Online Proceedings

Poster Session 1

Commons East, Easel 81

11:00 AM to 12:30 PM

Tapulator: A Non-Visual Calculator using Natural Prefix-Free Codes

Vaspol Ruamviboonsuk, Junior, Computer Science
Mary Gates Scholar
Mentor: Richard Ladner, Computer Science & Engineering

A new non-visual method of numeric entry into a smartphone is designed, implemented, and tested. Users tap the smartphone screen with one to three fingers or swipe the screen in order to enter numbers. No buttons are used—only simple, easy-to-remember gestures. A preliminary evaluation compares the method to a standard accessible numeric keyboard with a VoiceOver-like screen reader interface. Preliminary results indicate that users can enter numbers faster and with higher accuracy when using this method. The Tapulator, a complete calculator based on this non-visual numeric entry that uses simple gestures for arithmetic operations and other calculator actions is described.

Poster Session 1

Commons East, Easel 82

11:00 AM to 12:30 PM

Appliance Reader: A Smartphone Application for Blind People to Read Their Digital Displays

Antonius Denny (Denny) Harijanto, Senior, Computer Engineering
Michael Thanh Hung (Mike) Hotan, Senior, Computer Engineering
Mentor: Richard Ladner, Computer Science & Engineering
Mentor: Bryan Russell, CSE, Intel

Jane is blind and in need of a new wall oven for her apartment. As she shops for the new oven she discovers that almost all ovens have digital displays for showing the state of the oven, its temperature, and the time left on its timer. They also have buttons flush with the display panel for providing input. She knows from past experience that she can put Braille labels on the buttons and will eventually memorize which button does what, so input will not be a problem. However, how will she read the digital display? Through a friend she learns of a smartphone application called Appliance Reader, which can read digital displays for various appliances including some wall ovens. In order to implement the appliance reader, we have to utilize computer vision algorithms particularly those that match one image against another. The two images can be taken from different perspectives so geometric transformation needs to be inferred from the two images. Once the two images are matched properly, the digital data has to be interpreted to form what would be told to the user. The interpretation requires a machine learning algorithm because standard optical character recognition is not sufficient. A user interface suitable for blind users is needed to make the Appliance Reader practical. As a part of CSE Mobile Accessibility group, which develops smartphone application to aid handicapped people, the two of us have been developing Appliance Reader from the ground up. Under the guidance of Professor Ladner and our mentor, an Intel researcher, Bryan Russell, we have been researching and building Appliance Reader, which is a computer vision-based smartphone application.

Poster Session 1

Commons East, Easel 80

11:00 AM to 12:30 PM

Making Public Transit Accessible via Crowdsourcing

David Kawai (David) Wong, Junior, Computer Engineering
Mentor: Richard Ladner, Computer Science & Engineering
Mentor: Shiri Azenkot, Computer Science & Engineering

Public transit plays an important role in the daily lives of blind individuals. Unable to drive, blind commuters must rely on buses, subways, and other methods of public transportation to travel from one destination to another. Although transit schedules and even real-time arrival information are now readily available and accessible, there are still barriers to overcome. For example, a blind bus rider may find it difficult to locate a designated bus stop. Prior research has shown that, given a set of information about a bus stop, blind commuters will find their bus stop with significantly greater ease and lesser time spent. Our goal is to create a database of useful and reliable information regarding every bus stop in Seattle and make that information accessible to blind individuals. In order to collect such a vast amount of data, we have turned to crowdsourcing, a process that involves dividing work amongst a large group of workers over the Internet. By providing workers with Google StreetView images

(ground-level visuals on the appearance of a location) and simple, multiple-choice questions, we hope to collect accurate information to answer questions such as "How many benches are at the bus stop?" or "In what direction is the bus stop relative to the intersection?" After acquiring a sufficient set of bus stop data, we will test the usefulness of this information to blind commuters by providing them with access to our database and analyzing their feedback. We believe that our work in this field will improve the public transit experience for blind individuals by providing them with a feeling of confidence and independence during their commute. Our work in this field may lead to further research in utilizing crowdsourcing to solve accessibility issues.

SESSION 1Q

TOMORROW'S TECHNOLOGICAL SOLUTIONS AND APPROACHES FOR TODAY'S PROBLEMS

Session Moderator: Marc Dupuis, Computing and Software Systems

389 MGH

1:15 PM to 2:45 PM

* Note: Titles in order of presentation.

The Needle in the Haystack: Automatic Idiom Identification

Felix (Grace) Muzny, Senior, Computer Science, English
Mary Gates Scholar
Mentor: Luke Zettlemoyer, Computer Science & Engineering

Natural Language Processing is a subfield of Computer Science that inspects the intersection of computation and human language. Idiomatic language is in turn ubiquitous in human language—when someone is a "diamond in the rough", they are not literally an uncut diamond, rather, they are a person whose goodness is hidden by their surface appearance. However, a "diamond in the rough" is a phrase that can literally mean an uncut diamond—this is to say, it has two different senses, one that is idiomatic and one that is literal. Given a dictionary entry, a human can easily distinguish between idiomatic and literal definitions, however, doing this in an automatic fashion is difficult because it requires asking whether the meaning represented in the definition corresponds with the literal meaning of the phrase. This research leverages Wiktionary, an extremely large, collaboratively authored dictionary, to perform idiom identification in a scalable manner through machine learning algorithms. To do this identification, we are developing two sets of features—traits that we describe a definition-phrase pair with—selectional preference features and graph-based features. Selectional preference features describe traits of language that are used in a way that violates its literal meaning while graph-based features describe where the entry occurs in relation to the other pages within Wiktionary (what links does this page contain?). We are using two kinds of machine learning algorithms for idiom identification—supervised methods and semi-supervised methods. Supervised methods allow us to verify the validity of the features that we develop, and semi-supervised methods allow us to use these features for knowledge discovery because they are ideal for discovering definitions that are not yet marked as idiomatic but that ought to be.

SESSION 2A

GRAPHS AND GEOMETRY

Session Moderator: Werner Stuetzle, Statistics

085 MGH

3:45 PM to 5:15 PM

* Note: Titles in order of presentation.

Stable Marriage when You are Not Sure Who Would Make a Good Long-Term Partner

Eric Christopher Lei, Senior, Computer Science, Economics, Mathematics
Mary Gates Scholar
Mentor: Anna Karlin, Computer Science and Engineering

Can a group of men and women be paired together to guarantee no divorces or scandals? In 1962, two economists posed this question mathematically and offered a solution. A 2012 Nobel Prize went to two of the key researchers that worked on this problem. Their solution took the form of an algorithm, a systematic procedure that could be performed by a computer. The solution to the "stable marriage problem" has both interesting theoretical properties and real-world applications, such as the assignment of medical students to hospitals. In the original version of the problem, each man and woman know their preferences over all their potential partners. We are considering a variant in which they only know some of their preferences. The algorithm not only needs to figure out whom to match with whom, but also what information to gather on preferences. We present our preliminary results on efficient algorithms for this problem. This is joint work with Anna Karlin and Jamie Morgenstern.

SESSION 2G

MICRO- AND NANO-MATERIALS IN ACTION

Session Moderator: John Berg, Chemical Engineering

242 MGH

3:45 PM to 5:15 PM

* Note: Titles in order of presentation.

Battery-Free Gas Sensor Nodes Utilizing Ambient Radio Frequency Energy
Chen Shi, Senior, Bioengineering, Electrical Engineering
Mary Gates Scholar
Mentor: Joshua Smith, Computer Science & Engineering, Electrical Engineering
Mentor: Aaron Parks

Gas sensors are widely used in daily life and industry. An important application of gas sensing is the monitoring of environmental factors affecting health, such as the concentration of carbon monoxide, in populated areas. However, most conventional gas sensors are powered by batteries, which need periodic replacement. The goal of this project is to integrate a Wireless Ambient Radio Power (WARP) energy harvesting platform with a new class of amperometric electrochemical gas sensors, provided by KWJ Engineering, Inc, to create novel battery-free gas sensor nodes. The WARP sensing platform, developed in Dr. Joshua Smith's group, utilizes ambient radio frequency (RF) energy from common sources such as cellular towers and TV broadcast stations, which provide a reliable and pervasive 24-hour power source. The gas sensors manufactured by KWJ Engineering possess the advantages of low power, low cost, high sensitivity, and high selectivity. Particularly, the low power requirements of the gas sensors make it possible for them to be powered by the RF energy harvested by the WARP platform. Currently the potentiostat circuit needed to properly bias the gas sensor and acquire the gas concentration is being developed. The gas sensor circuitry will be integrated with the WARP platform to produce the battery-free RF-powered gas sensor nodes, followed by system optimization for reliable and efficient operations. With such gas sensor nodes, long-lived wireless sensor networks with zero maintenance cost could be deployed in continuous toxic gas monitoring applications, including air quality monitoring in cities and process control in industry.

POSTER SESSION 3

Commons East, Easel 72
2:30 PM to 4:00 PM

Studying Logged Behavior with Inferred Models: Utilizing State in Model Inference
Roykrong Sukkerd, Senior, Computer Science
Mentor: Michael Ernst, CSE
Mentor: Ivan Beschastnikh, Computer Science & Engineering

Software systems are often difficult to comprehend. When a system behaves in an unexpected manner, or when developers must modify the code written by someone else, they have the challenge of understanding the system behavior. A typical way developers use to gain insight into a system is logging. Logging is a method that captures program activities and states during an execution into a log file (a text file), which developers then inspect. Unfortunately, the size and complexity of logs are often too large for manual inspection. Synoptic is a tool that supports the task of inspecting execution logs. Synoptic takes a log as input and infers a finite state machine model of the process that generated the log. The model that Synoptic produces is compact and captures essential behavior traits of the subject system. However, there is a missing piece of information. Synoptic is only interested in system activities, thus the models it infers are activity-based models. System states, which are also often captured in a log, are ignored. This project presents an approach to utilize state in model inference. We extend Synoptic to model data values in a subject system by annotating states in the inferred model with conditions on data values. We use Daikon, an implementation of dynamic detection of likely invariants, to determine those conditions from an input log that captures data values at each state of the system. We argue that this new model helps developers reason about their systems better.

POSTER SESSION 3

Commons West, Easel 19
2:30 PM to 4:00 PM

Enhancing Education Efficacy through the Utilization of Video Games
Geoff Phillips, Senior, Physics, Applied & Computational Mathematical Sciences (Engineering & Physical)
McNair Scholar
Mentor: Marty Stepp, Computer Science & Engineering

Video games represent a powerful educational paradigm; nevertheless, strong devotion to convention as well as the highly stigmatized reputation of games as a purely recreational medium, have led to an overarching reluctance to implement them within high-level course curricula. Emerging research on the use of real-time interactive simulations as educational tools consistently shows marked improvement in learning outcomes, however, a rigorous means of quantitative analysis is lacking. To date, all studies correlating learning outcomes with the intrinsic components of video games rely on one of three different methods of data acquisition: surveys, pre-gameplay versus post-gameplay evaluations, or recorded observation of play activity. Each of these methods has its own inherent limitations on the amount and quality of data obtainable. In addition, these methods prohibit researchers from being able to dynamically alter any of the key control variables such as: challenge, contextual motivation, or narrative. This study proposes an alternative model, capable of quantitatively evaluating the educational efficacy of implementing each variable. This is accomplished by embedding the data collection methods and algorithms directly into the framework of the game itself. This would further allow

for dynamic alterations of each variable according to the parameter requirements of the inherent statistical methods. We expect the result to be a direct and measurable correlation between student learning outcomes and the modified variables. Also central to this study is the imbedding of real-time data collection functions and processing algorithms, which can automatically save and upload data to a server. This would drastically increase the possible number of participants—thereby increasing the sample size—while simultaneously allowing for greater modularity of data analysis. The primary long-term goal of this research is to develop a model and educational game that implements this framework and affords future studies the tools to overcome the aforementioned limitations in data acquisition.

POSTER SESSION 4

Commons East, Easel 59

4:15 PM to 5:45 PM

Developing a Model for Noise in Stochastic Gene Expression

Shane A (Shane) Colburn, Senior, Physics: Comprehensive Physics, Electrical Engineering
NASA Space Grant Scholar
Mentor: Georg Seelig, Electrical Engineering and Computer Science & Engineering
Mentor: Alex Rosenberg, Electrical Engineering

Cellular noise is the variability in a biological quantity due to randomness. It is mathematically defined as the quotient of the quantity's standard deviation and mean. Systems of genotypically identical cells may exhibit noisy gene expression—there are often differences between cells in terms of protein abundance. We are interested in developing a model to better understand this noise. Many factors contribute to gene expression noise, one of which is alternative splicing. RNA splicing is when sequences of RNA are "cut out", leaving only the regions that will be translated into protein. Alternative splicing is when pre-mRNA transcripts are spliced differently, resulting in similar yet different specialized proteins. This is common in eukaryotes and increases protein diversity, allowing cells to take on particular functions and thereby playing a role in cellular decision making. To model the noise, we began by cloning different plasmids, each of which had two fluorescent reporter genes, citrine and mCherry. To better understand the alternative splicing mechanism, one of our plasmids contained an extra sequence of bases, an intron, in citrine. We used our two reporters to produce fluorescence readouts of correct versus incorrect splicing. We measured the fluorescence with a flow cytometer and the noise could be quantified by determining the spread and mean of the data. We hypothesize that as mRNA concentration increases, noise in expression will decrease, since variation decreases with increasing sample size. We also hypothesize that the plasmid with the intron will be noisier than that without since the alternative splicing mechanism will have a greater length sequence to splice. Alternative splicing will sometimes produce abnormal transcripts that are found to be in high concentration in cancerous cells. So in modeling the noise and gaining a better understanding of this mechanism, we may improve understanding of how cancer develops.

POSTER SESSION 4

Commons East, Easel 68

4:15 PM to 5:45 PM

Turing Machines and Computation

Mark Bennett, Junior, Mathematics, Seattle Central College
Reynaldo Maray Siu Chang, Junior, Pre Engineering
Mentor: Francois Lepeintre, Science and Math, Seattle Central Community College
Mentor: Bryan Johns, Mathematics, Seattle Central Community College

In the 20th century, mathematicians attempted to reduce mathematics to a set of axioms. One of the properties that David Hilbert, a famous mathematician, postulated was decidability. This refers to an algorithm that can determine whether or not another algorithm in the same system is actually computable. In his 1939 paper, Alan Turing proves decidability cannot occur in formal mathematical systems. Turing used an unconventional way to prove his argument—the Turing Machine. The machine is a theoretical computational device. This abstract device has revolutionized the way we think of computing and computers among other fields. It consists of four parts: a head that reads symbols, an infinite tape that stores inputs and outputs, a state register that holds the current state (or state of mind if you prefer), and a table that describes how the machine transitions from state to state. Turing machines can compute anything that is computable. Computability equates to being able to describe how to obtain the solution of some function with an algorithm. This abstract idea is interdisciplinary; which makes it a great educational tool. We have built a working Turing machine out of LEGOS. To both extend and improve functionality, we have added various components. These components cross over to other fields of study making the project great for students of various education paths to reproduce in a team setting. Our hope is that others can use our work and further devolop it. Besides providing a concrete example of an abstract idea, building the Turing machine allows us to reaffirm findings. It is in the spirit of scientific descovery to reproduce work. This research project allows us to follow in the foot steps of intellectual giants.

POSTER SESSION 4

Commons East, Easel 53

4:15 PM to 5:45 PM

3D Intuition of Computers Based on 2D Images

X (Fred) Li, Sophomore, Mathematics (Comprehensive), Physics: Comprehensive Physics
Mentor: Nell Tack, CSE

Current computers' knowledge in our 3D world is mainly based on "depth data" collected form reflection of radiations (like infared or laser). Such data give the measurement the distance between points in space and can be used to reconstruct the original environment in 3D, hence enable computers to get an idea of what our world "looks" like. The unsatisfying points would be it takes time and considerate amount of memory to get such depth data and computers don't "learn" during such process. I've been looking into an alternative way to teach computers to view our world in 3D, that is through 2D images. This is inspired by the fact that human can understand 3D world through 2D information collected by our eyes. In my research, I first train a computer to pull out geometric information in 2D images, then relate such 2D geometry to a 3D geometry data base to get 3D geometric structure. The decision making through the process is controlled by supervised learning. The goal of this reasearch is to let computers see our 3D world through normal cameras without the necessary help from extra visual aids and train them to become an intelligent helper intead of just a mere measuring tool.