

# Paper Abstracts

ACSW 2009  
Wellington, New Zealand

19-23 January, 2009

## Contents

|   |   |    |
|---|---|----|
| 1 | ACSC: The 32nd Australasian Computer Science Conference                         | 2  |
| 2 | ACE: Australasian Computing Education Conference                                | 10 |
| 3 | ADC: Australasian Database Conference   | 19 |
| 4 | AISC: Australasian Information Security Conference                              | 28 |
| 5 | APCCM: Asia-Pacific Conference on Conceptual Modelling                          | 33 |
| 6 | AUIC: Australasian User Interface Conference                                    | 40 |
| 7 | AUSGRID: Australasian Symposium on Grid Computing and<br>e-Research             | 44 |
| 8 | CATS: Computing - The Australasian Theory Symposium                             | 49 |
| 9 | HIKM: Australasian Workshop on Health Informatics and Knowl-<br>edge Management | 57 |

# 1 ACSC: The 32nd Australasian Computer Science Conference

## **ALARM: An Adaptive Load-Aware Routing Metric for Hybrid Wireless Mesh Networks**

*Asad Pirzada, Ryan Wishart, Marius Portmann and Jadwiga Indulska*

**Tuesday, 1:30pm, Soundings Theatre**

**Abstract** Hybrid Wireless Mesh Networks (WMN) are generally employed to establish communication during disaster recovery operations. The Hybrid WMN network is formed in a spontaneous manner when wireless nodes belonging to different agencies are operated in the disaster area. These wireless nodes generally have heterogeneous configurations in terms of number of transceivers, computational power, battery resources and mobility pattern. Internet Protocol (IP) acts as the common platform for integrating these heterogeneous devices. Routing protocols are engaged to determine paths between sets of wireless nodes. A number of routing metrics have been developed to achieve performance gains in multi-radio, multi-hop WMNs. However, most of these metrics need access to external information like link quality statistics, channel numbers, etc on a regular basis. This frequent information exchange coupled with the incessant variation in mobility and traffic load conditions causes degraded performance of these metrics in Hybrid WMNs. In this paper, we present a routing metric named ALARM specifically designed for Hybrid WMNs. ALARM is computed using the number of packets queued per wireless interface. This computed value offers an accurate representation of the traffic load, link quality, interference and noise levels. With the help of extensive simulations, we show that our routing metric outperforms well-know routing metrics like ETT and WCETT under varying mobility and traffic load conditions in Hybrid WMNs. We further show the practicality of the metric through a prototype implementation and provide performance results obtained from a small-scale testbed deployment.

## **The Impact of Quanta on the Performance of Multi-level Time Sharing Policy under Heavy-tailed Workloads**

*Malith Jayasinghe, Zahir Tari and Panlop Zeephongsekul*

**Tuesday, 2:00pm, Soundings Theatre**

**Abstract** The rationale for using the multi-level time sharing policy is that it can minimise both waiting time and slowdown of jobs that require relatively small service requirements. In a heavy-tailed distribution as there are large number of such small jobs, by improving the performance of these small jobs will improve the overall performance. In this paper, we investigate the effect of quanta on the overall performance of the multi-level time sharing policy under a range of workloads and task size variabilities. We measure the performance using slowdown and flow time. First, we show that for most workloads and task

size variabilities there exists a unique set of quanta (optimal set of quanta) that would result in the best performance. The set of quanta that would produce the best overall expected flow time, however, may not be unique for a few workloads and task size variabilities. Second, we investigate the performance degradation in one metric under the optimal parameters (i.e quanta) of other metric. We find that under high system loads and task size variabilities using the optimal set of quanta corresponding to overall expected slowdown can result in the overall expected flow time to deteriorate significantly (250%).

**Scheduling Parallel Applications on Utility Grids: Time and Cost Trade-off Management**  
*Saurabh Kumar Garg, Rajkumar Buyya and H. J. Siegel*  
Tuesday, 2:30pm, Soundings Theatre

**Abstract** With the growth of Utility Grids and various Grid market infrastructures, the need for efficient and cost effective scheduling algorithms is also increasing rapidly, particularly in the area of meta-scheduling. In these environments, users not only may have conflicting requirements with other users, but also they have to manage trade-off between time and cost such that their applications can be executed most economically in the minimum time. Thus, choosing of the best Grid resource becomes a challenge in such a competitive market. This paper presents two novel heuristics for scheduling parallel applications on Utility Grids that manage and optimize trade-off between time and cost constraints. The performance of the heuristics is evaluated through extensive simulations of a real-world environment with real parallel workload models to demonstrate the practicality of our algorithms. We compare our scheduling algorithms against other common algorithms used by current meta-schedulers. The results shows that our algorithms outperform other algorithms by minimizing the time and cost of application execution on Utility Grids.

**A Citation Analysis of the ACSC 2006 2008 Proceedings, with Reference to the CORE Conference and Journal Rankings**

*Raymond Lister and Ilona Box*  
Tuesday, 3:00pm, Soundings Theatre

**Invited Talk Abstract** This paper compares the CORE rankings of computing conferences and journals to the frequency of citation of those journals and conferences in the Australasian Computer Science Conference (ACSC) 2006, 2007 and 2008 proceedings. The assumption underlying this study is that there should be a positive relationship between citation rates and the CORE rankings. Our analysis shows that the CORE rankings broadly reflect the ACSC citations, but with some anomalies. While these anomalies might be minor in the larger scheme of things, anomalies need to be addressed, as the careers of individual academics may depend upon it. Rankings are probably here to stay,

and this paper ends with some suggestions on how the rankings process should now evolve, so that it becomes more transparent.

## **Privacy-aware Access Control with Generalization Boundaries**

*Min Li, Hua Wang and Ashley Plank*

**Wednesday, 9:00am, Soundings Theatre**

**Abstract** Privacy is today an important concern for both data providers and data users. Data generalization can provide significant protection of an individual's privacy, which means the data value can be replaced by a less specific but semantically consistent value and the personal information can be collected in a generalized form. However, over-generalized data may render data of little value. A key question is whether or not a certain generalization strategy provides a sufficient level of privacy and usability? In this paper, we introduce a new approach, called privacy-aware generalization boundaries, which can satisfy the requirements of both data providers and data users. We propose a privacy-aware access control model related to a retention period. Formal definitions of authorization actions and rules are presented. Further, we discuss how to manage a valid access process and analysis the access control policy. Finally, we extend our model to support highly complex privacy-related policies by taking into account features of obligations and conditions.

## **Microdata Protection Through Approximate Microaggregation**

*Xiaoxun Sun, Hua Wang and Jiuyong Li*

**Wednesday, 9:30am, Soundings Theatre**

**Abstract** Microdata protection is a hot topic in the field of Statistical Disclosure Control, which has gained special interest after the disclosure of 658000 queries by the America Online (AOL) search engine in August 2006. Many algorithms, methods and properties have been proposed to deal with microdata disclosure. One of the emerging concepts in microdata protection is k-anonymity, introduced by Samarati and Sweeney. k-anonymity provides a simple and efficient approach to protect private individual information and is gaining increasing popularity. k-anonymity requires that every record in the microdata table released be indistinguishably related to no fewer than k respondents. In this paper, we apply the concept of entropy to propose a distance metric to evaluate the amount of mutual information among records in microdata, and propose a method of constructing dependency tree to find the key attributes, which we then use to process approximate microaggregation. Further, we adopt this new microaggregation technique to study k-anonymity problem, and an efficient algorithm is developed. Experimental results show that the proposed microaggregation technique is efficient and effective in the terms of running time and information loss.

## **HOPPER: a Hierarchical Planning Agent for Unpredictable Domains**

*Maciej Wojnar and Peter Andreae*

**Wednesday, 1:30pm, Soundings Theatre**

**Abstract** Hierarchical Task Networks (HTNs) are a family of powerful planning algorithms that have been successfully applied to many complex, real-world domains. However, they are limited to predictable domains. In this paper we present HOPPER (Hierarchical Ordered Partial-Plan Executor and Re-planner), a hierarchical planning agent that produces partial plans in a similar way to HTNs but because it interleaves planning and execution it can handle unexpected events in unpredictable domains. HOPPER can detect and recover from unexpected events that invalidate the plan, and it can detect and exploit unexpected opportunities both serendipitously and by interleaving decompositions.

## **Inference of Gene Expression Networks Using Memetic Gene Expression Programming**

*Armita Zarnegar, Peter Vamplew and Andrew Stranieri*

**Wednesday, 2:00pm, Soundings Theatre**

**Abstract** In this paper we aim to infer a model of genetic networks from the time series data of gene expression profiles by using a new gene expression programming algorithm. Gene expression networks are modelled by differential equations which represent temporal gene expression relations. Gene Expression Programming is a new extension of genetic programming. Here we combine a local search method with gene expression programming to form a memetic algorithm in order to find not only the system of differential equations but also fine tune its constant parameters. The effectiveness of the proposed method is justified by comparing its performance with that of conventional genetic programming applied to this problem in previous studies.

## **Using Chronological Splitting to Compare Cross- and Single-company Effort Models: Further Investigation**

*Chris Lokan and Emilia Mendes*

**Wednesday, 4:00pm, Soundings Theatre**

**Abstract** Numerous studies have used historical datasets to build and validate models for estimating software development effort. Very few used a chronological split (where projects end dates are used so that training sets only contain projects that were completed before the start date of each project in the validation set), and only one compared chronological split to random split. Therefore the aim of this study is to investigate further and compare the use of chronological and random splitting. We do so in the context of comparing cross-company and single-company models for effort estimation. We used 450 single-company

projects and 741 cross-company projects from the ISBSG Release 10 repository, and estimates were obtained using manual stepwise regression. We found that with these data the use of chronological splitting, and different splitting dates, did not affect prediction accuracy. We were not able to obtain a converging set of findings when comparing cross- to single-company predictions given that different accuracy measures presented contradictory results.

## **Verification of the SIP Transaction Using Coloured Petri Nets**

*Lin Liu*

**Wednesday, 4:30pm, Soundings Theatre**

**Abstract** The Session Initiation Protocol (SIP) is one of the leading protocols for multimedia control over the Internet, including initiating, maintaining and terminating multimedia sessions. The protocol makes use of SIP transactions to complete the control tasks. In this paper we focus on the INVITE transaction, the transaction that is used to initiate a session. SIP is designed to operate over a transport protocol that is either reliable or unreliable. Our previous work has verified the functional properties of the INVITE transaction over a reliable transport medium, using Coloured Petri Nets (CPNs). In this paper, we use CPNs to model and analyse SIP INVITE transaction when the medium is unreliable. The verification reveals similar problem as that in the case of a reliable medium, i.e. the transaction may terminate in an undesirable state while one communication party is still waiting for a response from its peer. Moreover, there are undesirable terminal states in which retransmitted requests may lead to erroneous operation. This result provides theoretical evidence and timely support for the Internet Draft that has been recently submitted to the Internet Engineering Task Force to propose updates to SIP INVITE transaction.

## **Discovering Itemset Interactions**

*Ping Liang and John Roddick*

**Thursday, 9:00am, Soundings Theatre**

**Abstract** Itemsets, which are treated as intermediaries in association mining, have attracted signi-

cant research due to the inherent complexity of their generation. However, there is currently little literature focusing upon the interactions between itemsets, the nature of which may potentially contain valuable information. This paper presents a novel tree-based approach to discovering itemset interactions, a task which cannot be undertaken by current association mining techniques.

**A ConceptLink Graph for Text Structure Mining**  
*Rowena Chau, Ah Chung Tsoi, Markus Hagenbuchner  
and Vincent Lee*  
**Thursday, 9:30am, Soundings Theatre**

**Abstract** Most text mining methods are based on representing documents using the vector space model, where each document is modeled as a linear vector representing the occurrence of a bag of independent words. It is well known that using this vector-based representation, important information, such as semantic relationship among concepts, is lost. This paper proposes a novel text representation model called ConceptLink graph. The ConceptLink graph does not only represent the content of the document, but also captures its semantic structure in terms of the concept relationships. The ConceptLink graph is constructed in two main stages. First, we find a set of concepts by clustering conceptually related terms using self-organizing map method. Second, by mapping each documents content to concept, we generate a graph of concepts based on the occurrences of concepts using singular value decomposition. The ConceptLink graph will overcome the keyword independence limitation of the vector space model to take advantage of the implicit concept relationship exhibit in all natural language texts. As an information-rich text representation model, the ConceptLink graph will advance text mining technology beyond feature-based to structure-based knowledge discovery. We will illustrate the ConceptLink graph method using samples generated from benchmark text mining dataset.

**Evaluating the Dynamic Behaviour of Python Applications**  
*Alex Holkner and James Harland*  
**Thursday, 1:30pm, Soundings Theatre**

**Abstract** The Python programming language is typical among dynamic languages in that programs written in it are not susceptible to static analysis. This makes efficient static program compilation difficult, as well as limiting the amount of early error detection that can be performed. Prior research in this area tends to make assumptions about the nature of programs written in Python, restricting the expressiveness of the language. One may question why programmers are drawn to these languages at all, if only to use them in a static-friendly style. In this paper we present our results after measuring the dynamic behaviour of 24 production-stage open source Python programs. The programs tested included arcade games, GUI applications and non-interactive batch programs. We found that while most dynamic activity occurs during program startup, dynamic activity after startup cannot be discounted entirely.

## **Improving the Transparency of Proxy Injection in Java**

*Hendrik Gani and Caspar Ryan*

**Thursday, 2:00pm, Soundings Theatre**

**Abstract** Due to their flexibility, proxies have been used for various purposes in standalone and distributed applications. More specifically, object-level proxies support fine grained capabilities and offer the potential to transparently (i.e. with minimal human intervention) inject proxy-based functionality into an ordinary application. Consequently, several solutions based upon code transformation for addressing various limitations of transparency in existing approaches are considered and evaluated. Furthermore, the automation of the majority of the required code transformation has been implemented, which along with the deployment of various (proxy-based) adaptive and test applications, demonstrates the transparency, correctness, flexibility, and practicality of the solution. As the solution was presented in terms of Java, the discussion involves Java-specific characteristics; however some of the more general concepts should be useful for addressing similar issues in other object-oriented languages.

## **ParaAJ: Toward Reusable and Maintainable Aspect Oriented Programs**

*Khalid Aljasser and Peter Schachte*

**Thursday, 2:30pm, Soundings Theatre**

**Abstract** Aspect Oriented Programming aims to ease maintenance and promote reuse of software components by separating core concerns from crosscutting concerns: aspects of a program that cannot be connected to a single program component. In AOP languages such as AspectJ, this kind of concern is encapsulated in an aspect and connected to main classes using pointcuts. This removes extraneous code from the classes of the program, allowing them to focus on their core purpose and making them more maintainable and reusable. The implementation of each crosscutting concern, which would have been scattered throughout an object oriented program, is centralised in a single aspect. Unfortunately, due to the way aspects and classes are associated, and the lack of an explicit interface between them, these aspects may be tightly coupled to the classes of the program and so may not be as reusable or maintainable as might be expected. We propose ParaAJ (Parametric Aspects), as an extension to AspectJ. ParaAJ allows classes to specify which aspects should be applied, and allows applications to specify which aspects to apply to which classes in what ways. This makes it easier for classes and aspects to be developed and maintained independently, and encourages reuse of both.

## **A Domain Specific Language for Execution Profiling & Regulation**

*Peter Nguyen, Katrina Falkner, Henry Detmold and David Munro*

**Thursday, 3:00pm, Soundings Theatre**

**Abstract** Profiling consists of three stages: the collection of performance data, the processing of that data to infer performance information and the feedback of this performance information into the system. Feedback refers to using the information to change either the runtime system or the sampling parameters for subsequent profiling runs. This paper will concentrate on the latter approach for feedback. Existing profiling tools focus on data collection, requiring manual intervention during processing and feedback. The user must interpret the results presented to them to identify new profiling strategies. We introduce the concept of profiling regulation, whereby the processes of collection, processing and feedback are automated. We define a domain specific language, *sampspec*, that provides expressibility and control over the profiling process. The user provides a declarative specification of the information to collect, the computations to perform and the strategies to employ based on this information. This is in contrast to the manual inspection of results and restarting the profiler. Thus, the profiling process becomes one of specification of strategies for data collection and processing, and how these strategies can adapt over time. In this paper, we describe the system model and illustrate our language through a series of worked examples.

## **Fast and Compact Hash Tables for Integer Keys**

*Nikolas Askitis*

**Thursday, 4:00pm, Soundings Theatre**

**Abstract** A hash table is a fundamental data structure in computer science that can offer rapid storage and retrieval of data. A leading implementation for string keys is the cache-conscious array hash table. Although fast with strings, there is currently no information in the research literature on its performance with integer keys. More importantly, we do not know how efficient an integer-based array hash table is compared against other efficient hash tables for integers, such as bucketized cuckoo hashing. In this paper, we explain how to efficiently design an array hash table for integers. We then demonstrate, through careful experimental evaluations, which hash table whether it be a bucketized cuckoo hash table, an array hash table, or alternative hash table schemes such as linear probing offers the best performance—with respect to time and space—for maintaining a large dictionary of integers in-memory, on a current cache-oriented processor.

## 2 ACE: Australasian Computing Education Conference

### What Our ICT Graduates Really Need From Us: A Perspective From the Workplace

*Tony Koppi, Judy Sheard, Fazel Naghdy, Joe Chicharo, Sylvia L. Edwards, Wayne Brookes and David Wilson*

Tuesday, 1:30pm, ICON Foyer

**Abstract** A national Discipline-Based Initiative (DBI) project for Information and Communications Technology (ICT), funded by the Australian Learning and Teaching Council, has sought the opinions of recent graduates of ICT in the workplace to help inform the curriculum. An online survey was devised to question graduates on workplace requirements and university preparation for abilities categorized as: personal/interpersonal; cognitive; business and technical. The graduates in employment have highlighted broad mismatches between the requirements of their professional work in these categories and the preparation for employment they received from university. A regression analysis was used to determine influences on graduates opinions of the preparation they received at university. The quantitative and qualitative results from this survey could have far-reaching consequences for ICT education and this initiative will enable the development of curricula that ensures graduates are equipped with the skills required by the ICT industry.

### Intervention Programmes to Recruit Female Computing Students: Why Do the Programme Champions Do It?

*Annemieke Craig*

Tuesday, 2:00pm, ICON Foyer

**Abstract** This paper looks at intervention programmes to improve the representation of female students in computing education and the computer industry. A multiple case study methodology was used to look at major intervention programmes conducted in Australia. One aspect of the research focused on the programme champions; those women from the computing industry, those working within government organisations and those in academia who instigated the programmes. The success of these intervention programmes appears to have been highly dependent upon not only the design of the programme but on the involvement of these strong individuals who were passionate and worked tirelessly to ensure the programmes success. This paper provides an opportunity for the voices of these women to be heard. It describes the champions own initial involvement with computing which frequently motivated and inspired them to conduct such programmes. The research found that when these types of intervention programmes were conducted by academic staff the work was undervalued compared to when the activities were conducted by staff in indus-

try or in government. The academic environment was often not supportive of academics who conducted intervention programmes for female students.

**Issues Regarding Threshold Concepts in Computer Science**  
*Janet Rountree and Nathan Rountree*  
**Tuesday, 2:30pm, ICON Foyer**

**Abstract** Threshold Concepts deserve discussion and reflection in Computer Science Education; they provide a conceptual framework intended to re-empower tertiary educators. At this stage, the idea of Threshold Concepts raises plenty of questions, promises renewed learner and teacher engagement, and suggests a means of focusing on the key aspects of a discipline that will allow a learner to, for example, think more like a computer scientist. But what precisely are threshold concepts? Can we identify them? Can we agree on which concepts are threshold concepts and which are not? Can we validate them? If threshold concepts do exist, and can be identified and agreed upon, then how would they alter what we teach, how we teach, and how we assess? Do threshold concepts represent anything new or unexpected? The purpose of this paper is to set out issues for the Threshold Concepts model in Computer Science Education and encourage on-going discussion.

**Human Fallibility: How Well Do Human Markers Agree?**  
*Debra Haley, Pete Thomas, Marian Petre, Anne De Roeck*  
**Tuesday, 3:00pm, ICON Foyer**

**Abstract** Marker bias and inconsistency are widely seen as problems in the field of assessment. Various institutions have put in place a practice of second and even third marking to promote fairness. However, we were able to find very little evidence, rather than anecdotal reports, of human fallibility to justify the effort and expense of 2nd marking. This paper fills that gap by providing the results of a large-scale study that compared 5 human markers marking 18 different questions each with 50 student answers in the field of Computer Science. The study found that the human inter-rater reliability (IRR) ranged broadly both over a particular question and over the 18 questions. This paper uses the Gwet AC1 statistic to measure the inter-rater reliability of 5 markers.

The study was motivated by the desire to assess the accuracy of a computer assisted assessment (CAA) system we are developing. We claim that a CAA system does not need to be more accurate than human markers. Thus, we needed to quantify how accurate human markers are.

## Ten years of the Australasian Computing Education Conference

*Simon*

Tuesday, 4:00pm, ICON Foyer

**Abstract** The Australasian Computing Education Conference is now in its eleventh year. This paper charts the ups and downs of the conference from its origin in 1996, through its troubled years, to its recent apparently steady state. All 328 papers from the ten conferences are classified according to Simons system for classifying computing education papers, and features of interest are pointed out. Only one clear trend over time is observed, and that is a steady and distinct increase in the proportion of research papers. The analysis then moves from the papers to their 496 distinct authors, exploring where the authors come from, how many papers each has contributed to the conference, and which authors appear to have made this conference their home. A final look at the number of papers presented each year suggests that the conference might once more be experiencing difficulty, and speculates on its future.

## A Citation Analysis of the ICER 2005-07 Proceedings

*Raymond Lister and Ilona Box*

Tuesday, 4:30pm, ICON Foyer

**Abstract** This paper identifies the most commonly cited conferences, journals and books of the 43 papers within the first three ICER proceedings. A large array of conferences, journals, and books were cited. However, only a small set of journals and conferences were cited frequently, and the majority were only cited within a single paper, which is consistent with a power law distribution, as predicted by Zipfs Law. The most commonly cited books are concerned with education in general (29%) or psychology (20%), while 17% of books are concerned with computer science education and 12% with computing content. The citation results for ICER are contrasted with earlier published citation analyses of SIGCSE 2007 and ACE200507.

## A Taxonomic Study of Novice Programming Summative Assessment

*Shuhaida Shuhidan, Margaret Hamilton and Daryl DSouza*

Tuesday, 5:00pm, ICON Foyer

**Abstract** Learning to program is difficult, a situation that is largely responsible for high attrition rates in Computer Science schools. Novice programmers struggle to grasp an early understanding of programming, which can lead to frustration and eventually surrender. The problem has generated interest in a range of enquiries, and has given impetus to the need for a teaching-research nexus towards a better understanding of novice programming problems. We continue the trend in this paper and report on a study we have conducted

of novice programmers efforts in summative assessment. Our study involves multiple choice questions and coding question drawn from a programming examination. We analyse the answers provided by novices to final examination questions, and attempt to understand why students make such errors. We aim to categorise and classify the questions in the context of two well-known learning taxonomies: Blooms Taxonomy and the SOLO Taxonomy.

### **A Model of Internal Factors Influencing Student Learning of Programming**

*Angela Carbone, Ian Mitchell, Dick Gunstone and A. John Hurst*

**Tuesday, 5:30pm, ICON Foyer**

**Abstract** This paper presents a model of internal factors influencing student learning of programming. This is based on literature relating to student learning and learning of programming. The model comprises two dimensions: motivation and capability. It is used as a framework to explore the data gathered from a study of first year undergraduate IT programming students at an Australian University. The authors propose that this model is a useful tool to assist academics with developing tasks to facilitate student motivation and skills to learn programming.

### **A Focus Group Study of Student Attitudes to Lectures**

*Michael Hitchens, Raymond Lister*

**Wednesday, 9:00am, ICON Foyer**

**Abstract** This paper reports on the findings from focus groups, conducted at Macquarie University, on the attitudes of computing students to lectures. Students felt that two things were vital for a good lecture: (1) that the lecturer goes beyond what is written in the lecture notes; (2) that the lecture is interactive, by which students meant that the lecturer asks if students understand concepts and adjusts the delivery accordingly, and also the lecturer answers the students questions. The students in the focus groups also discussed what makes for a bad lectures: (1) lecturers reading straight from slides; (2) lecturers who blame the students, by saying that students dont work hard enough and are too lazy to turn up to lectures; and (3) lecturers who cover the material too slowly or too quickly. The most prominent reason given for not attending lectures was the timetabling of lectures in such a way that students had too few classes in one day to make the sojourn to university worthwhile. Any university seeking to improve attendance at lectures should perhaps look as much to improving its timetabling practices as it does to improving the practices of its individual lecturers.

**How Students Develop Concurrent Programs**  
*Jan Lonnberg, Anders Berglund, Lauri Malmi*  
Wednesday, 9:30am, ICON Foyer

**Abstract** This paper describes a qualitative, explorative study of how students approach developing and testing concurrent programs. The study is based on interviews with students working on the final programming assignment in a concurrent programming course. We discuss the effects of the students' approaches to constructing and testing programs on their work, how teaching can be improved to support the students in performing these tasks more effectively and how software tools can be designed to support the development, testing and debugging of concurrent programs.

**Quality of student contributed questions using PeerWise**  
*Paul Denny, Andrew Luxton-Reilly, Beth Simon,*  
Wednesday, 1:30pm, ICON Foyer

**Abstract** PeerWise is an online tool that involves students in the process of creating, sharing, answering and discussing multiple choice questions. Previous work has shown that students voluntarily use the large repository of questions developed by their peers as a source of revision for formal examinations and activity level correlates with improved exam performance.

In this paper, we investigate the quality of the questions created by students in a large introductory programming course. The ability of students to assess question quality is also examined. We find that students do, very commonly, ask clear questions that are free from error and give the correct answers. Of the few questions we examined that contained errors, in all cases those errors were detected, and corrected by other students. We also report that students are effective judges of question quality, and are willing to use the judgements of their peers to decide which questions to answer. We include several case studies of questions that are representative of the kinds of questions in the repository and provide insight for instructors.

**Losing Their Marbles: Syntax-Free Programming for Assessing Problem-Solving Skills**  
*Colin Fidge and Donna Teague*  
Wednesday, 2:00pm, ICON Foyer

**Abstract** Novice programmers have difficulty developing an algorithmic solution while simultaneously obeying the syntactic constraints of the target programming language. To see how students fare in algorithmic problem solving when not burdened by syntax, we conducted an experiment in which a large class of beginning programmers were required to write a solution to a computational problem in structured English, as if instructing a child, without reference to program code at all. The students produced an unexpectedly wide range

of correct, and attempted, solutions, some of which had not occurred to their teachers. We also found that many common programming errors were evident in the natural language algorithms, including failure to ensure loop termination, hardwiring of solutions, failure to properly initialise the computation, and use of unnecessary temporary variables, suggesting that these mistakes are caused by inexperience at thinking algorithmically, rather than difficulties in expressing solutions as program code.

### **A Perspective on the International Olympiad in Informatics for CS educators**

*Margot Phillipps and Leon Sterling*

**Wednesday, 4:00pm, ICON Foyer**

**Invited Talk Abstract** At the 2008 International Olympiad in Informatics held in Cairo, the Australian and New Zealand teams had their best ever performances. This talk will give details of the Informatics competition, and how teams are supported in Australia and New Zealand. Some sample informatics problems will be described. We argue that Informatics is an excellent basis for Computer Science at University and consequently it is important for CORE educators to understand and engage with the high school competition.

### **Easing the Transition: A Collaborative Learning Approach**

*Katrina Falkner, David S. Munro*

**Wednesday, 5:00pm, ICON Foyer**

**Abstract** Engaging first year students is a difficult problem, as students must develop independent study skills while concurrently mastering their chosen topic. At the same time, they find themselves in an alien environment, removed from their peer group and anonymised by University structures. Retention is of particular concern within ICT as across Australia, and globally, we have seen a recent dramatic drop in number of students applying for ICT degrees, and poor progression and retention rates.

Collaborative learning is a strategy that involves the students themselves in the ownership and direction of their learning; each student is responsible for not only their own learning but of the learning of the group. In this paper we describe our approach to student engagement based on applying a range of collaborative learning techniques within an introductory Computer Science course, addressing specifically the task of collaborative problem solving. Results from three years of adopting this change in teaching methodology indicate increased student confidence, participation and student ability.

## **Evolution of an International Collaborative Student Project**

*Cary Laxer, Mats Daniels, sa Cajander, Michael Wolowski*

**Wednesday, 5:30pm, ICON Foyer**

**Abstract** International collaborative student projects are inherently difficult for everyone concerned the students working on the projects, the faculty guiding the students, and the clients submitting the projects. With more and more schools recommending, or even requiring, that their students have some form of international experience in their degree programs, these projects will become more prevalent in helping to educate computing students in the 21st century. Understanding cultural differences between countries helps students have a better appreciation for the global aspects of computing and the issues faced in making software work in an environment they are not used to. This paper discusses the evolution over four years of collaborative projects between computing students at two schools, one in Sweden and one in the United States. The projects are based in courses at both schools that deal with computing in society. We discuss what the faculty teaching the courses and guiding the projects have learned and how they have improved the experience, what the students learn through these projects, and how the clients interact with the students and faculty. Suggestions for further development of these projects are also made.

## **Experiences in Teaching Quality Attribute Scenarios**

*Ewan Tempero*

**Thursday, 9:00am, ICON Foyer**

**Abstract** The concept of the quality attribute scenario was introduced in 2003 to support the development of software architectures. This concept is useful because it provides an operational means to represent the quality requirements of a system. It also provides a more concrete basis with which to teach software architecture. Teaching this concept however has some unexpected issues. In this paper, I present my experiences of teaching quality attribute scenarios and outline Bus Tracker, a case study I have developed to support my teaching.

## **A People-First Approach to Programming**

*Donna Teague*

**Thursday, 9:30am, ICON Foyer**

**Abstract** Students continue to find learning to program difficult. Failure rates from introductory programming units are high, as are attrition rates from IT courses. Case studies were conducted in 2007 involving Queensland University of Technology (QUT) introductory programming students who took part in weekly interviews and focus groups, and responded to questionnaires. Students divulged details relating to their attitude and approach to study, together with

the level of confidence they had in their ability to learn to program. Four of the case studies are included in this paper which portrays students with varying levels of confidence, motivation, determination, attitude and study ethic, and how they each struggle to learn to program. The purpose of the studies was to determine to what extent each of these factors has an influence on student learning outcomes. The studies focus on the people rather than the more traditionally studied cognitive difficulties of learning to program. The data collected from the case studies give some insight into the social barriers on many levels that students face and how they are dealt with and in some cases overcome. The paper concludes with a discussion on student programmer personas as a design taxonomy and pedagogical tool.

**Teaching and Assessing Programming Strategies Explicitly**  
*Michael de Raadt, Richard Watson, Mark Toleman*  
Thursday, 1:30pm, ICON Foyer

**Abstract** This paper describes how programming strategies were explicitly instructed and assessed in an introductory programming course and describes the impact of this curricular change. A description is given of how strategies were explicitly integrated into teaching materials and assessed in assignments and examinations. Comparisons are made between the outcomes of novices under the new curriculum and results of novices' learning under the previous implicit-only strategy curriculum, measured in an earlier study. This comparison shows improvement in novices' strategy application under the new curriculum.

**Surely We Must Learn to Read Before We Learn to Write!**  
*Simon, Mike Lopez, Ken Sutton, Tony Clear*  
Thursday, 2:00pm, ICON Foyer

**Abstract** While analysing students marks in some comparable code-reading and code-writing questions on a beginners programming exam, we observed that the weaker students appeared to be able to write code with markedly more success than they could read it. Examination of a second data set from a different institution failed to confirm the observation, and appropriate statistical analysis failed to find any evidence for the conclusion. We speculate on the reasons for the lack of a firm finding, and consider what we might need to do next in order to more thoroughly explore the possibility of a relationship between the code-reading and code-writing abilities of novice programming students.

**The BRACElet 2009.1 (Wellington) Specification**  
*Jacqueline Whalley and Raymond Lister*  
Thursday, 2:30pm, ICON Foyer

**Invited Talk Abstract** BRACElet is a multi-institutional computer education research study of novice programmers. The project is open to new members.

The purpose of this paper is to: (1) provide potential new members with an overview of BRACElet, and (2) specify the common core for the next data collection cycle. In this paper, BRACElet is taking the unusual step of making its study design public before data is collected. We invite anyone to run their own study using our study design, and publish their findings, irrespective of whether they formally join BRACElet. We look forward to reading their paper.

### 3 ADC: Australasian Database Conference

#### Large-scale Video Sequence Indexing: Impacts, Ideas and Trends

*Heng Tao Shen*

Tuesday, 1:30pm, Rangimarie Room 1

**Invited Talk Abstract** With the advances of hardware (e.g., wide availability of Webcam) and software (e.g., video editing or instant messaging software), the amount of video data has grown rapidly in many fields, such as broadcasting, advertising, filming, personal video archive, and medical/scientific video repository. In addition, Web has generated enormous impact by popularizing video publishing and sharing (e.g., social networking websites). Online delivery of video content has surged to an unprecedented level. The wide availability of video data fuels many novel applications, such as near-duplicate video detection, in-video advertising, video recommendation, web video search, etc. With these demanding applications, how to manage large-scale video databases and search similar video content is of uttermost importance. Although content-based video search has recently attracted plenty of attention, the high complexity of video data, coupled with large volume, poses huge challenges towards large-scale video sequence search. As the volume of video data continues to grow rapidly, the demand of efficient indexing on large-scale video databases from database community is increasingly imperative. In this talk, we will look at the problem of effective indexing supports for large-scale video sequence search in various forms, such as clip matching, subsequence matching and continuous stream matching. Its impacts and challenges will be discussed, followed by our recent ideas and developments to tackle this problem. Its future trends in next age will also be discussed.

#### The Incidence of Sparsity on Collaborative Filtering Metrics

*Jesus Bobadilla and Francisco Serradilla*

Tuesday, 2:30pm, Rangimarie Room 1

**Abstract** This paper presents a detailed study of the behavior of three different content-based collaborative filtering metrics (correlation, cosine and mean squared difference) when they are processed on several ratio matrices with different levels of sparsity. The total number of experiments carried out is 648, in which the following parameters are varied: metric used, number of k-neighborhoods, sparsity level and type of result (mean absolute error, percentage of incorrect predictions, percentage of correct predictions and capacity to generate predictions). The results are illustrated in two and three-dimensional representative graphs. The conclusions of the paper emphasize the superiority of the correlation metric over the cosine metric, and the unusually good results of the mean squared difference metric when used on matrices with high sparsity

levels, leading us to interesting future studies.

### **Solving the Golden Transaction Problem for ARIES-based Multi-level Recovery**

*Jayson Speer, Markus Kirchberg, Faizal Riaz-ud-Din and Klaus-Dieter Schewe*

**Tuesday, 3:00pm, Rangimarie Room 1**

**Abstract** Transaction throughput is a crucial issue for database systems. Multi-level transactions have been proposed in an attempt to offer improved concurrency of transaction processing by allowing operations that would otherwise be performed serially to take place concurrently. Therefore, it is vital that recovery algorithms do not impede this concurrency by artificially introducing restrictions that otherwise do not need to exist. The ARIES recovery algorithm has had a significant impact on the current thinking on database transaction logging and recovery. Its corresponding extension to multi-level transactions, i.e. ARIES/ML, preserves the unique features of ARIES but places significant restrictions on rollback processing. In this paper, we present an algorithm that solves the so-called ‘golden transaction’ problem of the ARIES/ML algorithm.

### **On Inference of XML Schema with the Knowledge of an Incorrect One**

*Irena Mlynkova*

**Tuesday, 4:00pm, Rangimarie Room 1**

**Abstract** The XML has undoubtedly become a standard for data representation and manipulation. But most of XML documents are still created without the respective description of their structure, i.e. an XML schema. Hence, in this paper we focus on the problem of automatic inferring of an XML schema for a given sample set of XML documents. Contrary to existing approaches we propose an algorithm that exploits additional input information – an incorrect XML schema. Consequently, we are able to exploit the information which was correct once and to infer the schema more efficiently.

### **ActiveTags: Making tags more useful anywhere on the Web**

*Stephan Hagemann and Gottfried Vossen*

**Tuesday, 4:30pm, Rangimarie Room 1**

**Abstract** The tags in social tagging systems store meaning for the taggers who entered them, and other users often share this understanding. The result of this, a folksonomy, is typically used in several ways, including information retrieval and clustering, serendipitous information access, or visualization of folksonomic characteristics. For these uses tags work pretty well; however, the ambiguity of tags makes it difficult to use them for more than searching and

browsing. This paper introduces examples of current programmatic support in the form of mashups and highlights its shortcomings. It identifies several types of tags based on their structure and language, and discusses how these types support programmatic uses. The main part is the presentation of the ActiveTags system, a browser extension with supporting server infrastructure. Using it the same community process that creates a folksonomy can be used to enhance tags with programmatic meaning. Users are enabled to create reliable mashups based on tags. Effectively, this leads to customized views of Web pages with tagged content. ActiveTags naturally increases the usability of social tagging systems and further extends the notion of user-generated content.

**Unified Q-ary Tree for RFID Tag Anti-Collision Resolution**  
*Bela Stantic and Prapassara Pupunwiwat*  
**Tuesday, 5:00pm, Rangimarie Room 1**

**Abstract** Radio Frequency Identification (RFID) technology uses radio-frequency waves to automatically identify people or objects. A large volume of data, resulting from the fast capturing RFID readers and a huge number of tags, poses challenges for data management. This is particularly the case when a reader simultaneously reads multiple tags and Radio Frequency (RF) collisions occur, causing RF signals to interfere with each other and therefore preventing the reader from identifying all tags. This problem is known as Missed reads, which can be solved by using anti-collision techniques to prevent two or more tags from responding to a reader at the same time. The current probabilistic anti-collision methods are suffering from Tag starvation problems so not all tags can be identified, while the deterministic methods suffer from too long Identification delay. In this paper, a "Unified Q-ary Tree Protocols" based on Query tree is presented. In empirical study compared with the Query tree and 4-ary tree, we show that the proposed method performs better, it requires less number of queries per complete identification, which results in less total identification delay.

**A Citation Analysis of the ADC 2006 - 2008 Proceedings, with Reference to the CORE Conference and Journal Rankings**  
*Raymond Lister and Ilona Box*  
**Tuesday, 5:30pm, Rangimarie Room 1**

**Abstract** This paper compares the CORE rankings of computing conferences and journals to the frequency of citation of those journals and conferences in the Australasian Database Conference (ADC) 2006, 2007 and 2008 proceedings. The assumption underlying this study is that there should be a positive relationship between citation rates and the CORE rankings. Our analysis shows that CORE conference and journal rankings broadly reflect the ADC citations, but we note some anomalies. While these anomalies might be minor in the

larger scheme of things, any anomalies need to be addressed, as the careers of individual academics may depend upon it. The concept of conference and journal rankings is probably here to stay, and this paper ends with some suggestions on how the rankings process should now evolve, so that it becomes more transparent.

### **S.E.A.L. - A Query Language for Entity-Association Queries**

*Edward Stanley, Pavle Mogin and Peter Andreae*

**Wednesday, 9:00am, Rangimarie Room 1**

**Abstract** The paper presents the S.E.A.L. query language and interpreter for entity association queries that allows such queries to be expressed very much more simply than can be done in SQL. S.E.A.L also supports EAV data, enabling users to write queries without having to know whether a particular attribute is a regular attribute or an EAV attribute. The language also allows parts of the query to be omitted if they are implied by the rest of the query, and the interpreter will infer the missing requirements, or ask the user to resolve the ambiguity. S.E.A.L. will make it easier for users of databases, particularly science and eCommerce databases, to make use of the valuable information in their databases.

### **Score Aggregation Techniques in Retrieval Experimentation**

*Sri Devi Ravana and Alistair Moffat*

**Wednesday, 9:30am, Rangimarie Room 1**

**Abstract** Comparative evaluations of information retrieval systems are based on a number of key premises, including that representative topic sets can be created, that suitable relevance judgements can be generated, and that systems can be sensibly compared based on their aggregate performance over the selected topic set. This paper considers the role of the third of these assumptions - that the performance of a system on a set of topics can be represented by a single overall performance score such as the average, or some other central statistic. In particular, we experiment with score aggregation techniques including the arithmetic mean, the geometric mean, the harmonic mean, and the median. Using past TREC runs we show that an adjusted geometric mean provides more consistent system rankings than the arithmetic mean when a significant fraction of the individual topic scores are close to zero, and that score standardization (Webber et al., SIGIR 2008) achieves the same outcome in a more consistent manner.

## Engineering Agile Systems

*Dimitrios Georgakopoulos*

Wednesday, 1:30pm, Rangimarie Room 1

**Invited Talk Abstract** The majority of today's software systems and organizational/business structures have been built on the foundation of solving problems via long-term data collection, analysis, and solution design. This traditional approach of solving problems and building corresponding software systems and business processes, falls short in providing the necessary solutions needed to deal with many problems that require agility as the main ingredient of their solution. For example, such agility is needed in responding to an emergency, in military command control, physical security, price-based competition in business, investing in the stock market, video gaming, network monitoring and self-healing, diagnosis in emergency health care, and many other areas that are too numerous to list here. The concept of Observe, Orient, Decide, and Act (OODA) loops is a guiding principal that captures the fundamental issues and approach for engineering agile information systems that deal with many of these problem areas. However, there are currently few software systems that are capable of supporting OODA. In this talk, we advocate a combination of complex event processing, service computing, and OODA principles for building agile systems, and provide a tour of corresponding research issues and state of the art solutions. We also provide specific examples of agile systems from the video surveillance, emergency response, and intelligence gathering domains.

## Elliptic Indexing of Multidimensional Databases

*Ondrej Danko and Tomas Skopal*

Wednesday, 2:30pm, Rangimarie Room 1

**Abstract** In this work an R-tree variant, which uses minimum volume covering ellipsoids instead of usual minimum bounding rectangles, is presented. The most significant aspects, which determine R-tree index structure performance, is an amount of dead space coverage and overlaps among the covering regions. Intuitively, ellipsoid as a quadratic surface should cover data more tightly, leading to less dead space coverage and less overlaps. Based on studies of many available R-tree variants (especially SR-tree), the eR-tree (ellipsoid R-tree) with ellipsoidal regions is proposed. The focus is put on the algorithm of ellipsoids construction as it significantly affects indexing speed and querying performance. At the end, the eR-tree undergoes experiments with both synthetic and real datasets. It proves its superiority especially on clustered sparse datasets.

## **Efficient XQuery Join Processing in Publish/Subscribe Systems**

*Ryan Choi and Raymond Wong*

**Wednesday, 3:00pm, Rangimarie Room 1**

**Abstract** Efficient XML filtering has been a fundamental technique in recent Web service and XML publish/subscribe applications. In this paper, we consider the problem of filtering a continuous stream of XML data against a large number of XQuery queries that contain multiple inter-document and value-based join operations in their where clauses. To perform efficient join operations, the path expressions from these queries are extracted and organized in a way that multiple path expressions can be joined simultaneously. The join operations are then pipelined to minimize the number of join operations and to share any intermediate join results as much as possible. Our system operates on top of many currently available XPath filtering engines as an add-on module to extend their features to support queries with join operations. Experiments show that our proposal is efficient and scalable.

## **Event-based Communication for Location-based Service Collaboration**

*Annika Hinze, Lisa Eschner and Yann Michel*

**Wednesday, 4:00pm, Rangimarie Room 1**

**Abstract** Location-based context-aware services for mobile users need to collaborate in disparate networks. Services come and go as the user moves and no central repository is available. The user's personal information and service usage statistics need to be protected. To support service collaboration we propose a service infrastructure that relies on an event-based service-oriented architecture. We implemented a basic version of the architecture and used it for a tourism information system. An advanced version has been modelled using formal methods to evaluate privacy aspects. This paper reports about both architectures and our experiences for their application to tourism-related services.

## **Mobile Information Exchange and Integration: From Query to Application Layer**

*Van Tran, Raymond Wong, William K. Cheung and J Liu*

**Wednesday, 4:30pm, Rangimarie Room 1**

**Abstract** Due to the popularity of mobile devices, more and more commercial applications have been developed on these devices. While commercial applications are mostly backed by relational database systems, numerous database engines have been ported to or built on these devices, for example, SQLite. Since connectivity can be unstable or slow, applications such as iAnywhere have considered offline operations while data can be synchronized with the database

server whenever the devices are online. On the other hand, while Web-based and XML content are very common these days, unfortunately, these mobile versions of database engines failed to fully support them. This paper considers a translation-based method with a decentralised versioning system in place to support offline operations. Web and XML contents are stored and versioned in a distributed manner and can be synchronized with each other without connecting to a server. The schema of these data can be automatically generated on device. With these schema, a translation engine which allows querying these data using SQL by translating the query to a corresponding XML query is facilitated. We believe this framework support mobile data applications on XML or Web data in a seamless manner. Finally, an initial prototype has been implemented and described in this paper.

### **Access Control: What is Required in Business Collaboration?**

*Daisy Daiqin He, Michael Compton, Jian Yang and Kerry Taylor*

**Thursday, 9:00am, Rangimarie Room 1**

**Abstract** Access control has been studied extensively, and there are a number of theories and techniques for handling access control for single or centralised systems. However, unique and challenging security issues concerning collaboration in the context of service oriented computing (SOC) have arisen due to the dynamic and loosely coupled nature of the environment in which these collaborations are conducted. Individual organisations usually define their access control policies independently. When a collaboration opportunity arrives, a number of issues may arise, for example, how to decide if the collaboration is possible given the access control policies, how to define the policy for the collaboration and deciding under what conditions a service is allowed to be forwarded to other parties. Furthermore, different types of collaboration, in terms of the way collaboration is carried out, require different access control support. In this paper, we propose a model encoded in description logic to capture all the necessary elements for specifying access control policy for collaboration. Based on the model, various inconsistencies between access policies from different business units are identified. The paper also shows how a description logic reasoner can be used to prove that two policies are suitable, or not suitable, for collaboration. The policy model and policies are encoded in a *SROIQ* knowledge base. We believe this work lays a foundation for access policy development, negotiation and enforcement for cross-organization collaborations.

## **CSC: Supporting Queries on Compressed Cached XML**

*Stefan Böttcher and Rita Hartel*

**Thursday, 9:30am, Rangimarie Room 1**

**Abstract** Whenever a client frequently has to retrieve, to query and to locally transform large parts of a huge XML document that is stored on a remote web information server, data exchange from the server to the client may become a serious bottleneck that simply limits scaling of the amount of information that can be processed locally on the client by a client-based application. We present Compressed Structure Caching (CSC) as a solution that reduces the amount of data exchange by a combination of the following techniques: compression of the XML document's structure, client-side caching of the structure and of already received XML content, inference and optimized loading of the content needed on the client to answer a given query. We provide a performance evaluation that demonstrates that our approach significantly reduces the amount of data exchange from server to client.

## **Information Retrieval in Structured Domain**

*Vincent W. L. Tam and John Shepherd*

**Thursday, 1:30pm, Rangimarie Room 1**

**Abstract** In this research, we investigate the effectiveness of utilizing the structure of a website to increase the quality of document retrieval within a structured domain. In particular we examine various methods to combine evidence within the website in order to improve the quality of pages returned.

## **Ranking-Constrained Keyword Sequence Extraction from Web Documents**

*Xue Li*

**Thursday, 2:00pm, Rangimarie Room 1**

**Abstract** Given a large volume of Web documents, we consider problem of finding the shortest keyword sequences for each of the documents such that a keyword sequence can be rendered to a given search engine, then the corresponding Web document can be identified and is ranked at the first place within the results. We call this system as an Inverse Search Engine (ISE). Whenever a shortest keyword sequence is found for a given Web document, the corresponding document can be returned as the first document by the given search engine. The resulting keyword sequence is search-engine dependent. The ISE therefore can be used as a tool to manage Web content in terms of the extracted shortest keyword sequences. In this way, a traditional keyword extraction process is constrained by the document ranking method adopted by a search engine. The significance is that the whole Web-searchable documents on the World Wide Web can then be partitioned according to their keyword phrases. This paper discusses the design and implementation of the proposed ISE. Four evaluation

measures are proposed and are used to show the effectiveness and efficiency of our approach. The experiment results set up a test benchmark for further researches.

## **Conditional Purpose Based Access Control Model for Privacy Protection**

*Md Enamul Kabir and Hua Wang*

**Thursday, 2:30pm, Rangimarie Room 1**

**Abstract** This paper presents a model for privacy preserving access control which is based on variety of purposes. Conditional purpose is applied along with allowed purpose and prohibited purpose in the model. It allows users using some data for certain purpose with conditions. The structure of conditional purpose based access control model is defined and investigated through a practical paradigm with access purpose and intended purpose. An algorithm is developed to achieve the compliance computation between access purposes and intended purposes. According to this model, more information from data providers can be extracted while at the same time assuring privacy that maximizes the usability of consumers' data. This model extends traditional access control models to a further coverage of privacy preserving in data mining atmosphere. Its interior is a new structure for managing collected data in an effective and trustworthy way. This structure helps enterprises to circulate clear privacy promise, to collect and manage user preferences and consent. The implementation of the idea in the paper shows the flexibility of the model, and finally we provide comparisons of our work to other related work.

## 4 AISC: Australasian Information Security Conference

### Faster Group Operations on Elliptic Curves

*Huseyin Hisil, Kenneth Koon-Ho Wong, Gary Carter and Ed Dawson*

Tuesday, 1:30pm, Rangimarie Room 2

**Abstract** This paper improves implementation techniques of Elliptic Curve Cryptography. We introduce new formulae and algorithms for Jacobi-quartic, Jacobi-intersection, Edwards, and Hessian curves. These formulae and algorithms perform faster group operations in suitable coordinate systems. To show this, a comparison is made with classic point multiplication algorithms using previous and current operation counts. Most notably, the best speeds are obtained from Jacobi-quartic curves which provide the fastest timings for most point multiplication strategies benefiting from the proposed  $2M + 5S$  doubling and  $7M + 3S$  addition algorithms for a suitably selected curve constant. In addition, the new addition algorithm provides a very efficient way to protect against side channel attacks which are based on simple power analysis (SPA).

### Strong Designated Verifier Signature in a Multi-user Setting

*M. Choudary Gorantla, Colin Boyd and Juan Manuel González Nieto*

Tuesday, 2:00pm, Rangimarie Room 2

**Abstract** The security of strong designated verifier (SDV) signature schemes has thus far been analyzed only in a two-user setting. We observe that security in a two-user setting does not necessarily imply the same in a multi-user setting for SDV signatures. Moreover, we show that existing security notions do not adequately model the security of SDV signatures even in a two-user setting. We then propose revised notions of security in a multi-user setting and show that no existing scheme satisfies these notions. A new SDV signature scheme is then presented and proven secure under the revised notions in the standard model. For the purpose of constructing the SDV signature scheme, we propose a one-pass key establishment protocol in the standard model, which is of independent interest in itself.

## Algebraic Analysis of LEX

*Muhammad Reza Z'aba, Haavard Raddum, Leonie Simpson, Ed Dawson, Matt Henricksen and Kenneth Wong*  
Tuesday, 2:30pm, Rangimarie Room 2

**Abstract** LEX is a stream cipher that progressed to Phase 3 of the eSTREAM stream cipher project. In this paper, we show that the security of LEX against algebraic attacks relies on a small equation system not being solvable faster than exhaustive search. We use the byte leakage in LEX to construct a system of 21 equations in 17 variables. This is very close to the requirement for an efficient attack, i.e. a system containing 16 variables. The system requires only 36 bytes of keystream, which is very low.

## A New Version of the RC6 Algorithm, Stronger Against $\chi^2$ Cryptanalysis

*Routo Terada and Eduardo T. Ueda*  
Tuesday, 3:00pm, Rangimarie Room 2

**Abstract** We analyze the  $\chi^2$  cryptanalysis, one of the most successful cryptanalysis technique against the RC6 algorithm. We apply this type of cryptanalysis as distinction cryptanalysis as well as key-recovery cryptanalysis. We present a modified version of RC6 by introducing a very simple data-dependent swapping function in its structure. The conclusions inferred by statistical experiments is that this modified version is stronger against the  $\chi^2$  cryptanalysis technique.

## Foundation for Systems Security

*Clark Thomborson*  
Tuesday, 4:00pm, Rangimarie Room 2

## Slotted Packet Counting Attacks on Anonymity Protocols

*Volker Fusenig, Eugen Staab, Uli Sorger and Thomas Engel*  
Tuesday, 5:00pm, Rangimarie Room 2

**Abstract** In this paper we present a slotted packet counting attack against anonymity protocols. Common packet counting attacks make strong assumptions on the set-up and can easily lead to wrong conclusions, as we will show in our work. To overcome these limitations, we account for the variation of traffic load over time. We use correlation to express the relation between sender and receiver nodes. Our attack is applicable to many anonymity protocols. It assumes a passive attacker and works with partial knowledge of the network traffic.

## **Towards a Decision Model Based on Trust and Security Risk Management**

*Baptiste Alcalde, Eric Dubois, Sjouke Mauw, Nicolas Mayer and Sasa Radomirovic*

**Tuesday, 5:30pm, Rangimarie Room 2**

**Abstract** From choosing the daily lunch menu to buying or selling stock options, decisions have to be made every day. In general, due to incomplete information, making a decision carries a risk. Typically, such risks are mitigated through risk management. However, risk is not the only element involved in the decision process. When the decision to be made concerns an interaction between two entities, trust plays an important role. Trust, in such an interaction, is a prediction of one entity's reliance on the other entity to perform a certain action. In this paper we formulate a trust reference model and take a first step towards a decision model by combining the trust model with an existing risk model. The decision model is illustrated by an example in the e-banking domain.

## **Passwords and Perceptions**

*Gilbert Notoatmodjo and Clark Thomborson*

**Wednesday, 9:00am, Rangimarie Room 2**

**Abstract** The security of many computer systems hinges on the secrecy of a single word – if an adversary obtains knowledge of a password, they will gain access to the resources controlled by this password. Human users are the weakest link in password control, due to our propensity to reuse passwords and to create weak ones. Policies which forbid such unsafe password practices are often violated, even if these policies are well-advertised. We have studied how users perceive their accounts and their passwords. Our participants mentally classified their accounts and passwords into a few groups, based on a small number of perceived similarities. Our participants used stronger passwords, and reused passwords less, in account groups which they considered more important. Our participants thus demonstrated awareness of the basic tenets of password safety, but they did not behave safely in all respects. Almost half of our participants reused at least one of the passwords in their high-importance accounts. Our findings add to the body of evidence that a typical computer user suffers from password overload. Our concepts of password and account grouping point the way toward more intuitive user interfaces for password- and account-management systems.

## **Preliminary Security Specification for New Zealand’s igovt System**

*Yu-Cheng Tu and Clark Thomborson*

**Wednesday, 9:30am, Rangimarie Room 2**

**Abstract** The New Zealand government has proposed an identity management system, to provide an effective and convenient alternative for citizens to access online government information and services. The proposed system is branded as “igovt”. The igovt proposal offers two types of authentication services. The first service provides people and businesses with logon identities. The second service provides semi-anonymised identities to government agencies, each of which carries a strictly limited amount of information about a logon identity along with an assurance that it corresponds to a living New Zealand citizen or a registered business entity. The New Zealand government has carefully designed the system, with clearly-articulated policy principles. It has also conducted several privacy impact assessments and public consultations. However, the New Zealand government has not published any security analyses for igovt, and we are not aware of any unpublished ones. In this paper, we propose a lightweight methodology for the elicitation of security requirements of a complex but incompletely unimplemented system, such as igovt. We illustrate the use of our methodology by developing preliminary security specifications for a portion of the igovt system.

## **Proposal for Effective Information Flow Control Model for Sharing and Protecting Sensitive Information**

*Masato Arai and Hidehiko Tanaka*

**Wednesday, 1:30pm, Rangimarie Room 2**

**Abstract** Information leakage has become a serious problem for computer systems that handle a company’s sensitive information, such as intellectual properties and manufacturing know-how. The majority of the causes can be attributed to loss or theft of information or worms and viruses. As a countermeasure, forbidding the sharing of information through removable media or the Internet is effective, but it also places restriction on the handling of general information that can be made public. Also, to segregate sensitive information from environments that can easily be infected by worms or viruses, the sandbox model can be used; however, even sensitive information is sent as email attachments to various locations within the organization, and this model cannot be applied to business cases where information must be stored and carried out on removable media. In this article, we propose an information flow control model that is suitable for both sharing and protecting sensitive information on computer systems in which general information that can be made public and sensitive information that cannot be exposed outside the company are mixed. In the proposed model, by segregating the environment for programs that use the Internet and the environment in which programs handling sensitive information are executed,

using existing techniques such as the sandbox model, sensitive information are protected from environments that can be easily infected by worms or viruses. At the same time, by combining automatic file encryption and encrypted file access control, sensitive information can be safely transmitted as encrypted text through removable media or the Internet as the need arises.

### **Open and Trusted Information Systems/Health Informatics Access Control (OTHIS/HIAC)**

*Vicky Liu, Luis Franco, William Caelli, Lauren May and Tony Sahama*

**Wednesday, 2:00pm, Rangimarie Room 2**

**Abstract** Information and Communications Technologies (ICT) globally are moving towards Service Oriented Architectures and Web Services (SOA/WA). The healthcare environment is rapidly moving to the use of SOA/WS systems interconnected via this global open Internet. Such moves present major challenges where these structures are not based on highly trusted operating systems. This paper argues the need of a radical re-think of access control in the contemporary healthcare environment in light of modern information system structures, legislative and regulatory requirements, and security operation demands in Health Information Systems. This paper proposes the Open and Trusted Health Information Systems, a viable solution to providing appropriate levels of secure access control for the protection of sensitive health data.

### **The Stateful Cluster Security Gateway (CSG) Architecture for Robust Switched Linux Cluster Security**

*Visham Ramsurrun and K. M. S. Soyjaudah*

**Wednesday, 2:30pm, Rangimarie Room 2**

**Abstract** This work presents a new cluster security model for securing switched Linux clusters. The stateful CSG improves upon the stateless CSG in the sense that it supports stateful firewalling, provides high availability, greater scalability & load balancing capability. This model combines various mechanisms like distributed sender-initiated Layer 2 per-packet firewall load balancing, firewall state synchronization, failover, MAC address takeover, Network Access Control using switch MAC ACLs & port security, and Layer 2 & Layer 3 packet filtering in order to provide robust, scalable and reliable cluster-level security. Experimental results of performance not only give an idea of the effectiveness of the new scheme at boosting firewall performance & reliability, but also at improving network performance & security. In addition, the response of the new scheme in the face of threats is assessed qualitatively and its salient characteristics like tamper resistance, anti-spoofing, anti-sniffing & low end-user host processing strain, are highlighted.

## 5 APCCM: Asia-Pacific Conference on Conceptual Modelling

### Modeling Natural Language Communication in Database Semantics

*Roland Hausser*

Wednesday, 2:30pm, Soundings Theatre

**Invited Talk Abstract** Database Semantics (DBS) is a computational model of how communicating with natural language works. Dened at a level of abstraction which may be applied to natural and articial agents alike, such a model is a precondition for achieving free human-machine communication in natural language, and thus has great potential for a wide range of practical applications. The basic functionality of DBS is a cognitive agents turn taking between the speaker mode (mapping content stored in memory into language surfaces) and the hearer mode (mapping surfaces into content for storage in memory). DBS is dened in terms of (i) the data structure of at feature structures, (ii) the algorithm of time-linear LA-grammar, and (iii) the database schema of a classic network database.

### Reverse Engineering of XML Schemas to Conceptual Diagrams

*Martin Necasky*

Wednesday, 4:00pm, Rangimarie Room 2

**Abstract** It is frequent in practice that different logical XML schemas representing the same reality from different viewpoints exist. There is also usually a conceptual diagram modeling the reality independently of the viewpoints. It is important to keep the XML schemas and conceptual diagram consistent as they are both utilized for different purposes. In practice, this is however rarely the case. In this paper, we propose a reverse engineering method as a solution to this problem. We provide a semi-automatic algorithm that produces mappings of components of the XML schemas to components of the conceptual diagram. The method only provides suggestions for the mapping and manual participation of a domain expert is therefore required.

### Modelling Web-Oriented Architectures

*Gunnar Thies and Gottfried Vossen*

Wednesday, 4:30pm, Rangimarie Room 2

**Abstract** Service-oriented architectures (SOAs) provide the basis of distributed application frameworks where software components are provided as modular and reusable services. Until today there is no generally accepted method for conceptual modelling of a SOA. Rather, there exist several procedural methods

which are used in practice. On the other hand, recent developments in the context of what is commonly termed Web 2.0 show how easy it can be to link or compose (mesh) IT components dynamically, so that original SOA goals like exhibity, reusability, or reduction of complexity can indeed be achieved by relatively simple means. An interesting concept in this context is the Web-oriented architecture (WOA), which represents a specialization of SOAs obtained by using simple Web 2.0 technologies and standards (e.g., HTTP, SSL, XML). This paper introduces a methodology for designing WOAs, where the big picture follows existing SOA models. In particular, this WOA methodology comprises conceptual as well as realization issues and breaks WOA design down into three distinct phases.

## **Conceptional Modeling and Analysis of Spatio-Temporal Processes in Biomolecular Systems**

*Andreas Schafer, Mathias John*

**Wednesday, 5:00pm, Rangimarie Room 2**

**Abstract** In life science, deeper understanding of biomolecular systems is acquired by computational modeling and analysis. For the modeling of several kinds of reaction networks, e.g. signaling pathways, information on intracellular space, like the locations and motions of molecules, has to be taken into account. In this paper, we introduce Labeled SpacePi, an extension of the  $\pi$ -calculus, in order to model spatio-temporal processes in cells. The formalism is tailored to the available data and knowledge about biomolecular systems. For the analysis, we employ model checking techniques known from the field of safety-critical systems. To this end, we develop a translation of Labeled SpacePi models into hybrid automata. Two use cases - one considering the activation of a signaling pathway and the other one concerning active transport in cells - demonstrate our concept by making use of the established analysis tools HyTech and HySat.

## **Towards Accurate Conict Detection in a VCS for Model Artifacts: A Comparison of Two Semantically Enhanced Approaches**

*Kerstin Altmanninger and Gabriele Kotsis*

**Wednesday, 5:30pm, Rangimarie Room 2**

**Abstract** In collaborative software development the utilization of Version Control Systems (VCSs) is a must. For this important task some graph-based VCSs for model artifacts already emerged. Optimistic approaches, which are nowadays the designated ones, allow parallel editing of one resource and therefore changes can result in conicts and inconsistencies. To be exible for the ever increasing variety of modeling environments and languages VCSs should be independent of the modeling environment and applicable on any modeling language. Those VCS characteristics implicate a lack of information for the conict detection method by virtue of rstly receiving solely the state of an artifact without

concrete editing operations and secondly due to unavailable knowledge about the semantics of a modeling language. In such VCSs inconsistencies would even arise more often. Hence, accurate conict detection methods are indispensable for the realization of optimistic, environment and language independent VCSs. This can be achieved by providing some understanding about the modelss semantics which is possible by specifying machine interpretable formal semantics. Therefore, in this work, a comparison of two semantically enhanced conict detection approaches is presented with respect to their suitability for the integration in an optimistic, environment and language independent VCS for model artifacts to achieve more accurate conict reports.

### **Conceptual Application Domain Modelling**

*Bernhard Thalheim, Klaus-Dieter Schewe and Hui Ma*

**Thursday, 9:00am, Rangimarie Room 2**

**Abstract** Application domain description precedes requirements engineering, and is the basis for the development of a software or information system that satisses all expectations of its users. The greatest challenge in this area is the evolution of the application domain itself. In this paper we address this problem by explicit consideration of application cases that are dened by user proles and intentions and the system environment, i.e. scope and context. User proles and intentions are captured through the concept of persona. We show how the application domain description can be mapped to requirements and discuss engineering of application domain descriptions.

### **Multi-Level Domain Modeling with M-Ob jects and M-Relationships**

*Bernd Neumayr, Katharina Grün and Michael Schre*

**Thursday, 9:30am, Rangimarie Room 2**

**Abstract** Using traditional semantic data modeling, multi-level modeling can be achieved by representing objects in di?erent abstraction hierarchies, namely classication, aggregation and generalization. This, however, leads to accidental complexity, complicating maintenance and extension. Several modeling techniques, like deep instantiation, powertypes and materialization, have been proposed to reduce unnecessary complexity in modeling objects at multiple levels. Multilevel objects (m-objects) and multi-level relationships (m-relationships) build on these results and provide a natural, intuitive representation of the concretization of ob jects and relationships along multiple lev- els of abstraction. By integrating aspects of the different abstraction hierarchies in a single concretization hierarchy, they improve readability and simplify maintenance and extension as compared to previous approaches. The discussion on conceptual modeling is complemented by a brief presentation of M-SQL, a data manipulation and query language for working with m-objects and m-relationships in an object-relational setting.

## A Semantic Associative Computation Method for Automatic Decorative-Multimedia Creation with Kansei Information

*Yasushi Kiyoki and Xing Chen*

Thursday, 1:30pm, Rangimarie Room 2

**Invited Talk Abstract** In the design of multimedia systems, one of the important issues is how to deal with Kansei of human beings. The concept of Kansei in Japanese includes several meanings on sensitive recognition, such as impression, human senses, feelings, sensitivity, psychological reaction and physiological reaction. This paper presents a new concept of automatic decorative-multimedia creation and a semantic associative computation method. This method realizes automatic main-media decoration with dynamic sub-media data selection for representing main-media as decorative multimedia. The aim of this method is to create a new field of automatic decorative-media art with semantic associative computing. This paper defines an automatic media decoration model with semantic spaces and media-decoration functions. Automatic media decoration is realized by applying the Mathematical Model of Meaning (MMM) to a media-transmission space for computing semantic correlations between main-media objects and sub-media. The process of this dynamic media decoration method consists of the following functions:

1. Extraction of semantic Kansei features of main-media object, such as music, image and video.
2. Mapping of the main-media object onto the media-transmission space between main-media and sub-media.
3. Semantic associative computation of correlations between the main-media object and the features of the sub-media space by MMM, and creating a vector of the main-media object with the features of the sub-media space.
4. Mapping of the vector of the main-media object to the sub-media space, and semantic associative computing between the main-media object and sub-media data. the semantic associative computation, and selects one of the sub-media objects with high correlation values to the target main-media object.
5. Automatic rendering of the target main-media object with the selected sub-media object for decorating the main-media presentation.

This paper shows several significant applications of the semantic associative computation method for automatic decorative-media creation.

## **Contrasting Classification with Generalisation**

*Thomas Kühne*

**Thursday, 2:30pm, Rangimarie Room 2**

**Abstract** Classification and Generalisation are two of the most important abstraction mechanisms in modelling, and while they share a number of similarities, they are unmistakably different with respect to their properties. Recently, a number of (meta-) modelling language design approaches de-emphasised the differences between classification and generalisation in order to gain various advantages. This paper aims to demonstrate the loss in precision and the loss of sanity checks such approaches entail. After a careful comparison between classification and generalisation, I identify problems associated with the above mentioned approaches and offer alternatives that retain a strong distinction between classification and generalisation.

## **Synthesis of Orchestrators from Service Choreographies**

*Stephen McIlvenna, Marlon Dumas and Moe Thandar Wynn*

**Thursday, 3:00pm, Rangimarie Room 2**

**Abstract** Interaction topologies in service-oriented systems are usually classified into two styles: choreographies and orchestrations. In a choreography, services interact in a peer-to-peer manner and no service plays a privileged role. In contrast, interactions in an orchestration occur between one particular service, the orchestrator, and a number of subordinated services. Each of these topologies has its trade-offs. This paper considers the problem of migrating a service-oriented system from a choreography style to an orchestration style. Specifically, the paper presents a tool chain for synthesising orchestrators from choreographies. Choreographies are initially represented as communicating state machines. Based on this representation, an algorithm is presented that synthesises the behaviour of an orchestrator, which is also represented as a state machine. Concurrent regions are then identified in the synthesised state machine to obtain a more compact representation in the form of a Petri net. Finally, it is shown how the resulting Petri nets can be transformed into notations supported by commercial tools, such as the Business Process Modelling Notation (BPMN).

## **Business Process Integration: Method and Analysis**

*Evan D. Morrison, Alex Menzies, George Koliadis and Aditya K. Ghose*

**Thursday, 4:00pm, Rangimarie Room 2**

**Abstract** In the study of business management, process integration has become an interesting area of research that affects analysts studying and working on existing system plans. Process integration aims to investigate relationships

across a business compendium to produce classifications and merge similar activities into a standardized system. Integration is the process of merging elements from two similar antecedent processes to create a single process that can be used to replace the original processes. This paper proposes a practical method for process integration and provides a theoretical framework and metrics for business process integration assessment. In the provision of metrics that take into account similarity of activities within processes we are able to offer solutions that provide minimal change reducing change costs, and minimizing change impact risks.

## **Conceptual Business Document Modeling using UN/CEFACTs Core Components**

*Philipp Liegl*

**Thursday, 4:30pm, Rangimarie Room 2**

**Abstract** Before two businesses can engage in a business-to-business process an agreement about the process execution order and the business documents exchanged in the collaborative process must be found. Although several initiatives from different industries have started standardization initiatives for business documents a set of shortcomings still remain. (1) The different standards do not have a common semantic basis causing inter-operability problems between them. (2) Furthermore, they try to include every possible element any industry might need into the business document standard. (3) Moreover, most of the standards are transfer syntax specific and do not provide a conceptual representation mechanism. In this article a new concept for the standardization of business documents called UN/CEFACTs Core Components Technical Specification is presented which solves these shortcomings. Using Core Components the business document modeler can unambiguously define documents with a common semantic basis on a conceptual level. In order to allow for a better integration into UML modeling tools we introduce the UML Profile for Core Components. With the UML based core component model and an XML schema generator the modeler can derive XML schema artifacts from the conceptual model.

## **Extracting and Modeling the Semantic Information Content of Web Documents to Support Semantic Document Retrieval**

*Shahrul Azman Noah, Lailatulqadri Zakaria and Arifah Che Alhadi*

**Thursday, 5:00pm, Rangimarie Room 2**

**Abstract** Existing HTML mark-up is used only to indicate the structure and lay-out of documents, but not the document semantics. As a result web documents are difficult to be semantically processed, retrieved and explored by computer applications. Existing information extraction system mainly concerns with extracting important keywords or key phrases that represent the content of

the documents. The semantic aspects of such keywords have not been explored extensively. In this paper we propose an approach meant to assist in extracting and modeling the semantic information content of web documents using natural language analysis technique and a domain specific ontology. Together with the users participation, the tool gradually extracts and constructs the semantic document model which is represented as XML. The semantic models representing each document are then being integrated to form a global semantic model. Such a model provides users with a global knowledge model of some domains.

## **Extracting Conceptual Graphs from Japanese Documents for Software Requirements Modeling**

*Ryo Hasegawa, Motohiro Kitamura, Haruhiko Kaiya and Motoshi Saeki*

**Thursday, 5:30pm, Rangimarie Room 2**

**Abstract** A requirements analysis step plays a significant role on the development of information systems, and in this step we produce various kinds of abstract models of the systems (called requirements models) according to the adopted development processes, e.g. class diagrams in the case of adopting object-oriented development. However, constructing these models of sufficient quality requires highest intellectual tasks and skills of human requirements analysts. In this paper, we develop a computerized tool to extract from a set of Japanese text documents conceptual information, called conceptual graph, which can be used as intermediate representation to generate software requirements models. More concretely, by applying the variation of text-mining techniques that we have developed, we extract significant words from text documents referring to the same problem domain and identify relevant relationships among them. The extracted words can be considered as concepts and they are constituents of a conceptual graph in the domain. This constructed graph can be used for generating requirements models, e.g. object oriented models, feature model, and even as a domain ontology that can be utilized during requirements analysis activities. We have made experimental analyses of our tool. This paper also includes the discussion on how the extracted conceptual graph can act as an object-oriented model, a feature model and a domain ontology, in order to show its wide applicability.

## 6 AUIC: Australasian User Interface Conference

### Using Remotely Executing Software via a Mobile Device

*Vipul Delwadia, Stuart Marshall and Ian Welch*

Tuesday, 1:30pm, Rangimarie Room 3

**Abstract** There are scenarios in mobile computing that may benefit from separating presentation from computation. Traditionally this separation can be achieved via tools such as VNC. However such factors as network latency and additional communication overhead can slow down the presentation of a remotely executing mobile application below acceptable performance levels, especially for domains like gaming where responses may need to appear to be instantaneous. We present RemoteMe, an architecture and Java-based prototype for mobile-client / server communication that only requires a very thin mobile client. We hypothesise that RemoteMe will support faster response times to user input than existing software solutions such as VNC. This paper presents a preliminary analysis of our first prototype, and experimentally compares it to an open-source mobile-based VNC system.

### A Citation Analysis of the AUIC 2006 - 2008 Proceedings, with Reference to the CORE Conference and Journal Rankings

*Raymond Lister and Ilona Box*

Tuesday, 2:00pm, Rangimarie Room 3

**Abstract** This paper compares the CORE rankings of computing conferences and journals to the frequency of citation of those journals and conferences in the Australasian User Interface Conference (AUIC) 2006, 2007 and 2008 proceedings. The assumption underlying this study is that there should be a positive relationship between citation rates and the CORE rankings. Our analysis shows that CORE conference and journal rankings broadly reflect the AUIC citations, but we note some anomalies. While these anomalies might be minor in the larger scheme of things, any anomalies need to be addressed, as the careers of individual academics may depend upon it. The concept of conference and journal rankings is probably here to stay, and this paper ends with some suggestions on how the rankings process should now evolve, so that it becomes more transparent.

### Augmenting Image Place AR 3D Interactions for Wearable Computers

*Thuong Hoang, Shane Porter and Bruce Thomas*

Tuesday, 2:30pm, Rangimarie Room 3

**Abstract** This paper presents a set of large object manipulation techniques implemented in a wearable augmented reality computer system that are opti-

mised for the outdoor setting. These techniques supplement the current image plane approach, to provide a comprehensive solution to 3D object manipulation in an augmented reality outdoor environment. The three extended manipulation techniques, Revolve, Xscale, and Ground plane translation, are focused on using what we determined to be the best coordinate system for object rotation, scaling and translation. This paper goes on to present the generalised plane technique for the constrained translation of graphical objects on arbitrary planes to enable more complex translation operations. The paper presents the techniques from both the user interface and software development perspectives.

## **Comparison of Techniques For Mixed-Space Collaborative Navigation**

*Aaron Stafford, Bruce Thomas and Wayne Piekarski*

**Tuesday, 3:00pm, Rangimarie Room 3**

**Abstract** This paper describes the results of two studies conducted to determine the role of visual cues for a collaborative navigation task in a mixed-space environment. Both studies required a user with an exocentric view of a virtual room to navigate a fully immersed user with an egocentric view to an exit. The first study compares natural hand-based gestures, a mouse-based interface and an audio-only technique to determine their relative efficiency on task completion times. The follow-up study compares natural handbased gestures against a mouse-based interface in a scenario in which participants are unable to communicate verbally.

The results show that visual cue-based collaborative navigation techniques are significantly more efficient than an audio-only technique. The results also show that natural hand gestures are more expressive and lead to quicker completion times in situations where verbal communication is not possible.

## **Sketching ER Diagrams**

*Paul Schmieder, Beryl Plimmer and Gill Dobbie*

**4:00pm, Rangimarie Room 3**

**Abstract** Hand-drawn diagrams are frequently used as the first visualization of a model. Converting these preliminary diagrams into a specific formal format is time consuming. Computer based sketch-tools can offer support during the informal sketching stage and automatic conversion to formal representations. Entity Relationship diagrams are particularly difficult to convert because of their characteristics such as cardinality notations. We extend the general diagram sketching tool InkKit with domain semantics to successfully recognize and automatically convert Entity Relationship diagrams. This approach takes advantage of sketching as the preferred initial design realization while minimizing the effort required to translate the initial design to a functional prototype.

## Assessing Usability for Input Operation using Frequency Components of Eye Movements

*Minoru Nakayama and Makoto Katsukura*

Tuesday, 4:30pm, Rangimarie Room 3

**Abstract** Dynamic usability-testing is required for the improvement of various Human-Computer interactive systems. This paper examines the indices of assessing usability using the frequency spectrum of eyemovements up to four Hz. An input operational task experiment was conducted using a mouse (Mouse), a keyboard (KeyBD) and a keypad (KeyPAD), and the conventional subjective system usability measurements (SU-scores) and error rates were measured. Most power spectrum densities (PSD) for eyemovements in the first second of the experiment followed the same order of the SU-scores or error rates. Cross spectrum densities (CSD) between horizontal and vertical eye-movements and coherence as standardized CSD also significantly correlate with the results of the SU-scores and error rates. To determine the frequency range of CSD and coherence for usability assessment, frequency components used as factors were extracted using factor analysis. According to the correlation coefficients between these and the performance of factor scores for predicting the conventional metrics, factor scores of CSD can be better indices for assessing usability than can indices of coherence.

## Multi-Platform Document-Oriented GUIs

*James Kim and Christof Lutteroth*

Tuesday, 5:00pm, Rangimarie Room 3

**Abstract** In recent years, increasing complexity of graphical user interfaces (GUIs) of applications has led to problems in GUI management, since there is no single layout to fulfill every user's needs. GUI editors have been developed to enhance end-user options but they commonly fail to preserve personalized GUIs. This paper presents an extension to the GUI editor built into the Auckland Layout Model (ALM) that can permanently store user-defined GUI layouts and reproduce them on different platforms. A novel technique called the document-oriented approach has been exploited to empower end-user customization, which allows GUI layouts to be dynamically edited, saved using a standardized XML-based GUI description language, and loaded in a platform-independent manner.

## Using Machinima to Promote Computer Science Study

*Christian Jones and Callum Munro*

Tuesday, 5:30pm, Rangimarie Room 3

**Abstract** The project develops a Machinima animation promotional film for the Department of Computer Science at Heriot-Watt University, and evaluates whether the promotional material is more engaging and entertaining than traditional materials; more informative about the provider (its values and fa-

ilities); promotes Computer Science as relevant to student lives; uses media immediately recognisable to the student; and is innovative and differentiates the provider from others in the marketplace. Real interviews with current students relating their likes and dislikes of the University were coupled with computer game animation to create an entertaining and informative multimedia advertisement. The multi-stage evaluation with potential applicants has shown that these students respond positively to the content and presentation of the innovative multimedia film, and are encouraged to choose Computer Science and Heriot-Watt University for Tertiary education.

## 7 AUSGRID: Australasian Symposium on Grid Computing and e-Research

### Stream-Components: Component based Stream computation on the Grid

*Paul Martinaitis and Andrew Wendelborn*

Wednesday, 9:00am, Rangimarie Room 3

**Abstract** This paper reports on an investigation into component-based design in object-based distributed systems. We focus on the modeling of stream processing in terms of components (as exemplified in the EU CoreGrid project ProActive), to show how a stream processing system can be built from objects, and components in composition. In particular, we explore mechanisms for dynamic reconfiguration and distributed management of streams in a grid context. In further work, we will examine the relationship of such design to grid workflows and web services.

### Optimizing Tunneled Grid Connectivity across Firewalls

*Jefferson Tan, David Abramson and Colin Enticott*

Wednesday, 9:30am, Rangimarie Room 3

**Abstract** Grids today generally assume that concurrent network connections are possible among many processors attached to high-capacity networks. However, inter-network boundaries dividing independent institutions often have firewalls, typically to restrict how many and which ports are accessible. In some cases, ports are opened indefinitely for Grid applications, but this compromises security significantly. On the other hand, solutions that manage port openings in an ad-hoc manner for applications are non-trivial to implement. An alternative firewall traversal technique is required that will provide manageable openings with less complexity involved. This is possible through proxies and managed tunnels using ports already authorized across the firewalls. We have developed a transparent connectivity mechanism for this, called Remus, which reroutes Grid connections through a tunnel on ports allowed across firewalls. However, a single tunnel presents a performance bottleneck. In this paper, we present the method by which Remus distributes several connections over multiple tunnels, improving throughput as a result. Rerouting wrappers hide the tunneling from applications, intercepting outgoing connections and rerouting them transparently. Well-known and mature tools and protocols, such as SSH and/or SOCKS, are utilized, instead of imposing untried and experimental mechanisms. Results of our experiments are also presented for large file transfers over a Globus-based Grid that uses Remus.

## Using Markov Chain Analysis to Study Dynamic Behaviour in Large-Scale Grid Systems

*Christopher Dabrowski and Fern Hunt*

Wednesday, 1:30pm, Rangimarie Room 3

**Abstract** In large-scale grid systems with decentralized control, the interactions of many service providers and consumers will likely lead to emergent global system behaviours that result in unpredictable, often detrimental, outcomes. This possibility argues for developing analytical tools to allow understanding, and prediction, of complex system behaviour in order to ensure availability and reliability of grid computing services. This paper presents an approach for using piece-wise homogeneous Discrete Time Markov chains to provide rapid, potentially scalable, simulation of large-scale grid systems. This approach, previously used in other domains, is used here to model dynamics of large-scale grid systems. In this approach, a Markov chain model of a grid system is first represented in a reduced, compact form. This model can then be perturbed to produce alternative system execution paths and identify scenarios in which system performance is likely to degrade or anomalous behaviours occur. The expeditious generation of these scenarios allows prediction of how a larger system will react to failures or high stress conditions. Though computational effort increases in proportion to the number of paths modelled, this cost is shown to be far less than the cost of using detailed simulation or testbeds. Moreover, cost is unaffected by size of system being modelled, expressed in terms of workload and number of computational resources, and is adaptable to systems that are non-homogenous with respect to time. The paper provides detailed examples of the application of this approach.

## A Min-Min Average Algorithm for Scheduling Transaction-Intensive Grid Workflows

*Ke Liu, Jinjun Chen, Hai Jin and Yun Yang*

Wednesday, 2:00pm, Rangimarie Room 3

**Abstract** Transaction-intensive grid workflows are attracting more and more attentions with the prosperity of e-business and e-government applications. They are workflows normally with a huge number of relatively simple concurrent instances, such as business transactions, whilst some of which may involve considerable communication overheads. However, there are almost no specific scheduling algorithms which deal with such workflows, and existing scheduling algorithms are not efficient enough for such a scenario if corresponding adjustments are not conducted. To address this problem, we propose a novel Min-Min-Average (MMA) algorithm for efficiently scheduling transaction-intensive grid workflows involving considerable communication overheads. The MMA algorithm is based on the popular Min-Min algorithm but uses a different strategy for transaction-intensive grid workflows with the capability of adapting to the change of network transmission speed automatically. The comparison based on

the simulation performed on SwinDeW-G, our peer-to-peer based grid workflow environment, demonstrates that the MMA algorithm can improve the scheduling performance significantly over the original Min-Min algorithm when scheduling transaction-intensive grid workflows with considerable communication overheads involved.

## **A Small-World Network Model for Distributed Storage of Semantic Metadata**

*Arno Leist and Ken Hawick*

**Wednesday, 2:30pm, Rangimarie Room 3**

**Abstract** The growing uptake of semantic web and grid ideas is raising the importance of optimising distribution algorithms for semantic metadata. While it is not yet clear how real-world metadata distribution patterns ought to evolve, practical experience of social and technical networks suggests that a small-world pattern is desirable and practical. We explore simulated small-world networks of semantic metadata and some graph parameters and metrics. We discuss the implications of inter- and intra-domain path lengths for semantic queries on web and grid structures.

## **Node-level Architecture Design and Simulation of the MAGOG Grid Middleware**

*Jeremy Cohen, Uli Harder, Fernando Martinez Ortuno, Colin Richardson and John Darlington*

**Wednesday, 3:00pm, Rangimarie Room 3**

**Abstract** The Middleware for Activating the Global Open Grid (MAGOG) is a middleware design developed by Dr Colin Richardson while working at the Internet Centre within the Department of Computing at Imperial College London, UK. The MAGOG middleware provides an elegant and novel solution to the problem of discovering remote resources in a globally interconnected environment such as the Internet, in situations where users want to gain access to such resources to carry out remote computation. While existing Grid middleware enables the building of Grid infrastructures within closed environments where all users are known to each other, or where there is some pre-existing relationship between resource providers and users, the true Grid model should enable any user at any location to access remote resources without any prior relationship with the provider. The MAGOG middleware provides an architecture to enable the discovery of resources in such an environment and to enable the agreement of pricing and Service Level Agreements (SLAs) for the use of these resources. This paper provides an overview of the MAGOG middleware design and early simulation work that has been carried out to verify this design. It then focuses on the initial design for the infrastructure that players in the market, enabled by the middleware, will need to deploy in order to become a node in the environment.

## **A Grid Resource Allocation Mechanism for Heterogeneous E-waste Computers**

*Timothy Lynar, Ric Herbert, William Chivers and Simon*  
**Wednesday, 4:00pm, Rangimarie Room 3**

**Abstract** While most grids and clusters are built from uniform nodes of new hardware, the notion of a grid of obsolete computers rescued from the scrapheap appeals in several ways such as lower environmental and financial cost. Existing resource allocation mechanisms, in assuming that nodes are equally capable, equally power efficient and equally reliable, fail to cater for the distinct features of an e-waste grid. This paper presents a resource allocation mechanism explicitly designed with the features of such an e-waste grid in mind, proposes an objective function to assess the mechanism, and tests the mechanism against two existing resource allocation mechanisms. In simulated tests of the three mechanisms, the new mechanism performs better than the other two, particularly under heavy load.

## **An Approach to Vickrey-based Resource Allocation in the Presence of Monopolistic Sellers**

*Pham Hai Nam and Yong Teo*  
**Wednesday, 4:30pm, Rangimarie Room 3**

**Abstract** Market-based approaches proposed recently proved to be promising for competitive resource sharing in peer-to-peer and grid computing. Many approaches leverage on the Vickrey-Clarke-Groves (VCG) mechanism to achieve incentive compatibility which embraces truthful bidding of participating agents. This paper addresses a deficiency of VCG that to the best of our knowledge has not been studied. When one or more agents possess a large portion of the market share of resource, a monopoly situation arises. Applying VCG mechanism does not lead to an allocation because the second price cannot be mathematically determined. Using both theoretical and simulation analysis, we show the importance of addressing this problem. Our results show that monopoly situation arises in many types of market settings, from auction to exchange, and with a relatively high occurrence rate. To address this, we propose a new pricing method suitable for many market settings that achieve budget balanced and economic efficiency but relax the strategy proof property.

## **Impact of Grid Computing in Structural Biology**

*Ashley Buckle*  
**Wednesday, 5:00pm, Rangimarie Room 3**

**Invited Talk Abstract** Structural biology aims to understand the function of proteins and other large molecules by determining their atomic structures in 3D. The technique of protein crystallography is typically used to solve this 10,000 piece 3D jigsaw puzzle. One such crystallography method, termed Molecular

Replacement (MR), uses known structures that are predicted to share some degree of structural likeness to the target, to kick-start the puzzle-solving process. This can be a trial-and-error procedure involving testing tens to thousands of starting structures in parallel, placing demands upon computational resources. In order to address this problem we have developed a hierarchical grid-based approach that leverages a range of distributed computational resources. Generally, the approach performs multiple MR calculations across a grid of networked computers, permitting high-throughput MR. We have leveraged three different classes of resource that were available to us, namely local research laboratory based computers; a 1000 CPU Condor pool at Monash University; and a World Wide Grid of machines leveraging computers in an Australian University Enterprise Grid, the PRAGMA testbed, the Open Science Grid and the Swiss National Grid. This has allowed us to perform high throughput MR calculations, thereby increasing the likelihood of determining the atomic structures of proteins.

## 8 CATS: Computing - The Australasian Theory Symposium

### Spreading of messages in random graphs

*Ching-Lueh Chang and Yuh-Dauh Lyuu*

Tuesday, 4:00pm, Angus Room

**Abstract** Chang and Lyuu study the spreading of a message in an Erdős-Rényi random graph  $G(n, p)$  starting from a set of vertices that are convinced of the message initially. In their strict-majority scenario, whenever more than half of the neighbors of a vertex  $v$  are convinced of a message,  $v$  itself also becomes convinced. The spreading proceeds asynchronously until no more vertices can be convinced. Following Chang and Lyuu, we derive lower bounds on the minimum number  $\text{min-seed}(n, p)$  of vertices that need to be convinced initially so that all vertices will be convinced at the end. Our main results are that  $\text{min-seed}(n, p) = \Omega(\min\{n, p^2 n^2\})$  and  $\text{min-seed}(n, p) = \Omega(n^{2/3})$  hold with high probability. We also consider the case of random seeds. For any sufficiently large constant  $c > 0$  and any  $s \leq n/(c \ln n)$ , we show that if one picks the set of seeds uniformly at random from the family of all  $s$ -sized sets, then with high probability, not all vertices will be convinced at the end.

### Minimum Cost Homomorphism to Oriented Cycles with Some Loops

*Mehdi Karimi and Arvind Gupta*

Tuesday, 4:30pm, Angus Room

**Abstract** For digraphs  $D$  and  $H$ , a homomorphism of  $D$  to  $H$  is a mapping  $f : V(D) \rightarrow V(H)$  such that  $uv \in A(D)$  implies  $f(u)f(v) \in A(H)$ . Suppose  $D$  and  $H$  are two digraphs, and  $c_i(u)$ ,  $u \in V(D)$ ,  $i \in V(H)$ , are nonnegative real costs. The cost of the homomorphism  $f$  of  $D$  to  $H$  is  $\sum_{u \in V(D)} c_{f(u)}(u)$ . The minimum cost homomorphism for a fixed digraph  $H$ , denoted by  $\text{MinHOM}(H)$ , asks whether or not an input digraph  $D$ , with nonnegative real costs  $c_i(u)$ ,  $u \in V(D)$ ,  $i \in V(H)$ , admits a homomorphism  $f$  to  $H$  and if it admits one, find a homomorphism of minimum cost. The minimum cost homomorphism problem seems to offer a natural and practical way to model many optimization problems such as list homomorphism problems, retraction and precolouring extension problems, chromatic partition optimization, and applied problems in repair analysis.

## Testing Square-Freeness of Strings Compressed by Balanced Straight Line Program

*Wataru Matsubara, Shunsuke Inenaga and Ayumi Shinohara*

Tuesday, 5:00pm, Angus Room

**Abstract** In this paper we study the problem of deciding whether a given compressed string contains a square. A string  $x$  is called a square if  $x = zz$  and  $z = u^k$  implies  $k = 1$  and  $u = z$ . A string  $w$  is said to be square-free if no substrings of  $w$  are squares. Many efficient algorithms to test if a given string is square-free, have been developed so far. However, very little is known for testing square-freeness of a given compressed string. In this paper, we give an  $O(\max(n^2, n \log^2 N))$ -time  $O(n^2)$ -space solution to test square-freeness of a given compressed string, where  $n$  and  $N$  are the size of a given compressed string and the corresponding decompressed string, respectively. Our input strings are compressed by balanced straight line program (BSLP). We remark that BSLP has exponential compression, that is,  $N = O(2^n)$ . Hence no decompress-then-test approaches can be better than our method in the worst case.

## On Process Complexity

*Adam Day*

Wednesday, 9:00am, Angus Room

**Abstract** Process complexity is one of the basic variants of Kolmogorov complexity. Unlike plain Kolmogorov complexity, process complexity provides a simple characterization of randomness for real numbers in terms of initial segment complexity. Process complexity was first developed by Schnorr. Schnorr's definition of a process, while simple, can be difficult to work with. In many situations, aevin. In this paper we define a variant of process complexity based on Levin's definition of a process. We call this variant strict process complexity. Strict process complexity retains the main desirable properties of process complexity. Particularly, it provides simple characterizations of Martin-Lof random real numbers, and of computable real numbers. However, we will prove that strict process complexity does not agree within an additive constant with Schnorr's original process complexity. One of the basic properties of prefix-free complexity is that it is subadditive. Subadditive means that there is some constant  $d$  such that for all strings  $x, y$  the complexity of  $xy$  ( $x$  and  $y$  concatenated) is less than or equal to the sum of the complexities of  $x$  and  $y$  plus  $d$ . A fundamental question about any complexity measure is whether or not it is subadditive. In this paper we resolve this question for process complexity by proving that neither of these process complexities is subadditive.

## Reasoning about a distributed probabilistic system

*Ukachukwu Ndukwu and J.W. Sanders*

Wednesday, 9:30am, Angus Room

**Abstract** Reasoning about a distributed system that exhibits a combination of probabilistic and temporal behaviour does not seem to be easy with current techniques. The reason is the interaction between probability and abstraction (local block), made worse by remote synchronisation. The formalism of process algebra has not so far provided much insight, and so the alternative of shared-variable concurrency has been explored. In this paper the recently proposed language PTSC (for probability, time and shared-variable concurrency) is extended by constructs for interleaving and local block. Both enhance a designer's ability to modularise a design; the latter also permits a design to be compared with its more abstract specification, by concealing appropriately chosen design variables. Laws of the extended language are studied and applied in a case study consisting of a faulty register-transfer-level design.

## Augmenting Edge-Connectivity between Vertex Subsets

*Toshimasa Ishii and Kazuhisa Makino*

Wednesday, 1:30pm, Angus Room

**Abstract** Given a graph  $G = (V, E)$  and a requirement function  $r : \mathcal{W}_1 \times \mathcal{W}_2 \rightarrow \mathbb{R}_+$  for two families  $\mathcal{W}_1, \mathcal{W}_2 \subseteq 2^V - \{\emptyset\}$ , we consider the problem (called *area-to-area edge-connectivity augmentation problem*) of augmenting  $G$  by a smallest number of new edges so that the resulting graph  $\hat{G}$  satisfies  $\delta_{\hat{G}}(X) \geq r(W_1, W_2)$  for all  $X \subseteq V$ ,  $W_1 \in \mathcal{W}_1$ , and  $W_2 \in \mathcal{W}_2$  with  $W_1 \subseteq X \subseteq V - W_2$ , where  $\delta_G(X)$  denotes the degree of a vertex set  $X$  in  $G$ . This problem can be regarded as a natural generalization of the global, local, and node-to-area edge-connectivity augmentation problems. In this paper, we show that there exists a constant  $c$  such that the problem is inapproximable within a ratio of  $c \log \alpha(\mathcal{W}_1, \mathcal{W}_2)$ , unless  $P=NP$ , even restricted to the directed global node-to-area edge-connectivity augmentation or undirected local node-to-area edge-connectivity augmentation, where  $\alpha(\mathcal{W}_1, \mathcal{W}_2)$  denotes the number of pairs  $W_1 \in \mathcal{W}_1$  and  $W_2 \in \mathcal{W}_2$  with  $r(W_1, W_2) > 0$ .

## A Simple Algorithm For Triconnectivity of a Multigraph

*Abusayeed Saifullah and Alper Ungor*

Wednesday, 2:00pm, Angus Room

**Abstract** Vertex-connectivity and edge-connectivity represent the extent to which a graph is connected. Study of these key properties of graphs plays an important role in varieties of computer science applications. Recent years have witnessed a number of linear time 3-edge-connectivity algorithms - with increasing simplicity. In contrast, the state-of-the-art algorithm for 3-vertex-connectivity due to Hopcroft and Tarjan lacks the simplicity in the sense of ease

of implementation though its time and space complexity is theoretically linear. In this paper, we propose a linear time reduction from 3-vertex-connectivity to 3-edge-connectivity of a multigraph. This reduction was previously unknown, while the reduction in the opposite direction already exists. We apply an existing linear time 3-edge-connectivity algorithm on the reduced graph for solving the 3-vertex-connectivity of the original graph. Hence, for a graph with  $V$  vertices and  $E$  edges, the proposed reduction turns into an  $O(V + E)$  time and space algorithm for 3-vertex-connectivity while enjoying the simplicity of the 3-edge-connectivity algorithms.

## **Type Checking and Inference for Polymorphic and Existential types**

*Koji Nakazawa and Makoto Tatsuta*

**Wednesday, 2:30pm, Angus Room**

**Abstract** This paper proves undecidability of type checking and type inference problems in some variants of typed lambda calculi with polymorphic and existential types. First, type inference in the domain-free polymorphic lambda calculus is proved to be undecidable, and then it is proved that type inference is undecidable in the negation, conjunction, and existence fragment of the domain-free typed lambda calculus. Secondly, their variants with multiple applications of the quantifier rules are introduced, and their type checking and type inference are proved to be undecidable. Finally, it is proved that we can reduce undecidability of type checking and type inference problems in the Curry-style lambda calculus in negation, conjunction, and existential fragment to undecidability of those problems in another variant of the domain-free polymorphic lambda calculus.

## **Transformation Rules for Z**

*Mark Utting, Petra Malik and Ian Toyn*

**Wednesday, 3:00pm, Angus Room**

**Abstract** Z is a formal specification language combining typed set theory, predicate calculus, and a schema calculus. This paper describes an extension of Z that allows transformation and reasoning rules to be written in a Z-like notation. This gives a high-level, declarative way of specifying transformations of Z terms, which makes it easier to build new Z manipulation tools. We describe the syntax and semantics of these rules, plus some example reasoning engines that use sets of rules to manipulate Z terms. We demonstrate the utility of rules by discussing two sets of rules. One set defines unfolding of the schema expressions in Z. The other set is used by the ZLive animator to preprocess Z expressions into a form more suitable for animation.

## Computing Nash Equilibria Gets Harder New Results Show Hardness Even For Parameterized Complexity

*Vlad Estivill-Castro and Mahdi Parsa*

Wednesday, 4:00pm, Angus Room

**Abstract** In this paper we show that some decision problems regarding the computation of Nash equilibria are to be considered particularly hard. Most decision problems regarding Nash equilibria have been shown to be NP-complete. While some NP-complete problems can find an alternative to tractability with the tools of Parameterized Complexity Theory, it is also the case that some classes of problems do not seem to have fixed-parameter tractable algorithms. We show that K-UNIFORM NASH and K-MINIMAL NASH SUPPORT are  $W[2]$ -hard. Given a game  $G=(A,B)$  and a nonnegative integer  $k$ , the K-UNIFORM NASH problem asks whether  $G$  has a uniform Nash equilibrium of size  $k$ . The K-MINIMAL NASH SUPPORT asks whether  $G$  has Nash equilibrium such that the support of each players Nash strategy has size equal or less than  $k$ . First, we show that K-UNIFORM NASH (with  $k$  as the parameter) is  $W[2]$ -hard even when we have 2 players, or less than 4 different integer values in the matrices. Second, we illustrate that even in zero-sum games K-MINIMAL NASH SUPPORT is  $W[2]$ -hard (a sample Nash equilibrium in a zero-sum 2-player game can be found in polynomial time (von Stengel, 2002)). Thus, it must be the case that other more general decision problems ter tractability in those decision problems regarding Nash equilibria seem elusive.

## Lower Bounds on Quantum Query Complexity for Read-Once Decision Trees With Parity Nodes

*Hideaki Fukuhara and Eiji Takimoto*

Wednesday, 4:30pm, Angus Room

**Abstract** We introduce a complexity measure for decision trees called the soft rank, which measures how well a given tree is balanced. The soft rank is a somehow relaxed variant of the rank. Among all decision trees of depth  $d$ , the complete binary decision tree (the most balanced tree) has maximum soft rank  $d$ , the decision list (the most unbalanced tree) has minimum soft rank  $\sqrt{d}$ , and any other trees have soft rank between  $\sqrt{d}$  and  $d$ . We show that, for any decision tree  $T$  in some class  $G$  of decision trees which includes all read-once decision trees, the soft rank of  $T$  is a lower bound on the quantum query complexity of the Boolean function that  $T$  represents. This implies that for any Boolean function  $f$  that is represented by a decision tree in  $G$ , the deterministic query complexity of  $f$  is only quadratically larger than the quantum query complexity of  $f$ .

**Longest Paths in Planar DAGs in Unambiguous Logspace**  
*Nutan Limaye, Meena Mahajan and Prajakta Nimbhorkar*  
**Wednesday, 5:00pm, Angus Room**

**Abstract** Reachability and distance computation are known to be NLOG-complete in general graphs, but within ULOG and co-ULOG if the graphs are planar. However, finding longest paths is known to be NP-complete, even for planar graphs. We show that with the combination of planarity and acyclicity, finding the length of the longest path (and also enumerating one such path) is also in ULOG and in co-ULOG. The result extends to toroidal DAGs as well. We also address the question of when the three problems are indeed equivalent on DAGs, and give partial bounds. When the number of distinct paths is bounded by a polynomial, counting the number of paths is known to be in NLOG. We show that for planar DAGs with this promise, counting can be done by an unambiguous AuxPDA. The UAuxPDA bound also holds if we want to compute the number of longest paths, or shortest paths, and this number is bounded by a polynomial (irrespective of the total number of paths). Along the way, we show that counting in general DAGs is possible in LogDCFL provided the number of paths is bounded by a polynomial and the target node is the only sink.

**Formal Model of a Protocol Converter**  
*Jing Cao and Albert Nymeyer*  
**Thursday, 9:00am, Angus Room**

**Abstract** Reuse of components is a burgeoning field in chip design. Shorter time to market and assured quality are just two good reasons to reuse previously engineered components. Problems arise however when chip designers need to interface these components as they typically conform to different standards, or no standards at all. The components that we deal with in this work are protocols. The popular model of interfacing is via a ‘converter’ that translates data between the components. We develop a theoretical model of a converter that will enable two arbitrary protocols to communicate. We formally define correctness conditions, and guarantee that the resulting converter satisfies these conditions. We suggest a number of conditions (expressed in CTL formulas) that any converter would be expected to satisfy, but also allow the designer to define his own (CTL) conditions. As well, we allow protocols to be non-deterministic, and we ensure only valid data is sent to the converter. The verification of the conditions is carried out by a model checker. We have implemented our theoretical model and we present experimental results.

## Boolean Affine Approximation with Binary Decision Diagrams

*Kevin Henshall, Peter Schachte, Harald Sondergaard and Leigh Whiting*

Thursday, 9:30am, Angus Room

**Abstract** Selman and Kautz's work on knowledge compilation has established how approximation (strengthening and essing, at the expense of completeness. In the classical approach, the knowledge-base is assumed to be presented as a propositional formula in conjunctive normal form (CNF), and Horn functions are used to over- and under-approximate it (in the hope that many queries can be answered efficiently using the approximations only). However, other representations are possible, and functions other than Horn can be used for approximations, as long as they have deduction-computational properties similar to those of the Horn functions. Zanuttini has suggested that the class of affine Boolean functions would be especially useful in knowledge compilation and has presented various affine approximation algorithms. Since CNF is awkward for presenting affine functions, Zanuttini considers both a sets-of-models representation and the use of modulo 2 congruence equations. Here we consider the use of reduced ordered binary decision diagrams (ROBDDs), a representation which is more compact than the sets of models and which (unlike modulo 2 congruences) can express any source knowledge-base. We present an ROBDD algorithm to find strongest affine upper-approximations of a Boolean function and we argue its correctness.

## Structural Properties of Random Graph Models

*Andras Farago*

Thursday, 1:30pm, Angus Room

**Abstract** Many different random graph constructions are used to model large real-life graphs. Often it is not clear, however, how the strength of the different models compare to each other, e.g., when does it hold that a certain model class contains another. We are particularly interested in random graph models that arise via abstract geometric constructions, motivated by the fact that these graphs can model certain wireless communication networks. We set up a general framework to compare the strength of random graph models, and present some results about the equality, inequality and proper containment of certain model classes, as well as some open problems.

## Linear Axis for Planar Straight Line Graphs

*Kira Vyatkina*

Thursday, 2:00pm, Angus Room

**Abstract** A linear axis is a straight line skeleton for a polygonal shape. The concept of a linear axis  $\epsilon$ -equivalent to the medial axis has been introduced

and studied for simple polygons and for those with holes. In this paper, we generalize the notions of a linear axis and of  $\epsilon$ -equivalence to the case of planar straight line graphs. We show that for some graphs, a linear axis  $\epsilon$ -equivalent to the medial axis does not exist, for any  $\epsilon > 0$ . However, if the graph vertices are in general position, a sought linear axis does exist for any  $\epsilon > 0$ , and can be computed in  $O(n \log n)$  time in the absence of certain correlations in the graph structure.

## **Edge-Selection Heuristics for Computing Tutte Polynomials**

*David Pearce, Gary Haggard and Gordon Royle*

**Thursday, 2:30pm, Angus Room**

**Abstract** The Tutte polynomial of a graph, also known as the partition function of the  $q$ -state Potts model is a 2-variable polynomial graph invariant of considerable importance in both combinatorics and statistical physics. It contains several other polynomial invariants, such as the chromatic polynomial and flow polynomial as partial evaluations, and various numerical invariants such as the number of spanning trees as complete evaluations. We have developed the most efficient algorithm to-date for computing the Tutte polynomial of a graph. An important component of the algorithm affecting efficiency is the choice of edge to work on at each stage in the computation. In this paper, we present and discuss two edge-selection heuristics which (respectively) give good performance on sparse and dense graphs. We also present experiment data comparing these heuristics against a range of others to demonstrate their effectiveness.

## **Distributing Frequency-Dependent Data Stream Computations**

*Sumit Ganguly*

**Thursday, 3:00pm, Angus Room**

**Abstract** For time-efficiency, data stream computations are often performed in a highly distributed fashion (e.g., internet applications and sensor networks). A distributed computation is modeled as a binary tree, whose leaf nodes contain the input stream fragments  $\sigma_i$ , each of which are reduced to a message  $\phi(\sigma_i)$  that is sent to its parent. Internal nodes compose the messages received from its children using a message composition function  $\odot$  and relay the composite message upwards in the tree. Finally, the root node obtains a summary composite message of the input stream fragments and processes it to return an answer. A maximally flexible distributed algorithm is one where all possible computation trees over the same set of leaf nodes compute a correct answer, and each tree is identified with a distinct distributed computation. A basic question here is: what kind of data stream computations can be distributed in a flexible manner?

## 9 HIKM: Australasian Workshop on Health Informatics and Knowledge Management

### Assessing Viewing Pattern Consistency in Mammogram Readers

*Anthony Maeder*

Thursday, 1:30pm, Rangimarie Room 3

**Abstract** Breast cancer screening programs typically require very large volumes of x-ray images (mammograms) to be viewed by highly experienced readers. The readers can recognise a wide range of different visible features indicative of abnormal situations, which they use as a basis to generate a report on their findings. Errors in reporting can occur if the readers fail to identify a particular feature of interest for further visual inspection during the viewing process. This risk is typically reduced by training readers to follow a particular viewing path through an image, which they should be able to apply consistently. Knowledge of the degree of consistency in this viewing behaviour within and between viewers would inform the development of an automated checking approach based on monitoring of viewer visual attention. This paper provides some experimental results on reader viewing pattern profiles obtained using eye tracking, and suggests a suitable consistency characterization model.

### A Classification Algorithm That Derives Weighted Sum Scores for Insight Into Disease

*Anthony Quinn, Andrew Stranieri, John L. Yearwood and Gaudenz Hafen*

Thursday, 2:00pm, Rangimarie Room 3

**Abstract** Data mining is often performed with datasets associated with diseases in order to increase insights that can ultimately lead to improved prevention or treatment. Classification algorithms can achieve high levels of predictive accuracy but have limited application for facilitating the insight that leads to deeper understanding of aspects of the disease. This is because the representation of knowledge that arises from classification algorithms is too opaque, too complex or too sparse to facilitate insight. Clustering, association and visualisation approaches enable greater scope for clinicians to be engaged in a way that leads to insight, however predictive accuracy is compromised or non-existent. This research investigates the practical applications of Automated Weighted Sum, (AWSum), a classification algorithm that provides accuracy comparable to other techniques whilst providing some insight into the data. This is achieved by calculating a weight for each feature value that represents its influence on the class value. Clinicians are very familiar with weighted sum scoring scales so the internal representation is intuitive and easily understood. This paper presents results from the use of the AWSum approach with data from patients suffering from Cystic Fibrosis.

## **Characterizing Image Properties for Digital Mammograms**

*A. Nguyen, J. Dowling, A. Maeder, P. Nguyen and E. Brunton*

**Thursday, 2:30pm, Rangimarie Room 3**

**Abstract** Adoption of computed radiology (CR) and direct radiology (DR) imaging for screening mammograms in many countries alongside digitally scanned film mammograms has resulted in a wide range of different intrinsic (physical) characteristics of images becoming common place. It is sometimes conjectured that viewer performance could be adversely affected by this wider variability, as compared with the variability that was formerly experienced with film only. This paper identifies several aspects of the image characteristics relevant to viewer perception, including intensity properties (such as contrast), spatial properties (such as texture) and structure properties (such as breast density). We then provide quantitative descriptions of the variability of these properties over a test set of 12 screening mammograms drawn from three different modalities and containing a typical mix of screening cases.

## **Privacy and Security in Open and Trusted Health Information Systems**

*Vicky Liu, William Caelli, Lauren May and Tony Sahama*

**Thursday, 3:00pm, Rangimarie Room 3**

**Abstract** The Open and Trusted Health Information Systems (OTHIS) Research Group has formed in response to the health sectors privacy and security requirements for contemporary Health Information Systems (HIS). Due to recent research developments in trusted computing concepts, it is now both timely and desirable to move electronic HIS towards privacy-aware and security-aware applications. We introduce the OTHIS architecture in this paper. This scheme proposes a feasible and sustainable solution to meeting real-world application security demands using commercial off-the-shelf systems and commodity hardware and software products.

## **Understanding The Implementation of an Electronic Hospital Information System in a Developing Country: a Case Study From Pakistan**

*Muzaffar Malik and Haroon Khan*

**Thursday, 4:00pm, Rangimarie Room 3**

**Abstract** Literature on implementation of hospital information systems is scarce, especially with regard to developing countries. Pakistan Institute of Medical Sciences (PIMS) is a large public sector hospital in Pakistan that implemented a hospital information system (HIS) successfully. This article studies how this success was achieved and examines the hurdles faced in the implementation of the HIS and how they were overcome. The article aims to provide a

better understanding of implementing HIS in a developing country setting, to add to academic knowledge in the area as well as to serve as a guide to anyone wishing to implement an HIS in such a setting.

### **GP Attitudes Towards Using HI Systems in Their Professional Role**

*John Knight, Margaret Patrickson and Bruce Gurd*

**Thursday, 4:30pm, Rangimarie Room 3**

**Abstract** This paper reports on a qualitative study of South Australian General Practitioner (GP) attitudes towards adopting Health Informatics (HI) technology. The study suggests attitudes are determined by GP perceptions of competing managerial, technological and political factors. Findings indicate increased exposure to HI use in performance of their role influences GP perceptions of the importance and certainty of implementation outcomes. However the prospect of such technologically facilitated change tends to manifest in resistance if perceived as uncertain, involuntary or not of demonstrable benefit to patients. The findings highlight the desirability of HI technology use being associated with benefits to GP patients and practices rather than with change to the GPs professional role and value.

### **Making 12,000 Healthcare Organisations Interoperate, and Other Challenges**

*Alan Hesketh - New Zealand Ministry of Health Deputy, Director-General*

**Thursday, 5:00pm, Rangimarie Room 3**