Spring Simulation MultiConference

WIZER: A Tool for Automating Model Improvement in Multi-Agent Social-Network Systems

Journal:	*Spring Simulation Multiconference*
Manuscript ID:	SMC-2005-ADS-010.R1
Topic Area Conference:	Agent-Directed Simulation
Submission Type:	Final Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Yahja, Alex; Carnegie Mellon University, Institute for Software Research International Carley, Kathleen; Carnegie Mellon University, Institute for Software Research International
Keywords:	knowledge-based search, probabilistic argumentation causal inference, model improvement, simulation control, validation

WIZER: A Tool for Automating Model Improvement in Multi-Agent Social-Network Systems

Alex Yahja and Dr. Kathleen M. Carley
Computation, Organizations, and Society Program, Institute for Software Research International,
Center for the Analysis of Social and Organizational Systems, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
ay@cmu.edu, kathleen.carley@cmu.edu

ABSTRACT

There has been a significant increase in the use of multi-agent social-network models due to their ability to flexibly model emergent behaviors in complex socio-technical systems while linking to real data. These models are growing in size and complexity which requires significant time and effort for calibration, validation, model-improvement, and to gain understanding as to why the models behave as they do. In this paper, we present our knowledge-based simulation-aided approach for automating model-improvement and WIZER, our tool implementing this approach. WIZER is capable of calibrating and validating multi-agent social-network models, and facilitates model-improvement and understanding. By employing knowledge-based search, causal analysis, and simulation control and inference techniques, WIZER can reduce the number of simulation runs needed to calibrate, validate, and improve a model and improve the focus of these runs. We ran a preliminary version of WIZER on BioWar – a city-scale social agent network simulation of the effects of weaponized biological attacks on a demographically-realistic population within a background of naturally-occurring diseases. The results demonstrate the efficacy of WIZER.

Keywords: knowledge-based search, probabilistic argumentation causal inference, model improvement, simulation control, validation

1. INTRODUCTION

Currently, a sea change is occurring in how we model and think about knowledge, individuals, teams, groups, networks, organizations, markets, institutions, and other societal systems due to developments in the field of computational modeling and analysis [Axelrod 1997][Carley and Prietula 1999][Epstein and Axtell 1996][Prietula et al., 1998][Taber and Timpone 1996][Gilbert and Troitzsch 1999][Ward 1985]. Computational modeling and analysis is becoming a useful scientific tool for addressing socio-technical problems which have complex interrelated dynamic parts. Societal problems such as natural disaster response and biological attacks are complex and do not happen in a vacuum but rather within a complex context of social, organizational, geographical, technological, regulatory, and other constraints.

There has been a rapid increase in the use of multi-agent models [Lucena et al., 2004] – as well as social network analysis [Wasserman and Faust 1994] – to address complex socio-technical problems. Model assessment – determining how valid, how explainable, and how robust a model is – is becoming a major concern [Cioffi-Revilla 2002]. Indeed, identifying reliable validation methods for complex systems such as electronic medical surveillance systems is a critical research area [Reifman et al., 2004]. Calibration and validation serve as a foundation for model improvement through simulation and inference.

Models contain both explicit and implicit assumptions about some portion of the real world. These assumptions form abstractions of reality and these abstractions may or may not be sound. Moreover, the real world changes continuously and in unexpected ways. The required fidelity of the model varies as a function of the research and/or policy questions being asked. Calibration, validation, and model-improvement are hard due to the changes in the real world, altered goals, inherent assumptions and abstractions, and human cognitive limitations (such as biases, bounded rationality [Simon 1957], and memory capacity).

Information exploitation is a technique that has yet to be fully employed to deal with the problem of calibration, validation, and model improvement. (The term “validation” will be used from now on to denote calibration, validation, and model-improvement.) Few multi-agent simulations have exploited the depth and breadth of available knowledge and information for validation that resides in journals, books, websites, human experts, and other sources. Simulations results are also normally designed solely for human analysis. While this may be sufficient for small-scale simulations, for large-scale simulations, automated help for validation and analysis is crucial. Most agent-based toolkits focus on providing a programming environment to build and execute the model; little work to date probes the important aspect of automating validation and analysis (this is conventionally left to humans to perform). To successfully automate validation and analysis, domain knowledge must be exploited, for example by an expert system inference engine. A simulation and inference engine that can do virtual experiments and knowledge inference in concert would facilitate focused search by using both the simulation engine's search space and the inference engine's knowledge space to arrive at better parameter and meta-model values for validation. This paper describes our approach for doing knowledge-based simulation-aided validation in multi-agent social-network systems, embodied in a tool called WIZER (**W**hat-**I**f Analy**ZER**). WIZER applies knowledge control of the simulation, inference and intelligent search in multi-agent social-network simulations.

The results presented in this paper are based on WIZER runs using BioWar. BioWar is a city-scale multi-agent social-network simulator capable of modeling the effects of weaponized biological attacks on a demographically-realistic population within a background of naturally-occurring diseases [Carley et al., 2003] [Carley et al., 2004]. BioWar currently runs a few thousand to several million agents. Unlike traditional models that look at hypothetical cities (such as the Brookings' smallpox model [Epstein et al., 2004] and the SARS model [Huang et al., 2004]), BioWar is configured to represent real cities by loading census data, school district boundaries, etc. It models both healthy and infected agents as they go about their lives, enabling observation of absenteeism, drug purchases, hospital visits, and other data streams of interest.

2. PROBLEM STATEMENT

Today, multi-agent social-network systems are increasingly employed in modeling due to their power, flexibility, and ability link to real data. These models are growing in size and complexity, resulting in a significant increase in the time and effort required for calibration, validation, and to understand why the models behave as they do. An automated and intelligent tool is needed that can be used to calibrate such models and to facilitate validation and model-improvement, thereby increasing model fidelity and freeing user time for policy analysis.

3. RELATED WORK

Multi-agent systems are usually "validated" by strictly applying requirements engineering. In software engineering terms [Pressman 2001], validation means the determination of the correctness of the final program or software produced with respect to the user needs and requirements – not necessarily the empirical data or the real world. Formal methods [Dershowitz 2004] used in software engineering for control and understanding of complex multi-agent systems lack an effective means of determining if a program fulfills a given formal specification [Edmonds and Bryson 2004]. Complex societal problems contain "messy" interactions, dynamic processes, and emergent behaviors, so it is often problematic to apply requirements engineering and/or formal methods.

Another validation method is evolutionary verification and validation or EVV [Shervais et al., 2004][Shervais and Wakeland 2003], which utilizes evolutionary algorithms, including genetic algorithms and scatter search, for verification and validation. While EVV allows testing and exploitation of unusual combinations of parameter values via evolutionary processes, it employs knowledge-poor genetic and evolutionary operators, not the scientific method, for doing experiments, forming and testing hypotheses, refining models, and inference, precluding non-evolutionary solutions.

Docking is another approach to validating multi-agent systems [Axtell et al. 1996]. Docking is based on the notion of repeating a scientific experiment to confirm findings or to ensure accuracy. It considers whether two or more different simulation models align (produce similar results), which is used in turn as a basis to determine if one model can subsume another. The higher the degree of alignment among models, the more they can be assumed to be valid, especially if one (or both) of them has been previously validated. The challenges in applying docking are the limited number of previously validated models, the implicit and diverse assumptions incorporated into models and the differences in data and domains among models.

One application of docking is to align complex multi-agent simulations against mathematical or system dynamics models. BioWar's anthrax simulation has been successfully docked against the Incubation-Prodromal-Fulminant (IPF) mathematical model, a variant for anthrax of the well-known Susceptible-Infected-Recovered (SIR) epidemiological model [Chen et al., 2003] and BioWar's smallpox model has been docked against a SIR model of smallpox [Chen et al., 2004]. While aligning a multi-agent model with a widely used mathematical model can show the differences and similarities between these two models, the validity is limited by the type of data the mathematical model uses. For example, the IPF model mentioned above operates on population-level data, so the result of the alignment will be only valid at the granularity of population-level data. Mathematical models also have difficulties representing non-numerical (symbolic) knowledge, including the knowledge base underlying complex context-sensitive agent interactions.

Validating complex multi-agent simulations by statistical methods alone [Jewell 2003] is problematic due to the coarse granularity required for statistical methods to operate properly and the insufficient representation of symbolic knowledge. Statistical methods are good at describing data and inferring distributional parameters from samples, but statistic methods alone are insufficient to handle the highly dynamic, symbolic, causal, heterogeneous, and emergent nature of societal systems.

Complex multi-agent simulations are not normally validated using expert systems (such as OrgCon [Burton and Obel 1998]) as it is thought that it is sufficient to let human experts alone perform the analyses, experiment design, and quantitative and symbolic reasoning. This view is especially prevalent as most simulations are in the realm of purely numeric models.

Human experts can do validation by focusing on the most relevant part of the system and thinking about the problem intuitively and creatively. These subject matter experts (SMEs) have the knowledge needed to judge model performance in their specialized fields. Applying learned expertise and intuition, SMEs can exploit hunches and insights, form rules, judge patterns, and analyze policies. Managed and administered properly, SMEs can be effective. Pitfalls include bounded rationality, implicit biases, implicit reasoning steps, judgment errors, and others.

Another approach to validation is direct comparison with real world data and knowledge. Validation can be viewed as experimenting with data and knowledge, using models as the lab equipment for performing computational experiments [Guetzkow et al., 1972][Bankes 2004]. Simulation models to be validated should reflect the real world and results from experiments in simulation should emulate changes in the real world. If results from virtual or computational experiments are compared to real world data and match sufficiently, the simulation is sufficiently valid. Simulation [Law and Kelton 2000][Rasmussen and Barrett 1995] has an advantage over statistics and formal systems as it can model the world closely, free of the artifacts of statistics and formal systems.

There is related work in engineering design methods using Response Surface Methodology or RSM [Myers and Montgomery 2002] and Monte Carlo simulations [Robert and Casella 1999] to do direct validation, but only with numerical data and limited to a small number of dimensions. RSM is collection of mathematical and statistical techniques (e.g., gradient descent search) for the modeling and analysis of problems in which a response of interest is influenced by several variables. It can include virtual experiments using Monte Carlo simulation. It usually tests only a few variables and operates to find the best fit equation so that the correlation of equation's predictions with real data is statistically significant.

4. OUR APPROACH: KNOWLEDGE-BASED SIMULATION-AIDED MODEL-IMPROVEMENT

WIZER (**W**hat-**I**f Analy**ZER**) is a coupled inference and simulation engine that extends the Response Surface Methodology to deal with the high dimensional, symbolic, stochastic, emergent, and dynamic nature of multi-agent social-network systems. Viewing simulation systems as knowledge systems, WIZER is designed for controlling and validating them directly with empirical data and knowledge using pattern analyses and knowledge inferences (mimicking those of SMEs) and virtual experiments (mimicking those of RSM).

WIZER integrates an inference engine and simulation virtual experiments to do calibration and validation for model-improvement and to provide explanations. It improves on RSM features by performing knowledge-intensive data-driven search steps via an inference engine constrained by simulation outputs, instead of just doing statistical and mathematical calculations. WIZER facilitates knowledge-based simulation control and simulation-assisted inference, enabling reasoning about simulations and simulation-assisted reasoning. It enables the management of model assumptions, contradictory or incomplete data, and increases the speed and accuracy of model validation and analysis. It is capable of explaining the reasoning behind inferences using both the simulation and inference engine. Searching in WIZER is performed using both simulation and knowledge inferences. The amount of searching is reduced as the knowledge inferences, empirical data and knowledge, and virtual experiments constrain the search space.

WIZER seeks to emulate scientists doing experiments and analyses via the scientific method, instead of simply emulating an experimental setup. While other toolkits such as Swarm (http://wiki.swarm.org) and Repast (http://repast.sourceforge.net) are designed with the goal of assisting the design and implementation of agent-based simulations, WIZER is designed to help with scientific experimentation, validation, analysis, and model improvement. WIZER is conceptually able to run on top of any simulation system, including those crafted using Swarm and Repast toolkits. WIZER is basically a logical reasoning, experimentation, and simulation control engine with statistical and pattern recognition capabilities. This is similar to techniques scientists employ when designing, executing, and analyzing experiments. WIZER differs from Evolutionary Programming [Fogel 1999] as it does not need a population of mutation candidates and the mutation operator. Instead, WIZER applies knowledge inference to simulations to design the next simulation run, based on scientific experimental method. If the result of inferences mandates a radical change, a revolution would occur. WIZER also differs from Evolutionary Strategies and Genetic Algorithms [Dianati et al. 2003] as it does not use recombination/crossover operators. In short, WIZER employs a unique logical reasoning, simulation control and scientific method for doing virtual experiments. WIZER emulates what scientists are doing to advance science.

As shown in Figure 1, WIZER includes Alert WIZER and the WIZER Inference Engine. Alert WIZER determines which data streams of the simulation outputs do not fall within the empirical data value ranges and how.

The WIZER Inference Engine takes the simulator's influence diagram of what parameter influences which output data and the empirical constraints and confidence intervals on parameters to make a judgment on which parameters to change and how (including causal links and the model or agent submodel itself, if necessary). This results in new parameters for the next simulation. This simulation in turn yields new outputs which are fed into Alert WIZER. This cycle repeats until a user-defined validity level is achieved.

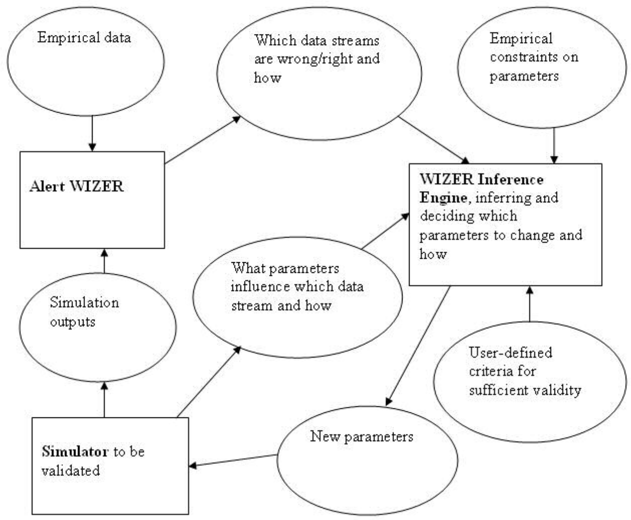

Figure 1. WIZER System Diagram. Rectangles are components or processes and ovals are data. The arrows denote the flow of data in the direction of the arrowhead.

In other words, WIZER consists of:

- A system for determining which outcome variables match or fall within the acceptable range of the real data – Alert WIZER. This system will create an "alert" when there is not a match. Inputs to Alert WIZER include real and virtual data. Real data include various types of data such as subject matter experts' (SMEs) estimation of behavior, 1^{st}, 2^{nd}, and 3^{rd} order statistics for data streams at the yearly, seasonal, monthly, and day of week level, and actual streams of data. Alert WIZER also includes statistical tools.
- An intelligent system for identifying which of the "changeable" parameters should be changed and how to improve the fit of the virtual to the real data – the WIZER Inference Engine. This component uses a database relating parameters to the variables and modules they impact. This includes assumptions about the expected range for parameter values (according to SMEs) or best guesses, thus placing confidence measures on parameters.
- A local response surface analysis feature that can run simple virtual experiments for parametric studies.

The knowledge bases in the inference engine are populated with the knowledge about the simulator, simulation outcomes, domain facts and knowledge, assumptions, ontology, problem solving strategies, information about statistical tools it employs and other data. The knowledge bases contain both knowledge (hard or certain rules and facts) and assumptions (soft or uncertain rules and facts). Simulation outcomes provide measurements of the degree-of-support an assumption has. These different types of knowledge are included to enable the inference engine to reason about its reasoning. For example, knowledge about the simulation allows the inference engine to back up its symbolic reasoning with simulation outcomes and also to reason about the simulation. Part of the knowledge base is portable between simulations, but users need to provide the remainder.

The emergence of causal links based on low-level interactions can be probed by the inference engine, including probes to see what an individual agent does in its life and what events affected this agent and why, in addition to sample based probes. For sample based probes, WIZER conducts inferences based on the application of its included statistical tests.

The WIZER Inference Engine was inspired by the rule-based Probabilistic Argumentation Systems (PAS) [Haenni et al., 1999] for handling assumptions. While a rule-based system is sufficient if knowledge engineers are able to check the causal relations inherent in rules, for large knowledge bases manual checks are cumbersome and prone to errors. Thus, there is a need for automated and formal causal checking. Fortunately, causal analysis has been treated mathematically [Pearl 2003]. WIZER uses a novel probabilistic argumentation causal system (PACS), which utilizes the probabilistic argumentation [Haenni et al., 1999] in causal analysis [Pearl 2000]. Users of WIZER specify which rules are causal in nature and WIZER is capable of suggesting causal links and performing empirical computations to provide justification for these causal links. Results from social network analysis form one silo of domain knowledge fed into the WIZER inference engine. The inference engine in turn, along with the execution of virtual experiments in simulations, provides knowledge-based grounding for the emergence and evolution of social networks from low-level agent behaviors and interactions. The causal mechanisms encoded in WIZER enable formal computation of interventions or actions, instead of mere observation. This allows WIZER to make changes in parameters, causal links, and meta-models, and to analyze the consequences. In other words, WIZER can emulate what scientists do by changing and analyzing experiments.

Causal analysis involves mechanisms (stable functional relationships), interventions (surgeries on mechanisms), and causation (encoding of behavior under interventions). Associations common in statistics can characterize static conditions, while causal analysis deals with the dynamics of events under changing conditions. Simply turning off some potential causal links and resimulating is insufficient and while counterfactual testing – checking would happen if (true) facts were false – can uncover causal effects, the method can fail in the presence of other causes or when other causes are preempted and it ignores the sufficiency aspect. These weaknesses of this (global) counterfactual test can be addressed by sustenance, providing a method to compute actual causation [Pearl 2000]. Sustenance means minimally supporting an effect. Actual cause is computed by constructing causal beams and doing local counterfactual test within the beams. Causal beams are causal links that have been pruned to retain a subset of causal links that sustains the occurrence of an effect. Dynamic beams are simply causal beams with a time dimension [Pearl 2000].

To account for the probability of causation, the causal model [Pearl 2003][Pearl 2000] specifies the use of Bayesian priors to encode the probability of an event given another event. It does not distinguish between different kinds of uncertainty. It is unable to model ignorance, ignores contradictions and is incapable of expressing evidential knowledge without the use of the probability distribution format. Since the intended use of WIZER is to do validation in environments with incomplete, contradictory, and uncertain knowledge and because WIZER needs to clearly delineate between assumptions and facts, we need an improved causal model, built by borrowing concepts from the Probabilistic Argumentation Systems (PAS). Table 1 shows the encoding of facts, assumptions, and rules for rule-based systems using probabilistic argumentation, while Table 2 shows the encoding of facts, assumptions, and causations for causal analysis enhanced with PAS-like assumption management. In both tables, let P_i be proposition *i, a_i* be assumption *i*, ***caused*** be the causation operator, and **=>** be the implication operator.

Table 1. Rule-Based Encoding

Type of Knowledge	Logical Representation	Meaning
A fact	P1	P1 is true
A rule	P1 => P2	P1 implies P2
An uncertain fact	a1 => P1	If assumption a1 is true, then P1 is true
An uncertain rule	a2 => (P1 => P2) or equivalently P1 ∧ a2 => P2	If assumption a2 is true, then P1 implies P2

Table 2. Causation Encoding

Type of Knowledge	Logical Representation	Meaning
A fact	P1	P1 is true
A causation	P1 caused P2	P1 caused P2
An uncertain fact	a1 => P1	If assumption a1 is true, then P1 is true
An uncertain causation	a2 => (P1 caused P2)	If assumption a2 is true, then P1 caused P2

We call Table 2's formalism the probabilistic argumentation causal systems (PACS). WIZER includes both rule-based and causal formalisms. PACS algorithmic details are derived from both PAS [Haenni et al., 1999] and causal analysis [Pearl 2003]. Simulation virtual experiments can be seen as a proxy for real world experiments when doing real world interventions would be unrealistic or unethical. Causal analysis uses computations based on real-world experimental and non-experimental data. WIZER adds another dimension to causal analysis: allowing quasi-experimental – that is, simulated – data. Additionally, WIZER enhances PACS by

- Providing a mechanism to arrive at probability distributions or profiles for assumptions related to causations.
- Automating causal analysis for simulations and enhancing it with virtual experiments. In particular, WIZER improves upon dynamic beams [Pearl 2003] by doing virtual experiments, and allows the estimation of sufficiency [Pearl 2003] by virtual experiments.
- Utilizing simulations or virtual experiments as a proxy of the real world when evaluating interventions and using causal beam calculations to uncover true causal relations. WIZER can modify causal links and infer missing ones.
- Allowing better inference by letting the inference engine run simulations (e.g., supports for dynamic beam) in the midst of causal inferences as needed. This allows the examination of the empirical claims of causal inferences.

The internal workings of the WIZER Inference Engine are complex, but its basic operations are simple. Let $P = \{p_1, \ldots, p_n\}$ be propositions, $A = \{a_1, \ldots, a_n\}$ be assumptions, h be the hypothesis and $K = c_1 \wedge c_2 \wedge \ldots \wedge c_n$ be the knowledge base of clauses, where c_i is an element of the set of all possible A and P clauses.

Let α be the (conjunctive) arguments supporting h

We have

$$\alpha \wedge K \models h$$

or equivalently

$$\alpha \models \sim K \vee h$$

or equivalently

$$\sim(\sim K \vee h) \models \sim\alpha$$

$$K \wedge \sim h \models \sim\alpha$$

In other words, if we know K and h, we can compute the supports, that is, the arguments supporting h. The hypothesis h is a clause produced by Alert WIZER after comparing simulation data streams with empirical data. After finding the arguments supporting h, the degree of support can be found, defined as

$$dsp(h, K) = prob(\text{a support } \alpha \text{ of } h \text{ is valid} \mid \text{no contradiction}, K)$$

Similarly, the degree of plausibility can be found, defined as

$$dpl(h, K) = prob(\text{no support of } \sim h \text{ is valid} \mid \text{no contradiction}, K)$$

These two measures are used to determine which arguments are the most relevant to the hypothesis at hand, pinpointing which parameter values, causal links, and/or submodels should be changed. In other words, hypothesis h is the input to WIZER Inference Engine and the arguments supporting h are the output, leading to changes in parameter and meta-model values.

The operations described above are performed for both rule-based and causal clauses. Then, for clauses denoted as causal, additional operations are performed to see whether and to what degree the causal relations are empirically correct, partially based on the degree of support and the degree of plausibility. Sustenance, causal beams and actual cause are also computed.

The causal computation capability of WIZER is useful in simulations to:

- Provide a formal computational means to convert simulation results or happenings to user-friendly causal sentences.
- Allow examination and perhaps modification of the implicit and explicit causal assumptions and links in the incomplete and often erroneous real world picture employed by simulation models, based on empirical data. For example, a model of biological attacks on a city should include empirical data for the city and realistic causal relations.
- Allow probing of potential causal links and examination of the robustness of causal links using empirical data and quasi-experimental data obtained by simulations based on other known mechanisms and data values. For example, a simulation may have modeled Washington DC and policy analysts would like to know the effects of quarantining certain city blocks or closure of some major roads to mitigate the spread of smallpox. The mechanisms, data values, and stochastic processes in the city model themselves do not contain direct answers to the above causal question. Utilizing causal computation would allow this question to be answered – based empirical data and quasi-experimental/simulation data.
- Allow the formal modeling of interventions in simulations. Again, if city officials want to close some major roads to mitigate the spread of smallpox, the effects of this action could be examined by causal computation. Furthermore, since human policy language is often vague, coarse, and incomplete. WIZER can help refine policies computationally. Evaluation of policy using statistics alone is unsound, as statistical methods observe the outcome but cannot capture the intervention policy itself.
- Allow symbolic values to be considered in determining causal relations. For example, the recent shortage of flu vaccine caused the CDC to recommend restrictions on who received the vaccine. The effect of this policy is that CDC now has a stockpile of unused flu vaccine. One cause of this outcome is that people who were eligible to receive shots did not do so, as they believed that none were available. WIZER would be able to probe similar kinds of cause and effect relationships.
- Allow experimentation and simulation control. As WIZER modifies, runs, re-modifies, and re-runs simulations, it uses causal mechanisms to keep track of and help inform what causes a certain series of modifications to work or fail and to suggest possible next steps.

5. RUN SETUP AND EMPIRICAL DATA

WIZER was used to validate BioWar. As mentioned earlier, BioWar [Carley et al., 2003] is a city-scale spatial multi-agent social-network model capable of bioattack simulations. BioWar has a large number of variables and interactions. Application of the Spiral Development model [Boehm 2000] to BioWar code development means that any previous validation of model predictions may no longer apply to a new version. Figure 2 shows the partial causal relationships among entities in BioWar used in WIZER Inference Engine. Note that the arrow direction implies "may cause or influence".

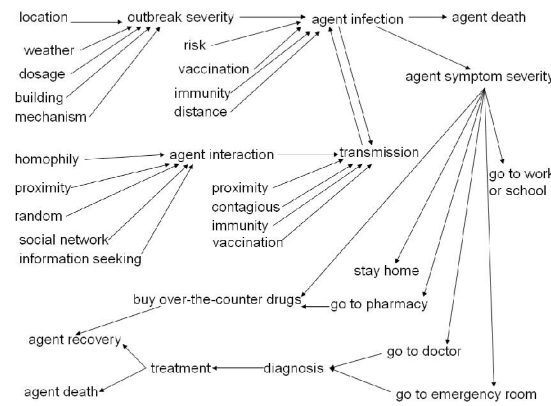

Figure 2. Partial Causal Diagram for BioWar

We have implemented Alert WIZER, which takes the empirical data on school absences, workplace absenteeism, doctor visits, emergency room visits, with additional emergency room visitation data from SDI (Surveillance Data Inc.), and over-the-counter drug purchase data. It also uses the outputs of the BioWar simulator and conducts minimum bound checking, maximum bound checking and mean comparison.

The following empirical data was used to compute the empirical bounds and means for the Alert WIZER:

- NCES Indicator 17, 2002 (Year 2000 data), for calculating school absenteeism http://nces.ed.gov/programs/coe/2002/section3/indicator17.asp
- CDC Advance Data, from Vital and Health Statistics, no. 326, 2002, for calculating ER visits http://www.cdc.gov/nchs/data/ad/ad326.pdf
- CDC Advance Data, from Vital and Health Statistics, no. 328, 2002, for calculating doctor visits http://www.cdc.gov/nchs/data/ad/ad328.pdf

- 1997 US Employee Absences by Industry Ranked for determining work absenteeism http://publicpurpose.com/lm-97absr.htm
- Over-the-counter (OTC) Drug Sales extracted from Pittsburgh Supercomputing Center's "FRED" data containing pharmacy sales data.

BioWar simulation outputs include:

- Number present and absent per day for each school
- Number present and absent per day for each workplace
- Number of visit records generated per day for each emergency room
- Number of visit records per day for each doctor office, based on insurance claims and assuming that each visit produces one insurance claim
- Number of units of seven types of over-the-counter drugs purchased per day at each pharmacy

6. PRELIMINARY RESULTS

WIZER was run on "Challenge 3" and "Challenge 4" data from BioWar [Carley et al., 2004] using an implementation of Alert WIZER. Challenge 3 data consists of 4 data streams with 10 simulation runs for each attack case (no attack, anthrax attack, and smallpox attack) for each of 4 cities. The city population and locations (buildings and facilities) were scaled at 20%. The parameters were adjusted following an execution of preliminary inference engine steps based on a partial causal diagram of BioWar. We present the means from three of the Challenge 3 simulation output data streams in Tables 3-5.

Table 3 shows that the simulated means of school absenteeism rates for normal simulation cases (no bioattack) fall between lower and upper empirical bounds for the simulations of Norfolk, Pittsburgh, San Diego, and "Veridian Norfolk" (a part of Norfolk specified by Veridian, Inc.). For anthrax attack cases, the simulated means are higher than normal means but still lower than the empirical higher bounds. This is plausible as the empirical higher bound contains (contagious) influenza outbreaks and other disease cases. For smallpox attacks, however, the simulation mean for one city – San Diego – is higher than the empirical higher bound. Smallpox is highly contagious so this is also plausible. For other cities, the simulated means of school absenteeism remain within expected bounds.

Table 3. School Absenteeism

City, percent of simulated population	Empirical lower bound	Empirical higher bound	No Attack (mean)	Anthrax (mean)	Smallpox (mean)
Norfolk, 20%	3.04%	5.18%	3.45%	3.75%	3.55%
Pittsburgh, 20%	3.04%	5.18%	3.52%	4.67%	4.46%
San Diego, 20%	3.04%	5.18%	3.78%	3.81%	5.57%
Veridian Norfolk, 20%	3.04%	5.18%	3.73%	4.05%	4.31%

Table 4 shows that for doctor visits the simulated means for the four cities fall within the empirical bounds for normal (no attack) cases. For anthrax attack cases, the simulated means are higher than those for normal cases for two cities, and slightly lower for two other cities. For smallpox attacks, the means are higher than those for normal cases for three cities and the same for one city. The results for attack cases are imperfect but indicate correct trends. More runs of WIZER are needed to pinpoint the causes. All means for anthrax and smallpox attacks are within the empirical bounds.

Table 4. Doctor Visits per Person per Year

City, percent of simulated population	Empirical lower bound	Empirical higher bound	No Attack (mean)	Anthrax (mean)	Smallpox (mean)
Norfolk, 20%	0.415	1.611	0.499	0.476	0.499
Pittsburgh, 20%	0.415	1.611	0.493	0.485	0.573
San Diego, 20%	0.415	1.611	0.726	0.753	0.796
Veridian Norfolk, 20%	0.415	1.611	0.707	0.821	0.738

Table 5. Emergency Room Visit per Person per Year

City, percent of simulated population	Empirical lower bound	Empirical higher bound	No Attack (mean)	Anthrax (mean)	Smallpox (mean)
Norfolk, 20%	0.056	0.232	0.112	0.108	0.112
Pittsburgh, 20%	0.056	0.232	0.109	0.106	0.129
San Diego, 20%	0.056	0.232	0.149	0.159	0.188
Veridian Norfolk, 20%	0.056	0.232	0.161	0.187	0.168

For emergency room visits (Table 5), the simulated means for four cities fall within the empirical bounds for normal (no attack) cases. For anthrax attacks, the simulated means are higher than those of normal cases for two cities and slightly lower for two others. For smallpox attacks, the simulated means are higher than those for normal cases for three cities and the same for one city. The results for attack cases are imperfect but indicate correct trends.

Challenge 4 data has 12 data streams: school absenteeism, work absenteeism, doctor visits, emergency room visits, emergency room visits using the Surveillance Data Inc. data, and seven drug type purchase data streams. Table 6 shows the percentage of validated data streams for six cities for the no attack case.

Table 6. Percentage of Challenge 4 Data Streams Validated

City	Streams Validated
San Francisco	5/12 = 41.67%
San Diego	7/12 = 58.33%
Pittsburgh	7/12 = 58.33%
Norfolk	6/12 = 50.00%
Hampton	4/12 = 33.33%
Washington DC	4/12 = 33.33%

7. DISCUSSION

Automation of simulation experiment control and analysis is rarely viewed as a critical feature of simulation systems; instead, experimental control, analysis, intervention, validation, and model-improvement are left for humans to perform. Most simulation platforms aim to provide tools to ease the coding of simulation systems, rather than automating the analysis, control, validation, intervention, and model-improvement. WIZER indicates that such automation can be very useful, especially when dealing with socio-technical and public health problems which have a high degree of uncertainty and interactions. Based on empirical data and knowledge, simulations can bound the inferences and allow the empirical claims of the inferences to be investigated. At the same time, knowledge-based inference and control of simulation can reduce the number of simulation searches and virtual experiments that need to be conducted. Simulations – and inferences on them – here act like a dynamic version space on both search and knowledge spaces.

The results presented in this paper are preliminary. More WIZER and simulation runs are needed to get better statistics – such as the median and variance –, and to evaluate error margins, the effects of sample choices, search space traversal, and the performance of combined simulation and knowledge search.

ACKNOWLEDGMENTS

This research was supported, in part, by DARPA for work on Scalable Biosurveillance Systems, the NSF IGERT9972762 in CASOS, the MacArthur Foundation, and by the Carnegie Mellon Center on Computational Analysis of Social and Organizational Systems. The computations were performed on the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputing Center. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA, the National Science Foundation, the Pittsburgh Supercomputing Center, the MacArthur Foundation, or the US Government.

REFERENCES

Axtell, R., Axelrod, R., Epstein, J., and Cohen, M. 1996. “Aligning Simulation Models: A Case Study and Results.” Computational and Mathematical Organization Theory, 1 no. 2: 123-141.

Axelrod, R. 1997. “Advancing the Art of Simulation.” In Proceedings of the International Conference on Computer Simulation and the Social Sciences, Cortona, Italy.

Bankes, S. C. 2004. “Models as Lab Equipment: Science from Computational Experiments.” in Proceedings of North American Association for Computational Social and Organizational Science (NAACSOS) Conference 2004. ISRI, CASOS, Carnegie Mellon University, Pittsburgh, PA.

Boehm, B. 2000. “Spiral Development: Experience, Principles, and Refinements, Spiral Development Workshop.” Special Report CMU/SEI-2000-SR-008. Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA.

Burton, R. and Obel, B. 1998. *Strategic Organizational Diagnosis and Design: Developing Theory for Application, 2nd ed.* Kluwer Academic Publishers, Dordrecht, the Netherlands.

Carley, K. M. and Prietula, M., eds. 1999. *Computational Organizational Theory*. Lawrence Erlbaum Associates, Mahwah, NJ.

Carley, K. M., Fridsma, D., Casman, E., Altman, N., Chang, J., Kaminsky, B., Nave, D., and Yahja, A. 2003. “BioWar: Scalable Multi-Agent Social and Epidemiological Simulation of Bioterrorism Events.” in Proceedings of North American Association for Computational Social and Organizational Science (NAACSOS) Conference 2004, Pittsburgh, PA, http://www.casos.ece.cmu.edu/casos_working_paper/carley_2003_biowar.pdf.

Carley, K. M., Altman, N., Kaminsky, B., Nave, D., and Yahja, A. 2004. “BioWar: A City-Scale Multi-Agent Model of Weaponized Biological Attacks.” CMU-ISRI-04-101 Technical Report, Institute for Software Research International, Carnegie Mellon University, http://reports-archive.adm.cs.cmu.edu/anon/isri2004/CMU-ISRI-04-101.pdf

Chen, L-C, Carley, K.M., Fridsma, D., Kaminsky, B., and Yahja, A. 2003. “Model Alignment of Anthrax Attack Simulations.” Decision Support Systems in the special issue on Intelligence and Security Informatics, in press.

Chen, L-C., Kaminsky, B., Tummino, T., Carley, K. M., Casman, E., Fridsma, D., and Yahja, A. 2004. "Aligning Simulation Models of Smallpox Outbreaks." in Proceedings of the Second Symposium on Intelligence and Security Informatics, Tucson, Arizona, June 10-11, 2004. Also in Springer-Verlag Lecture Notes in Computer Science Vol. 3073.

Cioffi-Revilla, C. 2002. "Invariance and Universality in Social Agent-Based Simulations." in Proceedings of the National Academy of Sciences of the U.S.A., 99 (Supp. 3):7314-7316.

Dershowitz, N., eds. 2004. *Verification: Theory and Practice*. Springer Verlag, New York, NY.

Dianati, M., Song, I., and Treiber, M. 2003. "An Introduction to Genetic Algorithms and Evolution Strategies." Technical Report, University of Waterloo, http://www.swen.uwaterloo.ca/~mdianati/articles/gaes.pdf

Edmonds, B. and Bryson, J.J. 2004. "The Insufficiency of Formal Design Methods: the Necessity of an Experimental Approach for the Understanding and Control of Complex MAS." in Proceedings of AAMAS 2004, ACM.

Epstein, J.M. and Axtell, R.L. 1996. *Growing Artificial Societies*. MIT Press, Cambridge, MA.

Epstein, J.M., Cummings, D.A.T, Chakravarty, S., Singa, R.M., and Burke, D.S. 2004. "Toward a Containment Strategy for Smallpox Bioterror: An Individual-Based Computational Approach." Brookings Institution Press, Washington,DC. http://www.brook.edu/press/books/towardacontainmentstrategyforsmallpoxbioterror.htm

Fogel, L.J. 1999. *Intelligence Through Simulated Evolution: Forty Years of Evolutionary Programming*. Wiley Series on Intelligent Systems, New York, NY.

Gilbert, N. and Troitzsch, K. G. 1999. *Simulation for the Social Scientist*. Open University Press, Berkshire, United Kingdom.

Guetzkow, H.S., Kotler, P. and Schultz, R.L. eds. 1972. *Simulation in Social and Administrative Science: Overviews and Case-Examples*. Prentice Hall, Englewood Cliffs, NJ.

Haenni, R., Kohlas, J., Lehmann, N. 1999. "Probabilistic Argumentation Systems." Technical Report 99-9, Institute of Informatics, University of Fribourg, Fribourg, Switzerland.

Huang, C-Y, Sun, C-T, Hsieh, J-L, and Liu, H. 2004. "Simulating SARS: Small-World Epidemiological Modeling and Public Health Policy Assessments." Journal of Artificial Societies and Social Simulation, 7, no. 4. http://jasss.soc.surrey.ac.uk/7/4/2.html

Jewell, N.P. 2003. *Statistics for Epidemiology*. Chapman & Hall/CRC, Boca Raton, FL.

Law, A.M. and Kelton, W.D. 2000. *Simulation Modeling and Analysis, 3rd ed.* McGraw-Hill, New York, NY.

Lucena, C., Carcia, A., Romanovsky, A., Castro, J., and Alencar, P., eds. 2004. *Software Engineering for Multi-Agent Systems II*. Lecture Notes in Computer Science 2940, Springer Verlag, New York, NY.

Myers, R.H. and Montgomery, D.C. 2002. *Response Surface Methodology: Process and Product Optimization using Designed Experiments*, *2nd ed.* John Wiley & Sons, New York, NY.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, United Kingdom.

Pearl, J. 2003. "Statistics and Causal Inference: A Review." Test Journal, 12 no. 2 (December): 281-345.

Pressman, R. S. 2001. *Software Engineering*. McGraw-Hill, New York, NY.

Prietula, M.J., Carley, K.M., and Gasser L. 1998. *Simulating Organizations*. AAAI Press/The MIT Press, Menlo Park, CA.

Rasmussen, S. and Barrett, C.L. 1995. "Elements of a Theory of Simulation." ECAL 95, Lecture Notes in Computer Science, Springer Verlag, Berlin, Germany.

Reifman J., Gilbert, G., Parker, M., and Lam, D. 2004. "Challenges of Electronic Medical Surveillance Systems." RTO HFM Symposium on NATO Medical Surveillance and Response: Research and Technology Opportunities and Options, Budapest, Hungary. http://www.rta.nato.int/Pubs/RDP.asp?RDP=RTO-MP-HFM-108

Robert, C. P. and Casella, G. 1999. *Monte Carlo Statistical Methods*. Springer Verlag, New York, NY.

Shervais, S., Wakeland, W., and Raffo, D. 2004. "Evolutionary Verification and Validation of Software Process Simulation Models." in the 5th International Workshop on Software Process Simulation and Modeling, extended abstract.

Shervais, S. and Wakeland, W. 2003. "Evolutionary Strategies as a Verification and Validation Tool." Portland State University draft paper, http://www.sysc.pdx.edu/faculty/Wakeland/papers/EvoVandVRevD.pdf

Simon, H. 1957. "A Behavioural Model of Rational Choice." Quarterly Journal of Economics (69): 99-118.

Taber, C. S. and Timpone, R. J. 1996. *Computational Modeling*. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-113. Sage, Newbury Park, CA.

Ward, Michael D. 1985. *Theories, Models, and Simulations in International Relations*. Westview Press, Boulder, CO.

Wasserman, S. and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, United Kingdom.

BIOGRAPHY

Kathleen M. Carley received her PhD in Sociology from Harvard University in Cambridge, MA in 1984. She received two S.B.'s, one in Political Science and one in Economics from the Massachusetts Institute of Technology, Cambridge, MA in 1978.

She is a Professor of Computation, Organizations and Society in the Institute for Software Research International in the School of Computer Science at Carnegie Mellon University, Pittsburgh, PA. She is the author or co-author of 5 books and over 100 articles in the area of computational social and organizational science and dynamic network analysis. Recent publications include - Designing Stress Resistant Organizations: Computational Theorizing and Crisis Applications with Zhiang Lin (Kluwer, 2003); Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers with Ron Breiger and Pipp Pattison (National Academy Press, 2003); Smart Agents and Organizations of the Future in The Handbook of New Media, edited by Leah Lievrouw & Sonia Livingstone (Sage, 2003). Her research combines cognitive science, social networks and computer science, especially in computational social and organization theory, group, organizational and social adaptation and evolution, dynamic network analysis, computational text analysis, and the impact of telecommunication technologies on communication and information diffusion within and among groups. Her computer simulation models meld multi-agent technology with network dynamics and are in areas such as BioWar – a city, scale model of weaponized biological attacks; OrgAhead – a model of strategic and natural organizational adaptation; and Construct – a model of the co-evolution of social and knowledge networks and personal/organizational identity and capability.

Prof. Carley is a member of the Academy of Management, Informs, the International Network for Social Networks Analysis, American Sociological Society, the American Association for the Advancement of Science and Sigma XI. In 2001, she received the Lifetime Achievement Award from the Sociology & Computers Section of the American Sociological Association.

Alex Yahja is a PhD candidate at the Computation, Organizations and Society program in the Institute for Software Research International in the School of Computer Science at Carnegie Mellon University, Pittsburgh, PA. He received his M.Sc. in Robotics in 2001 and M.Sc. in Engineering and Public Policy in 2003, both from Carnegie Mellon University.

Alex Yahja is interested in using computation to tackle the hard problems in the fields such as social science, management science, law, health science, organization science, economics, and public policy, among others. His focus is on molding computational concepts and software to better serve the needs of common folks and society. His specific research areas are model validation, knowledge-based control of simulation, model improvement, social simulation, knowledge inference, scientific design of computational experiments, and automatic programming. He is a member of a team that built BioWar – a city-scale model of biological attacks on a demographically-realistic population against the background of naturally-occurring diseases, melding multi-agent technology with network dynamics and knowledge-based simulation-aided inference.