Cross-Cultural Research 45(2) May 2011: 1–23.

Cross-Cultural Research
XX(X) 1–23

DOI: 10.1177/1069397110393894
http://ccr.sagepub.com
SAGE

Comparing Agent-Based Computational Simulation Models in Cross-Cultural Research

Claudio Cioffi-Revilla, PhD[1]

Abstract

Mel Ember was co-Principal Investigator in the Mason-HRAF Joint Project on Eastern Africa, a multiyear project aimed at developing and analyzing advanced computational agent-based models of human societies across 10 countries and 12 ecosystems. A major unsolved challenge in this kind of social science research is to devise a systematic way to compare, contrast, and communicate different models of social dynamics along relevant dimensions and characteristics, given the inherent complexity of most computational agent-based models. This article proposes a viable systematic framework for comparing models and illustrates its application using some of the models that Mel helped inspire and develop as senior project participant.

Keywords

cross-cultural research, computational social science, Eastern Africa, social simulation, agent-based modeling, comparative analysis, computational methodology

[1]George Mason University, Fairfax, VA

Corresponding Author:
Claudio Cioffi-Revilla, PhD, Center for Social Complexity, Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA 22030
Email: ccioffi@gmu.edu

Introduction: Motivation and Background

Model-to-model analysis is a special topic of comparative social science research, focusing on side-by-side comparisons of models of social systems and processes (Clark, Cumow, & Cole, 1975; Cioffi & Gotts, 2003; Deutsch, 1948; Hales, Rouchier, & Edmonds, 2003; Rouchier & Bousquet, 2001; Rouchier, Cioffi-Revilla, Polhill, & Takadama, 2008; Rouchier, Edmonds, & Hales, 2003; Grimm et al., 2005, 2006; Kahn, 2007). The topic of model-to-model (or M2M) analysis has deep and ancient roots in Western scientific thinking, reaching back to Aristotle's tripartite taxonomy of political systems in the 4th century BC—that is, monarchies, aristocracies, and democracies, and their associated "degenerative" forms (tyrannies, oligarchies, and ochlocracies [rule by the "mob"], respectively), as shown in Table 1.[1] Although contemporary social scientists use other taxonomic schemes for classifying and comparing social systems,[2] parsimonious taxonomies applicable to a broad spectrum of models—not just empirical social systems—provide systematic frameworks for advancing comparative social research.[3]

The increasing number and variety of simulation models in computational social science[4] motivates the need for developing a systematic framework for comparative analysis of social simulations. Such a diversity of models covers many types of social agents and dynamics, as relates to other computational social science models (e.g., social network models, event data models, social geospatial models, and other computational models). The need for a comparative framework for social simulations is particularly acute for spatially complex models—that is, large social simulations with significant geographic features—given their interdisciplinarity and number of participants as well as their longer life-cycle characteristics (Cioffi, 2010b). Comparing simple "toy" models is difficult; comparing complex simulation models poses additional challenges. The comparative approach in computational social science is still largely undeveloped.

This fledgling interdisciplinary field, at the intersection of social science, computer science, and related disciplines (e.g., geography, organizational science, and environmental science), needs rigorous concepts and robust methodologies capable of yielding deeper insights in support of scientific progress.[5] Comparative analysis of computational models in cross-cultural research can also extend new, valuable, theoretical, and methodological bridges between traditional social sciences and newer areas of computational social science (Cioffi, 2010a). This is an important task for promoting sustainable scientific development.

Comparing social simulations can be framed in set-theoretic terms as an examination of the union and intersection of models within a larger model universe, as illustrated in Figure 1. Each model constitutes a set consisting of

Table 1. Classical Aristotelian Classification of Polities

Normal ("stable")	Monarchy	Aristocracy	Democracy
Degenerative ("failed")	Tyranny	Oligarchy	Ochlocracy

attributes (variables) and behaviors (dynamics) within a much larger universe U of social models. Accordingly, *similarities* and *differences* are identified by the *intersection* and *union of moons*, respectively. The most interesting cases for comparative analysis are where both intersections (similarities) and moons (differences) exist. Two extreme (degenerative) cases occur when (a) there are no similarities between models (disjoint sets), such that their intersection is empty, or (b) one model is a strict subset of the other, such that the intersection is just a subset of the larger model, and their union also equals the larger model. Here the focus is on cases where both intersections and moons exist.

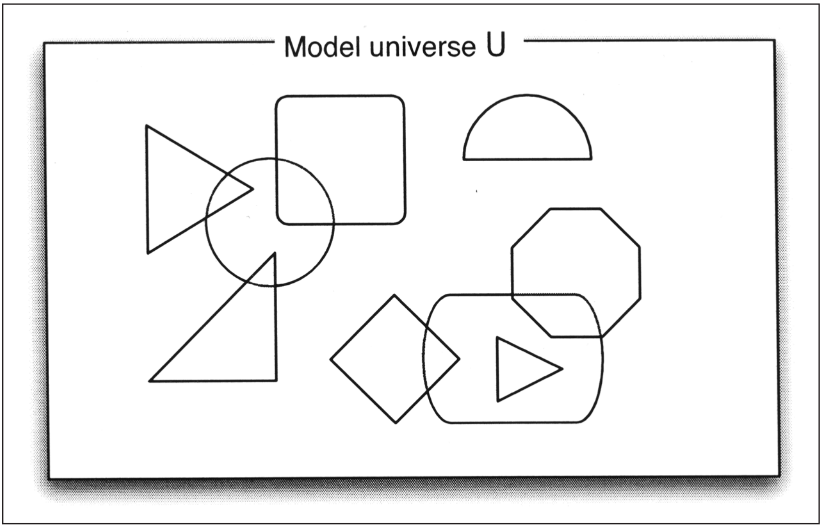

Figure 1. Different models represented as objects (shaped point-sets) that encapsulate attributes and behaviors, within a modeling universe, U
Note: Similarities and differences among models correspond to intersections, or "lenses," and nonoverlapping components, or "moons," respectively.

This article presents elements of a systematic framework for comparing social simulations, especially for models explicitly involving societies of interacting agents. The framework is based on two levels of comparison. Each level

Table 2. Agent-Based Models From the Mason-HRAF Joint Project on Eastern Africa

HerderLand	Herders interact with each other in a relatively small local landscape
RebeLand	Country-scale model with several provinces and ecosystems
AfriLand	Multicountry model of an international region and ecosystems
RiftLand	A version of AfriLand calibrated to the Eastern Africa region

consists of additional categories or dimensions that provide for increasingly specialized comparisons among models. The two "root" levels concern generic and specific comparisons, based on methodological and formal features, respectively, as discussed in the following sections. The two root levels are open to further development, an advantage at this stage of scientific infancy in the field of computational social simulations.

The focus in this article is on comparing two recent social simulation models—called "HerderLand" and "RebeLand" (see Table 2)—selected from the computational agent-based models developed by the Mason-HRAF Joint Project on Eastern Africa. (See project references in Note 2.) All four models are in the MASON system (Luke et al., 2005, Sullivan, Coletti, & Luke, 2010 (IN PRESS). Here, HerderLand and RebeLand are used for illustrative purposes, because they are the most published models from the Mason-HRAF project.

The focus is on comparative aspects that illustrate main similarities and differences between models since each model is separately and individually described elsewhere in technical detail.

Comparing Social Simulations

Social simulations consisting of agent-based models—or other kinds of computational models (e.g., microsimulations, queuing models, or system dynamics models)—pose some interesting and challenging opportunities for cross-cultural research. Social scientists accustomed to modeling have faced some of these issues before. For instance, statistical regression models are often compared in terms of the following features:

- Dependent and independent variables
- Operational measures and measurement error in each variable
- Data transformations applied to operational measures
- Functional form of specified models
- Regression plots and related graphics
- Parameter estimates and their statistical significance
- Stationary properties

- Goodness of fit
- Mathematical properties or theoretical inferences from estimated models—
- Treatment of missing cases

Similarly, know-how also exists for comparing mathematical models, such as game-theoretic models of social relations. For example, games representing sociocultural relations can be compared in terms of these and other dimensions:

- Number of players: 2-person vs. *N*-person games
- Length of play: finite vs. infinite games
- Information regime: complete vs. incomplete information games
- Number and types of strategies: pure vs. mixed
- Symmetry: symmetric vs. asymmetric games
- Embeddedness: simple vs. nested games
- Form: normal vs. extensive

Methodological experience gained during past decades enables comparative analysis of different types of sociocultural models, whether rendered in statistical or in mathematical form. This cannot be said for computational models where cross-cultural analysis is just as desirable. Comparisons among simulation models yield new insights that advance understanding by highlighting similarities and differences. New categories (dimensions) for comparison are needed, given the different nature of social computational models used in simulation. Which are the relevant categories for comparison? Which of them would be most insightful for purposes of cross-cultural research? How do different simulation platforms facilitate or hinder cross-cultural comparisons? The framework outlined in the next sections is intended to contribute answers to such questions. The main caveat is that a definitive framework is still premature in the area of complex agent-based simulations that comprise interdependent ("coupled") social and biophysical systems.

The comparative analysis of social simulations can be approached from two basically distinct but related ontological levels: *generic* and *specific*, as discussed in the next sections. Although the four models in Table 2 are agent-based, the same two levels should also apply to other types of social simulation models (e.g., systems dynamics).

Generic Comparison

Generic comparison of social simulations can be based on contrasting characteristics common to all social simulations, independent of specific formal language in which they are rendered or instantiated. A fundamental way to frame such a first, high-level comparative approach is to recall the defining features of all social simulations, in terms of the following:

1. Focal system, empirical domain, or focal region being simulated;[6]
2. Simulation system or computational model used;
3. The critical *modeling relation* linking a given focal system and a corresponding simulation system (model), comprising several key aspects, such as *abstraction*, *validation*, and *data* considerations; and
4. *Policy relevance* or applied value in terms of real-world significance, as some social simulations are intended for pure research and others are used for policy analysis.

Several technical terms included in these generic comparative dimensions include the following.

Definition 1 (Focal system): The focal system *F* of a social simulation is the empirical reality or focal domain being modeled. This includes real-world components consisting of human, social, natural (i.e., biophysical), or man-made (engineered or artificial) entities and their interrelations.

Definition 2 (Simulation system): A simulation system *S* is the computational model abstracted (see below) from a given focal system, consisting of a selection of human, social, natural (i.e., biophysical), or man-made (engineered or artificial) entities and relevant relations among them.

Definition 3 (Model abstraction): Formally, abstraction is a mapping *A* from a given focal system *F* to an instantiated simulation system *S*, such that $A: F \rightarrow S$. Abstraction refers to the selection or subset of referent real-world entities and relations that are to be formally included (i.e., formalized) in a given simulation system. Abstraction is a function of the set of questions a given simulation system is intended to answer.

Definition 4 (Model validation): Validation is the process of establishing an acceptable correspondence between focal system and simulation system in the context of modeling and simulation. A valid simulation model is one that has acceptable correspondence with a given referent or focal system of interest.

Comparative analysis of models along generic dimensions—focal system, formalization, validation, and policy relevance—yields important information and insights, such as cross-cultural features that are valuable for understanding and developing models. The fact that these dimensions are generic—or common to all social simulations—does not detract from their importance. Understanding these high-level features is essential for a meaningful appreciation of more specific features that are relevant in cross-cultural contexts. In contrast, misunderstandings are common when comparisons focus immediately on specific technical features without regard to high-level, general aspects. For instance, different social simulations can be about the same focal system (e.g., an agent-based model and a systems dynamics model of the same Eastern Africa region), while other social simulations model different focal systems (e.g., Eastern Africa and Central Asia) using the same type of computational model (such as agent-based models). Likewise, social simulations may be similar or different in terms of validation requirements (high-vs. low-fidelity; Kuznar, 2006).

Comparing Focal Systems: Empirical Domain

In terms of the four models in Table 2, the respective focal systems have increasing empirical scope, ranging from local (or within-country) at one end of the spectrum, to regional on the other (multicountry) although all four aim at the same, general focus area of investigation (Eastern Africa). Accordingly, the four models provide a micro-to-macro spectrum of socioecological dynamics in terms of layered multilevel complex systems.

Differences, similarities, and their dynamics in simulated cross-cultural patterns are fundamentally dependent on generic features of computational models. The HerderLand model was inspired by the Mandera Triangle in northeast Kenya (microscale model consisting primarily of herders and farmers). In contrast, the RiftLand model represents a much larger focal area consisting of most of the countries and main cultural groups of Eastern Africa (macroscale model). RiftLand includes the countries of Kenya, Uganda, Rwanda, Burundi, and associated borderlands with neighboring countries (in counterclockwise order): southern Somalia, southern Ethiopia, southeastern Sudan, eastern Democratic Republic of Congo (DRC), and northern Tanzania. A total of nine countries are represented in RiftLand (the same number as the more abstract AfriLand model). RebeLand, by comparison, is a meso-scale model of a single but politically complete country that is analytically situated between the two scales of HerderLand (most micro) and RiftLand (most macro). While HerderLand and RiftLand focus on empirically referenced regions (the Mandera Triangle and Eastern Africa, respectively), RebeLand and AfriLand represent abstract country-systems (comprising one and nine countries, respectively). Understanding generic similarities and differences is essential for evaluation and comparative analysis along other dimensions.

Comparing Simulation Systems: Formal Implementation

Computational models can be equation-based or object-oriented, depending on the primary ontology and language in which they are expressed. Variable- or equation-based models are high-dimensional mathematical systems that generally cannot be solved in closed form; hence the need for simulation. For example, systems dynamics models are equation- or variable-based systems that contain too many variables to obtain closed-form solutions. Simulation makes it possible to obtain time-dependent trajectories (e.g., Hanneman, 1988; Lowe, 1985; Ruloff, 1978, 1981; Sterman, 2000).

In Table 2 the simulation implementation system employed across all four models is object-oriented from an ontological perspective; the basic building blocks are classes of *social*, *natural*, and *human-built* entities, not variables or equations. Variables and equations are also used in a computer program (code), but in object-oriented modeling (abstraction) they are "encapsulated" within the social or physical entities represented in each model.

Many programming languages and simulation systems exist for formalizing object-oriented social simulation models (Nikolai & Madey, 2009). All four models are implemented in the MASON system.[7] Therefore, all four models contain social and ecological components "instantiated" (rendered) in the same programming language (Java) although the scope of each model (simulation system) differs according to focal system and scale. The MASON system has a number of defining features, including the following (Luke et al., 2005):

- Source code written in 100% Java;
- Fast, portable, and fairly small, to run on most laptop computers;
- Completely separate model (scheduler, data fields, random number generator) from visualization (Graphic User Interface, GIU), which can be added, removed, or changed at any time;
- Models can be check-pointed and recovered, and dynamically migrated across platforms;
- Identical results are produced across platforms (Mac OS, Windows, Lynux);
- Models are self-contained and can run inside other Java frameworks and applications;
- 2D and 3D visualization;
- Can produce PNG snapshots, Quicktime movies, charts, and graphs, and output data streams; and
- Easy compatibility with ECJ for use in evolutionary computation.

These and related features make MASON a powerful tool for cross-cultural research because the resulting models permit the representation of diverse groups and culture systems within a common simulation framework for data collection from simulation runs and comparative analysis of models and results. Comparing the conflict patterns and interaction dynamics of herder and farmer groups in the HerderLand model with those of households and government agents in the RebeLand model is greatly facilitated by the common MASON environment. The same template for representing different groups, belief systems, behavioral norms, and other defining components can be drawn from basic concepts and principles from cultural anthropology and implemented across models.

Comparing Modeling Relations: Focal-Simulation Link

Similarities and differences across models exist in terms of abstraction, validation, and data considerations as part of the overall modeling relation between focal system and simulation system. In terms of *abstraction* (see Definition 3 above)—the first aspect of a modeling relation—RebeLand and AfriLand are relatively abstract and generalized, whereas HerderLand and RiftLand are intended to be more empirical, particularly the latter. Nonetheless, RebeLand and AfriLand are designed with relevant empirical features, such as landscape characteristics and key socioeconomic properties that resemble those of the real world (e.g., Zipf-like distribution of population sizes, Pareto distribution of wealth, and Poisson distribution for the onset of public issues).[8] Since model abstraction is always a function of questions asked by a given simulation system relative to a given focal system, some models abstract more than others from empirical reality—in a way that is not different from statistical or mathematical modeling (e.g., regression models or game-theoretic models, respectively).

The degree of abstraction defines a scale that goes from low resolution (high-level abstraction; so-called “toy models”) to high resolution or high fidelity (low-level abstraction; empirical models). Accordingly, RebeLand and AfriLand are relatively low-fidelity models intended to answer high-level questions about political stability, insurgency, state failure potential, and societal effects of biophysical change (including climate change and variability). HerderLand and RiftLand are designed at a higher level of fidelity for answering more specific questions about environmental changes, intra- and intergroup conflict, migration patterns, refugee flows, and related dynamics.

Validation requirements—the second aspect of the modeling relation—vary across models. Since HerderLand and (particularly) RiftLand aim at a higher level of empirical fidelity than the other two (RebeLand and AfriLand), their validation requirements are more stringent. All four models contain some fundamentally valid features—such as a Zipf distribution of settlement sizes (intentionally jittered by some noise), Pareto distribution of wealth, and Poisson distribution for the onset of public issues. Beyond these basic features, RiftLand also has an additional set of empirical features that increase validation requirements in terms of outputs from simulation runs. Validation is a complex process involving calibration of model components as well as tests of simulation runs that match simulated and focal data (Cioffi, 2010c). In social science such a process is known as *external validation* (Kaplan, 1964), as opposed to *verification*, which refers to debugging (roughly corresponding to *internal validation* in social research). Some common validation requirements include comparison between the following categories of simulated and empirical data:

- Correspondence between main qualitative behaviors
- Distributions of onset times, intensity, and duration for significant events
- Size distributions of groups or areas affected
- Emergent social organization of relevant groups
- Resource depletion and sustainability patterns
- Long-term equilibria or lack thereof (oscillations, asymptotic behaviors)
- Coherence between short-term (high-frequency) phenomena and long-term (low frequency) trends

Each of these categories of validation comprises multiple key components. For example, focal-simulation comparisons of distributions (for both duration and size dimensions) comprise other aspects, such as distribution moments (mean, variance, kurtosis, skewness) and probability functions (density function, cumulative density function, intensity or hazard function). Simple correlations are useful, but generally inadequate due to the limited amount of information they convey compared to a broader array of qualitative and quantitative measures. A more in-depth discussion of validation in social simulations with specific reference to MASON East Africa models is provided elsewhere (Cioffi, 2010c; see also Cioffi, 2002 in the context of establishing universality). The precise set of validation requirements and strategies depends on observed characteristics of a focal system (and levels of measurement) and output data collected from runs of the simulation system.

Finally, *data requirements*—the third aspect of the modeling relation—should be determined by previously discussed features: abstraction level and validation requirements. A common misconception (error) in simulation in general, and in social simulations in particular, is to prioritize data-related issues over all others—including but not limited to validation requirements pertaining to each specific model—as if data requirements could be identified a priori, independent of abstraction level and validation requirements.

Similarities and differences exist among the four Mason-HRAF East Africa models in Table 2 in terms of data requirements. All models require data on the following categories:

1. Features of the focal terrain
2. Land-cover features (biomass distribution)
3. Natural resource distribution data relevant to modeled agents (economic resources, hydrology)
4. Climate (mainly bimodal seasonal rainfall patterns)
5. Population data (generally household-level data)
6. Basic social ontology (agent identity, basic needs, behavior rule set)

Beyond these requirements, different models use additional data determined by the focal system and focal-simulation relation (abstraction).

For example, the number and location of watering holes matters greatly in HerderLand (and RiftLand), because they play a significant role in emergent conflict patterns (Hailegiorgis, Kennedy, Balan, Bassett, & Gulden, 2010; Kennedy et al., 2010a; see also Kuznar and Sedlmeyer, 2005, for an earlier model of pastoralists and conflict inspired by the Darfur region, Sudan). Boundaries are a constituent feature in the other three models, and empirical data for borders are used in RiftLand (whereas provincial and state boundary data are only notional in RebeLand and AfriLand). RebelLand has features that are built on empirical data patterns pertaining to population settlement distributions and mechanisms for onset and duration of public issues, placing capacity demands on the polity. AfriLand's data requirements parallel those of RebeLand. RiftLand has the greatest data requirements, because it is the most empirically tuned model. In each instance, data requirements are determined by research questions being addressed by each model and the abstracted simulation system instantiated in code, not simply by a priori features of the focal system (which are infinite!). Selecting data requirements is a deductive exercise based on research questions, not an inductive process uninformed by theory.

Specific Comparison

A more *specific* comparative approach to social simulations is based on formal and technical aspects that are narrower than generic dimensions discussed earlier. A linguistic perspective is helpful for specific comparisons, given the formal nature of all computational models, including social simulations.

Social simulations can be compared according to their *semantics*, *syntax*, *pragmatics*, and *genetics*. The main technical comparison is arguably by syntax, which specifies the formal language and programming structure and functioning of a social simulation. For instance, social simulations can be syntactically classified and compared according to two categories: *variable-based models* (e.g., system dynamics simulations, queuing models) or *object-based models* (e.g., cellular automata, agent-based simulations, agent-based networks). Some would argue that the variable versus object orientation is ontological, not just syntactic (Barker, 2005; Cioffi, 2008; Lau, 2001). Beyond this classification, each social simulation model contains important implementation details that warrant specific comparisons.

Variable-based social simulation models can be classified by mathematical structure as *continuous*, *discrete*, or *hybrid*. By contrast, object- or individual-based models are constituted by entities (social, natural, artificial) specified by their *attributes* and *behaviors* (methods). In a variable-based social simulation, the formal structure is given by mathematical equations. By contrast, in an object-based social simulation (e.g., multiagent model) the formal structure is fully specified by code.

Social simulations that are object-based—such as all agent-based models—can be compared using the *Unified Modeling Language* (UML; Ambler, 2005; Eriksson, Penker, Lyons, & Fado, 2004), based on a standardized notation consisting primarily (but not exclusively) of *class diagrams*, *sequence diagrams*, and *state diagrams*. Each provides a complementary view of the same system—as is with different images of the same object (natural light, thermal imaging, ultraviolet).

Comparing Ontologies

Definition 5 (Class diagram): A UML class diagram (Figure 2) is a graphic representation of the main entities (classes, objects) included in a social simulation, such as human, social, natural (i.e., biophysical), or man-made (engineered or artificial) entities and their interrelations.

A class diagram describes the basic ontology of a model. The class diagram shown in Figure 2 illustrates some of the notational conventions for representing the main entities and relations in a social system in UML. As a social entity, a `Polity` (typewriter font denotes a computational object) consists of (diamond head links) a `Society` and a system of *Government*. The latter has three relations abstracted in this class diagram (there are also others, of course): (a) receiving support inputs (resources, taxes) from `Society`; (b) influencing

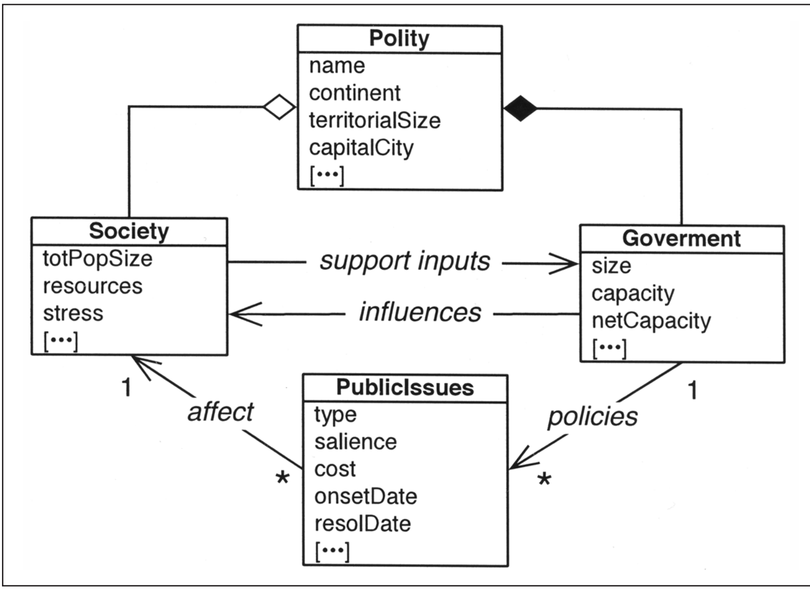

Figure 2. Example of UML class diagram of a polity

`Society` through information and other means, including government propaganda; and (c) enacting policies to solve emerging *Public Issues* that affect `Society`. Each entity/object also "encapsulates" its own attributes and "methods" (rules specifying how attributes change over time). For example, the attributes of *Government* might include *size* (variety of institutions), *capacity* (resources available for policies), and other institutional features. Other entities and relations (e.g., different identity groups and authority relations within Society) can also be represented in a higher resolution diagram. (Further examples are provided in Cioffi, 2008; Cioffi et al., 2007; Cioffi, Rogers, & Latek, 2010; Cioffi, De Jong, & Bassett, 2010; Taber and Timpone, 1996.)

Note that an ontology defined by a UML class diagram consists of both static and relational aspects. For example, the computational "class" named `Household` (a component within `Society`) may contain attributes such as *size* (number of individuals belonging to a household), `wealth` (net household income after expenses), *cultural identity* (ethnic membership of `Household`), and other relevant cultural features in the simulation model. Other classes instantiate parcels, government agents, and biophysical features. Methods, on the other hand, specify how attributes change, such a household becoming richer or growing in number, or *State* gaining or decreasing in capacity (a key attribute for financing public policies to address issues in all models except HerderLand). Thus, methods implement dynamics that determine the state of the model, whereas classes (modeled entities) and their attributes implement mostly static ontological features of the simulation model. The state of the simulation model refers to the current value of all attributes encapsulated within its classes as determined by associated methods.

The precise ontology of each model in the Mason-HRAF Project on Eastern Africa is described elsewhere and is too extensive to be covered here in great detail. The following summary describes the basic socionatural ontology in each model:

HerderLand. The main classes comprise locations, agents (herders and farmers), watering holes, and biomass, all within a local area of approximately 150 × 150 km in size, inspired by the Madera Triangle region in northeastern Kenya.

RebeLand. As a country-level model, the main classes consist of cities and smaller settlements, provinces or administrative districts, roads, resources, and social agents (in this case households, insurgents, government institutions, public administration, and security agents).

AfriLand. This multicountry model is comprised of nine RebeLand-like polities, including all the RebeLand entities (and dynamics), in addition to borders that can be traversed by agents.

RiftLand. Classes contained in HerderLand and AfriLand compose the entities in Rift-Land, in addition to other entities such as ecosystems, bodies of water, transportation infrastructure, and a social landscape that more closely resembles Eastern Africa.

From a methodological perspective, cross-cultural comparison across models is greatly facilitated by the fact that all entities and relations are rendered in terms of computational objects consisting exclusively of their encapsulated attributes and methods. Entity-to-entity comparisons (for example, as between different cultural groups or types of social relations) can be framed in terms of comparing specific (and unambiguous) code pertaining to the relevant classes and associations—a task that is generally difficult or impossible based on narratives or even detailed ethnographies.

Comparing Dynamic Processes

Beyond ontology, the next two types of UML diagrams describe dynamics and can be used for comparing micro- and macro-cultural processes within and across models.

Definition 6 (Sequence diagram): A UML sequence diagram (Figure 3) is a graphic representation of the schedule with which main events

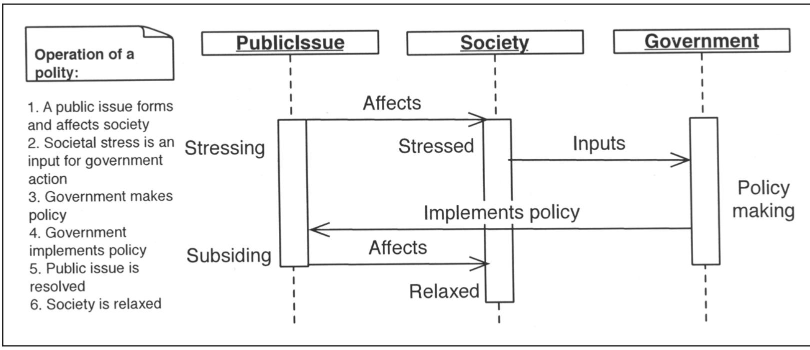

Figure 3. Example of a UML sequence diagram for the operation (functioning) of a basic polity

and interactions (e.g., situational changes, information flows, decision making, behavioral acts, lotteries, resource transfers) occur among entities (agents or other classes) in a social simulation.

The sequence diagram shown in Figure 3 illustrates more specific micro-interactions that occur among various entities within the `Polity` system. A sequence diagram is read from top to bottom, following the chronology ("schedule") of events listed down the left column. Each object/entity is denoted by a vertical "lane" and arrows represent interactions of various kinds. Unlike a class diagram, which is mostly static, a sequence diagram represents the main dynamics occurring in the system, and the precise order of interaction between components. Note that the `Polity` is not explicitly denoted, because all dynamics are internal to it. For comparative purposes, note that a sequence diagram is equivalent to a directed graph, in the sense of social network analysis. Therefore, a sufficiently well-specified sequence diagram can be quantitatively described by a vector of network metrics at both node and network levels.

The key internal process of a social simulation is arguably the so-called "main simulation loop."

Definition 7 (Main simulation loop): The main simulation loop of a social simulation is a discrete process that details what happens one step at a time between consecutive updates of the state of the model.

Several types of graphics can be used for describing the main simulation loop of a model. Flowcharts are the most common, but can be complicated.

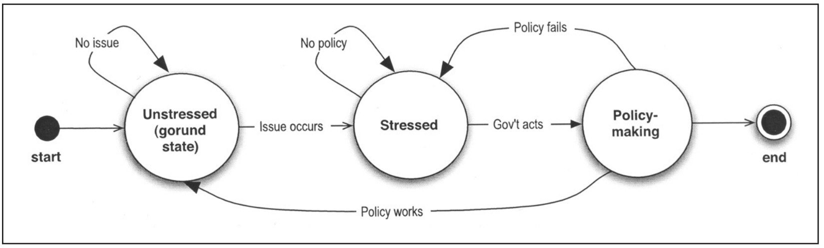

Figure 4. Example of a UML state diagram for a basic polity

UML sequence diagrams can be more informative, but sometimes are difficult to design when many classes of objects interact. Main simulation loops can be compared directly using one or more of these types of graphics (Cioffi & Gotts, 2003). From a cross-cultural perspective, the main simulation loop of a model highlights cultural similarities and differences among group relations—features that have significant implications for social dynamics. For example, social features and interaction patterns such as complementary opposition, cultural taboos, competing authorities, ethnic traditions, and other social conditions can be represented by unique patterns in sequence or flowchart diagrams.

A complementary dynamic representation for comparative analysis is provided by the UML state diagram.

Definition 8 (State diagram): A UML state diagram (Figure 4) is a graphic representation of the set of main states and all feasible (i.e., practically possible) interstate transitions of the system.

The state transition diagram in Figure 4 (also called "finite-state machine" diagram) illustrates the set of discrete conditions ("states") in which the simple `Polity` system can find itself while it operates. Unlike the earlier two diagrams, the state diagram is "systemic" or at the system-level, because it portrays the state of `Polity` as a whole, "in the aggregate," independent of the state of constituent components. A well-specified state diagram can also be described as a directed graph or as a matrix with states as rows/columns and elements denoting possible transitions.

UML diagrams for comparing systems are used in addition to more conventional *flowchart diagrams* for specifying a main simulation loop and other important processes within a given simulation model.

In the case of all four MASON models from the Mason-HRAF Eastern Africa project, these are rendered ("instantiated") as discrete event simulation models, meaning that basic processes driving social and biophysical dynamics are rendered as time steps, not continuous processes. For example, in RebeLand each household assesses its own well-being in terms of a wellness function, and based on such an assessment decides what to do next. Possible choices might include continuing to earn an income or, if greatly dissatisfied, joining an insurgency. Each of these steps is rendered as an event. Similarly, government agents (for instance, the State) assess their situation and decide what to do next. Governmental action may include producing additional security agents to fight insurgents. When encountering government security agents, insurgents might decide to flee or engage, depending on circumstantial factors. These dynamics are represented and compared using UML sequence and state transition diagrams such as those in Figures 3 and 4.

Comparisons of social simulations by other dimensions, such as semantics, pragmatics, and genetics focus specifically on issues of meaning, intended use (every model aims at answering some core question), and origin, respectively. For instance, whereas some simulations are primarily intended for pure scientific research, others are intended for policy purposes (pragmatics). Similarly, the meaning of certain terms (e.g., “agents”) may agree or differ across simulations being compared.

Summary

Melvin (“Mel”) Ember was co-Principal Investigator in the Mason-HRAF Joint Project on Eastern Africa, a multiyear project funded by the Office of Naval Research under the Multi-University Research Initiative (MURI) Program. The project aims at developing and analyzing new advanced computational agent-based models of human societies across 10 countries and numerous ecosystems in Eastern Africa, including Kenya, Uganda, Tanzania, Rwanda, Burundi, southern Somalia, southern Ethiopia, southeastern Sudan, eastern Democratic Republic of Congo, and northern Tanzania.

A major unsolved challenge in this kind of social science research is to devise a systematic way to compare, communicate, and contrast different models along relevant dimensions and characteristics, given the inherent complexity of most computational agent-based models. This article presented a viable systematic framework for comparing models and illustrated its application using some of the models that Mel helped inspire and develop as senior project participant.

The proposed framework is based on two kinds of comparisons: generic and specific. Generic comparisons pertain to comparative dimensions that are common to all social simulations, including comparisons among focal systems (empirical domain), types of computational models (programming language), process of abstraction, model validation (external validity), and data requirements (as determined by the former dimensions). Specific comparisons pertain to narrower formal issues, such as detailed ontology (main classes and associations) and dynamics. The latter are represented by UML class, sequence, state, and flowchart diagrams to enable comparative analysis.

Acknowledgments

The author thanks the members of the Mason-HRAF Joint Project on Eastern Africa (MURI Team) for helpful criticism and comments.

Author's Note

An earlier version of this article was presented at the special session in honor of Melvin Ember, at the Annual Conference of the Society for Cross-Cultural Research (SCCR) and Society for Anthropological Sciences (SASci), Albuquerque, New Mexico, United States, February 17 to 21, 2010. The opinions, findings, and conclusions or recommendations expressed in this work are those of the author and do not necessarily reflect the views of the sponsors. This article is dedicated in memory of Melvin ("Mel") Ember.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to authorship and/or publication of this article.

Financial Disclosure/Funding

Funding for this study was provided by the Center for Social Complexity of George Mason University and by ONR MURI grant no. N00014-08-1-0921.

Notes

1. It seems likely but unproven that much earlier, preclassical scholars of the 2nd or 1st millennium BC had developed comparative schemes for different types of social systems, given the lost libraries of Alexandria, the Near East, China, and elsewhere in the ancient world.
2. For example, Almond, Powell, Dalton, & Kaare Strom (2006); Bartolini, (1993); Dahl, (1984); Deutsch et al. (1981); Ember & Ember (2010); Goldstone (2003); Ito (1997); Landman (2008); Lave & March (1993); Lichbach & Zuckerman (1997); Przeworki & Teune (1970); Ragin (1987); Saberwal (1987); Sartori (1991).
3. Lave & March's (1993) evaluation framework—in terms of "truth, beauty, and justice" can also be used for comparing models. However, the intent here is more specialized.
4. For recent surveys of social simulation in computational social science, see, for example, Carley & Gasser (1999); Cioffi (2010a); Gilbert & Troitzsch (2005); Kuznar (2006); Taber & Timpone (1996); Takadama, Cioffi-Revilla, & Deffaunt (2010); Terano & Sallach (2007).
5. Comparative analysis of computational social science models is in its infancy, so there is a lack of consensus on a widely shared framework or comparative standard.

A unified framework may be premature, given the rapid pace of methodological developments in the area of computational modeling. This includes (but not exclusively) agent-based modeling. Examples of proposed frameworks include Cioffi & Gotts (2003), Grimm et al. (2006), Janssen, Barton, Alessa, Bergin, & Lee (2008), and Ostrom (2009). The framework in this paper draws on these proposals and additional ideas generated by the Mason-HRAF Joint Project on Eastern Africa (Cioffi, 2010c; Cioffi, De Jong, and Bassett, 2010; Cioffi & Rouleau, 2009, 2010; Hailegiorgis et al., 2010; Kennedy et al., 2010a, 2010b) and the Mason-Smithsonian Joint Project on Inner Asia (Cioffi et al., 2007; Cioffi, Rogers, & Latek, 2010; Rogers, 2007; Rogers & Cioffi, 2009).

6. The simulation literature also uses the term "target system" (Gilbert & Troitzsch, 2005).
7. The MASON system (Luke et al., 2005) is available at <http://cs.gmu.edu/~eclab/projects/mason>. The MASON web site contains numerous references, examples of MASON models, and links to other multiagent simulators, such as RePast, Ascape, Swarm, and NetLogo. The MASON System Project was initially funded by the Mason Center for Social Complexity and is a collaborative project with the Mason Evolutionary Computation Laboratory (ECLab). Additional funding for MASON has been received from the U.S. National Science Foundation, DARPA, and the Office of Naval Research.
8. These and other simulation features are described in the respective paper-of-record of each model. See also Cioffi, 2002, for a set of related comparative dimensions.

References

Almond, G. A., Powell, G. B., Dalton, R. J., & Kaare Strom. (2006). *Comparative politics today: A world view*. New York, NY: Pearson Longman.

Ambler, S. W. (2005). *The elements of UML 2.0 Style.* New York, NY: Cambridge University Press.

Aristotle. (1984). *The politics* (C. Lord, Trans.). Chicago, IL: University of Chicago Press.

Barker, J. (2005). *Beginning Java objects: From concepts to code* (2nd ed). Berkeley, CA: Apress.

Bartolini, S. (1993). On time and comparative research. *Journal of Theoretical Politics, 5*, 131-167.

Carley, K. M., & Gasser, L. (1999). Computational organizational theory. In G. Weiss (Ed.), *Multiagent systems: Modern approach to distributed artificial intelligence* (pp. 299-330). Cambridge, MA: MIT Press.

Cioffi-Revilla, C. (2002). Invariance and universality in social agent-based simulations. *Proceedings of the National Academy of Science of the USA, 99*(Suppl. 3, 14), 7314-7316. Retrieved from http://www.pnas.org/content/99/suppl.3/7314.full.pdf+html

Cioffi-Revilla, C. (2008). Simplicity and reality in computational modeling of politics. *Computational and Mathematical Organization Theory, 15*(1), 26-46.

Cioffi-Revilla, C. (2010a). Computational social science. *Wiley Interdisciplinary Reviews: Computational Statistics, 2,* 259-271. doi:10.1002/wics.95. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1708051

Cioffi-Revilla, C. (2010b). On the methodology of complex social simulations. *Journal of Artificial Societies and Social Simulations, 13*(1), 7. Retrieved from http://jasss.soc.surrey.ac.uk/13/1/7.html

Cioffi-Revilla, C. (2010c). MASON RebeLand and data aspects of agent-based simulation models. In D. Schmorrow & D. Nicholson (Eds.), *Advances in cross-cultural decision making* (pp. 291-301). Orlando, FL: CRC Press/Taylor & Francis.

Cioffi-Revilla, C., & Gotts, N. (2003). Comparative analysis of agent-based social simulations: GeoSim and FEARLUS models. *Journal of Artificial Societies and Social Simulations, 6*(4), 10. Retrieved from http://jasss.soc.surrey.ac.uk/6/4/10.html

Cioffi-Revilla, C., Luke, S. C., Parker, D. C., Rogers, J. D., Fitzhugh, W. W., . . . Amartuvshin, C. (2007). Agent-based modeling simulation of social adaptation and long-term change in inner Asia. In T. Terano & D. Sallach (Eds.), *Advancing social simulation: The first world congress in social simulation* (pp. 189-200). New York, NY: Springer.

Cioffi-Revilla, C., Kenneth De Jong, K. D., & Jeffrey Bassett, J. (2010). *Evolutionary computation and agent-based modeling: Biologically-inspired approaches for understanding complex social systems* (Working Paper, Mason-HRAF Joint Project on Eastern Africa). Manuscript submitted for publication.

Cioffi-Revilla, C., Rogers, J. D., & Latek, M. (2010). The MASON HouseholdWorld of Pastoral Nomad Societies. In K. Takadama, C. Cioffi-Revilla, & G. Deffaunt (Eds.), *Simulating interacting agents and social phenomena: The second world congress in social simulation* (pp. 193-204). Berlin, Germany: Springer.

Cioffi-Revilla, C., & Rouleau, M. (2009, June 23-24). *MASON AfriLand: A regional multi-country agent-based model with cultural and environmental dynamics*. Proceedings of the Human Behavior-Computational Modeling and Interoperability Conference 2009 HB-CMI-09. Joint Institute for Computational Science, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Cioffi-Revilla, C., & Rouleau, M. (2010). MASON RebeLand: An agent-based model of politics, environment, and insurgency. *International Studies Review, 12*(1), 31-46.

Clark, J., Cumow, R., & Cole, S. (1975). *Global simulation models: A comparative study*. New York, NY: John Wiley.

Dahl, R. A. (1984). *Modern political analysis* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Deutsch, K. W. (1948-1949). Some notes on research on the role of models in natural and social science. *Synthese, 7*, 532-533.

Deutsch, K. W., Dominguez, J. I., & Heclo, H. (1981). *Comparative government: Politics of industrialized and developing nations*. Boston, MA: Houghton Mifflin.

Ember, C. R., & Ember, M. R. (2010). *Cultural anthropology* (13th ed). Englewood Cliffs, NJ: Prentice Hall.

Eriksson, H.-E., Penker, M., Lyons, B., & Fado, D. (2004). *UML 2 Toolkit.* New York, NY: Wiley.

Gilbert, N., & Troitzsch, K. G. (2005). *Simulation for the social scientist* (2nd ed.). Buckingham, UK: Open University Press.

Goldstone, J. A. (2003). Comparative historical analysis and knowledge accumulation in the study of revolutions. In J. Mahoney & D. Rueschemeyer (Eds.), *Comparative historical analysis in the social sciences* (pp. 41-90). Cambridge, UK: Cambridge University Press.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot V., Giske J., . . . DeAngelis. (2006). A standard protocol for describing agent-based models. *Ecological Modelling, 198*(1-2), 115-126.

Grimm, V., Revilla, E., Berger,U., Jeltsch, F., Mooij, W. M., Steven F. Railsback, S. F., . . . DeAngelis. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science, 310*, 987-991.

Hailegiorgis, A. B., Kennedy, W. G., Balan, G. C., Bassett, J. K., & Gulden, T. R. (2010, July 5-8). *An agent based model of climate change and conflict among pastoralists in East Africa*. Proceedings of the International Congress on Environmental Modeling and Software (IEMSS2010), Ottawa, Ontario, Canada.

Hales, D., Rouchier, J., & Edmonds, B. (2003). Model-to-model analysis. *Journal of Artificial Societies and Social Simulation, 6*(4), 5. Retrieved from http://jasss.soc.surrey.ac.uk/6/4/5.html

Hanneman, R. A. (1988). *Computer-assisted theory building: Modeling dynamic social systems*. Newbury Park, CA: SAGE.

Ito, S. (1997). A framework for comparative study of civilizations. *Comparative Civilizations Review, 26*, 4-15.

Janssen, M. A., Barton, C. M., Alessa, L. N., Bergin, S., & Lee, A. (2008). Towards a community framework for agent-based modeling. *Journal of Artificial Societies and Social Simulation, 11*(2), 6. Retrieved from http://jasss.soc.surrey.ac.uk/11/2/6.html

Kahn, K. (2007, March 15-16). Comparing multi-agent models composed from micro-behaviours. In J. Rouchier, C. Cioffi-Revilla, G. Polhill, & K. Takadama (Eds.), *Proceedings of M2M 2007: Third International Model-to-Model Workshop, Marseille, France* (pp. 165-177). Retrieved from http://m2m2007.macaulay.ac.uk/M2M2007Kahn.pdf

Kaplan, A. (1964). *The conduct of inquiry*. San Francisco, CA: Chandler.

Keith S., Mark C., Sean L. (2010). GeoMASON. Department of Computer Science, George Mason University, 4400 University Drive MSN 4A5, Fairfax, VA 22030-4444, USA. Retrieved from http://cs.gmu.edu

Kennedy, W. G., Hailegiorgis, A. B., Rouleau, M., Bassett, J. K., Coletti, M., Balan, G. C., & Gulden, T. R. (March, 2010). *An agent-based model of conflict in East Africa and the effect of watering holes*. Proceedings of the 19th Conference on Behavior Representation in Modeling and Simulation (BRiMS). Columbia, South Carolina. Retrieved from http://brimsconference.org/archives/2010/papers/19%20- %20Kennedy.pdf

Kuznar, L. A. (2006). High-fidelity computational social science in anthropology. *Social Science Computer Review, 24*(1), 15-29.

Kuznar, L. A., & Sedlmeyer, R. (2005, October). Collective violence in Darfur: An agent-based model of pastoral nomad/sedentary peasant interaction. *Mathematical Anthropology and Cultural Theory: An International Journal, 1*(4).

Landman, T. (2008). *Issues and methods in comparative politics: An introduction* (3rd ed.). London, UK: Routledge.

Lau, Y. T. 2001. *The Art of Objects: Object-Oriented Design and Architecture*. Boston: Addison-Wesley.

Lave, C. A., & March, J. G. (1993). *An introduction to models in the social sciences*. Lanham, MD: University Press of America.

Lichbach, M. I., & Zuckerman, A. S. (Eds.). (1997). *Comparative politics: Rationality, culture, and structure*. New York, NY: Cambridge University Press.

Lowe, J. W. G. (1985). *The dynamics of apocalypse: A systems simulation of the classic maya collapse*. Albuquerque, NM: University of New Mexico Press.

Luke, S., Claudio C.-R., Liviu P., & Keith S. 2005. MASON: A Java Multi-Agent Simulation Environment. *Simulation: Transactions of the Society for Modeling and Simulation International, 81*(7), 517-527.

Nikolai, C., & Madey, G. (2009). Tools of the trade: A survey of various agent based modeling platforms. *Journal of Artificial Societies and Social Simulations, 12*(2). Retrieved from http://jasss.soc.surrey.ac.uk/12/2/2.html

Ostrom, E. (2009). A general framework for analyzing sustainability of socioecological systems. *Science, 325*, 419-422.

Przeworski, A., & Teune, T. (1970). *The logic of comparative social inquiry*. New York, NY: Wiley.

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley, CA: University of California Press.

Rogers, J. D. (2007). The contingencies of state formation in eastern Inner Asia. *Asian Perspectives, 46*, 249-274.

Rogers, J. D. and Claudio, C. R. (2009). Expanding Empires and the Analysis of Change. *Current Archaeological Research in Mongolia* (pp. 445-459). J. Bemann, H. Parzinger, E. Pohl and D. Tseveendorzh Bonn (Eds.). Germany: Bonn University Press.

Rouchier, J., & Bousquet, F. (2001). A multi-agent model for transhumance in North Cameroon. *Journal of Economic Dynamics and Control, 25*, 527-559.

Rouchier, J., Cioffi-Revilla, C., Polhill, J. G., & Takadama, K. (2008). Progress in model-to-model analysis. *Journal of Artificial Societies and Social Simulation, 11*(2), 8. Retrieved from http://jasss.soc.surrey.ac.uk/11/2/8.html

Rouchier, J., Edmonds, B., & Hales, D. (Eds.). (2003, March 31-April 1). Proceedings of the International Workshop M2M Model to Model, EHESS, Groupement de Recherche en Economie Quantitative d'Aix-Marseille GREQAM, Vielle Charit, Marseille, France.

Ruloff, D. (1978). Historical empires in comparison: Five computer-constructed ideal types. *Peace Science Society (International) Papers, 28*, 81-96.

Ruloff, D. (November, 1981). *Computer simulation in history: The case of the classic Maya collapse*. Paper presented at the 22nd Annual Convention of the International Studies Association, Philadelphia, Pennsylvania.

Saberwal, S. (1987). A comparative analysis of very long-term processes. *International Political Science Review, 8*, 307-318.

Sallach, D. L., & Macal, C. N. (2001). The simulation of social agents: An introduction. *Social Science Computer Review, 19*(3).

Sartori, G. (1991). Comparing and miscomparing. *Journal of Theoretical Politics, 3*, 243-257.

Sterman, J. D. (2000). *Business dynamics: System thinking and modeling for a complex world*. Boston, MA: McGraw-Hill.

Taber, C. S., & Timpone, R. J. (1996). *Computational modeling*. Thousand Oaks, CA: SAGE.

Takadama, K., Cioffi-Revilla, C., & Deffaunt, G. (Eds.). (2010). *Simulating interacting agents and social phenomena: The second world congress in social simulation*. New York, NY: Springer.

Terano, T., & Sallach, D. (Eds.). (2007). *Advancing social simulation: The first world congress in social simulation*. New York, NY: Springer.

Bio

Claudio Cioffi-Revilla, PhD, is professor of computational social science and founding director of the Center for Social Complexity, George Mason University. His research interests are in modeling human and social dynamics, especially conflict and origins and development of social complexity. He has written six books on quantitative and mathematical approaches to social science research, and more than 60 papers in peer-reviewed journals and edited volumes. His research is funded by NSF, DARPA, ONR. He serves as a National Academies Jefferson Science Fellow at the U.S. Department of State, Office of The Geographer, and is a research scientist at the Smithsonian National Museum of Natural History in Washington, DC.