CENTER OF EXCELLENCE FOR SOFTWARE TRACEABILITY (COEST) – AN UPDATE ON TRACELAB+, PLUS ASSESSMENT OF PROVIDED TRACE MATRICES

Jane Hayes, PhD
hayes@cs.uky.edu

Funding for the work discussed in this presentation provided by the National Science Foundation grant CCF0810924 and NASA grant # NNX06AD02G.
“The ability to describe and follow the life of a requirement, in both a forward and backward direction, e.g. from its origins, through its development and specification, to its subsequent deployment and use, and through periods of ongoing refinement and iteration in any of these phases”

Gotel & Finkelstein
How do we verify that all requirements have been met?
Is that all design elements exist to satisfy requirements?
Information Retrieval (IR) is a study of methods and techniques for finding relevant documents in document collections, given user queries.

- Intersection of Statistics, AI and Databases
- Proven success

Generated Traceability Matrix
Importance of TMs

- Some Software Engineering Activities that Require TMs
  - Criticality analysis
  - Requirement satisfaction
  - Change impact analysis
  - Hazard reachability analysis
  - Regression testing
  - Traceability analysis
  - Risk analysis
  - Test Plan and Test Case Generation
  - Interface analysis
  - Consistency checking
How Ensure TMs are Accurate?

- Manual review
  - Time consuming
  - Error prone
  - Mundane

- Spot check
  - Incomplete
  - Mundane

- Assisted checking – enter Trace Matrix Analysis
Trace Matrix Analyzer (TMA)

- Treat TM as graph (edges, nodes)
- Apply graph heuristics to TM analysis
  - Children with too many parents?
  - Parent with no child?
  - Possible missing links?
  - Possible bad link?
- Visualize possible issues

Advantages

- Greatly reduce workload of those performing requirements assurance
- Simple to operate
- Friendly User Interface (UI)
- Powerful expandability

DEMO of TMA
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 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
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
Benchmarks

- 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**
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
TraceLab

- Research environment designed to allow researchers to visually compose, execute traceability experiments using library of shared components.
- Components in any memory managed language such as Java, C#, etc. TraceLab also allows calls to tools such as Matlab, R, etc.
- TraceLab currently runs in Windows environment but designed to port to Linux.
TraceLab - The Role of Contests

- Define contests for community participation
  - Task (feature location)
  - Data set (benchmark)
  - Collection of “frozen” components with one “open”
- Prize to winner
- Permit practitioners to use “best” algorithms for specific benchmark (“player” from contest winner)
### Leader Board

<table>
<thead>
<tr>
<th>Technique Name</th>
<th>Technique Description</th>
<th>Candidate Username</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. VSM with basic preprocessing</td>
<td>Standard Vector Space Model that pre-processes data using stopwords removal, and term stemming.</td>
<td>Yonghee</td>
<td>0.039</td>
</tr>
<tr>
<td>2. VSM with splitter</td>
<td>VSM</td>
<td>Adam</td>
<td>0.031</td>
</tr>
<tr>
<td>3. Baseline</td>
<td>Basic Vector Space Model without any data preprocessing</td>
<td>Greg</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### Metrics Results

**Average Precision:**

- **Preceding:**
  - Minimum: 0.271
  - First quartile: 0.416
  - Median: 0.620
  - Mean: 0.638
  - Third quartile: 0.866
  - Maximum: 1.000
  - Interquartile range: 0.450
  - Standard deviation: 0.250
  - Sample size: 28

**Preceding Recall:**

- Minimum: 0.271
- First quartile: 0.416
- Median: 0.620
- Mean: 0.638
- Third quartile: 0.866
- Maximum: 1.000
- Interquartile range: 0.450
- Standard deviation: 0.250
- Sample size: 28
TraceLab Environment

The image shows a flowchart diagram of the TraceLab Environment. The diagram includes various components and processes, such as the Target Artifacts Importer and Preprocessor, Source Artifacts Importer and Preprocessor, TFIDF Dictionary Index Builder, Tracer Component, and Results Metric Computation. The flowchart starts with the Start node and ends with the End node, indicating the sequence of processes in the environment.

The components library on the left includes options like Algorithms, Control Logic, and Preprocessors, among others. The workspace view on the right displays output information, including global log levels and messages related to the Tracer Component and Results Metric Computation.
Standardized Metrics

Evaluation Results

Below you can see the evaluation metrics of your results for each dataset.

Select Dataset: CM1

<table>
<thead>
<tr>
<th>Metric</th>
<th>Standard VSM</th>
<th>Current Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Median</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Mean</td>
<td>0.3892348...</td>
<td>0.38923486774382077</td>
</tr>
<tr>
<td>Max</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.29</td>
<td>0.29</td>
</tr>
</tbody>
</table>
Creating Contest for Traceability Techniques

Eight areas of traceability challenges with 58 research projects

Task Data Sets
Collaborative data collection and verification
Guidelines for creation and usage of evaluation metrics
Example Contests

- **Contest 1**
  - **Task:** Trace retrieval from use cases to code
  - **Data Sets:** EasyClinic, eTour, Eanci, SMOS
  - **Metrics:** Average Precision

- **Contest 2**
  - **Task:** Reducing human effort for relevance feedback
  - **Data Sets:** EasyClinic, eTour, Eanci, SMOS
  - **Metrics:** Average Precision, Number of feedbacks provided by human analyst
TraceLab Timeline

- Currently in beta-use at 6 universities
- Planned public release in July 2012 in conjunction with launching 5-6 research contests – culminating in The Grand Challenges of Traceability at ICSE 2013
- Will be open-sourced towards the Fall of 2012
DEMO of TRACELAB

Demo
Questions?
Backup
Challenges

2. **Purposed** - Traceability is fit for purpose and supports stakeholder needs.

3. **Cost-Effective** - The return from using traceability is adequate in relation to the outlay of establishing it.

4. **Configurable** - Traceability is established as specified, moment-to-moment, and the rich semantics accommodate changing stakeholder needs.

5. **Trusted** - All stakeholders have full confidence in the traceability, as it is created and maintained in the face of inconsistency, omissions and change; they can and do depend upon it.

6. **Scalable** - More and more artifacts are supported by traceability, of varying types and at variable levels of granularity, as the traceability extends through-life, and across organizational and business boundaries.

7. **Portable** - Traceability is exchanged, merged and reused across projects, organizations, domains, product lines and supporting tools.

8. **Valued** - Traceability is a strategic priority valued by all, where every stakeholder has a role to play and actively discharges his or her responsibilities.
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?
Benchmarks

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?
Benchmarks

High water marks—

Will high benchmarks thwart innovation?

Is this a good or bad thing?

“I tripled my salary to give you all a good benchmark!”
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?