Abstracts: Undergraduate Research in Biomathematics
April 21-22, 2006, SUNY Geneseo

KEYNOTE ADDRESS

Kasiyanov, Alexander. Agroterrorism in the US: Security Challenge for the 21st Century. Exypnos Research Institute, Omaha, NE

Recent events of war and terrorist attacks have produced an increasing awareness of potential agroterrorism in the United States. Agriculture is one of the easiest sectors of the U.S. economy to disrupt, and its disruption could have catastrophic consequences for the economy. A number of different animal pathogens could cause loss of production or the death of livestock. An act of agroterrorism that does not result in the destruction of the food supply industry could have a psychological impact on people. This presentation gives an overview of dangerous diseases and practical mathematical applications for the prediction of disease distribution between humans and animals.

Callear, Christina, and Anthony Macula. Optimization and expansion of an approach to group testing. SUNY Geneseo.

Group testing; or pooling; is a widely used general procedure applicable whenever a large group of objects is to be subjected to the same test. In many cases; a number of objects or complexes of the objects to be tested will produce an undesirable result. Group testing offers an advantage over individual testing when the number of these target objects or complexes is small in comparison to the size of the group as a whole. Group testing has proved useful in a range of fields including medical diagnostics; industry and most recently in working with DNA. This work addresses the development; optimization and expansion of a probabilistic group testing method that leads to the identification of portions of cohorts or complexes that collectively produce an observable result.

Chynoweth, Mark, and Gregg Hartvigsen. Estimating wildlife home ranges with geographic information systems. SUNY Geneseo.

An animal's home range is the area generally used by an animal in its daily activities. These home ranges are useful to understand when conducting studies on specific species or addressing issues that wildlife scientists experience in national parks and on other public lands. Because these areas normally grow over time our estimates of home range sizes need to include estimates of variability. There are presently several methods available to describe the spatial extent of an animal's use of its range. However; most methods require extensive data collection of animal positions that can often be time consuming and costly. Using Geographic Information Systems; we estimated the asymptote of the function describing the increasing area used by three grizzly bears. In addition we used subsets of simulated random walk data sets to estimate changes in our confidence in home range sizes. Our estimate of the grizzly bear home ranges was relatively robust to the removal of sample data. As we continued to remove location points; home range estimates remained similar. Our method allows wildlife scientists to estimate animal home ranges with fewer location points than previous method and provides confidence levels for these estimates.

Darling, Michael, and Dr. Cheri Boyd. Why PAM works: an in-depth look at scoring matrices and algorithms. Nazareth College of Rochester.

We all know that PAM matrices and scoring matrices are used in pairwise sequence alignment; but do we know why they work? This talk will focus on the mathematics behind the many PAM matrices; the log odds matrices; and how they are used in scoring algorithms.

Darrow, Brett. A fitting approximation. SUNY Geneseo.

Even though computers can do much of the computational work for us; some problems are still computationally complex. One practical way to reduce the complexity of a problem is to use an approximation method. Approximation methods can be important if the simulation of a complex system is desired. In this talk; we discuss the fitting of output of one approximation method with complexity O (n2) to the output of a more complex (O(n3)) and more accurate approximation method that predicts the thermodynamic stability of DNA fragments. The goal is to then extrapolate the less complex method in the hopes of practically improving its accuracy without making it more computational complex.

Dresch, Jacqueline. The largest component in subgraphs of circulant-like graphs. SUNY Geneseo.

Circulant graphs have recently been used as starting points in the construction of small-world networks used to study disease dynamics. We began our study by taking random induced subgraphs of circulant-like graphs; representing the population susceptible to a particular disease after random vaccination. Then; by computing the expected size of the largest component in these subgraphs we obtained an upper bound on the number of individuals who could potentially contract the disease after one infection.

FitzGerald, Daniel, and Gregg Hartvigsen. The dynamics of cooperation in small-world networks. SUNY Geneseo.

Cooperation between individuals in both human societies and biological systems has long been studied. We introduce a new model of cooperation that tracks the probability that each individual carries of cooperating with another as they interact in the confines of a small world network. We then study the coexistence of clusters of cooperators and defectors that emerges in a small but intriguing portion of parameter space.

Hirschbeck, Sarah, Christian Volk, and Feizabadi Mitra Shojania. A biomathematical approach for investigating the evolution of tumors during a course of chemotherapy. Canisius College.

Links established between biology and other disciplines including computer science; physics; and mathematics have provided powerful tools to pioneer the study of biological systems in unique ways. One of these tools is mathematical and computational modeling of a biological system which provides an additional method to investigate real conditions in a systematic manner. The growth and control of tumors have long been subjects of medical and scientific scrutiny. One can apply physics and mathematics to analyze the evolution of tumor cells; as a fundamental step in understanding and improving treatment. The Webb-Gyllenberg and Gopmertzian models of tumor growth explain how a

tumor behaves as it ages. Though these two models both look at tumor growth over time; they do so from different perspectives. The Gopmertzian model investigates the behavior of the total cell population over time where the Webb-Gyllenberg model divides the tumor into two parts; the proliferating subpopulation and the quiescent cell population; and evaluates the kinetics of each subpopulation. In the model presented by Kozusko-Bajzer ; the above models are combined and analytical solutions for the evolution of proliferating cells and quiescent cells are obtained. In this work; the concepts of the Kozusko-Bajzer Model are presented. The modification for this model in the presence of a cell-cycle specific anti cancer drug (considered with an "on-off" type behavior) is introduced and the evolution of each subpopulation is simulated numerically.

Kyu, Shuya. Specialized herbivore feeding leads to increased speciation in plants. SUNY Geneseo.

We explore an agent-based; evolutionary; plant-herbivore model to understand how the presence of the herbivores influences the speciation in plants. Each organism carries a 31-bit interaction code; which serves as its genome and determines the organism's performance in intraspecific (mating) and interspecific (feeding) interactions. A species is defined as a collection of individuals that are able to mate with at least one other member in the species. The model was run for 1000 generations and the number of species that evolved and the rate of evolution of the dominant species (the movement of the center of mass of the species in sequence space per one generation) were analyzed. We found that specialized herbivore feeding led to an increase in plant speciation. The presence of herbivores also increased the rate of evolution of the dominant plant species.

Kremer, Colin, Kate Huggler, Gregg Hartvigsen, and Gary Towsley. Investigating control methods for *Dendroctonus rufipennis* outbreaks using computer modeling. SUNY Geneseo.

The spruce bark beetle, *Dendroctonus rufipennis,* is native to spruce forests (*Picea* sp.) across North America. Adult beetles primarily attack fallen spruce trees, especially white spruce, but at high densities they will also colonize live trees. Females bore galleries into the tree's bark and deposit eggs. Both adults and larvae damage the tree's phloem region, potentially resulting in the tree's death. Outbreaks are usually instigated by sharp increases in dead trees, caused by blow downs, forest fires, or biomass accumulation in fire suppression regions. Severe outbreaks of *D. rufipennis* harm forest ecosystems and adversely impact the timber industry, motivating the search for methods effective in limiting outbreaks.

We have developed a stochastic, spatial computer model that approximates the spread of *D. rufipennis* through spruce forest patches of varying density and composition. The probability of beetles spreading from one tree to a neighboring tree is:

$$P = \left(1 - \frac{N}{K}\right)\left(\frac{(\alpha + \beta S)D_t}{D_{\max}}\right)$$

where N is the population of beetles in the neighboring tree and K and D_t are the tree's carrying capacity and diameter based on data from the USDA Forest Service. D_{max} is the diameter of the largest tree, α is the probability of colonization faced by the largest unoccupied tree globally, and β provides the increase in colonization likelihood due to the tree being dead. S gives the state of the tree, alive or dead, 0 or 1 respectively.

Using our model we are investigating the effects of two management methods, the creation of trap trees and the removal of infested trees, on controlling *D. rufipennis* outbreaks; preliminary results indicate both methods are effective.

Marcus, Daniel, and Gregg Hartvigsen. The spread of influenza through a multi-city small-world network. SUNY Geneseo.

A description of the spread of influenza through a small world network which has been expanded to handle multiple cities. A small world network; introduced by Watts and Strogatz; is a network which has a parameter p which affects the randomness of the network. Each city is in itself a small world network of people; and the whole model is a small world network of cities. The model changes the dynamics of the epidemic curve of a single small world network by allowing the transition of influenze from one city to another.

McCarthy, Andrew, and Gregg Hartvigsen. The effect of network structure on influenza evolution. SUNY Geneseo.

A small-world network was used to model the spread of influenza through a population. To build our networks we created a circulant lattice with vertices having a degree of 16 with a probability $0 <= swnP <= 1.0$ that each edge is rewired to a randomly chosen vertex in the graph. In this model each strain was represented by a 31-bit binary number which can mutate during each new infection. Mutations can result in new strains. After infection hosts cannot become re-infected. If, however, the virus mutates (we tested this from one to four bits) then hosts become susceptible to the new strain. We found that the largest number of hosts infected and the largest number of strains evolved was positively related to *swnP* when R_o was relatively low ($1.0 <= R_o <= 2.0$). With higher R_o values (rapid epidemic spread), however, we found that the most strains evolved when $swnP = 0.05$.

Nimmo, Kayla, and Lauren Wood. Detection of mispaired oligonucleotides using SYBR green I fluorescence. SUNY Geneseo.

We are constructing a collection of single-stranded DNA sequences (a DNA library) for DNA computing; micro-arrays and nanotechnology. Since these sequences must not cross-hybridize with each other; it is important to have assays to detect the presence of mismatches. We designing such an assay using SYBR Green I; a dye whose fluorescence increases when bound to helical DNA . We have investigated detection limits of our method using single-stranded 16-mer oligonucleotides and have optimized concentrations and conditions. We have detected one cross-hybridization event in a pool of sixteen strands.

Priore, S., C. Schiano, W. Pogozelski; W., and A. Macula. Molecular computation: a DNA-based model for solving mathematical problems. SUNY Geneso.

DNA computing is a rapidly growing field that utilizes DNA's natural potential to spontaneously assemble and store information. DNA is also easy to manipulate which makes it adaptable to a wide range of applications. DNA computing uses this potential to solve computationally difficult mathematical problems. Creating a functional library of combinatorial strands of the required length is a prerequisite for a DNA computing model. This task is accomplished by using a computer algorithm that produces DNA sequences that will bind only to their complements. Validation of this algorithm

was performed on test strand using SYBR Green I dye and a Real-time PCR machine to measure florescence. After validation was completed construction of the library was able to begin. We used a DNA polymerase-based method to combine ten unique; 16 base pair; single-stranded oligonucleotides; to from a library of 32 double-stranded molecules of length 80 base pairs.

Quinzi, David, and Garrett Jones. Making computers do computational work for us. SUNY Geneseo.

The increasing amount of research being done in silico has made providing researchers access to computational research tools more important than ever. We will present our methods and examples of providing broad simplified access to in silico computational research. These methods can be used to make use of idle processor time on servers as well as provide a simple user interface for research applications run on distributed systems.

Reynolds, Sara. Evolutionary relationships: a mathematical concept? Nazareth College.

This presentation will explore the mathematics behind determining the evolutionary relationships amongst DNA sequences. We will then discuss how the resulting data is used in the formation and evaluation of phylogenetic trees.

Saracco, Jeff, and Garrett Jones. Distributed implementation and visualization of small-world networks. SUNY Geneseo

A program that models the spread of influenza in a SWN (Small World Network; a mathematical description of a population) is expanded to incorporate multiple cities. Since running a single computation of SWN becomes taxing on a computer's resources; a new optimized implementation is written with the expansion of data structures to accommodate multiple cities on multiple computers. The revised algorithm takes advantage of cluster computing. The number of SWN's is only limited to the number of available cluster nodes; each SWN is allocated its own computer. The program allows communication between multiple SWNs. A visualization of multiple SWN's has also been achieved; such that the spread of infection can be viewed; a distributed approach to visualization is also being pursued.

Shkalim, Sara, and Susannah Gal. Solving Satisfiability problems using DNA methylation. SUNY Binghamton.

DNA computing is a relatively new technology that has the potential to surpass the capabilities of today's supercomputers. DNA computers are many times smaller than today's computers; while at the same time are capable of holding more data. DNA computers can also perform calculations in parallel. This allows us to solve complex mathematical problems in hours; whereas it might take electrical computers years to complete. With its ability to store more data and process information at a quicker rate; it is possible to solve intricate mathematical problems. Satisfiability problems (SAT) are Boolean expressions that can be solved using DNA. We have set up an experiment using methylases and restriction enzymes to solve a 3-variable SAT problem. Methylases transfer a methyl group to a specific DNA residue which in some cases; blocks cutting of DNA at that site by a restriction enzyme. Thus we can convert a "1" (cuttable) to a "0" (not cuttable) on the DNA using DNA meythlases. By doing a chain of reactions involving separately methylating and then cutting the methylated DNA; we can solve a SAT problem; where only one answer could exist depending on the clauses established.

While there are limitations to such a design; it is possible that methylation is a valuable way to efficiently use and create DNA computers.

Stallard, Cynthia, and Anthony Macula. Duh^3:Doubley Dynamic DNA Programming. SUNY Geneseo.

The Maple programming language was used to create a program that generates DNA sequences based on certain stipulations input by a user. These include the probability of selecting a specific base at each entry in a sequence and the score of the strand as compared to previous strands that were selected. The strands are compared using scores of weighted stacked pairs and subsequently using a dynamic threshold created by the program and based user inputs. The standard approach is to use a static threshold; but use of a dynamic threshold makes better sense when one considers statisical thermodynamics.

Vyacheslav Rykov and Vladimir Ufimtsev. Efficiency of the Two-Stage Group Testing Algorithm for DNA Library Screening. University of Nebraska at Omaha

Group testing is concerned with finding a small number of defective units among a large population; efficiently. We present a two-stage group testing algorithm that is aplicable to DNA library screening. The efficiency of two- stage group testing is determined by the expected number of tests required to establish the random number of positive clones in a DNA library; which are distributed by the Poisson distribution with a given parameter mu. We will show that the expected number of tests N required to establish the positive clones; for a large size of library t asymptotically approaches mu*log(e)*log(t)*(1 + o(1)); if mu grows as ln(t).

Sponsored by the Biomathematics Career Initiative Program @ SUNY Geneseo and funded by the National Science Foundation
Gregg Hartvigsen, Bailey 4, 585.245.5448, hartvig@geneseo.edu
Tony Macula, South 325D, 585.245.5482, macula@geneseo.edu