Information Knowledge Systems Management 10 (2011) 25–49
DOI 10.3233/IKS-2012-0184
IOS Press

Chapter 2

Understanding change in complex socio-technical systems

An exploration of causality, complexity and modeling

William B. Rouse* and Nicoleta Serban

Abstract: This paper elaborates the conceptual underpinnings needed to understand and influence change in complex socio-technical systems. The nature of causality is first addressed, followed by consideration of the nature of complexity. It is argued that, at least from a practical perspective, the difficulty in understanding causality increases as complexity increases. The possibility of influencing change is addressed in terms of concepts, principles and models for analysis and design in a range of domains or contexts.

Keywords: Causality, complexity, modeling, economics, healthcare

1. Introduction

Understanding and influencing change in complex socio-technical systems requires a basic understanding of the nature of causality in complex systems. For engineered systems (e.g., an automobile or aircraft), one can often deduce causes of system states from the design documentation of the system. However, once one includes the people and organizations associated with these engineered systems, such deduction becomes less feasible and prone to error. One then may need statistical modeling methods to address the data created by the broader socio-technical system and infer underlying relationships. At some level of complexity, deduction and induction may be supplanted by abduction, i.e., educated guessing.

The difficulty of deducing, inferring, or abducing causes of system states increases with the complexity of the system of interest. This is due to many possible transformations between causes and changes, including the possibility of independent actors in the causal paths or networks between causes and changes. The systems through which these changes evolve are often dynamic and laced with uncertainties, including uncertainty about the nature of the system itself. Thus, to address causality, we must also consider the nature of complexity.

Our interests in this paper include not only the fundamental nature of causality and complexity, but also how one can influence the nature of change in complex systems. How, for example, might one influence the course of economic changes or military engagements, hopefully to facilitate better outcomes? One approach to influencing change is to model the relationships between causes and changes as mediated by

*Corresponding author. E-mail: bill.rouse@ti.gatech.edu.

the complex systems in which the cause-change relationships arise. This can enable, in effect, feedback loops from experienced or projected changes to interventions that influence causes directly or indirectly. The third major part of this paper explores approaches to modeling complex systems for the purpose of influencing change.

At the outset, it is important to elaborate our philosophy in addressing causality, complexity, and modeling. We are decidedly pragmatic in this exploration of these fundamental topics. Thus, our concern is with the practical utility of alternative definitions, frameworks, and so on. We are less concerned with theoretical correctness, in part because this is strongly influenced by the disciplinary lens one adopts. As the other papers in this special issue illustrate, the system context of interest and issues of concern have an enormous impact on the choice of lenses for addressing these issues.

This paper begins consideration of causality by first reviewing a variety of basic arguments, many of which are drawn from other papers in this special issue. We then consider several historical illustrations from a range of domains. These illustrations provide evidence of the difficulty of crisply ascribing causality in realistically complex domains.

We next consider the nature of complexity. A variety of definitions are discussed. The role of intentions in choosing among definitions is noted. Finally, the concept of complex adaptive systems is introduced and reviewed.

The next section considers the role of models in understanding causality and complexity. Holistic and reductionist approaches to modeling are contrasted. The historical illustrations are revisited, and approaches to modeling the phenomena of interest in each domain considered.

Finally, we consider how to influence change. A structured set of heuristics are proposed, drawn from the contrasting points of view presented throughout this paper. We hasten to note, however, that our goal was not to synthesize all these perspectives into a single point of view. Instead, we intended these heuristics as simply down-to-earth practical guidance.

2. Causality

Wikipedia defines causality as the relationship between an event (the cause) and a second event (the effect), where the second event is a consequence of the first. It is often easy to observe a temporal relationship between two events. However, inferring causation between events can be rather difficult, particularly if one is seeking the cause.

Aristotle differentiated four aspects of the chain of cause:

- Material – the fact that something exists
- Formal – the intentions associated with something
- Efficient – the initiating event
- Final – the purpose or end something serves

Hoffman and Klein (2011) summarize this categorization into a set of questions: “What is X made of? What does it mean for something to be an X? What produces an X? What is X for?”

Thus, for example, an error by a human pilot cannot “cause” an accident if the airplane did not exist, or if he were not intending to land, or if he were not distracted by air traffic control, or if airplanes were not intended to land.

Mumford and Anjum (2011a) review the concept of causality from an a historical perspective, starting with the above classic notion of causality drawn from Aristotle’s philosophical writings and bringing the discussion to modern perspectives that highlight the presence of powers (or dispositions) which are mediators in influencing causality.

It is useful to differentiate necessary, sufficient, and contributory causes. Oxygen is necessary to start a fire, but heat, fuel and oxygen are the sufficient conditions for a fire. An extended lack of rain may contribute to these sufficient conditions being manifested. It is often a confluence of factors that causes the consequences of interest (Mumford and Anjum, 2011a).

There is also the possibility of questionable causes. Feedback loops in complex systems make it difficult to infer what the cause is and what the effect is. Does the thermostat cause the furnace to heat the room, or does the temperature of the room cause the thermostat to turn the furnace on to heat the room? Obviously, without the thermostat, nothing happens. However, without the temperature differential, nothing happens either.

A classic problem in inferring cause is the fundamental attribution error (Jones and Harris, 1967). The error is as follows: If one is successful, one will tend to attribute this success to intelligence and hard work; if one fails, one will attribute this to bad luck. In contrast, if someone else succeeds, one will attribute their success to good luck; if they fail, one will attribute their failure to a lack of intelligence and hard work. This archetypal exemplar from the behavioral and social sciences is but one illustration of how people's perceptions can have enormous influence.

There are clearly many ways to make errors in attempting to infer causality. This is due in part to the lens with which we view the phenomena of interest. One type of lens relates to different disciplines. Somewhat simplistically, physicists tend to think terms of fundamental forces: gravity, the strong and weak nuclear forces, and electromagnetism. In theory at least, they view these four fundamental forces as they interact to cause all other events in the universe.

Engineering tends to view causal systems as entities with outputs and internal states that depend only on the current and previous input values. Contemporary views in biology are expressed at levels ranging from molecules to cells to systems to populations, with the inherent notion of cascading relationships among levels. For history, events are often considered as if they were agents (efficient causes) that can then bring about other historical events.

In general, causal connections are seldom articulated among scientific realms and levels of description, such as physics, chemistry, biology, psychology, and sociology. Beyond differing disciplinary views, there are also differing social, cultural, and religious views that have enormous impacts on the inference of causation.

Mumford and Anjum (2011b) recognize that causation is not just a chain of effect-cause relationships in a complex world: "The world is more of an interconnected web than a series of chains. . . . Such interconnectedness may be what defines a system . . . those interconnections will be causal." They argue that causality is context based and there are multiple interconnected causes and effects. There may be one dominant cause (or more) but also assisting and intermediary causes, which describe the context of change and influence causality. Actors, transformations, worldview, ownership and environment all form the context of causality (Haslett, Barton and Stephens, 2011). The sheer presence of causes is necessary but not sufficient. Their intensity, efficacy, typicality or unusualness, and very importantly, their aggregate response will lead to different effects with different outcomes (Hoffman and Klein, 2011). That is, it is not only important to account for causes but also to value their strength.

In most realistically complex systems, where there is a large number of possible states and high density of interconnectivity among system elements, there are typically multiple possible causal paths and interactions among paths, particularly when the system is laced with behavioral and social phenomena, i.e., systems elements include people and organizations. Freeman (2000) compares the human brain with the society we have created. The brain is an integrated assembly of cells, each of which is a semiautonomous (not independent as often specified in adaptive complex systems) agent. Each cell is like a person in a society infinitely complex and unable to function or even exist for long, except in the context of the whole.

An important notion here is path dependence. The nature of a change is often affected by the series of precursor steps to this change. Similar changes emerging from different paths can have different implications due to their antecedents, e.g., agreements negotiated long before the current state. Thus, the path from initiating events through the network of relationships among system elements, and to consequent changes can have an enormous impact on how causality in interpreted. The set of alternatives available in any given circumstance may be limited by the nature of the choices one has made in the past, even though past circumstances may no longer be relevant (Arthur, 1994).

2.1. Modeling causal relationships

The most common inductive approach to modeling causal systems is to use Bayesian networks sometimes called causal networks (Darwiche, 2009; Pearl, 2000; Spirtes et al., 2000). Pearl (1985) was the first to define Bayesian networks to emphasize the use of Bayes theorem in making inferences, the subjectivity of the information entering the network, and the extent to which it defines causal reasoning.

Formally, a Bayesian network consists of a structure in the form of an acyclic graph composed of nodes (vertices) and links (edges) with a probability value assigned to each edge defining the conditional probability of the nodes corresponding to the edge. In the world of causality, edges in the graph represent direct causal influences. The parents of a node denote the direct causes of the node and the descendants of a node denote the effects.

Various topics in Bayesian network analysis have been of interest, from modeling, variable selection to inference, reasoning and prediction. Important aspects in Bayesian network analysis for complex systems include evaluating the network properties locally and sensitivity analysis.

Sloman and Fernbach (2011) highlight the downsides of causal Bayesian networks, which model acyclical and hierarchical structures. Human-made systems commonly do reveal transparent hierarchical structures but less so natural systems. Therefore, the existing causal Bayesian networks are limited in characterizing human-made systems, which could be cyclical or natural systems that often do not have a hierarchical structure. Sloman and Fernbach (2011) also mention other approaches to describe causality including simplifying heuristics and simple linear combinations, which may be more flexible in manipulating patterns of causal systems but in the same time prone to misrepresentation as they are based on one realization of the data.

The decade of social networking, research and practice focused on modeling, designing and influencing social networks has advanced many opportunities to create a worldwide social space. Social and organizational networks operate on multiple levels, from the society as whole to the family and individual. Interactions exist within and between the network levels, which makes their modeling and prediction very challenging. One example we discuss in this paper is healthcare, which is an interconnected web with many actors such as people, insurance providers, government, medical equipment vendors, medical research labs, and many more.

McCrabb and Drabble (2011) introduce a model for analysis of social networks (called Williamsburg), which takes into account loose coupling in the network, dimensionality and validation without imposing boundary conditions. The underlying property of their model is transitivity, which is a common feature in social networks. Many theories have been formulated for social networks including theories of proximity (physical or electronic), contagion theories, cognitive theories (semantic networks and knowledge structures), and social support theories among many others (Jablin and Putnam, 2000).

At the heart of these graphical approaches lies statistical modeling and inference. For example, the conditionality of events reveals links in the network (Xie, 2011). Mediation and interaction analysis (Chan, 2011) create the sequence of links in chain-like causalities.

Tsallis (2011) raises an interesting point in terms of the *ethical aspect* of understanding and influencing a system. We must ask ourselves whether we wish to do that, why, under what conditions, to what extent, with what purpose and with what probability of success.

2.2. Illustrations of causal relationships

A few historical illustrations help to portray the way in which causality often involves a confluence of factors, particularly in complex socio-technical contexts. We primarily describe chains of events from a traditional perspective. Labeling one or more of these events as "causes" is often controversial.

It is important to reflect on and understand the purposes for which one attempts to uncover causes in chains of events. Examples include:

- To predict possible unintended consequences (e.g. Mumford and Anjum's (2011a) example of affirmative action to promote gender equality in higher education in Norway leading to an overload of female faculty)
- To control for antipathetic cases (e.g. in administrating medication)
- To intervene in case of irreversible damage (e.g. emergency response before large scale hurricanes)
- To appreciate the complexity of the system (e.g. how neurons interact in response to a stimulus)

The purpose for attempting to ascertain causality determines the extent to which one also must investigate the high-level context surrounding the chain of causality as well as the factors (powers) intervening in the chain of causality (Mumford and Anjum, 2011a, b)

2.2.1. The Great Depression

The economic depression, which started in 1929 and continued into the 1930s, began with a dramatic drop of the stock market. However, while this drop in the value of shares provided the backdrop for what followed, it had limited direct effect because less than 2% of Americans owned shares in companies listed on the stock market. The precipitating event has been claimed to be a rumor spread by a small merchant, a holder of stock in the Bank of United States (Gray, 2001). He claimed that the bank had refused to buy back his stock in the bank. Subsequently, a large crowd gathered at the bank's branch in the Bronx on December 10, 1930 seeking to withdraw their money. The bank was not prepared for this first bank run of the Great Depression and failed the next day. The public panicked and the American banking system shuddered. People flocked to withdraw their money from other banks. In turn, the banks called in loans and sold assets in order to stay liquid. In that month alone, over 300 banks around the country failed.

So what "caused" these consequences? In terms of Aristotle's characterization of causes (i.e., material, formal, efficient, and final), the material cause was under-capitalized banks. The formal cause was depositors' demands for funds. The efficient cause was the initiating event – the rumor and run on the first bank. The final cause was the banks' intended purpose to protect depositors' assets. Thus, one can argue that the rumor was only the last element in a chain of causation. Without the other elements of causation, the rumor would have been meaningless, would not in fact have happened.

The complexity of the banking system is due to how banks in the U.S. operate. (Of course, this in not necessarily how banks have to operate.) Depositors' assets are loaned or invested to provide the expected returns plus bank profits. Consequently, banks retain in reserve much less than the total value of depositors' assets. Loans and investments are used to enable businesses to earn profits and pay loans. However, demand for return of deposits results in calling of loans of businesses. As a result, businesses curtail operations leading to decreased revenue and employment. Decreased employment leads to decreased consumption and increased withdrawals of savings. Decreased consumption and increased withdrawals exacerbate banks' and businesses' problems, as well as unemployment. The result is an economic depression.

2.2.2. Housing collapse

U.S. government policy prompted low interest rates and programs to encourage home ownership that led to growth of demand. The increasing demand, including that due to speculation, resulted in seemingly unending higher house prices and homeowner equity. This enabled people to monetize increasing home equity to finance consumption, while also decreasing savings. This consumer spending drove economic growth. However, adjustable rate mortgages, that were used to attract low-income buyers, later resulted in increased mortgage interest and increased consumer defaults. Consequently, housing prices plummeted, leading to deterioration of banks' and other institutions' investment portfolios.

This scenario has a well-documented series of precursors. In 1971, President Richard Nixon addressed the trade deficit crisis by making the U.S. dollar a floating currency. The next year, trading in currency futures was launched on Chicago Mercantile Exchange. The Chicago Board Options Exchange formed in 1973 in the wake of explosive growth in use of option contracts. A seemingly unrelated event, the Community Reinvestment Act of 1977 resulted in more lending to poorer people and steadily rising house prices (Economist, 2008).

By 1986, cross-border capital controls had all but disappeared. The Glass-Steagall Act was repealed in 1999, allowing commercial banks into investment banking. Asian growth and savings by 2000 made external deficits easy to finance. Exponential growth of Credit Default Swaps had, by 2001, separated market risks from default risks. By 2008, the financial industry accounted for 23.5% of stock market value, versus 5.2% in 1980; and 41% of profits, versus 16% in 1980 (Economist, 2008).

Alan Greenspan (2007), then Chair of the U.S. Federal Reserve Bank, applauded these developments, "Being able to profit from the loan transaction but transfer credit risk is a boon to banks and other financial intermediaries which, in order to make an adequate rate of return on equity, have to heavily leverage their balance sheets by accepting deposit obligations and/or incurring debt. A market vehicle for transferring risk away from these highly leveraged loan originators can be critical for economic stability, especially in a global environment."

In summer of 2007, Ben Bernanke encouraged the financial market regulators to pay greater attention to principles than rules. Indeed, the real estate market collapse in 2008 was in part due to the ignorance of many banks and mortgage companies, which found that they were exposed to subprime lending because of the adherence to the rules but not to the principles of lending.

The complexity of the financial system involved a variety of factors. Risk-return tradeoffs had been transformed to return maximization independent of risk. Third-party risk management completely ignored conditional probabilities, i.e., correlated risks of "common mode" failures. The country's economic dependence on financial institutions required government intervention and support when they were at risk of failure. This support fixed these institutions' balance sheets but did not stimulate loans for business formation and growth. Economic growth had depended on consumer spending, which encouraged consumption and debt, both of which were no longer sustainable.

Was it greed or stupidity? Lowenstein (2001), the Economist (2009), and Salomon (2009) provide insightful explanations. Put simply, people became enamored with sophisticated mathematical models

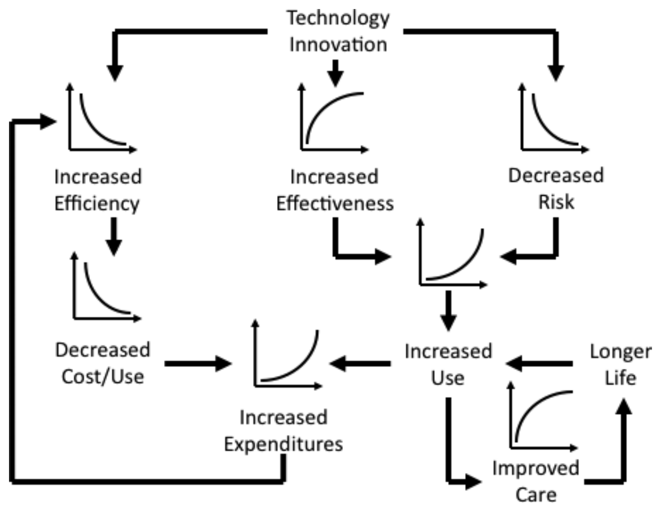


Fig. 1. The dynamics of escalating healthcare costs.

that were argued to enable optimal investment decisions. These models were powerful and valuable, but the actual decisions of interest, over time, transgressed the assumptions underlying the models, making model-based decisions of questionable validity. However, as a senior banking executive told one of the authors, "We were so busy, making so much money that is was socially unacceptable to critique the assumptions underlying the models."

Lissack (2011) points to the limitations of strict compliance with and submission to models at a higher management level without clear vision, creativity and innovation. "The assumption that models correspond to reality . . . is the curse of much modern management theory and practice." Creativity and innovation are critical as they allow one to be opened to the "possibility of self-organization and emergent behaviors, which are a source of unpredictability in the real world."

2.2.3. U.S. healthcare costs

Healthcare costs in the U.S. are the highest in the world, 50% to 100% greater than other developed countries, depending on the way costs are characterized. The quality of health provided is, however, rather poor, with the U.S. ranked between 30 and 40 relative to other developed countries. The system is sorely in need of transformation (Rouse and Cortese, 2010).

How did this happen? The chain of causation is quite clear. Advances in public health and medicine led to increased life expectancy. Consequently, chronic care needs came to dominate, relative to acute care needs. At the same time, and seemingly unrelated, U.S. farm subsidiaries for corn lead to low cost, low quality, high calorie junk food. The result has been an epidemic of obesity and diabetes, as well as other chronic diseases, especially among lower income populations.

In parallel, and again seemingly unrelated, technological innovation has led to new screening and treatments for chronic diseases. As shown in Fig. 1, the increased demand for these procedures has completely outstripped any technology learning curve efficiencies. The cost reimbursement system encourages increased use of procedures and does not pay for treatment outside facilities. Overall, there are limited market control mechanisms to impede these trends due to third-party payments for services (Rouse, 2009).

The complexity of this causal chain is influenced by a variety of factors. First and foremost, the complex adaptive system of healthcare amounts to a federation of millions of entrepreneurs with no on one in charge (Rouse, 2008). Major stakeholders have made enormous investments premised on the existing business model. There are limited market mechanisms for change with political processes for change effectively stalemated. Succinctly, changes of information systems (in terms of standardization and availability) and incentives (in terms of payment for services vs. outcomes) are the central issues; the former progressing very slowly; the latter stymied by all of the above. Aarts (2011) discusses health IT implementation and its use in healthcare along with the advantages and disadvantages of having a centralized health IT system.

2.2.4. Pickett's Charge

The battles of the American Civil War have received much attention. Perhaps no battle has received more attention than the Battle of Gettysburg. It led to Abraham Lincoln's immortal "Gettysburg Address" and, in conjunction with the Battle of Vicksburg, was the turning point in the war. After Gettysburg and Vicksburg, it was clearly only a matter of time before the Union forces prevailed over the Confederacy.

The basic chronology is well documented. The South was under-resourced, in terms of people and industry, and poorly governed due to State Rights, limiting coordination of forces and resources. The South's goal, by 1863 at least, was to secure peace and recognition of their independence – there was no expectation of victory over the North.

The South invaded the North in Pennsylvania. Morale among Northern troops was assumed to be weak; the South also urgently needed supplies. The Battle of Gettysburg was stalemated at Cemetery Ridge and Gen. Robert E. Lee wanted a "final battle" to end the war. Consequently, Lee ordered Gen. James Longstreet to attack the ridge using the Divisions of Generals George Pickett, Johnston Pettigrew, and Issac Trimble.

Not believing in the potential efficacy of the attack, Longstreet delayed, muting the effect of the South's bombardment of the ridge and allowing the North to muster reinforcements. Pickett's Charge resulted in the decimation of the South's forces, with causalities exceeding 50% of the 15,000 attacking forces. The loss at Gettysburg, as well as the surrender at Vicksburg the following day, was the turning point of the Civil War.

We can consider the causes of this result using Aristotle's framework. The material cause was the under-resourced and poorly governed South. The formal cause was Lee's desire for the "final battle" despite his inferior position. The efficient cause, the initiating event, was Lee's order and Longstreet's equivocation. The final cause was the South's invasion of North to sue for peace and independence.

Considering the complexity associated with this socio-technical system, a dominant factor was the South's inabilities to cover eastern (Gettysburg) and western (Vicksburg) fronts with limited resources. Despite this limitation, the South decided to act decisively to force the end of war against arguably superior forces. More tactically, the South had poor information on the magnitude, state, and intentions of opposing forces. Consequently, assumptions replaced fact-based assessments due to the lack of information. Overarching all of this, there were numerous quasi-independent agents leading both North and South forces.

2.3. Summary

There is a wide variety of views on causality, depending to a great extent on the disciplinary lens adopted. For realistically complex socio-technical systems, inferring the elements of causal networks can be quite difficult. In most situations, there are multiple factors that combine to precipitate the event of interest – the event for which one seeks the cause. This inference process is very much complicated by the complexity of the system of interest. This leads us to explore the nature of complexity.

3. Complexity

3.1. Definitions

There are many definitions of complexity (Rouse, 2007). The U.S. National Institute of Standards and Technology defines complexity as the intrinsic amount of resources, for instance, memory, time, messages, etc., needed to solve a problem or execute an algorithm (NIST, 2004). Similarly, Gell-Mann (1995) defines complexity in terms of how long it would take, or how much capacity would be required, at a minimum, for a standard universal computer to perform a particular task. From a different perspective, Gell-Mann (1995) defines complexity as the length of a concise description of a set of an entity's regularities. Carlson and Doyle (1995) argue that the essence of complexity is the elaboration of highly structured communication, computing, and control networks that also create barriers to cascading failure events. This definition focuses on the functioning of the entity.

Abbass and Petraki (2011) define complexity in the context of systems in which the topological structure abstracted through the connectivity of the interacting components is paramount, purely emphasizing the complicated structure of system ignoring its underlying features. Namatame and Komatsu (2011) relate complexity to robustness of a system and the risks associated with lack of robustness. Their view on complexity is one of a multi-level network.

Tsallis (2011) reviews various entropy-based measures of complexity from statistical mechanics. Using these measures, the author points to the difference between simple systems and mathematically complicated and complex systems. For example, the classical Boltzmann-Gibbs entropy is used to measure the entropy or energy in a simple system whereas q-entropy, for example, will be used to describe the energy in a complex system. He provides a series of characteristics of simple systems (short-range space-time correlations, Markovian processes, additive noise, etc.) and of complex systems (long-range space-time correlations, long-memory, etc.). All these features describe the degree of complexity of the system. The use of BG entropy (q-entropy with $q = 1$) or q-entropy ($q > 1$) is dictated by the system's microscopic probabilistic-dynamic nature. In some sense, q will measure the degree of complexity.

Thus, this selected set of definitions provide a contrast among: 1) how long it takes an entity to accomplish something, 2) how an entity functions (e.g., its structure, perhaps probabilistically) and, 3) a description of the entity. Obviously, there are multiple points of view on the nature of complexity.

Many commentators on the subject argue that the best way to define complexity is to focus on the characteristics of complex systems. For example, various online sources define complex systems as having a large number of mutually interacting dynamical parts that are coupled in a nonlinear fashion and involve feedback loops. For example, cells are organized structures of macromolecules, themselves built out of biochemical structures that have atomic constituents, which themselves have several layers of subatomic organization. Many of these definitions also indicate that complex systems tend to have indeterminate boundaries (i.e. are open), have a memory, may be nested, and often include humans and organizations.

Rouse (2003) defines a system as a group or combination of interrelated, interdependent, or interacting elements that form a collective entity. Elements may include physical, behavioral, or symbolic entities. Elements may interact physically, mathematically, and/or by exchange of information. Systems tend to have purposes, although in some cases the observer ascribes such purposes.

He then defines complex system as entities whose perceived complicated behaviors can be attributed to one or more of the following characteristics: large numbers of elements, large numbers of relationships among elements, nonlinear and discontinuous relationships, and uncertain characteristics of elements and relationships. Complexity is perceived, rather than absolute, because apparent complexity can decrease with availability of information and learning. One need only observe an expert auto mechanic or heart surgeon to witness this phenomenon.

This is not to say that complexity can inherently be decreased by information and learning. For example, we certainly have a much more informed and sophisticated view of cancer now then in past decades (Mukherjee, 2010). Interestingly, now that we better understand cancer, we see the disease as much more complex than we had previously. Thus, the causal paths are more complicated, but we now see the possibility of delineating these paths and better understanding the causality of the various versions of this disease.

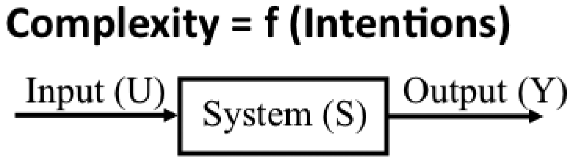
Rouse (2005, 2007) argues that system complexity tends to increase with the number of elements and number of relationships among elements. Complexity is also affected by the nature of relationships including logical (e.g., AND vs. OR and NAND), functional (e.g., linear vs. nonlinear), spatial (e.g., lumped vs. distributed), structural (e.g., feedforward vs. feedback), response (e.g., static vs. dynamic), time constant (e.g., not too fast vs. very slow), and uncertainty (e.g., known vs. unknown properties). Complexity tends to be greatly increased to the extent that the elements and relationships are human and social, respectively. Finally, the effects of these attributes on complexity depend on the knowledge, experience and skills, as well as the intent, of the humans for whom complexity is of interest.

We hasten to note that we are not arguing that behavioral and social phenomena can all be captured by the nature of relationships outlined above. Indeed, the fact that humans and organizations have interests and intentions beyond those of the system of interest, can be the essence of why causality is so difficult to attribute to observed changes. The earlier historical illustrations provide evidence of this.

Carlson and Doyle (2002) contrast the complexity of optimized and self-organized systems. Optimized systems are designed and, consequently, are heterogeneous in that they are organized into intricate and highly structured networks, with hierarchies and multiple scales. Self-organized systems are not, initially at least, "designed" and, hence, are not intentionally structured. Of course, social groups, for example, may eventually organize in terms of a hierarchy of component units, e.g., into a national function and component divisions, each of which includes several local units. The key point is that the nature of complexity of engineered systems could be rather different from the complexity of organizational and especially natural systems.

Csete and Doyle (2002) contrast the complexity of a Boeing 777 with 150,000 different subsystems structured in a complex control system consisting of about 1000 computers that can automate the vehicle to that of biological systems, probably the closest would be the fruit fly. They point out that not the similarities in comparing an aircraft and a fruit fly but the differences are relevant to their intrinsic complexity. The *E. coli* chemotaxis system has some remarkable design features, from the modularity and dynamics of the sensors and signal transduction network, to the construction and action of the flagella, to the stochastic search algorithm implemented by run and tumble to overcome the inherent limitations in gradient sensing at such small scales. Nothing like this system exists in engineering. (Csete and Doyle, 2002 supplemental material).

In the absence of centralized control, a system exhibits patterns of order despite the simple, sometimes even chaotic, local interactions (flows and dynamics) (Abbass and Petraki, 2011). Kelso (1995) defines self-organization in the context of the manifestation of chaos, which looks like noise but has hidden order and the capacity for rapid and widespread changes (just like our thoughts). Chaos is just like a random walk, you may think that it is predictable for a moment but it turns into a different other direction in the next moment.



Intention	Example
Classification	"It's an instance of type S."
Explanation	"It's type S because ...."
Prediction	"It's future output will be Y."
Control	"If input is U, it's output will be Y."
Detection	"It's output is not Y, but should be."
Diagnosis	"It's output is not Y because ..."

Fig. 2. Complexity as a function of intentions.

Other widely cited characteristics associated with complexity and complex systems are coordination or synchronized interactions (Kelso, 1995), multi-functionality (Freeman, 2000), stability (Wang and Chen, 2001), emergence (Abbass and Petraki, 2011; Balazs and Epstein, 2009), and robustness (Haslett, Barton and Stephens, 2011).

3.2. *Complexity and intentions*

The U.S. National Science Foundation sponsored a series of workshops on complex engineered, organizational and natural systems. These workshops involved 250 experts on complexity from a wide range of disciplines. A variety of conclusions were reached, many of which are discussed elsewhere in this paper (Rouse, 2007). One of the more controversial conclusions was that the complexity of a system depends on the intentions prompting the interest in assessing the complexity of the system. Over 90% of the participants in this series of workshops agreed with the conclusion that system complexity relates to the interaction of observer and system. Thus, the notion of absolute complexity is of little use.

For example, if one purchases a B747 aircraft to use as a paperweight, then it is just a large mass and not particularly complex. On the other hand, if one intends to fly and maintain this aircraft, it is a fairly complex system. Further, it is more complex if one does not have knowledge, experience and skill in aircraft operations and maintenance.

Molecules are too macroscopic for a particle scientists to study, and too minuscule (microscopic) for a physiologist. In economic networks, "micro" refers to individual system elements (i.e. the borrower and the bank) whereas "macro" refers to the financial system as a whole. Clearly, there is a significant disconnect between the two as the intentions from the macro to micro economics are not aligned because of the conflict between individual incentives and aggregated welfare, or their impact on the overall efficiency of the network at large.

Ottino (2004) reflects this philosophy, "Complex systems can be identified by what they do (display organization without a central organizing authority – emergence), and also by how they may or may not be analyzed (as decomposing the system and analyzing sub-parts do not necessarily give a clue to the behavior of the whole)."

Figure 2 illustrates the conclusion that complexity is a function of intentions. Classifying a system does not require addressing the system in much depth. Is the object an automobile or a plane? Explanation requires dealing with greater complexity. At the extreme, diagnosing why the plane will not fly correctly involves a significant increase in the complexity that must be addressed. The point is that the aspects of the system that one takes into account depend on the intent for which complexity is of interest. Being a passenger in an airplane is much less complex than flying and maintaining the airplane, because one does not need to be concerned with the nature of the relationships within the propulsion system, for instance, in order to be able to use the drop-down tray and put one's seat back.

This conclusion explains why different views of complex systems serve different purposes (Rouse, 2003). For example, one might view the airplane in terms of hierarchical mappings from system, to subsystem, to components, etc. This view reflects a notion of complexity typically being due to large numbers of interacting elements. Another view might be uncertain state equations where complexity is due to large numbers of state variables and significant levels of uncertainty. Yet another view is discontinuous, nonlinear mechanisms where complexity is due to departures from expectations of continuous, linear phenomena. One might also view the system in terms of autonomous agents where complexity is due to reactions of agents to each other's behaviors and resulting emergent phenomena.

Which view is correct? It all depends on one's intentions. Assume that the interest is in the effects of turbulent flow on aerodynamic behavior and vehicle performance in high-density traffic. One could employ hierarchical mapping for designing the vehicle; the uncertain state equations view for exploring vehicle dynamics; the discontinuous, nonlinear mechanisms view to model turbulence; and the autonomous agents view to understand traffic effects. Thus, the most appropriate view of complexity depends on the problem of interest, e.g., poor vehicle handling qualities vs. traffic congestion problems.

3.3. Complex adaptive systems

One class of complex systems includes engineered systems such as automobiles, aircraft, factories, power plants, and so on. The system elements and relationships among system elements are known because they were designed or engineered. Once these types of systems become part of a broader, socio-technical context, the nature of complexity changes. These broader systems are often referred to as complex adaptive systems (Rouse, 2000, 2008).

These types of systems are nonlinear and dynamic and do not inherently reach fixed equilibrium points. The resulting system behaviors may appear to be random or chaotic. Complex adaptive systems are composed of independent (or conditionally independent) agents whose behavior can be described as based on physical, psychological, or social rules, rather than being completely dictated by the dynamics of the system.

Agents' needs or desires, reflected in their rules, are not homogeneous and, therefore, their goals and behaviors are likely to conflict – these conflicts or competitions tend to lead agents to adapt to each other's behaviors. Agents are intelligent and learn as they experiment and gain experience, and change behaviors accordingly. Thus, overall system structure and behavior inherently changes over time.

Adaptation and learning tends to result in self-organizing and behavioral patterns that emerge rather than being designed into the system. The nature of such emergent behaviors may range from valuable innovations to unfortunate accidents. There is no single point(s) of control – systems behaviors are often unpredictable and uncontrollable, and no one is "in charge". Even if one knew all the facts, one still could not explain the working of a certain higher-level phenomenon in terms of lower-level phenomena of a system. Consequently, the behaviors of complex adaptive systems usually can be influenced more than they can be controlled.

Most, perhaps all, large public-private systems are complex adaptive systems. They are difficult to understand and generally impossible to control. However, understanding the agents or stakeholders in the system in terms of goals, objectives, strategies, tactics and plans can enable designing incentives and inhibitions that may lead to tipping points (Gladwell, 2002) or the notion of thresholds (Mumford and Anjum, 2011a) whereby agents embrace small changes, learn how to adapt and self-organize to take advantage of these changes, and consequently escalate adoption across the system. Models can provide the insights needed to accomplish these ends.

4. Modeling

There is a variety of approaches to modeling complex systems. Somewhat simplistically, two broad classes of approaches are holistic and reductionist. The holistic approach considers the characteristics and functioning of the overall system with little if any decomposition. In contrast, the reductionist approach, discussed later in this section, attempts to decompose a system into its structural elements to understand how these elements function together to yield the behavior of the system.

4.1. Holistic approaches

Large-scale public-private systems provide interesting opportunities to elaborate holistic views of complex systems. Such systems involve numerous private enterprises operating in a marketplace that is heavily influenced by government policy and, in some cases, by government funding. Examples include:

– *Defense*: Many interdependent private enterprises with integrated delivery of products and systems for public use, one source of payment, and substantial and integrated public oversight
– *Education*: A large number of independent, mostly public, enterprises with distributed delivery of products and services, many from private enterprises, as well as distributed payment and distributed public oversight
– *Finance*: Many interdependent private enterprises with integrated delivery of shared services, but distributed delivery of products and services to consumers, and distributed payment with integrated public oversight
– *Food*: A large number of independent private enterprises with integrated delivery systems, but distributed products and payment, with integrated public oversight of products, but less so services
– *Healthcare*: A very large number of independent private and public enterprises with distributed delivery of products and services but, for older and poor consumers, one source of payment, and integrated public oversight of products, but less so services

Note that *Defense* is a complex private sector product delivery system, embedded in a complex public sector service delivery system. *Education* and *Healthcare*, in contrast, are primarily complex service delivery systems, with both private and public sector service providers. *Finance* and *Food* predominantly involve private sector product and service delivery, with public sector oversight, albeit quite intense of late for *Finance*.

Table 1 summarizes the holistic characteristics of public-private enterprise discussed above. *Defense* is the most integrated while *Education* is the least integrated of these enterprises. Note that while oversight is integrated for *Finance*, *Food*, and *Healthcare*, the level does not approach that of *Defense*.

In order to assess and contrast the complexity of these five domains, consider the following notional model of complexity C.

$$C = f(NE, DI, PSI, PI, OI) \qquad (1)$$

Table 1
Characteristics of public-private enterprises

	No. of enterprises	Delivery	Products/services	Payment	Oversight
Defense	1,000	Integrated	Integrated	Integrated	Integrated
Education	100,000	Distributed	Distributed	Distributed	Distributed
Finance	10,000	Integrated	Distributed	Distributed	Integrated
Food	100,000	Integrated	Distributed	Distributed	Integrated
Healthcare	1,000,000	Distributed	Distributed	Integrated	Integrated

where *NE* is the number of enterprises, *DI* is the level of delivery integration, *PSI* is the level of product/service integration, *PI* is the level of payment integration, and *OI* is the level of oversight integration.

Delivery integration refers to the extent that the flow of resources across the value network is managed as a single or integrated entity. Product/service integration refers to the extent to which the consumer receives a single product/service. Payment integration refers to the extent that a single user pays for the products/services received. Oversight integration refers to the level of influence, management, and control of product and service delivery by a third-party constituent.

Note that integrated information systems are key to the other types of integration, particularly *DI* and *PI*. This is also the case for *PSI* when the product or service involves access to and use of information, such as in online financial services. The level of information integration differs substantially across types of enterprise. *Finance* has the highest level of information integration; *Healthcare* the lowest. The consequence is the well-known enormous paperwork burden experienced by *Healthcare*.

We would expect C to increase with *NE* and levels of integration – *DI*, *PSI*, *PI*, and *OI* – either required for success or imposed by oversight. *Education* is the least integrated enterprise and, hence, the least complex despite the large number of independent enterprises. It seems reasonable to argue that *Finance* is less complex than *Food* as it is a much less diverse industry and, until recently, oversight was less complicated; a case in point is the contrast of the Federal Reserve with the Food and Drug Administration.

Considering *Healthcare*, the fragmentation of provider enterprises and the third-party payment system, via either employers or government, contributes substantially to the complexity of this enterprise (Rouse, 2008). The lack of standard processes and practices can be contrasted with *Food* or, in general, *Retail* (Basole and Rouse, 2008). Hence, the complexity of *Healthcare* exceeds that of *Food*.

It could be argued that *Defense* has the greatest complexity due to the levels of integration imposed across all aspects of the enterprise. However, relatively few enterprises are involved and the single customer dictates standard processes and practices. Consequently, it can be argued that *Health* exceeds *Defense* in complexity.

Relationship (2) summarizes this notional analysis of the complexity of these public-private enterprises.

$$C_{Healthcare} > C_{Defense} > C_{Food} > C_{Finance} > C_{Education} \tag{2}$$

In summary, holistic approaches to modeling complex systems can enable qualitative analyses that provide insights into sources of complexity. Such analyses are particularly useful when they enable benchmarking one type of system versus another. For example, we found it useful when investigating what healthcare might learn from retail in terms of efficiency of supply chains (Basole and Rouse, 2008). We can see from the foregoing discussions why the complexity of various public-private systems differs.

4.2. *Reductionist approach*

A generalized objective with respect to a complex system is to determine its state, perhaps in order to influence or control the system. While achieving this objective is premised on the system being observable and controllable, consideration of these constructs is beyond the scope of this chapter (Sage and Rouse, 2009). Consequently, similar to Shannon (1948), we define complexity as the amount of information that must be processed to determine the state of a complex system, expressed in bits (or binary units) (Basole and Rouse, 2008).

In order to operationalize this definition, a model of the system of interest is needed. A complex system can be modeled as a highly interconnected and layered network of physical, economic, informational, and social relationships (Basole, et al., 2011). The conceptualization of systems as complex networks is not new. It is based on the fundamental thinking that individuals and organizations do not merely operate in dyadic relationships, but are deeply embedded in complex economic and social systems consisting of numerous inter- and intra-organizational relationships.

The reductionist approach to modeling a complex system requires specification of the entities and relationships that embody a system's structure and enable the dynamics of system behavior. When the model is represented as a network diagram, the basic building blocks of models of complex systems are nodes (entities) and links (relationships). (Note that if we were to adopt another representation such as differential equations or if-then rules, the building blocks used to depict the system would be quite different.) Nodes represent agents or actors (e.g., people or firms), while links represent relationships, or ties, between actors in a complex networked system. We note here that similar modeling approaches are employed in representing causal relationships with the exception that causal models do not only provide binary information about the relationships between nodes/actors but also the probability statements about the relationships.

4.2.1. Healthcare example

The *Healthcare* value network is one of the most complex of the five domains discussed in (Basole and Rouse, 2008). This network can be described as a loose federation of independent enterprises, all trying to optimize the market from their perspective and for their benefit (Fig. 3). No single enterprise or type of enterprise dominates. Further, enterprises from private and public sectors, as well as academia and nonprofit organizations, are laced throughout the value network (Rouse, 2008; Rouse and Cortese, 2010).

This can result in very confused customers, often receiving conflicting guidance from different players. However, this situation will inevitably change. The Internet has enabled highly informed customers to make well-informed choices, while it may have also contributed to information overload, e.g. patients who second-guess doctors or self-diagnose their illness. The vastness of the available information should eventually lead to a more extensive effort to keep it to a manageable level while making informed decisions in areas that have not yet been explored.

Of course, more data does not necessarily mean more information. For example, in the past 10-15 years, many financial derivatives have been constructed using enormous data sets. Use of this financial data prompted excitement from the statistical community. However, not long after, researchers in quantitative finance realized that the increase in the number of observations was offset by increased variation in the data, i.e., "microstructure noise" (Barndorff-Nielsen and Shephard, 2002; Zhang et al., 2005). There is a trade-off between sensitivity (represented by noise) and resolution (represented by the number of observations).

However, great value can be provided by information that is not necessarily that sophisticated. For example, as more information on provider performance – and availability – becomes accessible, health consumers will have greatly increased leverage. Consumers will know which providers perform the best in terms of outcomes and when they have openings on their calendars. Once such information becomes

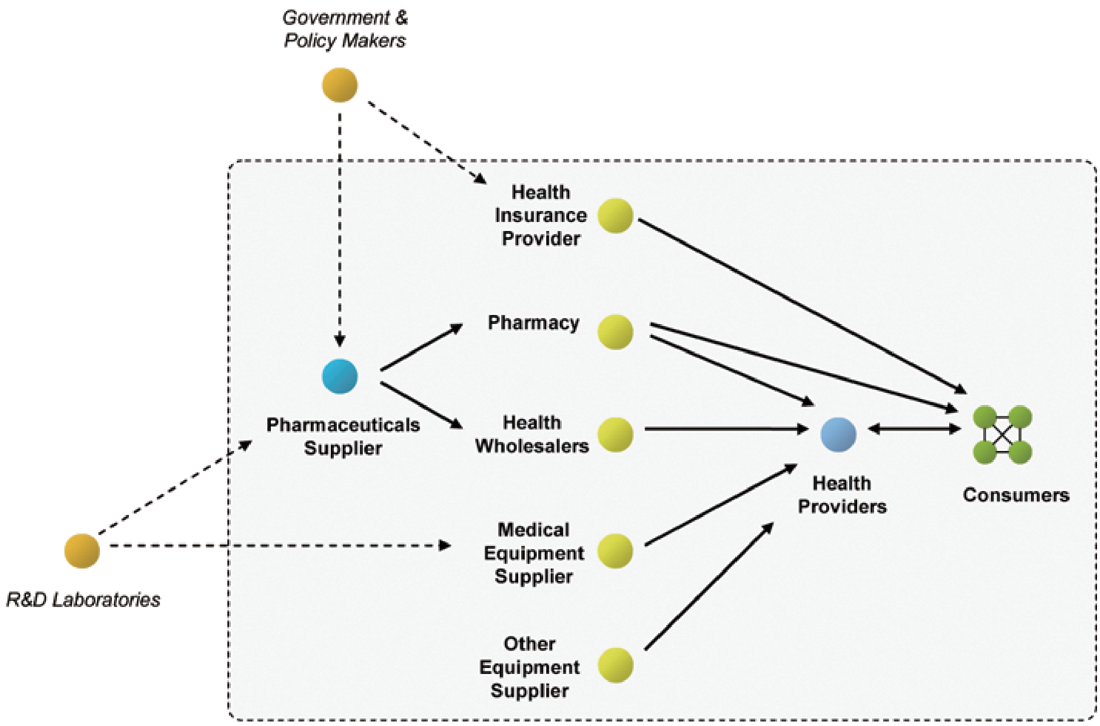


Fig. 3. Network model of healthcare enterprise.

widely available, it can be expected that the extreme fragmentation of the industry will not persist, if only because the projected economics of the industry as it is are not tenable.

In addition, there are a large number of providers of services with many dimensions. Consequently, service is uneven, costs are high, and consumers are often confused and frustrated. The providers and enablers that can fix the B2C (business to consumer) value proposition, while also reducing B2C complexity, are likely to reap enormous benefits. At the same time, the push for "consumer directed" healthcare may result in increased complexity for consumers, which has not proved successful in other markets such as retail and telecom. Innovations that increase B2B (business to business) complexity in order to reduce B2C complexity are more likely to be successful. Of course, these increases in B2B complexity need to be in arrears where it can be managed, e.g., supply chain optimization.

4.2.2. Complexity assessment

In order to assess the complexity of networks such as depicted in Fig. 3, this representation can be generalized as shown in Fig. 4. As discussed earlier, the objective for which complexity is to be assessed has to be specified. The objective of interest is the state of the network. In this section, we present an axiomatic model of the complexity associated with determining network state, based on the axioms of network, probability, and information theories.

The state can be defined as the identity of all nodes involved in any randomly chosen transaction, t_m, where $m = 1, 2, 3, \ldots, T$. Each type of transaction can be selected with probability pt_m. The complexity of the network can be defined as the amount of information that has to be collected to determine the state of the network, i.e., the identity of the nodes involved in the transaction of interest. To determine this, one needs to know the conditional probabilities that particular nodes are involved given the type of

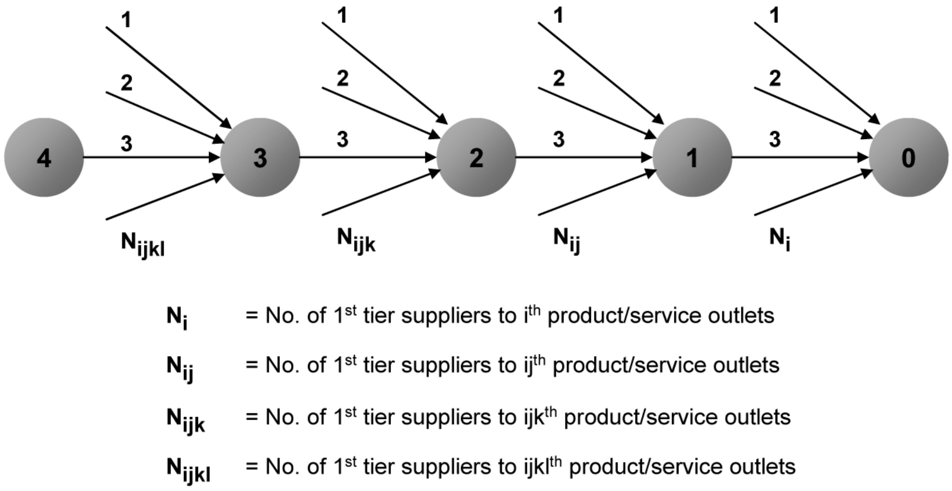


Fig. 4. General network model.

transaction of interest. From Fig. 4, one can see that the conditional probabilities cascade from right to left depending on which paths exist from left to right. In general, not all enablers are suppliers of all providers. Therefore, these conditional probabilities are not uniform.

Given knowledge of the conditional probabilities of interest, Eq. (3) shows how complexity C can be calculated using Shannon's calculation of entropy in information theory (Shannon, 1948). This measure has since been applied in domains ranging from failure diagnosis (Golay, Seong and Manno, 1989) to manufacturing (Deshmukh, et al. 1998; Kaimann, 1974) to sociology (Butts, 2000) as a measure of the observational and/or computational effort involved to assess the state of a system. Indeed, all measures of complexity are based on the characteristics of a representation of a system (Rouse, 2007), with network representations the most common (Casti, 1995).

$$C = \sum_{m=1}^{T} pt_m \left\{ \begin{array}{l} \sum\limits_{i=1}^{N_i} -p(n_i|t_m) \log\left[p(n_i|t_m)\right] \\ + \sum\limits_{i=1}^{N_{ji}} -p(n_j|n_i t_m) \log\left[p(n_j|n_i t_m)\right] \\ + \sum\limits_{i=1}^{N_{jik}} -p(n_k|n_i n_j t_m) \log\left[p(n_k|n_i n_j t_m)\right] \\ + \sum\limits_{l=1}^{N_{jikl}} -p(n_l|n_i n_j n_k t_m) \log\left[p(n_l|n_i n_j n_k t_m)\right] \end{array} \right\} \quad (3)$$

where N_i, N_{ij}, N_{ijk} and N_{ijkl} are the number of nodes at each "tier" of the network and $p(n_l|n_i n_j n_k t_m)$ is the conditional probability that a particular node is involved given the transaction is type t_m, and the logarithm is to the base 2.

The measure of complexity resulting from the above equation is binary digits, or bits. Intuitively, it represents the number of binary questions one would have to ask and have answered to determine the state of a value network. This measure is not without subtlety. For example, if one claims, as we do below, that the complexity of the entire *Retail* market is over 30 bits, there will undoubtedly be many skeptical responses. However, once one explains that this means that more than one billion binary questions

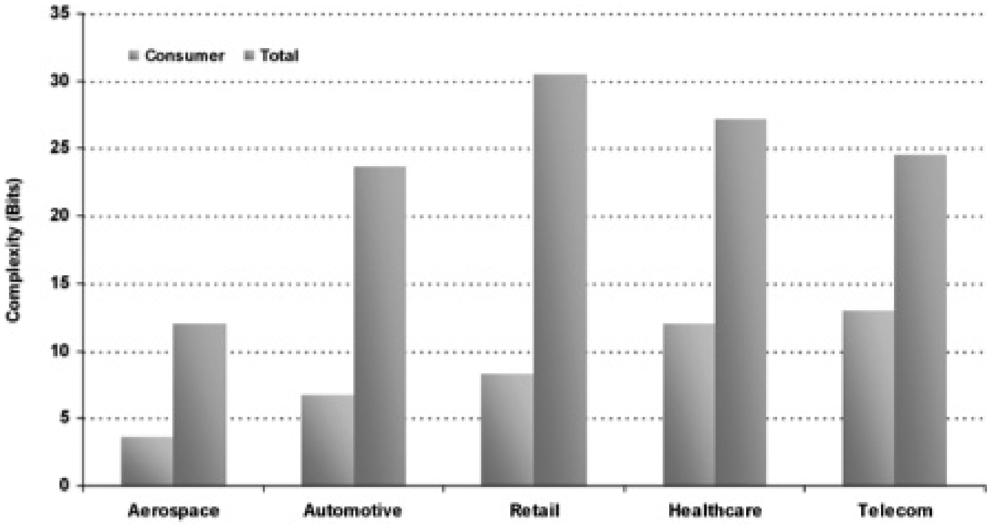


Fig. 5. Complexity assessments.

would be needed to determine the state of the system, people begin to understand the implications of this measure of complexity.

Note that Eq. (3) has repeated terms of the form – *p log p*. If the network of interest included only one upstream node, with probability p of being involved in the transaction and $(1 - p)$ of not being involved, then the complexity calculation would be of the form – *[p log p + (1–p) log (1–p)]*. This value is maximum for $p = 1/2$. In general, if there are N upstream nodes and the probability of each being involved in a transaction equals $1/N$, then uncertainty and, hence, complexity is maximized.

This observation implies that complexity, as we have defined it, can be decreased by greatly simplifying supply chains, i.e., having only one supplier for each element of the system. Unfortunately, this tends to reduce variety and can lead to increased risk of losing the sole supplier for an element of the system. A better strategy may be to allow increased complexity as long as it can be managed by, for instance, enhanced back office information systems. Indeed, this has been the strategy in *Retail*.

The trade-off between excess complexity and decreased variety translates into the trade-off between robustness and hypersensitivity of the system. Excess complexity can be defined as useful redundancy. The need for redundancy is most often for sustainability or robustness of the system. Csete and Doyle (2002) provide very insightful examples. Minimal cellular life is thought to require about ~300 genes, yet even *E. coli* have approximately 4000 genes. Gene knockout studies of *E.coli* confirm that about 90% of its genes are not individually essential for viability in the laboratory. But the "excess" complexity is not merely redundancy, but the presence of complex regulatory networks that effect robustness but not minimal functionality.

Using publicly available data from the Fortune 1000, we were able to identify the number of health-related companies in each node of Fig. 3, as well as for four other domains – aerospace, automotive, retail and telecom (Basole and Rouse, 2008). The probabilities associated with each company being involved in any given transaction were calculated in one of two ways. The predominant way was simply to estimate the probability as one divided by the number of supplier or manufacturers. In a few cases, we adjusted the probabilities to reflect the fact that a Fortune 1000 supplier must be supplying at least one Fortune 1000 manufacturer. The results are shown in Fig. 5.

Table 2
Phenomena and models at different levels

Level	Issues	Models
Society	GDP, Supply/Demand, Policy	Macroeconomic
	Economic cycles	System dynamics
	Intra-firm relations, Competition	Network models
Organizations	Profit maximization	Microeconomic
	Competition	Game theory
	Investment	DCF, Options
Processes	Patient, material flow	Discrete-event models
	Process efficiency	Learning models
	Workflow	Network models
People	Patient behavior	Agent-based models
	Risk aversion	Utility models
	Disease progression	Markov, Bayes models

Several observations are important. First, highly fragmented markets are much more complex than highly consolidated markets. There are relatively few aerospace and automotive providers compared to retailers and consumer products companies. While manufacturers of airplanes and automobiles are likely to claim that their products are complex, consumers do not have to address this complexity and these industries benefit from this. Many more people fly on airlines and drive automobiles than design and develop such systems.

Second, consumer complexity can be reduced by either market consolidation, so there are fewer choices, or by increased B2B efficiency that reduces B2C complexity. The aerospace and automotive industries are examples of the former and the retail industry is an example of the latter. Note that the telecom industry is clearly employing both mechanisms, while healthcare, via consumer-directed healthcare, is moving away from both mechanisms. This suggests that new intermediaries will emerge in healthcare to manage complexity for consumers.

Of particular interest is the comparison of *Retail* and *Healthcare*. *Retail* is the most complex domain because, as indicated earlier a very large number of companies are in the retail industry. However, the consumer does not experience this complexity because of a high degree of back office automation. *Healthcare* includes fewer enterprises, but the lack of integration results in consumers having to deal with much more of the network. If *Retail* operated the same way as *Healthcare*, buying a toaster or can opener at a retailer would result in the consumer receiving ten or more bills from suppliers of components, probably many months later, with little explanation of why this supplier was involved in creating the appliance. This would not make for happy consumers.

Note that this conclusion regarding the complexity of *Healthcare* is consistent with our earlier conclusions based on more holistic analyses. The fragmentation of this domain contributes greatly to its complexity, especially for consumers. Thus, we see that the qualitative and quantitative analyses can be quite complementary.

4.3. Multi-level models

It is often useful to represent complex systems at multiple levels. This allows for what might be characterized as a hybrid holistic-reductionist model. Figure 6 illustrates this concept for healthcare delivery (Rouse, 2009; Rouse and Cortese, 2010; Grossman, et al., 2011).

Healthcare is often viewed in terms of clinicians interacting with patients. This is the people level of the multi-level model in Fig. 6. These people interact in the context of "careflow" processes whereby capabilities and information enable delivering patient care to yield health outcomes. These processes

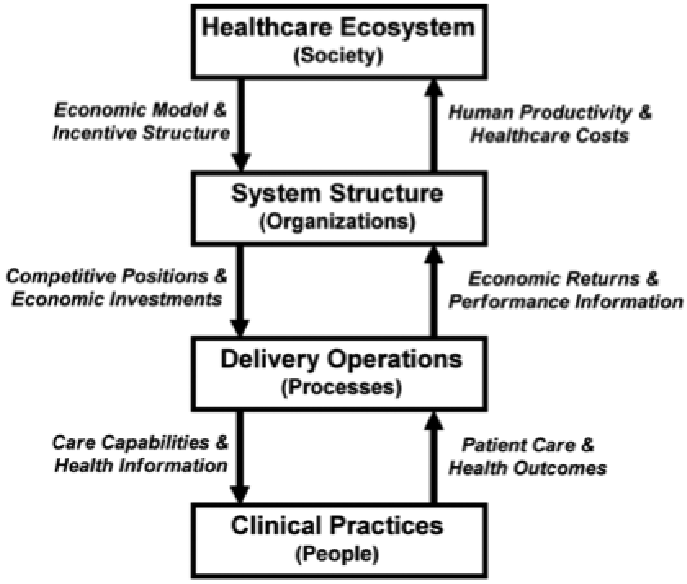


Fig. 6. Multi-level model of healthcare delivery enterprise.

exist within and across organizations that may provide one function or specialty or perhaps a whole process. These organizations – in the U.S., they are often independent businesses – are concerned that their investments yield returns. If they own an MRI machine, for example, they want to make sure it is very busy and generating revenue. This phenomenon is key to the escalating healthcare costs in the U.S. (Rouse, 2009, 2010).

The ecosystem level of this model is where the "rules of the game" are set in terms of incentives and inhibitions that govern how the organizations function. For instance, if enhanced productivity is rewarded then provider organizations will attempt to get patients healthy and back to work as soon as possible. In contrast, if the ecosystem uses price controls to restrain spending, providers are likely to find means to bundle sufficient procedures to remain profitable, whether or not these procedures are warranted from a health outcome perspective.

We have developed a version of the multi-level model in Fig. 6 for the wellness program at Emory University. At the people level, there is an agent-based model where each agent is an Emory employee with a health record and computed risks. The process level is represented as a discrete event model of the flow of program participants through the assessment and coaching processes. The organization level includes microeconomic models of the provider and payer organizations, both of which are organizational elements of Emory. Finally, at the ecosystem level are represented the rules of the game chose by Human Resources and the Benefits Committee.

This model has been useful to explore alternative revenue models for this wellness program. Their current capitated model is due to be replaced by either a fee-for-service model or a pay for outcomes model. In this context, the outcomes are reduced risk of diabetes, cardiovascular disease, and so on. The economic returns at the ecosystem level are reduced downstream healthcare costs and decreased employee productivity loss. Two of the central issues being explored with the model are how the ecosystem level should share savings with the organizational level, and how the organizational level should redesign itself as a function of how savings are shared.

Table 2 addresses the multi-level model more generally. A variety of phenomena can be represented at each level. Similarly, there is a variety of models that can be employed to represent the phenomena

Table 3
Approaches to modeling historical illustrations

Illustration	Representation
Great Depression	System dynamics representation of controls and feedback loops
Housing collapse	System dynamics representation of independent risk and return processes
Healthcare costs	Agent-based model of independent actors optimizing their returns on investments
Pickett's Charge	Agent-based model of each commander and their units, with limited shared information

of interest. There are many good choices among available models, but no single modeling paradigm is always the best choice.

One particular aspect of the multi-level modeling approach merits special attention. Once the multiple levels are elaborated, instantiated, parameterized, and validated, one might be tempted to integrate all the pieces into a single "mega" model. This is often a big mistake. The multiple levels of Fig. 6 provide alternative views of the system of interest. Decision makers often highly value the "inspectable" nature of these different views. A mega model might be more efficient computationally, but the inherent transparency of the multi-level model would disappear, reducing the model to a "magic box" that converts input data to projected outcomes with few, if any, ways to explore why particular outcomes arise.

4.4. Historical illustrations revisited

Table 3 revisits the four historical illustrations discussed earlier. The predominant representation that might be employed in each context is noted. It is likely that other levels of representation would be needed to assess the impact of the context in which these chains of events played out. In particular, it is quite clear that the ecosystems in which these events evolved had an enormous impact.

5. Influencing change

Given this broad review of the nature of causality and complexity, as well as how complex systems can be modeled, how do these concepts, principles, and models enable influencing change? We have found that the following five steps provide helpful guidance when attempting to influence change.

Understand relevant causal pathways: It is essential to understand the causal network from variations of levels of attributes or structural relationships among attributes to outcome changes. The process mapping often done during model development can help with this. Indeed, the discipline required to instantiate models is often a key benefit of a model-based approach.

Understand how interventions propagate: It is also very important to understand how interventions will propagate through the complex system to yield changes. Such propagation is often complicated by many-to-many relationships, feedback loops, redundancies, and so on. Computational models can provide an excellent way to gain this understanding

Global understanding for local interventions: Use global understanding to identify high-leverage local interventions. One cannot expect to change everything, or even very much, so make sure that small interventions have the potential to lead to larger outcome changes. Models can be very useful here.

Attempt to create tipping points: The idea is for small, easy to agree upon changes, to trigger, over time, escalating impacts. Such tipping points can leverage change far beyond their seemingly small starting points. Models can also be very useful for exploring this possibility.

Monitor propagation of interventions: Track how interventions propagate, paying particular attention to higher order and unintended consequences. Models will always be approximations. Thus, model projections and measured data will inevitably disagree. Large enough disagreements should prompt exploration of model inadequacies.

But, how can one predict unintended consequences? One possibility is to show computationally that an important variable is affected negatively rather than positively as expected, e.g., costs increase because more patients avail themselves of the less expensive treatments. This is an example of a phenomena evolving other than as expected.

Much more difficult is the emergence of unexpected phenomena. We have seen this when healthcare executives view our interactive models and say, "But you have forgotten . . . " When we subsequently have integrated such phenomena into these models, the computational models have shown unanticipated and unintended tradeoffs. This form of interactive discovery can be very powerful.

For important changes, for example, transformation of the US healthcare system, one should regularly cycle through the above five steps. A very interesting research question concerns how one might computationally support a process that embodies these steps. Computer-aided monitoring might be the best place to start. In contrast, automated modeling will be a very difficult challenge.

6. Conclusions

This paper has elaborated the conceptual underpinnings needed to understand and influence change in complex socio-technical systems. The nature of causality was first addressed, followed by consideration of the nature of complexity. It was argued that, at least from a practical perspective, the difficulty in understanding causality increases as complexity increases. The possibility of influencing change was addressed in terms of concepts, principles and models for analysis and design in a range of domains or contexts.

References

Aarts, J. (2011). Towards Safe Information Technology in Healthcare: A Survey of Issues and Causes. *Information • Knowledge • Systems Management*, In Press.

Abbass, H. & Petracki, E. (2011). The Causes of No Causation: A Computational Perspective. *Information • Knowledge • Systems Management*, In Press.

Arthur, W.B. (1994). *Increasing Returns and Path Dependence in the Economy*. Ann Arbor: University of Michigan Press.

Balazs, A.C. & Epstein, R. (2009). Emergent or Just Complex? *Science*, 325, 25 September, 1632-1634.

Barndorf-Nielsen, O.E. & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society, Series B 64, 253-280.

Basole, R.C., & Rouse, W.B. (2008). Complexity of service value networks: Conceptualization and empirical investigation. *IBM Systems Journal*, *47*(1), 53-70.

Basole, R.C., Rouse, W.B., McGinnis, L.F., Bodner, D.A., & Kessler, W.C. (2011). Models of complex enterprise networks. *Journal of Enterprise Transformation*, *1*(3), 208-230.

Butts, C. T. (2000). An Axiomatic Approach to Network Complexity. *Journal of Mathematical Sociology*, *24*(4), 273-301.

Carlson, J.N., & Doyle, J. (2002). Complexity and robustness. *Proceedings of the National Academy of Science*, *99*(1), 2538-2545.

Casti, J. L. (1995). The Theory of Networks. In D. Batten, J. Casti & R. Thord (eds), *Networks in Action: Communications, Economics, and Human Knowledge*, pp. 3-24. Berlin: Springer-Verlag.

Chan, D. (2011). Advances In Statistical Analytical Strategies For Causal Inferences In The Social And Behavioral Sciences. *Information • Knowledge • Systems Management*, In Press.

Csete, M.A. & Doyle, J.C. (2002). Reverse Engineering of Biological Complexity, *Science*, 295, issue 1 (March), 1664-1669.

Darwiche, A. (2009). *Modeling and Reasoning with Bayesian Networks.* Cambridge University Press.

Deshmukh, A. V., Talvage, J. J., & Barash, M. M. (1998). Complexity in Manufacturing Systems, Part 1: Analysis of Static Complexity. *IIE Transactions*, *30*(7), 645-655.

Economist (2008). Link by link. *The Economist*, October 16.

Economist (2009). In Plato's cave: Mathematical models are a powerful way of predicting financial markets. But they are fallible. *The Economist*, January 22.

Freeman, W. (2000). *How Brains Make Up Their Minds?* Columbia University Press, NY.

Gell-Mann, M. (1995). What is complexity? *Complexity*, *1*(1).

Gladwell, M. (2002). *The Tipping Point: How Little Things Can Make a Big Difference*. Boston, MA: Back Bay Books.

Golay, M. W., Seong, P. H., & Manno, V. P. (1989). A Measure of the Difficulty of System Diagnosis and its Relationship to Complexity. *International Journal of General Systems*, *16*(1), 1-23.

Gray, C. (1991). Streetscapes: The Bank of the United States in the Bronx; The first domino in the Depression. *The New York Times*, August 18.

Greenspan, A. (2007). Age of Turbulence: Adventures in a New World. New York: Penguin.

Grossman, C., Goolsby, W.A., Olsen, L., & McGinnis, J. M. (2011). *Engineering the Learning Healthcare System*. Washington, DC: National Academy Press.

Haslett, T., Barton, J., & Stephens, J. (2011). How to Watch the Right Butterfly: Some Guidelines for Organizational Design. *Information • Knowledge • Systems Management*, In Press.

Hoffman, R., & Klein, G. (2011). Naturalistic Investigations and Models of Indeterminate Causal Reasoning: Things Aristotle and Hume Never Told You. *Information • Knowledge • Systems Management*, In Press.

Jablin & Putnam (2000). *The New Handbook of Organizational Communication: Advances in Theory, Research and Methods*. Sage Publications.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*, 1-24.

Kaimann, R. A. (1974). Coefficient of Network Complexity. *Management Science*, *21*(2), 172-177.

Kelso, JA. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*, MIT Press, Cambridge, Massachusetts, London.

Lissack , M. (2011). Resisting Clueless Categories. *Information • Knowledge • Systems Management*, In Press.

Lowenstein, R. (2001). *When Genius Failed: The Rise and Fall of Long-Term Capital Management*. New York: Random House.

McCrabb, M., & Drabble, B. (2011). Determining Causality and Dependency in Loosely Coupled, n-Dimensional Social Networks. *Information • Knowledge • Systems Management*, In Press.

Mumford, S., & Anjum, R. (2011a). Fundamentals of Causality. *Information • Knowledge • Systems Management*, In Press.

Mumford, S., & Anjum, R. (2011b). Effects of Context. *Information • Knowledge • Systems Management*, In Press.

Mukherjee, S. (2010), *The Emperor of All Maladies: A Biography of Cancer*. New York: Scribner.

Namatame, A. & Komatsu, T. (2011). Management of Systemic Risks and Cascade Failures in a Network Society. *Information • Knowledge • Systems Management*, In Press.

NIST (2004). http://www.nist.gov/dads/html/complexity/html.

Ottino, J.M. (2004) *Engineering complex Systems*, *Nature*, 427, January, 399.

Pearl, (1985) Bayesian networks: A model of self-activated memory for evidential reasoning, *Proceedings of 7th Conference of the Cognitive Science Society*, University of Califormia at Irvine, pp. 329-334.

Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge, UK: Cambridge University Press.

Rouse, W.B. (2000). Managing complexity: Disease control as a complex adaptive system. *Information • Knowledge • Systems Management*, *2*(2), 143-165.

Rouse, W.B. (2003). Engineering complex systems: Implications for research in systems engineering. *IEEE Transactions on Systems, Man, and Cybernetics – Part C*, *33*(2), 154-156.

Rouse, W.B. (2005). *Thoughts on Complex Systems: An Initial Framework and Discussion Questions*. Atlanta, GA: Tennenbaum Institute, October.

Rouse, W.B. (2007). Complex engineered, organizational & natural systems: Issues underlying the complexity of systems and fundamental research needed to address these issues. *Systems Engineering*, *10*(3), 260-271.

Rouse, W.B. (2008). Healthcare as a complex adaptive system. *The Bridge*, *38*(1), 17-25.

Rouse, W.B. (2009). Engineering perspectives on healthcare delivery: Can we afford technological innovation in healthcare? *Journal of Systems Research and Behavioral Science, 26*, 573-582.

Rouse, W.B. (2010). Impacts of healthcare price controls: Potential unintended consequences of firms' responses to price policies, *IEEE Systems Journal*, *4*(1), 34-38.

Rouse, W.B., & Basole, R.C. (2010). Understanding complex product and service delivery systems, in: *Handbook of Service Science*, P. Maglio, ed., (pp. 461-480). London: Springer.

Rouse, W.B., & Cortese, D.A. (eds), (2010). *Engineering the System of Healthcare Delivery*. Amsterdam: IOS Press.

Sage, A.P., & Rouse, W.B. (eds), (2009). *Handbook of Systems Engineering and Management* (2^{nd} ed.), New York: Wiley.

Salomon, Felix (2009). Recipe for disaster: The formula that killed Wall Street, *Wired*, February 23.

Shannon, C. (1948). A Mathematical Theory of Communication. *Bell Systems Technical Journal*, *27*, 379-423.

Sloman, S. & Fernbach, P. (2011). How People Represent Complex Causal Systems. *Information • Knowledge • Systems Management*, In Press.

Sowa, J.F. (2000). *Processes and Causality*. http://www. jfsowa.com/ontology/ causal.htm.

Spirtes, P., Glymour, C., & Scheines, R. (2001). *Causation, Prediction, and Search*. Cambridge, MA: MIT Press.

Tsallis, C. (2011). Entropy: A Unifying Path for Understanding Complexity Concepts in Natural, Artificial and Social Systems. *Information • Knowledge • Systems Management*, In Press.

Wang, X.F. & Chen, G. (2001). Synchronization in scale-free dynamical networks: robustness and fragility. *IEEE Transactions on Circuits and Systems*, Vol. 49, 1, 54-62.

Watkins, N.W. & Freeman, M.P. (2008), *Natural Complexity*, *Science*, 320, April, 323-324.

Xie, Y. (2011). Population Heterogeneity and Causal Inference. *Information • Knowledge • Systems Management*, In Press.

Zhang, L., Mykland, P. & Aït-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100, 1394-1411.

Authors' Bios

William B. Rouse is a professor in the College of Computing and School of Industrial and Systems Engineering at Georgia Instituite of Technology. His research focuses on understanding and managing complex public-private systems such as healthcare, energy and defense, with emphasis on mathematical and computational modeling of these systems for the purpose of policy design and analysis. Rouse has written hundreds of articles and book chapters, and has authored many books, including most recently *Economic Systems Analysis and Assessment* (Wiley, 2011), *People and Organizations: Explorations of Human-Centered Design* (Wiley, 2007), *Essential Challenges of Strategic Management* (Wiley, 2001) and the award-winning *Don't Jump to Solutions* (Jossey-Bass, 1998). He has edited or co-edited numerous books including *Engineering the System of Healthcare Delivery* (IOS Press, 2010), *The Economics of Human Systems Integration* (Wiley, 2010), *Enterprise Transformation: Understanding and Enabling Fundamental Change* (Wiley, 2006), *Organizational Simulation: From Modeling & Simulation to Games & Entertainment* (Wiley, 2005), the best-selling *Handbook of Systems Engineering and Management* (Wiley, 1999, 2009), and the eight-volume series *Human/Technology Interaction in Complex Systems* (Elsevier). Among many advisory roles, he has served as Chair of the Committee on Human Factors of the National Research Council, a member of the U.S. Air Force Scientific Advisory Board, and a

member of the DoD Senior Advisory Group on Modeling and Simulation. Rouse is a member of the National Academy of Engineering and has been elected a fellow of four professional societies – Institute of Electrical and Electronics Engineers (IEEE), the International Council on Systems Engineering (INCOSE), the Institute for Operations Research and Management Science (INFORMS), and the Human Factors and Ergonomics Society (HFES).

Nicoleta Serban is an associate professor in the School of Industrial and Systems Engineering at Georgia Institute of Technology. She received her Ph.D. and M.S. in Statistics from Carnegie Mellon University. She also holds a B.S. in Mathematics and an M.S. in Stochastic Processes and Theoretical Statistics from the University of Bucharest. Dr. Serban's research crosses multiple disciplines including methodological statistics, molecular biology, healthcare, industrial engineering and socio-economics. Her primary methodological contributions in statistical research are for the analysis of multiple time-varying random functions. Her research has been published or accepted in the top journals in statistics, engineering and biology. Dr. Serban received the 2004 David Byar Young Investigator Award by the Biometrics section of American Statistical Association as well as other paper awards and an NSF Career Award in 2010.