An Ecological Approach to the Evolution of Organism Complexity

Mikel Maron
Evolutionary and Adaptive Systems
University of Sussex

August 26, 2004

Abstract

How evolution led to complex life is one of the great questions. This paper describes simulations that investigate the role of ecological interactions in the evolution of complexity. Webworld is a robust model of evolution in food webs. It is extended for variability of organism complexity under evolution. Statistical and network analysis indicates a clear tendency for complexification within the model, led by adaptations that initially disconnect the species from trophic interactions. This suggests a process where short term fitness is increased by less connection to the ecosystem, but long term fitness is insured by incorporation within the ecosystem. Certainly it suggests a greater role for ecosystems in the evolution of complexity.

Acknowledgments

Many thanks go to Ezequiel Di Paolo for guidance throughout this project, Anna for her support, and Flexifoil Kites for great distraction.

Contents

List of Figures

List of Tables

1 Introduction

Life is complex. There is no sure way to measure it, no definition of it, and disagreement on its importance, yet it is apparent organisms have become more complex over the course of evolution. The conjecture proposed is that ecological interactions have a core role in the process of complexification. Within an ecosystem, a species must adapt to numerous relationships with predators, prey, competitors, and mutualists. Surviving within such variety could seemingly require more complex adaptations. In turn, the increasing complexity of species could lead to more stable, intricate ecosystems. The feedback of this process of organism and ecosystem adaptation may lead to ever increasing complexity.

To investigate, the Webworld model of evolution in food web assembly is extended to account for species complexity. Food webs describe predatory interactions, one of the primary determinates of ecosystem composition. Webworld incorporates foraging, population, and evolutionary dynamics to construct model food webs with very good correspondence to real food webs. Species in Webworld have an abstract representation in a list of features, which are used in calculation to determine the strength of trophic interactions. The extension allows these features to vary in complexity under the force of evolution. With this extended simulation, statistical and network analysis is employed to determine the presence, extent, and causes of any ecosystem effects in the evolution of complexity.

Section 2 begins with a survey of the study of the evolution of complexity in biology, and then the history of food web research. The Webworld model is introduced in theoretical detail. Section 3 goes through the details of implementing and running Webworld, and the specific changes required for complexity variance in evolution. Section 4 reports metrics on food webs evolved in this extended Webworld, followed by further experiments suggested by the results. Section 5 discusses the meaning of the results, both within the model and to research generally. Section 6 concludes with future research directions.

2 Background

2.1 Evolution of Complexity

Defining complexity, especially in biology, is tricky. The most formalized definition and measure of complexity comes from physics and information theory, in entropy. The second law of thermodynamics defines entropy as an increasing process in the universe, where over time the amount of useful energy declines irreversibly into heat. Maxwell's Demon is able to decrease entropy by extracting an equal quantity of information from the system [23], making entropy a measure of the amount of uncertainty in a system. A very simple system has low entropy, while a completely random system has high entropy. Biological and other complex systems are somewhat subjectively placed in a region between the extremes of entropy measurement, where neither too simple or too random, they persist in 'interesting' ways. Chris Langton, from his studies on cellular automata, famously wrote that "life exists on the edge of chaos" [21].

Within evolutionary biology this idea has been rigorously opposed. Fisher's Fundamental Theorem of Natural Selection states that through evolution, fitness will increase [15]. The theorem itself applies only to quantifiable genetic variance within a species, yet in principle it holds for macroevolution as well. It does not necessarily follow that this increase in fitness requires an increase in complexity. Stephen J. Gould further argues that complexification is only an anthropomorphic view of humans as the pinnacle of an evolutionary sequence [18]. Bacteria are the most widespread and successful organisms, yet have remained simple for billions of years. As the simplest form of life, bacteria press close to what Gould calls the "left wall", which represents the structural barrier beyond which life is impossible. Organisms moving to the right, towards more complexity, are simply the result of a stochastic, statistical spread against this barrier. Further, if complexity was a force in evolution, there would be evidence of a consistent increase. Yet the fossil record shows that complexification doesn't occur steadily as an unrelenting process, but in mass extinctions or random bursts of punctuated equilibria. This view is the extreme, and Biology has addressed other possibilities from many approaches.

Maynard Smith and Szathmry agree in the "fallacy of progress", but do allow for the evolution of complexity through major transitions in the way information is transmitted between generations [25]. A common feature of these transitions is that replication can no longer take place independently, but only as part of a larger whole. For example, chromosomes are collectively reproducing genes, individuals in sexual populations must find a mate, cells within a multicellular organism only exist as part of the larger organism. Each transition must provide immediate advantage to the replicator. Often this takes the form of a division of labor, where replication becomes more efficient through differentiation and cooperation of sub-units. The concept of contingent irreversibility is introduced to explain why evolution continues to spread away from the "left wall". In many cases, though not absolutely, an entity that replicates as part of a larger whole for a time will accidentally lose the capacity for independent replication. In a sense, additional "left walls" are built repeatedly as organisms increase in complexity towards the right.

Beyond the means of reproduction, complexification has been studied in a number of ways, though not in such detail. Bonner defined complexity as the number of cell types within the organism [6]. This is very nearly countable for all organisms. It has a relation to function, unlike the number of genes in the organism, which has no consistent relationship to organism complexity. Gould's objections to complexity are addressed by focusing on the most complex living things, rather than on life in general, and its shown that under the cell type measure, the complexity of the most complex organisms has increased. To explain this consistent increase, number of cell types is correlated with body size. Its assumed that there is an advantage to larger body size. Larger size means a more complex organism, as basic processes like respiration and digestion require more internal structure. However, it's only necessary to look at the rise of mammals after the dinosaurs to see a problem with this argument. Every adaptation must have immediate short term advantage [16]. For mammals, though eventually more complex than the dinosaurs, survived whatever catastrophe lead to extinction of the dinosaurs because of smaller size.

Evolutionary Development has been actively extending this line of research, investigating the evolution of organism structure beyond number of cell types [7]. Cell type numbers plateau in lineages, while structure becomes more complex. Complexity in development is formally defined in measures of the number of different physical parts and the different interactions among these parts, at different temporal and spatial scales. In particular, the development and reuse of modular parts is significant to the development of greater complexity. To distinguish between "left wall" statistical increases in complexity, and complexity as adaptation, evolutionary trends are statistically evaluated as passive or active. This methodology has shown the evidence of actively increasing morphological complexity.

Biological subsystems involved in information processing, like genetic regulatory networks, the immune system, chains of hormonal reactions, and neural systems, are amenable to complexity analysis [4], but any trends haven't been placed within an evolutionary framework. Ethology could also offer a perspective on the evolution of complexity, since an animals' behavior is a prime factor in selection, but is also quite far from a macroevolutionary framework. Within microevolution of behavior, the most formal work has been in Evolutionary Stable Systems(ESS). However, complexity is not necessarily more adaptive here; in the Prisoner's Dillemma, the simple strategy of Tit for Tat is an ESS [2].

The role of ecosystems in evolution has been largely under examined, and rests on unchallenged assumptions. Bennett looks at the possible links, and in an argument similar to Gould's, states that ecosystems have little role in long term evolutionary change [3]. Ecological interactions may be responsible for microevolutionary change, but from the fossil record, random and significant changes in the climate wipe out much of the change accumulated through ecosystem development. Which organisms survive these massive climate changes is largely luck and not related to complexity.

Besides these conclusions from paleobiology, there is little empirical or theoretical work exploring the question of ecosystems and evolution of organism complexity. Like many questions oriented on such time scales, it is impossible to conduct direct experiments and the fossil evidence is patchy. Models and simulations are now nearing the stage where questions of this scale can be pondered more substantially.

2.2 Review of Food Web Research

Food webs are a representation of predation relationships within an ecosystem. They have been consistently employed as tools in the empirical and theoretical study of ecology, from its largely descriptive beginnings to the present day as a mathematical sophisticated science. A food web was originally formed as a matrix, where every species is aligned along the columns and rows of the matrix, with 1 assigned to each entry if there was a trophic relationship between the two species, and 0 otherwise.

This representation presented a great amount of data in compressed form, but was not wholly embraced by field ecologists. In constructing a food web, several delineations must be chosen. The spatial scale of the ecosystem, defining the edge of study, is porous and always open to influence beyond an arbitrary boundary. Not all species can be practically included in a food web; accounting for all bacteria would be extremely difficult. Considering migratory species and seasonal variations, the species composition of ecosystems are temporal and any study might miss this interactions. Central to empiricists objections were the great simplifications in focusing only on trophic interactions, ignoring the effects of competition, mutualism, parasitism, and all the other unique factors of ecosystems that made them interesting to study.

Still, food webs demonstrated for all that ecosystems were complicated, and this complexity was accepted as the reason ecosystems persisted in stable compositions of species over time. That is until Robert May employed food webs as a theoretical tool to show that stability isn't equated with complexity [24]. May took randomly constructed model food webs, and performed linear stability analysis. The food webs were assumed to be in an equilibrium state, then perturbed slightly and subsequent system behavior analysed. May found that the greater number of species, or greater connectance between the species, led to less stable equilbiria. This was a direct contradiction of accepted wisdom, and led to greater interest in mathematical analysis of ecosystems.

Many of May's methods were controversial, particularly the use of randomly generated food webs, and shed doubt on his conclusions. Since the world was full of incredibly complex ecosystems, what properties led to stability and persistence? Certainly food webs aren't constructed in random fashion, and the webs May constructed had little resemblance to properties of real food webs, in the number of and relationships between trophic levels [22]. When real food webs were subjected to May's analysis, they were more stable than randomly assembled matrices [38].

Population dynamics were absent from May's linear analysis, and it was thought population changes might increase stability, or at least have significant effect on trophic relationships. Models of population dynamics were well established for two species, so early work started with food chains, three species Lotka-Volterra models and the like. Lotka-Volterra in a three species chain resulted in unrealistic situations, with either both predators going extinct, or the top predator and bottom prey increasing indefinitely [8]. The logistic version of the equations fared better, but was only stable under restrictive conditions. However, it was found that with the additional links, so that the top predator had an additional prey, the Lotka-Volterra was stable. These "weak links" stabilised the food web under population changes [26]. This lead to investigation of alternate functional responses, the equation governing population interaction of two species, with dynamics that encouraged stable webs and biological realism. The Holling equation incorporates prey handling time, the Beddington response adds competition between predators, and adaptive foraging in particular makes "weak links" a natural occurence. The proper functional response is still a topic of debate [1].

Real food webs are not constructed randomly, so assembly methods have been employed with good results [36]. The first assembly methods were based on invasions, where the food web structure is built up over time in response to new species being introduced one by one. Simple rules are employed to introduce new species. In the cascade model, each species is assigned a number. A new species can only potentially prey on species with a lesser number, and be preyed upon by species with a greater number. In the niche model, new species can only potentially prey on species within a range slightly above and below its number. Besides these constraints, interaction strengths are assigned randomly. These models show some agreement with the structure of real food webs.

Evolutionary methods of food web assembly have received increasing attention. Evolutionary change in species within an ecosystem no doubt have an effect on the developing structure of the community. The extent of this influence is disputed, with some seeing invasions as a more significant force. Much of the research in evolution in ecosystems has been lead by physicists, with resistance from ecologists, which is reminiscent of Robert May's early research. The development of evolutionary models has reiterated research trends, with the first models involving linear analysis of evolving matrices [33], later incorporating population dynamics [20] [34], and even population age distributions [35]. By far the most prolific and developed program investigating food webs under evolution is Webworld [11]. This is the model employed here, and is described in detail in the next section.

It's important to note, as more complexity is added to the models themselves, detailed analysis becomes difficult and intractable. Statistical and network methods have been employed, and the results compared to natural food webs. Network analysis has placed food webs outside the classification of "small world" and "power law" networks, leaving open the prospect of a hidden pattern, or less hopefully no consistency among food webs [13] [37] [17]. A major difficulty is the lack of large numbers of high quality empirical food webs, as food webs have been produced with varying methodologies and rigor [9]. No food web incorporates the full population dynamics or interaction strengths, though theoretical results are driving empirical research in this area[5].

2.3 The Webworld Model

The Webworld model has been jointly developed by Caldarelli, Drossel, Higgs, McKane and Quince, during several years of research aimed at combining different approaches to ecosystem and food web study [10]. Theoretical ecology has mostly passed over the role of evolution in assembling food web structure, while statistical physics has ignored the role of population dynamics and adaptive foraging. By combining ecological and evolutionary time scales, Webworld has been used to examine questions on the stability and network structure of food webs, as well as long term evolutionary trajectories and extinction events in an ecological context [30] [31]. The resulting model is robust, and in particular, the predator response equation has been shown to enable construction of biological realistic food web structures [12]. While portions of the model have been criticized as too complex [14], the evolutionary process is much more elegant than the artificially restrictive assembly models based on cascades and niches.

In Webworld, a species is defined as a set of morphological or behavioral characteristics, taken from a set of all possible characteristics. The interaction of two species' sets of characteristics determines the strength and direction of their interaction in the food web. These characteristics can be thought of as features like "sharp claws", "camouflage body pattern", "fruit bearing", etc., but are simply represented in the abstract as integers. Speciation occurs by randomly taking one individual from a species, and modifying this feature list to create a new species. A species is represented as L features, from K possible features. In most of the published work, $L = 10$ and $K = 500$.

When starting a run, a KxK feature matrix of normally distributed random numbers is calculated. The matrix is antisymmetric, so $m_{\alpha\beta} = -m_{\beta\alpha}$. The entry $m_{\alpha\beta}$ describes how feature α performs against feature β. The score S_{ij} of species i against species j is calculated as

$$S_{ij} = max\{0, \frac{1}{L} \sum_{\alpha \in i} \sum_{\beta \in j} m_{\alpha\beta}\} \tag{1}$$

where α and β iterate over all features of species i and j. Essentially, the score is the sum of the performance values of each species' features against the other, and because the feature matrix is antisymmetric, one species will have a positive score, meaning it is the predator, while the other will have 0, meaning the prey. This score is later used in calculating the population dynamics.

The population dynamics are more accurately described as tracing the flow of energy resources through the community; for simplicity, population size and energy resources of a species are equivalent in the model. External resources are input to the system at a constant rate, R, and distributed to species depending on the scores S_{ij}. The external environment is represented as species 0, which is assigned a set of unchanging features. Species with a positive interaction score against the environment are primary producers. The value of R has been systematically tested to examine the effect on the resulting food webs [10], and there is good correspondence to the situation in natural ecosystems, where more resources result in larger, more diverse communities [28].

Most generally, the population of a species, $N_i(t)$, is updated according to this equation.

$$\frac{dN_i(t)}{dt} = -N_i(t) + \lambda \sum_j N_i g_{ij}(t) - \sum_j N_j g_{ji}(t) \tag{2}$$

The first term is the constant death rate, independent of interaction with other species. For simplicity, the death rate is equal for all species and set to unity; this essentially sets the time scale for the population dynamics. The second term represents the amount of resources transfered to species i, from predation on other species. $g_{ij}(t)$ is the 'functional response' equation, or the rate at which species i consumes species j. λ is the ecological efficiency, or the percentage of resources transferred for one species to another in consumption. Here, $\lambda = 0.1$, a value well supported by experimental evidence [28]. The final term represents predation by other species on species i.

g_{ij} is based on ratio-dependent functional response [1], for a single species

$$g_{ij}(t) = \frac{S_{ij} N_j(t)}{b N_j(t) + S_{ij} N_i(t)} \tag{3}$$

extended to cover multiple predators and multiple prey. It is called ratio-dependent, because both the size of the predator and prey populations are accounted for; when prey are relatively abundant, predation saturates at a level determined by b, and when predators are relatively abundant, predation levels are dependent on prey population levels.

With multiple predators on a single prey, the predator term in the denominator is modified to account for this competition.

$$g_{ij}(t) = \frac{S_{ij} N_j(t)}{b N_j(t) + \sum_k \alpha_{ki} S_{kj} N_k(t)} \tag{4}$$

The sum runs over all predators k which prey on species j. The new factor α_{ki} represents the fact that competition between the members of the same species is stronger than between individuals of different species. For members of the same species, α_{ki} is 1, and for different species < 1, depending on the degree of similarity, q_{ki}, between the species. q_{ki} is calculated as the percentage of features the species have in common. Also, it is assumed that there is some minimal level of competition, c, between any species. So, the equation for measuring competition is

$$\alpha_{ki} = c + (1 - c) q_{ki} \tag{5}$$

where c is a constant between 0 and 1. It has been shown that increasing the minimum competition decreases food web diversity, and vice versa [10].

When a predator feeds upon multiple prey, it is assumed that the predator must divide the total of its effort among those prey. In other words, it can not prey on each of these species separately as if they existed in isolation. Further, a predator will choose to divide its efforts according to which prey will afford the maximum transfer of resources; this is known as the Optimal Foraging Theory [28]. The efforts of predation on each prey, f_{ij}, must sum to 1. To maximize the return of efforts, it has been shown that

$$f_{ij}(t) = \frac{g_{ij}(t)}{\sum_k g_{ik}(t)} \tag{6}$$

is an Evolutionary Stable Strategy (ESS) [10]. That is, there is no other choice of efforts that can invade a population of predators with this strategy. Also note, f_{ij} must have some minimum value greater than zero (usually .00001), since an effort of zero could never increase within these iterative equations.

Taking efforts into account, the final form of the response function is

$$g_{ij}(t) = \frac{S_{ij} f_{ij}(t) N_j(t)}{bN_j(t) + \sum_k \alpha_{ki} S_{kj} f_{kj}(t) N_k(t)}. \tag{7}$$

The response function and efforts calculation are interdependent. In order to arrive at a value to update population levels, the two equations are iteratively calculated for all species, until they converge on the ESS. So, initially, all efforts are set at $1/number\ of\ prey$. These efforts are used to calculate g_{ij} for every predator and prey. This leads to new values for the efforts, which are then fed back to the response equation. In practice, this iteration continues until all changes are less than 10%.

With efforts at the ESS, population changes are calculated. If this leads to any extinctions, that species is removed from the web. With population changes, the ESS for predation efforts will also change, so the process is iterated. Populations will eventually converge to fixed values. At that point, a species is chosen for mutation and added into the web, and the entire process is repeated.

3 Methods

3.1 Running Webworld

As stated previously, Webworld begins by calculating the feature matrix, and choosing the value of the parameters. b, which controls maximum predator response in the response function, is set to the typical value of 0.005. c, the minimum competition between two species, was set to 0.5. R, the amount of resources input to the system, was set to $10,000$. Initially, 10 species are created, with randomly chosen features, and each with a population size of 10. The interaction scores for all species is calculated.

The model proceeds in three time scales: foraging, population change, and evolutionary dynamics. At each step, the functional responses and division of efforts are iteratively calculated according to the present population levels. Once near the ESS, the populations are then updated, using a discrete version of the dynamics.

$$N_i(t+\Delta t) = N_i(t)(1-\Delta t) + \Delta t[\lambda \sum_j N_i g_{ij}(t) - \sum_j N_j g_{ji}(t)] \tag{8}$$

The time-step $\Delta t = 0.2$ is large, but sufficient for calculating the stationary values of the populations identical to the results of the continuous equation. Responses and population levels are recalculated until reaching a stability, which in practice means no changes in population greater than 10%. If the population of a species is < 1, it is considered extinct, and removed from the food web.

Once the food web is stable, a speciation event is triggered. One species is randomly chosen to undergo speciation. A single individual is removed from that species, and one of that individual's features is randomly changed, becoming a new species. A check is performed to make sure the new species isn't already present in the web; if so, speciation is attempted again. The scores of this new species against all predators and prey are calculated, and the process of calculating efforts and population changes is started again.

There are a few possibilities for the web, with the introduction of this new species. It can immediately go extinct, under competition from its parent species. Sometimes, it can co-exist with its parent and its competitors, causing no subsequent extinctions. Other times, the new species out competes its parent or other competitors, over consumes prey, or causes even more indirect effects, leading a cascade of extinctions. Webworld is run for some number of speciation events, from 1,000 to 100,000, and the resulting webs are used for statistical analysis.

For clarity, the simulation process described here is illustrated by the flow chart in Figure 1, and the process of calculating the predation score, and the potential result of a speciation, is elaborated with an example. Consider two species, each with five features (rather than ten, for ease of illustration). The matrix of values for their feature interactions is shown in Table 1, with species 1 along the columns and species 2 along the rows. The values sum to 0.6538, meaning that species 1 is a potential predator on species 2; whether there is actual predation depends on the assignment of efforts in the ESS calculation.

Now consider the situation of a speciation, where species 1 is mutated into species 3. The values in the last column are different depending on the new feature, as in Table 2. The values now sum

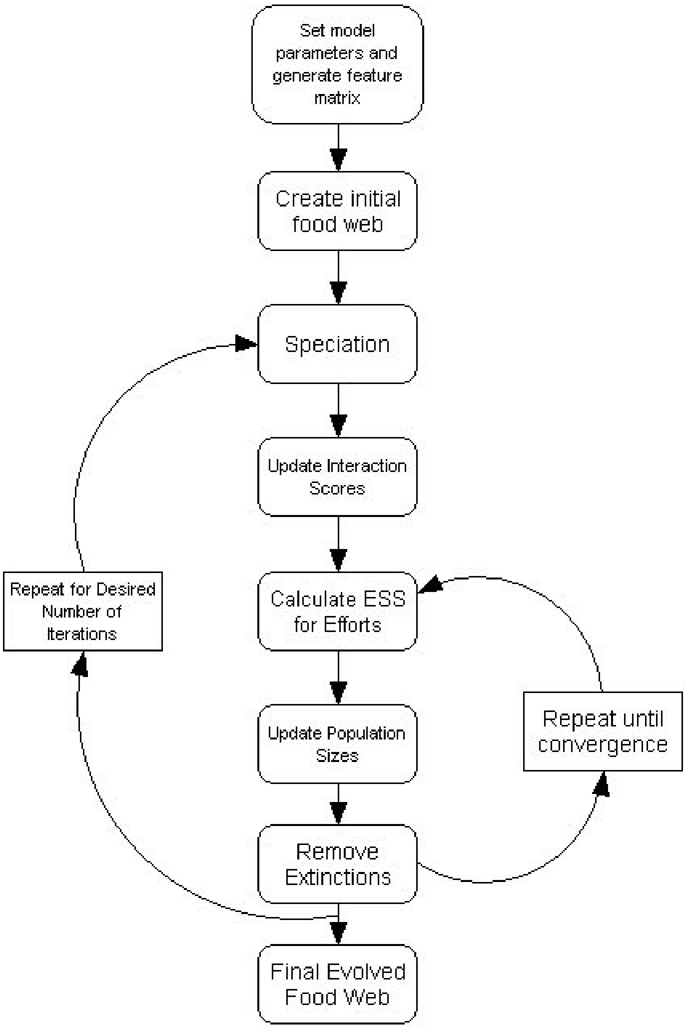

Figure 1: Flow chart illustrating the simulation process

Table 1: Feature interaction matrix of two example species

	33	59	172	399	402
12	0.4800	0.1949	0.9640	-1.0384	0.4889
73	1.0455	-1.4359	0.0501	-1.0833	0.5773
243	1.0470	0.0265	0.7417	-0.1366	-1.5233
409	-0.6576	0.9418	-0.4310	0.4283	-1.4277
411	0.2055	-0.7503	0.1001	1.1549	0.6913

Table 2: Feature interaction matrix after a speciation. Species 2 now preys on the descendent of Species 1.

	33	59	172	399	139
12	0.4800	0.1949	0.9640	-1.0384	-0.3967
73	1.0455	-1.4359	0.0501	-1.0833	-1.7239
243	1.0470	0.0265	0.7417	-0.1366	0.4392
409	-0.6576	0.9418	-0.4310	0.4283	-2.6573
411	0.2055	-0.7503	0.1001	1.1549	0.1238

to -2.3676, and because the feature matrix is antisymmetric, species 2 will have a score of 2.3676 against species 3, and will be a potential predator. Note, that most mutations won't cause a change in relationship, or even much change in the strength of interaction. Those that do cause significant changes modify the evolving structure of the food web as a whole.

3.2 Extending Feature Complexity in Webworld

Within the Webworld model, species' features, and score calculations from the feature matrix have received the least critical examination. Two species feature sets define a submatrix of normally distributed random numbers. By the Central Limit Theorem [27], the sum of any number of normally distributed independent random numbers will itself be normally distributed. So, typically, any species will have a normal distribution of interactions with other species: most will be quite insignificant, a few will specify predator and a few will specify prey.

The significance of the feature matrix can be tested by removing it altogether. Mutation occurs directly on the interaction scores. The most significant problem is that interactions are now independent of each other, and no longer normally distributed. This quickly leads to species with few predators and very high scores against prey. With unbounded increases, the web quickly collapses to total extinction. The feature matrix seems to be the simplest way to insure interdependence of interactions.

In this thesis, speciation events were extended for two possibilities. In addition to randomly changing one feature as before, a feature could be added or deleted from the species. This introduces two new parameters to the model: the probability of adding and the probability of deleting a feature during speciation. These probabilities, plus the remaining probability of only changing a feature, sum to one. The initial species still start with a feature list of size 10.

Differing feature list sizes leads to two additional changes. The number of feature performance values summed for the predation score is now variable. However, all scores for a species will still be normally distributed. Second, when calculating the competition of two species, the percentage of features in common will be in respect to either the shorter or longer feature list size. So, if a species with size 10 feature list shares 10 features in common with a size 15 feature list species, they will have 100% similarity with respect to the shorter list, and 67% similarity with respect to the longer list. In most of the experiments here, calculation was with the shorter list; however it's later shown that this choice has some significance.

How does varying the feature list length correspond to complexity? Since the species representation is not a genome, as the list directly interacts with other species, the problem of lack of correspondence between real genomes and species complexity does not apply. Within the bounds of the model, feature length does correspond to complexity.

To measure the strength of any observed effects, the ratio of probability of feature additions to feature deletions was varied from 0 to 1. There was always a mutation pressure towards a smaller genome. The probability of additions and deletions always summed to 40%. If there were no evolutionary advantage to increased complexity within the food web, the average feature list size would decrease exponentially at different rates depending on these ratios. However, if there were some significance to complexity, this decrease would be slowed or even reversed. Runs from each of 30 ratios in the range were performed 6 times, to account for statistical irregularities. The runs progressed for 2,000 speciation events, a conservative amount in light of the later results.

The final evolved webs from each run were used for analysis. The webs were reloaded into objects, from the logs, and various metrics performed to examine any trends or deviations from the regular Webworld model. Metrics initially included the average genome size, number of species, connectance, characteristic path length, clustering coefficient, and stability to extinction. For the network metrics, edges are treated as bidirectional and are considered present only when there is a transfer of more than one unit of resources. Connectance is the number of edges present, divided by the number of possible edges (the number of possible edges is $\frac{n*(n-1)}{2}$, where n is the number of nodes). The characteristic path length is the average distance between any two nodes in the graph. The clustering coefficient is the percentage of nodes, connected by an edge, which both have edges to a common third node. The stability to extinction measures the average percentage drop in species number, in response to systematic single removal of each node from the network [31].

3.3 Coding Practice

Webworld was implemented in C++, for execution efficiency and object oriented features. The core template class, **Graph**, implements a generic, directional graph with a list of **Node** and **Edge** objects. This representation was chosen, as its much simpler to add and delete nodes than in a matrix. Conversion to matrix form is possible, and necessary for some of the analysis methods in the class, such as the characteristic path length and connectance.

The **Webworld** class is a child of **Graph**, and implements the foraging, population, and evolutionary dynamics of the model, constructs the feature matrix, logs results during a run, and provides an interface for setting the parameters of the original and extended webworld. The **WNode** class, child of **Node**, stores the population size and feature list for a species, and the **WEdge** class, child of **Edge**, stores the response function and effort of the relationship it represents.

This implementation of Webworld was closely examined against the described model, and tested for correctness by comparing the results of runs against published results. The results corresponded very well, and to the best of my knowledge the implementation is correct.

Additionally, there's **wwsimple**, which runs a version of Webworld without features. **SGraph** runs simpler Lotka-Volterra type food webs, and **SGraphConf** configures experiments in these food webs through a simple XML protocol. There are further utility classes, for handling matrices, regular expressions, and random numbers.

4 Results

4.1 Testing the Webworld Model

To satisfy that this program correctly implements the regular Webworld model, several test runs were performed and the results compared to published results [10]. The parameters, as in all runs in this thesis, $b = 0.005$, $c = 0.5$, and $R = 10,000$. For these tests, runs were concluded after 10,000 speciations, as in the Webworld papers. These test runs corresponded well to those reported in the Webworld papers; the results from one test run are here presented in detail. Figure 2 displays the change in number of species over the run. Table 3 compares various network properties of this sample web. Figure 3 displays the size distribution of extinction events, following speciation. Values for these metrics are within published ranges of deviation and overall there is good correspondence.

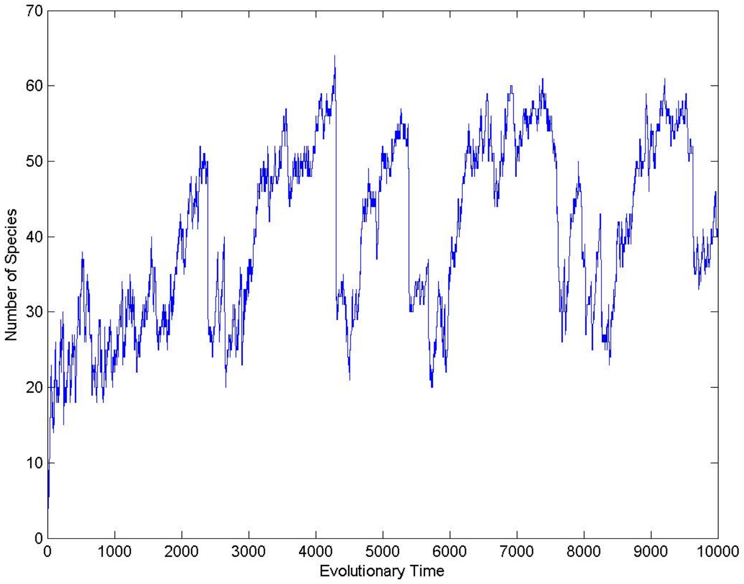

Figure 2: Number of species as a function of time, in the test run

Table 3: Network metrics of the test run, compared against results published in [10]

	Test Run	Published Results
No. of species	41	33
Links per species	2.2	1.76
Av. level	1.75	1.95
Av. max. level	3.0	3.0
Basal species (%)	24	18
Intermediate species (%)	72	80
Top species (%)	4	2

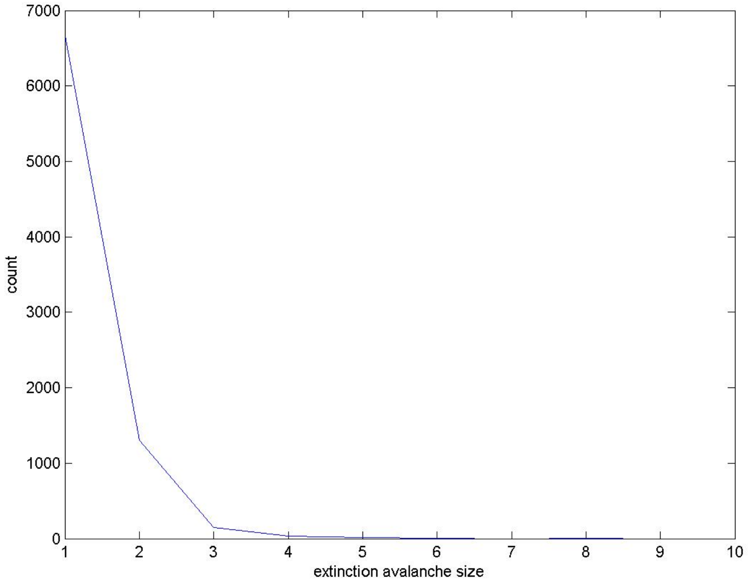

Figure 3: Size distribution of extinction events, in the test run

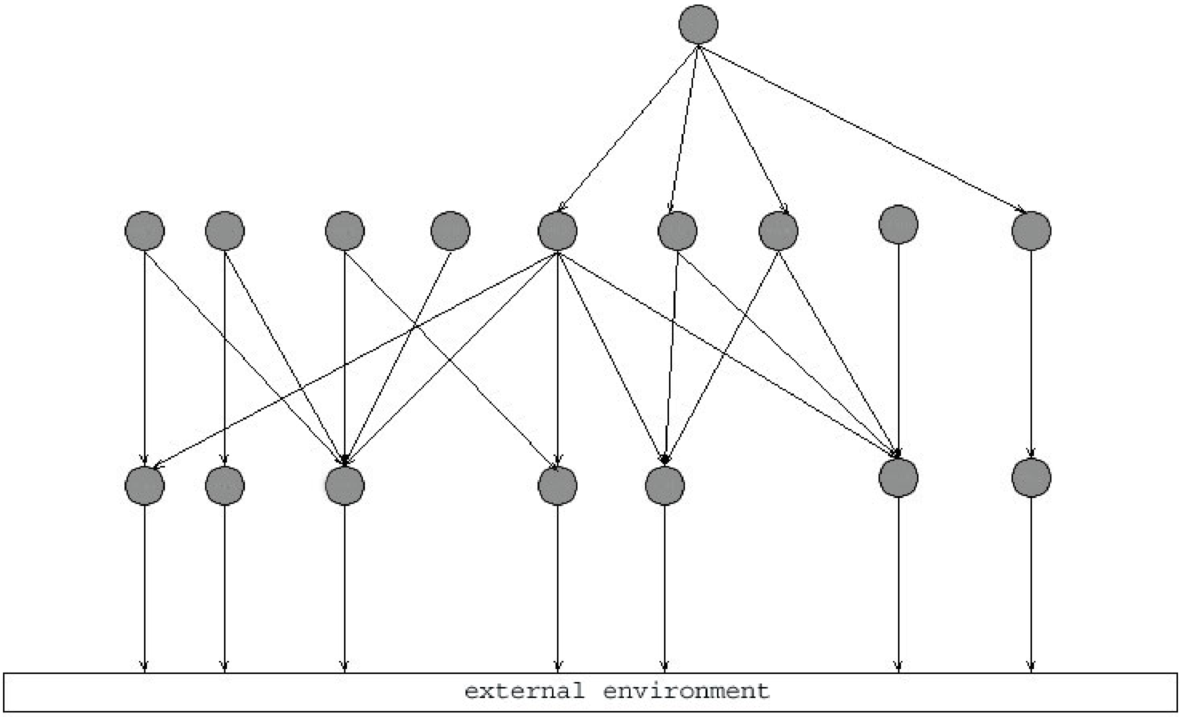

Figure 4: An example food web produced by Webworld

4.2 Response to Length Mutation

The main result to examine is how the average genome size responded to mutability. In the absence of any advantage for a longer genome, any probability pressure towards deletions would tend to push the average size towards minimum values. However, there was a clear response towards increasing genome size, a shown if Figure 5. All changes per ratio are roughly linear, with ratios above 1 : 2 resulting in increase, and below 1 : 2, decreases. Even when genome size decreased below 10, the decrease is less than would be expected. Figure 6 illustrates statistical outliers and standard deviation for the genome size results of all six runs for each ratio. Box plots were constructed for all metrics, but don't add any relevant information.

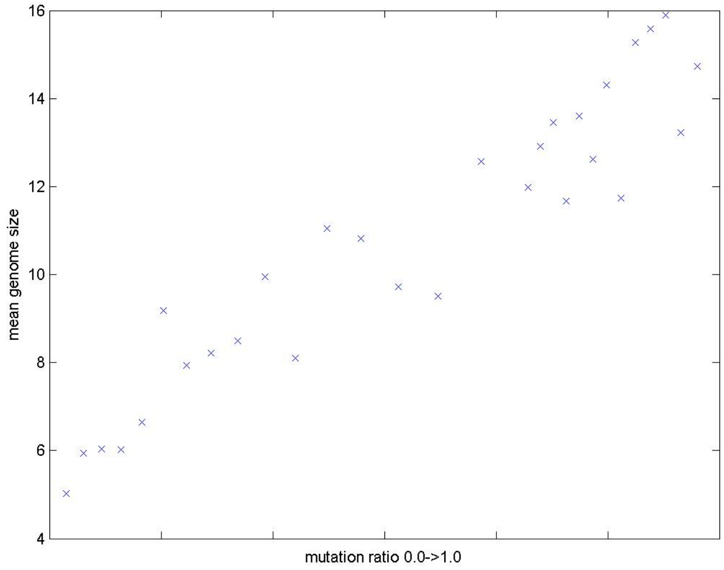

Figure 5: Mean genome size in response to variations in addition/deletion mutation ratio

Next to examine is how various network metrics on these evolved webs, standard in food web analysis, compared to the regular Webworld model and real food webs. The number of species, shown in Figure 7, also displays a clear linear response to mutation ratio. In published results [10], under the same parameters, the average number of species was 33. Again, near the 1 : 2 ratio the result matches, while there's just slight increase or decrease relative to change in ratio along the range.

Stability analysis is performed by duplicating the evolved web and counting subsequent extinctions after deleting a single species, each in turn. The metric here is the average percentage of species remaining after each species deletion test. Essentially, this tests how sensitively dependent the food web structure is on individual species. As shown in Figure 8, there was no relationship between stability and mutation ratio. Further, there is good correspondence with the regular Webworld model, where on the whole the food webs were very stable [31].

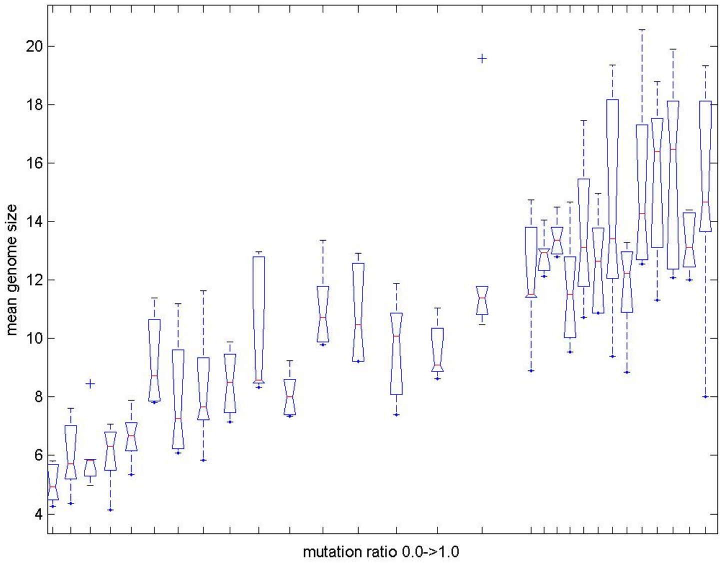

Figure 6: Box plot of genome size in response to variations in addition/deletion mutation ratio

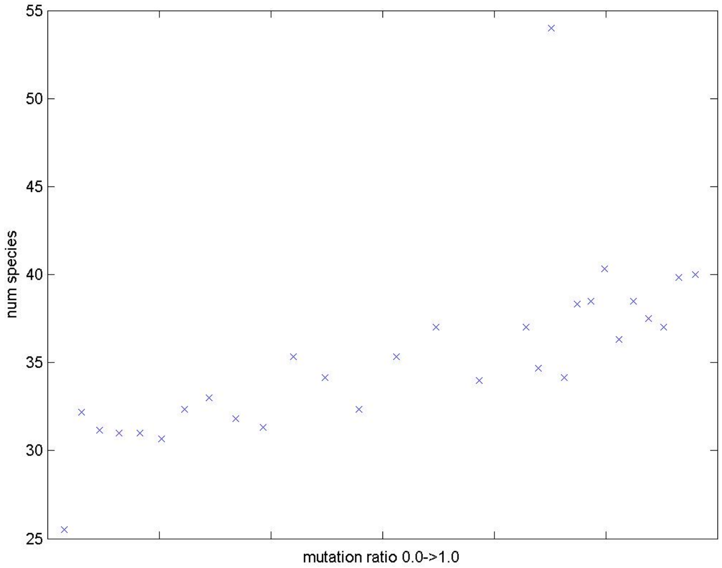

Figure 7: Mean number of species in response to variations in addition/deletion mutation ratio

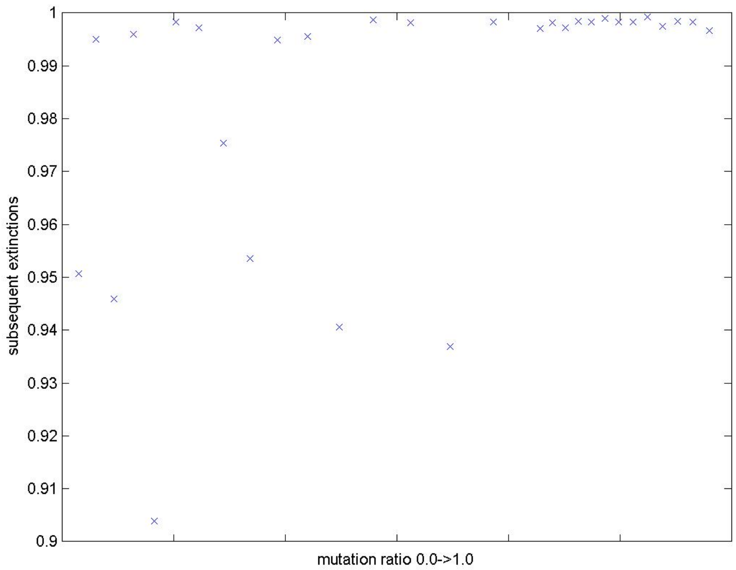

Figure 8: Mean stability in response to variations in addition/deletion mutation ratio

In network studies on real food webs [13], the metrics of connectance, characteristic path length, and clustering didn't show any particular constancy and there was no typical network structure, but these values did tend to fall in a particular range. Connectance varied from 0.026 to 0.315, characteristic path length varied from 1.33 to 3.74, and clustering varied from 0.02 to 0.43. Metrics for the webs evolved here all fell within the middle of these ranges. Both connectance and clustering, shown in Figures 9 and 10, tend to decrease as the mutation ratio is increased. Characteristic path length, in Figure 11, showed no relationship to the mutation ratio.

Considering the trend in connectance and clustering, and direct inspection of sample evolved webs, there seemed to be a tendency for increase in disconnected nodes. These species would be basal species, but without any active predators. This metric was added, and the result in Figure 12 demonstrates that there is in fact a correlation between increased disconnectance and increased mutation ratio. An increase in disconnected nodes would account for the decrease in connectance and clustering.

From these results, there is a clear tendency for increased species complexity in food webs evolved by Webworld. There's also a tendency for increased disconnectance; are these results connected? Can the result be separated from the ecological context, or from species evolution? These questions lead to further experiments, described in the next section.

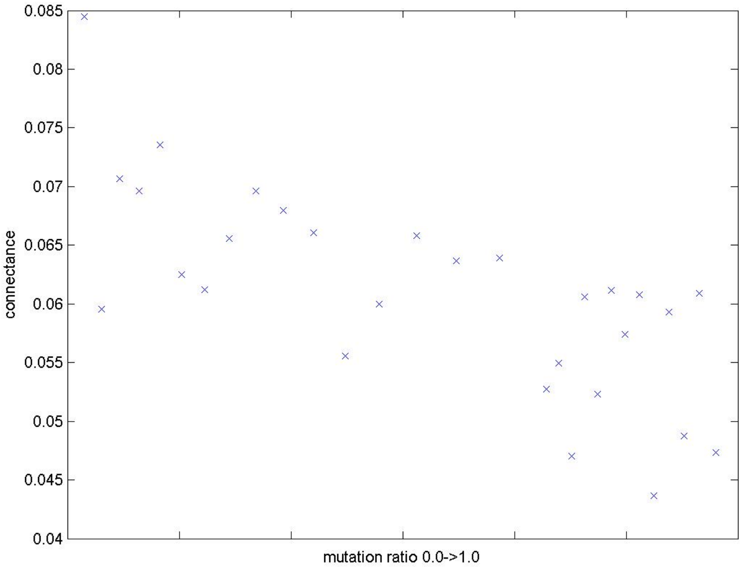

Figure 9: Mean connectance in response to variations in addition/deletion mutation ratio

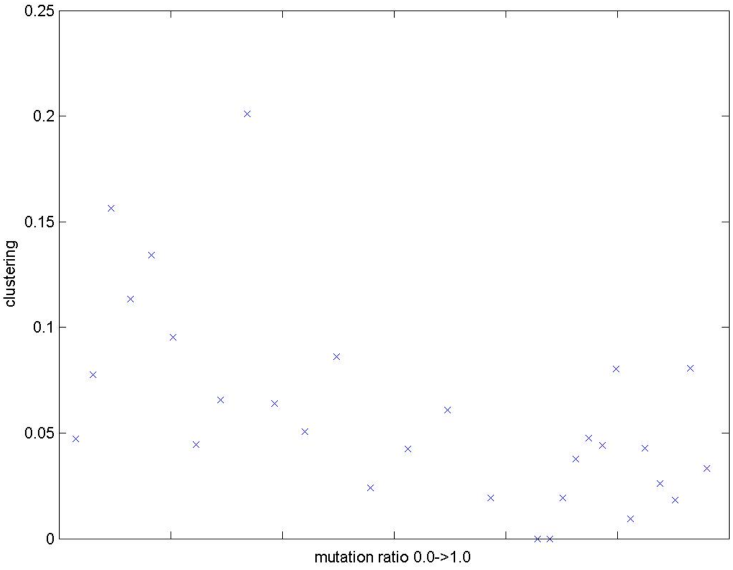

Figure 10: Mean clustering coefficient in response to variations in addition/deletion mutation ratio

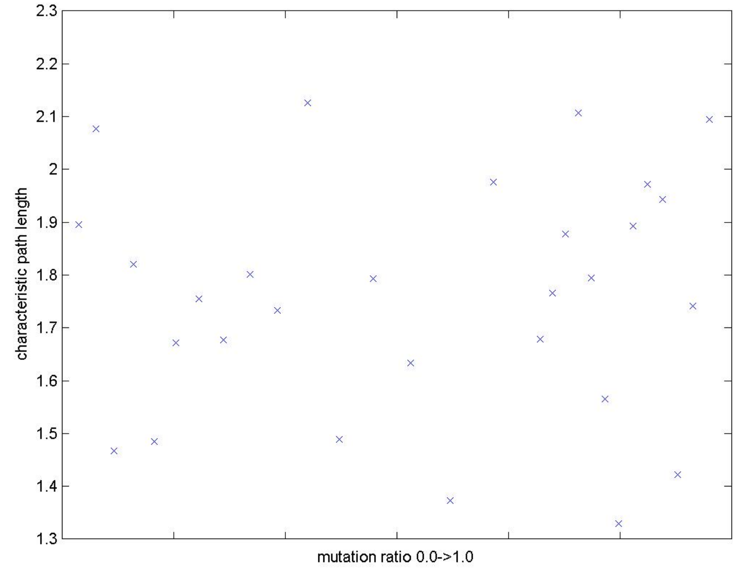

Figure 11: Mean characteristic path length in response to variations in addition/deletion mutation ratio

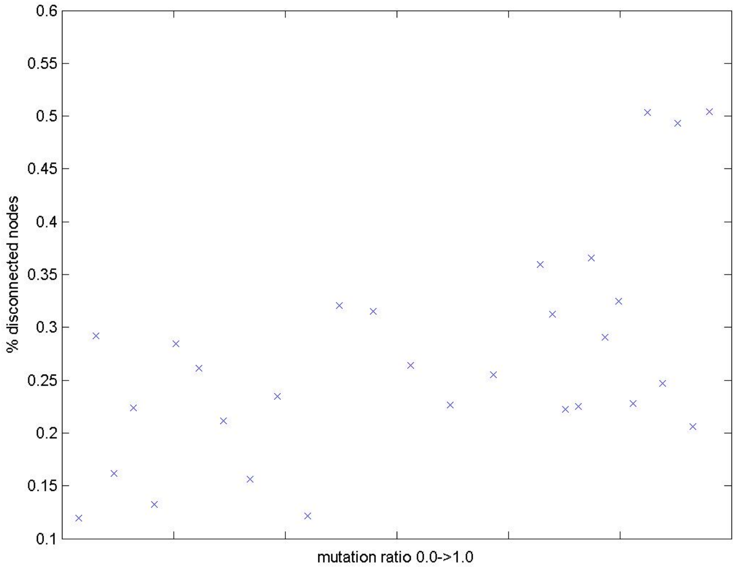

Figure 12: Mean percentage of disconnected nodes in response to variations in addition/deletion mutation ratio

4.3 Accounting for Increased Complexity

The following experiments have been designed to test various hypothesis for the observed increase in species complexity. Rather than exhaustively test the entire range of mutation ratios, only the 19% addition to 21% deletion ratio was used, in six runs each.

First, a precaution. The 2,000 mutation runs were conservative, the assumption being that any trend observed would continue in the same way. However, it is conceivable that the increases are temporary, and later in the run as the web stabilized further, the trend would reverse. So, a set of simulations were allowed to continue for 10,000 mutation events. In fact, the genome size continued to increase, to an average of 19.6, from an average of 15.8 in the eariler runs. Figure 13 plots the genome size change against time, and Figure 14 plots the number of species over time. The conservative approach is justified.

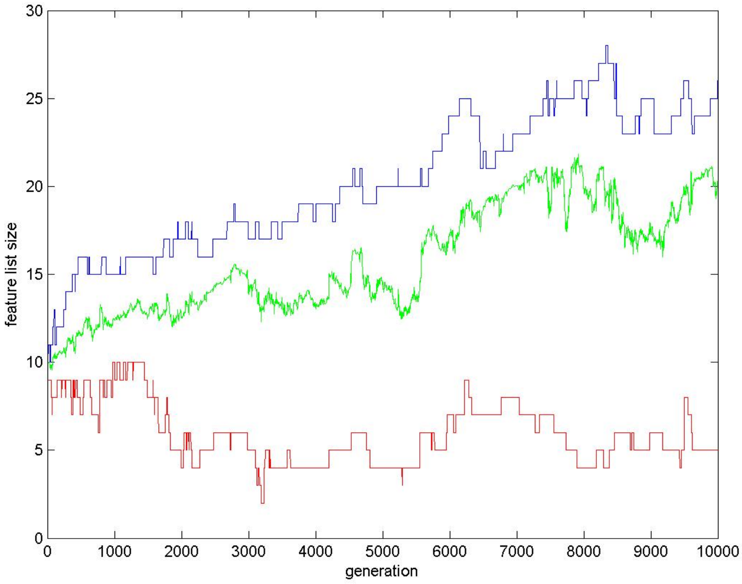

Figure 13: Average, largest, and smallest genome size plotted against time, for 10,000 mutation events

One possibility is that species complexity increased with no regard to its meaning to the adaptability of the species in the food web. In these runs, the calculation of the interaction score was modified to use only the first 5 features of each species, and the rest were ignored. Thus, there was absolutely no adaptive pressure for increase in the ecological context, and any affect species complexity had on the food web structure was removed. Any increase would be evidence for a sort of “genetic drift” in speciation (differing from the usual definition of genetic drift, within a single heterogeneous population). Relative to previous runs with this mutation ratio and use of the full feature list, the results were a consistent decrease in genome size, to an average of 6.3, as well as a decrease in the number of species, to an average of 30. Genetic drift could not account for the increase.

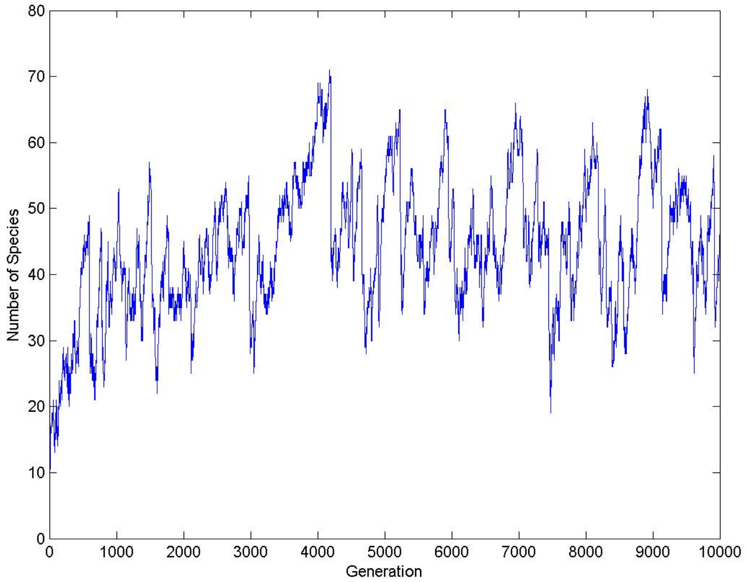

Figure 14: Number of species as a function of time, for 10,000 mutations, with length mutation

In contrast, could complexity increase without the consistent structure of the food web? How vital is interaction in developing complexity? In these runs, three levels of Gaussian noise was added to each edge in the food web. Noise with standard deviation of 1 and .5 resulted in complete collapse of the web. Noise with standard deviation of .25 allowed webs to persist and evolve. Complexity did increase, but at a much slower rate than before. The average genome size was 11.9, and number of species was 34.4. A stable web is necessary for increased complexity, but doesn't sufficiently explain it.

Is the increase in disconnectance related to the increase in complexity? To examine this question, in six more runs, after a mutation the new species was examined to see if it was disconnected; in other words, basal with no active predator. If it was a disconnected species, it was discarded and another mutation performed, until a predator or a basal species with a predator developed. The resulting webs had little change in species complexity from the initial size of 10, to a final average feature list size of 10.7. Examining a few samples from previous runs, disconnected species were usually the first to evolve longer features lists and persist. These species and their ancestors came to dominate, including ancestors that evolved to be predators.

Clearly, the complexity increase was driven by new disconnected species. Why this should be is not so clear. A possibility was that one of the modifications, that is how the similarity between two species is calculated, was freeing some species from predation. Throughout all the experiments, similarity was calculated with respect to the shorter feature list. Any increase in a species feature list would result in no change or increase to similarity measurements against shorter species. However, if similarity was calculated with respect to the longer feature list, an increase would result in no change or a decrease to the similarity.

Table 4: Summary of results from second set of experiments.

	Mean Feature List Size	Number of Species
2000 mutations	15.8	39.5
10000 mutations	19.6	42.3
limited feature list	6.3	30
noisy interaction	11.9	34.4
remove disconnected species	10.7	31.2
long list similarity	14.2	35.3

This was tested in a further six runs, by employing the longer list method. The result was mixed. There was still an increase in feature list size, to an average of 14.2, and still a high percentage of disconnected species, to an average of 21%, but the effect was diminished from previous runs. Decreasing competition made things harder, but not impossible, for disconnected species.

5 Discussion

The results show that within the interactions of the model food web, species complexity does evolve to increase. This increase can not be attributed to either the ecosystem or evolution of a single lineage alone, but rather the interaction and feedback of these processes. The ecosystems developed are in some sense more diverse than food webs without variable complexity, as they contain greater number of species, yet there is a different structure and less interactions. There is a relationship between the evolution of species complexity and community interaction, but in a somewhat different fashion than anticipated.

Specifically, increasing complexity encouraged basal species without predators. With the obvious advantages of no predation, these species led the increase in complexity for the rest of the web. Published results on Webworld found that species overturn is higher on higher trophic levels [10]. As basal species have the largest population size, they are less likely to go extinct and more likely to persist after speciation.

However, no rigorous explanation was found for why a longer feature list length results in greater numbers of these species. It was found that the method of calculating competition influenced the degree of this effect, but didn't wholly account for it. In the second round of experiments, competition was calculated with respect to the longer feature list, and on average this reduced interspecies competition. The principle of competive exclusion, based on the Lotka-Volterra model, states that the stable coexistence of two species is only possible where intraspecific competition was greater that interspecific competition [32]. With less interspecies competition, established species had a better opportunity to survive the introduction of new species with varying feature list sizes, effectively slowing the rate of complexity increase.

For real food webs, disconnected species are akin to species that manage to step out of the present ecological context into new niches. The movement of animals from sea to land and the emergence of social insects are dramatic examples that seem to support this. The results suggest this process is at work more generally in stable ecosystems, rather than just in the punctuated equilibria of these examples. Rather than being a process of mutual encouragement, there is a positive tension between the evolution of species complexity and ecosystem stability. Ecosystems must respond to these disconnected species by incorporating them within their structure, or face collapse and monoculture. The collective adaptability of species within an ecosystem to interact and stabilize preserves diversity. Once incorporated and stabilized, this sets the stage for another increase in complexity, by exploitation of some new niche. In a sense, disconnection is an advantage to short term fitness, while reincorporation is necessary for long term fitness.

However, these thoughts do step beyond particular aspects of the model. Certainly, evolution of complexity in real organisms has not been led by basal species, which then later evolve to become predators. The split of species lineages into the major kingdoms occurred in a remote time and most likely once. The process of speciation in Webworld may be too loose. The authors of Webworld admit as much, as primarily real speciation involves loss of reproductive ability between subgroups of species, due to morphological, behavioral, or spatial factors; it's not just the swapping of features. Speciation with some restraints could address this problem.

The lack of a developmental process in Webworld is related. One reason that small steps in evolution do not lead to huge leaps in function is that mutations occur directly on the genotype, while selection occurs on the phenotype. The constraints of the developmental process limits the rate of change in a species lineage. In Webworld, mutation occurs directly on the phenotype. Can species complexity be effectively modeled without some form of development? The challenge would be to find a representation with the elegance of the Webworld feature matrix, without being overly specific to the terribly intricate realities of biological development.

Disconnected species also represent a problem to the representation of the entire food web. Removing disconnected species in the follow up experiment seems justified, as they are not truly in the food web. The problem is that food webs, as isolated networks of energy flows, are themselves a convenience of study for ecology. The most complete studies of trophic interactions have been on lake ecosystems, since the lake provides a natural boundary in which to collect species and examine feeding. Yet, many interactions cross the boundary of the lake surface: some fish feed on insects at and near surface, birds and mammals live primarily on land but feed on fish, and naturally amphibians cross this boundary. In reality, no food web exists in isolation; they are more like dense subgraphs of a planet wide network. In this light, species with little connection to a single food web, may represent these species that crossover denser parts of the network. Could these species, which must adapt to multiple subwebs, be a source increased complexity?

To summarize, the model found that increased species complexity evolves within ecosystems. No definitive answer explains this effect. The way Webworld responded does suggest that organism development and the interaction 'between' ecosystems may have some role.

6 Conclusions

This work set out to explore the relationship between ecosystems and the evolution of organism complexity. The role of community interactions in evolution is considered limited by some, but it has not been investigated empirically or in models. Though complexity in biological systems is still being actively defined, the Webworld model provided a straightforward measure of complexity within an evolving food web. Webworld was extended so that organism complexity was also subject to mutation, and the resulting webs analysed by various network metrics.

The results demonstrated a clear tendency for increasing organism complexity. At the same time, the number of species disconnected from the food web increased. Further tests showed that interaction between the ecological context and species development were both required for complexification, and that the complexity increase was largely led by species temporarily disconnected from trophic relationships. This suggests a process of repeated seperation and reincorporation of new species into the ecosystem, leading to increased organism complexity.

From here, the results point in many directions. Primarily, if complexity was shown to evolve in another food web model, then the results here would be greatly strengthened. Also, Webworld should be studied with static feature list sizes, in systematically varied lengths, and the resulting networks compared to the results here.

The disconnected species shed light on the arbitrary boundary of food webs, suggesting one approach. Rather than a single Webworld, multiple webs would be evolved concurrently. When disconnected species arise, they are set to join one of the other foods webs, as an invader species. Evolution and invasion assembly, in combination, is worthy of testing in both the regular and extended Webworld model.

Considering the relationship of development and complexity, a simple model of development would be useful. One possibility may be to represent species as networks themselves, in some way analogous to gene regulatory networks. For this new model, predation scores are dependent on some non-linear interaction of the two species' networks. Defining these networks and the environment they operate in, to satisfy the response functions of Webworld, could be a fruitful line of research. Though, perhaps it's enough for the model to recognize that speciation does not involve the tremendous jumps in function currently possible. It would be straightforward to incorporate some minimal restrictions, both on the regular and extended Webworld model, so that from an initial stable web, the lineages of flora and fauna remain separate.

Current food web research is incorporating more ecological factors [29]; for example, age distributions and heterogeneous populations, spatial effects, non-trophic interactions, sex, external factors like seasons and weather, bacteria and decomposers. Any of these could conceivably have an effect on complexity. Seasonal change would be a straightforward modification to Webworld, by varying the resource parameter during the course of the run.

Whichever the approach, continuing research on the systematic interactions of ecosystems, evolution, and complexity is of urgent importance, as the Earth's evolved and complex ecological networks face the threat of unraveling.

References

[1] Abrams, P.A. and Ginzburg, L.R. (2000). The nature of predation: prey dependent, ratio-dependent or neither? *Trends in Ecology and Evolution* 15:337-341.

[2] Axelrod, Robert and Hamilton, William (1981). The Evolution of Cooperation. *Science* 211: 1390-96. AAAS.

[3] Bennett, K. D. (1997). Evolution and Ecology, The Pace of Life. Cambridge University Press.

[4] Bock, Gregory and Goode, Jamie (eds) (2000). Complexity in Biological Information Processing. John Wiley & Sons, Ltd.

[5] Borer, E.T., Anderson, K., Blanchette, C.A., Broitman, B., Cooper, S.D., Halpern, B.S., Seabloom, E.W., Shurin, J.B. (2002). Topological approaches to food web analyses: a few modifications may improve our insights. *Oikos* 99:2.

[6] Bonner, John Tyler (1988). Evolution of Complexity, by Means of Natural Selection. Princeton University Press.

[7] Carroll, Sean B. (2001). Chance and necessity: the evolution of morphological complexity and diversity. *Nature* 409:1102-9. Macmillan Magazines, Ltd.

[8] Chauvet, E., Paullet, J., Previtte, J., Walls Z. (2002). A Lotka-Volterra Three-species Food Chain. 75 No. 4. Mathematics Magazine.

[9] Cohen, J.E., Brian, F., Newman, C.M. (1990). Community Food Webs. *Biomathematics* 20. Springer-Verlag.

[10] Drossel, B., Higgs, P., McKane, A.J. (2001). The Influence of Predator-Prey Population Dynamics on the Long-term Evolution of Food Web Structure. *Journal of Theoretical Biology* 208, 91-107. Elsevier Ltd.

[11] Drossel, B., McKane, A.J. (2003). Modelling Food Webs. *Handbook of Graphs and Networks: From the Genome to the Internet* Bornholdt and Schuster (eds). Wiley-VCH, Berlin.

[12] Drossel, B., McKane, A.J., and Quince, C. (2004). The Impact of Non-linear Functional Responses on the Long-term Evolution of Food Web Structure. *Journal of Theoretical Biology* submitted. Elsevier Ltd.

[13] Dunne, J.A., Williams, R.J., Martinez, N.D. (2002). Food-web structure and network theory: The role of connectance and size. *PNAS* 99, no. 20.

[14] Dunne, J.A. Brose, U., Williams, R.J., Martinez, N.D. (2004). Modelling food-web structure and dynamics: implications for complexity-stability. *Sante Fe Institute Working Paper* 04-07-021.

[15] Fisher, Ronald A. (1930). The Genetical Theory of Natural Selection. Oxford University Press.

[16] Futuyma, Douglas J. (1998). Evolutionary Biology. Sinauer Associates, Inc.

[17] Garlaschelli, D., Caldarelli, G., Pietronero, L. (2003). Universal scaling relations in food webs. *Nature* 423:165-168.

[18] Gould, Stephen J. (1994). The Evolution of Life on the Earth. *Scientific American* 271, no. 4:85-91. SMS Meyer Periodical.

[19] Krebs, John R., Davies, Nicholas B. (1997). Behavioural Ecology, An Evolutionary Approach. Blackwell Science.

[20] Laessig, M. Bastolla, U., Manrubia, S.C., Valleriani, A. (2001). The shape of ecological networks. *Physical Review Letters* 86:4418.

[21] Langton, Christopher G. (1992). Life at the Edge of Chaos. *Artificial Life II* Langton, C.G., Taylor, C.E., Farmer, J.D., Rasmussen, Steen (eds). Addison Wesley.

[22] Lawlor, L.R. (1978). A comment on randomly constructed model ecosystems. *American Naturalist* 112:445-447.

[23] Leff, Harvey S. and Rex, Andrew F. (eds) (1990). Maxwell's Demon: Entropy, Information, Computing. Princenton University Press.

[24] May, Robert (1973). Stability and Complexity in Model Ecosystems. Princeton University Press.

[25] Maynard Smith, John and Szathmry, Ers (1995). The Major Transitions in Evolution. Oxford University Press.

[26] McCann, K., Hastings, A., Huxel G.R. (1998) Weak trophic interaction and the balance of nature. *Nature* 395, 794-98.

[27] Mendenhall, W., Beaver, R.J., Beaver, B.M. (2002). A Brief Introduction to Probability and Statistics. Duxbury.

[28] Morin, P.J. (1999). Community Ecology. Blackwell Science, Inc.

[29] Polis, G.A., and Winemiller, K.(eds) (1996). Food Webs: Integration of Patterns and Dynamics. Chapman & Hall.

[30] Quince, C., Higgs, P.G., and McKane, A.J. (2004). Topological Structure and Interaction Strengths in Model Foods Webs. *Ecological Modelling* submitted. Elsevier B. V.

[31] Quince, C., Higgs, P.G., and McKane, A.J. (2004). Deleting Species From Model Food Webs. *Oikos* submitted. Blackwell Publishing.

[32] Renshaw, Eric (1991). Modelling Biological Populations in Space and Time. Cambridge University Press.

[33] Sol, R. and Manrubia, S.C. (1996). Extinction and self-organized criticality in a Model of Large-Scale Evolution. *Physical Review E* 54:42-45.

[34] Standish, Russell K.(2004). Ecolab, Webworld, and self-organization. Proceedigs of the Ninth International Conference on the Simulation and Synthesis of Living Systems. MIT Press.

[35] Stauffer, D. and Chowdhury, D. (2003). Unified "micro-" and "macro-" evolution of eco-systems: self-organization of a dynamic network. *Physica A* 336:102-111.

[36] Williams, R.J., Martinez, N.D. (2000). Simple rules yield complex food webs. *Nature* 404:180-83.

[37] Williams, R.J., Berlow, E.L, Dunne, J.A., Barabsi, A-L, Martinez, N.D. (2002). Two degrees of separation in complex food webs. *PNAS* 99, no.20.

[38] Yodzis, P. (1981). The stability of real ecosystems. *Nature* 289:674-676.

A Code Listing

Webworld

Lotka-Volterra Webs

Utility Classes

Executables