SAND2006-1514P

March 10, 2006

From Pencils to Computers

Timothy G. Trucano

Optimization and Uncertainty Estimation, Organization 01411
Sandia National Laboratories
Albuquerque, NM 87185

tgtruca@sandia.gov

Pencil and paper (**PnP**) theoretical physics is highly correlated with mathematical subject matter and mathematical ability on the part of the researcher. One could rationally argue that the boundaries of mathematical technique in theoretical physics prior to the widespread use of large-scale computation were defined by the themes, content, and applications of Morse and Feshbach's classic Methods of Theoretical Physics (Morse and Feshbach, 1953). By the 1980's this situation had changed dramatically and permanently. The era of Computational Physics (**CmP**) was underway, and has proceeded with exponential growth to the present. We are now far more likely to find skill sets associated with developing and applying computational models in modern theoretical physics publications than skill sets mirroring the mathematical tradecraft in Morse-Feshbach (1953).

Of course, to the degree that **CmP** has replaced **PnP**, what is also implied is that the ***scope*** of problems now addressed in theoretical physics goes far beyond what analytic mathematical procedures can decisively attack. This is why computation in physics has become so important. Beyond this fact, however, there is the important observation that discourse associated with **CmP** is ***different*** than that associated with **PnP** and the difference resides more in the medium than in the message. It is this last observation that I presume is provocative. I think that we can also draw some additional conclusions that intersect concerns reflected in ongoing research on the influence of **C**ognition and **U**ncertainty in complex **D**ecisions (CUD) at Sandia. These conclusions, briefly discussed below, are loosely organized around the concepts of verification, validation and the evaluation of credibility for theoretical models and seem to imply complication, if not degradation, of computation-informed decision processes. An important current example of where these concerns may play out is Quantitative Margins and Uncertainty (QMU) for the NNSA Stockpile Stewardship Program.

In the context of physics, ***verification*** is the problem of establishing with required levels of rigor (say for purposes of publication) the mathematical correctness and accuracy of a theoretical physics model. For example, if the model is an algebraic equation, verification centers on the problem of ensuring that when outputs are calculated from inputs, the output values are mathematically accurate. This, in turn, requires that in the original equation, "plus" signs did not accidentally replace the required "minus" signs, and that "multiplication" was not accidentally performed instead of a required "division." In Morse and Feshbach, the problem of verification becomes increasingly complex over that for simple algebraic equations, for example to provide some guarantee that the mathematics of analytic procedures for investigating the solution of partial differential equations is (1) correctly generated and (2) accurately evaluated when applied to problems. Despite the potential complexity of the mathematics (a randomly chosen example is Biedenharn and Blatt, 1954, which has *no references*, hence an implied mathematical skill level on the part of both reviewers and readers), for **PnP** verification is logically strongly enabled and made relatively decisive by the social dynamics of colleagues, reviewers and readers of published material. In other words, if you know Morse and Feshbach then you are capable of critically reviewing, as well as re-creating, the key theoretical work and testifying to its underlying mathematical correctness and accuracy in a given theoretical publication. And there was a time when it was expected that a well-trained theoretical physicist would know Morse and Feshbach. This, of course, is not intended to suggest that gross mathematical mistakes, such as incorrect interchange of integrals and limits, might not creep into published papers. But this is based on the potential for human failure more than the unbounded nature of the verification problem.

The problem of verification is fundamentally different for **CmP**. This is true for many and varied reasons, but it centers on different paradigms for creation, aggregation, and communication of information before, during, and after publication; the greater uncertainty underlying the generated information; and more formidable cognitive barriers to understanding what the information is telling us for purposes of verification. The social dynamics that critically assesses and may need to re-create **CmP**, engaged in by colleagues, reviewers and readers of publications, is very different, with less definitive results about correctness and accuracy. I question whether the social dynamics and discourse that are imperative for scientific progress in **PnP** have much relation at all to their counterparts in **CmP**. As only one example, it is highly unlikely that any reviewer of a current article in the American Physical Society journal Physics of Plasmas can reproduce the content of multi-dimensional, multi-physics calculations that may be present, let alone provide independent evidence that the calculations are "verified" beyond the level of intuitive judgment (the calculations "look right"). The consequences of this differentiation, if it is as large as I claim, may be quite grave.

The differences certainly reflect different challenges for evaluating the credibility of results, and this must influence the confidence with which the results can be used. For one thing, verification differences between **PnP** and **CmP** become prominent in ***validation***, which is the evaluation of physical fidelity of theoretical physics constructs via comparison with experiment. The significant quality control and more definitive conclusions achieved for verification in **PnP** often effectively remove mathematical accuracy error bars from theoretical – experimental comparisons. This bases the interpretation and implications of such a comparison on nothing but physics. The opposite is often the case in **CmP**, where mathematical error bars may dwarf the comparison with experiment, making interpretation and inferring conclusions difficult and potentially dangerous. Even worse, it is possible that mathematical accuracy is unknown, hence a major contributor to uncertainty in the experimental comparison. This quickly destroys the purpose and results of performing validation. The social factors that we mentioned in the context of verification are also prominent in validation, and magnify the difficulties created by coupling poorly achieved verification into validation.

The reason verification and validation are prominent is that these are the accepted community techniques for evaluating the quality of theoretical physics (Oberkampf, Trucano, and Hirsch, 2003). They assess ***credibility*** of theoretical constructs. Poor performance of verification and validation creates fundamental uncertainty in **CmP** that corrupts conclusions and community discourse. There is little counterpart of this phenomenon in **PnP**. Sadly, poor credibility in **CmP** inevitably undermines the decisions that might use it. One may be inclined to ask the extreme question: “Why bother using **CmP** at all?” The answer, of course, is the perceived extreme usefulness of **CmP** even if verification and validation are lacking. Caution is suggested nonetheless.

Our inability to establish precise and rigorous components of credibility in **CmP** is interesting from the perspective of CUD. As one illustration, I stress that perceptual shifts in **CmP** from quantitative to qualitative discourse sometimes results from community self-awareness of inadequate credibility. This is summarized by statements such as “The calculation provides insight into the problem, but we don’t expect it to be rigorously mathematically accurate.” The shift of emphasis to “intuition” is often built on tools that emphasize the intuitive content of giant masses of computational data, such as elaborate 2-D and 3-D temporally sequenced color visualizations, and therefore create subtle content filters that may further enlarge cognitive challenges to communication.

This statement could not be made about a **PnP** model, with its attendant (in theory) well-understood and *expected mathematical accuracy*, although the value of such a model might be claimed to be mainly intuitive from the *physics fidelity* perspective. For a **CmP** counterpart, I believe it is actually fair to complain about the value of physics “intuition” that might result from a model in which mathematical accuracy is poorly understood or nonexistent.

Claiming “insight” as the purpose of highly quantitative but inadequately mathematically characterized calculations is a shift of content of the relevant theoretical construct from the analytic end of the cognitive spectrum toward the intuitive end. There is no particular reason to believe that such a procedure does not increase cognitive complexity for decision processes. Introducing intuition in a decision problem that is basically analytic, such as ignoring verification hurdles in **CmP** results, creates the same kind of confusion that results when overly analytic information is introduced into basically intuitive decision problems. These phenomena are well-recognized in the judgment theory literature (Cooksey, 1996; Hastie and Dawes, 2001), but may be poorly acknowledged by the **CmP** community. These difficulties are transformed in complex ways in the social dimensions that are associated with complex and high-consequence decision processes, and influence decision procedures in ways that we don’t really understand.

Uncertainty in the credibility of **CmP** is an example of a problem for decision making that we need to worry about. The uncertainty that I have briefly hinted at in **CmP** contains components of both fundamental variability (*aleatory uncertainty* in the jargon) and lack-of-knowledge (*epistemic uncertainty*). The resulting *heterogeneous uncertainty* (Helton, 1994) in important information places great stress on complex decisions that must use that information. It is wise to keep in mind that the most complex decisions we face in national security, for example problems of intelligence analysis (Heuer, 1999), must combine both analytic and intuitive judgment in an optimal way, and this is made ever more difficult by the presence of heterogeneous uncertainty. Is it really so obvious that **CmP** is better than **PnP** in these decisions, despite the promise of a richer and deeper engagement with "physical reality?"

Broad characteristics of the epistemological transition from **PnP** to **CmP** are of special interest when we consider computational social models. All of the cognitive challenges become much harder and more open-ended in computational social modeling, raising the level of complexity as to how to apply these models in critical decision making (Turnley, 2004). Beyond this, I stress that really important decisions are social (i.e. organizational) products. (In fact, it is probably correct to assume that *all* decisions are social processes.) Cognitive limitations and heterogeneous uncertainty components in computational information have the potential for disrupting essential community discourse and may fatally fragment decision processes that require a rich level of social organization.

While it is easy to believe that, in some respects, computing is indeed replacing pencil-and-paper (at least in physics), I recommend that the following conclusions be kept in mind. First, replacing pencils with computers is not in and of itself the introduction of a third leg of science. To the extent that a massive calculation is only conceived as a straightforward epistemological replacement of a pencil-and-paper result, there is going to be real trouble for any decision process that needs to understand the generated information. Second, the cognitive burden introduced by computing, certainly in physics and far more in fields such as social science, is dramatically greater than that associated with **PnP**. This means that "more" or "better" **CmP** doesn't necessarily yield "better decisions." Third, the richer presence of uncertainty in **CmP**, and the potential degradation of needed credibility, means that what "better" means can be hard to determine in important problems.

Decision processes that work effectively with experimental and pencil-and-paper theoretical information streams are not assured to work with complex computational information streams that have more poorly understood efficacy, salience and credibility. The challenges I have discussed here are graver for social science, which has less experience of clean information emerging from experiment and pencil-and-paper theory. The real task is then how to properly manage the attendant risk in decision processes. This is a topic that will continue to engage SNL as we further investigate the relationship of computational cognitive burden and uncertainty to complex decisions.

References

L. C. Biedenharn and J. M. Blatt (1954), "A Variation Principle for Eigenfunctions," Physical Review, Volume 93, Number 1, 230-232.

R. Hastie and R. M. Dawes (2001), Rational Choice in an Uncertain World, Sage.

P. Morse and H. Feshbach (1953), Methods of Theoretical Physics, McGraw-Hill.

R. Cooksey (1996), Judgment Analysis: Theory, Methods, and Applications, Academic Press.

J. C. Helton (1994), "Treatment of Uncertainty in Performance Assessments for Complex Systems," Risk Analysis, Volume 14, Number 4, 483-511.

R. J. Heuer (1999), Psychology of Intelligence Analysis, CIA, Sherman Kent School Publication.

W. L. Oberkampf, T. G. Trucano and C. Hirsch (2003), "Verification, Validation and Predictive Capability in Computational Engineering and Physics," Applied Mechanics Reviews, 2004, Volume 57, Number 5, 345-384.

J. G. Turnley (2004), "Validation Issues in Computational Social Simulation," Galisteo Consulting Group.

Acknowledgements

I thank Laura McNamara, Jessica Glicken Turnley, Gerold Yonas, George Backus and William Oberkampf for comments on this manuscript. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.