Understanding the Creative Mind*

Ashwin Ram
Linda Wills
Eric Domeshek
Nancy Nersessian
Janet Kolodner
College of Computing
Georgia Institute of Technology
Atlanta, Georgia 30332-0280

1 Computational Creativity

Margaret Boden, a master at bringing ideas from artificial intelligence and cognitive science to the masses, has done it again. In *The Creative Mind*, she has produced a well-written, well-argued review and synthesis of current computational theories relevant to creativity. This book seems appropriately pitched for students in survey courses and for the intelligent lay public. And if ever there were a topic suitable for bridging the gap between the researchers and the layperson, this is surely it: What is creativity, and how is it possible? Or, in computational terms (the terms that Boden argues ought to be applied): what are the processes of creativity?

Boden's stated goal is to explain how creativity is possible, where creativity is taken to be a psychological phenomenon, and an explanation of possibility is taken to be a computational process. As computationalists with active interests in creativity, we find this perspective congenial. But while offering many examples of creativity and surveying many approaches to creativity, the book leaves most details of the processes of creativity and their interactions unexplicated. Nevertheless, although Boden does not deliver a full-fledged computational explanation of the phenomenon, she does provide a strong argument that such an explanation is possible.

The early motivational sections of the book enthusiastically play up the notion that creativity, as opposed to "mere novelty," is somehow paradoxical. The

---

*To appear in *Artificial Intelligence*. A short version of this review appeared in *Behavorial and Brain Sciences*, 17(3):549, 1994.

middle section offers broad but shallow coverage of existing computational models with some emphasis on connectionist approaches. The exposition is studded with excellent examples of creativity drawn from the worlds of high culture and epochal science. Major chapters are devoted to "Unromantic Artists" and "Computer Scientists" – that is, to computer programs that have been built to simulate artistic and scientific creation and interpretation (e.g. AARON, BORIS, TALE-SPIN, ARCS, DENDRAL, BACON, AM, and so on). The final section dwells on a grab-bag of ancillary issues such as the relationship of randomness to creativity, the degree to which creativity is reasonably conceived as a special gift limited to the chosen few, and the nature of computational theories and explanations of phenomena such as creativity (including a round in Searle's Chinese Room).

Our interest is primarily in the examples, implementations, and theories that comprise the middle section of the book, but it is important to spend a bit of time understanding the problem as Boden has laid it out. As noted, Boden's goal is a computational account of the psychological phenomenon of creativity. Her achievement is to make the possibility of such a theory seem more probable (or, perhaps, at least conceivable). Both goal and achievement, however, must be contrasted with other possible ends. Boden designates the object of her study as "P-creativity" (or *psychological creativity*), distinguishing it from at least two other related concepts: "mere novelty" and "H-creativity" (or *historical creativity*).

P-creativity is a cognitive notion. By asking how some individual came up with an idea that seems beyond what they ought to be able to think, one concerns oneself with thought processes, and can deploy all the tools of computational modeling to understand these processes. In contrast, H-creativity refers to judgments that are made by a culture about the novelty and worth of ideas. Boden downplays the value of this standard, arguing convincingly that H-creativity is overly restrictive, and that P-creativity is the more significant in that H-creativity typically results from it. Boden chooses many H-creative ideas as glamorous examples, but the assumption remains that most instances of H-creativity must in the end be explained in terms of some individual's P-creative act. We agree that the important scientific question is how P-creativity could happen and that the right kind of answer to this question is a computational one. After all, no one has much of a handle on a computational model of culture. The key distinction between P-creativity and H-creativity is Boden's position that creativity is an attribute of mental *processes* rather than mental *products*. Although there is consensus that historically significant innovations are creative, Boden holds that what is creative when thought by one individual may not be so when thought by another. As computationalists, we like this emphasis on process over product in defining creativity.

Furthermore, we believe that a creative outcome is not the outcome of extraordinary mental processes, but of mechanisms that are on a continuum with those used in ordinary thinking. In our view (and Boden's), extraordinary out-comes arise from the application of ordinary mechanisms, enhanced and applied with conscious (strategic) control. For example, later in this review we describe Maxwell's use of analogy in deriving the electromagnetic field equations. In doing so, Maxwell constructed a hybrid analogical source model for electromagnetism that draws physical and mathematical constraints from two mechanical source domains: continuum mechanics and machine mechanics. It is not ordinary to construct a hybrid hypothetical analogy as Maxwell did, but analogy is an ordinary mechanism. To understand creativity, we need to understand what is different about the employment of ordinary mechanisms in creative problem solving. The focus on the outcome, for example, as in historical creativity, may provide criteria for what counts as a creative idea, but not an understanding of what is a creative reasoning process. (We will return to this point later.)

Boden's distinction between psychological and historical creativity is important (in fact, indispensable) in establishing the book's focus. Opposing P-creativity to mere novelty is also important. It serves to rule out easy, boring cases of new ideas that are not interestingly new. How Boden makes this distinction, however, strikes us as somewhat problematic. Boden argues that true creativity (as opposed to mere novelty) occurs when a person thinks a thought that is outside the space of thoughts that are even conceivable to that person – outside, as it were, their knowledge level (Newell, 1982). To clarify this idea, she invokes representations, rules, and search spaces, noting that fixing these constructs limits what can be generated by the thought processes of the reasoner. Creativity, then, requires the modification of these structures in order to expand their generative capacity.

Notice, however, that these clarifications have the effect of building aspects of a particular computational account of mental life into the definition of creativity. The effect is to limit the range of computational explanations up for consideration to those that are expressible within the particular computational paradigm chosen to model the mind. We believe the choice of constructs playing a role in mentation (and thus up for modification) are subject to debate; as will be elaborated below, we would invoke constructs such as cases, indexing structures, adaptation rules, and control strategies.

The balance of this review is organized as follows. In the next section we offer an initial critique of Boden's approach to characterizing creativity, and raise a set of questions we believe must ultimately be addressed (though we certainly do not claim to be able to answer all of them). In Section 3, we lay out our preferred framework for thinking about and modeling creativity – a framework in which much of mental life depends on the retrieval and manipulation of past experiences. Within this *case-based reasoning* framework, we focus on five major influences on cognition (and thus on the potential for creativity); each of these five influences is illustrated using examples of mechanical design, the first of three different domains we have studied, and is related to some of Boden's observations. Section 4 takes up some of the issues raised in our early critique of Boden's model, using examples of everyday creative interpretation (our second domain) to argue against the notion of special creative processes. Section 5 focuses on "constructive modeling," which integrates analogical reasoning with visual reasoning and thought experimentation. The value of this process and how it fits into our framework is illustrated by an example of historically (and psychologically) important scientific creativity (our third and final domain). Section 6 concludes this review by summarizing our approach to modeling creativity and relating it to Boden's position.

2 Characterizing the "Thinkable"

Although we disagree with Boden's choice of constructs, one needs *some* characterization of the space of thoughts that are *ordinarily thinkable* by the computational model, and the set of modifications to the thought-generating elements in the model that modify this space in an interesting manner. Ideally, what counts as an interesting modification should be specified in a manner independent of the particular computational modeling paradigm, although the modification mechanisms themselves can, of course, only be specified in the chosen formalism. In other words, the issue is: whatever the constructs involved in mentation, be they cases, rules, or search spaces, what counts as the "ordinarily thinkable," and what counts as a "creative" (as opposed to mundane) modification of the space of ordinarily thinkable thoughts?

We agree with Boden in that she refrains from defining the thinkable in terms of what is derivable through deduction from the reasoner's knowledge (as, for example, is often done in formalizations of "knowledge levels" (Dietterich, 1986; Newell, 1982)). Instead, the search space includes everything derivable from all the available reasoning operators (which could, and usually do, go beyond deduction). However, this leads to the paradox that, in some sense, every thought must be part of the set of thoughts that could be generated through available reasoning operators; if one comes to think a thought, it must have been thinkable. Boden's answer to this is that some operators carry out *conceptual change* (Carey, 1985; Nersessian, 1992; Ram, 1993; Thagard, 1992) and thus fundamentally modify the search space.

This account falls short in two ways: first, conceptual change is as elusive a notion as creativity itself (Nersessian, 1992), and, second, it is not obvious why the search space generated by application of conceptual change operators is not considered part of the thinkable. An independent (and operationalized) characterization of what makes these conceptual change operators different from all the other more ordinary inferential operations is needed.

In particular, consider Boden's formulation of thought as a search over a given search space defined by a set of constraints, operators, and representations. Boden implies that creative search involves changing or extending the constraints, operators, or representation, using an additional set of operators (with associated constraints and representations) whose job it is to modify the first set. Thus, ordinary thought is a search over an ordinary (albeit non-deductive) search space, whereas creative thought is a *meta-search* using a separate set of operators. While such an account, in principle, is perfectly acceptable, it is unclear what theoretical principles would license the placement of a given operator (or piece of knowledge) into one or the other of the search or meta-search categories. As we will elaborate below, we do not believe there are special meta-search operators that are different from ordinary inferential mechanisms.

Furthermore, we are skeptical that those individuals noted for producing many interesting ideas undergo radical conceptual change in order to produce each idea. Although this may be true of many historically significant ideas, we would prefer a model of *long-term conceptual development* in which the individual evolves a search space, that, when explored by *normal* thought processes, still includes many thoughts that would be considered *creative*.

These objections notwithstanding, we are fully sympathetic with Boden's goal of explaining creativity by appeal to computational processes. We were, therefore, most interested in the particular set of processes suggested: heuristic search (as in BACON), multiple levels of representation (as in BORIS), fuzzy matching (using an unspecified connectionist implementation), and most notable, conceptual change (unimplemented).

We agree with the idea of creativity emerging through multiple interacting processes, but we think that Boden's account leaves open several questions. First, the discussion of the mechanisms, though suggestive, is more descriptive than computational. Second, it is unclear what the overall process model is: How do all these mechanisms fit together? How do they interact? Do they operate on the same representations? If not, how do they communicate, and what do they communicate about? A third set of issues relates to Boden's suggestion that these processes are not unique to specially endowed individuals. It is never quite clear whether these processes are unique to creative thought, or, if not, what distinguishes those thoughts that are creative from those that are not, within a single individual.

3 Five Aspects of Thought

Parallel to and independent of Boden's analysis, we have been studying creative reasoning in several different domains, with a similar goal of producing computational process models of creativity. Much of what we have found concurs with Boden's observations and proposals, but we are seeking more specifics and more coherence in our models. We believe that in order to analyze creative reasoning, one needs a theoretical computational framework in which to model thinking. To this end, we propose using a computational approach rooted in case-based reasoning (Kolodner, 1993). This paradigm is fundamentally concerned with memory issues, such as remindings from partial matches at varying levels of representation and the formation of analogical maps between seemingly disparate situations – exactly the kinds of phenomena that researchers up to, and including, Boden have highlighted as central to creativity.

Accordingly, we see creative thought, like all thought, as involving processes of problem interpretation and problem reformulation, case and model retrieval, elaboration and adaptation, and ultimately, evaluation. Interpretation and reformulation are part of *situation assessment* – the process of redescribing a problem in the vocabulary of a memory's indexing scheme. Elaboration and adaptation include standard analogical processes as well as the more general process of constructive modeling, discussed at length in Section 5. Evaluation includes outcome determination, be it by simulation or by case-based prediction. All of these processes follow from our enriched case-based reasoning model (Kolodner, 1994), and fit together into a coherent whole within that framework. Research in case-based reasoning has provided extensive knowledge of how to analyze and reformulate problems, how to reuse solutions to old problems in new situations, how to build and search libraries of experience, how to merge and adapt experiences, and how to evaluate candidate solutions.

Our examples of creativity are drawn from three disparate domains: We are studying creativity in the everyday activities of average people by studying the design of mechanical devices (Kolodner & Wills, 1993a; Wills & Kolodner, 1994a, 1994b) and by looking at the processes involved in reading and understanding science fiction stories (Moorman & Ram, 1994a, 1994b; Ram, 1993). At the same time, we are examining and analyzing what led to the significant scientific discoveries of Maxwell and Faraday (Nersessian, 1984, 1992, 1993). Examples drawn from these studies, as well as Boden's own examples, will be used to illustrate our points.

Our research suggests that creativity is not a process in itself that can be turned on or off; rather, it arises from the confluence and complex interaction of *inferences* using multiple kinds of *knowledge* in the context of a *task* or problem and in the context of a specific *situation*. Much of what we think of as creativity arises from interesting *strategic control* of these inferences and their integration in the context of a task and situation. These five aspects – inferences, knowledge, task, situation and control – are not special or unique to creativity but are part of normal everyday thinking. They determine the *thinkable*, the thoughts that the reasoner might normally have when addressing a problem or performing a task.

To give a taste of what we mean by each of these five aspects, the next five sections give examples of each aspect in the context of design. Design is a pervasive form of thinking which most people do every day, not just in specific engineering contexts. All five aspects of thought are involved in design reasoning along the entire continuum from routine to creative design. The goal of this section is to give examples of the five aspects that determine the thinkable. The next section discusses what it means to go beyond the thinkable with respect to these five aspects.

3.1 Inferential Mechanisms

We have performed an exploratory study in which we observed a four-person team engaged in a seven-week undergraduate mechanical engineering (ME) design project (Kolodner & Wills, 1993a; Wills & Kolodner, 1994a). The task was to design and build a device to quickly and safely transport several eggs from one location to another. In this study, we observed that designers move fluidly between a variety of inferential methods. Typical ones include problem understanding, decomposition, elaboration, and redescription, as well as remembering, adapting, and merging design artifacts previously seen.

For example, while trying to think of ways of launching a heavy transport device, carrying several eggs from a pool of water, our ME designers recalled the behavior of a submarine submerging and launching a missile. This helped them to visualize the desired behavior of the device being designed and to elaborate the problem specification. While visualizing and acting out the missile launch, the students noticed that submarines launch missiles one at a time. This led to a redescription of the problem from launching a group of eggs in a single launch to launching each egg individually in multiple launches. The students went on to merge this idea with other ideas they had earlier, such as enclosing each egg in a tennis ball for protection (an adaptation of an earlier idea to enclose several eggs in a NERF football).

Such inferences are driven and guided by the evaluation of proposed design ideas through critical analysis, as well as by experimentation and mental simulation. The generative mechanisms, guided by critiques, respond to opportunities to create new alternatives by merging or adapting proposed ideas. The design specification is incrementally updated as ideas are tested and flaws or desirable features become apparent.

The types of inferential methods we observed (e.g., problem elaboration and redescription, solution remembering, adapting, and merging) were applied throughout the design process to produce routine (thinkable) as well as innovative ideas. They were applied in a flexible and highly opportunistic manner, with their application heavily influenced by the other four aspects of thought. Computational models of several inferential mechanisms exist, which exemplify the inferential aspect of thought. These include:

- reinterpretation of an idea in terms of a different but familiar idea (e.g., Jones (1992) shows how this can lead to useful problem reformulations which facilitate the operationalization of abstract advice (in the form of proverbs) during planning situations),
- visualization, mental simulation, and thought experimentation, which we have seen to be useful in evaluating and elaborating ideas, and in reformulating problems in design (Kolodner & Wills, 1993b) and scientific reasoning (Nersessian, 1992, 1993),

- constraint relaxation and substitution, which is useful in problem reformulation and elaboration (e.g., Moorman & Ram (1994a) show how new concepts can be formed or understood, while reading science fiction stories, by systematically tweaking constraints on known, familiar objects),
- relaxing constraints during memory search, which facilitates problem reformulation and retrieval (e.g., Turner (1994) calls this *imaginative retrieval* and shows how it can be used to retrieve ideas for writing short stories),
- relevance assessment, which is useful, for example, in retrieval and evaluation (Ram & Leake, 1991), and
- explanation of anomalies, which is also useful in retrieval and evaluation (e.g., (Ram, 1994; Schank, 1986)).

3.2 Knowledge Sources

Our second aspect of thought is knowledge. Designers draw on a variety of knowledge sources, particularly previous design experiences, accumulated from personally designing artifacts, studying case studies of designs in school, and observing artifacts designed by others. Designers typically work within a "design culture" (Navinchandra, 1992) of common engineering practices, design styles, techniques, and technologies. Innovation often arises when ideas from one culture are applied in another. In our ME design study, one designer drew much inspiration from automotive engineering, a design culture in which he is intensely interested. Many of his ideas came from recalling devices and concepts from the car domain, such as shock absorbers, unit-body versus single-frame construction, and air-bags.

A crucial part of what makes this transfer possible involves understanding, elaborating, and redefining the given problem specification to make connections to domains with which they are familiar. Designers often build on their knowledge of previous, similar problems (and their solutions) to derive new constraints and priority structures that improve or go beyond those stated in the original problem description. For example, our ME designers redefined their launch problem, based on recalling how submarines launch missiles. They derived evaluative issues and new criteria and constraints, based on their experiences with devices such as cars, toys and sports equipment, as well as designs for previous high-school egg-drop projects.

Many of the aspects of constraint exploration we observed in our designers can be experienced by Boden's reader when, in Chapter 4, she encourages the reader to play a game of necklace building within a set of rules. As Boden points out, the construction and exploration of conceptual spaces is often facilitated by drawing analogies to familiar concepts so that knowledge and reasoning techniques can be transferred to the current problem. As we will show later, the same sorts of redescription and construction of conceptual structures occur in the other two areas we have studied – science fiction reading (in which new concepts must be invented to understand the stories) and scientific discovery (in which new hybrid models are designed by merging pieces of knowledge from multiple source domains). We call this process *constructive modeling* (Clement, 1989; Moorman & Ram, 1994b; Nersessian, 1992, 1993, in press; Nersessian & Greeno, in process). Other existing mechanisms for accessing and manipulating knowledge sources include redescription and abstraction, such as reinterpretation of data at a higher level (for example, symbolic interpretation of numerical data (Ram, 1993; Kuipers & Byun, 1991)), and cross-contextual analogy (e.g., (Ram, 1993; Schank, 1982)).

Transferring knowledge from one design culture (or domain, in general) to another is not necessarily P-creative. However, identifying a domain as relevant, figuring out which pieces of knowledge or which strategies can be transferred to a new problem, and how to adapt and combine them to solve the new problem can be a creative process. These are important questions of focus which Boden does not address, but which are central to understanding what guides exploration within a generative system. (Boden is concerned more with how creativity is possible than with what guidance can make it more probable.) We believe many of the answers to these focus-related questions come from the task at hand and the situational context.

3.3 Tasks

A third aspect influencing what is thinkable is the task. Design is a complex task, involving several subtasks, such as brainstorming, critiquing, gathering information about and elaborating ideas, and finding, constructing, and integrating design pieces. Which aspects of a remembered design experience or a proposed design alternative the designer focuses on depend on what is relevant to the task at hand. This can greatly influence the strategic control of the design process, as well as which new constraints or criteria are added to the design specification and which elaborations or adaptations of ideas are suggested.

For example, there are numerous facts associated with submarines, but our designers were drawn to the fact that they launch missiles one at a time, as opposed to, for example, facts about how missiles are aimed at their target or about the cramped, claustrophobic interior. They were viewing the submarine missile launch from the perspective of trying to borrow its solution to the problem of initiating a powerful launch from water; thus, what was relevant was the detail that multiple, relatively small missiles are launched one at a time. This focus on individual launches helped suggest a new way of looking at the problem (Kolodner & Wills, 1993b).

3.4 Situation

Situation is our fourth aspect of thought. Design does not typically occur in a vacuum. Rather, designers usually try to experiment with their design (e.g., a mock-up, simulation, prototype, or partial construction) in a real-world situation (e.g., the typical operating environment, a potential maintenance situation, a worst-case scenario). This provides concrete feedback that can refine the problem specification to require any positive features noticed and to prohibit any flaws that were detected. At the same time, the evolving specification can be used to reinterpret entities in the environment and realize their relevance to the problem at hand.

Designers operate in a rich context of ideas, which are not only recalled and adapted from previous experiences, but also recognized in the current external environment. (That is, the environment can be a source of inspiration, in addition to knowledge and experiences recalled.) The continual elaboration and reformulation of the problem and desired solution primes the designer to recognize good ideas when they are stumbled upon. Problem redescription often enables the designer to overcome functional fixedness and notice new, alternative functions and uses for common design pieces. This leads to insights into new ways of solving pending problems (thus facilitating serendipity).

For example, at one point in the ME design project, the students were considering using a spring launch device, but had the problem that the springs bent when compressed. After generating, simulating, and critiquing a few proposals, they augmented their specification to require that each spring be enclosed in a collapsible tube. However, they could not immediately think of anything that could serve as a collapsible tube, so they temporarily gave up on designing the launch mechanism. Later, as they were looking for protective egg cushioning material, they came across toilet paper holders and immediately recognized them as the collapsible tubes they needed to keep springs straight (Wills & Kolodner, 1994b). By playing with the springs, noticing problems, and suggesting fixes, the designers formed a specific, concrete description of what they needed. This description was used to reinterpret the paper holder when it was seen and to recognize its additional function of preventing springs from bending upon compression.

Being situated facilitated the designers' discovery by bringing to their attention objects that could solve their problem without requiring the objects to be recalled as relevant solutions. Playing with the springs in a concrete situation also provided feedback to help the designers elaborate and refine their description of what they needed. The designers became immersed in the problem – redescribing it and viewing it from multiple perspectives, considering, comparing, and critiquing several options – so that when a relevant solution was spotted, the way it fit into the problem was immediately discerned.

The importance of becoming immersed in the problem situation is implicitly acknowledged by Boden when she interrupts Chapter 4 to encourage the reader to temporarily stop reading and to play the necklace-building game. She suggests that the reader practice building necklaces (with pencil and paper), play around with the rules, record any interesting things that are noticed, etc. Although Boden does not analyze why this is so important, constructing specific necklace-building situations does provide feedback that can help the reader understand the problem constraints, their implications, and ways of modifying them.

3.5 Strategic Control

Finally, the fifth aspect of thought is the strategic control of inferences. Designers must make many decisions over the course of a design: which idea to elaborate or adapt next, which constraint to relax, how to set priorities. They also move between various tasks, subproblems, and design processes in a flexible and highly opportunistic manner.

We observed a variety of strategic control heuristics used by our ME designers. Some were opportunistic. An example is letting extremes distract. When an alternative was proposed that satisfied some desired criteria extremely well compared to the other alternatives, our designers directed their efforts toward elaborating that alternative (Wills & Kolodner, 1994b). They optimistically suspended criticism or discounted the importance of criteria or constraints that were not satisfied as well. Suspending criticism during brainstorming is a common strategic ideation technique which involves taking a cognitive risk. A similar mechanism is seen in creative interpretation, in which the reader must suspend disbelief in unfamiliar aspects of a story in order to understand it (see below). Sometimes, as constraints are relaxed or placed at a lower priority, an opportunity to reformulate the problem is revealed (Kolodner & Wills, 1993b). Noticing invariants (Kaplan & Simon, 1990), as well as anomalies, can also aid in understanding a problem and reveal ways of redescribing it.

Some strategic control heuristics are more deliberate, based on reflection. For example, one heuristic our designers used was to try quick, easy adaptations of a proposed solution first before stepping back and reformulating the problem or relaxing constraints (Wills & Kolodner, 1993a, 1994a). Other deliberate heuristics include making non-standard substitutions (Kolodner, 1994; Kolodner & Penberthy, 1990), applying adaptation strategies in circumstances other than the ones they were meant for (Kolodner, 1994; Navinchandra, 1992), merging pieces of separate solutions with each other in nonobvious ways (Kolodner, 1994; Kolodner & Penberthy, 1990), and goal-directed inferential control (Nersessian, in press; Ram, 1991; Ram & Hunter, 1992).

Often, creativity arises when a set of "normal" strategies are applied to a situation in which a run-of-the-mill solution is not immediately forthcoming and the control heuristics allow the reasoner to devote more resources to the problem, looking further and further afield for possible knowledge and strategies until something results in a creative solution. Examples include a problem reformulation that takes several steps; an analogy to a far-off case or model; an analogy from a hybrid analog constructed incrementally from more than one source; a strategy imported from a different problem-solving culture; an unexpected and novel opportunity afforded to the reasoner by virtue of an unusual task context. Many of these could happen during "ordinary thought," but most thought does not allow enough leeway to look that far or to play with ideas for that long or it does not occur in a context that affords such an opportunity.

4 Beyond the Thinkable

Based on this view of creative thought, we offer a very pragmatic definition of the *normal* search space. It is not the deductive (or other) closure of everything that is known – an inherently uncomputable concept. Rather it is the *space of the thoughts one would usually explore in a pragmatic context.* There may be cases where important possibilities are outside the space of theoretically conceivable thoughts. (Perhaps rings of carbon atoms could never arise within the chemical theory prevailing at the time Kekule tackled benzene.) But, in other cases, thoughts that are within the theoretical space are nevertheless pragmatically inconceivable (e.g., the discoveries made by Swanson's (1990) program which are nevertheless H-creative). In creative individuals, even the usual search space may be interestingly different or expanded so as to provide the basis for creative thought using the very same mechanisms that on other occasions would produce more mundane thoughts.

Consider, for example, the problem of reading a science fiction story. Although creativity is usually thought of in the context of problem-solving or inventive tasks, we believe that creativity is an essential and ubiquitous component of other kinds of reasoning tasks as well, including explanatory and comprehension tasks. In point of fact, all these tasks involve understanding. Reading science fiction stories requires what we call *creative understanding,* in which the reader must learn enough about an alien world in a short text in order to accept it as the background for the story and simultaneously must understand the story itself. Creative understanding requires the extrapolation, modification, or extension of existing concepts and theories to invent new ones (Moorman & Ram, 1994a, 1994b; Ram, 1993). The extrapolation is constrained by the content of the story, by the system's existing concepts and theories, and by the requirements of the reading and understanding task.

As an example, consider the following short story, *Men Are Different* by Alan Bloch (1963).

> I'm an archaeologist, and Men are my business. Just the same, I wonder if we'll ever find out about Men – I mean *really* find out what made Man different from us Robots – by digging around on the dead planets. You see, I lived with a Man once, and I know it isn't as simple as they told us back in school.

> We have a few records, of course, and Robots like me are filling in some of the gaps, but I think now that we aren't really getting anywhere. We know, or at least the historians say we know, that Men came from a planet called Earth. We know, too, that they rode out bravely from star to star; and wherever they stopped, they left colonies – Men, Robots, and sometimes both – against their return. But they never came back.
>
> Those were the shining days of the world. But are we so old now? Men had a bright flame – the old word is "divine," I think – that flung them far across the night skies, and we have lost the strands of the web they wove.
>
> Our scientists tell us that Men were very much like us – and the skeleton of a Man is, to be sure, almost the same as the skeleton of a Robot, except that it's made of some calcium compound instead of titanium. Just the same, there are other differences.
>
> It was on my last field trip, to one of the inner planets, that I met the Man. He must have been the last Man in this system, and he'd forgotten how to talk – he'd been alone so long. I planned to bring him back with me. Something happened to him, though.
>
> One day, for no reason at all, he complained of the heat. I checked his temperature and decided that his thermostat circuits were shot. I had a kit of field spares with me, and he was obviously out of order, so I went to work. I pushed the needle into his neck to operate the cut-off switch, and he stopped moving, just like a Robot. But when I opened him up he wasn't the same inside. And when I put him back together I couldn't get him running again. Then he sort of weathered away – and by the time I was ready to come home, about a year later, there was nothing left of him but bones. Yes, Men are indeed different.

In order to understand this story, the reader must infer that the narrator is a robot, that robots are the dominant lifeform in the future, that humans have practically died out, that robots are capable of making factual errors such as the ones that the narrator made, and so on. The reader must construct an appropriate model of this world, and interpret the story with respect to this model even as the model evolves. The reader must also be willing to *suspend disbelief* (Coleridge, 1817/1926; Corrigan, 1979) to understand concepts which do not fit into a standard world view. This is another example of a strategic control mechanism that requires a willingness to take a cognitive risk.

In *Men Are Different*, robots, which in the real world are physical objects used as tools in manufacturing, are conceptualized as independent volitional agents. The new concepts are constructed by merging and extending the existing concepts representing human agents and robotic artifacts, resulting in a novel view of the situation at hand (Moorman & Ram, 1994a). The reader must adopt this view to build an appropriate story model. Interestingly, the irony in this story derives from the fact that the robot in the story performs what one might view as the reverse inference: conceptualizing the man as a physical object to be repaired in a manner that one might use to repair a physical robotic device (Moorman & Ram, 1994b).

It would, of course, be unreasonable to assume a special purpose "meta-search space" generator for science fiction story understanding. The creative understanding processes required to read *Men Are Different* are not unique to science fiction stories; understanding any fictional story requires similar kinds of processing. The same is true of nonfictional stories as well as unfamiliar real-world scenarios, although the types and degree of conceptual modifications required may be different.

Thus, reading a science fiction story is presumably accomplished within the same type of search space and using the same set of reading and comprehension operators as reading a mundane narrative. The example illustrates that these ordinary operators and processes can take the reasoner out of the space that would usually be explored. In fact, situations like this show just how fluid the movement is from the usual to the unusual.

The question, of course, is how the search space comes to be expanded to facilitate creative thought using ordinary mechanisms. If normal traversal of a search space depends on knowledge, inferential methods, and control methods, then interesting paths may result from modifying any of these three components. Most obviously, transformations of basic knowledge (e.g., conceptual change) can yield new results. But application of new inferential methods can also produce novelty; for example, adopting a heuristic from a different task context, such as an architect adopting the engineer's heuristic of "incorporate the obstacle." Finally, differences in control methods will produce differences in results; consider methodological differences between scientists, such as the willingness to take cognitive risks, the willingness to explore a "silly" idea, the ability to evaluate and prune unlikely candidates. For example, we would rate AM+Lenat as a creative combination even though AM by itself was not. Analysis of the task and situation influences the knowledge, inferential methods, and control strategies that are available.

5 Constructive Modeling

Reading and understanding *Men are Different* requires the invention of a system of concepts and theories that represent a sentient, humanoid robot, through the extension of one's prior understanding of multiple concepts, such as volitional agents, men, and industrial robots (Moorman & Ram, 1994a). In creative design, too, new conceptual structures are formed from multiple sources. Problem descriptions are incrementally elaborated and reformulated, typically by analogy to pieces of several similar problems. New design ideas are generated by combining several ideas from experiences with existing devices. The behavior of a proposed design is predicted, simulated, and visualized based on multiple pieces of knowledge of how related devices or design pieces work.

These are everyday instances of the constructive modeling process we have found to be central in significant scientific discoveries throughout the history of the sciences. For example, it figures centrally in the development of the field representation of electromagnetic forces by Michael Faraday and James Clerk Maxwell. Here we will illustrate our points by looking briefly at Maxwell's derivation of the electromagnetic field equations (Maxwell 1890). The Maxwell case reinforces Boden's contention that even in instances of H-creativity, explaining the episode demands an analysis of P-creativity.

This case shows constructive modeling to be a dynamic process involving analogical and visual modeling as well as thought experimentation (mental simulation) to create sources where no direct analogy exists (Clement, 1989; Nersessian 1992, 1993, in press; Nersessian & Greeno, in process). What distinguishes this process from the computational models of analogical reasoning Boden discusses is that they employ cases where the analogical base is ready to hand. Further, although Boden does note the importance of visual representation in some instances of analogy, neither she nor the computational models she discusses attempt to integrate it into their accounts. Indeed, we believe the constructive modeling processes identified in the Maxwell case show the need for an integrated account of analogy, visual representation, and mental modeling for understanding creative thinking.

Finally, this case points to something missing entirely from Boden's analysis. The social context is crucial to understanding a creative episode in science - and we presume in more ordinary cases, too. Maxwell's location in Cambridge led to his training as a mathematical physicist. This determined the nature of the theoretical, experimental, and mathematical knowledge and the methodological practices with which he formulated the problem and approached its solution. The work of Faraday and William Thomson (later, Lord Kelvin) contributed to these as well. Continental physicists working on electromagnetism at the same time employed quite different practices and drew from fundamentally different mathematical and physical representational structures. These kinds of social factors can be figured into the account without our being required to produce a computational model of culture.

Maxwell's constructive modeling process provides a good example of an instance in which all five of the aspects of creative thinking we have been discussing are employed. He used multiple knowledge domains and informational formats, in the context of solving a complex problem within specific cognitive and social situation. Maxwell exercised strategic control continually to evaluate the models and the inferences he drew from them, and to integrate the solutions to the sub-problems into a consistent mathematical representation. The modeling process involved adjusting multiple constraints drawn from

- the physics of elastic fluids,
- experimental data on electricity and magnetism,
- Faraday's hypotheses about the lines of force that form when iron filings are sprinkled around magnets and charged matter (Faraday, 1835-55),
- Faraday's visual lines of force model (shown in Figures 1a and 1b), accounting for continuous transmission and interconversion of forces (Maxwell, 1890, vol. 1, pp. 155-229),
- Faraday's interlocking curves model (shown in Figures 2a and 2b), representing the dynamical balance between electricity and magnetism (Maxwell, 1890, p. 194n), and
- William Thomson's hypothesis of rotational motion of magnetism and his analogies, and mathematical equations (Lamor, 1937).

Maxwell's goal (Maxwell, 1890, vol. 1, pp. 451-513) was to provide a unified representation of the continuous transmission of electric and magnetic forces that he hoped would encompass optical phenomena as well. The full model is an imaginary hybrid construction that integrates physical and mathematical constraints from two analogical source domains – continuum mechanics (fluids, elastic media, etc.) and machine mechanics – with constructs from magnetism and electricity. Unlike the cases customarily considered in the literature on analogy, where an existing problem solution in the source domain is transferred to the target domain, in this case, the source and target domains interact to create and modify a series of constructed models that become the objects with which Maxwell reasoned (Nersessian, in press; Nersessian & Greeno, in process). Further, reasoning with the models demands that they provide simulations and thus be animated in a manner similar to thought experiments (Nersessian 1993). In the text itself, Maxwell provided an extensive set of instructions for how the reader should visualize and animate the models.

Maxwell's model construction proceeded as follows. Maxwell first constructed a primitive model (Figure 3a) consistent with the constraints discussed above: a fluid medium composed of elastic vortices and under stress. With this form of the model he was able to provide a mathematical representation for several magnetic phenomena. Analyzing the relationships between current and magnetism required alteration of the model. We can see in Figure 3a that all the vortices are rotating in the same direction, which means that since they touch, friction is produced and they will eventually stop. Mechanical consistency, thus, requires the introduction of "idle wheels" (as in machine gears) surrounding the vortices, and Maxwell argued that their translational motion could be used to represent electricity. Figure 3b shows a cross section of the hybrid model. For the purposes of calculation, Maxwell now had to make the elastic vortices into rigid pseudospheres. We can see how the imaginary system provides a mechanical interpretation for electromagnetism: motion of the particles creates motion of the vortices and vice versa. In this model, as was known experimentally, electric current produces magnetic effects and changes in magnetic effects produce current. Using the model, he derived mathematical equations to represent these relationships.

It then took Maxwell nine months to figure out how to represent the final – and most critical – piece of the problem: electrostatic actions. He found that if he made the vortices elastic and identified electrostatic polarization with elastic displacement, he could calculate the wave of distortion produced by polarization. That is, adding elasticity to the model enabled him to show that electromagnetic actions are propagated with a time delay, i.e., they are field actions and not Newtonian actions at a distance. At this point, we have a fully mathematized representation of the electromagnetic field. There are significant sign "errors" in this part of Maxwell's analysis, but Nersessian (1984, in press) has argued that all but one (a minor substitution error) can be seen not to be errors when we view him as reasoning via the constructed model.

This case study illustrates that it was through a process of embodying physical and mathematical constraints in a series of constructed models and reasoning about and with these that Maxwell generated the field equations for electromagnetism – an *historically and individually* creative process.

6 Summary and Conclusions

Inference and the control of inference, knowledge representation and representational change: these are the main interrelated pieces of the creativity puzzle. Each relies heavily on episodic and semantic memory. Together, they fit into a model of reasoning that is recognizable as (but looser than) case-based reasoning. A creative individual is one in whom these factors combine to form a search space – a repertoire of thoughts – that is different from the usual and contains many creative ideas waiting to be constructed. Of course, the search space can only be explored in the context of a task or problem and a specific situation; thus, the repertoire is defined pragmatically, and serendipity (as Boden points out) plays an important role.

In a specific individual, more creative thoughts will likely result when these pieces come together in a novel way to yield an unexplored and unexpected path through the search space. Creativity, as Boden points out, is not an all-or-none phenomenon. Every new thought is creative to some extent. Every new thought results from those same processes that, on occasion, produce results we value as creative. The more the search space is varied in a given context (through representational change, novel inferential methods, or strategic control heuristics), the more creative the resulting thoughts are likely to be. Over time, an individual may become more expert as he or she acquires (or reformulates) knowledge, reasoning strategies, and methodologies that change the search space or how it is explored.

The framework we have sketched here is broadly compatible with Boden's, but is more specific in its suggestions for integrating multiple types of interacting and interactive processes in a task context. In accounting for creativity, we emphasize issues of control and the role of experience (or cases). By focusing on how mental activity is directed towards a task in some situation, we ensure that the resulting theory addresses pragmatic issues in thinking and control of thinking. As Boden would require, our approach is computational. We believe, in fact, that the greatest contribution of *The Creative Mind* is the clear case it presents for the legitimacy of computational theories of creativity. Boden leads the reader to an understanding of that goal, and, having framed the question, suggests how research might proceed towards a meaningful answer.

7 Acknowledgments

This paper is based in part on research by Kenneth Moorman, who also provided us with helpful suggestions for improving this paper. We are grateful to our editors Mark Stefik and Stephen Smoliar for their many insights and helpful suggestions for improving this review. This work has been supported in part by the Advanced Research Projects Agency, monitored by ONR under contract N00014-91-J-4092, by NSF Grant No. IRI-8921256 and ONR Grant No. N00014-92-J-1234, by NSF Scholars Awards DIR8821442 and DIR9111779, and by the Georgia Institute of Technology. All views expressed are those of the authors.

8 Bibliography

Bloch, A. (1963). "Men are Different" in *50 Short Science Fiction Tales*, Asimov, I. and Conklin, G. (eds.), New York: MacMillan Publishing Co.

Carey, S. (1985). *Conceptual Change in Childhood.* Cambridge, MA: MIT Press.

Clement, J. (1989). Learning via Model Construction and Criticism. In Glover, G., Ronning, R., & Reynolds, C. (eds.), *Handbook of Creativity: Assessment, Theory and Research*, pp. 341-381. New York: Plenum, 1989.

Corrigan, R.W. (1979). *The World of the Theatre.* Glenview, IL: Scott, Foresman and Co.

Coleridge, S.T. (1817/1926). *Biographia Literaria; or, Biographical sketches of my literary life and opinions*, chapter 14. New York: Macmillan.

Dietterich, T.G. (1986). Learning at the Knowledge Level. *Machine Learning*, Vol. 1, pp. 287-316.

Faraday, M. (1835-55). *Experimental Researches in Electricity.* Reprinted, New York: Dover.

Jones, E. (1992). Brainstormer: A Model of Advice Taking. PhD Thesis, Yale University.

Kaplan, C. & Simon, H. (1990). In Search of Insight. *Cognitive Psychology*, Vol. 22, pp. 374-419.

Kolodner, J. (1993). *Case-Based Reasoning.* San Mateo, CA: Morgan-Kaufmann.

Kolodner, J.L. (1994). Understanding Creativity: A Case-Based Approach. In S. Wess, K.D. Althoff and M.M. Richter (eds.), *Topics in Case-Based Reasoning*, selected papers from the *First European Workshop on Case-Based Reasoning.* Kaiserslautern, Germany. November 1993. Springer-Verlag.

Kolodner, J. & Penberthy, T. (1990). A Case-Based Approach to Creativity in Problem Solving. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, Cambridge, MA. August.

Kolodner, J. & Wills, L. (1993a). Case-Based Creative Design. In *AAAI Spring Symposium on AI and Creativity.* Stanford, CA. March 1993. Reprinted in *AISB Quarterly*, special issue (no. 85) on AI and Creativity, edited by Terry Dartnall, Autumn 1993.

Kolodner, J. & Wills, L. (1993b). Paying Attention to the Right Thing: Issues of Focus in Case-Based Creative Design. In *AAAI Case-Based Reasoning Workshop*, (pp. 19-25).

Kuipers, B.J. & Byun, Y.-T. (1991). A Robot Exploration and Mapping Strategy based on a Semantic Hierarchy of Spatial Representations. *Robotics and Autonomous Systems*, 8(1-2): 47-63

Lamor, J., ed. (1937). *The Origins of James Clerk Maxwell's Electric Ideas.* Cambridge: Cambridge University Press.

Maxwell, J.C. (1890). *The Scientific Papers of James Clerk Maxwell*, W.D. Niven, ed. Cambridge: Cambridge University Press.

Moorman, K. & Ram, A. (1994a). A Model of Creative Understanding. To appear in the *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94).* Seattle, Washington.

Moorman, K. & Ram, A. (1994b). Integrating Creativity with Reading: A Functional Approach. To appear in the *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society.* Atlanta, GA.

Navinchandra, D. (1992). Innovative Design Systems: Where are we, and where do we go from here?. Parts I and II. *The Knowledge Engineering Review*, 7(3) pp. 183-213 and 7(4) pp. 345-362.

Nersessian, N. (1984). *Faraday to Einstein: Constructing Meaning in Scientific Theories.* Dordrecht: Kluwer Academic Publishers.

Nersessian, N. (1992). How Do Scientists Think? Capturing the Dynamics of Conceptual Change in Science. In *Volume XV: Cognitive Models of Science* of R Giere, Ed., *Minnesota Studies in the Philosophy of Science* 15 . Minneapolis: University of Minnesota Press.

Nersessian, N. (1993). In the Theoretician's Laboratory: Thought Experimenting as Mental Modeling. In *Proceedings of the 1992 Meeting of the Philosophy of Science Association*, Vol. 2, D. Hull, M. Forbes, and K. Okruhlik Eds. East Lansing Michigan: Philosophy of Science Association.

Nersessian, N. (in press). Abstraction via Generic Modeling in Concept Formation in Science. In Cartwright, N. & Jones, M.R. (eds). *Idealization in Science.* Amsterdam: Editions Rodopi.

Nersessian, N. & Greeno, J. (in process). Dynamic mental modeling in scientific reasoning, manuscript.

Newell, A. (1982). The Knowledge Level. *Artificial Intelligence* 18(1): 87-127.

Ram, A. (1991). A Theory of Questions and Question Asking. *The Journal of the Learning Sciences*, 1(3-4): 273-318.

Ram, A. (1993). Creative Conceptual Change. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, pp. 17-26, Boulder, CO.

Ram, A. (1994). AQUA: Questions that Drive the Explanation Process. In R.C. Schank, A. Ram, and C.K. Riesbeck (eds.), *Inside Case-Based Explanation*, Chapter 7. Lawrence-Erlbaum Associates.

Ram, A. & Hunter, L. (1992). The Use of Explicit Goals for Knowledge to Guide Inference and Learning. *Applied Intelligence*, 2(1): 47-73.

Ram, A. & Leake D. (1991). Evaluation of Explanatory Hypotheses. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, pp. 867-871, Chicago, IL.

Schank, R. (1982). *Dynamic Memory: A Theory of Learning in Computers and People.* New York: Cambridge University Press.

Schank, R. (1986). *Explanation Patterns: Understanding Mechanically and Creatively.* Hillsdale, NJ: Lawrence-Erlbaum Associates.

Swanson, D. (1990). Medical Literature as a Potential Source of New Knowledge. *Bull. Med. Libr. Assoc.*, 78(1):29-37.

Thagard, P. (1992). *Conceptual Revolutions.* Princeton, NJ: Princeton University Press.

Turner, S.R. (1994). *MINSTREL.* Lawrence-Erlbaum Associates. Forthcoming.

Wills, L.M., & Kolodner, J.L. (1994a). Towards More Creative Case-Based Design Systems, to appear in the *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, pp. 50-55. Seattle, Washington.

Wills, L.M., & Kolodner, J.L. (1994b). Explaining Serendipitous Recognition in Design, in *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, pp. 940-945. Atlanta, GA.