CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)

Jess M. Bermudes

Project Number

S1201

Project Title

The Need for Speed: Testing the Execution Speed of Today's Popular Computer Languages

Abstract

Objectives/Goals

This project is designed to test the execution speed of popular computer languages such as QBASIC, Java, and C++ and determine which language will execute its code the fastest.

Methods/Materials

Seven different tests were written in each of the three computer languages. Each test was designed to measure execution speed of common programming tasks such as calculating variables or sorting a list. All tests used the system timer in calculating starting and ending times for each program while executing the given task for each individual test. Each test was executed on various platforms to validate my hypothesis of the fastest language.

Results

C++ was the fastest of the three computer languages being tested with Java as a close second with QBASIC in third. C++ and Java were able to consistently execute faster than QBASIC. On some tests, QBASIC took seconds longer than C++ and Java, which were able to complete their tasks in milliseconds.

Conclusions/Discussion

C++ is a computer language that compiles its source code straight into the native language of the machine. Because of this, compiled C++ programs are ready to be executed when instructed to. Java still compiles its source code, but it is first translated into generalized bytecodes which can be interpreted by any computer with the java virtual machine installed, and can then be executed. QBASIC turns out to be the slowest of the three languages because it reads each statement line by line, and then translates the statement into machine code that can be executed by the computer.

Summary Statement

This project tests three different computer languages to see which one can be executed the fastest.

Help Received

Father helped construct board.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Lisa A. Cheek

Project Number
S1202

Project Title
Weight and Number of Raisins per Ounce in a Box of Cereal

Abstract

Objectives/Goals
I wanted to find out how many raisins are contained in either, Vons Raisin Bran,Total Raisin Bran, Post Raisin Bran, Oatmeal Crisp Raisin Bran, Safeway Select Raisin Bran, and Kellogg's Raisin Bran.

Methods/Materials
First, I obtained all of my materials,which included: six different boxes of cereals, scale,and bowl. Then, I set up all the boxes,I opened all the boxes,and then I opened the packages. I placed the cereal into a bowl, then I removed all the raisins. I repeated this step for all boxes. Finally, I divided the total number of raisins by the total number of ounces, then I weighed to find the number of raisin per ounce.

Results
I found out that Total had the most number of raisins, Vons came in second, Safeway Select came in third, Post came in fourth, Kellogg's came in fifth,and Oatmeal Crisp came in sixth.

Conclusions/Discussion
As my hypothesis I thought that Post would come in first, because it had the most fiber, potassium, and iron. As listed above from the results my hypothesis was incorrect, because Total came in first and Post came in fourth.

Summary Statement
How many raisins are contained in different boxes of raisin bran cereals.

Help Received
Uncle hepled take pictures.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)

Charles C. Ciongoli, III; Jonathon Nostrant; Jon Tashman

Project Number

S1203

Project Title

Emulating 3D Spatial Audio

Abstract

Objectives/Goals

The goal of this project was to create a spatial audio system, or a surround sound effect, with only two stereo speakers. The human brain has developed complex auditory cues to decipher the approximate location of origin. The time delay between the arrival of a sound between the left and right ears, known as Inter-aural Time Difference (ITD), is the most significant auditory cue. A secondary goal was to create a phased array of sound. A phased array is multiple sound sources emitting sound with differing time delays to interfere constructively at a point.

Methods/Materials

To complete the project goals, much research on sound, its properties, and its perception, had to be gathered. A program was written in C++ that calculated the ITD time difference, played a sound file, and delayed either the left or right channel so as to hopefully trick the brain into believing the sound source is at a location it is not. To create the "phased array", another program, written with the aid of streaming media framework (GStreamer), was executed on two different computers. ITD calculations will be applied to the "phased array."

Results

On average, four out of five people tested on the ITD model felt the sound source was where the program was trying to "place" it, rather than the speakers right in front of the listener. However, a source could not be "placed" behind the listener. The "phased array" test returned garbled audio and the program abruptly shutdown.

Conclusions/Discussion

It is possible to create a surround sound effect with two speakers, as shown by the successful ITD model, but only in front of the listener; for full 360 degrees of audio around the listener, speakers must be behind listener. This is why the secondary goal of the "phased array" was developed. The "phased array" test has so far been a failure. One possible error could be network traffic. Due to time constraints and no reply to critical questions e-mailed to the developers of GStreamer, the "phased array" is making slow progress, but is still being developed.

Summary Statement

This project is aimed at creating a surround sound system through C++ coding and a "phased array" of sound.

Help Received

Mother helped with board; Computer teacher taught programming and math; Mother and Father helped with driving

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Kyra I. Davis

Project Number
S1204

Project Title
Is Money the Answer to Improving Education?

Abstract

Objectives/Goals
To determine if there is a correlation between money spent on middle school students and their performance in high school, using a microcosm of Tamalpais High and it's three "feeder" middle schools: Mill Valley Middle School, Bolinas/Stinson Middle School, and Martin Luther King Middle School.

Methods/Materials
To make this determination, I first collected and organized all of the students into each grade level (freshmen-senior); and the students of each grade level into the middle school they attended. I averaged the GPA for each category. I used the Chi Square Test (statistical analysis) to determine if there was a relationship between the middle school a student attended and their success (with regard to GPA) in high school. Once I determined that there was a relationship, I compared the average amount of money spent on a child without special needs at a particular school to how thr student from that school performed in high school in terms of GPA.

Results
Through the Chi-Square Test, the probability that there was NOT a relationship between money spent on students in middle school and their GPA in high school was under 1.5%, therefore it could be determined that there was a relationship between the two categorical variables. I discovered that MVMS spent $4,200 per student each year without special need, and Bolinas spent $6,915 while MLK spent $9,000. Freshmen through senior year, the average GPA of a student who attended MLK is a 2.25, the average of student who attended Bolinas was 3.01, while students who attended MVMS average GPA was a 3.25.

Conclusions/Discussion
The results were the inverse of what would be expected. These results taken to the extreme would mean that the best school would be a school without money. This is obviously false. While money is imperative to education in the US for things as basic as classrooms, desk, books and teacher; these results indicate that there may be a factor greater than money in determining a student's success. This factor may be the education of the guardians of the child; at what age the parents began reading to their child, or possibly the family's income. Further research would need to be done to determine what this factor is, but according to my results there is a factor greater than money.

Summary Statement
Analysis of stastical data regarding the relationship between money spent on students in middle and their sucess in high school.

Help Received
none.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Leslie Deckert; Cynthia Macias

Project Number
S1205

Project Title
Computer Simulation of Urban Sprawl in Norco

Abstract

Objectives/Goals
The objective is to determine the effects of sprawl in the next ten years.

Methods/Materials
The material used in this project is the Sim City 3000 simulation. The simulation was used to determine how the city changes over time. For example, the area located on Hamner and Hidden Valley. Ten years ago, this area was bare. All that could be seen was dirt and patches of grass. A few years ago, a Target was put in, followed by a Staples, and Albertson's. Now there is an entire shopping center.

Results
The simulation was used to predict the population growth and city growth in Norco within the years 2005 to 2015. It was determined that sprawl cannot be ceased. Within the years 2005 and 2015, the population could be at an estimated 75,000. 60% of which would be residence. The other 30% would consist of commercial and industrial areas, and recreational areas. There could be a problem with this. The city of Norco is not large enough to contain that population, and would become one of the most polluted cities in the United States.

Conclusions/Discussion
Sprawl is to spread out in an awkward or uneven way. In this case, sprawl is used to describe the state on which the city of Norco is growing. The way it is planned at the moment is spread out in an awkward way. Even though places are in good positions, they are spread out unevenly and far away. These data suggest that city planners plan out better positions to place the commercial areas and also happen to place them in more convenient areas.

Summary Statement
To find out the effects of urban sprawl using computer simulation.

Help Received
Andy McCue at UCR, and Laurie at Norco City Hall provided information on sprawl in Norco.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Arpi S. Emirzian

Project Number
S1206

Project Title
Music, Beethoven, Symphonies... MATH?

Abstract

Objectives/Goals
The objective of this project is to discover if a mathematical pattern can be derived from the repetitive patterns in Beethoven's Fifth Symphony. I hypothesize that if a pattern appears constant, then I can arrive at some kind of mathematical pattern among the successive notes.

Methods/Materials
Using Beethoven's Fifth Symphony (Transcribed for the Piano by Franz Lizst), in Microsoft Works Spreadsheet, I went through each measure and entered the frequency value of each note under "Frequency" (Column B), the count of how long each note was held according to the time signature under "Duration" (Column C), and the difference of the frequencies continuously under "Frequency Change" (Column D). When the "Sequence" (Column A) reached 79, I created a dotted line graph and observed to see if a pattern had formed from columns B, C and D.

Results
The final result was graphs that seemed to show a significant pattern in terms of successive notes throughout sections of the Fifth Symphony.

Conclusions/Discussion
I came to the conclusion that a mathematical pattern does not exist within Beethoven's Fifth Symphony; thought listening to the piece may give the idea that a mathematical pattern may exist within the repetitive chords and sections in the piece. The graphic form of the piece showed no significant mathematical pattern. There were many sources that listed many possibilities of mathematics detected in music. However, to derive the mathematical pattern was no where stated in my research. Each finding was concerned with deriving the frequencies, which I already had.

Summary Statement
I tried to derive mathematical patterns through the constructed graphs of Beethoven's Fifth Symphony.

Help Received
Mr. Platt helped with conclusion of results, Father helped construct board.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Evan M. Gates

Project Number
S1207

Project Title
Software for Autonomous Visual Tracking and Calculation of Trigonometric Parallax

Abstract

Objectives/Goals
The goal was to create software for an autonomous system using two cameras capable of tracking a rolling ball and determining its position through trigonometric parallax calculations.

Methods/Materials
Two Panasonic Ethernet cameras were used. A program, written in DrScheme, located the ball within the images, moved the cameras to center the ball, and calculated the position of the ball using trigonometric parallax. The images from the cameras were saved as jpeg files and then converted to xpm format for compatibility. The ball was found in the image using a randomized binary search algorithm. By using a variance function, outlying pixels were removed from the returned list of pixels representing the ball. The average of this list was used as the center of the ball. Using the angular position of the cameras, the point of intersection of the cameras' fields of view was calculated. This point represented the location of the ball.

Results
The ball search algorithm proved to be stable and immune to noise in the image. The visual tracking software functioned well, although slowly. The trigonometric parallax worked accurately within the limits of the cameras' six degree incremental rate of turn.

Conclusions/Discussion
The visual tracking software performed slowly due to the need to save every image first as a jpeg file and then as a converted xpm file. The resolution of the cameras angular movement was limited to six degree increments, in turn limiting the accuracy of the trigonometric parallax calculations. Planned future improvements include: eliminating the file read and write during image capture and conversion; rewriting the code in C++ for faster operation; and selecting a different camera and servo system with better angular resolution.

Summary Statement
Software for an autonomous system using two cameras capable of tracking a rolling ball and determining its position through trigonometric parallax calculations was created and tested.

Help Received
Parents helped format report; Mentor helped refine code

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Shannon K. Hextrum

Project Number
S1208

Project Title
Does the Universe Play Dice? The Question of Determinism vs. Indeterminism in the Examination of Aleatory Music

Abstract

Objectives/Goals
My goal was to deduce whether or not a selection of aleatory music really was random. I questioned if the composers, and perhaps the human mind, can truly operate on an indeterminate level.

Methods/Materials
I used two compilations of sheet music by John Cage: John Cage Piano Works 1935-48 and John Cage Prepared Piano Music - Volume 2 1940-47. After selecting 10 pieces within these books, I counted the distribution of notes in each. I used these figures to tell how close the occurences of each note were to yielding 1/12 of the whole. If each note appeared 1/12 of the time, I would assume that the pieces weren't biased in note selection and, thus, were random.

Results
The chi-square test for goodness of fit proved there was a 0% chance that my results followed the 1/12 ratios.

Conclusions/Discussion
The aleatory (or indeterminate) music was, in fact, caused (or determinate) because it held biases in notes. This could mean that the human mind cannot operate randomly, determinism may be quite applicable to our lives, or Einstein was correct when he stated, "God does not play dice."

Summary Statement
My project uses music as a model to test whether determinism or indeterminism can be applied to events in our world.

Help Received
My math teacher, Peter Foster, showed me how to perform the chi-square test for goodness of fit. My piano teacher, Jean Alexis Smith, gave me resources concerning music history.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Sarah J. Hurley

Project Number
S1209

Project Title
The Geometry of Close Packing Spheres

Abstract

Objectives/Goals
The objective is to determine whether different types of packing arrangements (hexagonal, face centered, or body centered close packing) affect the packing density of spheres. The specific purpose is to determine whether a hexagonal close packing arrangement produces the highest packing density in three dimensions, as this is true in two dimensions.

Methods/Materials
Three-dimensional models of each type of packing lattice were researched and created inside a cube in Lightwave (a 3D modeling program). The packing density of each situation was calculated by dividing the total volume of spheres in each cube by the volume of the cube using the formulas for the volumes of spheres, spherical caps, and cubes. The different packing densities were recorded and compared. All of the calculations of the packing density were repeated until they were consistent with current literature on packing density.

Results
Body centered cubic close packing calculated a packing density of 67.88%. Face centered cubic close packing had a packing density of 52.36%. Hexagonal close packing has a density of 74.46%.

Conclusions/Discussion
The results of this experiment were consistent with the hypothesis and the preliminary research done. The densest packing arrangement could be recreated to ship ping-pong balls or some similar application where packing density saves money and space. When studying crystals, if it is possible to determine the packing arrangement, then the crystal can be assigned an overall density due to the packing arrangement of it's atoms.

Summary Statement
The intuitive perception is correct that the closest packing of spheres in two-dimensions forms the basis for the closest packing arrangement of spheres in three-dimensions.

Help Received
Richard Hurley helped work in Lightwave. Bruce Rawles helped define packing arrangements by providing sources for reference.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Jeff L. Jensen

Project Number
S1210

Project Title
The HTH File Format

Abstract

Objectives/Goals
To create a simplistic file format for storing two dimensional images that supports compression, encryption, color mapping, and is easy to interpret.

Methods/Materials
Materials: 1 gigihurtz computer, Microsoft Visual C++ 6.0

Method: Design the file format and write out the specification, write the demo application showcasing a use of the hth file format, benchmark the file format, and publish the results and source code on my school's webserver. The method for writing the demo application is somewhat complex and will not be discussed here (as it would easily pass the 2400 character limit).

Public/private domain header files that were use include stdio.h, stdlib.h, iostream.h, gl.h, glu.h, windows.h, and math.h.

Results
The HTH File Format stores images using much less space than other formats, has encryption and color mapping, and is easy to read. It is also not computationally expensive to encode the images or read them. This is very beneficial when you are comparing it to other formats like JPEG which are extremely expensive in processing power.

Conclusions/Discussion
The HTH File Format is an excellent imaging solution, mainly for clip art images. RLE (Run-length) encryption is a proven technique for saving space in images, and caesar encryption supports a simple, effective, and scaleable encryption solution. Overall the HTH File Format is oftentimes, the best file format for the job.

Summary Statement
Producing the ultimate file format for storing two dimensional images.

Help Received
Jared Schiffman advised me on several theoretical topics concerning programming.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Shant H. Joukjian

Project Number
S1211

Project Title
Trapezoidal or Midpoint

Abstract

Objectives/Goals
The purpose of the experiment was to find out if the trapezoidal approximation or the midpoint approximation gives a more accurate result for the area under the curve. I hypothesized that the trapezoidal method will give a better approximation of the area under any curve.

Methods/Materials
1.Take the functions: y=0.5(x-2) 2 =1 intervals [0,4], y=2x-1 interval [0,4], y= cos x interval [0 p/2], y= e (x-2) interval [0,4], y= -3x+2 interval [0,4], y= sin x interval [0, p], y= -0.5 (x-3) 2 +3 interval [0,6], y= -e (3-x) interval [0,6].
2.Find the actual area under the curve using integration.
3. Divided the curve into 2, 4, 8, 12, 16, 18, 22, 24 subdivisions.
4. Using the trapezoidal method, find the approximate area under the curve using these subdivisions.
5. Using the midpoint method, find the approximate area under the curve using the same subdivisions.
6. Recorded the results and compared them to the actual area under the curve.

Materials:
·TI-83 Plus Graphing Calculator
·TI-83 Plus Computer Link and Program
·TI-83 Plus Trapezoidal and Midpoint Programs (on calculator)
·Graph Paper
·Ruler
·Pencil
·Computer

Results
After examining the data, I found that my hypothesis was incorrect. In the case of linear functions, such as y= -3x+2 and y= 2x-1 in the interval of [0,4], the area were always exact regardless of the number of intervals or the method we used. In the case of exponential functions such as y= e (x-2) in the interval [0,4] and y= -e (3-x) interval [0,6] the midpoint method gave a better approximation of the area under the curve. In the case of Trigonometric functions such as, y= cos x in the interval [0,p/2] and y= sin x, interval [0,p], the midpoint method gave a better approximation for the area under the curve. In the case of parabola such as y= -0.5(x-3) 2 +3, intervals [0,6] and y= 0.5 (x-2) 2+1, interval [0,4] again the midpoint method gave the better approximation to the area under the curve. In almost all the cases we found that,

Summary Statement
To find if the Trapezoidal or the midpoint method will give you a better approximation or area under any curve.

Help Received
Mother helped use TI calculator; mother and father helped with the board

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Nimi Katragadda; Simi Katragadda

Project Number
S1212

Project Title
Diagnosing Illness with a Statistically Significant Computer Program

Abstract

Objectives/Goals
Several illnesses, ranging from the common cold to bronchitis, plague today's society; the increasing shortage of nurses further prevents the treatment of these illnesses. Because of this problem, the purpose of our science fair project is to create a user friendly computer program that effectively diagnoses the illness that the user suffers from based on symptoms that the user inputs into the program. Prescription treatments and home remedies are also provided. Our hypothesis is that our program will be accurate at least 60% of the time.

Methods/Materials
We compiled a list of illnesses with their symptoms and treatments. Next, we created an algorithm to write the program. After writing the program, we made four different surveys, for four different illnesses. They directed the subjects to place a check next to the symptoms and treatments that they experienced. We placed these surveys in a physician's office so that the physician's patients could take them and mail them back to us with the addressed envelope with postage that we provided.

Results
We received 22 surveys (of 110). To determine if our results validated our hypothesis, we conducted a T test. The null hypothesis was that the program was accurate 60% of the time. The alternative hypothesis was that the program was accurate more than 60% of the time. The test was significant to the .01 level; therefore, we rejected the null hypothesis and accepted the alternative hypothesis. There is strong evidence that the program is accurate more than 60% of the time.

Conclusions/Discussion
We presented our program and the strong evidence of accuracy to physicians. We asked them their opinions.

Summary Statement
To counteract the current nursing shortage we have created a computer program that diagnoses a user's illness and provides treatments.

Help Received
Dr. Hitesh Shah gave us advice about illnesses and provided his office as the location of our surveys.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Kaitlin A. Kirk

Project Number
S1213

Project Title
Game Theory: Which Chess Move in Response to White's 1. e4 Maximizes Black's Outcome?

Abstract

Objectives/Goals
This project relates to a branch of science known as "combinatorial game theory." The objective of this experiment was to test my hypothesis: If e5 (pawn to king five) is played as the Black player's first move in response to White's first move of e4, then Black will be provided with the greatest possible initial advantage and thereby yield a greater percentage of wins for Black than other replies.

Methods/Materials
First, I determined 1. e4 was the most common opening move for White and therefore, arguably, the "best" move for White. Since e5 is symmetrical to e4 and has the same advantages, then 1...e5 could be considered the best response to e4.
PHASE I - Database Analysis to Determine the Historically "Best" Reply: I used the "industry standard" ChessBase 8.0 which is a database of 1.8+ million chess games to perform statistical analysis of the responses given to 1. e4. Thus, I determined which moves for Black yielded the best outcome and then proceeded to Phase II.
PHASE II - Modeling and Simulation to Validate Database Results: I programmed Chessmaster 9000 to play a series of 100 two-minute games against itself for each of the three Black moves selected from Phase I: 1...e5, 1...c5 and 1...g6.
I used the computer to play itself in order to reduce the variables that could be factors when two human opponents play. These variables include: players of different strengths and various physical and psychological factors. With the computer playing itself, each "player" would be of equal strength, hence the initial opening moves would be the largest factor in determining the outcome.

Results
The results from Phase I (Database Analysis) indicated that 1...c5 was the best move (49%), followed by 1...g6 (48%), and then followed by 1...e5 (44%).
The results from the Phase II (Database Validation) computer-versus-computer trials analysis confirmed that 1...c5 was still the best move (46.5%) followed by 1...g6 (45.5%) and then 1...e5 (42.5%).

Conclusions/Discussion
Although the scoring percentages differed somewhat between Phase I and Phase II results, the rank order and magnitude of effectiveness against 1. e4 remained essentially the same. Therefore, the information collected from my experiment indicates that my hypothesis is incorrect, and that in fact 1...c5 (not 1...e5) is the best response to the White opening move 1. e4.

Summary Statement
My experiment is about combinatorial game theory as it relates to modern chess opening theory.

Help Received
My parents helped purchase all project materials, and my mother helped edit my report.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)

Natalya Kostandova

Project Number

S1214

Project Title

The Battle of the Titans: Empirical Comparison of Sorting Algorithm Processing on Major Operating Systems

Abstract

Objectives/Goals

To compare the running times of sorting algorithms based on the operating system

Methods/Materials

Materials: Personal Computer, Distribution CD's of operating systems (Windows 98, Windows 2000, Windows XP, RedHat Linux, Mandrake Linux, and SuSE Linux), different C++ Compilers

Methods: - write C++ programs to perform Bubble, Selection, and Insertion integer sorts
- write C++ program to generate sets of numbers to be tested in the experiment
- create initial sets of numbers using above-mentioned program
- install the operating system to be tested
- compile the C++ source code for the sorting programs on OS that is currently analyzed
- run Bubble, Selection, and Insertion sorts on the generated sets of numbers for 5 times and record the running times required
- repeat the experiment for other operating systems
- analyze the results

Results

Results varied significantly for the different test cases. Linux and Windows operating systems performed about the same in the Bubble sort; Insertion sort was more in favor of Windows, and in Selection sort, Linux was the faster operating system.

Conclusions/Discussion

Efficiency of the OS is not its absolute quality; it is the application of OS and the process load that matters. Systems are designed to perform some tasks better than the others, but still can be beaten by the others in certain applications.

Summary Statement

I performed runtime analysis of classic sorting algorithm efficiency on major operating systems and examined the results to show how OS performance depended on the type of task chosen.

Help Received

Brother provided the distribution CDs of operating systems

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Haig J. Lafian

Project Number
S1215

Project Title
Applying the Normal Distribution

Abstract

Objectives/Goals
The project was done to discover if the normal curve applies to sets of large data. The project of this data was professional basketball player heights and weights, divided into three different teams. After finishing the project and analyzing all of the data, the large sets of data did show that all of the sets formed a normal curve, of which five of the six groups were concluded to be the same as predicted.
NOTE TO READER: Most of the data will be changed due to improvements suggested by professionals. PLease take into concideration, because the point of the project is the same, only instead of basketball player's heights and weights, other large sets of data will they be replaced by to result in more accurate results.

Summary Statement
The project will apply the normal distribution to large sets of data for the better analyzing.

Help Received
No Help Recieved

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Connie H. Leung

Project Number
S1216

Project Title
A Robust Web-Based Face Recognition System Based on Gabor Wavelets Matching

Abstract

Objectives/Goals
This study aims to develop a robust, Web-based face recognition system simulating human perception of objects and faces. After the September 11 terrorist attacks, the need for homeland security has considerably increased. An accurate face recognition system could greatly aid the identification of possible terrorists and other threats to security. The purpose of this study is to demonstrate that it is technically and economically feasible to employ Web-cams in public places, including airports, stadiums, and theme parks, to scan pictures of human faces and compare them with mug shots and ID photos hosted in a centralized database. In real time, the resulting system will identify potential suspects and alert law enforcement officials.

Methods/Materials
We collected frontal ID photos from individuals and stored them in a centralized database. Next, we set up a Web site to capture images of those same individuals with Web-cams. The Web images were taken with heads and faces in different sizes, poses, expressions, and alterations. Then these Web photos were automatically processed through the algorithms. Finally, using these Web images, we tried to find out how well our algorithms could identify each person from the database of 300 people.

Results
Our Gabor wavelets, Web-based system has demonstrated robustness in recognizing faces under four different conditions: poses, alterations, sizes, and expressions. Experimental data showed that low-resolution images captured by inexpensive Web-cams resulted in few penalties in recognition performance. In addition to the robustness, the proposed system possesses ease of implementation at a fraction of the costs of the existing commercial systems.

Conclusions/Discussion
This study has demonstrated the robustness of the Web-based face recognition system employing Gabor wavelet techniques. The performance of our system was found to be superior at a fraction of the costs compared with existing commercial systems using Principal Component Analysis (PCA) technologies. Future work with our system may include the optimization of speed and performance. A study of the tradeoff between recognition rates and time would yield a more cost-effective system. Moreover, our system has yet to be tested under a larger scale database with thousands of mug shots and ID photos.

Summary Statement
This Web-based face recognition system using Gabor wavelets improves on current commerical systems and can be implemented for commerical use to identify possible suspects.

Help Received
Used lab equipment at USC

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)

Ryan J. Liljestrom

Project Number

S1217

Project Title

Numerical Fulcrums and Prime of the Form k^2+1

Abstract

Objectives/Goals

It has not yet been proven in mathematics if there exists an infinite number of primes of the form k^2+1, where k is a positive integer. With the exception of the integer two, any prime of the form k^2+1 must also be of the form 4n^2+1, because k^2+1 must be odd so k must be even and k^2+1=(2n)^2+1=4n^2+1.

Results

This project deals with a special type of integers called "numerical fulcrums" and proves that the list of all positive integers which are not numerical fulcrums are integers n which yield a prime number in the function 4n^2+1.

Conclusions/Discussion

Numerical fulcrums could quite possibly be used some day to help solve the conjecture that there exist an infinite number of primes of the form k^2+1.

Summary Statement

Results from this project prove that numerical fulcrums, defined by the student, are related to the set of prime numbers of the form k^2+1

Help Received

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Mallory I. Matthews

Project Number
S1218

Project Title
The Effect of Capital Investment, Exports, Imports, GDP, & Industrial Production on the Performance of the S&P 500

Abstract

Objectives/Goals
The objective of this project was to determine the effect that investment capital (equipment & software) expenditures, Gross Domestic Product, exports, imports, and industrial production had on the performance of the S&P 500. The project also looked to determine the percentage correlation that each of these variables actually had on the S&P 500.

Methods/Materials
Data was gathered on the performance of the S&P 500, Gross Domestic Product, exports, imports, industrial production, and capital investment (equipment & software) expenditures for the past 52 two years. The percent change annually by quarter was calculated for each the above. Each variable was compared to the S&P 500 by deriving a correlation coefficient. Then the data was compared to see if there was a lead or lag affect.

Results
Industrial production had the largest correlation on the S&P 500, a 55.36% correlation after 3 quarters. Other variables also had a minor correlation with the S&P 500.

Conclusions/Discussion
Conclusions about the stock market's performance can't be drawn from these variables. However, the fluctuation in the stock market has a close correlation with historical events of the past 52 years.

Summary Statement
This project tests the effect different variables have on the performance of the S&P 500.

Help Received
Father provided some minor assistance with excel.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
David G. McIntosh

Project Number
S1219

Project Title
Scalable Encryption

Abstract

Objectives/Goals
Today, people can transfer information over high-speed Internet faster than ever. Unfortunately it also means that others can steal information faster than ever. A thief can steal thousands of credit card numbers in seconds. With the large number of businesses that perform transactions on the Internet, a portable, scalable, and fast program is needed to ensure privacy.

Methods/Materials
Development Machine was:
1.5 GHZ Athlon
20 GB Hard Drive
Red Hat Linux 8.0
Program uses:
make, gcc, GMU MP, bash, md5sum, sha1sum, dd, gawk, html, apache webserver for Linux, and perl

Results
My program, based on RSA encryption, generates secure keys, and then uses the two keys to encrypt and decrypt data, keeping it private. The strength of my program lies in its portability, speed, and scalability. The program uses the GNU MP Programming Library to manipulate numbers that are hundreds of digits long. Instead of writing my own library that was optimized or one computer, I used the free GNU MP which is optimized to work on computers ranging from Pentium's to Cray supercomputers. This means my program is extremely portable, and can run on almost any UNIX-based platform. The program uses several new approaches, including a simple prime checking method. Instead of using lengthy algorithms that guarantee a number is prime, my program checks the keys at the end of the process, to verify that they work.
Thirdly, my program is scalable. It can easily be configured to generate keys of any but number larger than 256. It can easily handle secure 1024 bit encryption in less than a second. It can be used for data authentication as well as data encryption

Conclusions/Discussion
My focus was on the core algorithm, making it portable, fast and scalable.

Summary Statement
The project contains five programs which allow users to protect their privacy.

Help Received
During the summer the Rocks NPACI team helped me become acquainted with linux.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Jamie L. Mountford

Project Number
S1220

Project Title
The Ability of Optical Character Recognition to Accurately Recognize and Identify Word-Processed Characters

Abstract

Objectives/Goals
To what extent can Optical Character Recognition accurately interpret individual word-processed characters? The objective is to determine if the characteristics of specific symbols on a document page affect the accuracy of an OCR output reading from that page.

Methods/Materials
Materials: computer, scanner, printer, Jet Suite Optical Character Recognition Software
Procedures:
A. Sample template was created on a word processor that included a variety of characters, including letters, symbols, numbers, and punctuation marks.
B. From this template, 13 document pages were formatted, each with a separate typeface characteristic.
C. Each of the separate pages was run through an OCR system and checked for accuracy.
D. Accuracy was measured for each individual character based on the number of errors produced by the OCR system from the original document to the image file on the computer.

Results
Alphabet and numerical characters produced the most accurate OCR output overall. Foreign characters and symbols had lower accuracy percentages and a higher number of average errors.

Conclusions/Discussion
The uniformity and familiarity of the numbers and letters aided the software in producing accurate output documents. Because of the lack of programmed memory and lexicon capabilities, the unusual symbols were generally misinterpreted and produced inaccurate output documents.

Summary Statement
My project analyzed Optical Character Recognition output documents to determine what qualities of individual characters affect the accuracy of the system identification abilities.

Help Received

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)

Hans H. Nielsen

Project Number

S1221

Project Title

Internetworking between Private Servers: Transparent Reverse Network Address Translation

Abstract

Objectives/Goals

This project created and tested a method to transparently route Internet Protocol (IP) traffic from the public Internet to hosts behind a network address translator (NAT). Most homes and businesses have a single public IP address, behind which sits an entire private network. When an unsolicited IP message is received at the public address, it is not currently possible to determine which of the many computers behind this public address to forward the message to.

Methods/Materials

To perform testing, a network was set up. There was a private segment and a public segment, with a NAT router in between. A program was designed and written that setup temporary connection paths through the NAT router. The novel aspect is that the destination is deduced just-in-time through the use of DNS hints. Six standard applications used to test the effectiveness of the program: SSH, telnet, VNC, WWW, SMTP, and NTP.

Results

A novel approach was discovered to dynamically determine which of the many private hosts a NAT should forward unsolicited requests from the Internet. This is accomplished transparently, meaning that existing applications are unaware that any reverse NAT is occurring. This has many practical applications, including replacing corporate VPNs, file transfers, or setting up public game servers. Best of all, this method is the simplest and easiest method to provide access for system administrators on the public internet who maintain remote computers on a business' private network.

Conclusions/Discussion

The trick here is to know which machine to send the request to. This is the first time this has been done using a general-purpose method. This method was successfully tested a second time, rewritten for industry standard Cisco routing equipment instead of an open-source Linux router proving its broad applicability.

Summary Statement

An algorithm for performing transparent Internet access to private networks.

Help Received

Parents helped review writings; Father mentored in understanding of network technology and issues.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Javid K. Pack

Project Number
S1222

Project Title
Geometrical Representations of the 24th Order Permutation Group

Abstract

Objectives/Goals
The objective of this study was to determine if the 24th order permutation group can be represented by geometrical symmetry operations of rotation and reflection. A group is a set of elements and operators that have the following properties: (1) closure, (2), associativity, (3) identity, (4) (inverse). A permutation group is a group that has n factorial elements, where n is an integer representing the number of objects being permuted.

Methods/Materials
Four points were found in three-dimensional space with the property that each is equidistant from the other three. These four points can be made to coincide with four of the eight vertices of a cube or with the four vertices of a regular tetrahedron. Matrices were constructed to represent the twelve rotation operators and twelve reflection operators of the permutation group. These matrices have the property that when applied to the vertices of a regular tetrahedron, all 24 possible permutations of these four vertices are obtained.

Results
The researcher found that the 24th order permutation group can be geometrically represented by the four vertices of a three-dimensional regular tetrahedron. The symmetry elements are: one identity element, three two-fold rotations, eight three-fold rotations, and twelve reflections. Matrices were found to represent each of the 24 symmetry operations.

Conclusions/Discussion
There does indeed exist a three-dimensional geometrical representation for the 24th order permutation group. It is intriguing to contemplate the next logical step: a four-dimensional hyper-solid having five vertices, all of which are equidistant from the other four. If such an object exists it will have 120 symmetry operators associated with it. Sixty of these will be rotations in four-dimensional space and sixty will be reflections. This possibility will be studied in the project I will do next year.

Summary Statement
This project finds geometrical representations of the 24th order permutation group.

Help Received

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Csaba Petre

Project Number
S1223

Project Title
One and Two-Dimensional Finite Element Analysis of Heat Transfer and Applications to Ferrofluid

Abstract

Objectives/Goals
The purpose of the research project was to develop a one and two-dimensional finite element analysis of heat transfer and applications for ferrofluid. In detail, the problem was to develop a fast and accurate numerical analysis which can calculate temperature distribution in a container of ferrofluid over time. The need for a ferrofluid heat transfer and convection simulation program arose during my team NASA Student Involvement Program (NSIP) project to develop a microgravity ferrofluid heat transfer experiment and a passive ferrofluid heat pump.

Methods/Materials
I used the implicit time marching finite element method to simulate heat transfer in any material, including ferrofluid. I started with a one-dimensional finite element solution, and after gaining experience with programming the finite element algorithm, I moved on to develop and program in C++ the two-dimensional finite element solution. I compared the simulation results with actual experimental data experiment for heat distribution in ferrofluid over time from my team NSIP experiment.

Results
I simulated the temperature distribution as a function of time in ferrofluid for both the one and two-dimensional models. I verified my simulation results with experimental data collected from my team project. I ran simulations and compared measured experimental data with simulated data from both a simulation of heat transfer by conduction only, and a simulation of heat transfer by convection and conduction. The simulation results were in good agreement with the measured data, indicating the power of this simulation method.

Conclusions/Discussion
In conclusion, I successfully developed the one and two-dimensional computer simulation program of heat transfer in ferrofluid, and my simulated results match measured results from an actual experiment.

Summary Statement
I successfully developed a one and two-dimensional finite element analysis of heat transfer and applied it to the problem of heat transfer in ferrofluid.

Help Received
Proffessor Finlayson at University of Washington supplied ferrofluid for experiments, and provided me with the ferrofluid convection equations.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
David A. Purpura

Project Number
S1224

Project Title
Can You Make Money at Blackjack?

Abstract

Objectives/Goals
My object was to see if a player can win consistently at blackjack.

Methods/Materials
Methods
1.Research blackjack playing and betting strategies.
2.Create a computer program to simulate thousands of games of blackjack to test the effectiveness of different playing and betting strategies while varying the numbers of players, number of card decks in the shoe, and the shoe utilization.
3.Simulate 5000 games of blackjack for each combination of 90 different game scenarios
4.Analyze the results from each blackjack game scenario.
5.Graph the results of blackjack simulations with winning balances
Materials
a)2.4GHz Pentium 4 Computer
b)Borland JBuilder 8 Personal
c)Microsoft Word 2000
d)Microsoft Visio 2000

Results
In the ninety (90) different scenarios that I ran only five (5) returned a positive balance, and only one (1) in the five (5) consistently had a winning balance. The fact that these scenarios won did not only depend on the player's strategies, but also the rules of the game (the number of decks, how often the deck is shuffled, and the number of players). Player 4 (the consistent winner), for example, did well in simulations 3014 and 3015 when there were six (6) decks and a reshuffle percentage of twenty-five percent (25%). But in simulations 3019 and 3020, with a reshuffle percentage of fifty percent (50%), Player 4 lost close to $2000 in each scenario. The only constant winner, no matter what the game specifics, was the dealer (won 98% (229/234) of the opportunities).
The best strategies were the one's demonstrated in simulation 3014, Wizard of Odds and True Count. In this scenario, after the 1000 hand mark the player begins to make an outstanding profit. What is important about this discovery is that the player started out rather unstable and lost $275, before he began making a profit. It is equally important to note that had Player 4 locked up money (put away after certain achieving profits of a specified amount), or stopped at his maximum, he would have made the most money in all of the scenarios, $1365.

Summary Statement
I tested popular playing/betting strategiesto see if it is possible to win consistently at blackjack.

Help Received
I used my dad for Java consultation

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Carey R. Shenkman

Project Number
S1225

Project Title
How Can It Be Proven That the Fibonacci Sequence Is Related to the Golden Mean?

Abstract

Objectives/Goals
To find, how it is possible to prove that the Fibonacci Sequence and the Golden Mean are indeed related. I hypothesized that my proof would be in showing that there is a horizontal asymptote at y=(1+5^(1/2))/2 for both the graph of f(x)/f(x-1) and the graph of [f(x-1)/f(x)] +1. I also hypothesized that I could show that the individual roots of the equation cancelled out to give (1+5^(1/2))/2 with very large domain values.

Methods/Materials
This project was centered on mathematical trials and analysis, and my materials were limited to pencils, note and graph paper, and a calculator for finding exact values for radicals, etc.
My control was the line y= approx. 1.62, to indicate my hypothetical asymptote. I manipulated the domain, or x, which gave me the values for the respondent range, or y. My graph would represent a series of dots (rather than a defined curve), and I would evaluate the progressive trend of these dots in respect to my set control in order to find the answer to my question.

Results
My data supported my hypothesis, as my resulting graph never touched the line of the Golden Mean. Although it continuously got closer and closer, it never quite reached my imaginary line. Towards thirty, it got so close that the graph became very difficult to draw. The graph had started out wide, but it quickly narrowed in scope and almost resembled a pulse monitor. Also, toward the higher domain values, the decimal seemed to match another digit with each one x increase. For example, 1.612 would hypothetically switch to 1.618, matching another digit with the next domain value. This is also noteworthy, as it shows how quickly the division almost reaches the Golden Mean. Another thing that's very interesting about the whole experiment, is how the Fibonacci Sequence, as it becomes tremendously large, manages to "stick by" this ratio. My results were identical with my second set of experiments, and the asymptote still existed in the exact same place when I took the reciprocal of the equation and added one.

Conclusions/Discussion
By finding the horizontal asymptote, that is the Golden Mean, I could prove its existing relationship with the Fibonacci Sequence. This asymptote ties the two together. It is the Mean, so that is the relationship to that side of the experiment, and it serves as the asymptote for the Fibonacci Sequence.

Summary Statement
Finding different ways to prove that the Fibonacci Sequence and the Golden Mean are related.

Help Received
none

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Tedd D. Smith

Project Number
S1226

Project Title
O(n) Sorting: A New, More Efficient Computer Sorting Algorithm

Abstract

Objectives/Goals
In this project, a new sorting algorithm was created using the language of C++. This sort would sort numbers digit by digit (by the one's space, then ten's space...) instead of as a whole. My goals were for this sort to run more efficiently (sort numbers faster) than any other sorting algorithm, and to have it run with an efficiency of O(n), which was previously said to be impossible.

Methods/Materials
To test the sort, I wrote a program in C++ that would have my new algorithm and five other commonly used sorting algorithms (insertion sort, selection sort, shell sort, bubble sort, and the quicksort) sort random arrays of numbers. The arrays of numbers would range in length from 10000 to 100000 numbers. The program would be run ten times for each size array (100 times total) to get accurate readings.

Results
My results were that the new sorting algorithm that I developed did run faster than all the other sorting algorithms that I tested it against. The quicksort was the next fastest, then the insertion sort, the selection sort, then the shell and bubble sorts.

Conclusions/Discussion
My conclusion was that my sort ran faster than all of the other sorting algorithms. I also concluded that it did run with an efficiency of O(n). Since the new sort only needed to make the same number of passes through the array regardless of it's size, that mathmatically proves it to be an O(n) sorting algorithm.

Summary Statement
I developed a new, faster computer sorting algorithm in C++.

Help Received
I had no help with this project.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Amit S. Vainsencher

Project Number
S1227

Project Title
Pattern Recognition Capabilities of Neural Networks

Abstract

Objectives/Goals
The aim of the project was to experimentally determine the pattern recognition capabilities of neural networks by presenting them with mathematically generated input of varying complexity and consequently measuring their performance as the network complexity and learning parameters were modified.

Methods/Materials
A computer application developed by the author was used to perform the experiment. The application used the neural networks back-propagation learning algorithm to conduct testing. The exclusive or (XOR) operation was used in preliminary testing followed by multidimensional geometric figures, which were used as a source of patterns presented to the networks.

Results
Results of testing with XOR revealed a greater number of neurons leads to the need for a higher learning parameter in order to maintain a lower rate of error. The particular method used in the experiment was unable to successfully classify the multidimensional figures. The experiment revealed that error decreases proportionally to the learning rate and is unrelated to the number of neurons when the network is incapable of successfully handling a pattern.

Conclusions/Discussion
The overall information gleaned from the experiment suggests what forms of behavior researchers should expect when experimenting with neural networks (specifically the back-propagation algorithm).

Summary Statement
Analyzes the learning capabilities of neural networks in specific situations and attempts to make broader generalizations of learning behavior from the resulting data.

Help Received
Work conducted independently.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Daniel A. Whisler

Project Number
S1228

Project Title
One:Four: Generalizing an Area Ratio for Related Quadrilaterals

Abstract

Objectives/Goals
Nearly 2,000 years ago, Hero related a formula for the area of any triangle. Brahmagupta's generalization included the areas of cyclic quadrilaterals. Bretschneider's theorem further encompassed any quadrilateral. In the spirit of these mathematicians, an area ratio for two isosceles trapezoids--drawn such that the smaller polygon is formed using two adjacent vertices and two diagonal midpoints of the larger trapezoid--will be generalized for any quadrilateral.

Methods/Materials
Proving the one to four area ratio for any trapezoid was accomplished through geometric relationships. Using trigonometry, the proof was extended into a formula for any convex quadrilateral where area is 1/2 times the diagonal p times the diagonal q times the sine of any intersecting diagonal angle ß, A=1/2pqsinß.

Results
Since the smaller quadrilateral has diagonal lengths exactly one half the diagonal lengths of the larger quadrilateral, multiply both lengths together and its area is always one fourth of the total area. Should the diagonals not intersect, however, (possible in highly complex concave polygons) then the area formula is undefined.

Conclusions/Discussion
To address the concave polygons, this research report generalized the area formula even further: for any infinitely concave polygon, there exists an infinitely larger convex polygon whereby both areas are related through the equation A=1/4^n(1/2pqsinß), where n is the number of successive quadrilaterals constructed through the midpoints of diagonals p and q intersecting at an angle ß. Should n=1 for related quadrilaterals, then a one to four area ratio always exists.

Summary Statement
Generalizing the area for any four-sided concave or convex polygon in relation to another polygon.

Help Received
Mr. J. Briggs provided the computer software program and helped edit the research report.

CALIFORNIA STATE SCIENCE FAIR
2003 PROJECT SUMMARY

Name(s)
Tony Wu

Project Number
S1229

Project Title
Developing a More Effective Ranking Algorithm with Query Analysis than Google's PageRank

Abstract

Objectives/Goals
The Researcher developed a new ranking algorithm of Web pages. This algorithm ranks Web pages by counting the number of query keywords and analyzing the distribution of the query keywords within a page. The objective was to see if the new algorithm is more effective than Google's PageRank in determining the relevance of a page.

Methods/Materials
Wrote a VB program that retrieves Google's top 500 -1000 Web pages with a query of 3 - 4 keywords and ranks them using the new algorithm. Randomly mixed the top 10 pages ranked by the researcher's algorithm with Google's top 10 pages and let 11 people evaluate these 20 web pages regarding their relevance to the query. Also, repeated above steps using other 10 queries in different subjects and let three people determine their relevance to the queries.

Results
The users' average evaluation scores showed that the researcher's algorithm retrieved 70.77% of relevant and partial relevant web pages while Google only 28.79% within top 10 pages.

Conclusions/Discussion
The researcher's algorithm is more effective than Google's PageRank in terms of the relevance of Web pages to search query for a certain group of people represented by the testing users and queries provided by these users, which was statistically significant at the .01 confidence level.

Summary Statement
For this project, the researcher created a new ranking algorithm of Web pages in Visual Basic code and compared it to Google's PageRank.

Help Received
Mr. Robert Ferazzi in University High School, Irvine, CA, Prof. Jeffrey D. Ullman, Prof. Gio Wiederhold, Dr. Jan Jannink and Ph.D. Student Glen Jeh in Computer Science Dept., Stanford University advised the experimental design and verified the research methodology.