

---

---

## Table of Contents

Message from the ACM Mid-Southeast Chapter Chair .....	3
ACM Mid-Southeast Chapter Officers.....	4
ACM Mid-Southeast Fall 2006 Conference Committee .....	4
Roster of Judges .....	5
Roster of Session Chairs .....	6
Conference Program.....	7
Keynote Address .....	17
Abstracts	
Undergraduate Students (2-year Colleges) .....	19
Undergraduate Students (4-year Colleges) .....	25
Master's Degree Students .....	37
Ph.D. Students .....	41
Professional.....	53
Panels and Roundtable Discussions .....	89
Tutorial.....	93
Conference at a Glance .....	95

---

---



## Message from the Chapter Chair

On behalf of the Association for Computing Machinery (ACM) Mid-Southeast Chapter executives, I welcome you to this year's Fall Conference and to scenic Gatlinburg, Tennessee. The Mid-Southeast Chapter of ACM takes pride in being a part of the world's oldest, largest and most reknown educational and scientific computing society. We, the ACM Mid-Southeast Chapter, are proud to be part of this tradition. Formed in 1959, the ACM Mid-Southeast Chapter is among the oldest chapters of the ACM. The chapter covers the U.S. mid-southeast region that includes the states of Kentucky, Tennessee, Georgia, South Carolina, Alabama and Mississippi. Chapter membership has been steady at around 100 people over the years, making it small enough for everyone to be known by everyone else, but diverse enough to accommodate nearly all areas of computing. Our annual conferences normally draw slightly bigger crowds, but they are still small enough for invigorating one on one discussions.

In keeping with the tradition of the Chapter, this year's conference is once again an opportunity for collegiality, renewal of friendships, meeting new people, and making new friendships likely to last for years. This year's conference has a cross section of presentations in scholarship, pedagogy, roundtable and poolside discussions. In these presentations, we are covering something you know but want to know more of, something you do not know and have been wishing to know, something interesting that you have been dying to hear about, and innovative pedagogical techniques that you may want to try out in your classrooms. As usual, we encourage students to join us and present papers, engage in dialogue with other students from across the region, and interact with us. So please make a date with us designating our Chapter conferences as the annual rejuvenation of your and your students' academic life.

Although we take pride in being a small Chapter, we still welcome new members to strengthen us and bring new ideas. So if you are not currently a member, I encourage you to become one. There are two membership categories at minimum membership fees. The Professional category at \$10.00 per a year, and student membership at \$5.00 per year; please do join us as we continuously explore and share the joy of learning and teaching Computer Science and Information Technology.

Once again, welcome and enjoy the conference and beautiful Gatlinburg.

*Joseph M Kizza*, Chair  
University of Tennessee–Chattanooga

---

## Chapter Officers

### **Chair**

Joseph Kizza  
University of Tennessee–Chattanooga  
joseph-kizza@utc.edu

### **Vice Chair**

Otha Britton  
University of Tennessee–Martin  
britton@utm.edu

### **Secretary**

Brenda Parker  
Middle Tennessee State University  
csbrenda@mtsu.edu

### **Treasurer**

Emery Gathers  
University of Tennessee–Martin  
egathers@utm.edu

### **Webmaster**

Judy Hankins  
Middle Tennessee State University  
csjudy@mtsu.edu

## Conference Committee

### **Conference Chair**

Otha Britton  
University of Tennessee–Martin  
britton@utm.edu

### **Program Chair**

Jim Clark  
University of Tennessee–Martin  
jclark@utm.edu

### **Student Paper Competition**

Randy Smith  
University of Alabama  
rsmith@cs.ua.edu

### **Hospitality Suite**

Kathy Winters  
University of Tennessee–Chattanooga  
Kathy-winters@utc.edu

---

## Student Paper Competition Judges

### **Chair**

Randy Smith, University of Alabama

### **Undergraduate 2-year**

Rick Wilkerson, Dyersburg State Community College

Jeff Gray, University of Alabama-Birmingham

Ramana Gosukonda, Fort Valley State University

### **Undergraduate 4-year**

Suk Jai Seo, Middle Tennessee State University

Anant Honkan, Georgia Perimeter College

Nancy Smithfield, Austin Peay State University

### **Master's**

Melissa Wiggins, Mississippi College

Jiang Li, Austin Peay State University

Zhijiang Dong, Middle Tennessee State University

### **Ph.D.**

Glenn Wiggins, Mississippi College

Cen Li, Middle Tennessee State University

S. Krishnaprasad, Jacksonville State University

---

## Session Chairs

### **Azalea Room**

- Session I: Joyce Crowell, Belmont University
- Session II: Li Yang, University of Tennessee–Chattanooga
- Session III: Nancy Smithfield, Austin Peay State University
- Session IV: Sylvia Colvin, Tennessee Valley Authority
- Session V: Glenn Wiggins, Mississippi College

### **Dogwood I**

- Session I: Cen Li, Middle Tennessee State University
- Session II: Allan Anderson, Northeast State College
- Session III: Brenda Parker, Middle Tennessee State University
- Session IV: Melissa Wiggins, Mississippi College
- Session V: Roundtable discussion (unmoderated)

### **Dogwood II**

- Session I: Tesfa Haile, Jackson State University
- Session II: Jim Vandergriff, Austin Peay State University
- Session III: Jeff Gray, University of Alabama–Birmingham
- Session IV: Richard Barber, Mississippi State College
- Session V: Panel Discussion Continuation (unmoderated)

### **Magnolia Room**

- Session I: Bob Bradley, University of Tennessee–Martin
- Session II: Wayne Walters, University of Southern Mississippi
- Session III: Ashraful Chowdhury, Georgia Perimeter College
- Session IV: Denise Williams, University of Tennessee–Martin
- Session V: Workshop/Tutorial on Alice (unmoderated)

### **Alternates**

- Bill Janeway, Eastern Kentucky University
  - Ramana Gosukonda, Fort Valley State University
-

---

---

ACM Mid-Southeast chapter  
2006 Fall Conference  
Gatlinburg, Tennessee  
Glenstone Lodge

**Thursday, November 16, 2006**

4:00 – 6:00 p.m.	Registration
6:00 – 7:30 p.m.	Social Meeting, Hospitality Suite
7:30 – 9:00 p.m.	Dinner — Individual Arrangements
9:00 – 12:00 p.m.	Social Gathering, Hospitality Suite

**Friday, November 17, 2006**

7:30 – 9:00 a.m.	Registration
7:30 – 8:15 a.m.	Morning Coffee Sponsored by Course Technology
8:00 – 8:15 a.m.	Welcome/Announcements — Azalea Room
	<b>Welcome</b> Chapter Chair
	<b>Conference Announcements</b> Conference Chair
	<b>Program Announcements</b> Program Chair
8:15 – 9:15 a.m.	Keynote Address
9:15 – 9:30 a.m.	Coffee Break Sponsored by Course Technology

---

---

---

**Session I: 9:30 – 10:30 a.m.****Azalea Room: Undergraduate 4-year Papers**

Session Chair: Joyce Crowell, Belmont University

- 9:30 – 9:50 *An Adaptive Self-Parking Vehicle*, Amos Smith, University of Alabama–Birmingham
- 9:50 – 10:10 *Effective Methods for Learning: a Study in Visualization*, Ian Mitchell, Middle Tennessee State University, (Advisor, Ms. Brenda Parker)
- 10:10 – 10:30 *Design and Implementation of Parallel Baccarat Simulation*, Lorenzo Jackson, Columbus State University

**Dogwood I: Doctoral Papers**

Session Chair: Cen Li, Middle Tennessee State University

- 9:30 – 9:50 *Mining Spam Email for Cyber Forensic*, Hun Wei, University of Alabama–Birmingham
- 9:50 – 10:10 *Software Architectures for Network-Centric Systems*, Amine Chigani, Virginia Tech
- 10:10 – 10:30 *A Modification to Esko Ukkonen's Edit Distance Calculating Algorithm*, Bruce Johnson, University of Tennessee–Knoxville

**Dogwood II: Professional Papers**

Session Chair: Tesfa Haile, Jackson State University

- 9:30 – 9:50 *Custom Designed Student Response System*, Sung Yoo, Middle Tennessee State University
- 9:50 – 10:10 *Intern Experiences in a Technology Program*, B. Wayne Walters, University of Southern Mississippi
- 10:10 – 10:30 *Teaching Java as a First Programming Language with Pair Programming: a Case Study*, Kathy Winters, University of Tennessee–Chattanooga
-

---

**Magnolia:****Professional Papers**

Session Chair: Bob Bradley, University of Tennessee–Martin

9:30 – 9:50

*Using Computer Science to Explore Life’s Deep Questions*, Steve Donaldson, Samford University

9:50 – 10:10

*Monsters and Masters Speak: What Literary A.I. Characters Think of the Question, “Can Machines Think?”*, Beth Walker, University of Tennessee–Martin

10:10 – 10:30

*Grading Using Subversion*, Roland H. Untch, Middle Tennessee State University

**Session II:****10:35 – 12:00 p.m.****Azalea Room:****Undergraduate 4-year Papers**

Session Chair: Li Yang, University of Tennessee–Chattanooga

10:35 – 10:55

*Sidestepping the Chinese Room: Aspects of Self-Awareness in a Humanoid Robot*, Jason M. Gruber, Birmingham-Southern College

10:55 – 11:15

*Frequent Webpage Access Patterns*, Ashley Plier, The University of Alabama–Huntsville

11:15 – 11:35

*Generating Smooth Models Using Subdivision Surfaces*, Shayan Javed, University of Alabama–Birmingham

11:35 – 11:55

Open

---

**Dogwood I:****Doctoral Papers**

Session Chair: Allan Anderson, Northeast State College

10:35 – 10:55

*MARS: a Metamodel Recovery System Using Grammar Inference*, Faizan Javed, University of Alabama–Birmingham

10:55 – 11:15

*JMLMP: a Formal Specification Language for a Multiparadigm Programming Language*, Matthew Thornton, Virginia Polytechnic Institute and State University

11:15 – 11:35

*A QoS-Driven Software Product Line Engineering Framework for Component-Based Distributed Real-time and Embedded Systems*, Shih-Hsi Liu, University of Alabama–Birmingham

11:35 – 11:55

*An Exploration of the Practicality of Naive Clustering Algorithms*, Trevor Jay, University of Alabama

**Dogwood II:****Professional Papers**

Session Chair: Jim Vandergriff, Austin Peay State University

10:35 – 10:55

*A Genetic Algorithm Approach to Course Scheduling Optimization*, Brian Toone, Samford University

10:55 – 11:15

*A Data Mining Project Implementation Model*. Vernon L. McGlone, University of the Cumberlands and Teresa A. McGlone, Eastern Kentucky University

11:15 – 11:35

*Establishing an Interdisciplinary Bioinformatics Curriculum at Fort Valley State University*, Ramana Gosukonda, Masoud Nagedolfeizi and Sanjeev Arora, Fort Valley State University

11:35 – 11:55

*Parental Views of Females in Computer Science and Technology*, Dr. Suzanne Smith, Prof. Kellie Price, Prof. Michael Laws and Ms. Rebecca Loyd, East Tennessee State University

---

---

<b>Magnolia Room:</b>	<b>Professional Papers</b> Session Chair: Wayne Walters, University of Southern Mississippi
10:35 – 10:55	<i>Extra Credit Game 2.0</i> , Denise Williams, University of Tennessee–Martin
10:55 – 11:15	<i>Teaching Database via Service Learning</i> , Greg Kawell, Samford University
11:15 – 11:35	<i>Web-Based RGB Color Exercises</i> , James M. Bateman, Austin Peay State University
11:35 – 11:55	<i>Integrating Alice into K-12 Outreach Events</i> , Jeff Gray, University of Alabama–Birmingham

**Lunch** **Poolside 12:00 – 1:00 p.m.**

**Session III** **1:00 – 2:40 p.m.**

<b>Azalea Room:</b>	<b>Undergraduate 4-year Papers</b> Session Chair: Nancy Smithfield, Austin Peay State University
1:00 – 1:20	<i>Instrumenting Xen for Security Using Aspect-Oriented Programming</i> , Hampton W. Haddock, Jr., University of Alabama–Birmingham
1:20 – 1:40	<i>VEMI: an Interface for Simultaneously Viewing Data in Disparate Formats</i> , William Hendrix, Furman University
1:40 – 2:00	<i>Development of a Humanoid Robot</i> , Scott Kubina, Samford University
2:00 – 2:20	<i>Extraction of 3D Geometry and Texture from 2D Photographs</i> , Anthony Snyder, Furman University

---

**Dogwood I:****Doctoral – Professional Papers**

Session Chair: Brenda Parker, Middle Tennessee State University

1:00 – 1:20

*A Knowledge Based Engineering Paradigm for Software Component Integration*, Varadraj Prabhu Gurupur, University of Alabama–Birmingham

1:20 – 1:40

*A Sensitivity Analysis of Initial Conditions for the Mumford-Shah Based Level Set Method of Image Segmentation*, Bruce Johnson and Yongshen Pan, University of Tennessee–Knoxville

1:40 – 2:00

*A Comparison of Approaches for Record Matching: Naïve Bayes Versus Distance in Database Systems*, Yan Liang, University of Alabama

2:00 – 2:20

*Evaluating Claims about the Impact of RFID on Privacy*, Syed Raza, Talladega College (Professional Presentation)

**Dogwood II:****Undergraduate 2-year Papers**

Session Chair: Jeff Gray, University of Alabama–Birmingham

1:00 – 1:20

*RFID*, Jackie Grindstaff, Northeast State Technical Community College

1:20 – 1:40

*Multi-point Ball Thrower*, Ariana Ohana, Chris Phillips and Isaac Thomas, Georgia Perimeter College

1:40 – 2:00

*Network Security: Server Scanning*, Robert Phillips and Derek Smallwood, Northeast State Technical Community College

2:00 – 2:20

*Mixer O.f.E.lixer (M.O.E.)*, Rizwan Babwani, Kyle McGuinn and Jordan Rowe, Georgia Perimeter College

---

---

**Magnolia Room:****Professional Papers**

Session Chair: Ashraful Chowdhury, Georgia  
Perimeter College

1:00 – 1:20

*Assessment and Accreditation: ABET's New Student  
Outcome Requirements*, Donald Sanderson, East  
Tennessee State University

1:20 – 1:40

*Decreasing Enrollments/ Increasing Manpower Needs:  
Solutions for Attracting Students to Computer Science*, Wayne  
Summers and Rodrigo Obando, Columbus State  
University

1:40 – 2:00

*Analysis of Student Performance in Programming Subjects of  
an In-house Exit Exam*, Singli Garcia-Otera, Masoud  
Naghedolfeizi and Nabil Yousif, Fort Valley State  
University

2:00 – 2:20

*Managing Complexity in Ethical Analysis*, Jeff Roach, East  
Tennessee State University

**Break****Poolside 2:20 – 2:35 p.m.****Sponsored by Course Technology****Session IV****2:35 – 3:35 p.m.****Azalea Room:****Professional Papers**

Session Chair: Sylvia Colvin, Tennessee Valley  
Authority

2:35 – 2:55

*Extending the "Bagging" Ensemble Approach for Classifying  
Highly Skewed Data*, Cen Li, Middle Tennessee State  
University

2:55 – 3:15

*Cheap, Flexible, and Secure Remote File Sharing with Samba,  
OpenVPN, and Apache*, Edwin Rudolph and Adel  
Abunawass, University of West Georgia

---

- 3:15 – 3:35      *Programming Language Support for XML Encryption*,  
David Frazier, East Tennessee State University
- 3:35 – 3:55      *Speculative Execution and Value Prediction Support for  
Distributed Systems*, Angkul Kongmunvattana,  
Columbus State University
- Dogwood I:**                      **Undergraduate 2 – Master’s Degree Papers**  
Session Chair: Melissa Wiggins, Mississippi College
- 2:35 – 2:55      *Virtual Machines in the Classroom*, Jackie Grindstaff,  
Northeast State Technical Community College  
(Undergraduate 2-year Presentation)
- 2:55 – 3:15      *Competitive Data Mining: an Educational Tool*, Michael  
Wade, University of Tennessee–Chattanooga
- 3:15 – 3:35      *Trust-based Routing in Wireless Ad-hoc Networks*, Alma  
Cemerlic, The University of Tennessee–Chattanooga  
(Advisor, Dr. Li Yang)
- 3:35 – 3:55      *Use of leJOS to Extend the Communication Capabilities of the  
LEGO Mindstorms*, Jeremy O. Blair, University of  
South Alabama

- Dogwood II:**                      **Professional Papers**  
Session Chair: Richard Barber, Pellissippi State  
College
- 2:35 – 2:55      *Autonomic Computing: Vision, Characteristics and  
Applications*, Srinivasarao Krishnaprasad, Jacksonville  
State University
- 2:55 – 3:15      *Sun Microsystems vs. Microsoft: the Controversy about  
Delegates*, Ken R. Adcock, Jr., UPS Supply Chain  
Solutions
- 3:15 – 3:35      *Writing Your Own Operating System – Hints, Tools and  
Resources*, Bob Bradley, University of Tennessee–  
Martin
-

---

3:35 – 3:55      Panel: *OS X as a Unix Environment*, Jeremy Ey, Justin Stinson and Eric Brown, Tennessee Technological University

**Magnolia Room:      Professional Papers**

Session Chair: Denise Williams, University of Tennessee–Martin

2:35 – 2:55      *Lessons Learned: a Year of Alice*, Will Lloyd, University of West Georgia

2:55 – 3:15      *The Case for Alice*, T. F. Higginbotham, Ph. D., Southeastern Louisiana University

3:15 – 3:35      *Introducing Programming to Entry-Level Computer Classes with Carnegie Mellon's ALICE Visual Programming Environment*, Dr. David Brown, Pellissippi State Technical Community College

3:35 – 3:55      *Alice: Learning to Program by Creating Movies*, Brenda Parker, Middle Tennessee State University

**Session V:              4:00 – 5:00 p.m.**

**Azalea Room:          Professional Papers**

Session Chair: Glenn Wiggins, Mississippi College

4:00 – 4:20      *Introducing Quantum Computing and Communication to Computer Science Curriculum*, Tzusheng Pei, Jackson State University

4:20 – 4:40      *A First Course in Information Security*, Nancy Smithfield, Austin Peay State University

4:40 – 5:00      *System Security Concerns in Component-Based Software Design*, Kathy Winters, University of Tennessee–Chattanooga

---

- Dogwood I:**                    **Roundtable Discussion**  
*Podcasting and Other Tools for Online Learning*, David Frazier, Michael Laws, Carolyn Novak and Jeff Roach, East Tennessee State University
- Dogwood II:**                    **Panel Discussion (continued from Session IV)**
- Magnolia:**                    **Tutorial**  
*Alice – a Gentle Introduction to Object-Oriented Programming*, Wayne Summers, Columbus State University
- 5:00 – 5:30 p.m.**                **Business Meeting, Dogwood II**
- 5:30 – 7:00 p.m.**                **Social Gathering, Hospitality Suite**
- 7:00 – 9:00 p.m.**                **Awards Banquet, Azalea and Dogwood I**  
**Entertainment:** Misty River Bluegrass Band
- 9:00 p.m.**                        **Social Gathering, Hospitality Suite**

**A “thank you” goes to the The University of Alabama at Birmingham, The University of Tennessee at Knoxville and Course Technology for lending financial support to this year’s conference.**

---

## Keynote Address

### Security and Privacy in Computer Forensics Applications

Prof. S. Srinivasan

Center for Information Assurance, University of Louisville

Computers are used daily by many people as a communications tool. Consequently, many applications generate and store several files. On many occasions such files contain confidential data that the user may not want divulged where as an investigator might want to analyze to see if any crime was committed using that computer. The end user would normally expect reasonable security precautions to be in place for protecting sensitive data, especially when the computer is part of a network. The criminal user would want to hide the information on the computer whereas a Computer Forensics investigator would want to find out the hidden information for evidence. My goal in this presentation is to lay out the Computer Forensics capabilities and how security and privacy aspects are important in this context. I would spell out a set of 10 policies that could be considered as the basis for considering the privacy aspects in light of Computer Forensics capabilities.

---

## About the Speaker

Srinivasan (nickname Srimi) is a Professor of Computer Information Systems and Director, Center for Information Assurance at the University of Louisville. He joined U of L in 1987. Prior to that, he taught at Austin Peay State University in Clarksville, TN, and the University of Akron in Akron, OH. His research interests are in Information Security. He has published several papers in both Mathematics and Computer Science. He spent his most recent sabbatical year (2000 – 2001) at UPS in Louisville, working on a major database project, and he is getting ready for his next sabbatical (2007 – 2008). He is heading the InfoSec program development at the University of Louisville, which was recently designated a National Center of Academic Excellence by the National Security Agency and the Department of Homeland Security. Also, he is leading U of L's Gifted Student Summer Program, which attracts bright students to a three-week summer academic program at U of L. This program has attracted students from all parts of Kentucky, Tennessee, Indiana, Georgia, New York, Texas, and India. Currently he concentrates his teaching in Information Security and Databases. He volunteers his time extensively for public education causes.

---

Student Abstracts  
Undergraduate 2-year Colleges

---

---

# RFID

Jackie Grindstaff

Northeast State Technical Community College

My presentation on RFID will go over the following:

The pros and cons of using RFID

Active vs. Passive RFID

RFID defined

How RFID Works: How it is being used outside of the computer science field

Potential Uses: Some of the many different types of RFID are in use today, with and without our knowledge

Advantages vs. Disadvantages

Current Users/Suppliers

Varying costs of RFID

Frequency Tags: What types of tags are available

Regulations and Standards

The Future of RFID

---

## Multi-point Ball Thrower

Ariana Ohana, Chris Phillips and Isaac Thomas

Georgia Perimeter College

The students at Georgia Perimeter College's Computing and Engineering Club designed and produced a robot that combines the trajectory physics of tossing a ball into various containers, similar to a basketball shot. Although this is a simple action, it involves various calculations of the dynamics of trajectory motion and then programming the individual servos to perform a single motion for each task.

The robot team modified the LYNXMOTION INC.'s robotic arm kit and programmed its actions using JAVA programming, a common programming language. This project uses a combination of the knowledge attained from the various Computer Science, Mathematics, and Engineering courses offered by Georgia Perimeter College.

---

# Network Security: Server Scanning

Robert Phillips and Derek Smallwood

Northeast State Technical Community College

When you have a website hosted, there are several questions you have to ask yourself. Are your sites being hacked? Are your servers safe? If there are any problems, how would you fix them? What programs are available that will help you fix the problems?

There are several programs out there that will accomplish what you need to make your site secure. One of the programs that is available is Nessus. Nessus can be executed on a Windows platform or a LINUX platform. The Nessus for Windows does cost you some money, but the LINUX version is freeware.

Nessus has been built so it can easily be scaled down to a single computer or to multiple computers with several processors. The more power you give Nessus the quicker it will scan the network. Each of the security scan that Nessus performs is written in as an external plug-in and written in NASL (Nessus Attack Scripting Language). Updating Nessus does not involve downloading untrusted binaries, which is better security for the program.

---

## Mixer O.f E.lixir (M.O.E.)

Rizwan Babwani, Kyle McGuinn and Jordan Rowe

Georgia Perimeter College

The students at Georgia Perimeter College's Engineering and Computer Science Club designed and produced a robot that performs the tasks similar to a bartender, mixing and serving various drinks. The robot team modified the LYNXMOTION INC.'s robotic arm kit and programmed its actions using JAVA programming, a common programming language. This project uses a combination of the knowledge attained from the various Computer Science, Mathematics, and Engineering courses offered to Georgia Perimeter College students. The implementation of robotics in the bartending and mixing of drinks by computer programming is applicable to various tasks in the real world, and it is an exciting aspect in the presentation of the capabilities of a simple robot. In real world applications, this robot could be modified to perform various utilitarian tasks useful in assembly lines, operations, and many other functions that use simple user input.

---

# Virtual Machines in the Classroom

Jackie Grindstaff

Northeast State Technical Community College

How and where VM are being used?

Why VM over just textbook and or paper labs?

How using VMs in the classroom help in the "real world."

Students being taught through VM vs. Students being taught without using VM

Virtual Machines are beginning to take hold of computer science classrooms.

Who is using this to teach students about real life situations, and how is this effecting the workplace?

A few questions that will be answered in this presentation.

---

Student Abstracts  
Undergraduate 4-year Colleges

---

---

# An Adaptive Self-Parking Vehicle

Amos Smith

University of Alabama–Birmingham

Many automobile manufacturers are integrating software into their vehicles in order to make the driving experience easier, safer, and more efficient. Examples of embedded automotive software include anti-lock braking and navigational control through GPS. These capabilities enable the motorist to rely less on outside resources such as paper maps, while having a more convenient and safer information source. A recent feature added by BMW is a car that can self-park. This capability allows the driver to park a car in a garage, whether or not the driver is in the car. The car uses a camera mounted on the windshield to measure the distance from a reflective lens that must be installed on the wall in front of the parking space. This form of autonomous parking has a few downsides, including the fact that the car can only park in a space with the lens. Furthermore, the BMW car cannot parallel park and is limited to moving forward into a fixed parking garage.

The main difficulty in obtaining a flexible autonomous parking capability is that the vehicle must be able to park in many various environments and configurations, where additional environmental installations such as the aforementioned lens are not available. This presentation will describe an investigation that resulted in a LEGO NXT robot that can perform autonomous parking in a parallel parking environment, without needing any specific environmental installations. The robot vehicle that will be presented uses various sensors to determine a valid parking space, and then navigates to parallel park into this space. This project serves as an initial investigation as an alternative solution to manual parallel parking, which is considered to be the most difficult form of parking by many motorists.

---

# Effective Methods for Learning: a Study in Visualization

Ian Mitchell

Middle Tennessee State University  
(Advisor, Ms. Brenda Parker)

The use of visualization techniques has recently been revitalized in the areas of science, technology, engineering and mathematics (STEM) education. In an effort to improve retention and reduce failure rates at our university a recent research study was performed to determine the impact of visualization techniques throughout our computer science curriculum. To what extent should a data structures and algorithms course, for example, use visualization techniques to improve student interest and comprehension? Which courses in our curriculum will benefit most with the use of visualization techniques? What visualization tools should be used? At what level should visualization techniques be emphasized? This paper presents a brief description of our findings related to these issues and concludes with information related to how our students learn.

---

# Design and Implementation of Parallel Baccarat Simulation

Lorenzo Jackson

Columbus State University

Baccarat is an ancient card game that can be traced back to the Middle Ages. A current version of this game requires a player to predict one of the three possible outcomes from a series of card draws. To gather statistical numbers on the chances of winning, I designed and implemented a program that simulates the game according to the rules used by casinos. The program is designed to play the game under every possible combination of card draws from different shuffling sequences. For example, one of the combinations may produce a “A, 2, 3, ..., K” shuffling sequence, whereas another combination produces a “2, 3, 4, ..., A” sequence. The number of possible combination increases drastically when the number of card decks increases. To speedup the execution time of the simulation, I ported a sequential version of my Baccarat simulation program to a middleware supporting parallel computing on a network of workstations called TreadMarks. This middleware provides a shared memory abstract to application programs, and therefore, allows me to develop a parallel version of my simulation program on a network of workstations with ease.

---

# Sidestepping the Chinese Room: Aspects of Self-Awareness in a Humanoid Robot

Jason M. Gruber

Birmingham-Southern College

Self-awareness is a sub-discipline of artificial intelligence that has received some theoretical attention, but few attempts at practical implementation. We present two features implemented in a humanoid robot that are intended as preliminary steps towards self-awareness. The primary feature is a self-awareness oriented exploration of self-recognition. Self-recognition is integral to self-awareness because it allows determination of that which is self and that which is not self. Our approach uses motion in the field of vision as a trigger to save a self image, which can later be compared to other images using a color histogram, pixel counts, and shape context descriptors. In addition, the robot is able to estimate the size of objects based on a previously implemented distance estimation algorithm. After explanations and preliminary results of these features, we close by discussing future extensions of the robot's functionality and the impact of existing features on the study of self-awareness in general.

---

# Finding Frequent Webpage Access Patterns

Ashley Plier and Jennifer Cuzzort

University of Alabama–Huntsville

In this project, we are investigating frequent webpage access patterns. We know each webpage has a unique URL. Therefore, the set of all URLs over a pre-specified domain can be viewed as a finite set from the alphabet  $\Sigma$ . When users access webpages within this domain, there are two scenarios: in the first scenario, users never use the Back button of their browsers. Then each internet access corresponds to a string over  $\Sigma$ . Different accesses correspond to different strings over the alphabet  $\Sigma$ . The history of all user webpage accesses can be considered a sequential database of such strings. In the second scenario, users do use the Back button. Then each internet access corresponds to a tree. Similarly, the history of all the webpage accesses can be viewed as a tree database.

Given a database of such an access history, we want to determine the webpage access patterns: that is, the most frequently accessed webpages, and how they are accessed together. These patterns can be used to analyze the utility and interrelatedness of these webpages. Using the above sequential database or tree database representing the webpage access history, what we really need to find are frequent substring or subtree patterns over the sequential database or the tree database. Our project goal is to implement algorithms which find all such frequent access patterns from the given databases.

We divided this project into two phases. In the first phase, we will implement an algorithm mining the frequent substrings over a sequential database. In the second phase, first we will use an encoding scheme to encode trees into strings. Then we apply our algorithm from the first phase to find frequent substring candidates. Finally we implement an algorithm to check whether these candidate strings indeed correspond to subtrees of the given tree databases.

---

# Generating Smooth Models Using Subdivision Surfaces

Shayan Javed and John K. Johnstone

University of Alabama–Birmingham

Methods for generating smooth models from subdivision surfaces are discussed. Subdivision surfaces have recently become popular as a modeling primitive in computer graphics. They are built from polyhedral meshes using repeated subdivision. They are the limit surface after an infinite number of subdivisions. Since the subdivision process for a point can be captured by a matrix multiplication, a limit point can be computed using an eigenvector of this subdivision matrix. Indeed, derivatives at the limit point can also be calculated using other eigenvectors. Each subdivision step involves a splitting step and an averaging step. In the splitting step, new vertices are introduced exactly in the middle of the edges, while the averaging step repositions the vertices using weighted averages of vertices around that vertex. We concentrate on the Loop subdivision scheme for triangular meshes, although other subdivision schemes are also considered. The mesh is stored using a winged edge data structure for efficient manipulation. A well-known issue with subdivision surfaces is the presence of extraordinary points, where continuity breaks down. These extraordinary points are evaluated and visualized, to determine whether they overly limit the subdivision surface as a smooth surface model.

---

Withdrawn

---

---

# Instrumenting Xen for Security Using Aspect-Oriented Programming

Hampton W. Haddock, Jr.

University of Alabama–Birmingham

Virtual Machine Monitors (VMMs) can provide a trusted computing base (TCB) to host production operating systems on recent hardware technology. The VMM can isolate the guest operating system from other instances running on the same hardware platform while sharing memory, CPU, I/O devices, etc. The objective of this research is to provide a secure computing platform using the VMM technology. Issues that are explored are defending the VMM code from attack and detecting intrusions and compromises of the guest operating systems. This work uses the Xen version 3.x hypervisor and takes advantage of the ring -1 technology provided by Intel in their VT feature of the Woodbridge Pentium processor. AspectC is used to instrument the Xen source code to provide logging and other monitoring capability. The project investigates Linux based systems and Microsoft operating systems.

This research uses Aspect Oriented Programming (AOP) with AspectC to provide several contributions to computer security. The research identifies cross-cutting concerns and uses AspectC to weave functionality into the Xen source code. The research identifies and implements logging functionality at target points in the Xen source code using AspectC. It also implements security monitoring at selected target points in the Xen source code using AspectC. The extensions to the Xen source code provides for independent development and maintenance because of the high cohesion and low coupling of AOP.

---

# VEMI: an Interface for Simultaneously Viewing Data in Disparate Formats

William Hendrix

Furman University

Scientific research often involves an analysis of data pertaining to some phenomenon, and many software packages have been developed for graphically displaying numerical data in an intuitive way. However, in order for inferences to be drawn about relationships between different data sets, more and more complex software has to be written to accommodate the display of data sets with different values, scales and formats. As an alternative to this, the Visualization Environment Manipulation Interface (VEMI) defines a standard interface whereby several smaller, simpler visualization environments can be manipulated simultaneously to reproduce important relationships between different data sets.

One of the major design issues with controlling multiple graphs simultaneously is deciding what sort of unifying features of graphs should be exploited to manipulate any kind of graph that a user might wish to display. As scientific data generally result from observing physical phenomena that are responding to some (known or unknown) stimulus, time is an integral aspect of the experimental results and should be reflected in the collected data. As such, time can serve as a universal control for which data are displayed, once the user specifies how time values are stored in a particular data set.

Ultimately, VEMI creates a truly flexible framework, where data of any variety may be assembled together in a single graph that facilitates the ability to form generalizations and draw conclusions. The greatest asset of VEMI is its extensibility; its independence from the underlying representation of data means that any data can be displayed, so long as a method for displaying the data is written according to its specifications.

---

# Development of a Humanoid Robot

Scott Kubina

Samford University

With robotics becoming a more integral cog in modern Computer Science, an increasing number of scholars and enthusiasts are looking to get their hands in the action. This talk presents the journey spent researching, designing, and implementing a primitive humanoid robot with functions such as reaching and motion tracking, and inform on the issues that can be expected when going through the process. Topics will include trials and tribulations of the mechanical design, computer vision in robotics, spatial awareness, the role of artificial neural networks in reaching for objects, forward kinematics, and real-world environments. The goal of the presentation is to inform which approaches did work in our project and which did not, and then to describe the triumph or downfall of the method. Information will be pooled from other projects dealing with the same topics to supplement the work.

---

## Extraction of 3D Geometry and Texture from 2D Photographs

Anthony Snyder

Furman University

Anyone who has ever attempted to create a realistic three-dimensional scene knows how difficult and time-consuming the process can be. A program that automatically generates a three-dimensional scene from a two-dimensional digital photograph is a 3D designer's dream. An extensive library of software has been developed in an attempt to move one step closer to achieving that dream. This software architecture created, consisting of numerous concepts, is tied together through a hierarchical tree structure concept, the nodes of which serve as a 3D scene graph hierarchy as well as a hierarchy of 2D images. The nodes of this tree structure also allow for the application of various grammars or "strategies" at different levels of the hierarchy. One such strategy is a concept that allows for placement, orientation, and projection of a set of orthogonal axes over an image. Subdivision of the image is then done through the use of a transparent 3D selection box, the face orientations of which are based on the orientation of the axes. Also provided is an extraction concept that generates dual Green's Function recognizers based on user-selected regions representing the edges of an object in the image. These recognizers, given an initial direction, can move through a selected sub-image tracing an outline of the object using a variation of Marr-Hildreth edge detection. The result is a set of points and scale values that are used along with a user-defined cross section to generate a VRML-like extrusion object. The spline's orientation in 3D space is determined by the user-defined axes mentioned above. GUI-based interaction allows for the application of interactive strategies to the tree nodes, for manipulation of the hierarchy, and for the editing, serialization, and reloading of image, extrusion, and transform data. A depth-first traversal of the tree is used to construct the 3D scene.

---

Student Abstracts  
Master's Degree Programs

---

---

# Competitive Data Mining: an Educational Tool

Michael Wade

University of Tennessee–Chattanooga

An introductory course in data mining teaches a variety of problem solving techniques; however, the real-world application of these techniques eludes traditional class projects and must be learned through hands-on experience. An educational tool that can be used to facilitate this experience is an academic competition in data mining, such as the KDD Cup. The KDD Cup adds to the educational value of an introductory course by providing a chance to apply newly learned techniques in combination with real-world data. More specifically, the KDD Cup is an academically fueled competition to build a classifier capable of meeting or exceeding problem specific requirements. Therefore, participants are involved in every aspect of data mining development which includes: data cleansing, statistical analysis, analytical problem solving, and algorithm development.

One metric of success for this competition are the advancements to the field of data mining that have been procured by participant research. Prima facie, the KDD Cup appears a flawless addition to an introductory education in data mining; in reality, there exist some negative aspects as well. One aspect that is unfavorable to the competition, as an educational tool, is that the final test data is not released in its entirety after the competition is complete. This prohibits a participant from analyzing the final test set for possible mistakes in their methodologies and competition entry – the submitted classifier. While the competition is not flawless, this paper evaluates the value of such competitive entries as an educational tool by documenting the experience of competing in this year's KDD Cup challenge.

---

# Trust-based Routing in Wireless Ad-hoc Networks

Alma Cemerlic and Dr. Li Yang

The University of Tennessee–Chattanooga

Mobile Ad-hoc Networks (MANETs) are extremely helpful in supporting and forming an instant network when no fixed infrastructure is available. MANETs can find applications in a variety of areas like emergency assistance, inter-vehicle communications. Most developed wireless ad hoc routing protocols are designed to discover and maintain an active path from source to destination with an assumption that every node is friendly. However, it is possible that the participating nodes are selfish or malicious. A mechanism to evaluate reputation and trust for each node is essential for the reliability of routing protocol in MANETs.

We integrate reputation and trust management into routing protocols in MANETs. Reputation and trust mechanism is based on constantly monitored and updated first-hand information and second-hand information. The nodes within the network are able to monitor their neighbors and obtain first-hand information based on the perceived behavior. Second-hand information is obtained from the sharing of first-hand information with other nodes. The nodes thus create total trust value by combination first-hand information and second-hand information. The total trust value is then available by all the nodes for routing decisions. Dynamic Source Routing Protocol (DSR) is selected to explore the possibility and benefits resulted from the integration of a reputation and trust management into routing protocol. Trust-based routing is designed to improve reliability in both route discover and maintenance in MANETs.

---

# Use of leJOS to Extend the Communication Capabilities of the LEGO Mindstorms

Jeremy O. Blair

University of South Alabama

Communication is a key element necessary for human cooperation. The same is true in a robotic environment. While humans can communicate with relative ease, robotic communication presents a more difficult problem. Various areas of computing (networks, data communications, robotics) have addressed the issues of distributed communication but often at a level of greater complexity. Often the robots used are expensive to create and utilize complex, high level communication protocols. Our research considered the use of a simpler communication protocol with less expensive robots. The LEGO MindStorms robots are complete robotic kits that contain a variety of sensors and inferred communication capabilities. Combined with the leJOS Java environment, the robots can be programmed to accomplish bidirectional communication between two robots and between a robot and computer. The straightforward platform allows for relatively simple communication as compared to that of more complex systems while maintaining the same benefits. Using the Mindstorms, an environment was created to establish both synchronous and asynchronous communication. The robots were also able to directly address each other or a computer independently while maintaining the ability to send a broadcast message. Like all communication mediums, efficiency still depends on a variety of factors. Specific to inferred communication, those factors include battery life, ultraviolet radiation, distance, line-of-sight relation, and message protocol. Once established, this intra bot communication can be used to coordinate the actions of multiple robots to accomplish a task. The communication protocols have served a variety of purposes including other summer research projects and course demonstrations. In this way, students at all levels of the curriculum have the opportunity to use a known language (Java) in a new and exciting way while considering advanced concepts of the curriculum.

---

Student Abstracts  
Doctoral Degree Programs

---

---

# Mining Spam Email for Cyber Forensics

Hun Wei

University of Alabama–Birmingham

Spam email has become a daunting problem as it affects almost everyone who uses email as an important method of communication. Most research on spam email focuses on differentiating spam from regular emails. However, little work has been done on analyzing the spam to reveal useful information about their creators. So far, there are millions of spam emails collected each day but not sufficiently analyzed. Traditional text mining techniques may not work ideally on spam emails because most spam emails apply some obfuscation techniques to avoid being detected by spam filters. Our goal is to build an automated system that will help humans to extract valuable information from those emails. To achieve this, the first step is to parse the spam emails and extract useful contents, such as sender's IP, hyperlinks, and date, and store them in a relational database that supports sophisticated queries. Then a survey of the data will reveal interesting attributes that can be used to categorize the emails by similarities. And data mining techniques can then be applied to the data to find patterns, such as spam source, ways of obfuscation, and eventually the spammers who author the emails. We hope that the research result will greatly benefit cyber crime investigators in tracing and hunting down the spammers.

---

# Software Architectures for Network-Centric Systems

Amine Chigani

Virginia Tech

In recent years, software development has evolved from the traditional platform-centric approach to a network-centric one. This evolution has reached its peak with the introduction of Network-Centric Operations (NCO) by the US Department of Defense. It has also resulted in conflicting interpretations of how to apply NCO concepts into the field of software architecture. Our work focuses on capturing the core concepts and goals of NCO, investigating the implications of these concepts and goals on software architecture, and providing an architectural style based on which network-centric software systems (NCS) can be developed.

In our preliminary research, we have identified at least four characteristics that distinguish network-centric systems from other systems. A network-centric system:

- Is a “system of systems” as opposed to being a “system”
- Has an underlying networked configuration that embodies the runtime environment on which the system’s components interact
- Has a dynamic runtime behavior, which means that the system’s actual interacting components are not necessarily known until runtime
- Has a fluid, dynamically-defined central control, which means that control over the system’s functionality is not necessarily owned by a particular component; rather, this control changes based on which function the system is performing and which component has initiated the system’s execution

In addition to these characteristics, we have identified several challenges that our proposed solution must address to achieve the goals of network centrality. These challenges include scalability, security, connectivity, interface standardization, testing, integration, and others.

In this presentation, we detail the preliminary results of our research focusing on the progress of our findings and the direction we intend for the remaining research work.

---

# A Modification to Esko Ukkonen's Edit Distance Calculating Algorithm

Bruce Johnson

University of Tennessee–Knoxville

Edit distance measures similarity between two strings (as the minimum number of change, insert or delete operations that transform one string into the other). An edit sequence  $s$  is a sequence of such operations and can be used to represent the string resulting from applying  $s$  to a reference string. We present a modification to Esko Ukkonen's edit distance algorithm based on representing strings by edit sequences. We conclude by demonstrating how much the performance of Esko Ukkonen's original algorithm can be improved after making our modification.

---

# MARS: a Metamodel Recovery System Using Grammar Inference

Faizan Javed

University of Alabama–Birmingham

Abstract: Domain-specific modeling (DSM) is a resource-efficient and expeditious alternative to the traditional software development process which traditionally involves a series of mappings from the domain idea, to the domain models and the source code. DSM enables subject matter experts to describe the essential characteristics of a problem at the same level of abstraction with the domain itself. The conventional approach to DSM involves creating a metamodel for a specific domain, from which instances pertaining to specific configurations of that domain can be constructed. From the defined models, software artifacts like design documentation and source code can be generated. However, as the metamodel undergoes evolutionary changes, repositories of instance models can become orphaned from their defining metamodel. This can result in instance models, which can contain important domain knowledge, failing to load into the modeling tool due to version changes that have occurred to the metamodel. In the model-driven software engineering realm, this problem highlights the need to have the capacity to recover the design knowledge in a repository of legacy models.

This presentation will discuss the Metamodel Recovery System (MARS), a semi-automatic inference-based system for recovering a metamodel that correctly defines the mined instance models through application of grammar inference techniques. MARS leverages the fact that a correspondence exists between the domain models that can be instantiated from a metamodel, and the set of programs that can be described by a grammar. The main gist of the project approach involves the use of an intermediate representation of the models in a textual language. A host of technologies ranging from XML transformation tools, to language development environments are used to facilitate the task. MARS has been successfully applied to various diverse domains with satisfactory results.

---

# JMLMP: a Formal Specification Language for a Multiparadigm Programming Language

Matthew Thornton

Virginia Polytechnic Institute and State University

Multiparadigm programming languages that incorporate imperative, object-oriented, logical, and functional programming paradigms liberate the developer from thinking about the solution to a problem in any particular mindset. Formal specification languages provide a method for mathematically specifying the behavior of a program, thereby eliminating the ambiguity of a natural language specification. Coupling a formal specification language with a multiparadigm programming language will give the developer the freedom of creativity to implement a particular program while, at the same time, giving him a concrete specification for what they are implementing. Unfortunately, there are no formal specification languages that provide mechanisms for modeling the structures of a robust multiparadigm programming language.

JavaMP is a multiparadigm programming language that is an extension to Java. It includes multiparadigm extensions to the Java programming language including the introduction of logical relations, a pass-by-name parameter passing mechanism (which facilitates returning values from calls to logical relations), lambda functions (including the ability to pass functions, return functions from functions, assign variables function values, and generally create functions on the fly.), and global functions and operators. The current implementation of JavaMP translates JavaMP code directly into Java code. Syntactically, the language is very close to native Java code and provides a good example of a multiparadigm programming language.

This presentation introduces JMLMP. It is an extension to the formal specification language for Java called the Java Modeling Language. A description of the language extensions and the syntax is provided, as well as a discussion as to where this research is headed.

---

# A QoS-Driven Software Product Line Engineering Framework for Component-Based Distributed Real-time and Embedded Systems

Shih-Hsi Liu

University of Alabama–Birmingham

Software complexity may be reduced and productivity may be increased by the synergy of Component Based Software Engineering (CBSE) and Software Product Line Engineering (SPLE). The synergistic techniques decrease complexity by uplifting software artifacts to a higher abstraction level, namely the component level, and defining the interfaces, interactions, and contexts of such artifacts for composing software systems. Such synergistic techniques also facilitate productivity increase by promoting feature reusability, leveraging component replacement and offering selectivity among product variants. When applied to Distributed Real-time and Embedded (DRE) systems, however, these technologies must fulfill such systems' time-critical missions and numerous functional and Quality of Service (QoS) requirements. New critical challenges to be solved are QoS sensitivity that influences functionality validity and performance quality and tangled requirements that increase requirements evaluation complexity.

The approach presented here introduces a QoS-driven SPLE framework, called QoSPL, for the construction of a set of DRE systems in the analysis and design workflows. To solve such challenges, the set of DRE systems are analyzed and designed in terms of a collection of QoS paths, each of which individually determines how well the services perform along the paths and as a whole represents a behavioral view of software architecture. QoSPL is expected to preserve the advantages of CBSE and SPLE, offer a less subjective and manageable QoS evaluation, reduce the evaluation complexity from the problem of tangled functional and QoS requirements, and avoid unnecessary construction workload by eliminating abundant infeasible design alternatives in the early software lifecycle. The adopted formalisms may also guarantee a high-confidence DRE product line construction.

---

# An Exploration of the Practicality of Naive Clustering Algorithms

Trevor Jay

University of Alabama

In the area of data mining it is common practice for performance discussions to center on the running time or computational performance of clustering algorithms. The algorithm is often applied in a fairly contrived situation. A common example has the algorithm being applied to a collection of documents containing "obvious" or "natural" clusters that vastly differ in some obvious way such as some being papers from medical research and others being software engineering journal articles.

This practice is understandable considering data mining is usually only brought to bear when the document corpus is far too large for human classification to be practical. It is extremely difficult to find appropriately accessible (for instance digital) document collections that are both large enough to well exercise the clustering algorithm and are understood enough to give a good gauge of an algorithm's practical performance. As a result much about clustering techniques suitability to the very application they were designed for is unknown.

We try to help emphasize the practical side of clustering technique evaluation. We examine common clustering techniques in the context of two large datasets that we have had the good fortune of having been clustered by human beings at an extremely small degree of granularity.

The first is a collection of some 600 NASA engineering documents formally categorized and filed by the engineers themselves. The second is a series of movie star ratings from some 480,000 Netflix users available as part of Netflix's Netflixprize initiative. We compare the clusterings derived from this second dataset to the human created categorizations found on the Internet Movie Database. Throughout we attempt to be naive in our implementations of clustering techniques.

Doing so we hope to get an insight into what is the base level practical performance of these techniques as compared to man-made clusterings.

---

# A Knowledge Based Engineering Paradigm for Software Component Integration

Varadraj Prabhu Gurupur

University of Alabama–Birmingham

All along the history of software engineering, traditional software development process has always been a labor intensive process. Software component integration has succeeded in solving this problem to a certain extent. The utilization of a Knowledge Based Engineering approach for software component integration may be a promising solution for this problem. Knowledge Based Engineering approach has provided well engineered solutions in the area of Systems Engineering. This approach could also be applied in the context of software systems. In order to achieve some level of automation in designing software systems various methodologies such as Design Structure Matrix, Markov Decision Process model and Markov Chain can be used for the process of automating the design process of a software system. Apart from the above mentioned methodologies SysML a new upcoming language recently approved by Object Management Group (OMG) is used for designing the architecture of a software system. By using SysML we now have a new approach for representing knowledge in the form of objects. Tools such as Telelogic TAU can be used for this purpose. The knowledge stored in the form of UML/SysML can be easily interpreted by using Telelogic TAU. We believe this approach will enable a considerable amount of reduction in the utilization of both cost and time involved in the development of software systems.

---

# A Sensitivity Analysis of Initial Conditions for the Mumford-Shah Based Level Set Method of Image Segmentation

Bruce Johnson and Yongshen Pan

University of Tennessee–Knoxville

The level set method provides a means of segmenting images. Fundamentally, the level set method for image segmentation is a search algorithm that determines where an evolving curve's boundary pixels – who are meant to encompass an image segment's perimeter – should be placed according to some criteria. A method has been devised that utilizes the Mumford-Shah functional as a means of establishing that criteria. It has been shown that this method for image segmentation has limitations that, while mentioned in the original research, were not quantified with examples. We present evidence of these limitations and discuss how they occurred. We conclude by offering a direction where this research might lead. This work is part of an ongoing research project.

---

# A Comparison of Approaches for Record Matching: Naïve Bayes Versus Distance in Database Systems

Yan Liang

University of Alabama

The interconnectedness of database system is constantly expanded. This interconnectedness places strong demands on the ability of database system to determine if two or more records refer to the same real-world entity, which is usually called record match or record linkage. Distance-based algorithms and probability-based algorithms are different ways to formulate the record matching problem. This paper makes an experimental comparison of a combined-distance algorithm and a Naïve Bayesian Classifier in supervised learning. The Naïve Bayesian Classifier is a kind of class probability estimation which encodes a set of conditional independence assumptions. To evaluate the performance of those algorithms, three measures, precision, recall and f-measure are utilized in this paper. Our study shows that for the given dataset, the Naïve Bayesian Model can achieve a higher f-measure on average than the combined-distance algorithm. For our dataset, precision and recall are not to be negatively affected. We also discuss the key parameters in algorithm implementations which have a key impact on the results. Accordingly, we present and analyze the experiment results from a subset of the most critical parameters.

---



## Professional Abstracts



# Custom Designed Student Response System

Sung Yoo

Middle Tennessee State University

A Student Response System (SRS) allows a teacher to ask a question and collect student responses in a classroom. Typically, the collected responses are analyzed to assess student understanding on a specific question. If a teacher wants to monitor the level of student understanding in a concept that is covered in a set of questions, it requires a manual analysis on collected responses from the corresponding questions. An intelligent system that automatically collects and analyzes student response may assist teachers to assess student understanding on a specific concept(s).

We have developed closed labs for CS-I and CS-II in the form of a web-based tutoring system, called AtoL (Adapted Tutor for On-line Learning). The AtoL system contains two major learning components for the students, a question tutor and a program tutor. The question tutor is designed to help a student learn a concept by solving a series of questions. AtoL was proven to be effective in teaching/learning in CS-I and CS-II.

The question sets used in AtoL are stored in the central database server. However, these questions cannot be utilized in classroom environment since most students in classroom do not have laptop computers. Since majority of the questions in the database are in multiple choices or true/false formats, we could use a cheaper SRS to utilize the existing question database in a classroom setting. We have chosen H-ITT clicker because its freely available software development kit (SDK) enables us to develop a custom designed SRS.

We have developed a SRS that utilizes questions from the AtoL database. A teacher may browse a list of questions available in a selected topic and present a question to a class using a classroom projection system. When a student with pre-assigned H-ITT clicker presses a button, our system collects the student response and records this activity in the AtoL database server. The SRS allows the teacher to engage a communication with whole students in a classroom simultaneously. Research and practice have shown this form of interactive lecturing to be successful. The SRS was proven to increase the student's attendance and provides immediate feedback on student learning. The custom designed SRS has enabled us to utilize the resources of the AtoL system in classroom. This has also enabled us to use the student progress monitoring tool in AtoL.

---

# Intern Experiences in a Technology Program

B. Wayne Walters

University of Southern Mississippi

Experience has shown that the effectiveness of internships can vary dramatically. An internship can provide excellent work experiences or on occasions may be less than satisfactory. It is quite rewarding when an internship comes along that provide an excellent opportunity for learning, leadership, and contributions to mankind. Just such an occasion occurred soon after Hurricane Katrina struck the gulf coast on August 29, 2005.

In response to the desperate need for support services, a volunteer organization known as Radio Response was organized by Mac Dearman, a wireless internet service provider (WISP), in Rayville, LA. Radio Response provided communication services to the center of the storm's landfall, Waveland and Bay St. Louis, MS. The call for help brought some of the most skilled communications specialists in the country to help rapidly build a WISP network that was strung from water towers and trees. Likewise, his call for equipment was responded to by many companies around the country.

Two very special University of Southern Mississippi students, Matt Justice and Peyton Gwinn fulfilled their University of Southern Mississippi's (USM) School of Computing internship requirements by assisting with the implementation of the WISP. By mid October, 2005, the volunteers from around the country had left and Matt Justice, a USM senior, was left in charge of managing and maintaining a vital system for the citizens of the effected area. Later, Peyton Gwinn stepped up to continue Matt's work. This opportunity has provided the interns with real-world experience and responsibility for managing a vital wireless network.

---

# Teaching Java as a First Programming Language with Pair Programming: a Case Study

Kathy Winters

University of Tennessee–Chattanooga

At the University of Tennessee at Chattanooga, we use Java as a first programming language. Like most other schools, one of the biggest problems we face is that of retention. Three interesting things happened this summer which lead me to use paired programming in our introductory Java course. First, I had a student tell me how frustrated he was with his programming attempts in lab. He understood the design aspects of the labs but the syntax was driving him nuts. Secondly, I attended a conference at which there was a poster session on using paired programming in an introductory course. Thirdly, I read the ACM article discussing how this technique was being used to successfully increase retention rates. I decided to give it a try. This presentation will discuss the details of that attempt.

---

# Using Computer Science to Explore Life's Deep Questions

Steve Donaldson

Samford University

Human existence appears to be shaped by a fascinating interaction of deterministic and random forces, but this very interplay presents a number of dilemmas for inquisitive individuals. For example: How can random processes produce meaningful results? In the large space of possibilities, how do significant creatures and behaviors arise? Can concepts of free will be reconciled with either determinism or randomness? To what extent can (deterministic) automata exhibit intelligent behavior? Is creativity primarily an illusion? Is it an inescapable consequence of human existence? How can structurally based operations (such as language and other aspects of intelligence) convey meaning? Are there universal principles of self-organization? What are the ramifications? This presentation describes how a variety of Computer Science techniques and demonstrations can be used to explore answers to questions such as these involving issues of chance and determinism and supports the claim that Computer Science can serve as a special window through which to view and understand the world as a collection of computational mechanisms operating on common underlying principles. Special attention is given to a previously-taught interdisciplinary honors seminar entitled (with apologies to Einstein) "Does God Play Dice with the World? - Explorations in Complexity, Creativity, Determinism, Randomness, Self-organization, Emergence, and Free-will." This course, team-taught with a statistician and composed of students from various major fields, provided a natural venue for utilizing various Computer Science methodologies to help explore its subject matter and suggests that this may be a viable mechanism for communicating other non-traditional subjects to their respective audiences.

---

## Monsters and Masters Speak: What Literary A.I. Characters Think of the Question, “Can Machines Think?”

Beth Walker

University of Tennessee–Martin

Frankenstein—the monster whose mind was made from more than the sum of its parts. Gort—the robot who was recognized as humanity’s real master. Hal 9000—the computer program that kills and dies. Moxon’s Turk—the mechanical chess-player that turns murderer. Non-human and artificial, these literary characters pass Turing’s test, but what have they gained by mimicking humanity? We’ll examine the moral and ethical questions that occupy the “thoughts” of these and other artificially intelligent characters. In the world of literature, the question is not “Can machines think?” but “What do they think about?” A non-technical, interdisciplinary approach to computer-science concepts, “Monsters and Masters Speak” explores the philosophy of artificial intelligence through literary characters.

---

# Grading Using Subversion

Roland H. Untch

Middle Tennessee State University

This paper describes GUS, an assignment management system suitable for use in both distance education and traditional campus settings. This system does not require a web server for network communication and runs on multiple platforms, including Windows, Mac OS X, and Linux.

GUS is constructed atop Subversion, an open-source revision control system. A key feature of this new system is that both assignment collection and return is facilitated; this two-way transport is essential in a distance education environment. Additionally GUS has been coupled, using the Portable Document Format (PDF), with Adobe Acrobat allowing pen-based grading using a Tablet PC. System design, grading experiences, and future work will be described.

---

# A Genetic Algorithm Approach to Course Scheduling Optimization

Brian Toone and Laurel Matherly

Samford University

Determining an optimal course schedule is an important problem that must be solved each semester by almost every college student. Similarly, faculty members must determine which courses their department will offer in a given semester in order to satisfy faculty preferences and benefit the maximum number of students. This work suggests an unconventional yet advantageous approach that uses a genetic algorithm to generate near-optimal student schedules and department schedules.

Genetic algorithms have been used in many areas of computer science to determine approximate solutions to optimization problems. We address two separate student scheduling problems. The first problem relates to the creation of a schedule for the upcoming semester based on actual catalog offerings and seat availability. The second problem relates to the creation of an extended schedule that spans a student's entire career. For the problem of deciding which courses a department should offer, we only consider the problem of determining a schedule for the upcoming semester based on the progress of student majors, faculty preferences, and classroom availability.

For all of these problems, our basic approach is the same. An initial population of student or department schedules is generated randomly. Then the operations of selection, crossover, and mutation are applied to this population to produce a new population of schedules. Selection is performed based on a fitness level of each schedule. Student schedule fitness is determined by fixed constraints (e.g., course pre-requisites) and student preferences (e.g., early morning courses). Department schedule constraints include fixed constraints (e.g., the number of classroom seats) and faculty preferences (e.g., desired days of week and times of day to teach). A key benefit of using a genetic algorithm in our approach is the scalability of the system to handle a large number of scheduling constraints, and the flexibility to add or remove individual constraints from the optimization algorithm.

---

# A Data Mining Project Implementation Model

Vernon L. McGlone and Teresa A. McGlone

University of the Cumberland and Eastern Kentucky University

Many businesses today have accumulated huge amounts of data in large databases. To convert this data into strategic business information, many companies are turning to data mining, which can be defined as the automated extraction of hidden predictive information from databases. It combines a variety of techniques, such as statistical analysis, visualization, decision trees, and neural networks. The purpose is to explore large amounts of data and discover relationships and patterns that shed light on business problems. Companies can use these findings for more profitable decision making and competitive advantage.

Although data mining tools have been around for many years, widespread use of data mining became feasible only after recent advances in hardware and software technology. However, many database/data warehouse managers take a wrong approach when getting started with data mining. Data mining is essentially a discovery process, with a high rate of project failures, although project failure is rarely due to poor implementation. Rather, data mining projects often fail due to a flawed project strategy.

It is quite common for organizations to structure their data mining project design as a version of standard corporate practice for evaluating and purchasing products and services. Such an approach rarely works. The data mining process must begin with an understanding of business opportunities. In fact, the key to successful data mining lies in a structured methodology to find problems, define solutions, set expectations, and deliver results. In this presentation, we describe a multi-step model based on such a structure. It is one which organizations can follow to make data mining projects more successful and profitable.

---

# Establishing an Interdisciplinary Bioinformatics Curriculum at Fort Valley State University

Ramana Gosukonda, Masoud Naghedolfeizi and  
Sanjeev Arora

Fort Valley State University

The department of Mathematics and Computer Science at Fort Valley State University (FVSU), an HBCU, has initiated developing a curriculum in bioinformatics that includes the following components: an elective course to introduce bioinformatics, training modules containing bioinformatics topics that can be integrated in science and engineering courses, and a bioinformatics laboratory. In addition, faculty involved in this project provides training to other faculty members who are interested in using bioinformatics tools in their courses.

The bioinformatics course provide students with an overview of biological databases including design, development and maintenance, algorithms for processing biological sequences, and use of industry standard software systems to access, retrieve, and analyze biological data available. This course has been cross-listed with the Department of Biology in order to be used as a major elective in both departments. The first offering of the course is scheduled for Spring 2007.

The training modules are currently being developed in collaboration with faculty of Computer science, Mathematics, Physics, Biology, Chemistry, and Engineering. The targeted courses for integrating these modules are: Database systems, Genetics, Biochemistry, Statistics, Robotics, and Biotechnology. The structure of modules will include hands-on approaches for accessing and analyzing biological data obtained from relevant databases.

The bioinformatics laboratory is being established through an externally funded project to carry out the implementation of the aforementioned curriculum. The lab will include a server equipped with various biological databases, industry standard software packages, 10 Windows based PCs and 10 UNIX workstations.

The bioinformatics curriculum developed at FVSU will not only expose students to the field, but also enhance their overall knowledge in computer science, mathematics, science and engineering as well as hands-on approaches. In addition, the curriculum will give students an opportunity to pursue career choices in this field.

---

# Parental Views of Females in Computer Science and Technology

Dr. Suzanne Smith, Prof. Kellie Price,  
Prof. Michaele Laws and Ms. Rebecca Loyd

East Tennessee State University

While it is predicted that 8 of the 10 fastest growing occupations from 2000-2010 will be in the IT/CS fields, it is expected that women will not be equally represented within these occupations (U.S. Department of Labor, 2003; Camp, 1997). “Girls in Science and Technology” (GIST) is a free science and technology camp at East Tennessee State University (ETSU) making efforts to change these trends. The primary goal of the “girls only” GIST camp is to introduce females to the fields of Information Technology, Computer Science and Math by providing discipline related activities, enhancing team work competency, connecting females with women mentors working in the field and creating a challenging yet fun atmosphere, free from male competition. The GIST camp has been offered over the last 4 years, gathering feedback from young females and their parents. One goal of the researchers is to discover trends in the parental views of females in technical fields. Influence at home is one factor which shapes the way a young female relates to technology, and the potential for her to choose a technical field as a profession (Verbick, 2002). Some of the research questions posed by investigators prior to starting the camp were the following: (1) Do parents perceive women as a minority in Computer Science/Technology? (2) Do parents perceive women will not be as successful as men in Computer Science/Technology? (3) Do parents perceive high paying jobs will be available to women with a degree in Computer Science/Technology? (4) Do parents perceive gender affects a child’s ability to succeed in Math and Technology?

A survey instrument was developed to target the research questions and gauge the feedback from parents. A modified Likert scale was used, eliminating the “undecided” option, so that some opinion had to be offered for each question. Information gained from the parental survey will be the topic of discussion for this presentation.

Camp, T. (1997). The Incredible Shrinking Pipeline, *Communications of the ACM*, 40, 103-110.

Verbick, T. (2002, February). Women, Technology, and Gender Bias. *JCSC*, 17(3), 240-250.

---

## Extra Credit Game 2.0

Denise Williams

University of Tennessee–Martin

The Extra Credit Game is a tool that provides an opportunity to review class material, concepts, and terms. Depending on one's perspective, the best thing about the Extra Credit Game varies. For students, it provides an opportunity to earn extra credit during the time class is scheduled. As a teacher, it creates a means for offering extra credit by asking students to focus on the material at hand, by rewarding students who study in advance as they are more likely to answer questions correctly, and by not worrying about asking students to attend events when they may have other obligations. The game is played in teams to add to the excitement. Students, with their teams, answer questions from the current class and recent classes. Like the Family Feud television show, it is possible to steal answers away from other teams. The Extra Credit Game offers a more exciting and an easy method to review class material, reinforce material, and offer extra credit while creating very little extra work for students or instructors.

---

# Teaching Database via Service Learning

Greg Kawell

Samford University

Samford University began in the small rural town of Marion Alabama in 1841. In 1887 the school moved to Birmingham Alabama. The town of Marion, which was located in the heart of Alabama agriculture, fell on hard times over the last 160 years. The economic impact on the area has gained the central part of Alabama the name the Black Belt. Over the last few years Samford University has began many programs to try and reach a helping hand back to the town of Marion. In the spring of 2006 I discussed the idea of developing a way to catalog all the technology the school district in Marion owns.

During the spring semester I taught our department's database course with a twist. We would still have lectures and assignments, but the course would also design a database to catalog the districts technology. The project consisted of taking the class the 75 miles down to Marion and conducting a focus group to discover the need for this database. Back in the classroom we agreed to divide the class into two groups with each developing their own database design. Presentations by each group were made several times during the semester explaining their design and getting class feedback. In late April we returned to Marion and each group had a chance to pitch their design. Feedback was given and by the end of the semester one design was sent for approval by the school district.

This presentation will more fully explain the project and also present the students' reaction to the project and the community that they were working with. Journals that were kept by the students showed that they not only enjoyed the project, but felt their learning was greatly enhanced by putting feet to the lecture.

---

## Web-Based RGB Color Exercises

James M. Bateman

Austin Peay State University

This interactive graphics program is written as a web page in client-side JavaScript and HTML and was developed by the author for teaching color theory. It is targeted for use in courses in computer graphics, web development, and multimedia. RGB is the hardware-related color model used for programming display unit graphics, HTML/XHTML, and CSS. Each color is thereby represented as a three-dimensional vector. Since vectorial operations in relation to color are unfamiliar to most students, the program is intended to facilitate their learning. Color concepts covered include complementation, saturation, scalar multiples of colors (brightening and darkening), color addition, decomposition into gray and saturated components, interpolation, and luminance. A series of problem types were developed to introduce concepts, provide visible examples, and test students' comprehension. Each problem type presents in three ways--numerically, with a three-bar chart, and with color swatches--one or more colors chosen as random triples of natural numbers less than 256. The student responds by entering RGB triples for requested colors that are functions of the given colors. Randomness ensures that each student gets his/her own problems since there are more than sixteen million triples. The program has two modes of operation: "Just practicing" and "Recording". Recording mode records in a table the successes or failures for attempted solutions of problems. Only the last two such scores are kept in the table for each problem. Problem types can be selected by menu or considered sequentially. New instances of a problem type can be requested. A "Check Answer" button provides scoring and feedback. A "Show Answer" button lets students see solutions for any number of instances while they are just practicing or after an attempt has been recorded. The program has been used in an on-line course and an on-campus course with favorable results.

---

# Integrating Alice into K-12 Outreach Events

Jeff Gray

University of Alabama–Birmingham

Much commentary has been offered over the past few years regarding the reasons for decreased enrollment in computer science across the country. As a response to the impact of credit hour reduction, many computer science departments have initiated K-12 recruiting efforts to attract students back to the discipline. The goal of such recruitment is to raise the awareness of computing among school systems that have a misconception about the nature of computer science (e.g., many state boards of education equate computing to teaching Microsoft Office<sup>1</sup>). This provides a unique challenge to outreach efforts. On the one hand, technology is deeply integrated into the lives of many students, but the understanding of the science behind the technology they use is perhaps near an all-time low.

This presentation will describe several ongoing recruiting initiatives at UAB within the state of Alabama (e.g., high school programming contest<sup>2</sup>, field trips<sup>3</sup>, visits to computer science classes in K-12<sup>4</sup>, workshops for K-12 teachers<sup>5</sup>, and mentoring of high school students for science fair projects<sup>6</sup>). Specifically, the presentation will provide a brief introduction to Alice and how it has been integrated into each of these outreach events.

Alice is a 3D graphical programming language from Carnegie-Mellon, which provides an “objects-first” approach to teaching programming. The Alice environment allows a student to create graphical animations and video games without the challenges that often accompany efforts to teach programming for the first time. For example, Alice has a “drag and drop” syntax-directed editor that forces a student to write syntactically correct programs.

1. <http://alex.state.al.us/standardAll.php?grade=9&subject=TC&summary=>
  2. <http://www.cis.uab.edu/programs/hspc/>
  3. <http://www.cis.uab.edu/field-trips/>
  4. <http://www.cis.uab.edu/gray/hs-outreach/hs-visit.ppt>
  5. <http://www.cis.uab.edu/programs/hsws/>
  6. <http://www.cis.uab.edu/news/details.php?event id=83>
-

## Evaluating Claims about the Impact of RFID on Privacy

Syed Raza

Talladega College

The proposed research project addresses the problem of representing and assessing claims about the effects of RFID technology on privacy. Sources will be scanned for claims about these effects, including contributions to assessing the claims and links to related claims. Clusters of claims will be formed and represented as part of an effort to provide useful information about the risks and safeguards for RFID deployment. The project will aim to emphasize the development of policy solutions as well as technological ones.

---

# Assessment and Accreditation: ABET's New Student Outcome Requirements

Donald Sanderson

East Tennessee State University

ABET's new pilot standards for CS, IS and IT contain a list of student attributes that each department must ensure their graduates meet. To do this requires most of us to make significant revisions to the way we have done assessment. This talk will cover the new A-I attributes common to all programs, and the specific attributes for CS, IS and IT programs. Ways to leverage assessment plans you already have will be discussed as will methods for planning new initiatives.

---

# Decreasing Enrollments/Increasing Manpower Needs: Solutions for Attracting Students to Computer Science

Wayne Summers and Rodrigo Obando

Columbus State University

Enrollments in Computer Science Departments across the U.S. are down over 30% from three years ago with the number of new majors down by 50% [Taulbee, 2006]. Recent data show that there are now more jobs in IT than any time in history. Projections from the Bureau of Labor Statistics indicate strong growth in computing related fields over the next decade [BLS, 2005] “Coupled with the declining representation of women in our undergraduate programs, our ability to produce a workforce that is sufficiently educated technically to meet the needs of the job market in computing is being severely challenged”[Taulbee, 2006]. We are facing a crisis in the U.S. and many other countries with decreasing enrollments in computer science and increasing demand for trained computer scientists.

The decrease in enrollments seems to be driven by two major factors – a) misinformation in the popular media, and b) lack of means to motivate students in our schools. “Computer science is an established discipline at the collegiate and postgraduate levels. Oddly, the integration of computer science concepts into the K-12 curriculum has not kept pace in the United States. As a result, the general public is not as well educated about computer science as it should be, and a serious shortage of information technologists at all levels exists and may continue into the foreseeable future” [ACM, 2003].

Universities working with the Computer Science Teachers Association (CSTA) need to build relationships and partnerships with K-12 schools to address these concerns. One mechanism is through Teacher Enrichment in Computer Science (TECS) workshops sponsored by the CSTA. Not only will this provide professional development for the K-12 computer teachers, but also help build a strong community of computer science educators who can share their knowledge and experiences. This will also provide a platform for

---

communicating the excitement of computer science to students and improve their understanding of the available opportunities.

Our university has made this partnership a high priority this year. We had a dialogue last Spring with a diverse group of K-12 teachers and administrators and have conducted the first of four workshops for a group of K-12 teachers and students. We have also conducted two workshops for 6-8 grade students in using an exciting programming environment called Alice (<http://alice.org/>). This presentation will highlight a number of these initiatives.

2004-2005 Taulbee Survey:

<http://www.cra.org/CRN/articles/may06/taulbee.html>

Bureau of Labor Statistics Monthly Labor Review, November 2005.

A Model Curriculum for K-12 Computer Science, ACM K-12 Task Force Curriculum Committee Report, October 2003.

---

# Analysis of Student Performance in Programming Subjects of an In-house Exit Exam

Singli Garcia-Otera, Masoud Naghedolfeizi  
and Nabil Yousif

Fort Valley State University

This paper analyzes the performance of students in Computer Science (CS) and Computer Information Systems (CIS) in the programming subjects (Principles of Programming I & II and Data Structure) of an in-house exit exam given in years 2003-2006. The student performance was measured in seventeen different categories of computer programming. The data for each category was collected based on student responses to questions related to that particular category. The data was analyzed with respect to whether a student was majoring in CS or CIS, and with respect to whether a student was a male or female.

The statistical analysis of the data indicated that approximately 80% of students performed above the passing grade of 70% in structured programming categories. This indicated that the retention of knowledge in such categories was satisfactory. However, the performance of nearly 70% students was weak in complex data structures and object oriented programming. The analysis also showed that students majoring in CS generally performed better than CIS students in most categories of computer programming. Additionally, both male and female students performed approximately at the same level.

The results of the analysis will be used to identify the problem areas and make necessary adjustments to both curriculum and teaching strategies in order to improve and enhance the long-term knowledge retention of students in computer programming subjects.

---

# Managing Complexity in Ethical Analysis

Jeff Roach and Don Gotterbarn

East Tennessee State University

SoDIS, an established method of ethical analysis, relates three different elements in the identification of ethical risks of a software project. These elements are the activities needed to complete the project, a list of possible ethical issues, and the stakeholders that the project affects. A complete risk analysis requires a consideration of all combinations of the activities, the issues, and the stakeholders to identify the ethical risks.

As the number of activities and stakeholders increase, the analysis results become progressively more complex to understand. A moderate sized project can require the presentation of 500,000 lines of results for the analysis. Even for small projects, the results for every triple (activity, issue, and stakeholder) may make the analysis impractical. This problem is further exacerbated when multiple concerns and possibly multiple solutions for each concern are recorded for each analysis triple.

The analyst must be able to easily view the important results. This was addressed by using a spreadsheet. Rows represented activity/stakeholder pairs and each column represented an issue group called a risk category. This two-dimensional view proved inadequate as each cell now had an additional dimension - the ethical issues for that particular risk category. This posed a presentation problem - how to display and manage more than four dimensions of data using a two-dimensional spreadsheet. The solution to the problem is the focus of this paper.

Another aspect of the problem was how to internally represent the analysis so that the spreadsheet could be quickly built and the results rapidly retrieved. This was solved with an object-oriented framework that utilized modified forms of existing design patterns.

This paper presents a detailed description of the user interface design and the object-oriented framework. We believe that this technique can then be further generalized to other problems requiring four or more dimensions of data.

---

# Extending the “Bagging” Ensemble Approach for Classifying Highly Skewed Data

Cen Li

Middle Tennessee State University

In many applications, data collected are highly skewed or imbalanced where data with one class label clearly dominates data from the other classes. In particular, the class of interest has much less data than that of the other classes. Successful classification methods, such as C4.5, Naïve Bayes (NB), Neural Network(NN), K Nearest Neighbor(kNN), and Support Vector Machine (SVM), that are effective on classifying balanced, and near balanced data, lose their classification ability when dealing with highly imbalanced data.

Many studies have been done to try to improve the quality of classification results on imbalanced data. These works may be categorized into one of the following three groups:

- Over-sampling approach which increases the size of data having the target class;
- Under-sampling approach which reduces the size of data having the non-target class(s); and
- Approaches that involves modification of the classification methods to counter the effect of dominating class on classification.

We propose a variation of the “bagging” classification ensemble approach for imbalanced data. In the classical “bagging” approach, data is first divided into  $N$  parts where an individual classifier is learned from each part. During classification, each of the  $N$  classifier generates one prediction, and the majority prediction is used as the classification for the test data. In our approach, when forming the  $N$  training data sets, the target data in each of the  $N$  training sets remains the same as the original data. The non-target data is randomly divided into  $N$  sets and are distributed into the  $N$  training sets. Majority vote is then applied on the  $N$  classifiers learned from the  $N$  training sets. Empirical studies are conducted to compare this approach with the existing approaches developed for imbalanced data classification. A case study is performed to show the effectiveness of this approach on real world data sets.

---

# Cheap, Flexible, and Secure Remote File Sharing with Samba, OpenVPN, and Apache

Edwin Rudolph and Adel Abunawass

University of West Georgia

As telecommuting becomes more commonplace, technical staff are increasingly asked to provide tools to support remote access to the organization's computing resources; access to files and file sharing being one of the most common. While email is frequently used for this purpose due to its ubiquity, it is less than ideal for secure and effective file sharing.

Although network file sharing is often as prevalent as email within the campus network, extending file sharing to remote users presents several challenges compared to email. Common network file sharing protocols such as CIFS and NFS were designed for use in "local" area networks. Making such services available remotely across the Internet raises additional and significant performance and security issues that must be addressed.

Many technical solutions exist to these problems, ranging in complexity, cost, and ease of use. Any solution will be a trade-off between these factors. For our relatively small department, a solution was needed that was primarily cheap, flexible, and easy to use; and hopefully limited in complexity. The target user would have a Windows PC or Mac with broadband Internet connectivity. The solution also needed to integrate well with a FreeBSD server environment and Kerberos authentication infrastructure. After investigating a variety of possibilities, we implemented a system consisting entirely of free software: Samba, OpenVPN, smb webclient, and Apache+SSL. Windows networking (Samba) as the base protocol provides flexibility and ease of use for the target user, allowing network shares to be accessed directly by the OS. Additionally, users may access files through a basic web interface. Authentication and encrypted encapsulation of data are provided by OpenVPN.

This talk will discuss challenges involved, alternative solutions that were explored, and an overview of our solution from technical and end-user perspectives.

---

# Programming Language Support for XML Encryption

David Frazier

East Tennessee State University

Encryption allows for either an entire XML document or selected elements to be encrypted. For example, if you are sending an XML encoded document of an employee record, you may only choose to encrypt their social security number. XML Encryption also allows for multilevel encryption.

This is extremely important in e-commerce application. Consider an ecommerce transaction from AcmeOnline.com. When a customer places an order, the entire order can be encrypted so that only employees of AcmeOnline.com can access the information. To add another layer of security, the credit card information element can be encrypted with a second key. Only those within AcmeOnline.com with a need to know the credit card information would have this second key.

This presentation will explore ways to implement XML Encryption in an ecommerce setting. I will specifically look at support for XML Encryption in some popular ecommerce development languages such as Java, .NET and PHP.

---

# Speculative Execution and Value Prediction Support for Distributed Systems

Angkul Kongmunvattana

Columbus State University

Advances in computer and network technologies have generated an explosive growth in the deployment of both large and small-scale distributed computing systems, such as clusters of PCs or workstations and symmetric-multiprocessor clusters. At the same time, the cost and complexity of application software development on these platforms have increased substantially because they presented programmers with two distinct memory interfaces: (a) shared memory for processors within the same node, and (b) distributed memory for processors that are not co-located. The programmer is thus faced with a dilemma. Performance advantages of using the native load/store interface for communication between co-located processors are apparent, but accesses to remote memory require an entirely different communication model. Employing two different communication models in the same program is problematic, especially in the case where the number of threads and processors per node is not known at compile time. It is imperative that the scientific community and the industry develop efficient system software to support seamless programming environment on these platforms. For this software to be employed in the emerging distributed computing systems that encompass both large and small-scale clusters with a broad range of system speeds and heterogeneity, spanning across local and wide-area networks, advanced techniques for tolerating the negative impacts of high network latency and non-dedicated network environments are crucial and inevitable. This presentation explains innovative mechanisms and policies to support speculative execution and value prediction. By allowing speculative execution, the software can better tolerate and even exploit the differences of communication latencies between accessing local and remote memories. As the opportunity cost of waiting for remote data to arrive increases, the potential benefit of speculative execution also increases. Additionally, highly accurate value prediction can dramatically improve overall performance without compromising the simplicity of application software development.

---

# Autonomic Computing: Vision, Characteristics and Applications

Srinivasarao Krishnaprasad

Jacksonville State University

Advances in the mass-fabrication of processing, memory, and communication hardware has led to the invasion of information technology and computing into all aspects of our lives. A significant trend in recent years is the increased interconnectivity of multitude of devices and the resulting complexity of the computing infrastructure. Managing, maintaining and troubleshooting these complex systems not only require highly skilled information technology workers but also may become impractical due to worker shortages and errors inherent in human interventions. Paul Horn of IBM Research proposed a solution in which computing systems monitor and regulate themselves in a way similar to our own autonomous nervous system: low-level body functions such as regulating body temperature and heart rate are done without a conscious effort from our brains. He coined the term *autonomic computing* for this technique. Some key features of autonomic computing include self-configuration, self-optimization, self-healing, and self-protection [2]. Designing systems to possess one or more of these aspects is a real challenge for researchers and engineers.

Recognizing the challenges involved in developing future computing platforms, IBM Research [1] has taken a leading role along with academic and other business partners. In this talk we present an overview of autonomic computing highlighting its vision, characteristics and implementation challenges. This evolving technology will benefit many aspects of computing as envisioned by researchers ([2], [3]). A brief survey of some of the applications and research activities based on autonomic computing at various research facilities will be presented.

## References

- [1] "Autonomic Computing," IBM Research, [www.research.ibm.com/autonomic](http://www.research.ibm.com/autonomic), accessed May 4, 2006.
  - [2] "The Vision of Autonomic Computing," J. Kephart and D. Chess, *IEEE Computer magazine*, January 2003.
  - [3] "Self-Repairing Computers," A. Fox and D. Patterson, *Scientific American*, May 12, 2003.
-

# Sun Microsystems vs. Microsoft: The Controversy about Delegates

Ken R. Adcock, Jr.

## UPS Supply Chain Solutions

The legal activity between Sun Microsystems and Microsoft regarding the licensing of Java is well-known. At issue was Microsoft's lack of support for certain key features of the Java specification along with their inclusion of additional features designed to allow the Microsoft implementation of the JVM to work uniquely well on the Windows platform. What is not commonly known even among technology professionals is that one aspect of this disagreement involves a vigorous debate between two great language architects of our time: Anders Hejlsberg of Microsoft and James Gosling of Sun Microsystems. Specifically, this debate centers on the issue of Delegates.

Although certainly not a new concept, the concept of the Delegate has existed in other programming languages, but is commonly referred to as being analogous to the C++ function pointer. The inclusion of the Delegate construct into Visual J++ produced considerable disagreement from Sun on pure technical merits and also served as part of the basis for Sun's lawsuit against Microsoft. Furthermore, Sun declared the existing specification could accomplish the same behavior in a more proper fashion and also stated the Delegate construct would probably never be made part of the Java specification. However, with regard to Microsoft, the Delegate went on to play a central role in the .NET framework.

This presentation will take a closer look at the specifics surrounding the debate about Delegates between Sun Microsystems and Microsoft. This presentation will demonstrate that arguments such as this can be interesting, complex, passionate, and best of all, instructive. The end goal of this presentation is to illustrate how our understanding can be made better by deeply examining the nature and evolution of these arguments rather than simply being passive consumers.

---

# Writing Your Own Operating System – Hints, Tools and Resources

Bob Bradley

University of Tennessee–Martin

There are many challenges and rewards involved in writing a new operating system. This presentation will give a list of hints, tips, tools and resources needed to develop your own multi-tasking operating system. First, we will talk about why anybody would want to write a new OS and what benefits can be gained by attempting this task. We will talk about all of the different areas that have to be mastered to write an OS, including booting, real mode, protected mode, interrupts, memory management, file systems, shells, input, output, multi-tasking, graphics, etc. Next, we will also talk about the free tools needed to develop an OS. Some of the free tools discussed include the Bochs PC Simulator, Microsoft Virtual PC (2004), VM Ware (Server and Player), Grub, asm, gcc and g++. Lastly, I will present some web resources that document, give sample source code and give hints for many of the tougher areas of OS development. This presentation should be useful to anybody that has to do (or teach) any type of low-level OS, diver or system programming and development.

---

# Lessons Learned: a Year of Alice

Will Lloyd

University of West Georgia

Alice, a programming language and environment designed for teaching, has been shown to offer many advantages to students beginning to learn software development. Alice allows students to focus on programming concepts and the development process without much concern about syntax. Most students also find that Alice is fun: they learn the basics of object-oriented programming and software engineering by creating and modifying interesting, challenging, 3D animations.

This presentation will draw on our experience in an introductory course aimed at both prospective CS majors and non-majors. We have used Alice in this "CS 0" course for two and a half semesters with over 200 students, and have learned a lot about how to use Alice effectively as well as about the strengths and weaknesses of Alice.

We will focus on creating effective lab exercises and programming assignments; using Alice to appeal to and meet the needs of both non-majors and prospective majors; encouraging students to focus on the broader issues of software development rather than on just programming-in-the-small; integrating an Alice-based CS 0 course into a more traditional introductory sequence; and discovering what not to expect from Alice.

---

## The Case for Alice

T. F. Higginbotham, Ph. D.

Southeastern Louisiana University

A new programming language by the name of Alice has recently become available. It is a 3D programming language in which a syntax error is impossible because everything is drag-and-drop. It was designed for beginning programmers and includes understandable object oriented programming, including inheritance. The programming support system (IDE) is excellent, but not perfect.

This paper is concerned with a course I taught this summer as an experiment to replace the contents of Computer Science 101, which we have tried to teach, without great success, something about Computer Science. The traditional course contain some history of Computer Science, the internal workings of computers, and some exercises in writing loops, assignment statements, and the like. We had intended to be a course in how to solve problems. No actual programming was done.

Alice, taught entirely over the Internet, went over very well with the students this summer. Many of the students had completed all eight programming assignments before the semester was half over. The only way this could happen would be because the students were enjoying what they were doing.

The SOT's (Student Opinion of Teaching) were very high. One student in the comment section said this course almost convinced him/her to change to a Computer Science Major.

Very few of the students were actual Computer Science majors. Most were from the humanities, the arts, or business. Did we recruit a few Computer Science majors this way? I hope so!

---

# Introducing Programming to Entry-Level Computer Classes with Carnegie Mellon's ALICE Visual Programming Environment

Dr. David Brown

Pellissippi State Technical Community College

ALICE is a visual programming environment developed at Carnegie Mellon University. It is intended to teach beginning programming and object-oriented concepts. Pellissippi State Technical Community College (PSTCC) introduced ALICE into their "Introduction to Computer Science" class during the fall 2006 semester. The impact on the students in this class has been dramatic. Student interest and retention in the course have improved and students are finding interesting and novel ways to utilize the software to express their inner creativity through programming.

This paper discusses the use of ALICE at PSTCC and how other institutions may be able to utilize this unique environment to reach students who may otherwise seem unreachable.

---

# Alice: Learning to Program by Creating Movies

Brenda Parker

Middle Tennessee State University

Alice is a free, open source programming environment that allows students to learn basic programming concepts while creating animated movies. It has been shown to help students succeed in Computer Science especially those that have little or no programming experience. Using Alice has also been shown to improve students' attitudes toward computer science. Alice uses drag and drop features to represent words in a programming language. This allows programmers to gain experience in learning logic and early programming concepts without the frustration of working with syntax errors. An introduction to the Alice programming environment will be introduced. Experiences involving its usage will also demonstrate how it can be used to introduce prospective students to the world of computing and therefore, foster interest in the Computer Science.

1 [www.alice.org](http://www.alice.org)

2 Moskal, M., D. Lurie, and S. Cooper, "Evaluating the Effectiveness of a New Instructional Approach" Proceedings of 2004 SIGCSE Conference, (Norfolk, VA), 75-79.

---

# Introducing Quantum Computing and Communication to the Computer Science Curriculum

Tzusheng Pei

Jackson State University

Quantum computing is inevitable. Under the current design and manufacturing technology, there are two ways to enhance computing power: running computers at higher speeds and implementing more parallelism on semiconductor computer chips. Because there is an upper bound of the speed of information transmission, it takes time to transmit information from one point of a CPU chip to another point. The clock rate for a computer CPU chip cannot be increased indefinitely, otherwise the chip will become out of synchronization at a speed high enough. Implementing higher parallelism means putting more computing units and/or more sophisticated pipelines on computer chips. This requires packing more electronic elements in limited spaces, either two dimensional or even three dimensional, and at the same time it meets the synchronization requirement. Consequently the electronic elements on computer chips will become smaller. At a certain point of shrinking electronic elements, quantum effects cannot be ignored, and the computer will not work properly. Quantum computing improves on “classical computers and classical complexity bounds by making use of quantum mechanical phenomena.” The excitement of quantum computing largely comes from the celebrated Shor's efficient (polynomial time) algorithm for factoring large numbers on a quantum computer and the Grover's efficient quantum search algorithm over the corresponding classical search algorithm. Discussions of quantum computing and communication courses at the undergraduate and graduate levels are presented. Additionally, suggestions are made for adding quantum computing materials in major undergraduate computer science courses, such as discrete mathematics, computer organization and architecture, theory of computation, and algorithms. In a broader view, the ultimate goal of research and education in quantum computing is to recast computer science in a quantum mechanical framework and pass the results to our next generation of students.

---

# A First Course in Information Security

Nancy Smithfield

Austin Peay State University

Information Security applies to all aspects of the protection of data as well as the systems and hardware that store, process and transmit the data. Viruses, phishing schemes, use of bots to attack other computers are realities that students need to understand and what policies, procedures and technology controls can be put into place to help control the risk.

We have recently developed two security courses designed for Information Systems, Internet and Web as well as our Database Administration students. The first course is an introductory course with prerequisites of the CS1 class or a computer hardware, software and programming class in the Information Technology Concentrations. The second class concerns network security and it is taken after the data communications and networking class. This presentation will provide details on the first course. The presentation will include objectives, topics covered, sample exercises, team projects and labs, and opportunities for the future.

---

# System Security Concerns in Component-Based Software Design

Kathy Winters

University of Tennessee–Chattanooga

Computer system security is of great concern to everyone especially those in the Information Assurance field. The United States economy is based on information. One must be able to authenticate code to make sure it is free from manipulation and unauthorized use. For truly secure systems the industry must go beyond systems such as intrusion detections and firewalls and deal with security in the design process. With the premise that the system is only as secure as its weakest link, one must look at the practice of code reuse and component based design and its affect on security. In both of these cases, the developer is using code with incomplete knowledge. The code is a plug and play black box which may or may not be thoroughly tested; may or may not be built and designed for the current use; and may or may not perform as intended under the current conditions. These issues should be addressed and considered when attempting to build secure computer systems. Of particular concern is the practice of code reuse and component based design in C and C++. This presentation will explore the pitfalls of code reuse and component based software design in C and C++ systems.

---



## Panels and Roundtable Discussions



## Podcasting and Other Tools for Online Learning

David Frazier, Michael Laws,  
Carolyn Novak and Jeff Roach

East Tennessee State University

We would like to propose a Roundtable discussion on the use of Podcasting and other tools to increase student participation in online classes. Keeping online students engaged is an ongoing struggle. We would like to discuss techniques that can, and have been used to help with this problem. In particular, we want to discuss the use of Podcasting, or the pushing of audio files to students. Other topics could include use of blogs, discussion boards, and collaboration tools such as Writely, an online word processor.

---

## OS X as a Unix Environment

Jeremy Ey, Justin Stinson and Eric Brown

Tennessee Technological University

As noted by [1, 2, 3, 4] there exists wide coverage of migrations to Mac OS X in academic environments. All of these migrations, however, are from prior versions of Mac OS to Mac OS X. With Mach and BSD roots, however, there is also interest in using OS X as a replacement to other UNIX operating systems. In our experience OS X serves this role reasonably well in both server and workstation roles.

Its selection as the platform has resulted in increased usage of our labs compared to the previous Linux configuration. Easy integration with the existing Windows environment, has provided users with a unified account structure. Using the Active Directory support built into OS X, all user accounts were moved into Active Directory and all user data hosted on an OS X RAID device.

This configuration, however, has not been without problems. During the initial migration it was desirable to provide many of the same applications as on the previous system. This was generally not a problem for command line applications, however, complex graphical applications, had difficulty. One recurring problem has been the use of network home directories. Many applications do not properly recognize and use such directories and either fail to run, or have problems while running.

A panel addressing the topic of OS X as a replacement for traditional UNIX environments would benefit three main groups. Users who are considering migrating to OS X would gain knowledge from the experience of existing users. Users who have already migrated to OS X for their UNIX environment would have an opportunity to share experiences and interact with other users. Users who replaced previous versions of Mac OS with OS X would gain knowledge about additional uses for this investment.

---

## References

- [1] Hanselman, S. E., and Pegah, M. Macintosh os x: a smooth migration. In SIGUCCS '03: Proceedings of the 31st annual ACM SIGUCCS conference on User services (New York, NY, USA, 2003), ACM Press, pp. 129–134.
  - [2] Houston, D. L. R. Putting osx in an open access lab: (or "the joy of x"). In SIGUCCS '03: Proceedings of the 31st annual ACM SIGUCCS conference on User services (New York, NY, USA, 2003), ACM Press, pp. 154–158.
  - [3] Jones, C. Should you upgrade to mac os x in your computing labs? In SIGUCCS '02: Proceedings of the 30th annual ACM SIGUCCS conference on User services (New York, NY, USA, 2002), ACM Press, pp. 256–257.
  - [4] Miller, C. R., and Bent, H. R. Os x: a ten-step program. In SIGUCCS '04: Proceedings of the 32nd annual ACM SIGUCCS conference on User services (New York, NY, USA, 2004), ACM Press, pp. 217–220.
-

## Tutorial Session



# Alice – a Gentle Introduction to Object-Oriented Programming

Wayne Summers

Columbus State University

We all know that writing computer programs can be difficult and intimidating. The programming environment Alice (<http://alice.org/>) makes programming easy and fun. Alice makes programming more accessible to girls as well as boys and to learners of all ages. It teaches object-oriented programming without getting bogged down in the syntax of a programming language.

This tutorial provides a brief overview of the Alice programming environment. A limited number of laptops will be provided, so participants may need to share computers. CDs with the software will be available for participants who have brought their own computers.

---

## Conference at a Glance



	<b>Azalea Room</b>	<b>Dogwood I</b>
<b>8:00 AM</b>	<b>Welcome</b>	
<b>8:15 AM</b>	<b>Keynote Address</b>	
<b>9:15 AM</b>	<b>Coffee Break – Poolside</b>	<b>Coffee Break – Poolside</b>
	<b>Session I – Undergrad 4</b>	<b>Session I – Doctoral</b>
	Chair: Joyce Crowell	Chair: Cen Li
<b>9:30 AM</b>	Amos Smith (26)	Chun Wei (42)
<b>9:50 AM</b>	Ian Mitchell (27)	Amine Chigani (43)
<b>10:10 AM</b>	Lorenzo Jackson (28)	Bruce Johnson (44)
	<b>Session II – Undergrad 4</b>	<b>Session II – Doctoral</b>
	Chair: Li Yang	Chair: Allan Anderson
<b>10:35 AM</b>	Jason Gruber (29)	Faizan Javed (45)
<b>10:55 AM</b>	Ashley Plier (30)	Matthew Thornton (46)
<b>11:15 AM</b>	Shayan Javed (31)	Shih-Hsi Liu (47)
<b>11:35 AM</b>		Trevor Jay (48)
<b>12:00 PM</b>	<b>Lunch – Poolside</b>	<b>Lunch – Poolside</b>
	<b>Session III – Undergrad 4</b>	<b>Session III – Doctoral/Pro</b>
	Chair: Nancy Smithfield	Chair: Brenda Parker
<b>1:00 PM</b>	Hampton Haddock (33)	Varadraj Prabhu Gurupur (49)
<b>1:20 PM</b>	William Hendrix (34)	Johnson/Pan (50)
<b>1:40 PM</b>	Scott Kubina (35)	Yan Liang (51)
<b>2:00 PM</b>	Tony Snyder (36)	Syed Raza (Professional) (68)
<b>2:20 PM</b>	<b>Break – Poolside</b>	<b>Break – Poolside</b>
	<b>Session IV – Professional</b>	<b>Session IV– UG2/Masters</b>
	Chair: Sylvia Colvin	Chair: Melissa Wiggins
<b>2:35 PM</b>	Cen Li (74)	Jackie Grindstaff (UG2) (24)
<b>2:55 PM</b>	Rudolph/Abunawass (75)	Michael Wade (38)
<b>3:15 PM</b>	David Frazier (76)	Alma Cemerlic (39)
<b>3:35 PM</b>	Angkul Kongnumvattana (77)	Jeremy Blair (40)
	<b>Session V – Professional</b>	<b>Roundtable Discussion</b>
	Chair: Glenn Wiggins	Podcast/Online Learning
<b>4:00 PM</b>	Tzusheng Pei (85)	Frazier/Laws/Novak/Roach (90)
<b>4:20 PM</b>	Nancy Smithfield (86)	
<b>4:40 PM</b>	Kathy Winters (87)	
<b>5:00 PM</b>		
<b>7:00 PM</b>	<b>Awards Banquet</b>	

	<b>Dogwood II</b>	<b>Magnolia</b>
<b>9:15 AM</b>	<b>Coffee Break – Poolside</b>	<b>Coffee Break – Poolside</b>
	<b>Session I – Professional</b>	<b>Session I – Professional</b>
	Chair: Tesfa Haile	Chair: Bob Bradley
<b>9:30 AM</b>	Sung Yoo (54)	Steve Donaldson (57)
<b>9:50 AM</b>	Wayne Walters (55)	Beth Walker (58)
<b>10:10 AM</b>	Kathy Winters (56)	Roland Untch (59)
	<b>Session II – Professional</b>	<b>Session II – Professional</b>
	Chair: Jim Vandergriff	Chair: Wayne Walters
<b>10:35 AM</b>	Brian Toone (60)	Denise Williams (64)
<b>10:55 AM</b>	Vernon/Teresa McGlone (61)	Greg Kawell (65)
<b>11:15 AM</b>	Gosukonda/Naghedolfeizi/Arora (62)	James Bateman (66)
<b>11:35 AM</b>	Smith/Price/Loyd/Laws (63)	Jeff Gray (67)
<b>12:00 PM</b>	<b>Lunch – Poolside</b>	<b>Lunch – Poolside</b>
	<b>Session III – Undergrad 2</b>	<b>Session III – Professional</b>
	Chair: Jeff Gray	Chair: Ashraful Chowdhury
<b>1:00 PM</b>	Jackie Grindstaff (20)	Donald Sanderson (69)
<b>1:20 PM</b>	Ohana/Phillips/Thomas (21)	Summers/Obando (70)
<b>1:40 PM</b>	Phillips/Smallwood (22)	Garcia-Otera/Naghedolfeizi/Yousif (72)
<b>2:00 PM</b>	Babwani/McGuinn/Rowe (23)	Jeff Roach (73)
	<b>Break – Poolside</b>	<b>Break – Poolside</b>
	<b>Session IV Professional</b>	<b>Session IV – Professional</b>
	Chair: Richard Barber	Chair: Denise Williams
<b>2:35 PM</b>	Srinivasarao Krishnaprasad (78)	Will Lloyd (81)
<b>2:55 PM</b>	Ken Adcock (79)	T. F. Higgenbotham (82)
<b>3:15 PM</b>	Bob Bradley (80)	David Brown (83)
<b>3:35 PM</b>	<b>Panel Discussion</b>	Brenda Parker (84)
	OS X as Unix Environment	<b>Tutorial</b>
	Ey/Stinson/Brown (91)	Alice/OOP
		Wayne Summers (94)
<b>5:00 PM</b>	<b>Business Meeting</b>	

# Notes



# Notes



# Notes



