Available online at www.sciencedirect.com
ScienceDirect
Simulation Modelling Practice and Theory 16 (2008) 242–256
SIMULATION MODELLING PRACTICE AND THEORY
www.elsevier.com/locate/simpat

Balancing the criticisms: Validating multi-agent models of social systems

Marcus A. Louie [a,*], Kathleen M. Carley [b]

[a] *Department of Engineering and Public Policy, Center for the Computational Analysis of Social and Organizational Systems, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[b] *Computation, Organizations and Society, Institute for Software Research International, Center for the Computational Analysis of Social and Organizational Systems, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

Available online 4 December 2007

Abstract

Using multi-agent models to study social systems has attracted criticisms because of the challenges involved in their validation. Common criticisms that we have encountered are described, and for each one we attempt to give a balanced perspective of the criticism. A model of intra-state conflict is used to help demonstrate these points. We conclude that multi-agent models for social systems are most useful when (1) the connection between micro-behaviors and macro-behaviors are not well-understood and (2) when data collection from the real-world system is prohibitively expensive in terms of time or money or if it puts human lives at risk.

Keywords: Multi-agent model; Simulation; Validation; Social systems

1. Introduction

Multi-agent models are touted as a method for analyzing "complex social systems," particularly those characterized by multiple interacting parts and non-linear behavior [11,13,14,29]. As a result, these models are being used to examine a variety of policy domains including civil violence [16], the spread of infectious disease [12], the effects of government policies on the transportation of goods [5], and the effects of mutual influence on domestic water demand [26].

Researchers and policy makers are turning to these models for reasons of ethics, cost, timeliness and appropriateness. In some systems, such as those modeling the spread of infectious disease, testing experimental conditions would put the safety of people at-risk, creating an ethical problem. In other cases, real-time evaluation of an existing system may be prohibitively long. Simulation allows for rapid assessment. Simulation is also used when the cost of collecting data on the dependent variable is prohibitively expensive, or there are a large number of experimental conditions to test. For example, in a disaster, simulation can be used to rapidly

* Corresponding author.
E-mail address: mlouie@andrew.cmu.edu (M.A. Louie).

1569-190X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.simpat.2007.11.011

evaluate many previously unexamined alternatives [12,27]. In all of these cases, since the real-world system under study is considered a complex, non-linear dynamic system, multi-agent simulations are often used as it is considered to have the appropriate level of complexity.

As the use of multi-agent models has become more prevalent, a growing concern has arisen with how to validate such models. From a history of science perspective it is important to note that the most advanced methods of validation were developed in engineering fields for assessing models of technical systems that followed fundamental physical laws. In contrast, these large-scale multi-agent systems are used for examining socio-political systems where the fundamental underlying laws are not known. Multi-agent models of social systems are difficult to validate because these models represent a new approach to simulation for which traditional validation methods are not always applicable. Given these challenges, we need to first ask what an appropriate validation process is for such models. Second, we need to know what value these models have even despite the challenges in their validation.

This paper proceeds in two parts. First, we synthesize previous work in simulation model validation to construct a validation strategy for models of social systems based on the purpose of a model. Second, we outline common criticisms related to validating multi-agent models of social systems. For each criticism, we seek to demonstrate that even if the legitimacy of the claim is granted, that the models can still be useful as a means for developing theories about the target system. A dynamic-network multi-agent model of intra-state conflict called the Regional Threat Evaluator (RTE) is used as an example. Note, this paper is not a presentation and validation of the RTE model. Rather, this paper is an analysis of the appropriateness of standard validation procedures for large-scale multi-agent models such as the RTE.

2. Designing a validation process

Previous work on validation processes for simulation models can be broken into two principal threads. One thread addresses what the *major steps* are in validation [4,9,30,35]. For example, Thomsen et al. [35] propose a trajectory of major validation steps for simulation models that are based on real data and whose purpose is to be used prescriptively. The other thread has focused on the specific *techniques* that might be used during each of the major steps of validation [3,20,23]. A well-recognized example from this thread is Law and Kelton's [23] treatment of statistical validation techniques for simulation models.

The validation process should be tied to the purpose and the context for which the model is being developed [6,8,9,30,35]. We distinguish a validation process from a validation technique. A process is a series of steps taken to validate different parts a model such as verifying that the model mechanisms are representative of the real-world or comparing model output to historical data. Techniques are the individual methods used to judge whether each part of the model is "valid." Statistical tests such as *t*-tests are examples of techniques.

Sargent [30] provides an overview of different validation techniques, each providing different types and levels of validity. He notes that the desired level of validity is determined on the purpose of the model, but does not attempt to describe in detail what different purposes are and how they relate to the validation process. Burton [8,9] complements Sargent's work by describing types of questions that are asked of simulation models while recognizing that the level of validation is still dependent on the question, or purpose, of the model.

In this section, we synthesize this related work and organize the types of questions that are asked of models and associate them with the types of validation that are appropriate for each type of question. To assist in giving structure to the synthesis, we use a conceptual description of the simulation model development and validation process given by Banks et al. [4]. Fig. 1 is a representation of the conceptual components and steps of the process, recreated from Sargent [30].

2.1. Types of validation

The boxes in Fig. 1 show three different parts of a simulation model that can be validated. This section identifies the different types of validation that can be performed for each of these model parts. Conceptual validity is determining the extent to which the model theories and the underlying assumptions are appropriate for the purpose of the model. Determining the validity of data involves making sure that the data are appropriate for the purpose of the model, that a sufficient amount of data exist to build and validate the model, and that the