

Parte XIII

Verificação e validação de modelos

- No processo de desenvolvimento de um sistema de simulação, primeiro construímos um “modelo conceitual” (ex.: fluxogramas, suposições) e em seguida utilizamos este modelo para programar o sistema.
 - Figura 5.1, Law. Retângulos representam estados do modelo ou do sistema, setas cheias são ações necessárias para mudar o estado, e as setas tracejadas indicam as avaliações feitas nas transições.
- Portanto, existem duas etapas na avaliação de correção do sistema:
 - Determinar se o modelo conceitual é uma representação precisa do sistema sendo estudado (Validação).
 - Determinar se o programa está de acordo com o modelo (Verificação ou Depuração). Ou seja, verificar a tradução do modelo conceitual para um programa de computador funcionando corretamente.

46 Verificação

- Embora o conceito de verificação seja simples, depurar um modelo de simulação de larga escala é uma tarefa árdua.
- Apresentaremos a seguir técnicas para qualquer programa, e técnicas específicas para simuladores.
- **Escreva e depure programas em módulos.**
 - Encontrar um erro em um programa grande é difícil.
 - Um programa principal simples é criado e testado, então módulos vão sendo inseridos e testados.
- **Mais de uma pessoa deve ler cada parte do código.**
 - Ex.: todos os membros da equipe se reúnem para analisar cada instrução de uma rotina (walk-through estruturado).
- **Execute a simulação sob várias configurações de entrada, e verifique se a saída é razoável.**
 - Em alguns casos podemos calcular exatamente o valor de saída. Ex.: tempo médio de serviço em uma fila para determinados parâmetros.
- **Trace**

- Em um trace, o estado do sistema (ex.: o conteúdo da lista de eventos, as variáveis de estado, contadores estatísticos) é impresso após a ocorrência de cada evento.
- Às vezes é necessário utilizar entradas determinísticas.
- Comparamos então o trace com cálculos feitos manualmente.
- É importante verificar como o sistema se comporta em condições extremas.
- Ao invés de imprimir, o analista pode utilizar um “debugger interativo”. Ou seja, parar a execução em determinado ponto no tempo, e analisar o estado do sistema.
 - * Está disponível na maioria das linguagens de simulação e programação.
 - * Este recurso também permite “forçar” o valor de variáveis, facilitando a ocorrência de alguns erros.

- **O modelo deveria ser executado (quando possível) sujeito a suposições simplificadoras.** Para facilitar o cálculo das saídas esperadas.
- **Para alguns tipos de simulação, uma animação pode ajudar a observar a saída.** Podemos também utilizar gráficos, que são mais fáceis de acompanhar que traces.
- **Calcule a média e variância de cada distribuição de entrada, e compare com a distribuição desejada (ex.: dos dados históricos).**
 - Isto é ainda mais útil se estavamos utilizando bibliotecas para geração de números aleatórios para determinada distribuição, pois podemos nos enganar quando a utilização exata dos parâmetros (isto é comum com a distribuição gamma).
- **Utilize pacotes de simulação para reduzir o tamanho do código.**
 - Mas cuidado com novas versões, pois podem conter erros.
 - Outra desvantagem é não saber exatamente como determinada função foi implementada, o que pode causar falhas de interpretação.
- **Teste de continuidade**
 - Pequenas variações nos dados de entrada geralmente provocam pequenas variações na saída. Se as variações forem grandes, o modelo deve ser investigado.
- **Independência de semente**
 - O modelo deveria produzir resultados semelhantes para sementes diferentes.

47 Validação

- Um processo de validação não garante um modelo totalmente válido, mas torna o modelo mais representativo do sistema real.

47.1 Determinando o nível apropriado de detalhes no modelo

- Determinar quais aspectos do sistema real devem ser incorporados no modelo é um fator crítico para a viabilidade (tempo de dinheiro limitado) e a validade do modelo.
- Apresentaremos a seguir algumas recomendações para esta decisão:
- **Defina claramente (i) as questões investigadas, (ii) as medidas de performance, (iii) como o modelo será utilizado, e (iv) as configurações (do sistema) de interesse.**
 - Os modelos não são de uso geral, mas projetados para propósitos específicos.
 - Um bom modelo para o problema errado nunca será utilizado.
- **Utilize a opinião dos especialistas e análise de sensibilidade.**
 - Pessoas mais familiares com o sistema podem indicar quais componentes têm mais chance de serem importantes.
 - Como o tempo é limitado, o analista deveria se concentrar nos aspectos mais importantes.
- **Comece com um nível moderado de detalhes, e aprimore se necessário.**
- **Não coloque mais detalhes que o necessário para responder as questões.**
 - Exceto detalhes que aumentam a credibilidade do sistema.
- **O nível de detalhes deve ser compatível com os dados disponíveis.**
 - Ex.: um modelo para uma nova fábrica geralmente tem menos detalhes que um modelo para ajustar uma fábrica existente.
- **Em alguns casos, a limitação de tempo e dinheiro é o fator mais importante na determinação do nível de detalhes.**
- **Se o número de aspectos de interesse é grande, um “protótipo” do modelo ou um modelo analítico podem ser utilizados para identificar os fatores importantes.**

47.2 Passos para o desenvolvimento de modelos válidos

- Descreveremos a seguir 3 etapas do processo de validação propostas por Naylor e Finger (1967). Inserimos algumas sugestões apresentadas em Law e Jain.

47.2.1 Passo 1: Desenvolva um modelo com alta validade aparente

- No início, o analista deve focar em tornar o modelo o mais razoável possível para as pessoas que conhecem o sistema.
- Para isso, ele deve utilizar toda informação disponível. A coleta destas informações já é proveitosa, mesmo que a simulação nunca seja realizada (afinal, o objetivo é resolver um problema). As principais fontes são:

Conversar com os “experts”. Um analista deve trabalhar junto com pessoas que são familiares com o sistema.

- Ex.: sistema de manufatura: operadores de máquinas, engenheiros de produção, gerentes, fornecedores, manuais.
- Raramente vai existir uma pessoa ou documento que contém toda a informação necessária.

Observar o sistema. Se existe um sistema similar ao sistema de interesse, então dados deste sistema deveriam ser utilizados.

- Deve-se tomar cuidado com a correte e representatividade (ex.: testes militares não correspondem a condições reais de combate: movimentação de tropas e fumaça).

Teoria existente. Ex.: se a taxa de chegada de clientes em um banco é constante, a teoria diz que existe grande chance do tempo entre chegadas serem observações independentes de uma VA com distribuição exponencial.

Resultados relevantes de modelos de simulação semelhantes. (Caso existam.)

Experiência/intuição. Frequentemente é necessário utilizar a experiência/intuição de alguém para supor como certos componentes complexos de um sistema funcionam.

- Principalmente se o sistema não existe em nenhuma forma.
- Agora discutiremos duas idéias que aumentam consideravelmente a chance de que o modelo seja utilizado para a tomada de decisões.
 1. O analista deve interagir regularmente o gerente (cliente) durante todo o processo. Benefícios:

- Frequentemente não existe no início uma idéia clara do problema a ser resolvido. Se o gerente acompanhar o estudo, ele vai tendo uma idéia mais clara da natureza do problema, e pode reformular o objetivo. O melhor modelo para o problema errado é claramente inválido.
 - O gerente mantém o interesse pelo estudo.
 - O conhecimento do gerente sobre o sistema contribui para a validade do modelo.
 - O modelo tem mais credibilidade, já que o gerente entende e aceita as suposições do modelo. É extremamente desejável que o gerente “assine embaixo” das suposições chave. “É um bom modelo, já que eu ajudei a desenvolvê-lo.”
2. O analista deve realizar um “walk-through estruturado” do modelo conceitual, com todas as pessoas chave, antes de iniciar a programação.
- No walk-through o analista apresenta o modelo conceitual e discute cada suposição.
 - O objetivo é garantir que as suposições estão corretas, completas e consistentes (sem contradições).

47.2.2 Passo 2: Teste empiricamente as suposições do modelo

- O objetivo agora é testar quantitativamente as suposições.
- Se uma distribuição de probabilidades teórica é utilizada, um teste de ajuste aos dados pode ser feito com gráficos (ex.: p-p plot) ou testes estatísticos (como Chi-quadrado ou Kolmogorov-Smirnov).
- Frequentemente coletamos várias amostras que julgamos ser produzidas pelo mesmo fenômeno aleatório.
 - Ex.: coletar o tempo de atendimento a clientes em diferentes dias.
 - Para determinar se podemos unir estas observações, podemos utilizar o teste de Kruskal-Wallis (K-W).
 - * Temos k amostras (consideradas independentes), cada uma com n_i observações independentes. Seja $n = \sum_{i=1}^k n_i$.
 - * Atribua rank 1 para a menor observação, rank 2 para a 2a menor, assim por diante. Seja $R_{i,j}$ o rank da j -ésima observação da amostra i .
 - * Calcule $R_i = \sum_{j=1}^{n_i} R_{i,j}$, $i = 1, \dots, k$.
 - * Então,

$$T = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1)$$

tem distribuição chi-quadrada com $k - 1$ graus de liberdade.

- Uma das ferramentas mais úteis nesta etapa é a “análise de sensibilidade”.
 - É utilizada para determinar se a saída da simulação muda de forma significativa quando alteramos (i) o valor de um parâmetro de entrada, (ii) uma distribuição de entrada ou (iii) o nível de detalhe de algum subsistema.
 - Se a saída é muito sensível a algum aspecto, então este aspecto deve ser modelado com cuidado.
 - Ex.: Suponha que os experts tenham indicado o valor 0.75 para um determinado parâmetro. Se a simulação utilizando o valor 0.7 ou 0.8 apresentar diferença significativa, devemos especificar melhor este parâmetro.
 - Ex.: Se uma simulação, que representa cada item produzido em uma fábrica, exige muito tempo de execução, podemos verificar se a saída é afetada pelo agrupamento de itens. Se não for, podemos tornar a simulação mais eficiente com o agrupamento.
 - É importante utilizar na comparação os mesmos números aleatórios.

47.2.3 Passo 3: Determine a representatividade dos dados de saída

- Se um sistema existente é similar ao que está sendo proposto, podemos validar um modelo do sistema existente e adaptá-lo para o sistema proposto.
 - Se os dados de saída do modelo forem “favoravelmente” (dependendo da função de utilidade do gerente) comparáveis com os observados no sistema real, então o modelo (do sistema **existente**) é considerado válido.
 - O modelo é então modificado para representar o sistema proposto. Quanto maior a semelhança entre os sistemas, maior a confiança.
 - Porém, não existe proposta definitiva para validar o modelo proposto (se existir, não há necessidade de um modelo de simulação).
 - Ex.: Em uma fábrica, determinar o efeito de substituir um humano por um robô, ou o efeito de uma nova política de escalonamento de produção.
 - Embora pareça desperdício de tempo e dinheiro, a modelagem do sistema existente traz algumas vantagens adicionais:
 - * Pode sugerir melhorias no sistema.
 - * Se este esforço for bem sucedido, aumenta a credibilidade do estudo.
 - * Quando poucos dados estão disponíveis para o sistema existente, este modelo facilita a comparação entre os sistemas.

- Vários testes estatísticos foram propostos para comparar a saída do modelo com os dados do sistema existente.
 - Porém, esta comparação pode ser complexa, pois geralmente os sistemas reais são **não estacionários** (as distribuições mudam com o tempo) e **autocorrelacionados** (as observações são correlacionadas com observações anteriores).
 - Assim, os testes clássicos (que assumem observações IID) não podem ser diretamente aplicados.
 - Além disso, como o modelo é apenas uma aproximação do sistema real, a hipótese nula de que os sistemas são iguais é claramente falsa.
 - Portanto, faz mais sentido verificar se a diferença é significativa o bastante para afetar as decisões baseadas no modelo.
- Além dos testes estatísticos, existe o chamado “teste de Turing”.
 - Este teste consiste em apresentar os dados do modelo e do sistema real (no mesmo formato), e pedir para pessoas que conheçam o sistema real indicar qual foi produzido pelo modelo.
 - Se eles conseguirem diferenciar, então suas explicações das diferenças ajudam a melhorar o modelo.
- Se não temos um sistema existente, ou se não temos como obter dados satisfatórios, então ainda assim é vantajoso que os experts revisem a saída do modelo.
 - Note porém que ninguém sabe exatamente como deve ser esta saída (visto que a simulação é necessária).
 - Uma animação também é uma forma efetiva dos experts avaliarem o modelo.
- Se as decisões a serem feitas com o uso do modelo são críticas, então testes de campo são utilizados em alguns casos (ex.: testes militares).
- Como geralmente o modelo muda com o tempo, e é utilizado para várias aplicações, é importante validar o modelo também com relação às observações futuras.
 - Uma vez tomada a decisão e construído o sistema proposto, podemos utilizar os novos dados gerados para validar o modelo.
 - Se nenhuma diferença significativa foi encontrada, aumentamos nossa confiança no modelo e podemos continuar usando. Se as diferenças encontradas sugerem objetivamente como melhorar o modelo, estas modificações devem ser aplicadas.
- Porém, se as modificações são aplicadas sem justificativa, dizemos que o modelo está sendo “calibrado”.
 - * Este procedimento é repetido até que a saída do modelo esteja de acordo com os dados.
 - * Porém, desta forma podemos questionar se o modelo representa o sistema ou um determinado conjunto de dados.
 - * Para responder esta questão podemos dividir os dados em um conjunto de calibragem e um conjunto de validação. Cada conjunto tem dados de entrada, e seus correspondentes dados de saída. Depois de calibrar utilizando o conjunto de calibragem, verificamos se a saída do modelo (aplicando os dados de entrada do conj. de validação) está de acordo com o conjunto de validação.

47.3 Algumas recomendações

- Experimentação com simulação é um substituto a experimentação com o sistema proposto real. Assim o objetivo ideal seria ser capaz de tomar as mesmas decisões se fosse viável testar no sistema real.
- A dificuldade do processo de validação depende da complexidade do sistema e da existência de um sistema semelhante.
 - Geralmente, quanto mais tempo (e portanto, dinheiro) gastarmos no modelo, mais válido ele deveria se tornar. Portanto, não é viável aumentar o nível de validade a partir de certo nível.
- Como os modelos de simulação são geralmente melhores para comparar alternativas do que para determinar valores absolutos, o simulador deve ser construído com um conjunto de objetivos em mente.
- Um documento com as suposições do modelo deveria ser atualizado regularmente, e em um determinado momento transformado em relatório.
 - Ou seja, o relatório não deveria ser escrito apenas no final do projeto, pois algumas suposições serão esquecidas.
 - Este relatório é útil no walk-through e para determinar outras aplicações para o simulador.
- O modelo deve ser validado relação com relação à métrica que será utilizada na tomada de decisão. Ex.: se o gerente está interessado no throughput, então avaliar a precisão na determinação do tempo de espera pode não ser importante.
- A validação não deveria ser realizada apenas no final do projeto, se sobrar tempo de dinheiro (o que ocorre com frequência). Ela deve ser considerada durante todo o estudo.