Chapter 31

COMPUTATIONAL LABORATORIES FOR SPATIAL AGENT-BASED MODELS[*]

CATHERINE DIBBLE
Department of Geography, University of Maryland, College Park, MD 20742, USA
e-mail: *cdibble@umd.edu* *url:* *http://myprofile.cos.com/cdibble*

Contents

---

[*] Dibble, Catherine "Computational Laboratories for Spatial Agent-Based Models," Chapter 31, pages 1511-1548 in: Tesfatsion, Leigh and Judd, Kenneth L. *Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics*, Amsterdam: Elsevier, 2006.

Abstract

An agent-based model is a virtual world comprising distributed heterogeneous agents who interact over time. In a *spatial* agent-based model the agents are situated in a spatial environment and are typically assumed to be able to move in various ways across this environment. Some kinds of social or organizational systems may also be modeled as spatial environments, where agents move from one group or department to another and where communications or mobility among groups may be structured according to implicit or explicit channels or transactions costs.

This chapter focuses on the potential usefulness of computational laboratories for spatial agent-based modeling. Speaking broadly, a *computational laboratory* is any computational framework permitting the exploration of the behaviors of complex systems through systematic and replicable simulation experiments. By that definition, most of the research discussed in this handbook would be considered to be work with computational laboratories. A narrower definition of computational laboratory (or comp lab for short) refers specifically to specialized software tools to support the full range of agent-based modeling and complementary tasks. These tasks include model development, model evaluation through controlled experimentation, and both the descriptive and normative analysis of model outcomes.

The objective of this chapter is to explore how comp lab tools and activities facilitate the systematic exploration of spatial agent-based models embodying complex social processes critical for social welfare. Examples include the spatial and temporal coordination of human activities, the diffusion of new ideas or of infectious diseases, and the emergence and ecological dynamics of innovative ideas or of deadly new diseases.

Keywords

agent-based simulation, computational laboratory, computational social science, computational economics, spatial economics, spatial social science, spatial networks, small-world networks, scale-free networks, synthetic landscape, inference

JEL classification: C63, C73, C88, C99, D43, I10, O33, Z13

1 Introduction

1.1 Overview of Chapter

> Research has been likened to warfare against the unknown. … The attacker will have a great advantage if he can bring to bear a new technical weapon.
>
> (Beveridge 1957, page 176)

An agent-based model is a virtual world comprising distributed heterogeneous agents who interact over time. In a *spatial* agent-based model the agents are situated in a spatial environment and are typically assumed to be able to move in various ways across this environment. Some kinds of social or organizational systems may also be modeled as spatial environments, where agents move from one group or department to another and where communications or mobility among groups may be structured according to implicit or explicit channels or transactions costs.

This chapter focuses on the potential usefulness of computational laboratories for spatial agent-based modeling. Speaking broadly, a *computational laboratory* is any computational framework permitting the exploration of the behaviors of complex systems through systematic and replicable simulation experiments. By that definition, most of the research discussed in this handbook would be considered to be work with computational laboratories. A narrower definition of computational laboratory (or comp lab for short) refers specifically to specialized software tools to support the full range of agent-based modeling and complementary tasks. These tasks include model development, model evaluation through controlled experimentation, and both the descriptive and normative analysis of model outcomes.

The objective of this chapter is to explore how comp lab tools and activities facilitate the systematic exploration of spatial agent-based models embodying complex social processes critical for social welfare. Examples include the spatial and temporal coordination of human activities, the diffusion of new ideas or of infectious diseases, and the emergence and ecological dynamics of innovative ideas or of deadly new diseases.

Consider a thought experiment to help motivate the usefulness of comp labs as complements to spatial agent-based models: Imagine that top decision-makers have asked you to apply an agent-based research model to avert a global pandemic, where the livelihoods and perhaps the lives of millions of people may depend upon the timeliness and quality of your results (Osterholm 2005, Aldous and Tomlin 2005). They need your answers within six months. Preliminary results even before then could provide crucial leverage for averting disaster, yet misleading results may do more harm than good.

Which comp lab tools would you wish you had available to assist with development, testing, and refinement of your model? Which simulations would you run first to explore the problem? How would you calibrate, test, apply, evaluate, and perhaps generalize your model and your results within six months? How would you adapt your model or your inference as the crisis begins to unfold, or as preliminary feedback from the success or failure of your advice begins to arrive? Which analytical tools would you most wish you had, given reasonable yet finite computational power and limited time for analysis of simulation results as they accumulate?

Comp labs provide the tools researchers need to perform such tasks. At the simplest level, a good agent-based model is capable of generating the phenomenon we seek to study. Yet generating a phenomenon is far from sufficient for effective agent-based research. What matters most is what we can *learn* from our models, and how much we can trust their results.

This chapter provides an overview of the comp lab capabilities most likely to be useful for spatial agent-based models, and explores the various ways they could be used effectively. This exploration is explicitly normative and does not presume to represent current practices in Agent-based Computational Economics (ACE).

Although this chapter specifically addresses comp lab tools for working with spatial agent-based models, many of the comp lab principles and tools discussed apply to aspatial models as well. Similarly, spatial landscapes may be interpreted quite broadly as anything that structures local context and interactions among a model's agents.

This chapter orients newcomers to comp labs by discussing and illustrating basic components and capabilities of comp labs for research with spatial agent-based models. The remainder of Section 1 highlights research challenges posed by richly structured distributed dynamic systems for standard economic modeling, and briefly summarizes how comp labs might aid researchers in addressing these challenges. Sections 2 through 4 examine three main categories of tools, such as comp lab support for controlling and testing models, for modeling agents, and for creating empirical or synthetic social and spatial landscapes. Subsequent sections 5 through 7 address finer points and more sophisticated methods for inference and for effective analysis of robustness and risk. Concluding remarks are provided in section 8.

1.2 Challenges Posed by Spatial Systems

> [N]ew tools … have removed crucial technical barriers and transformed a once inhospitable field into fertile ground for theorists.
>
> (Fujita, Krugman, and Venables 1999, page 2)

Local (micro) interactions among distributed dynamic agents generate global (macro) structures; diverse examples include market prices, market failures, organizational behavior, social norms, and regional settlement patterns. These dynamic spatial processes defy top-down modeling or deductive analytical inference due to the complex exogenous and endogenous boundary conditions arising from their micro-level interactions. Local interactions may be either spatial or aspatial; in general, the term refers to interactions among distributed subsets of agents.

Realistic geographic landscapes may generate conditions that violate one or more assumptions underlying the First and Second Welfare Theorems. The First Welfare

Theorem (efficiency) roughly states that if markets exist for all valued goods and services, if no firm or consumer can influence prices, if prices adjust perfectly so that all markets clear, and if all (price-taking) firms and consumers correctly anticipate these prices, then the market outcomes will be Pareto Optimal. The Second Welfare Theorem (equity) roughly states that if the First Welfare Theorem applies and if there are no externalities of consumption or of production, then any Pareto Optimal outcome may be reached, given appropriate transfers of wealth (Mas-Colell et al. 1995, page 308).

Yet uneven spatial distributions of goods and people across a geographic landscape may lead to conditions that violate these assumptions. For example:

- There may be too few local buyers and sellers to create local markets for goods, especially when there are high transportation costs.
- Small numbers of buyers or sellers may lead to thin markets where one side or both no longer acts as a price taker.
- Even with modern transportation and telecommunication systems, spatial distance continues to impose significant transactions costs such as imperfect information and severe coordination problems that affect market transactions.
- Geographic landscapes generate local environments where externalities of consumption or production are often the norm rather than the exception.

As Tesfatsion (2006a) demonstrates with respect to the implicit Walrasian Auctioneer assumed in competitive market models, tractable theoretical models often naively assume agent coordination or sophistication that may in fact not be feasible given the agents' contexts, information, or incentives.

The complexity that arises from interacting agents becomes even more interesting once we consider *strategic* interactions. Schelling's *Micromotives and Macrobehavior* (1978) summarizes the essential challenges for modeling strategic interactions among distributed agents:

> What we typically have is a mode of *contingent behavior*—behavior that depends on what others are doing. (page 17)
>
> [P]eople locate themselves voluntarily in some pattern that does not possess evident advantages even for the people who by their own choices form the pattern. (page 12)
>
> How well each does for himself in adapting to his social environment is not the same thing as how satisfactory a social environment they collectively create for themselves. (page 19)

Economic theory and game theory begin to provide formal theoretical frameworks for emergent macro effects of non-strategic and strategic micro-level interactions among agents. Yet it is nearly impossible to extend the fundamental theoretical results to realistically distributed systems of heterogeneous, dynamic, adaptive, and mobile agents when researchers are limited to thought experiments or top-down, equation-based computational models. Well equipped comp labs for spatial agent-based models can greatly extend our ability to explore beyond the bounds of purely analytical inference to establish new theoretical and applied results for important and interesting richly structured systems.

1.3 Addressing Spatial System Challenges with Comp Labs

A prototypical *spatial agent-based model* consists of a full specification for the following aspects:

- One or more classes of agents, and the types of interactions they may have with one another and possibly also with their environment.
- The nature of the spatial, social, or organizational environments within which these agents may or may not move around and which may structure their encounters.
- A specification of initial conditions for the simulation, generally including the initial locations of agents within their environment.
- A schedule of activities for the simulation, including a means for determining when each simulation should end.
- Means for observing and recording key data about the simulation's behavior.

In principle, the conceptual specification of a spatial model exists independently of any given *implementation* of the model in a particular computer language, simulation platform, or comp lab. Ideally, important models will be implemented in more than one computer language, model platform, or comp lab. In practice, implementation details such as the order in which agents take turns, nuances of their interaction structures, or the specific random number generators used often affect the model's behavior. So in practice the term "agent-based model"—whether spatial or aspatial—almost always refers to a specific implementation of the conceptual model. Essentially, for now, an agent-based model's implementation *is* its complete specification.

Tesfatsion (2006a) provides an excellent introduction to agents and an "ACE Trading World" that exemplifies a typical fully specified aspatial agent-based model. She includes pseudo-code outlines for the public and private data for each agent and its behavioral methods, and for the initial conditions, agent activities, and stopping rule for each simulation. The ACE Trading World has three types of cognitive agents: bean producers, hash producers, and consumer-shareholders who purchase beans and hash to consume at each simulation step (Tesfatsion 2006a, Tables 1-4). It is an interesting exploration of the operation of a simple market once agents are required to engage in explicit procurement rather than trading indirectly via a mythical Walrasian auctioneer.

Yet the ACE Trading World is strictly aspatial; each agent has perfect information about the prices posted by all other agents, and there are no transaction costs or spatial locations of the agents to structure their information or their interactions with one another. This is wise, as its behavior is already complex. Even models that seek to understand the effects of spatial or other interaction structures should be able to run aspatial control simulations where the same set of agents interacts in a null space, such as a perfectly mixed soup, in order to distinguish the effects of agent or other model specifications from the effects of their interaction structure. Thus, when simulations with the same populations of agents with the same random number seeds are run on richly structured spatial or organizational landscapes, the effects due to landscape structure can be thoroughly isolated from the effects due to other aspects of the model or of their interactions.

Richly Structured Spatial or Organizational Network Landscapes

Extending the ACE Trading World (Tesfatsion 2006a) even to simple spatial landscapes raises interesting questions. For example, imagine a network landscape of local villages, where each village is aspatial, per the original model, and links such as roads connect villages to one another in various patterns. This is likely to introduce several kinds of information costs (effort, noise) and transaction costs (shipping, time delay, tariffs). For now, consider simply Samuelson's (1952) simple iceberg model of shipping costs, where x% of the goods melt per mile.

If shipping costs are prohibitive even for the nearest neighbors, then we simply have islands, each of which operates as a separate ACE Trading World. Even so, it may nevertheless be interesting to explore the effects of population sizes and relative proportions of its three types of agents, per Tesfatsion's (2000, 2001) exploration of the effects of market power in labor markets. Alternatively, if shipping costs are zero and we have assumed perfect information and no transaction costs despite distance, then the distribution of agents across villages makes no difference and it functions as one global trading world.

In all other situations, imperfectly linked local markets for identical goods are distinguished only by their locations relative to each potential consumer. The landscape's network characteristics and its distribution of agents may have tremendously interesting effects on the adjustment dynamics and potential equilibria of local markets. We have not even begun to discuss related extensions such as information costs, local resource constraints, heterogeneous distribution of resources or production conditions, local externalities of consumption or production, or migration of agents from one node to another. ACE researchers have barely begun to explore such extensions, although see Wilhite (2001, 2006) and Dibble (2001b) for early work along these lines.

Networks are especially relevant for economics because almost all economic exchanges are mediated by transportation or communication networks of some kind. The structure of such networks is generally fixed within the time frame considered for most ACE models, although it can be even more interesting to consider the long-run co-evolution of economic processes, population distributions, and network infrastructure.

Such questions have been difficult to address in part because we have not had good tools to generate synthetic network landscapes and population distributions in order to explore their effects. While extensions to network landscapes will not be deeply explored within this chapter, the objective here is to motivate the importance of landscapes and other interaction structures for ACE research, as a key example of potentially useful capabilities provided by a well-equipped comp lab to support spatial modeling.

2. A Well-Equipped Comp Lab for Spatial Agent-Based Modeling

> In contrast, the real world is a single time-series realization arising from a poorly understood data generating process. Even if an ACE model were to accurately embody this real-world data generating process, it might be impossible to verify this accuracy using standard statistical procedures. For example, an empirically observed outcome might be a low-probability event lying in a relatively small peak of the outcome distribution for this true data-generating process, or in a thin tail of this distribution. (Tesfatsion 2006a, page 845)

In the real world, we almost never have an opportunity to rewind the historical tape to replay and explore the different outcomes that may result from chance events (Fontana and Buss 1994). In econometrics, we understand that each empirical observation contains some proportion of variation due to explanatory variables, inextricably entwined with an unknown but ideally well-behaved proportion of noise assigned to its error term and ascribed to chance or imperfect observation. Each implementation of an agent-based model is, by itself, a means to simulate one time-series realization at a time. With an agent-based comp lab, we have perfect control over both treatments and stochastic sources of variation. Thus we have the capability to simulate *exact* replicates of each treatment, to fully explore the effects of stochastic variation.

In my research on spatial systems, I have found the following three types of comp lab tools to be especially useful:

1. tools to facilitate implementation, testing, calibration, and basic operation of a spatial agent-based simulation model, including its landscapes, agent populations, initial conditions and spatial distributions, and model-specific rules and schedules;
2. basic tools to generate, control, and observe multiple realizations of the model, including separable realizations of landscape characteristics, agent characteristics, initial conditions and distributions, and stochastic events during the simulation's execution; and
3. advanced tools to provide especially effective control, search, optimization, and evaluation within an especially large or complex space of potential scenarios and associated stochastic realizations for each scenario of characteristics and initial conditions.

Multi-purpose spatial agent-based modeling platforms and comp labs are generally implemented in object-oriented programming languages. In an object-oriented language, each type of agent is defined by a *class*, which acts as a blueprint to define the basic private and public data fields and behavioral methods that each agent created with this class will possess. This class approach to the construction of agents offers two important advantages.

First, each agent instantiated from a class is *encapsulated* as its own separate entity, with private data and methods that may be accessed only by asking the agent for its answer or by telling it to do something that it knows how to do. Encapsulation seems intuitively obvious to social scientists. Moreover, an important advantage of agent encapsulation from a modeling perspective is that it supports modular programming. That is, agents only need to know what they can ask of each other. Any agent may alter its internal characteristics or methods without disrupting the public characteristics of its methods.

Second, classes can *inherit* characteristics from other classes, which provides tremendous advantages for developing agent-based simulation models and complementary laboratory tools. For example, we may define a *base class* of ruminant livestock, which we can use as the parent class from which we create *child classes* for sheep, goats, cows, camels, and llamas. Each child class inherits all capabilities of its base class, and may then redefine inherited methods or add new data and methods to the base class. When used wisely, encapsulation and inheritance enhance conceptual clarity and economical software design, development, and modifications.

A well-equipped comp lab provides a rich set of classes for cognitive and other agents, each provided with diverse capabilities to support use directly in models or as base classes for model-specific extensions. The laboratory should also provide at least one "landscape" class to provide structure for agent interactions, whether this is an aspatial institution, an organizational structure that may or may not include some metric for distances among teams, or a fully spatial landscape.

Even laboratories with sophisticated classes to generate spatial landscapes should always provide an aspatial randomly-mixed null landscape. This serves as a control to test for artifacts of the model unrelated to the structure of interactions among agents.

"Docking" refers to the practice of comparing results from matching simulation runs for different implementations of a given conceptual model. Docking can be an especially important approach to identify, isolate, and control for subtle differences that may be introduced even by apparently congruent implementations. See Axtell et al. (1995) or Axelrod (1997) for an interesting discussion of their experience docking their respective simulation models.

In this spirit, a comp lab's aspatial landscape can also be useful for docking spatial models with aspatial mathematical models of corresponding processes. For example, a new agent-based model of an infectious disease epidemic may be docked and calibrated against classical aspatial mathematical models or highly simplified spatial mathematical models for a particular population of agents. The spatial agent-based model can then be used with greater confidence to explore the unique effects of more realistic spatial structures on epidemic dynamics for the same disease as it unfolds within the same population of agents.

Finally, spatial comp labs that are well equipped to support both applied and theoretical work may also include tools to calibrate synthetic landscapes according to empirical characteristics, or to directly import various types of empirical landscapes from Geographic Information Systems (GIS) or satellite remote sensing observations (Dibble and Feldman 2004).

In order to simulate synthetic initial conditions, synthetic landscapes, stochastic events, and stochastic choices by cognitive agents, well equipped comp labs should provide classes to start and seed multiple, fully separable, ultra long-period (i.e. Mersenne Twister) random number series for a wide variety of distributions (e.g. uniform, Gaussian, fair or weighted coin flip, or roulette wheels that assign probability distributions derived from histograms, which are important for genetic algorithms). The long period before repeating random number patterns is essential for serious inference with any stochastic spatial agent-based model, in order to prevent spurious artifacts where observed cycles or patterns are driven merely by repetition in one or more of the model's random number series.

Time in simulation models may be absolute, relative, or both. These correspond, respectively, to agent actions that are triggered by a simulation clock (e.g. via simulation 'steps'), by other endogenous events or agent actions within the simulation, or by relative time elapsed since some endogenous triggering event. Simple examples of each would include aging, contracting an infectious disease, and becoming ill with the disease after some incubation period beyond infection.

Scheduling tools are important, not only for cognitive agents within each simulation, but also for other simulation agents (trees may grow, toxins may diffuse, water may accumulate), and especially for meta-activities such as data collection and other forms of simulation monitoring such as those related to stopping the simulation. Stopping rules may also be absolute, relative, or both. For example, a simulation may be set to stop after a certain number of steps have elapsed, after a persistent condition has been reached, or following a set delay after a specific event.

Ideally, each comp lab should have at least one fully-constructed proto-model, no matter how trivial, of a fully-implemented spatial agent-based simulation model on a simple landscape. Such proto-models serve several important purposes. For example, they provide simple working examples to show that the comp lab itself was correctly installed. They also provide a simple model from which researchers may learn, and a simple foundation that researchers may modify to create their own models.

Finally, at least one and ideally two important classes control simulations. First, a "main" class is generally responsible for setting up and running each simulation. It reads in the parameters and treatments for that simulation, sets up the landscape, agents, and initial conditions of the simulation, starts the simulation, and supervises data collection, visualization (if any), and stopping conditions. One or more secondary classes can supervise multiple simulations by allowing researchers to specify parameters and treatments for large batches of simulations.

2.1 Generating, Controlling, and Observing Many Simulations for each Model

The secondary classes mentioned above provide batch-control tools to specify fixed parameters, to specify lists of values for each of the parameters chosen for parameter sweeping for sensitivity analysis, or to specify levels of treatment variables in order to explore model behavior. For stochastic models, it is especially important to sweep across one or more sets of random number seeds for each of the separable random number series to be used by the model, in order to generate statistically significant realizations for each scenario. It is important to explore the variability of stochastic outcomes for each scenario's specific combination of model parameters and treatment variables.

Some of these parameters chosen for sweeping, including random number seeds, may be used to generate specific realizations of synthetic landscapes or other interaction structures, as discussed in greater detail in Section 4 of this chapter. Other parameters or series of random number seeds may in turn be used to generate synthetic agent populations and to control their initial conditions and stochastic decisions or events.

For example, one random number seed may control the initial distributions of agent characteristics and endowments. Another may be responsible for allocating agents to initial locations within a spatial landscape or related interaction structure such as an organization. Others may control various stochastic aspects of agent decisions or of landscape events as each simulation unfolds.

2.2 Advanced Tools for Effective Search, Control, Optimization, and Testing

An especially helpful tool for any comp lab is a supervisory genetic algorithm, simulated annealing, or other automated search and optimization heuristic to help discover combinations of scenarios and random number seeds that lead to particularly interesting outcomes. For example, see Section 7 of this chapter for a discussion of the ways in which a comp lab supervisory genetic algorithm can be used to support inference, optimization, and analysis of risks.

In turn, complementary analytical tools for network analysis, and analytical measures such as spatial statistics, can help to support generalization from specific spatial landscapes or populations to broader classes. Such analytical measures may characterize combinations of initial conditions, distributions, or outcomes, to relate common characteristics of each across multiple simulation outcomes.

Finally, spatial comp labs may also be used in conjunction with familiar econometric and statistical tools through systematic analysis of output when highly stochastic models generate especially large numbers of simulation outcomes. This can be especially helpful to support analysis of risk or resilience that may be associated with particular interventions such as for controlling epidemics or similarly normative modeling objectives.

These are classifications for comp lab tools that have been especially useful in my research with spatial agent-based simulations on richly structured organizational and spatial landscapes. As with agent-based simulation models, even general-purpose comp labs differ widely in their strengths and capabilities. However, this superset of capabilities that have proven useful for spatial agent-based modeling may provide helpful existence proofs and inspiration for evaluating the strengths and limitations of particular comp labs.

2.3 Principles of Comp Lab Engineering

Eventually, spatial agent-based models and comp labs may become so powerful and well equipped that serious researchers devote entire careers simply to studying the behavior of a well established and fully implemented model, or of model variants fully supported by the comp lab's existing classes. (This is one way to introduce students to comp lab research.) For now, even the simplest spatial agent-based modeling project generally involves the design and programming implementation of at least one new model-specific child class for agents, landscapes, or both. Thus, I highly recommend that comp lab researchers carefully study the O'Reilly *Extreme Programming Pocket Guide* (chromatic 2003). This graceful little book (81 pages) is as much about the clear communication and coordination of model specifications as about programming, so it is at least as important for researchers who hire programmers as it is for those who program models themselves.

Ideally, a comp lab provides a rich hierarchy of *well-factored* base classes, where the accumulation of a rich library of classes for agents, spatial landscapes or other interaction structures and related comp lab tools provides researchers with powerful research leverage. Factoring classes is similar to factoring numbers into primes; it refers to the clean division of class functionality into unique and irreducible units. To consider an example closer to the hearts of most researchers, factoring classes in an agent-based model is similar to factoring sections of a research paper to reduce unnecessary repetition.

3 Comp Lab Support for Heterogeneous, Mobile, Cognitive Agents

3.1 Hierarchies of Agent Classes

> Here "agent" refers broadly to bundled data and behavioral methods representing an entity constituting part of a computationally constructed world. Examples of possible agents include individuals (e.g., consumers, workers), social groupings (e.g., families, firms, government agencies), institutions (e.g., markets, regulatory systems), biological entities (e.g., crops, livestock, forests), and physical entities (e.g., infrastructure, weather, and geographical regions). Thus, agents can range from active data-gathering decision-makers with sophisticated learning capabilities to passive world features with no cognitive functioning. Moreover, agents can be composed of other agents, thus permitting hierarchical constructions. For example, a firm might be composed of workers and managers.
>
> (Tesfatsion 2006a, pages 835-836)

One of the most important distinctions between a spatial agent-based simulation model and a comp lab is the laboratory's provision of a powerful set of base classes for agents, for other model components such as landscapes (see Section 4 below), and for comp lab tools to support visualization, data collection, analysis, search, optimization, and control for simulation models.

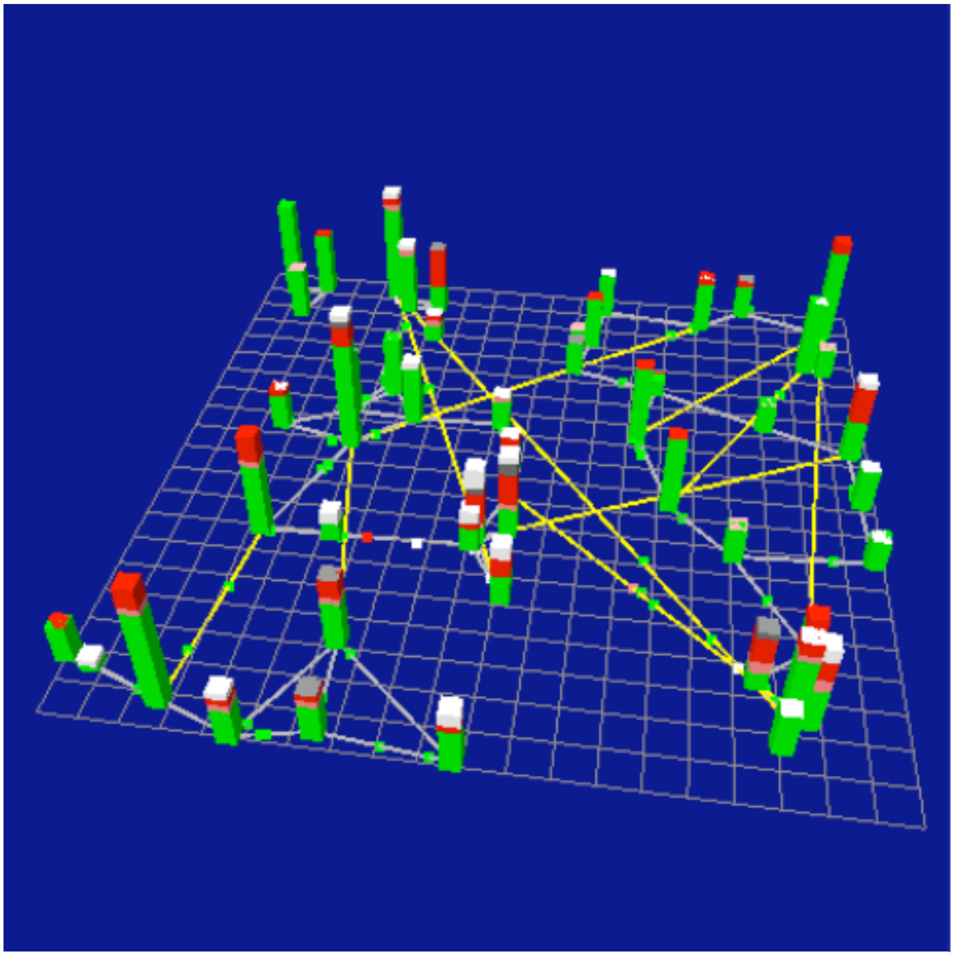

Figure 1: Synthetic or empirically derived organizational or geographic network landscapes, where individual agents travel between team or city nodes and bar charts summarize the current status of each node's population.

For example, the GeoGraph Comp Lab (Dibble and Feldman 2004) in use by my research group at the University of Maryland provides two base classes for agents, from which model agents may inherit specific geographic capabilities that correspond to classes of GeoGraph landscapes.

- **GeoAgent** is a network-enabled agent. It may either teleport from node to node or may be restricted to follow specific types of links within the landscape. It can evaluate nodes within a network-specific neighborhood, and can compute shortest paths from one node to another along multiple links. It is written to utilize the GeoGraph Node3D class, which is a network enabled node class used with GeoGraph network landscapes for modeling organizational structures or geographical landscapes such as networks of cities. Most of the simulations extend this class, including the epidemic model illustrated in Figure 1.
- **FreeAgent** is a free-roaming agent. This class provides the basis for 'flocking agents' such as villagers or wildlife on synthetic fractal terrain or empirical digital elevation models of natural landscapes, illustrated in Figure 2.

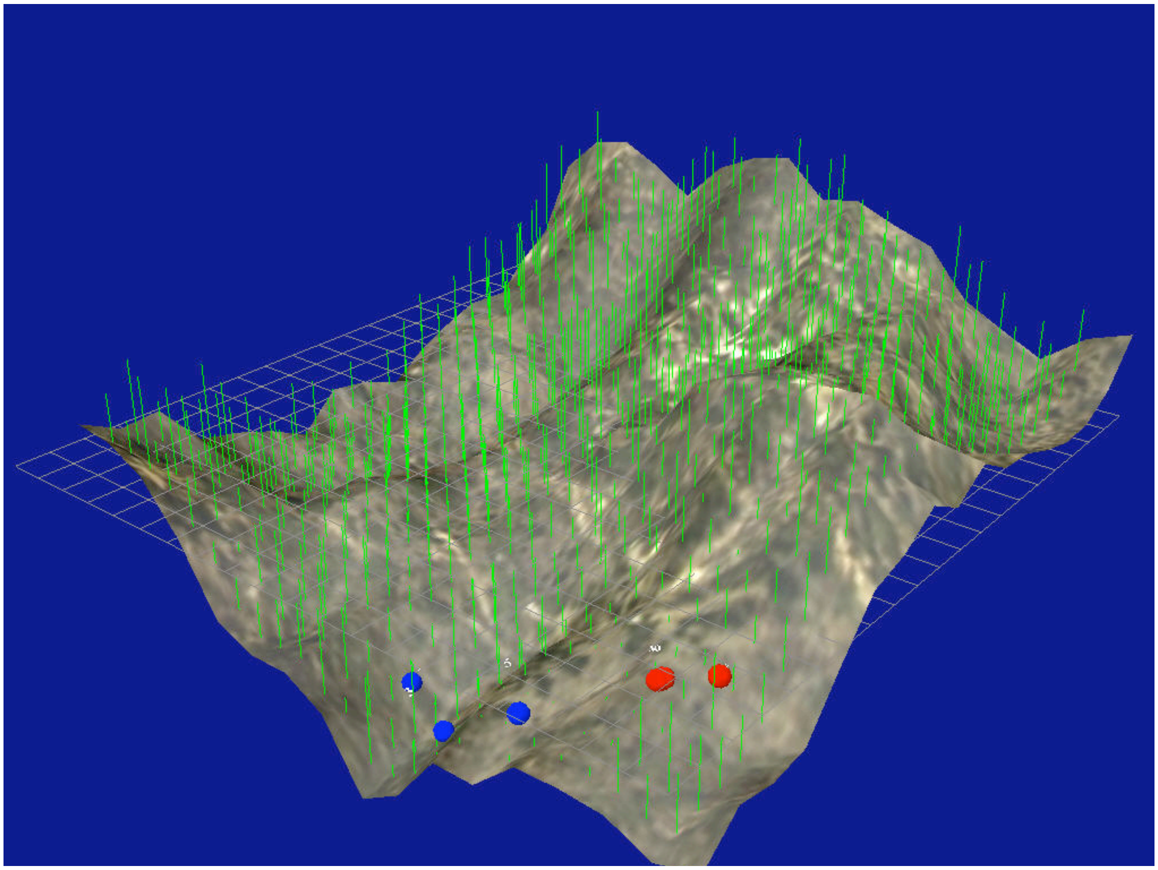

Figure 2: A synthetic three-dimensional fractal terrain landscape with parameterized renewable green "tree" agents and small flocks of "deforestation" round agents, shown toward the front. This could instead use empirically derived geographic terrain and vegetation, imported from a Geographic Information System or from Remote Sensing data.

As a second layer in our hierarchy of agent classes, we have developed two child classes, each of which serves as an agent class or base class for one of our primary research lines:

- **EconAgent** is a **GeoAgent** that knows how to collect information about its world in order to select from among its list of alternatives the one that provides sufficient (for satisficing) or optimal satisfaction of its objective function.
- **EpiAgent** is a **GeoAgent** that knows how to become infected with or immune to a disease, and how to progress through various stages of incubation, sickness, and death or recovery if it becomes infected. For many diseases, an EpiAgent becomes infectious for some duration that may overlap other stages, during which the agent "knows" how to transmit the disease to the landscape (e.g. via doorknobs or keyboards) or directly to other agents (e.g. via "sneezing").

For example, our EconAgent class could be used to implement Schelling's segregation game (Schelling 1978) directly, without further modification. We simply create instances for a simulation population of n+m agents, tell each of n agents that it is type "Blue" and each of m agents that it is type “Green,” and tell each what minimum percentage of neighbors of its own type it considers to be satisfactory.

In general, we develop a new child class only when we need to add new data fields or new capabilities (methods) to an agent. We simply use the class directly to create individual “instances” of agents for each simulation if all we need to do is to provide each individual model agent created from the class with values for its variables. In our case, the objective function for an EconAgent may be provided to the agent as a parameter rather than “hard-coding” each equation into a new class that differs only according to that equation. We create separate child classes only when their objective functions need to be defined according to radically different sets of variables.

3.2 Cognitive Agent Learning, Adaptation, and Evolution

Cognition refers to the methods agents use to make decisions about their behavior. Learning refers to their ability to modify their cognitive methods over time. Adaptation is generally distinguished from learning by being passive and biological rather than active and cognitive, although these two terms are often confused and used interchangeably. Alternatively, adaptation may refer to an agent selecting an alternative strategy that was already known to it, without requiring any cognitive effort to develop new methods. Evolution refers to the improvement of subsequent generations of agents as a result of natural selection (e.g. via survival and reproduction in proportion to relative fitness of the agents).

In general, learning may be modeled simply as imitation of the behavioral rules used by more successful agents. Alternatively, learning models such as genetic-based machine learning (Holland 1992, Goldberg 1989, Dibble 2001a) may apply “fitness selection,” “cross-over” and “mutation” operations to sets of competing behavioral rules in an attempt to evolve better-performing rules. Evolutionary models may not involve learning within individual agents, but may instead simply select for the agents who employ the most successful strategies. For example, evolution may select for firms that are able to compete most profitably in a given environment. Yet such models in principle provide an excellent demonstration of the importance of learning as opposed to evolution, as individuals or firms that cannot learn or adapt to changing circumstances are likely to die or go out of business when conditions change.

For more detailed discussions of models of learning, adaptation, and evolution, see the handbook chapters by Brenner (2006), Duffy (2006), and Young (2006).

4 Comp Lab Spatial and Organizational Landscapes

Sound generalizations based on scientific experiments require controlled conditions and sufficient experimental trials in order to distinguish fully their incidental effects from their systematic effects. A well-equipped agent-based comp lab provides such controls over the characteristics of agent populations for each model. Yet agents are only half of the story for realistically structured systems we wish to study.

Each agent's opportunities and constraints for interaction are determined not only by the characteristics of its own position in an organizational structure, or by its geographic location, but also by its structural *situation*; its access to other positions or locations, each with its current complement of agents and other characteristics. Historically, spatial and even social situations have been determined primarily by geographic distance, with the relative ease of access modified by natural features such as rivers, mountains, or coastlines. More recently, opportunities for social and economic interaction are driven by networks of transportation or communication *spatial technologies* (Couclelis 1994) that shrivel time and cost surfaces unevenly at all scales (Tobler 1999).

Many of the fundamental open questions at the frontiers of theoretical and applied economics and related social science research are driven by the analytical intractability of studying dynamic interactions among distributed, heterogeneous, mobile agents embedded or mobile within richly structured spatial and social networks. Formal theoretical analysis of the behavior of aggregate systems of such agents becomes intractable almost immediately. Agent-based comp labs can help theorists to explore the behavior of these spatially distributed socio-economic systems. Spatial structure is central to the dynamics of spatial processes such as the diffusion of innovations or of infectious diseases within a population of mobile individuals.

Section 4.1 introduces comp lab generation of synthetic landscapes. Section 4.2 briefly summarizes the original work on aspatial small-worlds that inspired the extensions by Wilhite (2001, 2006), Dibble (2001b), and Dibble and Feldman (2004) to spatial small-world synthetic landscapes. Section 4.3 introduces the contraction factor, distance decay, and positive feedback extensions that provide the conceptual foundation for generating geographically interesting spatial small-worlds for use in comp lab experiments. Finally, section 4.4 summarizes typical comp lab options for generating richly-structured synthetic landscapes for controlled experiments.

4.1 Introduction to Comp Lab Generation of Synthetic Landscapes

Ideally, we should be able to study the effects of exogenous landscape structure by exploring the ways in which selected local and global network characteristics affect the micro and macro evolution of systems of agents. Similarly, we should be able to study the effects of endogenous network structure by studying the co-evolution of agents and networks, especially the effects of positive feedback and of both micro and macro path dependence (Tesfatsion 1997). Systematic exploration of either is problematic when we are limited to observations and analysis of real-world geographies. For example, real-world network landscapes that have dissimilar characteristics may be inhabited by agents whose economic and social conditions are too different for meaningful comparison.

Similarly, a real-world landscape limits the generality of our results by offering only a single observation of an interaction structure and associated population distributions.

In contrast, parameterized families of synthetic landscapes can be coupled with parameterized families of synthetic population distributions to offer essential control for comp lab experiments, leading to far deeper and more generalizable understanding of the relationships between network structures and distributed dynamic processes.

In order for comp labs to reach their full potential as tools for theorists modeling distributed systems, we need to be able to generate parameterized families of richly structured synthetic landscapes that vary in the characteristics that we seek to study, yet that remain congruent in their other characteristics. For example, DeCanio *et al.* (2000, 2001) evaluate the effects of richly structured synthetic organizational networks on the efficiency of distributed processes among a collection of workers. Wilhite (2001) considers a trading economy modeled as a ring landscape divided into contiguous regions, then adds one or two random aspatial shortcuts to each landscape to explore the effects of network structure on local commodity markets.

Section 4 extends such approaches by addressing comp lab tools for controlled synthesis of landscapes and other interaction structures, ranging from simple fully-mixed aspatial random soups to richly structured network landscapes that structure opportunities for interactions among agents.

Our GeoGraph Comp Lab has been explicitly designed for use in controlling experimental conditions for spatial agent-based modeling through its ability to generate richly structured parameterized families of synthetic landscapes. These landscapes are useful for building and testing formal models grounded in interesting spatial structures, homogeneous or heterogeneous distributed mobile agents, and context-specific behaviors. To the best of our knowledge, this is the first and remains the only general purpose research comp lab for building bottom-up models that have large numbers of heterogeneous, spatially distributed, mobile individuals on richly structured synthetic network and terrain landscapes.

4.2 Aspatial Small-world and Scale-free Networks

> The goal before us is to understand complexity. To achieve that, we must move beyond structure and topology and start focusing on the dynamics that take place along the links. *Networks* are only the skeleton of complexity, the highways for the various processes that make our world hum. (Barabási 2002)

New formalizations of network structures have begun to revolutionize the study of everything from human social networks (Watts 1999) to the Internet (Albert, Jeong, and Barabási 1999), the error and attack tolerance of networks (Albert, Jeong, and Barabási 2000), and metabolic networks within a cell (Jeong et al. 2000). *Small-world networks* are networks characterized by a high degree of local structure, which nevertheless have surprisingly short average path lengths (e.g. “six degrees of separation”) due to the importance of random shortcuts (Watts and Strogatz 1998). *Scale-free networks* have a distribution of links per node that is exponential rather than normal or uniform, implying that that almost all nodes have very few connections while a few “hub” nodes are extremely well connected (Barabási and Albert 1999). See Strogatz (2001) for an excellent review.

Yet Barabási was right about the limitations of these network formulations. Each addresses merely the structural analysis of a static network, with limited capability to model dynamic processes among heterogeneous mobile individuals. Similarly, each network formulation has been purely aspatial, in the sense that it has structure but not yet spatial relationships or corresponding weights for the links; rather, each link has a uniform "distance" of one, and nodes have no natural location. Finally, small-worlds and scale-free networks represent only two dramatically distinct families of networks; until our GeoGraph Comp Lab (Dibble and Feldman 2004), there had not yet been a generalization to the synthesis and use of realistically hybrid spatial networks.

The original *Nature* paper on small-world networks (Watts and Strogatz 1998) reported the synthesis and analysis of a particular parameterized class of irregular networks. Each ring of n (= 1,000) nodes was initially configured with links to each node's k (= 4) nearest neighbors on either side. A small-world network was created from each k-connected ring by randomly rewiring each link in the network with a very small probability p. All networks in the paper were aspatial, where each link has its distance normalized to one, and the length of any given path is defined by simply counting the number of links it contains.

Two characteristics were measured for each small-world network. First, the Characteristic Path Length ($L(p)$) measured the average shortest-path distance between each pair of nodes in a network. Second, the Cluster Coefficient ($C(p)$) measured the number of links for each node that are still attached only to one of its k nearest neighbors (Watts and Strogatz 1998).

When compared across small-worlds, $L(p)$ falls precipitously and then levels off as p increases from 0 to 1. $C(p)$ falls extremely gradually for p close to 0, and only begins to fall precipitously as p converges to 1. Thus, the $L(p)$ and $C(p)$ curves leave a large lens-shaped gap for small values of p, for their example of 1,000-node rings. This gap corresponds to the network's small-world characteristics; where the $L(p)$ is low and it is relatively easy for signals to traverse the network, yet $C(p)$ is high as almost all links in the world remain local rather than shortcuts (Watts and Strogatz 1998).

Watts (1999) defines the *range* of a link to be the second-shortest path available between the two nodes. For example, the range of a shortcut connecting nodes separated by three base links would be three. Intuitively, small-world characteristics arise when the rewired links provide dramatically advantageous shortcuts by spanning especially large ranges of the base network.

Despite Watts's claims to the contrary (Watts 1999), the configuration of the base network does have an important effect on a network's small-world characteristics for any given n and p. To see this, consider the maximum range for alternative configurations of a network of n nodes, where each node is initially connected only to its nearest neighbors.

For a ring landscape, the maximum range is $n/2$. For a square grid landscape, the maximum range is $2n^{½}$. For $n = 100$, this corresponds to 50 versus 20. For $n = 10{,}000$, this corresponds to 5,000 versus 200. The larger the network, the more dramatic the small-world characteristics are on the ring landscape. This is true for aspatial networks, where each shortcut link has unit distance no matter how large its range.

To extend this analysis to spatial networks, consider the small-world characteristics of a grid lattice where each link has Euclidean distance. Without loss of generality, we can normalize to unit distance for each (orthogonal) base link. In such networks, maximum range remains as defined above, yet the shortcut links that correspond to each maximum range now have distances of n/π and $2^{½}n^{½}$, respectively. Small-world characteristics still exist for each, yet they are considerably less dramatic. For example, in unpublished small-world simulations conducted on Berkeley's Cray T3E super computer with n=1,000 for a spatial (Euclidean) grid, we found that $C(p)$ falls only linearly with respect to p.

Yet, although small-world effects are interesting and important, the truly profound innovation in Watts and Strogatz (1998) is the synthesis and study of parameterized families of irregular networks. In addition, small-world effects do seem to exist in everyday geographic landscapes, so it makes sense instead to turn the question around: what drives the small-world characteristics of a network, and how best can this be modeled in organizational networks or in spatial networks for geographic landscapes?

4.3 Spatial Small-worlds, Contraction Factors, and Modeling Globalization

In synthetic landscapes such as these GeoGraph spatial small-world networks, we can control structure as we improve shortcut technologies, or we can control technologies as we change structure. So we can explore separately the effects of changes in technology from changes in structure, in order to study the specific effects of each on geographic systems and processes. This is an important scientific advantage. Unlike real-world landscapes, GeoGraph's ability to synthesize stochastic families of spatial small-world networks allows us to control geographic structure to compare effects of different spatial technology regimes across landscapes that have equivalent numbers and arrangements of spatial technology shortcuts.

The driving force behind a network's small-world properties is the ratio between the length and the range of its small-world shortcuts. In large aspatial ring networks, the maximum of this ratio is high because $1 : (n/2)$ is high. In square grids, the maximum of this ratio is low because $(2^{1/2}n^{1/2}) : (2n^{1/2})$ simplifies to $0.71 : 1$, which is quite low. Yet real-world geographic landscapes do exhibit small-world properties, primarily due to the technological advantage of their shortcuts.

To model this, we generalize our synthesis of small-world networks and unify both aspatial and spatial small-worlds by introducing a *contraction factor* multiplier for the length (weight, time, cost, etc.) of each small-world shortcut. Let C denote the value of this contraction factor, which may be any real number between 0 and 1. Let x denote the uncontracted length of a small-world shortcut, and x' denote its contracted length. Then $x' = C \cdot x$.

The simplest such model would simply choose a contraction factor $C \in [0,1]$ and apply it to each small-world shortcut generated for the network. More complicated models might apply different values of the contraction factor to different shortcuts, or perhaps even assign the value of the contraction factor as a linear or non-linear function of properties of the shortcut itself. In principle, contraction factors could also have values greater than one, which would still represent a shortcut in many networks, merely a less effective shortcut than its corresponding aspatial or Euclidean value.

Contraction factors model improvements in spatial technologies as shortcuts become faster and cheaper over time. In geographic landscapes, the most natural interpretation of a contraction factor's value is its relative technological advantage in speed or cost with respect to the base technology. So a synthetic landscape that has one uniform contraction factor could be interpreted as a landscape spanned by two spatial technologies, one for the base and one for the shortcuts. Alternatively, the base could be interpreted as an isotropic plane, with a single network of spatial technology shortcuts super-imposed. The application of several different contraction factors to various sets of shortcuts would be interpreted as a selection of complementary or competing spatial technologies. Finally, the application of non-linear contraction factors could be interpreted as corresponding to various economic pricing schemes such as are often used to separate shipping costs into fixed and variable components according to the nature of the cargo.

Distance Decay

Distance decay is the usual geographic term reflecting a diminished effect or degree of interaction with respect to greater distance. In principle, this generalization could be used to alter the probability of assigning a shortcut's destination node either to nodes that tend to be closer or to nodes that tend to be farther away. For example, given a particular origin node for the small-world shortcut, its corresponding destination node could be interpreted to be either more or less likely to be chosen, according to some function of its distance from the origin node.

Positive Feedback

Similarly, define the degree of a node to be the number of links of any type attached to it. Then we may also define the probability of selecting a particular shortcut destination node according to its respective degree relative to the degrees of other nodes in the landscape. Again, we usually think in terms of positive feedback with respect to degree, but this is fully generalizable and could just as easily be modeled as negative feedback that would make a node less likely to be chosen as a destination if it were already well connected. Refinements of this principle could also pay attention to the type of links attached to a particular node, so that nodes are either more or less likely to be selected as shortcut destinations depending upon the degree to which they are or are not already well connected to a particular type of small-world shortcuts.

Addition of New Links

The net addition of new links could obviously affect the outcome of a particular model. The original small-world paper by Watts and Strogatz (1998) held the number of links constant by rewiring edge links rather than adding new links as shortcuts. For real-world landscapes, the base geography would generally be affected in different ways by the addition of new links, depending upon the scale of the model and upon the interpretation of the various types of spatial technologies. This last extension is important to note, yet it is not essential to the distinction between spatial and aspatial small-worlds.

4.4 Extensions and Calibration of Synthetic Spatial Small-world Networks

Section 4.3 presented a general conceptual model for the synthesis of spatial small-world networks that may also incorporate spatial analogs for the characteristics of scale-free networks. Yet this is only the simplest beginning. Each could be further generalized or calibrated, depending upon the modeling task at hand. Different contraction factors can be applied to different sets of shortcuts in order to model disparate spatial technology networks. Variable or non-linear contraction factors may be calibrated to represent more realistic technological and economic relationships among the various spatial technologies.

In turn, distance decay or expansion, as well as positive or negative feedback with respect to various types of node degree, may be tailored to actual or theoretical properties of specific organizational or geographic networks. Parameters for each may be tuned until the characteristics of the synthetic spatial small-worlds correspond to the characteristics of existing real-world networks suitable for each model. Finally, the spatial small-world extensions discussed above may be applied to either one or both nodes for each shortcut.

The following lists indicate GeoGraph agent-based comp lab options for modeling landscapes, initial population distributions, and agent travel or relocation decisions:

A. Network Landscapes (usually Spatial, but may be Aspatial)

Nodes may be distributed:

- as a circle
- as a square grid
- randomly
- from data (e.g. geographic coordinates read from a file)

Base links may be distributed:

- between immediate neighbors on the circle (i.e, as its circumference)
- between orthogonal immediate neighbors on the square grid
- as radial dendrites from a core node to peripheral nodes on square grid
- from network data read from a file (with optional attributes)

Base links may be weighted:

- 1 for all links (a binary/aspatial network, with no distance or weights)
- Euclidean distance for each link (determined by node coordinates)
- from data, providing distances or costs for each link (read from the link file)

Shortcut links may be distributed:

- Stochastically according to small-world logic (between randomly selected nodes)
- Stochastically according to scale-free logic (to a node according to its link count)
- Stochastically according to Dibble (2001b) and Dibble and Feldman (2004) spatial small-world logic, where both nodes are selected stochastically according to:
    - distance (with positive or negative weighting for distance decay)
    - degree (with positive or negative weighting for node's # of links)

B. Initial Population Distributions

Agents may be distributed:

- entirely to one node (whether there is one node or many in the landscape)
- uniformly across nodes (population must be an integer multiple of the # of nodes)
- stochastically across nodes, with variability ranging from 0 (1 node) to 1 (uniform)
- from data, providing integer populations for each node

C. Simple Models of Agent Travel Patterns

In a locational game, each agent may decide at each turn where it would prefer to be, which may or may not lead to a decision to relocate. In a simulation of an epidemic, an agent may decide whether to travel to another node based on endogenous perceptions of local risks or on a stochastic model parameter such as travelProbability. In turn, evaluations of potential destinations could be modeled as one of the following:

- random node—randomly choose any node in the landscape,
- random neighbor—randomly choose any node that is one link away from current,
- random base neighbor (randomly choose one node that is one base link away)
- local or global characteristics of a node, such as its current residents or neighbors,
- gravity model—choose nodes in the landscape according to a probability distribution across k destination nodes that is proportional to:

$$\text{gravityWeight} = (\text{jPop}^{\text{tau2}} \cdot \text{kPop}^{\text{tau1}}) / \text{distance}^{\text{rho}}$$

where:

gravityWeight = weight for the number of trips from node j to node k
j, k = nodeIDs for the from and to nodes, respectively
jPop = population of the origin node
kPop = population of the (potential) destination node
tau2 = scaling power for the origin node

tau1 = scaling power for the destination node
rho = distanceDecay

Note: tau2 and jPop can be ignored, as these will be the same for any given node. We only need to find the probability of a given destination node. Origin probability already scales with respect to jPop via travelProbability · jPop.

Here, a simple gravity model simplifies to:

gravityWeight = kPop / $\text{distance}^{\text{distanceDecay}}$
where: distanceDecay defaults to 2.0

When we assume that node populations remain constant, this means that each node calculates a roulette wheel of proportional probabilities for each of the n-1 other nodes in the landscape, where all probabilities sum to 1. Any agent who travels from that node thus uses that roulette wheel to decide where to go. The roulette wheel remains constant for each node during each simulation. When node populations fluctuate significantly during a simulation, each node would need to recompute its roulette wheel of population-weighted probabilities for each of the n-1 other nodes in the landscape.

5 Examples: Games, Diffusion, Innovation, and Globalization

This section introduces several spatial processes and related practical applications in order to illustrate the kinds of problems that may be explored via comp labs for spatial agent-based models.

Locational games and coordination problems include many familiar ACE examples such as the segregation tipping game, audience seating, and cafeteria selection first introduced by Schelling (1978). These also include the El Farol spatio-temporal bar-coordination problem, and the Standing Ovation Problem used by Miller and Page (2004) to teach agent-based modeling.

Diffusion processes become especially interesting when the diffusion occurs among mobile heterogeneous agents interacting on richly structured landscapes. Understanding diffusion dynamics can be especially important to encourage the diffusion of important information or to inhibit the diffusion of deadly infectious diseases.

Ecological innovation refers to the degree to which diffusion of heterogeneous ideas or viruses may encourage or inhibit opportunities for their recombination within agents to create new ideas, inventions, or deadly diseases.

Finally, increasing population densities and improvements in transportation and communication technologies facilitate globalization processes by profoundly altering the ease, frequency, and range of spatial interactions. In turn, such changes affect locational choices, spatio-temporal coordination, diffusion processes, and ecological innovation at all scales.

5.1 Locational Games and Spatio-Temporal Coordination Problems

Consider a class of models where agents play a locational game on the nodes of a network landscape. This is modeled as a game rather than as an individual optimization problem because the payoffs for each agent's locational choice are an endogenous function of the locational choices of the other agents. Persistent configurations reflect the Nash Equilibrium (Nash 1950, Nash 1951) aspect of settlement patterns: each agent's objective score is affected by the locations of other agents, so each agent's location (its strategy) is best subject to where everyone else has decided to locate (their strategies).

A Nash equilibrium need not be unique, and Nash equilibria may have very different degrees of stability with respect to small perturbations or errors in one or more of the strategies played. In non-zero-sum games such as these, some equilibria will generally have very different average payoffs than others. Mutually beneficial equilibria may evolve over time in repeated games, even when there exist no formal or even informal mechanisms to support cooperation among the agents. Analogs to many of these questions become especially interesting when considering the independent, localized actions of non-cooperative agents acting in space, with varying degrees of imperfect (local) versus perfect (global) information and coordination mechanisms.

If it were possible to logically derive all such results with respect to locational choices of individuals via spatial analysis or game theory, then we would solve such puzzles using only formal analysis, and there would be no need for simulation models. Simulation models are useful for modeling locational games precisely because of the degree of complexity introduced by the non-linear and widely divergent payoffs that arise from individual agent contexts, embedded in each unique spatial situation that may emerge.

Yet any persistent spatial configuration that emerges when agents are free to move is in effect a *generalized Nash equilibrium (gNE)*; given the configuration of agents in the landscape, individual agents may continue to make marginal moves among nodes between which they are indifferent, yet the spatial configuration generally remains unchanged (Dibble 2001b). The term "generalized" contrasts with a pure locational Nash equilibrium, in which no agent would move once the equilibrium was reached. In a generalized locational Nash equilibrium, the overall characteristics of the configuration remain unchanged and agents who move do so only on the margin. In other words, they are sufficiently happy with their locations that they would move only to approximately equivalent positions, and they would not move at all if we applied even a modest epsilon moving cost. Thus, we could formalize the notion of a gNE_{ε} that applies a moving cost of ε to induce a pure locational Nash equilibrium. Simulations here do not impose ε but such models could be used to study persistent configurations.

Similarly, we may define a *distributed Nash equilibrium* (*dNE*) according to the degree that local neighborhoods overlap by some percentage of each landscape's spatial diameter; a *range* treatment variable (Dibble 2001b). As *range* converges to 0, strategic interactions among agents converge to normal, isolated games among strictly local sets of agents. As *range* converges to 1, this becomes simply one global game involving all of the agents in the landscape. But for intermediate values of *range*, where neighborhoods for each local game overlap for some strategic positions or some players, we introduce the potential for tremendously interesting and empirically important percolation dynamics as strategic interactions from one semi-isolated local game ripple through the landscape to disturb or encourage equilibria in neighboring games.

5.2 Diffusion of Ideas, Information, or Infectious Diseases

Mathematical models of diffusion and percolation among static populations are well established. See for example Nowak and May (2001), and Stauffer and Aharony (1994). The new science of networks (see Barabási (2002) or Strogatz (2001)) continues to offer important breakthroughs on the immunization of complex networks (Pastor-Satorras and Vespignani 2002) and halting viruses in scale-free networks (Dezsó and Barabási 2002). Yet each of these models is limited to static social networks where the pattern of encounters between agents has been fixed *a priori* and remains static throughout the analysis.

Consider exploration of the fundamental characteristics of diffusion among populations of mobile, heterogeneous agents where the pattern of encounters is dynamic throughout the simulation and where the chance of transmission – whether of a virus or of an idea – is dependent upon the characteristics of each agent involved in each encounter. Encounters among agents are structured by two types of networks: social networks determine friendships, yet spatial networks determine the geographic distribution of social agents. For example, I may be able to catch an idea via email, but I am unlikely to catch a cold via email from a friend who is currently living far away.

5.3 Innovation and Ecological Emergence of New Ideas, Inventions, or Diseases

Even in the most sophisticated models, diffusion is treated as though there is a single element to be traced, and as though its effect upon each individual is independent of any competing or complementary elements that may be diffusing or have diffused in the past among members of the population in question. Yet the diffusion of complementary or competing infectious disease strains or ideas may make all the difference. Simply setting up multiple diffusions among a population is not especially difficult. What has been missing is some way to model the cumulative effects of multiple exposures (e.g. for anthrax or even for SARS) and the internal evolutionary effects of the competition or complementarity of multiple strains within each individual. What happens when otherwise relatively benign disease strains or ideas combine within affected individuals to evolve into an especially virulent strain?

This fundamental scientific question could have practical relevance such as:

- Globalization increases the range and frequency of interactions among distributed mobile agents, thus providing ideal conditions for exposure to multiple disease strains. Moreover, globalization facilitates the emergence of new diseases as strains from different species in distant parts of the world encounter one another and recombine to form novel and potentially lethal strains such as H5N1 Avian Influenza.
- During volatile social conditions, conditions favoring civil unrest are likely to be fostered by the diffusion of multiple competing or complementary rumors and ideas. Understanding their dynamics both among and within agents is important for understanding how best to respond to their emergent effects such as riots, looting, sniping and related resistance, or genocide against unpopular civilian factions. Sufficiently deep understanding of these models may lead not only to effective response but perhaps also to effective control or prevention, as peacekeeping teams diffuse competing information to calm the situation.

Research domains such as multi-scale diffusion of competing and complementary infectious diseases, behavioral norms, and social, economic, political, or technological ideas provide examples of multi-disciplinary research on fundamental ecological mechanisms and the dynamics of richly structured systems of interacting agents. Comp lab research for these domains benefits from our ability to simulate the full multi-scale range of dynamic processes, from the adaptive mechanisms within agents such as the internal genetic algorithms introduced in Dibble (2001a) and in Brenner (2006) and Duffy (2006) to the richly structured organizational and geographic landscapes introduced here and in Dibble (2001b), Dibble and Feldman (2004), and Wilhite (2006).

5.4 Globalization Processes and the Effects of New Technologies

Finally, consider a model of globalization where decreases in costs and other barriers to long-distance communication, exchange, travel, and migration facilitate increased ranges and frequencies of many types of spatial interactions among geographically separated agents. In the long run, reductions in the costs of spatial interaction generally lead to corresponding locational, socio-economic, institutional, and infrastructure adjustments as well, as agents adapt to the new costs of interaction.

Dibble (2001b) presents a simple example of a model of this type, which makes use of the GeoGraph Comp Lab spatial small-world landscapes described in Dibble and Feldman (2004). First, decreasing contraction factors model technological improvements or falling fuel prices, which reduce the impedance of distance along selected network shortcuts. Second, agents respond to the new contraction factors immediately via a change in their analysis of their current locations, by implicitly increasing their interaction range. In turn, such increases in their scales of interaction may result in direct or indirect incentives to relocate, as other nodes become more attractive either directly through their improved relative accessibility or indirectly through the relocations of other agents as disturbances ripple through the landscape.

6 Exploration and Analysis of Spatial System Behaviors

> always the more beautiful answer who asks the more beautiful question.
> (e. e. cummings, i: six non-lectures)

There is a fine art to designing any research model, not merely agent-based. There is of course the obvious necessary condition that it be capable of generating—either deductively or via simulation—the phenomenon of interest. See Epstein and Axtell (1996) and Epstein (1999, 2006). Yet generative models are necessary but not sufficient for the effective conduct of agent-based research (Epstein 2006). Undue focus on the model itself provides insufficient guidance for asking meaningful questions of our models, and for undertaking related refinements and further exploration. What matters most is understanding when a simulation model is required in order to generate insights that are not available in other ways, to be clear about what we can *learn* from the model, what is truly new about its results, and how much we can trust what we learn.

6.1 Parsimony and Predictive Accuracy

> Predictive accuracy concerns a model's fit to the population, whereas postdictive accuracy concerns a model's fit to a sample. (Gauch 2003, 280)

Parsimonious modeling means selecting the simplest possible model capable of generating a phenomenon of interest. Although Gauch (2003) addresses the importance of parsimony for statistical modeling of empirical data, his distinctions between predictive and postdictive accuracy and between signal and noise provide compelling arguments for the importance of parsimony in agent-based modeling as well.

As with econometric models, a highly complicated agent-based model that has many types of agents and a large number of parameters provides additional degrees of freedom that allow it to adjust to stochastic noise and thus to fit too closely to sample data. The costs of developing, calibrating, and running complicated agent-based models can be especially prohibitive. Yet complicated models risk overfitting to sample data, which provides merely *postdictive* accuracy that is least valuable for generalization beyond the sample data. In contrast, parsimonious models are more useful to the extent that they can generate crucial *predictive* accuracy regarding the full population of potential outcomes for whichever spatial processes we seek to understand.

Parsimonious spatial agent-based modeling provides two further advantages as well. First, parsimonious models are generally far easier, faster, and less expensive to develop, test, calibrate, and run. Second, systematic exploration of the interactions among key parameters affecting initial conditions or model behavior can be challenging due to the combinatorial explosion of parameter interactions even for simple models. Thorough exploration of model behavior quickly becomes prohibitive for highly complicated simulation models, as such models suffer exponentially from the effects of combinatorial explosion among their parameters, with the additional burden that each simulation generally takes far longer to run.

6.2 Preliminary Thought Experiments

> Theories help experimentalist[s] to design incisive experiments, and experiments yield results that guide the thinking of theoreticians. They coevolve. The indispensability of experiments in scientific research and the surprises experiments constantly throw up suggest that nature has many emergent characteristics in store. … Emergent characters mostly belong to the *structural* aspect of systems and stem mainly from the organization of their constituents. … Emergence is closely associated with microexplanations of macrophenomena.
>
> (Auyang 1998, pages 175-176)

As first mentioned in the introduction to this chapter, Schelling's *Micromotives and Macrobehavior* (1978) provides compelling examples of incisive questions about the behavior of distributed agent systems, where each question is framed sufficiently clearly to be answerable, of necessity, purely via thought experiments. Ideally, framing research questions for agent-based comp lab research does begin with questions answerable by thought experiments. Indeed, testing results from such a simulation model against logically derived results from thought experiments can be one of the most powerful methods both for aligning the simulation model with extant theory and for checking the implementation and reasonable behavior of the model. Both approaches are highly recommended. Yet neither is sufficient to establish new insights.

Thought experiments can save years of modeling and guide researchers toward essential questions that cannot be answered in any other way. Nevertheless, to establish new results, research with an agent-based model must extend to unknown territory, where macrophenomena emerge beyond what could be predicted by a careful thought experiment applied to knowledge of initial micro conditions. If you know the values for all of your explanatory variables, can you predict the values of your dependent variable? That's appropriate for early calibration and testing for a new model. Yet there is no need for simulation modeling if true emergence or surprises are never possible. Finally, can you tell from thinking in depth about your model that you are likely to find trustworthy treasure out there, for reasonable costs, risks, and opportunity costs of the research?

In order to be worthy of study, an emergent property needs to be more than merely something that is not programmed into the individual agents; it must meet the stronger standard that it arise from interactions among agents in a way that could not easily be predicted simply by knowing the micro specifications or objectives of the agents. For example, if you program agents to seek high concentrations of sugar in the landscape, their congregation at such places once the simulation runs is easily predicted and would not be considered an emergent property of the model.

6.3 Scaling Agent-Based Simulation Models

Well designed general purpose simulation modeling tools should be able to represent any spatial or temporal scale, depending for visual clarity only upon suitable cartographic generalization in the representation of the landscape and of its agents. For example, each GeoGraph node or agent can be scaled visually from the tiniest dots to a fully detailed graphic image. Similarly, nodes in a network landscape may be interpreted, and represented, as anything from tables in a café to world cities or even planets. Agents for each landscape may be scaled in proportion to the geographic scale, enlarged to facilitate visualization of the model's behavior, or reduced to avoid visual clutter.

The maximum number of nodes and agents for an agent-based simulation model depends upon the fidelity of the respective node and agent classes. Specifically, it depends upon the CPU time and memory requirements for the behavior and visualization of each agent and for each of the other components of the model. Interactive display of the landscape and its agents generally places a greater burden on the computer and thus limits the number of nodes and agents by approximately one order of magnitude, compared with running the equivalent simulation with visualization of real-time data charts but with the landscape graphics hidden.

The demographic and geographic resolution of an agent-based simulation model may be confused informally with scale, yet resolution refers in each case to the number of simulated units per unit in the real world. For example, to study diffusion processes in a true population of 1,000,000 people, we could run models that have resolutions of 1,000, 10,000, or 1,000,000 agents.

Although CPU time and memory requirements may determine the maximum number of nodes or agents for simulations on any given computer, it could be a grave mistake to assume that more of either is necessarily desirable. Higher resolution should be used only when it turns out to be important to the behavior of the model. Unnecessary demographic or spatial resolutions that extend the duration of each simulation impose an opportunity cost by limiting the number of research questions, scenarios, and stochastic replicates that can be explored. Sufficient lower bounds for the resolution may be specific to each model or to each research question. These can be established by comparing simulation outcomes from controlled simulations that are identical except for their levels of resolution.

6.4 Visualization and Data Collection

In our own laboratory work, we typically develop and debug each model using a 2-dimensional or 3-dimensional graphical user interface with a landscape and population of agents comfortable for visualization. After we are thoroughly comfortable with the model and have tested it extensively, we can turn off the visual landscape display and scale up the numbers of nodes and especially the numbers of agents for batch-mode experiments, if that turns out to be important scientifically for the model's results. Depending upon the model and upon the level of abstraction and resolution appropriate for rigorous results, very large numbers of agents can be available but are not necessarily required.

For batch-mode experiments, each simulation logs at least one record—and often several detailed files—to a text file for subsequent analysis. Log files can be inspected visually, but are best read into a more powerful database management system such as the

Statistical Analysis System (SAS www.sas.com). SAS's macro capabilities are especially effective for reading tens of thousands of log files at a time, and SAS's plotting, regression, data management, and data filtering capabilities support statistical inference, quality control, and exploration of anomalies and other surprises.

6.5 Practical Advice for Implementation and Testing

> This problem, too, will look simple after it is solved.
> (Charles Francis Kettering)

The ACE web site provides links to currently available comp lab platforms (Tesfatsion 2006b). Look for the comp lab that has the best comparative advantage for your purposes. In particular, look for class hierarchies that best support the agents, interaction structures, and laboratory tools you plan to employ. In many cases, existing demonstration models and prototypes may provide classes for types of agents that can be modified to develop early proof-of-concept models.

A comp lab development platform that is widely used serves an important role as standardized laboratory equipment, which provides scientific advantages for publication, evaluation, and replication of results. So far, Swarm (www.swarm.org), RePast (repast.sourceforge.net), and AnyLogic (www.xjtek.com) have played key roles in this regard. Agent-based comp lab research is still quite new, however, so it is important to balance the advantages of standardized equipment against important comparative advantages due to specialized or extended comp lab platforms such as the Trade Network Game (TNG) and SimBioSys (McFadzean and Tesfatsion 1999, McFadzean et al. 2001) or GeoGraphs (Dibble and Feldman 2004).

Testing Model Components

> The most exciting phrase to hear in science, the one that heralds new discoveries,
> is not 'Eureka!' (I found it!) but 'That's funny...' (Isaac Asimov)

Test *everything*. Put proto-agents and each model agent through its paces one at a time, to test their responses under controlled conditions. Then test agents and agent interactions again within the smallest possible controlled groups to test whatever interactions with other agents may be part of the model. Test their landscape or other interaction structure. Test the agents' ability to gather local or global info, to make appropriate choices, to move about within the structure, and, if relevant, to endogenously modify their landscape or interaction structure.

One of the essential yet often overlooked practices in comp lab work is to have a careful look at the raw data. Of course this is essential for superficial quality control: Did every simulation run correctly? Did anything go wrong, break down, turn up missing, or behave pathologically?

Yet examinations of raw data serve a far deeper scientific purpose than mere quality control. Although we may expect scientific surprises and insights to occur as a result of analytical procedures such as statistical analyses, they may also occur at the level of the raw data. Consequently, this basic step can be central to the scientific process of observation, insight, and understanding. Do the outcomes make sense? When they don't make sense, are we sufficiently alert to their "that's funny" intimations, however inconvenient they may be for our preconceived immediate purpose?

7 Genetic Algorithm Inference, Optimization, and Risk Analysis

Miller (1998) proposed to use a supervisory genetic algorithm to perform what he called "active nonlinear tests" (ANTs) by using the genetic algorithm to challenge each simulation model by seeking outcomes that provide exceptions or counter-examples to its usual results. This section briefly discusses a generalization of Miller's ANTs to the broader problem of providing effective search and optimization across both treatments and outcomes for a model.

7.1 Exploring Model Behavior

Systematic analysis of model behavior may involve millions of simulation runs, each controlled by sweeping across discrete lists of values for sensitive model parameters, for each treatment variable of interest, and for seeds to control one or more random number series for stochastic simulations. For example, consider even a very simple model that has four parameters, each with three levels to evaluate for sensitivity, and two treatment variables, each with five levels to evaluate for an experiment. This requires 3x3x3x3x5x5=2,025 simulation runs, even for a deterministic model where a single simulation is sufficient to evaluate each combination.

When stochastic behavior, synthetic landscapes, synthetic agent populations, and stochastic initial conditions are involved, even minimal evaluation of representative outcomes for each vector of parameters and treatment variables may require multiple seed values for one or more separate random number series. For example, in our GeoGraph framework, even a basic evaluation of epidemic outcomes requires analysis of combinations of three sets of random number seeds. For synthetic landscapes such as globalization networks, one random number seed determines landscape details (geoSeed). Another random number seed determines stochastic initial conditions such as population distributions and the locations of initial cases of the disease (iniSeed). A third controls all stochastic actions such as who travels where and who infects whom (actSeed).

Such controls enable us to replicate experiments to systematically evaluate the effects of stochastic events under controlled conditions. Yet even ten histories for each of the three random number seeds would require running 10x10x10=1,000 stochastic replications for each of the 2,025 distinct combinations of parameter and treatment levels in the example above, for a total of 2,025,000 simulations just for one simple experiment.

Of far greater importance scientifically, the standard focus on exploring model behavior via combinatorial sweeping across regularly spaced parameter values is a blind search for significant outcomes. As illustrated in Figure 3, regularly spaced parameter values may be entirely unrelated to the truly important parameter values where model behavior may reach significant extrema.

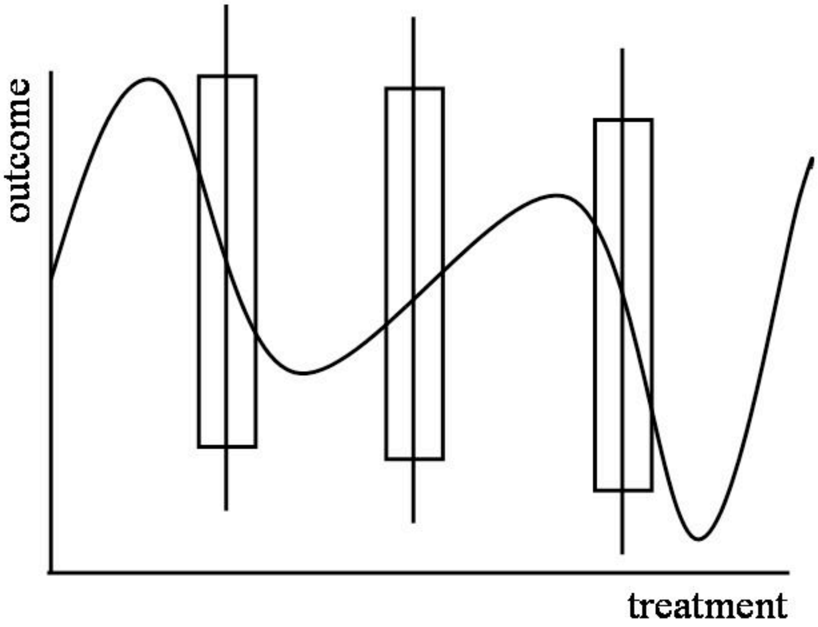

Figure 3: Running only a few stochastic replicates of each treatment level can result in variances so large that the signal becomes lost in the noise. Similarly, selecting treatment levels blindly via random or regular spacing may completely miss important local and global extrema.

Ideally, we would like to be able to search for interesting behavior in the outcome space rather than sweeping blindly in parameter space. As illustrated in Figure 4, a supervisory genetic algorithm allows us to do precisely that, and with far greater efficiency than brute force combinatorial sweeping of parameter spaces. To do so, the genetic algorithm can be set up to search across combinations of key parameters for extreme values of single or multiple combinations of outcome variables, based on results from one or more stochastic replications of the scenario that is associated with each combination of key parameters. In addition, the greater economy in searching for key scenarios releases computational resources that may in turn be used to simulate sufficient stochastic replications for each to be able to distinguish statistically significant differences among scenario outcomes.

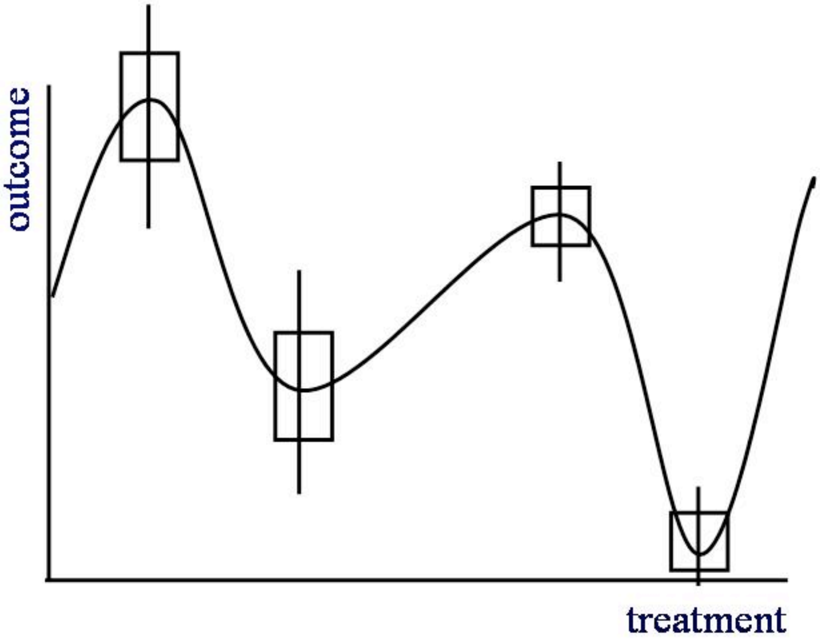

Figure 4: In contrast, an ideal experimental design runs enough stochastic replicates for reliable inference. Similarly, data-driven experimental designs may provide guidance for identification of key values for treatment variables and for basins of attraction leading to common outcomes.

7.2 Inference and Discovery of Key Exceptions

Using supervisory genetic algorithms to discover highly effective treatments or to search for exceptional or surprising simulation outcomes has the potential to profoundly enhance our ability to make the most effective use of limited computational and analytical resources. It permits us to discover and test incisive empirical insights, effective normative designs or interventions, and surprising heuristic insights. Once such treatments or outcomes have been identified by the genetic algorithm, subsequent ordinary batches of simulations can be carefully targeted in order to evaluate the accuracy, uncertainty, risk, and inference power of results obtained from any well-specified agent-based simulation model.

Finally, a supervisory genetic algorithm can be used to thoroughly explore the robustness with respect to risk of promising designs or interventions. When used to discover effective designs or interventions, the genetic algorithm searches across combinations of scenarios or intervention parameters for those that perform best across the population of simulations that is run to evaluate each string. In contrast, when the genetic algorithm is used to evaluate the stochastic risks associated with each intervention that corresponds to specific values for the treatment variables: it holds constant the model's sensitive parameters and treatment variables, and instead searches across random number seeds for

"Murphy's Law" worst-case combinations of stochastic events, where everything that could go wrong does go wrong. Some interventions may be far more resistant to worst case outcomes, which may be crucial to evaluate when stakes are high.

Figure 5 illustrates a supervisory GA where fitness for each GA string of model parameters is determined by the outcomes of large numbers of agent-based simulations based upon those parameters.

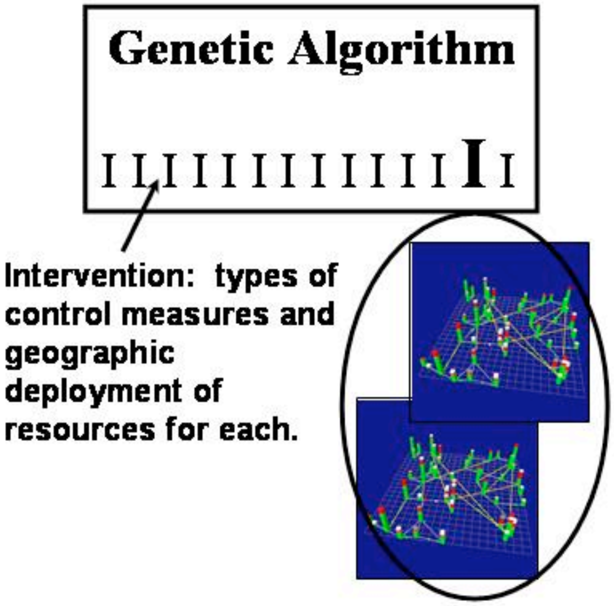

Figure 5: A supervisory Genetic Algorithm evolves populations of strings of simulation parameters by evaluating combinations of simulation parameters according to the results of one or more simulation runs based on those parameters.

8 Opportunities, Challenges, and Resources

Most of the comp lab issues discussed in this chapter have addressed support for relatively abstract theoretical research with spatial agent-based models. Yet comp lab models and tools such as supervisory genetic algorithms for exploration, optimization, and risk analysis can potentially support policy-relevant computational spatial social science. Spatial agent-based models are beginning to be used to provide decision-support for complex emergencies such as controlling the spread of infectious diseases.

In many cases, research on abstract landscapes provides sufficient insight to support real-world decisions. For this, synthetic landscapes such as those introduced in section 4 provide powerful scientific leverage. Comp labs would benefit from corresponding tools for populating agent-based models with initial distributions of carefully calibrated synthetic demographic populations.

Nevertheless, one of the most important methodological challenges facing comp labs for spatial agent-based models is to provide easy-to-use tools for importing real-world geographic landscapes from Geographic Information Systems (GIS) or satellite remote sensing data. In turn, closer interoperability between comp labs and GIS would provide spatial agent-based modeling researchers with access to the powerful tools for spatial analysis, visualization, and mapping that are incorporated in modern GIS.

A complementary methodological frontier, for both spatial and aspatial agent-based modeling, would provide comp lab tools to streamline the art of developing families of models ranging from highly abstract parsimonious theoretical models to highly calibrated models for policy and testing against empirical data.

We may also cross the ultimate boundary from *modeling* cognitive agents in our comp lab models to *incorporating* live human cognitive agents in our spatial agent-based models. This blurs the distinction between experimental economics with humans under controlled conditions in a laboratory versus experimental economics with humans under controlled conditions interacting within a virtual spatial environment in a comp lab (Macmillan 1996).

Finally, as hinted at in the thought experiment on pandemics in section 1, we may begin to combine policy-relevant comp lab simulations and inference with near real-time collection and relevance filtering of real-world surveillance data. This would permit us to improve our models in general and, more importantly, to learn how to signal when it is appropriate to begin, modify, or conclude a specific intervention. Pure comp lab experiments with spatial agent-based models on geographic landscapes may pave the way for new and effective uses of near real-time remote sensing satellite data or social systems sensing data (Gelernter 1992) as it arrives.

For example, comp lab simulations may be used to identify sensitive empirical indicators and corresponding critical thresholds for phase transitions in complex dynamic processes such as pandemics of infectious diseases, civil violence, or business cycles. When is an epidemic in danger of flaring into a full global pandemic? When and where should interventions to control the outbreak begin? What are the early warning signs that an outbreak of disease or civil violence or a downturn in the business cycle has turned the corner toward successful containment or recovery?

To return to our original thought experiment, global pandemics of infectious diseases have been infrequent events that create considerable disruption and are extraordinarily difficult to study, much less to control or to repeat in order to explore the effects of chance events. Spatial agent-based computational laboratories allow us to learn from virtual experience by creating and studying millions of simulated pandemics under controlled conditions, to evaluate risks and to explore effective methods for their prevention and control.

In turn, as illustrated by section 5, pandemics are merely one example of broad classes of complex social and environmental spatial dynamic processes. Effective spatial agent-based models and fully equipped computational laboratories provide opportunities to learn from far more extensive and better controlled virtual experience than could ever be generated by real-world systems: to understand the fundamental driving forces within such systems, and to discover and evaluate wiser designs, more effective interventions, and the risks or resilience associated with each.

Acknowledgements

For their thoughtful recommendations on earlier drafts I thank Leigh Tesfatsion, Allen Wilhite, and three anonymous reviewers. My thanks to Stephen Wendel for his research expertise and insightful comments. For financial support I thank the Office of Naval Research, the Environmental Protection Agency, and the National Institutes of Health, National Institute of General Medical Sciences.

References

Albert, Réka, Hawood Jeong, and Albert-László Barabási. 2000. Error and Attack Tolerance of Complex Networks. *Nature* 406: 378-82.

Albert, R., H. Jeong, and A. L. Barabasi. 1999. Internet - diameter of the world-wide web. *Nature* 401, no. 6749: 130-131.

Aldhous, Peter, and Sarah Tomlin. 2005. Avian Flu: Are we ready? *Nature* 435, no. 7041: 399.

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

Axtell, Robert, Robert Axelrod, Joshua M. Epstein, and Michael D. Cohen. 1995. Aligning Simulation Models: A Case Study and Results. *Santa Fe Institute Working Paper* 95-07-065.

Auyang, Sunny Y. 1998. *Foundations of Complex-System Theories in Economics, Evolutionary Biology, and Statistical Physics*. Cambridge, UK: Cambridge University Press.

Barabási, Albert-László. 2002. *Linked: The New Science of Networks*. New York: Perseus Publishing.

Barabasi, Albert-Laszlo, and Reka Albert. 1999. Emergence of Scaling in Random Networks. *Science* 286: 509-12.

Beveridge, W. I. B. 1957. *The Art of Scientific Investigation*. Third ed. New York: Vintage Books.

Brenner, T. 2006. Agent learning representation: Advice on modelling economic learning. *This Handbook.*

chromatic. 2003. *Extreme Programming Pocket Guide*. Sebastopol, CA: O'Reilly Media.

Couclelis, H. 1994. Spatial technologies. *Environment and Planning B-Planning & Design* 21, no. 2: 142-43.

DeCanio, Stephen J., Catherine Dibble, and Keyvan Amir-Atefi. 2000. Organizational Structure for the Adoption of Innovations. *Management Science* 46, no. 10: 1285-29.

———. 2001. Organizational Structure and the Behavior of Firms: Implications for Integrated Assessment. *Climatic Change* 48: 487-514.

Dezsó, Zoltán, and Albert-László Barabási. 2002. Halting Viruses in Scale-Free Networks. *Physical Review E* 65, no. 5: 4.

Dibble, Catherine. 2001a. Beyond Data: Handling Spatial and Analytical Contexts with Genetics Based Machine Learning. Chapter 3 in: *Spatial Evolutionary Modeling.* Editors Roman Krzanowski, and Jonathan RaperOxford: Oxford University Press.

Dibble, Catherine. 2001b. Theory in a Complex World: GeoGraph Computational Laboratories. Unpublished Doctoral Dissertation. University of California 2001.

Dibble, Catherine, and Philip G. Feldman. 2004. The GeoGraph 3D Comp Lab: Network and Terrain Landscapes for RePast . *Journal of Artificial Societies and Social Simulation <http://jasss.soc.surrey.ac.uk/7/1/7.Html>* 7, no. 1.

Duffy, J. 2006. Agent-based models and human-subject experiments. *This Handbook.*

Epstein, Joshua. 1999. Learning To Be Thoughtless: Social Norms and Individual Computation. *Brookings Institution, Center on Social and Economic Dynamics, Working Paper Number 6.*

———. 2006. Remarks on the foundations of agent-based generative social science. *this Handbook.*

Epstein, Joshua, and Robert Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Washington, D.C.: Brookings Institution, MIT Press.

Fontana, Walter, and Leo W. Buss. 1994. What Would be Conserved if "The Tape Were Played Twice"? *Complexity: Metaphors, Models, and Reality.* Editors G. Cowan, D. Pines, and D. Meltzer, 223-36. New York: Addison-Wesley.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venebles. 1999. *The Spatial Economy*. Cambridge, MA: The MIT Press.

Gauch, Hugh G. 2003. *Scientific Method in Practice*. Cambridge, UK: Cambridge University Press.

Gelernter, David. 1992. *Mirror worlds: Or the day software puts the universe in a shoebox: How it will happen and what it will mean*. New York: Oxford University Press.

Goldberg, David E. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley.

Holland, John H. 1992. *Adaptation in Natural and Artificial Systems*. Second ed. Cambridge, MA: The MIT Press.

Jeong, H., B. Tombor, R. Albert, Z. N. Oltval, and A. L. Barabasi. 2000. The large-scale organization of metabolic networks. *Nature* 407, no. 6804: 651-54.

Macmillan, Bill. 1996. Fun and Games: Serious Toys for City Modelling in a GIS Environment. *Spatial Analysis: Modelling in a GIS Environment.* Editors Paul Longley, and Michael Batty, 153-65. New York: John Wiley & Sons.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory*. New York: Oxford University Press.

McFadzean, David, Deron Stewart, and Leigh Tesfatsion. 2001. A Comp Lab for Evolutionary Trade Networks. *IEEE Transactions on Evolutionary Computation* 5, no. 5: 546-60.

McFadzean, David, and Leigh Tesfatsion. 1999. A C++ Platform for the Evolution of Trade Networks. *Computational Economics* 14: 109-34.

Miller, John H. and Scott E. Page. 2004. The Standing Ovation Problem. *Complexity* 9, no 5: 8-16.

Miller, John H. 1998. Active Nonlinear Tests (ANTs) of Complex Simulation Models. *Management Science* 44, no. 6: 820-830.

Nash, John. 1950. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences, USA* 36, no. 1: 48-49.

———. 1951. Non-Cooperative Games. *Annals of Mathematics* 54, no. 2: 286-95.

Nowak, Martin A., and Robert M. May. 2000. *Virus Dynamics Mathematical Principles of Immunology and Virology*. New York: Oxford University Press.

Osterholm, Michael T. 2005. Preparing for the Next Pandemic. *Foreign Affairs* July/August.

Pastor-Satorras, Romualdo, and Alessandro Vespignani. 2002. Immunization of Complex Networks. *Physical Review E* 65, no. 036104: 1-4.

Samuelson, Paul. 1952. The transfer problem and transport costs. *Economic Journal* 64: 264-89.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. First ed. Fels Lectures on Public Policy Analysis. New York: W. W. Norton .

Stauffer, D., and A. Aharony. 1992. *Introduction to Percolation Theory*. London: Taylor and Francis.

Strogatz, Steven H. 2001. Exploring Complex Networks. *Nature* 410: 268-76.

Tesfatsion, Leigh. 1997. A Trade Network Game with Endogenous Partner Selection. *Computational Approaches to Economic Problems.* Editors H. M. Amman, B. Rustem, and A. B. WhinstonKluwer Academic Publishers.

———. 2000. Concentration, Capacity, and Market Power in an Evolutionary Labor Market. *Evolution at Work for the New Millenium, Proceedings of the 2000 Congress on Evolutionary Computation.*, 1033-40. Vol. II. New Jersey: IEEE, Inc.

Tesfatsion, Leigh. 2001. Structure, Behavior, and Market Power in an Evolutionary Labor Market with Adaptive Search. *Journal of Economic Dynamics and Control* 25, no. 3-4: 419-57.

Tesfatsion, Leigh. 2006a. Agent-based computational economics: A constructive approach to economic theory. *This Handbook.*

Tesfatsion, Leigh. 2006b. "Agent-Based Computational Economics (ACE) Demos Web Site." Web page. Available at http://www.econ.iastate.edu/tesfatsi/acedemos.htm.

Tobler, Waldo. 1999. The World is Shriveling as it Shrinks. *ESRI User Conference*.

Watts, Duncan J. 1999. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton, NJ: Princeton University Press.

Watts, Duncan J., and Steven H. Strogatz. 1998. Collective Dynamics of 'Small-World' Networks. *Nature* 393: 440-442.

Wihlite, Allen W. 2001. Bilateral Trade and Small-world Networks. *Computational Economics* 18, no. 1: 49-64.

Wilhite, Allen W. 2006. Economic activity on fixed networks. *This Handbook.*

Young, H. P. 2006. Social dynamics: Theory and applications. *This Handbook.*