Models in Evolutionary Ecology and the Validation Problem

C. C. Maley*
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
545 Technology Sq.
Cambridge, MA 02139
617-253-6567
cmaley@ai.mit.edu

Abstract

All models of natural systems represent an abstraction and simplification of that system. Thus all models suffer a validation problem. Should we believe that the results of the model have any bearing on reality? This is a particularly acute problem for Alife models of evolution and ecosystems. The time scale of evolution and the complexity of ecosystems make controlled experiments difficult. If Alife is ever to contribute significantly to biology, we must find methods by which we can build confidence in our models. One alternative to experimental tests of a model is to validate it against previously verified theory. I have applied a series of ecological and evolutionary validation tests to a model of species diversification. Examination of the predator-prey dynamics, trophic cascades, competitive exclusion, adaptation, and the species-area curve in the model has shown that a course grained spatial structure was inadequate to capture the realistic dynamics of an ecosystem. Only when spatial structure was extended to the local patch dynamics did the model begin to behave realistically under a wide range of parameters. Validation of the ecological dynamics of the model provides indirect support for the evolutionary behavior of the species within the ecosystem.

The Problem of Validation

Every model is an abstraction and a simplification. The goal of a model is to capture the essence of a system in the real world such that the behavior of the model matches the behavior of the real system. Thus for any model we may ask if it is a valid representation of the real system. Answering this question is the problem of validation. Traditionally we can try to disprove the validity of the model by collecting data from the real system and comparing it to the predictions of the model.

In artificial life we rarely have that luxury. Artificial life models tend to be highly abstract and general because the field is striving to discover general properties of life. This makes experimental validation extremely difficult. The time scale of evolution tends to restrict experiments to observation of the fossil record (Benton 1990, for example) or manipulation of organisms with extremely short life-cycles in simplified environments (Krukonis 1996, for example). Similarly, the complexity and size of ecosystems makes ecological experiments cumbersome and difficult to control.

An alternative form of validation can be pursued indirectly through reference to ecological and evolutionary theory. Instead of asking if the model matches the experimental data, we can ask if the model matches our understanding of the dynamics of ecology and evolution. Then, to the extent that the theories of ecology and evolution have been validated by experimental observations, we can disprove the validity of a model when it fails to match those theories. What follows is an example of this technique applied to a model designed to examine the factors that impact the origin and maintenance of species diversity. While the purpose of this model is to explore new theoretical ground in biology, the ecological and evolutionary dynamics in the model have been validated against theories of predation, competition, adaptation and island biogeography.

Previous Work

While there are many studies that have examined the evolutionary or ecological dynamics of particular models, few have used those studies to try to validate the models for future use. However, three validation studies stand out as particularly relevant.

Hraber and Milne (1997) looked at genotype diversity under the presence or absence of selection and varying mutation rates in the ECHO model (Holland 1992; 1993). Mirroring Bedau et al.'s (1992) results, they found that genotypic diversity was greatest under

*This research was supported in part by NDSEG grant DAAH04-95-1-0557. I would like to thank Rod Brooks, Michael Donoghue, and Hal Caswell for their guidance and support.

high mutation rates or in the absence of selection on the genome. One weakness of this study was the fact that the organisms were constrained to 7-bit genomes, and so genome space was extremely constricted.

Jones et al. (1997) also examined genome diversity. They found that the ECHO model did indeed demonstrate the power law of the species-area curve but that the exponent (z) was unrealistically large. This could be due to any number of abstractions in the model, but can be explained in part by the attempt to substitute genome diversity for species diversity.

Booth has attempted to validate the ecological dynamics of a dramatically modified version of ECHO, called Gecko (Booth 1997). She demonstrated that the model exhibits a "trophic cascade." This is a phenomenon of plant or autotroph biomass in response to predation. In the absence of herbivores, plant biomass ought to be relatively high. In contrast, the introduction of herbivores tends to dramatically reduce plant biomass. However, the further introduction of carnivores to prey on the herbivores tends to restore plant biomass to some intermediate level. These macroscale dynamics can be observed in Gecko reassuring us that the model represents a reasonable abstraction of predation at some level.

The Problem of Species Diversity

Why has life diversified so dramatically over the last four billion years (Benton 1995)? Or more specifically, what are the most important factors that have influenced speciation and extinction in the history of life? Darwin (1859) gave us the basis for a solid understanding of microevolutionary dynamics, but so far, macroevolutionary questions like these lack the analogous grounding.

Investigations into the dynamics of species diversity depend fundamentally upon the concept of a species. Unfortunately, there does not exist a species definition that unambiguously partitions organisms into mutually exclusive groups (Mayr 1942; 1963; 1982; Paterson 1985; Templeton 1989; Simpson 1961; Wiley 1978; Valen 1976; Cracraft 1983; Mishler and Theriot 1997; Baum and Donoghue 1995). Yet, if you were to plot organisms as points in some high-dimensional space where each dimension represented a measurable characteristic, organisms would form recognizable species clusters. The predominate species definition, the "biological species concept," or more appropriately the "reproductive species concept," defines two organisms as belonging to the same species if they can potentially mate and produce fertile offspring (Mayr 1963; 1982).

Hypotheses for the origination and survival of a species typically hinge on both abiotic factors like geographic isolation of a subpopulation, and biotic factors like expansion into new ecological niches through evolutionary innovations. A model to examine species diversity grounded in our knowledge of microevolutionary dynamics must include representations of organisms, species, geography, and the dynamics of sexual reproduction, mutation, migration, death, and predation. Geography and migration are necessary to model allopatric speciation in the isolation and divergence of subpopulations. Mutation, reproduction with inheritance, and death are necessary conditions for the process of evolution, assuming an environment with finite space. Furthermore, reproduction must be sexual for the reproductive species concept to be relevant. And finally, predation, with the concomitant dynamics of specialization and generalization, introduces ecological niches to the model and so allows testing of hypotheses that make reference to ecological niches.

An Instantiation

One possible instantiation[1] of these requirements has been implemented. All italicized terms below are parameters to the model. Their experimental settings are generally parenthesized.

Geography: Space is organized into a two-dimensional grid of "patches." Each patch may contain many (2K) organisms, but within a patch there is little spatial structure. Up to half of a patch may be occupied by plants. The other half is reserved for animals. The representation of topology and climate are described below.

Organisms and Predation: In ecosystems, organisms provide resources for other organisms. That is, organisms eat each other. Furthermore, some organisms specialize on a small set of resources (e.g., Koala bears only eat eucalyptus leaves) while others are generalists and can subsist on a wide variety of resources (e.g., humans eat just about everything). To capture these dynamics and thereby set up a food web, we use three 32-bit genes:

Phenotype: A bit pattern representing the structure and makeup of the organisms.

Prey Template: A bit pattern specifying the phenotype of an organism that can serve as food.

Generalism: A bit pattern mask specifying which bits in the prey template are not necessary for the

[1] The current model was written in CILK, a parallel extension of C, and run on the Xolas cluster of Sun Ultra symmetric multiprocessors. See `http://xolas.lcs.mit.edu/` for information on the processors and links to the CILK distribution pages.

selection of a prey organism. These bits indicate "wild-card" positions.

The question of whether or not an organism can eat another organism is answered by applying the predator's generalism mask (consider a 4-bit example: 0011) to the predator's prey template gene (e.g., 1010) producing a prey template with wild-card positions (e.g., 10**). The modified prey template gene (10**) is then compared to the prey's phenotype gene (e.g., 1011) and the number of mismatching bits (n) is counted. Then, the chance that the predator may eat the prey is e^{-n} (in this case $n = 0$ so $e^{-0} = 1$). Making predation probabilistic on the degree of matching helps to smooth out the fitness landscape.

The bottom of the food web is populated by autotrophs, organisms that synthesize their own food, which I will call plants. All organisms include a flag to indicate whether or not they are plants. Plants may only mate with other plants. Similarly, an plant lineage may never mutate into a heterotroph (herbivore or carnivore) and vice versa. This is the only hard, extrinsic distinction forced on the evolution of the food web. Instead of matching their prey template and generalism genes against a prey organism, plants match them against the climate of their environment. The climate of a patch is represented by another 32-bit pattern. A partial match results in the consumption of a fraction of the energy available from the environment. A fractional amount of energy is translated into an integer by rounding up with probability equal to the fraction. Energy is modeled by a particular number of meals granted to the plant. In the experiments below, plants could acquire a maximum of one meal per time step.

Given this implementation of predation, one might expect all lineages to evolve towards greater generalism. The fact that we see specialists in biology seems to imply that there is some tradeoff such that specialists might be favored over generalists by natural selection under some conditions. In fact, there is very little evidence pointing to such a trade off. A competing hypothesis postulates that specialists evolve from generalists. Genes for utilizing resources that aren't currently present are not under selective pressure for maintenance and so may be destroyed by mutations. In other words, specialists may have no advantage over generalists. There is some evidence for this (Thompson 1994). The real story is bound to be a lot more complicated. Thompson (1994) takes a middle of the road approach to the problem arguing that any given lineage will have a collection of three different types of genes. 1. Genes that help utilize a resource but are independent of genes for utilizing other resources. 2. Genes that help to utilize more than one resource. 3. Genes that help utilize a resource at the expense of interfering with the utilization of other resources. Only the final form of gene results in a tradeoff between specialism and generalism. Since the jury is still out on this issue, we have used the simplest assumption, that there is no tradeoff between specialism and generalism.

Mutation: At the time of conception of a new organism, every bit in the genes of the new organism has a small probability of being flipped. This is modeled with a Poisson distribution with a mean of one bit flipped for every 10 new organisms. However, for ecological experiments the *mutation-rate* is set to 0.

Death: Unlike taxes, death only comes in three forms. An organism may be eaten by a predator, it may starve to death, or it may be crowded out through competition for space. Starvation is modeled by requiring all animals, but not plants, to eat every time step. We have avoided a more complicated mechanism for death that includes a metabolic cost for staying alive. The guiding intuition here is that one must simplify the life history of an individual in order to save enough processing power to simulate on an evolutionary time scale.

If a patch is filled to its carrying capacity, a new organism entering that patch has a 50% chance of randomly displacing one of resident organisms. The loser of this competition for space dies. However, only plants may crowd out other plants, and only animals may crowd out other animals.

Reproduction: Once *energy-conversion* number of meals (3) have been consumed a reproduction attempt is triggered. If an organism can locate a viable mate in its patch, they will produce a new organism. The new organism's genes are constructed by forming the parent's genes into two rings and then implementing two-point crossover between the two parental genomes. This avoids the bias towards separating genes at the ends of a linear genome. The new organism's genome then undergoes mutation. Finally, the number of accumulated meals of the original parent is reset to 0. The determination of which organisms are viable mates forms the representation of reproductive barriers.

Reproductive barriers are implemented by giving the organisms a fourth gene called the reproduction gene. The bits of the reproduction gene are assumed to represent the set of characters influencing reproduction such as mating season, reproductive morphology, mate choice, etc. Organisms may only mate if their reproductive genes differ in at most one bit. That is, potential mates must have reproductive genes within hamming distance one of their mate. Like earthworms, any pair of organisms of the same species may mate and reproduce sexually, but there are no male and female "sexes" per se.

Species: Using a hamming distance restriction on mating means that the relation of potential mating is not transitive. Organism A may mate with B, assuming a one bit difference, and organism B may mate with C, also assuming a one bit difference, but A may not be able to mate with C, since A and C probably differ in two bits of their reproductive genes. While this may create problems in classifying organisms into species, it reflects a similar phenomenon observed in biology. Sometimes individuals of a species may be well adapted to their environment across a gradient of resources in the environment. It occasionally happens that while individuals of the species may mate with neighboring individuals, individuals from more distant populations do not produce viable offspring (Oliver 1972; Fox 1951). Consider a graph of the potential mating relationships with nodes representing the states of the reproductive genes in the populations and links between the nodes represent potential mates. Then the dangling cluster (connected component) you would get if you picked up one of the nodes in this graph represents a gene pool and thus a species. Luckily, there is an efficient ($O(n)$) algorithm for identifying the connected components of a graph (Cormen *et al.* 1990).

One desirable effect of separating out a reproductive gene from the other genes (as opposed to considering genetic differences across the entire genome to determine potential mates) is that a species may exhibit a wide distribution of phenotypes. Under this implementation of reproductive barriers, individuals from opposite extremes of an environmental gradient would be allowed to mate as long as they shared similar reproduction genes. It is conceivable that individuals in a species might differ radically in their prey preferences. However, due to the recombination of the parental genes, the hybrid offspring may not be viable. This constitutes an implicit or *de facto* reproductive barrier called a "post-mating" reproductive barrier.

The reproduction gene has twice as many (64) bits as the other genes. If there were only a few bits in the reproductive gene, then the probability of two lineages randomly evolving to within a single bit would be relatively high. Conventional wisdom in Biology assumes that it is highly unusual for two species to coalesce into one.

Migration: At birth, a new organism performs a random walk of length between 0 and *dispersal* (3), starting at the location of the parent. The *dispersal* parameter is a system wide parameter that is fixed throughout a run of the model. If there is space available in the location at the end of the walk, the new organism is placed there. Otherwise it there is a 50/50 chance that it will replace a living organism of the same type (plant or animal). Failing this, it dies. Generally, the heterotrophs are constrained by food source limitations, not space. The random walk models the observed dynamic that the probability of dispersal over a distance falls off approximately quadratically with distance from the parent.

Topology: The random walk in the migration of a newborn organism is modified by the topological barrier value of each location. This value is the probability that the new born organism fails to enter the patch and must remain at is current location for that step of the random walk.

Life in a patch proceeds by working through the stack of living organisms from youngest to oldest, giving each a chance to locate and consume one prey organism in the patch. If there are no edible prey in the patch, the time of the predator's last meal is checked to determine if it starves to death. However, if the predator does find and consume a prey organism, the number of meals it has consumed may trigger a reproduction attempt.

Theoretical Validation

The problems of speciation include both ecological and evolutionary dynamics of the model. Thus, it is important to validate both the ecological dynamics and the evolutionary dynamics. Ecological theory provides qualitative predictions for predator-prey dynamics as well as competition. Evolutionary theory predicts changes in populations as they adapt to their environments. Ecological and evolutionary theory combine in the predictions of island biogeography.

Predator-Prey Oscillations

The Lotka-Volterra equations were an early abstraction of the dynamics of predator and prey populations. Predator and prey population sizes tend to oscillate. When predators are rare, prey populations will naturally expand, providing more food for the predators.

Once the predator population expands in response to this abundance, they cull back the prey population. Eventually the prey become hard to find and predators begin to starve to death. And so the process repeats.

At a minimum, a model of an ecosystem should exhibit these oscillations. The model specified above doesn't. As written, a predator can automatically find a prey organism in its patch. This means that even when prey organisms are rare, the predators can still find them. The prey are driven to extinction and the predator population crashes shortly thereafter.

A *prey-location* probability was introduced to the model to represent the chance that a particular predator might find a particular prey organism in its patch. However, a predator still gets one attempt for every potential prey organism in its patch. This results in stable herbivore-plant oscillations across a wide range of parameters. In particular, the *prey-location* probability can range from about[2] 0.15 to 0.3 and still produce stable oscillations. Higher than that and the herbivores overeat, killing off all the plants. Lower than that and the herbivores cannot find enough food to sustain life.

Trophic Cascades

While the introduction of the *prey-location* probability stabilized the two-trophic level dynamics, the difficulties multiply when we move to three trophic levels by the addition of a carnivore species. Unfortunately, there is only a knife-edge of parameter space that can sustain both a two and a three trophic level ecosystem. Since herbivores are rare, relative to plants, the *prey-location* probability must be high to allow the carnivores to find the herbivores. However, if it is too high, the carnivores will kill off all of the herbivores. All three trophic levels could only be sustained by lowering the *energy-conversion* parameter to the unrealistic levels, requiring the consumption of only two prey organisms to produce a new predator organism.

The problem comes when the predator population has expanded and culled back the prey population to low levels. For a prey organism to survive during a time step, it must evade location by *every* predator in the patch. Even a low *prey-location* probability cannot save a prey organism from a large predator population.

The spatial structure of the real world tends to soften intense predation. A prey organism is generally not directly threatened by all the predators in the patch, because some of the predators are busy stalking other prey or pursuing other interests. The model was elaborated with two mechanisms. First of all a predator *search-radius* was added, restricting the amount of the patch that a predator can patrol in one time step. Each predator starts searching in a randomly chosen location in the patch's organism array, and then looks for prey in the *search-radius* consecutive following locations[3] in the array. Second, a prey organism that survives an encounter with a predator, because of the *prey-location* probability, becomes more difficult to find by other predators. Specifically, the *prey-location* probability is multiplied by $e^{-(\text{number of encounters} \times \textit{predation-distribution})}$. When the *predation-distribution* is 0, there is effectively no spatial structure, and the model behaves as it did before. However, when the *predation-distribution* is positive, a form of spatial structure is imposed on the patch, and a small prey population has a better chance of surviving intense predation.

Results As long as the carnivores were able to survive, a trophic cascade was found. The addition of the *search-radius* had little effect. A small *search-radius* actually helps to support a larger herbivore population by preventing over predation in a two trophic level system. However, if we want carnivores to survive, they must find the herbivores scattered amongst the plants, and so require a large *search-radius* to survive. In a patch that could hold 2048 organisms (with a maximum of 1024 plants), carnivores required a *search-radius* of more than 128, and tended to be stable with a radius of 512.

The *predation-distribution* had a much more dramatic effect. Both herbivore and carnivore populations were stable with a *predation-distribution* ≥ 0.8. In fact, the average biomass of both herbivores and carnivores increased with the *predation-distribution*, up to the maximum value that was examined (2.0), as seen in the lower graph of Figure 1. Each data point indicates an independent run of the model, and the lines show the mean for all the runs. There are at least 40 data points for each parameter value. A trophic cascade can be seen when the two trophic level plant biomass is significantly less than the three trophic level biomass.

A similarly counter-intuitive result can be seen if we look at the effect of the *prey-location* probability on the populations of herbivores and carnivores as shown in Figure 2. The most dramatic trophic cascade appears when the *prey-location* probability is 0.2 (the setting used for future experiments). It is interesting to note that all three trophic levels thrive when the probability is at a minimum 0.15. Once again, as the param-

[2]Tweaking other parameters like *energy-conversion* or the amount of energy given to the plants can increase this range of stability.

[3]The array was formed into a "ring" so that the last location in the array is adjacent to the first location.

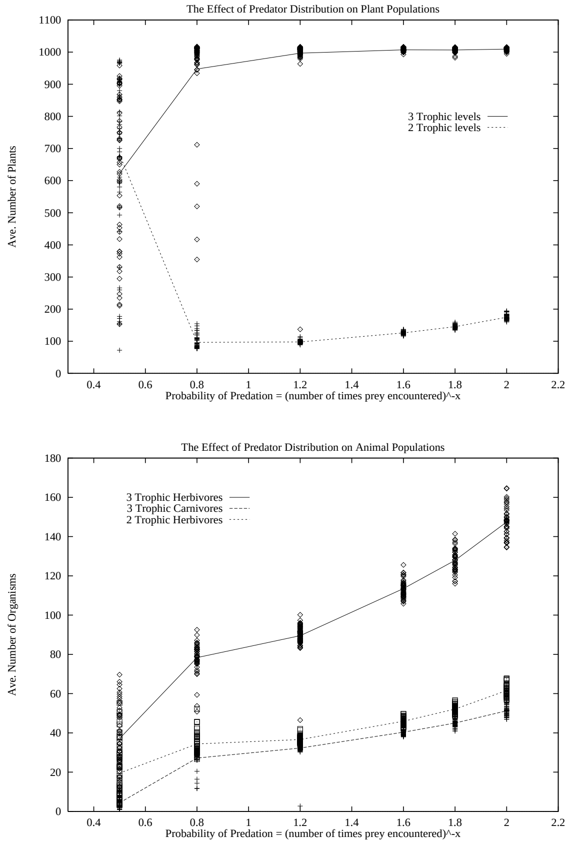

Figure 1: The upper graph shows the plant biomass in the two and three trophic level ecosystems. The lower graph shows biomass of the herbivores and carnivores in the two and three trophic level ecosystems.

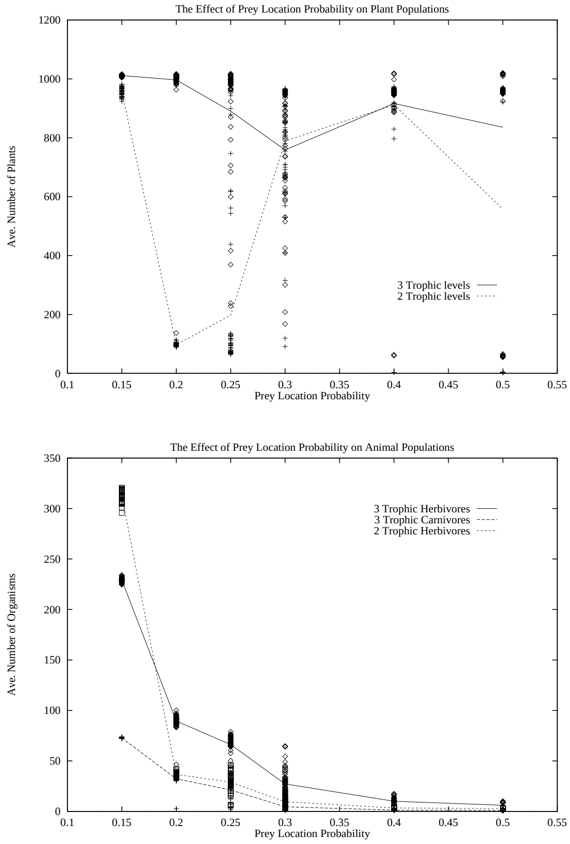

Figure 2: The effect of the prey location probability on the population sizes of plants (upper graph) and animals (lower graph).

eter values make predation more difficult, the average predator populations expand. This is probably due to a more efficient allocation of resources to the predators. The average maximum and minimum population sizes of the predators are shown in Figure 3 along with the means. When predation is easy, the predator populations can expand rapidly, reaching higher values than the maximum population sizes attained when predation is hard. However, the success of the predators has a dramatic effect on the prey, and so the predator populations collapse, reaching lower minima as compared to the minima in the runs where predation is more difficult. Ease of predation tends to destabilize the populations, causing larger fluctuations and resulting in lower average predator populations over time.

Competitive Exclusion

A common phenomenon in ecosystems is the competitive exclusion of one species by a species better adapted to the environment. While relative "adaptation" is often difficult to quantify independently of competitive exclusion experiments, it is easily identifiable in the model. A series of competitive exclusion experiments were designed wherein the environment was seeded with equal numbers of two species of plant. The "climate" of the environment had all its bits set to 0 and the prey template genes of the plant species differed by only one bit. That is, one of the two species was exactly one bit better adapted to the environment than the other. This should result in a slightly better energy absorption and a consequent slightly higher reproduction rate in the species with fewer bits set. A species was considered to have excluded the other species when it had expanded to 90% of the carrying capacity of the environment. 100 trials were run with 1, 4, 8, 16, 24, and 31 bits set to 1 in prey template gene of the inferior species. Mutation was turned off for these experiments. The superior species excluded the inferior in all the trials.

One might ask how long does it take for the superior species to take over 90% of the carrying capacity? If competitive exclusion tends to occur quickly, then we can infer that the slope of the fitness landscape in that region of the prey template gene is relatively steep. Figure 4 shows the average number of time steps re-

Figure 3: The average maximum and average minimum population sizes have been added to the herbivore and carnivore curves from the lower graph of Figure 1.

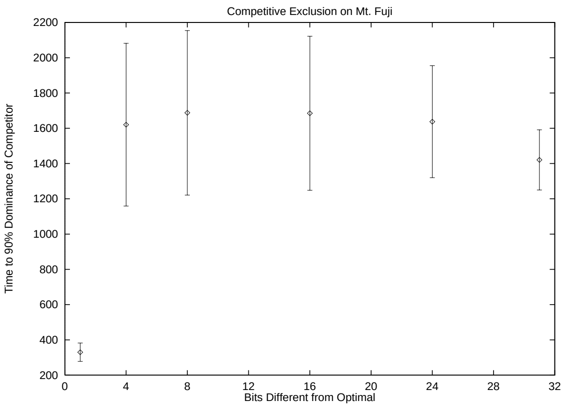

Figure 4: The average number of time steps, with standard deviation error bars, required for the superior plant species to take over 90% of the carrying capacity in a patch.

quired for the exclusion of the inferior plant. Figure 5 is a linear sketch of the fitness landscape with a single optimum surrounded by a steep slope that becomes more gentle by the time that the prey template gene is 4 bits away from the optimum. This is sometimes called a "Mount Fuji" fitness landscape.

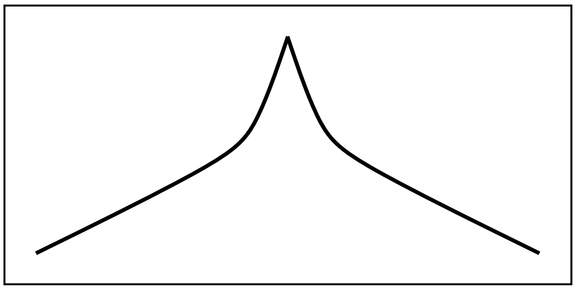

Figure 5: A sketch of the Mt. Fuji fitness landscape inferred from the data of Figure 4.

Adaptation

One of Darwin's (1859) key insights was that biological species are not immutable Platonic entities but rather they are dynamic conglomerations of lineages flowing over a fitness landscape. His theory of natural selection is in part a prediction that a species will adapt to its environment. That is, lineages should tend to climb the peaks in the fitness landscape.

A model of evolution should at least demonstrate that, all other things being constant, organisms evolve to be better adapted to their environment. This can be tested in the model when mutation is introduced into the dynamics. Consider a plant species evolving in the conditions of the competitive exclusion experiments. The optimal genotype matches the climate bit pattern of all 0's. This is achieved by having 0's in all the positions of the "prey template" that are not wildcards due to the generalism gene. If the prey template and generalism genes were evolving neutrally, with no natural selection, then we would expect one in four bits in the plant's prey template to have a 1 not masked by the generalism gene. We can thus detect natural selection in any significant reduction in that proportion of unmasked 1's.

The model was initialized with a plant species with exactly 8 out of 32 positions in its prey template set to 1 and no wild-card positions. After 5000 time steps the proportion of unmasked 1's was counted and averaged. The data for at least 32 runs in each setting of the *mutation-rate* parameter is shown in Figure 6. For all mutation rates, the resulting data is significantly below a mean of 25% unmasked 1's as tested by a one-sided t-test ($p < 0.001$ in all cases). There can be

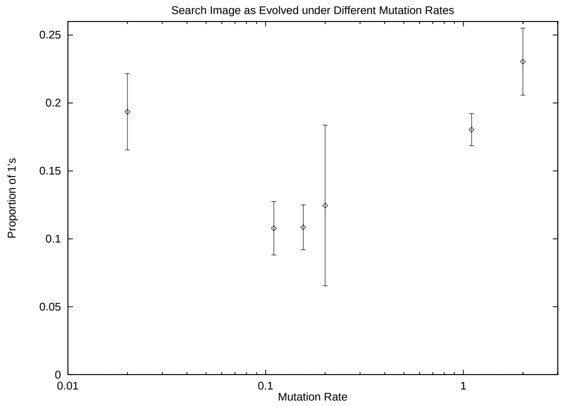

Figure 6: The normalized distance (with standard devations) of plant populations from the optimal genotype under different mutation rates. In all cases the null hypothesis of neutral evolution can be rejected.

little doubt that the plant species have adapted to their environment.

The Species-Area Curve

The previous experiments were conducted within a single patch. To investigate the relationship of species diversity to area, we must move to multiple patch experiments. The central result of island biogeography is that the number of species (S) on an island is a power law of the area (A) of the island (MacArthur and Wilson 1967). Specifically, $S = cA^z$, where c is a constant and the exponent z is particular to the group of organisms being studied. Begon et al. (1990, p.778) have summarized observed values for z which range from a low of $z = 0.10$ for English flowering plants to a high of $z = 0.43$ for mountainous mammals in the USA, with an outlier of $z = 0.72$ for cave dwelling invertebrates.

The model was run in four spatial configurations, 1 by 2, 2 by 2, 2 by 4, and 4 by 4 patches. Each patch could hold a maximum of 1K plants and 1K animals. The climates of the patches were made to differ by randomly flipping 2 of their 32 bits. There were topological barriers to migration of 0.9 between all patches. This meant that on the random walk which disperses new born organisms, a step of the walk had a 10% chance of entering a new patch. From the prospective of population genetics, where it only take one migrant per generation to prevent adjacent populations from diverging, this is a high migration rate. The mutation rate was set to 0.02 per 32-bit gene. Species diversity data were collected after 5000 time steps for 50 runs[4] of the model under each spatial configuration. A linear regression of the $\log_2$ species as determined by the $\log_2$ area gives $z = 0.45$ ($p < 0.001$, standard error $= 0.05$). While this is on the high end of the range found in nature, most island biogeographical studies have focused on a single group of organisms, such as birds, ants, or land plants, and do not look at diversity across multiple trophic levels, and so miss coevolutionary diversification effects. The z of 0.45 compares well to the values (0.7–0.95) found by Jones et al. (1997) for genotype diversity.

[4] 42 in the case of the 1 by 2 patch environment.

Conclusion

The initial failure of the model and subsequent elaboration of the predation algorithms illustrates an important benefit of validation studies. Failure in a validation study helps to sharpen our understanding of the essential features of the system and so helps to guide the further development of the model. It is also important to note that even when there exists a paucity of theory that can be used to verify the central results of a model, there generally exists a rich field of theory that impacts upon the dynamics of the model. In this case, a model of species diversification was designed for the very reason that there is a lack of established theory covering the topic. However, such a model must include both ecological and evolutionary dynamics and so we have validated it against both ecological and evolutionary theory. Because the dynamic of diversification intimately depends upon ecological and evolutionary interactions, these validation studies can help to support the end results of the model.

The fact that a model passes a series of validation tests against theory does not guarantee the quality of its results. The above model is only one instantiation of the requirements for modeling species diversity, and so stands as a single data point for theory. This perspective on models is particularly important to artificial life where seemingly trivial implementation details often manifest in artefactual results. In the end, the model does not provide Truth but rather a hypothesis that should be experimentally verified, similar to any other theoretical result. Although, it must be granted that such experimental verification is difficult. Meanwhile, by testing our models against theory, we may at least make progress toward significant theoretical insight.

References

D. A. Baum and M. J. Donoghue. Choosing among alternative "phylogenetic" species concepts. *Systematic Botany*, 20:560–573, 1995.

M. A. Bedau, F. Ronneburg, and M. Zwick. Dynamics of diversity in an evolving population. In R. Männer and B. Manderick, editors, *Parallel Problem Solving* *from Nature, 2*, pages 95–104. Elsevier Science Publishers, Amsterdam, The Netherlands, 1992.

M. Begon, J. L. Harper, and C. R. Townsend. *Ecology*. Blackwell Scientific Publications, London, 1990.

M. J. Benton. The causes of the diversification of life. In P. D. Taylor and G. P. Larwood, editors, *Major Evolutionary Radiations*, pages 409–430. Clarendon Press, Oxford, UK, 1990.

M. J. Benton. Diversification and extinction in the history of life. *Science*, 268:52–58, 1995.

G. Booth. Gecko: A continuous 2-d world for ecological modeling. *Artificial Life*, 3:147–164, 1997.

T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, Cambridge, MA, 1990.

J. Cracraft. Species concepts and speciation analysis. *Current Ornithology*, 1:159–187, 1983.

C. R. Darwin. *On the Origin of Species*. John Murray, London, 1859.

W. Fox. Relationships among the garter snakes of the *thamnophis elegans rassenkreis*. *University of California publications in zoology*, 50:485–530, 1951.

J. H. Holland. *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, MA, 1992.

J. H. Holland. Echoing emergence: Objectives, rough definitions, and speculations for echo-class models. Technical Report 93-04-023, Santa Fe Institute, 1993.

P. T. Hraber and B. T. Milne. Community assembly in a model ecosystem. *Ecological Modelling*, 1997.

T. Jones, P. T. Hraber, and S. Forrest. The ecology of echo. *Artificial Life*, 3:165–190, 1997.

G. P. Krukonis. Evolutionary potential of generalists and specialists: An experimental study with viruses. In *Bulletin of the Ecological Society of America*, volume 77 (3 SUPPL. PART 2), page 245, 1996.

R. H. MacArthur and E. O. Wilson. *The Theory of Island Biogeography*. Princeton University Press, Princeton, NJ, 1967.

E. Mayr. *Systematics and the Origin of Species*. Columbia University Press, New York, 1942.

E. Mayr. *Populations, Species, and Evolution*. Harvard University Press, Cambridge, MA, 1963.

E. Mayr. *The growth of biological thought : diversity, evolution, and inheritance*. Belknap Press, Cambridge, MA, 1982.

B. D. Mishler and E. Theriot. Monophyly, apomorphy, and phyologenetic species concepts. In Q. D. Wheeler and R. Meier, editors, *Species Concepts and Phylogenetic Theory: A Debate*. Columbia University Press, New York, NY, 1997.

C. G. Oliver. Genetic and pheontypic differentiation and geographic distance in four species of lepidoptera. *Evolution*, 26:221–241, 1972.

H. E. H. Paterson. The recognition concept of species. In E. S. Vrba, editor, *Species and Speciation*, pages 21–29. Transvaal Museum Monograph No. 4, Pretoria, 1985.

G. G. Simpson. *Principles of Animal Taxonomy*. Columbia University Press, New York, 1961.

A. R. Templeton. The meaning of species and speciation: A genetic perspective. In D Otte and J. A. Endler, editors, *Speciation and its Consequences*, pages 3–27. Sinauer Associates, Sunderland, MA, 1989.

J. N. Thompson. *The Coevolutionary Process*. University of Chicago Press, Chicago, IL, 1994.

L. Van Valen. Ecological species, multispecies, and oaks. *Taxon*, 25:223–239, 1976.

E. O. Wiley. The evolutionary species concept reconsidered. *Systematic Zoology*, 27:17–26, 1978.