CSC 509 Lecture Notes Week 8

Administrative Matters
CT Paper

I. Administrative Matters

A. *Remaining work:*

1. Assignment 6 -- last set of readings.
2. Assignment 7 -- in-class presentations.
3. Assignment 8 -- final project/paper.

Administrative Matters, cont'd

B. ***Presentation scheduling:***

1. See wiki.
2. Enter your prefered "no sooner than" date.

Administrative Matters, cont'd

C. ***Final Project Deliverables:***

1. See wiki.
2. Enter brief summary at today's brief meetings.

Now onto "CT" and "OP" papers;
First up is CT.

II. Some Starter Questions

A. What's CT got to do with formal modeling, i.e., the paper from assignment 2?

B. What's CT got to do with test adequacy criteria, i.e., the paper from assignment 5?

C. At what level of testing is CT applicable -- *system?*, *unit?*, *both?*

Some Starter Questions, cont'd

D. What's CT got to do with your 509 project?

E. What's CT got to do with your testing life?

III. **Noteworthy citations from the CT paper:**

A. Tai and Leim, 2002 TSE,
   -- *Pairwise Combos*

B. Kuhn and Wallace, 2004 TSE
   -- *Fault Interactions*

C. Cohen, Dwyer, and Sei, 2008 TSE
   -- *Next Level Stuff*

IV. **Eight classification categories.**

A. On page 3, near end of Intro.

B. Works as a "best practices" list for evaluating any testing methodology, not just CT.

C. The classification categories are:

Classification categories, cont'd

1. *Modeling*
2. *Test Case Generation*
3. *Constraints*
4. *Failure Characterization*
5. *Improvement Identification*
6. *Prioritization of Test Cases*
7. *Metrics of Efficacy*
8. *Empirical Evaluation*

V. Section 2 of the paper

A. Good coverage of the basics.

B. A bit technical here and there.

C. Bottom lines:

Section 2, cont'd

1. We're looking at all combos of parameter values.
2. Use fewer combos to avoid explosion.
3. Pairwise combos are surprisingly effective.
4. A covering array is a handy visualization.

VI. Section 3 of the paper

A. Doesn't follow 8 categories exactly.

B. Does thorough job of covering last 20+ years of research.

C. A typical time span for any kind of testing research

VII. Section 3.1 -- Modeling

A. They’re talking about a "model" for input parameters and their interactions.

B. This section says: *"To obtain the information on the interactions and constraints between parameters, we can study the requirement document, design document, codes, and other related documents."*

C. How does this idea of modeling compare with a formal predicative model?

VIII. Section 3.2 -- Generation

A. *3.2.1 -- Covering Arrays*

1. Lots of different possibilities

2. Mats et al. found that "Each Choice" wins, at least in their study.

3.2 -- Generation, cont'd

B. *3.2.2 -- Seeding*

1. Provide "hand selected" test cases.
2. Then let auto gen loose.
3. Fouche et al describe *adaptive* seeding, 2007 ACM FSE

3.2 -- Generation, cont'd

C. *3.2.3 -- Constraints*

1. Need to determine which combinations are valid.
2. SAT raises it's little head.
3. *Anyone remember SAT?*

3.2 -- Generation, cont'd

D. *3.2.4 -- Generation Technique*

1. "Classic" greedy algorithm.
2. "More Plodding" search algorithm.
3. *Anyone remember hill climbing, tabu search, simulated annealing?*
4. Genetic algorithm heuristic search also employed.

3.2 -- Generation, cont'd

E. *3.2.5 -- Generation Tools*

1. ***Lots*** of the out there.
2. Cohen et al. 1997 TSE -- AETG
3. Cohen et al. 2008 TSE -- AETG lives on
4. www.pairwise.org is interesting.

IX. Section 3.3. -- Test Case Prioritization

A. Formally, an ordering function.

B. Ordering criteria can be chosen adaptively.

C. E.g., higher priority cases are those that (initially) reveal more flaws.

D. This is not unique to CT.

X. Section 3.4 - Failure Diagnosis

A. Determine which forms of combination lead to failure.

B. Strengthen test by adding more tests of that form.

C. In adaptive approaches, test suites grow intelligently.

XI. 3.5 -- Metrics and Evaluation

A. *CT self-metric* -- combination coverage.

B. *CT versus the world metric* -- code coverage, mutation score

XII. 3.6 -- Applying CT

A. Interesting observation about using CT in the ag domain in 1926.

B. For modern software CT, there are empirical studies.

C. See, e.g., http://www.pairwise.org/results.asp

XIII. 3.7 -- Summary

A. Seminal work in 1985, 1992, 1994.

B. Figures 10 and 11 summarize results of ten research groups.

C. Marked increase in publication rate since 2002, still trending up.