Computational Ideation in Scientific Discovery: Interactive Construction, Evaluation and Revision of Conceptual Models

Ashok K. Goel & David A. Joyner
Design & Intelligence Laboratory, School of Interactive Computing
Georgia Institute of Technology
85 Fifth Street NW, Atlanta, GA 30338
ashok.goel@cc.gatech.edu; david.joyner@gatech.edu

Abstract

We present several epistemic views of ideation in scientific discovery that we have investigated: conceptual classification, abductive explanation, conceptual modeling, analogical reasoning, and visual reasoning. We then describe an experiment in computational ideation through model construction, evaluation and revision. We describe an interactive tool called MILA–S that enables construction of conceptual models of ecological phenomena, agent-based simulations of the conceptual model, and revision of the conceptual model based on the results of the simulation. *The key feature of MILA–S is that it automatically generates the simulations from the conceptual model*. We report on a pilot study with 50 middle school science students who used MILA–S to discover causal explanations for an ecological phenomenon. Initial results from the study indicate that use of MILA–S had a significant impact both on the process of model construction and the nature of the constructed models. We posit that MILA–S may enable scientists to construct, evaluate and revise conceptual models of ecological phenomena.

Background, Motivations, and Goals

We may adopt several epistemic views of ideation in scientific discovery. We have developed computational techniques and tools for supporting some of these epistemic views. Here we briefly present these epistemic views as background and motivation for the present work.

Conceptual Classification

Classification of data into concepts is ubiquitous in science. We all know about Linneas' classic work on classification in biology. Classification continues to be important in modern biology (e.g., Golub et al. 1999). Classification also is one of the most studied topics in cognitive science, artificial intelligence and machine learning (e.g., Langley 1996; Mitchell 1997; Stefik 1995; Thagard 2005; Winston 1993). The classic DENDRAL system (Lindsay et al. 1980) classified mass spectroscopy data into chemical molecules that produced the data. Chandrasekaran & Goel (1988) trace the evolution of early knowledge-based theories of classification.

We have studied both top-down hierarchical classification in which a concept is incrementally refined based on data (Goel, Soundarajan & Chandrasekaran 1987), and bottom-up hierarchical classification in which features of data are incrementally abstracted into a concept (Bylander, Goel & Johnson 1991). In recent work, we have developed a computational technique that grounds the concepts in bottom-up classification in perception and uses meta-knowledge of this perceptual grounding for repairing the semantics of the concepts when the classification results in an error (Jones & Goel 2012). The Augur system uses this meta-reasoning technique for revising almost-correct classification hierarchies in a variety of domains.

Abductive Explanation

Scientific theory formation often entails abductive inference (Magnini 2001), i.e., inference to the best explanation for a set of data. Artificial intelligence research has studied abduction from multiple perspectives (e.g., Charniak & McDermott 1985; Josephson & Josephson 1996). The classic BACON system (Langley et al. 1987) abduced physical laws, such as the gas law, from data. Bylander et al. (1991) have analyzed the computational complexity of the abduction task.

We have studied computational techniques for abductive explanation that assemble composite explanations for explaining a set of data from elementary explanations that explain subsets of the data (Goel et al. 1995). The RED system uses this technique for identifying red-cell antibodies in a patient's blood serum (Fischer et al 1991). RED uses domain-independent heuristics such as the *essentialness* heuristic for assembling a composite explanation from elementary explanations: this heuristic says that if some data item can be explained by only one elementary hypothesis, then the hypothesis should be directly included in the composite explanation.

Conceptual Modeling

Conceptual models too are ubiquitous in science (e.g., Clement 2008; Darden 1998, 2006; Nersessian 2008).

Conceptual models are abstract representations of the elements, relationships, and processes of a complex phenomenon or system. Research in both cognitive science and artificial intelligence has extensively studied conceptual modeling in several forms (e.g., Davis 1990; Johnson-Laird 1983; Lenat 1995; Novak 2010; Schank & Abelson 1977; Stefik 1995; Winston 1993).

We have developed conceptual models of complex systems that specify the way a system works, i.e., the way the system's structure produces its behaviors that achieve its functions (Goel & Stroulia 1996; Goel, Rugaber & Vattam 2009). We have used SBF modeling to model both engineering systems (Goel & Bhatta 2004) and natural systems (Goel et al. 2012) for supporting a variety of reasoning processes in design and invention as well as in science education (Goel et al. 2013; Vattam et al. 2011).

Analogical Reasoning

Scientific discovery often entails analogical reasoning (Clement 2008; Nersessian 2008). We all know about Bohr's famous analogy between the atomic structure and the solar system. Modern scientists too frequently use analogies in their everyday practices (Clement 1988; Dunbar 1997). Cognitive science and artificial intelligence research have developed several theories of analogical reasoning (e.g., Bhatta & Goel 2004; Hofstader 1996; Holyoak & Thagard 1996; Gentner & Markman 1997; Indurkhya 1992; Keane 1996; Mitchell 1993; Prade & Gilles 2014; Thagard et al. 1990). AI research on analogy however is not yet as mature as that on, say, classification.

We have studied analogical reasoning in scientific problem solving (Griffith, Nersessian & Goel 1996, 2000). Starting from verbal protocols of physicists addressing problems with spring systems (Clement 1988), we developed an AI system called Torque that emulates some of the problem solving behavior of the physicists. Torque addresses a key question about spring systems by making an analogy to bending beams. A critical ability of Torque is *problem transformation:* Torque imagines various transformations of the spring system so that a specific transformation (a spring stretched so much that it become linear) reminds it until it is reminded of the beam (Griffith, Nersessian & Goel 2000).

Visual Reasoning

Scientific discovery often engages visual representations and reasoning (Clement 2008; Magnini, Nersessian & Thagard 1999; Nersessian 2008). Although cognitive science and artificial intelligence research has explored visual representations and reasoning (e.g., Glasgow & Papadias 1992; Glasgow, Narayanan & Chandrasekaran 1995), AI research on visual representations and reasoning is not as robust or mature as on conceptual representations and reasoning. We have developed a language for representing visual knowledge, a computational technique for reasoning about visual analogies (Davies, Goel & Yaner 2008), and analysis of the use of visual analogy in Maxwell's construction of the unified theory of electromagnetism (Davies, Nersessian & Goel 2005).

Goals

In general, the AI methods of classification, abduction, analogy, conceptual modeling, and visual reasoning provide few guarantees of correctness of their results. Further, these methods do not by themselves evaluate their results. This raises an important question for developing AI techniques and tools for supporting the above epistemic views of scientific discovery: how may an AI method for scientific discovery evaluate its results or enable a human to evaluate the results? As an example, let us suppose a human scientist uses an interactive system to develop a conceptual model of a phenomenon; how might the system evaluate the model and provide useful feedback to the scientist?

In this paper, we describe an experiment in computational ideation through model construction, evaluation and revision. We describe an interactive tool called MILA–S that enables rapid construction of conceptual models of ecological phenomena, agent-based simulations of the conceptual model, and revision of the conceptual model based on the results of the simulation. *The key feature of MILA–S is that it automatically generates the simulations from the conceptual model.* We report on a pilot study with 50 middle school science students who used MILA–S to discover causal explanations for an ecological phenomenon.

Model Construction, Evaluation and Revision

Cognitive science theories of scientific discovery describe scientific modeling as an iterative process entailing four related but distinct phases: model construction, use, evaluation, and revision (Clement 2008; Nersessian 2008; Schwarz et al. 2009). Thus, a model is first constructed to explain some observations of a phenomenon. The model is then used to make predictions about other aspects of the phenomenon. The model's predictions next are evaluated against actual observations of the system. Finally, the model is revised based on the evaluations to correct errors and improve the model's explanatory and predictive efficacy. Present in this process are a number of significant challenges to students and scientists alike; constructing initial models may be straightforward, but developing methods for formally evaluating those models and generating usable feedback for subsequent revision is a significant open issue.

Scientific models can be of several different types, with each model type having its own unique affordances and constraints, and fulfilling specific functional roles in scientific inquiry (Carruthers, Stitch & Siegal 2002; Magnini, Nersessian & Thagard 1999). In this work, we are specifically interested in two kinds of models: conceptual models and simulation models. Conceptual models allow scientists to specify and share explanations of how a system works, aided by the semantics and structures of the specific conceptual modeling framework. Conceptual models tend to rely heavily on directly modifiable representations, languages and visualizations, enabling rapid iterations of the model construction cycle.

Simulation models capture relationships between the variables of a system such that as the values of input variables are specified, the simulation model predicts the temporal evolution of the values of other system variables. Thus, the simulation model of a system can be run repeatedly with different values for the input variables, the predicted values of the system variables can be compared with the actual observations of the system, and the simulation model can be revised to account for discrepancies between the predictions and the observations. A main limitation of simulation models is the complexity of the setting up a simulation, which makes it difficult to rapidly iterate on the model construction cycle..

AI research on science education has used both conceptual models (e.g., Bredeweg & Winkels 1998; Jacobson 2008; Novak 2010; vanLehn 2013) and simulation models (e.g., Bridewell et al. 2006; de Jong & van Joolingen 1998; Jackson, Krajcik, & Soloway 2000) extensively and quite productively. However, AI research on science education typically uses the two kinds of models independently from each other: students use one set of tools for constructing, using, and revising conceptual models, and another tool set for constructing and using simulation models. Cognitive science theories of scientific inquiry, however, suggest a symbiotic relationship between conceptual models and simulation models (e.g., Clement 2008; Nersessian 2008): scientists use conceptual models to set up the simulation models, and they run simulation models to test and revise the conceptual models.

We have developed an interactive system called MILA–S that enables construction of conceptual models of ecosystems, use of the conceptual models to automatically generate simulation models of specific ecological phenomena, and the execution of the simulation models. Thus, MILA–S facilitates simulating conceptual models, allowing human users to exploit the symbiotic relationship between conceptual models and simulations without learning to program simulations.

Project Description

MILA–S is an extension of MILA (the **M**odeling & **I**nquiry **L**earning **A**pplication).. MILA is an exploratory learning environment that allows middle school students to investigate and construct models of complex ecological phenomena. While both MILA and MILA–S use both conceptual models and simulation models, only MILA–S allows students to simulate their conceptual models.

MILA and MILA–S build on a line of exploratory learning environments dating back almost ten years, all supporting learning about complex ecological systems through model construction and revision. Its predecessors, the ACT (Vattam et al. 2011; Goel et al. 2013) and EMT (Joyner et al. 2011), were shown to facilitate significant improvement in students' deep, expert-like understanding of complex ecological systems.

For conceptual modeling, ACT used Structure-Behavior-Functions models that were initially developed in AI research on system design (Goel & Stroulia 1996; Goel, Rugaber & Vattam 2009). In contrast, EMT used Component-Mechanism-Phenomenon (or CMP) conceptual models that are variants of Structure-Behavior-Function models adapted for modeling ecological systems. Both ACT and EMT used NetLogo simulations as the simulation models (Wilsensky & Reisman 2006; Wilensky & Resnick 1999). The NetLogo simulation infrastructure is well suited to MILA–S because ecological systems are inherently agent-based. An example of a NetLogo simulation can be seen in Figure 1. Users interact with the simulation using the controls on the level; the simulation can be started, stopped, reinitialized, and the variables present within it can be modified. In the center portion is a spatial depiction of the simulation, presenting the interactions amongst the agents explicitly and visually. The right side provides information on the counts of different components.

Like most interactive tools for supporting modeling in science education (vanLehn 2013), both ACT and EMT provided one set of tools for constructing and revising conceptual models and another tool set for using simulations. Researchers constructed the simulations in ACT and EMT based on expert input; the students only experimented with the simulations. One version of ACT could simulate portions of a conceptual model (Vattam, Goel & Rugaber 2011). However, this integrated version was never actually used in classrooms.

The present intervention had two main parts. In the first part, 10 classes with 237 students in a metro Atlanta middle school used MILA for two weeks. During this time, students worked in small teams of two or three to investigate two phenomena: a recent massive and sudden fish death in a nearby lake and the record high temperatures in the local area over the previous decade. In the second part, two classes with 50 students used MILA–S to more deeply investigating the phenomenon of massive, sudden death of fish in the lake.

Technology Description

As mentioned above, MILA–S uses Component-Mechanism-Phenomenon (or CMP) conceptual models that are variants of the Structure-Behavior-Function models. In order to facilitate generation of simulations, MILA–S augments these models with a set of additional information necessary to create dynamic, interactive simulations, including ranges for variables and visual representations for components, leading to an extension we refer to as CMP*. In CMP models, mechanisms explain phenomena such as fish dying in a lake. Mechanisms arise out of interactions among components and relations among them. The

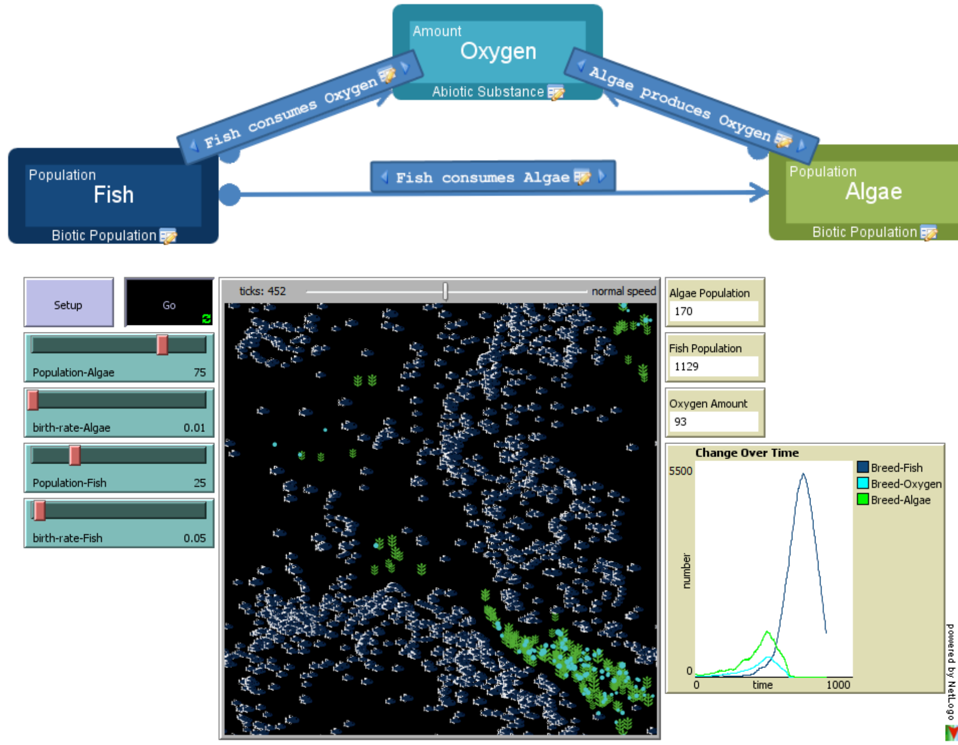

Figure 1: A conceptual model in MILA–S (top) showing relationships between fish, algae, and oxygen, and the simulation model (bottom) generated by MILA–S in NetLogo. This model was constructed by the team described in the third case study below; the simulation was generated and run from their model to obtain this screenshot.

representation of each component in CMP* includes a set of variables such as population, age, birth rate, and energy for biotic components, and quantity for abiotic components.

In the CMP model of a system, representations of components (and their variables) are related together through different kinds of relations. MILA–S provides the modeler with a set of prototype relations. For example, interactions between a biotic component like 'Fish' and an abiotic component like 'Oxygen' could be 'consumes', 'produces', or 'destroys'. Connections have directionality; a connection from 'Oxygen' to 'Fish' would have a different set of prototypes, including 'poisons'. Representations of relations are also annotated with parameters to facilitate the simulation, such as energy provided for 'consumes' and rate of production for 'produces'.

After constructing a CMP conceptual model, a student clicks a 'Run Sim' button to initialize MILA–S and pass their model for simulation generation. MILA–S iterates through some initial boilerplate settings, then gathers together all the components for initialization along with their individual parameters. At this point, each individual component is created according to a prototype for its category (biotic or abiotic); with this prototype comes a set of assumptions generalized to be true about all instances of that category. For biotic organisms, for example, it is assumed that all biotic organisms must eat, breathe, and reproduce. These assumptions do not hold true for all ecological systems, and thus they may limit the systems that MILA–S can address; however, the quick transition between conceptual models and simulation models is facilitated by these assumptions, and they have covered the vast majority of systems that have been examined within the tool.

After creating the components of the conceptual model within the simulation model according to the assumptions and parameters given, MILA–S writes functions governing interaction amongst components based on the relations specified in the CMP model. As before, these functions are created as instances of prototypes of interactions amongst agents, and those prototypes carry certain assumptions. For example, it is assumed that when one organism consumes another, that the target organism dies. Other relationships are assumed, but can be altered; for example, if one organism eats another, it is assumed that the target provides energy to the consumer, but the assumed relationship can be modified to provide instances of the target poisoning the consumer. MILA–S also assumes that species will continue to reproduce to fulfill their carrying capacity rather than hitting arbitrary limitations. These assumptions do limit the range of simulations that MILA–S can generate, but they also facilitate the higher-level rapid model revision process that is the learning objective of this project. Several common ecological relationships are not inferred or assumed by MILA–S, but may be supplied in the model.

This description provides merely a high-level overview of the way in which MILA–S generates simulation models directly from conceptual models; the intent here is to provide sufficient information about the nature of the conceptual and simulation models to understand the educational impacts of interaction with this tool. A much more complete description of the algorithm can be found in Joyner, Goel, & Papin 2014, including a more exhaustive list of assumptions made, a detailed description of the available prototypes, and a thorough account of the instantiation of component and mechanism prototypes.

Experiment Description

Prior to engagement with MILA–S, the 50 students in our study received a two-week curriculum on modeling and inquiry, featuring five days of interaction with CMP conceptual modeling in MILA. In the first part of the study using MILA, students also used pre-programmed NetLogo simulations that did not respond to students' models, but nonetheless provided students experience with the NetLogo interface and toolkit. Thus, when given MILA–S, students already had significant experience with CMP conceptual modeling, NetLogo simulations, and the interface of MILA.

A thorough, if initial, examination of the processes and results of model construction by the student teams provides two insights. Firstly, there exists a fundamental difference in the conceptual models that students constructed with MILA–S compared to the earlier models they constructed with MILA: while earlier models were retrospective and explanatory, models constructed with MILA–S models were prospective and dynamic. Secondly, the model construction process when students were equipped with MILA–S was profoundly different from their earlier process using MILA: whereas previously, conceptual models were used to guide investigation into sources of information such as existing theories or data observations, once equipped with MILA–S the students used the conceptual models to spawn simulations that directly tested the implications of their hypotheses and models thereof.

Model Construction Process

During prior engagement with MILA, we observed students engage in the model construction cycle. Model construction occurred as students constructed their initial hypotheses, typically connecting only a cause to a phenomenon with no mechanism in between. This was then used to guide investigation into other sources of information such as observed data or other theories to look for corroborating observations or similar phenomena. The conceptual model was then evaluated according to how well it matched the findings; in some cases, the findings directly contradicted the model, while in other cases, the findings lent evidence to the model. Finally, the conceptual models were revised in light of this new information, or dismissed in favor of stronger hypotheses.

During engagement with MILA–S, however, we observed a profound variation on the model construction process. The four phases of model construction were still present, but the nature of model use and evaluation changed. Students started by constructing a small number of relationships they believe to be relevant in the system, the model construction phase. After some initial debugging and testing to become familiar with the way in which conceptual models and simulations fit together, students generated simulations and used them to test the implications of their conceptual models. After running the simulation a few times, students then evaluated how well the results of the simulation matched the observations from the phenomenon. This evaluation had two levels: first, did the simulation accurately predict the ultimate phenomenon (in this case, the fish kill)? Once this basic evaluation was met, an advanced evaluation followed: did other variables, trends, and relationships in the simulation match other observations from the phenomenon? For example, one team successfully caused a fish kill by causing the quantity of food available to the fish to drop, but evaluated this as a poor model nonetheless because nothing in the system indicated a disturbance to the fish's food supply. Finally, equipped with the results of this evaluation, students revised their models to more closely approximate the actual system.

Thus, students still constructed and revised conceptual models, but through the simulation generation framework of MILA–S, the model use and evaluation stages took on the practical rigor, repeatable testing, and numeric analysis facilitated by simulations. Rather than speculating on the extent to which their model could explain a phenomenon, students were able to directly test its predictive power. Then, when models were shown to lack the ability to explain the full spectrum of the phenomenon, students were able to quickly return and revise their conceptual models and iterate through the process again.

Three Case Studies in Model Construction

We present three case studies from our experiment to illustrate the above observations about the model construction process. These case studies were chosen to demonstrate variations in the process and connections to the underlying model of construction and revision.

Case 1

The first team speculated that chemicals were responsible for killing the algae in the lake, which then caused the fish population to drop. They began this hypothesis by constructing a model suggesting that algae produces oxygen, fish consume oxygen, and harmful chemicals destroy algae populations. They then used MILA–S to generate and use a simulation of this model to mimic the initial conditions present in the system (i.e. a fish population, an algae population, and an influx of chemicals). This simulation showed the growth of fish population continuing despite the dampened growth of algae population from the harmful chemicals. The team evaluated this to mean that the death of algae alone could not cause the massive fish kill to occur. The team then revised their model to suggest chemicals directly contributed to the fish kill by poisoning the fish directly, as well as killing the algae.

The team then used MILA–S to generate another simulation. This time, when the team used the simulation under similar initial conditions, the fish population initially grew wildly, but the chemicals ate away at both the fish and algae. Eventually, the harmful chemicals finished eating away at the algae, the oxygen quantity plummeted, and the fish suffocated. Students evaluated that this simulation matched the observed phenomenon, but also evaluated that their model missed a relevant relation: based on a source present in the classroom, students posited that fish ought to consume algae. They revised their model, used their simulation again, found the same result, and evaluated that they had provided a model that could explain the fish kill.

Case 2

A second team started off by creating a simple set of relations that they believed was present due to their biology background and prior experience with MILA. First, they speculated that sunlight "produces" oxygen, and, then, that fish, in turn, consume the oxygen. Following these two initial relationships, they generated their first simulation through MILA–S and used it to mimic the believed initial conditions of the lake (i.e. a population of fish, available oxygen, available sunlight). Sunlight was inferred to be continuously available, and thus, at first, the population of fish expanded continuously without any limiting factor. However, when the population of fish hit a certain threshold, it began to consume oxygen faster than it was being produced. This led to the quantity of oxygen dropping, and subsequently, the population of fish dropping. The fish and oxygen populations instead began to fluctuate inversely, with oxygen concentration rebounding when fish population dropped, allowing the fish to recover.

The team ran this simulation multiple times to ensure that this trend repeated itself. In one instance, the fish population crashed on its own simply due to the suddenness of the fish population growth and subsequent crash. However, the team evaluated that this was not an adequate explanation of what had actually happened in the lake. The team posited that if this kind of expansion and crash could happen without outside forces, it would be more common. Second, the team observed that their model contained faulty or questionable claims, such as the notion that sunlight "produces" algae. This evaluation based on both the simulation results and reflection on the model led to a phase of revision. An 'Algae' component was added between sunlight and oxygen, representing photosynthesis. Students then used MILA–S to generate a new simulation, and used this new simulation to test the model. This time students found that their model posited that an oxygen crash would *always* occur in the system, and evaluated that while this successfully mimicked the phenomenon of interest, it failed to match the lake on other days.

Case 3

The third team began with an interesting hypothesis: algae serves as both the food for fish and the oxygen producer for fish. The team, thus, started with a simple three-component model with fish, algae, and oxygen: fish consume algae, fish consume oxygen, and algae produces oxygen. The team further posited that in order for algae populations to grow, they must have sunlight to feed their photosynthesis process. Sunlight, therefore, was drawn to produce algae. The team reasoned that if the fish population destroys the source of one type of 'food' (oxygen) in search for another type (actual food), it could inadvertently destroy its only source for a necessary nutrient.

The team used MILA–S to generate a simulation based on this model and ran it several times under different initial conditions. Each time, algae population initially grew due to the influx of sunlight. As a result, fish populations grew, due to the abundance of both algae (as produced via sunlight) and oxygen (as produced by the algae). As the fish population spiked, the algae hit a critical point where it began to be eaten faster than it reproduced, and the rate of sunlight entering the system was insufficient to maintain steady, strong growth. This caused the algae population to plummet, and in turn, the fish population to plummet as the fish suddenly lacked both food and oxygen. Sometimes, the algae population subsequently bounced back even after the fish fully died off, while in others both species died entirely.

Unlike the second team, this third team evaluated this to mean their model was accurate: under the initial conditions observed in the lake, their model predicted an algal bloom every single time. Thus, the third team provided two interesting variations on the model construction process observed in other teams: first, they overloaded one particularly component, demonstrating an advanced notion of how components can play multiple functional roles. Second, they posited that a successful model would predict that the same events would transpire under the same initial conditions every time.

Summary, Conclusions, and Future Work

Our hypothesis in this work was that affording students with the opportunity to automatically generate simulations from conceptual models and thus exploit the relationship between conceptual and simulation models will allow them to derive the benefits of experimentation and evaluation of the simulation models while maintaining the advantages of rapid construction and revision of the conceptual models, and thereby participate in the modeling process described in the literature of cognition of science. Initial results from this pilot study provide some evidence in favor of the hypothesis, although a controlled study is needed to conclusively verify these claims. Firstly, students approached the modeling process from a different perspective from the outset, striving to capture dynamic relationships among the components of the ecological system. These relationships promoted a more abstract and general perspective on the system. Secondly, the process of model construction, use, evaluation, and revision presented itself naturally during this intervention, with the simulations playing a key role in supporting the cyclical process of constructing conceptual models. By using the simulations to test their predictions and claims, and by subsequently evaluating the success of their conceptual models by matching observations from the actual phenomenon, students engaged in a rapid feedback cycle that saw rapid model revision and repeated use for continued evaluation.

We are presently engaged in a full-scale controlled investigation to test these ideas and tools with college-level biology students. The objective of this investigation is to identify the effect of supporting conceptual modeling and simulation on understanding of individual systems and of the process of scientific investigation as a whole. We expect the new controlled study with college-level students to confirm that our technology for construction, evaluation and revision of conceptual models supports productive ideation in scientific discovery.

Note that in addition to conceptual modeling, this project entails other processes of scientific discovery from the introduction. It engages abductive explanation as students explore multiple hypotheses for explaining an ecological phenomenon and construct the best explanation for the data. It also uses visual representations and reasoning: students construct a visual representation of their conceptual model of the ecological phenomenon (see Figure 1) and generate visualized simulations directly from the conceptual models.

Boden (1990) makes a distinction between H-creativity and P-creativity. H-creativity is historical creativity that has never been done by anyone before; P-creativity is psychological creativity that is personal. Following Boden, we distinguish between H-discovery and P-discovery. Our experiment with MILA–S is an example of P-discovery. We posit by enabling scientists to rapidly and easily construct, simulate, and revise conceptual models of ecological phenomena, MILA–S may also enable H-discovery.

Acknowledgments

We thank Nicolas Papin for his contributions to the development of MILA–S and Spencer Rugaber for discussions about this work. We also thank Vickie Bates and Lara Catall, the teachers in this study.

References

Boden, M. (1990) *The Creative Mind: Myths and Mechanisms.* London: Weidenfeld & Nicholson.

Bredeweg, B., & Winkels, R. (1998) Qualitative Models in Interactive Learning Environments: An Introduction. *Interactive Learning Environments* 5(1): 1-18.

Bridewell, W., Sanchez, J., Langley, P., & Billman, D. (2006) An Interactive Environment for Modeling and Discovery of Scientific Knowledge. *International Journal of Human-Computer Studies* 64(11): 1099-1114.

Bylander, T., Allemang, D., Tanner, M., & Josephson, J. (1991) The Computational Complexity of Abduction. *Artificial Intelligence* 49: 25-60.

Bylander, T., Johnson, T., & Goel, A. Structured Matching: A Task-Specific Technique for Making Decisions. *Knowledge Acquisition* 3(1):1-20, 1991.

Carruthers, P., Stitch, S., & Siegal, M. (editors, 2002) *The Cognitive Basis of Science*, Cambridge University Press.

Chandrasekaran, B., & Goel, A. (1988) From Numbers to Symbols to Knowledge Structures: Artificial Intelligence Perspectives on the Classification Task. *IEEE Transactions on Systems, Man, and Cybernetics* 18(3): 415-424.

Charniak, E., & McDermott, D. (1985) *Artificial Intelligence.* Addison-Wesley.

Clement, J. (1988) Observed Methods of Generating Analogies in Scientific Problem Solving. *Cognitive Science* 12: 563-586.

Clement, J. (2008). *Creative Model Construction in Scientists and Students: The Role of Imagery, Analogy, and Mental Simulation*. Dordrecht: Springer.

Darden, L. (1998) Anomaly-driven theory redesign: computational philosophy of science experiments. In T. W. Bynum, & J. Moor (Eds.), *The Digital Phoenix: How Computers are Changing Philosophy* (pp. 62–78). Oxford: Blackwell.

Darden, L. (2006) *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution.* Cambridge: Cambridge University Press, 2006.

Davies, J., Goel, A., & Yaner, P. (2008) Proteus: A Theory of Visual Analogies in Problem Solving. *Knowledge-Based Systems* 21(7): 636-654.

Davies, J., Nersessian, N., & Goel, A. (2005) Visual Models in Analogical Problem Solving. *Foundations of Science* 10(1).

Davis, E. (1990) *Representations of Commonsense Knowledge.* Morgan Kauffman.

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179-201.

Dunbar, K. (1997) How scientists think: Online creativity and conceptual change in science. In Ward, Smith, & Vaid (Editors), *Conceptual structures and processes: Emergence discovery and change* (pp. 461–493). Washington, DC: APA.

Fischer, O., Goel, A., Svirbely, J., & Smith, J. (1991) The Role of Essential Explanations in Abduction. *Artificial Intelligence in Medicine* 3: 181-191.

Glasgow, J., & Papadias D. (1992) Computational Imagery. *Cognitive Science* 16(3): 355-394.

Glasgow, J., Narayanan, N.H., Chandrasekaran, B. (Editors, 1995) *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, MIT Press.

Goel, A., & Bhatta, S. (2004) Use of Design Patterns in Analogy-Based Design. *Advanced Engineering Informatics* 18(2):85-94.

Goel, A., Josephson, J., Fischer, O., & Sadayappan, P. (1995) Practical Abduction: Characterization, Decomposition and Distribution. *Journal of Experimental and Theoretical Artificial Intelligence* 7: 429-450.

Goel, A., Rugaber, S., & Vattam, S. (2009). Structure, Behavior & Function of Complex Systems: The SBF Modeling Language. *AI for Engineering Design, Analysis and Manufacturing,* 23: 23-35.

Goel, A., Rugaber, S., Joyner, D. A., Vattam, S., Hmelo-Silver, C., Jordan, R., Sinha, S., Honwad, S., & Eberbach, C. (2013).

Learning Functional Models of Aquaria: The ACT Project on Ecosystem Learning in Middle School Science. R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognitive and Learning Technologies*. New York: Springer.

Goel, A., Soundararajan, N., & Chandrasekaran, B. (1987) Complexity in Classificatory Reasoning. In *Proc. Sixth National Conference on Artificial Intelligence (AAAI-87),* Seattle, July 1987.

Goel, A., & Stroulia, E. (1996) Functional Device Models and Model-Based Diagnosis in Adaptive Design. *AI for Engineering Design, Analysis and Manufacturing*, 10:355-370.

Goel, A., Vattam, S., Wiltgen, B., & Helms, M. (2012) Cognitive, Collaborative, Conceptual and Creative - Four Characteristics of the Next Generation of Knowledge-Based CAD Systems: A Study in Biologically Inspired Design. *Computer-Aided Design*, 44(10).

Golub, T., Slonim, D., Tamayo, P. et al. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 289(5439): 531-537.

Griffith, T., Nersessian, N., & Goel, A. (1996) The Role of Generic Models in Conceptual Change. In *Proc. Eighteenth Conference of the Cognitive Science Society,* San Diego.

Griffith, T., Nersessian, N., & Goel, A. (2000) Function-follows-Form: Generative Modeling in Scientific Reasoning. In *Proc. Twenty Second Conference of the Cognitive Science Society*, Philadelphia, pp. 196-201.

Hofstadter, D (Editor, 1996) *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought*, Basic Books, New York.

Indurkhya, B. (1991) On the Role of Interpretive Analogy in Learning. *New Generation Computing,* 8(4): 84-402.

Jackson, S., Krajcik, J., & Soloway, E. (2000) Model-It: A Design Retrospective. In M. Jacobson & R. Kozma (editors), *Innovations in Science and Mathematics Education: Advanced Designs for Technologies of Learning* (pp. 77-115). Lawrence Erlbaum.

Jacobson, M. (2008). A design framework for educational hypermedia systems: theory, research, and learning emerging scientific conceptual perspectives. *Educational Technology Research and Development*, 56(1), 5-28.

Johnson-Laird, P. (1983) *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness*. Harvard University Press.

Jones, J. & Goel, A. (2012) Perceptually Grounded Self-Diagnosis and Self-Repair of Domain Knowledge. *Knowledge-Based Systems, 27: 281-301.*

Josephson, J. & Josephson, S. (1996) *Abductive Inference: Computation, Philosophy, Technology*. Cambridge University Press.

Joyner, D., Goel, A., Rugaber, S., Hmelo-Silver, C., & Jordan, R. (2011). Evolution of an Integrated Technology for Supporting Learning about Complex Systems. In *Proc. 11th IEEE International Conference on Advanced Learning Technologies*.

Joyner, D., Goel, A., & Papin, N. (2014). MILA–S: Generation of Agent-Based Simulations from Conceptual Models of Complex Systems. In *Proc. 2014 International Conference on Intelligent User Interfaces*, Haifa, Israel.

Keane, M. (1996). On adaptation in analogy: Tests of pragmatic importance and adaptability in analogical problem solving. *The Quarterly Journal of Experimental Psychology*, 49/A(4).

Langley, P. (1996) *Elements of Machine Learning.* Morgan Kauffman.

Langley, P., Simon, H., Bradshaw, G., & Zytkow, J. (Editors, 1987) *Scientific Discovery: Computational Explorations of the Creative Process*. Cambridge, MA: MIT Press, 1987.

Lenat, D. (1995) CYC: A Large-Scale Investment in Knowledge Infrastructure. *Communications of ACM* 58(11): 32-38.

Lindsay, R., Buchanan, B., Feigenbaum, E., & Lederberg, J. (Editors, 1980) *Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project.* McGraw-Hill.

Magnini, L. (2001) *Abduction, Reason and Science.* Springer.

Magnini, L., Nersessian, N., & Thagard, P. (editors, 1999) *Model-Based Reasoning in Scientific Discovery*. Kluwer.

Gentner, D., & Markman, A. (1997) Structure Mapping in Analogy and Similarity. *American Psychologist,* 52(1): 45-56.

Holyoak, K., & Thagard, P. (1996) *Mental Leaps: Analogy in Creative Thought.* MIT Press.

Mitchell, M. (1993) *Analogy Making as Perception: A Computer Model.* MIT Press.

Mitchell, T., (1997) *Machine Learning.* McGraw-Hill.

Nersessian, N. (2008). *Creating Scientific Concepts*. Cambridge, MA: MIT Press.

Novak, J. (2010) *Learning, Creating and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations*. New York: Routledge.

Prade, H., & Gilles, R. (Editors, 2014) *Computational Approaches to Analogical Reasoning: Current Trends*. Springer.

Schank, R., & Abelson, R. (1977) *Scripts, Plans, Goals and Understanding.* Lawrence Erlbaum.

Schwarz, C., Reiser, B., Davis, E., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46*(6), 632-654.

Stefik, M. (1995) *Knowledge Systems.* Morgan Kauffman.

Thagard, P. (2005) Mind: Introduction to Cognitive Science. 2nd Edition, MIT Press.

Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial Intelligence*, 46, 259-310.

VanLehn, K. (2013). Model construction as a learning activity: a design space and review. *Interactive Learning Environments*, *21*(4), 371-413.

Vattam, S., Goel, A., & Rugaber, S. (2011). Behavior Patterns: Bridging Conceptual Models and Agent-Based Simulations in Interactive Learning Environments. In *Proc. 11th IEEE International Conference on Advanced Learning Technologies (ICALT-2011),* pp. 139-141. IEEE.

Vattam, S., Goel, A., Rugaber, S., Hmelo–Silver, C., Jordan, R., Gray, S, & Sinha, S. (2011) Understanding Complex Natural Systems by Articulating Structure- Behavior-Function Models. *Journal of Educational Technology & Society, 14*(1): 66-81.

Wilensky, U., & Reisman, K. (2006). Thinking Like a Wolf, a Sheep, or a Firefly: Learning Biology Through Constructing and Testing Computational Theories-An Embodied Modeling Approach. *Cognition and Instruction, 24*(2), 171-209.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, 8, 3-19.

Winston, P. (1993) *Artificial Intelligence,* 3rd edition, MIT Press.