
September 3, 2012

CVSM 2012 WG B&TE – Main Results

State of The Art: no benchmarks known, evaluation: ad hoc
in many papers, but only qualitatively

Proposed Section of a State of the Art Report: none.

Research Agenda: everything, domain must be built from scratch

Potential Benchmarks
• classical properties addressed by benchmarks: response time, throughput, etc.
make sense only for functions which require significant time / resources; here:
comparison of two models, batch merge of to models, ....??

• obstacles to benchmarks:
1. require common technology to represent models, e.g. EMF;
(EMF is not optimized for this purpose!)
2. require same (or very consistent) metamodels
3. require clear definition of expected result (if any!); problem: high vs. low quality differences
4. problem of long loading times and technologies measuring resource consumption (CPU time, memory)
achievable in the short term only for:
- model matching and difference derivation
- model types where stable tools are available

Evaluation / Technology Assessment
• based on more abstractly defined test cases, which leave room
for implementation decisions
• stemming from tool suppliers and major “real” users
• approach: “tool contest”

CVSM Tool Contest
1. open call for test cases
2. assessment (and improvement) of submitted test cases
3. “contest”: submission of solutions of the test cases
4. assessment and evaluation of submitted solutions
5. live presentation of the solutions
gold, silver and bronze models ....

CVSM Tool Contest: Proposals for Test Cases
• description of the development task (from a predefined list,
e.g. show difference between models, batch or interactive merging of two models, patching, ...)
• description of more detailed assumptions about the context and assumptions
• models involved in the test case provided as:
a) printable representation
b) Ecore/XMI representation, including meta model
c) additional other representations (e.g. tool-specific)

• expected result
• in general terms
• proposed “good” solution for the test models
• criteria to be used when evaluating submitted solutions, incl. scoring/grading scales, ....
• expected number of submitters of solutions
CVSM Tool Contest: Evaluation of Submitted Test Cases

- is the test case implementable / solvable? is the required effort feasible?
- is the test case relevant for / representative of the development task?
- what can the community learn from the solutions?

Udo Kelter, 2012-09-03

CVSM Tool Contest: Possible Realization

2012-10-15 submit workshop proposal to “Software Engineering 2013” in Aachen 2013-02-26ff
2012-11-15?? call for test cases published
2013-01-15 submissions due, evaluation starts
2013-02-26 workshop: discussion of test cases

Udo Kelter, 2012-09-03