Introducing Combinatorial Testing in a Large Organization: Pilot Project Experience Report

Jon Hagar[1],Rick Kuhn[2], Raghu Kacker[2], Tom Wissink[3]
embedded@ecentral.com;kuhn@nist.gov,
raghu.kacker@nist.gov, tom.wissink@lmco.com
[1]Grand Software Testing & Lockheed Martin
[2]National Institute of Standards and Technology,
[3]Lockheed Martin

This poster gives an overview of the experience of eight pilot projects, over two years, applying combinatorial testing in Lockheed Martin (LM), one of the world's largest aerospace firms. Lockheed Martin and NIST developed a Co-operative Research and Development Agreement (CRADA) to evaluate effectiveness and areas of suitable application for combinatorial testing in a real-world industrial setting with complex software requirements. (One of the ways in which NIST conducts joint research with US industry is through CRADAs, which allow federal laboratories to work with US industry and provide flexibility in structuring projects, intellectual property rights, and in protecting industry proprietary information and research results.)

Objectives in the pilot project evaluation included:

1) Investigate applicability of CT in a variety of application areas, including system, software, and hardware testing;
2) Determine effectiveness of CT for improving fault detection; and
3) Study potential for reducing test cost or overall lifecycle cost by finding errors earlier in the process.

Software Tools

The primary tool for most projects was ACTS. Additional tools with complementary capabilities were also used in the pilot projects.

- NIST & U. of Texas Arlington: ACTS
- Air Academy Associates: SPC XL, DOE KISS, DOE PRO XL, DFSS MASTER
- Phadke & Associates: rdExpert
- Hexawise: Hexawise tool

Application Areas

Eight pilot projects were identified, with the goal of using the new methods in areas with diverse testing needs:

- Flight Vehicle Mission Effectiveness (ME) – comparing CT with tests generated from a statistical analysis tool
- Flight Vehicle engine failure modes – compared CT tests with existing tests developed using previous practice
- Flight Vehicle engine upgrade – tests including combinations of flight mode factors; comparison with existing tests
- F-16 Ventral Fin Redesign Flight Test Program – application to problem analysis (system-level evaluation rather than software testing)
- Electronic Warfare (EW) system testing – evaluating and extending existing tests
- Navigation Accuracy, EW performance, Sensor information, and Radar detection – generating test cases for subsystems
- Electromagnetic Effects (EMI) Engineering compared CT tests with existing tests developed using previous practice
- Digital System Command testing – testing file functions with multiple parameters

Results and Evaluation

While results varied across the different pilot projects, overall it was estimated that CT would save roughly 20% of testing cost, with 20% - 50% improved test coverage. In some cases, significant, previously undetected bugs were discovered. Additional findings included:

Positive results - Demonstrated the ability to reduce test cost in a variety of areas; teams found many tools practical

Mixed results – Reluctance of many engineers to adopting new methods; some teams did not identify significant improvements

Lessons learned – Most critical factors affecting adoption: availability of education and training for the new method; clear demonstration of value.

Recommendations

- Develop and improve education and training materials
- Incorporate combinatorial methods into DoD guidance and industry standards; best practices
- Expand internal company guidance – developing a community of practice
- Greater availability of tools to support combinatorial testing – improved usability; matching tool to problem
- Modify approaches to using combinatorial testing – integrating combinatorial testing with other test practices; ability to adopt CT partially or gradually

Certain products may be identified in this document, but such identification does not imply recommendation by the US National Institute of Standards and Technology or other agencies of the US Government, nor that the products identified are necessarily the best available for the purpose.