THE INTERNATIONAL
C2 JOURNAL

VOLUME 4, NUMBER 3, 2010

SPECIAL ISSUE

Interagency Experimentation

GUEST EDITOR

R. Douglas Flournoy
The MITRE Corporation

Models in Multi-Agency C2 Experiment Lifecycles: The Collaborative Experimentation Environment as a Case Study

Anthony J. Bigbee
Jonathan A. Curtiss
Laurie S. Litwin
Michael T. Harkin

CCRP
Focus & Convergence for Complex Endeavors

THE INTERNATIONAL C2 JOURNAL

About the Journal

The International C2 Journal was created in 2006 at the urging of an international group of command and control professionals including individuals from academia, industry, government, and the military. The Command and Control Research Program (CCRP, of the U.S. Office of the Assistant Secretary of Defense for Networks and Information Integration, or OASD-NII) responded to this need by bringing together interested professionals to shape the purpose and guide the execution of such a journal. Today, the Journal is overseen by an Editorial Board comprising representatives from many nations.

Opinions, conclusions, and recommendations expressed or implied within are solely those of the authors. They do not necessarily represent the views of the Department of Defense, or any other U.S. Government agency.

For more information

Visit us online at: www.dodccrp.org
Contact our staff at: publications@dodccrp.org

CCRP
Focus
& Convergence
for Complex Endeavors

Models in Multi-Agency C2 Experiment Lifecycles: The Collaborative Experimentation Environment as a Case Study

Anthony J. Bigbee, Jonathan A. Curtiss, Laurie S. Litwin, and Michael T. Harkin (The MITRE Corporation, USA)

Abstract

We present the Collaborative Experimentation Environment (CEE) as a case study for the use of models to support multi-agency C2 experimentation lifecycles. The Collaborative Experimentation Environment (CEE) is a distributed capability and means for designing and conducting joint Net Centric Experiments (NETEXs) where the goal is to explore multi-agency coordination and mission effectiveness, such as national disaster response or responding to in-flight security incidents involving terrorism. From the inception of the project, the team has used models in deliberate ways across experiment lifecycles, from experiment conception and design, to post-hoc analysis. In this paper, our goal is to broadly describe our major model types—information flow/decision, event response, scenario, domain simulation, data collection and analysis, and architecture. Using this taxonomy, we review the elements of each model and provide examples. We then describe three completed experiments and identify crucial roles for models within those experiments. We conclude by offering some general lessons learned and identifying future work.

Introduction

This paper presents MITRE's Collaborative Experimentation Environment (CEE) as a case study for the use of models to support multi-agency C2 experimentation. We use the term *models* in the broadest sense: any artifact—computational or descriptive/diagrammatic—shared within the team and with stakeholders that represents a process, concept, or scenario. For some members of the C2 community, models are strictly associated with simulation; this perspective pervades prominent C2 experiment codes such as Kass (2006) and Alberts (2002); the MORS Experimentation Community of Practice *Experimentation Lexicon* does not define the term model. From the inception of the CEE project, the team has used models in deliberate ways across experiment lifecycles, from experiment conception and design, through experiment execution, to post-hoc analysis.

We believe that this paper makes a contribution to the practice of multi-agency C2 experimentation; it is intended as a report *from the field* resulting from two years of experimentation involving military, government, commercial, and other organizations that conduct operations in shared mission space. Our intent is to share a methodology and lessons learned in the spirit of the C2 community "[conducting] better experiments, develop a culture of experimentation, and sharing…the lessons learned" (Alberts and Hayes 2005). We briefly describe the CEE project and the experimental methodology, but focus the majority of this paper on what models we created, why we created them, and lessons learned. We do not advocate CEE experiments as the only or best ways to conduct multi-agency experiments.

MITRE's CEE is a distributed capability and means for designing and conducting joint Net Centric Experiments (NETEXs) where the goal is to explore multi-agency mission effectiveness. The NETEXs are human-in-the-loop discovery experiments (Alberts and Hayes 2005) that often use low to medium fidelity dynamic simulation, scaled to fit between tabletop and command-post experiments or events. Each NETEX environment is supposed to reflect real world coordination and collaboration issues where several agencies or organizations (including private and non-governmental) must execute overlapping missions; one example is an in-flight security incident over North American airspace. Because the missions involve multiple agencies, no single individual or group has complete knowledge of the relevant domain, procedures, and policies; rather, each agency holds a piece of the puzzle. Models depicting complex or complicated processes have proven useful for eliciting domain knowledge from stakeholders as well as documenting them for other team members. Because developing each experiment is a collaborative endeavor, attaining a shared understanding within the team is crucial. More importantly, imparting this larger, joint understanding of the domain has proven beneficial to participating agencies and resulted in operational procedure changes and policy deliberations.

Experimentation Approach

Expectations regarding close collaboration and effective coordination between agencies in many shared mission areas have grown. These expectations, major Federal organization evolution, and continued criticisms and identification of gaps in Federal coordination (GAO 2007) inspired the formation of CEE. A key feature of CEE is that each experiment features a new technology capability concept, new organizational structures and process, or both. Since each experiment involves human-in-the loop interaction within and between cells and organizations, and participants are allowed significant decision-making freedom and creativity, these discovery experiments (Kass 2006) do not completely follow classical precepts found in academic psychology laboratories.

The CEE experiment process includes events that are not pure experiment trials. In particular, a lightweight tabletop event usually takes place two to three months prior to the actual experiment during which the experiment concept, objectives, vignette phases, scenario elements, and roles and responsibilities are presented. This event is used to refine hypotheses, fill in gaps in the team's knowledge, build consensus, and elicit participation in the experiment itself.

Although we do strive for rigor via elements like hypotheses, independent or controlled variables, measurement and instrumentation, there are no repeated trials, only one group of subjects, and full factorial designs are impossible. This results in some loss of control and introduction of internal validity concerns. Many of these limitations stem from the use of subject matter experts with deep domain knowledge drawn from watch floors and operational cells—these participants are difficult to obtain, have limited time available, and may not be able to stay for the duration of a multi-day event. As a result, we often must choose a design that reduces one internal validity threat while increasing another. Relative to single group threats (Trochim 2006), for example, we usually present a new technology concept or a process first, and then remove it to see if performance/outcomes are approximately equal or less in order to address maturity/learning threats. Mortality threats increase, however, as we risk losing participants due to other commitments during later vignettes.

The role of cognitive performance in C2 behavior is not completely accessible to observers or the participants themselves. Thus, our measurement and instrumentation approach is a blend of objective and subjective techniques where we seek to triangulate on the causes and relationships between behavior and outcomes. Finally, one of the CEE goals is for participants and stakeholder observers to learn from the experience and modify local policies and procedures as they deem appropriate. This is a goal not represented in classical experiment designs.

Model Types

In this paper, we present a simple taxonomy comprising six major model types. We then summarize three NETEXs and discuss how models influenced the design, preparation, conduct, and post-hoc analysis phases.

Model Taxonomy and Development

The word *model* has strong connotations for certain communities; we use the term in a broad sense to mean any abstraction of a process, system, or behavior expressed in an artifact intended to be shared or as part of a system used to execute mission tasks in our experiments. Over the course of the CEE NETEXs, we have created six types of models:

1. Information flow/decision
2. Event response
3. Scenario
4. Domain simulation
5. Data collection and analysis
6. Architecture

Five of these types are descriptive—and can be construed as conceptual—and the sixth type, domain simulation, can be either predictive, descriptive, or both.

Figure 1 below depicts a typical flow of model development for each model type across the four major phases of CEE experimentation development.

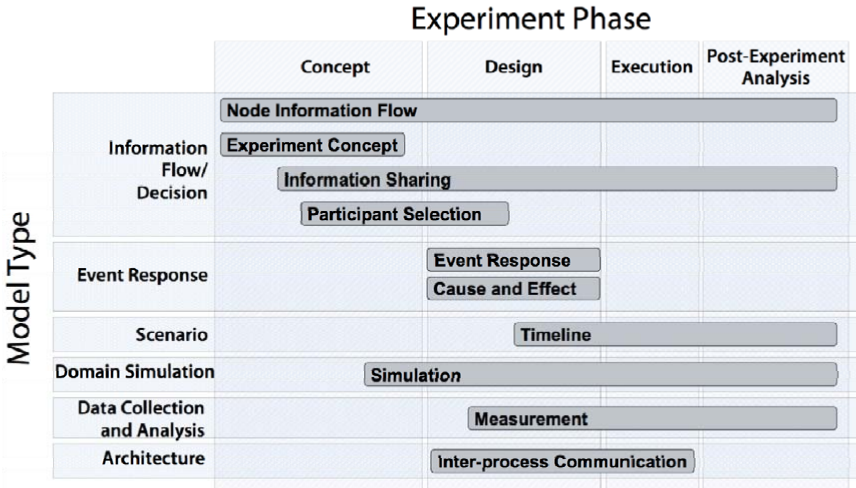

Figure 1. Model Development

A discussion of each of these model types follows. Illustrations for each model are included to give the reader a feel for the model's nature; detailed content is not important for the purposes of the paper. For additional details and descriptions, Maroney et al. (2009) describe a CEE experiment examining Unmanned Aircraft Systems (UAS) multi-agency coordination in hurricane events.

Information Flow Models

Information Flow Models are meant to depict relationships between potential participants. They are used to explore the domain, scope the experiment, refine hypotheses, and act as a reference for other models.

A *Node Information Flow Model* is used to capture a common understanding of the existing real-world relationships between organizations involved in the experiment. This model is created early in the experiment design process and is maintained throughout the duration of the experiment.

The format of Node Information Models is loosely based on *concept mapping* norms, where the nodes (boxes) are concepts, generally nouns, and the arrows connecting the nodes are the inter-relation between the nodes, generally verb phrases. For NETEX development, the nodes are people, organizations, physical locations, or other objects related to the experiment domain. The arrows indicate the flow of information (e.g., sharing, need) and responsibility relationships or other relations.

The model is developed and refined through meetings, often in real time, with internal domain experts and with the organizations expected to take part in the experiment, to ensure that it reflects the real world. Figure 2 describes the use of UAS during crisis response. The node colors correspond to major organizations, such as the Federal Aviation Administration (FAA).

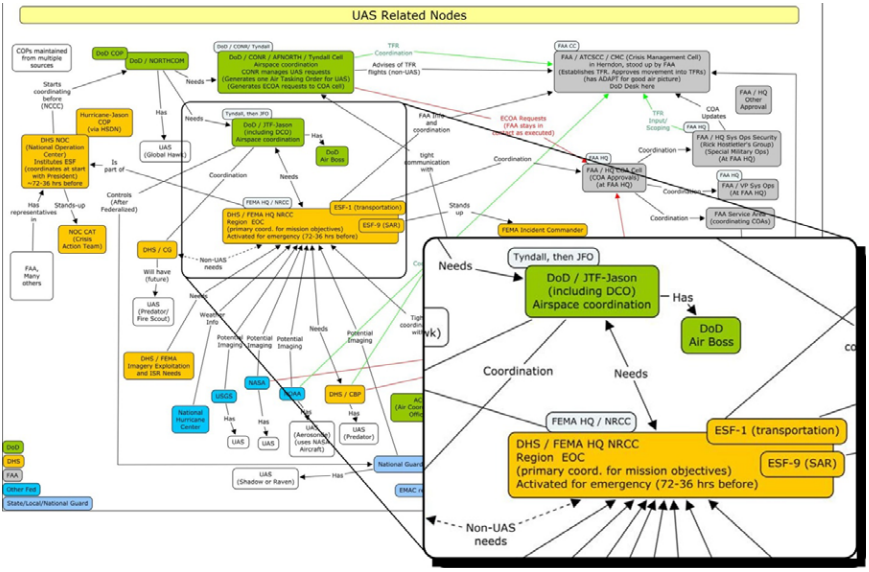

Figure 2. Node Information Flow

The Node Information Flow Model is often rearranged and simplified to illustrate a part of the experiment space.

Figure 3 provides an example of an *Information Sharing Model*. This diagram includes external sources of information from the experiment control cell/white cell, such as situation reports (SITREPS). The boxes are information systems used by experiment participants, and the arrows are the information passed into/between those systems.

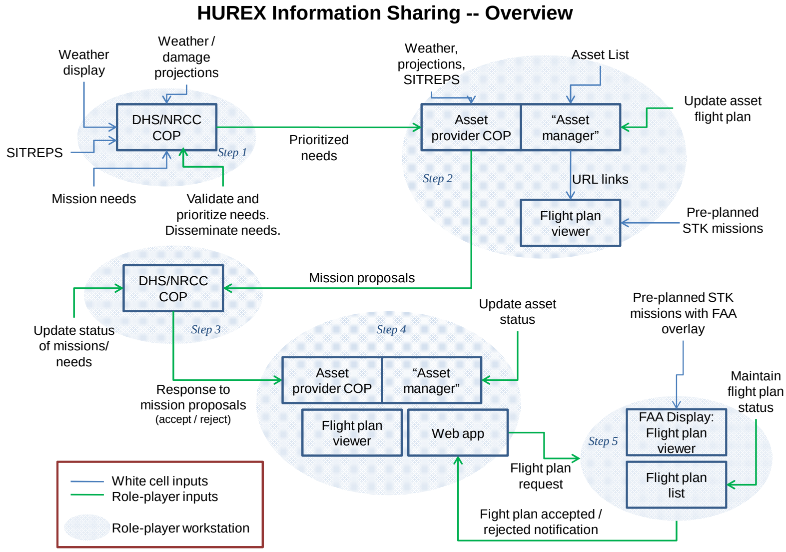

Figure 3. Information Sharing

For each experiment, the team also created an overall concept of experiment diagram that illustrated the key nodes and functions, shown in Figure 4.

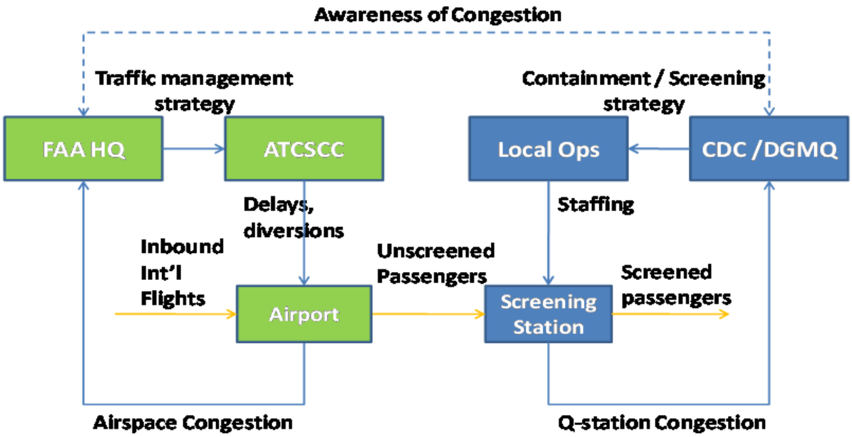

Figure 4. Experiment Concept

Once enough domain knowledge has been acquired and the experiment objectives have been formed, the team may use information from the Node Information Flow Model to create a *Participant Selection Model*. The purpose of this model is to show key relationships between hypotheses, operational concepts, and participants and determine which participants from the real world are required and which roles may be played by others. Figure 5 depicts operational concepts and participants for the in-flight security incident experiment, where the ovals are key multi-agency functions and arrows are associated with organizations, or sub-organizations, expected to participate in those functions.

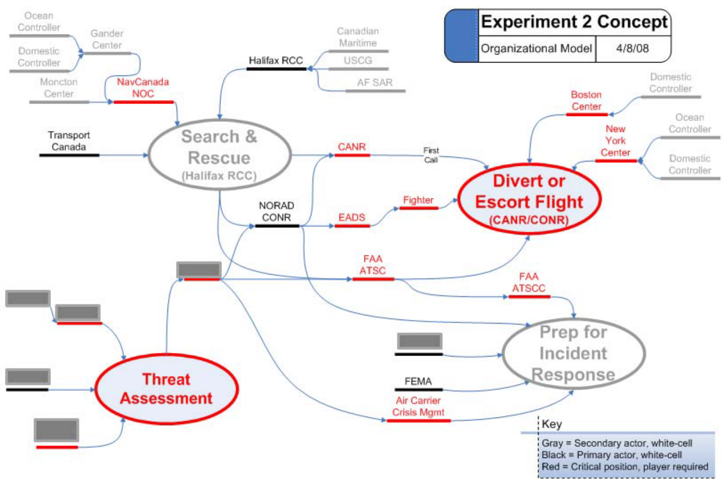

Figure 5. Participant Selection

Event Response Models

With the relationships and communications between the participants well understood, models are built to establish the effects of experimental injections on the participants. *Event Response Models* explore, for each potential inject, the actions the participants are likely to perform. They indicate which roles are played by the white cell, information necessary to perform the injection, potential actions that may result, and over what medium (e.g., phone, instant messenger, air traffic rerouting, collaboration tool) the inject and subsequent communications will occur. They may indicate measurement points or other parameters important to the experiment. Figure 6 is an example Event Response Model representing interactions that occur for an isolated ill passenger in the Pandemic Influenza Experiment (NETEX 09-02).

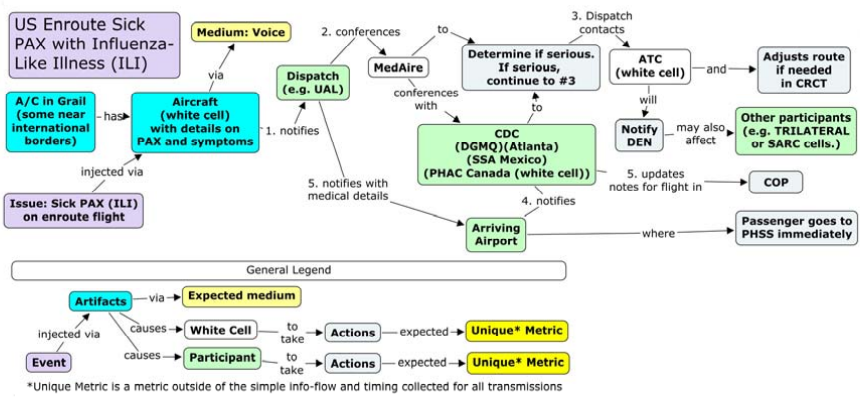

Figure 6. Event Response

Often, the experiments require participants to follow a chain of events in order to formulate an overall response. For these, a *Cause and Effect Model* (also known as an Ishikawa or fishbone diagram) is another type of Event Response Model, intended for planning experiment injects and exploring potential responses. An example from NETEX 08-02 is shown below in Figure 7, where multiple causes (arrows) combine to produce each expected effect (boxes).

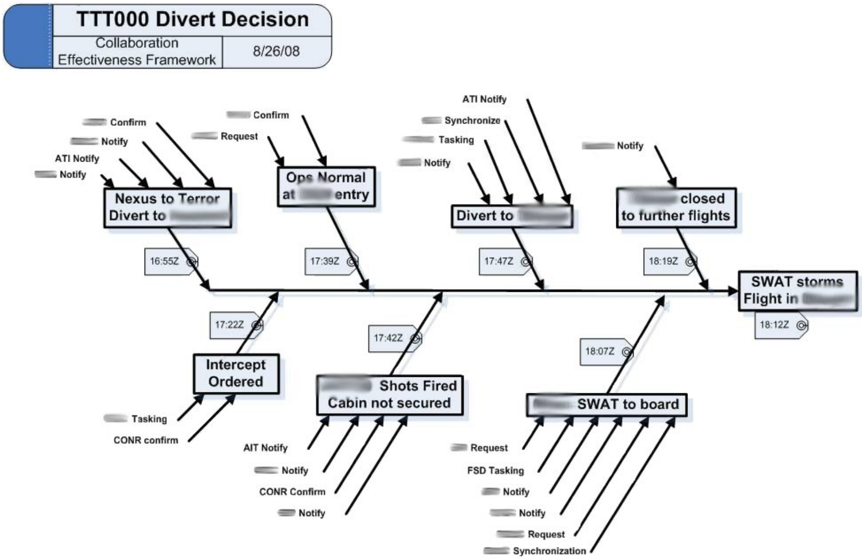

Figure 7. Cause and Effect

Scenario Models

Scenario Models form the timeline from which vignette scripts are written. Typical uses of the term scenario model can refer to mathematical models reflecting a system. In CEE, however, scenario models reflect all injects into the scenario and the actions expected from those injects. For scenarios involving an adversary, scenario models include a back-story, intelligence that was made up for the experiment, and other necessary background information.

To build each Scenario Model, the Event Response Models are placed in the order in which they will be executed for each experiment phase. The execution of these injects is timed to allow participants the ability to react to events and deal with them in a manner similar to real operations. Figure 8 depicts the overall model that integrates events into the scenario timeline. In this model, nodes (boxes) are actions, injects, or notes, and lines indicate probable flow of actions and interactions between the experiment participants.

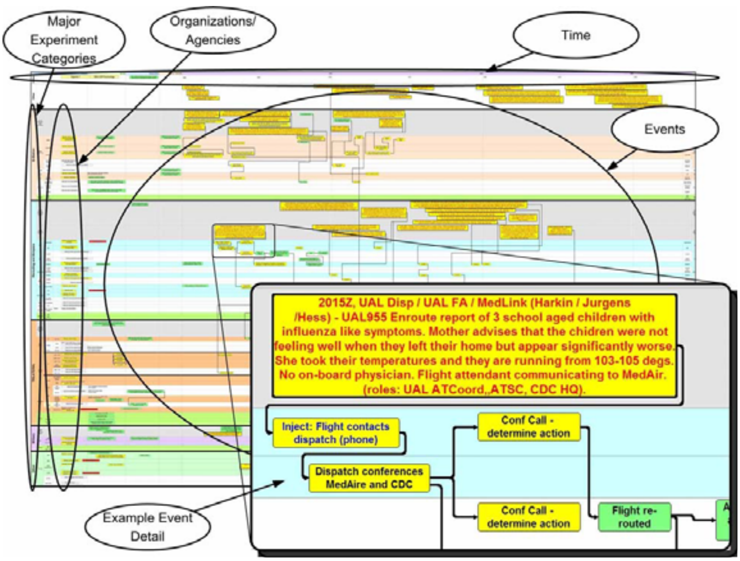

Figure 8. Scenario Timeline

To assist with envisioning the order of actions and underlying ground-truth in the vignettes, domain-specific illustrations are often included to help planners and white cell participants understand what is happening at various points. For NETEX 08-02, the illustrations centered on aircraft location, and on the complex web of terrorist relationships to each other, to other assets, and available intelligence information. Figure 9, an illustration from the HUREX, shows the hurricane track, mission-needs (green boxes), asset locations (pink boxes), and planned injects (red boxes).

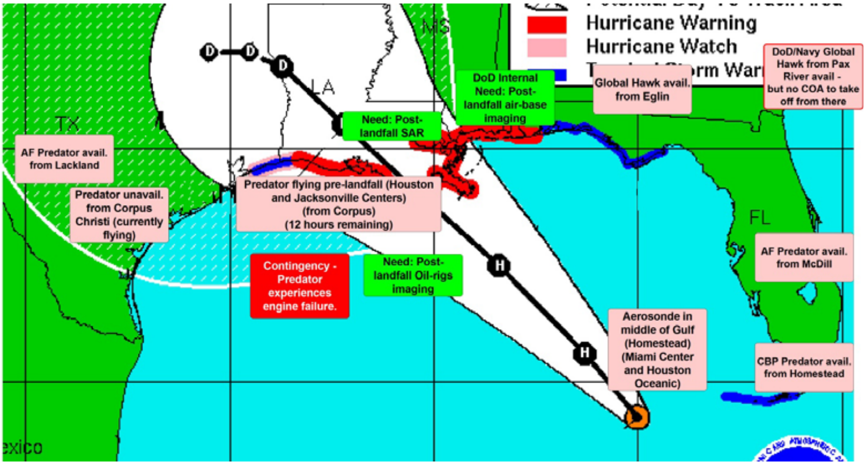

Figure 9. Scenario Illustration

Scenario models and Event Response models also reveal the content and ideal timing of event injects and artifacts needed. NETEX injects may be via voice technology, email, email attachments, text messages, automated systems (such as air traffic control simulations), or prototype tools. HUREX artifacts included hurricane reports from the National Hurricane Center, weather reports from the National Weather Service, and SITREPS from state and local governments.

Experience has shown that rehearsing participant actions, via the scenario models with a *what-if* approach, increases the likelihood that unanticipated, participant-generated events will be handled appropriately by the white cell.

Domain Simulation Models

Epstein (2008) identifies 16 reasons to use *Domain Simulation Models* for purposes other than prediction. The team uses simulation models when domain aspects are: complex, dynamic, or critical to the goals of the experiment and participants need a realistic enough representation of the domain.

We consider if the decision and coordination stimuli will be provided by *scripting* of inputs or not. If not, a search for government or commercial software follows. If no suitable or cost effective simulation models are available, the team considers developing one from scratch. This generally requires the process of interest be discrete enough to be modeled.

Further empirical domain knowledge is important for parameter settings. We conduct literature searches and interviews with experts. This may mean identifying abnormal or extreme values that would not normally occur in order to test some aspect of the hypothesis. An initial reaction from stakeholders/subject matter experts can be "this would never happen;" but, see Kliemt (1996) for a discussion about the utility of "unrealistic" models and "radical" assumptions.

Integrating simulations into the experiment environment often requires significant resources. This is often a high-risk area that should be dealt with as early in the planning and integration process as possible and by experienced simulation engineers.

In experiments conducted to date, we have brought to bear several different kinds of simulation models: spreadsheet-based, numerical computing environments such as MatLab, entity and agent-based, combat and sensor.

We have used combinations of entity/agent-based models and combat and sensor models. For example, in NETEX 08-02, we used the Joint Semi-Automated Forces Simulation (JSAF) and Air Warfare

Simulation (AWSIM) to simulate surveillance radars, and MITRE's GRAIL Real-time ATM Infrastructure Laboratory (GRAIL) simulation to represent airborne aircraft in the National Airspace System (NAS). In the Pandemic Influenza Experiment (PIE), we developed an airport screening model from scratch using the NetLogo agent-based modeling environment.

In our most recent experiment involving airport screening for pandemic influenza, the airport screening model developed by the team played a crucial role in all experiment development phases. The model screenshot in Figure 10 depicts a day's worth of screening activity at an airport and delays throughout screening. The analysis team used it as a post-event constructive simulation (Kass 2006) to answer questions such as: (1) how long can an airport wait when queues build before committing additional screening resources; and (2) what is the minimal amount of additional resources required to significantly reduce overall passenger wait times.

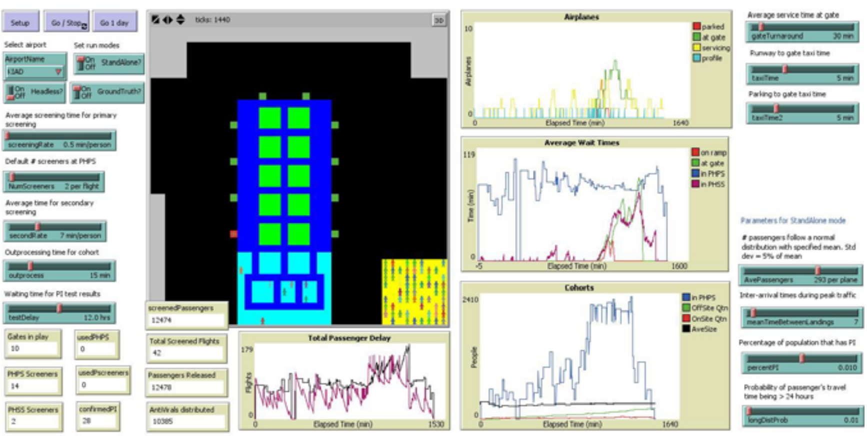

Figure 10. NetLogo Airport Screening Model Screenshot

Data Collection and Analysis Models

As experiment design and preparation unfolds, the data collection and analysis team begins decomposing the higher level models to identify metrics and collection points. Because the CEE focuses on multi-agency collaboration, the relevant indicators are a mixture of qualitative and quantitative measures.

For example, the Experiment Concept Model shown in Figure 4 depicts a control loop describing high-level causal links between information, awareness, actions, and the environment.

Awareness of airspace and airport congestion may trigger strategic or tactical actions which may or may not affect congestion. Awareness, in turn, is hypothesized to depend on the quality of information available—timeliness, relevance, and validity. Determining whether different experimental conditions result in different outcomes thus requires evaluating how the information differs between vignettes (the independent variable) as well as measuring awareness, actions taken, and their results (the dependent variables).

To pinpoint metrics and data collection sites, the team turns to the detailed domain simulation models and the experiment architecture. Simulations often generate important metrics for the system being simulated, although they are not always exposed as output. For example, measures of airspace congestion had to be derived from the airspace simulation. Since they depict information flow and the physical connections between nodes and systems, the models of information flow and experiment architecture are critical for determining whether all important lines of communications have been considered. More generally, models of social networks, information sharing, and communications are useful for developing quantitative measures of collaboration to supplement domain-specific metrics.

Assessing cognitive factors such as awareness, teamwork, or decision making in an environment where participants are allowed to be creative is challenging. Cognitive models and instrumentation techniques abound but some are applicable only for tightly scripted or highly instrumented experiments, while others are too intrusive. The CEE team includes a number of approaches in each experiment, typically building upon those that proved useful in earlier experiments. The metamodel in Figure 11 illustrates key considerations in our data collection planning; Maroney et al. (2009) describe in detail the measurement and instrumentation approach for the HUREX and analytical results. In general, a triangulation approach must be used due to hidden cognitive behaviors.

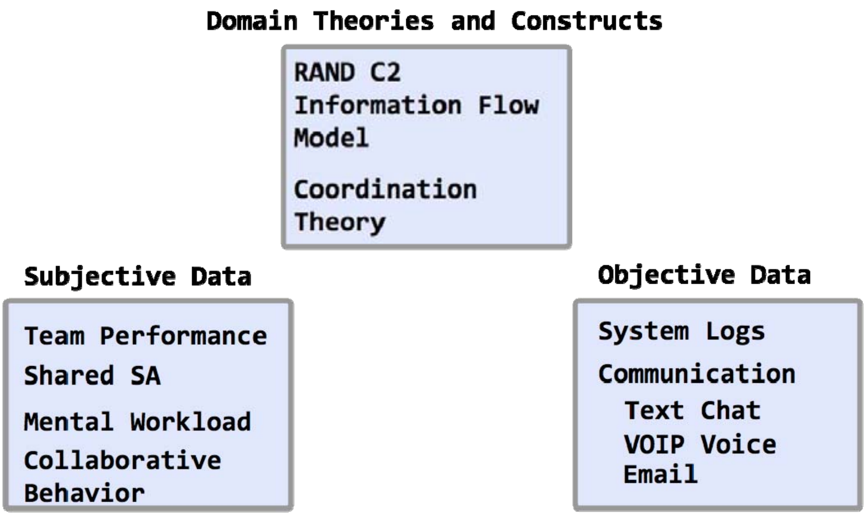

Figure 11. Measurement Metamodel

Architecture Models

Architecture, or *System Models,* represent the systems and components. Standards are available, such as Unified Modeling Language (UML) and Integrated Definition (IDEF), but we typically use simple block diagrams to represent (sub)systems and certain kinds of interactions or data flows. As C2 architecture modeling has received a great deal of attention in the last two decades, particularly in software engineering, we conclude this subsection by providing one of our Architecture Models in Figure 12; which characterizes interprocess communication between system components.

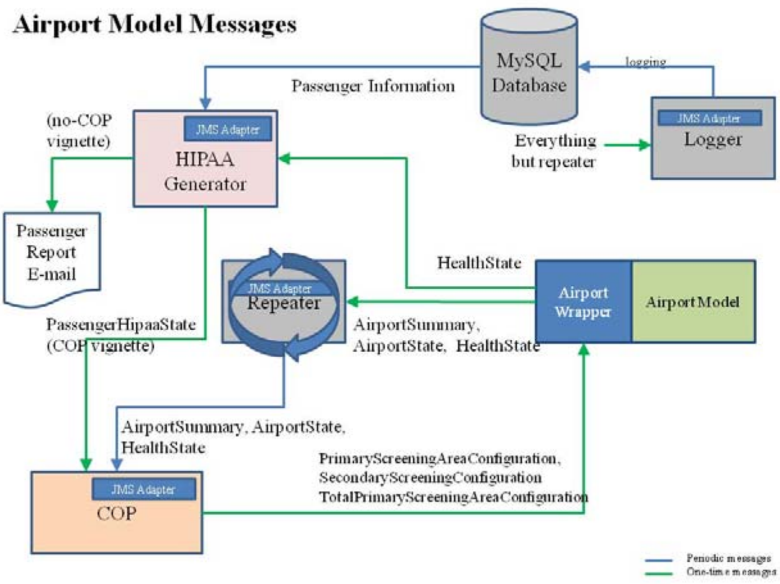

Figure 12. Interprocess Communication

Experiments

In this section, we describe three experiments as cases for how and why models benefitted from experimentation. Each experiment represents different domains, although all three feature the air domain and FAA operations.

NETEX 08-02 In-Flight Security Incidents

In-flight security incidents may result in mission execution requirements for many organizations. In a potential terrorist seizure of an aircraft bound for North America, for example, law enforcement, air traffic control, military air defense, intelligence community, emergency response, commercial air carriers, state and local government, and others may become involved.

In this experiment, it was critical to represent a significant amount of airborne traffic over the North Atlantic and the US and to provide *noise* events for the participants. In part, this was due to the terminal response phase of the incident where aircraft were judged hostile and tactical engagement decisions had to be made. Additionally, the joint collaboration space represented by the Domestic Events Network (FAA 2009) meant that representing the roles and responsibilities of each organization were important to build consensus and solicit participation prior to the experiment.

Since participants are allowed to behave creatively during the experiment, one of the challenges the team faced was to identify likely *and* significant courses of action and decisions participants would take. An outcome of scenario modeling and development is creating events that channel and create situations to meet experiment goals and to provide options if participants deviate from expectations (which usually happens). Participants' creative responses are desirable and sometimes the vignettes unfold in unexpected ways, but the overall goals of the experiment must be met. When experiments involve a thinking adversary—as played by elements of the white cell—event response models and scenario model artifacts are important to have in the experiment control/white cell to allow quick white cell response. Not all branches and sequels can be constructed or considered a priori, but these models can focus response and help the white cell members synchronize when trying to keep experiment objectives on track.

NETEX 09-01 UAS for Hurricane Crisis Response (HUREX)

The HUREX was conducted to evaluate alternative concepts for UAS usage and coordination in a national disaster response. With many possible assets and asset providers, and only one UAS allowed to fly per FAA facility, the experiment examined how to coordinate and optimally allocate UAS assets owned by different chains of command (DHS versus DoD) and how to keep the FAA in the loop (Maroney et al. 2009).

In this experiment, the phenomenon driving the situation was relatively slow compared to human decision-making and time available for each experiment vignette. As a result, the team made three major design choices. First, no domain simulation model would be required; the scenario team would be able to create scripted weather and state/local SITREPS, sensing collection requirements, and assets available; weather, flooding, and airborne inventory did not require explicit simulation. Secondly, to avoid the cost of acquiring and integrating flight planning tools, the team scripted flight plans in order to provide asset providers with enough options to meet collection requirements. Finally, to examine behavior during different phases of a weather disaster—pre-landfall planning, landfall situations, and post-landfall response, each vignette featured a time *jump* where the clock was advanced by 12 to 24 hours, and the team provided participants with new situation updates. Since no simulations were used, only the MITRE-developed prototype decision support tool required clock *advancement*. Time shifting within a vignette when multiple simulations and applications are part of the experiment environment is usually difficult, at best.

CEE often manipulates the nature of information sharing as an independent variable in its experiments. We modified the Information Sharing Model in Figure 3 to create Figure 13 and Figure 14.

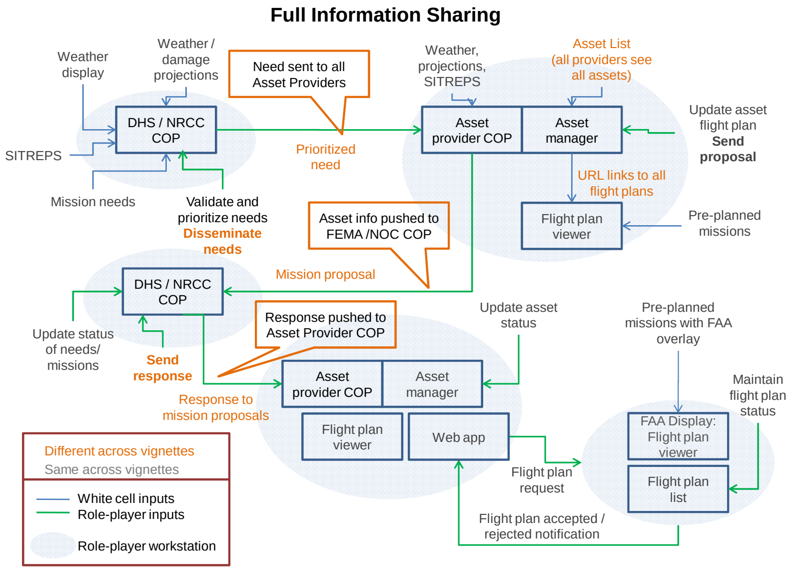

Figure 13. Information Sharing Vignette 1

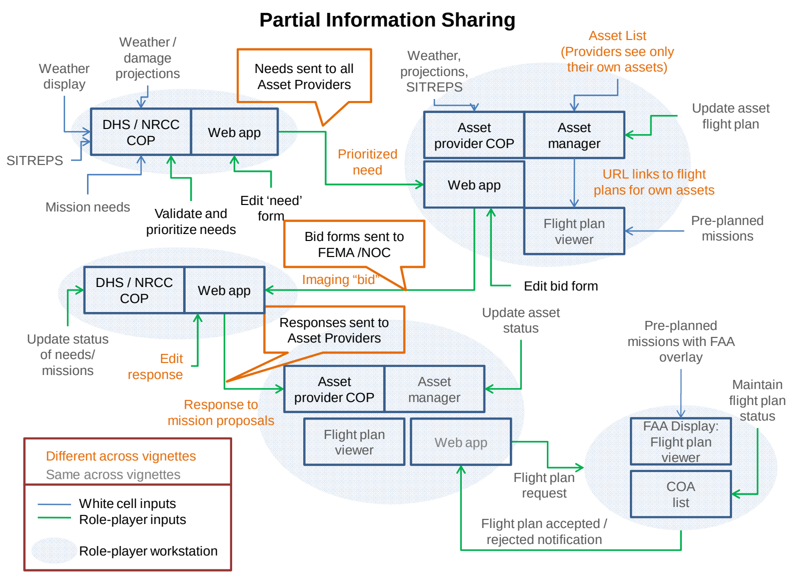

Figure 14. Information Sharing Vignettes 2 and 3

NETEX 09-02 Pandemic Influenza Experiment (PIE)

There are significant potential aviation impacts when executing a Risk-Based Border Screening Strategy (CIDRAP 2008) for detecting and mitigating pandemic influenza; and the US has never implemented a screening process like the one described by Schnirring (2008). Given both of these factors, MITRE implemented an airport screening simulation model. The behavior of this model and its required parameters elicited significant discussion within the health and aviation community experts that the CEE team consulted. It served to clarify requirements and also enable integration with the ground truth model (that provided input for international flight arrivals). A joint research project between MITRE and Research Triangle Institute provided a global disease spread model. This model was used to seed arriving international flights with passengers having the flu (influenza), flu-like illness, and other parameters. GRAIL was used to simulate the NAS for participant displays and to generate arrival information for the airport screening model.

During the experiment itself, the airport screening model provided the ground truth for screening process and state of international flight arrivals at almost 30 North American screening airports. Some passengers would be onboard awaiting screening, some would be in primary screening, and sometimes, a small number would be in secondary screening. The output of the model was fed to a prototype decision support tool that enabled participants to assess delays within and across the airports, predict future delays, and take actions, such as assigning more screeners to a flight in screening.

Finally, the analysis team used the airport screening model to:

- Compare participant behavior with other possible courses of action;

- Identify impacts to overall screening times and delays given various rates of influenza and other illnesses among passengers; and

- Identify if reactive behavior is sufficient to prevent severe delays or if proactive behavior and predictive decision support is necessary.

Lessons Learned and Future Work

Lesson Learned

From the inception of CEE, the team decided to make modeling a central part of the planning, execution, and post-hoc analysis phases of experimentation. We did this because of the fluid nature of multi-agency experiment planning as well as cost and time constraints. Since each experiment can have a different focus and include different domains, we needed reusable tools and techniques. As our concept for the experiment process evolved, we learned about the benefits and limitations of our process.

1. Models have value across experiment lifecycles even if they are designed for only one phase. The clearest example is when post-hoc analysis is conducted. The analysis team had all the modeling artifacts created prior to the experiment and could compare data collected with intended events and processes.

2. Unless descriptive models are created in one tool, there is little to no model *interoperability* or way to automatically link data collected with concepts and relationships created in the modeling tools. For example, there is no way to easily construct a *query* that would trace information created by one role to other roles that should have used that information as defined in a model. For certain kinds of experiments, the ability to quickly construct and execute these queries would make analysis more efficient and might lead to more or refined insights.

3. Models are useful stimulants for eliciting discussion. When a process or concept is concretely described by diagrams, stakeholders and participants become willing to provide clarification, ideas, or pose questions. Enabling interaction between stakeholders is one of the primary objectives of the CEE project.

4. Models do not take the place of active team communication and consensus building. With any written artifact, it is easy for authors to assume that content will be immediately considered and understood. Team members have multiple project responsibilities, are in different locations, and have different responsibilities. Active promotion of shared understanding, via periodic meetings to review content, is still required.

5. Choice of modeling tool matters. For descriptive models, it is often the experience and familiarity with a tool that influences a modeler's choice. One result of this choice is which kinds of descriptive models can be brought together in that tool. For domain simulation models, the impact is potentially more significant due to integration and interoperability requirements. Typically, this choice should involve the integration and architecture team(s).

Future Work

The team seeks ways to rapidly create, modify, and share models with a goal of reducing cost and improving on the ability to meet logical requirements of validity (Kass 2006). When scope or experiment designs change, the work necessary to refine models is unavoidable—this occurred during the latter half of the pandemic influenza NETEX as the novel H1N1 pandemic caused organizations to reevaluate their policies and understandings of novel influenza dynamics. We would like to further capitalize on completed modeling during post-hoc analysis, however, as discussed in the lessons learned. We believe that semantic technologies are a key part of the solution. A complete solution might include:

- Usage of ontologies by all models—the use of standard terms of references (concepts) with associated attributes and relationship types (arcs) in one or more ontologies;

- Data collected and tagged using the same ontologies;

- Intelligent queries on collected data using a semantically-enabled database. The technical language of the queries would be hidden by user interfaces; and

- The ability to pose queries and set alerts in near real-time for data collectors to increase their situation awareness and improve their judgments.

The use of simulation models during the experiment to provide a rich, dynamic environment to the participants takes time, effort, and expertise to bring to bear, no matter how automated the simulation model. Anything that can be done to reduce these costs is desirable.

The palette of techniques for multi-agency C2 experimentation and exploration is significant. Future events might focus on tabletop *heavy* events, where all the players are seated at the same table, have low to medium fidelity domain simulation(s) feeding information displays surrounding them, but still have facilitators stepping through events. Participants would be allowed to ask any questions they wanted and everyone would share in the questions and answers.

Acknowledgements

We are grateful for the expertise and time of our government and industry participants. We also thank our CEE teammates for their insights, as well as MITRE experts who lent their advice during planning and performed white cell roles during the experiments. This work was funded by internal MITRE research and development.

References

Alberts, David S., R. E. Hayes, J. Kirzl, D. Maxwell, and D. Leedom. 2002. *Code of Best Practice for Experimentation.* Washington, DC: CCRP. <http://www.dodccrp.org/files/Alberts_Experimentation.pdf>

Alberts, David S., and R. E. Hayes. 2005. *Campaigns of Experimentation*. Washington, DC: CCRP. <http://www.dodccrp.org/files/Alberts_Campaigns.pdf>

Epstein, Joshua. 2008. Why Model? *Journal of Artificial Societies and Social Simulation* 11(4)12. <http://jasss.soc.surrey.ac.uk/11/4/12.html>

Federal Aviation Administration (FAA). 2009. Domestic Events Network (DEN) and Presidential/United States Secret Service (USSS)-Supported Very Important Person (VIP) Movement. *Notice NJO 7210.724*. <http://www.faa.gov/documentLibrary/media/Notice/N7210.724.pdf>

Kass, Richard A. 2006. *The Logic of Warfighting Experiments*. Washington, DC: CCRP. <http://www.dodccrp.org/files/Kass_Logic.pdf>

Kliemt, Hartmut. 1996. Simulation and Rational Practice. In *Modeling and Simulation in the Social Sciences from the Philosophy Point of View*, eds. Rainer Hegselmann and Ulrich Mueller, 13-28. Dordrecht: Kluwer Academic Publishers.

Maroney, David R., M. Harkin, L. Litwin, J. Curtiss, A. Bigbee, R. Flournoy, and V. J. Gawron. 2009. HUREX – Experiment of Use and Prioritization of UAS in Hurricane Disaster Response. Paper presented at the *9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*, September 21-23, Hilton Head, SC.

Military Operations Research Society (MORS) Experimentation Community of Practice. 2008. *Experimentation Lexicon,* Version 1.0. <http://morsnet.pbworks.com/f/Experimentation+Lexicon+V1.1.doc>

Schnirring, Lisa. 2008. Feds Stage Airport Test of Plan to Slow Pandemic. *Center for Infectious Disease Research and Policy, CIDRAP News*, November 12. University of Minnesota, Minneapolis. <http://www.cidrap.umn.edu/cidrap/content/influenza/panflu/news/nov1208airport.html>

Trochim, William. M. 2006. *The Research Methods Knowledge Base,* 2nd ed. [online version current as of October 20, 2006]. <http://www.socialresearchmethods.net/kb/>

United States Government Accountability Office (GAO). 2007. *Aviation Security: Federal Coordination for Responding to In-flight Security Threats Has Matured, but Procedures Can Be Strengthened*. GAO-07-891R. <http://www.gao.gov/new.items/d07891r.pdf>