Chapter Nine

# EXPLORATORY ANALYSIS AND IMPLICATIONS FOR MODELING

***Paul K. Davis***

The theme that runs through this book is that real-world strategy problems are typically beset with enormous uncertainties that should be central in assessing alternative courses of action. In the past, one excuse for downplaying uncertainty—perhaps treating it only through marginal sensitivity analysis around some "best-estimate" baseline of dubious validity—was the sheer difficulty of doing better. If analysis depended on the time it took to set up and run computer programs, then extensive uncertainty work was often ruled out. Today, however, extensive uncertainty analysis can be done with personal computers. Better software tools are needed, but existing commercial products are already powerful. There is no excuse for not doing better.

A key to treating uncertainty well is *exploratory analysis*. Its objectives are to (1) understand the implications of uncertainty for the problem at hand and (2) inform the choice of strategy and subsequent modifications. In particular, *exploratory analysis can help identify strategies that are flexible, adaptive, and robust*.[1] This chapter describes exploratory analysis, puts it in context, discusses enabling technology and theory, points to papers applying the ideas, and concludes with some challenges for those building models or developing enabling technology for modeling and simulation.

---

[1]Paul K. Davis, *Analytic Architecture for Capabilities-Based Planning, Mission-System Analysis, and Transformation*, MR-1513-OSD, RAND, 2002.

## EXPLORATORY ANALYSIS

### Definition

Exploratory analysis examines the consequences of uncertainty. In a sense, it is sensitivity analysis done right. Yet because it is so different in practice from what most people think of as sensitivity analysis, it deserves a separate name. It is closely related to scenario space analysis,[2] which dates back to 1983,[3] and to "exploratory modeling."[4] It is particularly useful for gaining a broad understanding of a problem domain before dipping into details. That, in turn, can greatly assist in the development and choice of strategies. It can also enhance "capabilities-based planning" by clarifying *when* (e.g., in what circumstances and with what assumptions about other factors) a given capability (e.g., an improved weapons system or enhanced command and control [C2]) will likely be sufficient or effective. This contrasts sharply with establishing a base-case scenario and an organizationally blessed model and database, and then asking, "How does the outcome for this scenario change if I have more of this capability?"

Exploratory analysis is an exciting development with a long history, including work in the 1980s and 1990s with RAND's RSAS (RAND Strategy Assessment System) and JICM (Joint Integrated Contingency Model). It is, however, only one part of a sound approach to analysis generally—a point worth pausing to emphasize.

Figure 9.1 shows how different types of models and simulations (including human games) have distinct virtues. The figure is specialized to military applications, but a more generic version applies broadly to a wide class of analysis problems. White rectangles indicate

---

[2]Paul K. Davis (ed.), *New Challenges for Defense Planning: Rethinking How Much Is Enough,* MR-400-RC, RAND, 1994.

[3]Paul K. Davis and James A. Winnefeld, *The RAND Strategic Assessment Center: An Overview and Interim Conclusions About Utility and Development Options*, R-2945-DNA, RAND, 1983.

[4]Stephen C. Bankes, "Exploratory Modeling for Policy Analysis," *Operations Research,* Vol. 41, No. 3, 1993; and Robert Lempert, Michael E. Schlesinger, and Steven C. Bankes, "When We Don't Know the Costs or the Benefits: Adaptive Strategies for Abating Climate Change," *Climatic Change,* Vol. 33, No. 2, 1996.

“good”—i.e., if a cell of the matrix is white, the type of model indicated in the left column is very effective with respect to the attribute indicated in the cell’s column. In particular, analytic models (top left corner), which have low resolution, can be especially powerful with respect to their analytic agility and breadth. In contrast, they are very poor (have dark cells) with respect to recognizing or dealing with the richness of underlying phenomena, or with the consequences of both human decisions and behavior. In contrast, field experiments often have very high resolution (they may even be using the real equipment and people) and may be good or very good for revealing phenomena and reflecting human issues. They are, however, unwieldy and inappropriate for studying issues in breadth. The value of the type model for the particular purpose can often be enhanced a notch or two if the models include sensible decision algorithms or knowledge-based models that might be in the form of expert systems or artificial-intelligence agents.

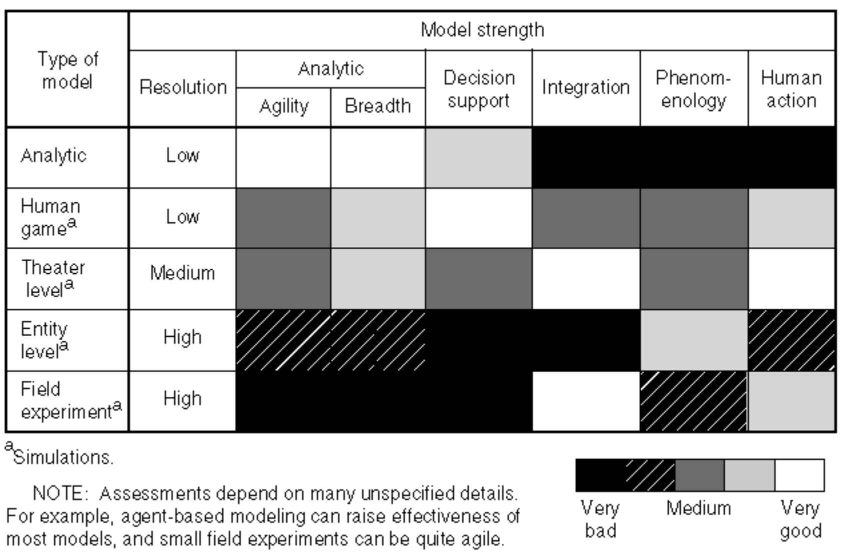



**Figure 9.1—Virtues of a Family of Models (Including Human Games)**

Figure 9.1 is an exhortation to the Department of Defense (DoD) regarding the need to have *families of models and families of analysis*.[5] Unfortunately, it is usual for government agencies to depend more or less exclusively on a single model, which is a serious shortcoming.

Figure 9.1 reveals a niche for exploratory analysis: the top left-hand corner of the matrix, which emphasizes analytic agility and breadth of analysis. However, the technique can be used hierarchically with a suitably modularized system model. That is, one can do top-level exploration first, and then zoom in to explore in more detail—but using the same techniques—the consequences of various details within particular modules. This is easier said than done, however, especially with traditional models. Specially designed models make things much easier, as discussed in what follows.

## Types of Uncertainty in Modeling

Exploratory analysis is about addressing uncertainty, but uncertainty comes in many forms. Parametric uncertainties arise from a model's inputs, from not knowing their precise values. They are not the same as structural uncertainties, which relate to questions about the form of the model itself: Does it reflect all the variables on which the real-world phenomena described by the model depend? Is the analytic form of the dependencies correct? Some uncertainties may be inherent because they represent what are called stochastic processes—the randomness of nature.[6] Some come from fuzziness or imprecision; some reflect discord among experts. Some relate to knowledge about the values of well-defined parameters; others refer to future values not yet known.

It is convenient to express uncertainties parametrically. Even when unsure about the correct form of the model, one can reflect uncertainty to some extent by having parameters that affect that form. For example, parameters may control the relative size of quadratic and exponential terms in an otherwise linear model. Or a discrete pa-

[5] Paul K. Davis, James H. Bigelow, and Jimmie McEver, *Analytical Methods for Studies and Experiments on "Transforming the Force,"* DB-278-OSD, RAND, 1999.

[6] The behavior of stochastic systems has a random component. Such systems are described with probabilistic equations.

rameter that is essentially a switch might determine which of a set of distinct analytic forms applies. Some parameters may apply to the fixed aspect of a model; others may apply to a random aspect. In taking this approach, one needs to keep straight how the different uncertainties come into play.[7]

## Types of Exploratory Analysis

**Ways to Conduct Exploratory Analysis.**[8] One form of exploratory analysis is input, or parametric, exploration, which involves running models across the space of cases defined by plausible discrete values of the parameters. This is done not one at a time, as in normal sensitivity analysis, and not around some presumed base-case set of values, but, rather, for all the combinations of values defined by an experimental design. The results, which may number from dozens to hundreds of thousands or more, can be explored interactively with modern displays. Within perhaps one half-hour, a good analyst can often gain numerous important insights that were previously buried. He can understand not just which variables "matter," but when they matter. For example, he may find that outcomes are insensitive to a given parameter for the so-called base case but are quite sensitive for other plausible assumptions. That is, he may identify *when* the parameter is important. For capabilities-based planning for complex systems, this can be distinctly valuable.[9]

---

[7]See the appendix for an example that uses Lanchester equations, chosen for familiarity rather than for current usefulness.

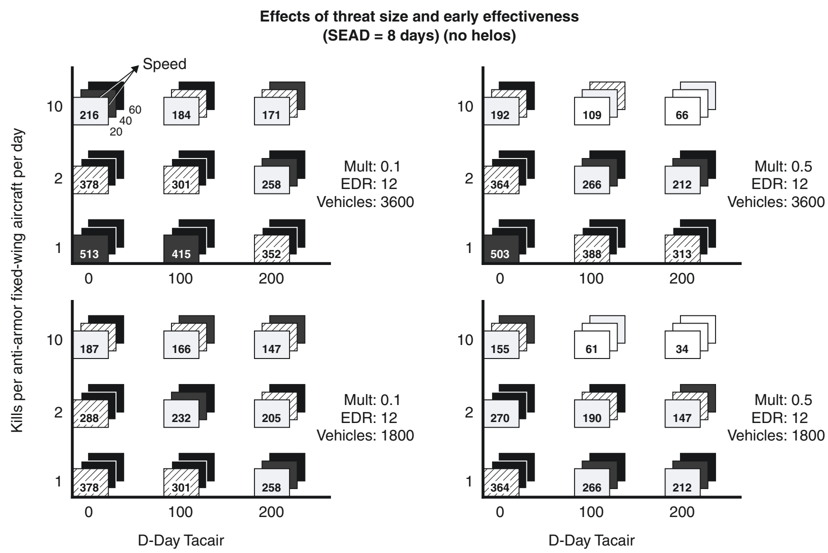
[8]See Paul K. Davis, David C. Gompert, and Richard L. Kugler, *Adaptiveness in National Defense: The Basis of a New Framework,* IP-155, RAND, 1996; Bankes, "Exploratory Modeling"; Arthur Brooks, Steve C. Bankes, and Bart Bennett, *Weapon Mix and Exploratory Analysis: A Case Study,* DB-216/2-AF, RAND,1997; Lempert, Schlesinger, and Bankes, "When We Don't Know"; National Research Council, *Modeling and Simulation,* Vol. 9, *Technology for the United States Navy and Marine Corps, 2000–2035,* National Academy Press, Washington, DC, 1997; and Granger Morgan and Max Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis,* Cambridge University Press, Cambridge, MA, 1992 (reprinted in 1998). The book by Morgan and Henrion is an excellent treatment of uncertainty in policy analysis generally.

[9]Many examples in military analysis involve warning time. Some capabilities, such as those of forward-deployed systems, are especially important only when warning time is short. If standard scenarios assume considerable warning time, such capabilities can be undervalued.

A complement to parametric exploration is probabilistic exploration, in which uncertainty about input parameters is reflected by distribution functions representing the totality of one's so-called objective and subjective knowledge. Using analytic or Monte Carlo methods, the resulting distribution of outcomes can be calculated. This can quickly give a sense for whether—all things considered—uncertainty is particularly important. In contrast to displays of parametric exploration, the output of probabilistic exploration gives little visual weight to improbable cases in which various inputs all have unlikely values simultaneously. Probabilistic exploration can be very useful for a condensed "net assessment." Note that this use of probability methods is different from using them to describe the consequences of a stochastic process within a given simulation run. Indeed, one should be cautious about using probabilistic exploration, because one can readily confuse variation across an ensemble of possible cases (e.g., different runs of a war simulation) with variation within a single case (e.g., fluctuation from day to day within a single simulated war). An unknown constant parameter for a given simulated war is no longer unknown once the simulation begins, and simulation agents representing commanders should perhaps observe and act upon the correct values within a few simulated days. Despite these subtleties, probabilistic exploration can be quite helpful.

After initial work with both parametric and probabilistic exploration, the preferred approach is *hybrid exploration*. It may be appropriate to parameterize a few key variables that are under one's control (purchases, allocation of resources, and so on) while treating the uncertainty of other variables through uncertainty distributions. Analysts might also want to parameterize a few of the principal variables characterizing the future context in which strategy must operate. In military affairs, one might parameterize assumed warning time and size of threat. There is no general procedure here; instead, the procedure should be suitable to the problem at hand. In any case, the result of such exploratory analysis can be a comprehensible summary of how known classes of uncertainty affect the problem at hand.

Consider the following examples of exploratory analysis. Figure 9.2 displays a mid-1990s parametric exploration of what is required militarily to defend Kuwait against a future Iraqi invasion by interdicting




Mult.: fraction of shooters used before SEAD time; EDR: equivalent deployment rate; Vehicles: AFVs to be killed for halt.

**Figure 9.2—Data View Display of Parametric Exploration**

the attacker's movement with aircraft and missiles.[10] Each square in the figure represents a particular model case (i.e., a specific choice of all the input values). Taken together, the four panels summarize parametric exploration in five variables (the x, y, and z axes of each panel, one for a row of two panels, and one for a column of two panels). With four such pages, one can cover results for seven variables. The outcome of a given simulation is represented by the pattern of a given square. A white square represents a good case, in which the attacker penetrates only a few tens of kilometers before being halted. A black square represents a bad case, in which the attacker penetrates deep into the region that contains critical oil facilities. The other patterns represent in-between cases. The number in each square gives the penetration distance in kilometers.

[10]Paul K. Davis and Manuel J. Carrillo, *Exploratory Analysis of "The Halt Problem": A Briefing on Methods and Initial Insights,* DB-232, RAND, 1997.

To generate such results for a sizable scenario space, RAND has often used a program called Data View.[11,12,13] After running the thousands or hundreds of thousands of cases corresponding to an experimental design for parametric exploration, one explores the outcome space by choosing interactively which of the parameters to vary along the x, y, and z axes. The remaining parameters then have values shown alongside the graph. These can be changed by clicking on one of them and selecting from the menu that comes up.

Figure 9.3 shows a screen image from some recent work with Analytica on the same interdiction-of-invader-forces problem treated in Figure 9.2. In this case, the graphical display of results is more traditional. Outcome is measured along the y axis rather than by a color or pattern, and one of the independent variables is plotted along the x axis. A second variable—D-Day (the day war commences) shooters—is reflected in the family of curves. The other independent variables appear in the rotation boxes at the top. As with Data View, one changes parameter values by selecting from a menu of values. Such interactive displays allow one to "fly through the outcome space" for many independent variables (parameters), in this case nine. More parameters could have been varied interactively, but the display was still quickly interactive for the given model and computer being used (a Macintosh PowerBook G3 with 256 MB of RAM).

---

[11]This was developed at RAND in the mid-1990s by Stephen Bankes and James Gillogly.

[12]Other personal-computer tools can be used for the same purpose and the state of the art for such work is advancing rapidly. A much improved version of Data View, called CARS™, is under development by Evolving Logic (www.evolvinglogic.com). For those who do their modeling with Microsoft EXCEL™, there are plug-in programs that provide statistical capabilities and some means for exploratory analysis. Two such tools are Crystal Ball® (www.decisioneering.com) and Risk® (www.palisade.com/html/risk.html). For a number of reasons, however (visual modeling, array mathematics, etc.), my colleagues and I have in recent times most often used the Analytica® modeling system. Analytica is an outgrowth of the Demos system developed by Max Henrion and Granger Morgan at Carnegie Mellon University; it is marketed by Lumina (www.lumina.com).

[13]For more recent exploratory analysis work, see Paul K. Davis, Jimmie McEver, and Barry Wilson, *Measuring Interdiction Capabilities in the Presence of Anti-Access Strategies: Exploratory Analysis to Inform Adaptive Strategy in the Persian Gulf*, MR-1471-AF, RAND, 2002.

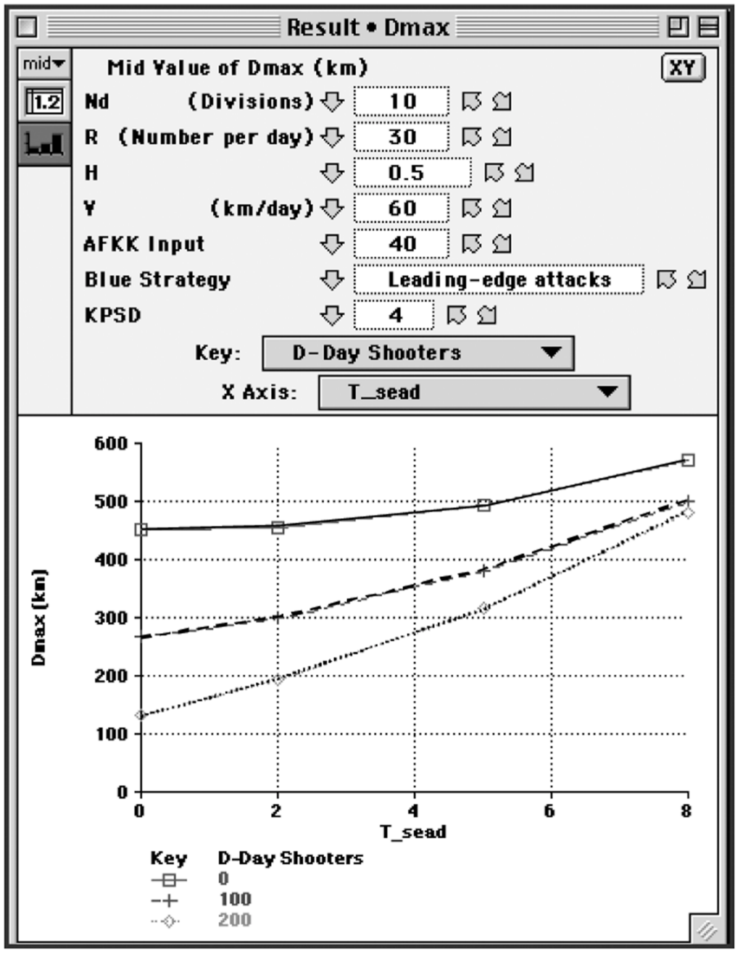



**Figure 9.3—Analytica Display of Parametric Exploration (Simultaneous Exploration of Nine Parameters)**

So far, the examples have focused on parametric exploration. Figure 9.4 illustrates a hybrid exploration.[14] It shows the distribution of

[14]Paul K. Davis, David C. Gompert, Richard J. Hillestad, and Stuart Johnson, *Transforming the Force: Suggestions for DoD Strategy,* IP-179, RAND, 1998.

simulation outcomes resulting from having varied most parameter values "probabilistically" across an ensemble of possible wars, but with warning time and the delay in attacking armored columns left parametric. The probabilistic aspect of the calculation assumed, for example, that the enemy's movement rate had a triangular distribution across a particular range of values and that the suppression of air defenses would either be in the range of a few days or more like a week, depending on whether the enemy did or did not have air-defense systems and tactics that were not part of the best estimate. That is, if the enemy had some surprises up its sleeve, suppression of air defenses would be likely to take considerably longer. We represented this possibility with a discrete distribution for the likelihood of such surprises. The two curves in Figure 9.4 differ in that the one with crosses for markers assumes that interdiction of moving

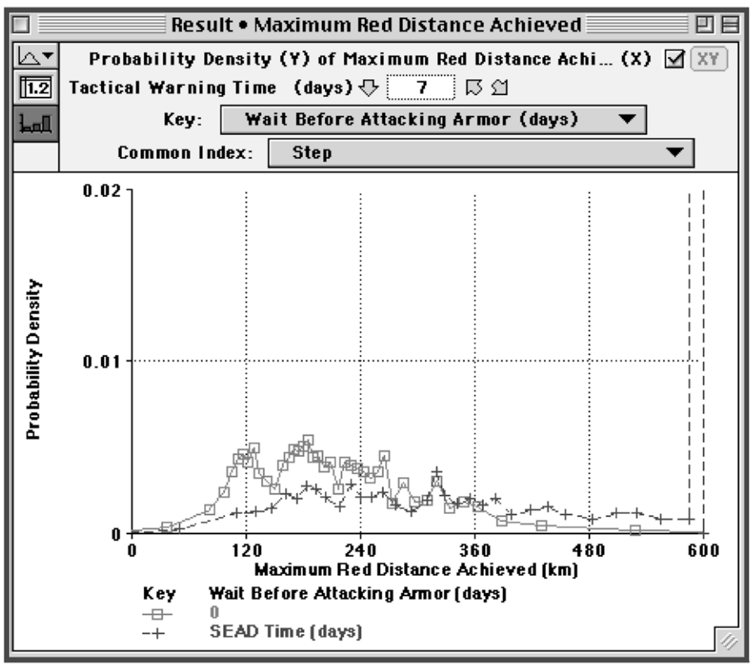



Figure 9.4—Display of "Probabilistic" Exploration

columns waits for suppression of enemy air defenses (SEAD). The other curve assumes that interdiction begins immediately because the aircraft are assumed stealthy or the defenses nonexistent.

This depiction of the problem shows in one display how widely outcomes can vary and how outcome distribution can be complex. The non-stealthy-aircraft case shows a considerable spike at the right end, where many cases pile up because, in the simulation, the attacker halts once he has reached an objective at about 600 km. Note that the mean is not a good metric: the variance is huge, and the outcome may be bimodal or even multimodal.

**Advanced Concepts.** The results just discussed are from analyses accomplished in recent years for DoD. Looking to the future, much more is possible with computational tools. Much better displays are possible, and, even more exciting, computational tools can be used to aid in the search process of exploration. For example, instead of "clicking" through the regions of the outcome space, tools could automatically locate portions of the space in which particular outcomes are found. Insights could be fine tuned by clicking around in that much more limited region of the outcome space. Or, if the model is itself driven by the exploration apparatus, the apparatus could search for outcomes of interest and then focus exploration on those regions of the input space. That is, the experimental design could be an output of the search rather than an input of the analysis process.

## ENABLING EXPLORATORY ANALYSIS

In principle, exploratory analysis can be accomplished with any model. In practice, it becomes difficult with large and complex models. If F represents the model, it can be considered to be simply a complicated function of many variables. How can we run a computerized version of F to understand its character? If F has M inputs with uncertain values, then we could consider N values for each input, construct a full-factorial design or some properly sampled subset, run the cases, and thereby have a characterization. However, the number of such cases would grow as NM for full-factorial analysis, which quickly gets out of hand even with big computers. Quite aside from issues of setup and run time, comprehending and communicating the consequences becomes very difficult if M is large. Suppose someone asked, "Under what conditions is F less than danger_point?" Given sufficiently powerful computers and enough time, one could create a database of all the cases, after which one could respond to the question by spewing out lists of the cases in which F fell below danger_point. The list, however, might go on for many pages, perhaps even thousands. What would be done with the list? This is one manifestation of what might be called the curse of dimensionality.

It follows that—even with a perfect high-resolution model and incredibly speedy computers—abstractions, such as aggregate representations, will still be necessary. In the most usual cases, in which the high-resolution model is by no means perfect, abstractions allow analysts to ponder the phenomena in meaningful ways, with relatively small numbers of cognitive chunks to deal with. People can reason with 3, 5, or, perhaps, even 10 such cognitive chunks at a time, but not with hundreds. If the problem is truly complex, ways must be found to organize the reasoning—i.e., the problem must be decomposed. One ends up using principles of modularity and hierarchy. To head off a rejoinder commonly made at this point by aficionados of networking technology, the need for an aspect of hierarchical organization is inescapable in most systems of interest, even if the system is highly distributed and relatively nonhierarchical. Everyone can observe this when interacting with the World Wide Web.

A corollary of the need for abstractions is the need for models that use the various abstractions as inputs. It is not sufficient to display the abstractions as intermediate outputs (displays) of the ultimate detailed model. The reasons include the fact that when a decisionmaker asks a what-if question using abstractions, there is a 1:n mapping problem in translating his question into the inputs of a more detailed model. That is, the decisonmaker asks, "What if we had 25 percent more capability?" but the detailed model represents many capabilities. What assumptions about these many capabilities would correspond best to the decisionmaker's question? In contrast, a more aggregated model may already have the concept of overall capability; it can then address the decisionmaker's question directly. That is, it accepts the decisionmaker's abstractions as inputs.

Given the need for abstractions, how do we find them and how do we exploit them? The approaches fall into two groups, one for new models and the other for existing, or legacy, models.

With new models, the issue is how to design, and the options of interest are

- Design the models and model families top down so that significant abstractions are built in from the start, but do so with enough understanding of the microscopics that the top-down design is valid.[15]
- Design the models and model families bottom up, but with enough top-down insight to build in good intermediate-level abstractions from the start.[16]
- Do either or both of the above, but with designs taken from different user or theoretical perspectives.

Note that this list does not include a pure top-down or bottom-up design approach. Only seldom will either generate a good design of a complex system. Note also the idea of alternative perspectives. This recognizes that many abstractions are not unique; to the contrary, there are different ways of viewing what the key factors of the problem really are (e.g., those in combat arms typically conceive military problems differently than logisticians do).

Only sometimes is there the opportunity to design from scratch. More typically, existing models must be used (or adapted and used). Moreover, the model "families" available are often families more on the basis of assertion or hope than lineage. What does one then do? Some possibilities are as follows:

- Given existing models developed at high levels of resolution, study the model and the questions that users ask of the model to discover useful abstractions. For example, one may discover that inputs X, Y, and Z only enter the computations as the product

[15]Paul K. Davis and James H. Bigelow, *Experiments in Multiresolution Modeling (MRM)*, MR-1004-DARPA, RAND, 1998.

[16]Paul K. Davis and James H. Bigelow, *Motivated Metamodels: Synthesis of Cause-Effect Reasoning and Statistical Metamodeling,* MR-1570-AF, RAND, 2003.

XYZ. If so, then XYZ may be a natural abstraction. Or, perhaps decisionmakers ask questions in terms of concepts such as force strength or force ratio, indicating that these are significant abstractions. For mature models, the obvious place to look is the list of displays that have been added over time to provide views into the internal workings of the model.

- Apply statistical machinery to search for useful abstractions. For example, if X, Y, and Z are inputs, such machinery might test to see whether the system's behavior correlates not just with X,Y, and Z, but with XY, XZ, YZ, and XYZ. If the computation does, in fact, depend only on XYZ, that fact will show up from the statistical analysis.
- Idealize the system and develop a "formal" mathematical representation (formal in the sense of being expressed symbolically without necessarily having the intention of computing the various terms and factors) that provides hints about the model's likely behavior. For example, such a representation might be much too complex to "solve," but, if coupled with some physical reasoning and a search for postulated simplifications, it might highlight the likelihood of an overall exponential decay, or an inverse dependence on one input, or various other nonlinearities that otherwise one might think to test for. It might suggest natural *aggregation fragments*, such as the product XYZ mentioned above. In practice, this approach is most powerful if one considers the problem from different perspectives that suggest different but plausible simplifications.[17,18]

This list is less straightforward than it first appears. The first approach is perhaps a natural activity for a smart modeler/programmer who begins to study an existing program—if, and only if, he is also a believer in higher-level depictions of the problem. The second approach seems to be favored by mathematically oriented individuals who lack enough class knowledge to take the first approach, or who

---

[17]One example of a simplification is the assumption that an integral is perhaps approximately equal to a representative value of the integrand times the effective width of the integration interval, and that this width is proportional to something physically straightforward.

[18]Davis and Bigelow, *Motivated Metamodels.*

believe—based sometimes on disciplinary faith—that such statistical procedures will prove successful and that those looking for more phenomenological abstractions are fooling themselves. The third approach is a hybrid. It argues that one should use one's understanding of phenomenology, and theories of system behavior, to gain insights about the likely or possible abstractions. Only then should one crank the statistical machinery. Where it is feasible, this is the stronger approach.

## Using Occam's Razor

An interesting tension arises in discussing how to form suitable abstractions. The principle of Occam's razor requires that one prefer the simplest explanation and, thus, the simplest model. Some, particularly enthusiasts of statistical approaches, tend to interpret this principle to mean that one should minimize the number of variables in a model. They tend to focus on data (natural or simulation generated) and to avoid adding variables for the purpose of "explanation" or "phenomenology" if the variables are not needed to predict the data. Instead, they prefer to "let the data speak." In contrast, subject-area phenomenologists may prefer to enrich the depiction by adding variables that provide a better picture of cause-effect chains. This, however, goes well beyond what can be supported with meager experimental data.

To not violate the Occam's Razor principle, one must remember the principle's longer form: Adopt the simplest explanation that truly explains all there is to explain—but nothing simpler! This should include phenomena that one "knows about," even if they are not clearly visible in the limited data (e.g., historical data on who won various battles with what overall attrition). I would add to this the old admonition (perhaps made first by Massachusetts Institute of Technology's Jay Forrester) to remember that to omit showing a variable one knows about may be equivalent to assuming its value to be 1 (as a multiplier) or 0.

It is sometimes useful to have a competition among approaches. For example, phenomenologists working a problem may be utterly convinced that it must be described with complex computer programs having hundreds or thousands of data elements. In such a case, it may be useful to study output behavior with a metamodel (also called a repro model and response surface).[19] In one instance with which I am familiar, such work by colleague James Bigelow showed that despite many man-years of effort building and feeding a complex ground-force model, results were strongly dominated by a single higher-level abstraction: theater-level force ratio. The implication was not that real combat is dominated only by theater-level force ratio, but, rather, that various assumptions and compromises made in developing the detailed model undercut any claims that its greater complexity and expense were adding predictive value relative to simpler models.

Although the discussion above distinguishes sharply between the case of new models and old ones, the two are connected. In essence, working with existing models should often involve sketching what the models *should* be like and how models with different resolution *should* connect substantively. That is, working with existing models may mean having to go back to design issues. If this seems suspicious, ask yourself how often you have found it easier to rederive a model and then decipher a program you have been given than to wade through the program on its own terms.

## Multiresolution, Multiperspective Modeling and Model Families

Abstractions, usually aggregations, are fundamental in exploratory analysis. Finding suitable abstractions, relating them, and conducting both high-level exploratory analysis and appropriate zooming into detail are greatly facilitated if models are designed in a special way. This is the subject of multiresolution, multiperspective modeling (MRMPM). Although the subject relates most directly to new models, it is also relevant to working with legacy models in preparing for exploratory analysis.

Multiresolution modeling (MRM) is building a single model, a family of models, or both to describe the same phenomena at different levels of resolution and to allow users to input parameters at those dif-

---

[19] A metamodel is a simple model that reproduces the aggregate behavior of a more complex model, as judged by statistical comparisons over many cases.

ferent levels depending on their needs.[20] Variables at level n are abstractions of variables at level n+1. MRM has also been called variable- or selectable-resolution modeling.[21] Figure 9.5 illustrates MRM schematically. It indicates that a higher-level model (model A) itself has more than one level of resolution. It can be used with either two or four inputs. However, in addition to its own MRM features, it has input variables that can be specified directly or determined from the outputs of separate higher-resolution models (models B and C,

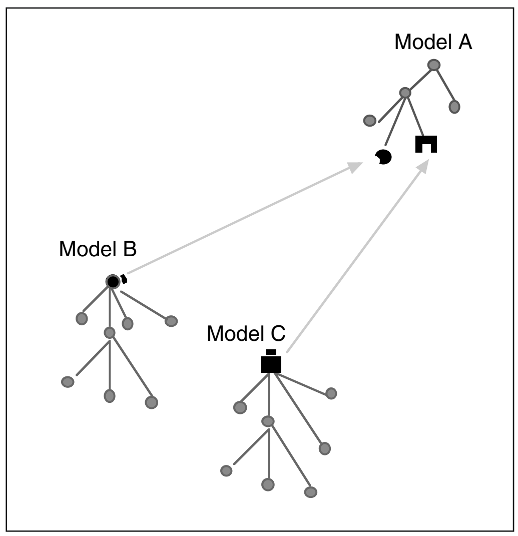



**Figure 9.5—Multiresolution Family of Models**

[20]Davis and Bigelow, *Experiments in Multiresolution Modeling*.

[21]Unfortunately, some authors use the term *multiresolution models* to mean only that there are *outputs* at different levels of detail or that a model happens to treat different phenomena asymmetrically (e.g., with detail for combat processes and a more aggregate depiction for logistics). Still other authors claim multiresolution features for models because their objects are hierarchical. The essence of multilevel resolution, as I use the term here (and as discussed in National Research Council, *Modeling and Simulation*), is having multiple levels of abstraction for key *processes*, such as attrition, movement, and communications. Achieving such MRM is quite challenging. RAND work on the subject dates back to the late 1980s.

shown as "on the side," for use when needed). In principle, one could attach models B and C in the software itself, creating a bigger model. However, in practice there are tradeoffs between doing this and keeping the more detailed models separate. For larger models and simulations, a combination single-model/family-of-models approach is desirable because it balances the needs for analytic agility and complexity management.

MRM is not enough, however, because, as noted earlier, different applications require different abstractions even if the resolution is the same—i.e., different "perspectives" are legitimate and important. Perspectives are distinguished by the conception of the system and the choice of variables; they are analogous to alternative "representations" in physics or engineering. Designing for both multiple resolution and multiple perspectives can be called MRMPM.[22]

With MRMPM models (single models or families), the concepts and predictions among levels and perspectives have to be connected (and reconciled). It is often assumed that the correct way to do this is to calibrate upward: treat the information of the most detailed model as correct and use it to calibrate the higher-level models. This, indeed, is often appropriate. The fact is, however, that the more detailed models almost always have omissions and shortcomings.[23] Models at higher levels, and from different perspectives, address some of them explicitly. Further, the different models of a family draw on different sources of information—ranging from doctrine or even "lore" on one extreme to physical measurements on a test range at the other. One class of information is not inherently better than another; it is simply different.

Figure 9.6 makes the point that members of a multiresolution model family should be *mutually* calibrated, with information flows in both directions. In the military domain, for example, low-resolution historical attrition or movement rates may be used to help calibrate more-detailed models predicting attrition and movement. This is not

---

[22]Paul K. Davis, "Exploratory Analysis Enabled by Multiresolution, Multiperspective Modeling," in Jeffrey A. Joines et al. (eds.), *Proceedings of the 2000 Winter Simulation Conference*, 2000.

[23]For example, detailed models often have a rich depiction of physical considerations but only a minimal representation of behaviors that adapt to circumstances.

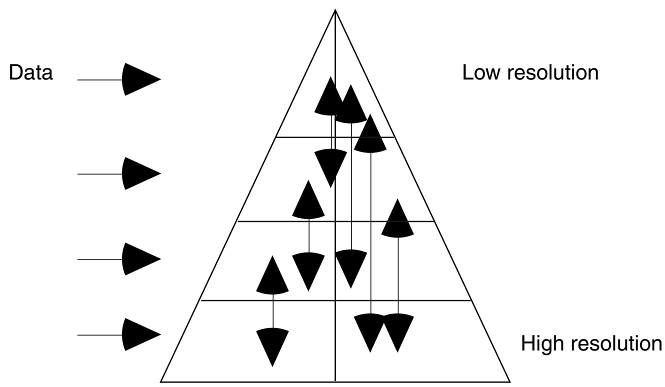



**Figure 9.6—Mutual Calibration of a Family of Models**

straightforward, because of the 1:n mappings. It is often done crudely, by applying an overall scaling factor (fudge factor) rather than correcting the more atomic features of the model, but it is likely to be something with which readers are familiar. However, much calibration is indeed upward. In a study using detailed order-of-battle information, for example, inputs on the number of "equivalent divisions" or "equivalent F-15 aircraft" used in abstract models can be computed from the data feeding high-resolution models. Furthermore, at least in principle, the attrition coefficients' dependence on situation (e.g., open versus wooded terrain for ground forces) should be informed by high-resolution work.

So, given their desirability, how should families of models be built? Or, given preexisting models, how "should" they relate before they are connected as software or used for mutual calibration? The first design principle may be to recognize that there are limits to what is feasible. In particular, there are limits to how well low-resolution models can be consistent with high-resolution models. *Approximation is a central concept from the outset.* Several points are especially important in thinking about this:[24]

---

[24]Davis and Bigelow, *Experiments in Multiresolution Modeling*.

- Consistency between two models of differing resolution should be assessed in the context of how the models are being used. What matters is not whether they generate the same final state of the overall system, but whether they generate approximately the same results in the application. That may be something as specific as summary graphs or rank ordering of alternatives. Another way to put this is to be practical, not theological, about how much detail is needed.
- The implications for consistency of aggregation and disaggregation processes must also be judged in context. Some disaggregation assumptions represent aggregate-level knowledge not necessarily reflected in the most detailed model.
- Comprehensive MRM is very difficult or impossible for complex modeling and simulation,[25] but having even some MRM can be far more useful than having none at all. MRM is not an all-or-nothing matter.
- The various members of an MRM family will typically be valid for only portions of the system's state space. As one moves from one region to another, valid description may require that parameter values or even the structure of the model itself be changed.
- Mechanisms are therefore needed to recognize different situations and shift models. In simulations, human intervention is one mechanism and agent-based modeling is another.[26]
- Valid MRM will often require stochastic variables represented by probability distributions, not merely mean values. Further, valid aggregate models must sometimes reflect correlations among variables that might naively be seen as probabilistically independent.

---

[25]For an excellent theoretical discussion of this, see Robert Axtell, *Theory of Model Aggregation for Dynamical Systems with Applications to Problems of Global Change,* Ph.D. dissertation, Carnegie-Mellon University (available from University Microfilms International, Ann Arbor, MI). Axtell's discussion, however, fails to emphasize the key significance of approximations. As a result, it is more pessimistic than my own work.

[26]Agent-based models include modules that represent (i.e., serve as agents for) adaptive entities, such as individual humans or groups. The basic ideas are discussed in most books dealing with "complex adaptive systems."

With these observations up front, the ideal for MRM is a hierarchical design for each MRM process, as indicated earlier, in Figure 9.5.

Models and analysis methods for exploratory analysis should have a number of characteristics. First, they should be able to reflect hierarchical decomposition through multiple levels of resolution and from alternative perspectives representing different “aspects” of a system. For example, one model might decompose a system into its organizational components, another might focus on different component processes, and yet another might follow component functions.

Less obviously, models and analysis methods should also include realistic mechanisms describing how the natural entities of the system act, react, adapt, mutate, and change. These mechanisms should reflect the relative “fitness” of the original and emerging entities for the environment in which they are operating. Many techniques are applicable here, including game theory methods and others that may be relatively familiar to readers. However, the most fruitful new approaches are those typically associated with the term *agent-based modeling*. These include submodels that act “as the agents for” specific entities—say, political leaders and military commanders or (at the other extreme) infantry privates on the battlefield. In practice, such models need not be exotic: they may correspond to some relatively simple heuristic, or intuitive, decision rules or to some well-known (though perhaps complex) operations-research algorithm. But to have such decision models is quite different from depending on scripts.

Because it is implausible that closed computer models will be able to meet the above challenge in the foreseeable future, the family of “models” should allow for human interaction—whether in human-only seminar games, small-scale model-supported human gaming, or distributed interactive simulation. This runs against the grain of much common practice, which imputes too much virtue to “closed models” that generate readily reproducible results.

The last item in the bulleted list above is often ignored in today’s day-to-day work, even by good analysts who have a family of models. Often, when they seek to use models at different levels of resolution analytically, they decide on a highest-level model to be used for excursions—i.e., for examining sensitivities. They then “calibrate” this highest-level model by using one or more detailed models. For example, they might use the Brawler model of air-to-air combat between small groups of aircraft in different groupings; they would then use results of that work to calibrate the air-to-air model of a theater-level depiction, such as in the TACWAR, JICM, Thunder, or START models. This is not easy. However, the analysts would sit down, talk, draw sketches, and so on, until they gained a sense of how to go about the calibration. Ultimately, for a particular study done on a limited budget and time scale (as most are) they might use expected-value outcomes of "representative" air-to-air engagements in Brawler to set attrition coefficients in the theater-level model. This might or might not be "correct," because the relationship between the engagement level and theater level is very complex: in a real air war, there may be thousands of engagements with a wide variety of characteristics, and how to aggregate is not so clear. For example, one might imagine that 80 percent of engagements are "normal" but have little effect on relative force levels, while 20 percent of engagements are of a different character and lead to one side annihilating the other's aircraft with no losses of its own. The overall time dependence of relative force levels, then, might be dictated by the unusual, nonrepresentative engagements. However, focusing on these unusual cases in doing the calibration might outrageously exaggerate one or both of the attrition rates. The "correct" way to go about the calibration would necessarily involve explicit, study-dependent integration over classes of engagement.

Sometimes, the higher-level model inputs need to be stochastic. Figure 9.7 illustrates the concept schematically for a simple problem. Suppose that a process—for instance, one computing the losses to aircraft in air-to-air encounters—depends on five inputs: Q, X, Y, a, and b. But suppose that the outcome of ultimate interest involves many instances of that process with different values of X and Y (e.g., different per-engagement numbers of Red and Blue aircraft). An abstraction of the model might depend only on Q, a, and b (e.g., overall attrition might depend on only numbers of Red and Blue aircraft, their relative quality, and some C2 factor). If the abstraction shown is to be valid, the variable Z should be consistent with the higher-resolution results. However, if it does not depend explicitly on X and Y, then there are "hidden variables" in the problem, and Z may

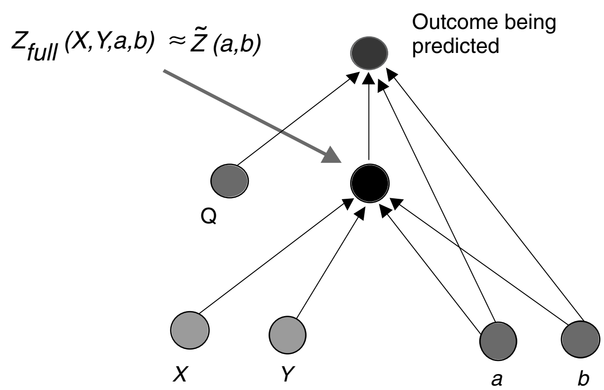



**Figure 9.7—Input to Higher-Level Model May Be Stochastic**

appear to be a random variable, in which case the predicted outcome would be a random variable. This randomness might be ignored if the distribution were narrow enough, but that might not be the case.

To compute what "should be," one would relate the probability density to a constrained integral over X and Y, appropriately weighting on the basis of likelihood and restricting the integration to regions where Z has the value of interest.

In the past, such calibrations have been rare, in significant part because analysts have lacked both theory and tools for doing things better. The "theory" part includes not having good descriptions of how the detailed model should relate to the simplified one. The tool part includes not being able to define the set of runs that should be done (representing the integral of Figure 9.7) and then to actually make those runs.

Ideally, such a calibration would be dynamic within a simulation. Moreover, it would be easy to adjust the calibration to represent different assumptions about command, control, communications, computers, intelligence, surveillance, and reconnaissance (C4ISR), as well as about tactics. We cannot do such things today, because modeling technology and practice are not up to it yet.

## LESSONS FROM RECENT EXPERIENCE

Both exploratory analysis and MRM/MRMPM are relatively new concepts, but there is a growing body of examples to illustrate their practicality for addressing problems—for instance, the problem of halting an invading army using precision fires from aircraft and missiles.[27] The most recent aspects of that work included understanding in some detail how the effectiveness of such fires are affected by details of terrain, enemy maneuver tactics, certain aspects of command and control, and so on. This provided a good test bed for exploring numerous aspects of MRMPM theory.

We developed a multiresolution personal-computer model (PEM),[28] written in Analytica, to understand and extend to other circumstances the findings from entity-level simulation of ground maneuver and long-range precision fires. A major part of this work was learning how to inform and calibrate PEM to the entity-level work. There was no possibility, in this instance, of revising the entity-level model. Nor, in practice, did we have a good enough understanding of the model to construct a comprehensive calibration theory. Instead, we had to construct a new, more abstract model and attempt to impose some of its abstractions on the data from runs of the entity-level simulation in prior work, plus some special runs made for our purposes. Had we had the intermediate-level PEM several years earlier, we could have used it both to define adaptations of the entity-level model that would have generated some of the abstractions we needed and to better define the experiments conducted with the high-resolution model. Instead, we had to make do with the situation

---

[27]Our work on precision fires is discussed in Davis, Bigelow, and McEver, *Analytical Methods for Studies and Experiments*; Jimmie McEver, Paul K. Davis, and James H. Bigelow, *EXHALT: An Interdiction Model for Exploring Halt Capabilities in a Large Scenario Space,* MR-1137-OSD, RAND, 2000; Paul K. Davis, James H. Bigelow, and Jimmie McEver, *Effects of Terrain, Maneuver Tactics, and C4ISR on the Effectiveness of Long-Range Precision Fires: A Stochastic Multiresolution Model (PEM) Calibrated to High-Resolution Simulation*, MR-1138-OSD, RAND, 2000; and Davis, Bigelow, and McEver, *Effects of Terrain*. Some of this work was also used in the summer study of the Defense Science Board and is reflected in Eugene C. Gritton, Paul K. Davis, Randall Steeb, and John Matsumura, *Ground Forces for a Rapidly Employable Joint Task Force: First-Week Capabilities for Short-Warning Conflicts,* MR-1152-OSD, RAND, 2000.

[28]Davis, Bigelow, and McEver, *Effects of Terrain*.

we found ourselves in. The result is a case history with what are probably some generic lessons learned.

Figure 9.8 illustrates one aspect of our multiresolution PEM approach. The figure shows the data flow within a PEM module that generates the impact time (relative to the ideal impact time) for a salvo of precision weapons aimed at a packet of armored fighting vehicles observed by C4ISR assets at an earlier time. Other parts of the PEM combine information about packet location versus time and salvo effectiveness for targets that happen to be within the salvo's "footprint" at the time of impact in order to estimate the effectiveness of the precision weapons. For the salvo-impact-time module, Figure 9.8 shows how the PEM is designed to accept inputs as detailed as whether there is enroute retargeting of weapons, the C2 latency time, and weapon flight time. However, it can also accept more aggregate inputs, such as time from last update. If the input variable "resolution of last update calculation" is set "low," then time from last update is specified directly as input; if not, it is calculated from the lower-level inputs.

Being able to depict the problem as in Figure 9.8, and to provide users the option of what inputs to use, has proven very useful—both for analysis itself and for communicating insights to decisionmakers in communities ranging from the C4ISR community to the programming and analysis community. In particular, the work clarified how the technology-intensive work of the C4ISR acquisition community relates to higher-level strategy problems and analysis of such problems at the theater level.

Another companion piece describes how, in developing the PEM and a yet more abstract model (EXHALT) used for theater-level halt-problem analysis, we experimented with methods for dealing with the multiperspective problem.[29] Perhaps the key conclusion of this particular work is that MRMPM rather demands a building-block approach that emphasizes study-specific assembly of the precise model needed. Although we had some success in developing a closed

---

[29]Jimmie McEver, Paul K. Davis, and James H. Bigelow, "Implementing Multiresolution Models and Families of Models: From Entity Level Simulation to Personal-Computer Stochastic Models and Simple 'Repro Models,'" *SPIE 2000,* Orlando, FL, April 2000.

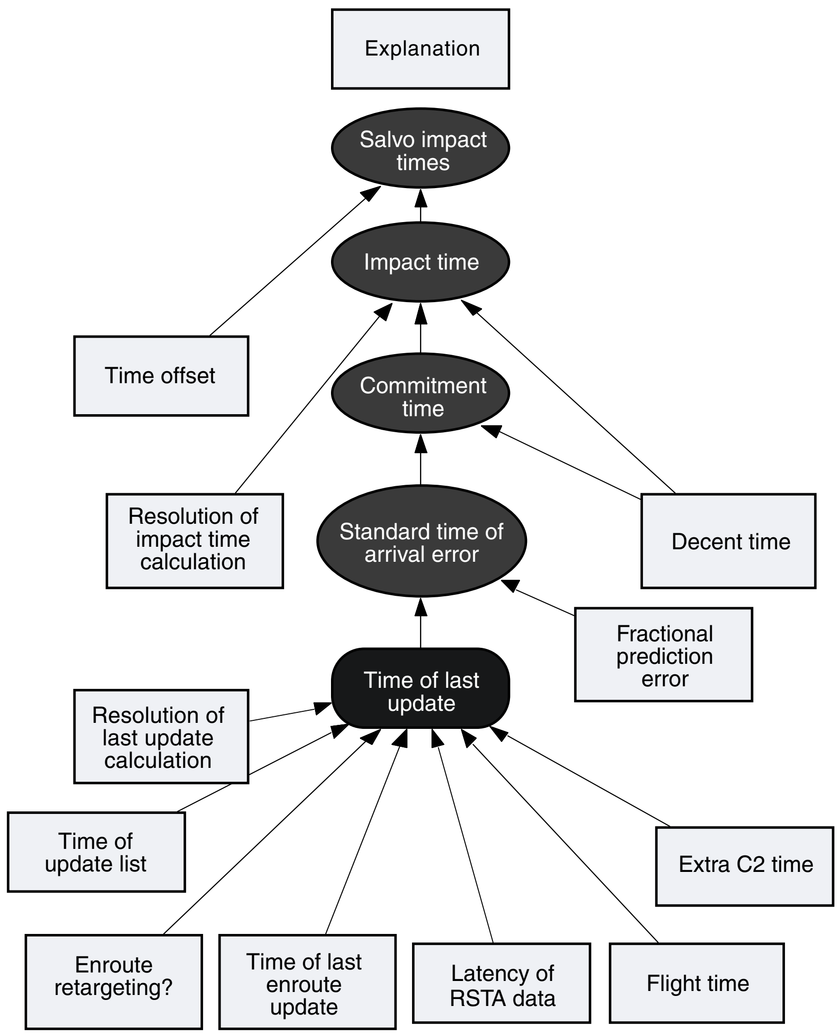



**Figure 9.8—A Multiresolution, Multiperspective Design for Salvo-Impact-Time Module**

MRMPM model with alternative user modes representing different demands for resolution and perspective (e.g., the switches in Figure 9.8), it proved impossible to do very much in that regard: the number of interesting user modes and resolution combinations simply precludes being able to wire in all the relevant user modes. Moreover, the explosion of complexity occurs very quickly. Thus, despite the desire of many users to have a black-box machine that can handle all the cases and perspectives of interest, it seems a fundamental reality that at-the-time assembly from building blocks, not prior definition, is the stronger approach. This was as we expected, but even more so.

The ultimate reason for the building-block conclusion is that even in the relatively simple problem examined, the real variable trees (akin to data-flow diagrams) are bushy rather than rigorously hierarchical. Furthermore, the different legitimate perspectives can simply not all be accommodated simultaneously without making the code itself very complicated to follow. In contrast, we found it easy to construct the model needed quickly—in hours rather than days or weeks—as the result of our building-block approach, visual modeling, use of array mathematics, and strong, modular design.

As powerful as current personal computer tools are in comparison with those of past years, they are still not up to the challenge of making the building-block/assembly approach rigorous, understandable, controllable, and reproducible without unrealistically high levels of modeler/analyst discipline. Also, the search models for advanced exploratory analysis are not yet well developed. Thus, there are good challenges ahead, not just for the model builders and users, but also for the community that builds the enabling technology.

## Appendix

### REFLECTING UNCERTAINTY WITH PARAMETERS, AN EXAMPLE

As an example, consider a model that describes the rate at which Red and Blue suffer attrition in combat according to a Lanchester square law:

$$\frac{d\tilde{R}}{dt} = -\tilde{K}_b\tilde{B}(t) \quad \frac{d\tilde{B}}{dt} = -\tilde{K}_r\tilde{R}(t) \ ,$$

where the attrition coefficients for Red and Blue have both deterministic and stochastic parts, each of which is subject to uncertainty. The equation for Red says that the quantity of Red capability decreases in proportion to the quantity of Blue (because Blue is shooting at Red):

$$\tilde{K}_b(t) = K_{b0}\left[1 + c_b\tilde{N}_b(t;\mu,\sigma_b)\right] \qquad K_{r0}\left[1 + c_r\tilde{N}_r(t;\mu,\sigma_r)\right] .$$

Here, $K_{b0}$ and $K_{r0}$ are average attrition rates for a given war. They may be highly uncertain (factors of 2, 3, or more), but they are constant within a war. That is, before the war, we may not know the sides' average effectivenesses, but they exist. This said, attrition will vary from battle to battle and from time period to time period within a given war. Such variation can be regarded as a stochastic process. These effects are reflected by the bracketed factors, above, where $N_r$ and $N_b$ are assumed to be normal random variables with means of m and standard deviations. Their parameters are also uncertain, perhaps strongly so, but it is a different kind of uncertainty than that about the average attrition for a given war.[30]

---

[30] I distinguish between deterministic uncertainty and stochastic processes, but both may be treated by the same mathematical tools, such as probability distributions. The distinction is important, however. For example, a commander discovering that his losses to attrition were three times what he expected on the first day of war—and ascribing that attrition to the unanticipated effectiveness of certain weapons—should not imagine that tomorrow is another day, that stochastic variation may result in very low attrition tomorrow, and that he therefore should continue as on the first day. Unfortunately for that commander, things won't "average out." He needs to change tactics.

So far the equations have represented input uncertainty. However, suppose that we do not know the correct equations of combat—except that, for some reason, we are convinced that the correct equations are Lanchesterian: either what aficionados call “Lanchester linear,” “Lanchester square,” or something in between. We could then reflect this uncertainty by rewriting the equation as

$$\frac{d\tilde{R}}{dt} = -\tilde{K}_b \tilde{B}^e(t)\tilde{R}^f(t) \quad \frac{d\tilde{B}}{dt} = -\tilde{K}_r \tilde{B}^g(t)\tilde{R}^h(t) \ .$$

Now, by treating the exponents e, f, g, and h as uncertain parameters, we can change the very structure of the model. Thus, by varying parameter values, we can explore both input and structural uncertainties in the model.

There are limits to what can be accomplished. Suppose that the correct equations of combat are indeed Lanchesterian but that the K-factors decay exponentially with time as combatants tire, lose efficiency, or husband ammunition. The consequences of different exponential decay times cannot even be explored if the phenomenon goes unrecognized. This is not an idle comment; we often do not know the underlying form of the system model: many aspects of phenomena are recognized, but not others. And they may not be observed except in unique circumstances. Despite this caveat, we can do a great deal with exploratory analysis to understand the consequences of uncertainties that can be parameterized.