INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 3, Number 3, Pages 255–272

THE RELIABILITY OF COMPUTER PREDICTIONS: CAN THEY BE TRUSTED?

IVO BABUŠKA AND J. TINSLEY ODEN

(Communicated by Zhimin Zhang)

Abstract. The issue of reliability of computer predictions of physical events is examined as the goal of verification and validation processes. It is argued that verification, both solution and code verification, can be carried out with a high degree of confidence, even though much remains to be done to improve and advance verification procedures. It is validation of mathematical models that stands as the major bottleneck of reliable computer predictions. Uncertainty of input data represents a major feature of validation processes and must be quantified if models are to be judged valid or invalid. Examples are given from solid mechanics and heat transfer that demonstrate validation processes employing stochastic models and fuzzy set theories.

Key Words. Verification, validation, computer simulations.

1. Introduction

Computational Science (CS) is the discipline concerned with the use of computers and computational methods to simulate physical events and to make quantitative predictions of physical phenomena. Such predictions are often used as a basis for critical decisions effecting the health, security, and well being of humankind. For this reason, CS has had a major impact on many fields in engineering, physics, chemistry, health sciences, economics, finance, politics, and other areas. The great promise that CS will be of immense value to all areas of science and technology depends on a crucial factor: the reliability of computer products and our ability to measure in some way this reliability [34].

In a recent report, Post [38] speaks of the coming crisis in CS arising from three major challenges: a) performance, b) programming, and c) prediction. We agree with Post that the performance and programming challenges have been met or will be met soon, but the prediction challenge will require considerable advancement and maturity in the way that simulation is done and interpreted.

Concerning performance, a look at the history of computer performance over the last one-and-a-half decades gives weight to the viewpoint that the so-called "performance challenge" of CS is well in hand. Using the 11/780 VAX as a unit measuring computer capabilities a decade before the end of the twentieth century (1989), the unit involves one megabyte of memory, a half gigabyte of disk storage, and one cpu with a speed of 0.1 megaflops. In 1992, three years later, the IBM

Received by the editors December 16, 2004 and, in revised form, July 1, 2005.
2000 *Mathematics Subject Classification.* 65N30.
The support of the U.S. National Science Foundation under ITR grant EIA-0205181, the Office of Naval Research under grants N0014-95-0401 and N00014-99-1-0724, and Sandia National Laboratory Project 268687 is gratefully acknowledged.

RS580 had 54 times the memory of the VAX and its speed was 1000 times greater. In 1997, eight years later, the SGI Power Challenge had 1000 times the memory of a VAX, six cpu units with a theoretical speed of 18,000 times that of the 1989 VAX. In November of 2004, IBM's Blue Gene/L supercomputer became operational with an expected peak performance of 36.0×10^{12} flops, over a hundred million times faster in unit capability than the VAX. Today, $1000 can buy a computer as powerful as the biggest and most capable computer available at any cost in 1990. Experts predict that Moore's law will continue to hold for 15-20 years so that this exponential growth in performance will continue [21]. These data convince us that the performance challenge is being met.

The programming challenge, supplying the software to keep up with hardware developments, is also being met, albeit at a slower pace. Significant progress in the development of new languages and new devices gives confidence that the programming challenge is being reasonably addressed and can be met even more vigorously in the years ahead.

The prediction challenge, which is at the heart of CS, is viewed as the most difficult challenge to be met in the future, and stands as a major bottleneck, perhaps a crisis, in CS. Again, the major issue is the reliability of computer predictions and their use as a basis for important decisions.

In the present paper, we will discuss the prediction challenge, comment on various mathematical aspects of it, and point to some serious limitations of contemporary methods of computer prediction. We will also address the question of what machinery must be developed in order to use such predictions to make meaningful decisions.

2. Verification and Validation in CS

Verification and Validation (V & V) has emerged in recent years as the intellectual and technological discipline that addresses the prediction challenge. Both are processes, verification being the processes addressing the quality of the numerical approximation of the mathematical model used as the basis for a prediction, and validation being the process addressing the reliability of the mathematical model as a faithful abstraction of reality. V & V has been the focus of much study and debate in recent years and a relatively large literature exists and is expanding (e.g.,[19], [23], [33], [39], [42]).

In [11], we present a detailed list of definitions and concepts related to V & V. Worthy of mentioning here are the following:

- *physical event*: an occurrence in nature, a fundamental entity of physical reality
- *simulate*: to build a likeness; in our case, the likeness is produced by the interpretation of prediction results produced by a computer
- *mathematical model*: a collection of mathematical constructions that provide abstractions of physical events, based on scientific theories covering the event, and available input information
- *computational model*: a discrete approximation of a mathematical model rendered in a form manageable by a computer or an appropriate computing device.

A typical example of a mathematical model is the set of equations and conditions characterized by a boundary value problem involving deterministic or stochastic differential equations together with functionals defining quantities of interest. These quantities of interest are the goals of the computer prediction, which, in turn, are the basis for decisions.