problem will be stated but the decision variables will be the quantities of species and the natural resources, instead.

Actually this problem is a different approach to the evolution problem that has been dealt formerly. Again the difference (as in the dual of the evolution problem) is confined to the instantaneous optimization part, otherwise the defined procedure is still valid.

The name *transposed* might falsely indicate that the constraint matrix is the transposed equivalent of the original problem, however here the meaning of the constraints is a bit different. The reason why it was called the transposed problem is that because there is a correlation between the submatrices of this problem with that of the original (dual) problem in a vague way. The insight that has led us to this approach was the apparent similarities with the original problem, that might indicate some interesting results between the two entirely different LP problems, as well as some advantageous points of the new model (when compared to the original one).

Now, the LP model can be defined in the following way;

$$\begin{array}{rl} \text{min} & dx' \\ \text{s.t.} & Ex' \geq f \\ \text{and} & x' \geq 0 \end{array}$$

where;

e_{ij} : benefit gained by species i through the existence of unit resource j.

f_i : accumulated benefit (fitness) required for species i.

d_j : energy dissipated (consumed) by one unit of resource j.

x'_j : the quantity of resource j.

Similar to the dual of the original model, this is Nature's problem, which tries to minimize the total energy dissipation in the universe, subject to minimal vital constraints of each species. However, here the decision variable x' represent the quantity vector of the resources; so it is a clearer formulation, than that of the dual problem that has been stated.

4.2.1 The Dual of the Transposed Problem

As a final LP model the dual of the transposed problem and it's significance will be stated. The problem, in opposition to the transposed problem represents the view of the species populations, just like the original problem.

$$\begin{aligned} \max \quad & f^T y' \\ \text{s.t.} \quad & E^T y' \leq d^T \\ \text{and} \quad & y' \geq 0 \end{aligned}$$

where;

y'_i : energy payable by nature per fitness level of unit species i.

The meaning of the dual variable y'_i, may look peculiar, however this is not confusing when the following fact is recalled. Each species obtains a certain level of fitness by the exploitation of certain resources, and this has some cost to nature. The argument can be validated by carrying out the dimension analysis. For the transposed problem the parameters are:

$$\begin{aligned} c_{ij} : & \quad \left(\frac{\text{fitness}}{(\text{res.j})(\text{spec.i})}\right) \\ f_i : & \quad (\text{fitness/spec.i}) \\ d_j : & \quad (\text{energy/res.j}) \\ x'_j : & \quad (\text{res.j}) \end{aligned}$$

Thus the dimension of variable y is:

$$y'_i : \left(\frac{\text{energy}}{\left[\frac{\text{fitness}}{\text{spec.i}}\right]}\right)$$

The objective in this case is to adjust the acceptable limits of energy consumable per fitness level of unit species of i (in a maximal way), so that the cumulative consumed energy of unit species per resource wouldn't exceed the restraints of energy dissipation for each resource. A careful observation of the problem will indicate the similarities of this LP to the primal problem of the first approach.

4.3 Comparison of the Two Approaches

Before starting the discussions on the two approaches, as a chronological convention we denote model I as the *primal problem* and model III as the *transposed problem*. We denote models II and IV as the duals of the former problems, respectively.

The main concern in comparing these 4 models is the apparent interrelationship between each problem. Models I and II are dual pairs; the relationship is obvious. So are the models III and IV. Models I and IV represent the same type of constrained optimization, though with entirely different variables; the same relation is true for models II and III. The most striking relation is that of models I and III (so also with II and IV), because they both represent the same problem in two different ways.

The two *opposing* problems remind the concept of duality (actually the similarity between models II and IV also are clear). The more interesting fact is that, these two problems (I and III) have the same decision variables; or at least the variable vector of model I is a subset of that of model III. This peculiarity arises from the fact that, the decision variables (quantities of the constituting species population) are also a subset of the resources of the problem.

In problem I, the relation $A'x \leq b'$ is a way of illustrating the food chains; actually the entities of the A' matrix are the consumption rates of resources by species. However, if a resource is consumed (without being eaten or utilized somehow) then this sort of behavior is not distinguished in this model. For example, if an animal kills another one without any apparent reason, this is considered as consumption as well. So the relation only maintains the availability of the resources and the average fitness values are assumed to be known and are dealt with the objective function.

Finally if the existence of species is favorable to any resources, i.e. has positive impact on the quantity of the resource, then this may be illustrated by negative consumption values.

In the other approach, (problem III); the mutual favorings and disfavorings in the ecosystem (without regarding consumption) are illustrated very clearly. This approach represents competition and struggle for life between the species in a better way.

The A' matrix in the first approach is sparser than the one in the next approach; actually the second model represents the mutual relationships in a detailed way. In the second approach however, there is not a mention about the limitations on resources in a direct way. On the other hand, the dual of the transposed problem provides a picture of limited universe with its constraints, where the energy dissipation of resources is limited, somehow.

Another difference between the two models is related to the algorithm given in the previous chapter. The result of the point optimization in the first model gives a targeting value of the population size vector and this result might be related to the amount of supplies of the species which act as resources, thus the RHS for the next equilibrium can be updated in this way. However we don't have any feedback of this sort for the natural resources. In the transposed model on the other hand, the optimal values determined can be used to update the whole RHS for the next equilibrium point.

So there is some parallelism with the two approaches when the dual models are taken into consideration as well. The similarities will be made clearer by a simple example. Of course this example is far from describing the complex structure of a real ecosystem, but it is sufficient for illustrating the structure of Lotka-Volterra type systems.

Example. Consider a simple ecosystem with the following resources.

1. lion.
2. wolf.
3. impala.

4. rabbit.
5. grass.
6. carrot.
7. lettuce.

Here the first 4 resources are species as well, the remaining ones are not considered as living entities of the ecosystem, but only inputing resources to the system.

Now according to the modeling approaches discussed, the problem may be formulated in the following ways.

1. (Primal Problem)

$$\begin{array}{lrrr}
\max & 10x_1 + 5x_2 & +6x_3 + x_4 & \\
\text{s.t.} & & & \\
 & & 0.1x_3 & \le 3 \\
 & 0.2x_1 & & \le 10 \\
 & x_1 & -0.5x_4 & \le 15 \\
 & 0.1x_1 + 2x_2 & & \le 48 \\
 & 4x_2 & +10x_3 + x_4 & \le 1000 \\
 & & 3x_4 & \le 300 \\
 & & 5x_3 + 4x_4 & \le 400 \\
 & & x_i \ge 0 &
\end{array}$$

- Time to time a lion may be killed by some impalas. (Constraint 1)
- Lions may kill the wolves time to time. (Constraint 2)
- Lions eat impalas, rabbits are beneficial to impalas. (Constraint 3)
- Rabbits are consumed by wolves mostly, and slightly by lions (Constraint 4)
- Grass is consumed by wolves, impalas and rabbits. (Constraint 5)
- Carrots are consumed by rabbits. (Constraint 6)

- Lettuce is consumed by rabbits and impalas. (Constraint 7)

III. (Transposed Problem)

$$\begin{array}{ll} \min & 20x_1 + 10x_2 + 10x_3 + 2x_4 + 0.1x_5 + 0.1x_6 + 0.1x_7 \\ \text{s.t.} & \end{array}$$

$$\begin{array}{rcl} x_3 + 2x_4 & \geq & 20 \\ -0.4x_1 + 10x_4 + 5x_5 & \geq & 30 \\ -2x_1 + x_4 + 20x_5 + 10x_7 & \geq & 10 \\ -0.5x_1 - 5x_2 + x_5 + 3x_6 + 4x_7 & \geq & 20 \\ x_i \geq 0 & & \end{array}$$

- Lions are favored by the existence of impalas and rabbits. (Constraint 1)
- Wolves are favored by the existence of rabbits and grass and disfavored by lions. (Constraint 2)
- Impala is favored by grass and lettuce and also by rabbits and disfavored by lions. (Constraint 3)
- Rabbit is favored by grass and lettuce, disfavored by lions and wolves. (Constraint 4)

The parameters are similar to those stated previously. They were generated in a rough way, bearing in mind some simple food chains, and symbiotical behavior. There are food chains like, lion-impala-lettuce or wolf-rabbit-carrot. There is competition between the two carnivores, the lion and the wolf. Symbiotical relations are clearer in the second model, for example impalas are favored by the existence of rabbits (though there is not a food chain including both species).

From the benefit-consumption relationships of the species and resources, the similarities between the two models become apparent. In matrix notation we can rewrite the problems as:

$$
\begin{array}{ll}
\multicolumn{2}{c}{\text{(I)}} \\
\max & cx \\
\text{s.t.} & \left[\frac{A}{B}\right] x \leq b \\
 & x \geq 0
\end{array}
$$

$$
\begin{array}{ll}
\text{(III)} & \\
\min & d[\tfrac{x}{x_0}] \\
\text{s.t.} & [\tilde{A}|(B')^T][\tfrac{x}{x_0}] \geq c' \\
 & x \geq 0
\end{array}
$$

To make clear the symbols in the second model, the following list of explanations is given:

- $c \rightarrow c'$: There is positive correlation between the entities.
- $B \rightarrow B'$: " "
- $A \rightarrow \tilde{A}$: The correlations are as follows; the positive (negative) A entity has a negative (positive) impact on the corresponding $\tilde{A}$ entity. Furthermore, the positive consumption entities in A have a positive impact on the corresponding $\tilde{A}^T$ entity. (Because of this additional feature, the $\tilde{A}$ matrix is more illustrative than the A matrix which is sparser.)

These correlations are not difficult to observe, the relation between c and c' seem to be clear. Since the resources consumed by the species are assumed to be beneficial to the species, there is an obvious correlation between the submatrices, B and B'. (This relationship is rather one sided, there is a transfer of benefit in one direction, from resources to species.) The bizarre relation between the mutual interaction matrices of the species arises from the following facts:

i) Two species may have prey-predator type of relationship. (One is consumed by the other, and provides a certain amount of benefit in this way.)

ii) Two species may be competitors, and therefore they may be *consuming each other* without obtaining direct benefits.

iii) Two species may be interacting in such a way that none of them consumes the other, but only one of them benefits from this.

iv) There may be symbiotical relation, none of them is consumed and both benefit.

As may be seen the mutual relations may be either antisymmetric (prey-predator), symmetric (symbiotical) or nonsymmetric at all. Consumption doesn't always mean benefits gained, so the correlation of the submatrices A and $\tilde{A}$ is a bit complicated.

So this peculiar sort of resemblance of these twin models (problems I and III) arouses further interest on the issue. The correlations between the two problem parameters make us ask the following question; is it possible to generalize this type of similarity to a certain class of optimization problems, and is it possible to find some simultaneous optimality conditions for the two problems?

The answer doesn't seem quite trivial. There is an obvious similarity between the two types of problems, but even if some relation holds it is not clear that this can be located easily. Nevertheless, some inferences are likely to be deduced from these problems, and this will possibly be a further topic of interest. The results might be useful for the general LP modeling theory and duality.

Chapter 5

Conclusion

In this research our basic motivation was to analyze certain mathematical models of evolution, and to generate some results by computer simulation, taking these mathematical models as the background. It was stated that there is a vast literature related to evolution studies and it is not claimed that a summary of the whole literature has been made in the second chapter. Evolution is a topic which has aroused so much attention that, scientists from nearly all disciplines have been concerned with it, and this is probably the reason, why such an immense literature has been accumulated which still doesn't have settled standards, axioms and fundamentals agreed upon. Despite certain developments in the evolutionary models, the topic matter is still in infancial state and interdisciplinary consensus hasn't yet formed [Feistel and Ebeling, 1989], [Brooks et. al. 1984]. So it is quite possible that some of the models, approaches and views have not been mentioned. A practical problem is the fact that most publications before the 80's are almost unobtainable, and I think that it would have been much better if certain articles and books were available during my thesis research. Especially it would be mostly invaluable if historical documents of interest had been at hand. Nevertheless I think that a brief summary of the work (at least a major portion of) that has been done until now, is given; and some views and opinions of certain scientists and philosophers are also discussed on certain issues of evolution.

The models that have been of interest were in two classes. The first class was that of a so-called autogenesis phenomenon, in which self-organization of certain (biological) molecules through a process of imperfect replication procedure was being analyzed. The model was adopted from a former research work on the issue *autogenesis* [Csányi and Kampis, 1985], [Kampis and Csányi, 1987]. In the context of this research, simulation analyses had been done for a single biosystem of molecules, and it's self-organization had been investigated independently. Apart from this, the main interest was to extend the single system model to a hierarchical multi-level system model, where each system is still concerned with its self-organization, however each are interrelated to each other by this hierarchical multi-level structure.

The results of our simulation runs indicate a similar formation of organization throughout time. The simultaneous appearances of these organizations in independent systems result with a supersystem (of higher level) which starts to self-organize itself; and at the same time the self-organization process continues on the independent systems. Finally a state is reached, where a sufficient amount of self-organization has been formed in the upper system (i.e. when similar supersystems have been self-organized as well, even higher level structure formations are possible) but the self-organization process has been stopped, since the self-organization of the upper level has destructed the existing organizations of the lower level systems. In this way most of the lower level systems have become *dead* and an upper level system (a more complicated organization) has formed on top of these. Thus the level of complexity of self-organizing systems can increase by this process.

The importance of an autogenesis process is that it represents the formation of self-organization of biological molecules (such as DNA and RNA chains) which is directly related to emergence of first living structures on earth (the single level model); and also the formation of more complex structures throughout the process of self-organizing of a larger system constituting of interactive self-organizing systems (the multi-level model).

The second class of the models that have been designed and experimented was the LP-type models or the optimization models. The essence of such models is that they represent the ecological population dynamics which exhibit evolutionary mechanisms, mutations, selection, and optimal search towards an equilibrium position.

In these population models, variations occur (dominantly) after a bottlenecking period, where population sizes are comparatively small, after a catastrophic event (here, optimization) has taken place and most of the species populations have become extinct. New species descending from the former ones have appeared and the ones that are viable will have chances to survive the next catastrophic event (another optimization). The dynamics is governed in this way; as a consecutive runs of LP optimization, where each denotes a catastrophic event in which the populations are affected drastically.

A subissue related to the LP model was situated in the final chapter. The former LP modeling has been done, based on the resource ↔ consumption balance of the populations. A different way of formulating the problem is suggested in this chapter for the same ecosystem, but with a different logic. Here the minimal survival limit ↔ accumulated benefit balance has been of concern. The idea emerged from the duality concept; the second approach objective and constraints are very similar to the objective and constraints of the dual problem of the first approach.

The objectives and the constraints for each problem can be given as:

I. Primal Problem.

Maximize cumulative fitness of the entire populations with respect to the limited resource constraints.

II. The Dual Problem.

Minimize total fitness (economize on the depletion of resources) with respect to the minimal benefit requirements of populations existing just before the time of the catastrophic incident.

III. The Transposed Problem.

Minimize total energy consumption subject to minimal accumulated benefit requirements for all populations.

IV. The Dual of the Transposed Problem.

Maximize total energy gain out of nature subject to maximal energy dissipation constraints of the resources.

A detailed analysis of these four problems will probably be of further interest to deal with. The striking similarity between the two approaches can be generalized to other models as well. The most important aspect of the evolution models is the fact that the objective variables are also defined as constrained resources of the problem.

Similar *twin models* may be constructed for other problems with the same aspect. Examples are not difficult to find, An autocatalytic chemical process, (in which there are some constraints) where several products are also inputing raw materials of the process, is one example. Another class of models is the inter-industry models (Leontief type), where each industry represents both a constraint and an objective variable.

Apart from the construction of these models and their implementations, three other areas of interest have been of major concern in this thesis research. The first one is that some discussions on the related topics of evolution have been dealt with, and the criticisms that have been placed, and some counter evolutionary theories were also of interest as well. The second one is that, although not directly related, some mathematical models and approaches of interest have been studied briefly. Finally, a relation between evolution and optimization is tried to be elaborated. The relationship between evolution and optimization is nothing new, and in fact Darwinian evolution and optimization has been resembled to a heuristic, an ascent algorithm in the phenotype space [Simon, 1981]. Both are myopic and converge to local optima (if possible).

The research study will not possibly terminate at this point, though. New optimization problems can be formulated from the mathematical models; and the graph theoretical models of phylogenesis are possibly topics of future research. Modified autogenesis models can be constructed, probably those models that can be more easily identified with the structural self-organization of DNA and RNA molecules. Simulation programs of Lotka-Volterra type ecosystems can also be written to model larger and more realistic ecosystems. And most important of all, the discussions in the latest chapter will be possibly extended, to account for some other economical or chemical models as well.

Appendix A

An Example Run of the LP Model

The dynamic structure of an ecosystem of primitive structures (such as certain bacteria) is illustrated through consecutive runs of LP optimization. 4 populations constitute the initial ecology, which are depicted as the core species. These core species also represent the first four resources in the problem. There are five more natural resources other than these. After each optimization step, new mutants appear from this basic core species. The descendants of the 1st population are the 1st and 5th, similarly for the 2nd population the descendants are the 2nd and 6th at the second run. If at any iteration, all of the descendants of a core population are extinct, then the corresponding resource is set to 0 and all the columns for the next iteration (those representing the species of the core population) with mean 0. Otherwise a weighted average column is calculated, for all mutants, and the columns for the next iteration are generated by perturbing this column randomly. To illustrate this process, after first iteration, the 3rd species remain alive so the 7th column for the second iteration is generated from the 3rd column. After second iteration, 3rd species become extinct whereas the 7th remain alive. So for the third iteration, columns 3, 7 and 11 are generated from the 7th column. It has been assumed that if a certain population has become extinct, then the descendants of this population is something quite different from the original species. Since the second population has been extinct by the end of the first optimization, the parameters of the 2^{nd} and 6^{th} columns are regenerated in the second simplex tableau with mean 0.

The optimal amounts of the population sizes are used to update the initial 4 RHS values in the next run. The procedure continues on by adding 4 new species descending from the original core populations, by random modification. The descendants of the same core are considered as similar resources, so the optimal values of all species of the same origin are summed up and the resulting value is set to the corresponding RHS value. For example, after the second iteration, the first RHS value (for the next iteration) is updated by adding x_1^* and x_5^*. So at the end of each optimization, the optimal values are added to modify the RHS of the first 4 constraints as described. The RHS values of the remaining constraints (natural resource constraints) are also modified randomly.

The output for 7 iterations is presented in the following pages. The problem formulations are written in LINDO format. As just stated, the problem at the first iteration has 4 variables and 9 constraints. At each iteration 4 new variables are added to the problem as described. The optimal solutions are also listed in LINDO format, just after the problem formulation of each iteration. The values are in the first columns of the output file. It has to be reminded that only these columns are of concern, and the rest (reduced cost columns) can be ignored.

```
MAX     3.7 X1 + 3.2 X2 + 5.7 X3 + 1.7 X4
SUBJECT TO
        2)   -2.0 X1 + 3.0 X2 + 2.0 X3 + 8.0 X4 <=   13.7
        3)   3.0 X1 - 4.0 X2 - 4.0 X3 + 7.0 X4 <=   6.1
        4)   0.0 X1 + 7.0 X2 + 7.0 X3 + 7.0 X4 <=   6.0
        5)   -2.0 X1 - 1.0 X2 - 2.0 X3 + 8.0 X4 <=   12.3
        6)   9.0 X1 + 6.0 X2 + 4.0 X3 - 1.0 X4 <=   12.1
        7)   7.0 X1 - 1.0 X2 + 3.0 X3 + 10.0 X4 <=   6.0
        8)   -3.0 X1 + 0.0 X2 + 4.0 X3 - 3.0 X4 <=   13.8
        9)   6.0 X1 + 0.0 X2 + 0.0 X3 + 0.0 X4 <=   10.3
        10)   -2.0 X1 + 7.0 X2 + 9.0 X3 - 1.0 X4 <=   10.8
END
           6.6979590      1.0000000    F  .10000000E+31
       X1  .48979590      .00000000    C  .10000000E+31
       X2  .00000000      .38571410    C  .10000000E+31
       X3  .85714290      .00000000    C  .10000000E+31
       X4  .00000000     7.7000000     C  .10000000E+31
MAX     3.1 X1 - 0.7 X2 + 5.2 X3 - 0.0 X4 + 3.3 X5 - 0.9 X6 + 5.9 X7
 + 0.4 X8
SUBJECT TO
        2)   -2.1 X1 + 0.2 X2 + 2.4 X3 + 0.2 X4 - 1.2 X5 + 0.8 X6 + 2.2 X7
 + 0.3 X8 <=   0.5
        3)   2.3 X1 + 0.5 X2 - 3.1 X3 - 0.8 X4 + 2.6 X5 - 0.2 X6 - 3.2 X7
 + 0.8 X8 <=   0.0
        4)   0.8 X1 + 0.8 X2 + 6.3 X3 - 0.6 X4 - 0.9 X5 + 0.2 X6 + 6.6 X7
 - 0.1 X8 <=   0.9
        5)   -2.2 X1 + 0.6 X2 - 1.7 X3 + 0.6 X4 - 3.0 X5 + 0.1 X6 - 1.9 X7
 + 0.2 X8 <=   0.0
        6)   9.6 X1 - 0.0 X2 + 3.7 X3 - 0.3 X4 + 8.3 X5 - 0.7 X6 + 3.7 X7
 - 0.0 X8 <=   12.3
        7)   6.4 X1 + 0.1 X2 + 3.3 X3 - 0.8 X4 + 7.9 X5 - 0.6 X6 + 2.7 X7
 + 0.4 X8 <=   14.0
        8)   -3.1 X1 + 0.7 X2 + 4.6 X3 + 0.5 X4 - 4.0 X5 - 0.2 X6 + 4.1 X7
 + 0.6 X8 <=   14.0
        9)   6.0 X1 - 0.9 X2 - 0.4 X3 - 0.6 X4 + 6.1 X5 + 0.5 X6 + 0.9 X7
 + 0.8 X8 <=   12.9
        10)   -2.8 X1 - 0.5 X2 + 9.6 X3 + 0.7 X4 - 2.7 X5 - 0.9 X6 + 9.6 X7
 - 0.6 X8 <=   10.9
END
           8.1543050      1.0000000    F  .10000000E+31
```

```
         X1  1.2407510       .00000000     C  .10000000E+31
         X2  .00000000       1.0779140     C  .10000000E+31
         X3  .00000000       .93041930     C  .10000000E+31
         X5  .76318620E-02   .00000000     C  .10000000E+31
         X6  .00000000       1.5632160     C  .10000000E+31
         X7  .49492010       .00000000     C  .10000000E+31
         X8  3.4069080       .00000000     C  .10000000E+31
         X4  5.0191900       .00000000     C  .10000000E+31
MAX      3.4 X1 - 0.4 X2 + 6.2 X3 + 0.6 X4 + 3.3 X5 - 1.3 X6 + 5.1 X7
 - 0.6 X8 + 3.7 X9 - 0.0 X10 + 6.0 X11 + 0.3 X12
SUBJECT TO
         2)    -2.7 X1 + 1.3 X2 + 1.5 X3 - 0.2 X4 - 2.0 X5 + 0.3 X6 + 1.5 X7
 + 1.1 X8 - 1.5 X9 + 1.3 X10 + 2.1 X11 - 0.3 X12 <=   1.2
         3)    1.3 X1 + 0.4 X2 - 4.0 X3 + 1.5 X4 + 2.6 X5 + 0.1 X6 - 3.0 X7
 + 0.2 X8 + 3.1 X9 - 0.2 X10 - 2.9 X11 + 1.4 X12 <=   0.0
         4)    1.3 X1 + 1.0 X2 + 6.3 X3 - 0.5 X4 + 1.6 X5 + 0.8 X6 + 6.7 X7
 - 1.0 X8 + 0.4 X9 + 0.7 X10 + 6.8 X11 + 0.3 X12 <=   0.5
         5)    -1.3 X1 + 0.9 X2 - 1.4 X3 + 1.3 X4 - 2.0 X5 - 0.1 X6 - 1.6 X7
 + 0.7 X8 - 2.7 X9 + 0.9 X10 - 1.3 X11 + 1.2 X12 <=   8.4
         6)    10.4 X1 - 0.5 X2 + 3.5 X3 - 0.8 X4 + 9.4 X5 - 0.7 X6 + 3.3 X7
 + 0.2 X8 + 9.8 X9 - 0.7 X10 + 4.5 X11 - 0.4 X12 <=   12.3
         7)    6.0 X1 - 1.5 X2 + 3.0 X3 - 0.8 X4 + 6.7 X5 + 0.2 X6 + 3.2 X7
 + 1.0 X8 + 6.6 X9 - 1.6 X10 + 3.5 X11 + 0.7 X12 <=   14.0
         8)    -3.5 X1 - 0.3 X2 + 3.3 X3 + 0.6 X4 - 2.1 X5 - 0.7 X6 + 4.6 X7
 + 1.0 X8 - 4.0 X9 - 0.5 X10 + 4.9 X11 + 0.9 X12 <=   14.0
         9)    6.4 X1 - 0.3 X2 + 1.3 X3 + 1.5 X4 + 6.9 X5 + 0.5 X6 + 0.2 X7
 - 0.1 X8 + 5.8 X9 + 0.4 X10 + 1.1 X11 + 1.5 X12 <=   12.9
        10)    -3.5 X1 - 0.1 X2 + 10.4 X3 - 1.1 X4 - 3.1 X5 - 0.1 X6 + 10.0 X7
 - 0.5 X8 - 2.3 X9 - 0.6 X10 + 10.5 X11 - 1.1 X12 <=   10.9
END
             2.5423220       1.0000000     F  .10000000E+31
        X1   .58298000       .00000000     C  .10000000E+31
        X2   .00000000       5.2793500     C  .10000000E+31
        X3   .29541930       .00000000     C  .10000000E+31
        X4   .00000000       1.8927310     C  .10000000E+31
        X5   .00000000       4.9848100     C  .10000000E+31
        X6   .00000000       3.7929580     C  .10000000E+31
        X7   .00000000       4.6480900     C  .10000000E+31
        X8   2.1190160       .00000000     C  .10000000E+31
        X9   .00000000       3.6124390     C  .10000000E+31
```

```
      X11  .00000000       4.9244730     C  .10000000E+31
      X12  .00000000       3.7123050     C  .10000000E+31
      X10  .00000000       2.6085200     C  .10000000E+31
MAX      2.5 X1 + 0.3 X2 + 6.1 X3 - 0.9 X4 + 2.5 X5 - 0.8 X6 + 5.2 X7
 - 0.7 X8 + 4.0 X9 + 0.1 X10 + 7.0 X11 - 1.0 X12 + 4.3 X13 - 0.9 X14
 + 6.5 X15 + 0.3 X16
SUBJECT TO
       2)   -3.6 X1 - 0.1 X2 + 2.1 X3 + 0.5 X4 - 3.6 X5 - 0.7 X6 + 0.5 X7
 + 0.8 X8 - 2.9 X9 + 0.6 X10 + 2.0 X11 - 0.4 X12 - 1.8 X13 + 0.2 X14
 + 2.1 X15 + 1.6 X16 <=   0.6
       3)    1.3 X1 + 0.2 X2 - 3.6 X3 + 1.5 X4 + 0.6 X5 + 0.6 X6 - 4.5 X7
 + 0.4 X8 + 1.9 X9 + 0.7 X10 - 3.9 X11 + 0.9 X12 + 1.2 X13 - 0.6 X14
 - 3.8 X15 + 0.0 X16 <=   0.0
       4)    1.6 X1 + 0.5 X2 + 5.6 X3 - 1.0 X4 + 2.2 X5 + 0.8 X6 + 7.2 X7
 - 0.6 X8 + 2.2 X9 + 0.7 X10 + 7.0 X11 - 1.4 X12 + 1.9 X13 - 0.4 X14
 + 5.3 X15 + 0.1 X16 <=   0.3
       5)   -2.1 X1 - 0.7 X2 - 1.7 X3 + 0.8 X4 - 0.7 X5 + 0.0 X6 - 1.8 X7
 + 1.3 X8 - 0.5 X9 + 0.3 X10 - 1.2 X11 + 1.9 X12 - 0.8 X13 + 0.1 X14
 - 0.9 X15 + 0.5 X16 <=   2.1
       6)   10.7 X1 - 0.8 X2 + 3.5 X3 - 0.9 X4 + 11.2 X5 - 0.4 X6 + 3.8 X7
 - 1.0 X8 + 9.7 X9 - 0.5 X10 + 3.4 X11 + 0.1 X12 + 9.8 X13 - 0.5 X14
 + 3.9 X15 - 0.9 X16 <=   12.3
       7)    5.7 X1 + 0.5 X2 + 2.1 X3 + 0.3 X4 + 6.9 X5 - 0.7 X6 + 2.6 X7
 + 0.3 X8 + 5.7 X9 - 0.3 X10 + 3.2 X11 + 0.9 X12 + 6.0 X13 + 0.8 X14
 + 3.8 X15 + 1.3 X16 <=   14.0
       8)   -2.5 X1 + 0.1 X2 + 2.9 X3 - 0.0 X4 - 4.1 X5 + 0.7 X6 + 4.0 X7
 + 1.5 X8 - 2.7 X9 - 1.0 X10 + 3.9 X11 + 1.1 X12 - 3.7 X13 + 0.7 X14
 + 3.2 X15 + 0.3 X16 <=   14.0
       9)    6.5 X1 + 0.6 X2 + 0.7 X3 + 0.6 X4 + 5.9 X5 - 0.1 X6 + 0.4 X7
 + 0.6 X8 + 6.0 X9 - 0.8 X10 + 1.7 X11 + 0.2 X12 + 5.8 X13 + 0.6 X14
 + 1.7 X15 + 1.4 X16 <=   12.9
      10)    -3.2 X1 - 0.6 X2 + 10.7 X3 - 1.0 X4 - 4.1 X5 - 1.0 X6 + 9.5 X7
 - 0.8 X8 - 2.7 X9 + 1.0 X10 + 10.9 X11 - 1.5 X12 - 3.4 X13 - 0.3 X14
 + 10.4 X15 - 0.6 X16 <=   10.9
END
           2.7526190       1.0000000     F  .10000000E+31
      X1   .00000000       .97585840     C  .10000000E+31
      X2   .00000000       .26434620     C  .10000000E+31
      X3   .00000000       .58368270     C  .10000000E+31
      X4   .00000000       .30748170     C  .10000000E+31
```

```
      X5   .00000000       2.1586620    C  .10000000E+31
      X6   .00000000       2.2825650    C  .10000000E+31
      X7   .00000000       3.5334260    C  .10000000E+31
      X8   .00000000       .36221680    C  .10000000E+31
      X9   .00000000       1.2031920    C  .10000000E+31
     X10   .00000000       1.3899640    C  .10000000E+31
     X11   .00000000       1.9214530    C  .10000000E+31
     X12  1.4674500       .00000000     C  .10000000E+31
     X13  1.3868540       .00000000     C  .10000000E+31
     X14  3.6330640       .00000000     C  .10000000E+31
     X15  .21186620       .00000000     C  .10000000E+31
     X16  .49740500       .00000000     C  .10000000E+31
MAX     3.4 X1 - 1.9 X2 + 6.5 X3 - 1.2 X4 + 5.3 X5 - 0.4 X6 + 6.6 X7
 + 0.1 X8 + 3.9 X9 - 0.6 X10 + 6.7 X11 - 0.5 X12 + 4.9 X13 - 0.0 X14
 + 6.5 X15 - 0.5 X16 + 3.8 X17 - 0.0 X18 + 7.4 X19 - 1.4 X20
SUBJECT TO
      2)   -1.0 X1 - 0.4 X2 + 3.0 X3 - 0.9 X4 - 2.7 X5 - 0.4 X6 + 2.0 X7
 - 0.3 X8 - 2.0 X9 + 0.2 X10 + 1.8 X11 - 0.1 X12 - 1.1 X13 + 0.1 X14
 + 2.4 X15 + 0.4 X16 - 1.8 X17 + 0.5 X18 + 1.9 X19 + 0.4 X20 <=   1.4
      3)    0.5 X1 + 0.1 X2 - 3.9 X3 + 0.2 X4 + 1.0 X5 - 0.6 X6 - 4.3 X7
 + 1.0 X8 + 1.3 X9 - 0.6 X10 - 3.2 X11 + 0.5 X12 + 1.9 X13 - 0.6 X14
 - 3.9 X15 + 0.4 X16 + 1.4 X17 - 0.9 X18 - 4.0 X19 + 0.2 X20 <=   3.6
      4)    1.1 X1 + 0.1 X2 + 4.8 X3 - 0.3 X4 + 0.9 X5 - 1.1 X6 + 6.2 X7
 - 1.3 X8 + 2.1 X9 + 0.6 X10 + 6.0 X11 - 2.0 X12 + 1.8 X13 - 1.0 X14
 + 5.3 X15 - 0.4 X16 + 0.9 X17 + 0.4 X18 + 5.0 X19 - 0.4 X20 <=   0.2
      5)    0.2 X1 + 1.0 X2 - 0.7 X3 + 2.5 X4 - 0.9 X5 + 0.2 X6 - 1.7 X7
 + 2.7 X8 - 0.8 X9 + 0.4 X10 - 1.3 X11 + 2.6 X12 - 0.2 X13 + 0.3 X14
 - 0.0 X15 + 1.0 X16 - 1.5 X17 + 0.2 X18 - 0.4 X19 + 2.2 X20 <=   2.0
      6)    9.4 X1 - 0.3 X2 + 4.0 X3 - 0.7 X4 + 8.9 X5 - 0.4 X6 + 4.2 X7
 - 0.9 X8 + 9.2 X9 - 1.1 X10 + 4.6 X11 + 0.2 X12 + 10.3 X13 - 1.3 X14
 + 3.4 X15 - 0.1 X16 + 10.2 X17 - 0.1 X18 + 4.5 X19 + 1.0 X20 <=   12.3
      7)    6.0 X1 + 0.4 X2 + 3.0 X3 + 0.9 X4 + 6.1 X5 + 0.1 X6 + 4.3 X7
 + 1.3 X8 + 6.7 X9 + 1.6 X10 + 2.9 X11 + 0.8 X12 + 6.4 X13 + 0.4 X14
 + 3.3 X15 + 0.9 X16 + 5.7 X17 + 0.1 X18 + 4.2 X19 + 1.2 X20 <=   14.0
      8)   -3.7 X1 + 1.3 X2 + 2.4 X3 + 0.0 X4 - 3.4 X5 + 1.4 X6 + 2.5 X7
 + 1.8 X8 - 2.8 X9 - 0.2 X10 + 3.0 X11 + 0.3 X12 - 4.7 X13 + 0.5 X14
 + 2.9 X15 + 1.7 X16 - 4.2 X17 + 1.5 X18 + 3.3 X19 + 1.9 X20 <=   14.0
      9)    5.5 X1 + 0.8 X2 + 2.2 X3 + 1.2 X4 + 6.1 X5 + 0.5 X6 + 1.3 X7
 - 0.5 X8 + 5.1 X9 + 1.1 X10 + 2.0 X11 + 0.4 X12 + 5.9 X13 + 0.9 X14
 + 1.8 X15 + 0.4 X16 + 6.0 X17 + 0.7 X18 + 2.2 X19 + 0.1 X20 <=   12.9
```

```
      10)  -3.2 X1 - 0.6 X2 + 9.4 X3 - 1.0 X4 - 4.4 X5 - 0.4 X6 + 9.4 X7
 - 1.3 X8 - 3.3 X9 - 0.8 X10 + 10.2 X11 - 1.1 X12 - 3.9 X13 - 0.8 X14
 + 9.8 X15 - 1.4 X16 - 4.1 X17 + 0.7 X18 + 10.2 X19 - 1.3 X20 <=   10.9
END
          14.742650      1.0000000     F  .10000000E+31
     X1   .00000000      2.2279050     C  .10000000E+31
     X2   .00000000      3.4941830     C  .10000000E+31
     X3   .00000000      .38393180     C  .10000000E+31
     X4   .00000000      3.4883310     C  .10000000E+31
     X5   .62201630      .00000000     C  .10000000E+31
     X6   5.5275220      .00000000     C  .10000000E+31
     X7   .00000000      .33902490     C  .10000000E+31
     X8   .49685880      .00000000     C  .10000000E+31
     X9   .00000000      1.9118740     C  .10000000E+31
    X10   .00000000      2.4124890     C  .10000000E+31
    X11   .00000000      .96781940     C  .10000000E+31
    X12   .00000000      .41254510     C  .10000000E+31
    X13   .00000000      1.4276510     C  .10000000E+31
    X15   .00000000      1.0512390     C  .10000000E+31
    X16   .00000000      1.2665060     C  .10000000E+31
    X17   .00000000      .93883200     C  .10000000E+31
    X19   1.8388220      .00000000     C  .10000000E+31
    X20   .00000000      2.6703590     C  .10000000E+31
    X14   2.8277340      .00000000     C  .10000000E+31
    X18   .00000000      1.3581660     C  .10000000E+31
MAX     3.2 X1 - 1.4 X2 + 7.0 X3 + 0.6 X4 + 4.4 X5 - 1.2 X6 + 8.3 X7
 + 0.5 X8 + 4.2 X9 - 0.8 X10 + 7.6 X11 + 0.6 X12 + 4.5 X13 - 1.2 X14
 + 7.0 X15 - 0.0 X16 + 3.5 X17 - 1.4 X18 + 6.5 X19 - 0.4 X20 + 5.1 X21
 - 0.3 X22 + 8.2 X23 + 0.4 X24
SUBJECT TO
       2)   -2.9 X1 - 0.4 X2 + 2.2 X3 + 0.6 X4 - 2.7 X5 + 0.5 X6 + 2.2 X7
 - 0.6 X8 - 1.7 X9 + 0.6 X10 + 1.8 X11 + 0.6 X12 - 1.4 X13 - 0.7 X14
 + 1.6 X15 - 0.4 X16 - 1.1 X17 + 0.1 X18 + 1.9 X19 + 0.1 X20 - 2.5 X21
 + 0.6 X22 + 1.5 X23 - 0.5 X24 <=   0.6
       3)   0.4 X1 + 0.3 X2 - 3.8 X3 + 0.9 X4 + 1.2 X5 - 1.3 X6 - 3.3 X7
 + 1.4 X8 + 1.2 X9 - 0.8 X10 - 3.6 X11 + 0.4 X12 + 1.3 X13 - 0.2 X14
 - 3.7 X15 + 0.1 X16 + 0.7 X17 - 1.5 X18 - 4.6 X19 + 0.3 X20 + 0.3 X21
 - 0.7 X22 - 3.2 X23 + 1.2 X24 <=   8.4
       4)   0.1 X1 - 0.4 X2 + 4.9 X3 - 1.0 X4 + 1.7 X5 - 1.2 X6 + 4.1 X7
 - 1.6 X8 + 1.0 X9 - 0.9 X10 + 4.2 X11 - 1.3 X12 + 1.3 X13 - 1.0 X14
```

```
 + 4.7 X15 - 1.7 X16 + 0.3 X17 - 0.9 X18 + 5.1 X19 - 2.1 X20 + 1.5 X21
 - 0.2 X22 + 4.9 X23 - 2.1 X24 <=   1.8
       5)   -1.7 X1 + 0.2 X2 - 0.2 X3 + 2.6 X4 - 1.3 X5 + 1.1 X6 - 0.2 X7
 + 1.9 X8 - 0.8 X9 + 0.1 X10 + 0.3 X11 + 2.4 X12 - 2.3 X13 + 0.3 X14
 - 0.6 X15 + 3.2 X16 - 1.6 X17 + 0.8 X18 - 0.8 X19 + 3.2 X20 - 0.8 X21
 + 0.6 X22 - 0.5 X23 + 3.4 X24 <=   0.5
       6)   10.0 X1 - 0.6 X2 + 4.9 X3 - 0.6 X4 + 9.7 X5 - 0.6 X6 + 4.0 X7
 - 1.8 X8 + 9.9 X9 - 1.2 X10 + 3.6 X11 - 1.5 X12 + 9.3 X13 - 1.0 X14
 + 3.9 X15 - 0.0 X16 + 10.3 X17 - 0.4 X18 + 3.7 X19 - 1.9 X20 + 9.5 X21
 + 0.1 X22 + 5.4 X23 - 1.6 X24 <=   6.0
       7)   6.4 X1 + 1.0 X2 + 4.5 X3 + 0.5 X4 + 5.4 X5 - 0.7 X6 + 4.9 X7
 + 1.7 X8 + 5.6 X9 + 0.1 X10 + 3.6 X11 + 1.1 X12 + 6.0 X13 - 0.5 X14
 + 3.5 X15 + 0.5 X16 + 5.7 X17 - 0.5 X18 + 3.6 X19 + 2.1 X20 + 5.6 X21
 + 0.1 X22 + 5.0 X23 + 1.8 X24 <=   5.6
       8)   -4.9 X1 + 0.5 X2 + 3.2 X3 + 0.9 X4 - 3.9 X5 + 0.7 X6 + 3.7 X7
 + 2.5 X8 - 5.0 X9 + 0.7 X10 + 4.3 X11 + 2.6 X12 - 4.8 X13 + 1.1 X14
 + 2.5 X15 + 1.9 X16 - 3.6 X17 + 2.4 X18 + 3.9 X19 + 0.9 X20 - 4.1 X21
 + 1.1 X22 + 4.0 X23 + 2.0 X24 <=   9.8
       9)   6.0 X1 + 0.2 X2 + 2.6 X3 - 0.6 X4 + 5.3 X5 - 0.4 X6 + 1.4 X7
 - 1.4 X8 + 5.2 X9 + 0.3 X10 + 2.4 X11 - 1.1 X12 + 5.7 X13 + 1.1 X14
 + 1.4 X15 - 0.5 X16 + 6.8 X17 + 1.2 X18 + 2.6 X19 - 0.6 X20 + 6.5 X21
 - 0.5 X22 + 1.8 X23 - 0.5 X24 <=   8.4
      10)   -4.7 X1 - 0.8 X2 + 9.8 X3 - 0.9 X4 - 4.1 X5 - 0.3 X6 + 9.9 X7
 - 2.2 X8 - 4.3 X9 - 1.1 X10 + 10.2 X11 - 1.7 X12 - 4.5 X13 + 0.2 X14
 + 11.0 X15 - 1.7 X16 - 4.4 X17 - 0.8 X18 + 11.0 X19 - 0.9 X20 - 3.4 X21
 - 0.7 X22 + 10.7 X23 - 1.6 X24 <=   12.9
END
          7.1395990      1.0000000    F  .10000000E+31
     X1  .00000000       .78368420    C  .10000000E+31
     X2  .00000000      1.8877320     C  .10000000E+31
     X3  .00000000      1.7671250     C  .10000000E+31
     X4  .00000000       .33010500    C  .10000000E+31
     X5  .00000000       .63413750    C  .10000000E+31
     X6  .00000000       .99766980E-01C  .10000000E+31
     X7 1.3243560        .00000000    C  .10000000E+31
     X8  .00000000       .31725800    C  .10000000E+31
     X9  .00000000       .59255930    C  .10000000E+31
    X10  .24114270       .00000000    C  .10000000E+31
    X11  .00000000       .73146430E-01C  .10000000E+31
    X12  .00000000       .35693160    C  .10000000E+31
```

```
       X13  .14821610       .00000000     C  .10000000E+31
       X14  3.6055130       .00000000     C  .10000000E+31
       X15  .00000000       .58383410     C  .10000000E+31
       X17  .00000000       .19542650     C  .10000000E+31
       X18  .00000000       .55388560     C  .10000000E+31
       X19  .00000000       1.4394840     C  .10000000E+31
       X20  .00000000       1.8811380     C  .10000000E+31
       X21  .00000000       .22778350     C  .10000000E+31
       X22  .00000000       .51058280     C  .10000000E+31
       X23  .00000000       .82742810     C  .10000000E+31
       X24  .00000000       .92974390     C  .10000000E+31
       X16  .00000000       .56158270     C  .10000000E+31
MAX      3.7 X1 - 1.9 X2 + 7.8 X3 - 0.2 X4 + 3.1 X5 - 0.3 X6 + 7.5 X7
 - 0.6 X8 + 4.1 X9 - 1.0 X10 + 7.6 X11 + 0.4 X12 + 3.5 X13 - 2.0 X14
 + 9.0 X15 + 0.5 X16 + 5.2 X17 - 1.2 X18 + 8.9 X19 + 0.0 X20 + 4.7 X21
 - 0.5 X22 + 8.1 X23 + 0.8 X24 + 4.9 X25 - 2.1 X26 + 7.9 X27
 - 1.2 X28
SUBJECT TO
       2)   -0.3 X1 - 0.2 X2 + 1.1 X3 - 0.2 X4 + 0.1 X5 - 0.8 X6 + 1.8 X7
 - 0.1 X8 - 0.9 X9 + 0.1 X10 + 1.4 X11 + 0.2 X12 + 0.0 X13 - 0.9 X14
 + 1.2 X15 + 0.6 X16 - 1.8 X17 - 1.4 X18 + 2.8 X19 - 0.2 X20 - 0.7 X21
 - 0.9 X22 + 1.7 X23 + 0.1 X24 - 1.6 X25 - 0.7 X26 + 1.5 X27
 + 0.8 X28 <=   0.1
       3)    0.7 X1 - 1.2 X2 - 3.1 X3 - 0.4 X4 - 0.4 X5 + 0.6 X6 - 3.7 X7
 + 0.8 X8 + 0.1 X9 - 0.2 X10 - 3.0 X11 + 1.3 X12 + 1.5 X13 - 0.6 X14
 - 3.1 X15 + 0.2 X16 + 1.1 X17 + 0.4 X18 - 3.9 X19 + 0.5 X20 + 2.3 X21
 - 0.2 X22 - 3.4 X23 + 1.2 X24 - 0.0 X25 + 0.6 X26 - 3.6 X27
 + 1.4 X28 <=   3.8
       4)    1.0 X1 - 1.5 X2 + 3.5 X3 - 1.1 X4 + 0.4 X5 - 2.0 X6 + 3.8 X7
 + 1.4 X8 + 0.8 X9 - 1.2 X10 + 3.9 X11 - 0.7 X12 + 0.7 X13 - 1.7 X14
 + 4.3 X15 - 0.2 X16 + 0.7 X17 - 1.8 X18 + 3.5 X19 - 1.0 X20 + 2.1 X21
 - 0.1 X22 + 4.0 X23 + 1.2 X24 + 0.9 X25 - 1.1 X26 + 4.4 X27
 - 1.0 X28 <=   1.3
       5)   -1.9 X1 + 0.5 X2 + 0.2 X3 + 0.7 X4 - 2.7 X5 + 0.5 X6 - 0.7 X7
 - 1.1 X8 - 3.3 X9 - 0.3 X10 + 0.7 X11 - 0.6 X12 - 1.6 X13 + 0.2 X14
 - 0.4 X15 + 0.4 X16 - 1.5 X17 - 0.1 X18 + 0.8 X19 + 0.8 X20 - 0.8 X21
 + 0.9 X22 - 0.6 X23 + 0.9 X24 - 3.0 X25 + 1.3 X26 + 0.5 X27
 - 1.0 X28 <=   0.0
       6)    9.1 X1 - 1.7 X2 + 2.9 X3 - 1.2 X4 + 10.5 X5 - 1.8 X6 + 3.1 X7
 - 1.2 X8 + 10.1 X9 - 0.6 X10 + 4.4 X11 - 0.9 X12 + 8.7 X13 - 1.7 X14
```

```
 + 4.2 X15 - 1.2 X16 + 9.9 X17 - 0.4 X18 + 4.1 X19 - 1.3 X20 + 8.8 X21
 - 0.4 X22 + 3.5 X23 - 0.9 X24 + 9.8 X25 - 1.0 X26 + 4.1 X27
 - 1.4 X28 <=   10.8
       7)   5.7 X1 - 1.4 X2 + 4.6 X3 + 1.1 X4 + 7.3 X5 - 0.6 X6 + 4.0 X7
 + 0.4 X8 + 6.8 X9 - 0.5 X10 + 5.2 X11 + 1.2 X12 + 4.8 X13 + 0.1 X14
 + 4.2 X15 - 0.9 X16 + 6.2 X17 + 0.0 X18 + 4.9 X19 + 0.9 X20 + 5.4 X21
 + 0.4 X22 + 3.4 X23 + 1.3 X24 + 7.6 X25 + 0.2 X26 + 3.8 X27
 - 1.5 X28 <=   5.4
       8)   -5.6 X1 + 0.9 X2 + 4.0 X3 + 0.6 X4 - 4.3 X5 + 0.2 X6 + 3.3 X7
 + 0.8 X8 - 6.5 X9 + 1.1 X10 + 3.1 X11 + 0.7 X12 - 4.6 X13 + 1.0 X14
 + 3.9 X15 + 0.1 X16 - 5.8 X17 + 1.4 X18 + 4.4 X19 - 0.0 X20 - 4.0 X21
 + 2.0 X22 + 4.9 X23 + 0.4 X24 - 5.5 X25 + 2.1 X26 + 4.0 X27
 + 0.6 X28 <=   14.6
       9)   4.2 X1 + 0.9 X2 + 1.9 X3 - 1.2 X4 + 4.9 X5 + 2.0 X6 + 1.3 X7
 - 0.5 X8 + 3.9 X9 + 0.8 X10 + 2.1 X11 - 0.5 X12 + 6.6 X13 + 1.4 X14
 + 1.2 X15 + 1.3 X16 + 4.6 X17 + 1.8 X18 + 1.7 X19 - 0.7 X20 + 6.9 X21
 + 1.9 X22 + 2.7 X23 - 0.9 X24 + 6.2 X25 + 1.9 X26 + 1.1 X27
 + 0.2 X28 <=   6.2
       10)   -5.5 X1 + 0.3 X2 + 9.9 X3 + 0.0 X4 - 6.2 X5 + 1.0 X6 + 10.2 X7
 - 0.8 X8 - 4.2 X9 - 0.3 X10 + 10.8 X11 - 0.9 X12 - 3.1 X13 + 0.9 X14
 + 9.1 X15 + 1.7 X16 - 5.3 X17 - 0.8 X18 + 10.8 X19 - 0.7 X20 - 2.8 X21
 + 1.1 X22 + 9.3 X23 + 1.3 X24 - 3.1 X25 + 1.1 X26 + 10.2 X27
 - 1.1 X28 <=   8.2
END
          10.866990      1.0000000     F  .10000000E+31
      X1  .00000000      2.0210970     C  .10000000E+31
      X2  .00000000      .66486670     C  .10000000E+31
      X3  .00000000      1.9438510     C  .10000000E+31
      X4  .00000000      .35063070     C  .10000000E+31
      X5  .00000000      4.4988230     C  .10000000E+31
      X6  1.3960290      .00000000     C  .10000000E+31
      X7  .00000000      2.0846520     C  .10000000E+31
      X8  .00000000      .65653850     C  .10000000E+31
      X9  .00000000      2.4139890     C  .10000000E+31
     X10  .00000000      .69058050     C  .10000000E+31
     X11  .00000000      3.3905760     C  .10000000E+31
     X12  .00000000      .29141740     C  .10000000E+31
     X13  .00000000      3.4013670     C  .10000000E+31
     X14  .00000000      2.0282250     C  .10000000E+31
     X15  1.1175050      .00000000     C  .10000000E+31
```

```
X16  .00000000    .20770720   C  .10000000E+31
X17  .43537760    .00000000   C  .10000000E+31
X18  .00000000    .39262150   C  .10000000E+31
X19  .00000000    2.6693820   C  .10000000E+31
X21  .00000000    2.6776340   C  .10000000E+31
X22  .00000000    1.3737110   C  .10000000E+31
X23  .00000000    1.2502720   C  .10000000E+31
X24  .00000000    .69141220   C  .10000000E+31
X25  .00000000    3.2531140   C  .10000000E+31
X26  .00000000    2.7782940   C  .10000000E+31
X27  .00000000    1.2562800   C  .10000000E+31
X28  .86309250    .00000000   C  .10000000E+31
X20  .15488430    .00000000   C  .10000000E+31
```

Appendix B

Glossary of Some Biological Terms

adaptation, an inherited, favorable variation, that improves the chance of survival of the individual.

allele, a certain gene form, which occurs on a certain gene locus.

chromosomes, the bunch of biological structures in a living cell which carry the genetic code, the genotypic configuration of the cell.

diploid, the case of having chromosomes in pairs.

DNA, deoxyribonucleic acid; a stable molecule in which the genetic information of an organism is stored.

genotype, one of the possible genetic codes, which characterizes the living structure.

haploid, the case of having single chromosomes.

Hardy-Weinberg equilibrium, the hypothetical state where the allele frequencies of a certain population reach an equilibrium, and remain stable. This is the end state in which no further evolution occurs. A population reachs the Hardy-Weinberg equilibrium when the following conditions hold:

- No further mutations are taking place.
- Population size is infinitely large.
- Population is isolated.
- No selection, equal viability.
- Random mating.

isolation, the separation of two or more populations of the same origin due to natural conditions, so that mating in-between is not possible. This phenomenon is the source of vast variability of many species which have descended from a certain archaical species.

locus, a position on the chromosome where a certain genetical characteristic is identified.

mutation, an unusual change, which causes the external distortion of the chromosomes which affects the genotype drastically.

phenotype, a set of genotypes that are not distinguishable by nature and selectionary mechanisms.

phylogeny, a hypothetical tree which illustrates the descendance relations between families of species.

RNA, ribonucleic acid; a molecule which is used in translating the genetic message of DNA into actual protein structure.

selection, the process of elimination of individuals, populations or alleles which are not viable, throughout time.

variation, formation of different phenotypes through mutations, fertilizations, cross-overs and other abnormal phenomena.

Bibliography

[1] P. M. Allen and J. M. McGlade. Optimality, adequacy and the evolution of complexity. In *Structure, Coherence and Chaos in Dynamical Systems*, pages 3–21. Manchester University Press, Manchester, 1988.

[2] Marjorie A. Asmussen. Evolution of dispersal in density regulated populations: A haploid model. *Theoretical Population Biology*, 23:281–299, 1983.

[3] Edward Beltrami. *Mathematics for Dynamic Modeling*. Academic Press, Orlando, Florida, 1987.

[4] Robert Boyd and Peter J. Richerson. The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, 132:337–356, 1988.

[5] Daniel R. Brooks, Paul H. Leblond, and D. David Cumming. Information and entropy in a simple evolution model. *Journal of Theoretical Biology*, 109:77–93, 1984.

[6] Reinhard Bürger. Dynamics of the classical genetic model for the evolution of dominance. *Mathematical Biosciences*, 67:125–143, 1983.

[7] Reinhard Bürger. Mutation-selection models in population genetics and evolutionary game theory. *Acta Applicandae Mathematicae*, 14:75–89, 1989.

[8] Vincent P. Crawford. Nash equilibrium and evolutionary stability in large- and finite population "playing the field" models. *Journal of Theoretical Biology*, 145:83–94, 1990.

[9] Tony Crilly. The roots of chaos—a brief guide. In A. J. Crilly, R. A. Earnshaw, and H. Jones, editors, *Fractals and Chaos*. Springer-Verlag, Berlin, 1991.

[10] Vilmos Csányi and George Kampis. Autogenesis: The evolution of replicative systems. *Journal of Theoretical Biology*, 114:303–321, 1985.

[11] Francis Darwin. *The Autobiography of Darwin and Selected Letters*. Dover Publications Inc., New York, first edition, 1958.

[12] Manfred Eigen. Selforganization of matter and the evolution of biological macromolecules. *Die Naturwissenschaften*, 58:465–523, 1971.

[13] Manfred Eigen and Peter Schuster. The hypercycle: A principle of natural self-organization, part a: Emergence of the hypercycle. *Die Naturwissenschaften*, 64:541–565, 1977.

[14] Manfred Eigen and Peter Schuster. The hypercycle: A principle of natural self-organization, part b: The abstract hypercycle. *Die Naturwissenschaften*, 65:7–41, 1978.

[15] Manfred Eigen and Peter Schuster. The hypercycle: A principle of natural self-organization, part c: The realistic hypercycle. *Die Naturwissenschaften*, 65:341–369, 1978.

[16] Péter L. Erdös and László A. Székely. Evolutionary trees: An integer multicommodity max-flow-min-cut theorem. Technical Report Memorandum No. 953, Faculty of Applied Mathematics, University of Twente, 7500 AE Enschede, The Netherlands, 1991.

[17] R. Feistel and W. Ebeling. *Evolution of Complex Systems*. VEB Deutscher Verlag de Wissenschaften, Berlin, 1989.

[18] L. R. Foulds. Testing the theory of evolution: A novel application of combinatorial optimization. *Discrete Applied Mathematics*, 15:271–282, 1986.

[19] R. Galar. Evolutionary search with soft selection. *Biological Cybernetics*, 60:357–364, 1989.

[20] John Haigh. The distribution of evolutionarily stable strategies. *Journal of Applied Probability*, 25:233–246, 1988.

[21] Hermann Haken. *Information and Self-Organization*. Springer-Verlag, Berlin, 1988.

[22] Guy Jumarie. Towards a mathematical theory of autopoiesis. *Cybernetica*, 30:61–89, 1987.

[23] George Kampis. Different forms of causation in dynamical systems. *Futura*, 8:21–35, 1989.

[24] George Kampis. *Self-Modifying Systems in Biology and Cognitive Science*. Pergamon Press, Oxford, 1991.

[25] George Kampis and Vilmos Csányi. A computer model of autogenesis. *Kybernetes*, 16:169–181, 1987.

[26] Samuel Karlin. Mathematical models, problems, and controversies of evolutionary theory. *Bulletin of the American Mathematical Society*, 10:221–273, 1984.

[27] Samuel Karlin and Howard M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, Orlando, Florida, 1981.

[28] Motoo Kimura. Some models of neutral evolution, compensatory evolution, and the shifting balance process. *Theoretical Population Theory*, 37:150–158, 1990.

[29] Gabriela Kirlinger. Two predators feeding on two prey species: A result on permanence. *Mathematical Biosciences*, 96:1–32, 1989.

[30] Fima C. Klebaner. Conditions for fixation of an allele in the density-dependent wright-fisher models. *Journal of Applied Probability*, 25:247–256, 1988

[31] P. Kůrka. Darwinian evolution in games with perfect information. *Biological Cybernetics*, 55:281–288, 1987.

[32] Sabin Lessard. Evolutionary stability: One concept, several meanings. *Theoretical Population Biology*, 37:159–170, 1990.

[33] U. Liberman. External stability and ess: Criteria for initial increase of new mutant allele. *Journal of Mathematical Biology*, 26:477–485, 1988.

[34] Eduardo M. Munoz and James F. Selgrade. Asymptotic behavior for a competitive model with genetic variation. *SIAM Journal of Applied Mathematics*, 49:1768–1778, 1989.

[35] Ilya Prigogine and Isabelle Stengers. *Order Out of Chaos*. Bantam Books, New York, 1984.

[36] Anatol Rapoport. Applications of game theoretic concepts in biology. *Bulletin of Mathematical Biology*, 47:161–192, 1985.

[37] Peter Schuster. A molekuláris evolúció elméleti modelljei. *Fizikai Szemle*, 39:292–303, 1989.

[38] Peter Schuster and Karl Sigmund. Replicator dynamics. *Journal of Theoretical Biology*, 100:533–538, 1983.

[39] Peter Schuster and Karl Sigmund. Fixation probabilities for advantageous mutants: A note on multiplication and sampling. *Mathematical Biosciences*, 95:37–51, 1989.

[40] Herbert A. Simon. *The Sciences of the Artificial*. The MIT Press, Cambridge, Massachusetts, second edition, 1981.

[41] Cecie Starr and Ralph Taggart. *Biology*. Wadsworth Publishing Company, Belmont, California, fifth edition, 1989.

[42] Nils Chr. Stenseth. On the evolution of cannibalism. *Journal of Theoretical Biology*, 115:161–177, 1985.

[43] T. L. Vincent. Evolutionary games. *Journal of Optimization Theory and Applications*, 46:605–612, 1985.

[44] Thomas L. Vincent and Joel S. Brown. Stability in an evolutionary game. *Theoretical Population Biology*, 26:408–427, 1984.

[45] Béla Vizvári. An experimental model for the evolution. Technical Report IEOR-9108, Faculty of Engineering and Science, Bilkent University, Bilkent, Ankara, 1991.

[46] Claude A. Welch, Daniel I. Arnon, et. al. *Biological Science*. Houghton-Mifflin, Boston, third edition, 1976.

[47] Mariusz Ziólko and Jan Kozlowski. Evolution of body size: An optimization model. *Mathematical Biosciences*, 64:127–143, 1983.