CATEGORICAL SOCIAL SCIENCE: THEORY, METHODOLOGY AND DESIGN

David L. Sallach

Center for Complex Adaptive Agent Systems Simulation
Argonne National Laboratory
sallach@anl.gov

Computation Institute, University of Chicago
sallach@uchicago.edu

ABSTRACT

Computational social science (CSS) has created innovative forms of modeling and analysis. However, now CSS appears to be approaching a significant barrier in the form of *complexity ceiling* that imposes constraints throughout the modeling lifecycle, from design to validation. Given the precedent of the physical sciences, a 'natural' solution is to reformulate the domain at ever-higher levels of mathematical abstraction. Following this example, an appropriate set of mathematical techniques is a prerequisite to establishing a foundation for CSS. The present explication makes the case that category theory will provide an effective mathematical framework capable of representing social complexities.

INTRODUCTION

The emergence of agent-based modeling and simulation (ABMS) has significantly contributed to the development of dynamic models in the social sciences, models the assumptions of which are explicit and explorable, and the evolution of which can be theoretically and/or experimentally investigated. Exemplars have been developed that illustrate notional conceptual models (Schelling 1978; Epstein & Axtell 1996; Axelrod 1997; Cederman 1997), rich empirical models (Kohler & Gumerman 2000) and nascent efforts at theoretically informed models (Cederman 2001; 2008; Sallach 2000; 2003).

The promise of social agent modeling and, more generally, CSS, is widely recognized (Gilbert & Troitzsch 2005; Epstein 2006; Miller & Page 2007; North & Macal 2007; Lin & Carley 2010). The prospect of representing social processes at multiple scales, at an arbitrary level of detail, is likely to produce rich, interwoven models with the potential to provide a more effective map of historical and policy-oriented dynamics. However, this potential will also raise deep challenges in validating such a model (Sallach 2011).

One approach to managing the complexity (both conceptual and computational) of social models is to move them toward a higher level of abstraction. Specifically, it is potentially important to more fully formulate the mathematical basis of CSS. The focus of the present paper is to consider how category theory (Mac Lane 1971; Landry 1998; Awodey 2010) provides a critical resource in formalizing computational models, as well as exploring their potential interaction.

Subsequent sections of the paper address: 1) the nature of category theory, 2) how categories can be used to integrate both domains and theories, 3) how categorical models can open up innovative forms of social analysis, and 4) what category theory can contribute to model design.

CATEGORY THEORY

Category theory requires specification of a type of mathematical object in terms of identity, composition, and a group of morphisms that preserves their structure. This definition of objects and morphisms, in conjunction with two standard axioms (associativity & identity), defines a framework that is simple enough to be broadly applicable. A collection of illustrative mathematical categories is listed in Table 1. These definitions and axioms constitute unique but similar regularities within each category type that focus inference.

CATEGORIES	**OBJECTS**	**MORPHISMS**
Set	Sets	All functions between sets
Poset	Partially ordered sets	Monotone functions
Top	Topological spaces	Continuous functions
Vect	Vector spaces	Linear maps
Field	Fields	Field homomorphisms
Groupoid	Diverse objects	Invertible morphisms
Topos	Set-like objects	Allows integration of algebra & logic
Log	Formulae	Proofs
Cat	Categories	Functors

Table 1. Illustrative Categories

Even with the limited set of examples in Table 1, it is striking to observe how broadly category theory is able to integrate diverse areas of mathematics: set theory, topology, linear algebra, lattice theory and formal logic (Goldblatt 1984; Lambek & Scott 1988), and the list is actually much longer. It not only allows them to be expressed in the same categorical terms, it also provides a basis for their possible integration. As Marquis (2009) carefully documents, category theory supports the formal integration of geometry, topology, algebra and logic. The potential of this integral formalism has yet to be fully realized, especially in the social sciences.

It should be noted that the object type of a category is actually secondary. As Lawvere (2005:2) writes, "[C]ategory theory does not rest content with mere classification … rather, it is the *mutability* of mathematically precise structures (by morphisms) which is the essential content of category theory." To the extent that appropriate representations can be identified, the emphasis on mutability can contribute to the dynamics that will be needed in social categorical analysis.

CATEGORICAL INTEGRATION

The forms of mathematical integration implied in Table 1 are important for CSS as well. Geometric categories can make GIS patterns available for deductive inference. But perhaps most interesting, is the integration of algebraic and logical models. The synthesis of algebra and logic, while not unique to CSS, has the potential to provide a basis for forms of analysis that are as yet unprecedented.

Integration of Multiple Substantive Domains

To achieve its potential, CSS requires strong forms of integration. From the outset, this emerging discipline has rested upon the effective fusion of the computational and the social sciences. It is thus relevant that category theory has already made significant contributions to computer science (Goguen 1991; Pierce 1991). Of particular relevance is the use of category theory in software engineering (Fiadeiro 2005). To the extent that substantive social science and software design can be integrated into the same mathematical framework, it has the potential to transform the process of software design and, most importantly, model validation.

For some models, the incorporation of neurological models is important. Here too, innovative work has been accomplished in the categorical representation of neurological processes (Ehresmann & Vanbremeersch 2007).

Theory integration

The second facet of categorial integration concerns the task of synthesizing social theory. Social science theories are distinguished by discipline, scale and other factors, and are often qualitative (Hage 1994). The varieties of theoretical fragmentation serve to undermine the completeness and coherence of social models, and reduce the possibility of their validation. The long-term success of CSS will require the coherent integration of theoretical insights in all their variety.

The characteristics of category theory that have allowed it to provide a common language to diverse mathematical areas (Landry 1998) have the potential to be applied to the synthesis of social theory in all of its diversity. However, to play that role, theorists and modelers will need to understand category theoretic tools, and how they can be applied. Table 2 suggests one way in which categorical structures can be translated into social applications.

Category Theory	Social Category Theory
Invariants	Ideal types
Natural transformations	Progressions
Symmetry	Reciprocity
Adjoint functors	Coupled processes, transitions
Categories	Social configurations & scenario templates
Morphisms	Actions, mechanisms & processes
Objects	Actors, aggregates

Table 2. Mapping from Categories to Social Categories

If, for a social model, we view the objects of category theory as social actors, then, depending on the type of behavior being modeled, the appropriate kinds of morphisms (e.g., actions or mechanisms) can be identified. Alternatively, the lifecycle of a social process, for example, the ecology of organizations (Hannan & Freeman 1989), could provide the starting point for scenarios, with characteristic interactions then being specified as the model definition proceeds. The hierarchy can be built either top down or bottom up, i.e., depending on the relevant theories and scales of interest, as long as it draws upon the strengths of category-theoretic insights.

Social Objects. Table 2 provides an overview of how a categorical hierarchy could translate into social equivalents, in skeletal form. Of course, to construct useful categorical models, social theory and social processes will need to be closely examined and refined in order to create the most effective alignment between category theory and social concepts.

Table 3 provides a view of one aspect of Table 2 (i.e., potential 'social' objects) in greater depth. The first point is that, for different theoretical or research purposes, level 0 may be filled using different types of social objects. In Table 3, there are three types of candidate social 'objects': 1) social actors, 2) social aggregates, and 3) geostrategic niche. These three examples are introduced, not because they are ideal, or an exhaustive set of social objects, but, rather, to suggest some of the considerations that may be involved in the selection and refinement of social objects.

The first example is social actor, the social object originally suggested in Table 2. Only, now it is possible to consider it in depth. The emphasis is upon assertive actors that make decisions and/or adapt to changed circumstances. Note that a social actor is defined as existing at multiple scales, and the third column illustrates examples of such actors.

All such candidates necessitate detailed conceptual analysis. Can social actors be modeled in the same way across scales? Will there need to be supplemental cross-scale morphisms? Will cross-scale interactions differ from homoscale interactions? Many such issues are distinct from those that are typically addressed in the substantive social science disciplines.

SOCIAL OBJECTS	**DEFINITION**	**EXAMPLES**
Social actor	Self-organizing social entity at any scale considered in an active mode. A social actor makes decisions and actively seeks to elicit coordination from internal and external actors.	Alliance, civilization, empire, nation, institution, state, organization, class-for-itself, social network, group, Median social self (I)
Social aggregate	Collection of social actors considered in a passive mode. A social aggregate is shaped by exogenous structures, events and processes, changing through proximate adaptation and evolution.	Population, aggregation, category, strata, class-in-itself, generation, Median social self (me)
GeoStrategic Niche	An integrated natural and social system considered in the interaction of its geological, biological , ecological, economic and strategic facets.	Bioneural, interactive and transportational movements

Table 3. Candidate Social Objects

The second social object is the social aggregate. It is designed to represent more passive responses to exogenous influences. The action generated may be adaptive, the result local intentionality and/or mixed motives. Often aggregate social objects are used to study the effects of largely exogenous structures or events upon largely broad diverse populations or population sectors.

The exogenous effects can range from the spread of disease to the creation of refugees by war or tyrannies. The distinction between social actors and social aggregates is that morphisms in the former are typically intentional and/or strategic while, in the latter, they are more passive or indirect. Both forms exist within human groupings, but their modes of transformation differ, making it analytically advantageous to distinguish them.

A third candidate social object is the geostrategic niche. This form is characterized by a co-definition of the social object with its environment (Sallach, North & Rivera 2012). Two examples are: 1) conflicting actors within geographical constraints and affordances (cf., Collins 1978), and 2) economic actors adapting to a given ecological niche (Peterson, *et al.,* 2011). More specialized examples may be expected to be utilized as well.

The three candidate social objects described in Table 3 are illustrative, not exhaustive. Different kinds of social objects can be defined for different purposes, and their definition is a salient along which more rigorous social theory can progress. Accordingly, the definition of increasingly specific types of social objects, in a categorical sense, is an essential line of research.

Social Morphisms. Social models are distinguished from those of the natural sciences by the need to represent intentional actors. As a result of this requirement, there must be: 1) a greater emphasis upon endogenous factors, which are 2) more dynamic, and 3) often interact with each other.

The relationship between exogenous and endogenous factors will comprise another important research area of categorical social science. In application, the former are apt to be analytical, while the latter are more likely to be simulative.

One of the most effective ways of balancing exogenous and endogenous forces is to distinguish the social objects to which they are applicable. This is what the distinction between social aggregates (exogenous) and social actors (endogenous) is designed to do.

It might seem that this could create a false separation that would undercut an analyst's ability to fully capture social dynamics. However, category theory is inherently flexible in its application. One of the first forms of integration that it affords is the integration among two or more types of social objects by mapping one to the other. The general strategy, then, is to identify social objects with coherent morphisms, and then reintegrate them as required by the particular form of analysis.

This strategy is suggestive, but a richer view of what is possible is necessary. Accordingly, this discussion now turns to the features that a categorical methodology will include.

CATEGORICAL METHODOLOGY

Categories exist at a high level of abstraction, and become highly specialized in applications (cf., Crane & Yetter 1993; Ehresmann & Vanbremeersch 2007). Thus, it is reasonable to assume that a categorical methodology will be required. Such a methodology can provide heuristics and guidelines about how category-theoretical insights can be incorporated into social science research. It also illustrates the type of contributions that can be expected by doing so.

An overview of various aspects of a categorical methodology for the social sciences is summarized in Table 4. The first row, categorical diagrams, notes that category theory is frequently expressed graphically. Given its topological and geometric influences, many shapes and structures are used as a basis for categorical reasoning. However, the prototypical diagram is the triangle, as depicted in Figure 1.

Under standard conventions, the capital letters are objects and the arrows are morphisms. Circular arrows are identity morphisms that also serve as the unit of the category. Straight arrows convey information from the Source (at the base) to the Target (at the tip). This particular triangle shows information being conveyed from A to C in either a one-step or two-step process. If the information at C is the same, the diagram is said to commute, and the two paths can be interpreted as two sides of an equation. A more concrete interpretation will depend, in part, on the nature of the (level 0) objects and their associated morphisms.

Heuristic	**Description**	**Source**
Categorical diagrams	Category theory employs graph-theoretic objects to organize (and sometimes chain) mathematical relations, and to develop proofs.	Fiadeiro (2005: 15-20)
Equivalence relations	An ability to define isomorphism up to equivalence facilitates analysis of ranges of social values.	Goldblatt (1984: 61-63)
Social dualities	Duality permeates categorical analysis. Given a duality with a proper connection between them, it is possible to define one concept in terms of the other.	Awodey (2010: 53-73)
Abstraction scaffolding	In category theory, there is a progression from mathematical objects, to morphisms, to functors to natural equivalences. This abstraction hierarchy provides a systematic way of introducing and controlling higher levels of analysis.	Baez & Dolan (1998)
Endogenous social metrics	A formal linear combination of the elements of some relevant set provides a natural basis for analysis, rather than in contrast to an imposed metric.	Baez (2001)
Process projections	There are a variety of category-theoretical techniques for projecting of empirical patterns.	McCleary (2001)

Table 4. Aspects of a Social Categorical Methodology

Category theory supplements strict equality with weakened definitions of equivalence relations and/or classes. The resulting ability to formalize and reason about regions will make a valuable contribution to social analysis. Many social concepts involve proximity assessments of, for example, stability, trust, legitimacy, preparedness, strength, income and other complex states. The potential is to address previously qualitative concepts with greater precision.

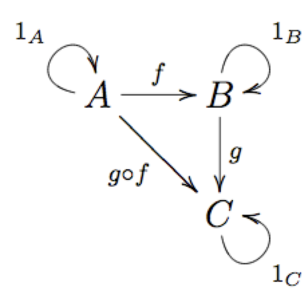

Figure 1. Prototypical Category Diagram

A specific example of is illustrated in Figure 2. It shows how reciprocal moves between two actors are defined by equivalence relations.

Specifically, for each game type (beneficent, instrumental and coercive, coded as blue, green and red, respectively), there is a range of possible moves, each of which is considered equivalent to the other. This coding allows for a variation of responses, anticipated and actual, that is understood to be equivalent by both actors A and B. It is also possible to introduce variances within the ranges, which may result in misunderstanding between the two actors.

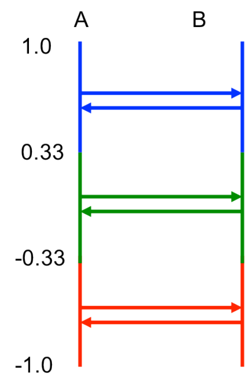

Figure 2. Equivalent Moves within Three Protogames

The third heuristic is the potential for the discovery and analysis of social dualities. When isomorphisms have an inverse, or adjoint functors exist between two categories, the resulting duality gives rise to innovative forms of analysis that are particularly applicable to coupled phenomena (Lawvere 2006; Marquis 2009). This form of analysis is likely to be especially relevant to various forms of co-evolution.

What Table 4 summarizes as abstraction scaffolding has been previously depicted in Table 2. Morphisms are applied to objects, which together define a category. Categories change through functors, and natural (social) transformations (when available) weave these functors into a fabric. In this way, a hierarchy of meta-level interaction is defined. It provides a coherent framework for analysis in many areas of investigation, and it seems likely that it will contribute to CSS as well.

The fifth row is not only a heuristic, but a caveat as well. Artificial metrics can obscure rather than clarify. This is an insight that has manifested itself in the social sciences as well, for example, when class distinctions have sometimes been attributed based at arbitrary breaks applied to continuous wealth or income data. However, as the analytical focus moves toward the identification of equivalence classes, it is a useful caveat to remember.

In category theory, there are a variety of techniques (here grouped together under the term sequence projection) for analyzing or projecting fine grain differences, including variation over time. These include fiber bundles, fibrations & cofibrations, sheaves and spectral sequences. Referencing such representations in the present discussion serves to note that, in addition to abstract theorization, category theory contains the means of addressing more detailed processes as well, and that these techniques are available to a methodology designed for CSS.

CATEGORICAL SOFTWARE DESIGN

Among the wide range of domains addressed by category theory, software engineering is particularly important. Not only are categorical designs natural recipients of social theories formulated categorically, the coherence of such categorical transformations will make a direct and significant contribution to the process of validating complex social models.

Software engineering based on category theory makes available the extensive range of categorical structures and processes that have been generated in abstract categorical analysis, including dual constructions, pushouts and pullbacks, limits and colimits, and adjunctions and indexed categories (Fiadeiro 2005; Iordache 2011). These techniques are applied to processes that engage in concurrent interaction. Moreover, Fiadeiro regularly invokes the social life of categories as a way of reasoning about the fact they are more fundamentally defined by interfaces with other processes than by internal characteristics. Although this "social life" is primarily an analogy, it underscores the broad compatibility of category theory with social dynamics.

ABMS Design. Given the fundamental role that agent-based systems play within CSS, the application of category-theoretic design to ABMS models is quite pertinent. An initial example is provided by Lloyd (2010), who empoys a cluster of empirically-based Bayesian techniques as a basis for generic ABMS design.

Specifically, Lloyd proposes using forward and inverse posterior probability densities to recognize structures and infer a relevant rule set, generate fractional realizations of such structures, from which dual forward/inverse cycles shape and realize agent actions. An example is applied to An's (2009) model of Systematic Inflammatory Response Syndrome (SIRS)

There are many ways that category-theoretic design techniques might be applied to specific domains (cf., Sallach 2012), some with a more social-theoretical strategy than the Lloyd's approach provides. Nevertheless, the Lloyd example serves as reinforcement of the primary premise of the present discussion. The ability of category theory to represent and address a diverse set of domains (including mathematics, computer science, social theory, social empirics and software design) documents the critical role that category theory can play in unifying and expressing the intertwining of a variety of complex scholarly domains.

CONCLUSION

The present exposition describes more than can currently be demonstrated. However, it does make a case for the position presented, and illustrates a variety of ways in which these conclusions are compelling.

Category theory provides a rigorous yet expressive formalism for representing and integrating challenging modeling domains. It is also a formalism that can support extensive theoretical and modeling syntheses, while still maintaining appropriate levels of exactness. These strengths make CT particularly appropriate for providing a mathematical foundation for computational social science.

Ultimately, it is important to return to the issue of the validation of social models. The practical use of historical and/or policy-oriented models is highly dependent upon their credibility. Category theory provides a formalism that can be precise when applied to theory, explicit during the design stage, and definitive at the assessment stage, while also attending to the consistency of intermediate stages. It deserves serious exploration and investigation.

Beyond what has been presented here, higher categories are an active area of mathematical research (Leinster 2004; Brown 2006; Simpson 2012). Because they seek to leverage weak equivalences (within various competing formalisms: Bergner 2010) to incorporate greater complexity, higher categories will be of great and continuing interest in the future of computational social science.

REFERENCES

An, Gary. 2009. Introduction of an agent-based multi-scale modular architecture for dynamic knowledge representation of acute inflammation. *Theoretical Biology and Medical Modeling* 5:11.

Awodey, Steve. 2010. *Category Theory.* New York: Oxford University Press.

Axelrod, Robert. 1997. *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration.* Princeton, NJ: Princeton University Press.

Baez, John C. & James Dolan. 1998. Categorification. Pp. 1-36 in E. Getzler & M. Kapranov, eds. *Higher Category Theory.* Providence, RI: American Mathematical Society.

Baez, John C. 2001. Higher-dimensional algebra and Planck-scale physics. Pp. 177-195 in C. Callender & N. Huggett, *Physics Meets Philosophy at the Planck Level: Contemporary Theories in Quantum Gravity.* New York: Cambridge University Press.

Bergner, Julia E. 2010. A survey of $(\infty, 1)$-categories companion. Pp. 69-83 in J.C. Baez & J.P. May, eds., *Towards Higher Categories.* New York: Springer.

Brown, Ronald. 2006. *Topology and Groupoids.* Deganwy, UK: BookSurge LLC.

Cederman, Lars-Erik. 1997. *Emergent Actors in World Politics: How States and Nations Develop and Dissolve.* Princeton, NJ: Princeton University Press.

_____. 2001. Modeling the democratic peace as a Kantian selection process. *Journal of Conflict Resolution* 45 (August):470-502.
_____. 2008. Articulating the geo-cultural logic of nationalist insurgency. Pp. 242-270 in S.N. Kalyvas, I. Shapiro & T. Masoud, eds., *Order Conflict and Violence.* New York: Cambridge University Press.
Collins, Randall. 1978. Some principles of long-term social change: The territorial power of states. *Research in Social Movements, Conflict and Change* 1:1-34.
Crane, Louis & David Yetter. 1993. A categorical construction of 4D topological quantum field theories. Pp. 120-130 in L.H. Kauffman & R.A. Baadhio, eds., *Quantum Topology.* River Edge, NJ: World Scientific.
Ehresmann, Andree C. & Jean-Paul Vanbremeersch. 2007. *Memory Evolutive Systems: Hierarchy, Emergence, Cognition.* Amsterdam: Elsevier.
Eilenberg, Samuel & Saunders Mac Lane. 1945. General theory of natural equivalences. *Transactions of the American Mathematical Society* 58:231-294.
Epstein, Joshua M. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling.* Princeton, NJ: Princeton University Press.
Epstein, Joshua M. & Robert Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up.* Cambridge, MA: MIT Press.
Fiadeiro, José Luiz. 2005. *Categories for Software Engineering.* Berlin: Springer.
Gilbert, Nigel & Klaus G. Troitzsch. 2005. *Simulation for the Social Scientist.* New York: Open University Press.
Goguen, J. 1991. A categorical manifesto. *Mathematical Structures in Computer Science* 1:49-67.
Goldblatt, Robert. 1984. *Topoi: The Categorial Analysis of Logic.* Mineola, NY: Dover.
Hage, Jerald. 1994. Sociological theory: Complex, fragmented and politicized. Pp. 52-65 in J. Hage, ed., *Formal Theory in Sociology: Opportunity or Pitfall.* Albany: State University of New York Press.
Hannan, Michael T. & John Freeman. 1989. *Organizational Ecology.* Cambridge, MA: Harvard University Press.
Iordache, Octavian. 2011. *Modeling Multi-level Systems.* Berlin: Springer.
Kohler, Timothy A. & George J. Gumerman, eds. 2000. *Dynamics in Human and Primate Societies: Agent-based Modeling of Social and Spatial Processes.* New York: Oxford University Press.
Lambek, J. & P.J. Scott. 1988. *Introduction to Higher Order Categorical Logic.* New York: Cambridge University Press.
Landry, Elaine. 1998. Category theory: The language of mathematics. *Philosophy of Science* 66 (September):S14-S27.
Lawvere, F. William. 2005. Taking categories seriously. *Theory and Applications of Categories* 8:1-24.
_____. 2006. Adjointness in foundations. *Theory and Applications of Categories* 16:1-16.
Leinster, Tom. 2004. *Higher Operads, Higher Categories.* New York: Cambridge

University Press.
Lin, Zhiang & Kathleen Carley. 2010. *Designing Stress Resistant Organizations: Computational Theorizing and Crisis Applications.* New York: Springer.
Lloyd, Kenneth A. 2010. A category-theoretic approach to agent-based modeling and simulation. Paper presented to Swarmfest 2010, Santa Fe.
Mac Lane, Saunders. 1971. *Categories for the Working Mathematician.* New York: Springer-Verlag.
Marquis, Jean-Pierre. 2009. *From a Geometrical Point of View: A Study of the History and Philosophy of Category Theory.* New York: Springer.
McCleary, John. 2001. *A User's Guide to Spectral Sequences.* New York: Cambridge University Press.
Miller, John H. & Scott E. Page. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life.* Princeton, NJ: Princeton University Press.
North, Michael J. & Charles M. Macal. 2007. *Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation*. New York: Oxford University Press.
Peterson, A. Townsend, Jorge Soberòn, Richard G. Pearson, Robert P. Anderson, Enrique Martìnez-Meyer, Miguel Nakamura & Miguel BastosAraùjo. 2011. *Ecological Niches and Geographic Distributions.* Princeton, NJ: Princeton University Press.
Pierce, Benjamin C. 1991. *Basic Category Theory for Computer Scientists.* Cambridge, MA: MIT Press.
Sallach, David L. 2000. Classical social processes: Attractor and computational models. *Journal of Mathematical Sociology* 24 (4):245-272.
_____. 2003. Social theory and agent architectures: Prospective issues in rapid-discovery social science. *Social Science Computer Review* 21:179-195.
_____. 2011. Herding concepts: The contextual validation of social agent models. Paper presented at the Computational Social Science Society of the Americas, Santa Fe.
_____. 2012. Socio-cultural structures: A categorical synthesis. Paper presented to the Midwest Sociological Society, Minneapolis.
Sallach, David L., Michael J. North & William A. Rivera. 2012. Scope and scale as an architectonic framework. Paper presented to the Fourth International Conference on Applied Human Factors and Ergonomics, San Francisco.
Schelling, Thomas C. 1978. *Micromotives and Macrobehavior.* New York: W.W. Norton.
Simpson, Carlos. 2012. *Homotopy Theory of Higher Categories: From Segal Categories to n-Categories and Beyond.* New York: Cambridge University Press.