Experiments in Multiresolution Modeling (MRM)

Paul K. Davis
James H. Bigelow

National Defense Research Institute

50th Year
RAND
A National Resource
1948 - 1998

The research described in this report was sponsored by the Defense Advanced Research Projects Agency. The research was conducted in RAND's National Defense Research Institute, a federally funded research and development center supported by the Office of the Secretary of Defense, the Joint Staff, the unified commands, and the defense agencies, Contract DASW01-95-C-0059.

Library of Congress Cataloging-in-Publication Data

Davis, Paul K., 1943–
Experiments in multiresolution modeling (MRM) / Paul K. Davis and James Bigelow
p. cm
"Prepared for the Defense Advanced Research Projects Agency by RAND's National Defense Research Institute."
"MR-1004-DARPA."
Includes bibliographical references.
ISBN 0-8330-2653-4
1. Combat—Mathematical models. I. Bigelow, J. H. II. United States. Dept. of Defense. Defense Advanced Research Projects Agency. III. National Defense Research Institute. RAND. (U.S.) IV. Title
U21.2.D269 1998
355.4 ' 8 ' 015118—dc21 98-33725
CIP

Published 1998 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1333 H St., N.W., Washington, D.C. 20005-4707
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information, contact Distribution Services: Telephone: (310) 451-7002; Fax: (310) 451-6915; Internet: order@rand.org

Experiments in Multiresolution Modeling (MRM)

Paul K. Davis
James H. Bigelow

MR-1004-DARPA

Prepared for the
Defense Advanced Research Projects Agency

National Defense Research Institute

RAND

Preface

This report describes final results of a small project sponsored by the Defense Advanced Research Projects Agency (DARPA), the intention of which was to identify useful directions for further work on multiresolution modeling (MRM). The work was largely accomplished in the Applied Sciences and Technology Center of RAND's National Defense Research Institute (NDRI), a federally funded research and development center (FFRDC) sponsored by the Office of the Secretary of Defense, the Joint Staff, the unified commands, and the defense agencies. Some of the modeling reflected here in simplified form was accomplished in the "Planning Future Forces" and "Transforming the Force" projects sponsored by NDRI's advisory board. NDRI also provided additional research-support funding to complete the work and present it at conferences. Comments are welcome and should be addressed to Dr. Paul Davis at RAND in Santa Monica, California (e-mail: pdavis@rand.org).

Contents

Figures

Tables

Summary

Definitions

Multiresolution modeling (MRM) is building a single model, a family of models, or both to describe the same phenomena at different levels of resolution.[1] For larger models and simulations such as the envisioned JWARS and JSIMS, a combination single-model/family-of-models approach is desirable. This would balance needs for analytical agility and complexity management.

Motivations for MRM

The motivations for MRM are many, and most are largely independent of computational power. Taking for granted the desirability of high-resolution models that describe underlying phenomena, the principal issue is why one needs lower-resolution, more aggregate models. The usual reason given involves economy: it is expensive and time-consuming to work always with high-resolution models with their voracious appetite for data, which is often unavailable or unreliable. However, other reasons may be even more fundamental. First, humans reason at different levels of detail and therefore *require* corresponding models cognitively. Second, we need models to exploit knowledge that comes at many different levels of detail. Indeed, separate laws operate at different levels of detail, as celebrated under the rubric of "emergent behaviors" in the study of complex adaptive systems, and as recognized militarily in the hierarchical nature of doctrinal rules. Third, real-world high-resolution models typically have limitations of scope that leave implicit some important determinants of higher-level behaviors. For example, most detailed models are "scripted" and omit crucial information on strategy and adaptive behavior. Last, but not least, low-resolution models are needed for *exploratory analysis* providing a broad high-level understanding of a problem and the potential value of alternative decisions. For related reasons, low-resolution models are needed for analytical agility.

[1]MRM is sometimes called variable- or selectable-resolution modeling. The word *fidelity* is sometimes used instead of resolution. MRM is closely related to model *abstraction*, that is, simplifying models in ways that still capture the essence of a phenomenon with respect to the application at hand.

A consequence of these considerations is that the models in an ideal hierarchical MRM family would be *mutually* calibrated using all available knowledge—working top-down, bottom-up, and sideways. The common impression that truth resides in detailed models is simply wrong.

Feasibility of MRM

However desirable MRM may be, there are limits to what is feasible. In particular, there are limits to how consistent lower-resolution models can be with high-resolution models. Several principles are useful here:

- 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 (e.g., something as specific as summary graphs or the rank ordering of alternatives). What matters, then, is the relationship between processed information derived from a mere portion of the two final states.
- The implications for consistency of aggregation and disaggregation processes cannot be judged in the abstract. Some aggregations appear to throw away information, but in fact discard only information of little significance to application results. Similarly, disaggregations that appear arbitrary may in fact represent aggregate-level knowledge rooted in doctrine, experience, or deeper theory.
- Comprehensive MRM is very difficult for complex modeling and simulation (M&S), but having even some MRM can be far more useful than having none at all.
- The various models and submodels in an MRM family will typically be valid for only portions of the system's state space. As one moves from one region in state space to another, valid description may depend not just on changing parameter values in models, but on changing the very structure of the models. For example, a force-on-force model driven by an adjusted force ratio might be respectable in a frontal assault, but a very different model would be needed to describe post-breakthrough exploitation or battles characterized by "collapsing" the enemy through attacks—by fires, forces, and information warfare—on his "center of gravity."
- Mechanisms are therefore needed to recognize different situations and to shift models or submodels. Human intervention is one mechanism; agent-based modeling is another. "Scripted models" have no such adaptiveness.
- Valid MRM will often require stochastic variables represented by probability distributions, not merely gross measures such as mean values. Further, valid aggregate models must sometimes reflect correlations among variables that might naively be seen as probabilistically independent.

Insights and Methods Related to MRM Design

Despite the motivations for and the feasibility of MRM, there are relatively few working examples. One purpose of this study was to better understand the obstacles to implementing MRM and the differences between designing for MRM and normal (non-MRM) practice.

Given these objectives and a number of hypotheses about how to proceed, we worked through a military problem in some detail, developing a series of abstractions (more-aggregate models) in sequence and noting with some care how various methods applied and what kinds of issues arose. The problem involved halting an invading army with precision fires from aircraft and missiles. Elsewhere we and our colleagues have used much richer and more-detailed models of this problem, but for this work we started at a level of detail comparable to that reflected in many current theater-level models. Although simpler than reality, this level was complex enough to involve dozens of variables and hundreds of input data elements. Thus, it was a good candidate from which to develop an MRM family.

In working through the problem in detail, we arrived at the following insights and methods, which we believe are more generally valid:

- *Using array formalism* (e.g., vectors) can greatly simplify what appears cognitively to be a very complex problem. It can be crucial to MRM.
- Exploiting arrays requires being willing to identify object classes that mix "apples and oranges" (e.g., aircraft and missiles as different types of "shooters"). This is at odds with the common object-oriented practice of limiting objects to physically and doctrinally distinct entities, rather than to an abstraction such as a "shooter." Further, exploiting the more abstract object classes *may* create problems for model users to whom the abstractions are inappropriate. Thus, there may be tradeoffs between MRM objectives and other objectives affecting model representation. This bears further study.
- *Using formal mathematics* to characterize the model can be quite powerful even if the related equations cannot be solved analytically. One can identify natural "aggregation fragments," i.e., natural aggregate variables, to use in MRM design. In some cases, the resulting MRM features involve no further approximations. In other cases, insight from formal theory can identify good approximate aggregate models that are quite different from what might naively be expected. In still other cases, insights from theory can greatly help in establishing the structure of a suitable "repro model" that can be tuned to represent the behavior of higher-resolution models. The resulting repro model can be more robust than models inferred naively from normal regression methods.
- A key element of MRM design, which virtually requires formal mathematics, is developing approximately *hierarchical trees of variables* as illustrated in Figure S.1 (terms are defined in the text). Implementation of such designs should allow users to start at any of the tree's levels. For example, in exploratory analysis, one might start high on the tree where there are fewer than ten independent variables. For more in-depth work, one would start lower on the tree, where there are dozens of independent variables (underlined variables are vectors). A properly developed tree also flags correlations that must be accounted for in calibrating upward (where branches connect). If the model is stochastic, the connections may imply the need for joint probability distributions.
- When considering approximate aggregations, much can be achieved by bringing to bear the methods of *estimation theory* (e.g., identifying good estimators rather than relying on simple arithmetic averages).
- In combat modeling, a key element of finding good estimators, and of developing sound aggregate models more generally, is sharply *distinguishing among phases* of operations. This practice is often not very difficult—if one has the appropriate formalism and visualization tools.
- A fundamental difficulty in MRM is the existence of many possible choices of aggregate variables. Any one set corresponds to a particular representation of the problem. Attempting to allow users to decide ad hoc which variables to specify as inputs would lead to extraordinarily complex and opaque computer code. Thus, we recommend that designers *identify user modes*—i.e., identify an appropriately broad range of representations (sets of aggregate variables), provide an interface in which users choose among them, and

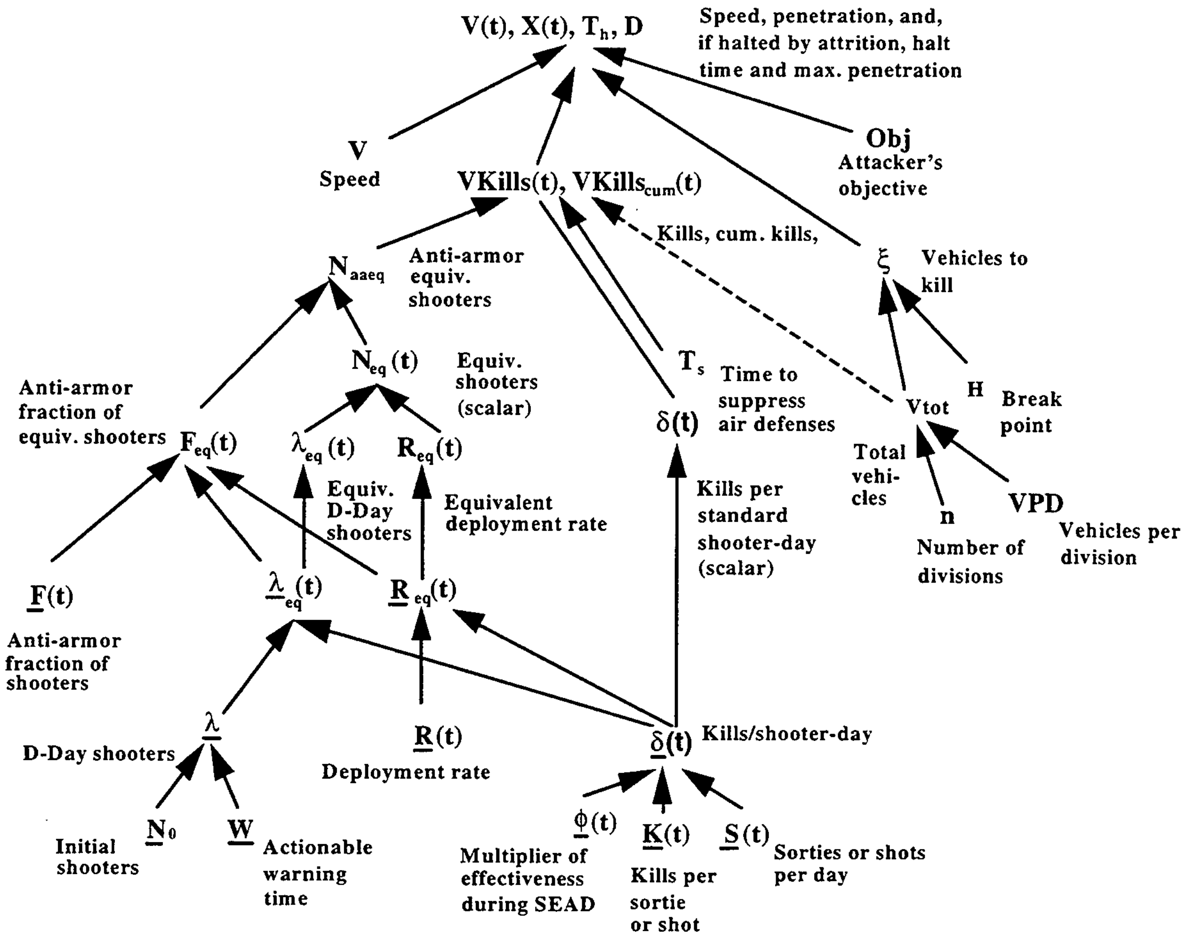

Figure S.1—Illustrative MRM Tree Design

modularize the model and program so that—in any given user mode—the associated structure and algorithms are as simple as possible. This affects clarity, reviewability, and model-explanation features.

- Introducing "abstractions of convenience," such as a constant multiplier of arrival rates for shooter A on days 1, 2, . . . n, can increase analytical agility by allowing quick responses to questions such as "What if deployment rates doubled?" Similarly, representing input-data vectors by time-dependent functions can reduce dimensionality a great deal with little loss of real-world accuracy.
- Even relatively standard multiple-regression packages can be used effectively in developing good repro models if sufficient theory is brought to bear to define the structure of the repro model one wants to tune. Other advanced statistical tools, when applied to output data generated by detailed models, can be helpful in recognizing natural "phase transitions" that should be exploited in MRM design.

Recommended Next Steps

MRM is still a frontier subject in modeling and simulation. Our current study has left us more bullish about its prospects, but much research is needed on issues as diverse as the basis of MRM in fundamental theory (e.g., general systems theory) and the development of computational and visualization tools that would simplify practical work in MRM. We believe that it would prove quite fruitful for the Department of Defense (DoD) to sponsor a number of studies analogous to the current one in which workers go through real-world problems in considerable detail, attempting to design MRM models and recording their insights and methods. Modeling is a mixture of art and science and benefits greatly from the shared experiences of diverse individuals. Such sharing, of course, depends on having model descriptions separate from the programs that implement the programs.

Despite the early state of MRM theory and practice, much can be accomplished now. We recommend that the DoD conceive and develop the JWARS and JSIMS programs in a way that includes research leading to MRM families, albeit families that would be imperfect and noncomprehensive. The payoffs for eventual users of these and other DoD models would be great. Such work, however, is not a mere matter of "programming," nor a matter of routine operations research. The need exists for in-depth research in military phenomenology that leads to sound theories and designs well before coding even begins. Building high-resolution simulations with broader scope and with agent-based adaptive decision processes can be very useful in this context, and the results could be used to better understand when various deterministic and stochastic aggregations are valid. In other cases research will need to exploit live or virtual simulations.

Acknowledgments

We greatly appreciate discussions in the course of the project with RAND colleagues Steve Bankes, John Adams, Richard Hillestad, and Jimmie McEver, and both discussion with and a review by Professor Bernard Zeigler of the University of Arizona.

Acronyms and Abbreviations

ATACMS	Advanced tactical missile system
BAT	Brilliant anti-tank munition
CAA	U.S. Army's Concepts Analysis Agency
CAS	Complex Adaptive Systems
CEM	Combat Effectiveness Model
C^4ISR	Command control, communications, computers, intelligence, surveillance, and reconnaissance
CVBG	Carrier battle group
DARPA	Defense Advanced Research Projects Agency
DIS	Distributed interactive simulation
DMSO	Defense Modeling and Simulation Office
DoD	Department of Defense
DSB	Defense Science Board
FLOT	Forward line of troops
IHVR	Integrated hierarchical variable-resolution modeling
JICM	Joint Integrated Contingency Model, a RAND model for work at the theater and operational levels, integrated to include strategic mobility and joint operations
JSIMS	Joint Simulation System
JTF	Joint Task Force
JWARS	Joint Warfare Simulation
M&S	Modeling and simulation
MLRS	Multiple-launcher rocket system
MRM	Multiresolution modeling
MTS	Multiple-time-scale methodologies
NRC	National Research Council
OOM	Object-oriented modeling
OOP	Object-oriented programming
OSD	Office of the Secretary of Defense
QDR	Quadrennial Defense Review
QJM	Quantified Judgment Model
RSAS	RAND Strategy Assessment System (a large-scale analytic wargaming system, a part of which has evolved into the JICM)
SAFOR	Semi-automated forces
SAIC	Science Applications International Corporation
SEAD	Suppression of enemy air defenses
VRM	Variable-resolution modeling

1. Introduction

Objectives

This report describes results of a project on multiresolution modeling (MRM) theory, which involves the art of constructing a model or a family of models that describe consistently the same system or process at different levels of resolution. The purpose of the project was to review key concepts and then to work through some concrete examples to gain tentative insights on (1) how going about MRM differs from normal modeling; (2) why MRM has been difficult in the past; (3) what is feasible in the future; and (4) what directions follow-on research by the research community should take.[2]

Although our work was quite limited in scope, it generated significant insights and practical results that are being used in ongoing projects. We believe that follow-on work by the community can be increasingly focused and productive as more is understood about the issues. One goal for future work should be a primer for MRM design and development. Other goals would be improved theories, methods, and computer environments to assist in MRM. We emphasize, however, that what has been accomplished to date is a set of small steps on a long and difficult journey. MRM theory is in its very early stages. It is a frontier problem in modeling and simulation (National Research Council, 1997, Chapter 6 and Appendix E).

Background

Origins of the Study and Earlier RAND Research

Much of the Department of Defense's (DoD's) interest in MRM (or what we have previously called variable-resolution modeling)[3] was stimulated by RAND work in 1990–1992, which drew

---

[2]A preliminary and much condensed version of this report has been published as Paul K. Davis and James Bigelow, "Introduction to Multi-Resolution Modeling (MRM) with an Example Involving Precision Fires," *Proc. SPIE*, Vol. 368, 1998.

[3]Our earlier work referred to *variable-resolution modeling*, which had the disadvantage of conveying to some readers the misimpression of continuously variable resolution, which is seldom feasible. We also referred to *cross-resolution modeling*, by which we meant connecting already existent models of different resolution—something that typically proves quite difficult to accomplish well, except for the relatively easy part of making the software connections "work." See Davis and Hillestad (1992a, b).

heavily on lessons learned from development of the RAND Strategy Assessment System (RSAS) in the mid-to-late 1980s (Davis and Huber, 1992, Davis, 1993). Issues raised by this work led RAND, under the sponsorship of the Defense Advanced Research Projects Agency (DARPA) and the Defense Modeling and Simulation Office (DMSO), to convene an international conference on variable-resolution and cross-resolution modeling in 1992. The proceedings were published (Davis and Hillestad, 1992b), as was a technical paper summarizing results (Davis and Hillestad, 1992a). Three of the conference's principal research papers were published separately in a special volume of *Naval Research Logistics*, later issued as a book (Bracken, Kress, and Rosenthal, 1995).

By 1996 there was broader-based interest in the subject, primarily because aggregation-disaggregation problems were arising in distributed interactive simulation (DIS),[4] but also because some public officials were coming to recognize the broad significance of the MRM problem.

In 1996 Dr. Judith Dahmann of DMSO and Mr. Dell Lunceford of DARPA convened a small group (Paul Davis and Richard Hillestad of RAND, Paul Reynolds of the University of Virginia, and Ben Wise of Science Applications International Corporation [SAIC]) to work on the MRM problem collaboratively. Work stopped because of an unanticipated funding cut. In the spring of 1997, DARPA began the current small project to rejuvenate the earlier investigation and to suggest directions for more-extensive follow-on work if that seemed appropriate.

Parallel Research Efforts

A few other developments are also significant as background information. During the late 1980s and 1990s several university-based authors began to publish on issues related to MRM, notably model abstraction.[5] Also, the Air Force's Rome Laboratory has recently supported related work on model abstraction as an enabler of advanced modeling and simulation.[6]

---

[4]In DIS experiments it frequently happens that two types of objects described at different levels of detail must interact (e.g., a group of individual tanks may encounter a platoon of enemy mechanized infantry). To describe their interaction, the high-resolution objects must be aggregated or the low-resolution object must be disaggregated. Then, after the interaction, it may be necessary to restore the level of resolution of both object classes as the simulation continues.

[5]See, e.g., Zeigler (1984, 1990), Fishwick (1988, 1989), Axtell (1992), and Fishwick and Lee (1996).

[6]See "Enabling Technologies for Modeling and Simulation," an account of a session chaired by Alex Sisti of the Rome Laboratory, in the *Proceedings of the International Society for Optical Engineering (SPIE)*, April 22-24, 1997, Orlando, Florida. For an overview see Caughlin and Sisti (1997). Abstracts of the conference papers are available on-line for *Proc. SPIE*, Vol. 3083, 1997, at http://www.spie.org/web/pubs_home.html. See also Vol. 3369, which has papers from a related meeting held in April 1998.

Two important events occurred in 1996–1997. Panels of a Defense Science Board (DSB) summer study and a National Research Council (NRC) study strongly endorsed the conclusion that MRM is one of the most fundamental challenges in modern modeling and simulation. Although this was not an independent event (Davis participated in both studies), the consensus achieved was significant given the broad expertise represented on the panels.[7] The consensus reflected the fact that one or another version of the MRM problem has arisen in many different disciplines over several decades. In each case it has been recognized as profoundly difficult—and important.[8]

Structure of the Report

The remainder of this report begins in Chapter 2 with definitions and some technical background. Chapter 3 describes reasons for MRM; Chapter 4 reviews some of the obstacles. In Chapter 5 we describe ingoing hypotheses and the approach used in this report to uncover principles and insights. Chapter 6 describes a worked-out example in some detail. Chapter 7 draws conclusions and recommends next steps for the research community and for the Department of Defense as it pursues models like JWARS and JSIMS.

---

[7]See NRC (1997) and DSB (1996). The NRC study has been formally published by the National Academy Press. Participants were drawn from many fields—e.g., engineering of complex systems (John Doyle, Caltech, and David McDowell, Georgia Tech); the DIS community (Duncan Miller, Lincoln Lab), high-resolution simulation (Don Blumenthal; Ben Wise, SAIC); academic modeling and simulation theory (Bernard Zeigler, University of Arizona, and Don Gaver, Naval Postgraduate School); statistics (John Lehoczky, Carnegie Mellon University); and DoD computer science and modeling (Richard Ivanetich, Institute for Defense Analyses and Richard Bronowitz, Center for Naval Analyses). Davis led report preparation. The DSB panel was chaired by Lt Gen. Jasper Welch (U.S. Air Force, retired). This report benefited significantly from the work of both panels.

[8]For a partial review of the academic literature through the 1980s, see Axtell (1992). See also Appendix E of NRC (1997) and Chapter 3 of Cloud and Rainey (1998). Some of the many terms under which MRM issues arise are model abstraction theory, problem decomposition, variable-resolution modeling, variable-fidelity modeling, hierarchical modeling, aggregation and disaggregation, chunking, and building of lumped, reduced, or simplified models.

2. Definitions and Technical Background

Prefacing Comments

All models are abstractions of reality, but some have more detail than others. Detail depends on *scope*: the extent of the system, input domain, and output range treated; and on *resolution*: the level of detail at which system components and their behaviors are depicted. Models also vary in the *perspective* that they embody (e.g., what a physicist might call *representation*) and in their user modes. As a result, system models in particular are typically *multifaceted* in the sense of Zeigler (1984). This report touches on all of these matters, but focuses primarily on issues of resolution.

What Is Resolution?

While fairly standard, the above definition of resolution is ambiguous because each of the components of a model may have its own resolution. Or, to put it otherwise, resolution has many components as Figure 2.1 suggests. For example, high resolution may refer to including fine-grained entities such as individual tanks rather than tank companies, and richer depiction of the entities (e.g., aspect-dependent vulnerability as well as position, velocity, and rate of fire), a richer depiction of how entities' characteristics depend on each other, a more detailed description of processes such as attrition and movement, or fine-grained scales in space or time.

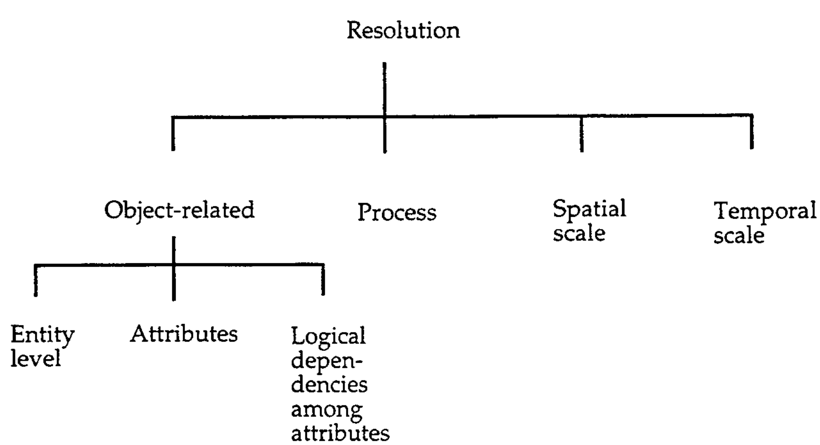

Figure 2.1—Dimensions of Resolution

In comparing two models A and B, it is often the case that model A has lower resolution in some respects and higher resolution in other respects. Indeed, some "high-resolution models" have remarkably low-resolution depictions of important phenomena. This, of course, creates confusion. Part of the art of multiresolution modeling is deciding what level of detail is needed in the various dimensions.

What Is Multiresolution Modeling?

Definition

If we now understand resolution, what constitutes "multiresolution modeling" (MRM)? We define multiresolution modeling as

(1) building a single model with alternative user modes involving different levels of resolution for the same phenomena;
(2) building an integrated family of two or more mutually consistent models of the same phenomena at different levels of resolution; or
(3) both.

Single-Model or Family?

Whether a single-model or family-of-models approach is preferable depends on circumstances and the extent of MRM desired. The theoretical model-design issues are very similar. Suppose a conceptual model decomposes hierarchically into subcomponents, where some higher-level variables are functions of lower-level variables, which are in turn functions of even lower-level variables. One could implement this all in a single program with nested functions. However, one could instead have separate programs for some of the low-level functions. Variables that previously were calculated from lower-level functions would now be input parameters. Those parameters could be calibrated from time to time by running the separate low-level programs (Figure 2.2), perhaps across a statistically weighted range of cases. In some cases the higher-level model's input might be a probability distribution (i.e., a random variable) rather than a simple parameter. That is, the calibration would generate a distribution function rather than a single number. In either case—whether the entire conceptual model is implemented as a single program or a family—what matters most theoretically is whether the connections between levels of differing resolution are well defined, meaningful, and understandable. If so, then we have a good MRM design.

This said, there are substantial *practical* differences between single-model MRM and family-of-models MRM at the level of software, interfaces, modularity, and burden of use and maintenance.

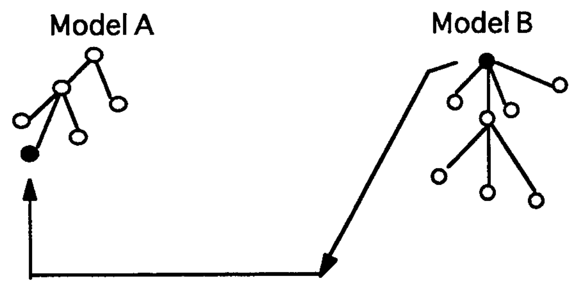

Figure 2.2—MRM Within Single Models and a Family of Two Models

Although a single model can have some MRM, the burden of complexity grows rapidly with numbers of levels unless interactions are sharply circumscribed.[9] In problems of interest it is not uncommon to have a dozen possible levels of resolution (e.g., from signal processing by a radar detector up to a force-on-force competition between aircraft and surface-to-air missile defenses). It would be folly to try to represent all of this in a single program. On the other hand, it is common in analysis and other applications of models to need to go up or down a level or two in resolution and to want the ability to do so within a given model. For example, one may want the option to represent strategic mobility explicitly in terms of airlift and sealift resources, to represent the net effect of those resources in measures such as millions of ton-miles per day, or to "script" the arrival of the systems being deployed.[10]

Our general conclusion on the issue of single-model versus family-of-models, then, is

- For larger models and simulations (e.g., JWARS and JSIMS), a combination approach is desirable: the basic model should have some selectable levels of resolution (MRM) , but for

[9]There are cases where a single model can reasonably have many levels of resolution. One involved artificial-intelligence models of national-level decisionmaking in analytical war gaming (Davis, 1986). These models were built and used briefly in the period just before the end of the cold war. In combat modeling, however, dealing with many levels of resolution in a single model is usually very complex because so many object types must contain algorithms for interactions with many other object types (Hillestad and Moore, 1996).

[10]Throughout this report "scripting" refers to specifying in the program or input data at what time various events occur. A script might specify that the 1st Brigade of the 82nd Airborne Division arrives on Day 4. Another script might specify precisely when particular units maneuver and to where; it might go on to specify when they would engage the enemy and for how long before breaking off battle. Such scripts substitute for decision models and other more explicit processes. In practice, workers often iterate the scripts until the simulation "tells a credible story" for baseline values of its input data. Sometimes, such scripts are useful. In other cases, scripting is so overused as to undercut any claim that the phenomena of interest are actually being simulated. Models are often sensitive to details of the scripts (especially details related to command decisions), but the scripts themselves may be held constant as other inputs are varied. This is dubious practice.

work crossing many levels of resolution there should be an integrated family of auxiliary models for occasional tuning and special purposes.

These auxiliary models would be hierarchically integrated—i.e., designed together with the basic model—so that, for example, experiments with a high-resolution auxiliary model could straightforwardly generate input parameters (or distribution functions) for the basic model—with different calibrations for substantially different regions of the problem space. Although a number of DoD and allied military organizations have developed hierarchical model families over the years, few have been so integrated.

Integrated models should have "good seams" in the sense that one should have cognitive and predictive consistency as one crosses levels of resolution. That is, if one moves from resolution A to resolution B in viewing a system, one should readily understand the transition and the relationship of higher and lower-resolution variables, and one should get consistent predictions. To put it differently, good seams imply cognitive and predictive consistency (Davis and Huber, 1992). It is extremely difficult to achieve this unless the models have been designed together. Table 2.1 indicates something of the scope of current DoD models. None of the models listed were designed together, making mutual calibration very difficult.[11]

The other point Table 2.1 allows us to make here is that whether resolution is "high" or "low" is in the eyes of the beholder. Looking at ground-combat examples, people using entity-level simulations such as Janus would consider models such as Vector II, TACWAR, or JICM (Joint Integrated Contingency Model) to be quite aggregated, while those using these latter models would instead associate "highly aggregated" (or "low resolution") with models like Trevor Dupuy's Quantified Judgment Model (QJM) (Dupuy, 1987) or the constant-coefficient Lanchester models used in historical analysis (Bracken, Kress, and Rosenthal, 1995). And, to close the circle, those who work with engagement-level models involving, for example, aspect angles and low-signature ground vehicles would regard even models like Janus to be aggregated. In this report we will frequently refer to high- and low-resolution pairs of models (or to base and aggregated models, base and lumped models, or base and simplified models) without indicating what their absolute levels of resolution are, except for the sake of examples.

[11]There are examples of one-way calibration. One is that the Army's Concepts Anaysis Agency (CAA) uses the ATCAL methodology to calibrate attrition parameters in its combat effectiveness model (CEM) with results of a higher-resolution model called COSAGE. The German IABG has a multilevel family for upward calibration (Scheckeler, 1992). Some of the difficulties with such families are discussed by Harshberger, Bennett, and Frelinger (1992).

Table 2.1

Illustrative Range of Nonintegrated DoD Models with Varied Scope and Resolution

Level of Model	Scope	Level of Detail	Time Span	Outputs	Illustrative Uses	Examples
Theater/ campaign	Joint and combined	Highly aggregated	Days to weeks	Campaign dynamics (e.g., force drawdowns, movement)	Evaluation of force structures, strategies, balances; wargaming	CEM, TACWAR, Thunder, JICM
Mission/ Battle	Multiplatform	Moderate aggregation, with some entities	Minutes to hours	Mission effectiveness (e.g., exchange ratios)	Evaluation of alterna-tive force-employment concepts, forces, systems; wargaming	Eagle, Vector II Suppressor, EADSIM, NSS
Engagement	One to a few friendly entities	Individual entities, some detailed sub-systems	Seconds to minutes	System effectiveness (e.g., prob-ability of kill)	Evaluation of alterna-tive tactics & systems; training	Janus, Brawler, ESAMS
Engineering	Single weapon systems and components	Detailed, down to piece parts, plus physics	Subseconds to seconds	Measures of system perform-ance	Design and evaluation of systems and subsys-tems; test support	Many, throughout R&D centers

SOURCE: NRC, 1997, Appendix E.

Can Higher- and Lower-Resolution Models Be "Consistent"?

A key issue in MRM is whether higher- and lower-resolution models can be "consistent." The answer is obviously yes in a practical sense, since we use low-resolution models in our daily activities and they prove to be "good enough." For example, we use classical mechanics rather than quantum mechanics and thermodynamics rather than statistical mechanics. In the military domain we use doctrinal planning factors rather than detailed wargaming, and aggregate wargaming rather than entity-level simulation in command post exercises. But what do we mean at a deeper level when we refer to consistency between two models? We shall repeat the standard discussion of this, point out its flaws, and then provide a revised and more useful version.

Traditional Discussion of Model Consistency

A Concrete Example to Define Issues

Before discussing issues at a generic level, let us consider a military example to illustrate key points. Consider the high-resolution model of a pair to be a stand-alone entity-level simulation model of combat between Blue and Red units of any size from a company to a brigade. Let the low-resolution model be an algorithm embedded in a corps- or division-level simulation, the purpose of which is to compute attrition and movement of the forward line of troops (FLOT) resulting from an encounter between a Blue battalion and a Red company, battalion, or brigade. The high-resolution model (e.g., Janus) might obtain its results by simulating duels between individual tanks or other weapons, consider terrain in such detail that it could calculate lines of sight, and consider time by the minute. By contrast, the low-resolution model might calculate a situationally determined combat-power score for each side, estimate attrition as a function of those scores, and then estimate FLOT movement as a function of the ratio of fractional attritions. These two models would be "consistent in the aggregate" (the usual meaning of consistency) if they produced the same aggregate results, or at least results that are sufficiently close when provided with inputs for the same case.

Because the two models have different levels of detail, an aggregation function would be needed to map high-resolution cases (both inputs and outputs) into low-resolution cases. Thus, the high-resolution model might compute the position of the FLOT as a line trace described by many points, while the low-resolution model might represent the corresponding part of the corps or division FLOT as a piston's edge, and denote its position by a single number. The aggregation function would specify how the high-resolution FLOT description is collapsed to that single number. On the input side, the high-resolution model would typically require many more inputs than would the low-resolution model, inputs such as inventories of individual weapon types rather than a measure of overall combat power.

More General Discussion

The diagram in Figure 2.3, or some variant, is a traditional way of making these points more generically.[12] Let us first review this description and then present what we believe is a better depiction. Figure 2.3 shows initial and final states in the corners of the rectangular diagram. Model A is a high-resolution model and appears in the forward plane; Model B is a lower-

---

[12]A number of workers (including one of us) have used similar diagrams (e.g., Davis and Hillestad, 1992a, b; Axtell, 1992; and John Holland in 1986 [referred to but not cited in Axtell's work]).

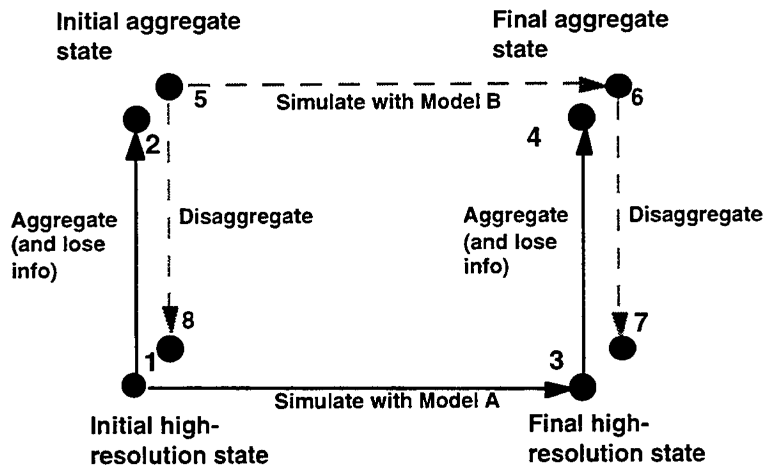

Figure 2.3—Usual Depiction of Weak and Strong Model Consistency in MRM

resolution model and appears in the rear plane. We assume that the initial aggregate states (2 and 5) are the same; that is, there is a valid way to aggregate from the initial high-resolution state.

In this depiction, it is usually said that A and B are "weakly consistent" if the final aggregate states 4 and 6 are the same, that is, if one gets the same result by starting with the high-resolution initial state, simulating with Model A, and then aggregating (1 to 3 to 4), or by starting with the initial aggregate state and simulating with Model B (5 to 6). "Strong consistency" is then usually said to exist if the final high-resolution states are equal (3 and 7), that is, if one gets the same results with paths 1 to 3, or with 5 to 6 to 7. The diagram highlights the fact that information is lost in aggregation, thereby leaving the impression that strong consistency is impossible or quite unusual. It would also be reasonable to come away from looking at Figure 2.3 with the belief that the higher resolution model is inherently superior.

Why the Usual Depiction Is Misleading

Although the diagram in Figure 2.3 is valid when properly presented, the image it conveys tends to focus thinking on the wrong issues. The problems here are subtle, and even insidious.

Figure 2.3 is unfortunate first because it conveys the notion of a closed system with all information contained in the detailed model. In practice, we more typically possess important information at different levels of resolution, and some of the lower-resolution information is crucial. In military domains this often relates to higher-level command-control processes (Davis, 1995). For example, a real-world combat system includes commanders who adjust strategy and tactics in response to circumstances. So it is that real-world commanders do not typically fight to the death in an obvious losing attrition battle. Instead, the outclassed commander will retire his forces to fight another day. Similarly, an Air Force commander suffering high attrition rates will not continue mindlessly, but will adjust strategy by such means as sending his aircraft to other targets or holding them back until air defenses are suppressed. That is, real military systems may be to some extent self-regulating, but these regulating processes are typically not represented in "high-resolution" models, because of their limited scope and dependence on scripting.

This frequent inability to understand macroscopic phenomena with bottom-up modeling has been celebrated in recent years as part of the discussion of complex adaptive systems (CAS). There are many examples, but one will suffice here to demonstrate that the issues are far more general. In recent years, researchers at Los Alamos National Laboratory made heroic efforts to simulate and understand traffic patterns in the city of Albuquerque. Their simulations were quite detailed, their computers extraordinarily powerful, and their results miserable. They then started over with a drastically simplified model of the traffic system, one based on a "cellular automata" description of automobile movements. Instead of carrying along a rich set of automobile and driver attributes, the researchers instead gave the automobiles very simple movement rules (e.g., speed up if the car ahead is drawing away, but slow down if the distance to the next car is too short). With a relatively primitive set of such rules, the simulation generated traffic patterns that could be readily calibrated to observed patterns. It was a few rules coupled with some gross features of the road network that determined results, not all the other myriad details that bottom-up modelers might think of.[13]

Because of such realities, it is better in system modeling to take a view like that suggested in Figure 2.4 (see also in NRC, 1997). Here the arrows indicate data that can be used to inform models at the different levels. Obviously, a model at any given level depends on data at that level (the horizontal arrows). However, if it is to be as realistic as possible, the model may also benefit from data pertaining to higher- or lower-resolution phenomena and to the models that represent them. Historical experience and military judgment, for example, may set bounds on plausible attrition levels as suggested above. Indeed, this macroscopic, aggregated, information is sometimes critical to higher resolution models. That is, calibration can be top-down as well as bottom-up. The image, then, becomes one of *mutually* calibrating a hierarchical family of models, rather than one of bottom-up calibration.

---

[13]Based on discussions with Chris Barrett and Darryl Morgeson at Los Alamos National Laboratories.

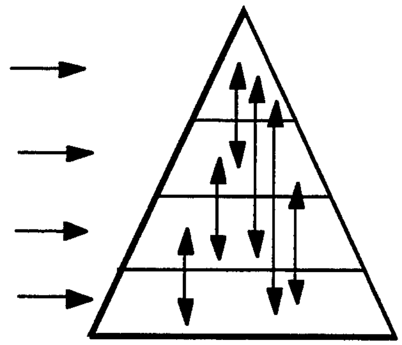

Figure 2.4—Mutually Calibrated Models in an Integrated Family

There are other problems with Figure 2.3. For example, it focuses on whether results with Models A and B are equal, rather than "close enough," whatever that might mean. And, finally, it conveys too negative an impression about the feasibility of respectable disaggregation by focusing on the loss of information during aggregation. We shall try to remedy these problems in the following section.

An Improved System Discussion of Consistency

With these items in mind, consider Figure 2.5. The first improvement is that the diagram includes funnel symbols denoting "projection operators" to emphasize that what matters in the context of an application is some processed version of some portion of the final state, not the full state itself.[14] For example, what matters may be the validity of a graph on a briefing chart, an object's position on a video monitor, or the rank order of some option in terms of cost-effectiveness. In each of these cases the "result" being focused upon involves much less information than the final state of a simulation. In Figure 2.5 we depict this in the top-right corner by indicating a final aggregate simulation state with Model B (6), which is then processed by a projection operator to create "aggregate results" (6b). Alternatively, one might have used

[14]Projection operators are used commonly in theoretical physics and some other domains, although under a variety of names. For example, when we characterize missile accuracies by a CEP, it is usually because—in the context of a particular application—it is unnecessary to worry about the detailed two-dimensional spatial distribution of impact points. We are "projecting out" (focusing on) an abstracted circular representation of impact points. An even more familiar example is when we "project out" the vertical component of acceleration of a moving body, because it is the vertical component that determines when the body will strike the ground.

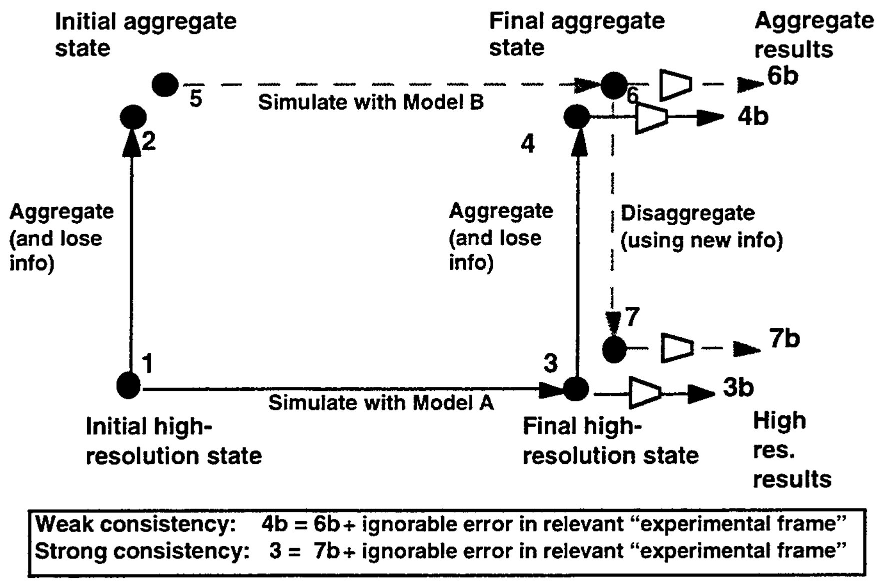

Figure 2.5—An Improved System Diagram

the higher-resolution Model A to end up with its version of the final aggregate state (4), and then processed that result to achieve result 4b. The question for "weak consistency" is whether 4b and 6b agree adequately. Do the corresponding graphs tell the same story? Do the video-monitor pictures look pretty much the same? And so on. The standards might be lower in program analysis, for example, than in mission rehearsal. The projection operators would also be very different (e.g., the former might involve a weighted ensemble average over future war scenarios, while the latter would be much more focused).[15]

Figure 2.5 also indicates cryptically (words along arrow from 6 to 7) that the disaggregation process may exploit additional heuristic information from history, doctrine, or other sources. Such information may be quite good and may not even be reflected in the "detailed model." Indeed, as noted above, it is common, not unusual, for the "detailed model" to have limitations of scope that omit some of the very information needed. A familiar example in combat modeling is that high-resolution models often depend on scripting to assure proper entity behaviors. They often do not explicitly represent the command-control reasoning that justifies that scripting. Ironically, the most important features of that reasoning sometimes *are* represented in more-aggregate models. At the same time, the information discarded during the aggregation process (1 to 2) may not be significant in determining subsequent states. For example, the detailed disposition of troops on the battlefield today may have nothing to do with their detailed

[15]These matters can be discussed rigorously in terms of *morphism requirements* within *experimental frames* (Zeigler, 1998).

disposition of troops on the battlefield today may have nothing to do with their detailed disposition tomorrow if the unit picks up its kit and maneuvers to a new location for tomorrow's battle. In contrast, if the unit is fatigued and demoralized today, then that may affect its ability to accomplish that maneuver and prepare for tomorrow's battle. That is, *some* of today's detailed state information may matter a good deal more than other information. In good aggregate modeling, one would try to reflect the essence of the relevant detailed information.[16]

Many other such examples exist, most of which are at first intellectually offensive. So it is that many theater combat models use an aggregate-level algorithm to compute attrition of either ground forces or air forces, following which they apply a disaggregation algorithm that allocates that attrition. Thus, "equivalent aircraft" may suffer attrition, which is then allocated among F-15s and F-16s from different bases. When modelers use these procedures, they are either bringing to bear additional information, making additional assumptions, or both. There are instances in which one would expect heuristically that all entities of a broad class would take attrition on a pro rata basis. In contrast, there are instances in which one might expect that the entities with larger scores (due to larger lethalities) would also be likely to have lower vulnerabilities and lower attritions. In such a case, one might allocate the attrition in inverse proportion to their scores. Our point here is not to argue for one or another assumption, but rather to argue for the under-appreciated point that such disaggregation heuristics can be legitimate and not a mere expedient by lazy modelers: they can reflect valid information.[17]

A Conflict Between Legitimate Paradigms

Finally, a word or two on the in-context comparisons of results. Here it is useful to distinguish sharply between two paradigms that commonly cause serious disagreements:

- Analyst's paradigm: what matters is whether the simplifications involved affect the study's conclusion (or the trainee's performance).
- Simulationist's paradigm: we are attempting to construct a virtual reality that will allow us to fight all kinds of simulated battles. To be sure, there are bounds on the circumstances that can be simulated, but we are concerned with far more than consistency with respect to any particular set of questions: we want the simulation to be as "real" as possible—at least in important respects.

Is it any wonder, then, that individuals disagree vociferously about the desirability of various model abstractions, including aggregation? We, the authors, grew up with the first paradigm

[16]Paul Reynolds and coworkers at the University of Virginia have been exploring software mechanisms for carrying along only the necessary detailed information. See Reynolds, Srinivasan, and Natrajan (1997).

[17]For discussion of how such techniques can be used, see Allen (1992).

and still consider it central to good analysis. However, as simulation technology has become more powerful and the opportunity to experiment with increasingly valid virtual worlds has grown, the simulationist's paradigm has become very important as well. Keeping straight which paradigm is relevant is not always easy. Let us note, however, that abstractions such as denoted in Figure 2.5 might be fully adequate for answering some specific study issue, but altogether inappropriate as part of a virtual-world simulation that workers want to use to represent the real system's dynamics.[18]

[18]An example of this is the abstracted (low-resolution) depiction of C^4ISR in otherwise high-resolution simulation of advanced concepts for using long-range precision fires to assist ground forces (Matsumura et al., 1997). The abstraction was quite adequate analytically, but would not have allowed simulators to "see" the C^4ISR platforms or their dynamics (e.g., orbiting).

3. Why Is MRM Important?

Workers in modeling and simulation (M&S) use both low- and high-resolution models, but the majority tend to place greater trust in high resolution. Many even say that as computing power grows, the need for low-resolution models will diminish. So also, military officers and civilian officials have sometimes taken this view, especially when they are more acquainted with computers, software, or even cut-and-dried "hard" engineering models than with higher level, squishy policy analysis. In any case, this impression is wrongheaded:

- The need for models at multiple levels of resolution, from high to low, will be nearly as strong a hundred years from now as it is today.

Here are some of the reasons.

Cognitive Needs

Although often underplayed, the most important issue is probably that humans need *models* (not just data displays) at many different levels of resolution and with many different perspectives on the world. After all, information exists and decisions are made at diverse levels. *Reasoning* occurs at different levels, and each level of reasoning has its own natural variables and concepts of cause and effect. To accommodate this, different models are *required*. One cannot readily use models from any one level to understand and accommodate information at a different level. When trying to work at any baseline level of detail, we need to look downward into greater detail, to understand what is going on, and also upward into greater abstraction, to see forests rather than trees and to make choices amidst uncertainty. That is, we are consistently challenged with the need to operate at different levels of resolution. Figure 3.1 suggests some examples.

Economy

Virtually every analyst agrees that it is sometimes necessary to use low-resolution models, because high resolution comes with a cost. Models become more complex, making them harder to program, debug, and validate. Inputs become more numerous, making cases harder to describe. Each case takes longer to run and, more important, to comprehend and analyze

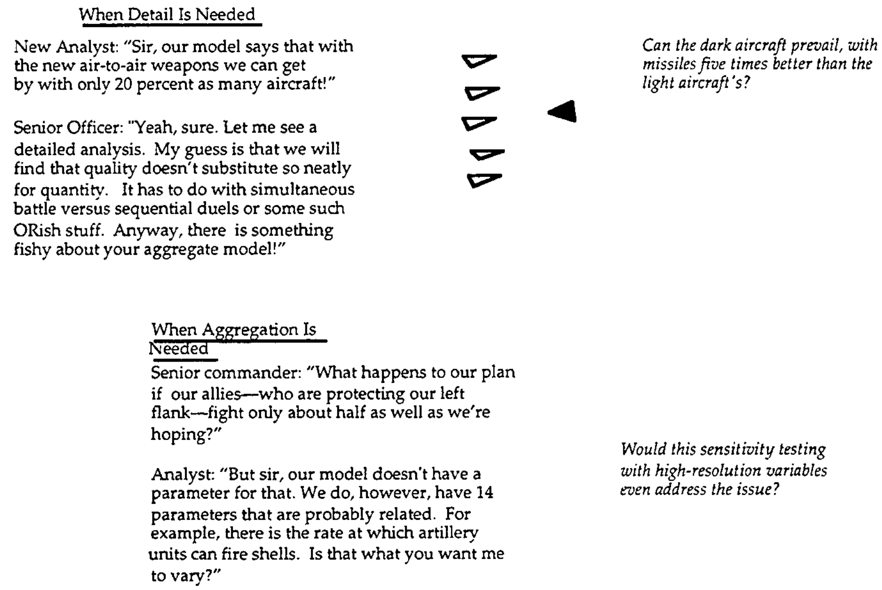

Figure 3.1—On the Need for Models at Different Levels of Aggregation

(cognitive issues again!) . Every analyst knows that the model does not do the analysis; rather, it merely provides inputs for analysis. Thus, project time and resource constraints may make it impossible to employ only high-resolution models.

Importantly, computing power is not the bottleneck in many of these steps. In fact, it is the main resource only in running the model. Building and debugging the model, collecting and validating inputs, analyzing output, and ultimately *understanding the problem and issues*—are still analyst intensive. Although new computer tools can help greatly, especially once the problem and model are properly structured, the rate of improvement is much less than that occurring in computational power. As a result, there will continue to be a premium on model simplification.[19]

Figure 3.2 indicates some of this notionally. As the model we use becomes more complex, the time and money required to develop, service, use, and understand its results grow, and eventually "take off" (A curves). Although better models, tools, designs, and displays can delay this effect (B curves), it does not go away.[20] On the other hand, if the model is too simplistic, then

[19]Related points are made by John Doyle in NRC (1997), Appendix B, page 142. He draws on engineering examples to discuss how costs grow faster with complexity than does error reduction.

[20]As examples of what might come into play for curve B, we think of models that better exploit multidimensional linear algebra, tools such as *Mathematica* ™ and high-level programming languages, visual display systems, automated search devices, and so on. There is a

the time required to understand it may actually grow, because it may take many false leads to recognize that the model is obscuring something important. Although schematic, the figure corresponds to the long-standing analyst's credo that the idea is to have a model that is as simple as possible, but not an iota more so. For the schematic of Figure 3.2, the optimum is roughly in the middle. The errors of being simplistic are avoided, and development and use costs are not too high.

Explanatory Power

Analysts frequently perform "back of the envelope" calculations (or, in today's world, simple spreadsheet calculations). Moreover, they worry when results thus obtained contradict the results from a highly detailed model, because low-resolution models can provide transparent and persuasive explanations of the results, which may stimulate decisions and actions. In contrast, high-resolution models are often so detailed as to be opaque. In such cases, one can only say,

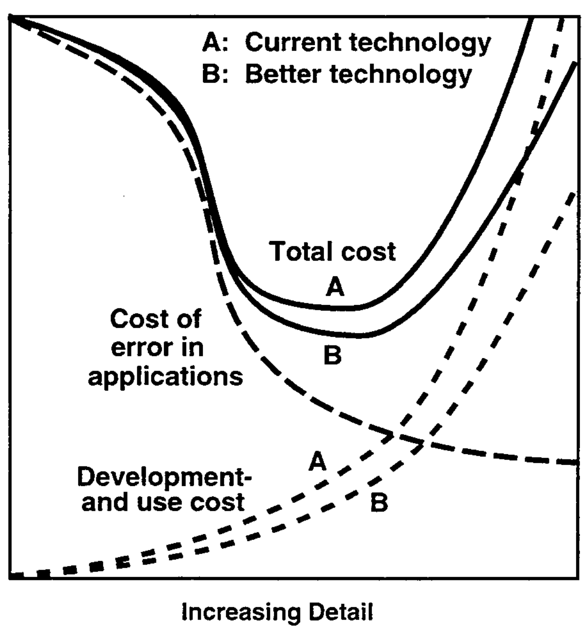

Figure 3.2—Time for Problem Solving Versus Model Complexity

"Well, the computer said so!"[21]—something anathema in good analysis except in rare cases when the model and data are known to be accurate.

Uncertainty, Ignorance, and Chaos

An important reason for MRM is that we often simply don't know enough about a system or process to develop a high-resolution model or, if one exists, to specify its inputs. A combat model, for example, may consider six or eight different kinds of fighter aircraft. But because one cannot know how many of each kind will be used in a particular future conflict, it may be just as accurate to use a single generic fighter. Or a cost model may estimate the cost of a new weapon system as the sum of hundreds of cost components, but the cost of each component will be an educated guess. Errors propagate upward and the net result can be quite wrong (and usually too low). Often in such cases one can obtain a better estimate by multiplying the anticipated weight of the new weapon by a historical cost per pound!

Even more profound issues relate to what information is and is not knowable and available. These issues are discussed under the rubric of *bounded rationality* by Nobel laureate Herbert Simon (Simon, 1996). Nonlinear dynamic models may also be *chaotic,* meaning that two cases with seemingly trivial differences in their initial conditions may evolve along very different trajectories.[22] Even very good knowledge of the initial state (i.e., with very small uncertainties) may still correspond to gross uncertainties about some aspects of later events (and yet other aspects may be well behaved and thus predictable in principle).

Thus, incorporating more detail in a model may merely spread the analyst's uncertainty across multiple factors.[23] Detail in a model provides a place to record information *if* it is available, but the mere presence of detail cannot create information.

---

great deal of potential here, but we see no evidence that analysis will ever become quick and simple for the difficult problems faced by policymakers.

[21]This is a subtle issue, however. As Donald Blumenthal, creator of Janus, has often noted, some high-resolution models are actually simpler to comprehend and explain than competitive low-resolution models. They may have many data elements and equations, but these may represent straightforward physics and other processes. In dealing with issues of strategy, policy, and force structure, however, the models needed—whether high- or low-resolution in nature—are inherently more complex. In these instances, cognitive and explanatory complexity usually increases rapidly with level of detail.

[22]This has been called the "butterfly effect," in which a butterfly flapping its wings in Hong Kong in December could in theory cause a storm to hit Oregon the next March (see Lorenz, 1993, although he used different cities).

[23]This point can be made precise using information entropy. It corresponds to the routine observation that—despite what a particular detailed model would allow given apparently plausible data—we *know* that system behavior will fall in a particular range. In such an instance we are reflecting macroscopic, aggregate-level knowledge.

Exploratory Analysis and Analytical Agility

Broad Observations

Yet another reason for MRM has been emphasized in our own research, and indeed is rather dramatically illustrated in an example that we work out in detail later in this report. For reasons just discussed, many models are only weakly predictive[24] and point predictions from such models have little value. To realize the value of such models, one needs to study effects of the uncertainties using something like what we call *exploratory analysis.*[25] One runs a multitude of cases with inputs, assumptions, and parameter values ranging over all the relevant circumstances. One is not searching for a policy that does well in a single base case. Indeed, *there is no meaningful base case because a wide range of cases are quite plausible.* Rather, one seeks a policy that does well in that whole range of circumstances. The policy must be robust and adaptive.[26]

Detail is the enemy of exploratory analysis. With more detail, not only does each case take longer to run, but there are more parameters to vary and the dimensionality of the case space increases. Each additional parameter tends to multiply the number of cases by a constant factor (e.g., if one defines N test values for each parameter, and wants to run all combinations, then adding a new parameter multiplies the number of cases by N). Careful experimental design can save a good deal of effort,[27] of course, but the explosion of work as a function of dimensionality is a fact of life.[28] Unfortunately, past models have generally not been designed with exploration or MRM in mind. RAND's JICM, a descendent of the RSAS, is a partial exception, as are some contemporaneous models such as the Air Force's Thunder.

One conclusion of our study is that MRM is *essential* for exploratory analysis examining problems across the many dimensions of uncertainty. This is so because it is neither computationally nor cognitively feasible to do serious exploration—including keeping track of causal relationships—if

---

[24]For more discussion of the "weakly predictive" concept see Dewar et al. (1996).

[25] Exploratory analysis has become a theme of RAND analysis for numerous DoD sponsors. For its roots in military analysis, see Davis (1994), particularly Chapter 4 and references there to 1980s' work with the RAND Strategy Assessment System. See also Bonder (1994) for an independent development of similar ideas. For broader discussion of the technological implications and challenges, see Bankes (1992).

[26]These notions are not altogether new, of course. Good analysts have always aspired to such broad-ranging investigation, and there are many examples of good sensitivity analysis in the literature. However, we believe that what we are promoting offers a good deal more: a fundamental shift away from "best estimates with some sensitivity analysis " toward a more profound appreciation of the manifest uncertainty in typical policy problems.

[27]See Cohen, Rolph, and Steffey (1998) for a review written in the context of testing. See also Cloud and Rainey (1998).

[28]We understand, of course, that Monte Carlo methods permit analysis with many uncertain variables. However, interpretation of results leads back to a need for fewer variables.

the number of variables is too large. Further, if ambitions for detail increase, the number of high-resolution variables explodes far faster than computational power is increasing.[29]

Applications of Exploratory Analysis

There have been numerous applications of exploratory analysis. One is the demonstration that "capabilities-based planning" in a scenario space of operational circumstances is both feasible and desirable (Davis, Hillestad, and Crawford, 1997). Another is recent related work in preparation for the Quadrennial Defense Review (QDR) (Davis, Gompert, and Kugler, 1996). OSD's Office of Program Analysis and Evaluation used the methodology to some degree in assessing for the QDR the value of different degrees and levels of modernization. Lempert, Schlesinger, and Bankes (1996) applied the concepts to the global-warming policy controversy. Dewar, Bankes, Hodges, Lucas, Saunders-Newton, and Vye (1996) discussed the implications of the concept for *credible* analytical uses of distributed interactive simulation (DIS) for the Army. Brooks, Bankes, and Bennett (1997) used exploratory analysis in a weapons-mix application. Other RAND colleagues have used the methods in examining Air Force logistics and Army division design (Jack Abel and Lou Miller, and Lou Moore, respectively). Indeed, exploratory analysis is seen as a key to the sound use of M&S for operations planning, program decisions, and many other purposes.

If we are correct, exploratory analysis will be a major element of future work, not only in policy analysis, but also in real-time decision support for commanders (NRC, 1997, 39ff). The relevance here is that we argue that

- Multiresolution modeling is essential for achieving the potential of exploratory analysis.

The reason, simply, is that MRM can enormously reduce the degrees of freedom, lessening the curse of dimensionality.

Finally, let us note that "simple models" (models with few degrees of freedom) also tend to increase analytical agility: it is easier and faster to change assumptions, to run the model, and to interpret the results. The trick, of course, is to assure that the model has not been overly simplified to the point of being misleading.

[29] We thank Professor Bernard Zeigler of the University of Arizona for a clear proof of this. See Zeigler, Praehofer, and Kim (1998).

Complex Adaptive Systems and Emergent Behaviors

An important motivation for MRM, one apparently not widely recognized a decade or so ago, is that complex adaptive systems (CAS) often exhibit regular, coherent behaviors at a macroscopic level that are not readily understandable in terms of the microscopic laws that govern the system. These are called *emergent behaviors*. Their existence is one reason that CAS research is so exciting today.[30]

The Second Law of Thermodynamics is an example of emergent behavior. It states, roughly, that an isolated thermodynamic system evolves from ordered to disordered states, never the reverse. Yet every movement, collision, or change of state of the individual particles that make up the system is completely reversible. This seems initially to be paradoxical. In this literature, life and intelligence are said to be *emergent behaviors* of biochemical systems. Although some of the paradoxes are now understood (see, e.g., Schrödinger, 1956) we don't yet know how to explain self-conscious, self-directed behavior, or consciousness itself, in terms of the chemistry of large molecules such as proteins and DNA.[31] Turning to military examples, we might consider trench warfare on the Western front during World War I to be emergent behavior. Nobody intended that it should happen, and the generals on both sides sought repeatedly to break open the combat and turn it into maneuver warfare. Yet some combination of circumstances (e.g., weapons available, density of troops) "conspired" against the wills of all concerned to force the fighting into the trench warfare mode. Another interesting example can be seen in recent Marine-sponsored research with agent-based simulation of small-unit operations. A relatively small set of low-level behavioral rules, when combined with an otherwise rather aggregated description of the units, can generate a wide range of interesting and doctrinally provocative behaviors at the force-on-force level (Ilachinski, 1996). No such behaviors would be observed in a traditional bottom-up detailed simulation.[32] It is noteworthy that the Marines have seriously exploited some

[30]The best broad and popular account is Waldrop (1992), which discusses and gives citations to the pioneering work of the Sante Fe Institute. An excellent overview by one of the field's pioneers is Holland (1995), which notes the importance of what we call aggregate-level models. For an accessible but reasonably technical survey, see Coveney and Highfield (1995). Nicolis and Prigogine (1989) is a more technically demanding tour of the horizon.

[31]Many scientists are addressing the issues, however. There even exists a *Journal of Consciousness Studies*. See also Crick (1994).

[32]Similar conclusions have been seen in fascinating work mentioned earlier on transportation modeling by Chris Barrett and Daryl Morgeson of Los Alamos National Laboratory. Some of this is described in a popular discussion of simulation and CAS at the Sante Fe Institute (Casti, 1997, Chapter 4).

of the more general concepts and many of the metaphors of CAS research in rethinking doctrine for the 21st century.[33]

Nobel prizes can be won for explaining emergent behavior from a reductionist perspective. But such explanations are often restricted to simple instances of the emergent behavior, or are partly qualitative in that they require calibration parameters to fit them to observation. Thus it is often necessary to construct models that deal with the behavior directly, rather than as a consequence of lower-level detail. That is, CAS phenomena are another motivation for aggregate models, or at least for partially aggregated models.[34]

Phenomena Difficult or Impossible to Model in Detail

As our final observation on why MRM is important, we note that it is not just the emergent phenomena of CAS that are difficult or impossible to capture in a high-resolution model. As a practical matter, there are many examples because high-resolution models must be limited in scope, lest they become so unwieldy as to be useless. The most obvious such military examples are command-control and information. Both may be represented in detailed, entity-level combat models only to the extent that analysts drive the models with finely tuned scripts meticulously constructed to assure that the entities behave sensibly given the circumstances (including command-control and intelligence). Because the information is implicit, however, the models are hard or impossible to use in calibrating the combat-multiplier effects of C^4ISR in more-aggregate models. This difficulty is perhaps not fundamental, in that detailed models could in principle have both scope and resolution, but it is extremely important in practice.

Friction in war is another example. Just as scientists do not typically calculate viscosities and conductivities from high-resolution models such as quantum mechanics, instead measuring them in their own macroscopic realm, so also analysts must use historical data and other information to scale parameters in high-resolution models so that they generate reasonable levels of attrition and

[33]Some of these issues are discussed in Alberts and Czerwinski (1997), which documents a conference on complexity held at National Defense University. See especially the chapters by Schmitt and Rinaldi, although in the former, one should recognize that the author's discussion of "Newtonian" physics is somewhat of a libel—Newtonian physics being by no means linear. See also USMC (1997). The influence of complexity on Marine doctrine is substantially due to the leadership of Lieutenant General Paul Van Riper (U.S. Marine Corps, retired).

[34]There is a poorly understood subtlety here. The adaptive "agent behaviors" so often critical to emergent behaviors may sometimes be regarded as high-resolution features. However, if a simulation describes an automobile (or an infantry platoon) with only some average parameters (speed, size) plus behavioral rules, does the model have high or low resolution? In many cases of interest the issue is moot, because the important behavioral rules apply to higher-level aggregate units (e.g., decisions of a corps or JTF (Joint Task Force) commander).

movement. The point here is one we made earlier: it is important to use all information, from whatever level, to *mutually* calibrate models at different levels of detail, rather than to imagine that the whole truth resides with the high-resolution models.

Let us end this discussion with another example that makes the points well, while simultaneously noting that the issues arise across the board—e.g., in logistics problems as well as in combat-modeling problems.

Some years ago, one of us attempted to estimate the required Air Force budget for recoverable aircraft components (Bigelow, 1984). These are aircraft components that can be repaired and reused, as contrasted with the consumable parts, which are discarded once they fail. The detailed model estimated how to manage individual recoverable aircraft components for Air Force Logistics Command item managers. It did an adequate job for perhaps 99 percent of the parts, and item managers dealt with the exceptions manually. The parts that were the exceptions changed over time, as one problem was fixed and another emerged. On the one hand, the model seems fully adequate. However, the detailed estimates for required buys of all the parts could be rolled up to get the total dollar requirement, and the resulting estimate of the budget proved to be completely wrong. It was precisely the exceptional parts, the ones that were not well handled by the day-to-day management system, that drove the overall budget, because the way to solve the problem of an exceptional part was usually to buy more of it. The budget for buying these parts had to be established two or three years before the parts it funded were delivered. And even though only a few parts were exceptional at any instant, a considerable number became exceptional sometime during that two or three years. The point is, the detailed model was adequate in its own context, but it omitted phenomena—such as the purchase mechanisms and time scales—that were crucial in the context in which the aggregate model was to be applied. To summarize:

- An aggregate model is often more than an aggregate version of a detailed model. It is a model with its own capabilities, its own uses. There is overlap between the aggregate and detailed models, but the set of phenomena represented by the aggregate model is often not included in the set of phenomena dealt with by the detailed model.

4. Obstacles to MRM

If MRM is so important, why isn't it practiced more? This was one of the questions we asked in the current study. Our conclusions on the matter follow.

Need for New Guiding Principles and Tools

One important reason is that model builders have no guiding principles for MRM design. Further, normal modeling practices are quite unfavorable to MRM. For example, good simulation modelers often leave room for generalization because the users they support invariably ask for embellishments. That is, the modelers avoid exploiting valid aggregate relationships because those relationships might fail in a more general version and the modelers want to avoid recoding. As we shall see, MRM designs also introduce considerable complexity—although it is complexity of a sort that can be greatly reduced by good software implementations. Our point here is that *some good non-MRM design practices are inappropriate if in fact we want MRM.* Another factor here is the relative absence of simulation environments that make MRM designs straightforward. A pressing need exists for new computational tools and improved environments.

Bad Examples Have Left Bad Tastes

Aggregate and other low-resolution models have a bad reputation in many circles—primarily because of the ubiquity of poor aggregate models that discard critical information. It does not follow, however, that they *needed* to be so foolish. Good aggregate models are often subtle, and even complex. Consider, for example, how an aeronautical engineer thinks about drag coefficients as a function of aircraft speed. Perhaps in freshman physics drag is proportional to velocity V, but in more advanced work it becomes a complex function with different regimes. (It can even vary inversely with V). But in each such regime the drag model is simple and aggregated. To take a military example, no one but an amateur believes that the 3:1 rule in ground warfare applies (if at all) to any but very special circumstances. Indeed, there are different rules of thumb ranging from 1:1 to about 6:1 for different levels and circumstances

(Davis, 1995).[35] In other cases, force ratio as normally interpreted is irrelevant. And, in air-to-air warfare, it is well known that the exchange ratio for dogfights between two qualitatively mismatched aircraft types can range from very large to something approximating unity, depending on the range of engagement and the numbers involved (n-on-1 versus 1-on-1 engagements).

Aggregating Detailed Models Often Produces Wrong Answers—but for Misunderstood Reasons

Paradoxically, one criticism of aggregate models should actually be turned on its head to "criticize" high-resolution models. Often if one aggregates a high-resolution model, the resulting model is clearly wrong in describing a range of aggregate-level behaviors, but the reason—alluded to earlier—is that information (e.g., about strategy or various constraints) critical to aggregate-level behavior was not in the "detailed model" in the first place and could therefore not be aggregated. Users of the detailed model avoid this problem (sometimes without even noticing it) by focusing narrowly on standard cases in which the various entities are meticulously scripted to behave reasonably. However, when either the detailed model or an aggregated version is used across a diversity of aggregate situations, the problems become obvious. At that point, the detailed modelers may claim that their level of detail is essential and may demand time (sometimes months) to build new scripted scenarios, when in fact the aggregate model could perform well if merely it were improved to include essential command-control rules (perhaps reflected in "agent models") and some other constraints. Such rules and constraints can also be included in detailed models, but such increases in flexibility and scope are not always desirable because of increased complexity and are certainly not always feasible with available resources and modeling methods.[36]

---

[35]We recognize, of course, that some statistically oriented analysts continue to publish articles about the alleged nonevidence for such force-ratio-related rules. However, such articles invariably reflect refusal to factor in a myriad of situational factors, including terrain, defenses, operational surprise, and the sides' qualitative fighting quality. Such factors are discussed in Dupuy (1987).

[36]This observation relates to a sensitive and often misunderstood point. Although it is true that traditional entity-level combat simulations have had extremely narrow scope and inadequate command-control, thereby making them inferior for some purposes to aggregate models, one goal of modern simulation should be to greatly improve these features of detailed models. This was the purpose, for example, of the CONMOD simulation at Livermore, which was never completed. It is a purpose of modern work with semi-automated forces (SAFOR). And it is a purpose of agent-based simulation such as that of Chris Barrett of Los Alamos National Laboratories or Andrew Ilachinski's work at the Center for Naval Analyses.

We shall return to this theme later, but this is also an appropriate place to emphasize one of our most important conclusions about aggregation:

- The strength and validity of aggregate models often depend fundamentally on their explicit representation of human interventions such as command-control decisions involving maneuver and adaptations. This phenomenon can correspond closely to what some workers call agent-based modeling.

These high-level interventions often dominate overall results of battle and can therefore be far more important than details of high-resolution interactions. To put it differently, they also reduce greatly the significance of the high-resolution information eliminated by aggregation.[37]

Honoring Approximations

Another reason that MRM is not more common is that people expect too much of it, set unreasonable standards, and then reject MRM when it fails to meet them. It is a *fundamental* point that even a good aggregate model will not be generally valid, and indeed may only very seldom be rigorously valid. The purpose of modeling, however, is to seek good approximations and to recognize that the approximations will be good only in some contexts (e.g., in some circumstances of scenario) to some degree of accuracy and precision. Different aggregate models will often be necessary for substantially different contexts. In normal activity all of us shift frequently from one mental frame to another, using related aggregate models to understand and deal with those frames (Figure 4.1). It can be difficult to estimate both the size of the error and the contexts in which an aggregate model is valid, but we get along reasonably well most of the time.[38] Representing such flexibility and common sense in models is not so easy.

Although the role of approximations may seem obvious, we concluded in our study that it was not. Many of the nonconverging and sometimes disputatious discussions about MRM, aggregation and disaggregation, and simple versus detailed models, has revolved around criteria of precise and general consistency. It is frequently easy to show that an aggregate version of a

---

[37]Examples here include (1) decisions to fall into march formation and maneuver to a different position, (2) decisions to reallocate air forces to sectors in trouble, (3) decisions to pull CVBGs back temporarily from an operating area with enemy submarines or cruise missiles, or (4) reworking the communication links in a network-centric Joint Task Force after losing one or two important platforms to attrition. In contrast, many users of old-fashioned models like TACWAR simulate peculiar wars in which the antagonists continue to "fight to the death" in pistons, with neither side adapting to circumstances, as in retiring to fight another day. Is it any wonder that warriors have so often found aggregate-level analyses unsatisfactory? The problem, however, has often been the poor representation of strategy and adaptation, not the aggregation.

[38]Axtell (1992) notes that one sign of a mature discipline is a set of commonly accepted heuristics for when to apply particular abstractions.

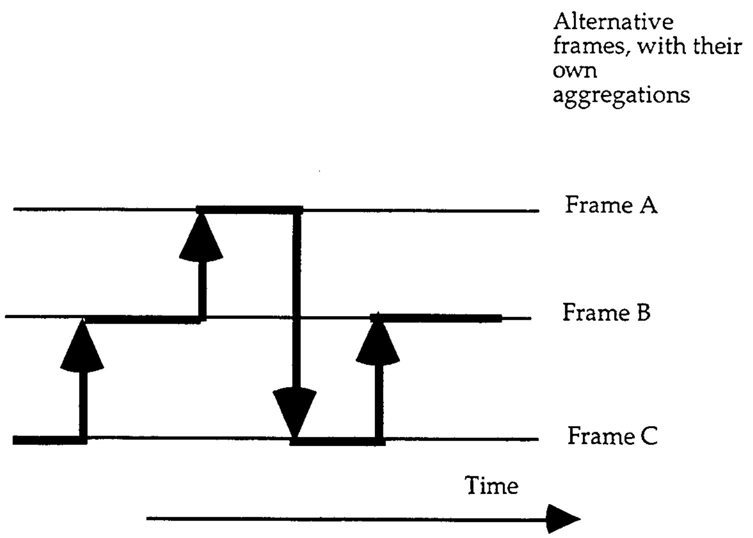

Figure 4.1—Frequent Shifts of "Frame" and Aggregation

model cannot be rigorously valid except in obscure and unphysical cases. Observers interpret this to mean that aggregation is therefore a flawed goal or downright bad. [39] However, the issue is the quality—in context—of the approximation. This cannot be assessed with purely "local" comparisons, because the approximation may be good only as an average over appropriate temporal or spatial scales.

- Most efforts to prove equivalencies (homomorphisms) between models at different resolution are probably doomed to failure unless the models being compared include projections onto the problem space and measure error rather than seeking exact equivalence.[40]

---

[39]As an example, consider responses to earlier RAND work (Hillestad and Juncosa, 1993), which derived the conditions under which the aggregate behavior of a system described by high-resolution Lanchester equations (a killer-victim-matrix level of description) could be described exactly by a scalar Lanchester equation. The conditions were stringent and some observers therefore concluded that the scalar equation was useless. This was a result of confusing necessary and sufficient and confusing exact and approximate.

A good deal of the academic literature on model abstraction (essentially MRM) focuses on cases of exact mathematical equivalence. See, for example, Simon and Ando (1961) or Courtois (1985). A major reason is that there was no real computational alternative in earlier years. However, we now know from research in physics and chemical physics, for example, that many important aggregations are merely good approximations of average behaviors on appropriate temporal and spatial scales.

[40]The important role of homomorphisms is discussed in Zeigler et al. (1998).

Probabilistic Analysis Is Often Needed but Is Nontrivial

A related point in understanding MRM and when aggregation is valid is recognizing that good MRM requires sophisticated treatment of probabilistic issues. We combine or drop variables and parameters when we aggregate, and these become "hidden variables" that contribute to unexplained variation in the aggregate model. Unfortunately, it is relatively unusual to find sophisticated treatment of probabilistic issues in the military M&S community—especially in work involving operational- or higher-level phenomena, or in higher-level force planning. To be sure, many high-resolution models are stochastic, but even they tend to focus on expected values rather than on probability distributions. Even worse, they often are stochastic in label only, with narrow distributions and "one-pass Monte Carlo" methods. Many aggregate models (e.g., TACWAR, JICM, and the basic version of CEM) are deterministic, even though they are sometimes used in contexts where stochastic representations could be important (e.g., at force ratios where a stochastic model would generate bimodal outcomes). In such cases, the models lack credibility and, again, give aggregation a bad name in the minds of those sensitive to more microscopic phenomena.[41] Finally, probabilistic analysis can be quite demanding when events are not probabilistically independent (Horrigan, 1991, 1992; NRC, 1997, Appendix J).

We do not discuss probabilistic issues much in this report but are preparing material on the subject for publication elsewhere.

Something Is Not Nothing

The last reason we note for MRM not having been common in the past is that people have sometimes viewed it as an all-or-nothing proposition. And, since designing a complete modeling system such as JWARS with sound and consistent MRM is well beyond the state of the art, if it is feasible at all, the tendency has been to fall back to a position of "it would be nice, but it's not feasible." This is wrongheaded because, in practice, it has often proven very useful to have some MRM features in a system even though a comprehensive treatment was too difficult. That is, the

[41] To be sure, good analysts can often work around the limitations of deterministic models, but this is often not well understood. Further, even good analysts often fool themselves on this matter by underappreciating the effects of probabilistic effects. See, e.g., NRC (1997), Appendices I and J; and NATO (1995). The Army's Concepts Analysis Agency (CAA) has noted the significance of stochastic effects in experiments with a modified version of CEM.

pragmatic approach is to do MRM on those model components where it *is* feasible. There is no all-or-nothing requirement.[42]

[42]Here we draw upon our experience with past RAND models (RSAS and JICM), which have had major advantages precisely because they had some MRM features, even though these were spotty and imperfect.

5. A Research Approach: Experimenting with MRM Design

Hypotheses About MRM Design

Because MRM theory is in its infancy, we cannot state definitively how to do it. Further, it—like modeling more generally—will continue to be as much art as science. We can, however, state the hypotheses that we have been exploring and testing in this project.

Adopt IHVR as a Design Principle

We know from past work[43] that the best kind of MRM, when it works, is integrated hierarchical variable-resolution modeling (IHVR). IHVR is quite different from approaches typically taken, even by excellent modelers. Key variables are described in hierarchical trees with the lowest-resolution variables at the top. Procedures necessary to relate variables at different levels can be cleanly specified and the calibrations accomplished on a modular basis (Davis and Huber, 1992; Davis, 1993). Aggregate variables (higher in the tree) are appropriate functions, not necessarily simple averages, of phenomena described by higher-resolution variables (lower in the tree). Significantly, some aggregate variables should be represented by probability distributions: over-averaging destroys essential information.

When we are viewing a problem that can be solved analytically, the tree can be thought of as a solution procedure. More typically, however, the tree will correspond to a model's data-flow diagram. Such diagrams are less commonly shown now because of the advent of object-oriented modeling (OOM). OOM typically focuses on hierarchies of objects, not processes, but it is the process hierarchies that cause much of the trouble in MRM. It is simple to define a division as the composite of its brigades but difficult to develop a division-level attrition equation that is consistent with the sum of attritions stemming from brigade-level equations—*unless* one assumes that complications average out.

[43]The best short introduction to RAND's past work is Davis and Hillestad (1992a). See Axtell (1992) for a bibliography of relevant academic work through the 1980s and an excellent discussion (a Carnegie-Mellon dissertation) of why MRM or "model abstraction" is difficult. His conclusions are a good deal more gloomy than ours for reasons we shall discuss later.

If one tries to sketch hierarchical trees for real models, one typically discovers that there are multiple interacting processes, each with its own tree, and that the various trees have cross-links between branches. That is, the model does not have a pure hierarchical structure. This reflects the "everything-affects-everything" aspect of the real world. This is why most people go no further.

In many practical problems, however, the variables can be so chosen that the cross-links between trees of variables are weak. They can either be ignored (perhaps bounds can be estimated on the error introduced thereby) or accommodated by second-order corrections. In this project, then, we were attempting to stretch the concept of IHVR to cover models that it does not fit in its pure form. Think of it, if you will, as an attempt to develop a theory of "integrated *nearly* hierarchical variable-resolution modeling."

Combine Computational and Analytical Approaches

We believe that progress in MRM depends on the combined use of both computational experiments and analytical thinking to help us *discover* and *calibrate* useful aggregations. Computational power is essential, because most real systems are too complex to treat with tractable closed-form mathematics. We know from our previous theoretical work that we should expect neither exact relationships nor universally valid ones. Instead, we should expect context dependence and dynamical behaviors. This is all the more reason for computational support. Tools are available in other disciplines that are at least relevant and perhaps directly useful. Of particular interest are methods from the "data mining" community, and other methods of statistical analysis, decision analysis, and evolutionary programming.[44]

On the other hand, we believe that much can be gained by using theory and mathematics to help the search process and by recognizing that a major obstacle to progress has been the unwillingness of most workers to step back from the form of thinking that simulation modeling involves—particularly "step-by-step thinking" when describing a system's evolution over time. That is, a strong dose of theory—even somewhat idealized—can help find those approximations that could be exploited in MRM design. We hope, of course, to make integrated hierarchical variable-resolution designs work—at least approximately.

[44]Our thinking on this matter has its origins in suggestions made to us by colleague Steven Bankes early in our project.

With this background, let us now discuss in more detail some of the experiments we conducted in the current project. The first set involved design; the second set involved computational efforts, which are continuing and will be reported elsewhere.

Approach

To test hypotheses and more generally gain insight about MRM, our approach—as in past efforts—has been to study a relatively simple problem in depth. In this case we used the problem both to consider the implications of designing for MRM and as a source of artificial data on which to do computational experiments in search of aggregate relationships. We shall describe the simple problem shortly, but before doing so, we want to highlight some useful distinctions.

Three Cases for MRM

Suppose we start with a valid "base model" and contemplate building a more abstract version by using aggregation and other means. There are three classes of problems worth describing separately. In Case A, the opportunity exists to rewrite the model without approximation in terms of aggregate variables that arguably clarify the model. In Case B, such rewriting is still feasible and useful to both understanding and analysis. In this case, however, the aggregations introduced either generate errors that must be proven ignorable or—to avoid serious errors—must include correction factors that may be fairly complex and may not be all that easy to fully comprehend. Finally, in Case C it is not even useful to reexpress the model to build in aggregate concepts, because doing so is difficult and tends to increase both conceptual and computational complexity unreasonably. An extreme example of this is empirical cost models—which are in some respects an insult to cost models developed bottom-up with great care and effort—because the empirical models have proven more reliable. In such cases it may be feasible to have model families that can be cross-calibrated to some degree, but the emphasis is more on model behaviors than on structural integration.[45] By and large, our focus in this report is on structural integration, but our discussion of Case C considers behavioral abstraction.

All three of these cases may arise in a given modeling system. That is, some model components may be amenable to Case-A MRM; others to Case B; and still others to Case C. As mentioned earlier, it is often a mistake to seek the Holy Grail of a comprehensive MRM design. It may be infeasible, and having MRM features in even some components can prove useful, particularly in

[45]This distinction between model behavior and structure is fundamental. It is addressed, for example, in Zeigler (1984), Fishwick and Lee (1996), and Caughlin and Sisti (1997).

modeling systems used for a diversity of problems. In some of those problems the MRM components will prove valuable, even though in other problems there is no recourse to the relatively detailed model. In this regard, we disagree with the frequently expressed adage that models should be kept "balanced." This position too often leads either to excessive detail or to avoiding the hard work of MRM because it cannot be comprehensive.

Case-A MRM: Rewriting Without New Approximations

Let us now elaborate on Case A, shown schematically in Figure 5.1. The initial model (contained by the outer, solid rectangle) is a function of high-resolution variables, but, upon inspection, one discovers that the model can be simplified by substituting aggregations. The result of rewriting the model is the lumped-model system indicated by the dashed rectangle. When the model is so rewritten, one can either use detailed variables or aggregate variables as inputs, depending on what information is most readily available and credible. This assumes, of course, that a transform exists to map the high-resolution inputs into the aggregate variables.[46]

There are many instances of Case A in practical problems. As discussed in detail in Section 6, in standard models involving aircraft attacking armored vehicles, the model's inputs include kills per sortie, sorties per day, and numbers of aircraft. However, the outcome depends only on *products* (or what others might call lumps and we shall call *aggregation fragments).* That is, what

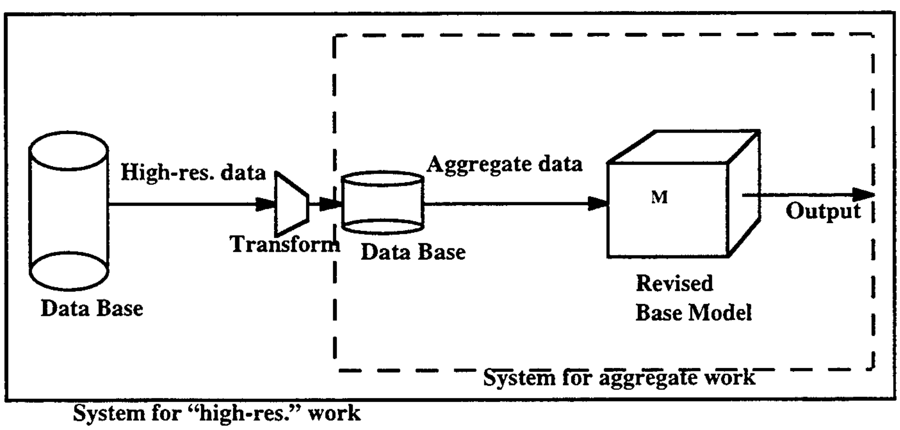

Figure 5.1—Case-A MRM

[46]Case A is one in which rigorous *parameter morphism* exists, to use the terminology of Zeigler, who is examining the underlying mathematical foundations of MRM. See, e.g., NRC (1997, Appendix E) and Zeigler, Praehofer and Kim (1998).

matters is the number of kills per day, which is the product of the three variables. It matters not to the mathematics whether one doubles kills per sortie, sorties per day, or aircraft; the kills per day is the same. To be sure, in some models these variables might appear independently.[47] Our point is that models frequently depend only on combinations of input variables and this can be exploited.

Sometimes this potential simplification of underlying model structure is not evident because of sheer complexity as the model grows to include various types of attack platforms (e.g., fixed-wing and helicopter aircraft, ground-based missiles, and sea-based missiles). Eventually, there are many equations and a great deal of input data. However, what appears complex is in this case still simple cognitively, if merely one uses vector/matrix notation (or, what is perhaps more convenient for computer work, array notation). Then the initial model can be compactly written in terms of straightforward array operations such as scalar products (i.e., "dot products") of vectors. The related algebraic and numerical calculations may be complicated and tedious, but the "chunks" are relatively few and easy to understand—just as it is easy to understand that F=ma generalizes to $\underline{F}=m\underline{a}$ where underlines denote vectors.

Case B: MRM with Approximate but Meaningful Aggregate Models

Basic Concept

Let us now turn to Case B. In this case, aggregation *does* introduce approximation, even if we take pains to make the aggregation "smart." The basic idea in this case is to express the model in terms of a meaningful abstraction (e.g., a more aggregate model) and, often, a correction factor that may not be fully "understandable," but that can be tolerated as a vestigial influence of higher-resolution complications. This practice is common in chemistry, physics, and engineering. For example, in dealing with the behavior of gases, scientists use equations predicting various thermodynamic quantities as the values they would have for infinitely dilute gases plus correction terms that are small but not entirely ignorable. Typically, the coefficients of correction factors are estimated or measured, and recorded, for a variety of type situations. In military simulation, we would expect correction factors to be different when the character of battle changes.[48] Figure 5.2 illustrates Case B schematically. The notation $M = M_0\,(1+M_1)$ is intended

---

[47]For example, aircraft losses as of a certain time might depend on sorties flown and not on kills per sortie.

[48]Although many trees have been sacrificed to support papers discussing the intricacies of Lanchester equations of attrition, the most important modeling issues relate instead to recognizing the many qualitatively different circumstances of battle and the need to use different equations (not merely different parameter values) for many of those different situations. This

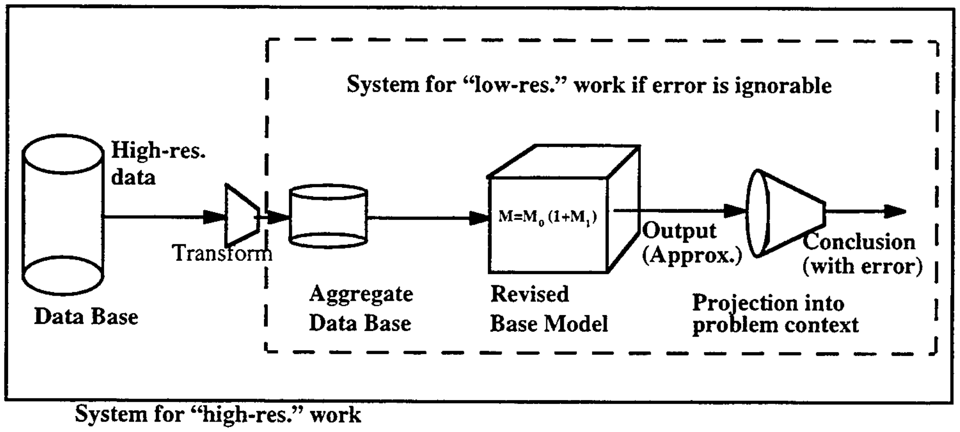

Figure 5.2—Case-B MRM

to suggest that the revised model is expressed as an attractive aggregate model plus a correction term.

Measuring Error in Context of Application

As noted in Section 2, the validity of an approximate aggregation depends not only on the domain and range of its input and output variables, but also on how the outputs will be used. That is, "validity" can be judged only in the context of the problem or application.

One way to think about this issue is in terms of a "projection operator" $\mathcal{P}$ and the actual result sought from an analysis. It is useful here to think in concrete terms. For example, the context in which an approximation may ultimately best be judged may be whether it changes a particular number or a particular curve on a summary viewgraph presented to decisionmakers. We can represent a simulation as an operator $\mathcal{G}$ that generates states $S(t;t_0)$ based on initial states $S(t_0)$. If the final state of interest is an aggregation of the final high-resolution state, then the projection operator is what maps the final aggregate state into those curves or numbers:

$$\text{Result at issue} = \mathcal{P}\,\mathcal{A}gg\,\mathcal{G}\,S(t_0) \qquad (5.1)$$

This is abstract, but in practice the projection operator $\mathcal{P}$ may be something familiar and mundane such as[49]

- selecting the limiting behavior for long times

was an important element of RSAS development in the 1980s. For a good discussion of issues, including distinctions seldom made in current models, see Allen (1995).

[49]The projection operators we describe can be related to the "transducer part" of the "experimental frame" emphasized by Zeigler (see Zeigler et al., 1998, and earlier work).

- averaging over time for the salient portions of the simulation
- time-smoothing by filtering out high-frequency components
- averaging final state over battalion-level battles within a corps sector
- averaging final state over units
- averaging over many possible variants of the "scenario."

These all reflect a defining feature of a linear projection operator, notably that

$$\mathcal{P}^2 = \mathcal{P} \qquad (5.2)$$

Now consider our simplified, aggregated model, represented by G_o. We then have

$$\text{Result at issue} = \mathcal{P}\mathcal{G}_o\mathcal{A}gg\,S(t_o) + \mathcal{P}(\mathcal{G}-\mathcal{G}_o)\mathcal{A}gg\,S(t_o). \qquad (5.3)$$

The question is whether the second term is ignorably small. This is quite different from asking whether the aggregate simulation is valid in general, or whether the final state generated by the simulation is numerically accurate in general.

Even though this is a purely formal treatment, it allows us to make the following observation:

- Often, the legitimacy of an approximation aggregation can be assessed by "understanding" its implications to the result at issue—without any requirement for exact solution or numerical computation of the error term.

For example, if an aggregation has the effect of doing violence to phenomenology only for short times t, then it will still be valid in context if the result at issue is concerned only about long-time behavior *and if* the errors introduced at short times do not propagate. Judgments on such matters can often be made without detailed calculations or complex mathematics.[50]

This point is not new, but it seems often not to be appreciated. A similar point was made recently by Iwaskai and Simon (1994, p. 170):

> The basic idea behind aggregation is this: if variables in a large dynamic system can be partitioned into subsets such that variables in each subset are more strongly connected to each other than to variables in other subsets, one can describe the

---

[50]This technique has long proven valuable in statistical mechanics (called statistical physics in Europe), where one of the principal challenges of theory is to deduce or motivate macroscopic equations (aggregate theories) in terms of first principles such as the Schrödinger equation. For many years workers attempted to use perturbation theory, which often led to ill-behaved infinite series that were dubious and difficult to interpret. A variety of methods were introduced to avoid such flawed approaches. One was a renormalization approach analogous to those developed earlier in quantum theory. Another was the projection-operator approach introduced and exploited by Robert Zwanzig , H. Mori, Irwin Oppenheim, and their students (e.g., Zwanzig, 1961; Mori, 1965; Davis and Oppenheim, 1971). Pozhar (1994) reviews much of this work. We believe that such methods could be applied to the MRM problem in combat modeling, but doing so would require significant work. Another class of theory that could be brought to bear here involves "multiple-time-scale" methods, on which there is a considerable literature.

> short-run behavior of each subsystem independently of the other subsystems. Furthermore, one can describe the long-run behavior of the entire system in terms of variables describing these subsets instead of individual variables, treating each subset as a black box.

The subtle feature in work of this type is being able to assess whether the various averages involved in the approximate model eliminate essential information at the aggregate level. This can easily happen. For example, if one averages over a number of mini-battles so as to generate deterministic outcomes, then probabilistic information is being discarded. If there are correlations from one time period to the next in the real world, then some of the "tail information" that has been discarded may be quite important, and even dominant. So it is that common theater-level models are fundamentally flawed unless it is reasonable to assume efficiency of command-control processes that have the effect of eliminating the correlations (e.g., allocating reserves to plug holes). Thus, the corollary to the above comment is

- Understanding the consequences of various averaging processes requires a deep understanding of the underlying phenomenology, including issues involving probabilistic dependencies (correlations).[51]

With this background, let us now return to the test problem and work it through in considerable detail. First we shall introduce a number of simplifications that are actually valid in the Case-A sense. We shall then introduce and discuss a number of approximate aggregations that illustrate more general principles.

Case C: MRM with Use of Low-Resolution "Repro Models"

Let us now turn to Case C. If the high-resolution and aggregate representations are sufficiently different structurally, cross-calibrating will probably depend on "curve fitting." That is, the low-resolution variables (or probability distributions) will generally have to be developed by running cases with the detailed model and analyzing the data thus generated using statistical and other data analysis tools. Figure 5.3 illustrates Case C schematically. The two models coexist, and there is some understanding about how to relate them as indicated by the transform function connecting the two data bases. Often, however, this is a very limited and imperfect connection.

With this background, let us now work through a problem in some detail. Section 6 describes design experiments pertaining to Cases A and B, and provides detailed motivation for many of

[51]This issue is discussed in NRC (1997, Appendix J). It is discussed under the rubric of "configural effects" in Horrigan (1992). It is discussed in the special context of correlated signals in Van Trees (1968).

the points made in Sections 2–5. We have done a number of computational experiments related to Case C, but will report those elsewhere.

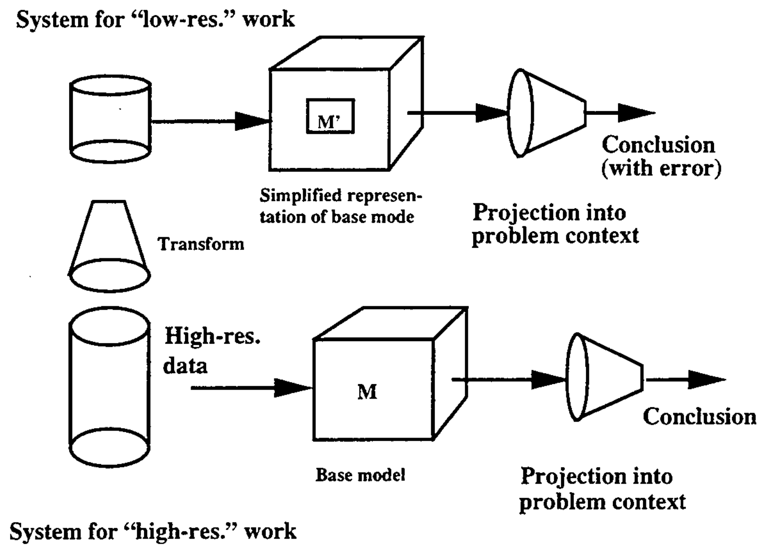

Figure 5.3—Case-C MRM

6. An Illustrative Problem: Halting an Invading Army with Precision Fires

In this section we work through a particular problem in considerable detail, abstracting the model in stages and noting the nature of the issues as we do so. For reasons of clarity, however, even our starting point is a good deal simpler than the full military problem. The Appendix discusses some of the more general issues.

Problem Statement

Consider the following problem (Davis and Carrillo, 1997): an invading mechanized army rolls into a country friendly to the United States and maintains a high rate of speed because it faces only weak opposition on the ground. The United States deploys forces to defeat or help defeat this invasion, but problems such as limited warning time and the absence of propositioned ground-force equipment mean that if the army is to be halted before it reaches its objective, the burden of effort may be on forward-deployed or rapidly deployable long-range precision fires. These munitions could be delivered by long-range bombers, Air Force or Navy fighter aircraft, Army attack helicopters, or ATACMS-like missiles launched from Army batteries or from surface ships. Let us assume that the terrain is open, that the United States enjoys good C^4ISR, and that the losses of U.S. forces are ignorably small during the period of interest so long as aircraft do not begin attacking ground forces in earnest until after air defenses are suppressed sufficiently (until after SEAD). Let us further assume that the invader can be halted only by destroying a substantial fraction of his armored vehicles. The model then reduces to a race: the invader is rolling toward his objective but taking attrition. Will the precision fires be able to halt the invader by causing the requisite percentage of attrition before the attacker reaches that objective? How soon could the invader be halted? How sensitively would this depend on forward-deployed forces, warning time, deployment rates, the effectiveness of the weapons, the duration of the SEAD period, and so on?

Clearly, this model is an abstraction of reality—one intended for use in thinking about "How much is enough?" and developing broad insights about matters such as the feasibility of halting an invasion with precision fires alone (the answer from analysis is that such a halt is feasible but very difficult). A real campaign would be much more complex (Appendix A). However, analysis at this nontrivial level of detail (Figure 6.1) has proven useful in a number of studies in the past

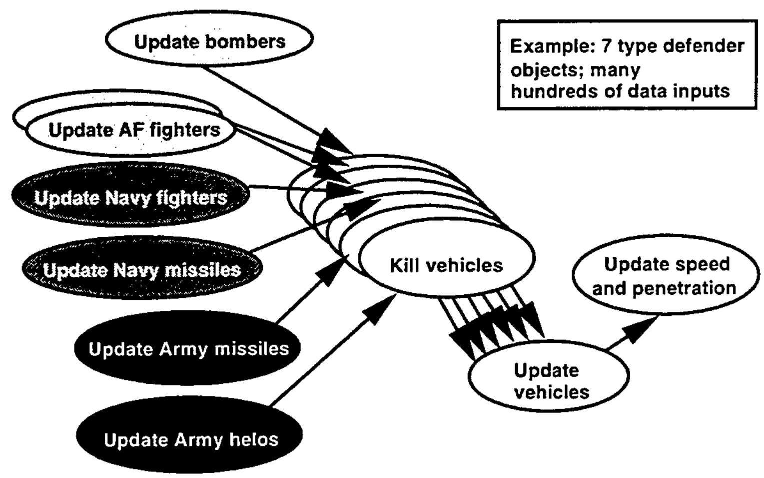

Figure 6.1—Data-Flow Diagram of Halt Problem Focused on Precision Fires

half-dozen years (Ochmanek, Harshberger, Thaler, and Kent, 1998). Indeed, this level of analysis can in some respects be more informative and insightful than studies using sophisticated theater models. It is especially informative, however, if the model is designed with multilevel resolution allowing analysts to conduct *exploratory analysis* on a relatively small number of variables (Davis and Carrillo, 1997). So, let us now see what is required to bring that about and how the process of building such a model may differ from normal practices in simulation.

Abstraction Through the Use of Vectors and Higher-Order Arrays

Cognitive Value of Chunking

Our first abstraction[52] does not reduce information, but it has extraordinary implications for cognition through "chunking." The idea is simple: treat the various types of anti-armor aircraft and missiles as components of a vector called "shooters." Attrition can then be understood in constructs such as number of shooters, times shots (or sorties) per day, times kills per shot (or sortie). This chunks the concepts efficiently, even though, in detail, a calculation would involve

[52]Some generic abstraction methods are described in Zeigler (1984), with some of that discussion reviewed in Caughlin and Sisti (1997). The generic methods we consider are dropping or simplifying components (e.g., entities, entity attributes, processes, or particular interactions); replacing a deterministic model with a stochastic representation (or vice versa); coarsening the ranges of descriptive variables, as in temporal or spatial smoothing; grouping components and aggregating; reducing the precision of calculations; and using mathematical abstractions such as vectors and arrays. This differs modestly from the other discussions.

summing over the types of shooters. Table 6.1 shows the results compactly using underlines to indicate vectors or matrices.

The model is cognitively simple, as indicated by Figure 6.2. At every time step, the model updates shooters to allow for deployments and allocation of shooters to anti-armor missions.[53] It then calculates kills of armored vehicles Vkills(t) and Vkillscum(t), which depend on number of shooters, sortie or shot rates, kills per sortie or shot and so on (as well as how much of the invading army remains). Finally, it calculates movement V(t) for the time period: 0 if the enemy

Table 6.1

Baseline Model

Variable	Meaning
Inputs	
$\underline{N}_O$	Initial number of shooters in theater
$\underline{R}(t)$	Deployment rate (shooters/day)
$\underline{W}$	Actionable warning time (days) before D-Day
$\underline{F(t)}$	Fraction of shooters allocated to anti-armor missions
$\underline{K(t)}$	Kills per shooter-sortie after SEAD
$\underline{S(t)}$	Sorties per shooter-day after SEAD
$\underline{\varphi}(t)$	Multiplier of $\underline{S}$ for times less than $\underline{T}_S$ ($\underline{\varphi}(t)$=1 for t≥T_S)
$\underline{T}_S$	Time required for SEAD (days)
n	Number of attacking divisions
VPD	Number of armored vehicles per division
H	Fractional attrition required to halt invader
Obj	Distance (km) to invader's objective
V	Speed of invader until halted (km/day)
Outputs	
Vkills(t)	Daily kills of armored vehicles
Vkillscum(t)	Cumulative kills
VH(t)	Remaining armored vehicles
V(t)	Movement rate (km/day)
X(t)	Penetration (km)

[53]Appendix A shows a broader version of the model that includes an explicit process for command functions. Leaving out the command process is itself an important abstraction in this model.

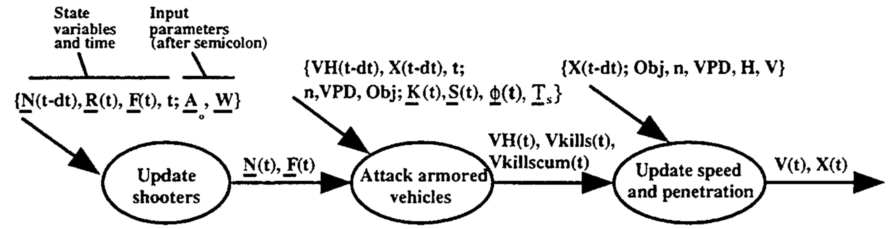

Figure 6.2—Data Flow in Array Version of Base Model

has reached his objective or been halted by attrition, and V otherwise.[54] X(t) is the penetration distance as of time t. The data requirements and degrees of freedom are still very large, however (hundreds or thousands of data elements).

Potential Conflicts with Object-Oriented Programming

Significantly, many and probably most simulations do not fully avail themselves of this type of simplifying formalism, perhaps because it requires greater mathematical and programming sophistication, and in some cases higher-level languages, but also because of the reluctance to treat "apples and oranges" as similar. Suppose, for example, that we are using object-oriented programming to implement our halt-phase model. Although the theory presented here suggests that various types of aircraft and missiles would be special objects within the more general object "shooter," this tactic would organize the model in such a way as to mix what many would see as very different kinds of things. An Army officer may not want to have his MLRS/ATACMs unit[55] represented as though it were another type of aircraft. Thus, in normal object-oriented programming as it is practiced within the DoD, the "shooter" object would probably not appear. In any case, *we believe that normal object-oriented programming may actually be a hindrance to MRM because such normal practice—in which objects are identified with physical named systems, not abstractions—fail to exploit simplifying abstractions.*

[54]The assumption of constant speed V is unrealistic. It is likely that ground forces would move faster or slower during a given time period depending on terrain, recent attrition, the status of the logistical tail, and so on. Units might also do a long-distance dash on the first day, depending on extra fuel canisters rather than refueling from fuel trucks. However, it is extremely unclear how movement rate would depend on many of these factors, particularly attrition and disruption. Rather than write an equation pretending to express knowledge that we do not have, we prefer in this particular report to use a constant movement rate, which we then vary parametrically. In other settings, of course, we treat movement rate as an assumed function of the various factors.

[55]MLRS/ATACMS is a multiple-launcher rocket system that can fire the advanced tactical missile system. The ATACMS missiles have advanced warheads such as the brilliant anti-tank munition (BAT), which have high lethality against armored targets over a substantial area.

This said, it is also possible that modeling and programming to exploit the concept of the shooter object would cause representational difficulties and annoyances for some users. Suppose, for example, that the model had "explanation capabilities" describing the reasons for various battle outcomes. If those explanations were couched in terms like "number of shooters," that might (or might not) be a distraction in applications involving only aircraft. Later in the section we discuss more fully the problem that choosing a model representation usually involves tradeoffs.

Using a Dose of Theory to Motivate MRM Design

Figure 6.1 helps in motivating an MRM design. Suppose we muse a bit about the underlying model structure. One way to do so is to introduce some formal mathematics. We say "formal," because we need not solve the problem or even compute integrals, but merely understand the conceptual structure. Let us further use the trick of thinking about where we want to go. Ultimately, we seek a solution in the form of D, the maximum penetration by the attacker (km). If we write down the equations defining D, we see that it depends on T_h, the time required to halt the invasion if it is halted at all before reaching its objective. To simplify discussion let us suppress vector notation temporarily (or, equivalently, consider the case of a single shooter type). We obtain an integral equation for T_h:

$$\begin{aligned} &D = MIN[Obj, VT_s] \\ &(n)(VPD)(H) = \int_0^{T_h} F(s)N(s)\phi(s)K(s)S(s)ds \end{aligned} \tag{6.1}$$

If we now plug in the equations for N(s) to reflect deployment, we find ourselves strongly motivated to introduce some aggregate variables as follows. Plugging in, we obtain

$$(n)(VPD)(H) = \int_0^{T_h} \left\{ F(s) \left[[N_o + \int_{-W}^{0} R(s')ds'] + \int_0^s R(s')ds' \right] \phi(s)K(s)S(s) \right\} ds. \tag{6.2}$$

This suggests the following definitions:

$$\lambda \equiv N_o + \int_{-W}^{0} R(s')ds'; \quad \delta(s) \equiv \phi(s)K(s)S(s) \quad \xi \equiv (n)(VPD)(H) \tag{6.3}$$

and leads to the following simplified form of the integral equation

$$\xi = \int_0^{T_h} F(s) \left\{ \lambda + \int_0^s R(s')ds' \right\} \delta(s)ds. \tag{6.4}$$

The vector analogs of the aggregate variables have precisely the same form because the problem is linear in the shooters.[56] The actual algebra could be unpleasant, because the integral equation is a sum over shooter types of integrals that depend on the shooter's T_S, K(t), S(t), etc. Further, there are a number of different cases, depending on whether T_h turns out to be greater than any of or all of the components of T_S. However, we need not bother with such matters here. We already know enough to be able to construct Figure 6.3.

Now we see that the problem solution could be computed by starting with the full set of problem inputs (the lowest items on the tree: V, $\underline{N}_o$, $\underline{W}$, $\underline{R}$(t), . . .). However, we could instead specify values of variables at a next-higher level. That is, we could specify values of $\underline{\lambda}$, $\underline{\delta}$(t), and ξ rather than the variables on which they depend. That is, we have substantially reduced the problem's dimensionality, which is essential for exploratory analysis. [Note: $\underline{T}_S$ may also be an input to F(t) and ϕ(t). It is also a top-level variable, as shown, because later $\underline{\delta}$(t) and $\underline{F}$(t) will be expressed as functions of $\underline{T}_S$ rather than data.]

While Figure 6.3 relates to the "solution," a closely analogous figure applies at any time during the simulation. That is, we are now motivated to conceive the simulation's data-flow diagram in a similar tree-like structure, as shown in Figure 6.4. This is substantively equivalent to the earlier data-flow diagram, but now with an MRM character. Furthermore, we have in mind here that

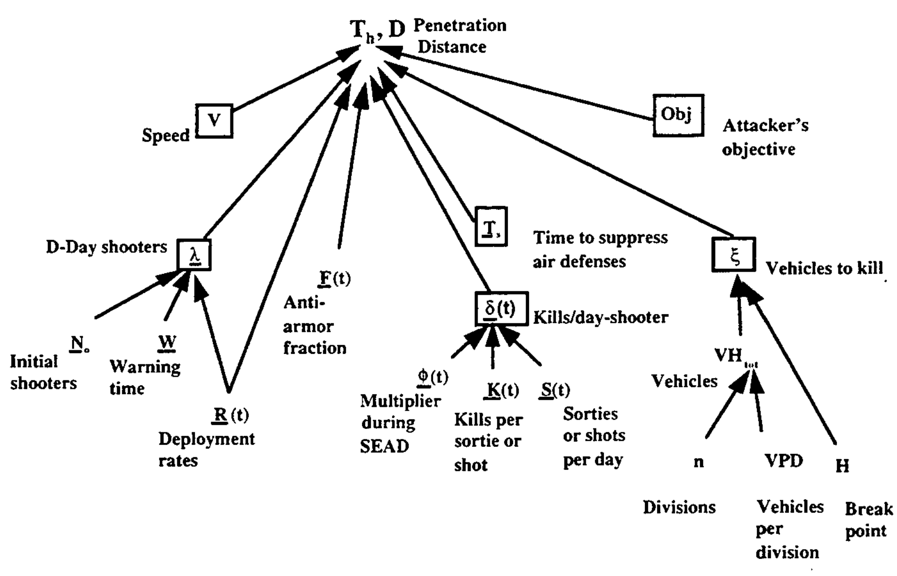

Figure 6.3—A First MRM Design

[56]Actually, the mathematical structure applies to a wide range of linear problems in which there are a number of "doers" contributing to completion of a task.

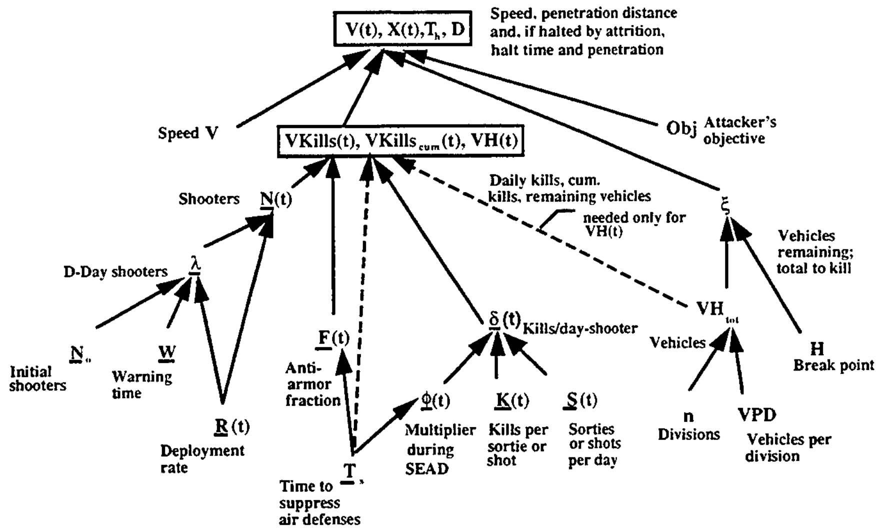

Figure 6.4—MRM Depiction of Simulation Problem

where a higher-level variable really depends only on an aggregate variable rather than lower-level variables, the algorithm in the implementing computer program will indeed be written in terms of the aggregate variable. This means that a user can equally well specify input at the lowest level, or at any of the higher levels of the tree. This is *not* standard practice, nor even a good practice if one is uninterested in MRM. It is only because we seek MRM features that it is desirable. Why? Because model generalizations may undo some of the neat aggregations. A wise programmer not interested in MRM would leave room for such generalizations rather than exploit "tricks" that might not always be valid. For example, if the original problem statement said to build a model of movement, then there would be no need for VH_{tot}: all dependence of V(t) and X(t) on VH_{tot} would be through ξ. However, if the model were generalized to cover killing vehicles after the halt, then VH_{tot} would be needed to avoid killing more vehicles than exist. In any case, *all of this shows how thinking with formal models—even if too difficult to solve—can help structure a simulation model for MRM.*

Identifying Aggregations

How does one find the *aggregation fragments* (e.g., the lumps such as $\underline{K}\ \underline{S}$) that simplify the problem rigorously as in Figure 6.4? There is no general answer, but we shall discuss some directions briefly:

- Watch for them during the design process, and subsequently.
- Consciously review variables in a search for aggregation fragments.
- Use idealized versions of the problem to suggest hypotheses about aggregation fragments.

- Conduct computational experiments to identify additional hypotheses.
- Where possible, work in computer environments facilitating these measures.

To elaborate, one way to find aggregation fragments is simply to *watch* for them, recognizing that aggregation fragments (such as $\underline{\delta}$ in the example) can be quite useful even if they are only part of what would comprise an overall aggregate model. This procedure is perhaps easiest in the design phase—assuming design is taken seriously and written down.

Another way is to review the variables with a view toward finding useful lumpings. In some instances the validity of the lumping may not be certain without looking into details of code, but good guesses can at least point the way.

Yet another approach is to consider extreme cases in which analytical solutions or steady-state conditions can be identified. In this case, variable lumping may be obvious, and it may (or may not) hold up in the actual problem.[57]

Finally, we mention computational experiments, which can reveal (i.e, help us "discover") aggregation fragments that are not at all obvious initially. We shall report separately on some experiments in this vein.

With some computer environments it is possible to do better. For example, the desktop environment of Microsoft Excel includes a tool for tracing inputs and outputs for specified cells. Also, visual programming environments such as iThink and Analytica are quite powerful in this regard, as well as for encouraging good design-and-documentation discipline.

Using Formal Theory to Introduce Equivalent-Shooter Concept

So far, we have merely exploited various "parameter morphisms" to restructure the problem conveniently (Case A). Let us now consider some approximations. Although vectors and higher-order arrays are elegant, analysts and their clients often prefer to think in simpler concepts such as "equivalent divisions," "equivalent F-15s," or "equivalent battle groups."[58] What we would like to do is reduce problem dimensionality by turning vectors into scalars and conceiving the

---

[57]There is a relationship here to the observation made by Herbert Simon some time ago to the effect that operations research has come to depend too heavily on simulation, which provides only one perspective into the problems (Simon, 1990).

[58]Even in entity-level ground-force modeling it is common to lump tanks of different types, artillery units of different types, and so on. Thus, in a higher-level model, an M-60 tank may be regarded as 0.5 "equivalent M-1/A-1s." In an entity-level model one M-1/A-1 might be inserted for each M-60 specified in the scenario. One of the ironies of the continuing debate about the merits of aggregated models is that all combat models are aggregate models, since none of them begin with the fundamental particles of the universe (whatever those are).

problem as though there were only one type of shooter. Can we do this, and at what price? In this problem, lethality is the measure of effectiveness for shooters. Thus, letting $\delta(t)$ represent the standard shooter's lethality at time t, we can define the number of equivalent shooters at time t as follows:

$$Vkills(t) = \sum_i N_i(t)\delta_i(t)$$

$$N_{eq,i}(t) \equiv N_i(t)\frac{\delta_i(t)}{\delta(t)} \tag{6.5}$$

$$Vkills(t) = \sum_i N_{eq,i}(t) \equiv \delta(t)N_{eq}(t)$$

This last equation has the desired form: a simple product. However, since $N_{eq}(t)$ depends on $\underline{N}(t)$ and $\underline{\delta}(t)$, no generality has been lost and the form's simplicity is a bit misleading. To elaborate, the scalar product $Vkills(t)=\underline{N}(t)\underline{\delta}(t)$ can always be represented as the product of scalars, $N_{eq}(t)\delta(t)$, for suitably defined $N_{eq}(t)$ and for $\delta(t)$ being the kills per day of the standard shooter against which others are compared in computing "equivalent shooters." The only issue is how useful it is to do so.[59]

Since it will make understanding Figure 6.5 and subsequent figures easier, let us note that if any data-flow diagram has a portion such as that of the left side in Figure 6.5, where a scalar C is

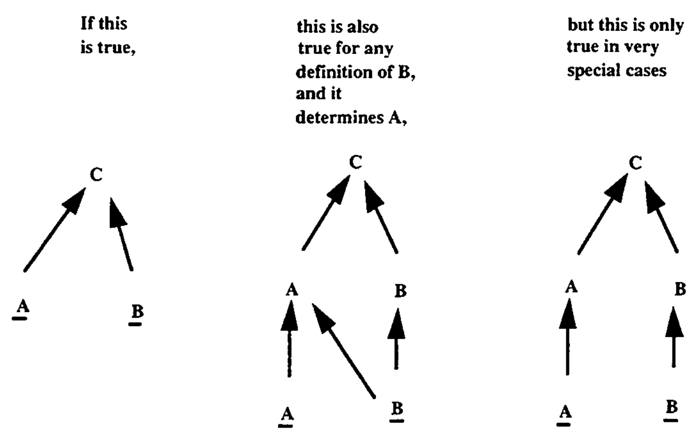

Figure 6.5—Generic Relationships

---

[59]In particular, note that the weights $(\delta_i(t)/\delta(t))$ are functions of time. Each actual shooter may be worth a different number of standard shooters in each time period. If this formal exercise is to be useful, we need to relate those weights not to time but to situation; and the fewer the situations we need to consider (and hence the fewer different weighting functions we need to use), the more this formalism will simplify our model.

determined by two input vectors $\underline{A}$ and $\underline{B}$, then we can always insert a convenient scalar representation of $\underline{B}$, B, and a scalar representation of A $\underline{A}$, such that C is determined strictly by A and B. However, we are not free to define A in any manner we choose: once we define B, then A must be whatever it takes to reproduce C. In general, at least, that will require A to be a function of both $\underline{A}$ and $\underline{B}$ as shown in the middle. Only in special cases will it be true that A depends only on $\underline{A}$. Also, both A and B will in general be functions of time, even if one of the original vectors was a constant.[60]

Still, we could in exploratory analysis specify $N_{eq}(t)$ and $\delta(t)$ directly, thereby greatly reducing dimensionality. The only problem in doing so is that at a deeper level the two variables are correlated (i.e., probabilistically nonindependent). Thus, interpretation of the results of exploratory analysis would have to be cautious.

Equivalent Deployment Rates

So far, we have still not been forced to make approximations above and beyond those inherent in the overall model. The calculation of $N_{eq}(t)$ merely introduces an interesting output and a simpler way to express Vkills.[61]

Going Further Takes Us to an Impasse

A more interesting challenge arises if we attempt to introduce equivalent deployment rates. That exercise might seem to be equally straightforward, because we could use the same equivalencies as in calculating equivalent shooters. However, there is a fundamental problem. If at time t_o we translate $\underline{R}(t_o)$ into a scalar $R_{eq}(t_o)$, we lose information about the mix of shooters being deployed. Because the equivalency factors depend on time-dependent lethalities, there is no single once-and-for-all equivalency factor. Mathematically, we have

$$N_{eq}(t) = \sum_i N_{eq,i}(t) = \sum_i \frac{\delta_i(t)}{\delta(t)} \{\lambda_i + \int_0^t R_i(s)ds\}, \tag{6.6}$$

and we see that $N_{eq}(t)$ is not so simple. If the equivalencies were constant, then everything would be simple, but that would be a very strong assumption. More generally, $N_{eq}(t)$ depends on the history (or memory) of the shooter mix, and not just on any constant equivalent

---

[60]The functional relationships suggested by the arrows in the several diagrams are different.

[61]This is a useful case because, in practice, simulationists may want to have a simple way to characterize the number of shooters, and even to express attrition to the adversary in such terms, while wanting to maintain full information on the number of each type of shooter present.

deployment rate. After all, early in conflict, a stealthy aircraft or a long-range missile would have far more lethality than a nonstealthy factor if air defenses were a problem. All of this means that the concept of "equivalent shooter" is fuzzy. For this reason among others, a simulation modeler uninterested in MRM would probably avoid the concept if he could.

Averting Impasse with an Approximate Two-Phase Abstraction

We *are* interested in MRM, but what do we do at this impasse? What follows is an example of Case-B aggregation in the sense of Section 3. We will be introducing a readily understandable approximation so that the model takes on an attractively simple structure.

Let us now *assume* that the equivalency factors are piece-wise constant in time and that all the air defenses are suppressed at the same rate. In that case, $\underline{T}_S$ is just a constant T_S and each $\underline{\delta}_i$ is a constant δ_i^0 during SEAD and a constant δ_i^1 thereafter. This is one example of a *cookie-cutter approximation*, of a sort familiar to all analysts. It greatly simplifies the history and memory issues. We shall use the same standard value for δ at all times t. As a kind of trick, we can now define two new functions and express $N_{eq}(t)$ as a function of those:

$$\begin{aligned} &N^0_{eq,i}(t) \equiv \frac{\delta_i^0}{\delta}\{\lambda_i + \int_0^t R_i(s)ds\} \quad N^1_{eq,i} \equiv \frac{\delta_i^1}{\delta}\{\lambda_i + \int_0^t R_i(s)ds\} \\ &R^0_{eq}(t) \equiv \sum_i \frac{\delta_i^0}{\delta} R_i(t); \;\; R^1_{eq}(t) \equiv \sum_i \frac{\delta_i^1}{\delta} R_i(t); \\ &\lambda^0_{eq} \equiv \sum_i \frac{\delta_i^0}{\delta}\lambda_i \quad \lambda^1_{eq} \equiv \sum_i \frac{\delta_i^1}{\delta}\lambda_i \end{aligned} \tag{6.7}$$

We now have the relationship in Eq. (6.8), which may seem a bit peculiar, because $N_{eq}(t)$ changes discontinuously at $t=T_S$, as though equivalent shooters somehow "materialize," but it is actually straightforward to understand: it merely says that in calculating equivalent shooters, we change from one curve to the other when SEAD is complete because the equivalency factors are different thereafter. Figure 6.6 illustrates the effect for a particular set of assumptions about deployment rates and the assumption that T_S is 4. The lower curve shows the number of equivalent shooters versus time if one uses the equivalency rates applicable during SEAD; the upper curve uses the rates applicable after SEAD. $N_{eq}(t)$ is taken as the lower curve until SEAD is complete, and then the upper curve.[62]

[62]If deployment rates were functions of time, and especially if there were upper limits to the deployment of some systems, the analytical formalism used here would require defining additional time regions and curves and having a somewhat more complex rule for which curve to

$$
\begin{aligned}
N_{eq}(t) &= \lambda_{eq}^{0} + \int_{0}^{t} R_{eq}^{0}(s)ds \quad \text{if } t \le T_s \\
N_{eq}(t) &= \lambda_{eq}^{1} + \int_{0}^{t} R_{eq}^{1}(s)ds \quad \text{if } t > T_s
\end{aligned}
\tag{6.8}
$$

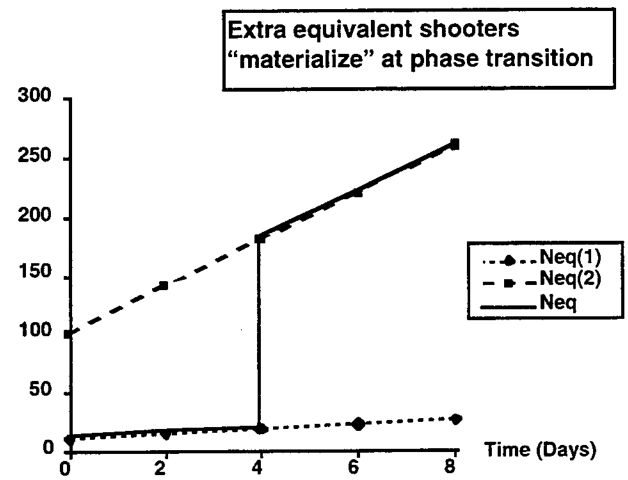

Figure 6.6—Equivalent-Shooter Calculation

To implement the equivalent-shooter formalism rigorously in a model, then, we must calculate both equivalent-shooter formulas at each time step, and shift from one to the other at the appropriate time.[63]

The significance for MRM theory is that the example illustrates something more general:

- A seemingly fuzzy and irrational concept (e.g., equivalent shooters) needed for MRM can often be given a clear meaning if merely we find the right simplifying assumption, which may be reasonably accurate *(in this case, that the SEAD period can be represented by a single number).* Such assumptions, however, might not come naturally in bottom-up entity-oriented simulation.

With this approximation, abstraction goes smoothly with the result shown in Figure 6.7. Note the multiple levels at which analysts could enter inputs, depending on their purposes and the information available. As an example, if there were six shooter types, and if the deployment rates and so on were constant, the abstractions could reduce dimensionality of inputs from 48 to between 5 and 8 depending on taste. The reduction would be from many hundred to tens if

---

use when in computing $N_{eq}(t)$. A better alternative in such a case would probably be to approximate the deployment curves with two-parameter functions of time.

[63]Introducing equivalent shooters is not always worthwhile. If there are N different situations or phases requiring N different weighting schemes to define equivalent shooters, then there must be more than N shooter types in the original model for the approach to reduce dimensionality or computational requirements.

many of the inputs were time-dependent. Note further that if we wanted to calibrate input values at an aggregate level based on higher-resolution simulation, the procedure for doing so would be rather obvious from the tree structure. Details would be complex because calibration would involve averaging over an ensemble of high-resolution simulations with appropriate probability distributions for the problem. Also, a higher-level variable might need to be represented by a probability distribution rather than by a single number, because its expected value might be a poor representation of the underlying phenomena. Still, how to proceed is fairly clear because of the tree relationship. See Lee and Fishwick (1997) for discussion of relevant software tools for achieving behavioral consistency.

The Problem of Multiple Representations

Figure 6.7 also demonstrates one of the most confusing issues of MRM. Although we would like to have neatly perfect hierarchical trees, we do not. Instead, a number of the branches interact. This means that the variables are in some sense correlated. Further, the tree may or may not be the "right one" for our purposes. There are many possible perspectives on, or representations of, the same problem—even if we hold assumptions constant and agree on what levels of resolution to consider generally. This reflects what Zeigler (1984) calls the *multifaceted* nature of models.

Correlations lower in a tree are not particularly serious when doing exploratory analysis higher on the tree. Consider $F_{eq}(t)$ and $N_{eq}(t)$. We could do exploratory analysis treating these as

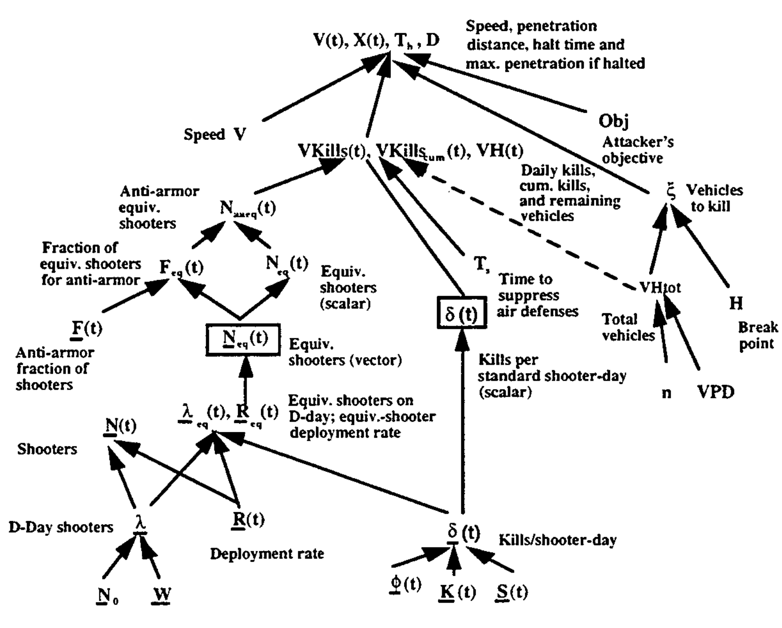

Figure 6.7—An MRM Design with Equivalent Shooters

independent inputs. The only problem would be that some portions of the input domain and outcome range would be more realistic than others (requiring bounds on the *experimental frame*). In practice, that could often be indicated with some shading in the final display, without accounting in detail for such correlations. On the other hand, if we were preparing a serious data base of values at this level of input, then the calibration experiments would need to account for the correlations: F_{eq} and N_{eq} could not be calibrated independently. This issue would simplify, of course, if the fractions in each category were constant over time, which would be true in some cases but not others.

The representation issue is more subtle and, we believe, a major obstacle to MRM. Suppose that the analyst wanted to have an equivalent deployment rate and an equivalent number of D-Day shooters as input variables—because that would make the problem relate directly to the understandable one-shooter problem. These combinations could certainly be done and Figure 6.8 shows such a design. It is no better or worse than the one above, but it uses different aggregate variables. The difficulty is that many combinations (i.e., representations or perspectives) are possible. If we try to draw the trees showing all of them, or if we try to write algorithms using all of them, the result becomes exceedingly complex. Further, we cannot have all representations all of the time. There must be choices. Our conclusion here is perhaps now obvious but we believe that it is an important guideline for MRM:

- Because MRM designs should provide a range of perspectives, and because doing so can be exceedingly complex, these perspectives should be modularized and a user-mode interface provided. These alternative perspectives would amount to alternative and equivalent submodels, each with their own data-flow diagrams, algorithms, and displays. Not all combinations can reasonably be provided in nontrivial problems, although adding an additional combination should not be difficult.

Assessing the Approximation

To obtain the MRM design of Figures 6.7 and 6.8 we used a cookie-cutter two-phase approximation. Although we obviously cannot assess its validity in general terms, intuition suggests that the approximation may often be a good one. The SEAD campaign might only be a matter of some days, and the difference between T_S across the various shooters may be small also (e.g., the range might be 2 to 5 days for a halt campaign lasting 10 days).

Figure 6.9 illustrates the issue for an imagined case in which Air Force and Navy aircraft are available on D-Day and continue to deploy subsequently. The particular illustration assumes that the Air Force aircraft are fully effective after two days, while the Navy aircraft are fully effective after six days. That is, because the Air Force fighters are more stealthy, they are able to

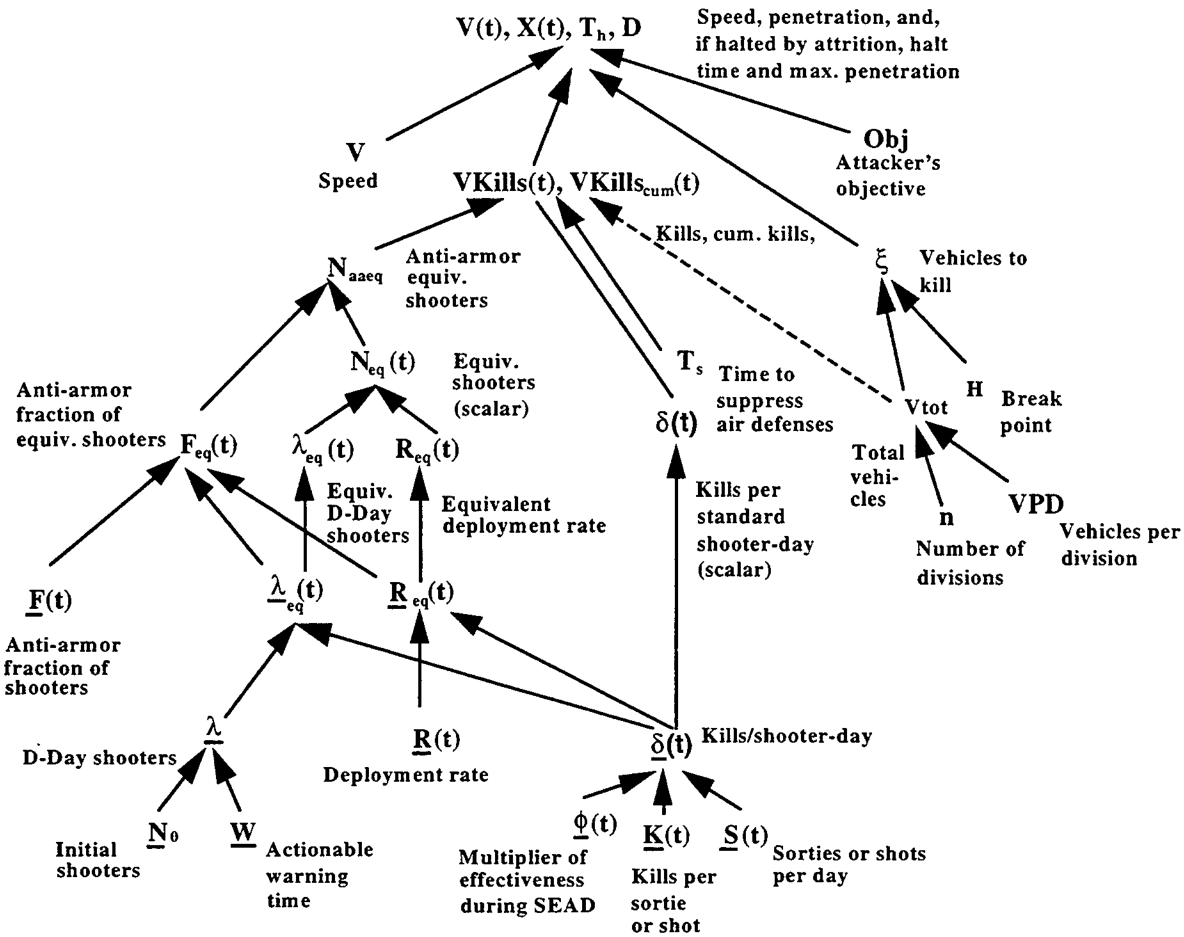

Figure 6.8—Alternative MRM Design with Equivalent Shooters and Deployment Rates

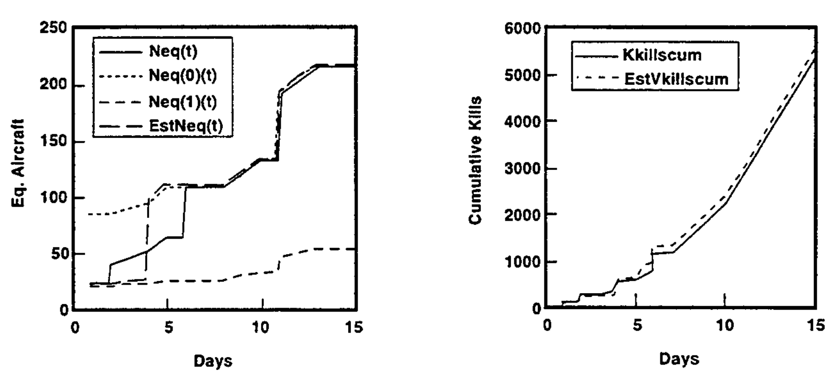

Figure 6.9—Example: Equivalent Aircraft Versus Time

operate sooner in the SEAD period. After the SEAD period, the aircraft are assumed equally capable in kills per day. While Air Force aircraft deploy at a reasonably constant although ragged rate, the Navy aircraft arrive in clumps with their aircraft carriers. As a result, the correct number of equivalent aircraft is given by the dark solid curve in the left chart. That curve shifts discontinuously when SEAD is complete (equivalent shooters "materialize"), because suddenly a number of aircraft that were previously ineffective become fully effective. It shifts abruptly again when the second carrier arrives. The dashed dark curve represents the results of a cookie-cutter approximation in which T_S is assumed to be four days. We see that there is an error between days 4 and 6. However, the effect on cumulative kills (right chart) is almost impossible to see. Thus, if our application-specific projection operator generated the cumulative kills versus time, the cookie-cutter approximation would be very good indeed.

At the same time, note that if we had discarded information on the phases, in this case the intra- and post-SEAD periods, no single constant would have sufficed to describe the buildup of equivalent aircraft. Further, because carriers arrive in chunks, a constant deployment rate would not be a good approximation in many cases. The approximation used in Figure 9, however, is rather good.

Time Smoothing and Estimation Theory

As our last example of MRM-related simplification using this model, let us now consider ways to drastically reduce the number of degrees of freedom by approximating the time dependence of variables such as $R_{eq}(t)$. As noted earlier, normal simulation modeling would probably define $\underline{R}(t)$ with a two-dimensional array, a given element of which would give the deployment rate on that day for a given shooter—$\underline{K}(t)$ and other variables would be similarly defined. This technique grossly inflates the data requirements. We have demonstrated above how equivalent-shooter methods can be introduced conveniently if it is reasonable to make the cookie-cutter assumption that there are two distinct phases: post-SEAD and intra-SEAD. The time dependence of the variables could then amount to checking which phase applies. Suppose, however, one wants to go further and replace the deployment rates, sortie rates, kills per sortie or kills per shot, and so on, by scalar constants (averages) rather than day-by-day numbers.[64] How would we go about that? There are many poor ways to do so, but "dumb aggregations" have given abstraction a bad name in many circles. Can we do better? We think so.

The main point here is familiar, but easy to forget when programming:

[64]Another alternative would be to approximate the deployment rates with simple functions of time. This approach would be preferable if, for example, the deployment of some systems was complete after a given time.

- When constants are substituted for functions, some care should be given to use "good estimators" rather than simple averages or standard values applicable in only some circumstances.

To illustrate, consider first a single shooter with sortie rate S and K kills per sortie after SEAD. During the SEAD phase the effectiveness K(t)S(t) is multiplied by ϕ. What approximations might we make for $\delta(t) = \phi(t)K(t)S(t)$? To make the example more interesting, assume an application (e.g., a war game) in which the user knows $\delta(t) = \phi(t)K(t)S(t)$ except for not knowing T_s and T_h. Assuming that K and S are probabilistically independent for simplicity, there are three obvious estimators for the user if he wishes to simplify the model. The first is to ignore the SEAD phase and just use estimated average values for K and S. The second is to use a weighted sum based on estimates of T_s/T_h and estimated averages of K and S during the two separate periods. The third is to reject the notion of an overall estimator and instead preserve the phases, using separate estimators. Suppressing subscripts, we have:

$$
\begin{aligned}
&1.\ \delta(t) = \delta = \overline{KS} \\
&2.\ \delta(t) = Est(T_s / T_h)\varphi\overline{K}^0\overline{S}^0 + [1 - Est(T_s / T_h)]\overline{K}^1\overline{S}^1 \\
&3.\ \delta(t) = \varphi\overline{K}^0\overline{S}^0 \qquad t \le Est\_T_s \\
&\quad\ \delta(t) = \overline{K}^1\overline{S}^1 \qquad t > Est\_T_s
\end{aligned}
\tag{6.9}
$$

Figure 6.10 illustrates the consequences graphically for a simple case in which two types of aircraft are deploying. The standard aircraft is more capable during the SEAD period (a ϕ of 0.5 rather than 0.1) but carries four rather than six weapons. There are ten of each aircraft initially. The standard aircraft deploys at a slower but constant rate (three per day rather than five per day). The actual time to complete the campaign is ten days, with a SEAD period of four days, but the estimates are based on a SEAD period of three days and a total period of seven days. We see that the first approximation is poor, whether measured by the dynamics or the final cumulative kills as of day 10. The second approximation is much better in predicting final cumulative kills but misses interesting early dynamics. The third approximation, although still based on an estimate of T_s different from the one assumed in the simulation, is pretty good on dynamics and is qualitatively correct; it is quite good for cumulative kills. How important the distinctions are would depend, of course, on details of the application.

One more general conclusion illustrated by the example is that even modest effort to find "good estimators" can materially improve the quality of abstraction or aggregation.[65] The other conclusion is that we should be cautious about eliminating distinctions between natural phases in

[65]Good texts are available on estimation theory. See, e.g., Van Trees (1968).

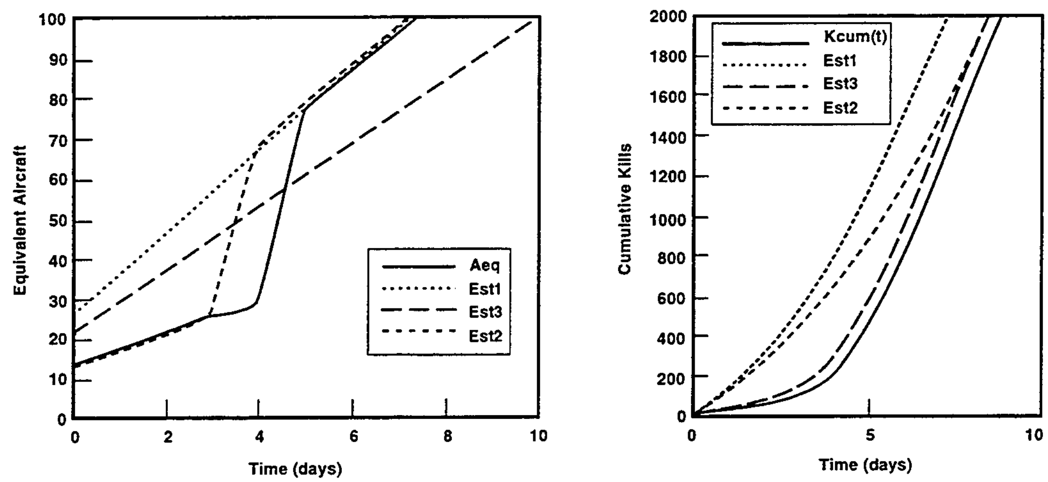

Figure 6.10—Comparison of Alternative Approximations

a problem. Doing so may or may not affect final states substantially, but it will discard information about even time-smoothed dynamics. Further, the resulting aggregate model may be quite misleading for sensitivity analysis. That is, while it may be natural to imagine that the sensitivity of an aggregate model to aggregate variables would be good, that is often not the case. Preserving phase information can improve results significantly.[66] Although this practice is second nature to simulationists, it is not necessarily so obvious to more reductionist modelers or to individuals more interested in statistical analysis than phenomenology.

This is far as we shall take the example in this report, but it should be clear that in cases like this, a great deal of model abstraction can be accomplished within an MRM design.

"Abstractions of Convenience"

As our last discussion of abstraction, we want to mention an MRM-relevant technique that is familiar to analysts, but that may not fit the usual taxonomies of abstractions (Zeigler, 1990; Caughlin and Sisti, 1997). If the base model's data are provided in the form of large arrays showing values by variable by day, then one can modify all of the data elements of a kind (e.g., all the daily deployment rates of a type shooter, or of all shooters) by appending multipliers. Then, one can examine questions such as "What if deployment rates double?" by merely changing the multiplier—a twenty-fold reduction of dimensionality for a given system in a twenty-day simulation. One can go further. For example, if one of the real-world concerns was

[66]Somewhat related points are made by Bruce Fowler (Fowler, 1992) in the context of an article discussing partial-differential-equation approaches to variable-resolution modeling.

that the initial part of the deployment might be slower than planned, then a one-parameter function of time could be applied to each data item. This procedure would allow the user to answer questions such as "What if it takes longer than planned for the initial deployment to get up to nominal performance?" with only a single parameter. Conceptually, we might speak of scaling, stretching, and distorting data vectors in useful but crude ways. All of these are quite useful in MRM systems. RAND has used such methods for about 15 years in the RAND Strategy Assessment System (RSAS) and its follow-on, the Joint Integrated Contingency Model (JICM). This work has proven invaluable in improving model flexibility and responsiveness. Such methods have been used by others in the community but often are *not* used.[67]

Observations About Uncertainty and Probabilities

In this report we do not discuss stochastic issues significantly—not because they are unimportant, but for lack of time and space. We are currently preparing a paper on the subject, but we want to post two observations here. First, parametric uncertainties are often so large that statistical uncertainties are only a minor portion of the problem. It then makes little sense to add the complexity of a stochastic representation for a few statistical uncertainties. It may be better to do parametric sensitivities with a deterministic model or to go to a full-fledged exploratory analysis that varies many parameters simultaneously over a wide range (Davis and Carrillo, 1997). On the other hand, there are problems in which probabilistic dependencies are fundamental to results. Working with such problems is quite different from working with statistical uncertainty akin to "noise." In such cases the aggregate variable in an MRM system should be represented by probability distributions that may not be at all Gaussian or otherwise pleasant in form. They may have large tails, they may be bimodal, and so on. It may be necessary to include joint probability distributions to reflect correlations. No general statement can be made on the matter. Probabilistic implications arise, for example, in cases of mine warfare and air defense (Horrigan, 1991 and 1992). They can also arise in ground warfare and in many other domains (see Appendices I and J in NRC (1997)). There is a major need for probabilistically sophisticated detailed simulations that could be used to assess when simpler representations are and are not valid.

---

[67]Whether these methods should be considered abstractions is subjective. An alternative would be to think of them as means for structuring the input space in defining the experimental frame. We see them as abstractions, because asking "What if we front-end-load the deployment?" or "What if things get bollixed up and everything slows down?" are abstractions of questions that might be posed in terms of alternative detailed operations plans.

Finally, let us note that probabilistic treatments are also quite useful in exploratory analysis of parametric uncertainty. We are preparing a paper on that subject currently, using a variant of the military problem treated here as our example.

Observations About Aggregations

The approach we have taken in this analysis was to work through a particular military problem in considerable detail, introducing additional abstractions and MRM levels sequentially. The purpose, of course, was to draw inferences about more general issues and principles. Among the observations we can make in the wake of our worked-out problem are these:

- Common aggregations may be quite misleading in their sensitivity to variables. Indeed, even their sensitivities with respect to aggregate variables may be misleading.
- However, common aggregations are often rather poor in comparison with other aggregations available—with some effort.
- One of the poorest approximations to make is to gloss over distinctions between qualitatively different phases of the phenomenon in question. If one does that, the aggregate approximation will likely give a poor picture of dynamics and poor estimates of sensitivities associated with phase-dependent variables.
- Better aggregations will involve using different constants for different phases. In the example we showed here, the phases involved time.

Some of these generalizations are familiar from other contexts. For example, all of the major DoD theater-level models of ground combat distinguish among some different phases directly or indirectly. They may be time-stepped, and at each time step they may readjust various parameters to reflect local circumstances of terrain and defensive preparations. They do not average over battles in well-prepared defenses in rough terrain and subsequent hasty battles in open terrain—except to the extent that the time step is too long (a day is often too long). Some models such as JICM and its predecessor the RSAS are strongly designed around the concept of "type battles" and the related concept of phase of battle. For example, in the JICM, a defender who is attempting to hold a line with a force-to-space ratio that makes such defense infeasible is adjudicated as suffering a "breakthrough," immediately after which he is penalized with high-attrition and fast-attacker movements in the "exploitation" phase. This representation made it possible to obtain battle dynamics that mirrored more closely the qualitative description of large breakthrough battles in World War II. In these battles initial movement rates were very low while the attacker was suffering high attrition, but—after paying the "entry price"—the attacker would break through and move rapidly, while also collecting prisoners.

7. Conclusions and Recommendations

Implications for MRM Design

Although we are far from having a primer for MRM design, the experience illustrated in this report suggest some differences between normal good design practice and what may be needed for MRM. Table 7.1 summarizes these.

The first distinction reflects a conflict between designing for "generality" and MRM (and related analytic agility). Normal design emphasizes using "fundamental" variables in a bottom-up approach. In contrast, MRM emphasizes finding and exploiting natural aggregate variables. The problem is that very useful aggregate variables in one context may not be so useful or valid in another context. This matter is part of a larger issue. In normal design practice DoD modelers

Table 7.1

Distinctions Between Normal and MRM Design

Normal Design	MRM Design
Design bottom-up and maintain generality (avoid using "trick variables" that might not be useful in a generalized version)	Exploit "aggregation fragments" in algorithms to provide multiple levels of operation
Design to reflect single best "reality"	Design to honor multiple representations, including reductionist representations. Design alternative user modes and corresponding model versions
Use object-oriented modeling and programming, building objects around physically similar entities (e.g., aircraft, missiles)	Extend "object" concept to "functionally" similar entities (e.g., shooters, whether aircraft or missiles)
Avoid structural approximations to maximize degrees-of-freedom flexibility	Seek approximations that simplify model structure sensibly, thereby increasing analytical flexibility for exploration and quick changes of assumption
Display aggregate variables only as needed	Choose perspective up front from possible modularized combinations of aggregate variables
Focus on simulation rather than needs for sensitivity tests	Insert multipliers, stretchers, and other "abstractions of convenience"
Design with bottom-up ethic	Design with ethic of using bottom-up, top-down, and sideways approaches in mutually supportive ways

seek a single valid bottom-up description, while in MRM there is need for alternative higher-level (more aggregate) depictions, which leads to the need to define alternative user modes and variants of the overall model corresponding to those user modes (variants differing only in their choice of aggregate variables).

We also observed that MRM design involves exploiting abstractions that may seem unnatural to those attempting to build a "virtual reality" simulation, and that this can lead to a conflict between *normal* DoD use of object-oriented methods and what MRM needs. Object-oriented modeling can certainly incorporate MRM abstractions such as "shooters," but at the expense of lumping together types of object that may be quite dissimilar. It is unclear to us whether designing around meta-objects such as "shooters" would cause problems for those who might use the resulting simulation less for joint-warfare analysis than, say, for training a particular Service's officers.

Normal modeling practice (especially by "modelers" who go directly to programming) seems to avoid structural approximations that would reduce the simulation's flexibility in incorporating complex data packages. In contrast, MRM design would presumably want to exploit structural approximations such as time smoothing—so as to increase analytic agility and to avoid encouraging users to "believe" the precision of overly detailed data such as daily arrival rates of platforms over a forty-day campaign.

Normal modeling practice seems to avoid introducing aggregate variables unless requested, usually in an output display. In contrast, MRM requires identifying and exploiting good aggregate variables from the outset of design.

Normal simulation modeling seems more focused on representing particular cases richly than on allowing for "What if?" questions of an aggregate nature. In MRM design we believe workers will want to build in "stretchers," "multipliers," and other "abstractions of convenience" that may not seem natural to the simulators.

Finally, but most fundamentally, we believe that normal DoD simulation modeling is driven by a bottom-up ethic and the notion—often implicit—that truth resides (or should reside) in the most detailed description. In contrast, enlightened MRM design involves modeling bottom-up, top-down, and every which way, with the intention of developing *mutually* calibrated models in a family.

As noted above, we do not yet have a primer for MRM, but we suspect that such a primer would have to address all of these issues.

Recommended Next Steps for Research

Despite the early state of MRM theory and practice, much can be accomplished now. We recommend that the Department of Defense conceive and develop the JWARS and JSIMS programs in a way that includes research leading to MRM families of models, albeit families that would be imperfect and noncomprehensive. The payoffs for eventual users of these and other DoD models would be great. Such work, however, is not a mere matter of "programming," nor a matter of routine operations research. There is need for in-depth research in military phenomenology that leads to sound theories and designs well before coding even begins. Building high-resolution simulations with broader scope and with agent-based adaptive decision processes can be very useful in this context, and the results could be used to better understand when various deterministic and stochastic aggregations are valid. In other cases research will need to exploit live or virtual simulations.

We also believe that it is evident that substantially more research is needed—research ranging from fundamental work on mathematics and modeling methods for MRM to applied work on improved design tools, including computational tools to assist MRM. The tools needed should integrate features such as visual programming,[68] very-high-level language features for analytical manipulations, statistical machinery suited to work with highly nonlinear models and able to accept theory-driven inputs, and probabilistic tools. Further, these tools should integrate well with the ultimate programming languages driving the simulations, and should reflect sound principles of software engineering, including the ability to design in an object-oriented modeling framework (see, e.g., Rumbaugh et al., 1991).

---

[68] *Analytica* and *iThink* are examples of systems with visual-programming features. *Visual Basic* is arguably a misnomer, although it is fast gaining popularity for its features and applicability. *Mathematica* and *Maple* are two well-known systems that include symbolic manipulation for analytical problem solving.

Appendix: Model of the Halt Problem

The starting model used in the text for the halt problem is simplified (abstracted) in many respects because we wanted to focus on particular methodological issues. Here let us merely mention briefly some of the issues that have been glossed over.

First, the model treats several key issues implicitly if at all, particularly command-control (C^2), C^4ISR, and the effects of and attacks on air defenses. Figure A.1 shows a more general data-flow diagram in which these factors are included explicitly.

These generalizations, however, are only a few of the many that could be added if the purpose were to represent combat operations with more fidelity rather than to provide a method for assessing how much could be accomplished by precision fires alone, in a variety of circumstances. Other important factors omitted or treated only indirectly include:[69]

- Engagement mechanisms by other than long-range precision fires (e.g., by regular or irregular ground units, including for purposes of delay)
- Details of terrain, road networks, and movement
- Enemy micro-strategy (e.g., spacing among vehicles, unit separations, dash tactics to minimize exposure time)
- Friendly counter-strategy (e.g., efforts to anticipate when and where vehicles will be on the road, or to exploit chokepoints and nodes)
- The likely nonlinear effectiveness of the overall shooter force as a function of threat size, shooter-force size, mix, details of terrain and tactics, and, of course, C^4ISR systems and related fusion capabilities)
- Detailed weapon characteristics and weapon-mix issues
- Weather.

Many others might be mentioned, but this list should suffice to underline the point that our starting-point model, which we treated as a "detailed model" for the purposes of this report, was already much abstracted.

[69]Some of these issues are touched upon in Ochmanek, Harshberger, Thaler, and Kent (1998) and Davis and Carrillo (1997). They are being explored in more depth currently in RAND work for the Air Force, Army, and OSD.

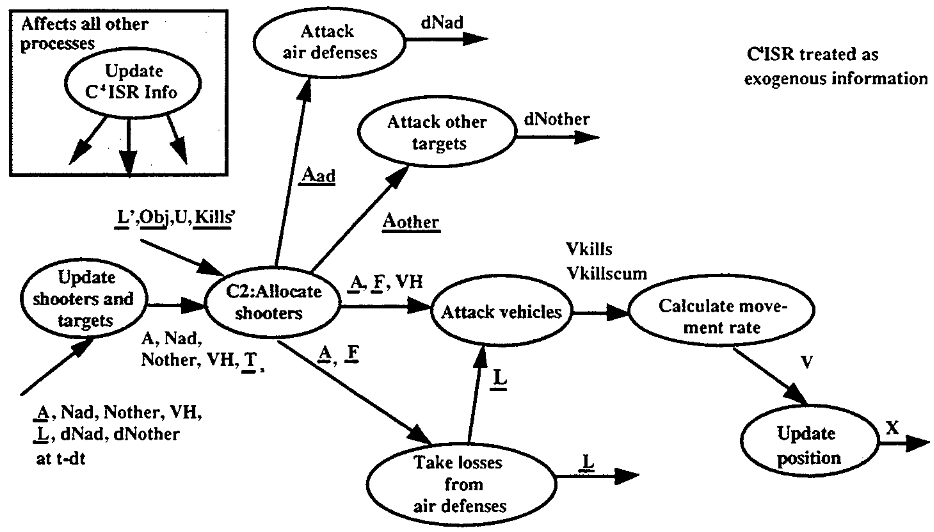

Figure A.1—An Expanded Data-Flow Diagram Showing Blue's Command Process, Role of C^{4}ISR, and Losses to Air Defense

Bibliography

Alberts, David S., and Thomas J. Czerwinski (eds.), *Complexity, Global Politics, and National Security*, National Defense University, Washington, D.C., 1997.

Allen, Patrick, "The Need to Represent a Wide Variety of Battle Types in Air-Ground Combat Models," *Military Operations Research*, Vol. 1, No. 3, 1995, pp. 19–25.

———, *Situational Force Scoring: Accounting for Combined Arms Effects in Aggregate Combat Models*, N-3423-NA, RAND, Santa Monica, Calif., 1992.

Asman, William, and Hans Jaarvield, "A Variable Resolution Transport Model Applied for NH2 in Europe," *Atmosphere Environmental*, 1992.

Axtell, Robert Lea, *Theory of Model Aggregation for Dynamical Systems with Applications to Problems of Global Change*, Dissertation, Carnegie-Mellon University, 1992; available from University Microfilms International, Ann Arbor, Mich.

Bankes, Steve C., *Exploratory Modeling and the Use of Simulation for Policy Analysis*, N-3093-A, RAND, Santa Monica, Calif., 1992.

Bigelow, James H., *Performance-Oriented Logistics Assessment (POLA): Executive Summary*, R-3823-A, RAND, Santa Monica, Calif., 1992.

———, *Managing Recoverable Aircraft Components in the PPB and Related Processes*, R-3094, RAND, Santa Monica, Calif., 1994.

Bonder, Seth, "Impact of the New Global Environment on U.S. National Security Planning—Challenges to the OR Community," *International Transactions of Operational Research*, Vol. 1, No. 1, pp. 31–39, 1994.

Bracken, Jerome, Moishe Kress, and Richard Rosenthal, *Combat Models*, Military Operations Research Society, Alexandria, Va., 1995.

Brooks, Arthur, Steve Bankes, and Bart Bennett, *Weapons Mix and Exploratory Analysis*, DB-216/2-AF, RAND, Santa Monica, Calif., 1997.

Casti, John, *Would-Be Worlds: How Simulation Is Changing the Frontiers of Science*, John Wiley & Sons, Inc., New York, 1997.

Caughlin, Don, and Alex F. Sisti, "A Summary of Model Abstraction Techniques," presented at the International Society for Optical Engineering (SPIE), April 22–24, Orlando, Florida, 1997.

Cloud, David J., and Larry B. Rainey (eds.), *Applied Modeling and Simulation: An Integrated Approach to Development and Operation*, McGraw Hill, New York, 1998.

Courtois, Paul J., "On Time and Space Decomposition of Complex Structures," *Communications of the ACM*, Vol. 28, No. 6, pp. 590–603, 1985.

Cohen, Michael J., John E. Rolph, and Duane L. Steffey (eds.), *Statistics, Testing, and Defense Acquisition: New Approaches and Methodological Improvements*, National Research Council, National Academy Press, Washington, D.C., 1998.

Coveney, Peter, and Roger Highfield, *Frontiers of Complexity: The Search for Order in a Chaotic World*, Fawcett Columbine, Ballantine Books, New York, 1995.

Crick, Francis, *The Astonishing Hypothesis: The Scientific Search for the Soul*, Scribners, New York, 1994.

Davis, Paul K., *A New Methodology for Modeling National Command Level Decisionmaking In War Games and Simulations*, R-3290-NA, RAND, Santa Monica, Calif., 1986.

———, "Planning for Adaptiveness," in Paul K. Davis (ed.), *New Challenges in Defense Planning: Rethinking How Much Is Enough*, RAND, Santa Monica, Calif., 1994.

———, *Aggregation, Disaggregation, and the 3:1 Rule*, MR-638-AF/A/OSD, RAND, Santa Monica, Calif., 1995.

———, *An Introduction to Variable-Resolution Modeling and Cross-Resolution Model Connection*, R-4252-DARPA, RAND, Santa Monica, Calif., 1993; also in Bracken, et al., 1995.

Davis, Paul K., and Manuel Carrillo, *Exploratory Analysis of the "Halt Problem": A Briefing on Methods and Initial Insights*, DB-232-OSD, RAND, Santa Monica, Calif., 1997.

Davis, Paul K., and Richard J. Hillestad (1992a), "Families of Models That Cross Levels of Resolution: Issues for Design, Calibration, and Management," *Proceedings, 1993 Winter Simulation Conference*, December, 1993.

Davis, Paul K., and Richard J. Hillestad (1992b), *Proceedings of Conference on Variable Resolution Modeling*, May 5–6, Washington, D.C., 1992, CF-103, RAND, Santa Monica, Calif., 1993.

Davis, Paul K., and Reiner Huber, *Variable Resolution Modeling: Issues, Principles, and Challenges*, N-3400-DARPA, RAND, Santa Monica, Calif., 1992.

Davis, Paul K., and Irwin Oppenheim, "Vibrational Relaxation in Gases," *The Journal of Chemical Physics*, Vol. 56, No. 1, 1971, pp. 86–94.

Davis, Paul K., David Gompert, and Richard Kugler, *Adaptiveness in National Defense: the Basis of a New Framework*, Issue Paper IP-155, RAND, Santa Monica, Calif., 1996.

Davis, Paul K., Richard J. Hillestad, and Natalie Crawford, "Capabilities for Major Regional Contingencies," in Zalmay Khalilzad and David Ochmanek, eds., *Strategic Appraisal 1997: Strategy and Defense Planning for the 21st Century*, RAND, Santa Monica, Calif., 1997.

Defense Science Board, *Task Force On Tactics and Technology for 21st Century Superiority*, report of the modeling and analysis panel, Appendix B in Volume 2, Department of Defense, Washington, D.C., 1996.

Dewar, James A., Steve Bankes, James Hodges, Thomas Lucas, Desmond Saunders-Newton, and Patrick Vye, *Credible Uses of the Distributed Interactive Simulation (DIS) System*, MR-607-A, RAND, Santa Monica, Calif., 1996.

Dupuy, Trevor N., *Understanding War: History and Theory of Combat*, Paragon House Publishers, New York, 1987.

Fall, Thomas, "Dynamic Focusing Approach to Mixed Level Simulation," unpublished paper presented to the International Society for Optical Engineering (SPIE), 1997.

Fishwick, Paul A., "The Role of Process Abstraction in Simulation," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 19, No. 1, 1988.

———, "Abstraction Level Traversal in Hierarchical Modeling," in *Modeling and Simulation Methodology: Knowledge Systems Paradigms*, B. P. Zeigler et al., eds., Elsevier North Holland, New York, 1989.

Fishwick, Paul A,. and Kangsun Lee, "Two Methods for Exploiting Abstraction in Systems," *Proceedings AI, Simulation and Planning in High Autonomy Systems Conference,* March 23–27, 1996, pp. 257–264.

Fowler, Bruce, "Resolution Changes and Renormalization for Partial Differential Equation Combat Models," in Davis and Hillestad, 1992b.

Frantz, Frederick K., "Analyzing Models for Abstraction," unpublished paper presented to the International Society for Optical Engineering (SPIE), 1997.

Harshberger, Edward R., Bart E. Bennett, and David R. Frelinger, "An Approach to Hierarchies of Models: Process Independence," in Davis and Hillestad (1992b).

Hillestad, Richard J., and Mario L. Juncosa, *Cutting Some Trees To Save the Forest: On Aggregation and Disaggregation in Combat Models*, MR-189, RAND, Santa Monica, Calif., 1993; also in Bracken et al., 1995.

Hillestad, Richard J. , and Louis Moore, *The Theater-Level Campaign Model : A New Research Prototype for a New Generation of Combat Analysis Model* , MR-388-AF/A, RAND, Santa Monica, Calif., 1996; MR-710 is the executive summary.

Hillestad, Richard, John G. Owen, and Donald Blumenthal, *Experiments in Variable Resolution Modeling*, N-3631-DARPA, RAND, Santa Monica, Calif., 1992; also in Bracken et al. 1995.

Holland, John, *Hidden Order*, Addison-Wesley, Reading, Mass., 1995.

Horrigan, Timothy, "Configural Theory and the Mathematical Modeling of Combat," Horrigan Analytics, HAS 91-17-1, 1991.

Horrigan, Timothy, "The 'Configuration Problem' and Challenges for Aggregation," in Davis and Hillestad (1992b), 1992.

Ilachinski, Andrew, *Land Warfare and Complexity*, 2 vols, Center for Naval Analyses, Alexandria, Va., 1996.

Iwasaki, Yumi, and Herbert Simon, "Causality and Model Abstraction," *Artificial Intelligence*, Vol. 67, pp. 143–194, 1994.

Lee, Kangsun, and Paul A. Fishwick, "Semi-Automated Method for Dynamic Model Abstraction," unpublished paper presented to SPIE, 1997.

Lempert, Robert, Michael Schlesinger, and Steven Bankes, "When We Don't Know the Costs or the Benefits: Adaptive Strategy for Abating Climate Change," *Climactic Change*, Vol. 33, 1996; reprinted as RAND RP-557, 1996.

Lorenz, Edward, *The Essence of Chaos*, UCL Press, London, 1993.

Matsumura, John, Randall Steeb, Thomas Herbert, Mark Lees, Scott Eisenhard, and Angela Stich, *Analytical Support to the Defense Science Board: Tactics and Technology for 21st-Century Military Superiority*, DB-198A, RAND, Santa Monica, Calif., 1997.

Mori, Hazime, "Transport, Collective Motion, and Brownian Motion," *Progress of Theoretical Physics*, Vol. 33, No. 3, March 1965, pp. 423–455.

National Research Council, *Modeling and Simulation*, Vol. 9 of *Technology for the Navy and Marine Corps, 2000-3030*, National Academy Press, Washington, D.C., 1997.

NATO, *Symposium on Coping with Uncertainty in Defence Decision Making*, two volumes, AC/243 (Panel 7) TP/9, report of a symposium held at TNO Physics and Electronics Laboratory, the Netherlands, January 16–18, 1995.

Nicolis, Gregorie, and Ilya Prigogine, *Exploring Complexity*, Freeman, New York, 1989.

Ochmanek, David A., Edward R. Harshberger, David E. Thaler, and Glenn A. Kent, *To Find, and Not to Yield: How Advances in Information and Firepower Can Transform Theater Warfare*, MR-958-AF, RAND, Santa Monica, Calif., 1998.

Pozhar, Liudmila A., *Transport Theory of Inhomogeneous Fluids*, World Scientific, Singapore, 1994.

Reynolds, Paul F., Sudhir Srinivasan, and Anand Natrajan, "Consistency Maintenance in Multiresolution Simulation," *TOMACS*, Vol. 7, No. 3, 1997, pp. 368–392.

Rinaldi, Steven M., "Complexity Theory and Air Power," in Alberts and Czerwinski (1997).

Rogers, David F., et al., "Aggregation Techniques and Methodology in Optimization," *Operations Research*, Vol. 39, No. 4, July–Aug, 1991.

Rumbaugh, James, Michael Blaha, Jim Rumbaugh, William Premerlani, and Frederick Eddy, *Object-Oriented Modeling and Design*, Prentice Hall, Englewood Cliffs, New Jersey, 1991.

Scheckeler, Günther, "A Hierarchy of Models at Different Resolution Levels," in Davis and Hillestad (1992b).

Schrödinger, Irwin, *What Is Life? and Other Scientific Essays*, Doubleday, New York, 1956.

Schmitt, John F., "Command and (out of) Control: The Military Implications of Complexity Theory," in Alberts and Czerwinski, 1997.

Sevinic, Suleyman, "Automation of Simplification in Discrete Event Modeling and Simulation," in *International Journal of General Systems*, Vol. 18, 1990, pp. 125–145.

Simon, Herbert, "Prediction and Prescription in Systems Modeling," *Operations Research*, Vol. 38, No. 1, 1990, pp. 7–14.

———, *Sciences of the Artificial*, 3rd ed., MIT Press, Cambridge, Mass., 1996.

Simon, Herbert, and A. Ando, "Aggregation of Variables in Dynamic Systems," *Econometrica*, Vol. 29, No. 2, 1961, pp. 111–138.

U.S. Marine Corps, *Command and Control*, FM 6-1, Washington, D.C., 1997.

Van Trees, Harry L., *Detection, Estimation, and Modulation Theory*, John Wiley & Sons, Inc., New York, 1968.

Waldrop, Mitchell, *Complexity*, Simon and Schuster, New York, 1992.

Wolfram, Stephen, *The Mathematica Book*, 3d ed., Cambridge University Press, New York, 1996.

Zeigler, Bernard, *Multifaceted Modelling and Discrete Event Simulation*, Academic Press, London and Orlando, Fla., 1984.

———, "Review of Theory in Model Abstraction," *SPIE*, 1990.

Zeigler, Bernard, H. Praehofer, T. G. Kim, *Theory of Modeling an Simulation*, 2nd ed., Academic Press, London and Orlando, Fla., 1998.

Zwanzig, Robert, *Physical Review*, Vol. 124, 1961, p. 983.