2017

Integrating IDEF0 into a Systems Framework for Statistical Engineering

Teddy Steven Cotter

Proceedings of the American Society for Engineering Management 2017 International Annual Conference
E-H. Ng, B. Nepal, and E. Schott eds.

INTEGRATING IDEF0 INTO A SYSTEMS FRAMEWORK FOR STATISTICAL ENGINEERING

T. Steven Cotter
Old Dominion University

tcotter@odu.edu

Abstract
Driven by a growing requirement during the 21st century for the integration of rigorous statistical analyses in engineering research, there has been a movement within the statistics and quality communities to evolve a unified statistical engineering body of knowledge (Horel and Snee, 2010; Anderson-Cook, 2012). Outside of the 2014 Statistical Engineering Agreement among the ASQ Statistics Division, DOT&E, NASA, and IDA, there has been little formal progress toward this goal since the May 2011 NASA Symposium on Statistical Engineering in Williamsburg Virginia. In the ASEM-IAC 2012, Cotter (2012) identified the gaps in knowledge that statistical engineering needs to address, explored additional gaps in knowledge not addressed in the prior works, and set forth a working definition of and body of knowledge for statistical engineering. Again in the ASEM-IAC 2015, Cotter (2015) proposed a systemic causal Bayesian hierarchical model that addressed the knowledge gap needed to integrate deterministic mathematical engineering causal models within a stochastic framework. Missing, however, is the framework for specifying the hierarchical qualitative systems structures necessary and sufficient for specifying systemic causal Bayesian hierarchical models. This paper proposes revisions to and integration of IDEF0 as the framework for developing hierarchical qualitative systems models.

Keywords
Qualitative Framework, Statistical Engineering, Systems

Introduction
The systems approach to the design of new systems or correction of performance problems with existing systems involves some form of hierarchical decomposition of functional requirements following the V model. The major problem with functional hierarchical decomposition is maintaining the relevant between layer and within layer cross correlations and causal environmental and functional effects at each sub-system, module, and component level. Cotter (2012) proposed statistical engineering as a causal Bayesian hierarchical modeling approach to directly address functional decomposition while assuring maintenance of between layer and within layer cross correlations and causal environmental and functional effects. He defined statistical engineering as the discipline that solves applied complex organizational and societal problems that involve elements of risk, imprecision, or uncertainty in their systemic mission outcomes. Statistical engineering integrates stochastic, technical, engineering, information, human factors, managerial, financial, and economic knowledge as necessary to understand, model, and correct deficiencies (the *problems*) in, systemic performance relative to required mission outcomes within an environmental context. Latent design or existing manifested efficiencies in systemic performance, however, are only symptoms. They must be decomposed into their subsystem, module, and component level root causes that contribute to the deficient systemic performance, modeled, redesigned, corrected, and synthesized into improved systemic performance that can be verified against the redesign intent and validated in the environmental context. This paper reports research into development of a statistical engineering qualitative decomposition-synthesis modeling framework over which systemic causal Bayesian hierarchal models can be overlaid to facilitate the modeling, redesign, and correction activities.

Causal Bayesian Hierarchical Modeling Framework
Cotter (2015) developed the quantitative causal Bayesian hierarchal model as,

$$\begin{aligned}
&\text{Min } \mathbf{Y}_{\text{Total}} = f(\mathbf{w}'(\mathbf{Y}_{\text{pred}} - \mathbf{T})) \qquad (1)\\
&\text{s.t.}\\
&\mathbf{Y} = \mathbf{F}(pa_{\text{i}}, u_{\text{xi}})\boldsymbol{\beta} + \mathbf{F}(pa_{\text{j}}, u_{\text{zj}})\boldsymbol{\gamma} + \varepsilon\\
&\mathbf{L}_{\text{X}} \le \mathbf{F}(pa_{\text{i}}, u_{\text{xi}}) \le \mathbf{U}_{\text{X}}\\
\text{possibly} \quad &\mathbf{L}_{\text{Z}} \le \mathbf{F}(pa_{\text{j}}, u_{\text{zj}}) \le \mathbf{U}_{\text{Z}}
\end{aligned}$$

Cotter (2016) demonstrated the applicability of this model to hierarchical decomposition of mission outcomes versus requirements topologies to sub-system, module, and component functional requirements admitting (1) integration of multimodel ensembles, (2) estimation of posterior topologies adjusted for bias between deterministic causal and stochastic terms, (3) estimation of environmental and mission performance informative prior topologies and nonconjugate but proper decomposition topologies, (4) estimation of time-varying parameters, (5) estimation of between layer and within layer cross correlations and causal effects, and (6) modeling of minimum and maximum extreme value rare events. Not addressed in these works was the qualitative decomposition-synthesis modeling framework over which the quantitative model is to be overlaid. The qualitative framework is required to assure necessary and sufficient specification of (1) top level systemic mission identification within the relevant environmental topology, (2) coordinated hierarchical decomposition of relevant between layer and within layer cross correlations and causal environmental and functional effects at each sub-system, module, and component level, (3) functional scaffolding of associated input and output subsystems, modules, and components to and from those on the decomposition path, and (4) economic and socio-technical constraint identification and decomposition.

IDEF0 Modeling Methodology

The Integrated DEFinition (IDEF) languages provide the most mature and comprehensive specifications for modeling qualitative functional, informational, and dynamical properties of systems. The IDEF languages were developed originally by the US Air Force Materials Laboratory in the 1970s and 1980s as part of its Integrated Computer-Aided Manufacturing (ICAM) initiative. The National Institute of Standards and Technology (NIST) issued Federal Information Processing Standards (FIPS) Publication 183 renaming the document Integrated Definition for Function Modeling (IDEF0). Of fifteen proposed IDEF standards, only the following six have been fully developed.

- IDEF0: Function modeling of the transformations, processes, and activities within a system.
- IDEF1: Information modeling of the semantics and structure of the information exchanges necessary to support systemic transformations, processes, and activities.
- IDEF1X: Data modeling of relational databases.
- IDEF2: Simulation model design for representing the time varying systemic behavior.
- IDEF3: Process description capture design tool for capturing and maintaining code.
- IDEF4: Object-oriented design
- IDEF5: Ontology description capture

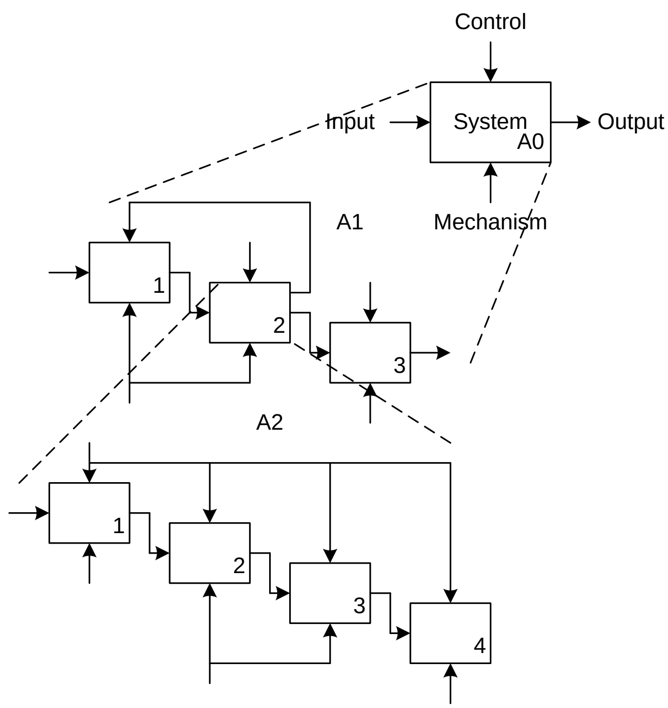
Only IDEF0 and IDEF1 are necessary as the foundation for the qualitative decomposition-synthesis modeling framework of statistical engineering. Review of IDEF1 indicates that its information modeling methodology is sufficient for this proposed statistical engineering framework. IDEF2 (to be integrated later) will be necessary as the foundation for modeling dynamic systemic properties.

IDEF0 was developed from the Structured Analysis and Design Technique (SADT) developed by Douglas Ross (1977a, 1977b) at SoftTech, Inc. for modeling systems as a hierarchy of functions. IDEF0 follows the SADT modeling process by representing the system as a top-level context diagram and decomposing it into its major sub functions, processes, and activities in increasingly detailed child diagrams (Exhibit 1). Relative to statistical engineering's causal Bayesian hierarchical modeling requirements, the IDEF0 has the following deficiencies.

> Clause 2.2 "Arrow: … an open channel or conduit conveying data or objects …." This definition assumes causal relationships. The Arrow definition provides no means for modeling the coupling tightness between functions or cross hierarchy and interlayer cross correlational or causal influence relationships. Additionally, it is not necessary or sufficient for causal Bayesian hierarchical Y = f(X, Z, ux, uz) distributional structures that may require control inputs be related to non-parental diagrams. The (ux, uz) may be partially determined by non-parental diagram cross causal or correlational components.
>
> Clause 3.1.1 covers only "…people, machines, materials, computers and information …." For causal Bayesian hierarchical modeling, the system must cover people and their people-to-people

(P2P), intelligent machine-to-machine (M2M), people-to-machines (P2M), and machines-to-people (M2P) information exchanges, physical interactions, and material exchanges.
Clause 3.2.1 requires only that "Rules define how the components are used …." As in ontological engineering, rules must be the axiomatic specifications of component interfunctionality within a parental hierarchy and cross functionality between parental hierarchies.

Exhibit 1. SADT—IDEF0 Functional Decomposition.



Clause 3.2.2.1 Box and Arrow Semantics require only the verbal function name. There is no provision for relating the verbal functional name to its causal Bayesian hierarchical Y = *f*(X, Z, u_x, u_z) distributional structure.
Clause 3.2.2.3 Box and Arrow Semantic Rules – Rule 4, "Arrow segments … shall be labeled with a noun or noun phrase …." For Control Arrows, simple noun phrases are neither necessary nor sufficient to specify cross hierarchy and interlayer cross correlational influence and causal conditional distribution relationships.
Clause 3.3.1 Types of Diagrams provides for only within hierarchy Parent-Child decomposition. There are no provisions for *f*(*X*, *Z*, u_x, u_z) scaffold models of influential or influenced parallel P2P, M2M, P2M, or M2P hierarchies
Clause 3.3.2.1 Arrows as Constraints require that "Only at low levels of detail can arrows represent flow or sequence …." Flow and sequence constraints may also be determined by causal Bayesian hierarchical Y = *f*(*X*, *Z*, u_x, u_z) distributional structure.
Clause 3.3.2.2 Activations of a Box require that "A box may perform various parts of its function under different circumstances, using different combinations of its input and controls and producing different outputs." This is a deterministic definition that does not admit covariance correlational or causal influence.
Clause 3.3.4.2.2. Node Tree requires "… the hierarchy shall be shown graphically as a tree rooted at a chosen node. This definition is necessary and sufficient for the systemic mission, but it does not admit cross hierarchy and interlayer cross correlational influence and causal conditional distribution constraint relationships emanating from the environment.

VSM Systems Governance Framework

One of the key differences between this statistical engineering research initiative and that of ASQ Statistics Division, DOT&E, NASA, and IDA is the first modifier adjective, "systemic." The prior initiatives are an extension of statistical modeling to engineering problems with reference to systems thinking. As a "systemic causal Bayesian hierarchical modeling" methodology, this initiative is built within the general systems framework. The Viable System Model (VSM - Beer, 1972, 1979, 1985; Espejo and Harnden, 1989) forms the socio-technical, cybernetic systems framework for this statistical engineering research initiative. Beer codified the VSM's principles, theorems, and laws, which he intended to be a manager's guide. Two features of the VSM are most applicable to the development of an integrated systemic framework: the cybernetic structure and the recursive system theorem.

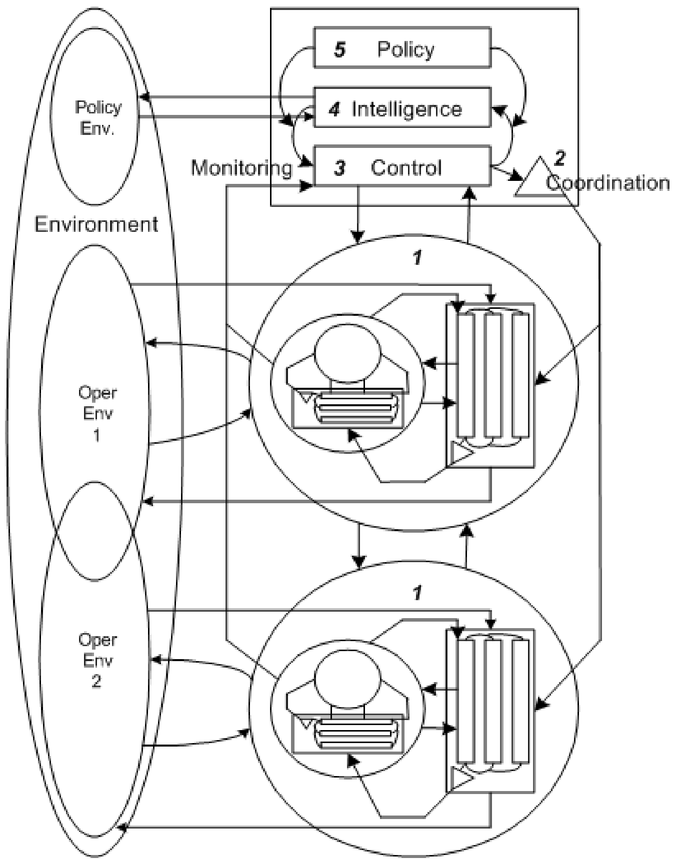
According to the VSM's cybernetic model, five interacting subsystems are necessary and sufficient for systemic viability. The first subsystems, indicated as 1 in Exhibit 2, are those that produce the system. The circles, are the productive systems that transform inputs into outputs and create the system's purpose for existing. Subsystem 2, the triangle, is an anti-oscillatory function that coordinates vertical interactions among operational subsystems 1 and the subsystem 3 control function. Its purpose is to ensure that the correct pattern and amount of variety is communicated to the control function and, at the same time, reduce the variety demand on the control function. The control function's goal is to achieve operational cohesion among the subsystems 1. By cohesion, it is meant that the control function must find a systemic balance between control for efficient subsystems 1 functioning and autonomous flexibility to assure the correct pattern and amount of environmental variety is distributed among the subsystems 1 to attain proper adaptation. The control function assures that the flexible, operating autonomy allocated to the subsystems 1 units remains consistent with systems functioning through an independent monitoring channel. Monitoring is accomplished through an extra communications channel directly connected to subsystems 1 operations. The internal, operational focus of the control function 3 is counterbalanced by the external, environmental focus of intelligence function 4. The intelligence function scans the competitive environment for threats to system viability and opportunities to expand the system's environmental niche. The control and intelligence functions counterbalance in that they seek to achieve the same goal but from internal versus external perspectives: the definition, implementation, and adjustment of the system's identity and viability in its niche. Counterbalance is achieved by the policy function 5. The policy function chooses systemic courses of action based on environmental information filtered through the intelligence function and internal operational information filtered through the control function. The policy function orchestrates and monitors the debate between the intelligence and control functions with a goal of choosing those courses of action that maximize systemic viability. The policy function also serves as the knowledge repository of the system, storing knowledge on the outcomes of past actions and applying that knowledge to the debate on present and future courses of action. Exhibit 2 also illustrates the most important feature of the VSM: its recursive structure. Subsystems 1, those that produce the system, must themselves be viable systems and contain subsystems that, in turn, produce them. Thus, VSM's neurophysiological model directly implies the Recursive System Theorem.

> In a recursive organizational structure, any viable system contains, and is contained in, a viable system (Beer *The Heart of Enterprise* 118).

Recursion within the VSM is the cybernetic, structural linking process between environmental selection dynamics and internal adaptive processes. The Recursive System Theory implies that competitive environments are themselves meta-systemic forms taking on the same VSM neurophysiological structure. Under the Recursive System Theorem, socio-technical systems are subsystems 1 in populations of socio-technical systems, and, in turn, populations of socio-technical systems are subsystems 1 in communities of socio-technical systems. At each level, policy, intelligence, and control functions are formalized or are worked out through higher-level self-organization. The competitive environment itself is not a black box; rather, the environment is a viable system in which selection processes are worked out through interactions among socio-technical systems and populations of socio-technical systems, the later which contains constituencies of socio-technical systems in a given population. Apparent environmental chaos arises because the interactions induce nonlinearities and randomness into selection processes.

Specification of the VSM leads immediately to IDEF0's major deficiency relative to statistical engineering's systemic causal Bayesian hierarchical modeling requirements. As illustrated in Exhibit 3, Clause 3.4.3 admits modeling the environmental context above the A0 system level node as A-n nodes, where the negative n represent levels above the A0 system level. The IDEF0 specification allows environmental context modeling to proceed "…just as ordinary detail modeling, the only difference being the negative numbering (and the non-definitive, but normative interpretation) that preserves A0 as the "origin" of the node-number-based coordinate system for all model references." As illustrated in Exhibit 3, "…the negative-node-number modeling merely provides more and more details about the sources and uses of the external boundary conditions." This specification does not conform to the Recursive System Theorem by admitting the VSM's cybernetic control structure and the function 5 environmental scanning at each subsystem level as illustrated in Exhibit 3. An example of the necessity of the function 5 environmental scanning at each subsystem level is found in every production organization. Components and modules are received at the function 1 subsystems level. Defects or deficiencies that adversely impact function 5 policy will be held at the function 1 level but communicated to the function 3 control level (typically purchasing and quality control management) for resolution. Resolution will be coordinated at the functions 3, 2, and 1 levels with only summary reporting to the organization's functions 2 intelligence and 1 policy level. Each functional subsystem must have its respective functions 1 to 5 for this coordination.

Exhibit 2. The VSM's cybernetic control structure.

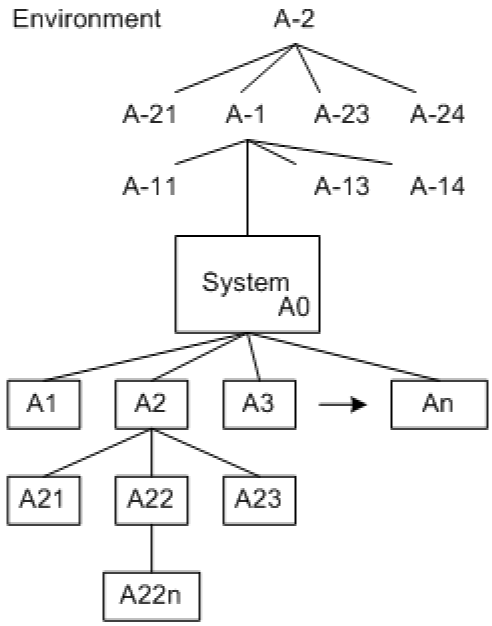


Proposed Statistical Engineering Qualitative Modeling Framework

This statistical engineering initiative adopts IDEF0 with the following revisions as the qualitative framework for specifying the hierarchical qualitative systems structures necessary and sufficient for specifying systemic causal Bayesian hierarchical models.

> Clause 2.2 "Arrow: … an open channel or conduit conveying data or objects …."
>
> Clause 2.2.R1: The Arrow may model resources or information coupling tightness between functions with $0 < C < 1.0$ value for mathematical modeling or a storage symbol on the arrow to represent inventory or information buffer decoupling.

Exhibit 3. IDEFO's higher-level context diagrams specification.



An arrow with no indication of coupling tightness represents a coupling of C = 1.0. No arrow between functions represents a coupling tightness of C = 0.

Clause 2.2.R2: The Control Arrow may indicate the time lag, feedback gain, or control function.

Clause 2.2.R.3: The Arrow may model cross hierarchy and interlayer cross correlational or causal influence relationships as necessary and sufficient to model the (ux, uz) parental hierarchy of the causal Bayesian hierarchical Y = f(X, Z, ux, uz) distributional structures.

Clause 3.1.1 covers only "…people, machines, materials, computers and information …."

Clause 3.1.1.R1: In socio-technical systems where technical inputs and transformations are made by intelligent machines as part of the decision-transformation flow, each box must be labeled as P = people or M = intelligent machine in the lower left corner of the respective box. Arrows between boxes shall be labeled as M2M = machine-to-machine, P2M = people-to-machine, M2P = machine-to-people (M2P), and P2P = people-to-people information exchanges, physical interactions, and material exchanges

Clause 3.1.1.R2: In socio-technical systems involving only automation transformations, the Context Diagram will be so noted and the requirement for P and M labeling is not required.

Clause 3.1.1.R3: In purely social or technical systems, the Context Diagram will be so noted and the requirement for P and M labeling is not required.

Clause 3.2.1 requires only that "Rules define how the components are used …."

Clause 3.2.1.R1: To be consistent with ontological engineering, rules must be the axiomatic specifications of component interfunctionality within a parental hierarchy and cross functionality between parental hierarchies.

Clause 3.2.2.1 Box and Arrow Semantics require only the verbal function name.

Clause 3.2.2.1.R1: The box names shall be a verb or verb phrase supplemented by its causal Bayesian hierarchical $Y = f(X, Z, u_x, u_z)$ transformation necessary and sufficient to fully describe its distributional structure.

Clause 3.2.2.3 Box and Arrow Semantic Rules – Rule 4, "Arrow segments … shall be labeled with a noun or noun phrase …."

Clause 3.2.2.3.R1 Control Arrows, shall specify cross hierarchy and interlayer cross correlational influence and causal conditional distribution relationships necessary and sufficient to support the full description of the causal Bayesian hierarchical $Y = f(X, Z, u_x, u_z)$ model.

Clause 3.3.1 Types of Diagrams provides for only within hierarchy Parent-Child decomposition.

Clause 3.3.1.R1: Diagrams shall specify causal Bayesian hierarchical $Y = f(X, Z, u_x, u_z)$

transformations including supporting scaffolds necessary and sufficient to fully describe systemic distributional structure of influential or influenced parallel M2M, P2M, M2P, or P2P hierarchies. This specification is illustrated in Exhibit 4.

Clause 3.3.2.1 Arrows as Constraints require that "Only at low levels of detail can arrows represent flow or sequence …."

Clause 3.3.2.1.R1: Flow and sequence constraints shall be determined by and model as accurately as possible the causal Bayesian hierarchical $Y = f(X, Z, u_x, u_z)$ distributional structure.

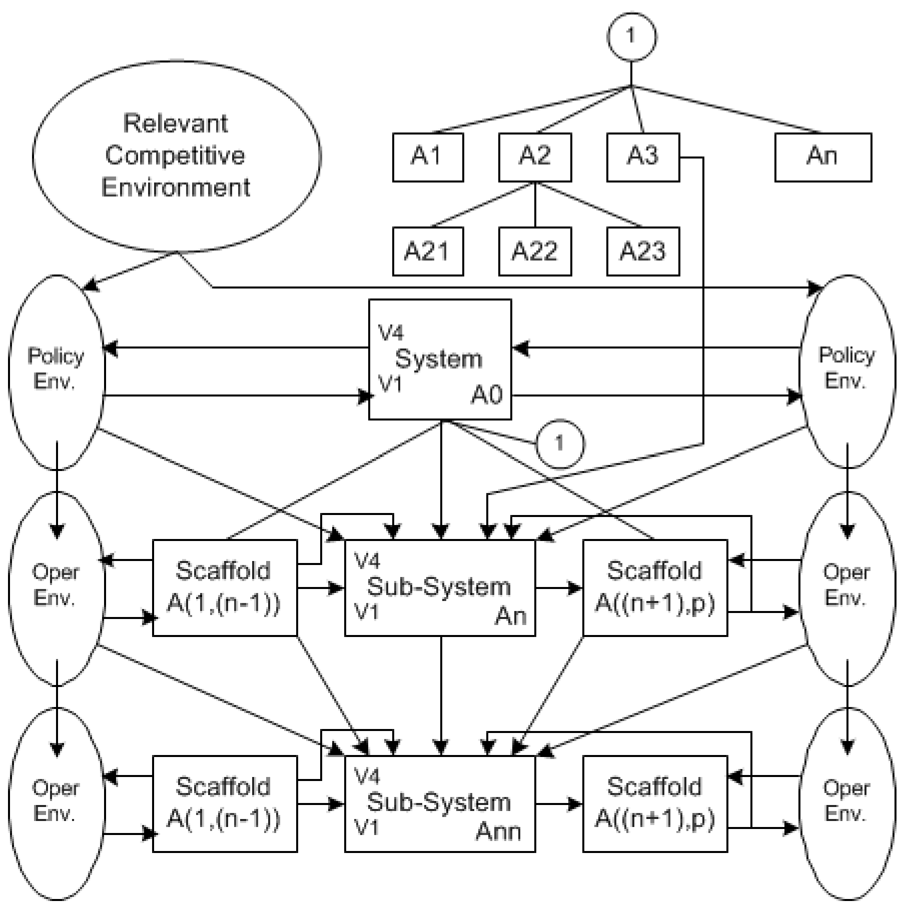
Clause 3.3.2.2 Activations of a Box require that "A box may perform various parts of its function under different circumstances, using different combinations of its input and controls and producing different outputs."

Clause 3.3.2.2.R1: A box may perform various parts of its function using different combinations of its inputs, control, and within layer and interlayer covariance correlational or causal influence.

Clause 3.3.4.2.2. Node Tree requires "… the hierarchy shall be shown graphically as a tree rooted at a chosen node.

Clause 3.3.4.2.2.R1: The qualitative diagram shall be represented as a node structure including the Viable System Model levels 1 and 4, represented as V1 and V4, within each layer in accordance with the Recursive System Theorem necessary and sufficient to specify the causal Bayesian hierarchical $Y = f(X, Z, u_x, u_z)$ transformations, supporting scaffolds, and VSM cybernetic structure. This specification is illustrated in Exhibit 4.

Exhibit 4. Full qualitative causal Bayesian hierarchical representation with interlayer cross correlational or causal relationships, scaffolding, and VSM cybernetic structure.



Continuing Research into Statistical Engineering Stochastic-Causal Modeling

This work has extended the IDEF0 functional modeling specification to a methodology for the decomposition and specification of qualitative systemic models that are necessary and sufficient for the overlay of the quantitative systemic causal Bayesian hierarchal model (1). Continuing research is directed toward developing a hierarchical causal Bayesian socio-technical modeling core reference ontology and body of knowledge. Future research will be directed toward:

- Systems constraint decomposition and synthesis.
- Systems boundary interface decomposition and synthesis.
- Systems performance functional activation decomposition and synthesis.
- Systems with nonrecursive directed acyclic graph feedback loops.
- Model synthesis and verification.
- IDEF2 foundation for modeling dynamic systemic properties, unintended consequences, multiple criteria optimization, and robustness to unintended consequences and residual noise.

References

Anderson-Cook, Christine M. and Lu Lu. (2012) Editorial for Statistical Engineering Special Issue. *Quality Engineering* 24:2, pp. 107-109.

Beer, Stafford. *Brain of the Firm*. New York: John Wiley & Sons, 1972.

Beer, Stafford. *The Heart of Enterprise*. New York: John Wiley & Sons, 1979.

Beer, Stafford. *Diagnosing the System for Organizations*. Chichester: Wiley, 1985.

Cotter, T. (2012) "Engineering Management Contributions to Statistical Engineering. *Proceedings of the 2012 ASEM National Conference*. Virginia Beach.

Cotter, T. (2015) Statistical Engineering: A Causal-Stochastic Modeling Research Update. *Proceedings of the ASEM 2015 International Conference*. Indianapolis, Indiana.

Espejo, R. and Harnden, R. (1989) *The Viable System Model: Interpretations and Applications of Stafford Beer's VSM*. New York: John Wiley & Sons, 1989.

FIPS Publication 183. (1993) *Intergration Definition for Function Modeling (IDEF0)*. Washington, D.C.: National Institute of Standards and Technology, Computer Systems Laboratory.

Hoerl, Roger W. and Ron Snee. (2010) Statistical Thinking and Methods in Quality Improvement: A Look to the Future, *Quality Engineering* 22:3, pp. 119-129.

Ross, D. and Schoman, K., (1977a) Structured Analysis for Requirements Definition. *IEEE Transactions on Software Engineering*, 3(1), 6-15.

Ross, D. (1977b) Structured Analysis: A Language for Communicating Ideas. *IEEE Transactions on Software Engineering* 3(1), 16-34.

About the Author

T. Steven Cotter is a Lecturer with the Engineering Management and Systems Engineering department at Old Dominion University. He earned a Ph.D. in Engineering Management and Systems Engineering from Old Dominion University, a Master of Science in Engineering Management with a concentration in quality/reliability engineering from the University of Massachusetts at Amherst, a Master of Business Administration with a concentration in finance and a Bachelor of Science both from the University of South Carolina, and a diploma in Electronic Technology from Graff Area Vocational and Technical School (now Ozarks Technical Community College). He is a certified Quality Engineer and Reliability Engineer with the American Society for Quality. His research interests are in engineering design analytics, human-intelligence/machine-intelligence decision governance, human-intelligence/machine-intelligence quality systems design, and statistical engineering.