SYSTEMS ENGINEERING
Research Center

Enterprise Systems Analysis

Technical Report SERC-2015-TR-020-4

January 14, 2015

Dr. Michael J. Pennock, Stevens Institute of Technology

Dr. William B. Rouse, Stevens Institute of Technology

Dr. Doug Bodner, Georgia Institute of Technology

Dr. Christopher Gaffney, Stevens Institute of Technology

Mehrnoosh Oghbaie, Stevens Institute of Technology

Pallavi Prasad, Georgia Institute of Technology

CENTER FOR COMPLEX SYSTEMS & ENTERPRISES

Tennenbaum Institute

Contract Number: HQ0034-13-D-0004 Task Order: 010, RT110

Report No. SERC-2015-TR-020-4

The Systems Engineering Research Center (SERC) is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Office of the Assistant Secretary of Defense for Research and Engineering (ASD(R&E)) under Contract H98230-08-D-0171 and HQ0034-13-D-0004 (TO 0110).

Executive Summary

This report documents the results of the SERC research task RT-110: Enterprise Systems Analysis. This task is a part of a larger SERC focus area: Enterprise and Systems of Systems (ESOS). The overarching goal of the ESOS focus area is to understand how, if at all, systems engineering methods can be adapted to address the challenges imposed by both enterprise systems and systems of systems.

As one aspect of addressing the challenges imposed by enterprise systems, an enterprise modeling methodology was developed as part of SERC research task RT-44a. The chief objective of this task (RT-110) was to evaluate that methodology critically and to identify areas for improvement. To that end, the method was applied to a case study of counterfeit parts in the DoD supply chain as test case. In parallel, the methodology was considered from a theoretical viewpoint based on the work conducted in adjacent domains such as complex systems, modeling and simulation, human factors, and economics. This two-pronged approach challenged the methodology both practically and technically. This report documents the findings of that evaluation.

In short, the methodology was found to be a useful construct, but some challenges and areas for improvement remain. In particular, some of the steps were found to be vague and require additional guidance to actually implement. The methodology makes extensive use of visualization, but the efficacy of visualization applied to enterprise systems remains unclear. Finally, the methodology relies heavily on the composition of computational representations of different aspects of an enterprise system. Such compositions impose substantial technical challenges when attempted for enterprise systems.

To address these issues, a number of recommendations were made. First, potential improvements and refinements to the methodology steps were developed. These culminated in a set of high level requirements for an enterprise modeling archive that would aid enterprise modelers in the execution of the methodology by capturing guidance and best practices.

Second, past work in visualization was examined to identify guidance for implementation. The output of this was a conceptual design for an interactive visualization based on this analysis. However, a number of unresolved difficulties were identified and would require further experimentation to address.

Third, a set of strategies was developed to mitigate the difficulties associated with model composition. These strategies explicitly recognize that model composition will not always be possible and compensates appropriately. Visualization was also considered as a potential approach to support the development of strategies when complete composition is not possible.

Beyond the modeling methodology itself, the findings of this research task have implications for the larger challenge of applying systems engineering approaches to enterprise systems. In particular, traditional systems approaches can only ever be applied to a subset of an enterprise problem. When one persists in applying only traditional systems approaches to an enterprise, one necessarily excludes all aspects of the enterprise system that are not “well behaved.” Unfortunately, this has the consequence of implicitly assuming that everything outside of the modeled subset is either fixed or has no impact on that subset. While this implicit assumption can be enforced to a degree via organizational constructs in traditional systems engineering, it is almost certainly invalid when attempting to engineer enterprise systems. Since enterprises are often highly interconnected (both internally and externally) and intrinsically adaptive, optimizing over a subset of the problem will often result in a fragile solution that can trigger a number of unintended consequences via second or higher order effects.

To address this larger challenge, the aforementioned strategy set is critical. In essence, it explicitly recognizes the modeling limitations and then allows one to consider how to adapt to or hedge sudden changes. Here multiple models become a potential asset as they may allow a decision maker to map out potential future scenarios and hedge them. Finally, application of these strategies can be facilitated through the identification of leading indicators of a major, impending change in an enterprise system. While there is some theoretical basis for these indicators, more research is required to assess their practicality.

While a great deal of progress was made over the course of RT-110, work remains. First and foremost, the model developed to analyze the counterfeit parts case study should be applied to determine if it actually yields useful insights. This is necessary to establish the ultimate utility of the enterprise modeling approach. Second, a single example is not sufficient to validate a methodology. Consequently, another case study should be initiated to continue to challenge the enterprise modeling methodology. Third, the efficacy of visualization to support enterprise decision making needs to be established. This means that one or more experiments should be conducted to determine the conditions under which visualizations actually aid enterprise decision makers and enhance decision making. Finally, additional work is necessary to implement the recommendations made in this report. This includes explicitly mapping the chain of phenomenon to representation to model, investigations into how to operationalize the strategy framework, as well as additional guidance for model composition and when it is appropriate to do so.

This Page Intentionally Left Blank

TABLE OF CONTENTS

Figures and Tables

1 INTRODUCTION

This report documents the results of the SERC research task RT-110: Enterprise Systems Analysis. This task is a part of a larger SERC focus area: Enterprise and Systems of Systems (ESOS). The overarching goal of the ESOS focus area is to understand how, if at all, systems engineering methods can be adapted to address the challenges imposed by both enterprise systems and systems of systems.

As one aspect of addressing the challenges imposed by enterprise systems, an enterprise modeling methodology was developed as part of SERC research task RT-44a. The chief objective of this task (RT-110) was to evaluate that methodology critically and identify areas for improvement. To that end, the method was applied to a case study of counterfeit parts in the DoD supply chain as test case. In parallel, the methodology was considered from a theoretical view point based on the work conducted in adjacent domains such as complex systems, modeling simulation, human factors, economics, and human factors. This two-pronged approach challenged the methodology both practically and technically.

What was found was that the methodology is a useful construct, but some challenges and areas for improvement remain. In particular, some of the steps were found to be vague and require additional guidance to actually implement. The methodology makes extensive use of visualization, but the efficacy of visualization applied to enterprise systems remains unclear. Finally, the methodology relies heavily on the composition of computational representations of different aspects of an enterprise system. Such compositions impose substantial technical challenges when attempted for enterprise systems. This report documents these findings and recommendations for improvement.

As the evaluation of the enterprise modeling methodology is a key driver of the work, it also drives the organization of this report. First, we revisit the enterprise modeling methodology developed during RT-44a (Section 2). This section contains a summary of the methodology (Section 2.1). From there, we review the lessons learned from applying the methodology to the counterfeit parts case study (Section 2.2). Based on the findings, we elaborate on the steps of the methodology in Section 2.3. This includes a summary of the research initiated to improve the steps as well as additional implementation guidance for some of the steps.

The remainder of the report organizes the ideas and issues raised in Section 2 into logical groups and elaborates on them. Section 3 presents the counterfeit parts case study both in terms of how the methodology was applied and a description of the resulting model. Section 4 captures the theoretical issues associated with actually accomplishing the goals of the methodology. To compensate for some of those theoretical issues, Section 5 presents the strategy framework. Since the methodology heavily relies on the computational composition of multiple views of the enterprise, Section 6 captures the technical issues with doing so. Similarly, the importance of visualization to the methodology led us to investigate the applicability of visualization to representing enterprise systems. This is captured in Section 7. As one of the findings from the counterfeit parts case study is that additional implementation support would improve the efficacy of the methodology, Section 8 documents requirements for a model archive to aid in identifying phenomena, visualizations, representations, best practices, and associated implications. Finally, Section 9 summarizes the conclusions drawn from this work and outlines future work.

As part of our investigation into visualization, a survey of off-the-shelf visualization tools was conducted. The results of this survey are documented in the Appendix. Furthermore, we developed a conceptual design for interactive visualization. While this is documented in Section 7.2, images of the design were necessarily reduced to fit into the body of the report. Consequently, a higher resolution version of the conceptual design is provided as a separate file.

2 REVISITING THE MODELING METHODOLOGY FOR ENTERPRISE SYSTEMS

The 10-step modeling methodology for enterprise systems was documented in the RT-44a final report. During RT-110, the modeling methodology was addressed in two ways. First, the methodology was applied to the counterfeit parts case study in order to identify areas of improvement. Second, the individual steps of the methodology were considered in greater depth and elaborated as needed.

In this section we will briefly review the 10 step methodology. Next, we will discuss the lessons learned from the application of the method to the counterfeit parts case study. (The full details of the application of methodology to counterfeit parts case study are provided in Section 3.3.) Finally, we will consider how particular steps can be elaborated.

2.1 REVIEW OF THE TEN STEP MODELING METHODOLOGY

To provide context for the subsequent discussion, below is a summary of the 10-step enterprise modeling methodology developed during RT-44a:

Step 1: Decide on the Central Questions of Interest

The history of modeling and simulation is littered with failures of attempts to develop models without clear intentions in mind. Models provide means to answer questions. Efforts to model socio-technical systems are often motivated by decision makers' questions about the feasibility and efficacy of decisions on policy, strategy, operations, etc. The first step is to discuss the questions of interest with the decision maker(s), define what they need to know to feel that the questions are answered, and agree on key variables of interest.

Step 2: Define Key Phenomena Underlying These Questions

The next step involves defining the key phenomena that underlie the variables associated with the questions of interest. Phenomena can range from physical, behavioral, or organizational, to economic, social or political. Broad classes of phenomena across these domains include continuous and discrete flows, manual and automatic control, resource allocation, and individual and collective choice. Mature domains often have developed standard descriptions of relevant phenomena.

Step 3: Develop One or More Visualizations of Relationships among Phenomena

Phenomena can often be described in terms of inputs, processes, and outputs. Often the inputs of one phenomenon are the outputs of other phenomena. Common variables among phenomena provide a basis for visualization of the set of key phenomena. Common visualizations methods include block diagrams, IDEF, influence diagrams, and systemigrams.

Step 4: Determine Key Tradeoffs That Appear to Warrant Deeper Exploration

The visualizations resulting from Step 3 often provide the basis for in-depth discussions and debates among members of the modeling team as well as the sponsors of the effort, which hopefully includes the decision makers who intend to use the results of the modeling effort to inform their decisions. Lines of reasoning, perhaps only qualitative, are often verbalized that provides the means for immediate resolution of some issues, as well as dismissal of some issues that no longer seem to matter. New issues may, of course, also arise.

Step 5: Identify Alternative Representations of These Phenomena

Computational representations are needed for those phenomena that will be explored in more depth. These representations include equations, curves, surfaces, process models, agent models, etc. – in general, instantiations of standard representations. Boundary conditions can affect choices of representations. This requires deciding on fixed and variable boundary conditions such as GDP growth, inflation, carbon emissions, etc. Fixed conditions can be embedded in representations while variable conditions require controls such as slider bars to accommodate variations – see Step 9.

Step 6: Assess the Ability to Connect Alternative Representations

Representations of phenomena associated with tradeoffs to be addressed in more depth usually require inputs from other representations and produce outputs required by other representations. Representations may differ in terms of dichotomies such as linear vs. nonlinear, static vs. dynamic, deterministic vs. stochastic, continuous vs. discrete, and so on. They may also differ in terms of basic assumptions, e.g., Markov vs. Non-Markovian processes. This step involves determining what can be meaningfully connected together.

Step 7: Determine a Consistent Set of Assumptions

The set of assumptions associated with the representations that are to be computationally connected need to be consistent for the results of these computations to be meaningful. At the very least, this involves synchronizing time across representations, standardizing variable definitions and units of measures, and agreeing on a common coordinate system or appropriate transformations among differing coordinate systems. It also involves dealing consistently with continuity, conservation, and independence assumptions.

Step 8: Identify Data Sets to Support Parameterization

The set of representations chosen and refined in Steps 5-7 will have parameters such as transition probabilities, time constants, and decay rates that have to be estimated using data from the domain(s) in which the questions of interest are to be addressed. Data sources need to be identified and conditions under which these data were collected determined. Estimation methods need to be chosen, and in some cases developed, to provide unbiased estimates of model parameters.

Step 9: Program and Verify Computational Instantiations

To the extent possible, this step is best accomplished with commercially available software tools. The prototyping and debugging capabilities of such tools are often well worth the price. A variant of this proposal is to use commercial tools to prototype and refine the overall model. Once the design of the model is fixed, one can then develop custom software for production runs. The versions in the commercial tools can then be used to verify the custom code. This step also involves instantiating interactive visualizations with graphs, charts, sliders, radio buttons, etc.

Step 10: Validate Model Predictions, at Least against Baseline Data

The last step involves validating the resulting model. This can be difficult when the model has been designed to explore policies, strategies, etc. for which there inherently is no empirical data. A weak form of validation is possible by using the model to predict current performance with the “as is” policies, strategies, etc. In general, models used to explore "what if" possibilities are best employed to gain insights that can be used to frame propositions for subsequent empirical study.

Summary

The logic of the ten-step methodology can be summarized as follows, with emphasis on Steps 1-7:

- Define the question(s) of interest
- Identify relevant phenomena
- Visually compose phenomena
- Identify useful representations
- Computationally compose representations

Note that this logic places great emphasis on problem framing and formulation. Deep computation is preserved for visually identified critical tradeoffs rather than the whole problem formulation. Steps 8-10 of the methodology are common to many methodologies.

Not all problems require full use of this ten-step methodology. Often visual portrayals of phenomena and relationships are sufficient to provide the insights of interest. As just noted, such views are also valuable for determining which aspects of the problem should be explored more deeply.

2.2 Lessons Learned from the Counterfeit Parts Case Study

The enterprise methodology was found to be useful overall, but does not provide enough guidance yet, especially for a non-expert modeler. Specific findings are the following:

1. Enterprise models require expert input and review from stakeholders representing different parts of the enterprise system being modeled. Expert review and input was done in an ad-hoc manner for this effort. Guidelines for effective use of such expert review and input need to be formulated. In particular, this input and review needs to be structured appropriately for maximum effectiveness for the first four methodology steps:
    a. Determining the central question of interest,
    b. Defining key phenomena,
    c. Developing and exploring visualizations, and
    d. Determining key trade-offs.
2. The scope of enterprise models needs to be bounded, since enterprises are complex phenomena. Additionally, complexity must be managed throughout the model lifecycle. Expert review and input can play a role here. However, guidelines based on experience in developing such models must be developed,

as well. For the current counterfeit parts model, the separation of the model into sub-models with an exogenous environment sub-model is a potentially useful way to begin addressing the issue.

3. Steps 6, 7 and 8 of the methodology (identifying alternate representations, assessing the ability to connect them, and determining consistent assumptions) are all part of an overall model architecture design effort, which includes model composition. In future versions of the methodology, it would be helpful to provide more guidance on model architecture design, as well as an explicit step or output relating to the model architecture.
4. The methodology is vague about what types of models best represent different aspects of an enterprise. Clearly, expert modelers over the years have compiled best practices for different aspects of enterprises, though mostly at an individual researcher level. What is needed is a catalog of such model types mapped to enterprise system elements, along with the conditions under which they are most applicable. This would support guidance for model architecture design.
5. Once representations are specified, it would be helpful for guidance on data formats or test datasets for the particular domain modeled by the representation.
6. The use of visualizations (step 3 of the methodology) is fairly limited here, since there is a limited understanding of how to do this effectively for enterprise modeling, and there do not seem to be readily available tools to support such efforts. Further research on visualization of enterprise behavior and outcomes is needed. One possibility is the following:
    a. An interactive but static influence diagram for enterprise exploration, coupled with
    b. An interactive and dynamic high-level system dynamics model, compatible with the existing model, with different "policy levers" that illustrate relationships and effects of policy decisions. The visualization output from such a model would be a series of graphs indicating magnitudes and trends of various variables of interest (e.g., counterfeit detections, capability losses).
    c. These are potentially useful types of visualization that could provide information on influences and interactions. Is this the best visualization approach, however, for this problem domain? Also, this raises the issue of configuration management between the models and visualizations.

Much of the remaining content of this report is intended to address one or more the issues raised here. In particular, elaboration of the methodology steps occurs in the following section (2.3). The challenges of managing complexity, the implications of establishing bounds, model selection, the interrelationships of those modes, and the strategies to address all of the above are covered in Sections 4, 5, and 6. The technical issues associated with visualizations are discussed in Section 7. Finally, requirements for a model archive to aid in identifying phenomena, visualizations, representations, best practices, and associated implications are described in Section 8. It is important to note that not all of the findings from the counterfeit parts case study could be addressed within the scope of RT-110, and thus, must be relegated to future work.

2.3 Elaboration of Methodology Steps

We can break the steps of the enterprise modeling methodology into three groups: those that required additional research, those that required additional methodological development, and those that are standard. Steps one through three culminate in the first visualizations of the problem and thus required research with regard to enterprise visualization. Steps six and seven involve computational composition, and thus required research with regard to model composition. Steps four and five required additional methodological development in that they were presented at a fairly high level in the RT-44a report and thus required additional definition to make them useable. Finally, steps eight through ten are standard for most modeling methodologies and thus do not require further elaboration.

2.3.1 Supporting Research for Steps One through Three

As noted previously, the objectives of steps one through three are to identify the phenomena relevant to the questions of interest and then develop visualizations that portray the relationships among the phenomena. The visualizations developed need to support the essentially cognitive activities of step four where decision makers, stakeholders, and analysts search for key relationships among the phenomena that warrant computational investigation. This is effectively an exercise in causal inference. The challenge, then, is how to construct visualizations that support causal inference for a potentially complex enterprise system.

While there is an extensive body of work on visualization, much of it is concerned with visualizing either individual phenomena or purely technical systems. There is some question as to how generalizable this work is to enterprise systems that tend to exhibit substantially greater complexity and multiple, potentially incompatible views. To that end, as part of this research task, an extensive investigation of the visualization literature was conducted. The key outputs of this effort are the identification of open issues in enterprise visualization that require additional research and a conceptual design for an interactive visualization that leverages what is known from the existing literature. These outputs are documented in Section 7 of this report.

Of particular concern with regard to the first three steps is the potential to trigger a specious assessment of the situation. This concern is driven by two factors. First, as the complexity of the visualized system increases, it is possible that decision makers could latch on to spurious correlations and draw incorrect inferences. Second, the group-based approach advocated by the methodology risks engendering a false sense of confidence in inferences drawn from the visualizations. In particular, Heath and Gonzalez [1995] investigated the effect of group interaction in interactive decision making. They found that interaction with others will increase the confidence in a decision but does not guarantee the accuracy or quality of the decision. Consequently, additional research, preferably experimental, is needed to assess and understand these risks to successfully implementing the methodology. The ultimate objective would be to provide guidance on how to properly construct visualizations to minimize these issues. Note any outputs of this effort would also support the execution of step four which will often involve refining the visualizations developed in step three. This notion is conveyed in the subsequent section.

2.3.2 Methodology for Step Four

Step four involves identifying key tradeoffs that warrant deeper investigation. This process will likely involve refining, augmenting, or perhaps even redoing the visualizations developed in step three. The question then becomes how one goes about doing so.

Fortunately, not all problems or questions require deep computational exploration. Visual portrayals of phenomena can lead to almost immediate recognition of where a mechanism might fail, or a process flow will lead to bottlenecks, or an incentive system may prompt unintended consequences. Visualization of phenomena can assist in identifying the portions of the overall problem that may merit deep computational modeling.

Computer technology, including graphics hardware and software, have made it possible to create impressive and pleasing visualizations of a wide range of phenomena. If one's sole purpose is to impress people and collect "wows," then the tools are available to achieve these ends. However, our goal is to create interactive visualizations that serve particular purposes.

We first need to consider users' purposes which visualizations are intended to support. Users' purposes seldom include using visualizations; these are simply the means to other ends such as:

- Problem Solving, perhaps using topographic rules [Rouse 1983] to explore structural relationships underlying the phenomena of interest
- Pattern Recognition, perhaps using symptomatic rules [Rouse, 1983], or recognition-primed decision making [Klein, 2003], to identify regularities and anomalies
- Procedure Execution, as illustrated for helicopter maintenance [Frey at al., 1992, 1993], which involves understanding how to execute procedural steps
- Navigation, involving maps and signs [Rasmussen, 1983] needed to move from one location to another, ranging from geographic locations to finding the subsystems of a power plant, for example

The methodology outlined below is intended to help one to design visualizations that will support users in their pursuits of these types of purposes.

Step 1: Identify information use cases, including knowledge levels of users

Use cases provide descriptions of alternative ways in which users will employ the visualizations to achieve their purposes – before the visualizations have been created. These high-level descriptions can be characterized in terms of six general tasks:

- Retrieve data and visualizations relevant to questions of interest
- Recognize characteristics of interest across chosen attributes
- Construct or select and parameterize representations provided
- Compute outputs of constructed representations over time
- Compare outputs to objectives or across output variations
- Refine constructed representations and return to Compute

Examples of the use of such descriptions are provided in the following subsection.

Step 2: Define trajectories in abstraction-aggregation space

Use cases define what information is needed and what actions are taken at every step of the task of interest. This includes the levels of abstraction and aggregation for requisite information elements and controls for each task.

Step 3: Design visualizations and controls for each point in space

Transform the outputs of Step 2 into what specifically users can see and do. There is a wealth of possibilities, which need to be compiled into a manageable set of choices. Note that many of the possibilities will be domain dependent.

Step 4: Integrate across visualizations and controls to dovetail representations

Visualizations and controls should not completely change for each step in a task. Integrated visualizations may be able to support more than one step. Individual controls may affect more than one view.

Step 5: Integrate across use cases to minimize total number of displays and controls

The set of visualizations may serve more than one purpose. For example, novices may use it to learn about a domain while experts use it to address real problems of interest. Experts may, for example, see the same central visualizations as novices, but have access to additional information and controls to enable manipulations of phenomena that novices would not understand.

2.3.3 Methodology for Step Five

Step five begins the development of a computational representation of the problem. We need to consider how we might address the tradeoffs resulting from Step 4 using mathematical and computational methods and tools.

Modeling paradigms potentially useful include dynamic systems theory, control theory, estimation theory, queuing theory, network theory, decision theory, problem solving theory, and finance theory. A multi-level modeling framework is used to illustrate how the different modeling paradigms can be employed to represent different levels of abstraction and aggregation of an overall problem.

Paradigms for Step 5

Table 1 summarizes the key assumptions underlying each modeling paradigm and the central phenomena predicted using these paradigms. Most of predicted phenomena in the right column of Table 1 could be applied to a wide variety of problems in many domains.

We will pursue step five in the context of the architecture of public-private enterprises. We use this architecture to organize thinking about how to use the theories summarized in Table 1 for each of the layers in the enterprise architecture. Issues and models associated with the four levels of the architecture are summarized in Table 2.

Table 1 - Modeling Paradigms, Typical Assumptions and Phenomena Predicted

Modeling Paradigm	Typical Assumptions	Phenomena Predicted
Dynamic Systems Theory	• Newton's Laws • Conservation of mass • Continuity of transport	• Response magnitude • Response time • Stability of response
Control Theory	• Known transfer function of state transition matrix • Stationary, Gaussian stochastic processes • Given objective function of errors, control effort	• Response time • Stability of response • Control errors • Observability • Controllability
Estimation Theory	• Known dynamics of process • Known ergodic (stationary) stochastic	• State estimates – filtering, smoothing, prediction • Estimation errors

	process • Additive noise inputs	
Queuing Theory	• Known arrival and service processes • Future state only depends on current state • Given service protocol, e.g., First Come, First Served, priority	• Number and time in queue • Number and time in system • Probability of balk or renege
Network Theory	• Discrete entities, e.g., agents • Decision rules of entities, e.g., agents • Typically binary relationships • Relationships only via arcs or edges	• Shortest distance between any two locations (nodes) • Shortest time between any two locations (nodes) • Propagation of sentiment among actors
Decision Theory	• Known utility functions • Comparable utility metrics • Known payoff matrix • Given voting rules	• Choice selected • Game equilibrium • Election results • Impacts of incentives
Problem Solving Theory	• Known human mental model • Known information utilization • Known repertoire of patterns • Known troubleshooting rules	• Time until problem solved • Steps until problem solved • Problem solving errors
Finance Theory	• Projected investments • Projected operating costs • Projected revenues and costs	• Net present value • Net option value • Net capital at risk

Table 2 - Levels of Modeling, Example Issues, and Potential Models

Level	Issues	Models
Ecosystem	GDP, Supply/Demand, Policy	Macroeconomic Models
	Economic Cycles	Dynamic System Models
	Intra-Firm Relations, Competition	Network Models
Organizations	Profit Maximization	Microeconomic Models
	Competition	Game Theory
	Investment	DCF, Options
Processes	People, Material Flow	Discrete-Event Models
	Process Efficiency	Learning Models
	Workflow	Network Models
People	Consumer Behavior	Agent-Based Models
	Risk Aversion	Utility Models
	Perception Progression	Markov, Bayes Models

Step 5 involves translating the model-determined representations of phenomena into computational form. There are several common computational frames.

Dynamic Systems

The dynamic systems frame is concerned with transient and steady state responses as well as stability of dynamic systems. Central constructs are stocks, flows, feedback, error, and control. Such systems are represented using differential or difference equations. Both have continuous states, but the former involves continuous transitions while the latter involves discrete transitions. Various elements of these models may be stochastic in nature, e.g., disturbances and measurement noise. A range of tools can be used to solve such equations (see below), but the essence of the computation with all approaches is the relationship between future system states and past system states over time.

Discrete-Event Systems

The discrete event frame is concerned with steady state responses in terms of average waiting time and time in the system, as well as average number of entities waiting and number of entities in the system. Central constructs are flows, capacities, and queues, as well as allocations of resources and time-based metrics. Such systems are represented using Markov chains with discrete states, and continuous transitions. Also important are probability distributions associated with arrival flows (e.g., Poisson) and service processes (e.g., exponential). A variety of tools can be used to compute the responses of discrete-event systems. As with dynamic systems, the essence of the computation with all approaches is the relationship between future system states and past system states, in this case averaged over time.

Agent-Based Systems

The agent-based frame focuses on large numbers of independent entities and the emergent responses of the collection of entities over time. Central constructs, for each agent, are information use, decision making, and adaptation over time. Such systems are represented by the sets of information sampling rules and decision making rules assigned to each agent. These systems may also incorporate network models of relationships among agents. A range of tools can be used to compute the responses of agent-based systems. The essence of all these approaches is simulation to compute the evolving state of each agent and collective patterns of behaviors. Of particular importance with such models is the notion that the observed emergent behaviors are not explicitly specified by the agents' rules.

Optimization-Based Frame

Another computational frame can overarch one or more of the above three frames. Beyond simply controlling a dynamic system, we may want to optimally control it as part of the ultimate enterprise strategy (See Section 5). Beyond predicting the queuing time in a discrete-event system, we may want to optimally allocate resources to minimize some criterion that differentially weights the queuing times of different types of entities. Beyond observing the emergent behaviors of an agent-based system, we may want to design optimal incentive systems that maximize the likelihood of desirable emergent behaviors.

Thus, the problem to be solved using the models discussed in this chapter may be determination of the optimal control strategy, the optimal allocation of resources, or the optimal incentive system. Such aspirations are pervasive in operations research and systems engineering, as well as economics, finance, etc. The Holy Grail is the "best" answer that maximizes benefits and minimizes costs.

We must keep in mind, however, that all of these pursuits must inherently solve their optimization problems in the contexts of models of the systems of interest rather than the complex reality of the actual systems. Indeed, the achievement of the "best" answer can only be proven within the mathematical representations of the system of interest. To the extent that the mathematical models of the system are good approximations, the best answer may turn out to be "pretty good." However, we always need to consider the model uncertainties that can arise in enterprise systems (Sections 4.4 and 4.5) and adjust our strategy accordingly (Section 5).

When there are humans in the system, we have to deal with not only our model of the system but also with humans' models of the system. The notion of "constrained optimality" is important here. Succinctly, it is assumed that people will do their best to achieve task objectives within their constraints such as limited visual acuity, reaction time delays and neuromotor lags. Thus, predicted behaviors and performance are calculated as the optimal behavior and performance subject to the constraints limiting the humans involved. If these predictions do not compare favorably with subsequent empirical measurements of behaviors and performance, one or more constraints have been missed [Rouse, 1980, 2007].

Determining the optimal solution for any particular task or tasks requires assumptions beyond the likely constraints on human behavior and performance. Many tasks require understanding of the objectives or desired outcomes and inputs to accomplish these outcomes, as well as any intervening processes. For example, drivers need to understand the dynamics of their vehicles. Pilots need to understand the dynamics of their aircraft. Process plant operators need to understand the dynamics of their processes. They also need to understand any tradeoffs between, for example, accuracy of performance and energy required, human and otherwise, to achieve performance.

This understanding is often characterized as humans' "mental models" of their tasks and context. To calculate the optimal control of a system, or the optimal detection of failures, or the optimal diagnoses of failures, assumptions are needed regarding humans' mental models. If we assume well-trained humans who agree with and understand task objectives, we can usually argue that their mental models are accurate, e.g., reflect the actual physical dynamics of the vehicle.

They key point is that optimizing the types of systems addressed in this report requires that we develop models of humans' model of reality. This is tractable if there are enough constraints on humans' choices and behaviors. Without sufficient constraints, however, the whole notion of optimality can become extremely ambiguous and often useless.

Most of the applications of the above computational frames have involved modeling and representation of the "physics" of the environment, infrastructure, vehicles, etc. These are certainly important elements of many overall multi-level models. However, the greatest challenges in developing such models for the types of problems of interest in this report are modeling and representation of the behavioral and social activities and performance throughout the system, especially when it cannot be assumed that the human elements of the systems will behave in accordance with the objectives and "rules of engagement" of the overall system.

Due to both the breadth and depth of the available set of representations that have been developed by the scientific, engineering, and business communities, it is effectively impossible for any enterprise modeler to be an expert in every potential representation that could be employed. Consequently, it would be very easy for one to miss a potentially useful representation or inadvertently misapply a representation. To address those issues, a useful mechanism for enterprise modelers would be a model repository that documents canonical representations for many phenomena and describes the strengths and limitations. While such a repository cannot replace true expertise, it can provide a starting point that will lead enterprise modelers to the appropriate domain experts. A working set of requirements for a model archive is described in Section 8 of this report.

2.3.4 SUPPORTING RESEARCH FOR STEPS SIX AND SEVEN

One we have identified potential representations for the phenomena of interest in step five, steps six and seven involve assessing the ability to computationally connect these representations and then establishing a consistent set of assumptions with which to do so. This is a non-trivial exercise. As discussed in Section 4.5, model composition is a topic that the modeling and simulation community has been struggling with for some time. This is particularly true when models of social phenomena are involved. Consequently, this is an area of active research.

It is not likely that there will ever be any universal approach to compose any two models. Consequently, one will only be able to compose models only for some or just portions of enterprise systems depending on the questions of interest. Ultimately, one would need some guidance as to when model composition is and is not appropriate.

As part of this research task, we investigated the nature of model composition issues in enterprise systems and developed some proposed guidelines that integrate the model composition approach with the overarching strategy approach. The findings of this investigation and the resulting modeling guidelines are documented in Section 6 of this report. Ideally, the proposed guidelines will be refined and improved through subsequent research.

3 COUNTERFEIT PARTS CASE STUDY

To illustrate a specific enterprise problem in depth, we have focused on the problem of counterfeit parts in the Department of Defense (DoD) supply chain as a case study problem to test and validate our enterprise modeling methodology. This problem involves an array of organizations ranging from DoD programs, agencies and services, to other federal agencies, to industry contractors. All are faced with the increasing problem of counterfeit parts infiltrating the DoD supply chain as replacement parts for deployed systems. Note that we are not studying a specific case study involving specific systems and counterfeit parts, but rather the broader problem of counterfeit parts.

3.1 Background

Counterfeiting is an age-old problem. Historically, counterfeiters produced either counterfeit goods or counterfeit money. The party receiving such goods or money could inspect the end-product to determine whether it was counterfeit. Good inspection procedures clearly are important in addressing counterfeit proliferation. Such proliferation typically causes economic harm to the consumer of the counterfeit goods or money, plus economic harm to the producer of the genuine articles, usually in the form of lost sales or brand debasement.

In the systems domain, though, counterfeiting has manifested itself differently. Here, the problem consists of an increasing production of and trade in counterfeit parts, or components of end-product systems such as airplanes, submarines or missiles. Thus, one cannot simply inspect the end-product system to determine whether one of its components is counterfeit. Under an inspection regimen, the individual parts themselves must be inspected before they become part of a larger system. In addition, systems with counterfeit parts may face performance or safety problems, so the problem to the consumer of such systems is not merely economic.

Counterfeit components typically are not inserted into new systems, since the original lead systems integrator takes care that all parts are supplied by an original component or equipment manufacturer (OCM/OEM). Rather, they are inserted as replacement parts as systems age. The biggest risk comes from obsolete parts no longer available from OCMs/OEMs. A counterfeit component may be embedded within several layers of sub-systems, making it difficult to detect when a sub-system is purchased for installation into an overall system. Due to the complex nature of today's advanced systems, the counterfeit component could be inserted into a sub-system well before the sub-system is inserted into the end-product system, and also at a different location such as the factory that manufactures the sub-systems.

It should be noted, however, that programs with long development times can have their sub-systems become obsolete before the overall system is deployed. Thus, there is risk for these types of programs that counterfeits can be inserted into new systems.

Concern centers around two types of counterfeiting – fraudulent counterfeits and malicious counterfeits. Fraudulent counterfeits derive from the traditional motivation of a counterfeiter to make a profit through fraud, by substituting an inferior product that is inexpensively produced relative to the cost of the genuine article. These counterfeits fall into several categories. Some parts are re-marked to appear as OCM/OEM. Defective parts can be passed as good OCM/OEM parts. Additionally, parts can be removed from scrapped assemblies and passed as new. Malicious counterfeits are designed to appear to perform correctly, but then malfunction at critical times or open security breaches so that adversaries gain advantage. Guin et al. [2014] detail a more complete taxonomy of counterfeit electronics, as well as counterfeit detection methods.

For the past decade, counterfeit parts are often electronics, such as integrated circuits. Much or even most of the counterfeit parts infiltrating the supply chain are imported from other countries [Business Insider, 2012; Economist, 2012; Pecht & Tiku, 2006; Villasenor & Tehranipoor, 2013]. Thus, there has been a substantial focus on import channels and imported goods. However, the problem of counterfeiting is increasingly occurring in other types of products, including batteries and automobile air bags.

3.2 ENTERPRISE PERSPECTIVE

Concerns about counterfeit parts, in particular electronics, have been aired for almost a decade [Pecht & Tiku, 2006]. Why is counterfeiting an enterprise problem, though? It could simply be treated as a technical problem of detecting counterfeits, eliminating them from the supply chain, and prosecuting the counterfeiters.

However, this type of counterfeiting occurs in an enterprise context. Consider that engineering of DoD's complex systems takes place in an enterprise context. For example, the Department of Defense (DoD) contracts with a prime contractor, or lead systems integrator, to design, develop and manufacture a new system such as a jet fighter, ship or unmanned aerial vehicle. This prime contractor, in turn, contracts with a host of sub-contractors, who likewise hire other sub-contractors. The government and industry organizations form a program enterprise that eventually delivers the new systems. Once delivered, the new systems require repairs, maintenance and part replacements. These are addressed by a supply chain enterprise that sustains these systems. The program enterprise essentially transitions from a design and development enterprise to this supply chain enterprise. This supply chain enterprise similarly involves contractors and DoD agencies. Many of these contractors are the same as those in the original design, development and manufacturing enterprise. However, as time progresses, some of these contractors exit the enterprise, and new ones enter.

In addition, the counterfeiting problem is addressed by policy-making and law enforcement agencies that add to the enterprise nature of the context, above and beyond the operational aspects of the program enterprise. Thus, the enterprise nature of the context has multiple decision-makers with differing perspectives and needs.

In addition to the technical aspects of counterfeiting such as production, distribution and detection, counterfeiting is also characterized by economic motivations and adaptive behavior. Suppliers and counterfeiters are motivated by economic profits. Counterfeiters may be motivated by strategic factors, if they are state-sponsored actors engaged in malicious counterfeit production. To avoid counterfeiting, the government and industry develop new production methods and features for goods that may be counterfeited, as well as new detection methods. Similarly, when counterfeiters discover that their goods are being detected, they adapt to new production and distribution methods.

Before studying the enterprise in more detail, we note that numerous contributing factors have influenced the increased occurrence of counterfeiting in the DoD supply chain. Typical of enterprise behaviors, there is no direct cause of increased counterfeiting, but rather a set of interacting influence factors. They are also socio-technical in nature, involving technical trends, business trends and socio-economic behavior [Bodner et al., 2013; McDermott et al., 2013]. These trends include the following [ABA, 2012; AIA, 2011; GAO, 2010, 2011, 2012a, 2012b; Pecht & Tiku, 2006; SASC, 2012; Stradley & Karraker, 2006]:

- Increased system complexity;
- Globalization of commerce and supply chains, especially in semiconductors and electronics;
- Globalization of DoD programs, causing inducements to use foreign suppliers;
- Outsourcing of design and manufacturing of major sub-systems by primes;
- Sub-system obsolescence caused by extended lifespan of systems and diminishment of OCMs/OEMs providing replacement parts over the lifecycle horizon (replaced by potentially unreliable independent distributors);
- Weak IP protection outside of U.S.;
- Increasing sophistication of design and manufacturing technology used by counterfeiters;
- Use of internet as a purchasing platform and its relatively anonymous nature;
- State subsidy, influence or control of potential foreign suppliers; and
- Decreased cost of counterfeits vs. genuine articles (e.g., movement toward environmentally-friendly electronics that are more expensive to produce).

We also note that DoD, in conjunction with industry, is developing a number of practices, strategies, policies and guidelines aimed at addressing the counterfeit parts problem. Many of these are new or under development; thus, it is not known how successful they will be. A summary of efforts includes [Cohen & Lee, 2014; DAU, 2013; DoD, 2011, 2012, 2013, 2014; Livingston, 2007a, 2007b; McFadden & Arnold, 2010; SAE 2013]:

- Acquisition regulations addressing supplier qualification, suspect counterfeit reporting, supplier penalties for counterfeits and pass-throughs;
- Use of a secure trusted foundry network of suppliers to reverse engineer and produce obsolete parts;
- Testing regimens to detect counterfeits at entry points in supply chain;
- Traceability of components throughout traversal of supply chain;

- Criticality analysis under Program Protection Plans to focus on parts/sub-systems deemed critical to mission;
- Industry standards for supplier qualification;
- Obsolescence management and re-engineering obsolete sub-systems; and
- Law enforcement to identify and remove counterfeiters.

To understand the enterprise context better, we engaged a set of enterprise stakeholders with a variety of perspectives from their positions within the DoD supply chain enterprise. The organizations involved included:

- DoD systems engineering – DoD systems engineering provides technical oversight for system design and development, plus technical support for anti-counterfeiting.
- DoD logistics – DoD logistics provides support for sourcing and distributing replacement parts for systems.
- DoD policy – DoD policy provides policy development for various issues, including anti-counterfeiting.
- A prime contractor – Prime contractors are ultimately responsible for counterfeit parts in operational systems. They have responsibility to flow-down rules and practices to suppliers to mitigate problems with counterfeiting.
- A component supplier – Suppliers are concerned with the integrity of their intellectual property, as well as potential safety issues involving counterfeits of their products.
- Obsolete parts manufacturing – Various companies agree to manufacture parts that the OEMs/OCMs have quit producing. These parts may need to be reverse-engineered if there is inadequate documentation on design and production, resulting in additional cost.
- An electronics industry consortia group – Industry consortia groups are concerned about issues that affect their members, and counterfeiting is one such issue for the electronics industry.
- Customs, law enforcement and counter-intelligence – These organizations detect potential counterfeits and counterfeiting organizations (typically involving imports), as well as pursue criminal investigations and prosecutions involving counterfeiting.
- Subject matter experts on counterfeit parts – There are a variety of subject matter experts in various aspects of the counterfeit parts domain.

These were largely unstructured discussions with new stakeholders often added for each session, based on discussion and relevance of the new stakeholder's perspective to the overall problem or to a topic that had been discussed. Nevertheless, these discussions served to validate and enrich the counterfeit parts model, plus determine potential users for model results. Summary findings from these roundtable sessions are captured in Table 3, Table 4, Table 5, Table 6 and Table 7.

Table 3. Enterprise take-aways

Area	Take-aways
Scope	• How does one bound the enterprise?
Stakeholder characteristics	• Services act on their own to address counterfeiting problems due to "on-the-ground" needs. • Different stakeholders have different perspectives on outcomes (e.g., regulatory versus operational). • Different agencies and different programs have differing levels of being able to invest in anti-counterfeiting efforts (wealthy versus poor stakeholders). Wealthy programs can continuously monitor their supply chains, while others rely on event-based monitoring. • Stakeholder behavior has to be studied to determine if risk is being transferred to others. • Primes are more likely to report counterfeiting incidents than suppliers in lower tiers.
Stakeholder relationships	• What are the trade-offs involved in using DoD regulations vs. having primes managing tiers and driving down terms & conditions?
Acquisition versus sustainment	• An important consideration is increasing design engineers' consideration of obsolescence for design refresh.
Risk characteristics	• Can a model detect when conditions have changed (e.g., increased risk) as opposed to when an event indicates an issue?

Table 4. Supply chain take-aways

Area	Take-aways
Tier characteristics	• Counterfeit intrusion typically occurs in tiers far away from prime. • One strategy is to identify the commercial/DoD boundary with respect to control in the supply chain and concentrate anti-counterfeiting efforts there. • Suppliers in tiers far from the prime typically lack knowledge about which components and sub-systems are critical.
Part flows	• Part flows are very complex and involve such phenomena as excess inventory and returns, which are not well-handled by monitoring. • Components and sub-systems classified as critical for one program may not be critical for another.
Evolution	• The supply chain is always changing. Assumptions that are

	good today may not be good in the future.
Design	• Supply chains are not designed to be resilient in the sense of mitigating counterfeiting (e.g., redundant suppliers, etc.).

Table 5. Economic take-aways

Area	Take-aways
Micro-economic	• Restricting sourcing via supplier qualification will have impacts on affordability and schedule. It is important that these impacts be considered in decisions. • For a relatively inexpensive part, testing can double or triple its cost. • False positives (genuine articles identified as counterfeits) can have a significant impact if they are rejected. Who pays those costs? • Suppliers are more concerned about safety impacts of counterfeiting than economic or intellectual property impacts. • Insurance is a traditional form of risk mitigation, but the threat is not well-enough understood for actuarial calculations.
Macro-economic	• It costs $8-10 billion to build a new fab. When there is a DoD need for relatively inexpensive electronic parts, the free market cannot meet that need due to investment cost. • In other countries, the state has made enormous investments to enable production. This could be used for malicious purposes. • One potential strategy for an adversary is to flood the market with counterfeits to drive up costs. • Given the size of the consumer electronics market, DoD is a small player, especially when the fragmented nature of programs is considered. It cannot address counterfeiting alone due to its market position. • Would it be possible to aggregate demand over all of DoD? • There is an inherent conflict in government policy between law enforcement (restriction of supply due to potential counterfeits) and trade and commerce (promotion of supply to reduce consumer costs). • Other industries face similar threats from counterfeiting – automotive, critical infrastructure, finance, internet-of-things.

Table 6. Counterfeiting take-aways

Area	Take-aways
Counterfeiter methods and characteristics	• Most counterfeiters are thought to be outside the U.S., with import channels through knowing U.S.-based partners. • Brokers are risk points in that they often import counterfeit parts. • Counterfeiters adapt to stay ahead of authorities. • Current adaptive behavior consists mainly of delivery channel modification rather than technology adaptation. • Counterfeiter tactics invalidate traditional quality control approaches to detection.
Counterfeit parts	• Fraudulent counterfeits are much more prevalent than malicious ones, at least in term of detections. • When supplies are tight, there tend to be more counterfeits. • When counterfeit incident reports decline, we do not know the cause. It could be that the government and responsible suppliers are winning (less counterfeiting), that the government and responsible suppliers are doing worse at detection, and/or that the counterfeiters have adapted to pass their product better.

Table 7. Strategy and policy take-aways

Area	Take-aways
Supplier qualification	• How do criteria affect supplier behavior? This is not known. • Can this effort derive insights for effective policies? • DoD is looking at industry standards for supplier qualification (e.g., SAE).
Testing and detection	• Components most vulnerable to counterfeiting are obsolete, and thus have small lot sizes. This makes statistical sampling in quality control difficult. • Testing is not 100% reliable and often is destructive. • Testing can be done based on part criticality to minimize cost.
Tracking and tracing	• DoD is moving in this direction, but it remains to be seen if this will be effective.

The critical question is what set of these strategies, policies and guidelines are best to address the counterfeiting problem, taking into account cost, effectiveness and the adaptive behavior of suppliers and counterfeiters. For example, test and detection is a standard methodology applied to identification of non-conforming product. However, as Cohen and Lee [2014] observe, obsolete parts are most vulnerable to counterfeiting, and these tend to have small lot sizes. They demonstrate that this works against the statistical effectiveness of standard test procedures. In addition, standard procedures can result in numerous false positives, not just identification of counterfeits (true positives). Testing itself can be quite expensive, and the cost of discarded good product from false positives increases the cost of a testing program significantly.

Another approach involves qualifying suppliers to those most likely to provide products without counterfeits. Of course, restricting suppliers tends to increase cost. Similarly, imposing penalties on suppliers who pass counterfeits even unknowingly or disallowing costs to remediate counterfeits can restrict the number of suppliers.

Finally, there are a number of trade-offs to be considered, such as the extent of those sub-systems deemed critical and the frequency with which sub-systems are re-engineered. Increasing the number of critical sub-systems or frequency of re-engineering reduces risk from counterfeit infiltration, but increases cost.

3.3 Applying an Enterprise Modeling Methodology

For the counterfeits parts case study, we utilize a recently specified methodology for enterprise modeling (Pennock & Rouse, 2014). This methodology has ten steps that start with the modeling question and end with the validation of model predictions.

In this report, we place significant emphasis on the first four steps that are largely conceptual in nature. Often these types of steps are glossed over in simulation modeling efforts. In a typical modeling effort for technical systems that serve the needs of one or two stakeholders, these conceptual issues may not be as important, because the answers are already well-understood. However, when modeling an enterprise problem with multiple stakeholders having different perspectives, and with socio-technical characteristics, these conceptual steps are important.

A recently updated version of this methodology is discussed in more detail in Section 2, along with the theoretical justification for the methodology. The remainder of this section describes the application of this methodology to the counterfeit parts problem. This is also discussed in Bodner [2014].

3.3.1 Decide on the Central Questions of Interest

The first step of the methodology is to decide on the central question(s) of interest. This question relates to the intended use of the model. In an enterprise problem context, this step also incorporates the perspectives of multiple stakeholders and potentially multiple uses. Thus, it may not be as obvious as for a model of a purely technical system with one or two stakeholders.

The central question for the counterfeit parts model is to determine an effective set of policies along multiple metrics to minimize adverse effects of counterfeit parts while accounting for adaptive behavior. There may not be one single policy that will address the problem. In addition, there is not one single decision point in the enterprise, so multiple actors have say over different policies. Thus, the question seeks a set of policies. This central question involves multiple questions at a lower level of detail explored in a set of trade-offs in a subsequent step.

3.3.2 Define Key Phenomena Underlying These Questions

The subject matter expert discussion was posed along a number of overarching themes, ranging from the enterprise aspects, to counterfeiting itself, to strategy and policy. These serve as a starting point for describing key phenomena. However, we want to go a step further than merely restating these items. We want to move toward synthesizing them into a useful representation from a modeling and systems engineering perspective. Therefore, we decompose the key phenomena into five primary modeling categories:

- Operational systems & constituents – this category includes the systems of interest plus counterfeit parts (Table 4, Table 6).
- Supply chain flows – this category includes the flow of parts that eventually are placed in systems (Table 4).
- Enterprise actors – this category includes the various operational decision-making organizations in the enterprise, including counterfeiters (Table 3, Table 5, Table 6).
- Policy – this category includes the various policy decision-making organizations in the enterprise (Table 3, Table 7).
- Exogenous effects – this category includes external influences to counterfeiting in DoD systems, many of which are macro-economic (Table 5).

In addition, the issues from Table 3 - Table 7 need to be enriched with other phenomena for completeness. Details of the key phenomena are shown below in Table 8.

Table 8. Key Phenomena in Counterfeiting

Category	Phenomena of Interest
Operational systems & constituents	• Work breakdown structures (major & minor sub-systems, components, etc.) • Vulnerabilities of system designs to counterfeiting • Mission profiles • System performance criteria (technical performance, availability, lifecycle cost, reliability and security)

	• Nominal system performance vs. counterfeit-induced performance • Maintenance and repair schedules • Technology upgrade policies and schedules • Configuration management • System characteristics over lifecycle • Counterfeit parts
Supply chain flows	• Globalized nature of DoD supply chain • Programs and supplier networks • Obsolete parts sources • Evolution of part flows over program lifecycle • Counterfeiting networks
Enterprise actors	• Programs and suppliers • Supplier behavior and adaptation • Supplier diminishment • DoD agencies (systems engineering, logistics & materiel readiness, policy) • Law enforcement • Counterfeiter motivations and capabilities • Counterfeiter risk and incentive behavior • Counterfeiter adaptation
Policy	• Extent of criticality analysis • Prevalence of testing • Use of tracking/traceability • Supplier qualification • Supplier penalties and disallowed costs • Obsolescence management • Law enforcement approach
Exogenous environment	• Technological progress over program lifecycle • Technology off-shoring • Threat profiles

3.3.3 Develop One or More Visualizations of Relationships among Phenomena

Previously, a number of visualizations were developed [Bodner et al., 2013]. They are largely consistent with the phenomena above; therefore, they were not redone as part of the modeling effort. They were posed along four levels of an enterprise system [Park et al., 2012]:

- The eco-system in which the enterprise operates,
- The networked structure of the enterprise as a system,

- The delivery operations by which the enterprise produces value, and
- The work practices that individuals and organizations use in delivery operations.

It should be noted that these four levels are one way to conceptualize an enterprise for modeling, and there may be alternate conceptualizations that are more useful under particular circumstances. Figure 1 illustrates the interaction between the different levels. Figures exploring details of each level are available in [Bodner et al., 2013]. Visualization is a key research thrust for enterprise modeling, since different stakeholders need ways to understand the enterprise and its complex relationships and interactions to support their own perspectives and decisions, and to understand perspectives of others. For this particular model, it would be of interest to have simpler visualization models that stakeholders can explore and then drill down to study complexities, interactions and secondary effects. Such models should be interactive. The topic of interactive visualization for enterprise modeling is discussed in Section 7.

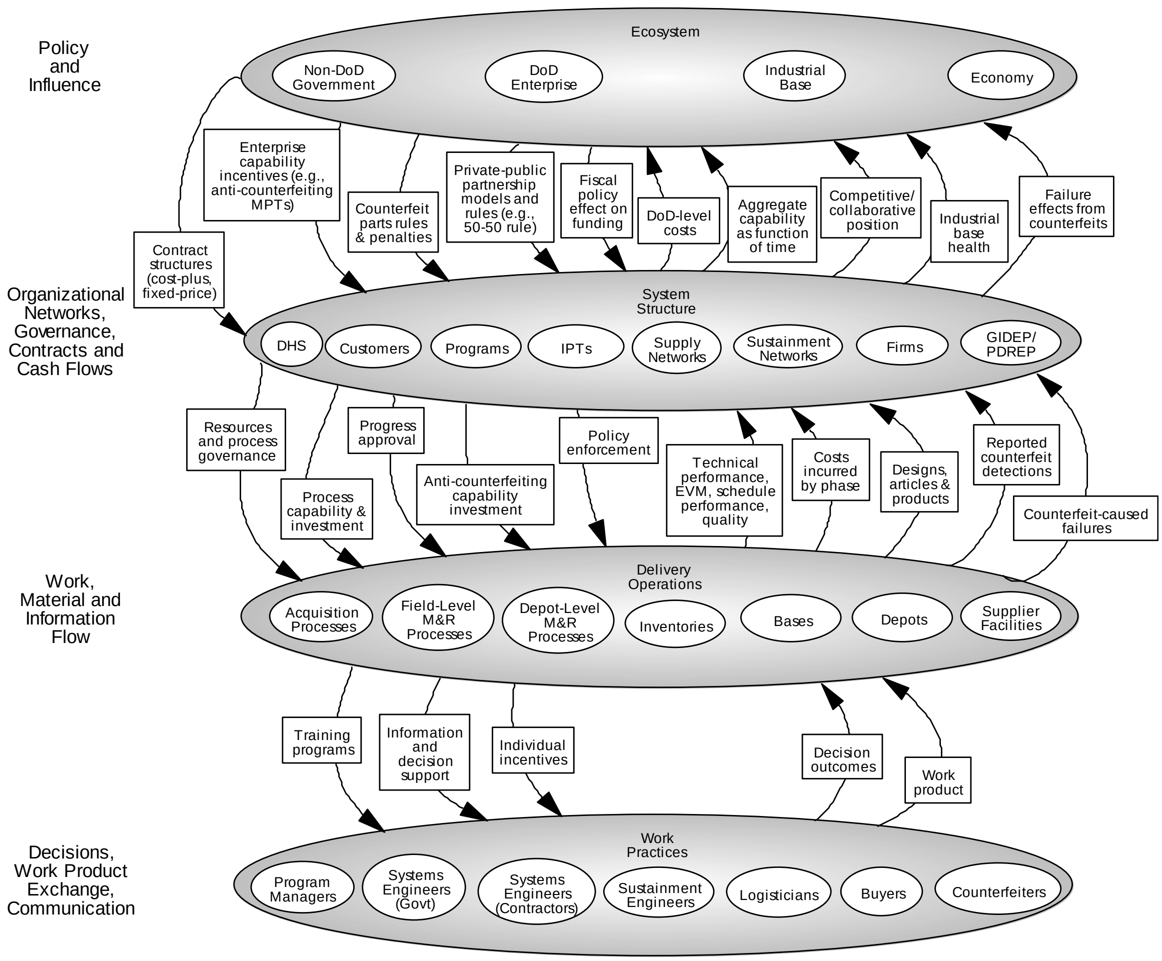


Figure 1. Interactions between different enterprise levels

3.3.4 Determine Key Tradeoffs That Appear to Warrant Deeper Exploration

From the anti-counterfeiting roundtables and literature reviews, a number of potentially interesting trade-offs were identified.

- What is the trade-off between the scope of liability and penalties for counterfeiting (including allowing pass-through counterfeits in sub-systems and systems) versus supplier availability across the program lifecycle?
- What is the trade-off between limiting foreign and/or non-trusted suppliers and the availability and cost of replacement parts in a restricted market?
- What is the trade-off between results of incrementally putting counterfeiters out of business, tolerating continued counterfeits and possibly enabling adaptation vs. waiting to put a network out of business?
- In defining critical sub-systems, what is the trade-off between the scope of the definition of criticality (i.e., wide versus narrow) and the resources needed to address that scope and performance impacts caused by that scope?
- What is the trade-off between the effectiveness of supply chain inspections for counterfeits versus costs of testing programs (including false positives) and delays caused by them?
- What is the trade-off involving cost and risk between stockpiling replacement parts (life-time buy) vs. re-engineering for new parts vs. sourcing from trusted foundry vs. buying obsolete parts?

3.3.5 Identify Alternative Representations of These Phenomena

Simulation modeling provides three primary paradigms – discrete-event (DE), agent-based (AB) and system dynamics (SD). Discrete-event models focus on events, processes that cause events, and new events triggered by executing events. Agent-based models focus on elements within a model, how they react to messages and state changes, and how system behavior emerges over time as a result of individual element behaviors. System dynamics models address rates of change, interdependencies, feedback loops and lags in system behavior. In addition, there are a variety of micro-economic models and macro-economic models that may be applied.

For the current model, we have focused on the simulation modeling paradigms. In particular, the enterprise nature of the problem, with its focus on different stakeholders and adaptive behavior, is well-suited for an agent-based representation. However, Table 9 presents more detailed reasoning for adoption of each modeling paradigm to each category. Then details of preferred representations are included in Table 10.

Table 9. Representation alternatives

Category	Representation alternatives

Operational systems & constituents	• Agent-based – AB models provide support for state transitions to model the system lifecycle, plus aggregation relationships (has-a) to model constituents. • Discrete-event – DE models provide support for systems moving through processes (e.g., operations, maintenance, repair), but limited support for aggregation. • System dynamics – SD models tend to combine detailed phenomena and provide no support for aggregation relationships. • Agent-based models are preferred due to flexibility and aggregation relationships.
Supply chain flows	• Agent-based – AB models have been used for supply chain modeling. Locations are modeled via states rather than processes as with DE models. • Discrete-event – DE models have been used extensively for supply chain modeling, since they provide excellent support for movement of parts through a network. • System dynamics – SD models can be used for supply chain modeling since they model inventory accumulation. However, the continuous representation is problematic for bulk (discrete) movements. • Agent-based models are preferred due to their general use in supply chain models and use here for system/constituents models.
Enterprise actors	• Agent-based – AB models have been used extensively to model interactions of individual units, as well as adaptive behavior. In addition, there is potential to embed micro-economic models in agents. • Discrete-event – DE models are not typically used for enterprise actor models. • System dynamics – SD models are not typically used for enterprise actor models. • Agent-based models are preferred due to their extensive use in modeling individual unit interactions, adaptive behavior and economic behavior.
Policy	• Agent-based – AB models have been used extensively to model interactions of individual units, as well as adaptive behavior. This could extend to policy units. • Discrete-event – DE models are not widely used for policy models as are AB and SD models. • System dynamics – SD models have seen extensive use for policy study, with the concept of variables being used as

	policy levers having resulting interaction effects. • Agent-based models are preferred do to the ability to represent adaptive behavior of policy-makers. Policies would be represented as variables, with the policy effects embedded in other sub-models.
Exogenous environment	• Agent-based – AB models likely require too much detail on individual elements of the exogenous environment, when the purpose of the exogenous environment is to aggregate external influences. • Discrete-event – DE models likely require too much detail, as well. • System dynamics – SD models can be used to aggregate behaviors and incorporate feedback loops, lags, etc. SD models have been used for macro-economic phenomena. • System dynamics models are preferred due to their representation of aggregate effects not requiring detail.

Table 10. Representation descriptions

Category	Phenomena of Interest
Operational systems & constituents	Agent-based model with constituents modeled as attribute objects in an object-oriented framework and with operational behaviors modeled via state-charts. Cohorts modeled rather than individual systems.
Supply chain flows	Either agent-based model of systems and constituent with locations and flows modeled via state-charts and attributes, or process-based discrete-event model with entities linked to agents representing systems and constituents. Supply network and counterfeiter network modeled to evolve over time.
Enterprise actors	Agent-based model with actors modeled as complex agents and relationships modeled by arcs (synchronized with supply chain flows for supplier relationships). Economic models embedded within agents to model supplier and counterfeiter adaptation.
Policy	Global variables set by analyst with an associated agent-based model to enable policy adaptation.
Exogenous environment	System dynamics model representing trends in technology progress, technology off-shoring.

3.3.6 Assess the Ability to Connect Alternative Representations

Primarily, the representations in Table 10 are agent-based. These discrete models operate via time-step advance. A system dynamics model of the exogenous environment can interoperate with them via such simulation platforms as AnyLogic™, where both formalisms are supported in underlying Java™.

The key is to design the interaction so that it is computationally efficient and scalable. For instance, such interactions can occur via condition-checking (e.g., a continuous system dynamics variable reaches a particular value and triggers an agent event). Such condition-checking may need to occur at each time-step. Therefore, careful selection and design of condition-checking protocols is needed.

3.3.7 Determine a Consistent Set of Assumptions

The assumptions used in the counterfeit parts model are the following:

- Agents representing enterprise actors and policy-makers react and adapt based on their knowledge of the enterprise state (as well as possible future states). In the current model, such knowledge is limited, and does not include any look-ahead capability (e.g., forecasts, etc.). This assumption limits adaptive behavior.
- Policy decision may be made at different levels of the enterprise. Interactions between different policies should emerge as effects in the model. However, certain relationships must hold, such as DoD-wide policies should override those of a particular program.
- The points at which the exogenous environment interacts with the remaining sub-models are the most relevant aspects of the external world affecting the enterprise addressing counterfeit parts.
- The existing dataset specifications reflect a reasonable set of parameters and variables that can be estimated.

These assumptions are subject to being revisited during later revisions of the model.

While not called for in the methodology, a model architecture was specified. This seems to be an output of Steps 5-7. The model architecture is described in Section 3.4.1. It should also be noted that many assumptions are implicitly made in the architecture and detailed design of the model, and these must be captured post-hoc.

3.3.8 Identify Data Sets to Support Parameterization

Due to the sensitive nature of the counterfeiting problem, there is a lack of published datasets on most elements. One source of counterfeiting data was investigated in detail, the GIDEP reporting database. This source contains reports on suspect counterfeits, announcements on end-of-life decisions for products used in DoD systems, and other useful information. However, the following limitations were noted:

- GIDEP is primarily designed for information exchange between stakeholders, not for statistical analysis that would support modeling.
- GIDEP is a self-reporting system, and not all counterfeit incidents may be reported. In addition, it reports suspect counterfeits, with care taken not to indicate actual counterfeits. This relates to the concern about the inability to deduce the cause behind a decrease in reported counterfeits based solely on the reported numbers.

Therefore, the approach taken is to construct a synthetic dataset that can be adapted with data from users. It should be noted that the synthetic dataset is designed to conform to the modeling paradigms used.

3.3.9 PROGRAM AND VERIFY COMPUTATIONAL INSTANTIATIONS

The model is implemented using AnyLogic. Description of the model is addressed in Section 3.4.

3.3.10 VALIDATE MODEL PREDICTIONS, AT LEAST AGAINST BASELINE DATA

This step will be conducted more fully as future work in conjunction with subject matter experts and potential users.

3.4 COMPUTATIONAL MODEL

This section describes the computational model for the counterfeit parts problem. The intent is not to provide a full set of model documentation, but rather to highlight key components of the model. Model documentation will be developed as part of a plan to transition the model to use by stakeholders. A preliminary version of this model is described in [Bodner et al., 2013]. The model is implemented in AnyLogic, a commercially available simulation software package that enables multi-method modeling, combining discrete-event, system dynamics and agent-based simulations.

3.4.1 MODEL ARCHITECTURE

The model architecture combines the five sub-models described previously, and is shown in Figure 2. The exogenous model is represented by the simplified systems dynamics causal diagram. Key variables influence the actor, supply chain and system models. For instance, an increasing system life trend increases the deployed lifetime of systems in the operational systems model. Similarly, technology progression impacts the generation of technologies available in the operational systems model, impacting obsolescence of currently deployed technologies. A data specification has been developed to support these relationships. Specific data to support them will be gathered from subject matter experts and from actual programs. In some instances, however, data will need to be parameterized so that the analyst can experiment with different scenarios.

The policy model includes agents that promulgate analyst-specified policies into the actor, supply chain and system models. Feedback provided from these models can influence the policy agents. For example, a policy of restrictive supplier qualification or severe penalties for counterfeit pass-through may result in too few suppliers or too much supply chain risk (e.g., sole-sourcing critical sub-systems). If this occurs, the policy agent may adapt to loosen restrictions.

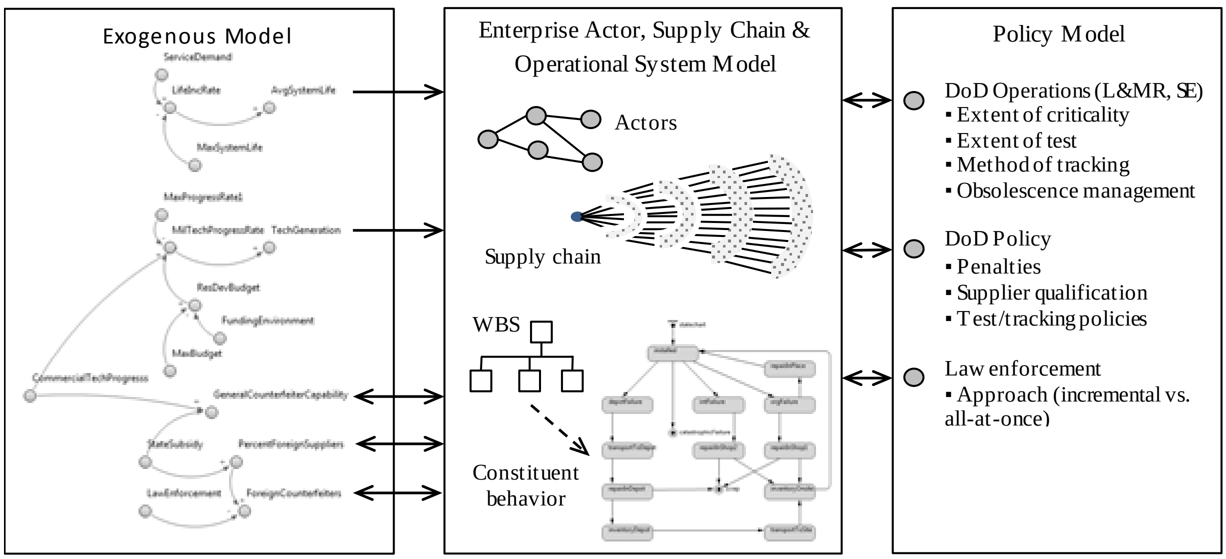


Figure 2 - Model architecture

3.4.2 OPERATIONAL SYSTEMS AND CONSTITUENTS

Operational systems and constituents address systems of interest to DoD that may experience counterfeit parts insertions. These include planes, ships, submarines, vehicles, etc.

3.4.2.1 Model Description

The operational systems model utilizes a work breakdown structure whereby the overall system contains sub-systems, which contain minor sub-systems, which in turn contain components. This is enabled by the object-oriented “has-a” aggregation relationship in AnyLogic agent objects. Sub-systems undergo maintenance, repair and replacement. Sub-systems and components are provided by suppliers in a tiered supplier relationship, such that a component supplier provides components to a sub-system supplier, which then provides a sub-system to another supplier, until it becomes a replacement sub-system to be inserted into a system. An example work breakdown structure is shown in Figure 3.

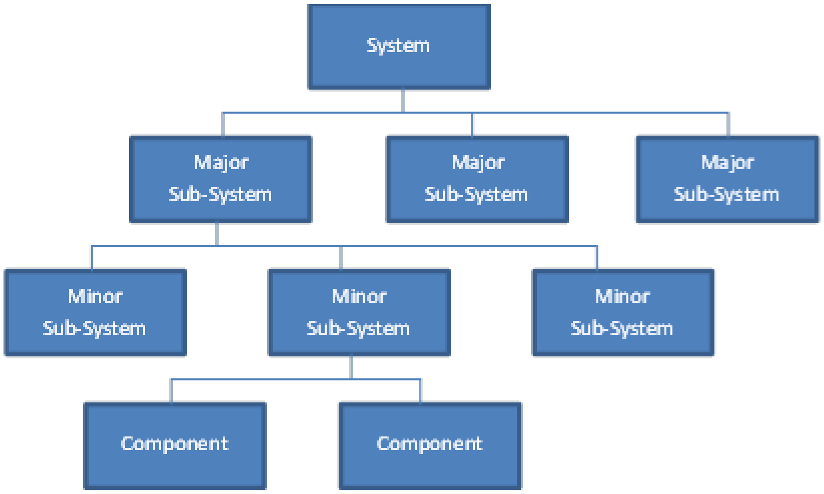


Figure 3 - Work breakdown structure

For illustration purposes, the model is based on an aircraft program for carrier-based jet fighters such as the F/A-18, where the aircraft have replaceable sub-systems termed “line-replaceable units” (LRUs). For maintenance and repair operations, these LRUs are removed from the overall system and serviced. At this point, they may be re-inserted into the system or replaced by another sub-system and put into inventory, depending on the type of repair. Without loss of generality, we will refer to replacement sub-systems as LRUs, since we assume that other types of systems may have similar part replacement operations.

The LRUs may have counterfeit components that are introduced into the LRU from unreliable suppliers in the supply chain. These counterfeit components have different reliability characteristics from genuine components, and thus the LRUs that contain them have different reliabilities than sub-systems that contain only genuine components. These reliability models are discussed in more detail in [Bodner et al., 2013]. Of course, different system designs may have different vulnerability characteristics to counterfeit parts. For instance, a work breakdown structure with redundant components or redundant sub-systems is in general more reliable than one without redundancy.

Systems and sub-systems undergo state transitions based on their repair and maintenance needs, as well as the locations at which maintenance and repair is conducted. Figure 4 shows such a state transition diagram for maintenance and repair operations of an LRU. Certain types of maintenance and repairs can be conducted only at certain types of locations. In this particular case, there are three levels of repair (organizational, intermediate and depot-level in order of increasing complexity/severity). After repairs and maintenance, the item is returned to inventory, where it will be available as a replacement. There is a chance, however, that the item may be scrapped.

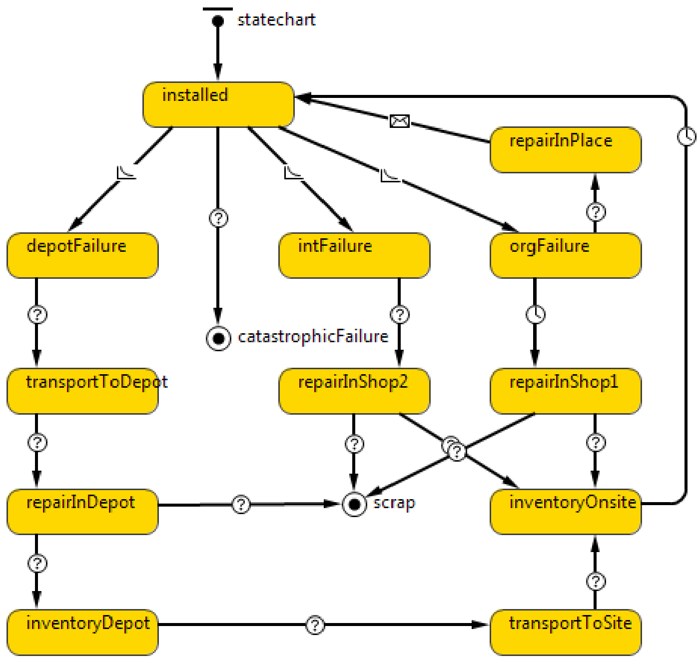


Figure 4 - State transitions

Minor sub-systems and components have their own state transition diagrams that influence the transitions of the LRU. These transitions are affected by the counterfeit status of the components.

Major sub-systems have a criticality level defined by the program. This is an integer value ranging from 1 to 5, where 1 is most critical. The criticality level is used for supplier sourcing and inspection decisions. The default is that only sub-systems with criticality 1 are subject to scrutiny in supplier selection and testing. However, having more than just a binary level of criticality allows the cut-off for criticality to be adjusted on a policy basis so that the effect of extending the number of sub-systems judged to be critical can be studied. It should be noted that components and minor sub-systems are not given a criticality level. Thus, two different major sub-systems could have the same components, sourced from the same place, and also have different criticality levels such that one is judged critical, while the other is not.

As systems age in the field, their sub-systems and components can become obsolete. This is primarily due to new technologies coming online. Extended lifetimes of systems make such obsolescence more likely to occur. Thus, sub-systems and components have an obsolescence attribute. The obsolescence of a sub-system depends on the obsolescence of its components. Some components are technology-critical, and if they become obsolete, then the sub-system is judged obsolete.

3.4.2.2 Relationship to Supply Chain Sub-Model

The operational system sub-model is strongly related to the supply chain sub-model. Here is a summary of relationships between the operational system and supply chain sub-models.

- Source locations in the supply chain model (i.e., factories) produce systems and their constituents. We are primarily interested in the production of new sub-systems and components used as replacements. Of course, some of these component source locations are counterfeiters.
- Supply chain locations in the supply chain sub-model receive and ship out systems and their constituents (warehouses and assembly factories). In addition, supply chain locations have inventories, which are collections of constituents.
- Supply chain locations are associated with a state in the state-transition of a system or constituent. For example, certain major repairs can take place only at a major repair facility. When a sub-system transition to a "major-repair-needed" state, it must be removed from the system and taken to a major repair facility. Other minor repairs may be able to be done on site at the system's current location.
- Demand for replacement major sub-systems influences demand for constituents via the work breakdown structure.

3.4.2.3 Relationship to Exogenous Environment Sub-Model

The operational systems sub-model is affected by trends in the exogenous environment sub-model. Here is a summary of relationships between the operational system and exogenous environment sub-models.

- Technology progress causes obsolescence of components, resulting in obsolescence of sub-systems.
- Trends toward longer system lifetimes increase the lifetimes of particular systems, making them more vulnerable to obsolescence in sub-systems.

3.4.2.4 Dataset Description

Here is a summary of the dataset specification for the operational system sub-model.

- Work breakdown structure
- Sub-system criticality level
- Sub-system and component obsolescence baselines
- Maintenance schedules (by type)
- Failure frequencies and repair durations (by type)
- Effects of counterfeit components on failure frequencies (by type)
- Mission intensity profiles

3.4.3 SUPPLY CHAIN FLOWS

The supply chain transports components and sub-systems to locations where they are needed to support repair and maintenance and upgrades. It should be noted that an acquisition and production supply chain supports transport of these items to a final assembly location where a new system is finished and then placed into deployment. The focus here, however, is on the sustainment supply chain that supports repairs, maintenance and upgrades. Thus, the acquisition and production supply chain is not modeled.

3.4.3.1 Model Description

The supply chain sub-model has a variety of location agents that model various nodes in the supply chain. These include the following generic locations:

- Factories – Finished systems emerge from factories ready to be deployed. Similarly, finished sub-systems and components emerge from factories ready to enter the supply chain. A factory can be OCM/OEM, contract manufacturer (CM), or trusted manufacturer of obsolete components (TMOC).
- Warehouses – Warehouses store sub-systems and components. Warehouses thus have inventory. A warehouse may be co-located with a factory if the factory ships components or sub-systems directly to a location at which a replacement
- Bases – Systems are stationed here for use in routine operations. Some repair operations can be performed there. Bases have inventories of certain sub-systems.
- Deployed system locations – Systems engage in missions at deployed locations. Some repair operations can be performed there. An example is a carrier for carrier-deployed jet fighters. Deployed locations have inventories of certain sub-systems.
- Repair facilities – Systems and constituents are taken to these facilities for maintenance and repairs that cannot be performed at a base or deployed location.
- Import points – Locations where imported components and sub-systems enter the U.S. Law enforcement (in the form of Customs) inspects imported products at these points.
- DoD program control points – Locations at which a component or sub-system becomes officially part of a DoD program. It should be noted that these are associated with warehouse or factory locations and are not a separate class of locations.

We assume that factories may be located outside the U.S., in other words those of component suppliers that are foreign. Components from such factories are transported to an import point by an importer, and then either to a warehouse or directly to a factory for insertion into a sub-system. Of course, this includes counterfeit components. It is assumed that these warehouses and sub-system factories are domestic.

At some point between the import points of foreign-sourced components and the prime contractor, there is a factory or warehouse at which the component or sub-system passes through a DoD control point, and it becomes a military part. Similarly, domestically-produced parts must pass through such as control point between their source and the prime contractor. Inspections may be done at this control point. We assume that inspections of sub-systems are not effective at identifying counterfeit components within the sub-system.

At both import points and DoD control points, inspections of components are conducted. The default behavior is a probabilistic process that identifies counterfeit suspect (i.e., true positives and false positives). For DoD control points, only critical components are tested.

One potential vulnerability in the supply chain occurs when excess inventory is sold [Livingston, 2014]. Often, it finds its way back into the supply chain, but there are no inspection systems to determine if it is in fact composed of genuine articles. While this is not currently modeled, it could be represented using the available building blocks for analyses that need to consider this possibility.

The current model would support analysis of different supply chain architectures via comparison by experiment. It may be of interest, though, to have a method to derive a suggested architecture based on given criteria. This is an area for future research.

3.4.3.2 Relationship to Enterprise Actor Sub-Model

The supply chain sub-model has significant relationships with the enterprise actor sub-model. Here is a summary of relationships between the two sub-models.

- Enterprise actors such as suppliers and the prime contractor own various locations in the supply chain (e.g., factories or warehouses).
- As suppliers enter and exit the market for particular components or sub-systems, new factories and warehouses become available, and existing factories become unavailable. Warehouses owned by suppliers exiting the market deplete their inventories.
- Enterprise actors can make decisions on purchase of system constituents that affect inventory levels, etc.
- Law enforcement agents in the enterprise actor model may decide to put an importer out of business if its suspect counterfeits hit a certain threshold.

3.4.3.3 Dataset Description

Here is a summary of the dataset specification for the supply chain sub-model.

- Demand
- Factory production rates
- Network structure (number of nodes of each type and arcs between them indicating constituent flow)
- Inventory levels
- Transport times
- Inspection probabilities parameters

3.4.4 Enterprise Actors

The enterprise actor sub-model provides the behavior of various organizations in the DoD anti-counterfeiting enterprise. The focus is on the operational aspects of the enterprise, so that the policy-making aspects of the enterprise are the focus of the policy sub-model.

3.4.4.1 Model Description

The enterprise actors are modeled as agents in the following categories:

- Programs – A program agent represents the DoD organization overseeing the sustainment of a particular type of system.
- Prime contractors – A prime contractor agent represents the firm that contracts with a DoD program for sustainment. In many cases, this is the lead systems integrator that handled acquisition.
- Suppliers – A supplier agent provides components or sub-systems for another supplier or for the prime contractor. A supplier may be an OCM/OEM, or it may be a CM. A supplier may also be an authorized seller (or franchisee with authorization from the OCM), an independent distributor, or a TMOC.
- Counterfeiters – Counterfeiters manufacture counterfeit parts, here components. The assumption is that counterfeiters are located outside the United States.
- Importers – Counterfeiters utilize importers to bring counterfeit parts into the United States. Importers also bring genuine articles from legitimate foreign suppliers into the U.S, though. Importers are independent distributors.

When a system is newly deployed, OEM/OCM suppliers provide components and sub-systems. As time progresses, these components and sub-systems move closer to obsolescence. The OEMs/OCMs exit the market, currently modeled as a probabilistic phenomenon. The eventual goal is to use an economic model of firm behavior. As OEMs/OCMs leave the market, primes must source from franchisees, contract manufacturers, or independent distributors. Note that franchisees and contract manufacturers may leave the market, as well. We assume that independent distributors are always an option. Another option is to use a TMOC. In addition to the sourcing decisions over time that gradually become more risky, a prime can elect to do a lifetime buy of remaining components. The trade-off here between risk and cost is a question of interest.

Counterfeiters gain capability in their methods via trends from the exogenous environment. These are discrete capability levels for each counterfeiter, which transition to the next level over time. If probabilities of successful detection via testing increase, then the counterfeiter will see its goods detected more frequently until it adapts by moving to the next capability level. This is a relatively simplistic model of adaptation that can be elaborated in the future. For instance, in the future, we plan to model counterfeiter adaptation to shutdown of its importers via reconfiguring import channels.

Currently, one program is modeled. However, this could be extended by adding additional program agents, as well as additional suppliers. There would likely be supplier overlap, with additional logic required to manage which programs have priority in provision of replacement parts.

3.4.4.2 Relationship to Policy Sub-Model

The enterprise actor sub-model has significant relationships with the policy sub-model. Here is a summary of relationships between the two sub-models.

- Supplier qualification causes suppliers to be excluded on a probabilistic basis, with the more risky type more likely to be excluded. Independent distributors are assumed to be excluded completely.
- Supplier penalties likewise cause suppliers to exit the market on a probabilistic basis.
- As counterfeiter capability trends upward, the baseline capability of counterfeiter agents increases, thus enabling them to elude detection via testing better.
- Law enforcement puts importers (independent distributors that import components) out of business once they hit a threshold for suspect counterfeits. This can be done on a one-by-one basis, or in a group. Future work involves modeling the effect on remaining counterfeit importers under each policy choice.

3.4.4.3 Relationship to Exogenous Environment Sub-Model

The enterprise actor sub-model has relationships with the exogenous environment sub-model. Here is a summary of relationships between the two sub-models.

- As a component becomes obsolete, OEMs and OCMs eventually exit the market.
- As the percentage trend of foreign suppliers increases, the percent of foreign component supplier agents increases.

3.4.4.4 Dataset Description

Here is a summary of the dataset specification for the enterprise actor sub-model.

- Set of enterprise actors and potential actors that can emerge during program evolution (with facilities to be specified in supply chain sub-model)
- Market exit probabilities due to obsolescence
- Counterfeiter baseline capability (range for agents)
- Probabilities that different categories of suppliers will not meet standards for qualification policy (by criticality)
- Market exit probabilities due to penalties for counterfeiting (depending on penalty level)
- Penalties for passing counterfeits (by criticality)
- Default criticality level for sub-systems
- Threshold at which an importer is closed down for counterfeit suspects

3.4.5 POLICY

The policy sub-model addresses the various policies that DoD and other players can bring to bear in anti-counterfeiting efforts.

3.4.5.1 Model Description

The following policy alternatives are currently modeled.

- Supplier qualification – One alternative is for DoD to qualify suppliers. This is the intent of such regulations as DFARS 2012-D055 (DoD, 2014). This would occur on a DoD-wide basis.
- Supplier penalties – Another alternative is to penalize suppliers, including the prime contractor, that pass counterfeit parts. This would occur on a DoD-wide basis.
- Criticality specification – A third policy option is to increase the level at which sub-systems are deemed critical (i.e., resulting in an increase in the number of sub-systems deemed critical in a particular system). This may be done on a program level or on a DoD-wide level.
- Law enforcement approach – It is assumed that law enforcement does not go after foreign counterfeiters due to jurisdictional issues. Rather, law enforcement concentrates on importers. Law enforcement may use one of two broad approaches. It can put importers out of business as they are identified, or it can wait to roll up a number of importers. This is independent of DoD policies.

Note that we distinguish between DoD-wide policies, program policies and law enforcement. In the enterprise actor model, specific programs may utilize policies within their spheres of influence. For instance, a program may decide to purchase obsolete parts from a firm in the trusted network that manufactures parts no longer made by an OCM/OEM. Likewise, a program may decide what level of testing is to be used. On the other hand, DoD may decide supplier penalties for passing counterfeits on a DoD-wide basis.

Each policy option is modeled as a variable that can be set by the user of the model. The variables are binary currently (except for criticality level), and they interact with model logic in the other sub-models for implementation of the policy. For example, if a policy variable is "turned on," it will trigger execution of logic that implements the policy and any consequences. If not, the default behavior of the sub-model will occur.

The following policy actors are modeled as agents in the policy sub-model.

- DoD (which controls supplier qualification, supplier penalties and DoD-wide criticality specification)
- Law enforcement (which controls law enforcement aspects of anti-counterfeiting)

Programs that implement their own policies are technically part of the enterprise actor sub-model, but the policy sub-model can be used to override default policies that the programs would follow.

It should be noted that policy agent adaptation is not currently modeled. That will be future work as the model continues to be validated by subject matter experts and potential users.

3.4.5.2 Dataset Description

The policy sub-model does not contain a dataset other than the variables that dictate policy options selected. As the adaptive nature of policy agents is explored, there likely will be a data specification to support this phenomenon. In addition, parameters to support policy specification will likely be added in the future (e.g., a set of penalty amounts for different levels of counterfeit passing).

3.4.6 Exogenous Environment

The exogenous environment sub-model addresses eco-system characteristics that affect the enterprise addressing the counterfeit parts problem.

3.4.6.1 Model Description

This model uses a system dynamics formalism to model trends over time that affect the DoD enterprise addressing counterfeit parts. These trends include the following:

- System lifetimes are increasing. More and more, systems continue to be used well beyond their intended lifetimes.
- Technological progress results in new generations of technology that can be used in systems. On the other hand, this progress impacts existing systems by making sub-systems obsolete over time. Also, suppliers eventually exit the market for obsolete sub-systems.
- Technological progress also enables counterfeiters to develop better methods to avoid detection. In the case of foreign counterfeiters, this is also enabled if there is state subsidy of technology development.
- For many years, off-shoring has led to an increase in the percentage of foreign suppliers, especially for electronics. This has been facilitated by state subsidy of industry.

These trends are modeled using the stocks-and-flows representation provided by system dynamics. Currently, the exogenous environment model is not very complex. Future work may involve extending it with the following phenomena:

- Macro-economic trends that affect the affordability of systems and sustainment
- Threat profiles that affect field operations and performance of deployed systems
- Geo-political trends that affect the extent of malicious counterfeiting
- Specializations of such phenomena as technological progress and system lifetimes so that different rates are applied to different categories of technologies and systems.

3.4.6.2 Dataset Description

The dataset for the exogenous environment is shown in Table 11.

Table 11. Exogenous environment sub-model dataset

Model Aspect	Data Elements
System lifetimes	Average baseline system lifetime, rate of change for average baseline system lifetime, service demand, ceiling on lifetimes
Technology progress	Technology level, rate of technological progress, military technology progress, commercial technology progress, R&D budget, funding environment
Foreign suppliers and counterfeiters	Percent foreign component suppliers, foreign counterfeiters, state subsidy, law enforcement level
Counterfeiter capabilities	General capability, commercial technology progress, state subsidy

3.5 Summary

This section has described application of an enterprise modeling methodology for the problem of addressing counterfeit parts in the DoD supply chain. A first-generation model has been developed that incorporates many of the features of the enterprise that addresses anti-counterfeiting and also includes the counterfeiters. Other features have been identified for future work.

This model composes system dynamics and agent-based modeling. The AnyLogic platform enables this in terms of a common model. However, there are computational issues to be studies as the model is scaled up. Agent-based models are discrete, and they operate via time-step advance. System dynamics models are continuous in nature. The key is to design the interaction between the two so that it is computationally efficient and scalable. For instance, such interactions can occur via condition-checking (e.g., a continuous system dynamics variable reaches a particular value and triggers an agent event). Such condition-checking may need to occur at each time-step. Therefore, careful selection and design of condition-checking protocols is needed.

As noted in Section 5, the counterfeit parts case study model could be used to demonstrate the strategy framework. Here, the strategy is heavily dependent on how far along a program is in its sustainment lifecycle.

Finally, next steps involve engaging stakeholders to review the current model and determine additions and modifications to suit requirements as a decision-making tool.

4 Theoretical Issues with Representing Enterprise Systems

While it is fairly well known that an enterprise system is more challenging to "engineer" than a traditional technical system, mitigating those challenges requires understanding the aspects of enterprise systems that lead to the associated difficulties. Only then can one devise methods and tools to address the challenges of enterprise systems. To that end, this section will discuss the differences between enterprise systems and traditional engineered systems. Next, we will consider how those differences lead to difficulties with modeling enterprise systems. These difficulties will be illustrated with a set of example enterprise systems. Finally, we will discuss the implications for the methods and tools required to mitigate these difficulties.

4.1 Enterprises As Systems

One way to view an enterprise system is to think of it as a hybrid between a social system and a traditional technical system. Thus, a natural way to consider an enterprise system is to begin with an examination of traditional engineering systems and then consider how they are impacted by adding a social layer to these systems.

Traditional engineering is purposeful. A system is engineered to provide desired functions or services. The system takes in resources and energy and produces a desired output, but it also produces waste that must be disposed. For it to be useful, a system should reliably provide these functions or services. In other words, for a given set of inputs and states, the system should always produce the intended outputs with acceptable levels of waste.

Ideally, we would like to engineer a system to be as efficient as possible, but environmental uncertainties force us to trade some of that efficiency in order to make the system robust. Prediction and control are central to striking this balance. The engineer designs the structure of the system to achieve the desired input/output relationship. The only way he or she can do so is if he or she can predict the behavior of the system in its operational environment. Note that prediction does not necessarily mean predictive modeling. System testing and design heuristics based on past experience can also be used to achieve a level of confidence in system performance. The critical point is that the engineer intentionally restricts the degrees of freedom within the system so that its behavior does not spontaneously change during operation. To do otherwise would be self-defeating.

An enterprise system shares several similarities with a traditional technical system:

- Enterprises take in resources and energy and produce desired outputs along with waste
- Enterprises are engineered to produce outputs in the most efficient way possible
- Enterprises operate in an uncertain environment so enterprise engineers must determine the appropriate trade between efficiency of production and robustness to uncertain events

These similarities have led several researchers to view enterprises as systems.

There are at least two broad perspectives with which one can view enterprises as systems. Architecting of enterprises is the perspective that is closest to systems engineering. This view focuses on analysis and design of functions, structures, and processes that enable an enterprise to deliver its products and services. Also of concern are modeling, simulation and visualization of the complex systems and networks associated with the architectural view [Gharajedaghi, 2007; Giachetti, 2010].

The other perspective, more social in nature, focuses on the management of enterprises [Rouse, 2005, 2006]. Here the concern is with understanding value from the perspectives of markets and competitors, developing value propositions for the enterprises' offerings, designing business processes associated with these offerings, and designing an overall organization. This view also focuses on how the enterprise learns and sustains knowledge and skills, including how people's time is allocated and managed across all of the above. Finally, this view pays careful attention to necessary investments in enterprise transformation.

This suggests that there are several key differences between an enterprise and a purely technical system:

- Enterprises must maintain and evolve themselves
- Enterprises have a significant social component
- Enterprises must operate in a societal ecosystem

When we consider an enterprise, we must introduce a social layer into the picture. A social system is almost the diametric opposite of an engineered system. It serves a multiplicity of purposes. It does not have any particular design. It does not have any particular outputs. The people within it can have an enormous number of different intentions and interactions, and the system can reconfigure itself. As a result, it has a large number of degrees of freedom, and its behavior can spontaneously change. In short, it is what many would call a complex system. Since a social system has no particular unified goal on its own, another way to look at an enterprise is to view it as the imposition of an engineered system onto a social system to control or at least influence its behavior.

The chief distinction is that, within an enterprise, the social system is far more adaptable than the engineered system. Ashby's law of requisite variety is instructive here. In the engineered system, we intentionally limit the number of degrees of freedom to achieve predictable behavior, but then we use this engineered system as a control mechanism for a social system with an enormous number of degrees of freedom. However, the law of requisite variety tells us that a control mechanism must have at least as many degrees of freedom as the system it controls [Ashby 1956]. As a result of this misalignment, the social system can adapt to the engineered system in unexpected ways to which the engineered system is unable to respond.

Yet this same adaptability also lends resilience to the enterprise system. The social component can alter, expand, and update the engineered component in response to experienced or anticipated events. Thus, the challenge of the enterprise engineer can be summarized as follows: The enterprise engineer would like to design the enterprise to produce outputs as efficiently as possible. However, the enterprise operates in a larger societal ecosystem that can introduce significant changes and perhaps shocks. Consequently, some of the enterprise resources must be allocated to making the system robust enough to survive changes and shocks and resilient enough to recover and adapt. Of course, determining the right mixture requires some degree of prediction, but the complexity of social systems limits our ability to do so. In the next section we will consider this dilemma in more depth.

4.2 SOCIAL SYSTEMS AND COMPLEXITY

In this section we will consider the assertion that the complexity of social systems introduces epistemic limitations that restrict an enterprise engineer's ability to predict the impact of his or her enterprise design decisions. The term complexity is widely used in the scientific, engineering, and business communities. But as Alderson and Doyle [2010] note, its usage by different communities of researchers can be diametrically opposed. To that end, we need to discuss just what is meant when we say a system is complex.

4.2.1 THE USAGE OF COMPLEXITY

Alderson and Doyle [2010] present a two-dimensional matrix of complexity (evolved from earlier work by Weaver [1948]). The first dimension indicates whether the system is described by a simple or a complex model. The second dimension indicates whether the system exhibits robust or fragile behavior. This results in four categories:

- Simplicity : simple model, robust behavior
- Organized complexity: complex model, robust behavior
- Disorganized complexity: simple model, fragile behavior
- Irreducible complexity: complex model, fragile behavior

In short, disorganized complexity describes the aggregate behavior of a large collection of randomly interacting entities while organized complexity describes the behavior of a large, highly organized collection of interacting entities. Examples of disorganized complexity would be the swarming behavior of insects or wave patterns in traffic. Examples of organized complexity include biological organisms or complex engineered systems. Systems that exhibit disorganized complexity may be described with a simple model but exhibit fragile behavior whereas systems that exhibit organized complexity may exhibit robust behavior but require a complex model to describe them. Alderson and Doyle assert, reasonably, that misdiagnosis of the type of the complexity exhibited by a system can result in erroneous conclusions about that system.

Another way to state the terminology problem is this: Modern science and engineering are predicated upon creating a compact description of a system that we can use to make accurate predictions about the future state of that system. The more compact the description and the more accurate the resulting predictions, the more useful the description. A description that is not compact is uneconomical in terms of making the prediction. It requires too much time and/or resources to obtain the prediction. A description that is inaccurate is uneconomical in terms of the response. The greater the spread of possible outcomes, the greater amount of resources and time we need to devote to covering them. Whenever we have difficulty achieving either of these objectives for a particular system, there is a tendency to refer to that system as complex.

Even so, some have argued that complex systems may exhibit regularities that can be exploited. For example, Boisot and McKelvey [2011] assert that complex systems can exhibit scale-free regularities that may be more expensive to exploit than traditional reductionist regularities because they require an adaptive response, but doing so can be better than guessing or waiting. However, Alderson and Doyle [2010] argue that scale-free concepts apply to disorganized complexity but not organized complexity.

In a similar vein, Poli [2013] makes a distinction between complex and complicated systems. He argues that complicated and complex systems are not differences of degree but are instead two different types of system that require different approaches. Decision makers often confuse the two and apply the wrong approach.

Thus, the problem with the term complexity is that there can be different reasons for why we cannot achieve compactness and predictive accuracy. When the blanket term complexity is applied for all of these cases, it leads some to mistakenly apply a method developed to address one cause of complexity to a system that is complex for a different reason.

4.2.2 THE IMPLICATIONS OF COMPLEXITY

Enterprises, as we will subsequently argue, likely exhibit a mixture of complexity types. Thus, as we move forward, it will be important to try to identify the underlying issue so that the right approach can be applied to the right problem. Even so, the implication of all forms of complexity is that there is a resulting epistemic limitation that must be explicitly recognized. Complexity limits what we can know about a system. Proceeding under the assumption that we know when we do not can result in counterproductive actions.

It is worth discussing some of the issues that can result from the various forms of complexity, particularly those that impact prediction and control. While a complete survey of the complexity literature is outside the scope of this report, what follows are some representative examples provided with the aforementioned caveat that these implications may be the result of one or more different types of complexity.

Casti [2012] provides a broad summary of the consequences of complexity with the following complexity principles:

- *Emergence:* Properties emerge at the system level that are not present at the level of the components
- *The Red Queen Hypothesis*: Competitive systems must keep evolving just to avoid extinction
- *No Free Lunch*: Generally speaking, the more efficient you make your system, the more susceptible it becomes to shocks

- *The Goldilocks Principle:* Complex systems operate on the edge of chaos: Not enough chaos and the system is frozen in place. Too much chaos and the system is destroyed.
- *Unpredictability/Incompleteness*: There are events that cannot be predicted by logical deduction. The structure of the system is not sufficient to predict all possible outcomes
- *The Butterfly Effect*: A seemingly insignificant event can cascade through the system with dramatic effect
- *The Law of Requisite Variety*: If the underlying system is more complex than the control mechanism, then control will be lost.

These principles are related to our discussion in three ways. First, they imply that complex systems are constantly adapting and changing. This makes it difficult to model them, and it makes it difficult for engineered control systems to keep up. Second, "optimizing" a system in the face of complexity may be the worst thing that one can do. Third, there are epistemic limitations. No theory or model will be able to predict all possible outcomes.

Helbing and Lämmer [2008] note that complex systems may exhibit:

- *History-dependence*: Starting the system with a different set of initial conditions may cause it to end up in a different state
- *Multiple local optima*: This means that it may be extremely difficult to find the global optimum
- *Instability and cascading effects*: When the system enters a critical state, failures may ripple through the system
- *"Guided self-organization is better than control*" [Helbing and Lämmer, 2008: 7]: The system may resist deliberate attempts to force change yet change radically following a seemingly insignificant event. Consequently, traditional control may be ineffective, and it may be better to try to guide the system's natural tendency to adapt.

Essentially, Helbing and Lämmer assert that (from a disorganized complexity standpoint) traditional prediction and control are largely futile. It is difficult to predict what the system is going to do. It is difficult to determine the best possible state for the system to be in, and it would be difficult to push the system to that state even if one knew what it was.

In contrast, when a system exhibits organized complexity, Alderson and Doyle [2010] assert that complexity results from attempts by the system to achieve robust behavior in response to a changing environment. In particular, "most of the complexity in highly engineered or evolved systems is in control processes that regulate the internal state and respond to external changes." [Alderson and Doyle 2010: 842] As a result, these systems tend to be "robust yet fragile" in that they can respond effectively to a wide range of events but are vulnerable to attacks on their signaling and control systems [Alderson and Doyle 2010; Doyle and Ceste, 2011]. A good example of this phenomenon is the emergence of cancer, which results from a disruption in the body's signaling and control system that suppresses uncontrolled growth.

Finally, Poli [2013] asserts that the difference between a complicated system and complex system is that in a complicated system, a problem can be associated with a particular cause. As a result, one can develop a permanent solution. For a complex system, problems result from multiple causes that interact. There is no way to identify a single cause. Consequently, one cannot control a complex system and, instead, one can only influence it.

What we can draw from the preceding is that whenever a system is complex (whatever the reason), modeling, prediction, and control become difficult. Consequently, traditional engineering approaches become problematic at best and counterproductive at worst. From an enterprise standpoint, the complexity that we are most interested in is the complexity of the social system, both in terms of the internal social system within the enterprise bounds and the external societal system that exists within the enterprise's ecosystem.

4.2.3 SOCIAL SYSTEMS ARE COMPLEX

Asserting that social systems are complex is hardly an original or controversial position. Rather our interest here is in understanding the implications of that complexity for enterprises. The interesting feature of social systems is that they exhibit aspects of both disorganized and organized complexity. Herding behavior and idea propagation are examples of disorganized complexity, yet human developed organizations, governments, and enterprises often evolve to exhibit organized complexity.

In his study of societal collapse, Tainter [1988] contends that a complex society is essentially a problem solving organization. However, each new problem it solves requires the addition of a new layer of complexity. Eventually, society becomes so strained by the burden of its own complexity that it becomes vulnerable to crises it would have previously found manageable. Enterprises may be no different in that they similarly avoid discarding previous investments.

Harvey and Reed's work, "Social Science as the Study of Complex Systems" [1996] examines the implications of social complexity on the study of social science. They postulate a layered, ontological hierarchy (Table 12) of a social system that starts from the "regularities of the physical universe" at the bottom to "societal evolution via historical modes of production" at the top. In essence, the complexity increases as we move up the layers of the hierarchy. As Harvey and Reed note, this results in "several epistemological breaks as we move along this abstraction dimension" [p.308].

Table 12 – Hierarchy of Complexity in Social System (Adapted from Harvey and Reed [1996])

Hierarchy of Complexity in Social Systems
14. Process II: Societal evolution via historical modes of production
13. Process I: Class struggle. Conflict over cultural hegemony
12. Values II: Hegemonic culture and subcultural bases of resistance
11. Values I: Struggle of hegemonic vs. subterranean world views
10. Norms II: Allocation of relative power among social institutions
9. Norms I: Personal conformity to general hegemonic standards
8. Roles II: Intraorganizational allocation of roles and resources
7. Roles I: Distribution of material rewards and esteem
6. Facilities II: Technical division of labor in productive spheres
5. Facilities I: Sociotechnical infrastructure of organization
4. Ecological organization of institutional time and space
3. Ecological organization of local biotic community
2. Biological evolution as a series of assisted bifurcations
1. Determinant regularities of the physical universe

This recognition is critical to the epistemic dilemma previously described. Any social system that we will deal with can be legitimately considered at multiple layers of abstraction. Furthermore, the behavior of the social system at each layer may be critical to the operation of the enterprise. Unfortunately, the behavior of the social system becomes increasingly complex as we move up the layers, and this limits what we can know about the system. Consequently, it becomes increasingly difficult to model and make precise predictions. It also means that a method we use to model the physical universe may be incompatible with a method to model the social universe. We should note that this inconsistency does not exist in the universe itself. Rather it is the direct consequence of the necessary but incompatible simplifications required to accommodate the epistemic limitations introduced at each layer.

Recognizing these limitations, Harvey and Reed introduced levels of modeling abstraction that map different modeling approaches to their appropriate ontological level. These include: predictive modeling, statistical modeling, iconological modeling, structural modeling, ideal type modeling, and historical narratives. They note that predictive modeling is really only appropriate on the lower levels of the ontological hierarchy. Given the importance of prediction to engineering, this would seem to be a significant inhibitor to engineering an enterprise unless we accept that the goal of design is to influence rather than control.

It is also interesting to note that while we might reasonably argue that the lower and middle levels involve mixtures of organized and disorganized complexity, the upper levels seem to approach irreducible complexity. They can only be described by a historical narrative, which is essentially a verbal description of what appears to have happened without any real prediction of future states.

To illustrate the impact of social complexity, consider a social system that is extremely important to most enterprises, the economy. Here, the work of the economist W. Brian Arthur in instructive. Arthur contends that the economy itself is a complex adaptive system [Arthur 1999]. Neoclassical economics assumes that the economy consists of only negative feedback loops that force economic agents to make decisions that lead to static equilibria. However the real economy also contains positive feedback loops. Consequently, neoclassical economic models can lead to inaccurate predictions. More importantly, many economic phenomena become path dependent. The result is that outcomes can be intrinsically unpredictable because they become sensitive to small changes in inputs.

For example, Arthur found that in circumstances with increasing returns, technology selection becomes path dependent and inferior technologies can actually be "locked-in" by incidental events [Arthur 1989]. Another study that employed an agent based model of the stock market found deviations from the standard rational expectations hypothesis and was able to replicate bubbles and crashes similar to what is seen in the real stock market [Palmer, et. al. 1994; LeBaron, et. al. 1999].

In an analysis directly applicable to enterprises, Christen, et. al. [2008] considered different control strategies at both the firm and economy level. Their results suggest that an over emphasis on efficiency may actually reduce the robustness of a firm by limiting the flow of information in the social network. They also found that, at the macroeconomic level, complex control schemes may actually exacerbate instabilities in a market versus a fixed-limiter control mechanism.

Kempf [2008] considers the illusion of control in managing enterprises. While traditional optimization approaches can be used to maximize the efficiency of production processes, an inability to forecast market behavior accurately limits the gain from this optimization. As the complexity of an enterprise increases, prediction becomes even more difficult, and decision makers may mistake noise for signal, resulting in wasted effort.

Thus, it would seem that most enterprises are embedded in a highly unpredictable social system, the economy. Adding cultural and political phenomena to the picture are unlikely to improve the situation.

Recognizing the difficulty of management decisions under such circumstances, Snowden and Boone [2007] developed the Cynefin framework to guide leaders in shifting their decision making approach depending on the circumstances. This framework consists of five domains: simple, complicated, complex, chaotic, and disordered. Leaders should alter their approach to problem solving depending on which of the domains they find themselves in.

While not phrased this way by Snowden and Boone, their framework essentially advocates adjusting strategy based on the epistemic limitations of the current situation. The idea of shifting strategy based on the context is certainly applicable to enterprise engineering. In particular, determining whether one is dealing with organized or disorganized complexity is critical. As we will consider in the subsequent sections of this report, it is likely that one will have to implement a mixture of strategies to engineer and maintain an enterprise.

4.3 Examples Of Enterprise Systems

It is very difficult to discuss enterprise systems in general. A major difficulty is that context matters. In the absence of context, the discussion is, at best, rather abstract. This section elaborates six examples of enterprise systems that were carefully chosen to stretch the overall intellectual framework presented in this report. [1]

We first discuss deterring or identifying counterfeit parts in aerospace and defense systems. In this case, the systems of interest were engineered, as was the organizational system for procuring these systems. The second example concerns financial systems and the bursting of bubbles. The investment products of interest were engineered or designed, as was the context of investment, although the context was not typically thought to be an example of engineering.

The next two examples focus on cities. First, we consider human responses to urban threats (e.g., hurricanes) and urban resilience. Then, we focus on one specific urban system in the context of traffic control via congestion pricing. In both cases, we have engineered networks of urban infrastructure embedded in the complex behavioral and social contexts of contemporary cities. For these examples, much less is designed in a formal sense. Many phenomena are emergent.

The final two examples address healthcare. First, we address the impacts of investments in healthcare delivery and how payment schemes affect investments and consequent health outcomes. We then address a particular health threat --human biology and cancer. This is an enterprise system in that the biological system succeeds or fails in the context of human lifestyles and environmental risks and consequences. Overall, these examples range from broad socioeconomic systems to individual humans functioning in a broader context.

Deterring or Identifying Counterfeit Parts

[1] An earlier version of the analysis of these six enterprise examples was presented in Pennock and Rouse [2014b].

Thousands of suppliers provide millions of parts that flow through supply chains to subsystem assembly and then final assembly of the overall system. Performance and reliability of these parts determines performance and availability of the overall system to serve its intended purpose, e.g., transportation, defense, etc. Downward pressures on suppliers' pricing of parts potentially undermine suppliers' profit margins, motivating them to cut costs somewhere. Leaning of materials and production costs reaches diminishing returns for one or more suppliers, which causes them to intentionally decrease quality of parts. Counterfeit parts are detected by increased and tightened inspection and/or inhibited by economic incentives for suppliers, both of which exacerbate cost problems.

Financial Systems and Bursting Bubbles

Demand for high-quality investments exceeds available supply. The financial sector is incentivized to increase supply by selling low-quality investments. Financial engineers create new derivatives that combine previously high-risk investments in a way that is intended to reduce the overall risk based on the assumption of low correlation among assets. These derivatives are sold as high-quality investments. This lowers the cost of capital to previously low-quality investments. This results in over-investment in low quality assets, which increases prices and consequently returns on investment for asset holders. Demand for additional low-quality assets increases. As the demand for low-quality assets exceeds supply, suppliers lower minimum standards and/or create fraudulent assets. The resulting positive feedback loop creates an asset bubble that introduces a systemic risk that increases the underlying risk of the "high-quality" derivatives, i.e., the bubble increases correlation among the assets. Eventually, the lowest quality assets begin to default, causing a chain-reaction resulting in a crash of financial markets.

Human Responses and Urban Resilience

A projected storm surge leads to predictions of flooding within a specific urban topography. Projected flooding leads to anticipated deterioration of infrastructure for transportation, energy, etc. Projections are communicated to inhabitants and subsequently communicated among inhabitants, resulting in altered perceptions. Perceptions (and later experiences) of impending deterioration lead people to adapt by planning to move to higher ground or to leave the area. Plans are shared among inhabitants, resulting in altered intentions. Intentions to move or leave enable projections of demands on urban infrastructure. Projections result in altered communications to inhabitants as well as among inhabitants. The results can range from resilient responses to complete gridlock.

Traffic Control via Congestion Pricing

Congestion in particular urban areas causes increased transit times in these areas. Time-varying time-unit pricing is adopted for use of these roads. Government likes the revenue. Business in these areas may be concerned about loss of traffic. Motorists respond by avoiding these areas and using other roads or modes of transportation. Increasing demands for alternatives affects congestion in these areas. Motorists communicate with each other in search of shortcuts and avoiding tolls. Thus, flows affect pricing, and pricing affects flows, with no guarantee of equilibrium.

Impacts of Investments in Healthcare Delivery

Demand for services (e.g., chronic disease care) and payment models (e.g., by Medicare) drive investments in capacities to provide services by healthcare providers. Capacities in the form of people, equipment and facilities are scheduled to meet demands. Use of capacities as scheduled results in outcomes, costs, and revenue. Quality of outcomes results in decreased demands for some services (e.g., reduced Emergency Department visits and in-patient admissions), but increased demands for others (e.g., out-patient chronic disease management). More subtly, decreased capacities to care for diseases with low payments can cause increased prevalence of other diseases – for instance, poor care for early diabetes mellitus leads to increases in coronary heart disease.

Human Biology and Cancer

Human genes express proteins that result in 50 trillion cells, with several hundred distinct types, that compose tissues that, in turn, compose organs, muscles, etc. within cardiovascular, pulmonary, vestibular, etc. systems. The nervous system, a network of specialized cells, coordinates the actions of humans and sends signals from one part of its body to another. These cells send signals either as electrochemical waves traveling along thin fibers called axons, or as chemicals released onto other cells. Signaling aberrations result in dysfunctions in the control of cellular processes, e.g., cell growth and death, resulting in diseases such as cancer. Targeted therapies, e.g., signal transduction inhibitors, can be used to treat cancers that result from aberrations to signaling pathways involved in cell growth. Cancers evolve in how they react to therapies.

Comparison of Examples

Table 13 compares the six examples in terms of historical narratives, ecosystem characteristics, organizations and processes, and people or basic elements. The framework of Harvey and Reed [1996] influenced the terminology chosen for this tabulation.

Table 13 - Comparison of example enterprise systems

Levels of Phenomena	Counterfeit Parts	Financial System	Urban Resilience	Congestion Pricing	Healthcare Delivery	Human Biology
Historical Narrative	Evolution of aerospace / defense ecosystem in terms of decision processes and incentives	Evolution of financial ecosystem in terms of investment instruments, regulations, etc.	Evolution of urban ecosystem in terms of social development, communities and neighborhoods	Evolution of transportation ecosystem in terms of technologies, demographics & expectations	Evolution of healthcare ecosystem in terms of ends supported and means provided	Evolution of humans in terms of genes, proteins, cells, tissues, organs, systems and signaling
Ecosystem Characteristics	Aerospace / Defense ecosystem – norms, policies, values and supplier economics	Financial ecosystem – what is assumed, allowed, illegal, and enforced	Urban ecosystem – norms, values and elements of social resilience	Transportation ecosystem – norms, values & expectations of convenience	Healthcare ecosystem – norms, values and resource competition	Human biological ecosystem, including factors such as lifestyle and environment
Organizations & Processes	System assembly and deployment networks and controls; test and evaluation	Commercial and investment banks, mortgage companies, and regulatory agencies	Urban infrastructure networks and flows -- water, food, energy, and people	Transportation infrastructure networks and flows, and control systems	Provider, payer and supplier organizations – investments, capacities, flows, outcomes	Cardiovascular, pulmonary, digestive, nervous, reproductive, et al. systems
People or Basic Elements	Flow of parts in supply chain to assembly and deployment	Investors, financial engineers, traders, and homeowners	Peoples' evolving perceptions, expectations and decisions, as well as shared beliefs	Individual vehicles and driver decision making in response to flows and controls	People's health and disease incidence, progression and treatment	Cellular processes and signaling mechanisms; therapy decisions

From these comparisons, we can see that these six examples have several common characteristics:

- All involve behavioral and/or social phenomena, directly or indirectly
- All involve effects of human variability, both random and systemic
- All involve economics (pricing) or financial consequences
- All include both designed (engineered) and emergent aspects

There are also important distinctions:

- Counterfeit Parts and Financial System involve deception by a subset of the actors
- Healthcare Delivery and Human Biology involve aberrant functioning by a subset of the actors
- Congestion Pricing and Urban Resilience involve aggregate consequences (e.g., traffic) of all actors

Another important distinction is between two classes of problems:

- *Bottom-Up*: Detection and remediation of aberrant actors involves stratifying actors and exploring behaviors of each stratum in different ways
    - Aberrant actors tend to react to remediation strategies, eventually undermining their effectiveness
- *Top-Down*: Economic strategies, e.g., pricing, payment models, procurement practices, based on aggregate behaviors
    - Individual actors tend to react to aggregate strategies, often undermining the desired consequences

Considering how the phenomena associated with these examples might be represented, three common features should be noted. First, the set of phenomena associated with a problem can be represented at different levels of abstraction, e.g., individual instances of counterfeiting versus macroeconomic policies that motivate counterfeiting. Second, each example has phenomena of interest that emerge within each layer of abstraction. This would suggest that a different representation of the enterprise system would be relevant for each layer. Third, each example exhibits feedback loops that cut across two or more layers. For example, the incentive to counterfeit increases with declining supplier profit margins. High-level policies designed to combat counterfeiting could raise costs at the lower levels. This could further erode profit margins and actually increase the incentive to counterfeit. Thus, the counterfeiting problem cannot be addressed without considering the relationships between the different layers of the enterprise system.

Thus, the analysis of these six problems is consistent with the difficulties identified during the discussion of complex social systems. Thus, we would expect challenges to modeling and prediction including but not limited to:

- Multiple, overlapping layers of abstraction
- Evolving behavior
- Path dependence

- Multiple potential equilibria
- Cascading effects
- Vulnerabilities in signaling and control

4.4 MODEL BIFURCATIONS

From a practical standpoint, the issues discussed above often manifest themselves as model uncertainty. In other words, you may have moved your model out of the zone where it is correct, but you do not know it. This is as opposed to aleatory uncertainty that is captured within our models. Aleatory uncertainty refers to uncertain events that can be represented as a random draw from a probability distribution. Generally speaking we can manage aleatory uncertainty, and it is the model uncertainty that presents difficulties.

The question is when does the behavior of a model start to substantially diverge from the behavior of the system it is representing? This is sometimes referred to as a bifurcation. The application of this term evolved out dynamical systems theory, and Robert Rosen [1978] makes extensive use of the concept in his discussions of the properties of system models. Formally, Rosen defines a model bifurcation using ε and δ neighborhoods, but for our purposes it is sufficient to say that two models diverge.

This concept is explained graphically in Figure 5. Suppose that the solid black line is the "true" system that we are trying to model. Suppose also that we have two models of the system that we can use, model 1 and model 2. Model 1 is represented by the dashed blue line and model 2 is represented by the dashed green line. Note that on left portion of the diagram, model 1 is the better representation of the system while on the right side of the diagram, model 2 is the better representation. There is a region of overlap in the middle where neither model 1 nor model 2 is perfect, but together they bound the true system. The two vertical lines indicate the bifurcation points. The left bifurcation point is where model 2 diverges from the true system while the right bifurcation point is where model 1 diverges.

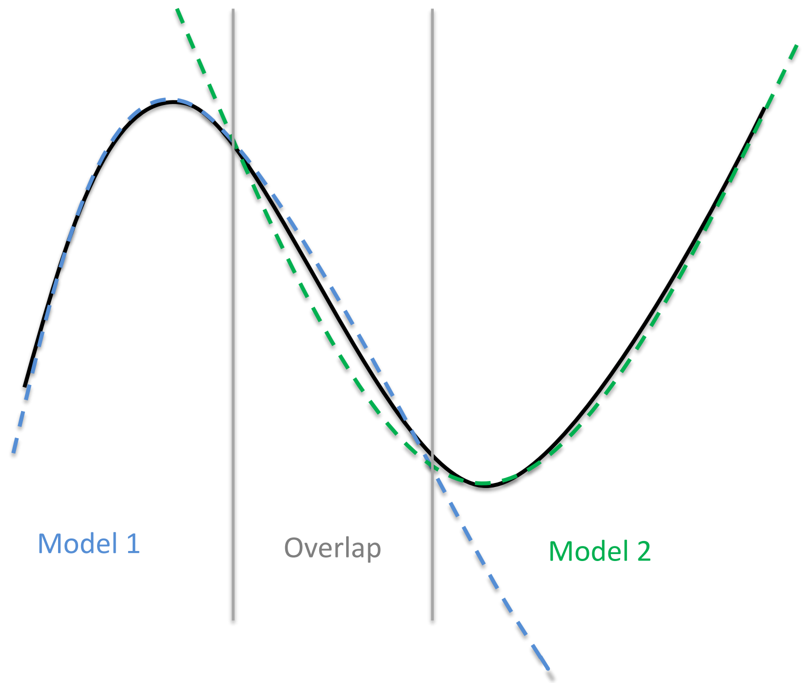


Figure 5 – Notional representation of model bifurcation

Perhaps the most important observation we can make about this example is that if we were to look at only model 1 alone or only model 2 alone, we would not detect the bifurcation points. If we were to look at both models 1 and 2 simultaneously we would see that the models bifurcate relative to each other, but we would not know which model is the better representation of the true system without additional information. This is the essence of the model uncertainty problem. We do not always know when our model is no longer valid since we cannot not find bifurcation points by simply running the model or performing a sensitivity analysis.

A more descriptive and intuitive term for a model bifurcation is "phase shift." A model bifurcation point is where the system of interest under goes a phase shift, and the model we were using is no longer accurate. For example, we can model ice as a rigid solid…until the temperature gets above 32°F. At that temperature (assuming standard atmospheric pressure) water under goes a phase shift from solid to liquid. So our model of water as a rigid solid is no longer valid. Thus, the melting point of water is a bifurcation point. In fact, one could make the argument that model switching is how we define phases.

We can extend this example by continuing to increase temperature. Once we get to 212°F, water undergoes another phase shift from liquid to gas. This necessitates another model shift. The challenge of modeling water is illustrated in Figure 6. Under a scenario of gradually increasing temperature we have to use three different, incompatible models of water: rigid solid, liquid, and gas. There are also two transition zones where the water is a mixture of two phases, and it is not clear what model we should use.

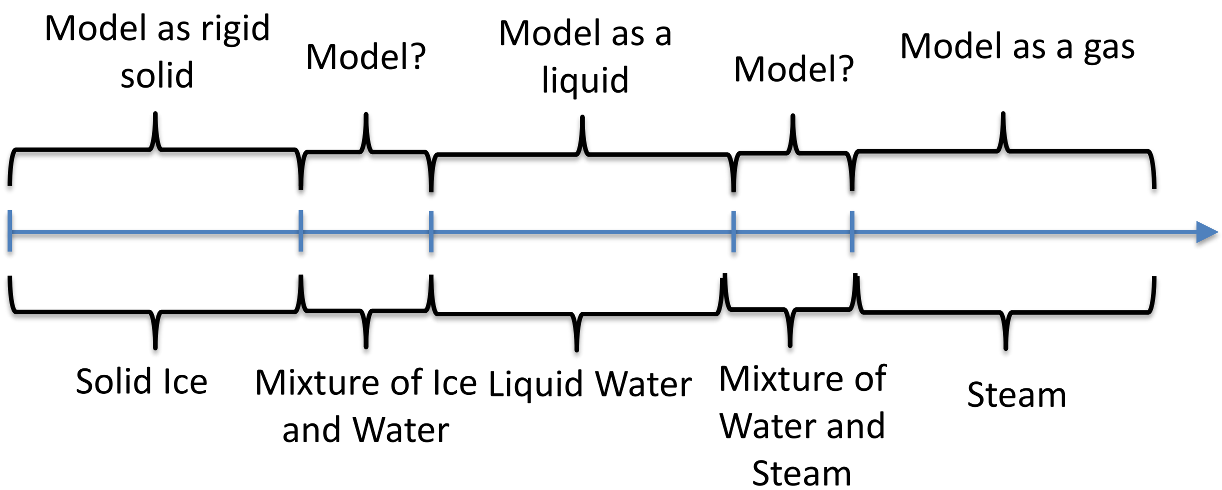


Figure 6 – Modeling water across phase shifts

Empirically, we know the bifurcation points of water, but these would not necessarily be evident using only models for a solid, a liquid, and a gas. For example, a standard model of a rigid solid may not include temperature as a parameter at all. If we were to model water vapor using the ideal gas law, reducing the temperature would certainly not reveal the condensation point. The ideal gas law would simply indicate that the volume would continue to drop (assuming fixed pressure). Thus, the problem is that if these three models were all we had, we may not ever find the bifurcation points. Thus, we would not know when the models are wrong. And this doesn't even factor in the transition zones where none of our models are accurate.

So how might we address this problem? The standard approach is reduction. We attempt to drill down a layer of abstraction and eliminate the bifurcation points. This concept is illustrated in Figure 7. For example we might be able to eliminate the bifurcation point between the liquid and gas models by going to a particle model (e.g., representing the water as collection of moving spheres). Using this model, we would avoid the problem of the volume of the gas getting arbitrarily small as the temperature drops because the particles themselves have volume. Thus we may potentially close this model gap. However, it is not clear that this model would capture all properties of liquid water (for example the surface tension of liquid water). It would certainly not be an accurate representation of solid ice, since the molecular structure of water is such that the density actually drops as liquid water freezes.

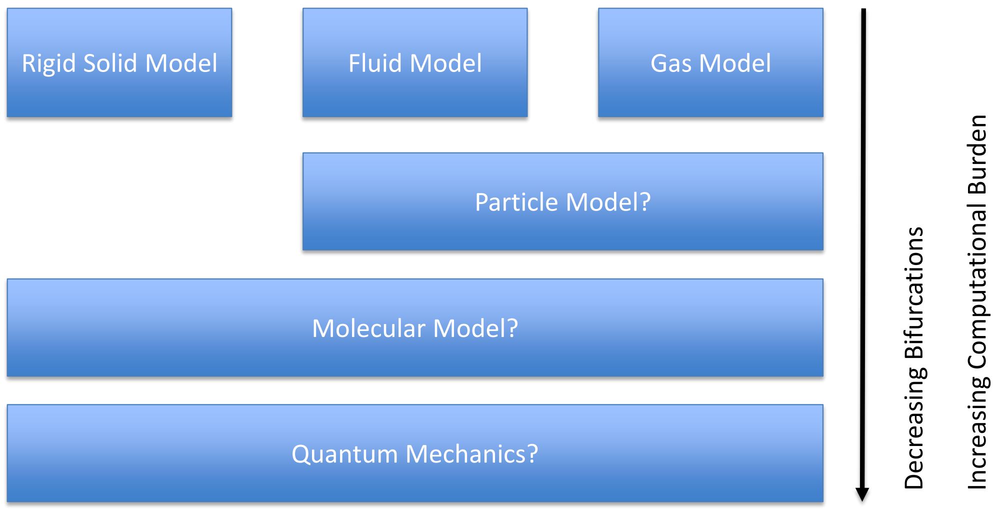


Figure 7 – Removing bifurcation points through reductionism

The next logical step would seem to be to employ a model that recognizes the molecular structure of water and the associated chemical properties. But even then, it is not clear that we would capture the all of the relevant properties. Perhaps we should just go straight to quantum mechanics to ensure that we have no bifurcations whatsoever?

Of course, it almost goes without saying that the computational and data collection burdens increase rapidly as we move toward quantum mechanics. At that point, modeling any macroscopic system is computationally prohibitive. As a practical matter, bifurcation points cannot be eliminated for most real world systems. (Even quantum mechanics bifurcates from both general relativity and the real world under certain circumstances). The question then becomes whether or not one crosses a bifurcation point for the situation and model of interest.

The issue we have with enterprise systems is that they undergo multiple phase shifts at multiple layers of abstraction. In fact, Casti [1985] notes that layers of abstraction themselves are effectively bifurcations. Casti applies the number of different models required to represent a system as a measure of its complexity. This would seem to comport with the discussion of complexity in previous sections. If, as we assert, social systems are complex, then they undergo many phase shifts, and it is effectively impossible to represent them with a single model.

A recent example of a phase shift in a social system is the emergence of high frequency trading (HFT) in financial markets. In essence, HFT works by leveraging high speed communication links to front run orders placed by other traders. This allows the high frequency trader to make a small amount on the transaction. But when this repeated many times, the amount of money can be substantial. This shift in behavior in a social system actually led to the installation of new infrastructure as high speed communication lines were laid just to allow this mode of trading to take place. It also led to countermeasures by those that were losing out to HFT. In one case, a new exchange was established that intentionally inserted a delay into the communications infrastructure so that it could not be exploited by high frequency traders. Thus, the entire behavior of the market place has shifted. This phase shift would not have been predicted by existing market models, and thus represents a model bifurcation. (If it had been, then the emergence of HFT would not have been such a surprise to financial community).

Consequently, our only course of action is to try find the phase shifts or bifurcation points so that we can deal with them. Returning to the water example above, knowing where the bifurcation points are allows us to know when to switch models. The problem is that we do not necessarily know the analog of melting point for a social system. In fact, social systems can be so complex that we have no idea how many bifurcation points to look for when we pose a question of interest. So how might we go about identifying phase shifts or bifurcation points?

The first thing to understand is the sources of the phase shifts. Some of these were identified at the end of the previous section including overlapping representations, feedback loops, and adaptive behavior. In some circumstances, we may be able to establish the existence of and possibly bound bifurcation points by modeling these sources. For example, agent based models (an example of a reduction approach) may be used to explore the potential consequences of adaptive behavior and feedback loops. The phase shifts revealed by these models are sometimes termed emergent behaviors.

Another possibility is to compare alternative models for the same system to attempt bound the phase shift. Notionally, this was represented by the overlap region depicted in Figure 5. A real life example of this approach is discussed in Section 6.1 in the congestion pricing case study where an agent based model and a differential equation model are compared to bound the bifurcation point.

As noted above, there are some phase shifts that cannot be detected by modeling. These can only be found empirically. If an experiment is feasible, it is the idea solution. However, for many enterprise problems, experimentation is either impossible or cost prohibitive (which is why we are considering modeling and simulation to address enterprise problems in the first place). In that case we would like to determine if there are leading indicators of an impending phase shift to give enterprise decision makers at least some time to prepare.

It turns out that finding leading indicators of phase shifts has been a line of investigation in complexity research for quite some time. For example, Bossomaier, et al. [2013] discussed how increases in metrics such as transfer entropy and mutual information can presage a phase shift. They illustrate their point by computing mutual information with historical market data and demonstrating that significant increases in mutual information correspond with major market corrections over the past decade. However, it is not clear how general and effective these types of metrics are, and additional investigation is required to fully understand their applicability to modeling enterprise systems.

4.5 Implications for Methods and Tools

Given that an enterprise is likely to be concerned with a number of phenomena both internal and external to the enterprise, one would ideally like to develop a family of strategies to address these phenomena. By recognizing the epistemic limitations, one can tailor modeling and simulation efforts to support strategy development. Of particular concern are the model bifurcations discussed in the previous section. These suggest that the traditional approach to engineering analysis will not be effective: building a model of the system of interest and searching the design space for the best options. The model bifurcations will likely interfere with finding an "optimal" solution. So how should the problem be approached?

There are essentially two objectives for this modeling effort:

- Explore the tradeoffs among strategy options to address a phenomena of interest
- Ensure the consistency of a family of strategies

The intent of the second objective is to ensure that the implementation of a strategy to address one phenomenon does not interfere with a strategy to address another phenomenon or generate counterproductive responses among other phenomena.

However the epistemic limitations have implications for realizing these objectives. In particular:

- There is no complete model of the enterprise or its ecosystem. We will always be leaving something important out.
- One cannot obtain conclusive answers from modeling and simulation. They can only be used to support decision makers and engineers in exploring trades.
- No single ontology can represent a whole enterprise problem. Multiple models will be needed to explore trades.
- Strategies developed to address phenomena in different ontological layers may interfere with each other. Model composition is needed to check the consistency of strategies.

In short, we will need to develop and compose models from multiple domains to support enterprise engineering and strategy development. This approach to modeling goes by several names such as multi-level, multi-scale, and multi-resolution modeling.

As an illustrative example of modeling an enterprise system, consider the health care delivery enterprise examined by Park, et.al. [2012]. In that work, the authors analyze the performance drivers of an actual employer-based prevention and wellness program.

First, they decompose the enterprise into four layers ranging from the overarching healthcare ecosystem down to the clinical practices used to treat individual patients. Each of these layers present different phenomena that the health care enterprise must contend with if it is to operate effectively. Second, the authors identified the potential design decisions that a would-be enterprise engineer would need to make to address each of the phenomena. Table 14 presents some representative examples of these design decisions for each layer. Ultimately, the authors built a simulation of this enterprise and analyzed the impacts of the design decisions. What we can draw from this example is that engineering an enterprise involves modeling a range of human and social phenomena at different layers of abstraction.

Table 14 - Examples of design decisions at each layer of the enterprise (Adapted from Park, Clear, Rouse, et. al. [2012]

Enterprise Layer	Enterprise Design Decisions
Healthcare Ecosystem (Society)	• Payment models • Participation policies
System Structure (Organizations)	• Investment decisions
Delivery Operations (Processes)	• Acceptable risk thresholds • Patient stratification • Business processes • Standard operating procedures
Clinical Practices (People)	• Acceptable clinical practices

Unfortunately, the very limitations that push us toward a multiple level approach also complicate the composition of these models. The symptoms of a deeper theoretical issue are evident in the modeling and simulation literature. As an example, Wang, et. al. [2009] posit that the High Level Architecture (HLA) standard for federating simulations (IEEE 1516-2010) has experienced issues in part because it neglects higher level conceptual interoperability between models. In fact, the *Journal of Simulation* devoted a special issue to the topic of composition entitled "Enhancing simulation composability and interoperability using conceptual/semantic/ontological models," (Vol. 5, No. 3). The use of ontology is key to understanding the challenges.

Coordinating ontologies among simulation models as a means to improve model composition has been examined repeatedly. [Hoffman 2011, Hoffman 2013, McGinnis et al 2011, Partridge, et al 2013, Tolk 2011]. Hoffman [2013: 77] asserts that, "For many technical domains and artificial systems, ontologies will be able to ensure the interoperability of simulation components developed for a similar purpose under a consensual point of view of the world."

However, when we expand the scope and attempt to integrate human and social phenomena into a traditional engineering model, difficulties arise. Hoffman [2013: 79] points out that, "…in many socio-technical and most social domains the specification of such 'well defined' domain ontologies (referential ontologies) will be impossible…Hence, in these cases, there is no easy mapping possible between referential and methodological ontologies…This mapping, if possible at all…would not be a technical matter, but a challenging and subjective task of selection."

The conclusion we can draw here is that it is unlikely that there is any final theoretical solution that will resolve all of these difficulties. Rather, a more promising approach is to develop methodologies to mitigate them. To that end, we have developed an over-arching methodology for modeling enterprise systems (Section 2), a strategy framework to mitigate the epistemic difficulties (Section 5), and finally guidelines for composing models when it is appropriate to do so (Section 6).

5 STRATEGY FRAMEWORK

Our examination of the six enterprise examples revealed that a would-be enterprise engineer may be concerned with any number of phenomena ranging from the adverse actions of external actors to aberrant employee behavior to a breakdown in communications. For any given enterprise, there are likely to be multiple such phenomena, each seeming to merit a response.

Two questions logically follow: what strategy options do the decision makers have available to them and what criteria determine when they should select one strategy over another? Two criteria seem immediately relevant: First is the ability of the decision maker to accurately predict the response of the enterprise and/or enterprise ecosystem to a design or strategy decision. Second is the ability of the decision maker to actually implement the decision. If the decision maker could predict all future responses to the decision, then there is no reason to ever make a change. He or she could determine the optimal policy for every possible circumstance and design the enterprise accordingly. No changes would ever be required. If the decision maker could respond to every new development instantaneously and costlessly, he or she would not need to predict anything. He or she could continuously adapt the enterprise as needed.

Obviously, real life lies in between these extremes. For any given enterprise and phenomenon of interest, there will be a different ability to predict and different ability to respond. The implication is that the decision maker's strategy will change based on the relative values of these two attributes. The predictability of social phenomena has already been discussed extensively. The ability to respond merits additional discussion.

Several factors influence a decision maker's ability to respond to a development in the social system. The first is the ability to detect that a change has occurred. This may seem trivial at first, but consider the difficulty of measuring macroeconomic phenomena. Metrics such as unemployment, gross domestic product, and inflation are often subject to long measurement time lags and multiple revisions. Furthermore, there are often biases inherent in the measurement techniques. Consequently, determinations such as whether or not an economy is in a recession are often contentious and not settled until long afterward. As discussed in Section 4.4, research into leading indicators of phase shifts may allow decision makers some forewarning of an impending change.

The second factor is the capacity to respond. In other words, does the decision maker know what to do, and is he or she capable of doing it? If there is no known way to respond to an event or the decision maker is not capable of implementing the response, the outcome is the same. For example, changes to a foreign country's import policies may have a substantial impact on a business enterprise, but there may be nothing that the enterprise can do about it.

The third factor is the speed of response. The decision maker must be able to implement the response before the consequences of the event are too severe to recover from. For example, if a competitor introduces a rival product that is vastly superior, an enterprise may have the ability to introduce a comparable product, but if it does not introduce it quickly enough, the enterprise may lose too much market share to recover.

The fourth and final factor is the cost of the response. The decision maker may be capable of responding quickly and correctly, but the response may be unaffordable. For example, an upstart competitor may be using a new, more efficient production technology that reduces costs, but the enterprise has such a large established production infrastructure that converting would be cost prohibitive.

Given different abilities to predict what could happen and respond to it, what options are available to enterprise decision makers? We contend that there are four basic options: optimize, adapt, hedge, and accept (Figure 8).[2]

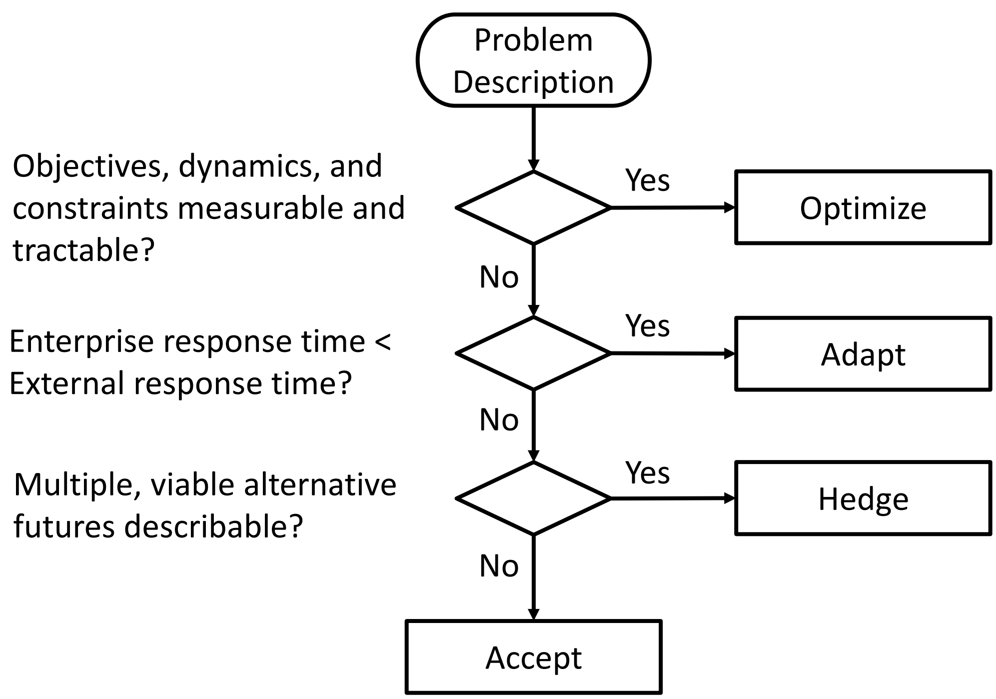


Figure 8 - Strategic framework for enterprise decision makers

[2] The optimize, adapt, hedge, framework was first presented in Pennock and Rouse [2014b].

If the phenomena of interest are highly predictable, then there is little chance that the enterprise will be pushed into unanticipated territory. Consequently, it is in the best interest of enterprise decision makers to *optimize* the enterprise architecture to be as efficient as possible. In other words, if the unexpected cannot happen, then there is no reason expend resources to make the system more robust, resilient, or flexible.

If the phenomena of interest are not highly predictable, but it is relatively straightforward to adapt the enterprise appropriately, it may be in the best interest of the enterprise to plan to *adapt*. For example, the decision makers may decide to implement the engineered system using a modular architecture that will allow them to change out modules as needed to respond to evolving circumstances. In this case, some efficiency has been traded to improve the ability to adapt. For the social system, the previously discussed guided self-organization approach may be appropriate. General constraints and guidelines may be set, but participants are allowed to adapt to local circumstances.

Models can help us explore the space of possibilities even if we cannot predict what will happen. However, for this approach to work the owners and/or operators of the enterprise must be able to identify and respond to potential issues faster than the external environment changes. For example, when we consider congestion management, new transportation projects often take years to plan and execute. By the time they are complete, traffic congestion may have changed or the project may simply serve to move the congestion around. Drivers can respond much faster than the government can make changes to infrastructure.

If the phenomena of interest are not very predictable and the enterprise has a limited ability to respond, it may be in the best interest of the enterprise to *hedge* its position. In this case we can use our models to aid in exploration of scenarios, but our system may not be able to handle sudden changes without prior investment. For example, a firm concerned about product obsolescence may choose to invest in multiple, potential follow on products. While the firm does not know which product will ultimately succeed in the marketplace, it will be ready to take advantage of whichever product ultimately does. If the firm were to take a wait and see approach, it would not be able to respond quickly enough, and it would lose out to its competitors.

If the phenomena of interest are totally unpredictable and there is no viable way to respond, then the enterprise has no choice but to *accept* the risk. Accept is not so much a strategy as a default condition. In other words, if one is attempting to address a problem where there is little ability to predict the efficacy of solutions and little ability to adjust the solution once implemented, then there is nothing that engineering can do to address the problem. Any engineered solution would be tantamount to the expenditure of resources on a wild guess.
To illustrate this point, let us consider an extreme example. A global nuclear war would likely destroy many enterprises. However, there is no basis for estimating the likelihood of such an event, and there is nothing that most enterprises can do to respond. Consequently, few non-governmental enterprises are likely to have a contingency plan for a global nuclear war.

To make the strategies more concrete, we can reconsider the six example enterprise systems. Table 15 presents examples of each type of strategy that could potentially be applied to each enterprise system. (Note: one additional example from the automotive industry based on another CCSE research effort is included in the table.) This table is intended to be illustrative not comprehensive. It is important to note that strategies will depend on the perspective and the scope of responsibility of the enterprise engineer and the issue of concern. Most real life enterprises will involve a mixture of strategies.

Table 15 - Strategies applied to example enterprise systems

Example System	Optimize	Adapt	Hedge	Accept
Counterfeit Parts	- Optimize supply chains assuming counterfeits are detectable	- Redesign obsolete systems - Employ modular architectures - Adapt inspection approaches as counterfeiters adapt	- Buy a lifetime supply of parts up front - Own the manufacturing process - Design for fault tolerance	- Buy on the open market and accept the risk of counterfeits
Financial System	- Write fixed laws and regulations governing financial markets	- Bail out firms considered too big to fail - Open market operations in response to changing financial situation	- Enforce reserve requirements - Stress test financial institutions	- Deregulate financial markets
Urban Resilience	- Optimize urban infrastructures and emergency response capabilities against environmental and demographic projections	- Adjust emergency response deployments as events unfold - Empower personnel on the ground to make decisions	- Apply zoning rules to limit damage exposure - Purchase spare capacity for both infrastructure and emergency response	- Accept that flood mitigation measures will fail for a storm surge greater than a certain level
Congestion Pricing	- Optimize road layout and tolls against traffic projections	- Change toll prices - Alter HOV policies - Reverse lanes	- Build road capacity in excess of projected usage	- Accept congestion pricing may not alleviate traffic due to accidents and disruptions
Healthcare Delivery	- Optimize facility capacities for expected utilization	- Reschedule patients to accommodate unexpected delays	- Maintain extra facility capacity to accommodate surges in demand	- Accept that variability degrades service
Human Biology	- Optimize therapeutic regime based on clinical results	- Adapt treatments to patient comorbidities	- Make lifestyle choices that reduce the risk of cancer	- Accept that cancer evolves
Automotive Industry	- Optimize manufacturing processes, inbound and outbound supply chains, and pricing of products and services	- Adapt by changing product mix (as Honda but not others did during recent recession)	- Hedge by investing in multiple technologies - Buy options on future consumables (fuel and raw materials)	- Accept by minimizing the other investments and keeping the, hopefully, cash cow functioning

When considering which strategies to employ, it is important to note that modeling choices and and strategy choices are interrelated as illustrated in Figure 9. If the system model does not bifurcate over the space of interest, then it is likely that one can employ an optimize strategy. If, on the other hand, the system model does bifurcate over the space of interest one is faced with a choice. One can take a reductionist approach and attempt to eliminate the bifurcation by creating a more detailed model. Consider the water modeling example presented in Section 4.4. One option was try to overcome the phase shifts by going to a molecular model of water, but this entailed substantial computational and data collection costs. So the price of attempting to preserve an optimize strategy could be high to prohibitive.

If one the other, hand one is willing to live with the model bifurcations, then one can take an adapt or hedge strategy to manage the bifurcations. This approach substantially reduces the modeling burden. Thus, one cannot consider the modeling choices and strategy choices in isolation.

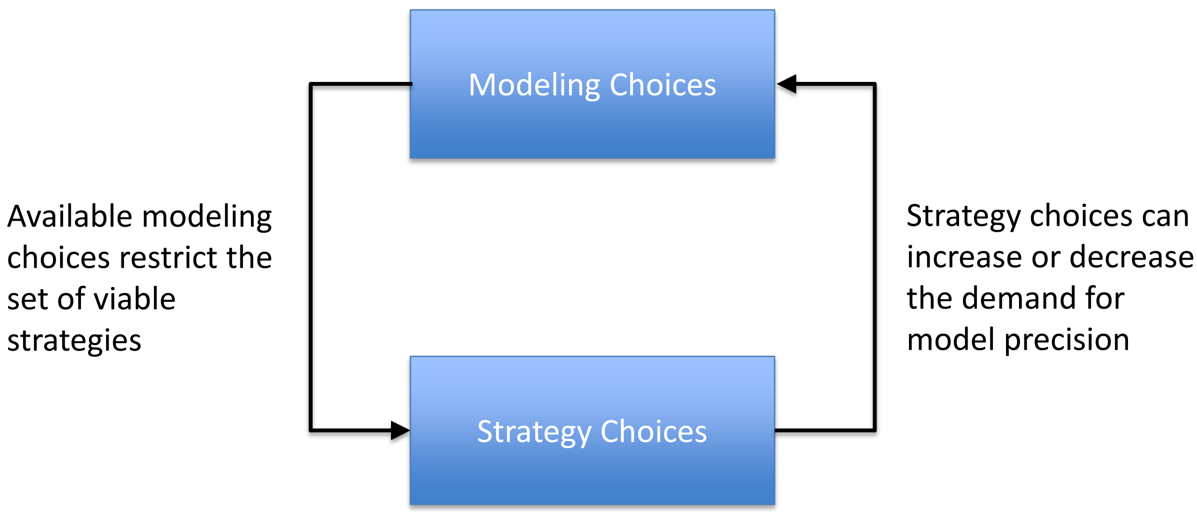


Figure 9 – Relationship between modeling choices and strategy choices

6 Technical Issues with Modeling Enterprise Systems

In the previous sections of this report, we have discussed the necessity of employing multiple models to represent enterprise systems. This poses a number of technical issues that we have largely discussed in a conceptual manner. In this section we will go into more detail on how these issues manifest themselves and how they might be remediated.

One of the major issues raised was that of model bifurcations or phase shifts. Given the importance of economic factors in the counterfeit parts problem, there was concern as to how robust the application of classical economic models are for enterprise problems. Consequently, a small scale case study was conducted on congestion pricing as this combines a classical economic model with a traditional engineering model. What was found was that the classical economic model bifurcates from the real system over the typical toll prices found on actual managed highways. (In other words, the classical economic model produces an incorrect result.) The finding reinforced the importance of identifying bifurcation points when developing enterprise models so that the risk can be addressed prior to making an erroneous decision. This case study is presented in Section 6.1.

The other major issue raised is the model composition issue. Depending on the strategy selected, it may be necessary to compose multiple models to represent an enterprise. (Note that this may not be feasible for many enterprise problems of interest). This topic was initially explored during RT-44a and the result was set of heuristics for model composition. During this task, we applied findings from investigations into the nature of modeling to understand when and how such heuristics are applicable. The results of this analysis are presented in Section 6.2.

6.1 CONGESTION PRICING CASE STUDY

As an illustration of a complex socio-technical enterprise problem that highlights many issues related to modeling and composition, we consider the case of traffic modeling and congestion pricing. Our aim is to show the high level of complexity involved in traffic management and to point out the related modeling issues. Congestion pricing provides a clear application of the optimize/adapt/hedge/accept methodology and highlights the issues of adaptive capabilities, phase transitions, overlapping representations, and composition of classical economic models with technical systems.

The problem of traffic congestion involves and affects a number of issues. As such, it is quite difficult to capture the entire traffic domain with a single model. In practice, multiple models must be used to gain a view of the entire situation, and so there is an increased potential for model error. The modeler must determine the objective that he is trying to achieve, and utilize the appropriate models. Depending on the objective, the problem can easily bleed into other domains, producing several unintended consequences that can include increased demand on alternate methods of travel (public transport, nearby roads, etc.), political and social issues surrounding the implementation of potentially unpopular tolls, allocation of toll revenues, and land use issues surrounding changing traffic patterns (e.g., property values, effects on business, or environmental impacts). That said, we focus here on two of the major issues directly involved with traffic management: modeling traffic flow and congestion pricing. As a practical matter, we discuss real-world congestion pricing in Singapore, Atlanta, and Minneapolis.

6.1.1 BACKGROUND

Let us begin by giving a broad overview of a typical traffic engineering problem: An area faces traffic congestion at some particular locations. To alleviate this the city must decide upon a congestion mitigation scheme. It is likely that one or several traffic models would be built to determine strategies that may lead to reduced congestion. Such models can incorporate many different strategies at different levels (e.g., changing a speed limit or installing a traffic circle on a particular street, adding an additional lane to a highway, incorporating a fixed or variable priced tolling policy for network entry, investing in public transit upgrades, etc.). The ultimate goal is for a reduction in traffic congestion, as can be measured by, for example, a decrease in average travel time, an increase in vehicle or passenger throughput, or a decrease in travel time variance.

6.1.1.1 Traffic Modeling

Two of the major types of systems modeling are differential equation based modeling and agent based modeling. Models that view traffic as similar to flow through a network include the classical partial differential equation model of Lighthill and Whitham (1955). Car-following models [e.g., Gipps (1981), Treiber et al. (2000)] utilize ordinary differential equations and base the behavior of one vehicle on that of the vehicle immediately in front. The bottleneck model of Vickrey (1969) and Arnott et al. (1993), and more recently the bathtub model of Arnott (2013) are other attempts to replicate flow through a network without explicit reference to individual driver attributes. These models can also be used to derive optimal tolling prices. Some newer work uses an agent-based approach, for instance, Zhang, Levinson, and Zhu (2008) give an agent-based model of congested networks, which includes the composition of several component models (user choice/demand, pricing choice, capacity choice). The benefit of an agent-based model is that individuals are given the capability to learn, interact, and adapt over time. For a thorough review of much of the work on traffic modeling see Helbing (2001).

There can be many causes of traffic congestion. Non-recurring congestion refers to cases that do not consistently repeat, for example, delays due to special events, weather related issues, or accidents. Recurring congestion is due to repeated issues such as high demand in morning and afternoon rush hours or physical issues such as lane merges or other road obstructions. Non-recurring congestion is generally unavoidable, and probabilistic methods can be used to model its occurrence. We focus here on recurring congestions and the methods that can be used to reduce it.

Both equation-based and agent-based models can be used to model traffic breakdowns. A phase transition or bifurcation in the traffic realm occurs when traffic changes from free-flowing to congested. The literature contains a number of studies concerning bifurcations. For example, Gasser, Seidel, and Sirito (2007) perform a bifurcation analysis of car-following models and explicitly consider driver reaction times and aggressive driving behaviors. The authors show how Hopf bifurcations can occur. Orosz, Krauskopf, and Wilson (2005, 2004) also consider a car-following model, discussing the stability of optimal velocity flow and showing how bifurcations can result in traffic jams. Daganzo, Gayah, and Gonzales (2011) show how bifurcations can be used to explain the instability of homogenous networks when density exceeds a critical level. The value of this critical level, however, increases as drivers become more adaptive. A phase transition model of traffic flow can be found in Colombo (2002, 2003), where congestion and free-flow phases of traffic are distinguished. This model is extended by Blandin et al. (2011), where an intermediary or “metastable” stage is introduced.

This intermediary stage, where demand is sufficiently high that a jam could occur under some set of circumstances, but it is not guaranteed to occur, is the source of a good deal of traffic uncertainty. The set of circumstances that forms a jam can be quite specific, with what seems to be the same set of events causing a jam in one case, but not in another. The difference can be a very minor event, such as a tap of the brakes or a lane change. The point is that, unless demand is very high, to the point in which a jam is inevitable, we must rely on probabilistic methods of jam formation.

6.1.1.2 Congestion Pricing

Common traffic management strategies can be placed in two categories; supply side and demand side. Supply side solutions generally affect the physical properties of the network – adding lanes, roads, or even public transportation options. Demand side strategies invoke methods to reduce or spread out demand on the network, reducing the likelihood and severity of traffic jams. Although supply side solutions can be effective (at least in the short-term), they are usually expensive, and future growth can result in further stresses to the system. Moreover, there are concerns that increasing the capacity of a roadway actually increases demand, and so the intended reduction to congestion may not be realized.

Although the road capacity provides an upper bound on traffic volume, jams can and do form on roads well under capacity. Because the probability of a jam increases with the percentage of road capacity in use, and because road capacity is in some sense fixed, the most common method used to reduce congestion is by reducing demand. This can be done in a variety of ways, but the most broadly accepted response to traffic congestion is to implement congestion pricing, i.e., to impose tolls designed to reduce demand.

Tolling has been studied, for example, in De Borger and Proost (2011), Lindsey et al. (2012), and Lindsey (2006). While tolling may reduce road demand, it also adds a layer of complexity to traffic modeling, as it requires additional insights into human behavior. For instance, we need to understand how people value time, and how that value might change under different circumstances. There are many forms of tolling, with the extremes being static pricing and dynamic pricing. A pure static pricing would set a flat price to travel for some predefined distance on a road, while a purely dynamic approach would change the price continuously. A dynamic tolling scheme that charges users the marginal cost of their choice to travel on the road will minimize system-wide trip time, and is therefore the ideal strategy. Implementation of this strategy is quite difficult, due to both physical limitations and driver response. There are several other tolling options in between these two extremes, such as step-tolling, where prices are changed at predefined time intervals. For an analysis of step-tolling see, for example, Lindsey et al. (2012), where the ability of drivers to slow down in anticipation of a toll decrease is considered.

The extent to which the optimal toll is used or attempted to be used in reality is unclear. In practice, traffic agencies often perform stated preference surveys to aid in toll rate determination, where road users are surveyed regarding the price that they would pay to avoid congestion. The results of these surveys are used to determine a distribution of price preferences, and ultimately can be used to help gauge the increase or decrease in demand corresponding with a toll change. The main presumption in tolling is that road demand will fall as toll prices increase.

A problem with such a model is that drivers may act under quite different circumstances in actual traffic than was considered in the survey. There is thus great potential to misapply the results of these surveys, and it is likely that the choices of drivers (e.g., enter or avoid a tolled area) is based both on what they see as well as what they know. This is especially a problem when there are many alternative choices or routes that one can take to avoid a toll. Examples of driver response to pricing include Brownstone et al. (2003) and Bento, et al. (2013).

One of the expected benefits of congestion pricing is to shift road demand to non-peak hours. Burris and Pendyla (2002) examine this expectation using real-world data from a tolled bridge in Florida. Their findings were much as expected: drivers with flexible working time, retired drivers, and low income drivers were most likely to alter their time of travel in response to dynamic pricing. If this is generalizable, it would indicate that the success of dynamic pricing is at least partially dependent on the demographics of the market.

Bonsall et al. (2007) summarize much of the existing literature on driver response to road pricing, concentrating on the feasibility of the optimal pricing strategy of continuously changing tolls. In particular, they focus on the ability of drivers to rationally respond to complex pricing (as opposed to the more commonly studied issue of physical implementation), ultimately concluding that people are unlikely, for various reasons, to react rationally to congestion pricing. Furthermore, their engagement with the system is expected to diminish with time, and their sensitivity to price changes is likely to decrease.

6.1.2 Traffic Management Examples and Strategies

In this section we discuss the issues involved in traffic modeling in the context of Pennock and Rouse (2014a, 2014b), who discuss several problems involved with enterprise system modeling. Two epistemic challenges to enterprise system modeling are given; overlapping representations and adaptive behavior, and a loose modeling methodology is given; optimize if possible, otherwise adapt or hedge, and finally, accept the consequences of events beyond control.

The adaptive capabilities of drivers and overlapping representations are pervasive in this domain and present an ongoing challenge to successful modeling. Moreover, these issues are closely related to the problems involved in phase transitions. The optimize/adapt/hedge/accept methodology can be used to provide guidance in traffic modeling. While certain parameters and situations can be optimized, the level of uncertainty inherent in the problem calls for adaption or hedging in many cases, and acceptance is required in the most uncertain situations. We begin this section with a discussion of real-world congestion pricing implementations in Atlanta and Minneapolis. These examples shed light on many of the issues faced in traffic modeling, and show how optimization and adaptive procedures are used in real-world scenarios.

6.1.2.1 Congestion Pricing Examples

Atlanta

We consider as an example the efforts made in metropolitan Atlanta to manage congestion. Traffic volume in Atlanta during morning and afternoon rush hours is sufficiently high that congestion delays were the norm. Atlanta had already made attempts at mitigating traffic congestion, offering bus routes with “park-and-ride” lots, and including high occupancy (2 or more people) vehicle lanes on the major highways in the region. Despite these attempts, high levels of congestion persisted in the region, and so the city came to consider additional measures.

Because heavy congestion was recurring, it appears that the root of the problem is the lack of sufficient capacity to meet traffic demand. Some relatively simple measures were taken, such as identifying areas where bottlenecks occurred, and engineered solutions for these locations were used when applicable.

A number of different studies were commissioned, studying the expected benefits of various traffic management strategies. These studies examined possibilities including high occupancy vehicle lanes, high occupancy toll lanes, express toll lanes, truck only lanes, and truck only toll lanes. The ultimate goal was to replace the current, high occupancy vehicle lanes with one or more of the above possibilities.

Ultimately, it was decided that the 2+ HOV lanes would be changed to 3+ HOT lanes, where vehicles with at least three riders could use the lane with no charge, and any other vehicle could enter the managed lanes by paying a time-varying toll. Other types of vehicles are also exempt from the toll; motorcycles, alternative fuel vehicles, transit vehicles, and emergency vehicles. Some of the project’s objectives appear to be (HNTB Corporation, 2010)

1) To protect mobility in the managed lanes (increase travel speed and decrease delays).
2) To maximize person/vehicle throughput in the managed lanes and decrease variability.
3) To minimize environmental impacts of managed lane construction.
4) To provide a financially feasible structure for the managed lane network.
5) To build a flexible infrastructure that allows for adaption to future network needs.

It is important to note that there is no objective explicitly stating that congestion should be reduced in general. Instead, the goals deal with maintaining a minimum speed (ultimately, 45mph) in the managed lanes.

To gauge consumer interest, a survey was given to determine driver sensitivity to tolling and travel time changes. Time values were shown to vary by time of day, trip purpose, and, within those segments, by household income and trip distance. The mean value of time for private drivers varied from $7 to $15 per hour, while commercial vehicles valued time at a rate of $23 per hour. Output from this survey was used in the generation of traffic and revenue forecasts.
The tolls implement variable pricing that updates every 15 minutes. Prices are based on the segment of the highway with the highest traffic volume and potential for speed decrease, and rise and fall with volume. Additional buses and bus routes were also added along with "park and ride" lots to access the buses. Investment was also made in an automated toll enforcement system and in a campaign to promote 3+ passenger carpools.

A discussion on the preliminary results of these efforts is given in [HTNB Corp 2010]. As a summary, we note that an increase in usage of express lanes was observed and the performance of those lanes slightly improved. However, the formation of 3+ car pools did not progress as expected, and so the increased usage of the lanes is a result of single drivers willingly paying the toll. Furthermore, congestion in the general-purpose lanes actually experienced a slight increase. Overall, the initial results seem to be inconclusive at best.

Minneapolis

Variable congestion pricing has also been implemented in Minneapolis, MN. HOT lanes are available for single drivers who pay the variable toll and for high occupancy vehicles who pay a reduced toll. Prices are based on the magnitude of traffic and the recent change in density, with higher prices occurring with higher traffic volume.

Using traffic data obtained from the Minnesota Department of Transportation, Goodall and Smith (2010) create a model of HOT lane demand. Their approach begins by segmenting the data into cases where observed speeds are above or below 55 mph in both the HOT lanes and the general purpose lanes. They found that HOT lane use increases from early-peak to mid-peak even when general purpose lanes had high speeds. Some drivers are therefore likely to use HOT lanes even when there is no evidence of congestion in the general-purpose lanes, perhaps in anticipation of oncoming congestion based on historical experience. Such drivers use the HOT lanes for their reliability, and are assumed to be insensitive to price.

A study by Janson and Levinson also investigated the situation in Minneapolis, showing that drivers use toll prices as a proxy for congestion, and thus increasing tolls correspond with increasing demand for HOT lanes. As evidence of this phenomenon, the pricing models were deliberately altered for a five-week period. Higher posted prices consistently resulted in increased HOT lane traffic. The typical economic assumption of higher prices leading to lower demand is thus violated in this case.

6.1.2.2 Optimize/Adapt/Hedge/Accept

The formation of traffic jams and driver response to tolling are both dependent on things outside of the modeler's control. Drivers can easily change their mind about whether or not they would pay a particular toll, and, if traffic is heavy enough to form (but not so heavy so as to guarantee) a jam, then any number of minor disturbances can set off a cascading traffic disruption. Due to the instability of these problems, an optimization approach is unlikely to work well here, and an adaptive or hedging strategy may provide a more robust solution.

The typical implementation might combine a dynamic pricing model (using some method for determining driver price preferences) with a congestion model. Great care must be taken in implementing these models. While a simple traffic flow optimization model might give a good indication of the effects of a change in speed limit or traffic light timing *ceteris paribus*, it is important to recognize that humans are operating the vehicles, and so their decisions as to whether or not to enter a system (or to obey speed limits and traffic signals) are dynamic. Optimization models cannot be expected to accurately model such a problem, although they may be helpful in evaluating responses to situations that could occur, i.e., they can be used in tandem with an adaptive or hedging strategy.

The first option that a modeler should consider is to optimize if possible. This is the traditional engineering approach, and is applicable when the problem is relatively simple and has well-defined parameters, or when the models needed to describe the system are compatible with respect to objectives, dynamics, and constraints. The optimize strategy is likely to be effective only for relatively small scale attributes of the system, in particular, those attributes that are weakly dependent on or independent from the social components of traffic. For example, the physical features of a highway (curvature, grade, etc.) may be optimized to provide the best or safest travel, while features more sensitive to the choices of drivers such as traffic light timing or toll pricing may actually create worse congestion if "optimized". In the example of congestion pricing in Atlanta we saw that, prior to deciding upon the tolling scheme, the roads were optimized where feasible.

In cases where optimization is ineffective, an adaptive strategy may be helpful. This builds into the system the ability to change in response to the overall system state. For example, the use of reversible roads that can change direction gives adaptive capabilities that can be used for rush hour management or special event management. Congestion pricing itself is an adaptive strategy, and its widespread usage provides clear evidence that the importance of adaption is understood by planners.

Another problem highlighted by traffic modeling is the difficulty that humans can have at replicating the actions that are required by a model. This can include clear cases of idealization such as perfect rationality assumptions, as well as less obvious problems such as physical barriers to action, where, even if a driver wishes to behave rationally, there exist certain limitations that preclude these actions from taking place. Optimization is not a good strategy in such cases, with adaption or hedging being better choices. (Note: one cannot decide to move to the HOT lane if you are stuck in the rightmost lane.)

For example, variable pricing has been in place in Singapore for nearly 40 years. Currently, entrance to the city center and some expressways are priced according to time of day. The prices are recalculated every three months to account for changes in traffic behavior. See Olszewski and Xie (2005) for a more detailed discussion. This type of implementation showcases an interesting strategy. As mentioned above, it is very difficult for drivers to react rationally (and according to model expectations) to dynamic pricing. Drivers in Singapore, however, are afforded the opportunity to gain experience with a price and develop a more rational response to tolling.

For cases where adaptive approaches are insufficient, a further strategy is to hedge, i.e., to pursue protective measures against any viable future states. For example, the authority in charge of highway planning can lease or buy land that is earmarked for future capacity expansion. The problem with hedging lies in determining a (relatively) complete set of viable future states. In the absence of such a set, the modeler's last resort is simply to accept whatever consequences arise from the systems operation. This strategy is a necessity in traffic management, as non-recurring congestion events (e.g., weather or traffic accidents) can occur anywhere in the network at any time, and so a level of acceptance of these events must be maintained.

6.1.2.3 Adaptive Behavior, Overlapping Representations, and Phase Shifts

Adaptive behavior is a fundamental challenge to traffic systems, as drivers are intrinsically adaptive, and their decisions directly impact the system. They can choose to enter a tolled area, or alternatively pick from a variety of alternative transit options (bike, walk, public transportation, carpool), or not make the trip. Furthermore, they can often adapt their behavior while in the system, choosing one route over another on a whim. Making matters worse is that drivers often make adaptive choices based on the behaviors of other drivers, who in turn, base their decisions off of other drivers. Moreover, the system is not only dependent on the decisions of individual drivers, but also on the aggregation of individual driver choices. The potential for feedback loops is thus very high, and the possibility of not reaching an equilibria is very real. Adaptive capabilities of drivers make an adaptive modeling approach more useful (and perhaps necessary).

A related issue that is likely unique to each situation concerns the decisions of drivers who choose to avoid a tolled area. What, if any, routes do they take? How likely are they to use public transit? What is the effect of additional demand on these alternate transportation modes? It is easy to see that the effects of mitigation schemes on one level (e.g. increasing the level of public transportation service) mean that the decisions on another level (e.g., install a traffic circle on Main Street) must not be considered in isolation. The takeaway here is that, although published traffic models may well give a good model of traffic flow through a network, the driver's option to dynamically avoid a tolled area can promote system instability, and provide grounds for the occurrence of feedback loops It is easy to see how they may occur in this context; a particular driver's decisions affect (at least) local traffic flow, which effects the decisions of other drivers, which further affects traffic flow and, in turn, the original driver's decisions. Other examples of feedback loops lie when moving between different levels of modeling. For example, in a congestion pricing situation, the toll price will affect traffic flow, which in turn affects the toll price, etc. It is not clear that an equilibrium will be reached.

Because traffic modeling is sufficiently complex to require different models for the different "layers" of the problem, overlapping representations present a potentially significant source of model vulnerability. Is there a clean partition between these models? If not, are the models compatible at their intersection? There are obvious cases, such as mismatched model assumptions, such as homogeneity of drivers (with respect to trip starting time, acceleration or braking habits, etc.) and the assumption of single lane roads in many common traffic models.

Another issue that may arise due to overlapping representations or adaptive behavior is the occurrence of phase transitions. A phase transition or bifurcation can occur in several ways: between two models, between a model and reality, or entirely within reality, and so the potential for error is quite large in this domain. A bifurcation between models would involve some phenomenon behaving differently in different system models. Depending on the domain of the models, this may or not present a problem. A bifurcation between a model and reality is similar, however, is more likely to result in an erroneous conclusion if the bifurcation occurs in a zone of interest. Phase transitions in traffic management occur in at least two areas: traffic jam formation and congestion pricing.

We discussed above the issues of traffic jam formation. The problem for modeling is that there is a large area where traffic jams may or may not form, and given the vast difference in system behavior between the two phases, a point estimate is likely to be a poor modeling methodology. A better approach would be to estimate the range where a jam can occur, and attempt to stay out of this range. This is consistent with an adaption or hedging approach.

With respect to congestion pricing, we recall the experiment performed in Minneapolis, where drivers responded to increasing tolls by increasing road demand. This is consistent with the findings of Brownstone et al. (2003), which studied data from a congestion pricing implementation in San Diego. The authors determined that drivers are willing to pay a higher toll to reduce travel time than was suggested by stated preference surveys and that higher tolls were viewed as an indication of higher congestion, and so the tolled lanes were more attractive in these cases. Thus, the pricing "signaled" more than just the price.

We take these studies to be evidence that consumer demand of toll roads does not strictly follow the classical economic idea of decreasing demand in response to increasing prices. At higher tolls, however, the classical assumption is likely to hold, and so it appears that the demand function undergoes a phase transition as some point, moving away from the classical demand model to its inverse. The lesson is that blind application of classical models can result in incorrect models in cases where phase transitions occur.

6.2 General Model Composition issues

A point that has been repeatedly emphasized in this report is that it is rare that enterprise systems can be represented using a single model. For most questions of interest, the complexity of the enterprise will necessitate several models, often at different levels of abstraction. The natural solution would seem to be to simply compose these models and use the resulting aggregation to search for the best answer to the question of interest.

Unfortunately, as illustrated by the congestion pricing case study presented in the previous section, this is not a trivial exercise. Model bifurcation or phase shifts are at the heart of the matter. For congestion pricing, the issue is that the classical economic demand model bifurcated from the actual behavior of drivers. So simply composing a standard economic model with a traffic flow model would not yield an accurate result in many common situations. Similarly, for counterfeit parts, we discussed the possibility that a policy that seems optimal in one level of the supply chain may induce the opposite of the intended behavior in another level. Unfortunately, it is not obvious that classical economic models would capture all of the relevant motivations of the actors in the system. Yet the appropriate non-classical views of economics may not be formalized in such manner that they can be combined with traditional engineering models. This is particularly true as we move into social and political theories that tend to be more conceptual in nature. The reasons for this were discussed in Section 4.2.3.

The only conclusion that can be reached based on the analysis documented in this report is that a general approach to composing disparate models for enterprise systems is impossible. However, this does not mean that it is never possible. Rather there is likely to be a spectrum of situations. For some questions and systems of interest, model composition will be a viable and beneficial approach. At the other extreme, for some questions, any attempt to compose models will be outright misleading even when computationally feasible. In between, there will be situations where some aspects of the problem can be addressed via model composition but others cannot. Thus, the challenges become understanding when it is and when it is not appropriate to compose models as well as understanding how the portions of the system that can be computationally modeled should relate with those that cannot in the context of strategy development. In fact, these are the chief motivations behind the modeling methodology (Section 2) and the strategy framework (Section 5).

In this section, we will address the more technical aspects of the model composition problem. That is when it is appropriate and when it is not appropriate to compose models and how one may go about it. This is by no means a complete solution. This is a domain of active research as indicated in section 4.5. To that end we will briefly discuss the work of other researchers on modeling in general and then move on to the issues specific to model composition. During RT-44a, a number of model composition heuristics were identified. Using the work discussed here, we will highlight the rationale that supports these heuristics but also reveals their limitations.

Finally, we will discuss how the model composition approach might be integrated with the strategy approach.

6.2.1 A Review of Modeling Theory

With regard to a theoretical discussion of modeling, there are several bodies of work that inform the model composition problem. First, there is the consideration of the nature and purpose of modeling. Second, there is the body of work from dynamical systems theory and bifurcation theory. Third, there are the complexity science investigations into emergence and phase transitions. Finally, and more recently, there is the explicit mathematical consideration of model composition.

As to the nature of modeling, probably the most critical work that informs the enterprise modeling problem is biologist Robert Rosen's *Fundamentals of Measurement and Representation of Natural Systems* [1978]. Rosen took a very abstract view of modeling systems. He framed the problem as one of systems inducing dynamics on meters and then investigated the implications of that setup. Outputs relevant to this investigation included the conditions for a mathematical model to represent a system, the notion of bifurcation of a representation (as discussed in section 4.4), and refutation of reductionism in the study of living organisms. In essence, he asserts that not everything can be captured by physics. What is critical is that he justifies this viewpoint with a mathematical argument as opposed to simply asserting it. It is that latter aspect that informs our position that an enterprise can rarely be captured by a single model.

This viewpoint was expanded by Casti [1985] who considered the importance of bifurcations to levels of abstraction and system complexity. In particular, Casti asserts that levels of abstraction are bifurcations, and casts complexity as the number of bifurcations possible when representing a system. Another interesting observation is that model parameters are slow changing states, decision variables moderately changing states, and outputs are fast changing states. This observation has direct implications for model composition. In particular, some of the difficulty comes from the perceived rate of change of a particular system state implicit in different models. (A slow changing state variable is easier to coordinate during composition).

The larger implication of Casti's work from our perspective is that there are many possible bifurcations for social systems; hence there are many different models and theories in social science. Each model or theory has an element of truth, but it is easy to over extrapolate, exceed the bifurcation point, and obtain bad predictions.

The idea of viewing a system from multiple perspectives has become a staple of systems engineering. Haimes [1981] was an early advocate for developing multiple models of a system, each from a different perspective. This approach was captured in the idea of hierarchical holographic modeling. Viewing a system from multiple perspectives has even become institutionalized in the US government via the Department of Defense Architecture Framework

(DoDAF) [DoD, 2014]. The importance of considering multiple perspectives is essentially unquestioned at this stage, but the challenge has been how to combine the perspectives as one moves beyond conceptual modeling and toward mathematical modeling.

Based on the above discussions, it seems that model bifurcations are what necessitate multiple models to represent a system. Bifurcation theory is a result of a very large body of work known as dynamical systems theory. An in depth discussion is not necessary for our purposes, and instead it is sufficient to note that bifurcation theory studies the circumstances when a small change in input results in a dramatic change in output for a dynamic system. In other words, there is a qualitative change in the system. Bifurcation theory has important implications for biology and ecology. In particular certain aspects of biological organisms are regulated by bifurcations (Think of it as mode or phase switching). The relevance of this work to modeling enterprise systems is that it informs an understanding of when and how complex systems undergo phase shifts.

In that spirit, researchers in complexity science have investigated bifurcations or phase shifts in more complex systems such as financial and social systems. This is sometimes referred to as emergence. This notion was discussed in section 4.4. In particular, the work of Bossomaier, et al. [2013] was presented as an example of an investigation into how empirical metrics such as transfer entropy and mutual information can presage a phase shift. As should be clear at this point, this informs our model composition problem as it provides indicators as to when we need to switch models.

Finally, some researchers have addressed the model composition problem directly via mathematical approaches. In the RT-44a final report we discussed the work of Zeigler, et. al. [2000]. They focused on establishing mathematically when one can combine discrete event and continuous time models. While important, this does not necessarily address the issue of conceptual interoperability.

More recently, Diallo, et. al. [2014] have advocated employing a branch of mathematics known as model theory to understand conceptual interoperability of models and simulations. Tolk, et. al. [2013] suggest that the implications of model theory are also applicable to the systems engineering process as a whole. While this area of research is in its infancy, it is likely to have implications for model composition, if only to provide some advice regarding what not to do.

6.2.2 TECHNICAL DRIVERS OF COMPOSITION ISSUES

While the theoretical work discussed in the previous section is certainly informative, it does not provide any practical guidance. To actually address the issues described, one must understand how they manifest themselves in real problems. To that end, we will discuss the technical drivers of model bifurcations and composition issues. Then we will discuss how this informs actual approaches to model composition.

As discussed at various points in this report, two key sources of model bifurcations in enterprise systems are the emergence of layers of abstraction and the adaptive behavior of enterprise participants. Consequently, it is important to understand how exactly these two drivers complicate composition.

Overlapping Representations

First, we shall consider the issues with overlapping representations. To facilitate the discussion, let us consider one very useful type of model, the venerable map. Maps support many different functions including: visualization, analysis, coordination, communication, data storage, and control just to name a few. Consequently, there are many types of maps such as political, topographic, population, vegetation, tactical, etc. It is a fairly obvious assertion that no one map can be made to satisfy all possible applications. Yet, people attempt this feat with mathematical models and simulations all the time!

For many real world problems, we need to take advantage of more than one map to solve a single problem. This is relatively straightforward when we can partition the problem space and obtain a clean divide between applicable maps. For example, imagine that we are planning a trip from New York to Washington, DC in the era before Google Maps. We would most likely use a low resolution map to plan our highway route between the cities and a higher resolution street map to plan our route to the hotel once we enter DC. There is no conflict between the maps because we are able to partition our problem in time and space such that there is no need to overlap the high-resolution map with the low resolution map. (Note: we are just changing level of aggregation, not level of abstraction.)

The difficulties arise when we cannot find a clean partition. Imagine that we are concerned with the possibility of climate change instigating genocide between ethnic groups. To consider this problem, we would want maps showing possible sea level rise, changes in rainfall patterns, agriculture, population densities and ethnicities, political boundaries, and even placement of military forces to name just a few. All of these viewpoints interrelate, and it is unlikely that a single partition will be found that will allow a clean separation between them. When we cannot find a clean partition, we are trying to simultaneously employ conflicting representations of the same underlying reality just as we saw in the archetypal examples (e.g., viewing traffic as both a flow and as individual drivers). Simultaneous representations can be problematic because each could, in principle, affect the state of the other, leading to vicious cycles.

The challenge of overlapping representations is exacerbated with enterprise systems because as we move up the layers of abstraction from physical systems to humans to organizations to enterprises, the number of potentially valid, relevant, overlapping representations increases. This is directly in line with the discussion in Section 4.2.3.

Adaptive Behavior

Enterprises are made up of people and organizations that are capable of prediction and adaptation. This can lead to positive feedback loops that can exaggerate tiny changes in the system state. The precise impact of these feedback loops can be difficult if not impossible to predict [Arthur 1999].

Consider the case of bubbles in financial markets. They are the result of investors getting caught in a positive feedback loop. In this loop, many investors think that market is going to go up and make investments accordingly, which causes the market to go up. This, in turn, causes investors to think the market will continue to go up. You can know that you are in a bubble market but have no way to know how high it will go or when it will burst. Classical economic models are built on negative feedback loops that push investors toward an equilibrium, and thus are poor predictors under such circumstances. In short, the financial market undergoes multiple phase shifts. Each would require a different model of investor behavior. For example, we might assert that bubbles burst when investors shift from leveraging their balance sheets, to deleveraging their assets. In some sense, their attention shifts from their income statement to their balance sheet. Unfortunately, it is difficult, if not impossible to know when to shift from one model to another.

Consequently one may be limited to establishing possible scenarios as opposed to specific predictions. (This is in line with the adapt and hedge strategies from Section 5.) For example, consider a case where a country is transitioning from an agriculture based economy to an industrial economy. People are beginning to leave their farms and move to cities. You may be able to predict that a few cities will get extremely large, a few more will be medium sized, and a large number will remain small. This is opposed to a set of equally sized cities. The problem is that it will be almost impossible to predict which cities will become extremely large. Consequently, when modeling enterprise systems to inform policy makers and decision makers, it may be impossible in many circumstances to make any sort of specific predictions. Given this intrinsic lack of predictability, it is again necessary to address this issue through an overarching approach to modeling enterprise systems as opposed to a particular modeling technique.

Bifurcations and Model Composition

The above discussions and the theoretical implications of section 6.2.1 would seem to suggest that when and how models bifurcate over the space of interest determines the amount of work required to compose them. To that end we need to relate the circumstances of model bifurcation to the composition approach.

In the RT-44a final report, we defined two types of model composition: coupling and combining. We will briefly summarize their definitions here:

- *Coupling* two models means that you can combine the existing models by coupling inputs and outputs

   - E.g., the output of an interest rate forecast model is an input into an investment portfolio model
- *Combining* two models means that you have to build a new model ontology (possibly from scratch) that encompasses the two models
   - E.g., using a molecular model for water as opposed to one model for liquid water and one model for steam

Of course, all else being equal, one would prefer to couple models as this is typically much less work. Many of the problems encountered during real world attempts to compose models are likely due to the inappropriate application of model coupling. Thus, we would like to understand how model bifurcations affect coupling and composition.

In short, if there are no model bifurcations over the space of interest, then one may be able to couple models. If there are model bifurcations, but they are well known and understood, one may be able to combine models. The combination of the models will require the introduction of state variables that drive the bifurcations. (In our earlier water example, these could include temperature, pressure, and possibly even dissolved solids depending on the scenario). The combination could take the form of an additional controller that regulates model switching when the bifurcation points are crossed, or it could be a reductionist approach that drops a layer of abstraction to capture the bifurcation.

If there are bifurcations over the space of interest but the models disagree on the bifurcation points, it is not possible to compose the models. However, one can still use multiple models and compare the results to support strategy development as will be discussed in section 6.2.4. Finally, if there are bifurcation points, but one cannot capture them computationally, then there is not really a lot that multi-modeling can do for you.

The key phase used above was “over the space of interest.” This requires some explanation. As suggested in the discussion about overlapping representations, when two models attempt to the represent the same thing but disagree on the representation, a model bifurcation is present. The driver of composition issues is overlap between the models (i.e., two models both attempting to represent some portion of the system or environment). While it may seem that the solution would be to eliminate the overlaps, overlaps are a necessity. If two models had no overlaps whatsoever, then there would be no reason to compose them, because they represent two completely isolated systems that cannot interact.

Even when we think about two spatially isolated systems exchanging messages (e.g., two command and control systems), the two systems must agree on what the messages mean about the world. Thus, their representations of that portion of the world overlap. So the key to understanding composition is to understand where two models overlap. This can be extremely subtle and difficult for many real world problems as there are often implicit assumptions in standard approaches that are so common that they are rarely questioned or thought about. Similarly, we can think about the “black box” problem of composing independently designed models as the case where there are overlaps but the representations of the overlaps are unknown. At the opposite extreme, a partition as described in the overlapping representation discussion is the case where the overlap is well-defined and the two models agree on the representation of the overlap. Thus, we can have two models and a well-defined interface between them. In between, we have the case where the two models overlap but a bifurcation occurs in the representation. Thus, it is the overlap that constitutes the space of interest in the composition discussion.

This concept is best illustrated using examples. Let us consider the relatively simple case of gravitation:

- Modeling the Earth going around the sun:
  - The Earth's mass is negligible compared to the sun so we can treat sun as a point mass and just compute the Earth's trajectory around it
  - The sun's mass is a parameter and the point of model overlap
- Modeling the moon going around the sun
  - The moon's path around the sun is affected by the Earth
  - If we are willing to sacrifice a little accuracy we can pre-compute the Earth's orbit around the sun and then feed that into another model that computes the orbit of the moon around the sun
  - The Earth and Moon models are sequentially coupled
  - The points of overlap are the are the sun's mass and the orbit of the earth
- Modeling a massive three body system
  - Each body affects the trajectory of the other
  - The trajectories must be computed simultaneously
  - The models of the three bodies' trajectories must be combined into a single model
  - Since each model affects the other, the representations completely overlap
  - If we were treat one of the trajectories as fixed, the models would quickly bifurcate from that fixed trajectory as the other bodies actually affect that trajectory
  - In essence, the combination resolves what would otherwise be a bifurcation

In the above examples, we see hints of Casti's partition of states into parameters, inputs, and outputs. In the first example, the Earth does affect the position of the sun, but for most practical purposes it is negligible. (Though this assumption highlights the importance of question of interest. This assumption may not be accurate for certain questions.) Thus, the sun is a slow changing state for our system and can be captured as a parameter.

In the second example, the earth's position changes, and this affects the position of the moon. Both the representations of the moon and the Earth (for this scenario) agree on this area of overlap. In our example, we assume (with some loss of accuracy) that the moon does not affect the movement of the earth. Thus both the model of the Earth and the model of the moon agree on the representation of the Earth's trajectory. It has a moderate rate of change (from a modeling perspective). Thus, the earth's trajectory becomes an input to the moon trajectory model, and we can sequentially couple the two models.

If the assumption that the moon does not affect the motion of the Earth is not correct, then we can no longer sequentially couple the models because they will bifurcate on the representation of the Earth's trajectory. This is the case we capture in the last example, the three body problem. In that case, the only thing we can do is create a unified model to rectify the bifurcations.

Common Composition Situations

To take these notions a step further, let us next consider three typical scenarios where multiple models are employed:

1. Using the best (or only) representation for each portion of the system space
2. Using models at different levels of fidelity for computational advantage
3. Using multiple models to bound a phenomenon

In each scenario we will consider how the model bifurcation impacts the modeling situation.

Scenario 1: Using the best representation for each part of the system space

This scenario is typical of many enterprise modeling situations. The counterfeit parts model presented in this report is one such example. A much simpler example would be modeling a health care system using an economic model to represent the health care input costs, a queuing model to representing the care process, and a Markov chain model to represent disease progression and patient responses to treatment.

In some cases, the motivation is simply to avoid reinventing wheel and use the best of breed. In other cases, multiple models are a necessity. In either case, the modeler wishes to minimize overlaps, but some are necessary so that the models can communicate with each other. (E.g., the economic model and the process model need to agree on what the input costs are). When models are developed independently and composed later, this risk of overlap issues emerging is greater since they are not explicitly controlled during development.

Composition difficulties occur when two models disagree about the representation of an area of overlap. In other words there is a risk of model bifurcation in overlapping regions. In the case of the counterfeit parts model, this was controlled through simultaneous and coordinated development of the component models (i.e., there is no black box). To consider how issues might arise, let us proceed with the health care example where the models were independently developed (i.e., off the shelf models)

For example, if the economic model and the care model must agree on the dynamics of a cost variable, the likelihood of a composition problem is much greater than if they just need to agree on the meaning of the cost variable. If the care model only needs to periodically get a cost value from the economic model, then they only need to agree on the semantics. The models may be coupled. If on the other hand, the care model implicitly assumes that costs are following a mean reverting process and the cost model is using a geometric Brownian motion process (two common stochastic price processes), then they are dynamically inconsistent. Even in they start in the same place, they will diverge as the model executes. The models bifurcate relative to the overlapping cost variable and cannot be coupled. One or both of the models will need to be modified to combine them.

A more prosaic, but still important example arises in the interface between the care model and the disease progression model. Sometimes the progression of a disease is captured as a discrete time Markov chain. This is logical given the circumstances as the patients are typically evaluated on a periodic basis (e.g., annually). The problem is that the care process is modeled using a discrete event simulation that operates in continuous time. Thus, if one wanted to use an off-the-shelf discrete time model for disease progression model, it has a fundamentally different representation of time. Each representation of time is logical in its own domain, but incompatible nonetheless. Thus the two models could not be coupled. One or both of the models would need to be modified to combine them. (The representation of time needs to be rectified to eliminate the bifurcation so that the two models can be combined).

Scenario 2: Computational Advantage

High fidelity models are often expensive in terms of computation. As a result, it is often impractical to use high fidelity models to search a large design space. In response, a common approach among engineers is to use the high fidelity model to parameterize a low fidelity model. The low fidelity model is then used to the search the design space for promising solutions. This is sometimes called multi-resolution modeling.

For example, a company exploring possible designs for a new car generates a parametric model that it allows it to trade off key design parameters. As part of the development of the parametric model, the company runs a high fidelity simulation of an engine over a wide range of parameter values and fits a response surface to the results. The response surface is used in the design trade model for computational advantage.

In this case, there is a deliberate and complete overlap of one model by another. This set up is viable as long as the two models do not bifurcate. In line with the theoretical discussions above, it is okay to use parameters in the low fidelity model as long as they are “slow changing.” In other words, changing the decision variables in the low fidelity model does not affect the value of the parameters. This allows the models to be sequentially coupled. If that were not the case, the models would bifurcate, and this setup would not be valid.

What makes this problem challenging is that “slow changing” and “fast changing” are relative concepts. Over a large enough time or space, even slow changing parameters change noticeably. This can happen in the multi-resolution modeling scenario when the analysis on the low fidelity model pushes the system out of the acceptable range for the pre-computed function or parameter value from the high-fidelity model. For example, imagine a case where the automotive company pre-computes the engine performance for a cold weather environment and parameterizes the low fidelity model. The low fidelity model is then run for a hot weather scenario. This may cause the engine model to bifurcate from the low fidelity model for that scenario. What is particularly concerning is that this may not be evident if one is just running the low fidelity model over a range of scenarios. (See the discussion in Section 4.4 about the inability to detect some bifurcations within the model).

This problem could be remedied by rerunning the high-fidelity model again but if this occurs too frequently, it starts to erode the computational advantage. In the extreme case, the models have a mutual state dependence that results in a model induced feedback loop.

For example, a decision model for an investor is used to model investment decisions based on the current price in a higher level market model (an example of adaptive behavior), but the current price depends on the decisions of all of the investors. So there is a feedback loop between the two models.

Scenario 3: Bounding phenomena with multiple models

Sometimes there may be a phenomenon of interest, say the failure temperature of a system, but none of the available models is completely accurate. Consequently, all of the models are run for the same scenario(s) and compared. This is effectively an instance of a bifurcation in the real system, and models are used in an attempt to bound the bifurcation.

We have already discussed two examples of this scenario in this report: The water example from Section 4.4 and the traffic jam density discussion from Section 6.1. To elaborate a bit more on the traffic modeling, [Daganzo, et. al. 2011] compare the jam densities predicted by a flow-based model versus an agent-based model. The jam density is a bifurcation point because it signifies a phase transition in the system: from free flowing to jammed. The researchers found that the flow-based model over-estimated the jam density while the agent-based model underestimated it. In essence, the two models bound the bifurcation point. This has important implications for strategy as will be discussed in Section 6.2.4.

6.2.3 COMPOSITION APPROACHES RELATED TO HEURISTICS

In the RT-44a report, we presented a number of heuristics that are used to compose models. In this section, we will revisit these heuristics in light of theory presented in this report. Generally speaking, the heuristics deal with coupling situations. Most of these revolve around converting potential two-way dependencies in the area of overlap into one way dependencies. In short these heuristics are predicated on the implicit assumption that the models will not bifurcate over the space of interest.

Heuristic 1: Assume that the impact of one model is negligible on the other

This heuristic assumes that the models do not bifurcate with regard to the portion where they overlap. Examples include the orbital mechanics example presented earlier (the Earth does not affect the movement of the sun) and the cost parameter from the care model. This heuristic is valid as long as one model is “slow changing” relative to the other. Of course, this depends on the question of interest. Consequently it is important to identify the model bifurcation points to ensure the validity of the assumption. Note that the challenge here is that these bifurcation points may not be discernable from the models themselves and need to be found empirically.

Heuristic 2: Assuming that there is a lag in the impact of an event in one level on the state of another

The example presented in the RT-44a report for this heuristic considered that fact that a large equity order by a major institutional investor will impact the market price of the equity, but it takes time for the news of the order to disseminate and the transaction to execute. Consequently, the institutional investor can make a decision based on the current market price, but the actual impact of the order on the market price will occur at a future time.

Interestingly, a bifurcation relative to this example has already occurred in real-life in the time since the RT-44a report. High frequency trading (HFT) is based on invalidating this assumption by front-running large institutional orders. (HFT already existed at the time the RT-44a report was written but was not widely known until Michael Lewis published *Flash Boys* in March 2014).

So the bifurcation here has to do with time. One can make the argument that nearly every phenomenon takes at least some miniscule amount of time to affect another phenomena (at least for most practical situations). Thus, this heuristic works by assuming that one system being modeled is sufficiently fast relative to the other that the other can be viewed as "slow changing." This allows the sequential coupling to occur without a bifurcation.

Once again, the challenge is to find the bifurcation points to ensure the validity of this heuristic. High frequency trading is a perfect example. The increased speed of high frequency traders allows them to see the institutional investor's order before it is complete … thus altering the price. Of course the emergence of HFT was not predicted by anyone. This example is also suggestive in terms how one might develop policies and countermeasures for enterprise systems.

Heuristic 3: Rolling up the state response behavior of the lower level to one that is compatible with the higher level

This heuristic was applied in the Scenario 2 in the previous section with regard to the automotive design model. Once again we see that fundamentally the issue is about assumptions of "slow changing" parameters and model bifurcation.

Heuristic 4: Externalizing state variables as decisions

The example of this heuristic presented in the RT-44a report involved a CEO that would like to assess the impact of his decisions on the operation of the company. Even though, in real-life the state of the company would impact his decisions and his decisions would impact the state of the company, the CEO would like to perform "what if" analyses on decision alternatives. This would make management representation external to the rest of the model and creates a one-way dependency. This heuristic obviously fits the established pattern as we see, yet again, that this heuristic involves essentially the same underlying assumptions as the others.

Thus the conclusion we can draw with regard to all of these heuristics is that, in a sense, they are all the same. They are all implicitly assuming that two models do not bifurcate relative to each other where they overlap for the question of interest. Alternatively expressed, one changes slowly relative to the other. In all cases, we are concerned with finding the bifurcation points so that we know whether or not these heuristics are valid for any given analysis.

However, we can turn this around a bit and infer that these heuristics are situations where modelers have ascertained through experience that the models typically do not bifurcate. This allows one to greatly simplify the modeling problem by assuming that one model changes slowly relative to the other. The challenge for enterprise systems is that many traditional modeling and simulation techniques were developed for simple physical or technical systems (simple at least relative to enterprises). As has been repeatedly pointed out in this report, enterprise systems are much more complex, which means that there are many more bifurcation points lurking about. This makes model composition for analyzing enterprise systems a challenging prospect. The HFT discussion is a perfect real world example of this challenge.

So we can only conclude that, for many enterprise problems, only part of the enterprise will be represented as a composed model. Thus, we need to consider how to integrate this with strategy development.

6.2.4 INTEGRATING STRATEGY AND COMPOSITION APPROACHES

If we consider the fact that both the strategy approach and the model composition approach depend on the presence and nature of model and enterprise system bifurcations, it stands to reason that the two can be approached in an integrated fashion. To that end, we propose the following guidelines for integrating the modeling approach with the strategy approach. These hypothesized guidelines will need to be investigated in greater depth in future work.

- If there are no model bifurcations over the space of interest
  - **Couple models and optimize**
- If the models bifurcate, but the bifurcations are known and understood
  - **Combine models and optimize**
- If there are bifurcations, but the models disagree on the bifurcation points
  - **Use multiple models and/or experimentation to bound the bifurcation points then adapt or hedge**
- If bifurcation points cannot even be bounded
  - **Use experiments or observational data to adapt or else accept**

We can consider the first two cases as the traditional engineering cases. If the system is fairly modular or we can safely apply some of the heuristics we discussed in the previous section, we can couple the off-the-shelf models, search the design space, and find the best answer to our question.

Alternatively, if there are bifurcations in the system, but they are well understood (often empirically and/or through more detailed models), then it is more work, but we can build a combined model of the system that captures the bifurcations internally to the model. Once again, we can use the model to search the design space for the best option.

While no real system ever truly meets the requirements for the first two cases, for many traditional engineered systems, they are close enough. As the complexity of the system of interest increases, we approach the third case. Here we know there are bifurcation points, but we don't know exactly where they are. The objective here is to use a combination of multiple models and/or experimentation to try to bound the bifurcation points. If one can at least bound them, then it makes it possible for one to consider taking a hedging approach in which one invests resources to address the event where the bifurcation point is crossed. Alternatively, one may feel that he or she can react quickly enough to employ an adapt strategy. Note that no model composition occurs for this case, though models will be compared. This precludes any algorithmic search for an optimal solution.

Finally, in the last case, we cannot even bound the bifurcation points. Computational models are not of much if any help at all. Thus, all one can really do is remain vigilant and adapt…or simply accept the risk. Note that the aforementioned work in complexity science to identify leading indicators of impending phase shifts could be of use in support of an adapt strategy here.

Note that enterprise problems tend to push into the last two cases more so than traditional engineered systems. In all likelihood, most enterprise problems will involve a mixture of these cases. As discussed in the traffic congestion case study, some portions of the larger problem are resolvable with traditional modeling and optimization approaches while others are not. Consequently, enterprise problems will often entail the simultaneous consideration of multiple independent models and data sets. If these models cannot be integrated computationally, there is only one place left for their implications to be integrated, the human mind. Consequently, in the next section we will consider the role of visualization in doing just that.

7 VISUALIZATION OF ENTERPRISE SYSTEMS

Visualization features prominently in the enterprise modeling methodology described in Section 2. In particular it serves as a means to facilitate communication and understanding among the typically diverse stakeholders of an enterprise. But, as discussed in the previous section (Section 6.2.4), enterprise problems will often necessitate the simultaneous consideration of multiple models and data sets that are not fully composable. Since they cannot be composed computationally, the burden of drawing inferences from these independent models falls on the human. As a practical matter, this can only be done using visualizations.

Consequently, this section considers the development of visualizations to support enterprise decision making. In particular, the emphasis is on supporting the decision makers and other enterprise stakeholders in terms of causal inference. Visualization is a well-established area of research with an extensive history. Thus, we begin with a review of the relevant literature, consider how this influences enterprise visualization, and identify outstanding questions that need to be resolved. Then we proceed to discuss a conceptual visualization design developed as part of this research task. (A larger version of this conceptual design is provided in a separate file.) The intent of this conceptual design is to illustrate how the principles discussed in the literature review can be applied to support enterprise decision making. In particular, it demonstrates the different components and functions required by a visualization to support causal inference for an enterprise system. (Note that as part of the investigation into visualization, we conducted a survey of off-the-shelf visualization tools. The results are included in the appendix to this report.)

7.1 REVIEW OF THE VISUALIZATION LITERATURE

Generally speaking, engineering interest in visualizations stems from a desire to identify casual relationships among phenomena. From an enterprise system perspective, the issue is the complexity of the enterprise. Not surprisingly, the difficulty of understanding causal relationships increases as a system becomes more complex [Rouse and Serban, 2011]. The application of a holistic model with multiple layers of abstraction can help decision makers to predict the behavior of a system, manage the current situation and make future decisions via the comparison of different strategies in a large enterprise [Nightingale and Rhodes, 2004]. The problem is that for complex cases, most of our understanding is derived from our tacit reasoning based on life experience as opposed to a particular model.

Thus, the issue at hand is how to develop visualizations in a principled fashion in order to support the human reasoning process effectively. Consequently, we need to understand how humans collect data and then derive conclusions from it. [Kodagoda, et. al., 2013]. Unfortunately, the academic literature on visualization seems to be lacking in this regard. Chen [2005] notes that there is a disconnect between studies that have been done on high-level user tasks and those evaluating the usefulness of visualization. Furthermore, “there are only few empirical studies that link cognitive processes, intelligence process and sense making into context of expertise.” [Pirolli & Card, 2005]

Information visualization has two major aspects: “Structural Modeling” and “Graphical Representation”. Structural modeling concentrates on detecting and understanding the underlying relationships that form a structure such as the complex network structure associated with enterprises, at levels ranging from the economy, the industry, a company, and its specific products and services. Graphical representation, in contrast, addresses portrayal of inputs and outputs of an enterprise such as production levels, market shares, revenues and profits. What we are interested in for this work is structural modeling. Chen [2006] identified several problems in structural modeling such as scalability, content similarity, state transition pattern, and structuring heterogeneous information, and he suggested using techniques like multidimensional scaling, link-reduction in graphs, clustering and classification, incremental clustering, and calculating proximity and connectivity to name a few.

Pirolli and Card [2005] developed an intelligence analysis model (see Figure 10) based on expert behaviors. The model consists of two major loops: one for data foraging and one for sense-making. The foraging loop includes searching for information, filtering, reading, and mining the information. The sense-making loop involves iteratively developing a mental model. Their model captures the feedback loops present in this process and is appropriate for both top-down or bottom-up approaches***.***

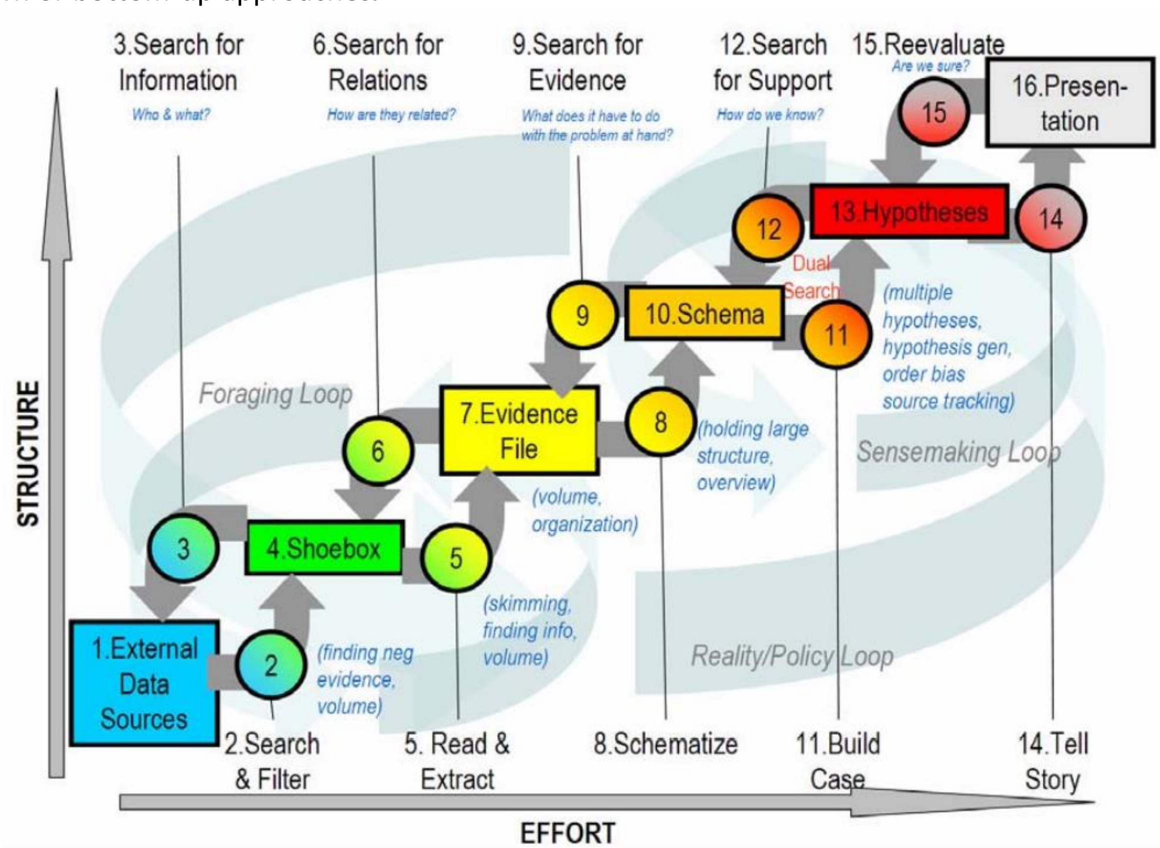


Figure 10 - Notional model of sense-making loop for intelligence analysis. [Pirolli and Card, 2005]

Pirolli and Card's model provides a natural framework to guide the development of a visualization as each activity within both the foraging loop and sense-making loop would need to be supported. Of course the activities are dynamic, and there are trades involved in how much effort to put into each activity. This tradeoff can vary depending on the question and system of interest. In particular, foraging is a tradeoff among the three processes of exploring, enriching and exploiting [Patterson, 2001]. For example, imagine that we have a large data set, and we want to draw inferences from that data set. We would likely start with highly abstract/low precision query and successively narrow it over time to achieve increasingly small data sets until we can reach a hypothesized relationship (see Figure 11).

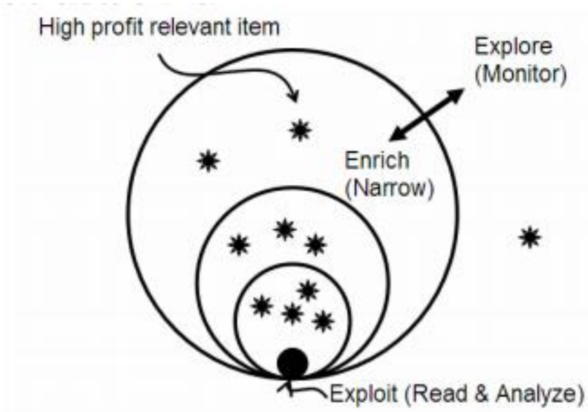


Figure 11 - The exploration-enrichment-exploitation tradeoff in information foraging. [Pirolli and Card, 2005]

The foraging loop feeds the sense-making loop. As we are interested in identifying causal relationships within enterprise systems, it is important to note that causal reasoning is a form of sense-making. Klein's [2007] Data Frame model is instructive for understanding sense-making. This model consists of four stages: connecting data to a frame, elaborating a frame, questioning the frame, and reframing. Any visualization developed would need to support these activities. If we apply this approach to Klein's model we come to the revised model that is shown in Figure 12. For causality, in particular, we are concerned with three attributes: propensity, reversibility and co-variation. Of course this alone does not tell us how the decision makers collect the pertinent information and how they decide to reject a coincidence. To this end, interactive visualization, as opposed to a static view of the system, may be useful in the sense that it helps the decision maker explore and investigate the plausibility of a hypothesized cause, understand the potential consequences, and consider what actions to take.

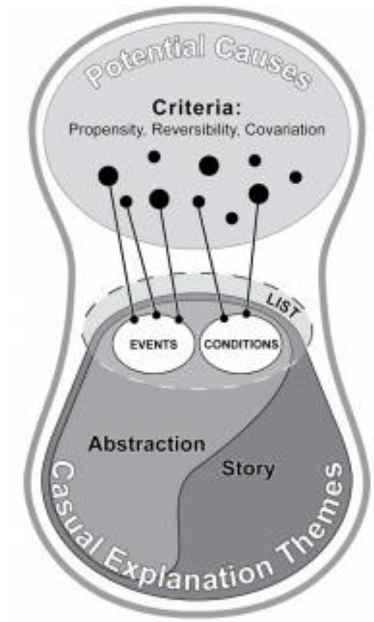


Figure 12 - Data Frame model revised to feature the causes and the frames (Klein & Hoffman, 2009)

The introduction of visualization into this space creates a situation where we have three systems in interaction: information foraging, sense-making, and the visual system. When we apply this trio of systems to a complex enterprise system, the situation becomes much more challenging. In particular, the complexity of the enterprise system and the difficulty of the problem of interest affect the abstraction model employed*.* Previous work found that the appropriate abstraction level of a display is affected by the type of task and the user's level of expertise [Frey, Rouse, & Garris, 1992]. The issue for enterprise systems is that we do not know if a given number of layers of abstraction faithfully presents the essence of our case. We also do not know if the same number of layers of abstraction should be used for every question, or how it will affect the accuracy and response times of users with different levels of expertise. Of particular concern is that some simplifications could result in negative consequences. Thus, how to determine the optimal number of layers of abstraction to visualize is an open question. It is also unclear how an "optimized" visualization with n-layers of abstraction will affect the time required to reach appropriate inferences regarding the enterprise system.

It stands to reason that a well-designed interactive visualization based on well-articulated use cases will result in more accurate situational assessments, more rapid problem identification, better causal inferences, and better strategic decisions. The question is how to integrate combinations of qualitative data along with quantitative data typical of enterprise systems in a single set of displays. How many different layers of abstraction (i.e., views) of the system should we have? How should we facilitate user movement among the views? How can we alleviate the cognitive load put on the user when having to deal with multiple system views? How can we assist the user in performing the tasks necessary to establish causality (e.g., testing hypothesized relationships)?

7.1.1 WHY VISUALIZATION?

Applying visualization to technical and business problem solving is essentially an attempt to leverage the impressive capabilities of the human visual system that has benefited from millions of years of evolution. For this reason, the literature sometimes employs a hunter-gather analogy (hence the term foraging). A hunter-gatherer or forager has to travel long distances to search for his or her food. Obviously the search process consumes time and energy but is also necessary to obtain more energy. This imposes an optimization problem of how to search to maximize the return on investment. In modern technological societies we are faced with the analogous dilemma of how to forage for information to maximize our return on investment.

Of course the basis of visualization is human physiology. It is these capabilities that visualization seeks to leverage. For example, the human visual system can recognize an animal in a complex scene in less than 150 ms. Thorpe, Denis, & Marlot [1996] performed an experiment where subjects had to decide after looking at an unseen picture for 20 ms if there is an animal in the picture or not. (Of course, this was very useful when predator versus prey recognition was important to decision making.) It is the rapid ability to filter and interpret complex situations that is of interest for enterprise decision making. Understanding how this system works informs how to best construct visualizations.

While a complete discussion of this extensive body of research is outside of the scope of this report, some representative findings are sufficient to illustrate the point. For example, the visual processing system can operate in two modes: a bottom up, task independent mode or a top-down task-dependent mode. The bottom-up approach which is also known as saliency-driven is faster than task-driven, and can be affected by shape, color and texture of the objects. [Pirolli and Card, 2005]. Understanding such principles feeds into visualization design choices. Marr [1982] portrayed the visual system as an information processing system that receives a two-dimensional array from the retina and develops a three dimensional explanation of the world. Biederman [1987] divided the object recognition process into two stages of segmentation at deep concavity area and organization of basic geometrical forms like cylinders, etc. He also estimated the number of objects that human can distinguish and concluded that an 18 year-old man can possibly distinguish around 30,000 objects. When we are concerned with interpreting multiple visualizations, short-term memory becomes a limiting factor. Miller's [1956] seminal work on short term memory indicated that the number of objects we can remember after a presentation is seven for digits, six for letters and five for words, and he calls it "The magical number seven plus minus two".

For the purposes of this effort, we take this body of work as given. We intend to leverage it to inform visualization design choices. However, visualization, per se, is not the primary focus of our effort as the design of visualizations is largely established at this point.

7.1.2 WHY INTERACTIVE VISUALIZATION?

The introduction of modern information technology has allowed us to consider interactive visualizations. Most modern interface designs are interactive. According to some researchers, this enhances analytical ability [Takken, 2013; Kirsh and Maglio, 1994]. To illustrate the value of being able to interact with a visualization, consider information search as a motivating example. In the traditional view, a user has to develop a single query that is understandable by a search engine that then produces a single output. Gerard Salton [1968] introduced the idea of using iterative feedback to improve user queries. While this model improved the way we satisfy an information need, it presumes we know what we need. In real life, users start with one feature of the topic of interest then browse through different sources. Each source brings up new ideas and directions and consequently leads to a series of individual sources. This methodology is called evolving search [Bates, 1989]. Bates proposed a model called berry picking. The model is inspired by picking huckleberries in a forest. Huckleberries are scattered, and they don't come in bunches. It is a semi-directed model that combines the traditional method with the browsing behavior of real humans. Similarly, when we consider a decision maker analyzing an enterprise, we must consider the fact that he or she may not know what he or she is looking for, e.g., the source of a performance deficiency. Thus any visualization must support browsing behavior, and an interactive visualization is particularly well suited to that task.

Returning to the foraging metaphor, a predator has two decisions: search or eat. When it decides to search, it must determine how much time to spend searching for prey. Any prey caught results in fixed amount of energy intake. However, the predator has no control over how long it takes to find the prey. There is a relationship between the time predator searches and the amount of energy intake, assuming that predator can recognize the type of prey and can decide to eat if the prey is large or to leave it if it is small [Stephens, 1986]. The information forager has a similar dilemma: whether to keep searching for better information or consume the information one already has. This analogy led a number of investigators to apply biologically inspired models to analyze the problem of information foraging.

Information visualization facilitates finding connections, developing hypotheses, and evaluating available evidence [Chen, 2005]. Piroli and Card's [1995] visual searching model has three components: decision assumptions (what problem to analyze), currency assumptions (how to evaluate our choices), and constraint assumptions. Constraint assumptions include factors such as task structure, interface technology, and users' knowledge of that technology. Newell [1994] outlined the timescales of activities to three bands: the cognitive band that mediates between 100ms and 10s, the rational band that can last from minutes to hours, and finally the social band that might take days to months to be understood by users. According to Piroli and Card's [1995] Information seeking and sense making spans from middle of the cognitive band, across rational band, and into social band. The implication is that foraging and sense making may occur across several periods of usage on interactive visualizations, e.g., after-work chats among maintenance technicians may lead to innovations in how best to use maintenance information systems.

Pirolli and Card [1995] used two biologically inspired models to analyze the scatter/gather document browser [Cutting, et. al., 1992]. The information patch model has a variety of food patches and an organism has to decide how to allocate its foraging time. The information diet model helps with the selection of the optimum mix of prey and adds a new model called dynamic foraging which takes into account both a person's choices and changes in state of mind. Pirolli and Card developed their model using the NIST Tipster corpus that contains millions of documents from different sources.

Another model is Holling's Disc equation [Holling, 1959] where R is the rate of currency intake (i.e., the value of information in our case) and U is the rate of information intake in the technology domain or energy intake in the biology domain. U is calculated by subtracting the expended energy or information from the gross amount of foraged energy or information. Both search and handling time are included in this model where Ts is the search time, s is search cost per unit, λ is the rate of encounter items per unit time, $\bar{u}$ is the average rate of currency intake and $\bar{h}$ is the average cost of handling.

$$R = \frac{U}{Ts+Th}\,.$$

$$\bar{u} = \frac{U\mathrm{f}}{\lambda\,\mathrm{Ts}}, \bar{h} = \frac{Th}{\lambda\,\mathrm{Ts}}$$

$$R = \frac{\bar{u}\lambda\,\mathrm{Ts}-\mathrm{sTs}}{Ts+\bar{h}\lambda\,\mathrm{Ts}} = \frac{\bar{u}\lambda-\mathrm{s}}{1+\bar{h}\lambda}$$

Using a stochastic model, Stephens and Charnov [1982] made a general model for the foraging process utilizing a Poisson process for path and prey encounter and found that the limiting distribution of energy intake reaches normality according to the central limit theorem as foraging time increases. This supports the use of Holling's Disc equation, and provides a quantitative way to consider foraging behavior.

While the biological metaphor for visualization and information search provides a theoretical perspective, other researchers like Rouse and Frey [Frey, Rouse, and Garris, 1992] conducted empirical studies to understand how people interact with and utilize visualized interfaces. They developed guidelines for selecting the appropriate levels of abstraction and aggregation for maintenance displays consistent with the maintenance personnel's mental models. More specifically they conducted five experiments by developing an information display system for the blade fold system for the SH-3 helicopter. The experiment was a three-factor design: level of abstraction, level of aggregation, and level of experience. Three levels of abstraction were used with the following highest to lowest order: flow, schematic and location diagrams, and the three levels of aggregation with the same order were systems, assemblies and subassemblies. The measurements were the list of displays that subject accessed, the order that each were used and length of time that they were viewed [Frey, Rouse, & Garris, 1992].

The results showed that the appropriate level of abstraction was significantly influenced by task type. Tasks were, problem solving, circuit tracing and following procedures. The high and medium levels of abstraction were used more in problem solving while the lowest level was used more for following procedures. The level of aggregation was mostly used from higher to lower order in circuit tracing and following procedures. The medium level of aggregation was used in problem solving and circuit tracing, and the low level of aggregation in following procedures and circuit tracing. In problem solving tasks, the high level of aggregation was mostly used in the first half of the experiment and medium level of aggregation was mostly used in the second half of the experiment. Finally, the experiments also showed that more experienced personnel spent most of the time on high abstraction displays and less time on medium abstraction displays for problem solving tasks [Frey, Rouse, & Garris, 1992]. It is these latter results that suggest the importance allowing users interact with multiple layers of abstraction within a system visualization.

Peters and Itti [2007] envisioned the future of interactive environments as “attention-aware” with the capability to predict, react and influence human visual attention. They emphasized the importance of understanding the different aspects of visual perception that identify locations that attract human gaze in a complex interactive visual interface. They compared several neurobiological inspired heuristics with gaze-tracking while five observers played a game. They determined that dynamic objects attracted more attention than static objects [Peters and Itti, 2007].

Other scientists investigated the impact of manipulation of information on sense-making activities and analytical reasoning. Their findings support the hypothesis that direct manipulation of an information display enhances the analytical reasoning [Takken, 2013]. Kirsh& Maglio determined that data manipulation can help to uncover information that is out of sight or hard to calculate [Kirsh & Maglio, 1994].

7.1.3 MENTAL MODELS AND CAUSAL REASONING

The ultimately goal of interactive visualization for enterprise decision making is to support causal reasoning. To that end, we need to understand how people build mental models. Donald A. Norman argues in his book “Things that make us smart” [Norman, 1993] that cognition has different dimensions, but we can categorize them into two main groups, experiential (look, see, respond mode) and reflective (learn, think, reflect mode). Effective decision making requires both experiential and reflective modes. What is important is to achieve the right proportion of reflectiveness. The reflective mode leads to advanced understanding and novel responses, which helps us to develop the best strategy. The experiential mode helps us with reactive, automatic decisions that require stored information. Reflective thought depends in our ability to draw inferences from stored data and to follow the chain of reasoning, which is time consuming. External aiding can facilitate reflective cognition by enabling a deeper chain of reasoning and also plays the role of external memory.

Rouse and Morris [1986] studied the mental models that we develop of the system that we interact with. They investigated different domains and identified the forms, structures and parameters of mental models. They defined mental models as the mechanism whereby humans become able to generate explanations of system objectives and functioning, observe system state and predict future state. They suggested that characteristics and forms of mental models are defined by the location of the task in the domain distinctions described by Newell and Simon [1972]. Rouse and Morris characterized domains in terms of the nature of model manipulation and level of behavioral discretion. They also described novice-expert shifts as transitions from representational to abstract models. Finally, they summarized more pragmatic considerations associated with identifying or inferring mental models.

Rasmussen [1983] studied behavior in familiar and unfamiliar environments. He argued that in a familiar situation we operate via a set of proven rules that orient us towards the goal. In contrast, in an unfamiliar situation we are goal-oriented, but we use trial and error to reach the goal. In his model, familiar situations are divided in two levels, skill-based and rule-based. Skill-based choices are based on previous training and skills that are acquired through years of experience. This level is completely unconscious and activated by signs. A sign is a situation or proper behavior from previous experience that activates predetermined actions or manipulations. Rule-based choices, on the other hand, operate based on stored rules triggered by received signals. “Signals are sensory data representing time-space variables from a dynamical spatial configuration in the environment, and they can be processed by the organism as continuous variables.” [Rasmussen, 1983]

Unlike Norman, Rasmussen limits routine cognition to familiar situations and called the unfamiliar level “knowledge-based”. In Norman’s interpretation, knowledge plays an important role in an unfamiliar environment. In contrast, Rasmussen [1983] states, “Symbols are abstract constructs related to and defined by formal structure of relations and processes which by convention can be related to features of external world.” Reflective cognition is a major component. Both agree regarding the role of mental-models in the knowledge-based level and the role of symbols in unfamiliar cases.

Rasmussen’s model has five levels of abstraction: physical form, physical function, functional structure, abstract function and functional meaning or purpose [Rasmussen, 1983]. The levels provide a potential framework for supporting the development of mental models via visualization in that any visualization design needs to consider when and how to portray these layers of absraction relative to the question and enterprise of interest.

Of course we not are interested in facilitating the development of just any mental model. The objective is to aid enterprise decision makers to develop correct and useful mental models. Consequently, representations of causality are of critical importance. In other words, any interactive visualization should facilitate the user testing his or her mental model for usefulness in causal reasoning.

While the exact meaning and conditions of causality are still matters of philosophical debate, in this report we take a pragmatic view. For our purposes, we are concerned with enabling a decision maker to have some confidence that he or she understands the outcome of potential courses of action. Thus, we do not intend to discuss every view on causality, but rather just those that are useful for our purposes.

Hume [1739-1740] suggested that the cause of an effect can be identified by three features: propensity, reversibility and covariation. Propensity is the requirement that it be possible for the proposed cause to lead to the effect. Reversibility tells us that the effect will disappear if the suggested cause is removed. Kahneman [1990] called this feature counterfactual reasoning. Covariation is the observed coincidence between causes and effects and can be investigated through statistical methods [Klein & Hoffman, 2009].

Thus, we can think of the mission of an interactive visualization as follows: The interactive visualization should allow the user to search through the available information and identify potentially important pieces of information. Next, the interactive visualization should aid the user in making sense of the information found by facilitating the identification of relationships among the pieces of information. This should result in the formation or elaboration of a mental model that represents one or more hypothesized causal relationships. Finally, the interactive visualization should enable the user to test the hypothesized mental model and determine whether or not it meets the requirements of causality.

7.1.4 VISUALIZATION DESIGN PRINCIPLES

Given that we have established goals for interactive visualization, the question remains as to how to go about designing a visualization that achieves those goals. A number of studies have evaluated the efficacy of particular design choices. For example, the use of large screens and especially multiple screens has become common despite the fact that they can increase the error rate in complex scenarios [Geiser, 1976]. However, other research has shown that using techniques such as grouping and integration can eliminate these types of errors [Mitchell and Miller, 1983]. The integration structure not only groups the data based on user tasks, but also shows preprocessed data based on the required level of abstraction.

Other researchers have found that among paging, scrolling and windowing design options, paging and windowing gave faster and more accurate results [Bury, et. al., 1982; Schwarz, et. al., 1983]. Henneman and Rouse [1984, 1986] compared different levels of hierarchies to find a tradeoff between the amount of information in each page and the number of pages. They found that increasing the number of levels of in a hierarchy or increasing the number of pages will decrease the task performance.

While there are a number of such studies, gaps remain. Chen [2005] addressed information visualization gaps in his paper, “Top 10 Unsolved Information Visualization Problems.” The gaps he identified are:

1. Usability
2. Understanding elementary perceptual-cognitive tasks
3. Prior knowledge
4. Education and training
5. Intrinsic quality measures
6. Scalability

7. Aesthetics
8. Paradigm shift from structures to dynamics
9. Causality, visual inference, and predictions
10. Knowledge domain visualization

Note that several of these gaps are the topics we have just discussed as critical to enterprise visualization. Thus there is a real concern that many common visualization techniques are not built on a firm foundation. For example, a common technique in visualization is the use of the rainbow spectrum to display quantitative information. Unfortunately, this technique can be misleading. Researchers have shown that the use of rainbow colors in most cases is obscuring, deceptive and perplexing [Borland & Taylor II, 2007]. In spite of this result, rainbow color maps are heavily utilized in data visualization. Another example of a common but distortionary technique is using 3D shapes on a 2D screen [Todd, 2004].

To summarize, past research suggests that interactive visualizations can potentially provide the aiding needed by enterprise decision makers to build mental models and make causal inferences. However, there appear to be substantial gaps in the how one should actually go about building an interactive visualization for an enterprise system. These lead to several outstanding questions and issues that must be resolved for interactive visualization to reach its full potential.

7.1.5 OUTSTANDING QUESTIONS AND ISSUES

From the above discussion, we know that it is important to organize visualized information into structures and groupings based on function and level of abstraction. This would seem to push us toward a modular approach to interactive visualization where we select and assemble the appropriate views based on the nature of the system and question at hand. Once assembled, we know that this visualization will need to support information foraging and sense making as well as provide the ability to test hypothesized mental models. The issue is that it is not clear how this modular approach should be architected both in terms of what the modules should be and how they should interrelate.

Of particular concern for enterprise systems is the level of complexity and ambiguity. While there are many definitions of complexity, the one that is relevant for this case is that of Casti [1985], which loosely defines the complexity of a system as the number of different models required to describe it. (Note that this can change depending on the question asked and is consistent with the model bifurcation discussion presented earlier in this report.) In one sense, this type of complexity is a defining characteristic of an enterprise system. A question logically follows: at what point does the complexity of the visualization overwhelm the utility of the visualization? In other words, as the complexity of the visualization increases, does the ability to draw causal inferences diminish?

In addition to complexity, we are also concerned with ambiguity. In real life, it is often difficult, if not impossible, to populate a model of an enterprise with a complete and highly accurate data set. Real enterprise decision makers are often faced with noisy and incomplete data on the current state of the enterprise. Consequently, there is a question of how increasing ambiguity coupled with increasing complexity affects the ability to draw causal inferences from a visualization.

We hypothesize the relationships depicted in Figure 13 below. For problems with low complexity and low ambiguity, interactive visualizations will be effective (green area) for making causal inferences. To link this to the organizational strategies discussed earlier in this report, this situation would allow an enterprise decision maker to follow an optimize strategy as the link between cause and effect is clear.

As complexity and/or ambiguity increase, we hypothesize that one will lose the ability to draw causal inferences. However, we suspect that interactive visualizations will be useful in identifying factors that influence outcomes (yellow area). In terms of organizational strategy, this means that one can map out possible scenarios and employ either an adapt or a hedge strategy.

Finally, as ambiguity and complexity continue to increase, the enterprise decision maker will simply have no basis for identifying any factors that consistently contribute to the outcomes of interest (the red area). The concern here is that an interactive visualization may actually be counterproductive in that it may create the false impression that the system is understandable when it is not. Decision makers may be prone to mistaking noise for signal. In this domain, the best organizational strategy is “accept” because any conclusions drawn from the interactive visualization are more likely to be wrong than right.

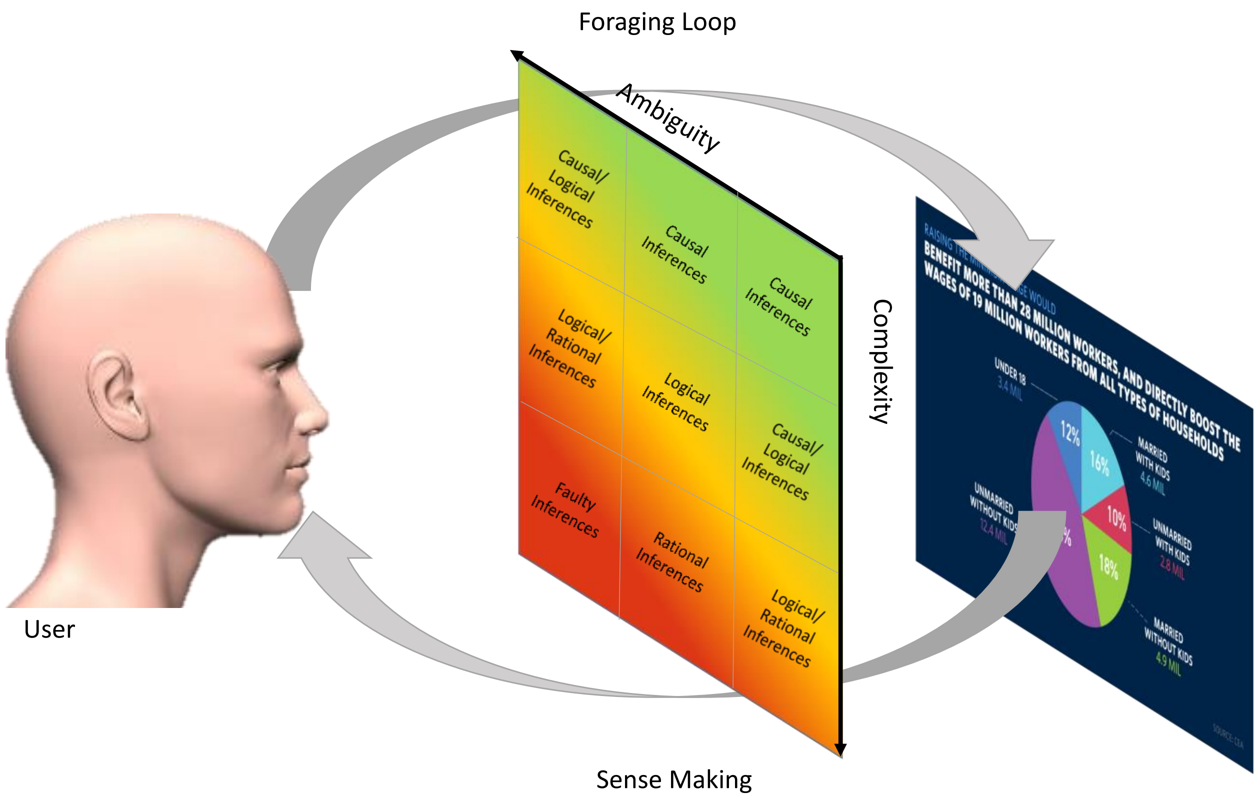


Figure 13. Hypothesized relationship between complexity, ambiguity, and visualization effectiveness

Of course these relationships are hypothetical at this point. Additional research would be required to test them. Regardless, the state of interactive visualization as applied to enterprise decision making can be summarized in the following manner: There are potential benefits to the approach, but there are still several outstanding issues that must be addressed in order establish meaningful guidance to developing effective enterprise visualizations. Without this guidance, there is a risk that interactive enterprise visualizations will mislead rather than inform decision makers.

7.2 CONCEPTUAL DESIGN OF THE INTERACTIVE VISUALIZATION ENVIRONMENT

Given the extant work on how visualization can be used to support causal inference, the next logical step was to adapt the findings of this work to enterprise visualization. As a test case, we developed a small-scale conceptual design for an enterprise visualization. The idea was to break up the interface into functional containers that support the foraging and sense-making processes necessary for causal inference as described in the previous section. Each container would include a visualization that supports one or more of the cognitive functions required for sense making. Furthermore, the containers were arranged in such a manner that that they could support a natural flow through the sense-making process.

Rather than conducting a purely abstract exercise in design, we leveraged an existing study of an enterprise problem that was performed by Liu et. al. [2014]. This study considered the factors that influenced twelve real-world decisions to withdraw an automobile from the commercial marketplace. This case is particularly interesting for our purposes because it is an enterprise diagnostic problem (Why did the car fail? Was it the car or other factors?), and it involves interactions among many different views of the system at different layers of abstraction (political, economic, market, consumer choice, etc.). This allows us to exercise all of the necessary aspects of the foraging and sense-making processes for a multi-view enterprise problem.

The visualization is designed to allow the user to investigate the question: Why did the car fail? The container design for this case is depicted in Figure 14. (Larger versions of all of the figures for the conceptual design are contained in a separate file that was provided in addition to this report). Note that each container has a description of its functions along with the intended user flow among the containers.

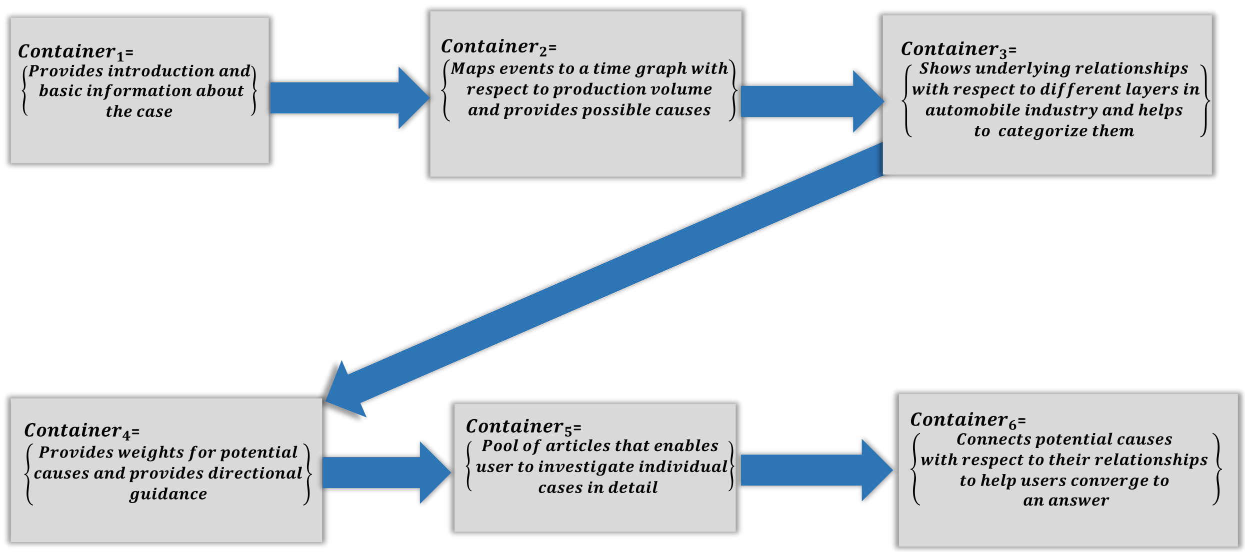
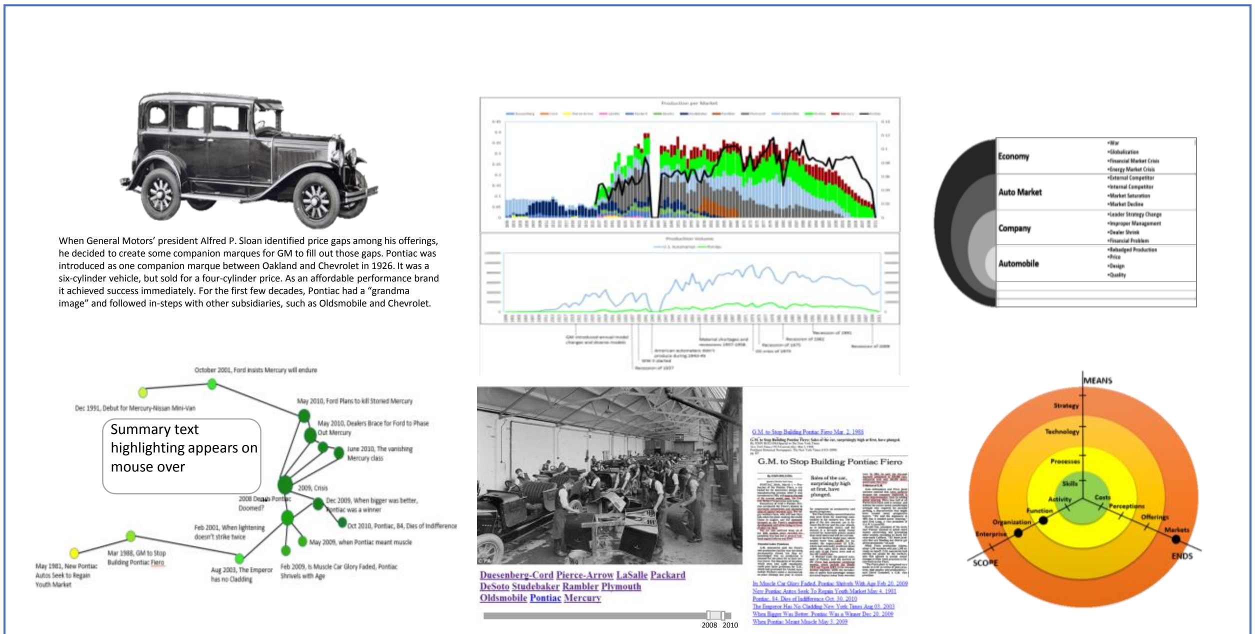


Figure 14 – Visualization container functions with flows

When these containers are instantiated with actual visualizations using real data from a particular failed automobile, we obtain Figure 15. The function of each container depends on the raw information available and the anticipated mental model of the user. When we say anticipated mental model, we mean the internal mental connection that users make based on the interactive form and purpose of each container. For example: the top, middle container has a graph with labels and hyperlinks. These link the time axis to events, and helps users make connections between different events and important factors such as production volume. This is first step to identifying possible causes. The lower left container has a network that relates relevant news articles with hyperlinks to the specific articles. These are actual news articles that describe events connected in some way to the car of interest. (Note that many of these articles were published as much as 100 years ago.) The selected news article is displayed in the bottom middle container. They can relate to any layer of abstraction of such news about the specific automotive company or the entire automobile industry, etc. Finally, the multi-layer, aiding model in the lower right helps the user keep track of where the information fits within the larger problem space.



- DFM Sense-making model (Klein, 2007)
- Intelligent Analysis (Pirolli & Card, 2005)
- Complexity (Casti, 1985)
- Ambiguity

Figure 15. How an aiding interface can help a non-expert go through 12 cases of automobiles that were withdrawn from the market

The role of each container in the sense making process is depicted in Figure 16. Note that at this point, the conceptual visual design is essentially a hypothesis informed by the literature. At this point its efficacy has not been established. This would require an experiment, which we plan to execute as future work. The intent is to capture the lessons learned from this small, scaled, controlled effort and feed it back into the design of visualizations for the counterfeit parts case study.

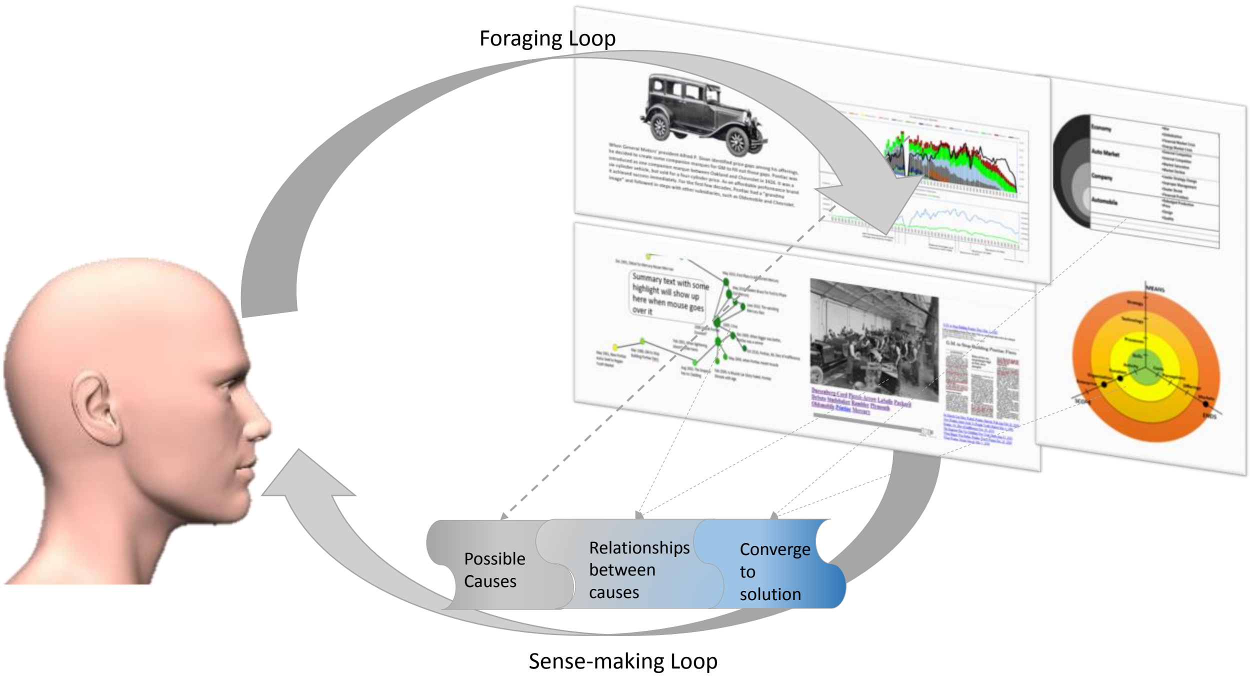


Figure 16 – Container support for the sense-making process

8 MODEL ARCHIVE REQUIREMENTS

In this report, much has been discussed regarding the selection of models to represent phenomena, understanding the relationships among those selected models, and, in some cases, the composition of those models. As discussed in Section 2.3.3, the breadth and depth of the available set of models that have been developed by the scientific, engineering, and business communities is too great to expect any enterprise modeler to master them all. To avoid both missing potentially useful models and misapplying models, a useful resource for enterprise modelers would be a model archive that documents canonical models for many common phenomena and describes the strengths and limitations of each. While such an archive cannot replace true expertise, it can provide a starting point that will lead enterprise modelers to the appropriate domain experts.

The intent here is to lay out a working set of high-level requirements for an enterprise model archive. As the enterprise modeling methodology (Section 2) is the driver behind the model archive, it also serves as the launching point for developing the requirements. To that end, we developed a small set of use cases that describe how an enterprise modeler would use the model archive while executing the ten-step methodology. We then derived a set of working, top-level requirements from these use cases.

While these requirements are based on both what we have learned as we has developed the modeling methodology and associated theory, the actual experience base informing these requirements still consists of a single case study, the counterfeit parts case study. Consequently, we do not consider these a validated set of requirements. To address that situation, at the end of this section, we discuss some potential approaches to refine and validate the requirement set.

8.1 Use Cases

The following are use cases that describe how an enterprise modeler would use the model archive while executing the 10-step methodology as well as adding to and updating the archive.

Use Case 1: Associating Phenomena with Visualizations (Steps 2 through 4)

Description:

During steps two through four of the modeling methodology, the user needs to take enterprise phenomena related to the question of interest (e.g., economic phenomena, engineering phenomena, social phenomena, etc.) and identify canonical visualizations (price diagrams, network diagrams, etc.)

Scenario:

1. The user identifies a phenomenon the he or she needs to visualize
2. The user executes a guided search of the archive to narrow down that set of possible visualizations to a candidate set applicable to the phenomenon
3. The archive provides the user with strengths and limitations of each of the candidate visualizations
4. The user selects the visualization technique that best fits the circumstances
5. The archive provides the user a list of related references and tools to aid in the application of the visualization

Use Case 2: Associating Phenomena with Representations (Step 5)

Description:

During step five of the modeling methodology, the user needs to identify potential computational models to represent the phenomena of interest. These may or may not be related to the visualizations selected during steps three and four. For example, a user may visualize a process initially using an IDEF diagram, but then model it using a queuing model. Of particular interest to the user will be finding representations that correspond to the tradeoffs of interest.

Scenario:

1. The user identifies a phenomenon that he or she needs to model

2. The user executes a guided search of the archive to narrow down that set of possible models to a candidate set applicable to the phenomenon
    a. The user chooses to guide the search based on the visualization technique previously selected
    b. The user then further restricts the search to correspond to the tradeoffs of interest
3. The archive provides the user with strengths and limitations of each of the candidate model
    a. The limitations will include known incompatibilities with other types of representations
4. The user selects the model that best fits the circumstances and constraints
5. The archive provides the user a list of related references, metadata, and tools to aid in the application of the model

Use Case 3: Composing Representations (Step 6)

Description:

Step six of the modeling methodology involves assessing the computational composability of a set of models. In essence the archive should provide the user guidance in implementing the guidelines described in Section 6.2.4. It may also recommend heuristics to consider such as those discussed in Section 6.2.3. In many cases, the best guidance that the archive could provide is pitfalls to avoid.

Scenario:

1. The user identifies a set of models the he or she needs to compose
2. The user enters the model types of interest into archive
3. The archive identifies potential sources of model bifurcation for each of the model types (See Section 4.4)
4. The archive identifies common pitfalls for each of the model types
5. The archive identifies potential conflict points among the select model types
6. The archive provides the user a list of related references, strategies, and tools to aid in the composition of the selected models types (when appropriate)

Use Case 4: Adding to the archive

Description:

Over time new representations and approaches will be developed. Users will be able to submit these new representations to the archive.

Scenario:

1. The user identifies a new representation to incorporate into the archive
2. The user submits a request to the archive to incorporate the new representation
3. The submission is peer reviewed for quality control
4. Suggestions for improvement are provided to the user
5. The user updates and resubmits the proposed representation
6. The new representation is accepted into the archive

Use Case 5: Updating an existing representation in the archive

Description:

Users will accumulate lessons learned from the application of the representations for actual enterprise problems. Users will be able to submit these lessons learns such approaches to compositions, limitations, pitfalls, etc. for incorporation into the archive.

Scenario:

1. The user develops a recommendation to improve a representation in the archive
2. The user submits a request to the archive to incorporate the recommendation
3. The submission is peer reviewed for quality control
4. Suggestions for improvement are provided to the user
5. The user updates and resubmits the recommended change
6. The archive is updated to reflect the recommended change

8.2 Requirements

The following are a proposed set of high-level requirements for the model archive that are derived from the use cases documented in the previous section.

1. The archive shall allow the user to identify visualization techniques that can be used to represent an enterprise phenomenon documented in the archive.
2. The archive shall provide the user with the known strengths and limitations of a user selected visualization technique.
3. The archive shall provide the user a list of references and tools associated with a user-selected visualization.
4. The archive shall allow the user to execute a guided search of mathematical or computational models that are documented in the archive.
    a. The archive's guided search shall allow a user to search for mathematical or computational models associated with enterprise phenomena documented in the archive.
    b. The archive's guided search shall allow a user to restrict results by a designated tradeoff documented in the archive.

c. The archive's guided search shall allow a user to search for mathematical or computational models associated with a particular visualization technique documented in the archive.

5. The archive shall provide the user with the known strengths of and limitation of a user selected mathematical or computational model documented in the archive.
6. The archive shall provide the user with a list of known incompatibilities between the user selected mathematical or computational model and the other mathematical or computational models documented in the archive.
7. The archive shall provide the user with a list of related references, metadata, and tools associated with the user selected mathematical or computational model documented in the archive.
8. The archive shall provide the user with a list of known sources of bifurcation for the user selected mathematical or computational model documented in the archive.
9. The archive shall provide the user with a list of known pitfalls in the application of the user selected mathematical or computational model documented in the archive.
10. The archive shall provide the user with a set of known incompatibilities for a user selected set of mathematical or computational models from among those documented in the archive.
11. The archive shall provide the user with a list of related references, strategies, and tools for composing a user selected set of mathematical or computational models from among those documented in the archive.
12. The archive shall allow a user to submit a new content item to the archive.
13. The archive shall allow a user to submit an edit to an existing content item in the archive.
14. The archive shall allow designated users to review a submitted new content item.
15. The archive shall allow designated users to review a submitted edit to an existing content item.
16. The archive shall allow a designated user to approve a submitted new content item.
17. The archive shall allow a designated user to approve a submitted edit to an existing content item.
18. The archive shall update its content to reflect a submitted new content item upon approval of that item.
19. The archive shall update its content to reflect a submitted edit to an existing content item upon approval of that item.

8.3 Refining and Validating

As noted previously, the requirements presented in this section are both high-level and preliminary. Consequently, they need to be both refined and validated. The most logical approach would be to review the requirements with potential users of the archive. The natural place to start would be peers within in the SERC ESOS area. This could be integrated with a larger peer review of the enterprise modeling methodology.

Beyond SERC, the Center for Complex Systems and Enterprises is currently in discussions with NSF to host a workshop on systems modeling education. This could provide another venue to consider how the archive could be used to support model builders. The intent of this workshop is to include a diverse group of representatives from multiple disciplines concerned with modeling. This would help guard against the risk of "group think" that could arise by limiting the review to SERC.

9 CONCLUSIONS AND FUTURE WORK

In this report, we considered the issues associated with the modeling and analysis of enterprise systems. In particular, we evaluated a methodology for modeling enterprise systems and identified potential avenues for improvement. A preponderance of the issues encountered reveals an underlying theme. Most, if not all, of the challenges associated with modeling, analyzing, and influencing enterprise systems revolve around qualitative shifts in the system as one moves along a dimension of interest. The problem with these shifts is that an approach developed to represent or influence the system on one side of the shift may be invalid on the other (hence the qualitative distinction). So critical are these shifts that their study has emerged in multiple domains. In dynamical systems theory these shifts are called bifurcations. In complexity science, they are called bifurcations or phase shifts. In physics, they are called phase shifts or symmetry breaks.

It is perhaps this last term that is the most meaningful for our purposes. As noted by Morrison [2008], traditional mathematical analysis is really about finding and exploiting symmetries. Symmetries allow us to reduce the computational burden when analyzing a system. Physics makes extensive use of symmetries. In classical mechanics, some of the assumed symmetries are that the rules of mechanics do not change over time and space. When these symmetries are coupled with Noether's theorem, they yield important conservations laws such as momentum and energy. The utility of such theories is undeniable. In essence, symmetries allow us to efficiently extrapolate the state of a system along dimensions of interest such as space and time.

But what happens when symmetries are broken? A solution that exploited a symmetry is not likely to be valid once that symmetry is broken. Under some conceptions, the number of symmetry breaks (or phase shifts) over the space of interest corresponds to the complexity of the system. Thus, the more complex the system, the fewer exploitable symmetries, and consequently, the lower the utility of any given model. We established in Section 4 that enterprises are complex systems. As a result, we expect there to be relatively few exploitable symmetries to predict the future behavior of an enterprise. Consequently, any purely analytical approach to modeling and influencing enterprises is built on a weak foundation.

A fanciful metaphor for a complex system is islands of order in a sea of chaos. Each of the islands has exploitable symmetries, and we can manage as long as we can stay on one island. The problem is that we may be hurled by unpredictable winds and currents from one island to another without warning – or, worse yet, between islands. Consequently, any reasonable strategy must explicitly recognize that the some of rules that one is exploiting today may be useless the next. Furthermore, each island is a little different, so there is not likely to be any one set of rules that will work on every possible island.

To put the results of this research task in perspective, the enterprise modeling methodology is implicitly attempting to manage the scope of the problem. A universal or complete model of the system would entail capturing every possible island and all of the ways to move from one to the next. This is impossible for most realistic problems. Rather, the methodology is designed to aid the decision maker in only considering those islands that one is likely to end up on in the near future, and even then only consider the features that may be different among those islands. Since we don't know when or how we might get blown between islands the strategy framework allows us to be ready to adapt to the extent that we can when that happens.

To put this in more technical terms, the traditional approach to systems modeling is to eliminate model bifurcations by increasing the complexity or fidelity of the model. We assert that this is either impossible or impractical – and certainly costly -- for most realistic enterprise problems. Thus, there is a larger implication for applying traditional systems engineering approaches to enterprise systems. If one cannot eliminate the model bifurcations, then traditional systems approaches can only ever be applied to subset of an enterprise problem. When one persists in applying only traditional systems approaches to an enterprise, one necessarily excludes all aspects of the enterprise system that are not "well behaved." Unfortunately, this has the consequence of implicitly assuming that everything outside of the modeled subset is either fixed or has no impact on that subset. While this implicit assumption can be enforced to a degree via organizational constructs in traditional systems engineering, it is almost certainly invalid when attempting to engineer enterprise systems. Since enterprises are often highly interconnected (both internally and externally) and intrinsically adaptive, optimizing over a subset of the problem will often result in a fragile solution that can trigger a number of unintended consequences via second or higher order effects.

Consequently, we assert that a traditional systems engineering approach (without modification) is impossible or impractical for most realistic enterprise problems. In a sense, the exhibition of many bifurcations is what makes an enterprise an enterprise. Thus, rather than attempt to eliminate them all, our approach is to actively hunt for them and attempt to manage them.

As far the utility of the methodology, it was found to be a useful construct, but some challenges and areas for improvement remain. In particular, some of the steps of the methodology were found to be vague and require additional guidance to actually implement. The methodology makes extensive use of visualization, but the efficacy of visualization applied to enterprise systems remains unclear. Finally, the methodology relies heavily on the composition of computational representations of different aspects of an enterprise system. Such compositions impose substantial technical challenges when attempted for enterprise systems. To address these issues, a number of recommendations were made. However, work remains.

First and foremost, the model developed to analyze the counterfeit parts case study should be validated and applied to determine if it actually yields useful insights. This is necessary to establish the ultimate utility of the enterprise modeling approach. Second, a single example is not sufficient to validate a methodology. Consequently, another case study should be initiated to continue to challenge the enterprise modeling methodology. Third, the efficacy of visualization to support enterprise decision making needs to be established. This means that one more experiments should be conducted to determine the conditions under which visualizations actually aid enterprise decision makers. Finally, additional work is necessary to implement the recommendations made in this report. This includes explicitly mapping the chain of phenomena to representation to model, investigations into how to operationalize the strategy framework, as well as additional guidance for model composition when it is appropriate to do so.

REFERENCES

ABA (2012), A White Paper Regarding Department of Defense Implementation of Section 818 of the National Defense Authorization Act for Fiscal Year 2012. Report issued by Task Force on Counterfeit Parts of the Committee on Acquisition Reform and Emerging Issues of the American Bar Association Section of Public Contract Law, October 5, 2012.

AIA (2011), Counterfeit Parts: Increasing Awareness and Developing Countermeasures. Arlington, VA: Author.

Alderson, D. L. and J. C. Doyle (2010), Contrasting Views of Complexity and Their Implications for Network-Centric Infrastructures, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(2): 839-852.

Arnott, R. (2013), A bathtub model of downtown traffic congestion, *Journal of Urban Economics*, 76: 110-121.

Arnott, R., A. de Palma, and R. Lindsey (1993), A structural model of peak-period congestion: A traffic bottleneck with elastic demand, *American Economic Review*, 83: 161-179.

Arthur, W. B. (1989), Competing Technologies, Increasing Returns, and Lock-In by Historical Events, *The Economic Journal*, 99(394): 116-131.

Arthur, W.B. (1999), Complexity and the economy, *Science*, 284(5411): 107-109.

Ashby, W. R. (1956), *An Introduction to Cybernetics*. London: Chapman & Hall.

Bates, M.J. (1989), The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, *13*(5): 407-424.

Bento, A., J. Hughes, and D. Kaffine (2013), Carpooling and driver response to fuel price changes: Evidence from traffic flows in Los Angeles, Journal of Urban Economics, 77: 41-56.

Biederman, I. (1987), Recognition by Component: A theory of Human Image Understanding, *Psychological Review*, 94(2): 115-147.

Blandin, S., D. Work, P. Goatin, B. Piccoli, and A. Bayeo (2011), A general phase transition model for vehicular traffic," *SIAM Journal on Applied Mathematics,* 71: 107-127.

Bodner, D. A. (2014), Enterprise modeling framework for counterfeit parts in defense systems, In *Proceedings of the 2014 Complex Adaptive Systems Conference*.

Bodner, D. A., P. Prasad, V. Sharma, A. Compagnoni, and J.E. Ramirez-Marquez, (2013). A socio-technical model of the problem of counterfeit parts. Technical Report SERC-2013-TR-020-3 Volume 2, Hoboken, NJ: Systems Engineering Research Center.

Boisot, M. and B. McKelvey, (2011), Complexity and Organization—Environment Relations: Revisiting Ashby's Law of Requisite Variety, in P. Allen, S. Maguire, and B. McKelvey (Eds.), *The Sage Handbook of Complexity and Management*, Los Angeles: Sage, 279-298.

Bonsall, P., J. Shires, J. Maule, B. Matthews, J. Beale (2007), Responses to complex pricing signals: Theory, evidence and implications for road pricing, *Transportation Research Part A: Policy and Practice*, 41: 672-683

Borland, D., and R.M. Taylor II (2007), Rainbow Color Map (Still) Considered Harmful. *IEEE Computer Graphics and Applications*, 27(2): 14-17.

Bossomaier, T., L. Barnett, and M. Harré, (2013), Information and phase transitions in socio-economic systems, *Complex Adaptive Systems Modeling*, 1(9).

Brownstone, D., A. Ghosh, T. Golob, C. Kazimi, D. Van Amelsfort (2003), Drivers' willingness-to-pay to reduce travel time: evidence from the San Diego I-15 congestion pricing project, *Transportation Research Part A: Policy and Practice*, 37: 373-387.

Burris, M, R. Pendyala (2002), Discrete choice models of traveler participation in differential time of day pricing programs, *Transport Policy*, 9: 241-251

Bury, K. F., Boyle, J. M., Evey, R. J., and Neal, A. S. (1982,. Windowing versus Scrolling on a Visual Display Terminal, *Human Factors: The Journal of Human Factors and Ergonomics Society*, 24(4): 385-394.

Business Insider (2012), Counterfeit Chinese microchips are getting so good they can't be identified. http://www.businessinsider.com/counterfeit-parts-from-china-raise-grave-concerns-for-both-us-companies-and-national-security-2012-6.

Casti, J. L. (1985), On System Complexity: Identification, Measurement, and Management, in J.L. Casti and A. Karlqvist (eds.), C*omplexity, Language, and Life: Mathematical Approaches*, Springer-Verlag: Berlin, 144-173.

Casti, J. L. (2012), *X-Events: The Collapse of Everything*, New York: Harper Collins.

Chen, C. (2005), Top 10 Unsolved Information Visualization Problems, *IEEE Computer Graphics and Applications,* 25(4):12-16.

Chen, C. (2006). *Information Visualization: Beyond the Horizon*. London: Springer-Verlag.

Christen, M., G. Bongard, A. Pausits, N. Stoop, and R. Stoop (2008), Managing Autonomy and Control in Economic Systems , in D. Helbing (Ed.), *Managing Complexity: Insights, Concepts, Applications*, Berlin: Springer-Verlag, 37-55.

Cohen, B. S., & Lee, K. (2014), On the limits of test in establishing products assurance, Institute for Defense Analysis, Arlington, VA, April 1, 2014.

Colombo, R. (2003), Hyperbolic phase transitions in traffic flow, *SIAM J. App. Math.*, 63: 708-721.

Colombo, R. (2002), On a 2 x 2 hyperbolic traffic flow model, *Math. Comput. Modeling*, 35: 683-688.

Cutting, D. R., Karger, D. R., Pedersen, J. O., & Tukey, J. W. (1992). Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections. *SIGIR'92 Proceedings of the 15th annual international ACM SIGIR conference on Research and development*, pp. 318-329, New York.

Daganzo, C., V. Gayah, and E. Gonzales (2011), Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness, and instability, *Transportation Research Part B*, 45: 278-288.

DAU (2013). Anti-counterfeiting. Defense Acquisition Guidebook. https://acc.dau.mil/CommunityBrowser.aspx?id=638350&lang=en-US.

De Borger, B. and S. Proost (2012), A political economy model of road pricing, *Journal of Urban Economics*, Elsevier, 71(1): 79-92.

Department of Defense (2014), DoD Architecture Framework Version 2.02, United States Department of Defense, http://dodcio.defense.gov/dodaf20.aspx (last accessed 2/13/14).

Diallo, S. Y., J. J. Padilla, R. Gore, H. Herencia-Zapana, and A. Tolk (2014), Toward a Formalism of Modeling and Simulation Using Model Theory, *Complexity*, 19: 56-63.

DoC (2012), Defense Industrial Base Assessment: Counterfeit Electronics, Report. Washington, DC: Author (Dept. of Commerce).

DoD (2011), DoD Supply Chain Materiel Management Policy, DoD Instruction 4140.01. Washington, DC: Author (Dept. of Defense).

DoD (2012), Protection of Mission Critical Functions to Achieve Trusted Systems and Networks (TSN), DoD Instruction 5200.44. Washington, DC: Author (Dept. of Defense).

DoD (2013), DoD Counterfeit Prevention Policy, DoD Instruction 4140.67, Washington, DC: Author (Dept. of Defense).

DoD (2014), Defense Federal Acquisition Regulation Supplement: Detection and avoidance of counterfeit electronic parts (DFARS Case 2012-D055), Federal Register, 79, 26091-26108.

Doyle, J. C. and M. Csete, (2011), Architecture, constraints, and behavior, *Proceedings of the National Academy of Science*, 108, supplement 3,15624-15630.

Economist (2012), Huawei: the company that spooked the world. Author, August 4, 2012.

Frey P.R., W.B. Rouse, and R.D. Garris (1993). Big graphics and little screens: Model-based design of large-scale information displays. In W.B. Rouse (Ed.), *Human/Technology Interaction in Complex Systems* (Vol. 6, pp. 1-57). Greenwich, CT: JAI Press.

Frey, P. R., W.B. Rouse, and R.D. Garris (1992), Big Graphics and Little Screens: Designing Graphical Displays for Maintenance Tasks, *IEEE Transaction on Systems, Man, and Cybernetic*s, 22(1):10-20.

GAO (2010), DoD should leverage ongoing initiatives in developing its program to mitigate risk of counterfeit parts, Report GAO-10-389, Washington, DC: Author (Government Accountability Office).

GAO (2011), Periodic Assessment Needed to Correct Parts Quality Problems in Major Programs, GAO-11-404. Washington, DC: Author (Government Accountability Office).

GAO (2012a). Additional Efforts Needed by National Security-Related Agencies to Address Risks. Report GAO-12-579T, Washington, DC: Author (Government Accountability Office).

GAO (2012b), Suspect counterfeit electronic parts can be found on internet purchasing platforms, Report GAO-12-375, Washington, DC: Author (Government Accountability Office).

Gasser, I., T. Seidel, G. Sirito, and B. Werner (2007), Bifurcation analysis of a class of car following traffic models II: Variable reaction times and aggressive drivers, *Bulletin of the Institute of Mathematics*, 2: 587-607.

Geiser, G. and W. Schumacher (1976), Parallel vs. Serial Instrumentation for Multivariable Manual Control in Control Rooms. In *Monitoring Behavior and Supervisory Control* (pp. 405-425). New York: Springer US.

Gharajedaghi, J. (2007), *Systems Thinking: Managing Chaos and Complexity – A Platform for Designing Business Architecture*. Burlington, MA: Morgan Kaufmann.

Giachetti, R.E., (2010), *Design of Enterprise Systems: Theory, Architecture, and Methods*, Boca Raton, FL: CRC Press.

Gipps, P. (1981), A behavioural car-following model for computer simulation, *Transportation Research Board Part B*, 15: 105-111.

Goodall, N. and B. Smith (2010), What Drives Decisions of Single-Occupant Travelers in High-Occupancy Vehicle Lanes?, *Transportation Research Record: Journal of the Transportation Research Board*, 2187(1): 156-161.

Guin, U., D. DiMase, and M. Tehranipoor (2014), Counterfeit integrated circuits: detection, avoidance, and the challenges ahead. *Journal of Electronic Test*.

Haimes, Y.Y. (1981), Hierarchical holographic modeling, *IEEE Transactions on Systems, Man, and Cybernetics*, 11(9): 606-616.

Harvey, D.L. and M. Reed (1996), Social Science as the Study of Complex Systems, in L.D. Kiel and E. Elliot (Eds.), *Chaos Theory in the Social Sciences: Foundations and Applications*, Ann Arbor: The University of Michigan Press.

Heath, C. and R. Gonzalez (1995), Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence against Information Collection Views of Interactive Decision Making, *Organizational Behavior and Human Decision Processes*, *61*(3): 305-326.

Helbing, D. and S. Lämmer, (2008) Managing Complexity: An Introduction, in D. Helbing (Ed.), *Managing Complexity: Insights, Concepts, Applications*, Berlin: Springer-Verlag, 1-15.

Helbing, D. (2001), Traffic and related self-driven many-particle systems, *Reviews of Modern Physics*, 73: 1067-1141.

Henneman, R. L., and W.B. Rouse (1984), Human performance in monitoring and controlling hierarchical large-scale systems, *IEEE Transactions on Systems, Man, and Cybernetics*, 2: 184-191.

HNTB Corporation (2010), Atlanta Regional Managed Lane System Final Report, available from http://www.dot.ga.gov/projects/studies/managedlanes/documents/finalreport.pdf.

Hoffman, M., (2011), Epistemic and normative aspects of ontologies in modeling and simulation, *Journal of Simulation*, 5(3): 135-146.

Hoffman, M., (2013) "Ontology in modeling and simulation: An epistemological perspective," in *Ontology, Epistemology, and Teleology for Modeling and Simulation*, A. Tolk, (Ed.), Heidelberg: Springer, 59-87.

Holling, C.S. (1959), Some Characteristics of Simple Types of Predation and Parasitism, *The Canadian Entomologis*t, 91(7):385-398.

Hume, D. (1739-1740), A treatise of human nature, London, UK: Reprinted by New Vision Publications.

Itti, L., C. Koch, and E. Neibur (1998), A model of Saliency Based Visual attention for Rapid Scene Analysis, *IEEE Transactions on Pattern Analysis and machine intelligence*, 20(11): 1254-1259.

Janson, M. and D. Levinson (2014), “HOT or Not, Driver Elasticity to Price on the MnPASS HOT Lanes”, available from http://nexus.umn.edu/Papers/HOTorNOT.pdf (last updated 2/21/14).

Kahneman, D. and C. A. Varey (1990). Propensities and counterfactuals: The loser that almost won, *Journal of Personality and Social Psychology*, 59(6):1101.

Kempf, K.G. (2008), Complexity and the Enterprise: The Illusion of Control, in D. Helbing (Ed.), *Managing Complexity: Insights, Concepts, Applications*, Berlin: Springer-Verlag, 57-87.

Kirsh, D. and P. Maglio (1994), On Distinguishing Epistemic from Pragmatic Action, *Cognitive Science*, 18(4): 513-549.

Klein, G, J. K. Phillips, E.L. Rall, and D.A. Peluso (2007), A Data-Frame Theory of Sensemaking. *Proceedings of the Sixth International Conference on Naturalistic Decision Making* (pp. 15-17). Mahwah, NJ: Lawrence Erlbaum Associates.

Klein, G. (2003), *Intuition at Work: Why Developing Your Gut Instincts Will Make You Better at What You Do*. New York: Doubleday.

Klein, G., and R. Hoffman (2009), Causal Reasoning: Initial Report of a naturalistic study of causal inferences, In *Proceedings of the 9th International Conference on Naturalistic Decision Making*. London, UK.

Kodagoda, N., S. Attfield, W. Wong, C. Rooney, and S.T. Choudhury (2013), Using Interactive Visual Reasoning to Support Sense-Making: Implications for Design, *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2217-2226.

LeBaron, B., W. B. Arthur, and R. Palmer (1999), Time series properties of an artificial stock market, *Journal of Economic Dynamics and Control*, 23: 1487-1516.

Lewis, M. (2014), *Flash Boys*, W. W. Norton & Company.

Lighthill, M. and J. Whitham (1955), On kinematic waves II: A theory of traffic flow on long crowded roads, *Proceedings of the Royal Society A*, 229: 317-345.

Lindsey, C., V. van den Berg, and E. Verhoef (2012), Step tolling with bottleneck queuing congestion, *Journal of Urban Economics*, 72: 46-59.

Lindsey, R. (2006), Do economists reach a conclusion on road pricing? The intellectual history of an idea, *Econ Journal Watch*, 3: 292-379.

Liu, C., W.B. Rouse, and Z. Yu, (2014), When Transformation Fails: Twelve Case Studies in the Automobile Industry, *Proceedings of the 4th International Engineering Systems Symposium*.

Livingston, H. (2014). Avoiding counterfeit electronic components – myths and unreliable facts, Presentation, BAE Systems.

Livingston, H. (2007a), Avoiding counterfeit electronic components, *IEEE Transactions on Component Packaging Technology*, 30: 187-189.

Livingston, H. (2007b), Avoiding counterfeit electronic components – Part 2: Observations from recent counterfeit detection experiences, Report, BAE Systems.

Marr, D. (1982), *Vision: A computational investigation into the human representation and processing of visual informatio*n, MIT Press.

McDermott, T., W. Rouse, S. Goodman, and M. Loper (2013), Multi-level modeling of complex socio-technical systems, *Proceedia Computer Science*, 16: 1132-41.
McFadden, F. E., and R.D. Arnold (2010), Supply chain risk mitigation for IT electronics, *Proceedings of the IEEE International Conference on Technologies for Homeland Security*, 49-55.
McGinnis, L., E. Huang, K. S. Kwon, and V. Ustun (2011), Ontologies and simulation: A practical approach, *Journal of Simulation*, 5(3): 190-201.
Miller, G. A. (1956), The magical Number seven, plus or minus two: Some limits on our capacity for processing information, *Psychological Review*, 63(2): 81.
Mitchell, C. M., and R.A. Miller (1983), Design Strategies for Computer-Based Information Display in Real-Time Control Systems, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 25(4): 353-369.
Morrison, F. (2008), *The Art of Modeling Dynamic Systems: Forecasting for Chaos, Randomness, and Determinism*, Mineola: Dover.
Newell, A. (1994), *Unified theories of cognition*, Harvard University Press.
Newell, A. and H.A. Simon (1972), *Human Problem Solving,* 104(9), NJ: Englewood Cliffs.
Nightingale, D. J., and D.H. Rhodes (2004), Enterprise Systems Architecting: Emerging Art and Science within Engineering Systems, *Proceedings of the ESD External Symposium*.
Norman, D. (1993), *Things that make us smart: Defining human attributes in the age of the machine*, Perseus Books.
Olszewski, P and L. Xie (2005), Modelling the effects of road pricing on traffic in Singapore, *Transportation Research Part A: Policy and Practice*, 39: 755-772
Orosz, G., B. Krauskopf, and R. Willson (2005), Bifurcations and traffic jams in a car-following model with reaction-time delay, *Physica D*, 211: 277-293.
Orosz, G., R. Willson, and B. Krauskopf (2004), Global bifurcation investigation of an optimal velocity traffic model with driver reaction time, *Physical Review E*, 70: 277-293.
Palmer, R., W.B. Arthur, J.H. Holland, B. LeBaron, and P. Tayler (1994), Artificial economic life: a simple model of a stockmarket. *Physica D: Nonlinear Phenomena*, 75(1): 264-274.
Park, H., T. Clear, W. B. Rouse, R. C. Basole, M. L. Braunstein, K. L. Brigham, and L. Cunningham, (2012), Multilevel Simulations of Health Delivery Systems: A Prospective Tool for Policy, Strategy, Planning, and Management, *Service Science*, 4(3): 253-268.
Partridge, C., A. Mitchell, and S. de Cesare (2013), Guidelines for developing ontological architectures in modeling and simulation, in *Ontology, Epistemology, and Teleology for Modeling and Simulation*, A. Tolk, (Ed), Heidelberg: Springer, 22-57.
Patterson, E.S., E. M. Roth, and D.D. Woods (2001), Predicting Vulnerabilities in Computer-Supported Inferential Analysis under Data Overload, *Cognition, Technology & Work*, 3(4): 224-237.
Pecht, M., and S. Tiku (2006), Bogus: electronic manufacturing and consumers confront a rising tide of counterfeit electronics, *IEEE Spectrum*, 43: 37-46.
Pennock, M. J. and W.B. Rouse (2014a), The challenges of modeling enterprise systems, *Proceedings of the 4th International Engineering Systems Symposium*.
Pennock, M.J. and W.B. Rouse (2014b), Why Connecting Theories Together May Not Work: How to Address Complex Paradigm-Spanning Questions, in *Proceedings of the IEEE*

International Conference on Systems, Man, and Cybernetics 2014, October 5-8, 2014, San Diego, CA.

Peters, R. J., and L. Itti (2007), Applying computational tools to predict gaze in interactive visual environments, *ACM Transactions on Applied Perception*, 5(2): 9.

Piroli, P., and S. Card (1995), Information Foraging in Information Access Environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 51-58), ACM Press/Addison-Wesley Publishing Co.

Pirolli, P., and S. Card (2005), The Sensemaking Process and Leverage Points for Analyst Technology as Identifies Through Cognitive Task Analysis, *Proceedings of International Conference on Intelligence Analysis*, 5: 2-4, McLean, VA: Mitre.

Poli, R. (2013), A note on the difference between complicated and complex social systems, *Cadmus*, 2(1): 142-147.

Rasmussen, J. (1983), Skills, Rules, Knowledge; Signals, Signs, and Symbols, and other Distinctions in Human Performance Models, *IEEE Transaction on Systems, Man and Cybernetics*, 14 (3): 257-266.

Rosen, R. (1978), *Fundamentals of Measurement and Representation of Natural Systems*, New York: North-Holland.

Rouse, W. B. (2005), Enterprises as Systems: Essential Challenges and Approaches to Transformation, *Systems Engineering*, 8(2): 138-150.

Rouse, W. B. and N.M. Morris (1986), On Looking Into the Black Box: Prospects and Limits in the Search for Mental Models, *Psychological bulletin*, *100*(3): 349.

Rouse, W. B., and N. Serban (2011), Understanding change in complex socio-technical systems: An exploration of causality, complexity and modeling, *Information, Knowledge, Systems Management*, *10*(1): 25-49.

Rouse, W., and R. Henneman (1986), On Measuring the Complexity of Monitoring and Controlling Large-Scale Systems, *IEEE Transactions on Systems, Man, and Cybernetics*, *16*(2): 193-207.

Rouse, W.B. (1980), *Systems Engineering Models of Human-Machine Interaction*, New York: North Holland.

Rouse, W.B. (1983), Models of human problem solving: Detection, diagnosis, and compensation for system failures, *Automatica, 19*(6): 613-625.

Rouse, W.B. (2007), *People and Organizations: Explorations of Human-Centered Design*, New York: Wiley.

Rouse, W.B. (Ed.) (2006), *Enterprise Transformation: Understanding and Enabling Fundamental Change*, Hoboken, NJ: Wiley.

SAE (2014), Compliance verification criterion standard for SAE AS6081, fraudulent/counterfeit electronic parts: avoidance, detection, mitigation, and disposition – distributors. http://standards.sae.org/wip/as6301/. Retrieved 6/4/2014.

Salton, G. (1968), *Automatic Information Organization and Retrieval*, New York.

SASC (2012), Inquiry into counterfeit electronic parts in the Department of Defense supply chain, Washington, DC: Author (Senate Armed Services Committee).

Schwarz, E., I.P. Beldie, and S. Pastoor (1983), Research Note: A Comparison of Paging and Scrolling for Changing Screen Contents by Inexperienced Users, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *25*(3): 279-282.

Snowden, D.J. and M.E. Boone (2007), A Leader's Framework for Decision Making, *Harvard Business Review*, 85(11): 68-76.

Stephens, D. W. (1986), *Foraging Theory*, Princeton University Press.

Stephens, D. W. and E.L. Charnov (1982), Optimal Foraging: Some Simple Stochastic Models, *Behavioral Ecology and Sociobiology*, *10*(4): 251-263.

Stradley, J. and D. Karraker (2006), The electronic part supply chain and risks of counterfeit parts in defense applications, *IEEE Transactions on Components Packaging Technology*, 29: 703-705.

Tainter, J. (1988), *The collapse of complex societies*, Cambridge: Cambridge University Press.

Takken, S. and W.B.L. Wong (2013), Tactile reasoning: Hands-on vs. Handsoff what's the difference? Accepted for the *11th International Conference on Naturalistic Decision Making*.

Thorpe, S., F. Denis and C. Marlot (1996), Speed of processing in the human visual system, *Nature*, *381*(6582): 520-522.

Todd, J. T. (2004), The Visual perception of 3D shape, *Trends in Cognitive Science*, *8*(3): 115-121.

Tolk, A. (2011), Enhancing simulation composability and interoperability using conceptual/semantic/ontological models, *Journal of Simulation*, 5(3): 133-134.

Tolk, A., S.Y. Diallo, J.J. Padilla, and H. Herencia-Zapana (2013), Reference modeling in support of M&S - foundations and applications, *Journal of Simulation*, 7:69-82.

Treiber, M., A. Hennecke, D. Helbing (2000), Congested traffic states in empirical observations and microscopic simulations, *Physical Review E*, 62: 1805-1824.

Vickrey, W. (1969), Congestion theory and transport investment, *American Economic Review*, 59: 251-260.

Villasenor, J., and M. Tehranipoor (2013), Chop-shop electronics, *IEEE Spectrum*, 50: 40-45.

Wang, W. G., A. Tolk, and W. P. Wang (2009), The levels of conceptual interoperability model: Applying systems engineering principles to M&S, *Proceedings of the Spring Simulation Multiconference*, Spring Sim 2009, San Diego, CA.

Weaver, W. (1948), Science and complexity, *American Scientist*, 36: 536–544.

Zeigler, B.P., H. Praehofer, and T. G. Kim (2000), *Theory of Modeling and Simulation*, 2nd ed, Amsterdam: Academic Press.

Zhang, L., D. Levinson, and S. Zhu (2008), Agent-based model of price competition, capacity choice, and product differentiation on congested networks, *Journal of Transport Economics and Policy*, 42: 435-461.

APPENDIX – SURVEY OF VISUALIZATION TOOLS

In order to support the development of enterprise visualizations, a survey was conducted of existing off the shelf visualization tools to assess their strengths and weaknesses for potential utilization in any enterprise visualizations. Eight popular tools were considered: d3, R, Microstrategy, QlikView, TIBCO Spotfile, Tableau, SAS, IBM Cognos. For each package, we considered its intended use, its advantages, and its disadvantages. Please note that the included references are links to websites providing additional information about the products.

Data-Driven Documents (d3.js)

Summary

Visualization package d3.js is a Javascript library intended to manipulate data-based documents [1] and generate a vast array of visual representations such as scatterplot matrices, node-link trees, bubble charts, and so on [2]. D3.js uses HTML, Scalable Vector Graphics (SVG) and Cascading Style Sheets (CSS) standards, avoiding proprietary representation [3]. It supports extensive datasets and its dynamic nature is suitable for interaction and animation.

Advantages

D3 also allows users to change particular elements and attributes of a representation, as opposed to changes of the representation itself, avoiding unnecessary computations and intermediate graphs (since the document is edited directly) [3].

Considering the fact that d3.js builds on existing web technologies, its flexibility evolves with that of the key standards used. The availability of different visualization opportunities is therefore an increasing matter. Moreover, the burden placed on the developer decreases significantly since the browser's inner capabilities are used to build-in functionality.

Disadvantages

Whereas it is relatively straightforward to evaluate d3's capabilities in terms of flexibility and performance, the steepness of potential users' learning curve is not a trivial matter. Different users have acknowledged different learning experiences, ranging from simple to complex ones [3]. Inevitably, the user's skills place an important role here, even though a vast documentation and practical examples are available on the web, as well as an active community able to help new users [1].

Finally, handling large data sets may result in a slower manipulation of the Document Object Model (DOM) [4]. Additionally, the Scalable Vector Graphics (SVG), the vector image format used in d3, also presents performance limitations when dealing with large data sets.

R

Summary

R is an open-source programming environment that facilitates data manipulation, calculation and generation of graphical representations. It provides a vast array of classical and modern statistical techniques (e.g. linear and nonlinear modeling, classical statistical tests, time-series analysis, clustering, and so on), being extremely popular among data miners and statisticians [5]. Some of these statistical techniques are part of the R environment, while others come from standard and recommended packages and from the CRAN family of Internet sites and other sources [6].

Advantages

R presents good data handling and advanced graphical capabilities, resulting in high-quality static graphs and easy inclusion of mathematical symbols and formulae where required. Additional packages allow for dynamic and interactive graphics.

Additionally, R provides all the basic and most important functions, while new ones may be developed, allowing for greater flexibility and customization. Given its open-source nature, new features are available quickly, but they are also more prone to development errors [7].

Disadvantages

Lastly, as a low level-programming environment, R is more suitable for users who have previous programming skills, or for those who are willing to learn and understand coding. Similarly to other packages, a comprehensive documentation is available online as well as a big community of users that may provide additional support.

Microstrategy

Summary

Microstrategy develops four software platforms: analytics platform, intended to analyze a wide variety of data and produce actionable conclusions; mobile platform, intended to develop code-free apps based on analytics, transactions, and multimedia content; identity platform, which adds an extra security layer to organizations by providing biometrical identification on smartphones; and finally, loyalty platform that gives companies a means for targeted marketing, commerce, and consumer engagement [8]. Microstrategy's business intelligence services have been used among different sectors, including education, technology services, communications and media, manufacturing, financial services, and others.

Advantages

One of the main advantages of Microstrategy resides on its cloud service for all its products, allowing customers to deploy in the cloud, on its premises or even adopt a hybrid solution. This provides a continuous access to information, regardless of organizations' privacy policies, location or devices used [9]. Microstrategy also adopts an open architecture that enables its partners to develop specific apps that can be easily integrated into the cloud.

Disadvantages

The growth of Microstrategy is, however, hampered by mega vendors that adopt aggressive strategies to increase their market share, decreasing the capability of smaller companies, such as Microstrategy, to grow and thus innovate [10].

At the same time, the robustness and flexibility of Microstrategy's services, result in a complex business intelligence platform, and therefore in a steeper learning curve for developers and users. Nevertheless, Microstrategy is betting on training sessions that show users how to create projects rapidly, something that stands for simple projects but not for complex ones, which require a sophisticated metadata model [10].

Lastly, Microstrategy uses a multitool development model to create and combine reports: the fact that individual report objects (e.g. tables and charts) are created on the Web or desktop and are then combined into documents or dashboards using another tool, leaves customers with a sense of a tiresome process [10].

QlikView

Summary

QlikView is a user-centric data discovery platform that generates dynamic views of the information needed for making decisions [11]. Users are able to ask and answer questions without following pre-defined paths or preconfigured dashboards, or without creating new reports or data visualizations. This information is kept in memory so that multiple users can access it quickly; for too large datasets, QlikView connects users to the data source. The more users using the software, the more information is generated, the more insightful the platform becomes.

Advantages

QlikView's intuitive nature allows end users to easily find relationships, trends and deviations in data as well as unrelated elements, without having to model the entire study in advance or write sophisticated SQL. [12].

This does not mean, however, that the analyses run using QlikView are simplistic; on the contrary, users are able to conduct complex analyses and take full advantage of QlikView's ample functional capabilities that include dashboards, interactive visualization, geospatial intelligence, big data support, mobile business intelligence and others.

Disadvantages

Despite the positive reviews, some concerns have also been raised in terms of QlikView's ability for handling security and managing a large community of named users [12]. Other aspects such as the management of metadata and the business intelligence infrastructure are described as weaknesses of the platform.

TIBCO Spotfire

Summary

The TIBCO Spotfire platform is intended to provide customers a flexible, user-centered data discovery and visualization tool, allowing for interactive dashboards, predictive analytics and analytic apps, collaborative environments and access to cloud services [13]. Spotfire integrates with more than 30 databases, such as Teradata Aster or DataSynapse GridServer, increasing customers' access to information [9]. Its predictive capabilities also enable the use of current predictive models from other applications, the creation of new ones using TIBCO enterprise runtime for R, R or S^+ or using Spotfire's predictive modeling tools [13]. Given its capacity to generate a wide range of graphics and visualizations (such as root cause analyses, dynamic control charts, energy surveys, and so on), Spotfire has a strong presence in distinct markets (e.g. financial services, energy, manufacturing, life sciences, etc.) [9,14].

Advantages

Aiming to have a platform for real-time and bidirectional integration with business processes [12], TIBCO has acquired three companies specialized in process monitoring and even stream analyses, geospatial analytics and real-time Key Performance Indicators (KPIs) and user experiences on mobile appliances. This, however, adds an additional challenge that is to enhance and manage Spotfire's usability, in particular for developers.

TIBCO is also building strong capabilities on the prediction of future markets by gathering and analyzing events, trends and deviances, all potential inputs to rich but simple data exploration interfaces. Indeed, the simplicity of the current software (specially from end users' perspective), and its functionality, even for complex analyses, are Spotfire's dominant factors. Plus, its stability, reliability and error-free qualities are attractive to potential users [12].

Disadvantages

Nevertheless, other aspects such as the creation of complex reports that use information from different sources (as opposed to simple ones) is a challenging objective, taking, on average, 8.7

days [12]. Another significant hindrance to the adoption of the software is its cost, even though the company has announced a new pricing model to tackle this subject.

Tableau

Summary

Tableau is a data exploration and visualization package that allows users to access relevant information and conduct different studies, such as quality and safety assessments, investment and sales strategies, customer service improvement, operations integration, and others [16]. Tableau's capabilities also include social media analytics, mapping software, mobile business intelligence, predictive capabilities and integration with R, and survey and time series analyses [15]. It has been used by government agencies, universities, small and medium businesses (SMB) and non-profit organizations [16].

Advantages

Tableau's main differentiator is its deep understanding of the market, allowing for simple to complex studies to be conducted, while maintaining its accessibility to ordinary users and developers. Its functionality, low implementation costs, data integration and visualization capabilities and geospatial agility are also highly praised.

The ever more critical enterprises' needs for reusability, scalability and embeddability seem to drive the company expansion strategy [12]. Indeed, the frequent but easy to migrate product releases may be in the core of such vision.

Disadvantages

Nevertheless, Tableau is frequently used as a complement to existing business intelligence platforms, which poses additional challenges in terms of governance and consistency when multiple tools are deployed. More importantly, such strategy tends to be risky, as business intelligence companies continuously look for new ways to fully address customers' needs, which may result in innovative products or simply the acquisition of other companies that are able to bridge such gap. In addition, Tableau lacks production reporting capabilities, and have high annual maintenance fees.

SAS

Summary

The visualization package SAS is intended to support visual statistics and analytics activities, combining a wide range of capabilities, such as data mining, descriptive modeling and predictive modeling, simulation and optimization [17,18]. Others features include auto charting, interactive dashboards and reports (with the option of being geospecific), cloud services, and a collaborative environment using mobile business intelligence apps. SAS incorporates analytical knowledge from different sectors (banking, communications, education, government, and so on), with an in-memory engine for data analysis, also supporting other platforms such as Teradata and Pivotal databases.

Advantages

In fact, this is SAS's main differentiator among other visualization packages: by producing domain- and industry- specific visual analytics applications, SAS focuses on particular business problems, adding more value to its users.

The platform also supports complex analyses using large volume of data, and is able to combine a simple production reporting with an interactive and visual data searching activity, that bring to light correlations, clusters and forecasts [12].

As important is SAS's strategy of using a unique platform for all its analytics applications, providing customers a single platform, but directly competing with leading-edge data discovery vendors.

Disadvantages

On the downside, SAS is considered difficult to use, limiting its large deployment. Moreover, the high ownership costs may preclude small businesses or entities from acquiring it. Other less satisfying aspects include some of SAS's functionalities such as reporting (it may take several days to create a report), dashboards, online analytic processing and interactive visualizations [12].

IBM cognos

Summary

The IBM Cognos family provides a full spectrum of business intelligence, performance management and analytics capabilities and solutions for any industry, domain and geography [19]. The platform allows for complex analyses, modeling and planning activities, as well as the generation of reports and a collaborative environment. Its applicability ranges from individual usage to small and medium businesses to large enterprises. IBM Cognos analytics solutions are built on open standards, allowing for single use, a combination with other products or even as part of a larger solution.

Advantages

The IBM Cognos platform deals with some of the largest deployments and is highly acclaimed by its completeness of vision [12]. Users also highlight its metadata management and business infrastructure capabilities, together with other features such as ad hoc reporting, dashboards and online analytic processing.

Additionally, IBM supports a web community of visualization experts, practitioners, academics and others, through the IBM Many Eyes community: using IBM Rapidly Adaptive Visualization Engine (RAVE), participants are able to describe their ideal visualization outcomes, and IBM patented methods present alternative visualizations ranging from line or bubble charts to heatmaps and phrase nets [20]. Other, to-be-released, IBM platform (Watson Analytics) will complement this software. Its main objective is to use algorithms to interpret natural language queries, access to databases, correlate information and finally draw conclusions able to be understood by users without technical or statistical expertise. New iterations may be run to further clarify the analysis.

Disadvantages

This new venture may address one of the disadvantages of IBM Cognos: the users' and developers' steeper learning curve when compared to other packages, especially when dealing with more complex analyses [12]. Other features that are possibly behind mixed reviews of IBM Cognos are related to geospatial intelligence, interactive visualization and embedded analytics and the reporting process that may take an average of 6 days to be produced. The high cost of the software is also considered a significant obstacle to users' adoption and software deployment.

[1] http://d3js.org
[2] https://github.com/mbostock/d3/wiki/Gallery
[3] Michael Bostock, Vadim Ogievetsky and Jeffrey Heer, D^3: Data-Driven Documents, 2011
[4] http://blog.visual.ly/why-d3-js-is-so-great-for-data-visualization/
[5] http://www.r-project.org
[6] http://CRAN.R-project.org
[7] http://www.analyticsvidhya.com/blog/2014/03/sas-vs-vs-python-tool-learn/
[8] http://www.microstrategy.com/Strategy/media/downloads/about-us/press-kit/Corporate-Brochure_PressKit.pdf
[9] http://www.gartner.com/doc/2647924?ref=SiteSearch&refval=&pcp=mpe#a149525237
[10] http://www.microstrategy.com/Strategy/media/downloads/products/Microstrategy-Gartner-SWOT.pdf
[11] http://www.qlik.com/us/explore/products/qlikview
[12] http://www.gartner.com/technology/reprints.do?id=1-1R2GHY1&ct=140224&st=sg
[13] http://spotfire.tibco.com
[14] http://spotfire.tibco.com/demos?Section=Hands-on+Demos
[15] http://www.tableausoftware.com
[16] http://www.gartner.com/doc/2647924?ref=SiteSearch&refval=&pcp=mpe#a-173691061
[17] http://www.sas.com/en_us/software/analytics/visual-statistics.html
[18] http://www.sas.com/en_us/software/business-intelligence/visual-analytics.html

[19] Cognos: BI: The IBM Cognos family - Analytics in the hands of everyone who needs it: http://www-01.ibm.com/software/analytics/cognos/
[20] http://www-01.ibm.com/software/analytics/many-eyes/