A Computational Infrastructure for Reliable Computer Simulations

J. Tinsley Oden[1], James C. Browne[1], Ivo Babuška[1], Kenneth M. Liechti[2], and Leszek F. Demkowicz[1]

[1] Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, TX 78712, U.S.A.
{oden, babuska, leszek}@ticam.utexas.edu
browne@cs.utexas.edu
http://www.ticam.utexas.edu
[2] Research Center for the Mechanics of Solids, Structures and Materials
The University of Texas at Austin
Austin, TX 78712, U.S.A.
kml@mail.utexas.edu

Abstract. The reliability of computer simulations of physical events or of engineering systems has emerged as the single most critical issue facing advancements in computational engineering and science. Without concrete and quantifiable measures of reliability, the confidence and usefulness of computer predictions are severely limited and the value of computer simulation, in general, is greatly diminished. This paper describes a mathematical and computational infrastructure for the systematic validation of computer simulations of complex physical systems and presents procedures for the integration of computer-based verification and validation processes into simulations. A general class of applications characterized by variational boundary-value problems in continuum mechanics is considered, but the approach is valid for virtually all types of models used in simulations. To simulate a particular feature or attribute of a physical event, four basic concepts are used: 1) hierarchical modeling and *a posteriori* estimation of modeling error, 2) *a posteriori* error estimation of approximation error, 3) quantification of uncertainty in the data and in the predicted response, and 4) the development of dynamic a data management paradigm, based on code composition of associated code interfaces and use of annotated source code. The result is a computational framework for computing bounds on, and estimating accuracy of, computer predictions of user-specified features of the response of physical systems.

1 Introduction

The systematic study of reliability of computer simulations of physical phenomena is generally divided into two basic processes,Verification and Validation (V and V). As a scientific discipline, V and V is in its very early stages of development. Verification is the process of determining the accuracy with which a

P.M.A. Sloot et al. (Eds.): ICCS 2003, LNCS 2660, pp. 385–390, 2003.

given mathematical model of physics is solved; validation is the process of determining that the mathematical model represents the actual physical system with sufficient accuracy. Here verification does not include software engineering issues of the logical correctness of the program which implements the simulations (this is so-called *code* verification). Verification is thus concerned with estimating and controling numerical approximation error. In recent years, significant progress has been made in this area (see Ainsworth and Oden [1], Babuska and Strouboulis [3], Oden and Prudhomme [17], Becker and Rannacher [9]).

Validation, until recently, has been primarily an experimental process in which data from physical observations are compared with computer simulations. The notion of hierarchical modeling provides a mathematical structure that can be useful in directing validation studies. In this structure, a class of models of events of interest is defined in which a "fine" model is identified which possess a level of sophistication believed to be high enough to adequately capture the event of interest with good accuracy. This model, however, may be intractable. The complexity of this model, may be such that one cannot solve it, even computationally. We then identify (families of) coarse models which are solvable. Using the fine model as a datum, the error in the solution of coarse models can be estimated and controlled, with the goal of obtaining a model best suited for the simulation goal at hand.

Physical data must be collected to characterize the fine model, and this must be furnished by data generated in actual laboratory tests. In general, these data will exhibit scatter (uncertainty) – and the basic fine-model data are then random functions. Thus the validation process is complicated by the fact that the models themselves are stochastic.

In the sections to follow, we outline briefly the program components that need to be integrated to produce a V and V computational system.

2 Problem Definition

As a representative physical situation, we consider a heterogeneous solid body D occupying a region in $\mathbb{R}^3$ and heated along its boundary ∂D and subject to an internal heat source $b - \sigma u$, u being the temperature field, b the heat supplied per unit volume and $\sigma > 0$ a reaction coefficient. The fine mathematical model is characterized by the equations,

$$\begin{aligned} -\mathrm{div} A\boldsymbol{\nabla} u + \sigma u &= b \quad \text{in } D \\ \mathbf{n} \cdot A\boldsymbol{\nabla} u &= g \quad \text{on } \partial D \end{aligned} \tag{1}$$

g being a prescribed heat flux and $\mathbf{n}$ a unit exterior normal. While the temperature field u is governed by this system, we are specifically interested in predicting certain features of the event, for instance, the temperature or the heat flux at a point $\mathbf{x_0}$ in direction $\boldsymbol{\nu}$ interior to D. These *quantities of interest* are functionals on u; e.g.

$$Q(u) = u(\mathbf{x_0}) \quad \text{or} \quad Q(u) = \boldsymbol{\nu} \cdot \mathbf{A}(\mathbf{x_0})\boldsymbol{\nabla} u(\mathbf{x_0}) \tag{2}$$

A weak form of this problem is

$$\begin{aligned} &\text{Find } u \in V \text{ such that} \\ &\qquad B(u,v) = F(v) \qquad \forall\, v \in V \end{aligned} \tag{3}$$

where V is an appropriate space of admissible functions, and

$$\begin{aligned} B(u,v) &= \int_D (A\nabla u \cdot \nabla v + \sigma uv)\, dx \\ F(v) &= \int_D bv\, dx + \int_{\partial D} gv\, ds \end{aligned} \tag{4}$$

Problem (2) (or, equivalently, (1)) is, in general, intractable. Typical complications that make it impossible or infeasible are:

1) the coefficient (thermal conductivity) A is a highly oscillatory function characterizing rapid oscillatory variations of mechanical properties at scales much smaller than the characteristic dimensions of D, and

2) the laboratory data needed to characterize A exhibit much scatter and, therefore, A is a random function (the data σ, b, and g may also be random). This uncertainty in the data must be quantified in some sense. Thus, it is desirable to develop statistical information sufficient to determine the probability density function (PDF) p for A, the variances, covariance, etc.

If p is the PDF, the mean or expected value of the random function f is $E[f] = \int_{-\infty}^{\infty} pf d\omega$. The solution u of (1) is, in this case, a random field, $u = u(\mathbf{x}, \omega)$, $\mathbf{x} \in D$ and $\omega \in \Omega$, Ω being an appropriate set of elementary events. Then, instead of (4), we use

$$\bar{B}(u,v) = \int_D E[A\nabla u \cdot \nabla v + \sigma uv]\, dx \tag{5}$$

$u, v \in V \times W$, W being a probability measure space.

3 Coarse and Discrete Models

Consistent with the theory of hierarchical modeling, we replace (2) by a simplified or coarse model,

$$u_0 \in V: \quad B_0(u_0, v) = F(v) \quad \forall\, v \in V \tag{6}$$

where

$$B_0(u_0, v) = \int_D (A_0 \nabla u_0 \cdot \nabla v + \sigma u_0 v)\, dx \tag{7}$$

Here A_0 is an effective thermal conductivity of the material obtained through homogenization methods and, for the purpose of this example, is assumed to be deterministic. The coarse model prediction of the quantity (or quantities) of interest is $Q(u_0)$.

In general, problem (7) is also intractable, but it can be solved approximately using numerical techniques such as finite elements. Thus, let V^h be a member

of a family of finite dimensional subspaces of V. We seek an approximation u_0^h of u_0 satisfying

$$u_0^h \in V^h : \quad B_0(u_0^h, v) = F(v) \quad \forall\, v \in V^h \tag{8}$$

Now, summing up our sources of error, we see that the error on the quantity of interest $Q(u)$ is,

$$\begin{aligned} Q(u) - Q(u_0^h) &= \underbrace{Q(u) - Q(u_0)}_{modeling\ error} + \underbrace{Q(u_0) - Q(u_0^h)}_{approaximation\ error} \\ &= \boldsymbol{\varepsilon}_{mod} + \boldsymbol{\varepsilon}_{approx} \end{aligned} \tag{9}$$

Estimating and controling the error component $\boldsymbol{\varepsilon}_{mod}$ is regarded as validation process (or as an aide to validation considerations), while controling $\boldsymbol{\varepsilon}_{approx}$ is a verification process. A general theory for controling this type of modeling error in quantities of interest Q through *a posteriori* error estimation and adaptive modeling has been developed by Oden and Prudhomme [19,17], and Oden and Vemaganti [15,16]. Techniques for deriving *a posteriori* error estimates for $\boldsymbol{\varepsilon}_{approx}$ and goal-oriented adaptive meshing have been advanced by Babuska and Strouboulis [3], Oden and Prudhomme [18,19], Rannacher, Becker and others [9]. Thus we assume that using the data available, it is possible to obtain computable lower and upper bounds on these components:

$$\boldsymbol{\gamma}_{mod}^{low} \leq \boldsymbol{\varepsilon}_{mod} \leq \boldsymbol{\gamma}_{mod}^{upp}, \quad \boldsymbol{\gamma}_{appox}^{low} \leq \boldsymbol{\varepsilon}_{approx} \leq \boldsymbol{\gamma}_{approx}^{upp} \tag{10}$$

Averages of these bounds can yield estimates of the errors (e.g. $\boldsymbol{\varepsilon}_{mod} \approx \frac{1}{2}(\boldsymbol{\gamma}_{mod}^{low} + \boldsymbol{\gamma}_{mod}^{upp})$). If these estimated errors exceed preset tolerances, the errors must be reduced by adaptive meshing (for $\boldsymbol{\varepsilon}_{approx}$) (see [17]) and adaptive modeling (for $\boldsymbol{\varepsilon}_{mod}$).

Another source of error relevant to validation is due to the randomness of the coefficients A. There are several approaches available to quantify this error. One direct approach is provided by the perturbation method described by Kleiber and Hien [14], which, for a first order approximation, takes the form

$$A(\mathbf{x}, \omega) = \bar{A}(\mathbf{x}) + \varepsilon A_1(\mathbf{x}, \omega), \quad u(\mathbf{x}, \omega) = \bar{u}(\mathbf{x}) + \varepsilon u_1(\mathbf{x}, \omega) \tag{11}$$

$\bar{A}$ and $\bar{u}$ being mean values of A and u. Then one treats first the deterministic fine problem (1) with the forms (4), involving $\bar{A}$ instead of A and $\bar{u}$ instead of u. Then $Q(u) = Q(\bar{u}) + \varepsilon Q(u_1)$. The random field $Q(u_1)$ is completely characterized in a stochastic sense, once the probability field of $A(\mathbf{x}, \omega)$ is provided and deterministic component $\bar{u}$ is known. Bounds on the truncation error inherent in the perturbation (11) can be calculated and used as additional measure of the modeling error [4]. The second approach is based on the theory of stochastic functions characterized by Karhunen-Loève expansion (see [5,11,6]). In [7], the problem of determination of the Karhunen-Loève expansion from experimental data is described. A brief survey [8] of the mathematics of verification and validation is presented when the experimental data [2] are used.

4 The Integration of Computational Components

Separate codes are written or are available for the component error estimations and adaptivity. The integration of these modules requires a major new dynamic data structure innovation to coordinate a component-oriented code development process, and requires a dynamic scaling of common data structures. An integrated compiler and runtime system is being developed to compose appropriate components and map data structures across interfaces (see [10,12,13]). Annotations of source code is to be used to guide compilation and enable interfacing different data structures. Different computational models provide for the specification of associative interfaces and the annotation language. The essential integration components are as follows:

1) Experimental data are collected to fully characterize the fine model, including statistics to give bounds on the data or PDF's and covariance matrices.

2) Quantities (or a quantity) Q of interest are (is) specified as the target physical entity to be predicted in the simulation (perhaps in the form of probability of the predicted values of the quantity).

3) The coarsest model is used to extract a preliminary estimate of Q and modeling and approximation errors are computed.

4) If the estimated error exceeds the prescribed tolerance, the model is enhanced and the calculation is repeated, each time adapting the computational model to reduce ε_{approx} until a model yielding results within the preset bounds is obtained.

5) The truncation error of the perturbation expansion is estimated; if the total error exceeds a preset tolerance, the data set and the fine model definition must be updated. If not, the predicted Q and the probability that it will takes on values in a given interval are produced as output.

In the absence of sufficient data to adequately characterize the random data fields, a worse-case analysis can be done giving upper and lower bounds on the predicted values.

The implementation of this and related algorithms is underway and will be the subject of forthcoming reports.

Acknowledgment. The assistance of Dr. Yusheng Feng and Dr. Fabio Nobile on this project is gratefully acknowledged. This work is done under ITR NSF grant 0205181. The support and encouragement of Dr. Frederica Darema of NSF is greatly appreciated by the authors.

References

1. Ainsworth, M., Oden, J.T.: A Posteriori Error Estimation in Finite Element Analysis. John Wiley & Sons, New York, (2000)
2. Babuška, I., Jerina, K., Li, Y., Smith, P.: Quantitative Assessment of the Accuracy of Constitutive Laws for Plasticity with Emphasis on Cyclic Deformation. Material Parameter Estimation for Modern Constitutive Equations. MD-Vol.**43**/AMD-Vol.**168** ASME (1993)

3. Babuška, I., Strouboulis, T.: Finite Element Method and its Reliability. Oxford Univ. Press (2001)
4. Babuška, I., Chatzipantelidis, P.: On Solving Elliptic Stochastic Partial Differential Equations. Comput. Methods Appl. Mech. Eng. **191** (2002) 4093-4122
5. Babuška, I., Tempone, R., Zouraris, G.E.: Galerkin Finite Element Approximations of Stochastic Elliptic Differential Equations. TICAM Report 02-38, The University of Texas at Austin, (2002)
6. Babuška, I., Liu, K-M.: On Solving Stochastic Initial-Value Differential Equations. Math. Models Methods Appl. Sci., (2003) In press
7. Babuška, I., Liu, K-M., Tempone, R.: Solving Stochastic Partial Differential Equations Based on Experimental Data. Math. Models Methods Appl. Sci., (2003) In press
8. Babuška, I.: Mathematics of the Verification and Validation in Computational Engineering. In: Proceeding of Int. Conf. on Mathematical and Computer Medeling in Sci. Eng., Prague Czech Republic (2003)
9. Becker, R., Rannacher, R.: An optimal control approach to *a posteriori* error estimation in finite element methods. In: Isereles, A. (ed.): Acta Numerica, Vol. 10. Cambridge Univ. Press (2001) 1-120
10. Browne, J.C., Berger, E., Dube, A,: Compositional Development of Performance Models in POEMS. Int. J. High Perf. Comput. Appl. **14**(4) (2000) 283–291
11. Deb, M.K., Babuška, I., Oden, J.T.: Solution of Stochastic Partial Differential Equations Using Galerkin Finite Element Techniques. Comput. Methods Appl. Mech. Eng. **190** (2001) 6359-6372
12. Edwards, C, Browne, J.C.: Scalable Dynamic Distributed Array and its Application to a Parallel hp-adaptive Finite Element Code. Proceedings of the 1996 POOMA Workshop, Santa Fe, New Mexico, (1996)
13. John, A., Browne, J.C.: Compilation of Constraints Programs with Noncyclic and Cyclic Dependencies to Procedural Parallel Programs, Int. J. Parallel Prog. **26**(1) (1998) 65-119
14. Kleiber, M. Hien, T.D.: The Stochastic Finite Element Method: Basic Perturbation Technique and Computer Implementation. John Wiley & Sons, (1992)
15. Oden, J.T., Vemaganti, K.: Adaptive Hierarchical Modeling of Heterogeneous Structures. Predictability: Quantifying Uncertainty in Models of Complex Phenomena. Phys. D., **133** (1999) 404-415
16. Oden, J.T., Vemaganti, K.: Estimation of Local Modeling Error and Goal-Oriented Modeling of Heterogeneous Materials; Part I: Error Estimates and Adaptive Algorithms. J. Comput. Phys., **164** (2000) 22-47
17. Oden J.T., Prudhomme, S.: Goal-Oriented Error Estimation and Adaptivity for the Finite Element Method. Comput. Math. Appl., **41**(5-6) (2001) 735-756
18. Oden, J.T., Prudhomme, S.: Estimation of Modeling Error in Computational Mechanics. J. Computat. Phys., **182** (2002) 496-515
19. Prudhomme, S., Oden, J.T.: On Goal-Oriented Error Estimation for Elliptic Problems: Applications to the Control of Pointwise Errors. Comput. Methods Appl. Mech. Eng., **176** (1999) 313-331