Research Conference on Information Technology

Hosted by

Zayed University - College of Information Technology

March 28, 2007 – Venue: ZU Dubai Campus, R-L1-015

Program

Coffee/ Registration: 8:30 am

Welcome – 9:00: M. Majewski and Z. Maamar, **Conference Co-Chairs, Zayed University**

Session 1 - 09:10 - 10:30 – Chair: N. Werghi, University of Dubai
1. **9:10 – 9:30** Bioinformatics: Protein Secondary Structure Prediction Present and Future Era, *Saad Subair, Al Ghurair University.*
2. **9:30 – 9:50** The Delphi Method for Graduate Research, *Gregory J. Skulmoski, Zayed University.*
3. **9:50 – 10:10** Education FOR and IN the Emerging Information Age, *Jeffrey Barker, Abu Dhabi University.*
4. **10:10 – 10:30** Integrating Policies into a Web Services Composition Approach, *Zakaria Maamar, Zayed University.*

Coffee Break - 10:30 - 10:50

Session 2 - 10:50 - 12:30 – Chair: L. Kaya, Sharjah University
1. **10:50 – 11:15** Information Technologies for Mathematical Visualization, *Miroslaw L. Majewski, Zayed University.*
2. **11:15 – 11:40** An unsupervised learning approach based on Hopfield-like network for assessing posterior capsule opacification, *Naoufel Werghi, University of Dubai & Rachid Sammoud and Fatma AlKirbi, University of Sharjah.*
3. **11:40 – 12:05** Methodologies & Approaches for The UAE Electronic Workforce, *Fahim Akhter, Ayesha Ghulam, Bushra Hassan, Sultana Awadhi, and Sumeya Hammadi, Zayed University.*
4. **12:05 – 12:30** Determining the Optimal Trade-Off between False Acceptance Rate and False Rejection Rate Using Conjoint Analysis, *Wendy Hui and Fahim Akhter, Zayed University.*

Lunch - 12:30 - 02:00 Informal discussion: How to boost research in the UAE?

Session 3 - 02:00 - 04:00 – Chair: J. Barker, Abu Dhabi University
1. **2:00 – 2:20** A Multi-Level Integration Framework for Software Components, *Leon Jololian, Zayed University.*
2. **2:20 – 2:40** Practical Research Towards Creating an IT Society, *Lami Kaya, Sharjah University.*
3. **2:40 – 3:00** Exploratory Study on an Innovative Use of COSMIC-FFP for Early Quality Assessment, *Manar Abu Talib, Zayed University.*
4. **3:00 – 3:20** Interactive On-Line Mathematics for Math Education, *Philip Burton, Zayed University*

Closing – 3:30 M. Majewski and Z. Maamar

Table of contents

Bioinformatics: Protein Secondary Structure Prediction Present and Future Era

Saad Subair

College of Computing
Al Ghurair University
Dubai, UAE

ABSTRACT: Advances in molecular biology in the last few decades, and the availability of equipment in this field have allowed the increasingly rapid sequencing of considerable genomes of several species. In fact, to date, several bacterial genomes, as well as those of some simple eukaryotic organisms –like yeast- have been completely sequenced. The Human Genome Project (HGP), aimed to sequence all of the human chromosomes, is almost completed with a rough draft announced in the year 2000. Known sequencing databases projects, such as GenBank, PDB, and EMBL, have been growing significantly. This surge and overflow of data and information have imposed the rational storage, organization and indexing of sequence information. The application of computer and mathematical sciences to the science of biology results in the Bioinformatics field.

The most fundamental tasks in bioinformatics include the analysis of sequence information which involves the following the prediction of the 3D structure of a protein using algorithms that have been derived from the knowledge of physics, chemistry and from the analysis of other proteins with similar amino acid sequences. Some researchers refer to this area with the name Computational Biology.

Protein structure prediction is categorized under Bioinformatics which is a broad field that combines many other fields and disciplines like biology, biochemistry, physics, statistics, and mathematics. Proteins are series of amino acids known as polymers linked together into contiguous chains. In a living cell the DNA of an organism encodes its proteins into a sequence of nucleotides (transcribed), namely: adenine, cytosine, guanine and thymine that are copied to the mRNA which are then translated into protein.

Protein has three main structures: primary structure which essential the linear amino acid sequence and usually represented by a one letter notation. Alpha helices, beta sheets, and loops are formed when the sequences of primary structures tend to arrange themselves into regular conformations; these units are known as secondary structure. Protein folding is the process that results in a compact structure in which secondary structure elements are packed against each other in a stable configuration. This three-dimensional structure of the protein is known as the protein tertiary structure. However, loops usually serve as connection points between alpha-helices and beta-sheets, they do not have uniform patterns like alpha-helices and beta-sheets and they could be any other part of the protein structure rather than helices or strands .In the molecular biology laboratory, protein secondary structure is determined experimentally by two lengthy methods: X-ray crystallography method and Nuclear Magnetic Resonance (NMR) spectroscopy method.

Since it had been concluded that the amino acid sequence is the only source of information to survive the denaturing process, and hence the structured information must be somehow specified by the primary protein sequence, researchers have been trying to predict secondary structure from protein sequence. the hypothesis suggests that an ideal theoretical model of predicting protein secondary structure form its sequence should exist anyhow.

There are two main different approaches in determining protein structure: a molecular mechanics approach based on the assumption that a correctly folded protein occupies a minimum energy conformation, most likely a conformation near the global minimum of free energy. Potential energy is

obtained by summing the terms due to bonded and non-bonded components estimated from these force field parameters and then can be minimized as a function of atomic coordinates in order to reach the nearest local minimum. This approach is very sensitive to the protein conformation of the molecules at the beginning of the simulation.

One way to address this problem is to use molecular dynamics to simulate the way the molecule would move away from that initial state. Newton's laws and Monte Carlo methods were used to reach to a global energy minima. The approach of molecular mechanics is faced by problems of inaccurate force field parameters, unrealistic treatment of solvent, and spectrum of multiple minima.

The second approach of predicting protein structures from sequence alone is based on the data sets of known protein structures and sequences. This approach attempts to find common features in these data sets which can be generalized to provide structural models of other proteins. Many statistical methods used the different frequencies of amino acid types: helices, strands, and loops in sequences to predict their location. The main idea is that a segment or motif of a target protein that has a sequence similar to a segment or motif with known structure is assumed to have the same structure. Unfortunately, for many proteins there is no enough homology to any protein sequence or of known structure to allow application of this technique.

The previous review leads us to the fact that the approach of deriving general rules for predicting protein structure from the existing data sets or databases and then applies them to sequences of unknown structure appears can be promising. Several methods have utilized this approach.

Artificial Neural networks have great opportunities in the prediction of proteins secondary structures. These methods are based on the analogy of operation of synaptic connections in neurons of the brain, where input is processed over several levels or phases and then converted to a final output. Since the neural network can be trained to map specific input signals or patterns to a desired output, information from the central amino acid of each input value is modified by a weighting factor, grouped together then sent to a second level (hidden layer) where the signal is clustered into an appropriate class.

The GOR method was first proposed by and named after its authors Garnier-Osguthorpe-Robson. The GOR method attempts to include information about a slightly longer segment of the polypeptide chain. In stead of considering propensities for a single residue, position-dependent propensities have been calculated for all residue types. Thus the prediction will therefore be influenced not only by the actual residue at that position, but also to some extent by other neighbouring residues. The propensity tables to some extent reflect the fact that positively charged residues are more often found in the C-terminal end of helices and that negatively charged residues are found in the N-terminal end.The GOR method is based on information theory and was developed by and the mostly known GORIV version uses all possible pair frequencies within a window of 17 amino acid residues with a cross-validation on a data base of 267 proteins. The GORIV program output gives the probability values for each secondary structure at each amino acid position. The GOR method is well suited for programming and has been a standard method for many years.

Determining the Optimal Trade-Off between False Acceptance Rate and False Rejection Rate Using Conjoint Analysis

Wendy Hui

College of Information Technology
Zayed University
Abu Dhabi, United Arab Emirates
wendy.hui@zu.ac.ae

Fahim Akhter

College of Information Technology
Zayed University
Dubai, United Arab Emirates
fahim.akhter@zu.ac.ae

ABSTRACT: Information security technologies that detect and prevent intrusions try to achieve a low false acceptance rate (FAR) and a low false rejection rate (FRR). However, very often, a trade-off between the two is involved. Increasing the sensitivity of a device or application would decrease the FAR, but increase the FRR. One could set the technology to a security level where the FAR equals the FRR, and this level is known as the crossover error rate (CER), which is commonly used to compare the accuracy of different security technologies. However, the optimal level of sensitivity for a piece of information security technology is dependent on the context in which it is applied. An organization may find a false rejection more tolerable than a false acceptance, and therefore the CER cannot be the optimal level of sensitivity for all situations. We propose the use of conjoint analysis to understand end-users' preference for FAR and FRR. This understanding of end-user preference can be used as an objective guideline the design of information security technologies. In this study, the technology we investigate is spam filtering. A pilot study has been conducted on a group of students at Zayed University, Abu Dhabi, at the beginning of Spring 2007. Preliminary results are presented.

Electronic Commerce Security: Methodologies & Approaches for the United Arab Emirates Electronic Workplace

Fahim Akhter; Ayesha Ghulam; Bushra Hassan; Sultana Awadhi; Sumeya Hammadi

College of Information Technology
Zayed University
Dubai, United Arab Emirates

{I7060, U2324548, U2207571, U2323266, U2422213}@zu.ac.ae

ABSTRACT: As electronic commerce (e-commerce) function in a more perplex and conjoint environment than traditional businesses, a higher degree of trust is required between the users and online businesses. Uncertainties inherent in the current e-commerce environment give rise to a lack of trust and reliability in e-commerce partnerships, thereby reducing confidence and creating barriers to trade.

The rationale why most users and businesses in U.A.E are still skeptical about e-commerce is the perceived security risks associated with conducting online business. Online users consciously or subconsciously analyze the provided level of security based on their experience in order to decide whether to conduct business with the specific company or else to move on to the next company. There is a need for a better understanding of hostile environment fuelled by financially motivated and targeted cyber threats that affect consumer's decision-making behavior. The purpose of the study was to identify the factors that support the implementation and acceptance of security in e-commerce among corporate United Arab Emirates (U.A.E.). The study has explored the common cyber attacks that threaten the U.A.E online businesses and described methodologies and approaches that can be developed to respond to those threats.

A descriptive web-based survey has adopted as an appropriate method to collect initial data from users due to its cost effectiveness, rapid turn-around, high response volume, and ability to cover a large geographical area. The combination of twenty-five questions of closed and open-ended has selected. The URL of the survey was electronically distributed among participant using mailing lists from Dubai chamber of commerce.

According to the statistics, 630 people from seven states of U.A.E accessed the survey and 410 actually filled out the form. Three hundred and seventy complete responses were chosen out of 410 anonymous responses for further analysis. A quantitative data was feed into a statistical package for the social sciences (SPSS) to understand and analyze the relationship among different responses.

Education FOR and IN the Emerging Information Age

Jeffrey Barker

Dean, CSIT, Abu Dhabi University

jeffrey.barker@adu.ac.ae

ABSTRACT: The world has moved through a number of information generations. For non-verbal information, we have progressed from cave paintings to hand-written manuscripts to printed documents to electronic documents and finally, with the global use of the world wide web, into a new emerging era of "dynamic information", termed Information Generation 3 (IG3).

Each of these information generations have effected our "systems"; our finance systems, business systems, social systems, legal systems, economic systems, management systems, manufacturing systems and, of course, information and communication (ICT) systems.

But what about education? Isn't it the most important system of all?

Education programs must be designed for IG3, the era of dynamic information, to adequately prepare students for this new volatile and dynamic era the world is now moving into.

Prof Jeff Barker's presentation will: give a simple non-technical definition of IG3 and its consequences discuss education for IG3; the importance of preparing students for working in the IG3 world discuss education in IG3; using IG3 technology to move education and teaching into the 21st century and the emerging information age.

A Multi-Level Integration Framework for Software Components

Leon Jololian

College of Information Technology
Zayed University
Dubai, United Arab Emirates
Leon.Jololian@zu.ac.ae

ABSTRACT: The lack of comprehensive set of standards to define rules of engagement between software components has been a major impediment to component integration. While several competing component architectures exists today, including JavaBean, COM, and CORBA, yet components belonging to different architectures lack the inherent compatibility necessary to integrate them within a single framework.

At the wiring level, components can be made to communicate with each other through the use of bridges and proxies that allow one component to request the invocation of a service from another component and then have that invocation translated into a format that the target component can accept. In this way, the problem defined at the wiring level has been addressed through mediation rather than through standardization. The solution is workable because in the presence of multiple standards, it is possible to convert from one set of standards into another, through the use of adapters.

In this presentation we argue that for achieving component integration, removing mismatches at the wiring level is only the first step. There are two additional levels above the wiring level that must be addressed. We refer to these two levels as the operational and semantic levels. The focus of this paper is to lay the framework for defining the type of mismatches that must be overcome at the operational level. It turns out that mismatches between components at the operational level are often traced to three basic elements: data, function, and control. This is due to the fact that in the design of components, developers use different models for the data imported/exported by the component, the services assumed to exist within the execution environment, and the control mechanism used in invoking services between components.

Through the principle of separation of concerns, we propose to separate the issues related to each of the three basic elements and then deal with each element independently. It is expected that standards can be established for describing the characteristic of each type of element (data, function, and control) within the component design model. Having these standards as meta-data for each component, it will become possible to mediate the mismatches between components in a way that is similar to what is possible at the wiring level.

Information Technologies for Mathematical Visualization

Miroslaw L. Majewski

College of Information Technology, Zayed University
Abu Dhabi, United Arab Emirates
mirek.majewski@mupad.com

ABSTRACT: Among other applications, Computer Algebra Systems (in short CAS) are widely used in scientific research and teaching. For both of those applications, visualization is an important operation and most modern CAS support a variety of plotting functions in two or three dimensions and sometimes also techniques to visualize certain objects in more than three dimensions.

For about eight years I am involved in development of MuPAD a general purpose Computer Algebra System that was invented at the Paderborn University in Germany and is developed by SciFace Software. MuPAD is available for Windows, Linux and Mac OS. As an informal member of the MuPAD research team I was participating in research on developing visualization tools of MuPAD and in particular the recent version of so called Virtual Camera.

In my presentation I will cover some of the aspects of developing a tool for mathematical visualization. I will describe three points of view – the user point of view on mathematical visualization, a technical point of view as well as programmer point of view. I will show how various recent developments in Information Technology influenced development of visualization tools. Finally, I will present some possible directions of future developments in mathematical visualization, and in particular development of highly intelligent tools combining dynamic geometry interfaces with symbolic computations.
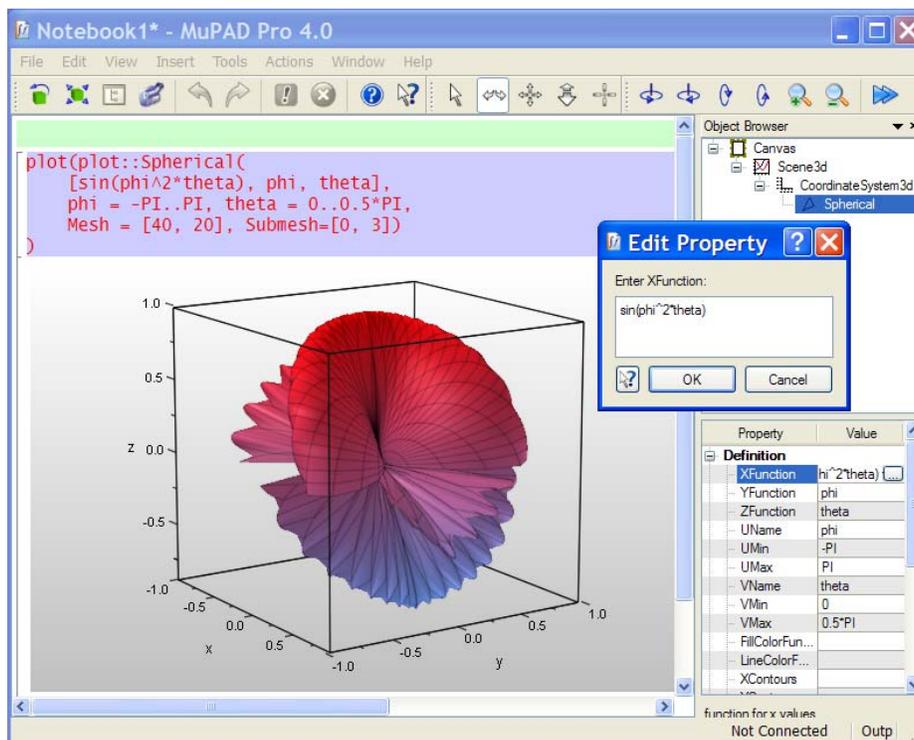

Fig. 1 – Modern interface for object oriented mathematical visualization

Exploratory Study on an Innovative Use of COSMIC-FFP for Early Quality Assessment

Manar Abu Talib

College of Information Technology, Zayed University
Abu Dhabi, United Arab Emirates

ABSTRACT: The functional size measurement method, COSMIC-FFP, adopted as the ISO/IEC 19761 standard in 2003, was developed by the Common Software Measurement International Consortium (COSMIC). It focuses on the "user view" of functional requirements and is applicable throughout the development life cycle. As some of the software systems targeted by COSMIC-FFP are large-scale and inherently complex, feedback on their functional complexity would facilitate containment of that complexity throughout the software life cycle.

In my research, a new early quality assessment of COSMIC-FFP models is proposed. The benefits of this work include earlier prediction of the functional complexity of the behavior of software in the COSMIC-FFP context, right from the requirements phase, as well as a mechanism for generating black-box test cases from the COSMIC-FFP model, test case prioritization and test set adequacy monitoring and optimization within given budget constraints, and an early prediction of reliability based on Markov chains. We also present a study of the scales, units and scale types of both COSMIC-FFP and the Entropy-based Functional Complexity Measure that forms the basis of the testing assessment method we propose here. Previous studies have analyzed the scale types of many pieces of software, but not the concept of scale itself, nor how it is used in the design of a measurement method. Two well-known case studies are introduced to demonstrate the applicability of the proposed methods: the Hotel Accommodation System and the Railroad System.

We include a formalized COSMIC-FFP definition in the AS-TRM context (Autonomic Systems Timed Reactive Object Model), a language for the formal design of autonomic reactive systems developed at Concordia University. We introduce the Steam Boiler case study to demonstrate the applicability of formalizing COSMIC-FFP in the AS-TRM context. Future work based on this research can include the development of AS-TRM specifications for several benchmark case studies, and the collection of COSMIC-FFP measurement data for both the theoretical and empirical validation of the proposed measurement method.

The testing method proposed here has been adapted to a specific class of projects, namely Enterprise Resource Planning (ERP) projects, which are perceived to be mission-critical initiatives in many organizations. They can be found in business transformation programs and are instrumental in improving organizational performance.

The major contributions in this research have been published in the following journals and papers:

1. Manar Abu Talib, Olga Ormandjieva, Alain Abran, Adel Khelifi, Luigi Buglione. Scenario-based Black-Box Testing in COSMIC-FFP: a Case Study. ASQ Software Quality Professional Journal 8 (3), June 2006, pp.23-33.

2. O. Ormandjieva, M. Abu Talib, A. Abran. Reliability Model for Component-Based Systems in COSMIC-FFP (a Case Study). Submitted to the International Journal of Software Engineering and Knowledge Engineering: Special Issue on Testing and Quality Assurance for Component-Based Systems.

3. Manar Abu-Talib, Alain Abran, Olga Ormandjieva. Markov Model and Functional Size with COSMIC-FFP. In Proceedings of the IEEE Conference International Symposium on Industrial Electronics (IEEE-ISIE3006), July 9-13 2006, Montreal, Canada.

4. M. Abu-Talib, O. Ormandjieva, A. Abran, L. Buglione. COSMIC-FFP & Functional Complexity (FC) Measures: A Study of Their Scales Types. In Proceedings of the 15th International Workshop on Software Measurement (IWSM2005), pp.209-226, September 12-14, 2005, Montreal, Canada.

5. Abu Talib, M., O. Ormandjieva, A. Abran and L. Buglione (2005). Scenario-based Black-Box Testing in COSMIC-FFP. Proceedings of the 2nd Software Measurement European Forum 2005 (SMEF 2005), Italy.

6. A. Abran, O. Ormandjieva, M. Abu-Talib. Information Theory-Based Functional Complexity Measures and Functional Size with COSMIC-FFP. In Proceedings of the 14th International Workshop on Software Measurement (IWSM 2004), September 2004. Montreal.

7. Manar Abu-Talib, Olga Ormandjieva, Alain Abran. AS-TRM and Functional Size with COSMIC-FFP. Submitted to the IEEE International Symposium on Industrial Electronics – ISIE 2007.

8. M. Daneva, A. Abran, O. Ormandjieva, M. Abu Talib. A case study of metric-based and scenario-driven black-box testing for SAP projects. In Proceedings of 16th International Workshop on Software Measurement (IWSM2006), Germany, November 2006.

The Delphi Method for Graduate Research

Gregory J. Skulmoski
College of Information Technology
Zayed University
Abu Dhabi, United Arab Emirates

Gregory.Skulmoski@zu.ac.ae

Francis T. Hartman
University of Calgary,
Calgary, Canada

Fhartman@ucalgary.ca

Jennifer Krahn
University of Calgary,
Calgary, Canada
Canadajenny.krahn@haskayne.ucalgary.ca

The Delphi method is an attractive method for graduate students completing masters and PhD level research. It is a flexible research technique that has been successfully used in our program at the University of Calgary to explore new concepts within and outside of the information systems body of knowledge. The Delphi method is an iterative process to collect and distill the anonymous judgments of experts using a series of data collection and analysis techniques interspersed with feedback. The Delphi method is well suited as a research instrument when there is incomplete knowledge about a problem or phenomenon; however it is not a method for all types of IS research questions. The Delphi method works especially well when the goal is to improve our understanding of problems, opportunities, solutions, or to develop forecasts. In this paper, we provide a brief background of the Classical Delphi followed by a presentation of how it has evolved into a flexible research method appropriate for a wide variety of IS research projects, such as determining the criteria for IS prototyping decisions, ranking technology management issues in new product development projects, and developing a descriptive framework of knowledge manipulation activities. To illustrate the method's flexibility, we summarize distinctive non-IS, IS, and graduate studies Delphi research projects. We end by discussing what we have learned from using the Delphi method in our own research regarding this method's design factors and how it may be applied to those conducting graduate studies research: i) methodological choices such as a qualitative, quantitative or mixed methods approach; ii) initial question degree of focus whether it be broad or narrowly focused; iii) expertise criteria such as technical knowledge and experience, capacity and willingness to participate, sufficient time, and communication skills; vi) number of participants in the heterogeneous or homogeneous sample, v) number of Delphi rounds varying from one to 6, vi) mode of interaction such as through email, online surveys or groupware, vii) methodological rigor and a research audit trail, viii) results analysis, ix) further verification through triangulation or with another sample, and x) publishing of the results. We include an extensive bibliography and an appendix with a wide-ranging list of dissertations that have used the Delphi method (including brief research description, number of rounds and sample size). The Delphi method is a flexible, effective and efficient research method that can be successful used by IS graduate students to answer research questions in information systems and to rigorously advance the IS body of knowledge.

Integrating Policies into a Web Services Composition Approach

Zakaria Maamar

College of Information Technology
Zayed University, Dubai Campus
Zakaria.maamar@zu.ac.ae

ABSTRACT. We present a four-layer approach to compose Web services. These layers are policy, user, Web service, and resource. We emphasize the role of the policy layer in concern separation between the specification of a composition and the functionalities of a Web service. A specification is about the operations that satisfy user needs, the order of these operations, and the way exceptions are handled. Functionality is about the capacity of a Web service to perform the operations of a specification. In addition, we discuss the way policies define the behavior of Web services based on the execution progress of a composition specification. This behavior is of four types, which are referred to as permission, restriction, dispensation, and obligation. A prototype illustrating the deployment of the multi-later approach to compose Web services is also presented.

Keywords: Behavior, Composition, Policy, and Web service.

An unsupervised learning approach based on Hopfield-like network for assessing posterior capsule opacification

Naoufel Werghi
College of Information Technology
University of Dubai
nwerghi@ud.ac.ae

Rachid Sammoud, Fatma AlKirbi
Department of Computer Sciences
University of Sharjah
{rsammouda,kfatma}@sharjah.ac.ae

Posterior Capsule Opacification (PCO) is a frequent post-operative complication that attacks the eye after undergoing a cataract extraction surgery in which the infected lens is removed and replaced by an artificial one (the intraocular lens or IOL). Such intervention is often followed by a complication, known as the Posterior Capsule Opacification, which manifests by the development of lens epithelium cells (LEC) that grow at the posterior half of the eye capsule behind the IOL. PCO is the commonest complication of cataract surgery occurring in up to 50% of patients by 2 to 3 years after the operation. There are different types of PCO, namely fibrosis, pearls and wrinkles. Theses names reflect the type of texture produced by the LEC distribution (Figure 1). Accessing the effectiveness of the
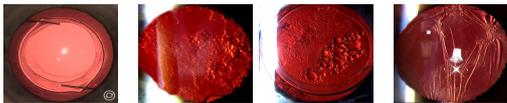

Figure 1: safe eye's capsule, eyes affected respectively with fibrosis , pearls and wrinkles PCO.

clinical trials performed to reduce or inhibit PCO requires a quantitative and qualitative analysis of the PCO, capable of measuring the amount of PCO in the eye's capsule with a reasonable reliability and accuracy. Human assessment is often corrupted by bias, subjectivity and lack of accuracy. This happens, for instance, when comparing PCO progression by studying images taken before and after treatment and also when comparing the severity of the PCO. This issue, fuelled the investigation of image processing technique that might overcome the limitations of manual assessment, particularly with regard to the subjectivity.

This paper presents some steps towards an objective and reliable assessment of PCO images. It proposes a methodology for defining measurement tool that would contribute in reaching this objective.

In the literature methods using digital images for assessing PCO addressed the problem either manually or automatically. In the manual approach [2, 1] the operator can trace manually the boundaries of the texture areas in the images. In the automatic approaches employed adhoc techniques [3, 4, 5] or texture analysis techniques [6, 7, 8].

A common factor in these approaches is the attempt to identify the textured areas that reflect PCO in the image, and eventually consider the relative size of these areas with respect to the whole image as a "measure" of the PCO amount. While theoretically this approach looks valid and justified, it raises serious problems practically. In effect the PCO texture characterized by a high level of irregularity in terms of both scale and frequency. Attempting the extraction of such areas via a segmentation approach would be a quite difficult task, bearing in mind that the segmentation of regular texture is already a hard problem.

The idea on which we build our approach, emanates from the concept of a region within a digital image, which we define in our application as a group of connected pixels sharing close chromatic value. In a smooth and clean image, corresponding to a safe eye, the number of regions is small. Ideally such image will show a single and compact region. In the opposite, an image corresponding to PCO case, would be characterized by a relatively larger number of regions, induced by the discontinuities and variations inferred by the PCO texture. Moreover the number of regions is expected to be function of the amount of PCO, and thus can reflect its level. Assuming this hypothesis ,we propose to investigate the suitability of the number of regions as measuring tool for assessing PCO from digital images.

the proposed paradigm consists in extracting and counting the regions in the image. The whole process involves a preprocessing stage whereby a circular central region is selected (Figure 2 (1.b)), covering about 40% of the whole image. This is because ophthalmologists focus on the central area, as it is the one affecting the most the vision acuity. The rest of the image is considered as background. Pixels in that region are set to the average value of the area of interest. Afterwards, an unsupervised clustering technique is applied to the image. In this stage the image pixels are assigned to different classes according to their chromatic values (Figure 2 (1.c)). Next, the image is filtered to remove artifacts and tiny-regions caused by over-clustering (Figure 2 (1.d)) . In the next stage, pixels in each class are grouped into sets of connected regions (Figure 2 (1.e)), so that the number of regions in the whole image can be determined.

The clustering aims to partition image pixels into K groups (clusters) sharing the same chromatic characteristics. The objective here is not to determine the optimal
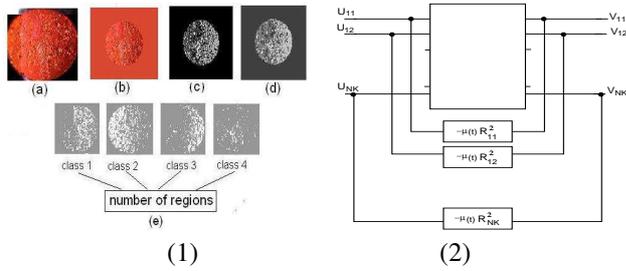
Figure 2: (1):Overview of the region extraction process: a: Original image, b: pre-processing and selection of the area of interest,c: image clustering using unsupervised classification technique, d:post-processing to eliminate artifacts and tiny regions. (e) extracting regions from each class and determining the total number of regions. (2): Architecture of the unsupervised classification network.

number of clusters, but rather to examine the variation of the number of regions with respect to the number of classes. We propose to use the standard *K-means* clustering technique [**?**]. This technique basically searches for the optimal partition of a set of objects (e.g. pixels) ,whereby each object is represented by a vector of feature (e.g. RGB components), into a K disjoint subsets (clusters) S_l, $l = 1..K$. The optimal partition is obtained by minimizing a given objective function, usually defined by the sum of the squared distances between the object feature vectors of each subset and their centroids $\mu_l, l = 1..K$. $E = \sum_{l=1}^{K} \sum_{X_i \in S_l} \|X_i - \mu_l\|^2$

The implementation of the algorithm deployed a Hopfield-like artificial neural network, inspired the architecture proposed in [**?**]. The structure of the network is composed of a $N \times K$ cells (Figure 2(2)). U_{il} and V_{il} represent respectively the input and the output of the neuron associated to i^{th} pixel and l^{th} class. The network has feed-back loops from the output to the inputs. After applying a new input, the network output is calculated and fed back to the input, then the output is calculate again. The process is repeated until the output becomes constant. The PCO image is considered as a multidimensional data where each pixel is represented by its three RGB components.

The classification was applied on a variety of 20 PCO images. For each classified image we extracted the number of regions, we remind here that a region refers to a group of connected pixels that belong to a same class. Figure 2, top depicts the distribution of the number of regions across the 20 images. We notice that the variations of the number of regions exhibit sensibly the same variation profile at each trial made with different number of classes.

We tested the validity of our approach by comparing the corresponding assessment with a medical expertise evaluation, that involved three observes. This was conducted as follows we brought the number of regions, extracted from all the images at each trial (with 3,4,5 and six clusters), down to the standard medical range [0-4]. This is a discrete range when the possible PCO level score can be either 0 (no PCO), 1 (mild), 2 (average) 3(strong) and 4
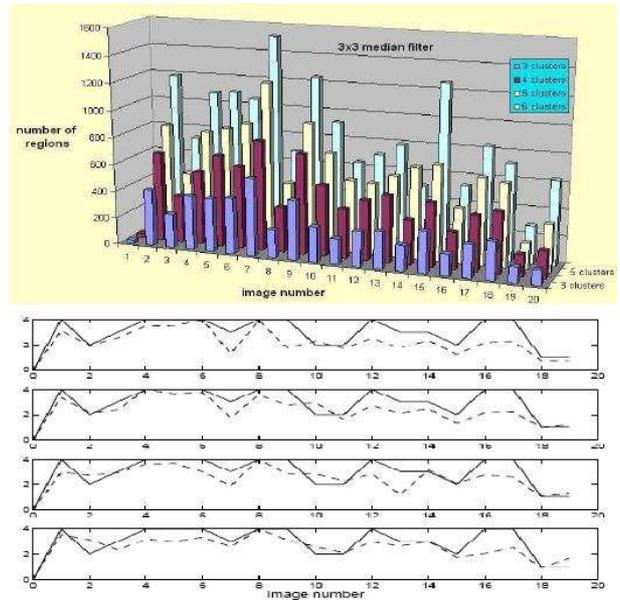
(severe).

Few samples of the results are depicted in Figure 2, bottom, where the medical expertise assessment is plotted together with each of the gradings obtained from the four trials. We notice that the variation profile of the medical expert grading and our approach's grading exhibit clear similarities across all the series of trials, despite some disparities at some instances. The the correlation factor between the two scorings is around 0.8 which is quite reasonable. As a future work we plan to investigate deploying fuzzy cluttering technique to best handle the boundary cases.

References

[1] K. Hayashi, H. Hayashi, F. Nakao, F. Hayashi, "In vivo quantitative measurement of posterior capsule opacification after extracapsular cataract surgery", American Journal of Ophthalmology, Vol. 125, No.6, pp. 837-843, 1998.

[2] M.R. Tetz, G.U. Auffrath, M. Speaker. "Photographic Image Analysis System, of Posterior Capsule Opacification". Journal of Cataract Refract Surgery, Vol. 23, pp.1515-1520, 1997.

[3] M.C. Wang, L.C. Woung, "Digital Retroilluminated Photography to analyze posterior capsule opacification in eyes with intraocular lenses". Journal of Cataract Refract. Surgery, Vol. 26, pp.56-51, 2000

[4] D.S. Friedman D.D. Duncan, B. Munoz, "Digital Image Capture and Automated Analysis of Posterior Capsular Opacification". Investigative Ophtalmolology Vision Sciences Vol. 40, pp.1715-1726, 1999.

[5] B. Barman E.J Hollick, J.F Boyce, "Quantification of Posterior, Capsular Opacification in digital Images after Cataract Surgery". Investigative Ophthalmology Vision Sciences, Vol. 41, pp.3883-3892, 2000.

[6] A.P. Paplinski and J.F. Boyce, "Segmentation of a class of ophthalmological images using a directional variance operator and co-occurrence arrays", Optical Engineering, Vol. 36, pp.3140-3147, November 1997.

[7] H. Siegl, A. Pinz, W. Bhl, M. Georgopoulos, O. Findl, R.Menapace, "Assessment of posterior capsule opacification after cataract surgery", Proceedings 12th SCIA, Scandinavian Conference on Image Analysis, Bergen, pp. 54-61 (2001)

[8] O. Findl et al , "Comparison of 4 methods for quantifying posterior capsule opacification" J Cataract Refract Surgery, Vol. 29, pp.106-111, 2003.

Practical Research Towards Creating an IT Society

Lami Kaya

Sharjah University
Computer Science Department
United Arab Emirates
LKaya@sharjah.ac.ae

ABSTRACT: The research we are conducting can be classified broadly into two categories, namely theoretical research and practical research. In the theoretical research category, we are working in the field of information theory, especially on multi-dimensional error detection/correction codes and effective decoding strategies. On the practical research, we are conducting some work that is geared towards helping to create an IT society. In this respect, we have research interests in a number of fields. We are developing some applications for mobile devices, such as mobile phones, PDAs, Pocket PCs with emphasis being on localization and Arabization. Some work is carried on interoperability issues of distributed geographic information systems (DGIS). Also, effective data stream processing in sensor networks employed in phenomena tracing applications and their security aspects are being investigated.

The availability of small low-power hardware and compact operating systems makes the deployment of large scale sensor networks (SN). Many applications of SN requires to process huge amount of continuous data. Sensor nodes are equipped with memory, processors, sensors, and short-range wireless communication. The usage of SN offers many opportunities for applications to observe phenomena in our physical world. Sensor devices have limitations in their memory, processing capability, battery power, and bandwidth. There are many issues needs to be investigated and improved, we are conducting work on effective processing of data streams in SN [1] and light-weight security solutions [2] suitable for such environments.

There has been great interest and demand on e-services in the region, especially in the UAE. There have been already some governmental initiatives to provide DGIS services in number of emirates, a good indication of interest in these field is creation of GIS centers in Dubai and Sharjah. The distributed GIS, is an example of the deployment of information systems (IS) from traditional architecturally closed and centralized-IS to more open and distributed-IS architectures [6]. The goal is to provide on-line GIS modelling and spatial analysis functions without running GIS servers a certain locality. The DGIS focus on open, distributed, task-centered services, which will broaden geographic information uses into an increasingly wide range of on-line geospatial applications. The implementation of DGIS can provide ubiquitous access for all different types of GIS applications, from environmental monitoring, emergency response, urban planning, spatial decision support systems, etc. We are studying some implementation and interoperability issues of DGIS systems [3].

Mobile gadgets, such as smart mobile phones, personal digital assistants (PDA), and pocket PCs are becoming widely available. These devices have constraints on their processing power, memory, batter power, communication capabilities, and screen size/resolution. There has been big demand on numerous applications, especially those with localized themes, such as Arabic language interface and related to Islamic issues. In these respect, we are encouraging our students to design and implement stand-alone or networked applications using various development tools that can run on various operating system and or device/platforms. We are gathering information on difficulties in developing and deploying such localized applications from technical, cultural, social, and commercial aspects.

In the theoretical research aspect, we are working in information theory field, especially on the construction and effective trellis decoding mechanisms for multi-dimensional error control codes. In this respect, an interactive visual debugging tool (VDT) has been designed and implemented for trellis related analysis of two-dimensional (2D) block codes [4]. The user interface provides number of useful interfaces for selecting code parameters, viewing encoded/received/decoded data sequences, introducing/analyzing errors at specific 2D code positions, running simulation, displaying results. The tool allows users to trace decoding process step-by-step and view detailed workouts at each node of the trellis. Since the VDT is now available, we are able to conduct numerous tests in constructing 2D codes and analyzing performances of resulting structures.

References:

1. Lin, X. et al: "Approximate processing of massive continuous quantile queries over high-speed data streams", IEEE Trans. on Knowledge and Data Engineering, Vol. 18, Issue 5, May 2006, pp. 683 – 698.

2. Yong W. et al: "A survey of security issues in wireless sensor networks", IEEE Communications Surveys & Tutorials Vol. 8, Issue 2, 2006, Page(s):2 – 23.

3. Jia Wenjue, et al: "GIS integration and interoperability based on GIS service chain", IEEE Int. Geoscience and Remote Sensing Symposium, 25-29 July 2005, pp. 4962 – 4965.

4. Kaya, L. and Al-Qudah, M. "A Visual Tool for Trellis Construction and Viterbi Decoder of Augmented RAC Array Codes", ACM Int. Conf on Wireless Communications & Mobile Computing Conference (IWCMC2006), Vancouver, Canada, July 3-6, 2006, pp. 1471-1475

Interactive On-Line Mathematics for Math Education

Philip Burton

College of Information Technology
Zayed University, Dubai Campus
Philip.Burton@zu.ac.ae

Recent developments in web technologies make it possible, for the first time, to write a mathematical formula or equation directly onto a web page, and to do it efficiently. The relevant web technology is the XML dialect MathML (Mathematical Markup Language) and, in particular, the ASCIIMath approach to it of Peter Jipsen (Chapman University, USA). At the same time, another XML dialect, SVG (Scalable Vector Graphics) is being adapted to display the graph of the equation. Some Math departments are already putting their math courses online using these technologies.

There are essentially two approaches: use open source (free) programs, or, alternatively, use proprietary solutions that can be bought off the shelf and require less individual technical knowledge. A good compromise between the two approaches seems to be to use the latest, and very inventive, open source methods to author mathematics for a web page and then to use a proprietary system to do the actual mathematics (such as algebraic manipulation, calculus, etc.). Any use of open source material requires some low-level implementation, that is, IT knowledge on the part of the instructor is required.

The effective use of MathML and SVG in web pages - either separately, or in conjunction with each other – can be a technical challenge, but examples do exist. The problems that occur with using MathML and SVG in synergy constitute a subset of the more general problem of making XML dialects work together on the same web page. The technical issues that must be confronted are concerned mainly with the use of XML namespaces and the differing behaviors of the various web browsers with respect to XML. Mixing XHTML, MathML and SVG in the same document requires a mechanism for validating the mixed-namespace document. Validation means checking the structure for conformity to a DTD or schema.

A working goal for such an interactive on-line math web page would be to allow the student to interact with every element, that is, mathematical expressions and their corresponding graphics, and be able to see the effect of changing one upon the other. In addition, by using an existing computational engine, like webMathematica or MapleNET, a student should be able to immediately evaluate (analytically or numerically) almost any math expression.

Moving the teaching and learning of mathematics from pen and paper to the computer is an essential development. This must happen soon to ensure that mathematics is seen as relevant by contemporary students as well as to allow the use of educational technology to the fullest extent possible. The purpose of this study is to examine the open source approaches to interactive on-line math and compare them with the proprietary approaches.