Synthetic Information Environments for Policy Informatics: A Distributed Cognition Perspective

Christopher Barrett, Stephen Eubank, Achla Marathe, Madhav Marathe, and Samarth Swarup
Network Dynamics and Simulation Science Lab, Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA 24061, {cbarrett, seubank, amarathe, mmarathe, swarup}@vbi.vt.edu

Christopher L. Barrett is the Director of the Network Dynamics and Simulation Science Laboratory at the Virginia Bioinformatics Institute, the Scientific Director of the Virginia Bioinformatics Institute, and Professor in the Department of Computer Science at Virginia Tech. For more than 25 years, he has led advanced research and development programs about computationally augmented decision-making in complex social and socio-technical systems, including detailed epidemiological simulation methods that use high performance computing.

Stephen Eubank is a Deputy Director of the Network Dynamics and Simulation Science Laboratory at the Virginia Bioinformatics Institute at Virginia Tech. His research interests are in building and analyzing mathematical and computational models of complex socio-technical systems. For the past decade, he has focused on providing epidemiological modeling support to policymakers as PI of an NIGMS-funded Modeling Infectious Disease Agents Study (MIDAS) research group.

Achla Marathe is the lead economist and social scientist at the Network Dynamics and Simulation Science Laboratory at the Virginia Bioinformatics Institute and Associate Professor in the Agriculture and Applied Economics Department at Virginia Tech. Her research interests are in network science, socio-economic systems, behavioral economics, and the economics of epidemiology.

Madhav Marathe is the Deputy Director at the Network Dynamics and Simulation Science Laboratory at the Virginia Bioinformatics Institute and Professor in the Department of Computer Science at Virginia Tech. His research interests are in network science, modeling and simulations of large societal systems, high performance computing, theoretical computer science, and policy and decision informatics.

Samarth Swarup is a research scientist at the Network Dynamics and Simulation Science Laboratory at the Virginia Bioinformatics Institute at Virginia Tech. His research interests are in artificial intelligence, machine learning, cognitive science, policy and decision informatics, and computational social science.
Abstract: Socially-coupled policy domains, such as public health, are influenced by the collective behavior of millions of individuals who participate in them and respond to policy plans and interventions. Policy-making for these systems is challenging because their complexity and interdependencies that can have cascading consequences.

The promise of policy informatics has been to provide computational tools that make managing information easier and to facilitate more efficacious policy planning by incorporating methodologies such as simulation. Here, we argue that we need to do more. Successful policy informatics has to take into account the differing perspectives and motivations of different stakeholders in the policy-making process. We describe synthetic information environments as a systematic solution to this problem by presenting a distributed cognition perspective.

Synthetic information is constructed by combining information from multiple unstructured data sources. A synthetic information environment provides models for forecasting and for in-silico experimentation with intervention policies, while also maintaining and propagating constraints between different stakeholder views. When viewed as cognitive augmentation technology, we see that such an environment draws multiple stakeholders into a single distributed cognitive system, thereby enabling efficacious interaction and planning.

We describe two different case studies in the context of pandemic preparedness planning which illustrate the use of our system.

Keywords: Synthetic information; Distributed cognition; Networks; Systems science; High-performance computing; Multi-theory; Cyberinfrastructure, Public Health, Epidemiology, Decision support.

1. Introduction

Policy-making is a distributed cognition problem. It almost always involves multiple social, economic, and infrastructural systems, and multiple stakeholders with different views of the problem and different goals. In addition, the systems involved change rapidly, and are influenced by the collective actions of large numbers of people. Thus any policy action can have cascading consequences, which affect different stakeholders differently.

Despite all this, we can generally view a policy problem sufficiently abstractly to conceive of it as a single problem. Consider the problem of responding to and containing an epidemic. We shall take this as a running example in this chapter. When an infectious disease outbreak occurs, society must respond to contain or eliminate it. This single statement encapsulates what needs to be a complex, distributed yet coordinated, and adaptive response from many individual and institutional actors. It also enables us to see it as a single cognitive problem. This is a cognitive problem because the policy-makers are situated in a perception-action feedback loop, which consists of continual monitoring of the status of the epidemic and adaptive implementation of policies to try to minimize the outbreak. However, it is a distributed cognition problem because no single policy-maker or other stakeholder has access to all the information, nor can they single-handedly take action. Policy-making and implementation is carried out through the collective reasoning and action of a large network of people who are spatially distributed and may have limited communication with each other.

Informally, distributed cognition refers to the coordinated cognitive activity of multiple individuals, and analysis is focused on the system of practice, not just the individuals [Hutchins 1995, 2001].
It has been argued that in some cases, the failure of public health policy has been due to a dysfunction of distributed cognition [Wallace & Fullilove, 2008]. For example, Morales & Fullilove [1992] found that the community response to the AIDS crisis was slow and ineffectual because of the stigma associated with AIDS, and led to Alameda County's treatment and prevention organizations competing with each other rather than integrating into a cohesive whole [Wallace & Fullilove, 2008].

In a similar vein, Comfort [2007] has analyzed the aftermath of Hurricane Katrina to argue that the failure to respond in an appropriate and timely manner was a cognitive failure. She frames crisis management as a complex adaptive system and posits that a cognitive frame is necessary to understand the interaction between communication, coordination, and control, and that “without cognition, the other components of emergency management remain static or disconnected and, as shown by the record of operations during Hurricane Katrina, often lead to cumulative failure.”

It is commonly said that to process the large amount of data and information involved in policy-making, an informatic approach is required. We argue here that this is not sufficient. The approach must be one that enables and facilitates distributed cognition. The goal is not simply to get stakeholders to talk to each other. They already do that. There is already information being passed between them, if in no other way, then through the consequences of their actions. However, by then it is too late. The goal, therefore, is to make their decision-making processes interdependent in the right way.

We do this by building a synthetic information environment (SIE). This is a simulation environment that combines data from multiple sources dynamically, and when used for decision-making, creates a coherent set of views for different stakeholders, thereby enabling coordination and distributed cognition.

Viewing the problem as a problem of distributed cognition changes the focus of analysis from the individual policy-makers to the system of which the policy-makers are a part. As we will discuss, it brings to the fore the problem of developing and maintaining a coherent and consistent picture of the ongoing epidemic (or other policy problem) between a team of decision-makers, even if they are separated in space and time. For an SIE to be a tool for distributed cognition, it must be integrated into the cognitive

A Synthetic Information Environment (SIE) is a collection of software systems, data sets and protocols for policy informatics that provides

- **adaptability**, through support for multiple views and multiple optimization criteria, for multiple stakeholders;
- **extensibility**, through the capability to incorporate multiple sources of data;
- **scalability**, through the capability to model very large, interacting, networked systems; and
- **flexibility**, by allowing evaluation of a large class of possible interventions and policies.

An SIE of a given socio-technical system is built by combining data from multiple sources of information, including the Census, infrastructure elements, time-use data, and geo-spatial data, to create a disaggregated model of the population for a region of interest. This “synthetic” information preserves the structural and causal relationships between interacting elements that constitute the underlying socio-technical system.

An SIE can anonymize individual elements (people or otherwise) so that they are not identifiable, if so desired. This is important for anonymity, privacy, and security. In addition to nominative data, individual elements as well as appropriate sub-systems are assigned numerical, declarative and procedural information, resulting in a realistic and complete representation of the social information environment. SIEs have been used to study topics ranging from the spread of epidemics, to human behavior in the aftermath of a nuclear explosion.
loop of the policy-makers. Then, if multiple policy-makers are part of the same SIE, it effectively induces an overlap in their cognitive systems and provides a structured and consistent state. In this chapter, we elaborate on this perspective.

We proceed with this chapter as follows. We describe the complexity and coupling between behavior and infrastructure during an epidemic, and then argue that policy-making in this context can be abstracted as a cognitive problem. We describe its distributed nature and the constraints induced by this perspective. Then we describe our attempts to address this problem using synthetic information technology, and describe two case studies where our system was used: in the first instance for pandemic response training, and in the second for actual pandemic response during the H1N1 outbreak of 2009.

2. An epidemic as a complex system

An epidemic, as with most policy problems, is a temporally extended event, and often, by the time the problem is recognized, it is too late to prevent it entirely.

For example, in the H1N1 “swine flu” pandemic of 2009, the first case was diagnosed in Mexico on March 17, 2009, though there is evidence that it had been spreading several months before that. From that point on, the virus spread very rapidly. In the United States, the first case was reported in California towards the end of March. Starting in April, authorities in Mexico City started to take control measures by closing down schools and public places, as the severity of the outbreak was recognized. The same was being done in Texas, though data about transmissibility and mortality rates were still unavailable. It was early May when the first epidemiological evaluation of the outbreak was published [Fraser et al., 2009]. By then, it was too late to try to prevent the rapid spread of H1N1 across the world. By June, the outbreak size was large enough that the WHO raised their pandemic alert phase to six (“global pandemic”).

The question of how most effectively to respond to pandemics such as H1N1 is complicated, involving public health systems, regional and urban population dynamics, economic effects, critical infrastructure availability and public policy. An integrated decision-support system in which all of these factors can be studied simultaneously is needed to identify effective response strategies.

It is well understood that planners must take individual behavior into account when preparing for crises. However, it is not as well appreciated that social responses to public policy can significantly impact the efficacy of public policy and disaster response [Sterman, 2006]. In the context of a large-scale crisis, when people’s reactions may produce very unusual patterns of social contacts, it is exceedingly complicated to predict and prepare an optimal response [Comfort, 2007]. Human response, public policies and specific crisis situations are intricately intertwined with one another, making it impossible to obtain a clean, simple formal model and solution.

Consider the following simple example: During the early stages of the recent H1N1 pandemic influenza outbreak, health officials implemented a number of non-pharmaceutical intervention policies aimed at reducing its spread. One important policy was temporarily closing schools. The policy was applied differentially, from closing individual schools all the way to closing all the schools in a given district. New York City ordered the temporary closure of 60 schools in order to reduce overall transmission until an effective pharmaceutical intervention became available.

Although school closures reduce student-student contact, they potentially increase contact between students and the community at large. This could reduce transmission within schools but potentially increase it in the community. Furthermore school closure has potential economic impact, since in many
cases closing a school implies that caregivers have to stay home from work as well. Compliance is another issue. Recent empirical analysis has shown that children did not necessarily stay at home during this period [Borse et al., 2011]. As the pandemic progressed, revised CDC guidelines did not advise across-the-board school closures, but encouraged individual schools to make decisions based on the perceived severity in their school.

Let us examine the factors involved in this decision more closely. It is now generally accepted that closing schools did help reduce the overall spread. There is still an ongoing debate on its implementation: should the schools have been closed earlier? Should they have been closed across the board or based on local conditions? What measures should have been taken to increase compliance and reduce the economic impact? Neither the severity of illness caused by H1N1 nor the susceptibility of the general public was clear in the early stages, further complicating policy decisions. The counter-factual scenario wherein the schools were closed later and H1N1 was more lethal could have produced a disaster.

The factors in these example decisions – uncertain consequences and conflicting motivations between micro and macro levels for individuals, federal, state and local authorities – are at the heart of issues such as non-compliance with public policy and, more generally, breakdown of the rule of law in society. The examples, though complex, are amenable to analysis.

How can mathematical models help in formulating and evaluating policies for such complex problems? Developing models for such complex systems that aim to make a point prediction (such as exactly how many people will be sick on some day) is of very little use. On the other hand, use of models for decision support and policy making that aid in planning, analyzing counter-factual experiments, promoting compliance, and targeting response strategies with limited resources are much more reasonable goals. Over the last 100 years, researchers have developed sophisticated mathematical and computational models aimed at supporting many of these objectives [Hethcote, 2000; Eubank et al., 2004, e.g.]. These range from simple mental models to complex agent-based models supported by high-performance computing.

However the models are represented, they have to be capable of taking information about the state of the epidemic as input, and analyzing the effect of one or more policy decisions on the course of the epidemic. If the models are able to provide a real-time or better response, they can be embedded in the decision cycle of the policy-makers. We refer to this decision cycle as the measure-project-intervene cycle (figure 1).

![Figure 1: The measure-project-intervene cycle in epidemiological policy-making](image)
The measure step consists of performing surveillance to estimate the current state of the epidemic, the project step consists of using models to estimate the possible future evolution of the epidemic, and the intervene step consists of implementing some policy to try to control the epidemic.

The effectiveness of this approach depends crucially on the models used in the projection step. It is important to note here that everyone uses models, whether they are explicit computational or mathematical models, or just mental models based on experience and intuition [Sterman, 2006]. Thus the measure-project-intervene cycle is not specific to our approach. It is an abstraction of the process followed in all policy-making. There are many advantages, however, to using explicit computational and/or mathematical models [Epstein, 2008]. Primarily, they make assumptions explicit, can be much more complex than intuition, and allow rigorous testing.

To complete the loop, feedback from the measurement step should be incorporated into the models to create a projection for the next step. A policy or intervention, therefore, must be an adaptive response to the state of the problem. In the case of the H1N1 pandemic, e.g., the initial policy to close schools temporarily was relaxed during the later phase of the pandemic. This is an example of how policies were adapted during the course of the measure-project-intervene cycle. Indeed this change was based on a better understanding of H1N1 but also feedback from the first round of school closures. Recent work by scientists using a combination of real data gathered via surveys and detailed computational models have provided further understanding of this issue; this will no doubt enhance our ability to better respond in the event of the inevitable next pandemic.

3. Policy-making as distributed cognition

To understand how computational tools can enable more efficacious responses to complex policy problems, we develop a distributed cognition perspective. There are three essential points we wish to make here: a) Mental models are susceptible to systematic errors when faced with complex problems, b) if properly designed, computational tools can function as cognitive augmentation to address these problems, and c) our approach, in particular, facilitates interaction between multiple policy-makers in a consistent manner by unifying them into a single, distributed cognitive system.

3.1 Mental models and cognitive errors

It is well accepted that cognition involves making mental models of the environment and the self [Neisser, 1967; Picton & Stuss, 1994; Grush, 2004]. These include models of dynamic processes [Gentner & Stevens, 1983], and are based on a combination of experience, intuition, and formal instruction. For instance, Vosniadou & Brewer [1992] showed that children’s mental models of the shape of the earth are based on their intuition and experience, modified by what they learn in the classroom. They found that children often start out with a rectangular-earth or disc-earth model. When told that the earth is round, they modify this naive model in various ways, e.g., by assuming that the earth is a hollow sphere where we live on a flat surface deep inside.

Thus new information is incorporated into an existing mental model instead of replacing it. This may lead to subtle errors in reasoning and planning because the inconsistencies in the mental models only come to light upon detailed examination. In the above experiment, the children themselves were unable to completely verbalize their mental model. On being asked the shape of the earth, most would claim that it is round or spherical, but when asked detailed questions like what would happen if we were to keep
walking in a straight line, some would respond that we would reach the end of the earth, but wouldn't fall off because we are inside the earth, which implies the hollow-earth mental model.

Similarly, people make systematic errors in their mental models of dynamic processes. For example, Sterman [1989] studied decision-making in a simulated “beer distribution game”, which involved inventory management in the face of uncertain demand. His simulation of this process contained multiple simulated actors, delayed feedback, and nonlinear effects. He showed that subjects in this game exhibited systematic errors of judgment in accounting for delayed effects of their control actions and tended to attribute changes in dynamics to exogenous events rather than their own actions. This suggests that in the face of complex dynamics, subjects tended to adopt an “open-loop” mental model, which can hinder learning and efficient decision-making.

These and other cognitive and neurological studies have been formalized into mathematical models based on predictor-corrector filters (such as Kalman filters) [Rao, 1999; Grush, 2004]. A schematic of such a model is shown in Figure 2.

This model assumes that the brain maintains an internal estimate of the state of the world, and is constantly generating predictions about observations through a forward model. These predictions are compared with actual observations and are used to “correct” the estimate of the state and to generate new actions. Thus effective performance depends crucially on the quality of the internal forward model and how feedback is incorporated into this model.

We can compare this schematic with the measure-project-intervene cycle in figure measure-project-intervene. The action is equivalent to the policy, and the forward model is equivalent to the combination of the computer simulation and the epidemiologists' mental models. Figure 2 serves to highlight the importance of updating the forward model and the controller based on the feedback.

However, as we see from the discussion above, relying solely on mental models can be hazardous, especially when dealing with complex socially-coupled systems, because people either fail to utilize feedback when faced with complex dynamics, or just try to modify pre-existing, erroneous models to incorporate new information. The latter is especially difficult to determine because such conceptual errors are not immediately obvious.

This is where computational and mathematical models become very important. If built correctly, they can incorporate nonlinearities, delayed effects of actions, and complex dynamics, and can be rigorously tested to analyze their phase space, their sensitivities to parameters and initial conditions, and the variance due to stochasticity. These models, thus, can extend the cognitive abilities of policy-makers and enable them to adapt more effectively to changing circumstances.

![Figure 2: A schematic of a cognitive system that uses a forward model.](image-url)
3.2 Extended cognition

It is important to distinguish this extended cognition from simple, temporary tool use. Clark & Chalmers [1998] discuss at length the conditions that must apply for a cognizer + artifact to constitute an extended cognitive system.

They rely in part on the notion of epistemic actions [Kirsh & Maglio, 1994], which are actions performed to aid cognitive processes (as opposed to actions performed to change the environment to a preferred state). Kirsh & Maglio [1994] studied subjects playing the computer game Tetris and showed that the subjects’ physical rotations of blocks on the computer screen were faster than they could perform mental rotations of the same blocks, and that, while playing the game, the subjects use these physical rotations instead of mental rotations.

Thus, some cognitive activity (rotation of blocks) has been offloaded to the physical system. These actions are desirable for the information they generate (where the blocks will fit), and not for any intrinsic benefit in the Tetris game to rotating the blocks while they are falling. Therefore these are epistemic actions.

Under certain conditions, Clark & Chalmers [1998] argue, it is natural to see the cognitive system as extending to encompass the physical artifacts that are being used to perform epistemic actions. They discuss an example of an Alzheimer’s patient using a notebook to keep track of information. They argue that this person + notebook should be considered a single, extended cognitive system because it satisfies four conditions, which we paraphrase below:

First, the artifact (notebook) is a constant for the particular cognitive activity (remembering). Second, the information provided by the artifact is readily available. Third, the information obtained from the artifact is automatically endorsed on retrieval. Fourth, arguably, the information obtained from the artifact must have been endorsed in the past and in fact, put there as a consequence of this endorsement.

If these conditions are all met, they argue, then the notebook occupies exactly the same position in the life of the Alzheimer’s patient as (internal) memory does for a person not afflicted with Alzheimer’s.

If we accept their thesis, it has multiple implications for the design and use of a tool for policy informatics. First, the tool must be designed in a way to make the relevant information “readily available”. This involves both good software and interface design and good training of users. Second, the users must trust the output of the tool, so that they “automatically endorse” its results. Third, the tool must be integrated into the cognitive loop of the policy-makers, so that it is “a constant” for the activity of policy-making. Fourth, arguably, they must have some understanding of the assumptions in the models and the technology behind the tool. This corresponds to “endorsement in the past”.

In reality, the picture is much more complicated since multiple stakeholders are involved in both policy planning and policy implementation. The cognitive abstraction is helpful in developing a picture of the overall process, and we can now unpack the effect of multiple actors by taking a distributed cognition view.

3.3 Distributed cognition

In epidemic policy, as in other policy spheres, decision-making is divided between federal and local levels. While decisions such as how much vaccine to manufacture, and when, where, and to whom to
make it available are centralized, decisions such as when and for how long schools should be closed are made at the local level (based on federal recommendations).

At both levels, decisions may be affected by considerations other than the epidemic itself. For example, decisions about vaccine manufacture may be affected both by cost and by logistics of distribution. Decisions about school closure may be affected by considerations such as how many children depend on school-provided lunches, and how many class hours would be lost.

These interdependencies between the epidemic, infrastructure, and socio-economic systems not only make decision-making complex and distributed, they result in an implementation of policy that is different from what is intended. In addition, policy implementation requires interpretation of directives within the local context. This can often mean that even if there is complete “buy-in” at the local level about federal policy recommendations, the implementation can still turn out other than intended as local authorities discuss, interpret, and apply directives. This has been analyzed from a distributed cognition perspective in the realm of education policy by Spillane, Reiser, & Gomez [2006]. They have found that policy implementation requires extended “sense-making” activities that are carried out through discussions in formal and informal settings, and are influenced by actual classroom experiences in an iterative fashion. The distributed cognition perspective thus leads to new insight by changing the focus of analysis.

In this view, individual cognition is not constitutive of distributed cognition, rather it is constituted by it. The unit of analysis is the system of practice, not the individual [Hutchins 1995, 2001]. To make sense of cognitive activity that is distributed across multiple individuals and their artifacts, it is necessary to understand how information is represented, communicated, and used in this system.

Hutchins [2001] takes a very broad view of distributed cognition, being willing to consider individual cognition at one extreme, and all of human culture at the other extreme as being instances of distributed cognition.

We take a narrower view, based on the extended cognition perspective discussed above. In our view, a system of multiple individuals and artifacts constitutes a distributed cognitive system precisely if it satisfies the same four conditions, but with information distributed across all the individuals and artifacts instead of just one individual and one artifact.

Essentially, it is the difference between a group of people and a team. A group of people can interact with each other, but don't have a clear division of cognitive activity, don't anticipate each other’s actions and don’t work in a coordinated manner toward some goal.

Optimizing team performance has been a subject of research for many decades. The role and importance of shared mental models in team performance has been recognized for some time now [Cannon-Bowers, Salas, & Converse, 1993; Artman & Garbis, 1998]. However, attention has been focused mostly on social and interpersonal interactions between team members and their effect on performance.

The role of distributed representation on problem-solving performance has received relatively less attention, though Zhang [1998] has showed that when decision-makers only have partial views of a system, there can be a net performance loss because they have to expend time and effort on communication, cross-checking, and constraint propagation. In such a situation, a team can actually have worse performance than an individual expert.

In the case of epidemic policy, different actors (and different institutions) have access to different sources of data, and make decisions under different constraints. For example, policy-makers at HHS may not have
access to all the data available to epidemiologists at the CDC, and may not need it. The epidemic system, in any case, is too complex for a single individual to be an expert in all aspects of the problem. Thus, we are always in a situation where multiple policy-makers have partial views of the problem. Therefore, it is necessary for the different groups to be able to coordinate their decision-making so that they are not working at cross-purposes.

An informatic approach to this policy problem must be able to integrate information from multiple sources and present each group with a view that is relevant to them. Such a system would provide a substrate for distributed cognition, even allowing the individual actors to be separated in space and time, but still binding them into a team.

Referring back to figure 2, the use of an explicit computer model as the forward model allows multiple stakeholders to access a common representation, and to access expected outcomes of interventions along the dimensions most relevant to them.

For several years now, we have been developing synthetic information environments to accomplish this objective. These tools implement complex network-science based models of epidemics on high-performance computing architectures, but keep these details hidden from users, presenting only a clean, easy-to-use interface that allows epidemiologists to construct experiments to evaluate various interventions based on the most current data available. They reduce or eliminate the necessity for explicit crosschecking and constraint propagation between different stakeholders by carrying out these processes automatically, and if necessary, opaque.

**Extended cognition** [Clark & Chalmers, 1998]: An individual and an artifact (e.g., an Alzheimer's patient and the notebook he uses to keep track of information) are said to constitute an extended cognitive system when,

- The artifact is a constant in the individual’s life for a particular activity.
- The artifact makes information readily available to the individual.
- The individual automatically endorses information obtained from the artifact.
- Information was endorsed when it was put into the artifact.

**Distributed cognition:**

Multiple individuals and artifacts are said to constitute a distributed cognitive system when they together satisfy the four conditions above.
As shown schematically in figure 3, when multiple policy-makers use an SIE as part of their extended cognitive system, it creates an overlap in their cognitive states, effectively integrating them into a single, distributed cognitive system.

Our approach aims to facilitate distributed cognition by providing better than real-time evaluations so that policy-makers can use them as part of their decision-making loop. The models used are capable of representing interventions flexibly, which can account for implementation differences on a location-by-location basis.

We next describe how we construct these synthetic information environments, and how they have been used for policy-making during epidemics as well as in other situations.

4. An approach rooted in synthetic information, interaction modeling and multi-perspective decision-making

We have discussed policy-making as a distributed cognition problem. We also motivated the need to develop a holistic computational decision making framework to support public health epidemiology. In this section, we describe an approach developed by our group over the last 15 years. The Comprehensive National Incident Management System (CNIMS), which embodies this approach, is currently being actively used to support public and defense policies as they pertain to large-scale crisis. As a first step we develop synthetic multi-theory, multi-layered (MTML) networks. The coupled multi-networks are composed of social contact networks that serve as the substrate for disease transmission, social friendship networks, and information and economic networks that facilitate information dissemination and economic activity as it pertains to epidemics. This MTML network synthesis is achieved using a first principles approach that combines diverse sources of data and information with well-accepted social and behavioral theories to synthesize MTML networks [Barrett, Bisset, Leidig, Marathe, & Marathe, 2010; Barrett et al., 2011; Barrett, Eubank, & Marathe, 2006; Monge & Contractor, 2003; Barrett et al., 2009].

Available observations and knowledge are normally not structured specifically for a particular question. We overcome this data problem by using available, sometimes imperfect information in the form of data and procedures, to synthesize an integrated representation of what is known in the context of the decision to be made. Additionally, the synthetic data created (population, contact networks, activities, etc.) by our approach have the following features: (1) they are statistically equivalent to the real data, (2) they are anonymous, which helps overcome issues related to human subjects, (3) they are comprehensive and provides justification for certain kinds of policy decisions, (4) their components can be replaced with real data as available, and (5) they represent integrated interaction-related information from multiple sources. It is important to note that these networks are not available explicitly -- we assert that it is simply impossible to construct representations of such networks based solely on measurements.

In the second step we develop high-performance computing enabled models to study dynamical processes on the MTML networks synthesized in the first step. Traditionally researchers have primarily focused on epidemic reaction diffusion processes. But as we discussed in the introduction, planning for and controlling epidemics requires one to understand not just disease propagation but also public policies and the intricate behavioral responses. We have developed a system called Simdemics to support this. It contains models of various reaction diffusion processes pertinent to supporting public health epidemiology; see [Bisset, Chen, Feng, Vullikanti, & Marathe, 2009; Barrett et al., 2010] for details on model description and [Chen, Marathe, & Marathe, 2010] on how such a system can be used.
The final step involves overlaying a decision support environment that allows decision-makers to carry out what-if experiments. The pervasive computing based environment allows decision-makers to design experiments and provides analysts automatic methods for ranking interventions and studying interaction effects as well as contrasts and main effects. The SIE system is schematically illustrated in figure 4.

Different stakeholders or policy-makers can have restricted, partially overlapping views of the system, as desired. The SIE system takes care of maintaining coherence between their views and updating structures as new information becomes available. It acts as cognitive augmentation by allowing users to offload the cognitive activity of forecasting onto the system. Since the same system is used by all the stakeholders, it draws them into a single distributed cognitive system without the need for explicit communication, constraint propagation, and negotiation.

Figure 4: A synthetic information environment (SIE) is constructed through a multi-step process that combines information from multiple unstructured sources, driven by a query and context. The resulting information structure supports multiple stakeholder (policy-maker) views and maintains constraints between them.

**Synthetic Information Environments (SIEs) enable distributed cognition in the following ways:**

- They provide rigorous models that can incorporate nonlinearities, delayed effects of actions, and complex dynamics, and can be rigorously tested to analyze their phase space, their sensitivities to parameters and initial conditions, and the variance due to stochasticity. An SIE, thus, extends the cognitive capabilities of a policy-maker by allowing him to partially offload the cognitive activity of reasoning about a complex system.
- If well designed, they can be incorporated into the mental life of policy-makers to extend their cognitive abilities. This requires designing a flexible and intuitive user interface, appropriate user training, and embedding the SIE into the cognitive loop of the policy-makers.
- When the same SIE is used by multiple policy-makers, it provides a consistent (although not necessarily identical) state for all of them, even if they are concerned with different aspects of the problem. This means that the cognitive systems of multiple policy-makers can be brought into concordance, which facilitates distributed cognition.
5. Applications

In our final section, we outline two recent epidemiological studies/exercises that were conducted using the synthetic information technology discussed above. Each of these studies highlights the role of realistic modeling environments in developing and analyzing public and military policy.

The first study was a tabletop exercise in which officials in two teams used the modeling environment for collective situational assessment and decision support. Policies to control epidemics and maintain force readiness were evaluated in these dynamically evolving scenarios.

5.1 Unified Combatant Command Pandemic Study

This work was conducted as part of an exercise hosted by a Unified Combatant Command to prepare defense decision makers for the information and response environment likely to be encountered in future influenza pandemics [Barrett, Beckman, et al., 2011]. This tabletop exercise sought to provide the participants with a realistic course of events, information flows and stakeholder interests that will be involved during nationwide spread of influenza. Specifically we constructed and ran a multi-scale simulation of the spread of influenza, which served as the “ground truth” of the event, guided the exercise scenario and informed the exercise white team. The results of the spread based on a national model are shown in figures 5 and 6. The simulation was run for a total of 300 days. The figures show how Influenza spreads from the west coast to the eastern parts of the US. Figure 5 shows the state of cumulative spread on day 91 and figure 6 shows the state of cumulative spread on day 193 of the simulation. Our high-resolution models enabled enhanced situational assessment, and leveraged surveillance information available to local public health authorities.

Using synthetic surveillance system techniques an estimate of what surveillance signals the epidemic would create was calculated. These results were interpreted from several different stakeholder perspectives. Additionally the nation-wide simulation was used to run simulations on a more localized (and at a highly-detailed) scale enabling more realistic estimations of the epidemic’s impacts on realistic areas of potential interest (Ft. Lewis area, Ft. Sam Houston area, and Ft. Carson / Peterson AFB / USAF Academy area). Figure 7 shows attack rate with and without interventions for the cities of San Antonio, Texas and Seattle, Washington where Fort Sam Houston and Fort Lewis are located respectively.
The actual state of the epidemic, surveillance reports, detailed assessment of the impacts on the areas of interest, the interpreted actions of both public health and DoD officials, and simulation support events were all organized on a single timeline. This timeline was used to drive the tabletop exercise. The timeline’s interactive display concisely conveys multiple levels of information tailored to the particular interests and influence of different stakeholder communities while still enabling a vision of the overall progress of the exercise.

Requirements defined during support of this exercise led to several novel uses of existing analytical tools and techniques and to significant advances in very large scale, highly detailed social epidemiological simulation technology. This marks the first time that knowledge of existing surveillance systems was applied to simulation results to provide a realistically obscured situational awareness. This enhanced the realism for the tabletop exercise for the participants. The exercise required that we integrate simulations occurring at different scales (nation-wide and regional). Technologically, this had not been done before and led to the development of a framework to enable this linkage. The use of the timeline to organize data from different sources and actions of different stakeholders while presenting a unified view of the events was important for the success of the exercise.

### 5.2 Federal Interagency Support during the Emergence of H1N1 Influenza during April and May 2009

This effort was in direct response to the initial reports of the emergence of the H1N1 Influenza virus that eventually caused a global pandemic. At first, infections were confirmed in Mexico, California, and Texas and, a few days later, New York. The rapid spread combined with initial overestimates of its mortality rate raised serious concerns of a repeat of the 1918 influenza pandemic. Initial reports about the disease characteristics were unreliable, with wide variations placed on important parameters like proportion of symptomatic individuals and duration of infectious period. Having developed a web-based front-end to our SIE called DIDACTIC, for just this purpose, we were able to quickly run a series of studies to learn the likely impact of the variations in these parameters on the US population. A snapshot of this interface is shown in figure 8. A quick report was drafted about the impact of disease characteristics on the size and shape of the expected epidemic curve. Several variants of disease models were added to the DIDACTIC tool.
As H1N1 Influenza continued to spread in the US, the Department of Health and Human Services (DHHS) teamed up with the Defense Threat Reduction Agency to place the DIDACTIC tool in the hands of US government analysts to provide day to day modeling results. At a daily interagency meeting at DHHS, led by the Office of the Assistant Secretary for Preparedness and Response (ASPR) and the Biomedical Advanced Research and Development Authority (BARDA), analysts and policy-makers discussed overnight projections of the pandemic and possible interventions. These projections were generated using DIDACTIC as one of just two tools that were available and capable of providing detailed forecasts and analyses on such a rapid timescale.

This integration inside the 24-hour decision cycle running the federal government’s response to this emerging crisis would not have been possible without the development of highly optimized modeling software as well as the web-enabled interface. The analysts at DHHS were able to perform course-of-action analyses to estimate the impact of closing schools and shutting down workplaces.

DIDACTIC, in conjunction with other modeling tools, was used to study a number of important policy questions, including: (i) the role of anti-virals as prophylactic drugs to reduce the severity of the pandemic and delay its peak, (ii) the effect of school closures in slowing the spread, (iii) the role of travel restrictions, especially international travel restrictions. Officials were able to use modeling tools such as DIDACTIC to evaluate these scenarios. This experience demonstrates the importance and feasibility of placing sophisticated modeling tools in the hands of public health decision makers and highlights the role that highly detailed modeling can play during a response to an emerging crisis. A distributed cognition viewpoint is important here.

We continued to interact with BARDA officials as a part of the NIH MIDAS project to support pandemic response as the pandemic continued to unfold. During the later stages of the H1N1 pandemic a rapid-turnaround study was undertaken to try and understand what conditions would need to occur to create a “third wave” of the pandemic as had been experienced in previous pandemics (1918, 1958). BARDA was charged with procuring the vaccines and ensuring their appropriate distribution, making the prospect of a third wave particularly relevant to their activities.
The study [Eubank & Lewis, 2009] showed that with significant increases in pathogen transmissibility and an underestimation of the current number of infected (or immunized through vaccination), a large third wave could be created. Alternatively, if the pathogen were to significantly change its antigenicity, thus weakening the level of immunity people had achieved through infection or vaccination and some increase in transmissibility as well, a large third wave could occur. Any one of these changes alone however did not seem likely to create a third wave of significant size. These results are illustrated in figure 9. This study along with others, helped alleviate concerns of a third wave and allowed the Department of Health and Human Services to refocus their efforts. Together these studies demonstrated the utility of synthetic information systems and associated high performance computing oriented pervasive modeling environments for situation assessment and public health logistic problems during an ongoing world-wide crisis.

6. Conclusions

This article describes a synthetic information based approach for policy-making and analysis as it pertains to public health and military preparedness.

The creation and implementation of policy is a complex, authorized, meta decision-making process. Policies, institutions, individuals, and associated common resources (e.g., use of societal infrastructures) interact as composed processes and constraints. The dynamics are complex everywhere in the composed system and, themselves, produce new properties and constraints. Socially coupled policies, such as public health, influence and are influenced by the collective behavior of millions of individuals who participate, interact, and respond to policy plans and interventions. Because of their complexity and interdependencies these policies can have cascading, unexpected consequences that are often beyond the unaided cognitive capacity of an individual to understand.

Advances in computational and communications technology, network and algorithm science, complex system dynamics, and cognitive science are creating a new revolutionary change in these sorts of distributed, socially-coupled system dynamics. In such settings, the ideas of an objective observer with access to a representational “world in a bottle”, a single “decision-maker” with purview over that world, and a compliant population become evermore obsolete. Our approach, thus, involves improving
understanding of, and producing decision support environments for, real world, complex, distributed, socially coupled decision systems. Here we have developed a distributed cognition perspective to explain our methodology.

There are multiple aspects to the policy-making problem that are addressed by our approach. First, public health as a policy domain, as with other domains, is very complex due to the scale of the system, its nonlinearities, feedback loops, and social coupling. Mental models are unable to deal with the scope of such domains, and can often contain subtle, unspoken assumptions and errors. SIEs externalize the models and can account for their complexity and stochastic variability, and can act as cognitive augmentation by allowing users to offload the cognitive activity of projecting forward the consequences of interventions in such a complex system.

Second, by virtue of a clean and intuitive interface design (DIDACTIC) and fast response enabled by a high-performance computing back-end, the SIE can be integrated into the decision cycle of policy makers, as was done by ASPR and BARDA analysts during the H1N1 pandemic response.

Third, an SIE is capable of maintaining different views for different stakeholders, and internally ensuring coherence between these views. This reduces the need for explicit communication and constraint propagation between stakeholders, which have been shown to be the main factors affecting team performance when representations are distributed. Thus, an SIE draws multiple stakeholders into a single distributed cognitive system by inducing an overlap between their cognitive states. This helps to make their decision-making processes interdependent and reduces the likelihood that they will end up working at cross-purposes.

Our systems have been used by multiple government agencies both for training and for response. This process has taught us that policy informatics is not simply a matter of organizing information, but that it must take into account the cognitive processes of the policy-makers and stakeholders as well. The perspective presented here motivates the need for developing pervasive synthetic information architectures and provides a natural way to understand distributed decision and policy-making.

Acknowledgments
We thank our external collaborators and members of the Network Dynamics and Simulation Science Laboratory (NDSSL) for their suggestions and comments. This work has been partially supported by NSF HSD Grant SES-0729441, NSF PetaApps Grant OCI-0904844, NSF NETS Grant CNS-0831633, NSF CAREER Grant CNS-0845700, NSF NetSE Grant CNS-1011769, NSF SDCI Grant OCI-1032677, DTRA RD Grant HDTRA1-09-1-0017, DTRA Grant HDTRA1-11-1-0016, DTRA CNIMS Contract HDTRA1-11-D-0016-0001, NIH MIDAS Grant 2U01GM070694-09, NIH MIDAS Grant 3U01FM070694-09S1 and the George Michael Scholarship from Lawrence Livermore National Laboratory.

References


Neurobiology, 4, 256–265.


