Spatial Modeling of Epidemics

Project Summary

We are interested in an interdisciplinary research on simulating ecological systems on massively parallel computers. Our long term goal is to investigate formalizations of the probabilistic ecological model, to design software tools for model implementation and to prototype the developed software on sample applications. Due to the time limitation of the MasPar Challenge we have designed and implemented two simple but yet interesting models fitting our general scheme; the model of simple epidemics with and without a vector parasite.

There is a significant interest in determining the population densities and associated spatial patterns in ecological models. To simulate the spatial and temporal dynamics of a multi-species eco-system, we partition its habitat into a grid of sites so small that each site is capable of supporting at most one host. Each site's state indicates the presence or absence of species involved in a simulation. A probabilistic model can then describe a site's state transition as a function of the states of its neighboring sites (also referred to as ecological stencil). Typically, the stencil is defined as a rectangular collection of sites with the affected site located inside it (but not necessarily at the center). The size of the ecological stencil and the location of the affected organism within it are based on the biotic and abiotic characteristics of the species and the habitat.

Until recently, ecologists have not had access to sufficient computational power to run such models on a realistically large scale, since parallel processing is necessary to achieve acceptable run-times. The implemented model of epidemics was done on the departmental MasPar computer with 2048 processing nodes. Our initial success with this implementation, encourages us to undertake the development of tools that for the given description of the habitat (e.g. size of the site, size of the ecological stencil, states of the sites and functions for the transition probabilities) will generate a computational model for a parallel machine. Executing a model, the ecologist is able to measure the number of sites in each state at the end of simulation, the distribution of species in the habitat, pathogen contagion, species' spatial clusters and their size and frequency, as well as fractal dimensions of such clusters.

1 Biological Background

Pathogenic infections often reach sufficient prevalence among members of a host population that average survival and fecundity rates of the host are significantly reduced. Pathogens consequently can govern a host species' population dynamics and can indirectly influence the host's ecological interactions with competitors, predators or prey [9, 18]). On a longer time scale, pathogens can affect biogeographic distributions of host species and may generate strong selective/co-evolutionary pressures in host populations [16].

Our interests lie in the epidemiological ecology of a class of four-species eco-systems composed of two interacting hosts, a parasite on both hosts, and a pathogen that employs the parasite as a vector between host individuals. The developed models specify the occurrence of the various species in both time and space. Ecologists have long recognized that the spatial context of ecological interactions is fundamentally important to understanding population dynamics, community stability, and biodiversity [8, 10]. However, the computational power necessary to deal with the complexity of spatially explicit models has only recently been available to ecologists ([22]).

1.1 Biological Examples

Most epidemic models assume that a pathogen is transmitted when an infective host contacts a susceptible host. Recently, however, greater attention has been directed to vector-borne diseases where individuals of a parasitic species sequentially exploit many host individuals, and can transport the pathogen from an infected to an uninfected host. The direct effect of the parasite on the host population is sometimes small, but the indirect effect (via the pathogen) is often severe.

We assume that a vector-borne pathogen can infect individuals of two host species. The host populations compete for space, and once an individual of one host species occupies a site, its mortality (freeing the site) is independent of the density, local or global, of the other species.

A single parasitic species can occupy a site only if an individual of either host population already occurs at that site; the parasite's range cannot exceed the range of its hosts. If the parasite exploits both a host individual that is infected with the pathogen and a nearby uninfected host, the parasite may carry the pathogen from the former to the latter host individual. We allow the parasite to transmit the pathogen both within and between host species. The parasite generally will reduce a host individual's survival and reproduction to a lesser extent than will the pathogen.

Landscape ecology emphasizes the generation of spatial patterns and the consequences of spatial heterogeneity in natural communities and human-dominated eco-systems. Following Turner [22] (see also [10]), we can conveniently dichotomize research perspectives in landscape ecology. One view assumes spatial pattern in the environment and investigates consequences for ecological processes, ordinarily the growth and dispersal of species exploiting the environment. For example, habitat fragmentation may influence local population dynamics, extinction probabilities, and the stability of interspecific interactions. The other perspective assumes particular ecological processes within and among species, and asks how these biotic effects govern variation in the occurrence and co-occurrence of the species. The latter models include the few studies analyzing spatial characteristics of epidemic processes [14, 23].

The implemented models combine elements of both approaches to landscape ecology, but emphasizes the latter. The implemented simulations will address the question of how behavioral selectivity influences the dynamics of an epidemic and so governs spatial pattern in an interactive biotic community.

2 The Model

The model presented here is broader than strictly necessary for implemented experiments. The reason for this is to provide a firm theoretical foundation for the future simulations.

2.1 Ecological Stencil

Consider an environment E consisting of a finite number J of sites. The vector $N_t = [n_1\ n_2\ \ldots n_J]^T$ describes the state of the environment at time t: $n_k \in \eta_k$ is the value of the attribute that characterizes the site k, and η_k is a finite set of all possible values of this attribute. In the model n_k reveals identity of the species occurring at each site k. The initial condition of the environment E is described by the vector N_0.

The state space S of the vectors N_t is $S = \eta_1 \times \eta_2 \times \ldots \times \eta_J$. The function $\pi : S \to [0, 1]$ is a random field if $\sum_{N_t \in S} \pi_t(N) = 1$ at any time t. The random field is a distribution assigning state probabilities. The properties of $\pi(N_t)$ at any time t will necessarily depend on the probabilities of transition between the elements of S.

Site k is surrounded by neighboring sites δk; $\delta k \subseteq (E - \{k\})$. Let $N_{\delta k}$ be a subvector of N_t that contains values of all attributes characterizing sites in δk at time t. We refer to δk as the ecological

stencil for site k, implying that the state of the sites in δk, as well as the state at side k, govern probabilities of transition among the states η_k. That is, the abiotic and biotic processes driving the stochastic variation in n_k will be described by a function of both n_k and the state vector $N_{\delta k}$

Let $Pr[n_k = m \mid N_{E-\{k\}}]$ denote the probability that site k is in state m at time $(t + \triangle t)$, given that site k was in state n_k, and the rest of the environment was in state $N_{E-\{k\}}$, at time t. $Pr[n_k = m \mid N_{E-\{k\}}]$ is the one-step transition probability for change in the species composition at site k. In general, change at any particular site may depend on N_E, the state of the entire environment. But such strong interdependence among sites seldom will be reasonable biologically. Imposing the ecological stencil, we have

$$Pr[n_k = m \mid N_{E-\{k\}}] = Pr[n_k = m \mid N_{\delta k}]$$

The stochastic dynamics of the species composition at any site k will depend only on the state at this site and the state of the ecological stencil.

The size and shape of δk may vary with abiotic factors. For example, the strength and direction of prevailing winds may imply that only particular sites can disperse propagules that could become established at a given site. Biotic processes will also influence the ecological stencil. Dispersal capacities of hosts and parasites, in both absolute and relative senses, may regulate their interaction and, consequently, the effect of a pathogen on the host species.

2.2 The Transition Probabilities

We assume a four-species community. Two host species are exploited by a single parasite species. We identify the host species by a binary variable $h \in \{0, 1\}$. $h = 1$ implies "good" host (from the perspective of the selective parasite), and $h = 0$ implies "bad" host. The parasite can expect to live longer and more readily acquire new hosts when it exploits a good host. The parasite acts as a vector, transmitting a pathogen from infected hosts to uninfected hosts. The eco-system is assumed to be closed to immigration, so that extinction of one or more species is possible.

The environment contains J sites arrayed as a rectangular lattice. Each site k is in one of nine possible states and $\eta_k = \{1 \ldots 9\}$. A particular site may be empty (state 1), or may be occupied by one of the host species, but not both. If a host h occupies site k, it may occur (i) alone: state $3 - h$ (ii) with the parasite: state $5 - h$, (iii) with the pathogen: state $7 - h$, or (iv) with both the parasite and the pathogen: state $9 - h$.

state	host	parasite	pathogen
1	absent	absent	absent
3-h	h	absent	absent
5-h	h	present	absent
7-h	h	absent	present
9-h	h	present	present

Let $p_k(i, j)$ represent the one-step transition probability from state i to state j. To define $p_k(i, j)$ we need to consider all sites in δk, the ecological stencil for site k. Let the symbol $\sigma_j k$ indicate the number of sites in δk that are in state j, and P_i denotes the set of states to which transitions from the state i have positive probabilities, i.e. $P_i = \{j \in \eta | p_k(i, j) > 0\}$. Since $\sum_{j \in P_i} p_k(i, j) = 1$ then we can always express the probability of the transition preserving the state as $p_k(i, i) = 1 - \sum_{j \neq i \in P_i} p_k(i, j)$

The process of spreading species from site k to some other site i can be logically split into three stages: reaching the site i, succeeding at this site, and winning a competition there. Correspondingly, we can define three probabilities associated with each stage:

1. $p_{kr}(i)$ probability of species from site k reaching (with a seed for a host plant, or a coil for a dodder) to site i. Often this probability will be a function of a distance from k to i. We assume initially that in the ecological stencil this probability is constant [1] and equals p_{kr}.

2. p_{ks} – probability that a single arrival (assuming that the site was reached from exactly one outside site) succeeds at site k. For host plants this probability is often independent of the site[2] in which case we will denote it by p_s. Selectivity of parasites makes the rate of success a function of the states of involved sites (i.e. the function of the types of hosts residing at sites k and i).

3. p_{kc} probability of the host that succeeded on site k to win competition with the different type of hosts in competition for space on site k.

Under these assumptions the summary probability P_{ks} that species will succeed at site k in a single time unit is:

$$P_{ks} = \sum_{i=1}^{m} \left(\begin{array}{c} m \\ i \end{array} \right) (1-p_{kr})^{m-i} p_{kr}^{i} * (1-(1-p_{ks})^{i}) = 1-(1-p_{kr}p_{ks})^{m}$$

where m denotes the number of sites in the ecological stencil δk hosting this species. If all sites in the ecological stencil have the same probability of reaching site k, the probability of an organism succeeding there is just a function[3] of the probability $p_{ke} = p_{kr}p_{ks}$. Since competition occurs only between hosts we will discuss its influence on state transitions below.

Considering the death of the species, we assume that it takes places with the constant rate μ and the average life time span[4] of the species is $\frac{\mu}{(1-\mu)^2}$.

Applying this model to the considered four-species eco-system we obtain the following description:

organism	site	stencil	time tick	p_{ke}	states in m	μ
host h	any	δk	$l\triangle t$	ρ_h	$3-h$	μ_h
parasitized host h	any	δk	$l\triangle t$	ρ_{h+2}	$5-h$	μ_{h+2}
infected host h	any	δk	$l\triangle t$	ρ_{h+4}	$7-h$	μ_{h+4}
parasitized & infected host h	any	δk	$l\triangle t$	ρ_{h+6}	$9-h$	μ_{h+6}
parasite on host h	with host g	δk	$\triangle t$	α_{hg}	$5-h$	μ_p
virus	with host h		$\triangle t$	β_h	$7-h$ & $9-h$	μ_v

p_{kc} – the competitive advantage of good host over bad host at site k

Table 1: Parameters of the model

First we consider an empty site (state 1). Between time t and $(t+\triangle t)$, the site may be colonized by a host of identity h, or no change may occur, hence $P_1 = \{2, 3, 1\}$. The probability of the host h succeeding at site k is:

$$P_{khs} = 1 - \prod_{i=0}^{4} (1-\rho_{2i+h})^{\sigma_{2i+h+2}k}$$

[1] We later consider also a more realistic model in which the probability decreases with each site traversed by a factor $w < 1$ (e.g., seeds usually disperse over many sites as they move across the field). Hence, in the more advanced model $p_{kr}(i) = p_{kr} * w^{|k-i|}$

[2] In diversified habitats this probability can be a function of the site position in the field.

[3] In more general cases we need to know p_{hr} and p_{hg} separately.

[4] However the maximum life time is unlimited in this model, an assumption that is reasonable only for viruses. If necessary, we can represent the death rate as $\mu^{(t)}$ with assumption that $\mu^{(t)}$ is a fraction of the population at time $t-1$ that is dead at time t. If M is the maximum life span of the species then $\sum_{t=1}^{M} \mu^{(t)} = 1$ and for $t > M : \mu^{(t)} = 0$.

Since there is competition for a space between good and bad hosts, the total probability of the host h surviving at site k can be expressed as:

$$P_{khw} = P_{khs}(1 - P_{k1-hs}) + P_{khs}P_{k1-hs}(1 - h + (1 - 2h)p_{kc})$$

where $p_{kc} > 0.5$. From that the probability of transition can now be expressed as

$$p_k(1, 3 - h) = P_{khw}$$

Suppose that only the host of identity h occurs at site k. The parasite may attack, the host may die, or no change may occur, hence $P_{3-h} = \{5 - h, 1, 3 - h\}$, thus:

$$\begin{aligned} p_k(3 - h, 5 - h) &= (1 - \mu_h)P_{hkw} \\ p_k(3 - h, 1) &= \mu_h \end{aligned}$$

When a site is occupied by a host (of type h) and the parasite, the host may become free of the parasite (often through parasite mortality), the host may die (so that the parasite no longer occupies the site), the parasite may transmit the pathogen to the host, or no change may occur. Thus, $P_{5-h} = \{9 - h, 3 - h, 1, 5 - h\}$ and:

$$\begin{aligned} p_k(5 - h, 9 - h) &= (1 - \mu_{h+2})\beta_h \\ p_k(5 - h, 3 - h) &= (1 - \mu_{h+2})\mu_d \\ p_k(5 - h, 1) &= \mu_{h+2} \end{aligned}$$

Consider site k is occupied by an infected host (of type h) without the parasite. The parasite may attack the host, the host may recover from the pathogen, or the host may die (leaving the site empty), hence $P_{7-h} = \{9 - h, 3 - h, 1, 7 - h\}$. By evaluating the probability of any parasite succeeding in an attack on site k with a host of type h we obtain:

$$\begin{aligned} p_k(7 - h, 9 - h) &= (1 - \mu_{h+4})(1 - (1 - \alpha_{0h})^{\sigma_4 k}(1 - \alpha_{1h})^{\sigma_5 k}) \\ p_k(7 - h, 3 - h) &= (1 - \mu_{h+4})\mu_v \\ p_k(7 - h, 1) &= \mu_{h+4} \end{aligned}$$

Note that the virus does not have any direct effect on the parasite.

If a site is occupied by a host that has been both attacked by the parasite and infected by the pathogen, then the parasite may die or the pathogen may die, or all three species may be lost in a single time interval, consequently $P_{9-h} = \{7 - h, 5 - h, 3 - h, 1, 9 - h\}$ and

$$\begin{aligned} p_k(9 - h, 7 - h) &= \mu_d(1 - \mu_v)(1 - \mu_{h+6}) \\ p_k(9 - h, 5 - h) &= \mu_v(1 - \mu_d)(1 - \mu_{h+6}) \\ p_k(9 - h, 3 - h) &= \mu_v\mu_d(1 - \mu_{h+6}) \\ p_k(9 - h, 1) &= \mu_{h+6} \end{aligned}$$

2.3 Modeling Simple and Vector-Born Epidemics.

The classic simple epidemic [4] assumes that the individual is either susceptible of infective. Susceptibles acquire a pathogen through direct contact with infectives; the only allowable transition is from susceptible to infective. The population size is fixed during the course of the epidemic. The epidemic progresses until (asymptotically) each susceptible acquires the pathogen.

By replacing parasitism of adjacent hosts by a contact between individuals, the described model is used to depict a stochastic, spatially explicit characterization of the simple epidemic [3]. Since the simple epidemic excludes recovery, the pathogen loss was eliminated (i.e., set $\mu_v = 0$). Parameters governing the gain and loss of the parasite allow us to generalize the notion of "contact" between host individuals. The host grid might represent plants, animals with fixed territories, or stable home ranges [14], or households in human populations. For this submodel, the state space is reduced by comparison to the full 4-species model and for n_k the state space is just $\{1, 3, 5\}$. The relative frequency of state 1 (empty sites) remains constant in this submodel. Empty sites represent locations which the host species cannot occupy, or, alternatively, sites where host individuals cannot be infested by the parasite. A site empty at time $t = 0$ remains so, and a site with a host never becomes empty. The only transition in this simple model is from state 3 to state 5 (infection) which has a probability P_{kds}, where $P_{kds} = 1 - (1 - \alpha_{00})^{\sigma_5 k}$ (see the previous section for the explanation of the notation). The discussion of the result of simulations of this submodel is presented in the following section.

Our model was used also to simulate a vector-borne disease including stochastic, and spatially explicit aspects. Following the assumptions of the simple epidemic, we assumed no host mortality and no pathogen mortality. Hence, the initial density and spatial array of hosts have an important influence on the rate of epidemic progress.

The parasite represents a vector of the pathogen; the parasite can acquire a host individual and thereafter be lost from the host without the host being infected with the pathogen. This assumption generalizes the role of the parasite entity. That is, parasite infestation of a host amounts to exposure to a condition that may result in pathogenic infection. Exposure might amount to direct contact (of random duration) with neighboring individuals (within the ecological stencil) as well as infestation by a biological parasite vector.

For any site k in the array of J sites, the state n_k varies at the species composition as k varies. For this submodel, the state space is reduced by comparison to the full 4-species model and for n_k the state space is just $\{1, 3, 5, 7, 9\}$. The probabilities for the allowable transitions for this version of a simple epidemic are shown in the table below (see the previous section for the explanation of the used notation).

Allowable Transition	Transition Probability
State 3 to State 5	P_{kds}
State 5 to State 3	μ_d
State 5 to State 9	β_0
State 7 to State 9	P_{kds}
State 9 to State 5	μ_d

3 Computational Challenges

The model presented in the previous sections must be evaluated over many different initial vectors N_0. For each N_0 computations must be repeated many times in order to obtain probabilisticly reliable results. The size of the simulated field must be large enough to make the influence of the border regions of the field negligible. (The border regions behave differently than the field's interior because part of

their ecological stencil δk is absent). Consequently, the volume of the needed computation is large, and the use of parallelism is necessary to achieve acceptable modeling times.

3.1 Selected Architecture

It is clear from MasPar architecture description that MasPar Processor Arrays are a specific and sophisticated (note the role of the router) hardware realization of a two-dimensional *Cellular Automata*, the computing device proposed originally by von Neumann

- A neighborhood function $g : I \times I \to 2^{I \times I}$, where I is the set of integers, and $g(x) = (x, x + n_1, \ldots, x + n_k)$ for all $x \in I \times I$.
- a finite automaton (V, v_0, f), where V is a set of *cellular* states, $v_0 \in V$ is the quiescent state, and f is the local transition function from the $(k+1)$-th cartesian power of V into V, that satisfies the following condition: $f(v_0, \ldots, v_0) = v_0$.

This definition matches our model of the eco-system very closely. This match prompted us to select the Single Instruction Multiple Data (SIMD) MasPar computer for implementation of our model. Clearly, the sites of the field can be naturally mapped onto the processor array. Local changes of state at each site can easily be programmed as operations on local variables whereas the global statistics about the temporal and spatial developments in the field can be kept on the front end. The described match between our model and the underlying abstract definition of the architecture used for this model execution resulted in an efficient and reliable implementation.

3.2 Algorithmic Challenges

The ecological model requires different algorithms than those designed for numerical applications. For example, the ecological stencils can easily extend over dozens of sites, whereas numerical algorithms use stencils that are relatively small (in the range of few to several array elements [15]). Thus, it was necessary to design new algorithms for crucial operations involved in parallel model execution on distributed memory machines.

3.2.1 Data and Operation Placement

Optimal placement of data and operations in large scientific computations executed on a distributed memory machine can significantly reduce the total computation time. Let n_1, n_2 stand for the lengths of the sides of the affecting rectangle of sites, p_1, p_2 be the offsets (measured from the lower left corner of the rectangle) of the affected site, and $s_{i,j}$ be the current state of the site (i, j). Then, the state transition probabilities of site (m, q) are defined as:

$$\sum_{i=m-p_1}^{m+n_1-p_1} \sum_{j=q-p_2}^{q+n_2-p_2} f(s_{m,q}, s_{i,j})$$

The above summation is evaluated for each site in the field and for each step of the simulation. Hence, it is likely to dominate the total execution time of the modeling. This is just one example of a reduction evaluated simultaneously over many overlapping continuous sections of an array. Other examples of usage of such operations are likely to be found in cluster recognition, fractal dimension computation, or in modeling physical phenomena (e.g. solvers of partial differential equations characterizing fluid flow).

On a distributed-memory parallel computer, the efficiency of the implementation is strongly affected by the way data structures and operations in the program are distributed over the processors. Operations involving many operands can be performed most efficiently if the executing processor is selected as "close" as possible to the processors that store these operands (closeness is defined by the computer architecture and topology of the network involved). We refer to the problem of distributing data structures and operations as *data alignment*. This problem is particularly acute when the communication is synchronous as in the case of SIMD machines that we plan to use in our implementation.

Optimization of data alignment for regular iterative algorithms has been analyzed by us in [19]. We have also proposed there an approximate solution for this case. In the ecological modeling simultaneous reduction operation is of particular importance. As usual, we assume that the binary operator used in reduction is associative and commutative. Some algorithms proposed by us require also that this operation has an inverse, a condition satisfied by the most important reductions involving addition or multiplication. We also assume that an m-dimensional grid, where m is the dimensionality of the array involved in the reduction operation, can be nested in the computer communication network, as this is the case for our implementation on the two-dimensional grid of MasPar. Under this assumption, we need to consider only reduction operations evaluated simultaneously for an interval (i.e. a one-dimensional, continuous section of a linear array) over the intervals of the same array. The more general m-dimensional interval case can be solved by applying a one-dimensional interval algorithm to each dimension separately. The proposed algorithms are parameterized by the size of the reduced interval n and the position of the final result p. We assume that these parameters change from application to application, but are known at compile-time.

Simultaneous reduction is evaluated over a one-dimensional consecutive section of an array, called here an *array interval*. As explained above, we call this operation simultaneous reduction, because each array element is used as an operand to many reductions evaluated simultaneously over different overlapping intervals. This is distinct from what is usually referred to as *Parallel Reduction*, which involves parallel evaluation of a single reduction [1] or its variants. We designed and implemented several different algorithms for simultaneous reduction (see [20]) and measured their efficiency on MasPar (see [13]). As the result we selected a row-column algorithm that takes advantage of fast "xnet" MasPar communication.

3.2.2 Higher Level Algorithms and Code Generation

Data reduction is a crucial component of another algorithm that we investigated in the context of our ecological model. Some of the ecological measurements required the cluster recognition and cluster size computation. With the volume of data gathered in these experiments we needed to generate summary measurements of the spatial geometry. Ecologists use two measures which required some algorithm development:

- *Fractal Dimension* uses image component processing techniques, due to the nature of our data. For rapid prototyping we used a singular array for ease of simulating a PRAM architecture. Our current implementation uses the more efficient xnet constructs.

- *Relative Patchiness* uses parallel local computations and communications. We relied heavily on the xnet and data reduction features for this computation.

Below we briefly describe each measurement and the algorithms used to compute these values. The ecological significance of these measures is covered in the following sections of this document.

Fractal dimension could be computed in many ways; we chose to compute it by using an image component analysis technique, detailed by Krummel, et al. [11] and Sugihara , et al. [21], which uses

a linear regression of the logarithms of the area and perimeter of each image component. Two sites n_{i_1,j_1}, n_{i_2,j_2}, where $(i_1, j_1), (i_2, j_2)$ denote the coordinates of these sites, are considered to be *adjacent*, or *contiguous*, if and only if their coordinates differ by at most one, that is,

$$|i_1 - i_2| \leq 1 \wedge |j_1 - j_2| \leq 1$$

Clearly, this is the same notion as the eight nearest neighbor connectivity of the MasPar, when boundaries of the processor array do not have toroidal wrap. Two sites are m-adjacent if they are adjacent and share a common state, m, and the usual transitive closure of this relation is referred to as m^*-adjacency. An *image component* or a *cluster* of state m, denoted by C_m, is any of the maximal set of sites that are in m^*-adjacency relation to each other, i.e. $s_1\ m*-adjacents_2 \in C_m\ \Rightarrow s_1 \in C_m$. The *image component analysis problem* consists of finding all image components for the given state m. The *graph connected component problem* is a generalization of this problem, where given a graph, the vertices map to sites, and adjacency is determined by the edges in the given graph. This is a well known problem in image processing applications, and many algorithms exist for both the graph connected component problem, and the image component problem.

In our solution we wanted to have a natural mapping of the problem to the architecture, take advantage of its communications primitives (xnet), and limit (or eliminate altogether) use of shared memory. Many algorithms proposed for the image component problem rely on remapping the state of each site to the processors [6, 17], or on introduction of "dummy" sites to prevent algorithmic failure [2]. However, the considered architecture had shared memory. Thus, we used a more appropriate algorithm developed by Cypher [7], for mesh connected SIMD machines. This algorithm assigns a unique number to each cluster C_m. In our applications clusters contain sites which have states indicating the presence of the pathogen.

The fractal dimension measurements involve computing the perimeter and area for each cluster, which can be done in parallel using data reduction. A site is on the *perimeter* of a cluster if it maps to a PE on the boundary of the PE array, or if any of its neighbors has a different state (a neighbor being any PE within one xnet hop in any direction). The processor having processor number $iproc = label$ is called the *root* processor of the cluster. Each label is decoded into a pair of coordinates x_{pos} and y_{pos}, defined as $x_{pos} = label \bmod nxproc$ and $y_{pos} = \lfloor \frac{label}{nxproc} \rfloor$. The pair x_{pos}, y_{pos} uniquely identifies a processor responsible for computation of the cluster with the given unique label (this is the processor with $ixproc = x_{pos}$, $iyproc = y_{pos}$. For a site outside of any cluster these values are set to -1. Since on the MasPar, $nyproc \leq nxproc$ each processor with $ixproc = x_{pos}$ will also process initially data for all clusters with the labels that decode into x_{pos}. First, local cluster data (size increment equal one for each site, perimeter flag (1 for sites on the perimeter, 0 otherwise) are circularly shift $nxproc$ times (along x dimension) and recorded by the proper processor (i.e. the one with $ixproc = x_{pos}$ of the label of the site originating the data). Similar shift of this information along y dimension enables the processor responsible for the given label to obtain the needed values. If n_c denotes the maximum number of clusters (clearly $c_n < N/4$) then the maximum number of shifts in the second stage, as well as the size of shifted data is equal to $c_n/nxproc$. The total number of steps is $0(\sqrt{N})$.

Relative patchiness is easy to compute on the MasPar thanks to its mesh topology. For each state, m, we compute the relative patchiness of that state, $R_p(m)$ by performing a parallel computation followed by a data reduction. There are two cases, some site in the field, $n_k \in N_t$ at a given time has state m, or no site in the field n_k has state m. In the latter case relative patchiness of that state is obviously zero. Otherwise let $N_m = \{n_k \mid n_k \in N_t \wedge n_k = m\}$. There is a set of edges connecting sites to adjacent sites and the spatial boundaries, denoted $adj(N_m)$. The number of edges per site is the same as the connectivity of each processor on the MasPar as we do not make a special case of the boundary, giving us $|adj(N_m)| = 8|N_m|$. We also need to compute the number of adjacent sites

in dissimilar states, $D(n_k, m) = \{n_i | n_i \in adj(N_m) \wedge n_i \neq m\}$. To get the final result we compute the ratio:

$$R_p(m) = \frac{|D(n_k, m)|}{|adj(N_m)|}$$

The complexity of this computation is $O(m \log N)$ since the data reduction is performed $O(m)$ times with $O(\log N)$computational cost per invocation.

4 Performance Results And Interpretations

To compare efficiency of our modeling on MasPar with efficiency on other machines we selected one serial (A Sun Sparc Station 1) and one parallel machine (shared memory Sequent S81 with 15 processors) to run the modeling.

On lightly loaded machines we ran the job 100 time steps sampling every 10 time steps. We used a 2×2 stencil. The version of the software ported is essentially the same as implemented on MasPar, except that it outputs the contents of each site at every sample, and does not have the relative patchiness/image component analysis measures installed. We also ran the models on maspar.maspar.com site with 8192 processor MasPar, thanks to the cooperation of Helen Asher and the MasPar Corp. The maspar.maspar.com site was significantly more heavily loaded than our local machines.

Let N represent the number of sites, P the number of processors for a given architecture. Our benchmarks provide a fair comparison for the MasPar, because we were careful of used algorithms efficiency. The slowest algorithm we used, the image component analysis subprogram, is of complexity $O(\sqrt{N})$. In the other architectures the software runs in no better than $O(N)$ time, as each individual element is touched. We timed our benchmarks using the Unix 'C' shell time command. We ran the Sequent S81 in both single processor mode, and with all 16 of its processors switched on, and these values are indicated in parentheses after in the Machine Column. The results are shown in the following table.

Machine	Compiler	Optimization	Sites	User	System	Elapsed	Util
MasPar MP-1	mpl 2.1.17	Default	2048	66.5	1.3	1:12	93 %
Sparc Station 1	gcc 2.1	Default	2048	701.4	1.2	12:02	97 %
Sparc Station 1	gcc 2.1	-O	2048	62.3	0.1	1:02	99 %
Sequent S81 (1)	cc	Default	2048	546.7	0.9	9:07	99 %
Sequent S81 (1)	cc	-O	2048	474.7	0.7	7:55	99 %
Sequent S81 (16)	cc	Default	2048	594.4	0.7	0:40	1457 %
Sequent S81 (16)	cc	-O	2048	517.8	0.8	0:35	1450 %
Sparc Station 1	gcc 2.1	-O	4096	143.2	2.4	3:11	76 %
Sequent S81 (16)	cc	-O	4096	1058.5	0.8	1:12	1463 %
Sparc Station 1	gcc 2.1	-O	8192	261.1	0.2	4:22	99 %
Sequent S81 (16)	cc	-O	8192	2164.6	3.4	2:27	1472 %
Sparc Station 1	gcc 2.1	-O	16384	526.1	0.5	8:48	99 %
Sequent S81 (16)	cc	-O	16384	4947.9	22.0	5:36	1473 %

Timings For Differing Architectures And Number Of Sites

One thing that is surprising is the amount of difference that the optimizer made on the Sparc Station 1. It is also interesting to note that as the number of sites increased, the inherent limitations of the single processor architecture resulted in a performance degradation.

On the Sequent the timing is taken across all of the processors, for the user and system timings, so they are really about $P = 16$ times too high. The S81 manages processes by having one processor

acting as the master, and forking off slave tasks, and joining the separate tasks together at the end of their execution. This causes the master to be inactive during the duration of the forks, so the Sequent loses $\frac{1}{P}$ of its processing power during the parallel portion of its execution.

Unfortunately, we could not run an optimized version of the MasPar software, because our site does not have any user selectable options for optimization. Notice that the MasPar implementation would scale up with small performance penalty to a larger configuration, as communication is local and regular in our computation, and the most frequent operation, summation of the stencil is performed very efficiently in parallel, while the Sparc and Sequent cannot handle the larger number of sites as efficiently. Unfortunately we could not get good timings off of the machine at maspar.maspar.com, as the system operator has the machine swapping jobs to disk, each swap taking "Minutes to complete" (according to the operator), and is mutually exclusive of processing. Our benchmark could not run to completion without a swap.

5 Ecological Results And Interpretations

The classic simple epidemic [4] assumes that each individual in a population is either susceptible or infective. Susceptibles acquire a pathogen through direct contact with infectives; the only allowable transition is from susceptible to infective. Population size is fixed during the course of the simple epidemic. Spread of the disease occurs without reference to spatial structure; the classic model assumes homogeneous mixing. The model is deterministic, so that the epidemic progresses until (asymptotically) each susceptible has acquired the pathogen.

Our model differs from the simple epidemic in several, biologically significant respects. We consider a stochastic, spatially explicit model of a vector-borne disease. Individuals of a single host species can be infested by a single parasitic species. The parasite serves as the pathogen's vector (without effect on the parasite). Infestation of a host by the parasite depends on the host's local environment (the ecological stencil), as does transmission of the pathogen from infective to susceptible host via their common parasite.

Our model resembles the simple epidemic to the extent that neither hosts nor pathogens suffer mortality. Hence the initial density and spatial array of hosts can influence the spatio- temporal occurrence of the pathogen.

Since our model is stochastic, and spread of the infection must be preceded by successful propagation of the parasite, it need not be true that every susceptible host ultimately becomes infective. Clearly, if the parasitic species goes extinct before the entire population is infected, no further susceptibles can acquire the pathogen.

The biological environment at the initiation of each simulation is relatively simple. Most sites in the environment (90.2%) are occupied by hosts that are neither infested by the parasite nor infected by the pathogen. A thin barrier of sites that cannot be occupied by host individuals divides the environment into left and right halves. Two "islands" within the barrier region join the halves of the environment. The parasite, and hence the pathogen, can disperse across the barrier only by occupying one or both of the islands first. At the far left edge of the left half of the environment, we place a relatively small number of parasitized hosts, hosts with pathogenic infection and no parasite, and hosts with both the parasite and the pathogen. Each of these states occupies close to 2% of the environment's sites. Essentially, the epidemic spreads from left to right across the environment. The overall frequency of infected hosts, as well as both large and small scale spatial patterning in the pathogen's occurrence, depend on the simulation parameters.

To analyze the simulation results biologically, we examine the following quantities:

1. the overall frequency of diseased hosts (i.e. with and without the parasite) in the environment

at each of 12 sampling points (including the initial condition);

2. the spatial extent of the epidemic at the final sampling point;

3. the overall biodiversity of the environment at each sampling point; we use an entropy-based formulation for diversity; and

4. the fractal dimension of the ensemble of patches of infected hosts; we used the perimeter-area regression method at the final sampling point to estimate the fractal dimension of the epidemic.

The environmental frequency of diseased hosts allows us to track the epidemic's temporal evolution. Clearly, the extent of infection within a population is the most fundamental biological attribute of the process. Similarly, the spatial extent of the epidemic (following simulation) provides simple comparative descriptions of different results. Over a given number of time steps, some epidemics may fail to grow significantly, some may progress to the barrier and stop, and others may infect every host in the environment.

Biodiversity indices address the collective, qualitative variability of the environment. Diversity, by definition, increases as the number of distinct biological states increases; states are identified by the presence/absence of host individuals, the parasitic species, and the pathogen. Diversity, hence, is very sensitive to the extinction of the parasite. Diversity also increases as the "evenness" of the proportional representation of the various (extant) states increases. Diversity will be low if, for example, uninfested and uninfected hosts dominate the landscape. Similarly, diversity will be low if the pathogen reaches most or all sites in the environment. Diversity will increase as the community approaches an even mix of hosts without parasite or pathogen, hosts with the pathogen only, and hosts with both parasite infestation and pathogenic infection. To calculate diversity we use the quantity H:

$$H = -\sum_{i=1}^{S} p_i \ln(p_i)$$

where S is the number of distinct states in the environment ($1 \leq S \leq 5$, for this model) and p_i is the fraction of the sites in the environment where the i^{th} state occurs.

Finally, the fractal dimension of the collection of patches where the pathogen occurs provides a quantification of the spatial complexity of the epidemic. If D is the fractal dimension of an ensemble of patches, we expect $1 < D < 2$; patches are more complex than linear ($D = 1$), but are sufficiently irregular in shape that $D < 2$. Increasing fractal dimension indicates increasing spatial complexity. We defined patches by presence of the pathogen at a focal site and any of the 8 surrounding sites. We estimated D by linearly regressing $\ln(patchperimeter)$ on $\ln(patcharea)$; D is twice the slope of the resulting regression line.

To present the initial biological results, we compare epidemics where the host species' susceptibility to infection is low ($\beta = 0.001$) with epidemics where susceptibility is high ($\beta = 0.5$). Within a given level of susceptibility, the results differ predictably as the ratio $(\frac{\alpha}{\mu_d})$ varies. α quantifies the rate at which a focal, uninfested host is attacked by the parasite infesting a nearby host individual. μ_d is the probability that the parasitic infestation of a host individual is lost in a single time step. Essentially, α is the parasite's birth parameter, and μ_d is the parasite's mortality parameter.

For the low-susceptibility case, we let the parasite's attack rate α be $0.002, 0.015$ and 0.125. For each value of α, we let μ_d be $0.001, 0.01, 0.1$ and 0.5. Four clear patterns emerge in the simulated biotic community; none includes extensive infection of hosts by the pathogen.

1. When $\alpha = 0.002$, the parasite cannot spread fast enough (in 100 time steps) to provide the conduits for the potential spread of the pathogen. Consequently, the vector-borne disease is essentially confined to the area where the pathogen was introduced.

2. When $\alpha = 0.015$ (and $\mu_d < 0.5$), most hosts in the left half of the environment acquire the parasite in 100 time steps. However, the parasite is not able to disperse across the barrier. Successful propagation of the parasite in the left region provides a vector for the pathogen; the constraint identified in the first case is removed. However, we observe few diseased hosts because the susceptibility to infection (β) is so low.

3. When $\alpha = 0.125$ (and $\mu_d < 0.5$), all or nearly all host individuals in the entire environment acquire the parasite within 100 time steps. The incidence of disease remains low, since susceptibility is so low.

4. When $\mu_d = 0.5$, the second and third cases change in that the parasite's spatial distribution becomes fragmented by mortality. Large patches of parasitized hosts become a mix of uninfested and infested host individuals, so that a spatial "flicker" emerges.

Table 2 shows the average frequency of infected hosts and mean biodiversity indices for a set of simulations producing pattern 1 above. Table 3 shows results for pattern 2, where the parasite reaches, but does not cross, the environmental barrier. Table 4 presents numerical results for pattern 3, where the parasite spreads across the entire environment. Results verify our intuitive understanding of the dynamics. Vector-borne disease may be spatially constrained because the pathogen's vector is constrained, and/or the host's susceptibility is low. For the high-susceptibility case ($\beta = 0.5$), we let the parasite's attack rate α be 0.002.0.015 and 0.5. For each value of α we let μ_d range from 0.001 to 0.5 as above. Four clear patterns again emerge; we number them 5 through 8.

1. When $\alpha = 0.002$, the parasite cannot spread. The pathogen, therefore, remains at low frequency despite the host's high susceptibility to the disease.

2. When $\alpha = 0.015$, most hosts in the left half of the environment are parasite-infested and then infected by the disease. However, the barrier stops the spread of the epidemic (through 100 time steps), because the parasite fails to disperse into the right half of the environment.

3. When $\alpha = 0.5$ and $\mu_d < 0.1$, all or nearly all host individuals acquire the disease within 100 time steps. Most host individuals are infested by the parasite as well as infected by disease.

4. If $\alpha = 0.5$ and $\mu_d \geq 0.1$, the environment contains all (nearly all) infected hosts. However, many hosts have lost the parasite through mortality.

Table 5 presents the average frequency of infected hosts and biodiversity indices for a two simulations producing pattern 5. Table 6 shows average numerical results from simulations producing pattern 6, and Table 7 gives results from a set of simulations producing pattern 7. The numerical results again indicate the importance of the vector for this type of disease, even when susceptibility among hosts is high. We see that a large-scale epidemic for a vector-borne disease requires an ecologically successful vector species (the parasite) and a minimal level of susceptibility among hosts. Further research will detail the interactive effects of parasite population growth and host susceptibility to the pathogen on spatio-temporal patterns in epidemics.

In most of our simulations patches of diseased host individuals tended to be relatively few in number, but often large in size. Essentially, the disease either failed to spread very far (or at all), or else quickly spread to the barrier (or across the entire environment). Fractal dimensions of patches describe large-scale spatial complexity. The greatest estimates of D (fractal dimension of the ensemble of patches of diseased hosts) occurred when the disease did not spread very far from its initial locations. For these cases we observed $1.9 > D \geq 1.7$; "patchiness" was relatively complex. When the disease spread over the entire environment (or the left half in some cases), spatial complexity was reduced. For these cases

we observed $1.7 > D \geq 1.34$. Growth of the epidemic reduced spatial complexity; spatial heterogeneity declined as the incidence of the disease increased.

Most of our biological results make intuitive sense. What is exciting is that we can produce the full range of epidemic conditions by manipulating only a few, easily understood, probabilistic parameters. Hence we can explore significant biological questions in a computationally convenient manner.

5.1 Table 2

Low susceptibility; neither the parasite not the pathogen spread during 100 time steps. $\beta = 0.001$ and $\alpha = 0.002$ for each simulation. Entries are means across μ_d values $0.001, 0.01, 0.1$ and 0.5. P represents the overall frequency of the disease in the environment; H is community biodiversity. Sampling time 0 is the initial condition; each positive sampling time occurs after 10 time steps in the simulation.

Time	P	H
0	0.042	0.45
1	0.043	0.41
2	0.043	0.41
3	0.043	0.41
4	0.043	0.42
5	0.046	0.42
6	0.049	0.43
7	0.048	0.44
8	0.049	0.45
9	0.05	0.45
10	0.052	0.47
11	0.051	0.47

Table 2: Run with Parameter Set 1

The standard error for each P-value is 0.02. The minimal biodiversity value observed at the final sampling point was 0.32 (for $\mu_d = 0.5$); low diversity reflected the loss of hosts with both the parasite and the pathogen, and loss of hosts with only the parasite. The maximal biodiversity value observed in this set of simulations was 0.63 (for $\mu_d = 0.001$); low parasite mortality preserved community diversity.

5.2 Table 3

Low susceptibility; the parasite reaches, but does not cross, the barrier. The pathogen does not spread. $\beta = 0.001$ and $\alpha = 0.015$ in each simulation. Entries are means for $\mu_d = 0.001$, 0.01 and 0.1. The parasitic species was extinct by time step 20 for $\mu_d = 0.5$. P and H are as above.

Time	P	H
0	0.042	0.45
1	0.044	0.52
2	0.045	0.64
3	0.049	0.75
4	0.052	0.82
5	0.053	0.89
6	0.058	0.93
7	0.06	0.96
8	0.063	0.99
9	0.065	1.0
10	0.067	1.01
11	0.067	1.01

Table 2: Run with Parameter Set 2

Although the disease remains confined as above, biodiversity is greater than in pattern 1; at a broad spatial scale, uninfested and hosts infested by the parasite occur at roughly equal frequency. Little variation in H values was noted among runs.

5.3 Table 4

Low susceptibility; the parasite's distribution spans the environment. $\beta = 0.001$ and $\alpha = 0.125$. Despite the ecological success of the parasite, disease frequency (P) remains low because of low susceptibility. Symbols are as in preceding tables. Entries are means for $\mu_d = 0.001, 0.01$ and 0.1.

Time	P	H
0	0.042	0.45
1	0.044	0.86
2	0.045	0.97
3	0.048	0.95
4	0.053	0.65
5	0.055	0.62
6	0.06	0.62
7	0.062	0.63
8	0.065	0.62
9	0.066	0.64
10	0.068	0.64
11	0.068	0.64

Table 4: Runs With Parameter Set 3

The decline in average biodiversity at sampling time 4 reflects a rapid decline in the frequency of pathogen-infected, but unparasitized hosts occurring in the results.

5.4 Table 5

High susceptibility; spread of the disease is constrained by the parasite's extinction or slow growth. $\beta = 0.5$ and $\alpha = 0.002$. For $\mu_d = 0.1$ and 0.5, parasite extinction occurred before the simulation was through half the time steps. The numbers below are for $\mu_d = 0.001$ and 0.01. Symbols are as above.

Time	P	H
0	0.042	0.45
1	0.063	0.43
2	0.066	0.44
3	0.072	0.46
4	0.078	0.46
5	0.079	0.48
6	0.092	0.49
7	0.097	0.51
8	0.105	0.53
9	0.111	0.54
10	0.121	0.56
11	0.121	0.56

Table 5: Runs With Parameter Set 4

5.5 Table 6

High susceptibility; the disease reaches, but does not cross, the barrier. The effect of the barrier on the epidemic occurs because the parasite fails to cross the barrier. $\beta = 0.5$ and $\alpha = 0.015$. For $\mu_d = 0.5$, the parasite went extinct within 20 time steps. Entries are means for $\mu_d = 0.001, 0.01$ and 0.1.

Time	P	H
0	0.042	0.45
1	0.084	0.49
2	0.138	0.63
3	0.197	0.73
4	0.256	0.81
5	0.312	0.86
6	0.362	0.92
7	0.41	0.92
8	0.45	0.94
9	0.47	0.95
10	0.48	0.95
11	0.48	0.95

Table 6: Results From Parameter Set 5

5.6 Table 7

High susceptibility; the disease occurs throughout the environment. $\beta = 0.5$ and $\alpha = 0.5$. Entries are means for $\mu_d = 0.001, 0.01, 0.1$ and 0.5.

Time	P	H
0	0.042	0.45
1	0.33	1.0
2	0.55	1.12
3	0.88	0.66
4	0.96	0.4
5	0.97	0.4
6	0.97	0.4
7	0.96	0.4
8	0.96	0.38
9	0.97	0.4
10	0.96	0.39
11	0.96	0.39

Table 7:Runs From Parameter Set 6

The decline in biodiversity following sampling time 2 occurs as both hosts without parasites or the pathogen and hosts with parasites, but without the pathogen decline to very low frequency or are lost.

6 Conclusion and Impact on Biological Sciences

The distribution and abundance of all species exhibit some degree of spatial variation. Spatial heterogeneity in abiotic factors or biotic processes may govern population dynamics and the resulting characteristics of ecological communities. Despite a long-standing recognition of the importance of spatial variation, analytical and computational models of spatially detailed ecological interactions have only recently become available [8, 10]. Our spatially explicit model addresses the population dynamics of (as many as) four species through simulation of the epidemiological landscape of a carrier-borne disease. Related ecological questions can be approached by modifying the model.

At a general level, our work should benefit ecologists since it combines two current, challenging elements of ecology. The simulations will investigate the ways parasites and pathogens may regulate host populations by exploring the spatio-temporal dynamics of the interacting species at the landscape level [22].

Our simulations should provide several, more specific benefits to ecologists. Each potential benefit arises as a result of an advantage of parallel computing over available analytical models and methods.

Our model links individual-level behavior directly to population dynamics and, consequently, to community-level properties. Integrating principles from behavioral ecology into population dynamics has been an important, but elusive, research objective. Our model will generate predictions concerning effects of behavioral selectivity on a consumer's expected logarithmic growth rate in a random environment.

Most spatially explicit, analytical models of ecological interactions rely on reaction-diffusion methods. In contrast, we plan to present our ecological results in terms that should appeal to experimentalists. To do so, we will use methods of landscape ecology [22]. Larger-scale patterns will be characterized by frequencies of patches of different sizes, etc. Smaller-scale patterns will be quantified by nearest-neighbor probability distributions, etc.

Some spatial stochastic processes can be solved for a statistical equilibrium. However, these analytical methods have a restricted transition structure and yield only equilibria - rather than time-dependent behavior. Our simulations, of course, describe the temporal and spatial evolution of our models. Epidemic theory generally requires analysis of the time course of the spread of disease. Hence simulation may prove more valuable than analytical modeling of random fields.

The benefit of simulating of the temporal and spatial evolution of the eco-system over finding its statistical equilibrium is particularly important if there are systems highly sensitive to its initial conditions (i.e. value of N_0). In such eco-systems extinction of some of the species may be highly probable for certain vectors N_0 but unlikely for others. It will be of great interest to ecologists to find out what combination of eco-system characteristics can result in such initial condition sensitivity. Since catastrophic events (e.g., volcano eruption, meteoric impact) can drastically change initial condition of an eco-system, existence (or nonexistence) of initial condition sensitivity is of importance in studies of extinction.

Our computational research resulted in novel parallel algorithms for problems created in implementing the discussed models. In particular we have researched the data and operations distribution among the processors to minimize execution time. In the future, our research will also investigate parallel code generation techniques that will customize the generate code to the required input parameters to avoid run-time condition execution that is expensive in SIMD architectures.

The future simulations will be based on more complex interspecies interactions than most analyses of spatial pattern in ecological dynamics offer. Our results will be directed toward hypotheses allowing experimental tests; this is perhaps the most important potential benefit to ecologists.

References

[1] Andrews G. 1991. *Concurrent Programming: Principles and Practice*. Benjamin/Cummings Publishing Company, Inc. Redwood City, CA.

[2] Baase, S. 1988. *Computer Algorithms: Introduction to Design and Analysis*. Addison-Wesley Publishing Company, Inc. Reading, MA.

[3] Bailey, N.T.J. 1975. *The Mathematical Theory of Infectious Diseases and Its Applications*, 2nd Edn. Charles Griffin, London.

[4] Bartholomew, D.J. 1983. Stochastic Models for Social Processes, 3rd Edn. Wiley: New York.

[5] Burks A.W. 1968. *Theory of Self-Reproducing Automata, John von Neumann*. University of Illinois.

[6] Cypher, R. & Sanz, J. L. 1989. Hypercube and Shuffle-Exchange Algorithms for Image Component Labeling. Journal Of Algorithms 10:140-150.

[7] Cypher, R. & Sanz, J. L. 1990. Algorithms for Image Component Labeling on SIMD Mesh-Connected Computers. IEEE Transactions on Computers, 39:277-281.

[8] Hastings, A. 1990. Spatial heterogeneity and ecological models. Ecology 71:426-428.

[9] Hochberg, M.E., M.P. Hassell & R.M. May. 1990. The dynamics of host-parasitoid-pathogen interactions. American Naturalist 135:74-94.

[10] Kareiva, P. 1990. Population dynamics in spatially complex environments: theory and data. Philosophical Transactions of the Royal Society, London, B, 30:175-190.

[11] Krummel, J. R., Sugihara, G., O'Neill, R. V., & Coleman, P. R., 1987. Landscape patterns in a disturbed environment. Oikos, Copenhagen, 48: 321-324.

[12] Leighton, F. T. 1992. pp 201-210, 267-268, 339, 382, Introduction To Parallel Algorithms and Architectures: Arrays · Trees · Hypercubes, Morgan Kaufman Publishers: San Mateo, California.

[13] Maniatty, B., B. Sinharoy & B.K. Szymanski. 1993. Efficiency of Data Alignment on MasPar. SIGPLAN Notices, **28**, 1:48-51.

[14] Murray, J.D. & W.L. Steward. 1992. On the spatial spread of rabies among foxes with immunity. Journal of Theoretical Biology 156:327-348.

[15] Reed, D.A., Adams, L.M. & Patrick, M.L. 1987. Stencils and Problem Partitionings: Their Influence on the Performance of Multiple Processor Systems. IEEE Trans. on Computers C-36:845-58.

[16] Schall, J.J. 1983. Lizard malaria: parasite-host ecology. pp 84-100, in *Lizard Ecology: Studies of a Model Organism* (ed by R.B. Huey, E.R. Pianka & T.W. Schoener). Harvard University Press:Cambridge.

[17] Shiloach, Y. & Vishkin, U. 1982. An $O(\log n)$ Parallel Connectivity Algorithm. Journal Of Algorithms, 3: 57-67.

[18] Schall, J.J. In press. Parasite-mediated competition in *Anolis* lizards. Oceologia.

[19] Sinharoy B. & B.K. Szymanski. In review. Data alignment in SIMD machines. IEEE Trans. Parallel and Distributing Programming, also Technical Report 91-9. Rensselaer Polytechnic Institute. Department of Computer Science, 1991.

[20] Szymanski B.K., B. Maniatty & B. Sinharoy. 1993. Simultaneous Parallel Reduction. Rensselaer Polytechnic Institute. Tech. Rep. CS 92-31, also submitted to Parallel Programming Letters.

[21] Sugihara, G., & May, R. M., 1990. Applications of Fractals in Ecology, Trends in Ecology and Evolution, New York, Volume 5, Number 3: 79-86.

[22] Turner, M.G. 1989. Landscape ecology: the effect of pattern on process. Annual Review of Ecology & Systematics 20:171-198.

[23] Yachi, S., K. Kawasaki, N. Shigesada & E. Teramoto. 1989. Spatial patterns of propagating waves of fox rabies. Forma 4:3-12.