CS 475/575
VV&A

M. Overstreet
Old Dominion University
Spring 2005

VV&A

- Based on
  - "Verification, Validation and Accreditation of Simulation Models," by Osman Balci, WSC '97. "Verifying and Validation Simulation Models," Robert Sargent, WSC '02

some terms

Quality metrics: accuracy, efficiency, maintainability, portability, reusability, usability (CHI). here only accuracy is of concern

Model Verification: model is correctly transformed from one representation to another

Model Validation: the model, within a particular domain, behaves with satisfactory accuracy consistent with M&S objectives.

Model testing: checking for inaccuracies or errors

Accrediation: US military term: "the official certification that a model or simulation is acceptable for use for a specific purpose." DoD Directive 50000.29, http://triton.dmso.mil/docslib/mspolicy/directive.html

DoD effect

US DoD is big player in this area.

On the web Defense Modeling & Simulation Office, (www.dmso.mil) but each branch (Army, Navy, Air Force) also on Web.

"Properly" done, can be expensive. Few organizations can afford "proper" V&V

General Comment

Large complex problem.

One approach not appropriate for all.

Much business simulation is short and quick (gain insight).

Industry (like Boeing, NASA) builds "engineering models" as design tools.

Common Approaches

Development does V&V.

- adv: they understand assumptions, model, code
- disadv: objectivity? blind spots?

Independent V&V: done by independent group.

- adv. objectivity (maybe)
  - independent of schedule, budget, biases
- disadv:
  - costs.
  - independent of schedule, budget, objectives
  - human tendency to find something regardless

V&V Principles (Balci)

- V&V must be conducted throughout life cycle
- V&V outcome not binary
- Based on M&S objectives
- Requires independent assessment
- Requires creativity & insight
- Early error detection better
- Complete testing not feasible
- Type I, II, & III errors must be avoided
- Successful V&V of all submodels does not imply Model V&V
- Validity does not imply credibility and acceptability
- Having well-formulated problem essential

V&V Techniques

- Informal
- Static
- Dynamic
- Formal

Informal

- audit
- desk checking
- documentation checking
- face validation
- inspections
- reviews
- Turing tests
- walkthroughs

Static

- Cause-effect graphing
- Control analysis
  - calling structure analysis
  - concurrent process analysis
  - control flow analysis
- Data analysis
  - data flow and data dependency analysis
- Interface analysis
- Structural analysis
- Traceablity assessment

Dynamic

- Acceptance testing
- Assertion checking
- Compliance testing
  - performance, security standards
- Execution testing
  - execution monitoring, profiling, tracing
- Software testing
  - Black box: equivalence partitioning, extreme input, trace-drive input, stress testing, ...
  - White box (structural): branch, condition, data flow, ...
- Statistical techniques
- Visualization/animation

Formal

- Induction
- Lambda Calculus
- Predicate Calculus
- Proof of Correctness

Some pragmatic approaches

- Compare with other models if possible
- Use degenerate tests - (where you know the answer)
- Event validity - does the right sequence of events occurs when triggered? Compare model generated events with system events
- Use fixed values (constants) - similar to degenerate tests, but requires code change.
- Drive model with real system traces to compare model output with system output.
- Determine internal variability of model. If high, may require model revision
- Sensitivity analysis/accuracy requirements
- Predictive validation: model used to forecast system behavior which can be checked (e.g. weather models)

data validity

- not really different, but often a time-consuming, complex issue
- outputs no better than data used to build model
- on many projects, the data costs much more than the model
- many models are “data driven” ModSAF reads about 300,000 of data on start-up.
- data are needed to:
  - build conceptual model
  - validate model
  - perform experiments to model

need to validate:

- ⌘conceptual model
  - ⍁underlying theories and assumptions are correct/acceptable
- ⌘computerized model
  - ⍁all software testing techniques apply, and then some
- ⌘operational model
  - ⍁the world changes over time. are assumptions/behaviors still valid? (is the world still flat?)

Sample form

Category/ Item	Technique used	Justification for technique	Reference to supporting report	Result/ Conclusion	Confidence in result
Pgm correctness	Static testing Dynamic testing				

Strengths	
Weaknesses	