Working Paper Series
ISSN 1170-487X

PROCEEDINGS OF THE SECOND COMPUTING WOMEN CONGRESS (CWC 2006): STUDENT PAPERS

Editors:
Annika Hinze
Doris Jung
Sally Jo Cunningham

Working Paper: 02/2006
February 2006

Department of Computer Science
The University of Waikato
Private Bag 3105
Hamilton, New Zealand

Proceedings of the Second Computing Women Congress (CWC 2006): Student papers
Hamilton, New Zealand, 11-19 February 2006
Editors: Annika Hinze, Doris Jung, Sally Jo Cunningham

Preface

The Second Computing Women Congress was held at the University of Waikato, Hamilton, New Zealand from February 11th to 19th, 2006. The Computing Women Congress (CWC) is a Summer University for women in computer science. It is a meeting-place for female students, academics and professionals who study or work in Information Technology. CWC provides a forum to learn about and share the latest ideas of computing related topics in a supportive environment. CWC provides an open, explorative learning and teaching environment. Experimentation with new styles of learning is encouraged, with an emphasis on hands-on experience and engaging participatory techniques.

Each year, the CWC invites women in computing, IT, and related interdisciplinary fields to submit proposals for lectures, courses, and seminars. Contributions from the whole spectrum of computer science and also IT-related gender research are welcome. Participants from all levels of experience and with diverse backgrounds are welcome to present or teach.

The congress aims to provide role models for those early in their computing careers and also a meeting place for those well into their careers. Students at undergraduate and graduate level, and professionals from academic, scientific or commercial backgrounds are all welcome. By learning from each other's skills and experiences, we seek to form a community that shares interests and knowledge. The CWC is organised with the support from organisers of similar European events—we aim at establishing a series of annual congresses for Australasian women in IT.

In 2006, the CWC for the first time accepted student paper submissions. We invited students to submit papers describing their research or significant individual projects to a special session of the 2006 CWC. The purpose of this session is to recognize excellent work being conducted by CS/IS/IT students, and to offer a friendly forum for students to showcase their work and receive constructive feedback.
Submissions were reviewed by the organizing committee of the CWC, and the authors of accepted papers are given the opportunity to present their work to CWC attendees. This volume of accepted papers is published as a working paper of the Computer Science Department of the University of Waikato. The best student paper receives the *Margaret Jefferies Award*.

The CWC 2006 Proceedings contains the following student papers:

- Kathryn Hempstalk: *Hiding Behind Corners: Using Edges in Images for Better Steganography*
- Supawan Prompramote, Kathy Blashki: *Playing to Learn: Enhancing Educational Opportunities using Games Technology*
- Judy Bowen: Celebrity Death Match: *Formal Methods vs. User-Centred Design*
- Liz Bryce: *BECOMING INDIGENOUS: an impossible necessity*
- Tatiana King: *Privacy Issues in Health Care and Security of Statistical Databases*
- Nilufar Baghaei: *A Collaborative Constraint-based Intelligent System for Learning Object-Oriented Analysis and Design using UML*
- Sonja van Kerkhof: *Alternatives to stereotypes: some thoughts on issues and an outline of one game*

The selection of papers shows a wide range of projects in computer science: from formal methods over databases and healthcare to games and media/arts.

The *Margaret Jefferies Award* for the best student paper at the CWC was given to Kathryn Hempstalk for her paper (*Hiding Behind Corners: Using Edges in Images for Better Steganography*) about her Honours Thesis project.

We would like to thank the organizing committee of the CWC 2006 for their help in preparing the Computing Women Congress. We would also like to thank the University of Waikato, New Zealand; the Department of Computer Science; Google Incorporated, USA; and KT Design, New Zealand for their active support for the CWC 2006.

Annika Hinze, Doris Jung, Sally Jo Cunningham
Hamilton, February 2006

Hiding Behind Corners: Using Edges in Images for Better Steganography

Kathryn Hempstalk
University of Waikato
Department of Computer Science
Hamilton, New Zealand
kah18@ waikato.ac.nz

ABSTRACT
Digital steganography involves taking an electronic file and hiding it inside another electronic file. Current digital techniques do not tend to take the cover (what the message is hidden in) into account, and thus leave telltale marks on the stego-object (what the cover becomes after hiding information). Since these marks will cause people other than the intended recipient to suspect a hidden message, it is important to make them as inconspicuous as possible. This paper investigates using the cover's original information to avoid making marks on the stego-object, by hiding raw electronic files inside digital colour images. Steganalysis and machine learning is then used to evaluate the hiding process in order to ensure the information is hidden in the best possible way.

Categories and Subject Descriptors
I.4.9 [**Image Processing and Computer Vision**]: Applications – *Steganography.*

General Terms
Algorithms, Design, Security.

Keywords
Steganography, steganalysis, LSB embedding, image filtering.

1. INTRODUCTION
Hiding information by embedding secret data into an innocuous medium is often referred to as steganography. Steganography can be applied electronically by taking a message (a binary file) and some sort of cover (often a sound or image file) and combining both to obtain a "stego-object". The stego-object is essentially the cover with its redundant information replaced with the message. Unfortunately, replacing the redundant information with structured message data often results in the stego-object being an augmented version of the cover "marked" by the hidden data – and this makes it easy to suspect the presence of steganography. Most of these marks can be attributed to the hiding algorithm's lack of concern for the cover's content. If the cover's original content were taken into account then the message could be concealed in areas of the cover where it would be less likely to leave marks.

Previous attempts at achieving this goal have required the user to provide the original cover as well as the stego-object. The best areas to hide are first identified in the original cover, then these areas are mapped across to the stego-object and the hidden information is retrieved. The original cover must be provided because the information overwritten in the message hiding process may have been used to identify the best hiding areas. However, to provide the original object is not secure, because taking the differences between the two objects would be enough to suspect the existence of (and in some cases, recover) the hidden information.

This paper investigates an approach that eliminates the need for providing the original object. We use images as a cover medium and introduce two new algorithms based on using image filters to determine effective hiding places. The next section describes these algorithms. Section 3 briefly describes some steganalysis methods used to test the effectiveness of the new algorithms. Section 4 describes the experimentation performed using the algorithms and the steganalysis techniques and Section 5 describes the results of this experimentation. Finally, the paper is concluded in Section 6.

2. ALGORITHMS
The simplest way to hide binary data on an image is to use a lossless image format (such as a Bitmap) and replace the *x* least significant bits of each pixel in scan lines across the image with the binary data. This is not secure as an attacker can simply repeat the process to quickly recover the hidden information. This technique, known here as "BlindHide" because of the way it *blind*ly *hide*s information, is also not good at hiding – the initial portion of the image is left degraded while the rest remains untouched.

A tool known as "Hide and Seek for Windows 95" [1] attempts to get around the security issues in BlindHide by randomly distributing the hidden information across the image. A more modern version of this algorithm, dubbed "HideSeek", is used here. HideSeek uses a random seed (provided by hashing a password) to pick the order in which it will write to the pixels. HideSeek is much more secure than BlindHide, but does not necessarily leave the image in a better condition. The noise introduced by HideSeek is randomly placed and often causes the resulting stego-image to look speckled.

The noise left behind by both HideSeek and BlindHide is much more noticeable to the naked eye in large blocks of colour – where a single modified pixel stands out amongst its uniform neighbours. This is expressed explicitly by the Laplace formula [2]. The Laplace formula simply measures the difference between a pixel and its four touching neighbours. The magnitude of the formula increases with the colour variation and this can be used to detect steganography by counting the number of pixels at a given magnitude. Untouched images are more likely to contain a large number of pixels with zero magnitudes since there is no reason for small random variations to occur in large blocks of colour. Stego-images often contain small variations, and can be detected easily by examining Laplace magnitude counts. Therefore it seems

reasonable to suggest if a hiding algorithm were able to use the Laplace formula during embedding, it would be able to hide the information in a less noticeable way.

To do this, we introduce the FilterFirst algorithm. FilterFirst uses an edge-detecting filter, such as the Laplace formula, to find the areas of the image where there are pixels that are the least like their neighbours. It hides in the highest values of the *filter first*. Since we are only changing the x least significant bits for steganography, we can use the y most significant bits for the filter. Here x and y are integers where $1 <= x <= 7$ and $y = 8 - x$. We can guarantee that FilterFirst will be able to retrieve the information from the same pixels it hides in because the bits used in filtering are not changed by the hiding process. FilterFirst eliminates the need to provide any extra information, such as the original image, yet ensures the same pixels are used for hiding and retrieval.

However, FilterFirst is similar to BlindHide in that it is not a secure algorithm. An attacker can repeat the filtering process and retrieve the hidden information with very little effort. Whilst it should be more difficult to identify steganographic images using FilterFirst, it is much easier to retrieve the hidden information than with HideSeek.

To create an algorithm that both hides effectively and securely we combine HideSeek and FilterFirst to create BattleSteg. BattleSteg stands for *Battle*ships *Steg*anography and is based on playing an augmented game of Battleships to determine the best places to hide. In this algorithm, the h% of highest filter values is designated as 'ships'. 'Shots' are randomly picked as pixel positions on the cover image, until a ship is found (known as a 'hit'). When a hit occurs, the next series of shots are clustered around that hit in the hope there are more ships in that area. After i initial shots we return to shooting randomly to prevent huge expanding clusters of shots – which may cause noticeable visual degradation on the stego-image. BattleSteg is overall more likely to avoid pixels in large blocks of colour than HideSeek, yet has a similar amount of security as without the random seed it is impossible to know where to place the shots.

3. STEGANALYSIS

Just as clever techniques have been devised for hiding information, an equal number of clever techniques have been designed to detect the hidden information [3]. These techniques are collectively known as 'steganalysis'. As introduced earlier, the Laplace formula is one such steganalytic method. Two other popular techniques are RS Analysis [4] and Sample Pairs Analysis [5].

RS Analysis makes small modifications to the least significant bit plane in an image then uses these modifications and a discrimination function to classify groups of pixels. The counts of the groups based on the modifications allow the calculation of an estimated embedding rate. Images that do not contain steganography often have a natural embedding rate of up to 3%, whereas images containing hidden information usually have estimated embedding rates which accurately reflects the amount of hidden information.

RS Analysis is a special case of Sample Pairs Analysis, which also uses least significant bit modifications to help calculate an estimated embedding rate. Sample Pairs Analysis utilises finite state machines to classify groups of pixels modified by a given pattern. Both steganalysis techniques are very accurate at predicting the embedding rate on stego-images using least significant bit embedding. Since the two proposed techniques, FilterFirst and BattleSteg, both use least significant bit embedding, we can use RS Analysis and Sample Pairs Analysis to compare them against more traditional techniques such as BlindHide and HideSeek.

4. EXPERIMENTAL DESIGN

Rather than evaluating a set of images by calculating the steganalysis information and checking the values by hand, machine learning is used. The idea behind this is simple – if a machine learner has trouble accurately predicting whether steganography is present, then the steganographic method is more effective for that picture than a method where the machine learner can correctly classify the image. Across many images, the accuracy of all predictions should indicate the effectiveness of one algorithm over another.

For testing in this paper, 100 images were combined with 2 messages. Both messages were of set length, s, one containing random data and the other the text of Moby Dick. Each algorithm was set to write to only the least significant bit of each colour (red, green, blue). 200 stego-images were then obtained by embedding each message in each image using a given steganography technique.

The 200 stego-images and 100 original images then had all their steganalysis values calculated and added to a file. This file contained the estimated embedding rates from RS Analysis and Sample Pairs Analysis, as well as counts of occurrences of each Laplace formula value encountered in every image. The file was then used to train SMO [6][7], a support vector machine learning algorithm from the WEKA Machine Learning Workbench [8].

SMO begins by identifying features that distinguish untouched images from stego-images. The machine learner runs through 10 times 10-fold cross-validation, giving the accuracy values that indicate the effectiveness of the algorithm being tested. More effective algorithms will have lower accuracy rates through cross-validation, as SMO will be unable to determine features that distinguish whether an image contains steganography. This process was repeated for each algorithm and for different values of s.

5. EXPERIMENTAL RESULTS

Using embedding rates varying from 5% to 55% (by which time all algorithms were scoring ~100% accuracy), the four algorithms were evaluated using WEKA's SMO. BattleSteg and FilterFirst were run with two variations – once using the Laplace formula for filtering and once using the Sobel filter [9] (a gradient based edge detection algorithm). As the results in Figure 1 show, the machine learner could do no better predict steganography (the majority for the image set) for FilterFirst for embedding rates up to 10%. The three other algorithms, BlindHide, BattleSteg and HideSeek, all had accuracy rates up at 90% before they even got to a 10% embedding rate.

Figure 1 also shows that FilterFirst is much more effective at hiding than the two traditional algorithms, HideSeek and BlindHide. Unfortunately, BattleSteg did not perform as well as expected, beating only HideSeek. BlindHide appears to sit between BattleSteg and FilterFirst, but should be avoided as a steganographic technique due to its lack of security and ability to often be detectable by the naked eye.

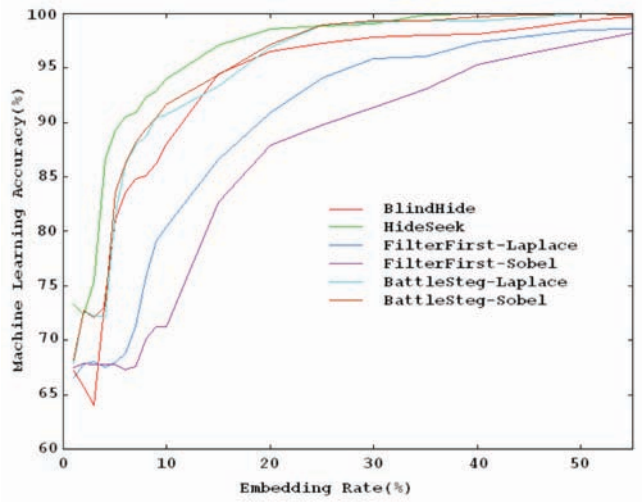


Figure 1: Graph of Embedding Rate versus Machine Learning Accuracy

These results seem to suggest that FilterFirst should be the algorithm of choice for effective steganography. By simply shuffling the order in which the highest filter bits are written to, it becomes a secure and effective hiding technique. A possible alternative to shuffling is to use a fixed proportion of the highest-filter valued pixels and then select the remaining pixels randomly. BattleSteg does attempt to do this, but it cannot guarantee the proportion of high filter valued pixels it writes to. It is possible for BattleSteg to never have a 'hit' whilst using this algorithm.

6. CONCLUSION

This paper has introduced two new techniques for image steganography, FilterFirst and BattleSteg. These two techniques attempt to improve on the effectiveness of the hiding by using edge detection filters to produce better steganography. These techniques have been tested against more traditional image steganography techniques, BlindHide and HideSeek, by using steganalysis methods.

Using a machine learner to predict the use of steganography on images, we have shown that the proposed FilterFirst algorithm is very effective at hiding information. FilterFirst beat all the steganalysis techniques until embedding rates became greater than 7% and performed better than all other steganography algorithms tested. The results indicate that using features of the cover, such as edges, is a better way of hiding information than in scan lines or randomly across the image.

The steganalysis and steganography algorithms discussed in this paper have been implemented into a tool for ease of use during this research. This tool has been made available for free under the Gnu General Public License at: http://diit.sourceforge.net

7. REFERENCES

[1] Maroney, C. Hide and Seek 5 for Windows 95, computer software and documentation, originally released in Finland and the UK, n.d. Downloadable from: http://www.rugeley.demon.co.uk/security/hdsk50.zip

[2] Katzenbeisser, S. and Petitcolas F.A.P. *Information hiding techniques for steganography and digital watermarking*. Artech House, Norwood, MA 02062, USA, 1999.

[3] Wang, S. and Wang, H. Cyber Warfare: Steganography vs. Steganalysis, *Communications of the ACM Volume 47, Number 10*, pp 76-82, 2004

[4] Fridrich, J., Goljan, M. and Du,R. (2001) Reliable Detection of LSB Steganography in Color and Grayscale Images, *Proceedings of ACM Workshop on Multimedia and Security,* Ottawa, October 5, 2001, pp. 27-30.

[5] Wang, Z., Dumitrescu, S. and Wu, X. Detection of LSB Steganography via Sample Pair Analysis, *Information Hiding 2002, LCNS 2578*, pp 355-372, 2003.

[6] Platt, J. *Fast Training of Support Vector Machines using Sequential Minimal Optimization*. Advances in Kernel Methods - Support Vector Learning, B. Schoelkopf, C. Burges, and A. Smola, eds., MIT Press, 1998.

[7] Keerthi, S.S., Shevade, S.K., Bhattacharyya, C. and Murthy, K.R.K. *Improvements to Platt's SMO Algorithm for SVM Classifier Design*. Neural Computation, 13(3), pp 637-649, 2001.

[8] Frank, E. and Witten, I.H. *Data Mining*, Morgan Kaufmann Publishers, San Francisco, USA, 1999.

[9] Green, B. Edge Detection Tutorial, 2002, Available at: http://www.pages.drexel.edu/~weg22/edge.html (May 2005)

Playing to Learn: Enhancing Educational Opportunities using Games Technology

Supawan Prompramote Kathy Blashki
School of Information Technology
Faculty of Science and Technology
Deakin University, Melbourne, Australia
{sprom,kblashki}@deakin.edu.au

Abstract: There is a marked absence of research on the development of games to enhance the learning process of young people in educational settings. Importantly, games can be usefully and successfully deployed as the primary learning vehicle and not simply for the purposes of review and reinforcement as is the current ubiquitous use. Game-based learning has considerable capacity for integrating a diversity of subject matter, pedagogic approaches and learning styles. To this end, this paper describes a programming tool, GBuilder, with the express purpose of enabling and empowering the students to develop their own learning programs in survival literacy within a game environment. Unlike previous programming environments for children, GBuilder supports a transpose of knowledge in the model to a standard programming language, which is helpful later when young children need to modify programs written in conventional languages.

1. Introduction

Play is high-valued attribution as a component of learning experiences when it comes to computers and play. The term of 'computer' and 'play' automatically brings to mind the computer-based game, which means fun and non-education to children (Roussou 2004). To encourage and enhance children learning experiences, it is essential to exploit their natural recreational enjoyment of computer games and transpose those levels of engagement to the learning environment.

The concept of game-based learning is indeed not original. There have been a number of pioneers attempted to find new ways of combination educational content and computer games since the last two decades. Unfortunately, the games were not successful in market. The reason why the market totally crashed was the software was pretty boring – kids were not interested (Prensky 2001:7).

The involvement of children in designing and developing their own games allows children to feel as they are truly inventors, so that they feel enjoyed, motivated and more willing to learn. Creating computer games is fun, creative, and empowering development; on the other hand it can be a challenge and complexity for children. Developing computer games involves programming language, but research showed that most children do not understand what computer programming is and how programs are produced (Greening 1998; Smith, Cypher & Schmucker 1999).

During the past 30 years, many researchers have attempted to build a number of environment systems with the intention of making programming accessible to young children, for example, Stagecast Icicle (Sheehan 2002) and PatternProgrammer (Wright, 2006). However, these environments use rules to allow their users easily building things that are tailored to their own needs such as games, simulations and educational software. Therefore, the fact of how well users can translate knowledge from these systems to a standard programming language is ignored. In addition, even though those rule based programming environments are easy to use and learn, once a learner has mastered the environment there is nowhere to progress to the next advance level. In contrast, object oriented systems allow their students to learn further and to build their own specific-built software and tools.

This research intends to explore and investigate the capabilities of rule based and objected oriented systems, in a special education setting with the express purpose of enabling and empowering students to develop their own learning programs within a game environment. By combining rule based and object oriented systems, we have created a generalised visualisation and interaction computer based game making tool, called GBuilder, in which students can extend and transpose the knowledge from the system to a standard programming constructs. The successful of this project will enhance the learning environments of young people and consequently increase their potential for successful and fulfilling learning experiences.

2. The GBuilder Environment

GBuilder is an object oriented event based environment, which combines the capabilities of rule based and object oriented systems with the express purpose of enabling and empowering students to develop their own learning programs within a game environment. The environment is therefore fun, enjoying, and engaging to use.

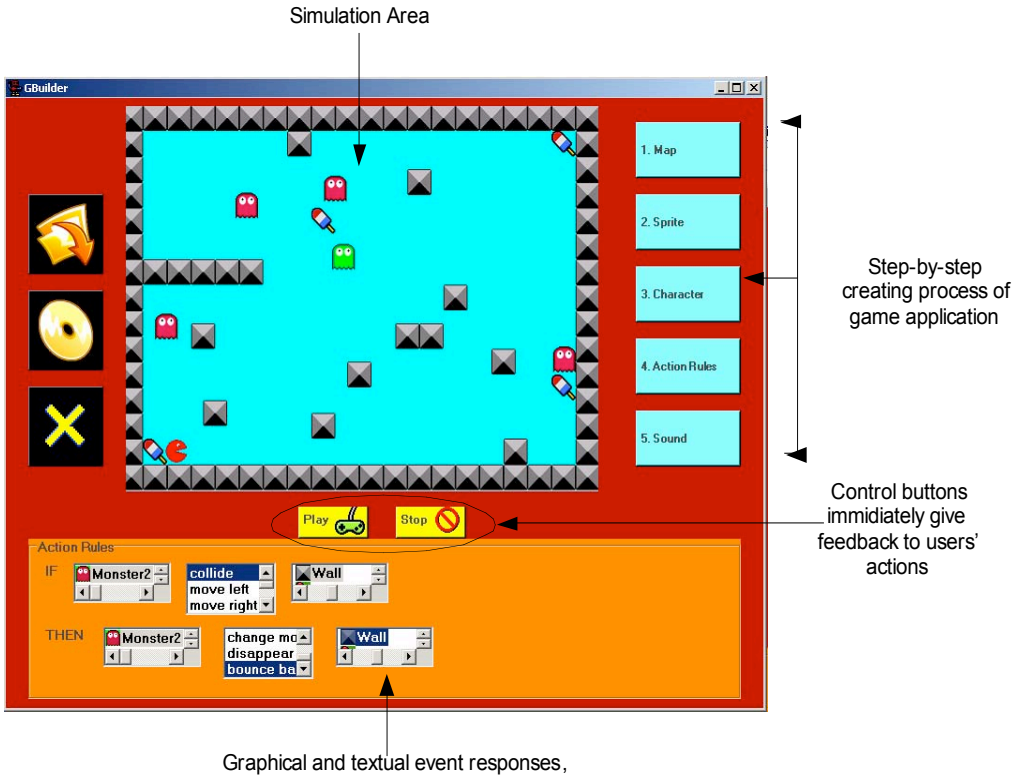


Figure 1: Screen snapshot of GBuilder

In GBuilder, the industry-standard programming concepts including classes and objects, event handling, decision making, data types, and encapsulation are used. This is to allow the learners to transfer learning experiences for the programming paradigm from the system to a standard programming language and to exploit the availability of programming textbooks and other teaching material.

GBuilder provides particularly suitable design and implementation of flexible visualization and interaction frameworks to its users. It is easy to use as children can use the drag-and-drop and click-and-scroll to explore and manipulate the environment. The system supports interactivity by providing the 'simulation area' to show immediately the results of connection between the commands given to the computer and its behavior. Graphical output is programmed implicitly, provided by the environment.

Creating game application in GBuilder involves 5 steps: draw a map, create sprites (classes), create characters (objects), write actions (event based responses), and insert sound. As these steps are in numeric order, they are simple and straightforward for a student to follow to produce interesting game examples.

Text and graphics are used to create action rules. Whilst textual event responses are used for the acquisition of requisite survival literacy skills, graphics are used for making programming more like thinking. That is children idea are analogous to graphical objects and they can directly interact with those objects in the process of programming. Both event based actions and graphical output are visible to children for the purpose of investigation and experimentation.

The framework is written in Visual Basis.NET language for the following reasons. Firstly, the .NET framework provides a common language runtime (CLR) which includes compilers for over 20 programming languages. Therefore, a system developer can choose their own privilege languages for writing program. This benefits to migration of Gbuilder to other development environments. The .NET applications can only run on windows-platform based workstations, however, this fits to GBuilder requirements as most primary schools use computer on a window platforms.

3. Conclusion

GBuilder is a visualisation and interaction framework which exploits a children's enthusiasm for games for enhancing their programming learning. With an effective design, the framework allows young children easily building things that are tailored to their own needs and support their imaginative play. GBuilder is an explicit model providing users visible source code and what is happening. Children gain some control and understanding of programming environment through event responses. These event responses comprise of graphics and texts, which are a better way for children to understand program and what is happening.

As the study is still in the initial experimentation, more work has to be done. For example, a defining process of action rules and a creation of event responses are still quite naïve. However good feedbacks and responses during prototype testing are an encouraging sign of the possibilities a fully functional programming environment for children might finally engender.

References

Greening, T. (1998). Building constructivist toolbox: An exploration of cognitive technologies. *Educational Technology, 38,* 23–35.

Maria Roussou (2004). Learning by doing and learning through play: an exploration of interactivity in virtual environments for children, Computers in Entertainment (CIE), 2,(1).

Prensky, M. (2001). Digital game-based learning, New York : McGraw-Hill.

Sheehan, R. (2005). The Icicle Programming Environment for Children, PhD thesis, University of Auckland.

Smith, D. C., Cypher, A., & Schmucker, K. (1996). Making Programming Easier for Children. Interactions, September/October 1996, 59-67.

Wright, T. (2006). PatternProgrammer: Yet Another Rule-Based Programming Environment for Children. In Proc. Seventh Australasian User Interface Conference (AUIC2006), Hobart, Australia.

Celebrity Death Match
Formal Methods vs. User-Centred Design

Judy Bowen

Department of Computer Science
University of Waikato
New Zealand
jab34@cs.waikato.ac.nz

Abstract. Formal methods and user-centred design are often seen as opposite ends of the spectrum of software design methodologies. On one hand we have a mathematical approach where everything is described in a precise formal language and rigorous proofs are carried out before refinement processes lead us to code. On the other hand we have designers who talk to users, build models out of paper and post-it notes, draw pictures and concern themselves with notions of usability and aesthetics. While these approaches may appear very different, in fact the rationale behind their use is the same: both aim to produce correct, working, usable software. The basis of my research is to examine formal methods processes and informal user-centred design methods and find ways of integrating them so that we are able to enjoy the benefits of both.

1 Introduction

The use of formal methods such as specification languages like Z [1], B [2], CSP [5], verification of such specifications, refinement calculi, model checking *etc.* benefits the software development cycle by helping to ensure that the software we build is reliable and correct. The process of creating a specification or building a formal model ensures that we think carefully about the requirements of the system before we start trying to code it, and we can subsequently use such models to prove properties and behaviours about the systems and satisfy ourselves that what we propose to build is correct and will do the right thing in all circumstances.

Similarly, user-centred design methods aim to ensure that all of the users' needs and requirements are properly considered before we build any system and by using methods such as task analysis, prototyping and usability testing we can also ensure that the users will be able to correctly interact with the system and it will behave as they expect.

While the methods used by these two approaches are very different, it is clear that many of their aims are the same, primarily to build better software by design. Because the approaches seem so different it is often the case that they are considered as opposing methods of design - either you can take the

formal approach and use specifications and proofs, or you can take the human-centred approach. This polarised view is further compounded by the fact that formal methods generally focus on functionality of systems, *i.e.* what they can do and how they do it, whereas human-centred design methods focus more on the interfaces to the system and how people will interact with them. The reality is, of course, that we want the systems we build to be reliable functionally as well as having visually appealing interfaces which are also correct and usable.

Rather than designing systems as if they can be clearly separated into two independent parts: the underlying functionality and the interface, I am interested in looking at fully integrated design methods which treat the system as a whole and give the same importance to all parts of it throughout the design process. This 'holistic' approach to system design and implementation will still allow for modularity, whereby we can consider different parts of the system in different ways (using different design methods), but underpinning it will be clear links between the formal and informal processes. The aim is to create a rigorous process which gives us all of the benefits of both user-centred design methods and formal methods and which ensure that at all stages of the design process everyone is working towards the same provably correct implementation.

2 Do You Speak My Language?

Integration of software engineering methods, both formal and informal, is not a new idea. However, one of the common problems with this practice is that typically it either results in a brand new method which practitioners of both original methods are expected to learn and use, or it 'absorbs' one method into another, which again requires one group to give up their preferred working methods in order to to use the new improved version.

Rather than trying to persuade different groups of designers working on different parts of the system design to adopt the languages and conventions of the other, we are more interested in allowing both groups to continue using their own processes and instead build bridges between the two sets of processes. In this way we retain the benefits that different methods offer at different stages of the design process.

The aim then is to develop ways of linking both the formal and the informal at all stages of the design process. In this way information gained during, *e.g.*, prototyping of interface designs can be fed back into the formal refinement process and *vice versa*. Formal practitioners and interface developers can continue to use their preferred, tried and tested method, but we will introduce a well-defined communication channel between the different processes which ensures everyone is working toward the same goal.

An additional benefit is that it allows us to overcome some of the weaknesses of a particular design method which may be apparent if we try to use just one approach for the design of the entire system.

3 Early Stages of the Research

The first stage of the research was to look at different ways of modelling graphical user interfaces (GUIs) and consider how such models may be incorporated into a formal process. We have already shown in previous work [3] how we can include the GUI of a system in a formal specification using the formal language Z [1] and some of the benefits this can produce. The intention of our current research, however, is not to try and persuade interface designers to start using Z as their main design tool, but rather to find ways of working which encompass their existing methods.

One of the problems we have identified is the way we consider implementations and refinements in formal processes. Such refinement processes aim to move from an abstract formal description of a system to an actual working implementation whose description is given in the code that creates it. Along the way we use refinement processes to move in steps from the abstract to the concrete. When we consider informal design artefacts, such as paper-based prototypes for example, it is not obvious how we can clearly capture the intent of such artefacts. We cannot have code as our refinement target as a paper-based prototype is not described in any programming language. At the same time, however, it is important to be able to incorporate such designs into a formal process to ensure that they correctly capture the requirements and constraints of the system given in the formal specification. In this way we can prevent errors being introduced at this early stage in the design process.

We began by developing a model of the GUI, which we call the *presentation* model. This relies on identifying the components, or *widgets* of a GUI or GUI design and using this as the basis of any description of that GUI. The presentation model does not rely on an implementation of a GUI or a computer-based prototype but rather it captures the visual elements of a design even if it is a low-fidelity artefact such as a paper prototype. So far we have been able to use this presentation model to show that a prototype is a correct implementation of a particular specification using traditional data-refinement methods [4].

Currently we are focussing on refinement between different designs. So, rather than trying to relate designs in an iterative process back to some formal description each time we change the design, instead we are looking at how we can ensure the refinement relation is preserved by describing what it means for one informal design to correctly refine another. Using the presentation model again we have started by identifying different types of equality between designs, such as isomorphism, functional equality, component equality, as the first step in identifying a refinement relation. This also has interesting applications in the research area of plasticity of GUIs. As ubiquitous computing becomes more and more pervasive, designing software applications that will run on multiple platforms in many different contexts of use is increasingly important. We hope to show that our work on refinement between designs is useful in this area where we want to ensure that several different versions of the same application are equally correct.

4 Conclusions

While we are still only in the early stages of this research, the work we have done so far suggests that this is an interesting and worthwhile area to pursue which has a number of different possible applications. Ultimately we hope that we can show that not only do formal methods and user-centred design methods complement each other, but it is possible to use both in a fully-integrated design process which is both rigorous and creative.

References

1. ISO/IEC 13568. *Information Technology—Z Formal Specification Notation—Syntax, Type System and Semantics*. Prentice-Hall International series in computer science. ISO/IEC, first edition, 2002.
2. J.-R. Abrial. *The B-book: assigning programs to meanings*. Cambridge University Press, New York, NY, USA, 1996.
3. Judy Bowen. Formal specification of user interface design guidelines. Masters thesis, Computer Science Department, University of Waikato, 2005.
4. Judy Bowen and Steve Reeves. Formal refinement of informal gui design artefacts. In *To appear: ASWEC 2006: 17th Australian Software Engineering Conference*. IEEE Computer Society Press, 2006.
5. C. A. R. Hoare. Communicating sequential processes. *Commun. ACM*, 26(1):100–106, 1983.

"BECOMING INDIGENOUS: an impossible necessity"

WHAKAPAPA

The first Elizabeth Bryce in New Zealand was born in the Wanganui barracks in 1862, during Titokowaru's war with colonial settlers and disputes over confiscated land in Taranaki.
That Elizabeth Bryce was not only my great grandmother but also my Great Aunt Lizzie. She became Elizabeth McKenzie by marriage.
Her younger brother Jim married Agnes Gow and settled in Cambridge (around 1910-11) where I lived as a child.

Jim Bryce was my paternal grandfather because my mother, Elizabeth (Betty) McKenzie – married his son Maurice Bryce and became Elizabeth (Betty) Bryce!!

Elizabeth Bryce-McKenzie is interred in the Rotorua cemetery with my maternal grandparents and my parents,
My Bryce grandparents are there in Cambridge, in the Hautapu cemetery.

Bryces, Campbells and McKenzies who had all come to New Zealand from Scotland in the 1840s, settled first in Wellington, then Wanganui. The Gows, he a Free Church of Scotland minister, arrived in Dunedin after 1840.

****

Figure Number range Figure Boat People, art emergency BOG. 2005

The MFA at Otago School of Art (incidentally the first Art School in NZ) is a 2-year research degree. Candidates choose a 25/75% or 50/50% split between practice and written component. The written and theory component is rigorous, as it has been set up

to meet international standards of research.
The 'theory ' is critical thinking and candidates read extensively in contemporary thinkers who are relevant to their practice. Many candidates' work is multidisciplinary – or does not 'fit' into a traditional compartmentalisation of say sculpture, painting or photography. This is in keeping with contemporary art practice.

MFA ABSTRACT
My Masters project, is entitled:
"BECOMING INDIGENOUS: an impossible necessity"
or
"*Becoming* Indigenous – an impossible desire"
from a quote by Terry Goldie *Fear and Temptation*: The Image of the Indigene in Canadian, Australian, and New Zealand Literatures.[ii]

The project considers the concepts inherited by Pakeha through the desires of their ancestors who were New Zealand's Colonisers. It speculates on the possibilities of this apparent absurdity (to *become* indigenous).
It suggests that 'becoming' *any* other is a slow evolution that requires not only desire but an unconditional commitment, and, as Braidotti[iii] with reference to Irigaray, Deleuze and Guattari suggests a female 'becoming, needs to reference a point not based on the already referenced 'original'.
In the feminist view that is a patriarchal model. From the point of view of those colonised - and those who are post-colonial subjects – it is also a patriarchal model. The Coloniser" is '*the original'*. This project is from the position (uncomfortable though it may be) of a descendant of colonisers – an ethnic 'Pakeha' named ("othered") by the very group they colonised. It is a project that because of the deconstruction process of post colonialism is at times difficult, contradictory, confrontational and ambiguous.

Description of art practice

The practical component of my practice uses sculpture in the public arena – not commissioned public art but art in public places – art in 'non art' places –on these occasions the act of putting the work there could be likened to an unsolicited performance. An interference or intervention in a public place.
This does not exclude gallery spaces, artist gatherings, or studio presentations as these are chosen because they are the 'fringe' of the art world or public gathering type event. Blue Oyster Gallery in Dunedin is an experimental art space for 'new' work only, Terminus05[iv] was a gathering (symposium) of international artists working in public spaces not set aside as 'art' places. The first 'outing' of e.i.kapai[v] was a rather ambiguously advertised performance in a shop window during the Dunedin 2002 Fringe Festival.[vi]

My practice has more recently used multiple wax casts of small cupie dolls – nostalgia of the 1950s –70's, this refers to British origins – "Kleeware England' is just noticeable on some of the original dolls. They became synonymous with the N.Z. AMP shows – the

fairy doll that you could win at the side-show (pink). Later in the 60s and 70s some were *painted* brown, to be the baby'on the back of a Maori souvenir doll.

Later still, when dressed in 'cloak' and bits of raffia they were 'Maori dolls', a boy and a girl, given as mementos on your trip via Air New Zealand to Fiji – in the '70s.
It was a cheap, plastic doll to represent many ethnicities.

Figure Number range Figure. cupie and 'souvenir' dolls

I call the latter type of souvenirs 'Coloni ana';[vii]they are not 'Maori' souvenirs.
The original dolls were 'actors' in their commercial script. Dressed with nylon and glitter and on a stick in the country Shows, then dressed in 'cloak' and bits of raffia to be Maori dolls and again wrapped with a piece of jagged plastic to be indigenous Fijian.
In *my* scripts they are made of wax in their own *solid* color ready to act out my strategies.
In Terminus 05 a few hundred waxes occupied places officially and unofficially in the manner of 1860's arrivals from Britain.
In the 'Island (S)hopping'[viii] group show at the Blue Oyster Gallery they try ways of 'becoming' by becoming immersed in land[ix] or being sold as a commodity.


Figure Number range Figure &4 'Island (s)hopping' B.O.G. Dunedin Nov 2005

.Usually the work continues to develop or continue its life after its public appearance. Digital media can capture and suggest work that has happened in other places or is unseen.
Acts of 'disobedience,' the people and stories that are not documented in the 'official' story of a place.


Figure Number range Figure mutilated babies from 'Terminus05 Dunedin', & 5 'becoming sugar' after 'Island (S)Hopping.

Work that exists only as a virtual reality – or maybe hasn't happened yet – or perhaps occurred in the past.
These are concepts easily accepted by young people but are radical thought for those who grew up when space and time media were things of fantasy.

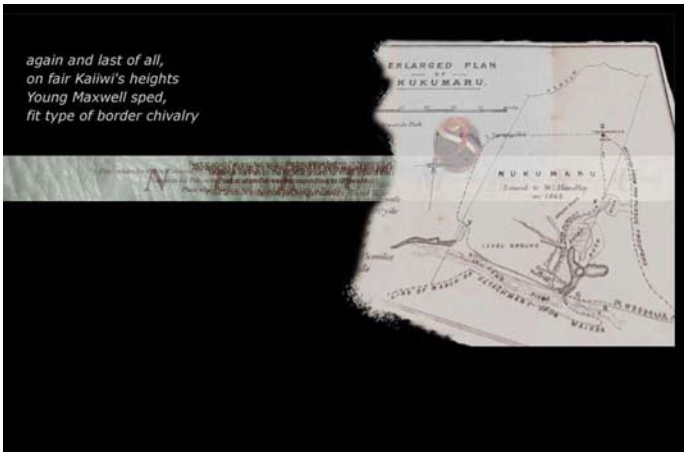


Figure Number range Figure rollovers on "Handleys Woolshed" proposal

Contemporary artists notoriously use tools from other disciplines.
For example Ralph Hotere's use of a metal grinder, roofing nails, corrugated iron has now become accepted materials for artists. Digital media and the Internet are media that are used in art now although not necessarily as their inventors had intended.
For example Angela Main and Caroline McCaw presented "Animalia" at the 'Converge' arts fellowship with Hitlab (Human Interaction Technology) as Caroline said they 'added content, concept and context to a demo of software.'

I am hoping future work can use Upstage – to gauge public reaction or unsolicited public opinion about some of my work.
Intervention and 'disobedience' is a concept in contemporary art that software like upstage (and HITlab) allows – you never know who is who, what is real, whether the actions are genuine or just anarchist. There is no longer a debate but a practice around 'interventionist art'.[x] For example Critical Art Ensemble, Sub Rosa, and in Wellington a group of artists had a collective 'project' – Interventions – in the 2004 Wellington Arts festival[xi]

HOW I BECAME A GEEK AND DISLOCATED MY SHOULDER.
During a Post Graduate Dip Design Studies paper we were to write an essay to be read on the web. The intention I think was to create a long scrolling document with links down the page.
I'd researched reading on the web, which seemed not only difficult but also boring. I returned to talk to *Susan* Ballard in the Theory department at Art School - her interests and specialisation included art and technology, cyborgs, automata and writing practices. She suggested that the computer system could be considered *female*. We discussed origins

of the web -as a database needed and used by librarians and the recording and patterning of weaving.
This reminded me of recording *taniko* patterns as a child, and being the only one who owned graph paper in my class because I swapped patterns with my cousin.
At the same time another sculptor sent me a URL of a delightful piece of writing.
. http://ensemble.va.com.au/water/dialogue.html

The small recollection of *taniko* weaving and this writing became the starting point of a way to use my art images. It suited my way of thinking and writing – adding asides and going off at a tangent. More importantly for that essay it could reflect the content and concepts of the web's beginnings with Ada Lovelace [xii]who apparently had conflicting art and science talents.
The web is not linear but rhizome[xiii] like.

In my MFA project, ideas of becoming, desire and invisibility can be better suggested by digital work using photography and time to suggest space distance and the 'unseen'.

I am intending using the hypertext writing as a way of 'publishing' my dissertation – with interspersed images and digital pieces that will work along side but not completely illustrate the text.
I have found in the past that problems with 'limited access' information can be partially solved with digital media.
For example–for cultural security a CD can be issued in 'limited edition' to family members only – or like a thesis, with limited access.

Excerpts from the appended text will be used to illustrate how I intend combining images and small digital clips for my dissertation. This can be 'searched through' depending on time available.

The Net is fundamentally and profoundly anti-spatial...
You cannot tell a stranger how to get there ...
but you can find things in it without knowing
where they are.'

(Michael Joyce)

ref: http://ensemble.va.com.au/water/dialogue.html

[i] J.E. Traue *Ancestors of the mind: a Pakeha Whakapapa.* J.E. Traue Chief Librarian Alexander Turnbull Library Wellington 1973-1990 describing the European tradition of identifying oneself (male) according to job description and learning

[ii] Terry Goldie in *Fear and Temptation*: The Image of the Indigene in Canadian, Australian, and New Zealand Literatures. "And the first felt need for indigenisation came when a person moved to a new place and recognised an Other as having greater roots in that place……. a process which I have termed 'indigenisation'. A peculiar word, it suggests the impossible necessity of becoming indigenous."

[iii] Rosi Braidotti, "Metamorphoses"

[iv] Terminus05 a sculpture symposium in October 2005 curated by Ali Bramwell (who at the moment is presenting a report *on* Terminus05 at another symposium in Bosnia) involving 24 international artists responding to two sites (Haulashore Island in Nelson and Customhouse Quay in Dunedin.) e.i.kapai was a participant and Eliz Bryce made the website for the promotion and duration of the symposium.

[v] e.i.kapai is the artist name, registered company and domain name for Elizabeth Bryce working in a particular (usually 'colonising') way.

[vi] "Tea with Hinemoa", performance, CD 'Children Songs of Maoriland' and "Coloni ana Collection" in Outre shop window by e.i.kapai. Dunedin fringe Festival 2002.

[vii] "Coloni ana Collection" in Outré shop window by e.i.kapai. Dunedin fringe Festival 2002.

[viii] "Island (s)hopping" B.O.G. November 2005. Group show on ethnic diversity curated by Di Halstead.

[ix] An allusion to John Pym a U.K. artist who in *Rural Myths* "buries" then photographs home owners in front of their houses (or castles) at a depth according to their apparent commitment to their home (land?).

[x] The Interventionists: User's manual for the creative disruption of everyday life. Mass MoCA

[xi] Art New Zealand online http://www.art-newzealand.com/Issue111/exhibitions111wn.htm

[xii] *Dr. Betty Toole* ***Ada Byron, Lady Lovelace (1815-1852)***.
http://www.cs.yale.edu/homes/tap/Files/ada-bio.html

[xiii] Deleuze and Guattari, A Thousand Plateaus

Privacy Issues in Health Care and Security of Statistical Databases *

Tatiana King
Discipline of Computer Science
University of Newcastle, Australia
Tatiana.King@newcastle.edu.au

ABSTRACT

Information is the 'currency' of modern health care. Patient data is becoming extremely important for management and planning in health care. However, with increasing dependency on information, together with the positive aspects of this trend, new challenges arise. The problem of individual rights and privacy, as well as the protection of the sensitive information, becomes very important. This problem is not less important in the context of statistical databases. In this paper we give the overview of the problem. We then present our work in progress on developing a practical statistical security control technique, tailored specifically towards health record databases. Special attention is paid to a social study surveying the public opinion about privacy in health care, which is an intrinsic part of our project and which is expected to reveal what kind of privacy is needed.

1. THE BACKGROUND OF THE PROBLEM

Information is the 'currency' of modern health care [22]. The amount of data stored in databases and its accessibility is increasing as a result of growing computer network applications. It is happening both within health care institutions and nationwide. Patient data are becoming extremely important for management and planning in health care. To a greater extent it is used to support the assessment of care quality as well [29]. At all levels of health care, decision makers have an interest in patient data or aggregation of it. However, at present data are often scattered across different sites, such as different hospitals or organisations, or different departments within the same organisation. In such case data may not be readily available to all interested parties. Consequently, there is lot of interest in the electronic personal health record (PHR) system, as it is expected to significantly improve the accessibility of health information [28, 8]. The PHR is a single, person-centred system, designed to track health of an individual and support health care activities across one's entire life experience. The last decade has brought unprecedented levels of investment in the development of new initiatives and the implementation of PHR. A number of programs have been established in order to support the implementation, including Connecting for Health in USA, Canada Health Infostructure Partnership Program, Information for Health 1998-2005: an Information Strategy for the Modern NHS in UK and National Health Information Agreement (NHIA) in Australia. Among the key challenges in this area is creating a transparent mechanism for governing a PHR system, so as to address public concerns about ownership, security, and privacy. The strategy for overcoming this challenge will have a profound effect on innovation, regulation and implementation of information technology in health care.

In other words, with increasing dependency on information, together with the positive aspects of this trend, new challenges arise. The problem of individual rights and privacy, as well as the protection of the sensitive information, is becoming very important. The patients' concern about the potential risks associated with the automation and sharing of their medical information, is growing [10, 16]. This problem is not less important in the context of statistical databases, even though at first it might appear that personal information is protected by making it anonymous or aggregated. Nevertheless, by correlating different statistics, a 'clever' user [13] may be able to deduce confidential information about some individuals or organisations.

It is essential to understand that we have different aspects of privacy and security in Health Informatics. Firstly, the protection of privacy is addressed by law, rules and regulations. Secondly, the laws and regulations can be enforced considerably by technical means. However, it is important to note that currently there is only a very limited set of techniques and regulations addressing specifically the statistical security. Finally, we need to recognise that without addressing human aspect of privacy it is impossible to provide appropriate techniques for its effective protection.

2. SECURITY OF STATISTICAL DATABASES

A statistical database (SDB) differs from an ordinary database primarily in its very limited querying interface. Data in a statistical database are usually organised according to a relational scheme. SDB is a database that returns statistical values such as averages and sums, derived from the records, rather than the values of the attributes in the records them-

*This research was supported by ARC Discovery Project grant DP0452182.

selves. Some attributes are very often confidential and so should be protected from disclosure. Hence, a statistical user who does not have access rights to individual records should be limited to statistical queries based on more than one individual. However, it is possible to carefully construct a series of queries and correlate different statistics to reveal particular values contained within the database. If that happens, then we refer to the database as *compromised*.

Threats to security may result in exact or partial disclosure [1]. Exact disclosure occurs when from the answer to one or more queries a user is able to derive the exact value of a confidential attribute. Partial disclosure refers to the situation in which some inference about a confidential attribute can be made, even if the exact value cannot be determined. Different types of compromises are considered and the classification of them is given [26, 3, 23].

A constant conflict of interests exists whenever a statistical database is used. First of all, it is expected that the collected data are of the high quality. It is essential in order to keep updating the national policy and enable statistical research. Furthermore, all statistics are in the public domain. However, the contents of individual records are not public. The sensitive information of an individual must stay confidential. Even though access control methods provide some protection for a single record, it is still easy to infer the contents of specific records from statistical data. Therefore, in absence of security control mechanisms a SDB can not be secure.

Any security control mechanism should meet two conflicting requirements: security and richness of provided information [13]. In other words, on one hand, it has to satisfy the user's request of aggregate statistics related to non-confidential and confidential attributes. On the other hand, it should be secure enough to withstand a user's ability to infer any confidential information related to a specific individual represented in the database.

Different techniques for the prevention of a statistical database compromise are suggested. By and large all of them can be divided into two groups: query restriction mechanisms and noise addition techniques [1]. Security methods which are based on the query restriction approach reject a query that could lead to a compromise. Nevertheless, the answer to the query is always exact. Such methods provide protection to a SDB through one of the following procedures: restricting the query set size to satisfy some conditions defined by the DBA [15]; controlling the overlap among queries to restrict the number of overlapping elements in the successive queries [14]; cell suppression to eliminate from the released statistics all cells that may cause disclosure of confidential information [11]; partitioning in order to group individual elements of a statistical database in a number of mutually exclusive subsets and release macro statistics that contain either an entire subset or none of the records from it [9]; and auditing that involves keeping up-to-date logs of all queries made by each user and checking them against the previous ones in order to monitor for possible compromise [24, 3]. The problem with most query restriction techniques is that they are overly restrictive which greatly affects the "usability" of the database. We define the *usability* as the ratio of the answerable queries to the total number of queries. A lot of work has already been done on establishing the theoretical bounds on usability in different scenarios, that is, the maximum percentage of queries that can be answered by any control mechanism without causing the database compromise. For example, in the database of n records, the maximum usability U for general SUM queries is of the order $\Theta(n_1^{-\frac{1}{2}})$ [5], and for k-compromise the order is $O(n^{-1-\frac{k}{2}})$ [18]. The usability for range queries have been studied in [19, 4, 21], and for range queries and k-compromise in [6, 7, 25]. In the above, k-compromise refers to the disclosure not of a single confidential value but rather of a statistic based on k or less values. An intruder typically has some knowledge about confidential values in the database and can potentially learn more from such statistic.

Noise addition techniques, which are used for prevention of data inference, involve changing the original data of a SDB [13] or the result of a statistical query [2, 12] by introducing some noise to it. As a result of implementing this approach a user of a SDB can be provided with answers to all queries but the answers to them would be only approximate. Most results refer to numerical data; however there are some recent techniques developed specifically for categorical data [17, 27, 20]. The main drawback of noise addition techniques is that they unfortunately adversely affect the data quality. Measuring data quality is a challenging task, and different approaches have been studied by various authors [1, 30].

3. TOWARDS A PRACTICAL SOLUTION TO STATISTICAL SECURITY PROBLEM IN HEALTH CARE

The aim of our research is to offer a practical solution for protection of sensitive information suitable for health records used in statistical analysis. In order to achieve this it is essential to discover the public attitude towards privacy in health care, to gain insight into what kind of privacy is needed. If it is necessary to ensure privacy of all individuals, the techniques will be different compared to the techniques which can be developed to protect the privacy of some groups of people; the techniques will also differ accordingly to the size and the structure of groups that require higher levels of privacy. The same is true in regard to the amount of personal information that is deemed confidential. In other words, it is presumed that not all personal information is equally sensitive and the choice of appropriate techniques will depend on the amount of information that requires a high level of protection.

Therefore, a multi-strategy research is designed. It will encompass two stages and combine qualitative and quantitative research. The Stage One of the study comprises arranging and facilitating several focus groups. The Stage Two is a social survey. We believe that the attitude of individuals towards privacy is highly dependent on their gender, age, level of education, region, ethnic background and so on. Therefore for the Stage One of the project (the Focus Group) participants will be selected to cover a range of these criteria: age, gender, different background, level of education and health condition. It will allow for examination of the ways in which people in conjunction with one another see the issues of privacy. The result of the discussion will be analysed

and used to identify the most important characteristics that influence privacy attitude. Qualitative data will provide the privacy perspectives of the people under the study and assist in constructing the questionnaire for a social survey. For the Stage Two of the project (the social survey) the target population is the adult population nation-wide. The participants of the survey will be chosen randomly to ensure that the results of the survey represent the entire population of Australia. The main aim is to determine the patterns in the data. Our intention is to extract the sub-populations of the collected data into individual classes on the basis of some chosen attributes and retrieve the classification rules for it. Another aim of the survey is to define the size of different partitions of statistical database in health care accordingly to the required (by patients) level of confidentiality. As the result we need to identify the structure of the partition with the highest expectation for privacy of their personal information; the group of people who do not necessarily have high requirement for privacy of all their information but who would require keeping confidential the sensitive health information about themselves; and the possible group which would gather people with no or very little concern about privacy of their personal information. In order to achieve it we shall use a classification data mining technique such as Decision Tree Induction. Additionally, descriptive methods of analysis will be used. Hence the quantitative data will be used in statistical analysis to extract the patterns of interest and answer our research questions and will be a base for constructing the model of privacy in a statistical database.

This is a part of a larger project that aims to develop a comprehensive set of techniques for protection of confidential individual health data used for statistical purposes. It comprises the conceptual framework for the protection system. We are going to partition a database horizontally and vertically accordingly to the level of required privacy. The horizontal partitioning will be defined by the patient's attitude to privacy, while the vertical one will depend on the sensitivity of the attribute. The number of partitions in the database will depend on the number of security levels of the system, which will be determined based on the outcome of the survey and existing privacy laws. In practice a database will be partitioned accordingly to the required level of security, which can be captured from the consent of a patient. The new method will be a combination of query restriction and noise addition techniques which will be applied simultaneously. Different levels of security will be provided for different users. It is considered so far that in different partitions of the database different query restriction methods will be used simultaneously with a noise addition technique. It will depend on the required level of security for the particular partition. For example, for the partition of the database, which consists of the records with the highest security requirement, we can apply the protection by k-compromise free method. Another option is to release macro statistics about atomic populations only. The choice will depend very much on practical results of the survey and further investigation of the existing protection techniques.

4. CONCLUSION

In conclusion, we note that the aspiration of all PHR initiatives and programs is to enable the electronic interchange of medical information, which promises to improve many aspects of medical care through better management and planning. At the same time, individuals must be reassured that their privacy will be protected at any stage, regardless of how their information will be used. Hence, the security and privacy issues are important considerations in the future implementation of these initiatives. Generally speaking, noise addition techniques can provide a high level of released statistics, but they suffer from approximate compromise and low quality of released statistics. On the other hand, query restriction mechanisms provide exact statistics, but could be overly restrictive, leading to a low usability of statistical databases. Our plan is to combine the existing query restriction methods with noise addition techniques in order to achieve the practical result.

5. REFERENCES

[1] N. R. Adam and J. C. Wortmann. Security-control methods for statistical databases: A comparative study. *ACM Comp. Surveys*, 21(4):515–556, 1989.

[2] L. L. Beck. A security mechanism for statistical databases. *ACM Trans. Database Sys.*, 5(3):316–338, 1980.

[3] L. Brankovic. *Usability of Secure Statistical Databases*. PhD thesis, The University of Newcastle, Australia, 1997.

[4] L. Brankovic, P. Horak, and M. Miller. An optimization problem in statistical database security. *SIAM Journal on Discrete Mathematics*, 13(3):346–353, 2000.

[5] L. Brankovic and M. Miller. An application of combinatorics to the security of statistical databases. *Australian Mathematical Society GAZETTE*, 22(4):173–177, 1995.

[6] L. Brankovic, M. Miller, and J. Siran. Range query usability of statistical databases. *Int. J. Comp. Math.*, 79(12):1265–1271, 2002.

[7] L. Brankovic and J. Siran. 2-compromise usability in 1-dimensional statistical databases. In *Proc. 8th Int. Computing and Combinatorics Conference, COCOON2002*.

[8] I. R. Cheong. Privacy and security of personal health information. *Journal of Informatics in Primary Care*, March:15–17, 1996.

[9] F. Y. Chin and G. Özsoyoğlu. Security in partitioned dynamic statistical databases. pages 594–601, 1979.

[10] R. Clarke. A history of privacy in australia. *Xamax Consultancy Pty Ltd*, (http://www.anu.edu.au/people/Roder.Clarke/DV/OzHistory.html), Accessed Feb 2004.

[11] L. H. Cox. Suppression methodology and statistical disclosure control. *J. Am. Stat. Assoc.*, 75(370):377–385, 1980.

[12] T. Dalenius. A simple procedure for controlled rounding. *Statistik Tidskrift*, 3:202–208, 1981.

[13] D. E. Denning. *Cryptography and Data Security*. Addison-Wesley, Reading, Mass., 1982.

[14] D. Dobkin, A. K. Jones, and R. J. Lipton. Secure databases: Protection against user influence. *ACM Trans. Database Sys.*, 4(1):97–106, 1979.

[15] I. P. Fellegi. On the question of statistical confidentiality. *J. Am. Stat. Assoc.*, 67(337):7–18, 1972.

[16] H. Giggins and L. Brankovic. Ethical and privacy issues in genetic databases. 2002.

[17] H. Giggins and L. Brankovic. Protection privacy in genetic databases. In *5th International Congress on Industrial and Applied Mathematics*, pages 73–78, 2003.

[18] J. R. Griggs. Database security and the distribution of subset sums in r^m. *Graph Theory and Combinatorial Biology, (Balatonlelle, 1996)*, pages 223–252, Bolyai Soc. Math. Stud, 7, 1999.

[19] P. Horak, L. Brankovic, and M. Miller. A combinatorial problem in database security. *Discrete Applied Mathematics*, 91(1-3):119–126, 1999.

[20] Z. Islam and L. Brankovic. Detective: A decision tree based categorical value clustering and perturbation technique in privacy preserving data mining. In *3th International IEEE conference on Industrial Informatics (INDIN2005)*, August 2005.

[21] S. Jajodia L. Wang, D. Wijesekera. Cardinality-based inference control in data cubes. *Journal of Computer Security*, 12(1):655–692, 2004.

[22] W. Lowrance. Privacy and health research. *A Report to the U.S. Secretary of Health and Human Services*, (http://www.aspe.hhs.gov/datacncl/PHR.htm), Accessed Feb 2005.

[23] J. Sebbery M. Miller. Relative compromise of statistical databases. *Australian Computer Jounal*, 21(2):56–61, 1989.

[24] J. Simpson M. Miller, J. Roberts. Prevention of relative compromise in statistical databases using audit expert. *Bulletin of Institute of Combinatorics and its Applications*, (10):51–62, 1994.

[25] L. Mathieson, T. King, and L. Brankovic. 2-compromise: The security of statistical databases with range queries. In *Proceedings of the 15th Australasian Workshop on Combinatorial Algorithms (AWOCA 2004)*, Ballina, Australia, 2004.

[26] M. Miller. A model of statistical database compromise incorporating supplementary knowledge. *Databases in 1990's*, pages 258–267, 1991.

[27] P. Kooiman P. Wolf, J. Gouweleeuw and L. Willenborg. Reflection on pram. In *Conference Programme of Statistical Data Protection '98*, March 1998.

[28] A. J. Siman. Protecting personal privacy: A first principle for canadian health healthcare information. *Management & Communications Canada*, XIII(4):12–13, 1999.

[29] J. H. van Bemmel and M. A. Musen (Eds). *Handbook of Medical Informatics*. AW Houten, Netherlands, 1997.

[30] P. Barnaghi Z. Islam and L. Brankovic. Measuring data quality: Predictive accuracy vs. similarity of decision trees. In *6th International Conference on Computer & Information Technology*, pages 457–462, 2003.

A Collaborative Constraint-based Intelligent System for Learning Object-Oriented Analysis and Design using UML

Nilufar Baghaei
Department of Computer Science and Software Engineering
University of Canterbury
n.baghaei@cosc.canterbury.ac.nz

ABSTRACT

Automatic analysis of interaction and support for group learning through a distance collaborative learning system is at the forefront of educational technology. Research shows that collaborative learning provides an environment to enrich the learning process by introducing interactive partners into an educational system and creating more realistic social contexts.

This paper presents COLLECT-*UML*, a constraint-based ITS that teaches object-oriented design using Unified Modelling Language (UML). UML is easily the most popular object-oriented modelling technology in current practice. Constraint-Based Modelling (CBM) has been used successfully in several tutoring systems, which have proven to be extremely effective in evaluations performed in real classrooms. We have developed a single-user version that supports students in learning UML class diagrams. The system was evaluated in a real classroom, and the results showed that students' performance increased significantly while interacting with the system. We are now extending the system to provide support for collaboration. An overview of both single-user and collaborative versions of the system is presented. A full evaluation study has been planned for April 2006, the goal of which is to evaluate the effect of using the system on students' learning and collaboration.

1. INTRODUCTION

E-learning is becoming an increasingly popular educational paradigm as more individuals who are working or are geographically isolated seek higher education. As such students do not meet face to face with their peers and teachers, the support for collaboration becomes extremely important [3]. There have been several definitions for collaborative learning. The broadest (but unsatisfactory) definition is that it is a *situation* in which *two* or *more* people *learn* or attempt to learn something *together* [4]. A more comprehensive definition states as follows: "… a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem".

Effective collaborative learning includes both learning to effectively collaborate, and collaborate effectively to learn, and therefore a collaborative system must be able to address collaboration issues as well as task-oriented issues [6].

In the last decade, many collaborative learning environments have been proposed and used with more or less success. Researchers have been exploring different approaches to analyse and support the collaborative learning interaction. However, the concept of supporting peer-to-peer interaction in Computer-Supported Collaborative Learning (CSCL) systems is still in its infancy, and more studies are needed to test the utility of these techniques. Some particular benefits of collaborative problem-solving include: encouraging students to verbalise their thinking; encouraging students to work together, ask questions, explain and justify their opinions; increasing students' responsibility for their own learning and increasing the possibility of students solving or examining problems in a variety of ways. These benefits, however, are only achieved by active and well-functioning learning teams [11].

This paper describes COLLECT-*UML,* an Intelligent Tutoring System (ITS) that uses Constraint-Based Modeling (CBM) approach to support both problem-solving and collaborative learning. The CBM approach is extremely efficient, and it overcomes many problems that other student modeling approaches suffer from. CBM has been used successfully in several tutors supporting individual learning [7]. We provide a brief overview of the single-user version which we have finished developing [1, 2] and describe extensions being made to this tutor, to support multiple students solving problems collaboratively.

2. RELATED WORK

This section provides examples of three types of CSCL systems, in the context of the collaboration management model [6, 9]:

- **Reflecting Actions:** The most basic level of support a system may offer involves making the students aware of the participants' actions. Actions taken on shared resources, or those that take place in private areas of a workspace may not be directly visible to the collaborators, yet they may significantly influence the collaboration. Raising awareness about such actions may help students maintain a representation of their teammates' activities. The system described in [8] is an example.
- **Monitoring the State of Interactions:** Systems that monitor the state of interaction fall into two categories: those that aggregate the interaction data into a set of high-level indicators, and display them to the participants, and those that internally compare the current state of interaction to a model of ideal interaction, but do not reveal this information to the users. In the former case, the learners are expected to manage the interaction themselves, having been given the appropriate information to do so. In the latter case, this information is either intended to be used later by a coaching agent, or analysed by researchers in an effort to understand and explain the interaction. EPSILON [11] and MArCo [12] are examples of such systems.
- **Offering Advice:** This will include the CSCL systems that analyse the state of collaboration using a model of interaction, and offer advice intended to increase the effectiveness of the learning process. The coach in an advising system plays a role similar to that of a teacher in a collaborative learning classroom. The systems can be distinguished by the nature of the information in their models, and whether they provide advice on strictly collaboration issues or both social and task-oriented issues. Examples include LeCS [10] and COLER [3].

Although many tutorials, textbooks and other resources on UML are available, we are not aware of any attempt at developing a CSCL environment for UML modelling. However, there has been an attempt [11] at developing a collaborative learning environment for OO design problems using Object

Modeling Technique (OMT) – a precursor of UML. The system monitors group members' communication patterns and problem solving actions in order to identify (using machine learning techniques) situations in which students effectively share new knowledge with their peers while solving OO design problems. The system dynamically assesses the group interactions and determines when and why the students are having trouble learning the new concepts they share with each other. The system does not evaluate the OMT diagrams and an instructor or intelligent coach's assistance is needed in mediating group knowledge sharing activities. In this regard, even though the system is effective as a collaboration tool, it would probably not be an effective teaching system for a group of novices with the same level of expertise, as it could be common for a group of students to agree on the same flawed argument.

3. COLLECT-*UML*: SINGLE-USER VERSION

COLLECT-*UML* is a problem-solving environment, in which students construct UML class diagrams that satisfy a given set of requirements. It assists students during problem-solving, and guides them towards a correct solution by providing feedback. The feedback is tailored towards each student depending on his/her knowledge. COLLECT-*UML* is designed as a complement to classroom teaching and when providing assistance, it assumes that the students are already familiar with the fundamentals of object-oriented design. For details on system's architecture, functionality and the interface refer to [1, 2]; here we present only the basic features of the system.

At the beginning of interaction, a student is required to enter his/her name, which is necessary in order to establish a session. The session manager requires the student modeller to retrieve the model for the student, if there is one, or to create a new model for a new student. Each action a student performs is sent to the session manager, as it has to link it to the appropriate session and store it in the student's log. Then, the action is sent to the pedagogical module. If the submitted action is a solution to the current problem, the student modeller diagnoses the solution, updates the student model, and sends the result of the diagnosis back to the pedagogical module, which generates appropriate feedback.

COLLECT-*UML* contains an ideal solution for each problem, which is compared to the student's solution according to the system's domain model, represented as a set of constraints. The system's domain model contains 133 constraints that describe the basic principles of the domain. In order to develop constraints, we studied material in textbooks, such as [5], and also used our own experience in teaching UML and OO analysis and design.

We conducted an evaluation study in May 2005 [2]. The study involved 38 volunteers enrolled in an introductory Software Engineering course at the University of Canterbury, which teaches UML modelling as outlined by Fowler [5]. The students learnt UML modelling concepts during two weeks of lectures and had some practice during two weeks of tutorials prior to the study.

The study was conducted in two streams of two-hour laboratory sessions. Each participant sat a pre-test, interacted with the system, and then sat a post-test and filled a user questionnaire. The pre-test and post-test each contained four multiple-choice questions, followed by a question where the students were asked to design a simple UML class diagram. The participants spent two hours interacting with the system, and solved half of the problems they attempted. The average mark on the post-test was significantly higher than the pre-test mark ($t = 2.71$, $p = 4.33E\text{-}08$). The students spent on average 90 minutes interacting with the system.

We also analyzed the log files, in order to identify how students learn the underlying domain concepts. Figure 1 illustrates the probability of violating a constraint plotted against the occasion number for which it was relevant, averaged over all constraints and all participants. The data points show a regular decrease, which is approximated by a power curve with a close fit of 0.93, thus showing that students do learn constraints over time. The probability of violating a constraint on the first occasion of application is halved by the tenth occasion, showing the effects of learning.

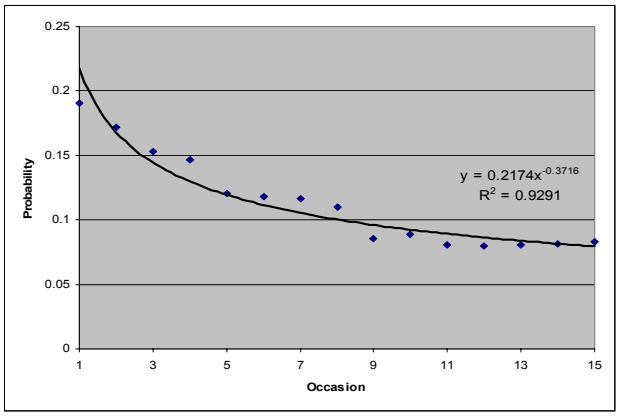


Figure 1. Probability of constraint violation

The results showed that COLLECT-*UML* is an effective learning environment [2]. The participants achieved significantly higher scores on the post-test, suggesting that they acquired more knowledge in UML modelling. The learning curves also prove that students do learn constraints during problem solving. Subjective evaluation shows that most of the students felt spending more time with the system would have resulted in more learning and that they found the system to be easy to use.

4. COLLECT-*UML*: MULTI-USER VERSION

The collaborative version of COLLECT-*UML* is designed for sessions in which students first solve problems individually and then join into small groups to create group solutions. The system has a distributed architecture, where the tutoring functionality is distributed between the client and the server.

The interface, which is an extension of the single-user interface, is shown in Figure 2. The problem description pane presents a design problem that needs to be modelled by a UML class diagram. Students construct their individual solutions in the private workspace (right). They use the shared workspace (left) to collaboratively construct UML diagrams while communicating via the chat window (bottom).

The private workspace enables students to try their own solutions and think about the problem before start discussing it in the group. The group area is initially disabled. When all of the students indicate readiness to work in the group by clicking on *Join the Group* button, the shared workspace is activated, and they can start placing components of their solutions in the workspace. The *Group Members* panel shows the team-mates already connected. Only one student, the one who has the pen, can update the shared workspace at a given time.

The chat area enables students to express their opinion regarding objects added to the shared area using sentence openers. The student needs to select one of the sentence openers before being able to express his/her opinion. When the student clicks on *Agree* or *Disagree* buttons for example, the sentence "I

agree …" or "I disagree …" appears in the chat window and the student may complete the sentence. The contents of selected sentence openers are displayed in the chat area along with any optional justifications.

The group moderator can submit the solution, by clicking on the *Submit Answer* button on the shared workspace. The feedback messages on the individual solutions as well as contribution to the group solution and collaboration will appear on the frame located in the right-hand side (Figure 2). The system gives collaboration-based advice based on the content of the chat area, students' participation on the shared diagram and the differences between students' individual solutions and the group solution being constructed. The task-based advice is given to the whole group based on the quality of the shared diagram.

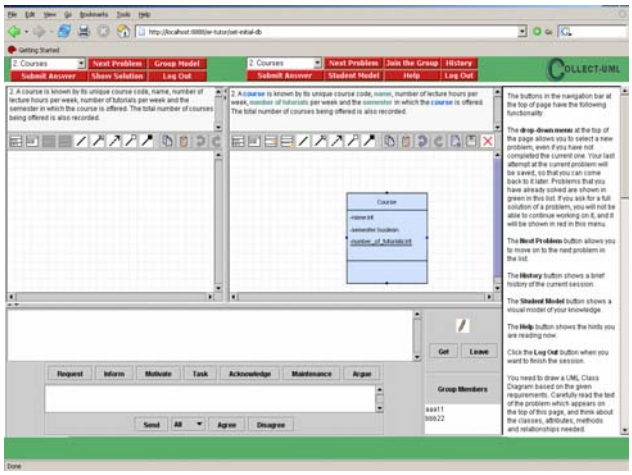


Figure 2. COLLECT-UML interface

Much research on learning has shown the potential effectiveness of collaboration for improving student's problem-solving skills. These benefits, however, are only achieved by active and well-functioning learning teams. An intelligent educational system therefore needs to provide support not only on the domain level, but also explicitly on collaboration.

The ultimate goal of COLLECT-UML is to support collaboration by modelling collaborative skills. The system is able to promote effective interaction by diagnosing students' actions in the chat area and group diagram using a set of 22 meta-constraints, which represent an ideal model of collaboration. These constraints have the same structure as domain constraint, each containing a relevance condition, a satisfaction condition and a feedback message. The feedback message is presented when the constraint is violated. In order to develop meta-constraints, we studied existing literature on characteristics of an effective collaboration. Figure 3 illustrates an example of a meta-constraint. This constraint makes sure that the student does take part in solving exercises and/or chatting.

5. CONCLUSIONS AND FUTURE WORK

This paper presented the single-user version of COLLECT-UML, and the results of the evaluation study performed. The results of both subjective and objective analysis proved that COLLECT-UML is an effective educational tool. The participants performed significantly better on a post-test after short sessions with the system, and reported that the system was relatively easy to use.

We then presented the multi-user version of the same intelligent tutoring system. We have extended COLLECT-UML interface, and developed meta-constraints, which provide feedback on collaborative activities. The goal of future work is to complete the implementation of the multi-user version and conduct a full evaluation study with second-year University students enrolled in an undergraduate software engineering course.

```
(240
"Would you like to contribute to the group discussion?"
T
(or-p (match SC CLASSES (?* "@" ?class_tag ?*))
      (match SC METHODS (?* "@" ?method_tag ?*))
      (match SC ATTRIBUTES (?* "@" ?attr_tag ?*))
      (match SC RELATIONSHIPS (?* "@" ?rel_tag ?*))
      (match SC DESC (?* "@" ?tag ?*)))
"descriptions"
nil)
```

Figure 3. An example of a meta-constraint

CBM has been used to effectively represent domain knowledge in several ITSs supporting individual learning. The contribution of the project presented in this paper is the use of CBM to model collaboration skills, not only domain knowledge. Comprehensive evaluation of the collaborative version of COLLECT-UML will provide a measure of the effectiveness of using the CBM technique in intelligent computer-supported collaborative learning environments.

6. REFERENCES

[1] Baghaei, N., Mitrovic, A. and Irwin, W. *A Constraint-Based Tutor for Learning Object-Oriented Analysis and Design using UML.* In Looi, C., Jonassen, D. and Ikeda M. (eds.) ICCE 2005, pp.11-18.

[2] Baghaei, N., Mitrovic, A. and Irwin, W. *Problem-Solving Support in a Constraint-based Tutor for UML Class Diagrams*, Technology, Instruction, Cognition and Learning Journal, 4(1-2) (in print), 2006.

[3] Constantino-Gonzalez, M., and Suthers, D. *Coaching Collaboration in a Computer-Mediated Learning Environment.* (CSCL 2002), (NJ, USA, 2002), pp.583-584.

[4] Dillenbourg, P. *What do you mean by "Collaborative Learning".* In Dillenbourg, P. (Eds.), Collaborative Learning: Cognitive and Computational Approaches, Amsterdam: Elsevier Science. pp.1-19, 1999.

[5] Fowler, M. *UML Distilled: a Brief Guide to the Standard Object Modelling Language*. Reading: Addison-Wesley, 3rd edition, 2004.

[6] Jerman, P., Soller, A. and Muhlenbrock, M. *From Mirroring to Guiding: A Review of State of the Art Technology for Supporting Collaborative Learning*. European Perspectives on CSCL (CSCL 2001), (Netherlands, 2001), pp.324-331.

[7] Mitrovic, A., Mayo, M., Suraweera, P. and Martin, B. *Constraint-based Tutors: a Success Story*. (IEA/AIE-2001), (Budapest, 2001), Springer-Verlag Berlin Heidelberg LNAI 2070, pp.931-940.

[8] Plaisant, C., Rose, A., Rubloff, G., Salter, R. and Shneiderman, B. *The design of history mechanisms and their use in collaborative educational simulations*. (CSCL 1999), (California, USA, 1999), pp.348-359.

[9] Reimann, P. *How to support groups in learning: More than problem solving*. In Aleven, V. (Eds.), Artificial Intelligence in Education (AIED 2003), (Sydney, Australia, 2003), pp. 3-16.

[10] Rosatelli, M., Self, J., and Thirty, M. *LeCs: A collaborative case study system*. 5th International Conference on Intelligent Tutoring Systems (ITS 2000), (Montreal, Canada, 2000), pp.242-251.

[11] Soller, A. and Lesgold, A. *Knowledge acquisition for adaptive collaborative learning environments*. AAAI Fall Symposium: Learning How to Do Things, Cape Cod, MA, 2000.

[12] Tedesco, P. and Self, J. A. Using meta-cognitive conflicts in a Collaborative problem solving environment. (ITS 2000), (Montreal, Canada,2000), pp.232-241.

Alternatives to stereotypes: some thoughts on issues and an outline of one game

by Sonja van Kerkhoff,
masters student in Media Technology,
LIACS, (Leiden Institute for Advanced Computer Science), Leiden University, The Netherlands.

www.sonjavank.com sonja@bahai-library.com

Abstract:

The computer game is only a little over 40 years old, and the range of online computer games is so narrow that 'action games of combat' sums them up! To quote **Chris Crawford's** Online book *The Art of Computer Game Design*[1] "Without a wide array of games there is little opportunity to choose between games; without choice there can be no natural selection." He goes on to create a taxonomy, distinguishing "skill-and-action games (emphasizing perceptual and motor skills) from strategy games (emphasizing cognitive effort)". My focus here is more on the content and in particular on today's popular MMOG (massively multi-player online games), which incorporate both of types of play. I do not aim to reach any conclusions, but rather to highlight a few issues concerning internet games as a background to the description of a project I am working on as part of my masters study in media technology at the University of Leiden.

A very brief computer game history

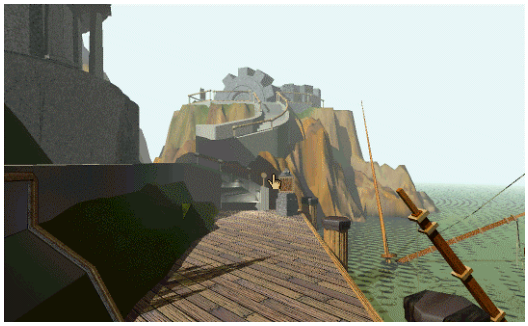
The arcade game *Spacewar!*, developed in 1962 at MIT (Stephen Russell et al.) is widely regarded as the first computer game. In it two players each control a spaceship, can shoot each other, turn their ships, and accelerate. The goal is to hit the other player before being hit yourself.
Eleven years later in 1973, Atari brought out another arcade computer game, Pong. White rectangles on a black background were manipulated to hit a ball or 'baddie'. Then came the game I remember as the first computer game, *Space Invaders*. This was brought out in 1977 by Taito: it was a more sophisticated action game.

Then computer game play changed to being primarily played in the home. This also meant more children were exposed to games on the computer. Another effect of this change of location and the introduction of cheap or flat-rate internet connection was the development of games of more complexity and with more duration. From that came games which you continue to play from where you have left off and, parallel to that, the development of player communities, where players meet online, chat or exchange information.

Single-player computer games were dominant until about 1993 when the first multi-player game, Doom (a violent *3d-shooter* or *first-person- killer game*) was brought out by ID Software. Here players at several connected PCs could play in the same game world.

I'll jump now to 2006, where on January the 14th, the top 200 massively multi-player online games (MMOG) were combat games, set variously in the military, the underworld, science fiction or Tolkien-like fantasy scenarios [2]. Even a game on the list called Frog Life, listed at 41, and with this description:
"Players compete as individuals against all the other players for speed, strength, intelligence, wealth, power, and etcetera. Over time you will rise above the rest." 2, bore the following martial image on the homepage:

One popular game that did not focus in eliminating the "enemy" was *Myst*, developed in 1993 by Cyan.



Here the player explores a world in a language of instruction that has a literary quality. For example it begins with:

> *"You have just stumbled upon a most intriguing book, a book titled Myst. You have no idea where it came from, who wrote it, or how old it is. Reading through its pages provides you with only a superbly crafted description of an island world… Suddenly your own world dissolves into blackness, replaced with the island world the pages described. Now you're here, wherever here is, with no option but to explore... (The Myst manual.) "3*

To quote further from Jesper Juul's thesis: "Myst also tries to differentiate itself from the action game with its excessive use of violence and death:

> *Myst is real. And like real life, you don't die every five minutes. In fact you probably won't die at all. [...] The key to Myst is to lose yourself in this fantastic virtual exploration and act and react as if you were really there. (Ibid.)* "[3]

If such a game of interactive fiction and exploration has been around for the for the past 15 years, why is it that no games of this type are in the top 200 list of multi player games for January 14th 2006?[4]

I do not have a cut and dry answer. My supposition is that game development has been dominated by a developer group (and expected audience) of young males with a taste for action games with varying degrees of gore and fantasy. The characters, and thus the role-models, offered by the games are drawn from a relatively limited cultural repertoire.

Before describing my attempt, with limited means, to contribute towards a greater diversity of internet play, I should mention some characteristics of the current state of internet gaming in relation to the real world of the player.

A few issues of Role Play

Live-action role-playing (LARP) blurs the boundaries of what is real and what is fantasy (a game) in two major ways. One is by blurring the location of the game itself, such as the "A.I" game "reportedly designed and operated by Microsoft as a viral marketing campaign for the film A.I: Artificial Intelligence."[5] in 2001. The game began with an enigmatic credit containing a set of mysterious symbols, at the end of the preview trailer for the film. When this name was entered into an internet search engine the viewers -now players- began a journey through a series of linked websites, where real-information was blended with the fictive world of the film. In the "months leading up to the film's premiere thousands of players took part in the game." One of the most active websites, developed by players themselves, was www.cloudmakers.org, where numerous responses were recorded as well as hints, as players helped each other to find the clues (sometimes hidden in code), such as the exchange of sounds or images. The 'unidentified game masters' also phoned players with invitations to attend real-world events. Events where attendees then phoned the clues they found to other attendees at other events.[6]

Most popular multi-player games do have clear boundaries for entrance, usually indicated by logging in and out of the game, but the fantastic 3D graphics or adrenaline-pumping action create another type of blurring of boundaries, that of game addiction which appears on the surface[7] to be an overwhelmingly male phenomena.

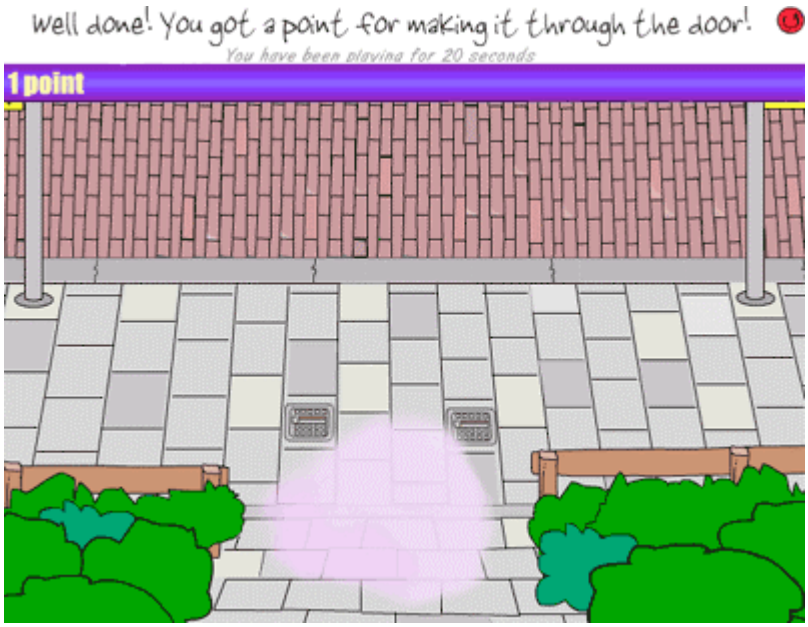
But for me a more serious issue is that a world, or worlds upon worlds of internet gaming are developing along the gender divide. My teenage sons currently live in the world of runescape8, albeit, not full of blood and gore, but you need to build weapons in order to play. A world most of their male classmates and none of their female classmates communicate with each other in. A world where as a female character you can choose not to have Lara Croft sized breasts but where your character is never-the-less one suited for some level of combat.

Game Structure
What I find lacking with such games is the linear style of play. Games such as runescape do involve some amount of strategy but the level of interaction with another live player is: I'll give you this if you give me that. Another common feature of such games is that you pick up qualities (often super or fantastical or archaic) by means of move-and-click types of responses.

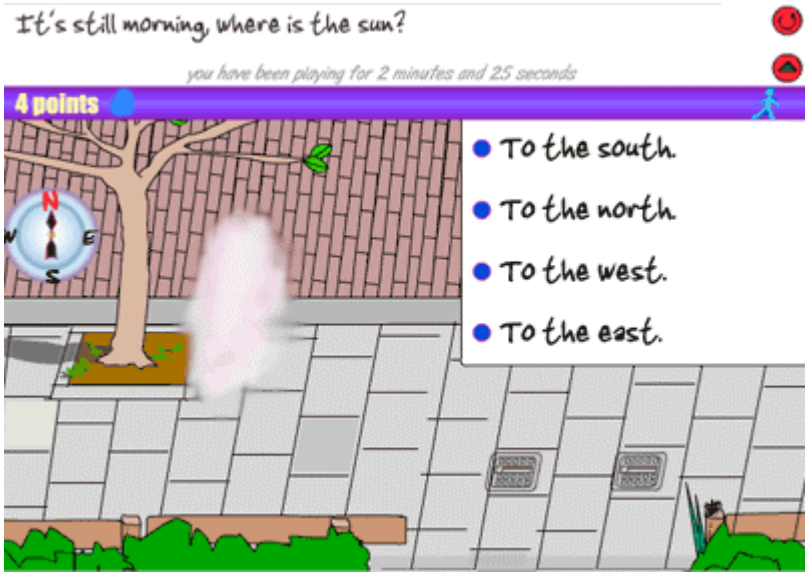
Choice – a guide to becoming human
Now I'll dive in and outline how I am attempting to overcome some of the problems of stereotypes in role play and representation, while aiming to keep the character interesting for the player.

My course focuses on practical projects. One project I am developing with another student (Ralph Kok [9]) is intended for children at the age of 8 years – below the age at the younger end of the multi player age group. It is only developed as a single player game in order to have an achieveable starting point and to be able to learn with the product. It will also be in use in a Dutch form on the website www.schooltv.nl from September 2006 onwards. The first aspect of the game we wanted to play with was to present players with an evolving, changing run-time avatar. A character whose appearance is affected by the players' behaviour.
The stated goal is that it is a game to test yourself. You are asked if you are good enough to get a hundred points and the title of the game for the player is "Choice".

We are aiming for an explorative approach to the game.

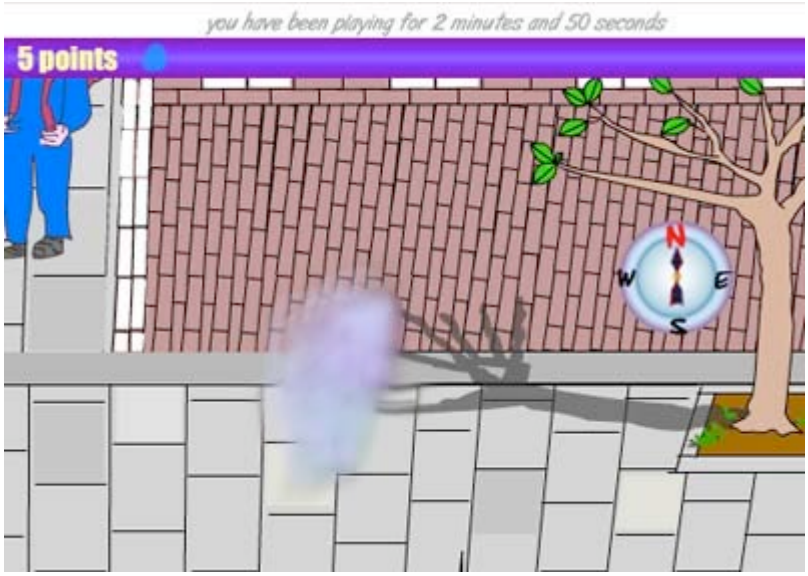


Initially, your character is unformed and amoeba-like in its movement. Mouse movement back and forth moves the background around you and in the beginning you learn that coming close to some objects such as a sign, a tree, or the street activates what we call "choice moments".



Like many of the multi player games already mentioned, interaction at this level is a question of clicking a correct answer to gain points. However all the choice locations and options are randomly built. So sometimes, if you try to cross the road, you just lose a few points, at other times you die. Sometimes there is a tree, sometimes there is something else or nothing. We aim to encourage curiosity within certain limits. Here the player is going to school, as indicated in the question: "Which way is school" in an earlier frame:

Time stands still while the avatar in the "choice moment" so as not to discourage deliberation and to provide moments for real-world discussion or time out of the game. Once you click on an option, time resumes. If you have not dawdled too much on the way to school, there will be time at the school gate for more "choice moments" and if you have sufficient sight, you could see more things. If you are late, you might lose some points for being late.



When new ways of interaction are introduced, the player gets hints that rise up from the bottom of the screen during play, such as the hint “Move closer to that person to find out more”.
Progress in location is indicated in the blue bar above – the light blue form moves towards the left as the player proceeds and the shape ‘evolves’ just as the character does.

By the time the player has reached the stoplights to cross the road, they will have passed 3 possibilities for gaining an eye. Getting an eye would depend on the number of points gained from 1 of the 3 possible nature/environment-related “choice moments” as well as points from the human-related interaction. There’s a reasonable chance for a curious player to have gained an eye (and in response, visuals or details seem brighter / clearer). This is a way of teaching the player to collect characteristics as well as points.

Once the player passes these points, they will have learnt the basics through play, and now the game becomes more complex in content (the options are not so ‘black and white’) and context plays a greater role in gaining points. Sometimes all the options will be ‘wrong’, for example because you have just bumped into someone. Here dilemmas such as whether to be nice or rude or fair or reasonable come to the fore. Also issues such as when playing becomes too rough and turns into fighting. Interpretation of others' feelings is another issue that arises in these interactions. After all, learning how to relate to others is what it is all about in becoming human, or in some cases superhuman!

This page maintains a current status of the development of this project: www.sonjavank.com/choice.htm

References:

1. *The Art of Computer Game Design* by **Chris Crawford,** http://www.vancouver.wsu.edu/fac/peabody/game-book/

2. http://www.gamesites200.com/mpog/

3. *A history of the computer game*, thesis by Jesper Juul http://www.jesperjuul.dk/thesis/2-historyofthecomputergame.html

4. I need to state the date because the ordering of the popularity of these games is democratically allotted and the ordering changes on a daily basis.

5. "This is not a game: Play in cultural environments" by Katie Salen and Eric Zimmerman, pp 17, in *Level up: Digital Games Research Conference*, 4-6 November 2003, Utrecht, The Netherlands.

6. The rest of this paragraph is paraphrased from the above.

7. A few links to news about game addiction: http://news.com.com/2100-1040-881673.html, http://www.wired.com/news/holidays/0,1882,48479,00.html,

8. www.runescape.com. It would be interesting to see how many players were really female in the game. My sons informed me that a number of their male players' characters are female.

9. Ralph Kok's website is: www.rockabit.com