Complexity: Definition and Reduction Techniques

Some Simple Thoughts on Complex Systems

Jon Wade[1] and Babak Heydari[1]

Abstract. Complexity can mean many things to many people. This paper presents an empirical, relativistic definition of complexity relating it to the system of interest, the behavior which is to be understood, and the capabilities of the viewer. A taxonomy of complexity is described based on this definition. Complexity and complication are compared and contrasted to provide some context for these definitions. Several methods of reducing or managing complexity are presented, namely abstraction, transformation, reduction and homogenization. Examples are given for each of these.

Keywords: System complexity, system complication, system complexity management

1 The Age of Complexity

"I think that the next century (21st) will be the century of complexity"
- Stephen Hawking

The 20th century was certainly a time when we witnessed technological advances on a number of fronts including agriculture, transportation, communication, computation, energy, medicine and the like. However, the 21st century is one of complexity in which the interaction between these technologies, human behavior and the forces of nature form new and evolving systems.

A number of systems trends have been driving the exponential increase in system complexity. The notable reasons for this are an increase in both the scale and scope of interconnectivity and the increased participatory role of human agents. The dramatic increase in Software and Networking has had the major impact on interconnectivity. No longer are interactions limited by physical connectivity as they are in electro-mechanical systems. It is not possible to clearly define the impact of changes in software as was done in electro-mechanical systems. Hence, it is not as easy as it once was to determine what makes a car stop and go. Note that the cost of software in an automobile is greater than the cost of the steel.

Networking greatly impacts the quality of interconnectivity among agents of the same type. Notably, both the speed and richness of communication has increased with examples such as high-frequency trading and communication in the academic

[1] Stevens Institute of Technology, Castle Point on Hudson, Hoboken, New Jersey 07030, USA.
jon.wade@stevens.edu, babek.heydari@stevens.edu

community. These factors also increase the quantity and connectivity between agents that previously had no connections. Socio-technical systems are driven by the human element. In the past, connectivity between humans was largely based on geography (according to one estimate the average American in the 1800's traveled an average of 50 meters per day [1].

Today, with the internet we are rapidly increasing the number of people who communicate and interact at a distance. In addition, through search technology we are discovering the shortest paths between two points. The hidden small world is now becoming much more visible. The six degrees of separation are being reduced, made accessible and turbocharged. The participation cost has been greatly reduced so that more people can participate which is an increase in scale, but is also an increase in scope as those who were once bystanders are becoming active participants as with Web 2.0. Finally, the impact of human behavior has risen to the degree that affects nature on a global scale. The world has truly become interconnected. The notion of systems with fixed boundaries, with agents playing fixed roles is quickly disappearing.

Complexity is the challenge, but what is complexity? How do we define it?

2 Complexity and Complication: Definitions

While many use the terms complexity and complication interchangeably, they have very different meanings. The literature follows these three basic categories of defining complexity.

Behavioral Definitions: The system is viewed as a black-box and the measures of complexity are given based on the outputs of the system. Behavioral measures include complexity as entropy in which the Shannon entropy of an output message from the system is regarded as a relatively objective measure of complexity [2]. Another definition is the effective complexity of the system in which the output of the system is divided into two parts: regularities and randomness. The effective complexity is the information content of the regularities whose determination is subjective and context dependent [3,4]. Statistical complexity defines complexity as the minimum amount of information from the past outputs of the system necessary to predict the future outputs [5]. This approach is problematic with respect to contexts which involve non-linear state changes to the system.

Structural Definitions: A measure or definition of complexity is given based on the structure/ architecture of a system. Many refer to the complexity of a system based solely on size. This is an objective definition that is perhaps the easiest to quantify. While complex systems quite often have a large number of components, complexity is more about how these components interact and are organized. For example, there are 45K protein coding genes in rice and 25K in Homo sapiens, but few would argue that rice is the more complex of the two. Another approach is to look at fractal dimensions [6]. While this definition is insightful, it is limited to certain types of structures. There are also hierarchical measures [7]. Simon claims that all complex systems have some degree of hierarchy and making building blocks on various levels is an important way that nature creates a complex system. However, the determination of the building blocks is arbitrary and context dependent.

Constructive Definitions: The complexity of the system is determined by the difficulty in determining its future outputs. The logical depth approach [8] shows how difficult it is to construct an object and regards the difficulty from a computational perspective, translating complexity into the number of steps needed to program a Turing machine to produce the desired output. This is a computational approach in which everything in the system needs to be digitized. Another approach is using thermodynamic depth [9] which is a more general form of logical depth which measures the amount of thermodynamic and informational resources necessary to construct an object. This method attempts to mimic the structure of a system by regenerating the output, but as this approach views the system as a black-box it can result in an unnecessarily large depth for the system.

There are also more general definitions of complexity [10, 11]. While all of these approaches have their merit, they do not seem to answer the essential question of what we mean when we use the word *complexity*.

The word *complicated* is from the Latin *com: together, plicare: to fold*. The adjective meaning of 'difficult to unravel' was first used in 1656 [12]. Interactions in complicated systems are often restricted with respect to interconnection, and can often be unfolded into simpler structures. In this case, decomposition works, while complex systems cannot be so easily unwoven. Complexity is related to the structure of the system.

Complexity is from the Latin *com: together*, *plectere: to weave*. The adjective meaning of 'not easily analyzed' was first recorded in 1715 [13]. Thus, from its first usage, complexity was synonymous with the ease of understanding something. The essence of complexity is interdependence. Interdependence implies that reduction by decomposition can't work, because the behavior or each component depends on the behaviors of the others.

Reductionism which alters the structure of a system cannot be used effectively as an analytic tool for a system whose behavior is critically dependent on these details. The structure often defines the system.

One can imagine complicated systems which are not complex, and complex systems which are not complicated. Figure 1 shows some examples of the possible permutations. The low complexity, low complication quadrant is populated with relatively simple inanimate objects, generally of a mechanical design. Systems engineering has traditionally been most successful in the high complication/low complexity quadrant, and system science in the low complication/high complexity quadrant. However, due to the need to engineer increasingly complex systems such as Systems of Systems and Socio-Technical systems, it is necessary to move systems engineering capabilities from the high complication/ low complexity quadrant, up to the high complication/ high complexity one.

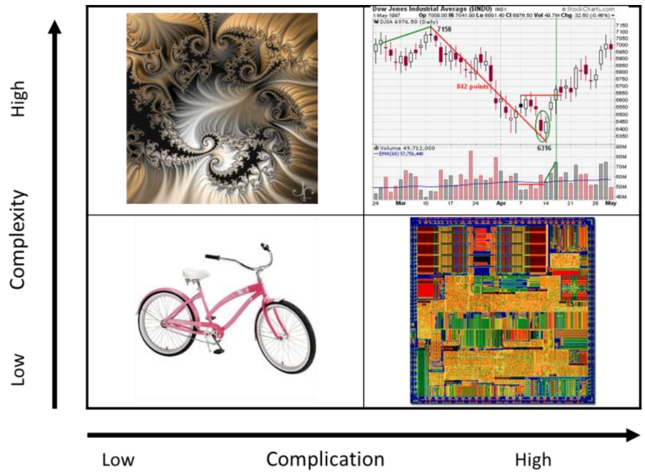

Fig. 1 Complication and Complexity: a) Mandelbrot Set, b) stock market, c) bicycle, d) processor chip.

Kurtz and Snowden [14], in the formulation of the Cynefin Framework, divide the decision making space into four domains, as shown in Figure 2. Roughly speaking, the "known" and "knowable" domains translate into the low-complexity, low-complication quadrant and the low-complexity high-complication quadrants, respectively, while the "complex" and "chaotic" domains are reflected in the high-complexity half of Figure 1.

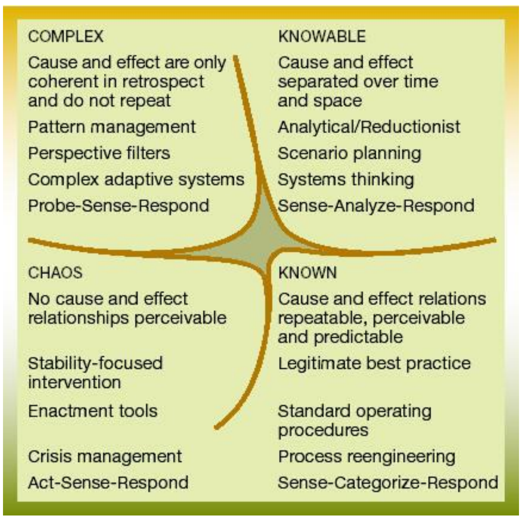

Fig. 2 Cynefin Framework (source: Kurtz & Snowden [14])

One might say that complexity is the degree of difficulty in understanding how a system works and thus how it behaves, but this might be too strong of a statement. With systems that are constantly evolving, it may not be possible to understand all the elements in a system, let alone how they interact, but it might be possible to predict how the system behaves. If we are to embrace complexity, then we need to accept the fact that understanding exactly how a system works may not be possible and we should focus on trying to understand how a system behaves. Thus, *embracing complexity involves a shift of emphasis from how something works to how it behaves.* This is major paradigm shift.

So what are the elements that make the behavior of a system difficult to predict?

One could ask the same question about something else which seems to be just as nebulous, such as 'beauty'. This is just as difficult to define and there probably isn't consensus on examples of beauty, let alone a consensus on the properties of an object, phenomenon or idea that imbues something with beauty. Just how do we objectively measure beauty? What are the common traits between things that are beautiful? Perhaps the same is true about complexity to some degree.

This is particularly difficult if one assumes that beauty is an intrinsic property independent of context and the observer. Rather than try to define beauty in terms of the characteristics of the object, perhaps it would make more sense to define it in terms of the effect that it has on the system which includes the observer. Such a definition might be, "beauty is something that brings pleasure to the observer." With this definition, it is clear that the beauty is dependent on the observer and context and one could imagine the means of perhaps measuring it through an electroencephalogram (EEG) or some other such device.

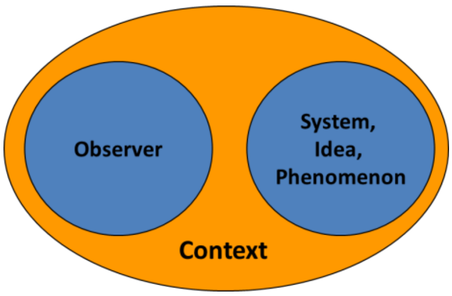

Fig. 3 Complexity – Relationship between Observer, System and Context.

Many others have discussed the critical importance of the observer on the system of interest. For example, philosophers such as John R. Searle [15] have divided the world into the ontologically objective and subjective; and into the epistemically objective and subjective. The Soft Systems Methodology (SSM) proposed by Checkland is a systematic approach [16] that is the result of continuing action research that is used to analyze real-world problems by treating systems as epistemological rather than ontological entities, thus being dependent on human understanding. This is particularly important in the case of complex systems in which the analysis lacks a formal problem definition. This view is supported by the constructionistic epistemologists (first used by Jean Piaget [17]) who maintain that natural science consists of mental constructs that are developed with the aim of explaining sensory experience (or measurements) of the natural world. Some contributors to this philosophy include: dialectic constructivism (Piaget [18]), radical constructivism (Glaserfeld [19]) (Watzlawick [20]), (von Foerster [21]), (Bateson [22]), and Projective Constructivist Epistemology (Le Moigne [23]).

Taking the same approach that we took with 'beauty', as shown in Figure 3, 'complexity' may be defined as a relationship to the observer and context as:

"the degree of difficulty in accurately predicting the behavior of a system over time."

Thus, the degree of complexity is not only related to the system, idea or phenomenon of interest, but also is dependent on the context, the behavior in which the observer is interested, and the capabilities of the observer. Thus, there are a number of means by which the complexity of the system can be reduced without changing the system itself.

We appear to have these same issues with the definitions of other key terms in systems including such things as what is a "system" and the "-ilities" such as security, availability, flexibility, adaptability, etc. To avoid confusion, one should remember that the notion of 'systems' is a model that is employed to make sense of reality and the context and observer are all critical to this model building. Certainly the phrase 'system of interest' makes this point explicit.

Context is a critical aspect in the analysis of systems and is often neglected when discussing complexity. Context has three distinct faces, as described below, and the term is often used to refer to one or a combination of them depending on the situation [24].

Computational Context: When analyzing systems in the space-time domain, the initial and boundary conditions are quite important in determining the state of the system. Context in this sense can change by moving the boundaries of the system or changing the time reference.

Interpretative Context: As an observer, one can have different interpretations of the state of a system, based on the perceptual frame work s/he is using. The state of the system (or parts of it) can be interpreted as order/ signal or disorder / noise depending on the view point of the observer. A particular shape of a termite mound could be viewed as a magnificent structure, if the observer has seen a castle or some similar structure before. Otherwise, the shape can be completely meaningless.

Paradigmatic context: In some complex systems, especially those with human elements, a notion of context emerges as a result of the combination of the internal states of the agents and their interactions. This notion of context includes a set of rules, standards, collective perceptual framework or a value structure. This can be thought of as a generalization of what Thomas Kuhn calls “paradigm” [25] specifically for the scientific community. This is also aligned with the notion of "socially constructed phenomena" that we have already talked about.

In most of the discussions about the context of a system, people refer to the first and sometimes the second form, but rarely the third. The important point is that in human-centric complex systems, there is a cyclical causation between the last two forms of context. In a way, the paradigmatic form shapes the internal interpretative context which itself influences the paradigm of the system.

3 Factors of Complexity

Complexity is far too, well complex, to be described with a scalar quantity. Rather there are several dimensions which reflect the overall difficulty in accurately predicting the future behavior of a system. The following are a set of factors that relate to the overall system of observer, context and system which is consistent with the definition of complexity that we have established. These factors consists of two major components. The first relates to desired accuracy and scope of the prediction and the second relates to the degree of difficulty in obtaining the desired predictive capability.

Prediction quality can be determined to depend upon the achievable precision, timescale and breadth of context. The following are some of the ranges for each of these which are relative to the system of interest.
The precision of predictive capability ranges from:

- Exact (approximate) state is deterministic
- Exact (approximate) states have stochastic probabilities
- Exact (approximate) states have stochastic ordering
- Future (current) states are ill-defined
- Future (current) states are largely unknown

The timescale of predictive capability ranges from:

- Beyond the expected life of the system
- Accepted life of system
- Significant fraction of life of system
- Small fraction of life of system
- Only for small deviations from current state

The breadth of context for the predictive capability ranges from:

- All imaginable contexts
- All likely contexts
- Some contexts
- Only current context

The desired quality level of prediction can be created by specifying a vector in this space.

Prediction difficulty is determined by three critical factors. The first factor is the degree of difficulty in understanding the relationships that govern the interactions and behaviors of the components. The second factor is the degree of difficulty in knowing the current state of the system to the level necessary to apply the relationship knowledge. The final factor is the degree of difficulty in knowing or computing the behavior of system. One of the most challenging aspects of this computation is ability to discover and predict unforeseen emergent behaviors. Quite often these emergent behaviors are dependent upon relationships that are not well understood and may be critically dependent on the system's initial conditions. The following are some of the ranges for each of these which are each relative to the system of interest.

The difficulty in understanding relationships governing interactions and behaviors:

- Essential relationships are well understood quantitatively
- Essential relationships are well understood qualitatively
- Essential relationships are not well understood
- It is unknown which are the essential relationships

The difficulty in acquiring necessary information needed to make a prediction:

- Essential information is known
- Essential information may be acquired with significant effort
- Essential information may not be acquired in that it is not measurable or the act of measuring it causes it to substantially change
- It is unknown what constitutes essential information in the future (currently)

The difficulty in computing the behavior of the system:

- Behavior of the system is evident through mental analysis
- Behavior of the system may be calculated in the desired time on a personal computer
- Behavior of the system may be calculated in the desired time on a super computer (1000x PC)[2]
- Behavior of the system may be calculated in the desired time on a foreseeable super computer (1Million x PC)
- Behavior of the system may be calculated on a theoretical quantum computing system
- Behavior of the system may not be calculable

For example, the relationship of factors is fairly well known in a weather system, but the challenge is to understand the current state to the necessary level of accuracy and being able to calculate the resultant weather more quickly than the actual phe-

[2] A factor of 1000x in computing is approximately equal to 15-20 years into the future.

nomenon. Climate change is much more difficult as the relationships between the relevant factors are not well understood.

This approach embraces complexity in that the taxonomy is not based on how the system works, but rather how it behaves. While other taxonomies may be used to describe the physical characteristics of the system, this may lead to erroneous conclusions about the systems complexity per our definition. For example, a simple cellular automata system may be composed of few agents, have well defined communication and simple rules for behavior, yet result in behaviors that are very difficult to predict. The converse is true as well.

4 Complexity Reduction

What can be done to reduce complexity, that is, to make system behavior more predictable? While some such as A. Berthoz [26] have proposed a set of organizing principles based on biological systems for "simplexity", the means to provide complementary relationships between simplicity and complexity, this paper is intended to describe approaches by which to reduce the difficulty in predicting the possible future behaviors of systems. Four possible approaches described in this paper to reduce complexity are: reduction, homogenization, abstraction, and transformation, each of which is described below.

4.1 Reduction

Reduction is the process of removing superfluous elements from the system, either in practice or in implementation, and/or limiting the context under which the system is allowed to operate and reducing the state space to something which is understood. For example, when using a subway system, most riders are interested in how to travel from point A to point B, making the necessary connections. A map, as shown in Figure 4, provides just this amount of information, by eliminating elements that are not relevant to understanding this particular behavior. It should be noted that reductionism in this case does not eliminate structure, but rather makes the essential structure much more visible.

Reduction in context can be used when a system is moving into a regime in which its operation is not valid, such that steps are taken to move it back into a known space. For example, an integrated circuit's operation is well understood within certain temperature, voltage and frequency constraints and it is not allowed to operate outside this regime where it becomes far less predictable and perhaps chaotic. Thus, a potentially complex system is transformed into one that while being complicated is highly predictable.

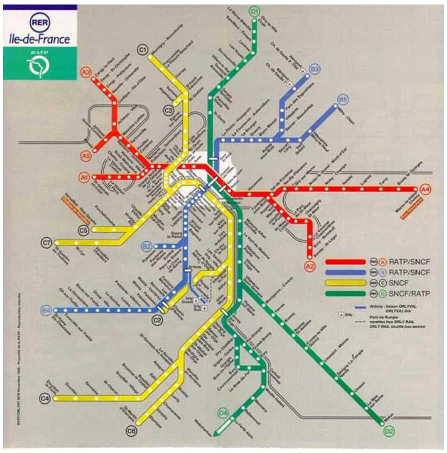

Fig. 4 Paris Metro Map (Source: http://www.ionbee.net/media/parismetromap.jpg)

4.2 Homogenization

Homogenization is somewhat related to reduction in that it provides the possibility to reduce the types of elements or agents by classifying them into sets that are relatively indistinguishable or homogeneous. This is the technique that allows statistics to be applied to situations rather than being forced to understand the behavior of each element. For example, it would be intractable to predict the behavior of more than a few molecules of air, yet the aggregate behavior of 10^{27} such molecules, namely pressure, volume and temperature, can be predicted with a simple ideal gas model if each molecule is treated as being indistinguishable. One should remember that if the behavior of interest is that of the individual molecules, then the system is highly unpredictable, and highly complex. Hence, the same system can be highly complex or very simple depending on the type of behavior of interest and the context of operation.

One must be very careful when applying the technique of homogenization not to overly simplify the model of the system to the point where it is not useful in predicting the desired behavior. For example, one part in a billion can make a big difference in certain reactions. In semiconductors doping levels on the order of 1 part per 100,000 can increase the conductivity of a device by a factor of 10,000 times. There are many systems in which a small amount of inhomogeneity can create starkly different behaviors. For example, pure water in isolation at 1 atmosphere pressure will freeze at -42 degC or even as low as -108degC if cooled sufficiently quickly, while water in the presence of dust or other impurities that can serve as crystallization sites freezes at the familiar 0 degC.

4.3 Abstraction

Abstraction is essentially the ability to decouple elements in a system and transform it from a woven to a folded statement in which interactions are restricted. A good example for this part is language and thought: the more abstraction we enter in our language by encapsulating a notion into a word, the more we will be able to deal with the complexities of a conceptual problem. In fact, the creation of jargon in a scientific field, is a form of abstraction that serves to reduce the complexity of that field. Mead and Conway's book, *Introduction to VLSI Systems*, published in 1980 [27] codified this layering, as shown in Figure 5, and helped to transform complexity to complication in VLSI systems. This success has allowed the creation of incredibly complicated systems with deterministic behavior which has driven software complexity and networking which has driven us to very complex systems.

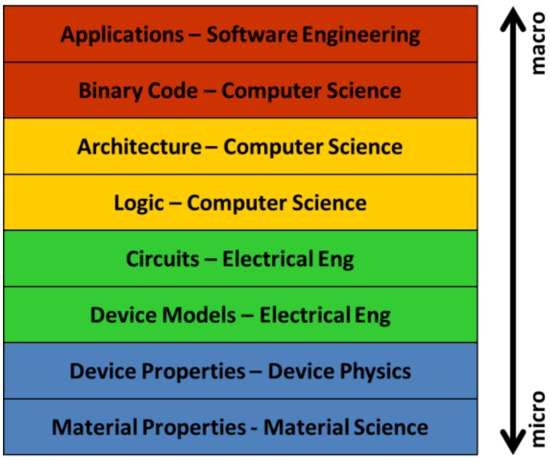

Fig. 5 Layering within Computing Systems utilizing VLSI Technology

It is also interesting to note that abstraction reduces the complexity at the existing boundary of a system, but it also creates a new level of complexity. In fact, this is one of the main mechanisms behind the progress of various fields in human knowledge: Efforts to reduce complexity results in creation of new level of abstractions. The resulting abstractions create a new boundary for the system and generate a new form of complexity, and the cycle continues.

4.4 Transformation

Transformation is a technique in which the problem space is altered such that it becomes more tractable and predictable. An example of this is taking a system that is very difficult to understand in the time domain and performing analysis on it in the frequency domain. Moving from systems governed by rules to ones governed by principles may be seen as a form of transformation. Sometimes perspective can have an enormous impact on one's ability to understand a system's behavior.

One of the important studies in systems science is that of networks. In this case, the system is analyzed with a transformation of its precise structure, to one that is characterized by local and non-local connectivity and diameter (degrees of separation). This transformation enables a significant reduction in the number of factors that need to be addressed to understand the behavior of the system. Each of the systems shown in Figure 6 is composed of networks of systems that experienced evolutionary processes and as a result have a similar network structure with respect to connectivity and diameter. In this case these are composed of 'scale free' networks whose degree distribution follows a power law, such that a small number of nodes have a large number of interconnections, while most have a small number of interconnects.

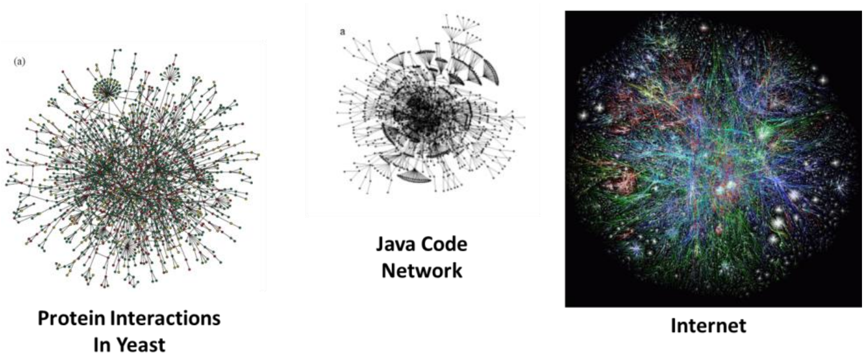

Fig. 6 Network Structures in Evolving Systems

It is known that these types of systems are rather resilient to random faults or attacks, yet are very susceptible to failure in the "too big to fail" nodes. These systems also involve tipping points which when tipped places the system in a different state such that it is usually not easy to return to the prior state. Thus, much can be understood about the system based on a small amount of information.

5 Conclusion

In summary, the following are some of the significant points made in this paper. First, system complexity is increasing exponentially due to increases in both the scale and scope of interconnectivity and the role of human agents in the system. Embracing complexity requires a paradigm shift from attempting to deterministically understand how a system works to how a system stochastically behaves. While one should not give up on understanding the inner-workings of a system, it cannot be assumed that complete knowledge of the system will be possible.

Complexity can defined as: "the degree of difficulty in accurately predicting the behavior of a system over time." This definition includes the critical framework of the system, observer and context. Thus, the complexity of a system can be simultaneously very high or very low depending of the type of behavior that the observer is trying to predict. Complicated systems may have many parts, but the scope and behavior of these interactions are generally well constrained, and their behavior is deterministic. Complex systems, on the other hand, have a much richer set of interactions, and have behaviors that are impossible to accurately predict. System complexity can be viewed from a multi-dimensional taxonomy including precision of prediction, time scale of prediction, difficulty in acquiring necessary information, and breadth of context.

Complexity can be reduced through reduction, homogenization, abstraction and transformation. A final general note to make, which seems obvious, is that when using any of these techniques, some information about the system is lost. Whether that piece of information is crucial or superfluous depends on the context and that particular application of the system. It is always essential to have the assumptions behind each of these four techniques in mind. Many systems failures are the result of a particular simplification technique being used successfully in one context and then being misapplied in another context in which the missing information is critical.

The challenge of science as Einstein put it, is to make things "as simple as possible, but no simpler."

References

1. Urry, John. 2007. *Mobilities.* Cambridge, Polity.
2. Shannon, C. E. (2001). A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review, 5(1), 3-55.
3. Gell-Mann, M., & Lloyd, S. (1996). Information measures, effective complexity, and total information. Complexity, 2(1), 44-52.
4. Heydari, B., & Dalili, K. (2012). Optimal System's Complexity, An Architecture Perspective. Procedia Computer Science, 12, 63-68.
5. Feldman, D., Crutchfield, J. (1998) Measures of statistical complexity: Why?, Physics Letters A, Volume 238, Issues 4-5, February 1998,
6. Mandelbrot, B. B. (1977). Fractals: form, change and dimension. San Francisko: WH Freemann and Company.
7. Simon, H. A. (1962). The architecture of complexity. In Proceedings of the American Philosophical Society.
8. Bennett, C. H. (1995). Logical depth and physical complexity. In The Universal Turing Machine A Half-Century Survey (pp. 207-235). Springer Vienna.
9. Bennett, C. H. (1995). Logical depth and physical complexity. In The Universal Turing Machine A Half-Century Survey (pp. 207-235). Springer Vienna.
10. Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.
11. Horgan, J. (1995). From complexity to perplexity. *Scientific American*, *272*(6), 104-109.
12. complicated. Dictionary.com. *Online Etymology Dictionary*. Douglas Harper, Historian. `http://dictionary.reference.com/browse/complicated` (accessed: April 1, 2014).
13. complex., *ibid.*
14. Kurtz, C.& Snowden, D. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. IBM Systems Journal Vol 42 No 3, 2003.
15. Searle, J. R. (1990). The Mystery of Consciousness, New York Review Books.

16. Checkland, Checkland, Peter B. & Poulter, J. (2006) Learning for Action: A short definitive account of Soft Systems Methodology and its use for Practitioners, teachers and Students, Wiley, Chichester.
17. Piaget, J. (1967). Logique et connaissance scientific [Logic and scientific knowledge]. Paris: Gallimard.
18. Piaget, J. (1970). L'épistémologie génétique. PUF, Paris.
19. Glaserfeld, von E. (1995), Radical Constructivism. A way of knowing and learning. The Falmer Press, London.
20. Watzlawick,P. (1977). How real is real? New York, Vintage.
21. Foerster, von H. (1984) Observing systems. Intersystems Publications Seaside, C.A. 2nd ed.
22. Bateson, G. (1972). Steps to an ecology of mind. Ballentine Book., New York.
23. Le Moigne, J.L. (1995). Que sais – je?, Les épistémologies constructivistes. PUF, Paris.
24. Heydari, B., Mitola, J., & Mostashari, A. (2011, April). Cognitive context modeling in the socio-technical systems. In Systems Conference (SysCon), 2011 IEEE International (pp. 272-277). IEEE.
25. Kuhn, T. (1962). The Structure of Scientific Revolutions. The University of Chicago Press, Chicago.
26. Berthoz, A. (2012). Simplexity: Simplifying Principles for a Complex World. Yale University Press.
27. Mead, C., & Conway, L. (1980). *Introduction to VLSI systems* (Vol. 1080). Reading, MA: Addison-Wesley.