COEST’s TRACELAB: INFRASTRUCTURE FOR TRACEABILITY RESEARCHERS AND PRACTITIONERS

Jane Hayes, PhD
hayes@cs.uky.edu

Jane Cleland-Huang, PhD
jhuang@cs.depaul.edu

Funding for the work discussed in this presentation provided by the National Science Foundation grants # CNS0959924, CCF0810924, CCF 0447594, CCF 0306303, and Siemens Corporate Research, and NASA grant # NNX06AD02G.
Our roots...

- **Early work**
  - TEFSE community
  - An idea
  - Towards a grand challenge
  - Beyond the challenges
  - TraceLab & Benchmarks

- **Seminal work in the mid 1990s** highlighted the traceability problems faced in industrial practice
Beginnings of a community…

- George Spanoudakis and Andrea Zisman launched the TEFSE series of workshops to encourage research in this area
  - TEFSE 2009, May 18, 2009, Vancouver, British Columbia, Canada
  - TEFSE 2007, Also referred to as Grand Challenges of Traceability: GCT, Slade, Kentucky
  - TEFSE 2005, November 8, 2005, Long Beach, CA
  - TEFSE 2003, October 7, 2003, Montreal, Canada
A forward looking idea...

- **Idea:** Form a **Center of Excellence for Software Traceability (COEST)**

- The first CoEST meeting occurred under the St. Louis Arch at ICSE 2005. It was attended by Jonathan Maletic, Giulio Antoniol, Alex Dekhtyar, Jane Cleland-Huang, Jane Hayes, and several students.
COEST Organization

Officers

Director:
Jane Huffman Hayes
Associate Professor, University of Kentucky

Vice Director of Europe:
Andrea Zisman
Professor, City University, London

Vice Director of the Americas:
Jane Cleland-Huang
Associate Professor, DePaul University, Chicago

Secretary/Treasurer:
Alexander Egyed
Professor, Johannes Kepler University, Linz, Austria

Body of Knowledge Coordinator:
Alexander Dekhtyar
CalPoly.

Grand Challenges Coordinator:
Olly Gotel
Independent Consultant

Publications Coordinator:
Jonathan Maletic
Professor, Kent State University

Student Coordinator:
Giulio Antoniol
Ecole Polytechnique Montreal, Canada

NEW BOOK
Software and Systems Traceability
Andrea Zisman
Jade Cleland-Huang
Olly Gotel
Springer Verlag
to be released in Fall of 2011
CoEST’s vision

- The **vision** of the COE for Software Traceability is to provide leadership for traceability research, education, and practice; promoting the pursuit of excellence from research idea to practice, based on a foundation of innovative, ethical, collaborative work.
- Seed funding was provided by NASA and NSF.

Everyone is welcome to join!
CoEST’s vision will be achieved by:

- Identifying the **grand challenges** of traceability
- Fostering international research **collaborations**
- Developing a repository of **benchmarks** for traceability research
- Delivering **tutorials** in the areas of requirements engineering and traceability
- Constructing a **Body of Knowledge** for traceability
- Cultivating **partnerships** with industry and academia
- Providing the **infrastructure** needed to support the above
Lessons from a wise cat

In *Alice in Wonderland*, Alice encounters the Cheshire Cat and asks for directions. The cat responds that it depends upon where she wishes to go. Alice says she doesn't know, whereupon the cat tells her that it doesn't matter which way she walks.
Towards the Grand Challenges...

- A group of CoEST members met at NASA’s IV&V facility in Fairmont, WV, to flesh out the grand challenges of traceability.
- We produced GCT 1.0.
- Over the past 2-3 years, a smaller group of people have been working to transform GCT 1.0 into a more cohesive research roadmap.
The Traceability Grand Challenge

- Traceability is always there, without having to think about getting it there. Traceability is neither consciously established nor sought; it is built-in and effortless. It has effectively ‘disappeared without a trace’

- Is our challenge feasible? Do practitioners agree that it is the right challenge?
The Goals of Traceability

Olly Gotel, Jane Cleland-Huang, Alex Dekhtyar, Jane Huffman Hayes, Andrea Zisman, Alex Egyed, Giulio Antoniol, Jonathan Maletic, Stephanie Ferguson, Ken McGill, Tim Menzies, Marcus Fisher, Lisa Montgomery, Brian Berenbach, Paul Gruenbacher, & numerous PhD students….

For each Challenge
- Definition
- Description of desired outcome
- Problem addressed
- High level goals (3-4)
- Impact of challenge area on trace strategy, use, creation and maintenance
- Research Projects (1-2)
- Sub Research Projects (7-10)
- Industry Practice objectives (3-5)
GCT Summary

- Each of the 8 challenge areas have been fully mapped to goals, research projects, and industrial adoption strategies.
- Grand Challenges will serve as a roadmap for guiding traceability research and advancing the state of practice.
**Ubiquitous Traceability**

- **Major Research Project:** RP1.1 Provide automation such that traceability is encompassed within broader software and systems engineering processes, and is integral to all tool support

- **Supporting Research Projects:** RP1.2 Embed traceability into all the software and systems engineering techniques and methods that it facilitates, and transfer this into industrial tool support

- RP1.3 Total automation of trace creation and trace maintenance, with quality and performance levels superior to manual efforts
Tracking our progress...

Center of Excellence for Software Traceability

Developing automated tools to minimize the cost and effort of traceability.

What is Traceability?

The Grand Challenges

About TraceLab

For Researchers

For Companies
Research Projects

Project 1: Traceability Automation and Integration
Providing automation such that traceability is encompassed within broader software and systems engineering processes, and is integral to all tool support.

Supporting Research Projects

Research Project 1.2
Embed traceability into all the software and systems engineering techniques and methods that it facilitates, and transfer this into industrial tool support.

Importance Difficulty Progress Comments Rating

Research Project 1.3
Total automation of trace creation and trace maintenance, with quality and performance levels superior to manual efforts.

Importance Difficulty Progress Comments Rating

Research Project 1.4
Investigate novel ways to define the traceability strategy, such as in an executable way so that the traceability solution simply follows from the specification of need, as per model-driven or formal development.

Importance Difficulty Progress Comments Rating

Please login first in order to leave a comments or provide ratings. Login Now Register
How do we get there?

- So now we know where we want to go, but how do we actually get there?
- TEFSE (ongoing)
- CoEST
  - Grand Challenges
  - Benchmarking
  - TraceLab
  - Education and outreach
- Early work
- TEFSE community
- An idea
- Towards a grand challenge
- Beyond the challenges
- MRI funding
Equipping a community

- Establishing a research environment takes time and effort
- Comparing results and evaluating progress towards the traceability goals is difficult

Early work
TEFSE
community
An idea
Towards a grand challenge
Beyond the challenges
MRI funding
A benchmark is a point of reference by which something can be measured

- A program that is specially designed to provide measurements for a particular operating system or application
- A set of performance criteria which a product is expected to meet
- A set of conditions against which a product or system is measured
Benchmarks

- **Define a task**
  - Retrieve/Generate traces from high level to low level requirements

- **Provide datasets**
  - CM1, HIPAA to World Vista, IBS

- **Agree on a core set of metrics**
  - Recall, Precision, Lag, Average Precision (??)

- **Capture/Report benchmarked results**
What about Qualitative Studies?

- Several identified challenges are more qualitative in nature
- How do we comparatively evaluate processes and methods which don’t lend themselves to quantitative metrics?
What about Qualitative Studies?

- Several identified challenges are more qualitative in nature
- How do we comparatively evaluate processes and methods which don’t lend themselves to quantitative metrics?

A framework enables us to collect metadata on each project, and then as a community move towards answering higher-level questions such as “under what conditions does technique X work most effectively?”
TraceLab- The Vision

- Build a tool, similar to MatLab, but designed specifically for the traceability community
- Equip new researchers with basic algorithms and components
- Make it easier to perform rigorous comparative evaluations
  - Datasets
  - Benchmarks
  - Repeatable experiments
- Permit practitioners to use “best” algorithms for specific benchmark
Major Research Instrumentation

- Currently has capabilities to design and run experiments
- Next features will integrate with benchmarking

TraceLab Version 0.1
TraceLab developers: Ed Keenan, Adam Czauderna, and Greg Leach
Traceability Research Timeline

1995 & earlier
Seminal work in traceability

Rudimentary tools

2nd Generation of Trace features in RM tools

Technology transfer pilots

1995-2010
Numerous researchers work on various traceability topics receiving funding from NASA, NSF, & Industry

What next?
The Grand Challenges provide a roadmap for future research efforts and the mechanism for tracking progress towards our goals

Grand Challenge Workshops held, GCT 1.0 released

2010: MRI funded by NSF for $2M

GCT 2.0 released, Jan. 2011

funded by NASA & NSF 2006/7
COEST’s TRACELAB: INFRASTRUCTURE FOR TRACEABILITY RESEARCHERS AND PRACTITIONERS

Jane Hayes, PhD
hayes@cs.uky.edu

Jane Cleland-Huang, PhD
jhuang@cs.depaul.edu

Funding for the work discussed in this presentation provided by the National Science Foundation grants # CNS0959924, CCF0810924, CCF 0447594, CCF 0306303, and Siemens Corporate Research, and NASA grant # NNX06AD02G.
Backup
What is a grand challenge?

What makes this a good Grand Challenge?
Is “Traceability” a grand challenge?

- Is traceability **important**? Why?
- Is traceability **difficult** to achieve?
- Do we have a **clear vision** of where we want to go?
Recall vs. Precision problem – small changes in thresholds can have inordinate impact upon recall vs. precision – creating zigzag graphs. For benchmarking metrics, how do we overcome this?
High water marks—

Will high benchmarks thwart innovation?

Is this a good or bad thing?
Benchmarks

Trust –
What kinds of checks and balances do we need to put into the process to make sure that benchmarks are fair?

How do we make comparisons anyway?
Benchmark issues

Early work
TEFSE community.
An idea.
Towards a grand challenge.
Beyond the challenges
TraceLab & Benchmarks

Yonghee’s work
Benchmark insights

- What is the purpose of benchmarking our community?
  - What do we hope to accomplish from benchmarking?

- What are the major pitfalls of benchmarking in the traceability community?
  - How can we avoid them?