**Natural Resources and Environmental Issues**

Volume 8 *SwarmFest 2000* Article 8

1-1-2001

# Getting "results": The pattern-oriented approach to analyzing complex systems with agent-based models

Steven F. Railsback
*Lang, Railsback and Associates, Arcata, CA*

# GETTING "RESULTS": THE PATTERN-ORIENTED APPROACH TO ANALYZING COMPLEX SYSTEMS WITH AGENT-BASED MODELS

STEVEN F. RAILSBACK

*Lang, Railsback & Associates, 250 California Avenue, Arcata, California 95521, Email: LRA@Northcoast.com*

## INTRODUCTION

Many Swarm users have had the experience of demonstrating an agent-based model (ABM) to colleagues and immediately being asked "but how do you get 'results'? — it's just a stochastic simulator." This response to ABMs is surprisingly common, especially from modelers that use conventional differential equation-based approaches. This question may at first seem presumptuous, but it deserves our attention for two reasons. First, of course, we need to respond to the question intelligently each time we hear it again. Second, we do indeed need ways to develop general results and conclusions from ABMs. This paper presents a general approach to using ABMs and similar simulators in a hypothesis-testing framework to draw general conclusions about the agents in our models.

Whereas the rest of this paper addresses ABMs, it is worth reminding Swarm users that this "how do you get results" question is just as easily applied to differential equations models. Ecology, economics, and other sciences are rife with differential equation models that include numerous restrictive and unrealistic assumptions so the equations can be solved. Some users of such models appear to assume that because their models represent aggregations of individuals, the model results are applicable to individuals in general. The opposite is often true– a hard look at all the assumptions in such models and the conditions under which these assumptions really apply to may indicate that the differential equation model's results are rarely if ever really applicable. My colleagues and I (Railsback et al. 1999) illustrated an example in ecology, the "result" of a mathematical derivation showing the fitness of animals to be maximized when they select habitat that minimizes the ratio of mortality risk over energy intake. Upon examination, this "result" (which, ironically, has been used in a number of agent-based models) was found to be based on a number of assumptions that are rarely if ever met in nature. (The most remarkable of these assumptions is that no habitat fails to provide positive growth rates; in reality, animals are continuously faced with the challenge of finding adequate food.)

Complex systems analysis and ABMs are intended to overcome the limitations of conventional modeling approaches, especially their dependence on restrictive and unrealistic assumptions. To do so, we need ways to test hypotheses and reach general conclusions.

## THE PATTERN-ORIENTED APPROACH

### *Microanalysis Of Complex Systems*

Getting "results" from ABMs is essentially the fundamental issue in analysis of all complex systems: determining the relations between the characteristics of individuals and system-level responses. The only thorough, explicit discussion of this issue I am familiar with is by Auyang (1998). Auyang rejects both the reductionist and holist approaches to science– examining only the parts or only a whole system is inadequate. Instead, Auyang recommends an approach called *synthetic microanalysis* because it integrates synthesis at the system level with analysis of the individuals that compose the system.

Synthetic microanalysis begins with synthesis of a broad conceptual framework of the system. The boundaries of the system to be studied are delineated along with the external driving variables by which the rest of the world affects the otherwise-isolated study system. System-level measures of state are defined; these can be aggregate measures of the state of the system's individuals (e.g., total number or biomass of animals) or statistical distributions of individual states (e.g., distributions of species, age, or size). Synthesis also includes defining the individuals and their characteristics, using information gathered from smaller-scale analyses like laboratory studies of individual animals. For agent-based modelers, this synthesis phase constitutes designing the ABM; it is highly system-specific and I do not address it further in this paper.

In the microanalysis phase, system-level concepts are posed and the individual behaviors and mechanisms ("traits") that explain the system-level concepts are determined. In our research, this phase includes determining how, in ABMs, to model individual agents and their environment so that realistic population-level response patterns emerge from individual traits. This phase provides the mechanistic understanding of the links between individual traits and system responses that allows prediction of how system dynamics respond to factors acting at the individual level.

### *Pattern-Oriented Microanalysis*

My colleagues and I (Railsback and Harvey, in review) developed an approach to microanalysis that allows us to test and possibly reject models of agent traits, and therefore infer what characteristics of agents can produce realistic emergent responses at the system level. The idea for this approach came from the "pattern-oriented" approach to ecological modeling recommended by Grimm et al. (1996). A similar approach was proposed by Clark and Mangel (2000) for testing models of decision making by individual animals.

Grimm et al. (1996) suggest that ecological models should be designed to address specific patterns observed in nature. Modelers should identify important observed patterns and attempt to understand and model the mechanisms that cause the patterns. This is a way to focus a modeling study, providing guidance on such issues as how aggregated the model should be and what scales are relevant. To us, the key benefit to pattern-oriented modeling is that it assures that the model produces predictions suitable for testing: models can be tested by whether they reproduce the patterns the model was designed to address.

Pattern-oriented analysis can be applied to complex systems by using it as a way of identifying and testing hypotheses about how individual agents behave. The following steps can be used; these constitute the microanalysis phase defined by Auyang (1998).

(1) Define a set of "testing patterns", observed patterns of system-level (or individual) responses to known stimuli that the ABM is designed to explain and reproduce. These patterns may be identified before the model is designed as a way to guide model development; or, if the model has already been designed, the patterns are identified before model testing begins so the modeler cannot be accused of picking only those testing patterns that confirm the model's formulation.

Useful patterns are often easily extracted from existing literature on the system being studied. Picking test patterns driven by a wide range of stimuli allows microanalysis to be more comprehensive and conclusive, but the model must include the mechanisms driving the patterns.

(2) Build a model that includes the mechanisms and agent traits believed to drive the testing patterns. The ABM must be designed to reproduce the test patterns as an emergent result of simpler agent behaviors. Modelers *must be sure that the test patterns are emergent responses of the model and are not hardwired in!*

(3) Pose alternative rules for agent behavior as hypotheses that will be tested. It can be most interesting to pose as alternative agent behaviors (a) maximizing the kind of simple, direct measures of fitness that are possible in ABMs vs. (b) the kinds of derived surrogates for fitness often found in conventional models.

(4) With the ABM, simulate the conditions under which each test pattern has been observed to occur. Repeat the simulations with each

alternative agent behavior rule. Reject rules that do not cause the test patterns to emerge from the model, and accept rules that do.

Completion of these steps allows the modeler to at least reject model rules that fail to produce emergence of the test patterns, and, perhaps, identify rules that succeed in producing realistic emergent system-level behavior. These results are likely to allow very meaningful and general inferences to be drawn.

### *An Unattractive Alternative: Testing Against Response Magnitudes*

A key advantage of the pattern-oriented approach is that it allows testing of a model against *patterns* of observed system responses instead of against the *magnitude* of observed responses. In other words, it allows us to test a model without having to go through the expensive, tedious, and often impossible steps necessary to attempt simulation of the exact conditions under which the test behavior was observed, and allows us to avoid the statistical pitfalls of comparing model results to observations. The following are reasons why the alternative to pattern-oriented analysis of testing ABMs by direct comparison to the magnitude of observed system responses is unattractive.

The first reason is the well-known "butterfly effect". Even models that accurately capture the mechanisms driving a complex agent-based system may fail to reproduce observed responses if the system is highly sensitive to initial conditions.

Secondly, it is often very difficult to simulate a system precisely enough to expect response magnitudes to be testable. It can be relatively simple and easy to modify an ABM to simulate the conditions under which a well-understood response pattern should emerge. For example, if it widely observed in business that people spend more money on video games in winter than summer, then a model to explain this pattern need only include the mechanisms driving spending on video games and be run simulating winter vs. summer conditions. On the other hand, testing a model against response magnitudes requires (1) reproducing in the model the magnitude of all the forces driving the response and (2) correct estimation of all the parameters linking the driving forces to the agent response. Such quantitative information is rarely available with sufficient completeness and accuracy. As opposed to testing the magnitude of responses, the pattern-oriented approach allows us to test an ABM even when input data and parameter values are relatively uncertain. The focus remains on getting the mechanisms right instead of on unnecessary detail and calibration.

Third, attempting to reproduce observed response magnitudes generally requires calibration of the model, and often the only data available for calibration are the same data used to test the model. Rarely are all of a model's parameters known well enough to test the model without calibration; usually, calibration is used to estimate values for important but uncertain parameters by comparing model results to observed data. Data used for calibration are no longer useful for testing the model, so typically the data available for comparison to model predictions are split between those used for calibration and those used for testing. Splitting the data in this way reduces the power of both calibration and testing. Because they do not depend on accurately predicting the magnitude of observed responses, pattern-oriented tests can be useful with little or no calibration, preserving more of the data for model testing.

Fourth, identifying a comprehensive range of independent tests is much more likely with the pattern-oriented approach. A wide range of system or individual response patterns are often readily obtained from the literature, along with sufficient understanding of the stimuli causing the pattern to allow the pattern to be reproduced in an ABM. Given the difficulty and expense of testing an ABM's ability to reproduce the magnitude of system responses, it is often unlikely that a diverse and comprehensive set of such tests could be performed.

The pattern-oriented approach also avoids the temptation to compare model results to observed responses statistically. Conventional hypothesis testing statistics are often unhelpful for testing ABMs. First, statistical comparisons are inappropriate between data sets where uncertainty results from different sources. Results from replicate ABM simulations have variability due only to the model's stochastic components, which is not comparable to the variability typically observed in biological or human systems, which also results from external driving forces and initial conditions. Statistical comparisons (e.g., t-tests) are useful for comparisons among ABM results, but rarely valid for comparing ABM results to observed data. Secondly, ABMs can produce output with arbitrary and high sample sizes that exaggerate statistical significance. We can easily observe each of a model's thousands of agents, so sample sizes are very large. Virtually any difference becomes statistically significant with such sample sizes, even differences that are clearly not otherwise meaningful. The pattern-oriented approach avoids these problems.

## EXAMPLE: HABITAT SELECTION BY STREAM TROUT

We applied the pattern-oriented approach to test model rules for how fish select habitat among patches varying in probability of surviving mortality (predation, etc.) and food intake (Railsback and Harvey, in review). This experiment is summarized here as an example pattern-oriented model analysis.

The primary individual behavior simulated in our trout model is movement to select new habitat, in reaction to changes in river flow, temperature, and trout density. These factors affect the spatially explicit food availability (which is affected by hydraulics and competition for food with other trout) and mortality risks. Habitat selection is simulated at a daily time step and spatial scale of a stream reach (~200 m length) and patch size of several $m^2$. Habitat selection is simulated by letting each fish move to the habitat patch providing the highest fitness, where fitness is defined as the expected probability of surviving and growing to reproductive size over a future time horizon. The probability of surviving over the time horizon is a function of food intake as well as predation risks– if a fish does not obtain sufficient food, it risks starving to death within the time horizon. (We found this approach to modeling habitat selection to have numerous advantages over methods used in previous IBMs; Railsback et al. 1999.)

The trout model's ability to simulate realistic emergent habitat selection was analyzed by testing six patterns of habitat shifts observed in real trout. From the fisheries literature, we identified these patterns:

- Where food availability is highly variable over space, fish exhibit "hierarchical" feeding behavior. The dominant (largest) fish gets the best feeding site, and if the dominant fish is removed the next-dominant fish moves into the best site and other fish move up in the hierarchy.
- Adult trout respond to flood flows simply by moving to the stream margins where velocities are low.
- Juvenile trout use higher velocities on average when in competition with another, larger, trout species.
- Juvenile trout use faster and shallower habitat on average in the presence of predatory fish.
- Trout user higher velocities on average when temperatures are higher (metabolic rates increase with temperature, so more food is needed to avoid starvation).
- When food availability is decreased, trout shift to habitat with higher food intake and higher mortality risks.

We simulated the conditions under which these response patterns have been observed. We posed three alternative rules as hypotheses for how trout select habitat. In addition to our approach (maximizing predicted survival and growth to reproductive size over an upcoming time horizon), we tested (1) maximizing the current growth rate and (2) maximizing the current survival probability (which is *not* a function of food intake because fish do not starve to death in one day).

Simulations with movement to maximize current growth reproduced three of the six habitat selection patterns, maximizing survival reproduced two patterns, and our "maximize predicted survival and growth" approach reproduced all six patterns. Two patterns (shifts in habitat with temperature and food availability) were not reproduced by models that consider only current growth and risk, but were explained by our objective that simulates prediction of how future starvation risk depends on current energy reserves and food intake. Ecological models of habitat selection have almost all been based on current growth or risk, and our experiment provides strong evidence that these conventional models cannot explain important observed behaviors. Our experiment provides some of the first evidence that realistic models of animal behavior need to consider how animals make predictive decisions. (This result is no surprise to those familiar with complex systems research– predictive ability is key to the emergence of complex, lifelike behavior from adaptive agents; e.g., Holland 1995.)

This experiment illustrates that we can use ABMs and pattern-oriented analysis to identify individual traits that produce realistic emergent population-level responses. (Animations of some of the experimental simulations are at: http://math.humboldt.edu/~simsys/EcolArch/.)

## CONCLUSIONS

Meaningful results can be obtained from agent-based simulations when the simulations are designed to test model assumptions against observed patterns of system-level response. Comparing alternative models of agent traits by their ability to cause emergence of realistic system-level behavior patterns can be a powerful way to infer the mechanisms by which agents respond to each other and their environment. Testing models against patterns of response instead of against the magnitude of observed responses allows model mechanisms to be tested comprehensively with a reasonable level of effort and expense.

## ACKNOWLEDGMENTS

This work was supported by EPRI, the Electric Power Research Institute, under agreement WO6953-02 with Lang, Railsback & Assoc., and by the USDA Forest Service, Redwood Sciences Laboratory, under research joint venture agreement no. PSW-99-007-RJVA with Lang, Railsback & Assoc. Software support was provided by Steve Jackson, Jackson Scientific Computing, McKinleyville, CA; and by Glen Ropella and Chris Langton, The Swarm Corporation, Santa Fe, New Mexico. Collaborators in developing the trout model and testing scenarios were Walt Duffy (California Cooperative Fisheries Research Unit), Roland Lamberson (Department of Mathematics, Humboldt State University), and especially Bret Harvey (Redwood Sciences Laboratory).

## LITERATURE CITED

Auyang, S. Y. 1998. Foundations of complex-system theories in economics, evolutionary biology, and statistical physics. Cambridge University Press, New York, USA.

Clark, C. W., and M. Mangel. 2000. Dynamic state variable models in ecology. Oxford University Press, New York, USA.

Grimm, V., K. Frank, F. Jeltsch, R. Brandl, J. Uchmanski, and C. Wissel. 1996. Pattern-oriented modelling in population ecology. The Science of the Total Environment **183**:151-166.

Holland, J. H. 1995. Hidden order: how adaptation builds complexity. Perseus Books, Reading, Massachusetts, USA.

Railsback, S. F., R. H. Lamberson, B. C. Harvey, and W. E. Duffy. 1999. Movement rules for spatially explicit individual-based models of stream fish. Ecological Modelling **123**:73-89.

Railsback, S. F., and B. C. Harvey. In review. Analysis of salmonid habitat selection rules using an individual-based model. In review at Ecology.