CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Ruchir Baronia

Project Number

J1401

Project Title

Rescuer: A Hands-Free Mobile App for Emergencies with Easy Access for the Physically Impaired

Abstract

Objectives/Goals

The objective of this project was to create a mobile application that can send predefined text messages with the current location when the user inputs a volume key pattern on his/her mobile device (by pressing the volume buttons in a specific sequence) or when he/she speaks a user defined voice recognition key word without having to launch the application or unlock the phone.

Methods/Materials

A computer, emulation software, android phones, android studio, Pocketsphinx (voice recognition library), and java were used to create this mobile app.

I experimented with multiple API's to achieve the most accurate voice recognition and location. I was also able to reduce CPU usage by multi-threading my application. During the development phase, I created 21 different app builds. By the end of 16th build, I was able to achieve the functionality that I wanted. After this, I polished my user interface (UI) to simplify it, and to provide more customization for the user. I was finally satisfied with the app in my 21st build.

Results

I created an efficient mobile app that quickly contacts for help in emergency situations. My app runs in the background, so it can be used without launching it, even when the device is locked. An SMS with the location of the user can be sent just by saying a keyword or pressing the volume buttons in a specific pattern. My application runs on approximately 97.3% of android devices, with a minimum android version of API 14, or Android 4.0.3/Ice Cream Sandwich, which means that my app can run on almost all Android devices efficiently.

Conclusions/Discussion

My final app was far more effective than envisioned in my original blue prints.
It has many applications, for example:
- Dangerous Situations: When user needs to signal for help secretly (e.g. At gunpoint, kidnap, assault, etc.)
- Medical Emergencies/Accidents: When people don't have time to launch the application/call/text (e.g. Heart attack, car crash, etc.)
- Physically Challenged: When a user can't move over to the phone to signal help (e.g. Visually impaired people can use voice recognition), or when the user can't speak (e.g. Speech impaired people can use hardware key response)

Summary Statement

I made a mobile app that runs continuously in the background (even when the phone is locked), and can send an SMS with the user's location after hearing the user's personal keyword or upon receiving a volume key pattern.

Help Received

Parents bought my project board and test phones from the store. Parents drove me around the city so I could test the location feature of the application.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Cameron N. Cage

Project Number

J1402

Project Title

A Sense of Stylometry

Abstract

Objectives/Goals

Stylometry, which has been around since at least the 19th century, is the analysis of linguistic style in writing, generally used to determine the author of an unknown piece, but could produce patterns about genre, gender, and relationships between pieces of the same author. With the Internet becoming even more prevalent in our lives, stylometry has growing potential in the field of deanonymizing authors for criminal investigation. My question was whether stylometric analysis is correct more often than random chance in attributing authorship.

Methods/Materials

My materials were a self written Python script that automates the procedures running on a computer with Ubuntu Linux. The Python script utilized 34 writing sample sets, each containing one 4,000 word excerpt from C. S. Lewis, Nevil Shute and George Orwell, and another 4,000 excerpt from one of the authors that is #unknown# to the computer. These writing sample sets were generated by another computer program from a selection of out of copyright books.

1. Determine average characters per word in the writing sample.
2.Repeat 1 for each writing sample.
3. Compare each writing samples# average characters per word to the unidentified piece of writing and choose the piece with closest average characters per word to the unidentified piece.
4. Record whether the closest average characters per word to the unidentified piece is correct.
5. Repeat 1 - 4 individually comparing word occurrences, frequency of pronouns, frequency of word pairs, frequency of word triples, frequency of word quadruples and weighted combination metric to the unknown text instead of average characters per word.
6. Repeat 1 - 5 with each writing sample set.

Results

In each sample set or trial there were three authors making random chance 1/3. All of the metrics were more effective than random chance. The correct prediction rate of frequency of pronouns, average characters per word, frequency of word quadruples, word occurrences, word triples, word quadruples, and weighted combination was 35.29%, 38.24%, 50.00%, 52.94%, 55.88%, 61.76% and 61.76%, respectively.

Conclusions/Discussion

My project provided clear evidence that stylometric analysis is more effective than random chance, given that all seven metrics were more effective than random chance and that my mean correct prediction rate was 50.84% in addition to the fact that my median prediction rate being 52.84%.

Summary Statement

I created a program that could attribute authorship from text samples more effectively than random chance.

Help Received

None. I designed and programmed my Python script all by myself, with only the Python documentation to assist me.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Galen M. Cholbi

Project Number

J1403

Project Title

Spaceship Soup: Predicting the Emergence of Moving Objects from Random Configurations of Cells in Cellular Automata

Abstract

Objectives/Goals

I wanted to find what conditions were necessary for moving objects, known as #spaceships#, to be produced in cellular automata with varying birth rules (how many of an empty cell#s eight neighboring cells are required to be on to turn it on) and survival rules (how many of an already-on cell#s eight neighboring cells are required to be on for it to stay on). I hypothesized that rules in which cells are born with three neighbors (Conway#s Game of Life and DryLife) would produce more spaceships than rules in which cells are born with two or four neighbors (Move, 2x2, and Seeds).

Methods/Materials

I used the open-source computer program Golly to run tests of five cellular automata: Conway#s Game of Life, Seeds, Move, 2x2 and DryLife. Using Golly#s #random fill percentage# feature, I created three random initial starting configurations, called #soups#, which are 16x16 grids of cells in which 25%, 50% or 75% of the cells are already #on#, with the #on# cells distributed randomly. I watched these soups evolve for 1,000 generations in each of the five cellular automata, and then counted the number of spaceships that had appeared.

Results

Seeds (which requires two neighboring #on# cells to turn an #off# cell on) produced by far the most spaceships of any rule, with about 500 ships produced. Conway#s Game of Life and DryLife (whose birth rules are three neighboring #on# cells and three or seven neighboring #on# cells, respectively) also produced spaceships. Move (which requires three, six, or eight neighboring #on# cells to turn an #off# cell on) and 2x2 (which requires three or six neighboring #on# cells to turn an #off# cell on) produced no spaceships.

Conclusions/Discussion

My hypothesis was partially correct. I hypothesized that rules in which cells are born with three neighbors would produce more spaceships than rules in which cells are born with two or four neighbors. My hypothesis was correct in that rules in which cells are born with three neighboring #on# cells did produce many spaceships, but incorrect in that the rule in which cells are born with two neighboring #on# cells would produce the most spaceships. I did predict that Move and 2x2 would produce fewer spaceships.

Summary Statement

I tested five different cellular automata to see which ones would produce the most moving objects after 1000 generations, starting with random configurations of cells on a 16x16 grid.

Help Received

I designed and performed the experiment myself. My parents proofread the final project report and display board.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Erin M. deCastongrene

Project Number

J1404

Project Title

Pegs and Programming

Abstract

Objectives/Goals

The objective of my project was to write a computer program that models peg solitaire. I wanted to make the program learn to get better at the game over time.

Methods/Materials

Laptop computer with Snap! programming language, which is based on Scratch. I wrote a program using Snap! that models peg solitaire and learns through trial and error, then tested it to see if it improved.

Results

My program was very successful at learning to get better at peg solitaire. There was a substantial increase from its starting win percentage to its win percentage after testing it with the learning software.

Conclusions/Discussion

I built a software model of peg solitaire that successfully learns through trial and error to get better at the game. After many trials, the program performed better than it did initially. Since it was improving, my program is proved to be working.

Summary Statement

I wrote an effective computer program that models peg solitaire and learns to improve.

Help Received

I designed and built the program myself after an overview from my father, a computer scientist, on his programming process.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Peter L. Eckmann

Project Number

J1405

Project Title

Are Genetic Algorithms Effective for Computationally Intense Problems?

Abstract

Objectives/Goals

Efficient energy use can help to combat global warming. Transportation of people and goods has increased with the expansion of global markets, leading to greater energy consumption. Finding the shortest transportation routes reduces energy needs. This challenge is represented by the traveling salesman problem (TSP), which attempts to find the shortest travel distance to visit each city of a given list once. A brute-force method, which calculates all possible routes between cities, is optimal for small city numbers, but it takes a powerful computer 30,000 years to solve a modest 25-city problem. Faster methods are needed. In this project, I have tested the hypothesis that genetic algorithms, which emulate the principles of evolution and natural selection, can find better solutions to the TSP than other common algorithms.

Methods/Materials

Java programs were written to address the TSP, and were tested with model sets of cities. Specifically, a virtual population of organisms representing specific travel routes was generated. An organism was allowed to produce offspring when its route was among the shortest in the population. Each offspring was mutated by swapping pairs of cities relative to its parent, and the algorithms were optimized for population sizes and survival parameters.

Results

By testing different model city sets, I found that a genetic strategy outperformed brute force and random algorithms, while it was inferior to the nearest neighbor algorithm in most cases. Furthermore, to optimize the parameters of the genetic algorithm, such as population size and survival characteristics, a nested genetic strategy was implemented as an additional level of genetic selection. This nested genetic strategy was found to outperform the simple genetic strategy. Taken together, my results demonstrate that genetic strategies can be effective at solving the TSP.

Conclusions/Discussion

My findings suggest that applying biological principles to algorithm design has great potential for improving computational performance and finding solutions that minimize energy consumption in transportation challenges.

Summary Statement

I found that genetic algorithms, which emulate the principles of evolution and natural selection, can be effective at solving the traveling salesman problem as an example of a computationally intense problem in computer science.

Help Received

I designed, wrote, and tested the Java programs. I was helped by my parents with the interpretation of the results and proof-reading my report.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Madison A. Elliott

Project Number

J1406

Project Title

Less Is More: A Geometric Analysis of the Environmental Impact and Cost of Inefficient Product Packaging

Abstract

Objectives/Goals

The purpose of this project was to determine if product packaging designs keep production materials and costs to a minimum, while also reducing the products' environmental footprint. We live in a world that has a exponentially growing human population and a limited amount of natural resources. Simple changes to the way we package materials can have a tremendous impact on limiting the usage of resources and also limiting the amount of waste that needs to be dealt with.

Methods/Materials

I tested the efficiency of various commercial items (rectangular prisms and cylinders) by changing the dimensions of the objects. The new dimensions allowed the object to have the least amount of waste, and retain the same volume as the original. My two main focuses were the Coca-Cola can and Kellogg's Cereal box, but I also tested many other commercial items.

Results

After I found the most efficient dimensions of each item, I found that each item wasted a significant amount of materials. The Coke can wasted 4.5 percent material, and the Kellogg's box wasted 20 percent material. (These items were my main focuses of this project)

Conclusions/Discussion

I proved that my hypothesis was correct through a series of mathematical tests of each of the commercial items, and discovered that many companies enlarge their surface area to get maximum advertisement space. Simple changes to the way we package materials can have a tremendous impact on limiting the usage of resources and also limiting the amount of waste that needs to be dealt with. A 4.5 percent savings many not sound like a lot, but when multiplied by the billions of products sold annually, this translates to enormous savings of resources and waste. My ultimate goal i to, at a minimum, change the way one product is packaged.

Summary Statement

By testing each commercial item, I found that product packaging today is enlarged for advertisement and can be improved.

Help Received

My math teacher assisted me with the beginning mathematical concepts/formulas

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Kathryn C. Forrest

Project Number

J1407

Project Title

Using Zipf's Law to Analyze Word Complexity through the Past Three Centuries

Abstract

Objectives/Goals

The objective is to determine whether Zipf's law applies to state of the union (SOU) speeches given throughout U.S. history.

Methods/Materials

Get the full text of 57 SOU speeches (the 1st speech from every 4-year term) from the UC Santa Barbara archives. Write computer code in R to organize the words into a data frame that includes each word's frequency (the total number of times each word appears) and rank (its position in the list of words ordered by frequency, with the most frequent word first). For example, the word "the" might appear 1,000 times (The frequency is 1,000.) and might be the most common word in a speech. (The rank is 1.) Once this data frame is created, make graphs (using a log10 scale on each axis) to illustrate whether Zipf's law applies to the speech.

Results

The equation that describes Zipf's law is y=k/x^a (where the exponent "a" is the slope (on the log-log scale), and "k" is the total number of words in the speech). Zipf#s law says that a should equal 1, so that the frequency of any given word is inversely proportional to its rank. The law was approximately true for most of the addresses. None of the graphs in this project had a slope of exactly 1, but many came close. The graphs of SOUs with slopes closest to 1 were given between the years 1845 and 1934. On average, these slopes were about 0.9. The steeper the slope is in the graph of a speech, the less complex the speech is.

Conclusions/Discussion

Zipf's law approximately applied to most of the speeches, although it was not applicable to speeches given before the 1820s. The law reveals how speech complexity has changed over time. It was concluded that human speech has grown less complex since the 1800s.

Summary Statement

Using R code, I proved that Zipf's law (which states that the frequency of a word in a text sample is inversely proportional to its rank) applied to state of the union speeches given since 1790.

Help Received

Father helped in learning R code. I did the programming, analyzing, and assembly of the project myself.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Anusha Ghosh

Project Number
J1408

Project Title
A Device to Detect Diabetic Retinopathy

Abstract

Objectives/Goals
As of 2013, 126.6 million people around the world have developed the eye disease diabetic retinopathy, and this number continues to grow every day. In many areas where people have contracted this disease, there is not enough access to eye care, and these people do not know that they have contracted the disease until they go blind. The goal of my project was to create a device that could easily take pictures of the retina, and then scan those pictures to check for signs of diabetic retinopathy.

Methods/Materials
The materials I used for my project included an iPhone 5, a 3D printer, ABS filament, a double-convex lens with a diameter of 50 millimeters, the Tropicamide Ophthalmic Solution eye drop, the FilmIc Pro app, and the Matlab program. I also used the online services Tinkercad to design my 3D models, and Fictiv and Shapeways to print them out.

A patient's eye was first dilated using the Tropicamide eye drop. I then used the FilmIc Pro app and my device to take a video of the retina. I selected images of the retina from the video, then cropped them so that they only showed the retina and processed the resulting image using my MATLAB program.

Results
I created a device using 3D printing which can be attached to an iPhone, and used it to successfully take images of the retina of several people. I then used the Matlab program I wrote to process the images, and found that my program gave accurate diagnoses of diabetic retinopathy.

Conclusions/Discussion
It is possible to create a low cost device to allow the easy detection of diabetic retinopathy. This device can be used around the world, especially in poor or remote areas of the world to allow easy and early detection of this disease to help prevent blindness among millions of people. Because of its relative ease of use, its portability, and low cost, it can be used by people without medical training to provide care at any location.

Summary Statement
I was successfully able to create a low cost 3D printed device and Matlab image processing program to help detect the disease diabetic retinopathy.

Help Received
I would like to thank the Vision Care Center of the Palo Alto Medical Foundation for providing me with advice and the permission to test patients. I would also like to thank my parents for guiding me throughout my project.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Swati Goel

Project Number
J1409

Project Title
An Efficient Assignment Algorithm, Given Preferences and Capacity Constraints

Abstract

Objectives/Goals
The goal of my project is to design an efficient algorithm that creates assignments that maximize overall satisfaction, and to then make software specifically applying the algorithm to scheduling in schools.

Methods/Materials
I used a laptop computer with a django setup and a sqlite3 database. I tested the algorithm against already existing methods and created a web app to make interaction easier.

Results
The algorithm I created works in less than a minute when making seven assignments per person for 500 people. It does not guarantee optimality, as the problem it solves is known to be np hard, but it is efficient, running in time that is, in the worst case, a quadratic in the number of people.

Conclusions/Discussion
I built a heuristic that takes the preferences of many people and then tries to output assignments that will create the highest overall satisfaction level. This is done efficiently using a combination of limited back tracking, random sampling, and pairwise switchings, as opposed to the known optimal but inefficient methods of simply trying all permutations or full back tracking. The algorithm I created can also be used to make better and more efficient decisions in many other aspects of public life.

Summary Statement
I designed an algorithm that takes many people's preferences and then creates assignments that maximize overall satisfaction.

Help Received
I created and coded both the algorithm and the web app by myself. My computer science teacher helped me understand how to use django.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Ilene J. Hoffman

Project Number

J1410

Project Title

Genre Differentiation Using Timbral Analysis

Abstract

Objectives/Goals

Music has always fascinated me - especially genres. What classifies a particular song into a certain genre or subgenre? Do these subgenres actually differ? I decided to try to determine whether there was any mathematical difference between often confused genres using MATLAB. I hypothesized that when subgenres were analyzed, there would be no clustering by genre, revealing that the classifications are subjective.

Methods/Materials

I analyzed over 200 songs from 10 different genres, divided into two groups. Group 1 (learning) was made up of rock, modern pop, 80's pop, EDM, blues, and classical music. Group 2 (testing) was made up of indie rock, alternative rock, emo, and pop-punk. I extracted the MFCCs of 250ms windows throughout entire songs, recording the Multivariate Likelihood Estimate mean vector and covariance matrix of these MFCC matrices into separate files. I then read these files into another script, which compiled all of these into struct arrays by genre. I then plotted the clustered struct arrays, using the first parameter of the .m vectors for the x coordinate and the second parameter for the y coordinate. This way, I was able to plot the clustering.

Results

My hypothesis that subgenres were classified subjectively rather than mathematically was supported by the resulting plots. All four subgenres in Group 2 were overlapping and indistinct since they were clustered very closely together. However, I was surprised to find the true roots of the umbrella genres in Group 1, finding rock and 80's pop completely intermingled, while classical music was completely separate, with blues somewhere in the middle.

Conclusions/Discussion

For further comparative analysis, my recommendations would be to implement timbral vector voting and convert my algorithms into a compiler language so that they will run faster. These algorithms could be put to use in softwares such as Pandora or iTunes.

Summary Statement

I used MATLAB to analyze & plot 200 songs from various genres using the Mel frequency cepstrum to try to find the mathematical differences between musical subgenres.

Help Received

I wrote the algorithm myself with minimal help from Mrs. Gontar, a senior designer at Via Telecom. I received a lecture on k-means clustering from Professor Gontar of Ben-Gurion University.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Knight A. Jarecki

Project Number
J1411

Project Title
The Science Behind Moneyball

Abstract

Objectives/Goals
The object of my project is to test which method works the best to predict run production.

Methods/Materials
computer, taking baseball statistics from the last 10 years and testing different equations to predict run production and using data analogy.

Results
My hypothesis was correct and world series contenders did have better run production then bottom of the barrel teams, Wild Card contenders, and division winners. I also found out that the run production in the last 10 years has gone down. Another factor that I found out was that the American league(AL) had better run production then the National League(NL). This is another theory of mine which I call the DH(Designated Hitter) factor. My theory of the DH factor is that the American League will always have better run production then the National League because the have the DH which is an extra hitter then a pitcher having to bat. This is a huge factor in baseball.

Conclusions/Discussion
In my project I tested the run production between World Series contenders, division winners, Wild Card contenders, and last place finishers(which I refer to as Bottom of the Barrel teams) in the last 10 years. In my Hypothesis I stated that world series finishers would have better run production then division winners, Wild Card contenders, and bottom on the barrel teams. My hypothesis was correct and world series contenders did have better run production then bottom of the barrel teams, Wild Card contenders, and division winners. I also found out that the run production in the last 10 years has gone down. Another factor that I found out was that the American league(AL) had better run production then the National League(NL). This is another theory of mine which I call the DH(Designated Hitter) factor. My theory of the DH factor is that the American League will always have better run production then the National League because the have the DH which is an extra hitter then a pitcher having to bat. This is a huge factor in baseball. If I could improve my project I would test more years and do players and teams and test the free agency factor as show in the movie/book Money Ball where you take no name players with the same statistics as big time players and have the same possible outcome.

Summary Statement
My project is about predicting run production with different equations.

Help Received
My parents Pamela and David Jarecki and ,my mentors Lisa Arreola and Thomas Hurst

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Sanjana S. Jilla

Project Number
J1412

Project Title
Tempus Genie: A Time Budgeting and Planning App for Students

Abstract

Objectives/Goals
The Tempus Genie is an App that analyzes all outstanding homeworks, projects, tests, and any extra-curricular activities that are due over the next few days, weeks, and months to generate a To-Do list and time budgets for the tasks that are due the next day.

Methods/Materials
I used Swift programming to code the App - the software included development of an algorithm that generated time budgets and priorities for various tasks. The App was tested by multiple students for 2 weeks. They used the App to budget and prioritize their daily tasks. After two weeks, feedback was collected on the accuracy and completeness of the Apps's recommendations.

Results
All the participants were asked to rate the various features on a scale of 1-5 for completeness and accuracy. The average scores were: To-Do list completeness: 4.33, time budget accuracy: 4.78, usability: 4.22, feature completeness: 3.33 and overall rating: 4.11.

Conclusions/Discussion
Today, there isn't an App or utility that helps students budget and prioritize their time across the various home works, tests, projects and extra-curricular activities. My App has shown that it is possible to automate the planning process and help students budget and plan better.

Summary Statement
I created an App that hepls students budget and prioritize their time efficiently. The App creates a To-Do list for the various task that are due the next day.

Help Received
I designed and implemented the algorithim and user interphase using Swift. A family member with programming background reviewed my work and gave me reccomendations to improve the implementation.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Solomon A. Kazmie

Project Number
J1413

Project Title
LifeLine: Thinking Outside the Black Box

Abstract

Objectives/Goals
To develop, demonstrate and test an iPhone-based data logging, position tracking and navigational path retracing system.

Methods/Materials
I used Apple#s XCode and Swift to design and program an iPhone app which uses GPS signals to track and log the phone#s position and orientation. I used MapKit to add a realtime map display and used MessageUI to program the app to send the logged path coordinates in an email report.

I added a stack-based #retrace mode# that calculates and guides the user back along their path to the apps point of origin. I deployed the app on numerous devices to test the iPhone#s position tracking, logging and retrace capabilities under a wide range of conditions.

Results
Using my LifeLine app on several devices in a variety of regions, terrains and natural conditions, I collected more than 10,000 data points across a 350 mile radius.

Conclusions/Discussion
My iPhone-based LifeLine systems functioned reliably and consistently, on land and on water, in all tested environments, including desert, lake, mountain, canyon, forest, fields, swimming pool, park, freeways and city streets. I had to add a slider to adjust the #retrace mode# to function for cars as well as pedestrians. In addition to its use in flight recovery, #Retrace mode# alone could be a life-saving tool in hiking, camping, sailing, scouting, recon, search and rescue situations.

Summary Statement
I wanted to use off the shelf iPhone#s to enhance the traditional black box flight data recording system by linking the black boxes with a distributed network of smartphone-based data storage nodes, which can also function when jettisoned a

Help Received
A family friend introduced me to the Swift programming language and helped explain Apple#s map annotation protocol.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Jimin Kim

Project Number

J1414

Project Title

A Combinatorial Proof for the Geometric Series, Binomial Theorem, and the Square of a Polynomial with Tiling

Abstract

Objectives/Goals

Provide a visual proof for complex mathematical identities.

Methods/Materials

Paper and pencil.

Results

The three formulas I proved using a visual method called tiling helped me understand the combinatorial concept behind these mathematical identities.

Conclusions/Discussion

I proved the formulas for the geometric series, binomial theorem, and the square of a polynomial with an inductive and combinatorial approach. To do so, I used a method called tiling. This allows many visual learners to understand proofs more easily.

Summary Statement

I visually proved the formula for the geometric series, binomial theorem, and square of a polynomial using a method called tiling.

Help Received

After I had done quite some research on tiling, I stumbled upon a concept within tiling, so I reached out to the UCI Math Department and received help from a PhD, Hayan Nam.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Katherine G. Knapp

Project Number

J1415

Project Title

The Glorious Unfolding: Similarities between Flexagons, Mobius Loops, and Klein Bottles

Abstract

Objectives/Goals

Study how shape and dimensionality determines the number of sides on a shape and investigate the similarities and differences between flexagons, Mobius loops, and Klein bottles.

Methods/Materials

Create paper models of loops with various number of half twists, create various types of flexagons, and crate a Klein bottle. Determine number of sides. Cut models in half and observe results. Create transition diagrams for flexagons.

Results

Through my experiments, I found that Mobius loops with an odd number of half twists are one sided, while those with an even number have two sides. Hexaflexagons with an odd number of faces are one-sided but hexaflexagons with an even number of faces and all tetraflexagons have two sides. Mobius loops cut into half split into two interlocking loops, except for Mobius loops with one, five and fifteen half twists, which became one longer loop. Trihexaflexagons cut in half resemblance to a Mobius loop with five half twists. I was unable to find a Mobius loop that resembled the hexahexaflexagon when it was cut in half.

When a Mobius loop was cut unevenly, double-sided loops would produce two interlocking loops of equal length. One-sided Mobius loops also produced two interlocking loops but with one of the loops double the length of the other.

Maps were created for the hexaflexagons, showing how to transition from one face to another in the fastest way possible using a Tuckerman Traverse diagram.

Conclusions/Discussion

All one-sided shapes have an odd number of twists. Because Klein bottles are one sided, they can only be made from loops with an odd number of twists. Every hexaflexagon with an odd number of faces has only one side, which means that an odd-faced hexaflexagon is a flat, one-sided shape with multiple faces.

Summary Statement

Are there similarities in flexagons, Mobius loops and Klein bottles?

Help Received

None other than from listed reference materials.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Elizabeth Kravtchenko

Project Number
J1416

Project Title
How Math Helped Me Become an All-American Athlete

Abstract

Objectives/Goals
The purpose of this project was to find out if completing the most challenging triathlon races as a youth athlete would help you become an All-American triathlete (rank top 10% nationally in your gender and age group). At the beginning, my hypothesis was: If I complete the longest and the most difficult races, then I will get a very high ranking score. After my research and analysis, I revised my hypothesis to: If I complete the shorter and easier races, then I will get a very high ranking score.

Methods/Materials
In my theoretical research and analysis, I used mathematical modeling and numerical methods to derive and solve complex exponential equations. First, by hand, using paper and pencil. Then, using the Desmos graphing calculator on my phone. Finally, using the advanced Excel functionality on the computer. In my experimental method, I completed multiple long and short distance triathlon races. My dependent variable was my race score. My independent variable was the time it took me to finish the race of a certain distance. I used the basic materials and equipment needed to complete my races.

Results
In my project, I combined the knowledge from sports sciences with my analysis of the race scoring algorithm to gain valuable insights and adjust my competition strategies. My mathematical analysis showed that shorter and easier races would produce higher race scores. Indeed, in the shorter and easier races, my scores were higher.

Conclusions/Discussion
My initial hypothesis was incorrect. My revised hypothesis turned out to be correct and it was strongly supported by my experimental results. Based on my analysis and results, I formulated several recommendations for USA Triathlon: how to improve the existing scoring algorithm and how to balance out the triathlon races. Understanding the science and data behind the races allowed me to choose the right racing strategies and achieve my goal. I received a high overall ranking score, which got me an All-American Honor.

Summary Statement
My project is a story of how applying a scientific method and choosing the right racing strategies based on math and data helped me become an All-American triathlete.

Help Received
Marsenne Kendall (science teacher), Tatiana Seletskaia (physics teacher), and Adrian Land (mentor) reviewed my project and provided feedback. Vladimir Kravtchenko (my Dad) taught me how to use the advanced Excel functionality and drove me to all of my triathlon competitions.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Lev Kruglyak

Project Number

J1417

Project Title

Inductive Reasoning Algorithm

Abstract

Objectives/Goals

The objective of this study is to understand whether or not the algorithm I created could be used to solve geometry problems.

Methods/Materials

Computer with Eclipse IDE and JRE 1.8.

The algorithm using a recursive function to build an inverse inductively defined tree and then traverses it to find the optimal solution.

Tested the algorithm on varying types of triangle congruence problems.

Results

Algorithm found the optimal solution in a short amount of time for standard triangle congruence problems. Algorithm was less efficient in solving problems with lots of misleading turns in the logic. Algorithm was incredibly good at separating the optimal solution from a problem with lots of extra information.

Conclusions/Discussion

The Inductive Reasoning Algorithm however, will never work on problems that require the use of algebraic properties, such as transitive property, substitution property, etc. The reason for this is the way the algorithm solves problems. The algorithm proves the problem by building a tree from the prove statement to the given using rules that describe a transformation from one set of object relationships to another, and the input can be generated from the output. Properties also describe a transformation from input to output, but the input cannot be generated from the output. Because of this, the Geometry Bot Algorithm cannot use properties to solve a problem. An example of this is the transitive property. The transitive property states that if a = b, and b = c, then a = c. Since the algorithm solves everything backwards, the algorithm would be asked to prove a = c. The algorithm has no way of knowing that b is the common term, so it would try to prove that a and c are congruent to every other value in the problem, leading to an inﬁnitely large problem.

Summary Statement

I devised an algorithm for solving problems using Inductive Reasoning, that can quickly find an optimal solution using a database of rules.

Help Received

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Rachana Madhukara

Project Number
J1418

Project Title
Devising a Secure and Efficient Hybrid Cryptosystem

Abstract

Objectives/Goals
The objective of this project is to devise a cryptosystem that is both efficient and secure.

Methods/Materials
Paper, pencils and a computer for initial research. Wolfram Mathematica software was used to validate final mathematical operations. After researching, determined that combining an asymmetric cryptosystem with a key exchange is beneficial. This was mathematically verified later.

Results
The combination suggested was more secure and efficient since the mathematical equations proved that the system is less susceptible to attacks.

Conclusions/Discussion
Combining the two systems did produce a stronger, more secure cryptosystem in the end. The cryptosystem is stronger and efficient as it can withstand more attacks while also optimizing computations. My cryptosystem is also able to eliminate most problems from both RSA and the Diffie-Hellman Key Exchange methods. Hence I conclude that my method produced a strong cryptographic system to encode messages.

Summary Statement
I devised a more efficient and secure hybrid cryptosystem to better encode messages.

Help Received
None. I designed and performed the project by myself. Then Dr. Nagabhushan reviewed my results.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Alexander T. McDowell

Project Number
J1419

Project Title
Fractal Learning: A Better Approach to Neural Networks?

Abstract

Objectives/Goals
Are Fractal Neural Networks more efficient than Feed-Forward Neural Networks in their accuracy of outputs, RMS error values, and overall learning growth and progression? My hypothesis is that #Fractal# Neural Networks are more efficient because of their sub-network properties and changing of weighted links. I think this is because the weighted links adjust to the output values of the sub-networks and make the net learn more efficiently.

Methods/Materials
For my experiments, I started with a pre-existing Feed-Forward Neural Net as a Control group, and then modified it to become a Fractal Neural Net. The nets were written in C++. To test the nets, I prepared nine different data sets to train the neural nets. The data sets ranged in difficulty and topic and were downloaded from the UCI Machine Learning Repository.

Results
After learning each data set, the Fractal Neural Net showed similar learning growth to that of the Control, and also had similar output and RMS error values as well. Even though these values were similar, the Control had values that were more efficient more often. Overall, the Control beat the Fractal Net four times, tied with the Fractal Net four times, and lost to the Fractal Neural Net once.

Conclusions/Discussion
My final conclusion is that the Fractal Neural Network is less efficient than Feed-Forward Neural Network.

Summary Statement
I compared two types of neural networks, a Fractal Net, and a Feed-Forward Net, at a variety of datasets to determine which was more efficient.

Help Received
None received. I made the Networks off of a code-base online, and based off of a paper done by Roderick Murray Smith, a professor at the University of Glasgow

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Dominic K. Olson

Project Number

J1420

Project Title

Are You in Tune? A Program to Make You a Better Musician

Abstract

Objectives/Goals

When a musician is learning to play an instrument, they must learn to hit all of the notes in tune. This can be more difficult on instruments that can produce continuous variation in pitch (e.g. Violin, Trombone). This project demonstrates a software program to make and analyze a recording, showing which notes were out of tune.

Methods/Materials

The Python language was used to develop this program. In order to find out the tuning of notes, an algorithm was needed to calculate frequency. To do this, a modified zero-cross algorithm was used. The standard zero-cross technique only works on "pure" tones, not those with complex harmonic components. The custom version identifies other features of the waveform and uses those features plus the zero crosses to correctly calculate the wavelength and frequency. The program also interpolated cross points and averaged multiple waves to increase precision.

Results

The program was tested with different musical instruments and different notes. The program can correctly identify the pitch of a note, and if it is sharp, flat, or in tune.

Conclusions/Discussion

A modified zero-cross algorithm is able to determine the frequency of notes within recorded musical passages. The program is able to identify if a note is out of tune, and how much. This program could potentially help train beginning musicians to improve their intonation (ability to play notes at the correct pitch).

Summary Statement

I wrote a computer program to analyze a musical recording and display which notes were in our out of tune.

Help Received

None. I wrote my software program myself.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Deepro F. Pasha

Project Number

J1421

Project Title

Programming a Smarter Planner for Students to Prioritize Their Daily School Work

Abstract

Objectives/Goals

As a very busy Middle school student with homework, tests, class projects etc. I regularly encounter the problem how to decide which class task to do first? To solve this problem, I wanted to test if it is possible to create a smarter Planner utilizing a programming language that can prioritize students# classwork each day.

Methods/Materials

Background Information about courses and tasks were collected through internet research and informal conversations with Students and Teachers. Early design was developed and possible logics were constructed. The program is written in Python 3.4.0 programming language. All the possible inputs were created in the Python 3.4.0 programming environment. The codes in the program were developed to accommodate a customizable database interface where students can input their list of tasks and courses according to their daily need. A customizable ranking system of courses and tasks were also developed where students can rank their tasks and courses depending on their priority. The codes were run and Trial 1 outputs were generated. Necessary adjustments and changes were made in the codes to make it more efficient. The process of Trial and Adjustment was repeated until finally generating desired outputs and was tested by other potential users. Program was refined and finalized.

Results

In my program, the user can just go through couple of question answers and get their daily work prioritized. It is Voice activated so the question answer part is more like talking with someone who is able to help. It is customizable, so students can input their list of tasks and courses according to their daily need and also can rank their tasks and courses depending on their priority. All these flexibility makes this program unique and applicable for any student with different kinds of priority choices for daily work.

Conclusions/Discussion

My hypothesis was proved. I created a smarter Planner utilizing python programming language that can prioritize students classwork each day. It is applicable for any student of any grade level. It talks to them, so they can feel like someone is helping them to prioritize their daily classwork. In the field of Mathematics and Computer Science my work is a good example of how developing program can help to solve different kinds of daily life problem. It is an example of connecting the power of programming to a practical problem solving.

Summary Statement

I programmed a smarter planner utilizing a programming language that can help to prioritize students# daily school work.

Help Received

I did the programming myself after learning to code from books and internet sources. I demonstrated my program to my school's honors science teacher.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Albert W. Qin

Project Number

J1422

Project Title

An Application of Artificial Neural Networks to Decentralized Cooperative Navigation

Abstract

Objectives/Goals

Swarms of robots offer an alternative to the conventional single-robot approach in navigation and exploration problems. In this project, I show the viability of using Artificial Neural Networks (ANNs) to control swarms of robots. I train ANNs to execute two tasks involving "exploring" area in a virtual environment: Separation - "dividing and conquering", and Cohesion - staying close together.

Methods/Materials

The project was written in Python.

The virtual environment consists of a 30x50 cell region, where each cell can either be an obstacle or an explorable area. Each robot has 3 proximity sensors that gives the distance to the nearest obstacle and 2 "swarm sensors" that give the distance and direction of the swarm member in the best position. The ANNs output the direction in which the robot turns. To evaluate the ANNs, the program determines how much area 3 robots (controlled by identical but independent ANNs) could explore under a time limit. In Cohesion, the robots have to stay together for the explored area to count. I used a simple genetic algorithm (GA) to train the ANNs. The GA keeps a population of 80 ANNs - each is evaluated, and is given a fitness score. In each iteration (generation), the GA selects the best ANNs through a process called roulette selection and creates a new population based on those ANNs. For each task, I ran the GA 5 times for 20-30 generations each. I then tested the best ANN out of all 5 runs in my tester program. Each ANN was tested at least 150 times. I also varied both obstacle density and swarm size to test the ANNs' versatility.

Results

The best ANNs stayed together very well in Cohesion, and separated immediately in Separation. I compared the best ANN scores from both tasks to a baseline - randomly wandering robots. The ANNs performed much better than the baseline - almost two times better in Separation, and over 50 times better in Cohesion. These differences were statistically significant ($ps < 0.001$). In addition, my tests showed that both ANNs were versatile.

Conclusions/Discussion

This project met my design goal of creating two intelligent ANN controllers. On both tasks, the ANNs perform well and are versatile. Although there is room for improvement (e.g. robots occasionally get confused), this project suggests that ANNs may be used to control swarms in real world navigation and exploration problems.

Summary Statement

This project demonstrated the viability of using ANNs to control swarms of robots by training ANNs to perform swarm navigation tasks in a virtual environment.

Help Received

I design and coded the project myself after searching the internet for methods pertaining to neural networks. My mother introduced me to Python. I also got coding help from Stack Overflow.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Arvind Ramachandran

Project Number

J1423

Project Title

TrackBar: Tracking Food Expiration to Reduce Wastage

Abstract

Objectives/Goals

According to the United Nations Environment Programme (UNEP), one third of the food produced in the world for human consumption every year is wasted. Wasting food also wastes resources such as water and energy used to produce this food.

I believe that one of the primary reasons for food wastage in the US is that people are unaware of the expiration dates of the food in their refrigerator and pantry. My project explores a novel solution to reduce food wastage in homes by tracking the expiration dates and alerting consumers before the food expires.

Methods/Materials

My solution is to encode product#s UPC code and expiration date in a QR code that can be part of the product labeling, and then use a smartphone application to scan the QR code and track expiration dates. The design criteria for my app were: input data quickly, allow users to edit the data manually, and alert users of the expiration states of products visually.

I created an Android smartphone application using Android Studio, Java and the Android SDK. My app uses the ZXing library to scan QR codes using the smartphone#s camera, SQLite to persist the items, and RecyclerViews, Activities and layouts to render the UI.

Results

My app, TrackBar, worked in accordance with the design criteria # it is quick and easy to use and alerts consumers before food expires. While testing at home, my app warned us of items about to expire, allowing my mom to use the items before expiration.

Conclusions/Discussion

In conclusion, my goal of reducing food wastage by alerting people to the expiration dates of food in the refrigerator and pantry was validated by my solution. I believe that my app will help in reducing food wastage and help save money and the environment.

Summary Statement

I created a smartphone based solution to track food expiration dates to reduce food wastage.

Help Received

I wish to thank Mr.Sagar Waghmare who guided me with Android programming, and my science teacher Mrs. D#Souza for her help and guidance throughout this project.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)
Sage C. Saling

Project Number
J1424

Project Title
Limiting Drug Interactions through an Online Dashboard Tool

Abstract

Objectives/Goals
The objective of this study is to determine if using a dashboard on the 30 most common medications for children under the age of 13 helps limit the number of harmful drug interactions reported for in the future.

Methods/Materials
Access to a computer, internet, test subjects (parents with children under the age of 13), pencil, paper, printer, red pens, black pens, and erasers. Gave out 2 assessments. One taken without dashboard for ten minutes, the second assessment taken with dashboard for ten minutes. Compared test scores from before and after using the dashboard.

Results
Several parents with children under the age of thirteen were given two identical assessments on harmful drug interactions, one to take before seeing the dashboard and the second to take while using the dashboard. Repeated trials were run to determine if using the dashboard would increase the subjects test scores by 30%. The average test score for the first assessment was 32.8%. The average test score for the second assessment was 90.75%. The total improvement score among the two scores was 59.05%. Proving my hypothesis to be correct.

Conclusions/Discussion
Repeated trials with multiple people revealed a 59.05% increase in test scores from before and after perusing the dashboard. It is concluded that using a dashboard on harmful drug interactions for children under the age of 13, can limit the number of harmful interactions in the future.

Summary Statement
I created a dashboard on harmful drug interactions, of the thirty most common medications for children under the age of 13.

Help Received
I designed, built, and tested on my subjects by myself. I got help in finding differnt programs I could program my dashboard in.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Kirtan U. Shah

Project Number

J1425

Project Title

Image Compression: Which Algorithm Performs the Best?

Abstract

Objectives/Goals

The goal of this project is to develop, test and compare multiple image compression algorithms to find out which image compression algorithms/methods are the most effective and time efficient.

Methods/Materials

Computer, Processing IDE (Java), several test images. I developed three algorithms and tested them; one was an existing one (LZ77), the other two were original ideas (Most Common Pixel, Lines). Next, I compared efficiency and effectiveness to see which performed the best.

Results

I found that the Lines algorithm was the fastest and most efficient, often coming up with 90% compression in less than < 100ms. The Most Common Pixel algorithm took more time, usually < 200ms, and could compress images up to 60%. Finally, LZ77 (an existing algorithm) performed well, with 85-95% compression, but was the slowest (~5-6sec per image).

Conclusions/Discussion

This project showed that algorithm that looks for patterns is efficient and effective most of the time. Even though effectiveness is important, data should be easy to uncompress. For example, this is important for web applications because users need to not only send data quickly, but also uncompress it quickly.

Summary Statement

I developed and compared multiple algorithms to find out which algorithms/methods are the most efficient and effective.

Help Received

I researched image compression algorithms and techniques online.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Rushil Srivastava

Project Number

J1426

Project Title

BliPC: A Low Cost Computer for the Blind Using a Raspberry Pi and Cloud Technologies

Abstract

Objectives/Goals

Create and program a computer for the blind community at a low price using inexpensive modern technologies. The computer should be able to allow the blind to 1) View and Edit Word Documents, 2) Send an email, and 3) Use Search Engines to browse the web.

Methods/Materials

Material: Raspberry Pi Model 2B running modified version of Debian OS. Wireless Keyboard with Braille Stickers. Wi-fi Module for Internet Connectivity. Open Source Libraries: Chromium, ChromeVox, iSpeech TTS, Cepstrial TTS. Developer#s Guide for Chrome Extensions, Apps, and Plugins(developer.chrome.com/extensions/devguide) outlining steps to developing applications for Chromium.

Methods: 1) Connecting Hardware Components. 2)Programming and Developing Hardware Components. 3) Testing and Modification of Software Components. 4) Final Usability Testing.

Results

What Worked:

Developed computer in the $300 range which is roughly $1000-5000 cheaper than other braille computers. Users are able to create documents, send emails, and use popular search engines such as Google to browse the web. The 4 main BliPC software and applications for Chrome worked smoothly. BliPC Screen Reader worked perfectly, with no errors at all and the BliPC WebText worked smoothly as well, converting most websites to all text and removing all images. The BliPC File Manager worked every now and then, and the BliPC TTS engine was the backup incase the Cepstrial TTS engine failed. BLIC, a brand new computer language for the blind, worked perfectly with the compiler compiling programs correctly.

Improvements to be Made:

The TTS engine malfunctioned every now and then, making it harder for the user to understand what was happening on the computer.

Conclusions/Discussion

BliPC met all of the original expectations the computer should be able to perform. Reliability of BliPC#s TTS Engine continues to be a problem and needs work to be done. All components of the BliPC aside from the TTS engine work smoothly. Since BliPC shows promise as an inexpensive computer for the blind, I have released the project#s source code on GitHub for the community to help build a stable version of BliPC in the future.

Summary Statement

I created an inexpensive computer for the blind using modern technologies and the cloud, and developed a computer language that developers and the blind can use to enhance user experience on the BliPC.

Help Received

I designed and programmed the project myself. I read quite a bit of articles on the Internet, and researched on the Google Developer Database. I found the motivation to create my project from my Science Teacher and Parents.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Rohan K. Vanheusden

Project Number

J1427

Project Title

Analyzing the Effects of Hidden Neuron Quantity on Neural Network Performance

Abstract

Objectives/Goals

The objective of this experiment was to determine how the amount of hidden neurons in an artificial neural network affect its performance in the game "Dots and Boxes." The opponent competing against the neural network makes a random decision each turn. The neural network will use an evolutionary/genetic model to learn how to play the game with no prior experience.

Methods/Materials

To investigate the objective, I made a simulation using an artificial neural network. The neural network competes against an opponent that makes random decisions. The neural network uses sigmoid neurons and an evolutionary/genetic algorithm. The game, "Dots and Boxes" uses a 4x4 grid in this simulation. Each turn, the owners of the lines in the game are fed to the neural network as inputs. There are three output neurons that represent the line the neural network chooses. If a line that is already claimed is picked, the game will restart. The fitness of the neural network is calculated by the expression pl+t, where p is the score of the neural network, l is the number of lines on the grid, and t is turns taken.

Results

The overall effect of increasing the number of hidden neurons in the artificial neural network is a decrease in the average highest fitness score of the network over 10,000 generations.

Conclusions/Discussion

The results show that adding more hidden neurons to the artificial neural network is counter-productive. One reason for this may be that larger neural networks take more generations to increase fitness. It is possible that because of the generations limit in the experiment, larger neural networks were under-developed. If this is the case, it is possible that adding more hidden neurons results in a slower fitness increase. It is also possible that additional training would have a positive impact on the performance of larger neural networks over many generations. The way that I chose to calculate the fitness of the neural network likely had a significant effect on the results.

Summary Statement

I analyzed how the number of hidden neurons used in an artificial neural network playing "Dots and Boxes" against a simulated opponent that made random moves affected the artificial neural network's performance and found that increasing amounts of hidden neurons caused the performance of the artificial

Help Received

I designed and programmed the algorithms used in the simulation for this project after doing research on how artificial neural networks work.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Johan D.S. Vonk

Project Number

J1428

Project Title

Pitch Detection on Arduino Using Autocorrelation

Abstract

Objectives/Goals

This project aims to create a device that reliably converts sound from an instrument into an array of notes.

Methods/Materials

Arduino Uno running C++, microphone with automatic gain control, credit card size printed circuit board, push button, 1.8″ TFT LCD screen, iPad running Garage Band (as synthesizer). Built a device to reliably calculate frequency of sound sample from microphone input, convert to notes and send to synthesizer as output.

Results

Evaluation criteria are: (a) no noticeable delay between the incoming audio signal and the digital MIDI output; (b) the correct pitch should be detected for the notes produced by a B flat clarinet and samples tested from the University of Iowa Electronic Music Studios; (c) the beginning and duration of each note should be consistently identified (called segmentation).

Device results are that pitch detection on the test set was accurate from Eb3 (155.5 Hz) to G6 (1568 Hz). It missed the top three notes on clarinet, mainly played by professional musicians. Note level segmentation worked well and was verified by playing a clarinet while listening to the output of a MIDI synthesizer.

Conclusions/Discussion

While the project pushed the limits of the Arduino in processing power and available memory, it can be made for about $20. Excluding the optional display, the product fits on a very small device that can be clipped onto an instrument. Using a synthesizer, the notes generated by the device can be played simultaneously on different instruments. Or, using music notation software, these notes can be printed as sheet music. It was a fun, interesting and educational project that might benefit others.

Summary Statement

I created a device that uses the mathematical principle of autocorrelation to convert sound it detects from a musical instrument into an array of notes.

Help Received

I researched and developed the algorithms, built the device and performed experiments. My father, Coert Vonk, helped in reviewing my code to keep it organized and maintainable. He contributed a driver to digitize the analog input signal.

CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY

Name(s)

Patrick Z. Yu

Project Number

J1429

Project Title

A Novel Approach to Image Recognition with Leaves

Abstract

Objectives/Goals

Image recognition technology is a hot field in computing that is currently applied in areas including health, security, gaming, autonomous vehicles, and virtual reality. The purpose of this project is to design an efficient image recognition algorithm using leaves as prototype objects. The algorithm will have the ability to learn and classify leaves by inputting a picture of a query leaf and comparing it with existing leaf samples. After mastering leaf identification as demonstrated by this project, the algorithm can be further incorporated into the many applications of image recognition.

Methods/Materials

A desktop app was created as a proof-of-concept to demonstrate the algorithm, and approximately 30 images of each leaf type were introduced into databases. Then, a series of query leaves was presented to the algorithm for comparison with each of the databases# existing leaves. This comparison is done by calculating the distances from the query leaf#s centroid to its edges and plotting histograms of leaf distances to form an average similarity percentage between the query and the database leaves. The individual comparisons with the database leaves were recorded to graph the data.

Results

All four leaves were identified correctly, with the first leaf as acer ginnala at 76% similarity, the second leaf as betula alleghaniensis at 87% similarity, the third leaf as castanea dentata at 90%, and finally the cercidiphyllum japonicum at 91%.

Conclusions/Discussion

The algorithm accurately detected and classified the four given leaves; however, the occasional outliers when comparing the query leaf with database leaves shows the importance of keeping a complete database of leaves. The more data points (or leaves) collected into a database, the better the algorithm will be at recognizing leaves. Consequently, defective samples will contribute less while different variations will be accounted for more in the database.

Summary Statement

The goal of this project was to create a novel and fast image recognition algorithm to classify leaves and other objects such as street signs, etc.

Help Received

My mentor Dr. Ismail provided guidance for parts of the project I wasn't too clear about, such as the Correlation method, and helped improve my slides and overall presentation.