DePaul University – College of Computing and Digital Media

School of Computing - Research Colloquium Spring 2011

organized by Jose P. Zagal

April 1

Adam Czauderna and Marek Gibiec, DePaul University

A machine learning approach for tracing regulatory codes to product specific requirements

Regulatory standards, designed to protect the safety, security, and privacy of the public, govern numerous areas of software intensive systems. Project personnel must therefore demonstrate that an as-built system meets all relevant regulatory codes. Current methods for demonstrating compliance rely either on after-the-fact audits, which can lead to significant refactoring when regulations are not met, or else require analysts to construct and use traceability matrices to demonstrate compliance. Manual tracing can be prohibitively time-consuming; however automated trace retrieval methods are not very effective due to the vocabulary mismatches that often occur between regulatory codes and product level requirements. Our research introduced and evaluated two machine-learning methods, designed to improve the quality of traces generated between regulatory codes and product level requirements. The first approach uses manually created traceability matrices to train a trace classifier, while the second approach uses web-mining techniques to reconstruct the original trace query. The techniques were evaluated against security regulations from the USA government's Health Insurance Privacy and Portability Act (HIPAA) traced against ten healthcare related requirements specifications. Results demonstrated improvements for the subset of HIPAA regulations that exhibited high fan-out behavior across the requirements datasets.

Adam Czauderna
Adam graduated from DePaul University in March 2010 receiving Master degree at Computer Science. For last two years worked as Research Assistant at Systems and Requirements Engineering Center (SAREC) directed by prof. Jane Cleland-Huang. During this time Adam had two publications together with Jane Cleland-Huang and Marek Gibiec, one at ICSE 2010 in CapeTown "A machine learning approach for tracing regulatory codes to product specific requirements", and the other at ASE 2010 in Belgium "Towards mining replacement queries for hard-to-retrieve traces". Both paper received the distinguished paper awards. Currently Adam works at CDM DePaul as researcher and application developer in development of a Software Traceability Instrument to Facilitate and Empower Traceability Research and Technology Transfer funded National Science Foundation Award.

Marek Gibiec
Marek graduated from DePaul in November 2010 and received a master's degree in Computer Science. During his studies he worked as a Research Assistant at Systems and Requirements Engineering Center directed by prof. Jane Cleland-Huang. His research

involves utilizing machine learning, data mining and online information retrieval techniques to enhance automated traceability. Currently, Marek continues his work at SAREC in addition to working as Android Developer at local mobile app software shop.

April 8

Fabian Bustamante, Northwestern University

Crowdsourcing Broadband Service Characterization to the Network Edge

Evaluating and characterizing Internet Service Providers (ISPs) is critical to subscribers shopping for alternative ISPs, companies providing reliable Internet services, and governments surveying the coverage of broadband services to its citizens. Ideally, ISP characterization should be done (i) at scale, to capture the diversity of available providers and their services, (ii) continuously, to capture dynamic changes due to network management policies and unscheduled events, and (iii) from end users, to ensure accuracy. While there has been significant progress toward this end, existing approaches exhibit an apparently unavoidable tradeoff between coverage, continuous monitoring and capturing user-perceived performance.

In this talk, I argue that network-intensive applications running on end systems avoid these tradeoffs, offering an ideal platform for ISP characterization. Based on data collected from 500,000 peer-to-peer BitTorrent users, I show the feasibility of this approach to characterize the service subscribers can expect from a particular ISP, discuss some of the remaining open issues, and present an implementation of our ideas as an extension to a popular BitTorrent client. Our extension, code named Dasu, has been available for a few months and has already been adopted by over 40,000 users worldwide.

Bio:
Fabián E. Bustamante is an associate professor of computer science in the EECS Department at Northwestern University. He joined Northwestern in 2002, after receiving his M.S. and Ph.D. from the College of Computing at Georgia Tech. Fabián leads the AquaLab group investigating systems issues with Internet-scale distributed computing. Fabián is a recipient of the US National Science Foundation CAREER award and the E.T.S. Watson Fellowship Award from the Science Foundation of Ireland, and a senior member of both the ACM and the IEEE. For more detailed information and a list of publications, please visit: http://www.aqualab.cs.northwestern.edu.

April 15

Massimo Di Pierro, DePaul University

Looking Inside Protons

In this talk we present an overview of Lattice Quantum Chromodynamics (LQCD), a computational approach to the mathematical model that describes quarks and gluons, the basic constituents of protons, neutrons and the majority of known matter. LQCD consists of discretizing a small portion of space and time and simulating all possible time evolutions of the fields within. LQCD can successfully predict properties of composite particles (such as their mass and average lifetime) within two standard deviations. LQCD is the most computational expensive project in Physics. It is an international effort. In US it is supported by the Department of Energy.

Bio:
Massimo is an Associate Professor of Computer Science in the School of Computing of DePaul University in Chicago. He has taught Object Oriented Programming, Foundations of Computer Science, Analysis and Design of Algorithms, Information Assurance, Secure Electronic Commerce, Network Programming, Parallel Algorithms, Scientific Computing, Monte Carlo Simulations and Web Frameworks.

Massimo holds a PhD in High Energy Physics from the University of Southampton in UK. He is the lead developer and maintainer of fermqcd (fermiqcd.net) and web2py (http://www.web2py.com). Massimo is a reviewer for the American Mathematical Society and an Editor of Computing in Science and Engineering (ACM). His expertise is in numerical algorithms and application to Computational Physics and Computational Finance.

April 22

University Closed

April 29

Leilah Lyons, University of Illinois Chicago

Facilitating Shared Learning with Shared Interfaces

Dr. Lyons' work has explored how different user interface strategies and different activity designs affect how learners externalize the knowledge they gain by interacting with shared complex system simulations and complex data sets. She will give an overview of several of

her past and ongoing research projects, which make use of strategies like role playing, embodiment, and tangible user interfaces to encourage learners to coordinate their actions and share what they know while engaged in collaborative, exploratory learning experiences. These learning experiences have been designed for use in both informal environments like science museums and history museums, and formal environments like classrooms, and have engaged learners in interactions with a variety of activities like participatory simulations, complex system simulations, and rich interactive historical data sets.

Bio:
Leilah Lyons is an Assistant Professor at the University of Illinois at Chicago, where she has a dual appointment in Computer Science and the Learning Sciences. Her main area of research interest is how technology can be used to mediate face-to-face collaborative learning in informal settings, like museums and zoos. Dr. Lyons obtained a PhD in Computer Science and a Graduate Certificate in Museum Studies from the University of Michigan.

May 6

Amon Millner, Olin College of Engineering

Creating Toolkits that Enable Youth to Design with Physical and Digital Media

Dr. Millner's research has involved developing tools and activities that enable young people to design tangible interfaces - projects that combine physical and digital media. In this talk, he will give an overview of his ongoing research, which he conducts at the intersection of human computer interaction, tangible user interface design, community organizing, and the learning sciences. He will present the trajectories of two of his research platforms: the Scratch Sensor Board and Modkit. The talk will focus on approaches to developing and deploying computational construction kits in informal learning environments. The presentation will feature example work from youth that demonstrate how the platforms (and associated activities) are evolving as tools to support young people in creating graphical computer programs and making sensors to control those programs from a variety of physical materials.

Bio:
Dr. Amon Millner is a Computing Innovation Fellow at the Franklin W. Olin College of Engineering. He completed his Ph.D. at MIT's Media Lab. He helped design the award-winning NSF-supported Scratch programming language, leading the efforts to incorporate sensors and software. He has designed, deployed, and evaluated computational construction kits, honing expertise in developing technological tools and crafting engaging activities that enable young people from diverse backgrounds to create novel tangible user interfaces. His degrees include an M.S. from Georgia Tech in Human Computer Interactions and a B.S from USC in Computer Science.

May 13

Iyad Kanj, DePaul University

On the independence number of graphs with maximum degree 3

Let G be an undirected graph with maximum degree at most 3 such that G does not contain any of the three graphs shown in Figure 1 as a subgraph. We prove that the independence number of G is at least n(G)/3 + nt(G)/42, where n(G) is the number of vertices in G and nt(G) is the number of nontriangle vertices in G. This bound is tight as implied by the well-known tight lower bound of 5n(G)/14 on the independence number of triangle-free graphs of maximum degree at most 3.

We show an algorithmic application of the aforementioned combinatorial result to the area of parameterized complexity. We present a linear-time kernelization algorithm for the independent set problem on graphs with maximum degree at most 3 that computes a kernel of size at most 140 k/47 < 3k, where k is the given parameter. This improves the known 3k upper bound on the kernel size for the problem, and implies a 140k/93 lower bound on the kernel size for the vertex cover problem on graphs with maximum degree at most 3.

This is a joint work with Fenghui Zhang.

Bio:
Iyad Kanj is an associate professor of computer science at the school of computing (CDM). He obtained his Ph.D. in computer science from Texas A & M University in 2001. His research interests include parameterized complexity, graph theory and algorithms, combinatorial optimization, computational geometry and its applications to wireless computing, and computational biology. More information about his research can be found at: http://facweb.cs.depaul.edu/ikanj/

May 20

CP Jois, Higher Education Company

Testing the limits of cluster computing in the development of low-cost flight simulation environments

The concept of simulation is very prevalent in many different fields. Simulation is the process of modeling some real thing, state of affairs, or process. The act of simulating something generally entails representing certain key characteristics or behaviors of a selected physical or abstract system. Its compelling power has been leveraged for a variety of reasons including modeling new concepts, prototyping new models, experimenting with

new algorithms, creating non-destructive, cost-effective environments for training and, more recently for an entirely new domain known as gaming and entertainment.

Flight simulation as a concept has existed since the Wright brothers invented a physical artifact that allowed them to demonstrate the dynamics of their invention. That invention later came to be known as a flight simulator. In the field of aeronautics, a flight simulator is used to train pilots on the ground. Over the past 3 decades, flight simulators, like many other forms of simulation, have slowly but surely made their entry into the domain of entertainment. However, simulators in the gaming and entertainment realm have lacked fidelity and traditionally lacked the levels of realism that commercial simulators provide. Two key drivers of gaps in experience have been found to be related to (a) computing power and (b) the constraints around provisioning the requisite visual and haptics environment.

Two important criteria that impact fidelity are levels of visual and control realism that a simulation environment renders. This study was largely empirical and driven by design experiments. It was guided by important questions around whether high fidelity simulation can be achieved through harnessing the power of clustered computers - essentially the cost/fidelity tradeoffs that apply.

This presentation focuses on the work that has been done till date. It concludes by introducing threads for future research.

Bio:
Dr. C P Jois is the Vice President of Enterprise Architecture at a higher education company. He also heads up product engineering, quality and CEC's center for accelerated research and prototyping – the Labs. He has 20+ years of technology management experience spanning multiple industry verticals performing various roles in software engineering, program management, IT strategy and architecture.

Dr. Jois focuses on software economics, enterprise architecture, platform interoperatibility, component/services driven product development, enterprise integration strategies, portfolio governance and information intelligence maturity. He is a technology evangelist who passionately pursues the power of innovation for business value, writes regularly on the compelling power of transformational technologies and speaks at various forums.

He has spent many years helping organizations deal with challenges around software asset reuse, software quality and portfolio refactoring.

He bears a keen interest in simulation technology and has spent most of his personal time over the past 9 years researching the potential for leveraging cluster computing as a means of provisioning high-fidelity flight simulator appliances.

Dr. Jois has broad-based experience with Fortune 500 organizations and has worked with clients in Banking, Aviation, Transportation, Hospitality, Insurance, & Manufacturing and has extensive experience in offshore outsourcing engagements.

May 27

Brian Dorn, University of Hartford

Supporting Informal Computing Education for End-User Programmers

Software development is no longer a task limited to professionally-trained computer programmers. Increasing support for software customization through scripting, the availability of application programmer interfaces on the Web, and a growing need for domain-specific application support have all contributed to an increase in end-user programming. Unfortunately, learning to program remains a challenging task, and the majority of end-user programmers lack any formal education in software development. In this talk I will provide an overview of our ongoing research to understand the unique challenges faced by self-taught programmers in the Web and Graphic Design disciplines and our efforts to design tools that scaffold their informal, Web-centric learning practices.

Bio:

Dr. Dorn is an assistant professor at the University of Hartford, holding a joint appointment with Computer Science and Multimedia Web Design & Development. His research interests lie at the intersection of HCI and Computing Education. He received his Ph.D. in computer science from Georgia Tech's School of Interactive Computing where he was also a member of the GVU Center. He also holds an M.S. from Iowa State University and a B.S. from Northwest Missouri State University.

June 3

Jose Zagal, DePaul University

Computing for Games (Research)

Videogames and computers are undeniably and unambiguously connected. Some have argued, in fact, that videogames have been the drivers for computer science research in a variety of fields including graphics, artificial intelligence, networking, and more. In this talk I will present a brief overview of work that explores a different relation between videogames and computer science: using computing to support research in and about games. What are some of the ways that computing can help us learn more about games as a creative medium, as a form of culture, and as a driver for new research methods and tools in games research? I will describe joint work with Dr. Noriko Tomuro using natural language processing for analyzing game reviews written by fans and videogame

aficionados as well as preliminary work developing tools and systems for providing meaningful support for game designers.

Bio:
Dr. José P. Zagal is a game designer and scholar. He serves on the faculty at DePaul University's College of Computing and Digital Media where he teaches a variety of courses on game design and analysis, online communities, and ethics. His research work explores the development of frameworks for describing, analyzing, and understanding games from a critical perspective. He is also interested in supporting games literacy through the use of collaborative learning environments. His book on this topic, "Ludoliteracy: Defining, Understanding, and Supporting Games Education" was published in 2010. Zagal is a member of the executive board of the Digital Games Research Association (DiGRA) and serves on the editorial boards of Games & Culture, the International Journal of Gaming and Computer-Mediated Simulations, and Loading, the Journal of the Canadian Gaming Studies Organization. José received his PhD in computer science from Georgia Institute of Technology in 2008, his M.Sc. in engineering sciences and a B.S. in industrial engineering from Pontificia Universidad Católica de Chile in 1999 and 1997.

Further details on his work are available at: http://facsrv.cs.depaul.edu/~jzagal/