Considering complexity in healthcare systems

Thomas G. Kannampallil *, Guido F. Schauer, Trevor Cohen, Vimla L. Patel

Center for Cognitive Informatics and Decision Making, School of Biomedical Informatics, University of Texas Health Science Center, Houston, TX, USA

A R T I C L E   I N F O

Article history:
Received 12 April 2011
Accepted 24 June 2011
Available online 7 July 2011

Keywords:
Complexity
Complex systems
Functional decomposition
Interrelatedness
Healthcare

A B S T R A C T

A recent trend in the literature has been to characterize healthcare activities in terms of complex systems. Complexity has often been loosely and variously defined, with meanings ranging from “not simple” to “complicated” to “intractable.” In this paper, we consider various aspects of complexity and how they relate to modern healthcare practice, with the aim of developing research approaches for studying complex healthcare environments. We propose a theoretical lens for understanding and studying complexity in healthcare systems based on degrees of interrelatedness of system components. We also describe, with relevant caveats, how complex healthcare systems are generally decomposable, rendering them more tractable for further study. The ideas of interrelatedness among the components of a system as a measure of complexity and functional decomposition as a mechanism for studying meaningful sub-components of a complex system can be used as a framework for understanding complex healthcare systems. Using examples drawn from current literature and our own research, we explain the feasibility of this approach for understanding, studying, and managing complex healthcare systems.

1. Introduction

The term “complexity” is often used in the scientific literature to define tasks or systems ranging from complicated to intractable, with a general meaning of being “not simple.” As noted by a Nobel laureate and co-founder of the Santa Fe Institute, “a variety of different measures would be required to capture all our intuitive ideas about what is meant by complexity and by its opposite, simplicity” [10]. What is generally acknowledged in prior research is that definitions of complexity are ambiguous, context-dependent [19], and subjective [10]. While researchers have referred to complexity of healthcare practice as an important consideration for patient safety and quality [1,17,22,24,32,36], it is important to note that some of this work has been met with skepticism (e.g., [21,26]), provoking responses that the key ideas of complexity theory used in healthcare are often distorted ideas “trotted out in the guise of complexity” [21] and are merely the “emperor’s new toolkit” [26].

Complexity theory has been used to study different aspects of healthcare, including management [24], continuity of care [33], nursing [18], and decision-making [6]. For example, Bar-Yam [1] built his analysis of the entire US healthcare system around an organizational definition of complexity, in which complex organizations (such as those engaged in healthcare practice) were distinguished by their designated tasks. These tasks were numerous, diverse, and performed by unique individuals. In contrast, Innes and colleagues [15] considered the individual patient consultation as their unit of analysis and highlighted features of this encounter that were analogous to those exhibited by complex adaptive systems, such as non-linearity (leading to uncertainty) and adaptation to the influence of outside agencies [23]. Most of the prior work on healthcare complexity is descriptive and provides limited insights for researchers and practitioners on how to study and understand complex systems.

In this paper, we propose a theoretical lens for understanding and studying complexity in healthcare systems, in particular in dynamic settings such as intensive and emergency care environments. This theoretical lens, predicated on the interrelatedness between the components of a complex system, provides both researchers and practitioners an approach to understanding and managing complexity. Drawing from prior studies and our own research, we provide examples as to how this perspective can be used to identify, study and potentially solve current, relevant problems in the healthcare domain. For this purpose, we consider the various senses of the term “complexity” and how they relate to modern healthcare practice, with the aim of facilitating better-informed research approaches to studying complex healthcare settings.

2. Defining complexity and its properties

Early characterizations of complexity emerged from physics (e.g., chaos theory, network complexity), computer science (e.g.,
computational complexity, cellular automata [13,34], economics, biology, and philosophy (e.g., time-space dimensions of complexity) [10]. Other related characterizations have described complexity as arising from the nature of problems in relation to their solution. For example, "wicked" [27] or "ill-structured" [30] problems have been attributed to have characteristics of complexity. Newer scientific disciplines have appropriated and adapted characterizations of complexity to fit their fields of study, as well. While there is considerable overlap among the many definitions [10], significant disagreement about the notion of complexity and the nature of complexity science remains [23].

We define complexity in terms of one of the most commonly accepted notions—the interrelatedness of components of a system [29–31]. By interrelatedness we mean influence of system components on each other. In this sense, complexity is relative: it increases with number of components in a system, number of relations between them, and uniqueness of those relations. This latter notion of uniqueness reflects the idea that expanding a system by mere repetition or simple transformations of relations and components does not substantially contribute toward its complexity. While the sheer number of components of a system may make it "complicated," it is the degree and number of relationships between the components, both manifest and latent, that make it inherently complex. This interrelatedness among components of complex systems manifests as properties, or features, of the system, such as non-decomposability and emergence, nonlinear behavior, and in some cases self-organization. Several researchers (e.g., [27,29–31]) have described these properties as identifying characteristics of complex systems. These properties, however, can also be understood as consequences of the interrelatedness of system components. Accordingly, we describe two particularly noteworthy properties of complex systems in terms of interrelationships: non-decomposability and non-linear behavior.

Non-decomposability is often a consequence of the web of interrelationships between complex system components. It means that such systems cannot be understood by attending to their individual components in isolation. That is, to the extent that components of complex systems are examined in isolation—by ignoring component interrelations—less can be understood about the system. It is important to consider, however, that some interrelations are usually more substantial than others, such that non-decomposability is not absolute. Non-decomposability does not mean that complex systems cannot be studied; rather, it implies that the focus or granularity in studying such systems needs to accommodate the constraints of interrelations.

An important behavioral outcome of interrelations, a sort of "non-decomposability of actions," is that of emergence [8,9,16]. Interactions between components of complex systems, due to their interrelations, often lead to unexpected behavioral properties of such systems. These properties typically cannot be predicted from the behavioral characteristics of individual system components. A particular form of emergence, self-organizing system behavior, may also occur.

Interrelations between system components tend to complicate responses of complex systems to external influence. That is, as systems become more complex, they tend toward increasingly nonlinear behavior. Linearity is characterized as predictability and proportionality of behaviors in response to external influences; increasingly complex systems tend to behave less predictably and proportionately [24]. When very small influences effect large changes, such behaviors are termed "chaotic." One implication of the nonlinearity of complex systems is their relatively greater detachment, or "freedom," from direct response to environmental influences. This ability of complex systems to maintain certain characteristics despite environmental influences, often termed robustness [4], has inspired the resilience engineering approach [14], which aims to improve the ability of error-critical systems, including the healthcare system, to tolerate human error.

3. Effects of complexity: challenges for "computability"

One of the critical effects of system complexity is on its "computability." In other words, there is a cost involved—in terms of cognitive, computational, temporal, or physical resources required or expended—when working within or on such systems. For example, individuals working within complex systems often expend substantial cognitive or physical effort in performing tasks, or they may employ heuristics (i.e., mental shortcuts) to cope. Likewise, for external observers trying to understand complex systems, the interrelatedness of components introduces significant challenges to effectively understanding the system, and important aspects of such systems are sometimes ignored to their peril. Understanding, describing, predicting, and managing are fundamental goals for individuals who work within complex systems, as well as for those who study them.

How the number of components and (unique) interrelations between components instantiates system complexity can be understood by considering representatively low and high combinations of both. Based on number of components and degree of relatedness, several combinations are possible, with each of these combinations exhibiting different computability challenges. The computability of a complex system (or a component of a complex system) is important to consider, as it is a measure for determining the extent to which one is able to describe, predict, and possibly manage a complex system. We consider four conditions to characterize the range of system complexity. These four conditions are not in any way comprehensive. Several intermediate conditions are possible, depending on numbers of components and degrees of interrelatedness. Our primary purpose is to show the significant range of complexity that can manifest in systems. A summary of these conditions is provided in Fig. 1. In the description that follows, we provide an example in which each condition occurs in a clinical practice setting.

1. Few components, low interrelatedness. These systems are simple, with low computational costs, making them relatively easy to understand, describe, predict, and manage under various circumstances. Moreover, such systems are readily decomposable and exhibit near-linear behavior under most circumstances. An example of few components (physician, note, and computer interface) and relations (inputs and computer responses) would...
be a physician simply copying patient medical information from their hand-written note into an electronic medical record (EMR) interface.

2. Many components, low interrelatedness. Such systems are complicated, in the sense that they contain a large number of components, but the relations between those components are few. This adds computational costs only to the extent that more system components must be considered. As such, these systems can be described, predicted, and managed, albeit at a linearly higher computational cost in comparison to simple systems. For example, consider an EMR system used by multiple personnel (physicians, nurses, pharmacists, and billing administrators), each interacting with the system in a limited manner for their specific role-based tasks.

3. Few components, high interrelatedness. Such systems are relatively complex and require significant computational costs. The relatively small number of components makes them more amenable to description but significantly more difficult to predict or manage. High interrelatedness within the system leads to lesser decomposability. We can possibly study such systems as a “whole” (due to their relatively small number of components), rather than decompose them into functional subcomponents. For example, consider the interactions between team members in a critical care unit during a trauma situation (e.g., a “code blue”). The team responding to a code may include only a few members, but the interaction between team members can be extremely divergent, depending on the situation. Shetty et al. [28], for example, found considerable divergence in the performance (deviation from protocol, errors) between two similar teams for identical resuscitation simulation scenarios, showing the significantly varied behavior of highly interdependent small critical care teams.

4. Many components, high interrelatedness. Such systems are complex and often present very high computational costs. Due to the high interrelatedness between their large numbers of components, such systems are challenging to describe and much more challenging to predict or manage. For example, multiple critical care teams attending to traumas from a mass casualty event have to deal with multiple patients with different conditions, with a significantly changed work environment (e.g., trauma protocols). Such a scenario would be a compounded case of the example presented in condition 3 above, resulting in a significant number of components (i.e., large numbers of patients and patient care teams) and high interrelatedness (within and across team members) to manage workflow in this critical care setting.

In order to study the behavior of systems, one has to understand the nature of organization of the components [35]—that is, identify components and interrelationships between components. In conditions 1 and 2, where the degree of interrelatedness is low, it is possible to describe, predict and manage the behavior of the system. In conditions 3 and 4, the significant number of interrelationships between components can cause highly erratic and unpredictable system behavior.

4. Studying complex systems

Complex systems can appear very different, depending on the aspects, granularity, and circumstances the researcher chooses to focus on. It is important to note, however, that such changes in perspective do not change the complexity of the system, as the system’s complexity is intrinsic and defined by the interrelations of all of its components, regardless of whether those interrelations are examined. Some aspects of interest, levels of analysis, or situations may reveal significant system complexity, whereas others may not. Thus, challenges that come with studying complex systems depend very much on the research questions that are posed and levels of detail and precision to which answers are sought.

As researchers, we cannot satisfactorily characterize complex systems from a global perspective alone. As presciently observed by Herbert Simon decades before the emergence of complexity science as a unified field, one cannot study the complexity of a system without “specifying the content of complexity” [31]. In other words, Simon made a case for complex systems to be decomposed, wherever possible, into smaller functional components and the relations between them. Decomposition, in this context, is the process of characterizing a system in terms of subsystems, or components, and the relations between them in order to characterize them in terms of discernable interrelationships between relatively simpler components. Groen and Patel [12] utilized similar perspectives on interrelatedness for analyzing the nature of coherence in medical text.

The challenge, then, is to identify the right “sizes” of components and the interrelationships that exist across these components. In other words, to extract information from a complex system, “one must focus on the right level of description” [11]. The task of cutting a system at its seams requires significant study of behaviors of components, component interrelationships, and most importantly, whether the isolated subset of the system is representative and appropriate for studying the problem at hand. The key concern here is the identification of the appropriate granularity and seams of functional components that can be further studied. Appropriate decomposition must be based on the nature of the problem being solved, the purpose of studying the complex system (i.e., describe, understand, predict, or manage), and the expected implications of studying the system [11].

In order to decompose a complex system into its constituent components, one has to first identify components (at an appropriate level of granularity). Next, the degree (or strength), uniqueness, and number of relationships between the various components must be determined. Degree of interrelatedness between components affects the system functioning. For example, tight coupling between two components means that their behavior is strongly linked. At the same time, a weak relationship indicates lesser dependence. Influence between components may also be characterized as probabilistic, correlational, or directional. Slicing, or disregarding, strong or unique interrelationships may have significantly greater effects than disregarding weak or redundant relationships in overall system behavior. For example, for studying physician handoff practices within emergency care settings, it is likely that issues related to the transfer of patient-related information are more important to consider than those related to resource availability or bed management. In other words, while creating a functional slice for studying emergency care handoff practices, it is more important to consider the patient-information component than the resources or bed management component.

The nature of the problem being solved and the purpose of studying the complex system are often driven by the research objectives and the context within which the problem is studied. In the example presented above regarding physician handoff practices in critical care settings, the creation of a functional slice was driven by the research question (i.e., studying handoff practices). Accordingly, researchers and clinicians should make appropriate, conscious decisions regarding which relationships to ignore and which to preserve, rather than implicitly making such decisions. While it is often impossible to make a “perfect” functional decomposition of a complex system, understanding relationships between various components should be utilized to decompose the system at its internal boundaries, or “natural seams.”
Toward that aim, the possibility of “latent” interrelationships between components should also be considered. System behavior often varies under different conditions that may or may not expose the nature of relationships between the various components within the system. Some relationships remain latent under most conditions and appear only under certain specific conditions. Such conditions can be considered as states of system “perturbation.” For example, the functioning of an ED changes when there is a sudden influx of critical patients from a mass casualty event. Accordingly, workflow, clinician activities, group behavior, and handoff processes tend to be significantly different from normal working conditions. Several aspects of the ED workflow changes, including that off-service personnel are brought in for clinical support, trauma protocols are adopted, and teaching (i.e., in teaching hospitals) is suspended [25]. In such situations, apparently new dependencies arise or weak ones may be strengthened (such as the activation of trauma protocols or suspending teaching, in the example). Depending on the complex problem being studied, it is important to consider varied system conditions and relationships that such conditions can expose.

Until now, we have been considering systems that may potentially be decomposed into smaller functional components for evaluation and analysis. Such complexity is referred to as organized complexity. Some systems are not amenable to decomposition: for example, those that have all their components interconnected. Weaver [35] describes this as disorganized complexity. In systems (or their functional components) in which disorganized complexity is substantial, the focus has to be on describing these systems in terms of statistical or limiting parameters (e.g., averages, limits, or boundaries). There are many cases of such disorganized complexity in clinical systems. For example, consider an ED where the arrival of new patients on a specific day is less predictable but it is possible to ascertain averages of patient arrival rates over a fixed period of time (say, one month). As researchers and practitioners, it is important to be cognizant of disorganized complexity and that appropriate variables must be considered for their description.

5. Examples of studying complex systems

One of the purposes of this paper is to develop an approach for the study of complex systems. As previously mentioned, our perspective is based on the principle of functional decomposition of a complex system: the problem to be solved has to be specified, followed by delineating the components of the system and the relationships among them, and, finally, isolating appropriate components of the system for study. While there is no set of general heuristics that can be applied for functional decomposition of every system, the above-mentioned stages can be used as a high-level road map. As examples to further explain our perspective, we draw on three research papers. These example papers were selected under the following criteria: (a) the authors describe the environment that they are studying as a complex healthcare setting, and (b) they use some form of functional decomposition process to identify components and specific relationships between components.

5.1. Workflow modeling in critical care

Malhotra et al. [20] developed a model of intensive care unit (ICU) workflow by decomposing the workflow activities to that of individual clinicians. In order to develop a cognitive model and to identify the potential sources of errors in workflow, they decomposed the ICU workflow into temporally organized (a) critical zones of activities, (b) knowledge resources that were used while a clinician (attending, nurse, resident) was in one of the critical zones, and (c) the errors that occurred during these activities. In other words, they decomposed ICU workflow into specific work activities both at the individual and collaborative level, identified the potential relationships between the work activities by a temporal sequencing of critical zones, and, finally, used the relationships to identify the sources of errors or breakdowns. The authors use the metaphor of a picture puzzle to explain the conceptual underpinnings of their work. By breaking down the problem into individual pieces that have unique characteristics and relationships, a model of components of the ICU workflow system was developed. The components of the model that were combined were those that were “sensible and informative” [20]. In this study, the authors decompose the problem of ICU workflow into individual activities and the cognitive requirements associated with those activities. By focusing on individuals (attending physicians, residents, nurses) and related work activities, the authors innovatively reduce the complex network of activities in the ICU to a comprehensible scale. Using layers of other variables including location of activities and temporal sequences, they developed and made intelligible a complex model of ICU workflow.

5.2. ED to ICU transfer

Chalfin et al. [5] investigated the association between the length of ED boarding and outcomes for critically ill patients. Using patient data on transfers from the project IMPACT database, they found that patients with longer ED boarding time (>6 h) had increased hospital length of stay (LOS) and higher mortality rates. In other words, the authors found an association between ED LOS and ICU patient outcomes, taking into account several available objective patient characteristics (e.g., age, gender, resuscitation status, and Acute Physiological and Chronic Health Evaluation II score).

The authors describe a study where they considered ED to ICU patient transfers based on specific patient parameters: patient characteristics, patient acuity, medical conditions, LOS, and outcomes (e.g., mortality). The slice of the selected research problem was limited to certain available and, indeed, relevant parameters. The important point to note here is that there are several more parameters that can potentially influence ED to ICU transfers—for example, role of non-clinical teams (e.g., bed management teams), ED crowding (leading to delays), ICU census, physician expertise, and availability of clinical resources (e.g., X-ray techs). While, most of these factors can influence ED to ICU patient transfers, the paper attempts to decompose the problem into studying only a set of patient characteristics (age, gender, APACHE II score) and its influence of ICU outcomes. As such, in this study, the authors decompose the problem (i.e., ED to ICU transfer) to a simplified version, taking into account only patient condition, ignoring several “interrelated” components that potentially could affect ED to ICU transfers. This may be an appropriate strategy, but it is important to be cognizant of how it may have limited the potency of its conclusions for understanding, describing, predicting, or managing actual critical care environments.

5.3. Collaborative work in a psychiatric ED

Cohen and colleagues [7] characterized the cognitive processes underlying decision making in a psychiatric emergency department (PED) using the theoretical framework of distributed cognition (DC). The PED shows many of the characteristics of complex systems that we have described: work is inherently collaborative and depends on relations within and across teams of diverse clinicians, ranging from physicians to social workers, specializing in substance abuse. Decisions in the unit are influenced by
relationships outside the unit, including relationships to other hospital units, social services, patients’ families, and, at times, local law enforcement agencies. Consequently, the information supporting decision-making in this context is derived from a highly interconnected network of diverse agents. Restricting the focus to the core functional unit responsible for care (usually consisting of an attending physician, resident, nurse, and social worker) allowed for characterization of the ways in which cognition within this unit is distributed across teams, time, space, and artifacts, highlighting latent flaws in the system that were predictive of errors observed during the course of this research.

6. Conclusion

In this paper, we describe a rational approach for understanding and studying complexity, specifically in healthcare settings. Using degree of interrelatedness between system components as an indicator of system complexity, we describe how complex systems can typically be considered in terms of functionally smaller components and the relations between them, based on theoretical, rational, and practical considerations. Such functional decomposition entails a comprehensive understanding of the context of work, its components, and principles that govern the actions of various components within the system.

The specific nature of modern healthcare work renders it particularly amenable to functional decomposition, as work is distributed between actors (physicians, nurses, residents, and other clinical support staff) and artifacts (information technology, machines, paper notes) (e.g., see [2,3,7]). There is often a structure in the relationships that exist between care providers, artifacts, and patients. While some relationships are apparent, others manifest only under certain conditions. As such, it is possible to characterize it as a network of actors, where (at a high level of decomposition) the nodes are actors (or artifacts) and the edges are their relationships. For example, the ED can be considered as a complex network of clinicians (attending physicians, residents, nurses), patients, and information technologies that are used to manage patient care. To study handoff activities in the ED, one has to consider the clinicians involved (actors), artifacts used (paper and electronic records) and information being transferred. Handoff activities can be considered as a sub-network within the larger ED network. In short, the distributed and fairly structured organization of health care settings makes the functional decomposition approach viable.

As with any research approach, there are potential disadvantages to functional decomposition. First, the process of selectively including some components or interrelations and disregarding others may lead to oversimplification of the problem. Second, creating progressively smaller slices of a complex system imposes greater demands toward understanding components and their intricate web of interrelationships to other components. Moreover, using a microstructure level of explanation may be difficult for people outside the field to conceptualize. In spite of these limitations, we believe that our approach is a useful and systematic mechanism for understanding complexity in healthcare settings.

Acknowledgments

This work was supported by a James S. McDonnell Foundation (JSMF) Grant (JSMF 220020152) to Vimla L. Patel. We would like to thank Joanna Abraham for her thoughtful comments on earlier versions of this paper.

References