CALIFORNIA STATE SCIENCE FAIR
2010 PROJECT SUMMARY

Name(s)

Tushar Alladi; Parag Kikla; Arun Kuchibhotla

Project Number

S1601

Project Title

Real-Time Markerless Hand Computer Interaction

Abstract

Objectives/Goals

In present day, the computer mouse is the prime method of interacting with a computer. Though effective, the mouse's single point of interaction limits the user from obtaining complete control. One should have the freedom to control applications as if one were using one's own hands. This type of freedom is presented with the concept of multi-touch; multi-touch users have the freedom to manipulate and interact with the applications with the aid of multiple points of contact. The downside of multi-touch is that the software, as well as the hardware limits the user. This type of ease should be available for users without the need to worry about the expense. Touch-less interactivity is considered to be the next-generation form of communication for computer users and therefore new approaches are being introduced into the market constantly. Multi-touch will enable computer programs to be more complex function-wise without taking away from the simplicity the user experiences. Our aim is to take that functionality that multi-touch provides, without the "touch," and make it affordable for everyone.

Methods/Materials

We started with a very simple approach that only required a camera (proliferating on all laptops and desktops) and computer software coded in Visual C++. Since the software was the driving force behind the project, we had to heavily research algorithms since we had no previous knowledge of their functionality. After analyzing over 20 algorithms and trying each one using a brute-force method, we discovered over time that the best way to approach this was to combine multiple algorithms.

Results

It is necessary to have a good camera (high Frames Per Second) for this program to work. The program tracks a contrast in color and uses the tracked contrast as a binary 1 (positive) and everything else as 0 (negative). If the code were to track a shape rather than a color, it may be more effective in helping the camera understand the contours of the hand and it will not delay as much. The results indicate an optimal position, lighting, and color for the current program.

Conclusions/Discussion

There is a huge impact that the environment has on the program, it won#t work as well with constantly varying background colors and will not work at distances over five feet meaning a background independent code would be preferable. In order for our project to have more potential, the method for tracking must change.

Summary Statement

The design and construction of a computer program to aid a user to control and manipulate applications on their personal computer by using the movements of their hands in 3-D space.

Help Received

CALIFORNIA STATE SCIENCE FAIR
2010 PROJECT SUMMARY

Name(s)	Project Number
Ryan P. Batterman	**S1602**

Project Title

Malware Identification by Statistical Opcode Analysis

Abstract

Objectives/Goals
This project determined the efficacy of statistical analysis of program assembly instruction (opcode) frequencies to identify Malware from Goodware.

Methods/Materials
Malware and Goodware binaries were obtained and a python script was created to extract opcode frequencies from specific parts of these files. Naive Bayes models and Kmeans based models were then trained using these executables. These models were tested using a different set of programs to determine their efficacy at identifying Malware from Goodware.

Results
The best Naive Bayes model had a recall of 1 for Malware and .8 for Goodware.

Conclusions/Discussion
Differences in opcode frequencies can differentiate Malware from Goodware. Certain instructions occur much more frequently in one group than in the other; these differences can be used to identify the two types of programs.

Summary Statement

This project examines models that differentiate Malware from Goodware using the frequencies of program assembly instructions.

Help Received

Communicated with mentor Joshua Kroll ; Pamela Durkee proofread papers and guidance

Name(s)	Project Number
Dylan Freedman	**S1603**

Project Title

A Novel Approach to Text Compression Using N-Grams

Abstract

Objectives/Goals

This project investigated a novel approach to text compression using principles of natural language. After observing traditional algorithms compress small text files and yield inefficient results due to a lack of useful redundant content, I wondered whether compression exploiting redundancies inherent in human language could offer any advantages.

Methods/Materials

To implement this concept, I first obtained a copy of Google's English N-Gram database, a comprehensive linguistic model for examining how often commonly observed sequences of words occur. To extract useful information, I optimized this database and sorted it alphabetically and by frequency so that information could be retrieved efficiently through a binary search. I then wrote an undemanding program able to quickly deduce the relative probability of a word occurring given a few preceding words as context. Compression was achieved by first converting each word of an input file into a ranking sorted by the word's respective probability of occurrence compared to other words that could have occurred. Then, preexisting compression algorithms were applied to the resultant rankings of the encoded file.

Results

This algorithm significantly outperformed multiple existing compression algorithms, working particularly well in comparison with other methods when compressing small English text files. Even files less than a few hundred bytes compressed to an average of 25% of their original size, an unprecedented ratio.

Conclusions/Discussion

With increased global dissemination of small text files such as cell phone text messages, emails, and chats, this method implemented properly could significantly reduce the environmental strains of data storage worldwide.

Summary Statement

My project investigates an effective, novel natural language approach to small text compression using n-grams from a comprehensive linguistic database

Help Received

CALIFORNIA STATE SCIENCE FAIR
2010 PROJECT SUMMARY

Name(s) **Tony Ho; Ritik Malhotra**	Project Number **S1604**

Project Title

Novel Genetics-Based Early Disease Detection Using Ontology-Driven Microarray Semantics with Optimized Mass Spectrometry

Abstract

Objectives/Goals

Doctors currently use symptoms to determine whether certain diseases are present, which is prone to error and subjectivity. Genetics allow for earlier and accurate diagnosis of diseases. Microarray experiments hold data describing genes that cause specific diseases, but current databases are disorganized and lack semantic enforcement. Mass spectrometry techniques to analyze tissue samples exist as a viable means to use data from microarrays to find diseases, but the current mass spectrometers aren't sensitive enough to find the diseases at low concentrations, before they become lethal. The purpose of this project is to identify key biomarkers for diseases by detecting the genetic changes caused by diseases. First, an efficient database to query microarray experiments without inaccurate data is needed. Second, sensitivity in mass spectrometers needs to be improved in order to facilitate the early detection of low abundance biomarkers.

Methods/Materials

A cluster of 8 quad-core 64 bit servers and an Agilent G6460 triple-quadrupole mass spectrometer with Agilent MassHunter software were used. First, a MySQL database using text mining techniques for the microarray experiment data was designed. Second, the inaccurate data was removed from the database by developing two novel algorithms: KNN-Delta and Semantic Outlier Factor. Third, thermal gradient focusing technology in the mass spectrometer for ion confinement was implemented. Fourth, conditions inside the mass spectrometer were optimized for efficient ion transmission. Fifth, the ion scattering within the electrospray ionization unit of the mass spectrometer was analyzed to detect areas of high densities. Finally, an inlet capillary to capture ions in the high density areas was designed.

Results

The microarray database was successfully created and was able to automatically annotate over 400,000 experiments using common semantics. The KNN-Delta and Semantic Outlier Factor algorithms were able to increase the accuracy of the database, removing 40,000 inaccurately annotated experiments. The mass spectrometry efficiencies resulted in a signal-to-charge ratio increase from 1000:1 to 6800:1, detecting samples as low as 150 femtomoles.

Conclusions/Discussion

Based on the successful implementation of the microarray database and the sensitivity improvements of the mass spectrometer, it is possible to detect diseases at small concentrations, before they become lethal.

Summary Statement

By developing novel algorithms for a microarray database and improving the sensitivities of a mass spectrometers, we are able to facilitate the early detection of diseases, before they become lethal.

Name(s)

Dhiraj R. Holden

Project Number

S1605

Project Title

An Analysis of the Primitive Cycles Existence Conjecture

Abstract

Objectives/Goals
The objective of this project is to make progress toward the proof of the Primitive Cycles Existence Conjecture. Also, the project intends to present an analysis of total stopping time graphs for 3x+d and an application of the 3x+d function to cryptography.

Methods/Materials
The number of iterations k takes before the kth iteration is equal to the k+nth iteration for any n for any k, i.e. total stopping times of the 3x+d function, were analyzed using a Java program to find stopping times for 1 to 9999 for d = 1,5,7,11,13,17 and plotted it.

Results
The first theorem details conditions for a number divisible by a number of a certain form that is necessary for it to be a primitive cycle, and the second theorem builds on the first theorem to determine under what conditions a possible cycle can exist. These cycles are a subset of all cycles for all d.

Conclusions/Discussion
The resulting graph demonstrated a logarithmic relationship between the number and the stopping time. Also for further research, these theorems may be generalized to assist in proving the Primitive Cycles Existence Conjecture.

Summary Statement

This project conducts an analysis of the Primitive Cycles Existence Conjecture concerning a generalization of the 3x+1 problem to 3x+d.

Help Received

Mother looked over report and abstract; Father helped with poster formatting and also looked over report; Dr. Haxell critiqued theorems.

Name(s)	Project Number
Kaavya N. Jayram	**S1606**

Project Title

Crank 0 Partitions and the Parity of the Partition Function

Abstract

Objectives/Goals
The formal properties of integer partitions have been investigated for over 200 years by some of the brightest minds in mathematics such as Euler, Hardy, and Ramanujan, with surprising applications to modern physics and computer science. The partition function p(n) denotes the number of ways in which an integer n can be written as an (unordered) sum of other integers. Motivated by Ramanujan's investigations into the modular properties of p(n), this project aims to make progress on the parity problem of p(n) by means of deriving generating functions for cranks and ranks.

Results
Berkovich and Garvan (2002) showed that there is always a bijection between the crank k and crank -k partitions of n for every k>0. Consequently, the parity problem for p(n) reduces to studying crank 0 partitions. I obtained the following results:

(1) I derived a generating function for crank 0 partitions of n, which is similar to a generating function for p(n). I also obtained a general form for the crank k generating function.

(2) I described an involution on crank 0 partitions of n, whose fixed points are called invariant partitions. I then derived a generating function for crank 0 invariant partitions.

(3) Finally, I derived a generating function for rank 0 self-conjugate partitions.

Conclusions/Discussion
The proof techniques are based on identifying and manipulating the key combinatorial objects underlying cranks and ranks, and avoid the analytic techniques inherent in previous methods.

A paper describing the above results has been accepted to the International Journal of Number Theory.

Summary Statement

I derived generating functions for the crank and rank of an integer partition to make progress on the parity problem for integer partitions.

Help Received

Dr. Laurens Gunnarsen was my mentor for the project.

Name(s) **Meredith P. Lehmann**	Project Number **S1607**

Project Title

Transportation Networks and the Propagation of Novel H1N1 Swine Flu-like Epidemics

Abstract

Objectives/Goals

The recent SARS and novel H1N1 pandemics have rekindled interest in the spread of epidemics. Recently, Colizza et al. simulated epidemics on a 3100 airport network across 220 countries. Their analysis shows that epidemics propagate primarily through large hub airports as preferred epidemic pathways. This study uses different US air travel data and, for the first time, long distance auto travel data to examine three main hypotheses suggested by their work: (H1) County travel flows per capita should be increasing in population size and highly connected counties should experience higher average county to county traffic flows; (H2) County infected evolution should differ markedly from that of a Susceptible-Infected-Removed model without travel; and (H3) Air traffic should be the main force driving epidemics.

Methods/Materials

The model is a network comprised of all 3076 counties in the continental US. Each county is a separate node in which epidemics propagate according to an SIR model subject to daily air and auto travel inflows and outflows. Roughly 1 billion long distance auto and 262 million air round trips from county-to-county comprise the travel model. Epidemic parameter values and initial conditions are based on swine flu data from spring 2009. This large scale simulation model is used to simulate US epidemics under different transportation network configurations.

Results

These simulated epidemics behave nothing like the Colizza et al. simulations. Long distance auto travel is the main driver of epidemics contra H3. In contrast to H2, travel essentially disperses infecteds across the country and county infected evolution behaves like an SIR model without travel once the incidence of infection reaches $1/100,000$. Per capita travel is not higher in large counties contra part one of H1 but more connected counties experience disproportionately larger travel flows consistent with the second part. However, such hub counties are not preferred epidemic pathways in the simulations.

Conclusions/Discussion

More effective travel restriction and quarantine policies and targeted pharmaceutical interventions can be designed when large hub airports are the primary epidemic pathways through which epidemics propagate but there are no preferred epidemic pathways if these simulations are correct. The social importance of the resulting inability to materially alter epidemic outcomes in this fashion cannot be overstated.

Summary Statement

This project shows that long distance auto, not air, travel is the main epidemic propagation mechanism in the US and that no travel mode is a preferred epidemic pathway that can be used to improve prediction and policy interventions.

Name(s)	Project Number
Jonathan F. Li	**S1608**

Project Title
Effects of Motility and Contact Inhibition on Tumor Viability: A Discrete Simulation Using the Cellular Potts Model

Abstract

Objectives/Goals
We studied the effect of motility and contact inhibition on tumor growth using mathematical modeling.

Methods/Materials
Based on the Celluar Potts Model, tumor cell properties such as motility, contact inhibition, volume constraints, and adhesion energies are incorporated in my model of tumor growth. The simulation is coded in COMPUCELL3D with Python. Conditions influencing the contact inhibition, motility, and compressibility of cells were changed for different data sets. MATLAB is used to generate growth constants using generalized logistic curves. To streamline this process, I created a macro so that the simulation could be run with minimal human presence. About 100 sets of data (4 GB) were recorded.

Results
The data closely followed a generalized logistic curve, initially growing exponentially and then leveling off at carrying capacity. Plots of volume vs. time graphs revealed that low motility cells had a slower growth rate. In addition, cells that exhibited a high response to contact inhibition showed the same pattern of a slower growth rate and even amplified the effect of motility. Also, less compressible cell lines exhibited faster growth. These observations were quantitatively confirmed by the growth constants of curves fitted to the data.

Conclusions/Discussion
Our results showed that higher migration leads to faster growth rates. Motile cells create space for cells in the center of the tumor to grow and proliferate. We also observed that cells with high response to contact inhibition amplified the effect of motility. By increasing the contact inhibition restraint, clustered cells were less likely to proliferate. Cells in the center of the tumor were completely surrounded by homotypic cells and thus were not able to grow.

These results call into question the effectiveness of chemotherapy. Some treatments administer drugs that kill healthy and cancer cells indiscriminately. When those cells die, both cell lines must compete for the newly created space. Our research shows that cells with high motility and less response to contact inhibition will have increased proliferation rates. Since cancer cells often have these attributes, chemotherapy may in fact speed the growth of tumors in certain cases. The results from the simulation can be used to identify specific attributes that are associated with high growth rates and lead drug production to target those characteristics.

Summary Statement
We developed a mathematical model for tumor growth using the Celluar Potts Model and studied the effect that motility and contact inhibition on tumor viability.

Help Received
Dr. John Lowengrub of the University of California, Irvine is my mentor.

Name(s)	Project Number
Randall Lin	**S1609**

Project Title

Adaptive Radiation by Asexuals Driven by Resource Competition in silico

Abstract

Objectives/Goals

Sympatric speciation faces much controversy in evolutionary biology, whereas allopatric speciation is widely accepted. We demonstrate that sympatric speciation based only on resource competition is possible for certain parameters in computer simulations. Every generation a set number of individuals are killed, followed by asexual reproduction of the remaining organisms based on their fitness, which is dependent on how well they utilize the available resources. Finally, every offspring has a chance of mutating, which changes its affinity to a resource. This simulation results in diversification of genotypes. We confirm speciation through the existence of at least one resource dominance per existing genotype, and run invasion experiments to confirm that negative frequency-dependent selection maintains these species. Our results verify that sympatric speciation is possible through competition for limiting resources alone and that cost decides the resulting equilibrium.

Summary Statement

Our results verify that sympatric speciation is possible through competition for limiting resources alone and that cost decides the resulting equilibrium.

Help Received

Name(s)

David C. Liu

Project Title

Continual Adaptation of Acoustic Models for Domain-Specific Speech Recognition

Abstract

Objectives/Goals

Advances in automatic speech understanding bring a new paradigm of natural interaction with computers. The Web-Accessible Multi-Modal Interface (WAMI) system developed by MIT provides a speech recognition service to a range of lightweight applications for Web browsers and cell phones.

However, WAMI currently has two problems. First, to improve performance, it requires continual human intervention through expert tuning--an impractical endeavor for a large shared speech recognition system serving many applications. Second, WAMI is limited by its global set of models, suboptimal for its variety of unrelated applications.

Methods/Materials

In this research I developed a method to automatically adapt acoustic models and improve performance. The system automatically produces a training set from the utterances recognized with high confidence in the application context. I implemented this adaptive system and tested its performance using a data set of 106,663 utterances collected over one month from a voice-controlled game.

To solve the second problem, I also extended the WAMI system to create separate models for each application.

Results

The utterance error rate decreased 13.8% by training with an adaptation set of 32,500 automatically selected utterances, and the trend suggests that accuracy will continue to improve with more usage.

The system can now adapt to domain-specific features such as specific vocabularies, user demographics, and recording conditions. It also allows recognition domains to be defined based on any criteria, including gender, age group, or geographic location.

Conclusions/Discussion

This research has enabled the WAMI system to automatically learn from its users and reduce its error rate. The extended WAMI can create customized models to optimize performance for each application and user group. These improvements to WAMI bring it one step closer towards being an "organic," automatically-learning system.

Summary Statement

I extended MIT's speech recognition system to make it learn on-the-fly as more people use it. The system serves many Web and mobile applications simultaneously. My work brings it closer to being an "organic" and self-learning system.

Name(s)	Project Number
Ian J. Mallett	**S1611**

Project Title

Salmonids by Numbers II: A Predictive Simulation-Analysis of San Lorenzo River Fish Populations

Abstract

Objectives/Goals
I constructed a statistical model that considers streamflow in the San Lorenzo to predict fish populations migrating past Henry Cowell Redwood State Park. I then created, from scratch, a computer simulation that predicts fish populations at any point in their life-cycle based on many of the factors that affect salmonids. I then created a website that lists all the data collected by the Monterey Bay Salmon and Trout Project.

Methods/Materials
-Copies of raw data about the fish, taken daily by many observers from San Lorenzo Valley High School and the Monterey Bay Salmon and Trout Project at the fish trap in Felton during spawning season since 2000
-Data available online from the United States Geological Survey about daily local average stream flow near Felton
-Information on salmonid lifecycles and the factors that affect them, from various sources, including Don Alley et al.
-Fathom statistical software, (student ed.)
-OpenOffice Calc/Writer/Impress
-Microsoft Excel
-VUE mind-mapping software
-TI-89 Platinum graphing calculator
-Python 2.5.4 programming language

Results
The effort to simulate fish populations has been successful. A rough sensitivity analysis of the simulation indicates that the most important factors that affect salmonid migrations are streamflow and sedimentation. The results obtained from the simulation are not more accurate than the statistical model that considers only streamflow. However, the simulation has the advantage of being able to predict fish populations at all stages of their lifecycle. In addition, the simulation could also be recalibrated fairly easily to a wide variety of other rivers and streams.

Conclusions/Discussion
Statistical methods, such as those that I presented in my research last year, accurately predict the relative frequency of the timing of fish populations migrating upstream when the explanatory variable, streamflow, is simple. By simulating the fish populations instead, a reasonable degree of accuracy is

Summary Statement

The project is a summative statistical and simulation analysis of critically endangered fish populations in the San Lorenzo Valley River.

Help Received

Brenda Mallett for formatting and data entry; Jane Orbuch and David Bernick for review and assistance; Don Alley, the MBS&TP, and the California Department of Water Resources for background research; Terry Umstead, with cooperation from the Monterey Bay Salmon and Trout Project for migration census

Name(s)	Project Number
Akshay Nathan	**S1612**

Project Title

A Super-Encryption Standard for Large Data Using Elementary Chaotic Cellular Automata

Abstract

Objectives/Goals

Cellular Automata are arrays of bits that evolve according to a rule. Some automata exhibit chaotic and random behavior which indicates that they have potential for encryption. Many other attempts at building an encryption system have been vulnerable to certain types of attacks. The goal of this project was to create and implement a novel encryption scheme based on cellular automata, and to evaluate its randomness, efficiency, and strength.

Methods/Materials

Each preliminary algorithm was implemented in C and tested using government recommended statistical tests. The final algorithm passed all of the tests multiple times, and exhibited better randomness qualities than some supposedly "true" random number generators. The algorithm was also timed, and growth analysis showed that with optimization, the scheme would be as fast as or faster than industry standard stream ciphers such as RC4.

Results

The final algorithm takes an input of a 3-part key and a plaintext. A unique aspect of this scheme is that the plaintext itself is run through a CA and decrypted through a designed inverting algorithm. The final ciphertext can only be broken knowing all 3 parts of the key.

Conclusions/Discussion

By using super encryption through a repeated sub-algorithm and by using a larger key, the scheme bypassed many of the attacks that are used against stream ciphers today. Although they display very complex behavior, cellular automata operations are very simple, and can be easily integrated into hardware. Additionally, this stream cipher is extremely conducive to parallel processing, making it ready for future computers. The results of this project demonstrate the practicality of cellular automata based stream ciphers by presenting a simple but elegant prototype that is secure and efficient.

Summary Statement

I created an encryption scheme using chaotic and random systems called cellular automata.

Help Received

Name(s)	Project Number
Ivan R. Pyzow	**S1613**

Project Title

Invisible Infinities: Determining the Fraction of Lattice Points Visible from the Origin in the Third Dimension

Abstract

Objectives/Goals
The objective was to determine the fraction of "visible" and "invisble" lattice points from the origin out of the entire set of lattice points in an infinite three dimensional space. I hypothesized that it would involve pi^3, because the known fraction of visible points in the second dimension is 6/pi^2.

Methods/Materials
I examined this problem in the second dimension to give insight into the third. After constructing a sample grid of 10x10 and manually determining visible and invisible point by drawing lines (representing lines of sight) from the origin to those points, I was able to recognize patterns which allowed me to contruct larger graphs with more efficiency. I was then able to construct 3D grids and plot visible points using patterns recognized in two dimensions. I created various representations of three dimensional graphs which I analyzed for patterns which might yield methods for approximating the fraction of visible points.

Results
I discovered that 2D visible points have coordinates (a,b) where a and b are relatively prime. In 3D, visible points have cooridinates (a,b,c) where a, b, c are relatively prime. I was able to reason that the fraction of visible points is 1/(riemann zeta of 3) ~ 0.83 because of the knowledge that the probability of m numbers being relatively prime is 1/(riemann zeta of m). By recognizing the grid of "invisible" points as the summation of an infinite number of grids with patterns, I was able to fashion an approximation of the fraction of invisible points with a series. I compared my series to 1/(riemann zeta of m) for the 2nd, 3rd, and the mth dimension and evaluated the error.

Conclusions/Discussion
My hypothesis was wrong. The fraction that I achieved was 1/(riemann zeta of 3), which is irrational, and based on my research of literature, does not have a representation involving pi^3. However, I am now able to hypothesize that the fraction of visible points in the mth dimension will be 1/(riemann zeta of m). This also suggests (by evaluating the fraction) that in the mth dimension as m goes to infinity, all points would be visible.

Summary Statement

My project deals with methods of determining the fraction of visible lattice points from the origin in the third dimension and their relation to prime numbers, the Riemann Zeta function, and other infinite series.

Help Received

My project advisor suggested books and concepts to research

Name(s)

Justin W. Rajewski

Project Title

Utilizing Artificial Neural Networks to Create a Learning Robot

Abstract

Objectives/Goals
As artificial intelligence continues to improve, artificial neural networks (ANN) are on the cutting edge. An ANN simulates a biological neural network using a mathematical model. The network consists of interconnected neurons. The problem with ANNs is that while they are able to find patterns in a set of data, they need a set of data to be trained on before they are useful. In the real world the data is not always available. The goal was to design a program that would be able to create a set of data for the network to be trained on by making mistakes and using them as experiences.

Methods/Materials
The platform used consists of a simple tank-style chassis, ultra-sonic distance sensor, accelerometer, and control electronics. The code is run on a Gumstix Verdex XL6P Pro which is an embedded Linux computer. It was chosen as it has enough power to run the code in real time.

Results
When the robot is first powered on it only knows how to drive forward. After a few short minutes of driving, crashing, and learning, it is apparent that the robot has learned to navigate around walls and other obstacles. During some sessions the robot performs better than others, it just depends on what solutions it finds and in what order. However, overall the robot performs very well.

Conclusions/Discussion
As these results indicate, this approach for training and running artificial neural networks shows potential in applications where human-like behavior is desired, such as in video game's AI. This method seems to be only limited by the ability to detect failures and to determine what should have happened.

Summary Statement

This project is to explore a method for adding human-like learning to robots.

Help Received

Everything was done without help with a small exception of online help for setting up the Gumstix to run the code on boot.

Name(s)	Project Number
Anna K. Simpson	**S1615**

Project Title

Local Layering of Images with a Natural User Interface

Abstract

Objectives/Goals

Local layering is a new concept giving users the ability to "weave" images as if they are strips of paper or bring regions of interest on several different images to the forefront simultaneously. Conventional stack-based image editing programs (such as Adobe Illustrator) handle a relative depth ordering by compositing a series of images at different layering levels based on a global order. This provides no support for local layering, reduces the possible image complexity, and limits the ability of the user to edit the image. My goal was to create a program to allow users to locally layer images in just one mouse click.

Methods/Materials

I wrote the program in C++ using the Qt graphics library. My program structures the composite image into groups for efficiency in selection and comparison. To relayer, my program locates the nearest intersection of two selected images and then uses integer vector manipulations to reorder the layers. All relayering is done in such a way that the composite remains physically plausible, as if the images were sheets of paper that can be woven but not cut.

Results

My program allows local editing of image layers with only one or two mouse clicks - a significant improvement on both commercial image-editing programs and previous work on the subject. Almost all operations are under real time. Users have several options for selecting and relayering groups using the mouse or keyboard. Unlike anything previously published, the program calculates the nearest overlap of two clicked images and relayers them, making it unnecessary for the user to locate exact intersections.

Conclusions/Discussion

I successfully created a program to locally layer images that will allow users to quickly and easily change image layering among multiple images. An application of my work is in the analysis of scans or microscope slides in medicine or neuroscience, which would give doctors a more complete picture when making a diagnosis. Other applications include compositions of satellite images for military intelligence and manipulations of images in entertainment and engineering.

Summary Statement

I created a program to allow users to quickly and easily edit image depth ordering (layering) in specific locations, a significant improvement over techniques commercially available.

Help Received

Worked in the Computer Graphics Laboratory at the Massachusetts Institute of Technology under Dr. Daniel Vlasic. My research was sponsored by the Department of Defence.

Name(s)	Project Number
Zachary M. Sohn	**S1616**

Project Title

The Effect of Encryption Key Combination on Data Security

Abstract

Objectives/Goals

As data technology advance and computer networking becomes more and more rampant, data protection and intact transit of authentic data become ever more important and challenging. Increasing security risks forces the data industry to respond mainly in two ways: find another way to cleverly hide unpack-able data or use a longer key. My goal is to present an encryption method that is more secure than simply lengthening keys.

Methods/Materials

For this experiment, I wrote both a symmetrical key string cipher, and a brute force algorithm to break it. I tested appending keys for security against my own method, which was to overlap encryption using prime number keys (compound key) for optimum security and compared it to a simple appended key (simple longer key).

Results

Compound key outperformed the appended key in securing the data when tested with a brute force decoding program.

Conclusions/Discussion

Data presented supports my hypothesis that a compound key offer a higher degree of security.

Summary Statement

Method of encryption that will ensure a higher degree of security without making a more cumbersome key.

Help Received

Some help with organization.

Name(s)	Project Number
Eli N. Weinstein	**S1617**

Project Title

Cooperation and Punishment: A Look at the General Phenomenon of Retribution through Evolved Strategies for a Modified Pr

Abstract

Objectives/Goals
The model of the Iterated Prisoner's Dilemma has been studied closely by many researchers in how it applies to the development of cooperation. In sociology this has been extended to how punishment - essentially hurting your opponent at a smaller cost to yourself - works to maintain cooperation. This project takes a more concrete approach, evolving strategies in an effort to see whether retribution is a robust phenomenon like cooperation has been shown to be.

Methods/Materials
This is an entirely computer-based experiment. It uses a cellular-automata based simulation (coded in C++) to model interactions between hundreds of "organisms" playing the Prisoner's Dilemma over and over again. Strategies evolved based on which were the most successful
Data analysis was done in MATLAB, looking at the strategies of many different organisms over time in a number of sophisticated ways using tools developed by the author.

Results
Punishment was not used in any clear strategies for any extended period of time in any of the runs under a variety of conditions.

Conclusions/Discussion
Part of the reason punishment did not arise may have been because of the volatile environment that was created. However, we can conclude with almost complete certainty that retribution is not a robust phenomenon, and does not arise simply from a situation in which players attempt to maintain cooperation.

Summary Statement

This project uses an evolutionary simulation to understand whether the general phenomenon of retribution can be explained easily through game theory.

Help Received

Name(s)	Project Number
Sophia R. Williams	**S1618**

Project Title

World Population Model: War and Disease, Good or Bad?

Abstract

Objectives/Goals
Determine how fast the world's population would grow if there was no violence, war or disease, and find how would that growth would impact the earth's natural resources (natural gas, oil, and electricity) using a mathematical model. The prediction was made that when there were no deaths from war, violence, or disease the population would increase much faster than current population rates and that this larger population would consume oil, electricity and natural gas at a more accelerated rate.

Methods/Materials
The number of deaths and their causes calculated from data from the World Health Organization and The World Factbook were inputted into the Cohort-Component Method which was made efficient by incorporating sub-variables and separating the formula into six formulas determined by sex and age. Using Microsoft Excel, the variables in the equation were experimented with to show how much population growth would change depending on given conditions. The information on the earth's supply of nonrenewable resources data was analyzed to find the average consumption of natural gas, electricity and oil per person. Then a linear function was created where the input variable represents the population.

Results
It was found that the population would grow faster when the appropriate changes were made, which caused an increase in resource consumption. Further investigation, revealed that when the birth rate was equal to the death rate the population decreased slightly when research says that it should stay at a constant rate. This showed a degree of error within the formula.

Conclusions/Discussion
The hypothesis was confirmed. The conclusion was that when there are no deaths from war, violence, or disease the population would increase much faster and would consume natural resources at a more accelerated rate. And although further experimentation revealed error, it also showed that increases in population when there are no deaths from war, violence, or disease can be offset if the death rate and the birth rate within the population were equal.

Summary Statement

Using a mathematical model, this project shows how world population would be affected if there was no war, violence, or disease, and how that change would impact the earth's natural resources.

Help Received

Mom, Karen Miller, helped me edit my work and, along with my teacher, Patricia Williamson, helped me think through my ideas and methodology.

Name(s)	Project Number
Dillon Zhi	**S1619**

Project Title

Computer Modeling of Altruistic Punishment and the Fitness of Cooperative Behavior

Abstract

Objectives/Goals

The purpose of this project was to develop a program in Java to investigate altruistic punishment and its effect on the fitness of cooperative behavior, where altruistic punishment refers to the punishment of non-cooperators, or defectors, at the expense of the punisher. It was believed that the use of altruistic punishment by the cooperators on defectors would significantly increase the relative fitness of the cooperators and encourage greater cooperation overall.

Methods/Materials

The Java program was developed on a computer using the Eclipse integrated development environment, an open source development platform. In the computer model, the simulated individuals played a series of games, the word "games" being in the mathematical, game-theoretic sense. The decisions of the individuals were dictated by pre-programmed logic. In the simulation, players were each given the same amount of "points" (standing for fitness), of which they could "contribute" some to their group or keep some for themselves. Next, only in the version of the game with punishment, players had the opportunity to individually "punish" other members of the group, based on how much they contributed. Fitness was determined by the number of points an individual had at the end of the simulation, and these values were compared between the cooperators and defectors. Besides the ability to punish, a secondary factor tested was the ratio of the number of cooperators to the number of defectors in the population.

Results

In all cases without punishment, the cooperators earned 20 to 25% less than the defectors. With punishment and equal numbers of cooperators and defectors, the cooperators averaged with 2.3% more than the defectors; with punishment and a population of 75% defectors, the cooperators averaged 21.62% less; and finally, with punishment and a population of 75% cooperators, the cooperators averaged 28.2% more.

Conclusions/Discussion

The original hypothesis was supported; in this computer simulation, the use of altruistic punishment greatly increased the relative fitness of the cooperators. With altruistic punishment present, cooperation may be a stable strategy.

Summary Statement

This project used a computer simulation to investigate altruistic punishment and its effect on the fitness of cooperative behavior.

Help Received

My parents helped me get started in learning Java and producing a program like this.

Name(s)

Anastasia L. Reshetikhin

Project Title

The Non-Linear Optics of Interfaces in R^3

Abstract

Objectives/Goals
To model the distortion of an image through a spherical interface.

Methods/Materials
Assuming that the object in question is a photograph of a cat composed of pixels, I used trigonometry, geometry, and Snell's Law to obtain a function (the distortion function) that gives the perceived position of the pixel given the original position of the pixel. Then I used Mathematica (a computer program) to model the distorted image of the cat according to how a fish in a sphere of water would perceive it.

Results
Assuming that the object in question is a photograph of a cat composed of pixels, I used trigonometry, geometry, and Snell's Law to obtain a function (the distortion function) that gives the perceived position of the pixel given the original position of the pixel. Then I used Mathematica (a computer program) to model the distorted image of the cat according to how a fish in a sphere of water would perceive it.

Conclusions/Discussion
Interfaces of different materials distort images. I demonstrated that, using geometry, programming, and laws of optics, one can model the distortion of an image. As a result, one can find the actual location of an object on the other side of an interface using the perceived (or distorted) location. This can be applied to marine observation.

Summary Statement

I modeled the distortion of an image through a spherical interface.

Help Received

My brother introduced me to the Mathematica program. My mother proofread the text of the project.