Verification and Validation of Simulation Models

One of the most important & difficult tasks
for a model developer.

V & V

- The goals of V & V:
  - Product a model that represents true system behavior closely enough to be useful.
  - Increase the credibility of the model, sot that the model will be used for management and predication.
- Validation should be an integral part of model development, not an isolated task. Iterative in nature.
- Verification:
  - Building the model right?
  - Is the model implemented correctly?
- Validation:
  - Building the right model?
  - Is the model an accurate representation of the real system?

V & V and Modeling Building

- Methods discussed in this chapter:
  - Informal subjective comparisons.
  - Formal statistical procedures: used together with output analysis.
- Model building:
  - Observe the real systems: components and their interactions; and collect data
  - Construct a conceptual model: assumptions, structure, hypotheses on the parameter values.
  - Translate the model into software.
- Important things to remember:
  - Domain knowledge
  - Stakeholders: operators, technicians, managers, users, …
  - Iterative process.

Verification of Simulation Models

- Areas involved: program testing & simulation knowledge
- Program testing:
  - Independent tester.
  - Walkthrough or trace the program.
  - Good programming practices: style, variable names, documentation, and etc.
  - Debugging aids or print out a wide variety of input parameters and output statistics for debugging.
  - Graphical interfaces or animation techniques.
- Simulation knowledge:
  - Reasonableness of model output.
    - Example: The queue length is unreasonably high. Some possible reasons?
    - Figure 10.2

Calibration and Validation of Models

- Validation is the overall process of comparing the model and its behavior to the real system and its behavior.
- Calibration is the iterative process of comparing the model to the real system, making adjustments to the model, comparing the revised model to reality, making additional adjustments, comparing again, and so on.
- Subjective tests: involve people and experience.
- Objective tests: requires data on system's behavior and the corresponding data generated by the model.
- Collection of more than one set of system data.
- Trade-offs: cost/time/effort versus accuracy (or mark).

Validation Process

- Naylor and Finger formulated a 3-step approach:
  1. Build a model that has high face validity.
     - If model is reasonable on its face to model users & experts?
  2. Validate model assumptions.
     - Structural assumptions: how the system operates?
     - Data assumptions: collection of reliable data & correct statistical analysis of data. (Chap. 9)
  3. Compare the model input-output transformations to corresponding input-output transformations for the real system.

Validating Input-Output Transformations

- The objective test of a model is its ability to predict the behavior of the real system.
- The structure of the model should be accurate enough for the model to make good predictions, not just for one input data set, but for the range of input data sets which are of interest.
- In this phase, the model is treated like a “black box” which accepts values of the input parameters and transforms these inputs into outputs.
- Make use of historical data.
- Use the main responses of interest as the criteria for validating the model.
- If the system is in the planning stages, other types of validation should be used. Example: subsystems may exist and partial input-output validation can be conducted.

Validating I-O Transformations (2)

- The model can also be used to compare alternative system designs.
- If some version of the system is operating and has been validated, the validity of the model of a nonexistent proposed system can be evaluated:
  - The responses of the two models under similar input conditions can be used as the comparison criteria.
  - If the proposed system is a modification of an existing system; changes range:
    - Minor changes of numerical parameters: # of servers
    - Minor changes of the form of a statistical distribution: distribution of service time
    - Major changes in the logical structure: scheduling policy
    - Major changes involving a different design for the new system

Validating I-O Transformations (3)

- As shown in Figure 10.3, validation should be viewed in the context of calibrating a model.
- If the current version of the model produces estimates that are not close enough to real system behavior, the source of the discrepancy is identified, and the model is revised in light of the new knowledge. This iterative process is repeated until model accuracy is judged adequate.
- Example 10.2
  - Model of input-output transformation: Figure 10.5
  - Input and output variables for the model: Table 10.1
  - Table 10.2
  - The estimate is not close enough. Identify potential flaws in model assumptions or structure.
  - If found, revise the model and estimate again. Table 10.3

Statistics Stuff & Power of Test

- Sample mean (Y_2 bar)?
- Sample standard derivation (S)?
- Statistical test: t-test
  - $t_0 = (Y_2 \text{ bar} - \mu_o) / (S/n^{1/2})$, μ_o is the specified value in the null hypothesis
  - For two-sided test, if $|t_0| > t_{\alpha/2, n-1}$, reject H_0
- Power of test:
  - probability of detecting a departure from H_0: $\mu = \mu_o$ when in fact such a departure exists.
  - probability of detecting an invalid model for V&V.
  - $1 - \beta = P(\text{rejecting } H_0 \mid H_1 \text{ is true})$
  - $\beta = P(\text{Type II error}) = P(\text{failing to reject } H_0 \mid H_1 \text{ is true})$.
  - β depends on the sample size n and on the true difference between $E(Y_2)$ and μ_o.
  - $\delta = |E(Y_2) - \mu_o| / \sigma$, σ is estimated by S. So, δ hat $= |E(Y_2) - \mu_o| / S$.
  - Find β (δ hat) from Table A.10 or A.11.

I-O Validation: Using Historical Input Data

- Event-scheduling without random-number generation; rather, use historical data, if available.
- It is expected that the simulation will duplicate as closely as possible the important events that occurred in the real system.
- It's important that all the input data and all the system response data be collected during the same time period.
- This technique is difficult for a large system or some applications.
- Example 10.4

I-O Validation: Using a Turing Test

- If no statistical test is readily applicable, persons knowledgeable about system behavior can be used to compare model output to system output.
- Simulation output data are sued to produce "fake" reports in exactly the the same format of the system output.
- The managers and engineers are asked to to decide which reports are fake and which are real.
- If a substantial number of fake reports are detected, the model builder questions the engineers and uses the information to improve the model.
- If engineer cannot distinguish between fake and real reports with any consistency, then there is no evidence of model inadequacy.

Summary

- Validation of simulation models is vital.
- Quite often simulations appear realistic on the surface because simulation models can incorporate any level of detail about the real system. It's best to make many comparisons.
- Some possible validation techniques: (see page 393)
- Part of the model builder's task is to choose those validation techniques that are most appropriate.