The Role of Reflection in Scientific Exploration

Todd W. Griffith (griffith@cc.gatech.edu)
J. William Murdock (murdock@cc.gatech.edu)
College of Computing
Georgia Institute of Technology
801 Atlantic Drive
Atlanta, Georgia 30332-0280, USA

Abstract

In this paper we explore the idea of reflection in the context of scientific exploration. How does an agent reflect upon its behavior in order to enable productive exploration? We outline an abstract cognitive architecture for combining reflection and exploration. To achieve this we present a language for modeling cognition: the Task-Method-Knowledge (TMK) language. We further present a computational model based on this language, ToRQUE2 (Griffith et al., 1997; Griffith, 1997). ToRQUE2 is a model of exploratory reasoning in the domain of scientific problem solving. We claim that the TMK language supports both reflection and exploration, and enables them to benefit from one another.

Introduction

One outstanding issue in scientific discovery is: how do scientists decide when to abandon one reasoning strategy to pursue another? Or from a cognitive perspective: how do we model a process for multi-strategy exploration? In our research we show evidence from a computational system, ToRQUE2, which suggests that "reflection," an agent's ability to reason about its own reasoning, plays an important role in the selection of reasoning methods and in the abandonment of reasoning paths.

For the past 20 years researchers have attempted to computationally model the processes of science. In this pursuit they have investigated tasks including: experimental data interpretation (Langley et al., 1987; Zytkow, 1990), hypothesis formation (Karp, 1989), and theory revision (Darden, 1991; Rajamoney, 1990; Falkenhainer, 1990). In order to understand these tasks, researchers have modeled historical cases of scientific reasoning (Kulkarni and Simon, 1988; Darden, 1991; Thagard and Nowak, 1990). One difficulty with such cases is that they rarely show paths of reasoning in which the scientist was unsuccessful, yet it is these paths which make up the majority of the cognitive effort (Nersessian, 1993). So researchers mistakenly place emphasis on the sources that have survived. This, however, presents a problem. It results in computational models which are cognitively implausible and are of little use in the development of scientific agents.

For this reason some would have us throw out the baby with the bath water by giving up on historical cases of scientific reasoning altogether. We argue that this abandonment is premature. The trouble is not with the research program but with the methodology. In order to construct accurate computational models, one must model each of the strategies available to the scientist as well as the control processes for deciding which strategy to use when. In our research, we have chosen to model reasoning found in problem-solving protocols (Clement, 1989), because these protocols show both successful and unsuccessful attempts at solving a problem using multiple reasoning strategies. The computational model developed through the protocols can then be applied to historical cases. The key is to acquire protocols which address similar issues to those found in the historical case (Nersessian and Greeno, 1990). Only when we can construct models of scientists which capture their many reasoning strategies, will we be able to construct agents capable of real scientific activity.

In the ToRQUE2 system we have modeled four reasoning strategies for exploratory problem solving: model-based search (MBS), model-based analogy (MBA), structure-based model transformation (SBMT), and limiting-case analysis (LCA). These strategies are represented in the Task-Method-Knowledge (TMK) language (Murdock and Goel, 1998). In the remainder of this paper we describe a control process which enables the selection of strategies through reflection. We claim that this process is sufficient for modeling exploration in science.

Reflection

A theme which runs through a vast range of research paradigms in cognition is the idea of reflection: humans are able to explicitly reason about their own cognition. There are a great many aspects of cognition which seem to be related to the notion of reflection. A few of these are:

- **Adaptation** People are able to adapt to an extremely broad range of novel situations in the world. It is clear that they are able to modify their reasoning abilities to some extent. It seems likely that the ability to modify one's own reasoning may depend on knowledge about that reasoning.
- **Explanation** People are able to explicitly state what they are doing and why. While it is clear that this ability is limited and is often inaccurate, it is certainly present. The ability to even partially describe one's own reasoning seems like conclusive evidence of at least a partial self-understanding.
- **Multi-strategy reasoning** Explicit representations of functional elements potentially enables dynamic, flexible selection of these elements within a reasoning episode.
- **Prediction** People are able to predict the actions of other people. It seems plausible that this ability is tied to some knowledge of their own mental processes.
- **Error Prevention** In some situations, people are able to predict cognitive errors that might make and act to prevent

them. For example, someone doing a complex arithmetic problem, as in (Brown and VanLehn, 1980), might recognize that there is a significant chance of error and might do the problem twice to confirm that the correct answer was obtained; such reasoning clearly depends on an understanding of what constitutes complex, error prone reasoning.

Error Recovery Many researchers in cognitive science have observed that explicit representation of processes can be particularly useful in recovering from mistakes after they occur (Suchman, 1987; Kirsh, 1991). When someone reaches a reasoning state which is clearly erroneous, they are motivated to reason about how they got to that state, even if the process which got them their was largely non-reflective at the time.

Much of the past work on TMK models has focused on the first three topics above: adaptation (Stroulia, 1994; Stroulia and Goel, 1995; Murdock and Goel, 1998), explanation (Goel et al., 1996; Goel and Murdock, 1997), and multi-strategy reasoning (Punch et al., 1996). In this paper we consider yet another possible application of reflection: that of exploratory reasoning. This idea extends the notion of multi-strategy reasoning; instead of simply selecting "*the*" best method or *a* satisfactory method, we might explore a variety of possible methods for accomplishing a given task. Tasks such as scientific reasoning, in which an agent clearly does not have an *a priori* account of what constitutes optimal or even satisfactory methods for solving problems are ideally suited to this kind of reasoning. Exploratory reasoning involves a great many complex decisions. Some of the kinds of decisions that an exploratory reasoner might make include: what method to consider next, when to attempt a new method for solving a problem, when to concede that a problem cannot be solved, etc. Reflection provides the knowledge which may be used to guide these decisions.

TMK Models

Our development of an account of reflection originates in a line of research which examines computational representations of physical devices (Goel, 1989; Goel et al., 1996; Goel et al., 1997). This research shows that effective adaptation of physical devices can be supported by models which are causal (i.e., that show the mechanisms by which effects occur), compositional (i.e., that show how the effects of elements of a device are combined), and functional (i.e., that take an intensional stance toward describing why elements are arranged as they are). The basic idea behind these models is that the relationship between the physical construction of a device (i.e., the structure) and the intended effect of that device (i.e., the function) is described by a flow of causal interactions (i.e., the behavior).

There have been a number of theories which have viewed elements of cognition as inherently device-like in nature. The theory of Generic Tasks (Chandrasekaran, 1988) suggests that there are primitive "building blocks" of cognition such as classification and recognition and that complex reasoning strategies can be viewed as complex devices resulting from the synthesis of these components. This perspective suggests that perhaps past work in adaptive redesign of physical devices may be applicable, by analogy, to the adaptation of these proposed abstract cognitive "devices."

It is this intuition which lead to the development of the TMK language. TMK models have been used for both explanation and adaptation of the capabilities of a variety of AI computer programs. TMK models are very much an extension of our modeling framework for physical devices. The division of reasoning into tasks (i.e., functional elements) and methods (i.e., behavioral elements) very much duplicates the functional and causal features of our physical device models, and the explicit modeling of knowledge states duplicates the compositional aspects of these device models.

The extensive AI research on TMK models has conclusively shown that they do provide at least *some* support for a computer system to adapt its own capabilities. Does this necessarily make them relevant to the study of human cognition?[1] Consider the space of all *possible* mechanisms for representing self-knowledge for exploration. It is intuitively obvious that this space is extremely large; virtually any kind of knowledge could potentially be in memory. In contrast, consider the space of all *sufficient* mechanisms for this phenomenon. This space is clearly far more restricted; the mechanism must be shown to be both powerful enough to allow exploration to occur, parsimonious enough to form a credible theory, and general enough to support a wide range of exploratory reasoning. To the extent that these are strong restrictions, it is apparent that the space of sufficient solutions is very small with respect to the space of possible solutions. Furthermore, as the breadth and depth of the sufficiency is increased, it is conceivable that the space of sufficient solutions becomes so small that any sufficient solution approximates a necessary solution. From this reasoning, we argue that because TMK models have been shown to provide a broad and powerful range of capabilities, they constitute a plausible hypothesis for a cognitive account of reflective self-knowledge.

Exploration

We define exploratory reasoning as reasoning which is not directed by some goal. This does not mean that goals must be absent from the reasoning. On the contrary, most interesting cognitive activities involve goals. It does mean, however, that some goals are either too abstract to direct the control, e.g., one might have a goal to discover the cure for cancer, or the agent may simply lack the knowledge to suggest possible alternatives, e.g., a robot which does not know what cabinets are wanders into a strange room full of cabinets with the goal to find a hammer. In the first case the goal does not by itself suggest a course of action. Discovering a cure for cancer is simply too broad to suggest a course of action. In the second case, since the robot has never seen a cabinet it is unlikely that its goal to find a hammer would direct it to open a cabinet. Certainly discovering the cure for cancer can be broken down into many subgoals, but a tremendous amount of knowledge is needed to both form these subgoals and to decide which of these subgoals to pursue. The issue is not whether or not

[1]Note that this question is fundamental to the role of AI and computational modeling in the study of human cognition. The argument presented here is present, either implicitly or explicitly, in an enormous range of research, e.g. (Newell and Simon, 1963; Minsky, 1975; Anderson, 1983). Furthermore, it is largely isomorphic to traditional philosophical views on the nature of all scientific theories, e.g. (Kant, 1781).

we should give agents more knowledge which is clearly true, but rather a control issue: what should an agent do when appropriate knowledge is not available?

Much of the distinction we are making between goal-directed and exploratory reasoning can be clarified by referring to an AI principle known as the principle of rationality (Newell, 1981). The principle of rationality states that an intelligent agent will act in a manner to achieve its goals. For this reason AI has focused on goal-directed behavior in which the actions of the agent are always directed toward some goal. One problem with this focus is that it does not say what an agent should do when the goals are not sufficient to direct reasoning as in the examples above. Newell often referred to such situations as impasses. Newell's solution to the impasse problem was to set up more goals (often these goals were learning goals). In some situations, however, an agent may not know in advance what the relevant sub-goals are nor how to choose between them. The point here is that rational agents cannot be purely goal-directed but must also be capable of making decisions when presented with abstractly specified goals or insufficient knowledge. Intelligent agents must have methods for pushing beyond the bounds of their current knowledge. In other words, agents must be able to refine their own state space. An early example of such an exploratory system is AM (Lenat, 1978).

In this research we propose that exploratory reasoning is achieved by making structural transformations to a model and then generating the behaviors which follow from the changes to the structure. The result is a model which may be either further or closer to the goal than the original model. This new model can then be transformed in a similar manner. An agent's ability to reflect on the resulting knowledge in working memory constrains this exploration. It is through this kind of exploration that imaginative reasoning occurs, and we posit that this is one form of human imagination which has led to discoveries and inventions in the face of seemingly unsolvable goals.

A Case of Exploratory Problem Solving

In this research we have investigated imaginative reasoning evidenced in a scientific problem solving protocol taken by John Clement (Clement, 1989). In the protocol a subject known as S2 attempts to solve a problem about springs:

> ... 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)

On our interpretation, S2 began the problem-solving session with the understanding that the stretch of a spring is due to it's flexibility. Then he derived a new understanding that a spring maintains constant slope when stretched through both twisting and bending. So, although this is a more modest outcome of imaginative reasoning than evidenced in historical cases, for S2 it was an instance of highly imaginative problem solving. To find a satisfactory explanatory model for the

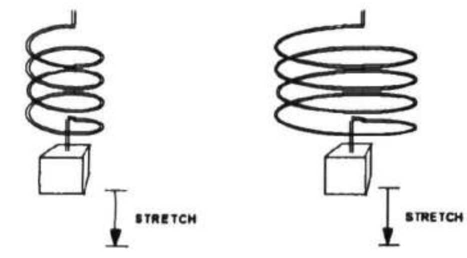

Figure 1: Initial Problem

problem solution, S2 had to construct a novel representation for himself of how a spring works.

To solve this problem, we claim that S2 used a number of transformations we call Generic Structural Transformations (GSTs) to construct a variety of models (Griffith, 1997; Griffith et al., 1996). We refer to these transformations as "function-follows-form" transformations (Griffith et al., 1997) because the form of the model is transformed first and then the behavior of the model is changed in response to this change. The progression of these transformations is shown in Figure 2. The figure shows how each target model is explored and how analogs are retrieved and generated based on that exploration. In a later section we describe how working memory queues are employed to generate this sequence.

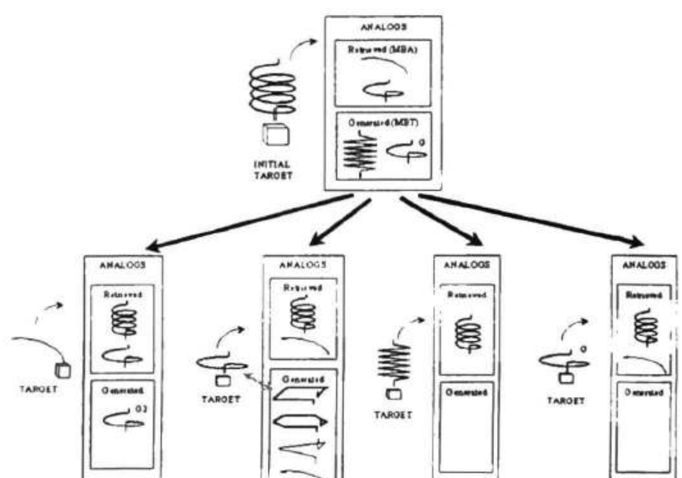

Figure 2: Progression of Target and Source Analogs

The TMK hierarchy shown in Figure 3 presents an architecture for addressing the task of relating some structural feature of a physical system with some behavioral or functional outcome. In the S2 case the structural feature is the diameter of the spring and the behavioral feature is the amount the spring will stretch. He needs to determine how the diameter will affect the behavior of the spring. The hierarchy presents four strategies which S2 appeared to use to address this task:

Model-Based Search, Model-Based Analogy,Structure-Based Model Transformation, and Limiting-Case Analysis.

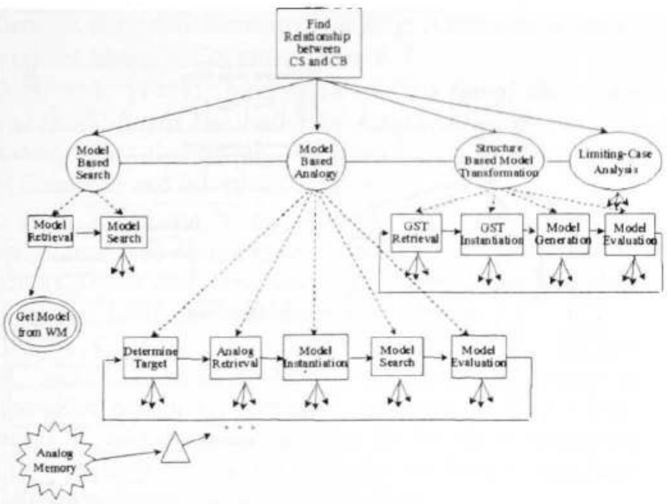

Figure 3: TMK Hierarchy

Instantiation

The value of TMK to a theory of exploration is twofold. First, TMK provides an architecture in which the exploratory processes of multiple subjects can be represented. Second, TMK allows for flexible experimentation with this representation. We claim that both of these features are important for modeling agents capable of exploration.

The TMK representation is more than a program that encodes a subject's reasoning. It is a flexible *model*, which captures the reasoning for some set of subjects. A TMK model shows different behaviors (i.e., reasoning strategies) depending on the knowledge of the individual subject. In experiments with ToRQUE, we have taken the protocols of eleven subjects (including S2) and modeled them.[2] ToRQUE shows eleven unique behaviors depending upon the knowledge of the subject. For example, one subject, S3, is concerned with the material properties of the spring, so he uses a method by which he can explore the properties of the materials. S3 explores properties of the material by considering what would happen if the materials were very flexible or very rigid, i.e., he instantiates a spring where the rigidity of the material is very small and one where it is very large. This is an example of limiting-case analysis. The TMK architecture in ToRQUE models this behavior through a limiting-case method. In ToRQUE the limiting-case method is dependent upon the agent's knowledge. The ToRQUE model hypothesizes that if an agent is unable to retrieve analogs and does not have any transformations available in working memory, then the agent may choose a limiting case strategy which involves instantiating a model with one of its values taken to the limit. The central issue here, however, is not the specific behavior of limiting case, but the fact that given a different knowledge condition the agent may use a different strategy for exploration. Thus the TMK architecture is able to model the reasoning of multiple subjects by providing the system with different knowledge conditions. This allows us to focus on the knowledge and methods which are essential to exploration processes.

[2]Five of these protocols formed an initial set for which ToRQUE was built. The other six were set aside for testing and evaluation

The second reason that TMK is valuable for modeling exploration is that it provides a means for conducting experiments. TMK representations are completely declarative. The declarative nature of these representations provides both the experimenter and the program itself with a means for adapting the control of the process. So as an experimenter I am able to change the ordering of methods or subtasks simply by reorganizing a list. This flexibility is vital when attempting to model the reasoning of subjects. It is clear that subjects are not limited to a single reasoning strategy. Researchers in cognitive science have identified analogy (Gentner, 1983; Holyoak and Thagard, 1989), limiting-case analysis (Nersessian, 1992; Weld, 1986), search (Newell and Simon, 1963) and transformation (Griffith, 1997; Griffith et al., 1997; Murthy and Addanki, 1987) as significant reasoning processes, and there are certainly others. It is also the case that human reasoners are able to choose particular reasoning strategies depending upon what knowledge they have available, e.g., finding one's way in a familiar city requires different strategies than finding one's way in a unfamiliar city. Thus, it is important to be able to account for both a subject's knowledge and strategies in using that knowledge. This is a stark contrast to the modeling found in the Protocol Analysis of (Ericsson and Simon, 1984) which imparts a rigid set of production rules to the system and a single reasoning process (heuristic search) based upon those rules. In TMK both the knowledge and the methods can be manipulated based on evidence from the protocols. The advantage is that experiments can uncover interesting relationships between knowledge and process.

To achieve exploration in TMK requires reflection over a working memory of target models (WMT), analogs (WMA), and GSTs (WMGST). As an agent addresses it's task it may come to a point where it does not know how to proceed. Past reasoning stored in working memory allows the agent to pick a GST which is related to the reasoning at hand or to reasoning which has occurred recently. This serves to constrain the randomness of the selection of a GST. In ToRQUE working memory is captured in a queue data structure which has a last-in-first-out (LIFO) structure. Figure 4 shows two snapshots of working memory queues. The snapshot labeled (A) shows the WM during the first model-based analogy process, prior to attempting any transformations. Snapshot B shows what transformations are placed on the queue when the circular coil becomes the target model. The transformation queue between A and B contains the transformations performed between these snapshots. All the transformations which are retrieved are ordered and placed onto this WMGST queue. Thus one can think of this queue as using the last transformation which the agent was thinking about but did not apply. Not all transformations can be used on all models so many transformations may be rejected quickly prior to being applied, e.g., a circle-to-square transformation is only possible if the target model is circular. Also, previously explored target models are removed from the queue such as when a coil retrieves a spring as an analog.

The exploration process proceeds through the interaction of Model-Based Analogy (MBA) and Structure-Based Model Transformation (SBMT) with the working memory queues, WMA and WMGST. MBA retrieves a set of analog models to solve the particular problem, one of which is selected and the rest of which are placed on the WMA queue. The answer which is produced from these analogs is evaluated by attempting to reduce the differences between it and the target model. One method of reducing these differences is to apply SBMT to the source or target analogs. Similarly GSTs are indexed and retrieved by these differences and one GST is applied while the remaining are placed onto the WMGST queue. As reasoning progresses a collection of transformations have been placed into WM. In this way WM is not being used as a repository for knowledge which is currently being addressed, but as a repository for knowledge which has been retrieved but which has not been considered.

Results

The Clement protocols provide some evidence that subjects need to reflect on their reasoning processes in order to carry out effective exploration. The difficulty with such evidence is that it is still unclear what aspects of reasoning are consciously accessible to the reasoner. While it is clear that there is some relationship between reflective knowledge (of the sort represented by TMK models) and consciousness, it is not clear exactly what that relationship is; we believe that this knowledge is partially but not fully consciously accessible. In the S2 protocol there are two clear indications of conscious reflection. The first is a quote in which S2 has become frustrated that his reasoning has cycled to a past reasoning state.

> I feel as though I'm reasoning in circles. And I think I'll make a deliberative effort to break out of the circle..somehow (057) [3]

The idea is that in order to solve the problem, S2, must deliberatively (i.e., consciously) consider what reasoning methods are available and decide upon a method that does not lead to the same unproductive answers. In the ToRQUE2 system this reflection is realized by looking at the current TMK state and recognizing that this state has been visited in the past. ToRQUE2 recognizes that particular methods have lead to the same set of models by looking at its working memory queues (i.e., the K in TMK). Because it is able to reflect on its own knowledge it can choose a new reasoning method. The choice of a new method is a consequence of identifying a past method or set of methods as inadequate given the current knowledge, i.e., reflecting on the state of its reasoning. Later in the S2 protocol we see the same pattern of reflective behavior:

> I keep circling back to these same issues without getting anywhere with them..I-I really haven't had a new idea..in a long time..about this...Ummm...I think I need to, again, to think about it in some radically different way, somehow...(117)

We claim that the need to find reasoning methods which lead to novel models is a crucial ability in exploratory reasoning. It is clear that in order to make progress, S2 needs to find reasoning methods which provide insight into the issues with which he is faced. The ToRQUE2 system provides a computational account of such reflective reasoning.

[3] The numbers indicated are numbers taken from the protocol.

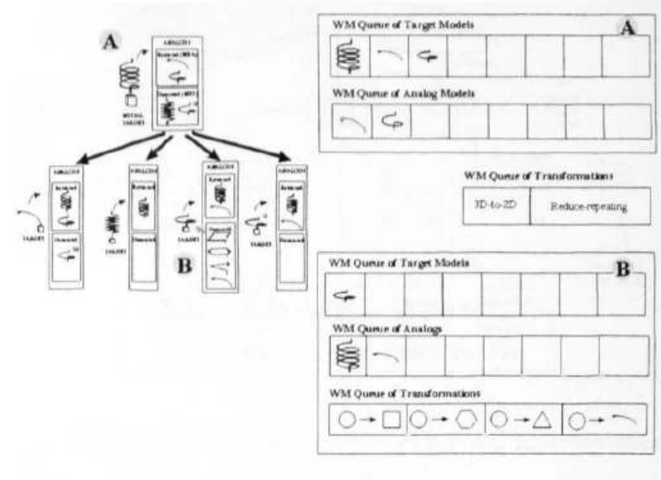

Figure 4: Working Memory Queues

Acknowledgments

The authors would like to acknowledge the extensive contributions of Ashok Goel, Nancy Nersessian, Dean Allemang, and Eleni Stroulia. Portions of this effort were sponsored by the Defense Advanced Research Projects Agency, and the United States Air Force Research Laboratory, Air Force Materiel Command, USAF, under agreement number F30602-96-2-0229. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Defense Advanced Research Projects Agency, Air Force Research Laboratory, or the U.S. Government.

References

Anderson, J. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22.

Brown, J. S. and VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science*, 4:379–426.

Chandrasekaran, B. (1988). Generic tasks as building blocks for knowledge-based systems: The diagnosis and routine design examples. *Knowledge Engineering Review*, 3(3):183–219.

Clement, J. (1989). Learning via model construction and criticism. In Glover, G., Ronning, R., and Reynolds, C., editors, *Handbook of Creativity: Assessment, Theory, and Research*, pages 341–281. Plenum, New York.

Darden, L. (1991). *Theory Change in Science: Strategies from Mendelian Genetics*. Oxford University Press, New York.

Ericsson, K. A. and Simon, H. A. (1984). *Protocol Analysis: Verbal Reports as Data*. MIT Press, Cambridge, MA.

Falkenhainer, B. (1990). A unified approach to explanation and theory formation. In Shrager, J. and Langley, P., editors, *Computational Models of Scientific Discovery and Theory Formation*. Morgan Kaufmann Publihsers, Inc., San Mateo, CA.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, 7.

Goel, A. K. (1989). *Integration of Case-Based Reasoning and Model-Based Reasoning for Adaptive Design Problem Solving*. PhD dissertation, The Ohio State University, Dept. of Computer and Information Science.

Goel, A. K., Bhatta, S., and Stroulia, E. (1997). Kritik: An early case-based design system. In Maher, M. L. and Pu, P., editors, *Issues and Applications of Case-Based Reasoning to Design*. Lawrence Erlbaum Associates.

Goel, A. K., Gomez, A., Grue, N., Murdock, J. W., Recker, M., and Govindaraj, T. (1996). Explanatory interface in interactive design environments. In Gero, J. S. and Sudweeks, F., editors, *Proceedings of the Fourth International Conference on Artificial Intelligence in Design*, Stanford, California. Kluwer Academic Publishers.

Goel, A. K. and Murdock, J. W. (1997). Meta-cases: Explaining case-based reasoning. In Smith, I. and Faltings, B., editors, *Proc. Third European Workshop on Case-Based Reasoning*, Lausanne, Switzerland. Springer.

Griffith, T., Nersessian, N., and Goel, A. K. (1996). The role of generic models in conceptual change. In *Proc. Eighteenth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates.

Griffith, T., Nersessian, N., Goel, A. K., and Clement, J. (1997). Exploratory problem solving in scientific reasoning. In *Proc. Nineteenth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates.

Griffith, T. W. (1997). *A Computational Theory of Dynamical Modeling in Scientific Reasoning*. PhD thesis proposal, Georgia Institute of Technology, Cognitive Science Program, College of Computing.

Holyoak, K. J. and Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13:295–355.

Kant, I. (1781). *The critique of pure reason*. On-Line Edition: Virginia Polytechnic Institute and State University. translated by J.M.D. Meiklejohn.

Karp, P. D. (1989). Hypothesis formation as design. Technical Report KSL-89-11, Knowledge Systems Laboratory.

Kirsh, D. (1991). Today the earwig, tomorrow man? *Artificial Intelligence*, 47(1-3):161–184.

Kulkarni, D. and Simon, H. A. (1988). The processes of scientific discovery: The strategy of experimentation. *Cognitive Science*, 12:139–175.

Langley, P., Simon, H., Bradshaw, G., and Zytkow, J. M. (1987). Discovering problems and representations. In *Scientific Discovery: Computational Explorations of the Creative Process*. MIT Press, Cambridge, MA.

Lenat, D. B. (1978). The ubiquity of discovery. *Artificial Intelligence*, 9:257–285.

Minsky, M. (1975). A framework for representing knowledge. In Winston, P. H., editor, *The Psychology of Computer Vision*. McGraw-Hill.

Murdock, J. W. and Goel, A. K. (1998). A functional modeling architecture for reflective agents. In *Proc. AAAI98 Workshop on Functional Modeling and Teleological Reasoning*, Madison, Wisconsin.

Murthy, S. S. and Addanki, S. (1987). PROMPT: An innovative design tool. In *Proceedings of the Sixth National Conference on Artificial Intelligence*. Morgan Kaufman.

Nersessian, N. J. (1992). How do scientists think? capturing the dynamics of conceptual change in science. In Giere, R. N., editor, *Cognitive Models of Science*, pages 3–44. University of Minnesota Press, Minneapolis, MN.

Nersessian, N. J. (1993). Opening the black box: Cognitive science and history of science. Cognitive Science Laboratory Report 53, Princeton University.

Nersessian, N. J. and Greeno, J. G. (1990). Multiple abstracted representations in problem solving and discovery in physics. In *Proc. Twelfth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates.

Newell, A. (1981). The knowledge level. *AI Magazine*, 18(1):87–127.

Newell, A. and Simon, H. (1963). GPS, a program that simulates human thought. In Feigenbaum, E. A. and Feldman, J., editors, *Computers and Thought*. R. Oldenbourg KG.

Punch, W., Goel, A. K., and Brown, D. (1996). A knowledge-based selection mechanism for strategic control with application in design, diagnosis and planning. *International Journal of Artificial Intelligence Tools*, 4(3):323–348.

Rajamoney, S. (1990). A computational approach to theory revision. In Shrager, J. and Langley, P., editors, *Computational Models of Scientific Discovery and Theory Formation*. Morgan Kaufmann Publihsers, Inc., San Mateo, CA.

Stroulia, E. (1994). *Failure-Driven Learing as Model-Based Self Redesign*. PhD dissertation, Georgia Institute of Technology, College of Computing.

Stroulia, E. and Goel, A. K. (1995). Functional representation and reasoning in reflective systems. *Journal of Applied Intelligence*, 9(1):101–124. Special Issue on Functional Reasoning.

Suchman, L. (1987). *Plans and Situated Actions: The Problem of Human Machine Communication*. Cambridge University Press.

Thagard, P. and Nowak, G. (1990). A computational approach to theory revision. In Shrager, J. and Langley, P., editors, *Computational Models of Scientific Discovery and Theory Formation*. Morgan Kaufmann Publihsers, Inc., San Mateo, CA.

Weld, D. (1986). The use of aggregation in causal simulation. *Artificial Intelligence*, 30:1–34.

Zytkow, J. (1990). Deriving laws through analysis of processes and equations. In Shrager, J. and Langley, P., editors, *Computational Models of Scientific Discovery and Theory Formation*. Morgan Kaufmann Publihsers, Inc., San Mateo, CA.