Agent-based complex systems are dynamic networks of many interacting agents; examples include ecosystems, financial markets, and cities. The search for general principles underlying the internal organization of such systems often uses bottom-up simulation models such as cellular automata and agent-based models. No general framework for designing, testing, and analyzing bottom-up models has yet been established, but recent advances in ecological modeling have come together in a general strategy we call pattern-oriented modeling. This strategy provides a unifying framework for decoding the internal organization of agent-based complex systems and may lead toward unifying algorithmic theories of the relation between adaptive behavior and system complexity.

What makes James Bond an agent? He has a clear goal, he is autonomous in his decisions about achieving the goal, and he adapts these decisions to his rapidly changing situation. We are surrounded by such autonomous, adaptive agents: cells of the immune system, plants, citizens, stock market investors, businesses, etc. The agent-based complex systems (ACSs) around us are made up of myriad interacting agents. One of the most important challenges confronting modern science is to understand and predict such systems. Bottom-up simulation modeling is one tool for doing so: We compile relevant information about entities at a lower level of the system (in “agent-based models,” these are individual agents), formulate theories about their behavior, implement these theories in a computer simulation, and observe the emergence of system-level properties related to particular questions (2, 3).

Bottom-up models have been developed for many types of ACSs (4), but the identification of general principles underlying the organization of ACSs has been hampered by the lack of an explicit strategy for coping with the two main challenges of bottom-up modeling: complexity and uncertainty (5, 6). Consequently, model structure often is chosen ad hoc, and the focus is often on how to represent agents without sufficient emphasis on analyzing and validating the applicability of models to real problems (5, 7).

A strategy called pattern-oriented modeling (POM) attempts to make bottom-up modeling more rigorous and comprehensive (6, 8–10). In POM, we explicitly follow the basic research program of science: the explanation of observed patterns (11). Patterns are defining characteristics of a system and often, therefore, indicators of essential underlying processes and structures. Patterns contain information on the internal organization of a system, but in a “coded” form. The purpose of POM is to “decode” this information (10).

The motivation for POM is that, for complex systems, a single pattern observed at a specific scale and hierarchical level is not sufficient to reduce uncertainty in model structure and parameters. This has long been known in science. For example, Chargaff’s rule of DNA base pairing was not sufficient to decode the structure of DNA—until combined with patterns from x-ray diffraction of DNA and from the tautomeric properties of the purine and pyrimidine bases (12). Thus, in POM, multiple patterns observed in real systems at different hierarchical levels and scales are used systematically to optimize model complexity and to reduce uncertainty.

POM was formulated in ecology, a science with a long tradition of bottom-up modeling. Ecology, in the past 30 years, has produced as many individual-based models as all other disciplines together have produced agent-based models (13), and has focused more on bottom-up models that address real systems and problems (14).

We describe here how observed patterns can be used to optimize model structure, test and contrast theories for agent behavior, and reduce parameter uncertainty. Finally, we discuss POM as a unifying framework for the science of agent-based complex systems in general.

Patterns for Model Structure: The Medawar Zone

Finding the optimal level of resolution in a bottom-up model’s structure is a fundamental problem. If a model is too simple, it neglects essential mechanisms of the real system, limiting its potential to provide understanding and testable predictions regarding the problem it addresses. If a model is too complex, its analysis will be cumbersome and likely to get bogged down in detail. We need a way to find an optimal zone of model complexity, the “Medawar zone” (Fig. 1).

Modeling has to start with specific questions (15). From these questions, we first formulate a conceptual model that helps us decide which elements and processes of the real system to include or ignore. With complex systems, however, the question addressed by the model is not sufficient to locate the Medawar zone because ACSs include too many degrees of freedom. Moreover, the conceptual model may too much reflect our perspective as external observers, with our specific interests, beliefs, and scales of perception.

A key idea of POM is to use multiple patterns observed in real systems to guide design of model structure. Using observed patterns for model design directly ties the model’s structure to the internal organization of the real system. We do so by asking: What observed patterns seem to characterize the system and its dynamics, and what variables and processes must be in the model so that these patterns could, in principle, emerge? For example, if there are patterns in age structure, sex ratio, and spatial distribution, then age, sex, and space should be represented in the...
model; if we know that agents behave differently at high densities (e.g., are more aggressive), behavior variability should be in the model. This use of patterns might force us to include state variables and processes that are only indirectly linked to the ultimate purpose of the model and are not part of our initial conceptual model. Ideally, the patterns used to design a model occur at different spatial and temporal scales and different hierarchical levels, because the key to understanding complex systems often lies in understanding how processes on different scales and hierarchical levels are bound to each other.

Multiple patterns were key to modeling spatiotemporal dynamics of the beech forests of central Europe (Fig. 2). Natural beech forests are characterized by a spatial mosaic pattern of successional stages. A cellular automaton model that focused on this pattern only (16) was too poor in structure to reveal the forest's internal organization. But the forests have more characteristic patterns. Different successional stages have different patterns of vertical structure: e.g., the climax stage has closed canopy and little understory, and the decaying stage has canopy gaps and an understory of young beech. Therefore, a newer model (17, 18) includes four height classes (from seedlings to upper canopy) (Fig. 2). The model also explicitly represents individual big trees because canopy gaps are caused by windthrow, an individual-level process. The model’s structure was thus determined by the multiple characteristic patterns: The mosaic pattern determined horizontal spatial scale and resolution, the vertical patterns determined the need for height classes, and canopy gaps determined that large beeches must be described individually.

When designed to reproduce multiple patterns, models are more likely to be “structurally realistic” (10). In particular, model components (e.g., individuals) correspond directly to observed objects and variables, and processes correspond to the internal organization of the real system, so that the model “not only reproduces the observed real system behavior, but truly reflects the way in which the real system operates to produce this behavior” (10, p. 5).

Structurally realistic models can make independent and testable secondary predictions. The beech forest model, for example, delivered independent predictions of forest characteristics that were not considered during model development and testing (20). Predictions of age structure in the canopy and the spatial distribution of very old “giant” trees were in good agreement with observations, considerably increasing the model’s credibility and justifying a completely new application: tracking woody debris (21). Complexity in pattern-oriented bottom-up models is not simply a burden but can provide rich opportunities to increase model credibility, gain understanding (18), and address more questions.

In an example from ecological epidemiology, multiple patterns guided the stepwise design and calibration of a model describing the spread of rabies among red foxes in central Europe (22). Observed patterns included the large-scale wave of rabies prevalence, disease pockets ahead of the wave, and temporal oscillations of prevalence at local and regional scales. The resulting model reproduced these patterns, but not by simply applying a preconceived model structure and then fitting it to the patterns; instead, one pattern after another was used to gradually refine model structure (23). Structural realism of this model is indicated by the striking match between model predictions and a long-term data set of hunted foxes, which combines aspects of rabies epidemiology (before the onset of foxes control), fox ecology (after control), and their interaction (during control).

In other ACS disciplines, we found only a few models explicitly addressing multiple patterns, although many models were implicitly based on multiple patterns. A model of consumer markets (24) addresses three patterns: (i) The statistical distribution of weekly sales of fast-moving consumer goods has fatter tails and thinner peaks than normal distributions; (ii) there are clusters of high sales volatility; and (iii) market shares of different stores follow power-law distributions. Exactly how these patterns influenced the design of the model is not clear, but pattern (iii) appears to be why the model is spatially explicit: Consumer agents only visit stores that are nearby.

Patterns for Contrasting Alternative Theories

Agents continuously make decisions to reach their goals—e.g., survival and reproductive success, profiting in a stock market, finding their best place to settle—in an ever-changing environment. How do we model these decisions? What information do agents have, what alternatives do they consider, and how do they predict the consequences of their decisions? Many studies of ACSs try only one model of decision-making and attempt to show that it leads to results compatible with a limited data set. This practice, however, may lead to the impression that bottom-up models include so many parameters that they can be fitted to data whether or not their structure and processes are valid.

A more rigorous strategy for modeling agent decisions, or other bottom-up processes, is to use “strong inference” (25) by contrasting alternative decision models, or “theories” (3, 6). First, alternative theories of the agent’s decisions are formulated. Next, characteristic patterns at both the individual and higher levels are identified. The alternative theories are then implemented in a bottom-up model and tested by how well they reproduce the patterns. Decision models that fail to reproduce the characteristic patterns are rejected, and additional patterns with more falsifying power can be used to contrast successful alternatives. Rigorous techniques can be used to design experiments and analyze data (6, 26).

As an example, consider the well-known “boids” model (27) that produces schooling-like behavior from a simple theory: Individual boids try to avoid collisions, match the velocity of neighboring individuals, and stay close to neighbors. The emergence of aggregations resembling fish schools from this theory (Fig. 3), however, does not prove that boids explains schooling in real fish.

To define theory for schooling of real fish, Huth (28) used observed patterns and contrasted alternative theories for fish behavior. Two patterns characterizing fish schools were defined and quantified: polarization and nearest neighbor distance (Fig. 3). Eleven alternative theories for how fish adapt swimming speed and direction were formulated. In the first nine theories, the influence of neighbors is averaged; but in two theories, fish adjust their swimming to only one neighbor—e.g., the one
closest in front. These two “priority” theories failed to reproduce realistic polarization values (Fig. 3), eliminating them as valid theory.

This example shows that looking at one pattern may not be sufficient to falsify weak theory: Looking at nearest neighbor distance alone suggested that both types of schooling model produce similar results, but in fact the priority theories produce schools only as compact, but not as polarized, as real schools. Moreover, the nine theories based on averaging differed widely in assumptions, but the fish school’s properties turned out to be robust to these assumptions. Demonstrating robustness is also key to a bottom-up model’s credibility, because it indicates that we captured the most important mechanisms. Huth and Wissel’s model also reproduced several additional patterns not considered during model development, providing further support for its structural realism.

This pattern-oriented theory development approach is increasingly used in models of ACS. Railsback and Harvey (9) used a stream trout model to contrast three theories for how individual fish select habitat. Only a new theory that assumes that fish select habitat to maximize expected survival over a future period reproduced observed patterns of feeding hierarchy, response to competing species and predatory fish, seasonal habitat shifts, and response to reduced food availability. Although these patterns are each qualitative, or “weak,” together they were able to falsify all but one theory of habitat selection.

In a model exploring what determines the access of nomadic herdsmen to pasture lands owned by village farmers in north Cameroon, herdsmen negotiate with farmers for access to pastures (29). Two theories of the herdsmen’s reasoning were contrasted: (i) “cost priority,” in which herdsmen only consider one dimension of their relationship to farmers—costs; and (ii) “friend priority,” in which herdsmen remember the number of agreements and refusals they received in previous negotiations. Real herdsmen sustain a social network across many villages through repeated interactions, a pattern reproduced only by the “friend priority” theory.

In economics, agent-based model experiments have been used to identify characteristics of artificial stock market investors that reproduce patterns well known from real stock markets (30). These patterns include continual and unpredictable stock price volatility, high skew and kurtosis in the distribution of profits among investors, and an inverse relation between current investment profits and future price instability. Two assumptions were contrasted about how much historic data investors use to predict the outcome of their investment decisions: (i) Investors all use 25-year memories of market data, versus (ii) memory varies from 0.5 to 25 years. Although none of the simulations reproduced all the observed market patterns, the assumption that all investors use 25-year memories failed to reproduce the most basic pattern: price volatility. This pattern-oriented analysis indicates that individual variation in investment decision-making is crucial to stock market dynamics.

Testing and contrasting alternative theories or decision models has several benefits. We are forced to be explicit about how decision models are formulated and tested; we can demonstrate how important the specific formulation of a decision—or any other low-level—model is; we can explore null models; and we can continuously refine models by applying additional patterns.

**Patterns for Parameters: Coping with Uncertainty**

Pattern-oriented modeling can reduce uncertainty in model parameters in two ways. First, it helps make models structurally realistic, which usually makes them less sensitive to parameter uncertainty (31). For example, an individual-based coyote population model reproduced an array of observed patterns with no fine-tuning of parameter values taken from the literature (32). The trout model (9) had four parameters that were particularly uncertain but important: each had relatively independent effects on four different outputs (size versus abundance, for juveniles versus adults), so they could be calibrated manually and independently.

Second, the realism of structure and mechanism of pattern-oriented models helps parameters interact in ways similar to interactions of real mechanisms. It is therefore possible to fit all calibration parameters by finding values that reproduce multiple patterns simultaneously. This technique is known as “inverse modeling” (33). For a spatially explicit individual-based model of brown bear dispersal from Slovenia into the Alps (34), a global sensitivity analysis of the uncalibrated parameter set revealed high uncertainty in
model output. To reduce this uncertainty, two data sets were used to identify five patterns. Quantitative criteria for the agreement between observed and simulated patterns were developed. The indirect modeling analysis started with 557 random parameter sets covering the plausible ranges of all parameters. The five observed patterns were used as filters: Only 10 of the 557 parameter sets reproduced all of them. This parameter filtering reduced the model’s global sensitivity by a factor of 4 (fig. S1).

Indirect parameterization is routine in physical process models (i.e., in chemistry, hydrology, and climate modeling), but rare so far in models of ACSs. An encouraging exception is the agent-based model of an ancient society, the Kayenta Anasazi, who occupied the Long House Valley in northeastern Arizona (United States) until 1300 A.D. Paleoenvironmental and archaeological records permitted the development of a detailed, spatially explicit agent-based model of this society and its history (35). These data include estimates of annual potential maize production for each hectare in the study area for the period 400 to 1400 A.D. and records of human settlement in the valley. Theories for agent decisions, for example, splitting households and moving, were based on detailed regional ethnographies.

The model includes variability in mortality, fertility, splitting of households, and maize harvest rates; with eight unknown parameters. To evaluate these parameters indirectly, the time series of the number of simulated households was compared to the historical record. The best parameter set reproduced all important trends and population sizes in the archaeological record. This parameter set also reproduced important features of the spatial distribution of the settlements (Fig. 4) and the gradual northward movement of the population. These spatial patterns can be considered independent predictions, strong indicators of the model’s structural realism.

Implications and Future Directions
Patterns are widely used by many modelers, particularly in disciplines where the low-level entities are physical objects such as atoms and stars, or are relatively easy to represent, such as flocking birds, pedestrians in a panicking crowd, or car drivers (“Brownian agents” (36); see also table S1). However, POM is the first attempt to explicitly formulate a rigorous and comprehensive strategy for modeling ACSs. The POM strategy is a...
way to focus on the most essential information about a complex system’s internal organization. Multiple patterns keep us from building models that are too simple in structure and mechanism, or too complex and uncertain. Using patterns to test and contrast alternative theories for agent behavior or other low-level processes is a way for the science of ACS to get beyond clever demonstration models and on to rigorous explanations of how real systems are organized and how they respond to internal and external forces. POM is just taking root, and we expect to see its rapid development in the near future.

Bottom-up models are virtual laboratories where controlled experiments distinguish noise from signal in the system’s organization. In particular, experiments contrasting noise from signal in the system’s organization indicates that more than 90% of them are "individual-based" records and 1036 "agent-based" records. A review of the 1023 individual-based patterns indicates that more than 90% of them are ecological (38).

Examples include individual-based models for forest management (39) and population viability analysis (40); see also the POM references and the ecological bottom-up reviews in the Supporting Online Material.

References and Notes

1. “Agent-based complex systems” (ACSs) are often referred to as “complex adaptive systems,” but we prefer ACS because it is not clear whether systems are really adaptive, which would require that they have a goal.


4. See the list of recent reviews of bottom-up models provided in the Supporting Online Material.


11. Patterns are observations of any kind showing nonrandom structure and therefore containing information on the mechanisms from which they emerge. Complex systems contain patterns at different hierarchical levels and scales. Ecosystems, for example, contain patterns in primary production, species diversity, spatial structure, dynamics of component species populations, behavior of individual organisms, resource dynamics, and response of all these to disturbance events and stress. Useful patterns need not be striking; qualitative or “weak” patterns can be powerful in combination. For example, we can easily identify a person in a crowd even without a strong pattern (e.g., a photograph) by using a set of weak patterns: sex, approximate age, hair color, size, etc. Each of these characteristic patterns excludes many individuals.


13. A search for "(individual-based or individual-oriented) and model" and for "(agent-based or agent-oriented) and model" in the Web of Science database (title, abstract, keywords) yielded 1023 "individual-based" records and 1036 "agent-based" records. A review of the 1023 individual-based records indicates that more than 10% of them are ecological (38).

14. Examples include individual-based models for forest management (39) and population viability analysis (40); see also the POM references and the ecological bottom-up reviews in the Supporting Online Material.


44. See www.rug.nl/biologie/onderzoek/onderzoekgroepen/biotheoretics/.


46. We thank L. Tesfatsion, P. Johnson, B. LeBaron, and T. Kohler for hints to relevant literature; S. Winter, M. Flade, and C. Rademaker for providing photos; H. Hildenbrandt for Fig. 3A; and three anonymous referees for valuable comments.

Supporting Online Material

www.sciencemag.org/cgi/content/full/310/5750/987/DC1
SOM Text
Fig. S1
Table S1
References and Notes

10.1126/science.1116681