

The Role of Generic Models in Conceptual Change¹

Todd W. Griffith, Nancy J. Nersessian, and Ashok Goel

College of Computing
Georgia Institute of Technology
Atlanta, Georgia 30332-0280
(404) 853-9381

{griffith,nancyn,goel}@cc.gatech.edu

Abstract

We hypothesize generic models to be central in conceptual change in science. This hypothesis has its origins in two theoretical sources. The first source, constructive modeling, derives from a philosophical theory that synthesizes analyses of historical conceptual changes in science with investigations of reasoning and representation in cognitive psychology. The theory of constructive modeling posits generic mental models as productive in conceptual change. The second source, adaptive modeling, derives from a computational theory of creative design. The theory of adaptive modeling uses generic mental models to enable analogical transfer. Both theories posit situation independent domain abstractions, i.e. generic models. Using a constructive modeling interpretation of the reasoning exhibited in protocols collected by John Clement (1989) of a problem solving session involving conceptual change, we employ the representational constructs and processing structures of the theory of adaptive modeling to develop a new computational model, ToRQUE. Here we describe a piece of our analysis of the protocol to illustrate how our synthesis of the two theories is being used to develop a system for articulating and testing ToRQUE. The results of our research show how generic modeling plays a central role in conceptual change. They also demonstrate how such an interdisciplinary synthesis can provide significant insights into scientific reasoning.

1. Conceptual Change in Science

In many instances, solving novel or difficult problems leads to conceptual change. Such conceptual change can range from minor changes in existing concepts to the radical kind of change one associates with "scientific revolutions". A significant issue in modeling conceptual change is how existing knowledge can be used in creating genuinely novel understandings. We hypothesize that generic models play a key role in creating these new understandings. These models encompass domain properties, relations, principles, and mechanisms.

To explore this hypothesis we analyze the role of generic models in a problem solving protocol collected by John Clement (1989). Our analysis makes use of the "cognitive-historical" theory of constructive modeling (Section 3) to provide a conceptual interpretation of the

problem-solving session (Section 4). We then join this analysis with the computational theory of adaptive modeling (Section 5) that we believe provides the representational constructs and processing structures necessary to model the protocol as so analyzed. Together, the conceptual interpretation and the computational theory enable the development of a new computational theory we call ToRQUE (Theory Revision through Questions, Understanding, and Evaluation) and a system which instantiates this model. (Section 7).

2. The Clement Protocol

The problem posed in the Clement protocol is as follows:

"... a weight is hung from a spring. The original spring is replaced with a spring made of the same kind of wire; with the same number of coils; but with coils that are twice as wide in diameter. Will the spring stretch from its natural length more, less, or the same amount under the same weight? (Assume the mass of the spring is negligible compared to the mass of the weight.) Why do you think so?" (Figure 1 a & b)

In the study, subjects were asked to assess their confidence in their answer and in their understanding. We focus on one subject, S2, who changed his concept of a spring by incorporating the physical principle of torque into his understanding of how springs function.

Unable to solve the problem directly, S2 began by reasoning that a spring when it is unwound is like a flexible rod (Figure 1c). He then reasoned that a spring of twice the diameter can be unwound into a longer rod, which will bend farther given equal force (Figure 1d). From this he concluded (correctly) that a spring of twice the diameter will stretch farther given equal force. S2, however, unlike most of the participants in the study, was not confident of this answer. He noticed that a significant difference between the stretched spring and the bent rod is that the bent rod has a varying slope, while the spring has a constant slope, i.e., the space between the coils is uniform both before and after the spring is stretched. At this point S2 constructed the models that are the primary focus of our modeling effort (Figure 1e-i). These models were constructed based on salient differences between the spring

¹ This research was funded in part by NSF Grant No. IRI-92-10925 and in part by ONR Grant No. N00014-92-J-1234. We thank John Clement for the use of his protocol transcript, James Greeno for his contribution to developing our constructive modeling interpretation of it, and Ryan Tweney for his helpful comments

and the flexible rod, and are designed to resolve what S2 regarded as an anomaly: the nonuniform slope of the bending rod (see Darden 1991 on anomaly resolution). He eventually constructed a model of a hexagonal coil (Figure 1g) that led to the understanding that a spring maintains its constant slope through the twist of the coil wire during stretching. The notion of torque was not present in S2's original model of spring, so we contend that S2's concept of a spring is changed in the problem solving process. Although we are modeling the whole protocol, given space limitations we will focus on just this final piece of reasoning and how we interpret it as employing "generic models".

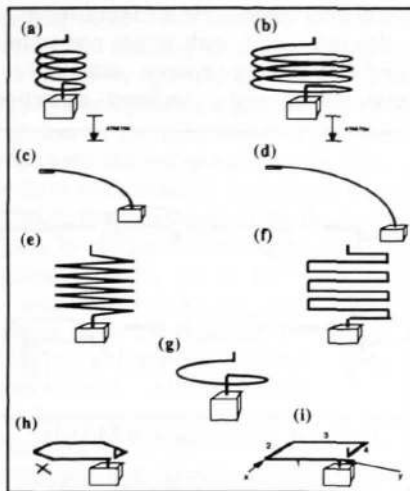


Figure 1: Clement Figures

3. Constructive Modeling

Nersessian (1992, 1995, in press) has argued that general modes of reasoning such as visual reasoning, thought experiment, analogy, and generic abstraction play significant roles in scientific conceptual change. These various modes often are employed together in an iterative reasoning process we call "constructive modeling." Constructive modeling is a semantic process in which the models produced are proposed as interpretations of the target satisfying specific constraints. Figure 2 provides a schematic representation of such a process. Constructing a model starts with properties and relations of a target system that serve as constraints to be satisfied by the initial model. A source domain satisfying some initial target constraints is selected. From this domain an initial analog model is retrieved or is constructed in the case where no direct analogy exists. This initial model - and each constructed model - serves as a source of additional constraints that interact with those provided by the target system to create an enhanced understanding of the target, in particular by making explicit further target constraints. The constraints can be supplied in different informational formats, including equations, texts, kinesthetic, diagrams, pictures, maps, and physical models. The model construction process involves different forms of abstraction (limiting case, idealization, generalization, generic abstraction), constraint satisfaction, adaptation, simulation, and

evaluation. Additional source domains may be called upon throughout the iterations. This cycle is repeated until a satisfactory representation of the target problem is achieved. This representation is a model of the same type as the target problem with respect to the salient target constraints. We interpret S2's reasoning to be a case of constructive modeling.

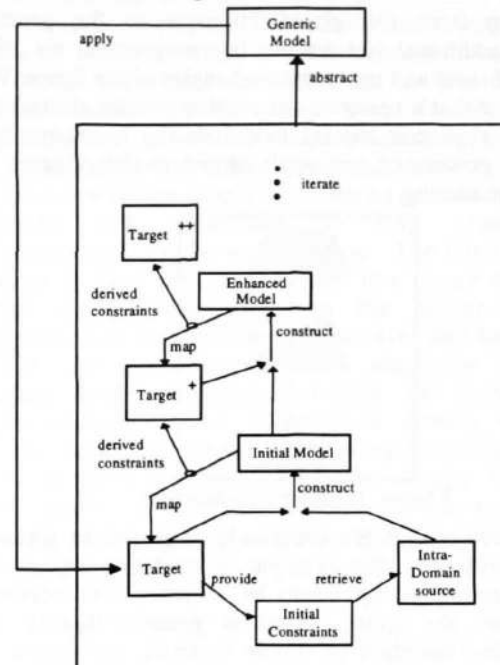


Figure 2: Constructive Modeling

Clearly, to engage in constructive modeling the reasoner needs to know the generative principles and constraints for physical models in one or more domains. This is why analogy plays such a significant role in the constructive modeling process. On our account, the function of analogies is to provide constraints and generative principles for building models. This view is in contrast to the direct transfer view of most computational models (See for example Falkenhainer *et al.*, 1989; Holyoak & Thagard 1989) Thus we view relations between domains in terms of the constraints they share. These constraints and principles may be represented in the different informational formats and knowledge structures that act either as explicit or tacit assumptions employed in constructing and adapting models during problem solving. Since these constraints are domain-specific they need to be understood at a sufficient level of abstraction in order for retrieval, transfer, and integration to be possible. We call this level of abstraction "generic".

What we mean can easily be conveyed by looking at a simple example taken from Polya (1954). Polya considered two cases, abstracting from an equilateral triangle to a triangle-in-general and from it to a polygon-in-general (Figure 3). Loss of specificity is the central aspect of this kind of abstraction process. We call this process "generic abstraction." The generic triangle created in this abstraction process is understood to represent those features

that all kinds of triangles have in common. Although the figure entertained by the mind is specific, some of its salient features, the lengths of the sides and the degrees of the angles, must be taken by the reasoner to be unspecified. In contrast to this, a logical generalization from one equilateral triangle to all equilateral triangles maintains the specificity of these salient aspects of "equilateral". In abstracting from the generic triangle to the generic polygon, additional features are left unspecified, viz., the number of sides and the number of angles of the figure. We hypothesize that a reasoner can employ generic abstraction to create a generic mental model during a constructive modeling process or can apply stored models created in previous reasoning.

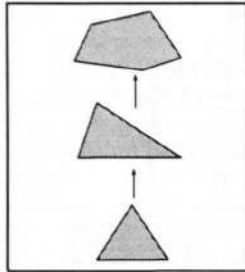


Figure 3: Generic abstraction

Generic models are commonly employed in solving physics problems. For example, in modeling a problem about a pendulum by means of a spring, the scientist understands the spring model as generic, that is, as representing the class of simple harmonic oscillators of which the pendulum is a member. We interpret much of the research in expert physics problem solving as demonstrating this (see for example Chi *et al.*, 1981). Further, we believe generic models facilitate analogical retrieval, mapping, and adaptation in the constructive modeling process. This is exemplified in the psychological literature by Holyoak and collaborators (see for example Gick & Holyoak 1983). Through the mediation of generic models, knowledge from multiple domains can be brought to bear on a problem and can be transformed to such an extent that something truly novel emerges, as is the case in conceptual change.

Goel has developed a theory of generic models in the context of design (see Stroulia & Goel 1992 and Bhatta & Goel 1993). In his work, generic models are learned from specific domain experiences and are used for analogical transfer across design domains (Section 5).

There are several ways in which we interpret generic models as playing a role in S2's constructive modeling process: generic abstraction is employed to create models that incorporate constraints from multiple domains; generic adaptation strategies are employed to make changes to models, and knowledge of generic transformations and principles is used in model construction and adaptation.

4. A Constructive Modeling Interpretation of S2's Reasoning

S2 was a computer scientist with extensive training in topology. In the protocol session, he spent considerable

time considering his "physical imagistic intuition" (025)² about the slope of the bending rod. We begin here at the point he claimed to have a visual experience that "expressed what [he was] thinking" (049) With the rod one "is always measuring in the vertical -- maybe somehow the way the -- the coiled spring unwinds, makes for a different frame of reference." (049) This insight would lead, though not immediately, to a model of the spring as an open horizontal (3-d) coil (Figure 1g). This part of the session generated a target constraint that was salient in this and the final two models (1e,i): spring coiling is in the horizontal plane.

At this point S2 was seeking to reconcile the rod (1c) and circular coil (1g) models. He achieved reconciliation by integrating the rod model with target constraints derived during the problem solving process: circularity, lying in the horizontal plane, and uniform distortion during

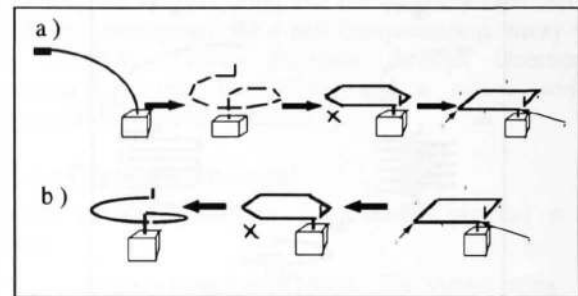


Figure 4: Progression of Models

stretching. S2 recognized that transmitting the force incrementally along the circle in the horizontal plane stretches it bit by bit, as though it had joints, but with even distribution. He now recalled an earlier idea that a "square is sort of like a circle". (117) We interpret him to mean that squares, considered generically are polygons and polygons approximate circles in the limit. He immediately considered bending up the rod into an approximation of the circle to create "a continuous bridge" between the two paradigmatic cases. We take this as his attempt to ascertain if a rod bent in a joint-like fashion in the horizontal plane and a circle bending under a force transmitted incrementally are of the same type with respect to the mechanism of bending. This interaction between the enhanced target (unfolding circle) and the initial source model (flexible rod) led to his constructing a series of generic polygonal models we have represented in Figure 4.

S2 first drew a picture of a horizontal hexagon (Figure 1h) and saw immediately that the hexagonal model is a model of a different type from any considered before for how the constraints would interact in the dynamic case where the spring is stretched. S2's next statement described a simulation that provided a crucial insight: "Just looking at this [1h] it occurs to me that when force is applied here, you not only get a bend on this segment, but because there's a pivot here ['X' in 1h], you get a torsion effect -- around here." (121) He went on, "Aha! -- Maybe the behavior of the spring has something to do with the twist forces

² These numbers are line numbers from the original protocol.

that might be the key difference between this [flexible rod], which involves no torsion, and this [hexagonal coil]." (122) Finally, S2 constructed the last model, drawing a square coil (1i) in order to exaggerate the torsion effect and considered the possibility that torsion is what "stops the spring from -- from flopping." (126).

We interpret these steps in S2's problem solving as employing generic models of the relational structures and physical properties of the polygonal models. Both the hexagon and the square models incorporate features of the rod because the straight-line segments can bend. However, in this orientation any polygonal model will localize the torsion at the corners, so that the motion in stretching is that of twisting rather than bending at the joints. Thus there is torsion plus bending in this stretching process. The square coil model or the hexagonal coil model or any polygonal model will provide a generic model of the spring coil with respect to the mechanism of torsion. The key difference between the polygonal models (1g-i) and earlier models we have not discussed here (1e,f) is that when the wire is coiled in the horizontal plane the bending segment does not have to change directions, so the bend is in the same relation to each piece and the springiness is distributed evenly, satisfying the target constraints. That the distribution of the twist would be even can be seen by extrapolating the polygon to the limit of a circle, where bending goes to zero. Although these steps are not in the protocol, we interpret generic models as having enabled S2 to grasp immediately the move backwards from the square coil to the hexagon through the intermediate extrapolations to the limit of the circular coil in which the torsion that is localized at the corners spreads itself out in such a way that it becomes a uniform property of the spring (Figure 4b).

5. Adaptive Modeling

Since we view model construction and adaptation as central in conceptual change, we have chosen to start with an AI theory that views design in a similar fashion for identifying representational constructs and processing structures for building computational accounts of constructive modeling in science. In an independent line of research, Goel and collaborators have developed an AI theory of conceptual design of physical devices that views device design as model construction and adaptation. This theory, called "adaptive modeling", arose from work on the Kritik project (Goel 1991). A designer's comprehension of the functioning of a known device is represented in the form of a structure-behavior-function (SBF) model that provides a functional and causal explanation of how the structure of the device delivers its functions. Figure 5 illustrates the main elements of an SBF model and the interdependencies between them. The computational system designs new devices by constructing SBF models for them, and new device models are constructed by adapting the SBF models of known devices. The SBF models of the new device designs are verified through a form of qualitative simulation, and, if needed, revised.

Recent work along this line of research has led to a theory of creative conceptual design. This theory extends

and expands Kritik's theory of adaptive modeling by incorporating analogical transfer as another family of adaptation strategies. It posits generic models for mediating the analogical transfer. In particular, it identifies two kinds of generic models: generic teleological mechanisms (GTM) and general physical processes (GPPs) (Stroulia & Goel 1992; Bhatta & Goel 1993). A GTM specifies a pattern of functional and causal structure such as feedback while a GPP captures a pattern of behavioral and causal structure such as heat flow. The generic models are abstracted from the SBF device representations of a known design situation, indexed by the functional/behavioral abstractions, and stored in memory. Given a new design situation, the stored generic models are accessed and instantiated to help create SBF representations for the new situation. The IDEAL system (Bhatta & Goel 1993) instantiates this theory of *model-based analogy*. Depending on the design situation presented to it and its relation to the available knowledge, IDEAL can use different model adaptation strategies ranging from incremental revision of known SBF representations within the problem domain to cross-domain analogical transfer of modeling knowledge in the form of generic models. The SBF theory of device comprehension and the adaptive modeling theory of solving design problems together provide us with the representation and processing structures for beginning to build a computational account of the constructive modeling reasoning process in science.

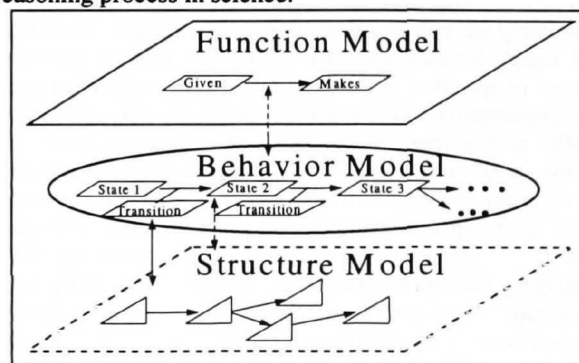


Figure 5: SBF Model

6. Synthesis of Theories

By itself "constructive modeling" provides an outline for a process of scientific reasoning that results in conceptual change. In order to acquire a more specific understanding we have been developing a computational theory based on the principles of adaptive modeling to explore and test our interpretation of the Clement protocol. This collaborative effort engages a problem central to cognitive science as an interdisciplinary research field: How can theories from different disciplines be synthesized to provide a richer understanding of reasoning processes? And how might a synthesis be utilized to develop computational systems for experimentation? In this project we have a cognitive-historical theory of constructive modeling paired with the computational theory of adaptive modeling. The result of this pairing is that we are provided

with 2 kinds of constraints for the choices we make in modeling. The first are cognitive constraints drawn from a "cognitive-historical" synthesis of philosophical, historical, and psychological studies of human reasoning. These include both interpretive constraints for analyzing data and processing constraints in the form of coarse-grained commitments. The second are computational constraints drawn from computer science and theories of cognition which include tractability, inferencing capability, and representational adequacy. Thus the choices we have made in developing ToRQUE garner support from both theories and the interaction between them. In the next section we explain and justify some of the choices that we have made in the development of ToRQUE with respect to the computational and cognitive constraints of these theories.

7. Computational Analysis

In our computational analysis we have developed a preliminary computational model of S2's reasoning. This analysis models a smaller piece than our constructive modeling interpretation of the protocol i.e. we have not focused on every aspect of S2's reasoning but have focused instead on specific issues such as his use of generic models. Our computational model is described in the SBF language of the theory of adaptive modeling. Thus far in our research the computational model, ToRQUE, has been instantiated in a partial experimental system.

For S2's reasoning the choice of adaptive modeling is particularly apt computationally for two reasons: there is a good match between the SBF formalism and the physical systems in question (i.e. springs, flexible rods, etc.) and, more importantly, SBF representations provide significant benefits with respect to the kinds of inferences available, and the speed with which those inferences are carried out. The structure (S) of S2's initial model of a spring is clearly one of multiple coil components that interact with one another. This interpretation is supported by S2's simplifying the representation by reducing the spring to a single coil: "It occurs to me that a single coil of a spring wrapped once around is the same as a whole spring." (023) The inference is not that a coil is equivalent to a spring, but that it has the same basic function (F) as a spring, because in most respects a coil is not the same as a spring. (e.g. it does not look like a spring or have the same structure as a spring). This inference provides evidence that S2 used separate notions of function (F) and structure (S). A spring and a coil can be "the same" functionally while not being the same structurally or topologically. It also shows that S2 considered the spring as divided into multiple coil components.

The task that S2 completed involves assessing the behavior (B) of a particular physical system with regard to its structure (S). Given a particular property of the spring's structure, e.g. the diameter value, how will the behavior of the spring be affected? S2's attempt to solve this problem requires having a representation of the behavior in question or being able to generate one quickly. One of the advantages of adaptive modeling is that the explicit storage

of this behavior provides a significant computational advantage over the generation of the behavior. The kind of inferences that can be made given the stored behavior are also important. For example, when S2 noticed the difference (change in slope in the flexible rod vs. uniform slope in the spring), he did so because the behavior shows this difference to be salient. By separating structure, behavior, and function into separately analyzable units, the SBF formalism prunes away differences that are irrelevant to the task, and makes it easier to target areas of significant difference. Thus once the model is paired with a task it is possible to see the salient differences without being distracted by ontologically distinct kinds of differences.

Once S2 considered a single coil in place of an entire spring we see that he began to focus on the topological feature of circularity. At this point in the protocol he has already considered the behavioral and structural differences, and has made some adaptations with respect to these parts of the model. The difference between the behavior of the flexible rod and the spring provided the initial set of salient differences, and the structural adaptation from many coils to one coil allowed S2 to focus his attention on what turned out to be the most important differences: circularity and orientation.

At this stage in the protocol ToRQUE's SBF model of a coil and the SBF model of a flexible rod each have a single component which has the function of providing a restoring force. Because "Structure" refers to components and the connections of components, the structures of two devices with a single similar component are necessarily the same.

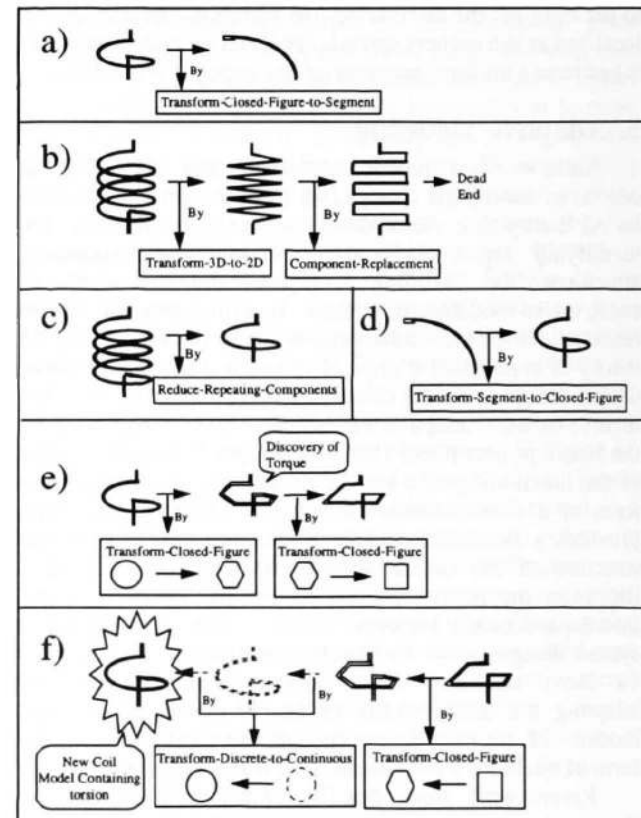


Figure 6: Sequence of Generic Adaptations

The topologies of these devices, however, may still be significantly different. That S2 addressed the differences in this order provides further support that SBF structures are a useful ontology for focusing inferences. Problems such as S2's that involve behavioral aspects of the physical system are handled best by focusing on behavioral differences first. Thus S2 is required to make use of the topological differences between the coil and the flexible rod, only after he has pruned away those differences which are presented by the behavior and structure.

Just as IDEAL uses GTMs and GPPs in adapting models, ToRQUE uses generic topological transformations (GTTs) for adapting models. Here we describe the use of these transformations with respect to S2's reasoning in the final insight section interpreted in Section 3. In ToRQUE, the "Reduce-Repeating-Components" transformation is used to reduce the spring to a single coil (Figure 6c). The "Transform-Segment-to-Closed-Figure" and "Transform-Planar-Orientation" bend the rod into a coil (6d). We assume here with S2 that a coil "is a circle with a break in it". Figure 6e shows the progression of closed-figure transformations, which leads to the hexagonal coil, the discovery of torque, and the exaggeration of the effect by the square coil model. By adapting the coil from a circle to a polygon, S2 was able to introduce new components into the model structure. Each side of the square, e.g., could now be treated as a flexible rod component, but with the significant change in orientation that now makes for twisting rather than bending at the joints. Thus a small topological change can result in a fairly large behavioral change, making new knowledge available from which to make inferences.

The most important inference occurs in evaluating the square coil. S2 had recognized the generic physical principle (GPP) of torsion in the hexagonal coil and constructed the square coil to examine it. He was reminded of this principle because of the behavioral and structural similarities between the GPP and the polygonal models. In Section 3 we interpreted S2 as making a final series of inferences only implicit in the protocol that involve the generic abstraction of the square coil with respect to torsion. To be satisfied that he had solved the problem, he needed to hypothesize that if torsion is true of square coils, perhaps it is true of all coils and to make the appropriate extrapolation. ToRQUE incorporates the GPP into the circular coil model through the "Transform-Discrete-to-Continuous" GTT, which depends upon a knowledge of limits which we know S2 possesses: A continuous shape such as a circle can be thought of as containing an infinite number of infinitesimally small segments. Figure 6(f) shows the transformations from the square coil back to an adapted model of the circular coil that capture our interpretation.

Conclusion

Our conceptual analysis provides a plausible interpretation of S2's reasoning as relying significantly on generic models. Our computational analysis shows how generic models such as GPPs (e.g., torque) and GTTs (e.g.,

Transform-Planar-Orientation) can help to achieve conceptual change. Here we highlight two significant conclusions that show the synergy of our interdisciplinary collaboration:

- An important issue in generic modeling is how to make the right inferences at the right times. SBF models enable and constrain these inferences.
- In analyzing protocol and historical data there are places where the reasoning process is not explicit, as in the portion of S2's reasoning we examined here. Interpretations of these gaps gain plausibility through computational models, which like ToRQUE have developed out of an interdisciplinary analysis of creative reasoning.

References

- Bhatta, S.R. & Goel, A.K. (1993) Learning Generic Mechanisms from Experiences for Analogical Reasoning. In *Proc. Fifteenth Annual Conference of the Cognitive Science Society*, Boulder, Colorado, July 1993, pp. 237-242, Hillsdale, NJ: Lawrence Erlbaum.
- Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981) Categorization and Representation of Physics Problems by Experts and Novices, *Cognitive Science*, 5, pp. 121-152.
- Clement, J. (1989) Learning via Model Construction and Criticism: Protocol Evidence on Sources of Creativity in Science, In *Handbook of Creativity: Assessment, Theory and Research*, Glover, G., Ronning, R., & Reynolds, C. (Eds.), chapter 20, pp. 341-381. New York, NY: Plenum.
- Darden, L. (1991) Anomaly Driven Redesign of a Scientific Theory: The TRANSGENE.2 Experiments, Technical Report, Ohio State University.
- Falkenhainer, B., Forbus, K.D., and Gentner, D. (1989) The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41, 1-63. (University of Illinois Technical Report UIUCDCS-R-87-1361, July, 1987).
- Gick, M.L. & Holyoak, K.J. (1983) Schema induction and analogical transfer. *Cognitive Psychology*, 12(3), 306-355.
- Goel, A.K. (1991) Model revision: A theory of incremental model learning. In *Proc. of the Eighth International Conference on Machine Learning*, pages 605-609, Chicago.
- Holyoak, K. & Thagard, P. (1989) "Analogical Mapping by Constraint Satisfaction: A Computational Theory." *Cognitive Science*, 13:295-356.
- Nersessian, N.J. (1992) How Do Scientists Think? Capturing the Dynamics of Conceptual Change in Science, In *Cognitive Models of Science*, ed. R.N. Giere. pp. 3-44. Minneapolis, MN: University of Minnesota Press.
- Nersessian, N.J. (1995) Constructive Modeling in Creating Scientific Understanding, *Science & Education*, 4: 203-226.
- Nersessian, N.J. (in press) Abstraction via Generic Modeling in Concept Formation in Science. In *Idealization in Science*, M.R. Jones & N. Cartwright, eds. (Rodopi).
- Polya, G. (1954) *Induction and Analogy in Mathematics*, Vol. 1, Princeton University, Princeton.
- Stroulia, E. & Goel, A.K. (1992) Generic Teleological Mechanisms and their Use in Case Adaptation, In *Proc. of the Fourteenth Annual Conference of the Cognitive Science Society*, 319-324, Lawrence Erlbaum, Hillsdale, N.J.