Paper #63

Addressing System Boundary Issues in Complex Socio-Technical Systems

CSER 2007

Joseph R. Laracy
Engineering Systems Division
Massachusetts Institute of Technology
70 Pacific St. #241 A
Cambridge, MA 02139
laracy@mit.edu

Abstract

Systems engineering researchers are familiar with a variety of challenges associated with doing foundational research in complex socio-technical systems. Some foundational issues have been avoided by focusing on applied research questions and ignoring the "socio" of the engineering system under development. Considerations of large-scale engineering systems often present a dilemma of where to draw the line between a system and its environment. How are social, political, economic, and institutional issues addressed? The lack of suitable methodologies for understanding the interface between a technical system and the human and organizational it exists within is a stumbling block. The author suggests a way ahead drawing on the ancestral disciplines of systems science. This approach led to the development of a system safety engineering methodology, System-theoretic Accident Models and Processes (STAMP), which has had significant impact on industry and the practice of safety engineering.

Introduction

Researchers and practitioners in the field of systems engineering occasionally refer to the systems they develop as *socio-technical*.

> The socio-technical concept arose in conjunction with…several projects undertaken by the Tavistock Institute in the British Coal Mining Industry. The time [1949] was that of the postwar reconstruction of industry… The second project was led, through the circumstances described below, to include the technical as well as the social system in the factors to be considered and to postulate that the relations between them should constitute a new field of inquiry. (Trist 1981)

The inclusion of human factors in the design of engineering systems was revolutionary at that time and still is today in some academic and industry circles. In 1930, MIT President Karl Compton initiated a movement to make the practice of engineering more scientific, thereby initiating the approach of *engineering science*. Engineering science – applied physics, chemistry, and mathematics – proved to be very successful in the Second World War. The development of Radar is

often cited as a product of the engineering science approach (Mindell 2004). Immediately following the war, the creation of the National Science Foundation revived the question of what it meant to do basic research in an applied field such as engineering (Kline 2000).

As systems continued to grow in size and complexity, the aerospace industry responded with what is now called systems engineering. The program for America's first ICBM, the Atlas missile, served as a test-bed for this new approach to interdisciplinary engineering system design. The Semi-Automatic Ground Environment (SAGE) air defense system, which enabled the North American Aerospace Defense Command (NORAD) to track, and if necessary coordinate a military response to Soviet strategic bombers, also made use of early systems engineering practices (Hughes 2000; Hughes 1998).

In parallel, system theorists in academia were considering many of the same concepts as industry engineers such as feedback, dynamics, flows, block diagrams, human-machine interaction, signals, simulation, and computers (Mindell 2004). However, as Kroes et al. point out, both groups encountered a serious problem:

> The field of systems engineering has inherited a conceptual problem from systems theory. Just as systems theory since its beginnings has been plagued by the question how to separate a system from its environment or context, the field of systems engineering has been confronted with a similar question about engineering systems. How are the boundaries of [engineering] systems to be drawn? What belongs to the [engineering] system under consideration and what to its environment? *For engineering systems this problem manifests itself conspicuously with regard to the status of non-technical elements, such as social, political, economic and institutional ones.* [emphasis added] To what extent are these, or ought these elements to be considered to belong to engineering systems or to the environment or context? (Kroes 2004)

Unfortunately, engineering science lacked the tools to address these fundamental questions in the new field of systems engineering.

The Boundary Problem

Catastrophic failures are associated with ignoring social, political, economic, and institutional elements. Mindell writes:

> It is highly significant that the Columbia Accident Investigation Board identified 'history and culture' as a major contributing cause of the accident. History and culture are not mysterious, inhibiting forces that act on the technological development; they are just as integral to technology as are Newton's laws and Fourier transforms. (Mindell 2004)

Another great defeat of the systems approach is associated with Robert McNamara's "Whiz Kids." "Through systems analysis, McNamara and his staff felt empowered to replace the complexity of real life with simplified models that lent illusory precision by their quantitative bases." (Jardini 1998) By dismissing many human variables and approaching the Vietnam War only as a national defense production problem, decisive factors in the outcome of the conflict, such as the fighting will of the North Vietnamese, were ignored.

Civilian problems such as housing, health care, education, poverty, and transportation were also studied with the systems analysis approach. Programs that modeled human factors and left room for compromise and negotiation were much more successful than those that left them out. For the unsuccessful programs, Mindell points out that "in retrospect, the engineers would often point to the detrimental effects

of politics, which stifled or derailed their projects. But in doing so, they pointed to the limitations of their models, which excluded politics and the social world as *external* variables." (Mindell 2004)

Clearly, the "socio" of socio-technical systems cannot be ignored. The work of Thomas Hughes is useful in considering large technical systems as a seamless web of social and technical elements where one distinguishes between physical artifacts, organizations, scientific components, legislative artifacts, and natural resources (Bijker ed. 1987). This view leads systems engineering researchers to ask the question of where to draw the boundary of the system and its environment. Furthermore, if social elements are considered, how are they to be analyzed?

One of the most conspicuous problems facing the systems engineer is the lack of formal education or on-the-job training to rigorously analyze the social forces that influence a system. An ABET accredited program does not require coursework in designing stakeholder surveys, conducting human experiments (human factors engineering), designing meaningful interviews, and other useful skills for engineering large scale, complex systems. Systems engineering researchers working on safety problems at MIT are assisted in this regard by the System Safety Working Group. The group has scholars in fields such as aerospace engineering, social psychology, computer engineering, organizational behavior, civil engineering, industrial relations, physics, and of course systems engineering. In this endeavor, a careful balancing act must be carried out. The systems engineer should acknowledge that "a systems approach is centered around the human being" and "the efficient design of systems is influenced decisively by the people who have to operate them." (Jenkins 1971) Nevertheless, he must also appreciate the field's scientific roots in dynamical systems theory, control theory, and biology (Emes 2006).

In essence, the problem comes down to methodology. How can the techniques of engineering science be connected with a modern understanding of human decision making, organizational behavior, and institutional inertia?

A Way Ahead – The Ancestral Disciplines

The good news is that many people have made significant progress at answering this question. The ancestral disciplines of systems science have much to offer 21st century systems engineers. Unfortunately, the term "systems thinking" has been so abused and misused that it has been reduced in many circles to a consulting buzzword. However, true systems thinkers — or those that take a systems approach — should expose themselves to the richness of:

1. General System Theory
2. Cybernetics
3. System Dynamics
4. Complex Adaptive Systems
5. Control Theory.

The ancestral disciplines are useful in two ways:

1. Scholars in the respective fields have confronted the human-machine problem directly and quite successfully.
2. New theories of socio-technical systems can be developed by creatively integrating the techniques of the ancestral fields.

In his *General System Theory*, Von Bertalanffy presents the concept of an open system: "An open system is defined as a system in exchange of matter with its environment…"(von Bertalanffy 1969) The concept of an open system is an important one in applied science, because often the

pure sciences (i.e. chemistry) make an assumption of a closed system – one isolated from its environment. This assumption has been implicitly imported into engineering design mental models. However, the large scale, complex systems that are the concern of the systems engineer are inherently open.

Often, engineers draw the boundary between system and environment in their models when they encounter variables that they cannot control. However, abstracting away variables that are beyond one's control does not mean they are being handled correctly. Cybernetics offers many insights into modeling human-machine systems (Ashby 1956; Wiener 1965). Cybernetic systems are inherently purposeful, goal-directed systems. The most fundamental model of control is shown below in Figure 1.

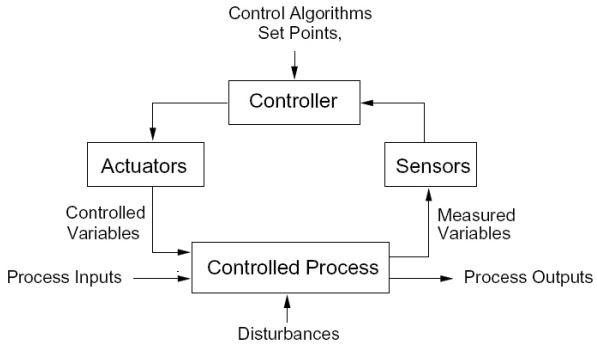

Figure 1. A Feedback Control System

Image Source (Leveson 2002)

Perturbations to the controlled process change the process in such a way that the sensors report the change to the controller which issues orders to the actuator to move the system toward the goal condition. While this model may seem trivial, it is useful to look deeper and realize that the entire model can be inverted. The environment has its own goals. The "external" disturbances of the environment attempt to impose its own set points for the process. In a symmetric scenario, such a process will never reach a stable equilibrium (Heylighen 2001). Through the IEEE Systems, Man, and Cybernetics Society as well as through some European faculties, cybernetics research continues to this day, albeit not nearly as pervasively as its founders would have hoped. With the closing of Heinz von Foerster's Biological Computer Laboratory at the University of Illinois, and other similar cybernetic research communities, the field deliquesced into computer science, decision and control engineering, artificial intelligence, robotics, and bioengineering (Hutchinson 2006).

Jay Forrester's System Dynamics (Forrester 1961) builds on the ideas of General Systems Theory and Cybernetics. von Bertalanffy's notion that complex systems can be modeled by systems of nonlinear differential equations and Wiener's notions of feedback and control are central themes of System Dynamics modeling. System Dynamics addresses concepts such as dynamic complexity, bounded rationality, flawed mental models, policy analysis, nonlinear (unintuitive) behavior, causal loops, delays, stocks and flows, and many concepts relevant to socio-technical system modeling (Sterman 2000).

System Dynamics does not distinguish between "hard" and "soft" variables as is the case with traditional engineering models. For example, a system safety engineer can develop a technical model of the physical system (i.e. a nuclear power plant) as well as the supporting human and organizational factors. The model shown in Figure 2, developed by Dulac and Leveson, captures important dynamic phenomenon such as "pushing the limits," "doing more with less," "delays cause pressure," and other feedback loops encountered in real world complex systems.

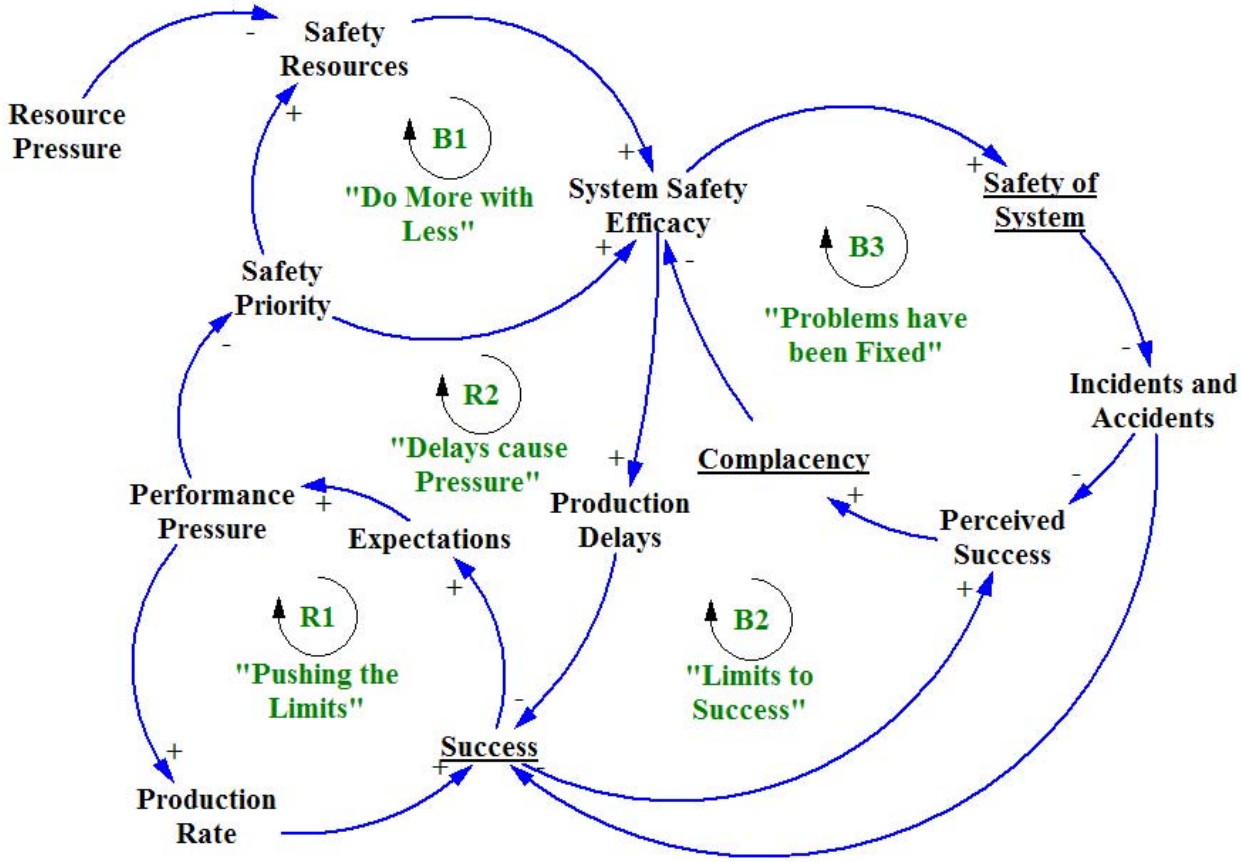

Figure 2. High Level Abstraction of a System Dynamics Model for Safety in Operations

Image Source (Dulac 2005)

Another ancestral discipline relevant to this discussion is the area of Complex Adaptive Systems (CAS). CAS such as the human brain, ecological systems, artificial neural networks, and some parallel distributed computing systems are characterized by the emergence of complex behaviors "as a result of often nonlinear spatio-temporal interactions among a large number of component systems at different levels of organization." (Chan 2001) Attributes of CAS include a reliance on distributed control, sensitivity to interconnectivity of components, co-evolution of the system with its environment, sensitivity to initial conditions in the case of mathematical chaos, and avoidance of equilibrium conditions. Engineering systems that exhibit properties of CAS cannot be separated from their environment. Chan states:

> CAS are dynamic systems able to adapt *in* and evolve *with* a changing environment. It is important to realize that there is no separation between a system and its environment in the idea that a system always *adapts to* a changing environment. Rather, the concept to be examined is that of a system *closely linked with* all other related systems making up an ecosystem. Within such a context, change needs to be seen in terms of *co-evolution with* all other related systems, rather than an *adaptation to* a separate and distinct environment. (Chan 2001)

Therefore, it is important for systems engineers to identify whether their system may exhibit CAS properties, and if so, ensure that their models acknowledge the intimate connection between the engineered system and environment. Agent-based modeling has been shown to be a valuable technique for understanding complex adaptive systems (Krenzke 2006).

Finally, control theory must be re-examined for its applicability to socio-technical systems. While many engineers have taken courses in this area and some have developed specialization in it, engineers tend to assume that the central ideas are limited to purely electrical and mechanical systems. Notions of feedback, stability, controllability, observability, and robustness can be applied creatively to improve the design and analysis of socio-technical systems.

System theorists generally acknowledge three types of structural organization. *Organized simplicity* is exhibited in traditional deterministic systems that can easily be decomposed into subsystems and components such as in structural mechanics. Systems that exhibit *unorganized complexity* on the other hand cannot be decomposed into parts. However, statistical techniques are applicable because of the regularity and randomness that characterize the system. The Law of Large Numbers becomes applicable and average values can be computed such as in statistical mechanics. The "new" complexity, *organized complexity*, describes systems that are too complex to be modeled with analytic reduction but not random enough to be modeled using statistics (Owens 2006). Figure 3 shows the relationship between the three types of organization.

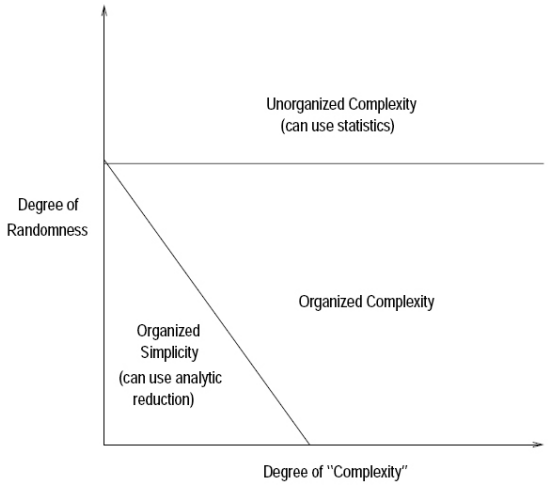

Figure 3. System Organization and Complexity.

Image Source (Weinberg 1975)

Systems characterized by organized complexity often exhibit strong, non-linear interactions and coupling between subsystems and components. Therefore, these systems must be studied holistically. Two underlying concepts provide insight into these complex systems: emergence & hierarchy and communication & control.

Abstractions for complex systems often involve layers. In the case where hierarchy exists, the level of organization increases as one moves toward higher layers. Additionally, the step from level n to $n + 1$ yields new properties that are not discernable at level n. This phenomenon is referred to as emergence, or emergent properties (Leveson 2002). As the next section will illustrate, reliability techniques that are effective for systems exhibiting organized simplicity are not necessarily applicable to systems exhibiting organized complexity.

System-Theoretic Accident Models and Processes (STAMP)

Traditional models of accident causation are rooted in a chain-of-events perspective.

Whether part of a preliminary hazard analysis or an accident reconstruction activity, the engineer attempts to understand the potential or actual accident by identifying the events or faults that could initiate the accident. Such fault and event trees are usually part of a method called probabilistic risk assessment (PRA). The goals of PRA are to estimate both the likelihood and severity of a risk. PRA was developed in the mid 1970s to improve nuclear power plant safety. Professor Norm Rasmussen of MIT chaired the Reactor Safety Study that was the first real probabilistic risk assessment (Apostolakis 2000).

A probabilistic risk assessment is a four step process:

1. Identify undesirable events.
2. Identify accident scenarios (sequences of events).
3. Estimate the probability of each scenario either based on statistical testing data, or expert judgment if scenarios are rare.
4. Rank the accident scenarios according to likelihood.

The framework yields a probability for each undesirable event identified in stage 1.

PRA turned out to be very successful for assessing risks in nuclear power shut-down systems. Such systems were historically very simple, electro-mechanical systems designed to minimize unnecessary complexity and used proven analog electrical technologies. PRA carries with it a number of important assumptions:

1. The events or faults in the trees are collectively exhaustive — all possible events are identified.
2. The events or faults in the trees are mutually exclusive — they cannot occur simultaneously.
3. The probability of each scenario is accurate enough to be useful to decision makers.

In the reactor shut-down system, nuclear engineers with decades of experience can probably develop trees that satisfy the first two assumptions due to their intimate knowledge of reactor design and operation. Furthermore, component technologies such as electrical relays could be extensively tested in the laboratory to compute reliability metrics such as mean time between failures (MTBF).

However, when complex systems like the Space Shuttle are considered, serious questions arise regarding the appropriateness of PRA. For instance, how does software change the picture? How can the MTBF of unique digital electronics be estimated? How many events or faults must be accounted for? Herein lies the problem of applying PRA to software-intensive systems. Software does not wear out and fail; it only implements a set of requirements that may or may not be correct. Subjective probability (expert judgment) must be used when thousands of laboratory MTBF tests cannot be carried out. If a spacecraft computer has 128 MB of memory, or 2^{30} bits, then it has $2^{\text{number of bits}}$ or $2^{2\text{^}30}$ states. Clearly, each state cannot be analyzed.

Before the Space Shuttle Challenger disaster, NASA headquarters reported the probability of a failure with loss of vehicle and human life as 10^{-5} (Feynman 1986). Before the Space Shuttle Columbia disaster, the reported probability was 1/250 (Stamatelatos 2002). According to NASA space operations spokesman, Allard Beutel, the post-Columbia figure is now 1/100 (Scottberg 2006).

Formal methods have also been proposed as a solution to the software safety problem. However, the complexity of formal specifications can quickly become unmanageable in large systems. In fact, it is possible for a formal specification to be longer and more error prone than the source code it specifies (Leveson 2002).

Additionally, a graduate degree in applied mathematics (formal logic) is required to rigorously apply formal methods.

A new model of accident causation is needed that recognizes the influence of software in the dynamic nature of accidents as well as the human and organizational factors. According to Leveson, "The hypothesis underlying the new model, called STAMP, is that *systems theory is a useful way to analyze accidents*, particularly system accidents." [emphasis added] (Leveson 2004) Component failures associated with hardware reliability engineering are not the only causes of accidents. Accidents often occur in complex systems when external disturbances or dysfunctional interactions among system components are not adequately handled by the *control system.* Inadequate control of safety constraints on system development and operation is the fundamental problem. "Safety then can be viewed as a *control problem*, and safety is managed by a *control structure* embedded in an *adaptive socio-technical system*." [emphasis added] (Leveson 2004) As shown in Figure 4, STAMP utilizes ideas from the ancestral systems science disciplines as well as traditional systems engineering.

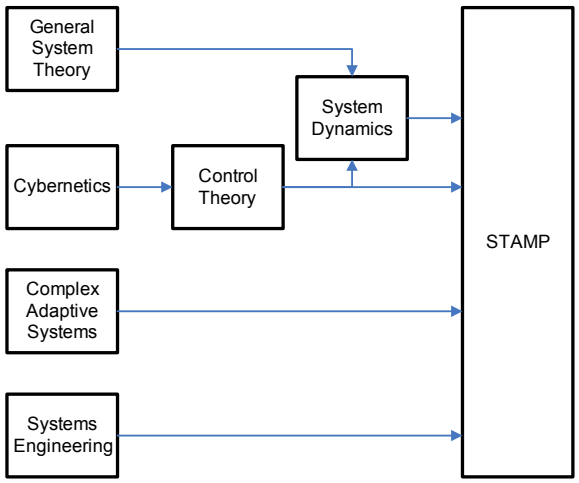

Figure 4. Ancestral Roots of STAMP

A STamP-based Analysis, or STPA, has 5 steps.

1. Identify the system hazards.
2. Identify system-level safety constraints.
3. Define the control structure.
4. Identify instances of inadequate control that could lead to a hazard.
5. Model the behavioral dynamics of the system with System Dynamics.

An example of system-level hazards for an air traffic control system is given in (Leveson 2002):

1. Controlled aircraft violate minimum separation standards (NMAC).
2. An airborne controlled aircraft enters an unsafe atmospheric region.
3. A controlled airborne aircraft enters restricted airspace without authorization.
4. A controlled airborne aircraft gets too close to a fixed obstacle other than a safe point of touchdown on an assigned runway (CFIT).
5. A controlled airborne aircraft and an intruder in controlled airspace violate minimum separation.
6. A controlled aircraft operates outside its performance envelope.
7. An aircraft on the ground comes too close to moving objects or collides with stationary objects or leaves the paved area.
8. An aircraft enters a runway for which it does not have a clearance.
9. A controlled aircraft executes an extreme maneuver beyond its performance envelope.
10. Loss of aircraft control

It is important to note that this approach is "top-down" as opposed to the "bottom-up" approaches like event trees that must identify every undesired event and trace it up to the unsafe state, or hazard. Consistent with other systems engineering activities, hazards are decomposed to the point where

they can be managed. This top-down approach produces a manageable number of hazards, rather than an unmanageable number of undesirable events.

Safety constraints are simply negative requirements. For example, the constraints for hazard 3 are "a) ATC must not issue advisories that direct an aircraft into restricted airspace unless avoiding a greater hazard. b) ATC shall provide timely warnings to aircraft to prevent their incursion into restricted airspace." (Leveson 2002) System safety engineers are very familiar with writing requirements so safety constraints are a natural extension.

Utilizing the principles of control theory, a control structure is developed for the socio-technical system. Constraints are assigned to individual components in the structure, and control actions are defined to implement the constraints. The generic model of control is provided below in Figure 5.

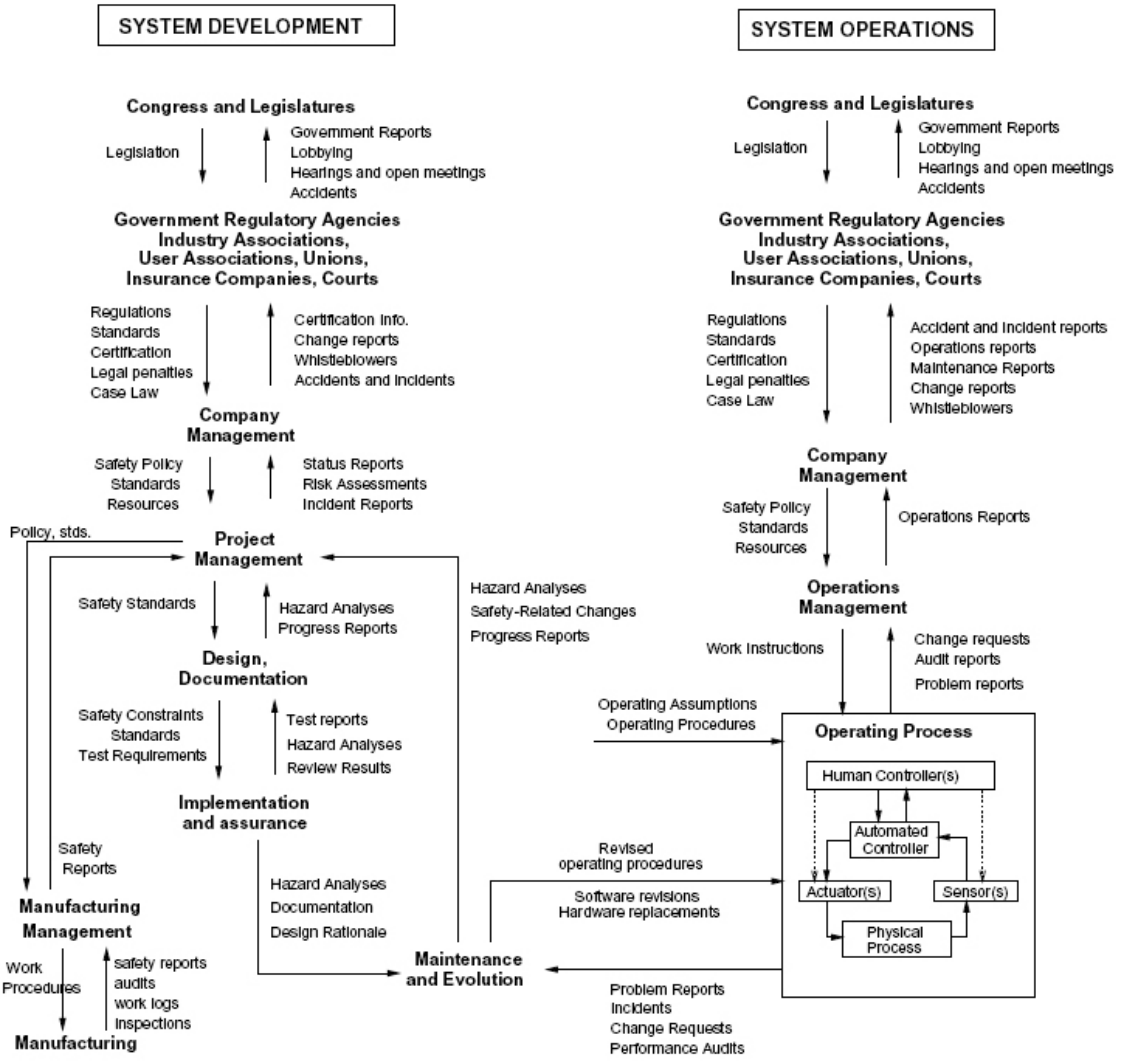

Figure 5. Generic Model of Socio-technical Control

Image Source (Leveson 2002)

Many accidents are not associated with component failure. Instead, they are the result of a slow degradation of the safety culture supporting the system and the development or operations enterprise. Systems migrate toward a state of greater risk in such a way that the evolution is not appreciated until an accident occurs. This notion of evolution is well understood with the techniques of Complex Adaptive Systems.

Identifying instances of inadequate control is a process of studying the control structure for ways that feedback, or more generally control, could be disrupted. A hierarchical taxonomy of such risks has been identified with the following three types at the highest level:

1. Inadequate Enforcement of Constraints (Control Actions)
2. Inadequate Execution of Control Action
3. Inadequate or Missing Feedback

This idea of studying feedback in socio-technical systems originates in the Cybernetics movement.

Finally, System Dynamics modeling is used to understand the behavioral dynamics of the system (Dulac 2005). Inadequate controls previously identified can be prioritized by quantitatively assessing their impact on key system safety variables. Additionally, response mechanisms can be tested, and their effectiveness judged (Laracy 2006).

Conclusion

Modeling large scale, complex systems is not an easy task. Addressing boundary issues between the technical system and the environment are particularly difficult. Often interdisciplinary expertise is needed to address the spectrum of challenges present in socio-technical systems. At MIT, the System Safety Working Group's unifying methodology, STAMP, draws from the ancestral systems sciences. By studying the ideas of the earlier systems scientists and developing new theories of socio-technical systems from them, systems engineers can hope to live up to the standards of General Bernard Schriever of the Air Force Research and Development Command. General Schriever once remarked that a systems engineering contractor should be staffed by "'unusually competent' scientists and engineers to direct the technical and management control over all elements of the program." (Hallam 2001)

Acknowledgements

I would like to thank my advisor, Professor Nancy Leveson, and the Columbia System Safety Working Group for sharing their ideas with me. I also appreciate the review and feedback received from my colleagues, Brandon Owens and Justin Colson, on this paper.

References

Apostolakis, G. (2000). "The Nuclear News Interview - Apostolakis: On PRA." Nuclear News, 27-31.

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

Bijker ed., W., Thomas P. Hughes ed., Trevor Pinch ed. (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* MIT Press, Cambridge.

Chan, S. (2001). "Complex Adaptive Systems." MIT, Cambridge.

Dulac, N., Nancy Leveson. (2005). "Risk Analysis of NASA Independent Technical Authority." MIT, Cambridge.

Emes, M., Alan Smith, Douglas Cowper. (2006). "Confronting an Identity Crisis - How to 'Brand' Systems Engineering." *Systems Engineering*, 8(2).

Feynman, R. P. (1986). "Rogers Commission Report: Appendix F - Personal observations on the reliability of the Shuttle." NASA.

Forrester, J. (1961). *Industrial Dynamics*, Productivity Press, Cambridge.

Hallam, C. R. A. (2001). "An Overview of Systems Engineering - The Art of Managing Complexity." MIT, Cambridge.

Heylighen, F., Cliff Joslyn. (2001). "Cybernetics and Second-Order

Cybernetics." Encyclopedia of Physical Science & Technology (3rd ed.), R. A. Meyers, ed., Academic Press, New York.

Hughes, A. C., T.P. Hughes. (2000). *Systems, Experts, and Computers: The Systems Approach in Management and Engineering, WWII and After*, MIT Press, Cambridge.

Hughes, T. P. (1998). *Rescuing Prometheus: Four Monumental Projects that Changed the Modern World*, Vintage Books, New York.

Hutchinson, J. (2006). "ECE Publications Manager." J. R. Laracy, ed., Email communication September 1.

Jardini, D. (1998). "Out of the Blue Yonder: The Transfer of Systems Thinking from the Pentagon to the Great Society, 1961-1965." RAND.

Jenkins, G. M., P.V. Youle. (1971). *Systems Engineering: A Unifying Approach in Industry and Society*, Watts, London.

Kline, R. R. (2000). "The Paradox of Engineering Science." *IEEE Technology and Society Magazine*, 19(3), 19-25.

Krenzke, T. (2006). "Ant Colony Optimization for Agile Motion Planning," MIT, Cambridge.

Kroes, P., Maarten Franssen, Ibo van de Poel, Maarten Ottens. (2004). "Engineering systems as hybrid, socio-technical systems." Engineering Systems Symposium, MIT.

Laracy, J. R. (2006). "A Systems Theoretic Accident Model Applied to Biodefense." *Defense and Security Analysis*, 22(3), 301-310.

Leveson, N. (2002). *System Safety Engineering: Back to the Future*, Cambridge.

Leveson, N. (2004). "A New Accident Model for Engineering Safer Systems " *Safety Science*, 42(2).

Mindell, D. A. (2004). "Historical Perspectives on Engineering Systems." Engineering Systems Symposium, MIT.

Owens, B. D., Nancy G. Leveson. (2006). "A Comparative Look at MBU Hazard Analysis Techniques." Proceedings of the 9th Annual Military and Aerospace Programmable Logic Devices International Conference (MAPLD).

Scottberg, E. (2006). "NASA Says Shuttle Risk Overstated; Yet Some Risk Unavoidable " Popular Mechanics.

Stamatelatos, M. G. (2002). "New Thrust for PRA at NASA." NASA, ed.

Sterman, J. (2000). *Business Dynamics: Systems Thinking for a Complex World*, Irwin McGraw-Hill, Boston.

Trist, E. (1981). "The Evolution of Socio-Technical Systems: A Conceptual Framework and an Action Research Program." Ontario Quality of Working Life Centre, Toronto.

von Bertalanffy, L. (1969). *General System Theory*, George Braziller, Inc., New York.

Weinberg, G. (1975). *An Introduction to General Systems Thinking*, John Wiley.

Wiener, N. (1965). *Cybernetics, Second Edition: or the Control and Communication in the Animal and the Machine* MIT Press, Cambridge.

Biography

Joseph Laracy is a PhD student in Engineering Systems and research assistant in the Complex Systems Research

Laboratory at MIT. His interests are in system safety and security. He has held engineering positions with Lucent Technologies, Ball Aerospace and Technologies, and Light Source Energy Services. Laracy is a member of the INCOSE, IEEE, and AIAA.